Geometry & Topology

Volume 27 (2023)

Issue 3 (pages 823–1272)

msp

GEOMETRY & TOPOLOGY

msp.org/gt

See inside back cover or msp.org/gt for submission instructions.

The subscription price for 2023 is US $740/year for the electronic version, and $1030/year (+$70, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Geometry & Topology is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

Geometry & Topology (ISSN 1465-3060 printed, 1364-0380 electronic) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840.

GT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

msp

Geometry & Topology **27**:3 (2023) 823–924
DOI: 10.2140/gt.2023.27.823
Published: 8 June 2023

A calculus for bordered Floer homology

JONATHAN HANSELMAN
LIAM WATSON

We consider a class of manifolds with torus boundary admitting bordered Heegaard Floer homology of a particularly simple form; namely, the type D structure may be described graphically by a disjoint union of loops. We develop a calculus for studying bordered invariants of this form and, in particular, provide a complete description of slopes giving rise to L–space Dehn fillings as well as necessary and sufficient conditions for L–spaces resulting from identifying two such manifolds along their boundaries. As an application, we show that Seifert-fibred spaces with torus boundary fall into this class, leading to a proof that, among graph manifolds containing a single JSJ torus, the property of being an L–space is equivalent to non-left-orderability of the fundamental group and to the nonexistence of a coorientable taut foliation.

57M27

1 Introduction

This paper is concerned with developing Heegaard Floer theory with a view to better understanding the relationship between coorientable taut foliations, left-orderable fundamental groups and manifolds that do not have *simple* Heegaard Floer homology. Recall that manifolds with simplest possible Heegaard Floer homology, called L–spaces,

are rational homology spheres Y for which $\dim \widehat{HF}(Y) = |H_1(Y;\mathbb{Z})|$ (all Heegaard Floer–theoretic objects in this work will take coefficients in $\mathbb{Z}/2\mathbb{Z}$). On the other hand, a group G is left-orderable if there exists a nonempty set $P \subset G$, called a positive cone, that is a closed subsemigroup of G and gives a partition of the group in the sense that $G = P \sqcup \{1\} \sqcup P^{-1}$. This auxiliary structure is equivalent to G admitting an effective action on $\mathbb{R}$ by order-preserving homeomorphisms. The work of Boyer, Rolfsen and Wiest is a good introduction to left-orderable groups in the context of three-manifold topology, including the interaction with taut foliations [4]; for closed, orientable, irreducible, three-manifolds it is conjectured that being an L–space is equivalent to having a non-left-orderable fundamental group; see Boyer, Gordon and Watson [3]. This conjecture holds for Seifert-fibred spaces and, as a natural extension of this case, graph manifolds are a key family of interest; see Boileau and Boyer [1], Boyer and Clay [2], Clay, Lidman and Watson [5], Hanselman [8; 9] and Mauricio [21]. Towards establishing the conjecture for graph manifolds, we prove:

Theorem 1.1 *Suppose that Y is a graph manifold with a single JSJ torus; that is, Y is constructed by identifying two Seifert-fibred manifolds with torus boundary along their boundaries. Then the following are equivalent:*

(i) *Y is an L–space.*

(ii) *$\pi_1(Y)$ is not left-orderable.*

(iii) *Y does not admit a coorientable taut foliation.*

The equivalence between (ii) and (iii) is due to Boyer and Clay [2]; the focus of this paper is understanding the behaviour of Heegaard Floer homology in this setting. To do this, we make use of bordered Heegaard Floer homology, a variant of Heegaard Floer homology adapted to cut-and-paste arguments. Briefly, this theory assigns a differential graded module over a particular algebra to each manifold with torus boundary. A chain complex for the Heegaard Floer homology of the associated closed manifold is obtained from a pairing theorem due to Lipshitz, Ozsváth and Thurston [20].

Our approach to this problem is to work in a more general setting. We consider a particular class of differential graded modules which we call loop-type (Definition 3.2), and introduce a calculus for studying loops; the bulk of this paper is devoted to developing this calculus in detail. Given a three-manifold M with torus boundary, M is called loop-type if its associated bordered invariants are loop-type up to homotopy (Definition 3.13). Recall that a slope in ∂M is the isotopy class of an essential simple closed curve in ∂M, and denote by $M(\gamma)$ the closed three manifold resulting from Dehn

filling along a slope γ. The set of slopes may be (noncanonically) identified with the extended rationals $\mathbb{Q} \cup \{\frac{1}{0}\}$. Of central interest is the subset $\mathcal{L}_M$ consisting of those slopes giving rise to L–spaces after Dehn filling. We prove:

Theorem 1.2 (detection) *Suppose that M is a loop-type rational homology torus. Then there is a complete, combinatorial description of the set $\mathcal{L}_M$ in terms of loop calculus. In particular, $\mathcal{L}_M$ may be identified with the restriction to $\mathbb{Q} \cup \{\frac{1}{0}\}$ of a connected interval in $\mathbb{R} \cup \{\infty\}$. Moreover, if M is simple loop-type, this interval has rational endpoints.*

We remark that this is the expected behaviour from the foliations/orderability vantage point, at least for graph manifolds. It is interesting that the analogous behaviour on the Heegaard Floer side appears to be intrinsic to the algebraic structures that arise. Namely, this is a statement that makes sense for loop-type bordered invariants, without reference to any three-manifold. The introduction of the technical notion of a *simple loop* (Definition 4.20) also allows us to state and prove a gluing theorem. Let $\mathcal{L}^\circ_M$ denote the interior of $\mathcal{L}_M$.

Theorem 1.3 (gluing) *Suppose that M_1 and M_2 are simple loop-type rational homology tori and neither is solid torus-like. Then, given a homeomorphism $h\colon \partial M_1 \to \partial M_2$, the closed manifold $M_1 \cup_h M_2$ is an L–space if and only if, for every slope γ in ∂M_1, either $\gamma \in \mathcal{L}^\circ_{M_1}$ or $h(\gamma) \in \mathcal{L}^\circ_{M_2}$.*

Note that Theorem 1.3 does not hold if either M_1 or M_2 is a solid torus. If M_1 is a solid torus with meridian $m = \partial D^2 \times \{\text{pt}\}$, then, according to the definition of an L–space slope, $M_1 \cup_h M_2$ is an L–space if and only if $h(m) \in \mathcal{L}_{M_2}$; equivalently, the statement of Theorem 1.3 holds if we use $\mathcal{L}_{M_i}$ in place of $\mathcal{L}^\circ_{M_i}$. When both M_1 and M_2 are solid tori, this simply amounts to the construction of lens spaces interpreted in our notation. More generally, however, there are bordered invariants that arise in the loop setting that behave just like solid tori with respect to gluing. We will need to deal with these explicitly; this amounts to defining a class of manifolds and loops which are referred to as solid torus-like (Definition 3.20).

Theorems 1.2 and 1.3 follow from working with loops in the abstract. Towards the proof of Theorem 1.1, and in the interest of establishing an existence result, a key class of loop-type manifolds is provided by Seifert-fibred spaces.

Theorem 1.4 *Suppose M is a rational homology solid torus admitting a Seifert-fibred structure. Then M has simple loop-type bordered Heegaard Floer homology.*

Let N denote the twisted I–bundle over the Klein bottle, and recall that the rational longitude λ for this manifold with torus boundary may be identified with a fibre in a Seifert structure over the Möbius band. As a Seifert-fibred rational homology solid torus, N is a simple loop-type manifold; compare [3]. The twisted I–bundle over the Klein bottle allows for an alternative detection statement for L–space slopes.

Theorem 1.5 (detection via the twisted I–bundle over the Klein bottle) *Suppose that M is a loop-type rational homology torus. Then $\gamma \in \mathcal{L}^\circ_M$ if and only if $N \cup_h M$ is an L–space, where $h(\lambda) = \gamma$.*

This answers a question of Boyer and Clay in the case of connected boundary; compare [2, Question 1.8]. In particular, our notion of detection aligns precisely with the characterization given by Boyer and Clay [2]. Indeed, we prove that, more generally, the twisted I–bundle over the Klein bottle used in Theorem 1.5 may be replaced with any simple loop-type manifold for which every nonlongitudinal filling is an L–space; see Theorem 7.3. There are many examples of these provided by *Heegaard Floer homology solid tori*; for more on this class of manifolds, see Hanselman, Rasmussen and Watson [12, Section 1.5].

Note that the interior of $\mathcal{L}_M$, denoted by $\mathcal{L}^\circ_M$, is the set of strict L–space slopes. The complement of $\mathcal{L}^\circ_M$, according to Theorem 1.5, corresponds to the set of *non-L–space (NLS) detected slopes* in the sense of Boyer and Clay [2, Definition 7.16].

According to Theorem 1.4, the exterior of any torus knot in the three-sphere gives an example of a simple loop-type manifold (indeed, this follows from work of Lipshitz, Ozsváth and Thurston [20]). More generally, if K is a knot in the three-sphere admitting an L–space surgery (an L–space knot), then $S^3 \smallsetminus \nu(K)$ is a simple loop-type manifold. In this setting it is well known that $\mathcal{L}_M$ is (the restriction to the rationals of) $[2g-1, \infty]$, where g is the Seifert genus of K, whenever K admits a positive L–space surgery. Specializing Theorem 1.3 to this setting, we have:

Theorem 1.6 *Let K_i be an L–space knot and write $M_i = S^3 \smallsetminus \nu(K)$ for $i = 1, 2$. Given a homeomorphism $h\colon \partial M_1 \to \partial M_2$, the closed manifold $M_1 \cup_h M_2$ is a L–space if and only if $h(\mathcal{L}^\circ_{M_1}) \cup \mathcal{L}^\circ_{M_2} \cong \mathbb{Q} \cup \{\frac{1}{0}\}$.*

Note that this solves [2, Problem 1.11] by resolving (much more generally and in the affirmative) one direction of [9, Conjecture 1] (a special case of [2, Conjecture 1.10] or [5, Conjecture 4.3]). It should also be noted that, owing to the existence of hyperbolic L–space knots, Theorem 1.6 provides additional food for thought regarding

[2, Question 1.12]. Namely, one would like to know if the order-detected/foliation-detected slopes in the boundary of the exterior of an L–space knot K coincide with the complement of $\mathcal{L}^{\circ}_{M}$, where $M = S^3 \smallsetminus \nu(K)$. For integer homology three-spheres resulting from surgery on a knot in S^3, there has been considerable progress in this vein; see Li and Roberts for foliations [18], Boileau and Boyer for left-orders [1], and also Hedden and Levine for splicing results [14].

Related work

Nonsimplicity is somewhat sporadic; however, it can be shown explicitly that nonsimple graph manifolds exist. On the other hand, it is clear from examples that certain graph manifolds with torus boundary do not give rise to loop-type bordered invariants, at least not in any obvious way. These subtleties seem particularly interesting when contrasted with the behaviour of foliations and orders: for more complicated graph manifolds, more subtle notions of foliations and orders (relative to the boundary tori) need to be considered [2]. On the other hand, it is now clear from work of Rasmussen and Rasmussen [27] that simplicity (for loops) is more than a convenience: the class of simple loop-type manifolds is equivalent to the class of (exteriors of) Floer-simple knots; see Hanselman, Rasmussen, Rasmussen and Watson [10, Proposition 6] in particular. As a result, Theorem 1.3 may be recast in terms of Floer-simple manifolds [10, Theorem 7]. This observation gives rise to an extension of Theorem 1.1 to the case of general graph manifolds; this is the main result of a joint paper of the authors with Rasmussen and Rasmussen [10], illustrating a fruitful overlap between these two independent projects. We note that the methods used by Rasmussen and Rasmussen are different from those used in our work. They appeal to knot Floer homology instead of bordered Heegaard Floer homology. This leads to somewhat divergent results and emphasis: while Rasmussen and Rasmussen give a clear picture of the interval of L–space slopes in terms of classical invariants, our machinery is better suited to gluing results.

Structure of the paper

Section 2 collects the essentials of bordered Heegaard Floer homology and puts in place our conventions. In particular, the definitions of L–space and strict L–space slope are found here.

Section 3 is devoted to defining loops and loop calculus. This represents the key tool; loop calculus provides a combinatorial framework for studying bordered Heegaard Floer homology. Note that we define and work in an a priori broader setting of abstract

loops. It seems likely that many of the loops considered do not represent the type D structure of any bordered three-manifold. This calculus is applied towards two distinct ends: detection and gluing.

Section 4 gives characterizations of L–space slopes and strict L–space slopes in the loop setting. In particular, we prove Theorem 1.2. The proof of the second part of this result — identifying the set of L–space slopes as the restriction of an interval with rational endpoints — requires a detailed study of non-L–space slopes. While interesting in its own right, this fact is essential to the gluing results that follow.

Section 5 proves Theorem 1.3 by first establishing the appropriate gluing statements for abstract loops. The section concludes with a complete characterization of L–spaces resulting from generalized splices of L–space knots in the three-sphere, proving Theorem 1.6.

Section 6 turns to the study of loop-type manifolds. We describe an algorithm for constructing rational homology sphere graph manifolds by way of three moves, and determine the effect of these moves on bordered invariants. Using this, we establish a class of graph manifold rational homology tori, subsuming Seifert-fibred rational homology tori, for which we now have a complete understanding of gluing in Heegaard Floer homology according to the material in the preceding two sections.

Section 7 collects all of the forgoing material in order to prove our main results. This section includes the proofs of Theorems 1.4 and 1.5, and from these results we prove Theorem 1.1.

Addendum: additional context and further developments

This paper was first posted to arXiv on August 22, 2015. In the intervening years there have been many closely related developments, depending on or growing directly out of this project. In hopes that this aids the present-day reader while maintaining the paper in its original form, we will expand on these here.

As described at the start of the paper, the genesis for this work was an interest in understanding the so-called L–space conjecture in the presence of an essential torus [2; 3]. While a comprehensive list of references detailing the support for this conjecture now seems too long to compile, we point to Dunfield [6] for more recent computational evidence. Theorem 1.1 provided a first step towards establishing the conjecture for graph manifolds. Indeed, the synchronous work of Rasmussen and Rasmussen [27] provided a key result that allowed the "simple loop-type" hypothesis to be replaced by

"Floer-simple"; then the proof of Theorem 1.1, which is the focus of Section 7, provides the scaffold used in our four-author project to establish the L–space conjecture for graph manifolds in full [10]. (Note that S Rasmussen gave a different, independent proof [28].)

One idea that is central to this work, but perhaps somewhat hidden from view, is the following: given a property of a bordered invariant that is unchanged under Dehn twists, one obtains a property associated with a three-manifold. This is used in Section 3, where we establish our class of interest, namely loop-type three-manifolds; see Definition 3.13.

Our strategy, at the beginning of this project, was to ignore the pathology of possible non-loop-type manifolds — based on empirical evidence that such manifolds never arose. The result of this, on the one hand, was the presence of an annoying technical assumption, but, on the other, provided the correct lens through which to view the bordered invariants for a manifold with torus boundary. Indeed, on completing this project and [10], it became clear that bordered invariants, subject to the loop-type condition, could be interpreted as immersed curves in a once-punctured torus. Moreover, gluing could now be stated in terms of Lagrangian intersection Floer homology in a once-punctured torus, giving rise to a gluing theorem that was then best-possible: Theorem 1.3 is restated for Floer-simple manifolds in [10, Theorem 7], and then further upgraded to loop-type manifolds with no simplicity hypothesis at all in the initial arXiv version of [11].

When viewed in these terms, loop calculus can be extended to general bordered invariants, giving rise to a graphical interpretation of the box tensor product in terms of intersections between train tracks — these latter are immersions of the graphs considered in this paper into the torus. Combined with new work of Haiden, Katzarkov and Kontsevich [7], this indicated that in the general setting our curves should carry local systems, which provides a means of handling graphs that cannot be converted to curves. So, once interpreted in the bordered setting, the loop-type hypothesis is simply the restriction to the trivial local system case. This is explained in detail in [12]. Note that it remains unclear what the role local systems play in general is — we are still not aware of an explicit example. (It is worth noting that local systems do play a role in Khovanov homology, which admits an analogous immersed curves description for decompositions along Conway spheres; see Kotelskiy, Watson and Zibrowius [16].) Nevertheless, these are easy to manage in practice, and up to minor modifications proofs go through. In particular, an alternative graphical approach to our work in Sections 4 (compare [11, Section 5]) and 5 (compare [11, Section 6]) now exists.

Despite these advances, however, it remains the case that, given a bordered invariant, one needs an algebrocombinatorial tool in order to work with and/or feed it into a

computer. And, indeed, some operations are still only manageable using the tools described here. This is the case for annular gluing, for example, as this uses the "merge" described in Section 6, which can be adapted to the non-loop-type setting (that is, to handle nontrivial local systems). This is explained in more recent work, and, as an application, we establish a complete description of the effect of cabling on bordered invariants for knots in S^3 in terms of immersed curves in [13]. The proofs depend on the merge operation and loop calculus in an essential way.

As described in Section 3, then, "loop calculus" is meant to describe the tools needed to study bordered invariants for manifolds with torus boundary. In particular, we view this calculus as a toolkit for studying type D structures over the torus algebra in general. With hindsight, this is a viable algebrocombinatorial machinery for working with objects in the Fukaya category of the once-punctured torus. This was our attempt to standardize the seemingly ad hoc techniques that appear in earlier work of the authors; see [3; 9], for example.

Acknowledgements

We are very grateful to Jake and Sarah Rasmussen for their interest in this work and for their willingness to share their own [27]. This strengthened the present work considerably and facilitated the joint paper at the intersection of the two projects [10].

Hanselman was partially supported by NSF RTG grant DMS-1148490. Watson was partially supported by a Marie Curie Career Integration Grant (HFFUNDGRP).

2 Background and conventions

We begin by briefly recalling the essentials of bordered Heegaard Floer homology; for details see [20]. We restrict attention to compact, orientable three-manifolds M with torus boundary. Let $\mathbb{F}$ denote the two-element field.

2.1 Bordered structures

A bordered three-manifold, in this setting, is a pair (M, Φ) where $\Phi\colon S^1 \times S^1 \to \partial M$ is a fixed homeomorphism satisfying $\Phi(S^1 \times \{\mathrm{pt}\}) = \alpha$ and $\Phi(\{\mathrm{pt}\} \times S^1) = \beta$ for slopes α and β in ∂M satisfying $\Delta(\alpha, \beta) = 1$. Recall that a slope in ∂M is the isotopy class of an essential simple closed curve in ∂M or, equivalently, a primitive class in $H_1(\partial M; \mathbb{Z})/\{\pm 1\}$. The distance $\Delta(\cdot, \cdot)$ is measured by considering the minimal

geometric intersection between slopes; thus, the requirement that $\Delta(\alpha,\beta)=1$ ensures that the pair $\{\alpha,\beta\}$, having chosen orientations, forms a basis for $H_1(\partial M;\mathbb{Z})$.

As a result, any bordered manifold (M,Φ) may be represented by the ordered triple (M,α,β), with the understanding that (M,α,β) and (M,β,α) differ as bordered manifolds (that is, these represent different bordered structures on the same underlying manifold M). We will adhere to this convention for describing bordered manifolds as it makes clear that bordered manifolds come with a pair of preferred slopes.

Definition 2.1 For a given bordered manifold (M,α,β), the slope α is referred to as the standard slope and the slope β is referred to as the dual slope.

2.2 The torus algebra

The torus algebra $\mathcal{A}$ is generated (as a vector space over $\mathbb{F}$) by elements

$$\iota_0,\quad \iota_1,\quad \rho_1,\quad \rho_2,\quad \rho_3,\quad \rho_{12},\quad \rho_{23},\quad \rho_{123}$$

with multiplication defined by

$$\rho_1\rho_2=\rho_{12},\quad \rho_2\rho_3=\rho_{23},\quad \rho_1\rho_{23}=\rho_{123}=\rho_{12}\rho_3$$

(all other products $\rho_I\rho_J$ vanish) and

$$\iota_0\rho_1=\rho_1=\rho_1\iota_1,\quad \iota_1\rho_2=\rho_2=\rho_2\iota_0,\quad \iota_0\rho_3=\rho_3=\rho_3\iota_1,$$

so that $\iota_0+\iota_1$ is a unit. Denote by $\mathcal{I}$ the subring of idempotents in $\mathcal{A}$ generated by ι_0 and ι_1. This algebra has various geometric interpretations; see [20]. The bordered Heegaard Floer invariants of (M,α,β) are modules of various types over $\mathcal{A}$, as we now describe.

2.3 Type *D* structures

A (left) type D structure over $\mathcal{A}$ is an $\mathbb{F}$–vector space N equipped with a left action of the idempotent subring $\mathcal{I}$ so that $N=\iota_0N\oplus\iota_1N$, together with a map

$$\delta^1\colon N\to\mathcal{A}\otimes_{\mathcal{I}}N$$

satisfying a compatibility condition with the multiplication on $\mathcal{A}$ [20]. The compatibility ensures that the map

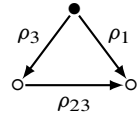
$$\partial\colon \mathcal{A}\otimes_{\mathcal{I}}N\to\mathcal{A}\otimes_{\mathcal{I}}N,\quad a\otimes x\mapsto a\cdot\delta^1(x),$$

promotes $\mathcal{A} \otimes N$ to a left differential module over $\mathcal{A}$ (in particular, $\partial^2 = 0$), where $a \cdot (b \otimes y) = ab \otimes y$. While we will generally confuse type D structures and their associated differential modules, the advantage of the type D structure is in an iterative definition

$$\delta^k \colon N \to \mathcal{A}^{\otimes k} \otimes_{\mathcal{I}} N,$$

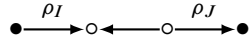
where $\delta^{k+1} = (\mathrm{id}_{\mathcal{A}^{\otimes k}} \otimes \delta^1) \circ \delta^k$ for $k > 1$. The type D structure is bounded if all δ^k are identically zero for sufficiently large k.

Given a basis for N, this structure may be described graphically. An $\mathcal{A}$–decorated graph is a directed graph with vertex set labelled by $\{\bullet, \circ\}$ and edge set labelled by elements of $\mathcal{A}$ consistent with the edge orientations. The labelling of the vertices specifies the splitting of the generating set according to the idempotents, while the edge set encodes the differential. For example, the $\mathcal{A}$–decorated graph

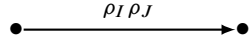


encodes the fact that there is a single generator x in the ι_0–idempotent with $\delta^1(x) = \rho_1 \otimes u + \rho_3 \otimes v$ (or $\partial(x) = \rho_1 u + \rho_3 v$). The higher maps in the type D structure can be extracted from following directed paths in the graph; for example, in the graph above we have $\delta^2(x) = \rho_3 \otimes \rho_{23} \otimes u$. By convention we drop the label on edges labelled by the identity element of $\mathcal{A}$.

An $\mathcal{A}$–decorated graph determines a type D structure with respect to a particular basis. In general we do not have a preferred choice of basis and we care about type D structures only up to homotopy equivalence, so there are many $\mathcal{A}$–decorated graphs which we should deem to be equivalent. Choosing a different basis for N leaves the vertex set unchanged and changes the arrows in a predictable way; for example, the basis change replacing x with $x + y$ adds an arrow out of the vertex labelled by x for each arrow out of the vertex labelled by y, and an arrow into the vertex labelled by y for each arrow into the vertex labelled by x. We can also replace an $\mathcal{A}$–decorated graph with one for a homotopy equivalent type D structure by using edge reduction (or its inverse) as described for example by Levine [17, Section 2.6]. Briefly, any segment of a graph of the form



may be replaced by a single edge



where the edge is simply deleted if the product $\rho_I \rho_J$ vanishes (we have chosen specific vertex labelling for illustration only). Again, we follow the convention that an unlabelled edge represents an edge labelled by the identity element of $\mathcal{A}$. Note that, by repeatedly cancelling such edges, we may always find a graph with no unlabelled edges.

Definition 2.2 An $\mathcal{A}$–decorated graph is reduced if no edge is labelled by the identity element of $\mathcal{A}$.

It turns out that any two $\mathcal{A}$–decorated graphs for homotopy equivalent type D structures can be related by a sequence of basis changes and edge reductions or insertions.

Given a bordered three-manifold (M, α, β), Lipshitz, Ozsváth and Thurston define a type D structure $\widehat{CFD}(M, \alpha, \beta)$ over $\mathcal{A}$ that is an invariant of the bordered manifold up to quasi-isomorphism [20]. As explained above, we will sometimes regard this object as a differential module over $\mathcal{A}$. This invariant splits over spin^c structures of M; that is,

$$\widehat{CFD}(M, \alpha, \beta) = \bigoplus_{\mathfrak{s} \in \mathrm{Spin}^c(M)} \widehat{CFD}(M, \alpha, \beta; \mathfrak{s}).$$

Note that $\mathrm{Spin}^c(M)$ can be identified with $H^2(M) \simeq H_1(M, \partial M)$.

2.4 Type A structures

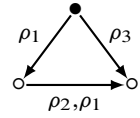
A (right) type A structure over $\mathcal{A}$ is an $\mathbb{F}$–vector space M equipped with a right action of the idempotent subring $\mathcal{I}$ so that $M = M\iota_0 \oplus M\iota_1$, together with maps

$$m_{k-1} \colon M \otimes_{\mathcal{I}} \mathcal{A}^{\otimes k} \to M$$

satisfying the $\mathcal{A}_\infty$ relations; see [20]. That is, a type A structure is a right $\mathcal{A}_\infty$–module over $\mathcal{A}$. A type A structure is bounded if the m_k vanish for all sufficiently large k. Given a bordered three-manifold (M, α, β), Lipshitz, Ozsváth and Thurston define a type A structure $\widehat{CFA}(M, \alpha, \beta)$ over $\mathcal{A}$ that is an invariant of the bordered manifold up to quasi-isomorphism.

There is a similar graphical representation for type A structures. Indeed, owing to a duality between type D and type A structures for three-manifolds, $\widehat{CFA}(M, \alpha, \beta)$ may be deduced from the graph describing $\widehat{CFD}(M, \alpha, \beta)$ by appealing to an algorithm described by Hedden and Levine [14]. This algorithm takes subscripts $1 \mapsto 3$ and $3 \mapsto 1$ while fixing 2, with the convention that a conversion of the form $23 \mapsto 21$ must be parsed as $2, 1$ (this example is shown below, as it occurs in the conversion of the sample graph shown previously). In this type A context, sequences of directed edges

must be concatenated in order to obtain all of the multiplication maps. For example, labelling the generators as before, the graph



encodes operations $m_2(x,\rho_3)=v$, $m_2(x,\rho_1)=u$ and $m_3(u,\rho_2,\rho_1)=v$, as well as $m_3(x,\rho_{12},\rho_2)=v$.

2.5 Pairing

Consider a closed, orientable three-manifold Y decomposed along a (possibly essential) torus so that $Y=M_1\cup_h M_2$ for some homeomorphism $h\colon \partial M_1\to\partial M_2$. If h has the property that $h(\alpha_1)=\beta_2$ and $h(\beta_1)=\alpha_2$, then we will write this decomposition as $Y=(M_1,\alpha_1,\beta_1)\cup(M_2,\alpha_2,\beta_2)$. The reason for this convention is to ensure compatibility with the paring theorem established by Lipshitz, Ozsváth and Thurston [20]. In particular, they prove that

$$\widehat{CF}(Y)\cong\widehat{CFA}(M_1,\alpha_1,\beta_1)\boxtimes\widehat{CFD}(M_2,\alpha_2,\beta_2),$$

where $\widehat{CF}(Y)$ is a chain complex with homology $\widehat{HF}(Y)$. As a vector space over $\mathbb{F}$, this chain complex is generated by tensors (over $\mathcal{I}\in\mathcal{A}$) of the form $x\otimes_{\mathcal{I}} y$, where $x\in\widehat{CFA}(M_1,\alpha_1,\beta_1)$ and $y\in\widehat{CFD}(M_2,\alpha_2,\beta_2)$, with differential

$$\partial^{\boxtimes}(x\otimes_{\mathcal{I}} y)=\sum_{k=0}^{\infty}(m_{k+1}\otimes\mathrm{id})(x\otimes_{\mathcal{I}}\delta^k(y)),$$

which is a finite sum provided at least one of the modules in the pairing is bounded.

As a particular special case, consider the pairing theorem in the context of Dehn filling. Given a three-manifold M with torus boundary, write $M(\alpha)$ for the result of Dehn filling M along the slope α; that is, $M(\alpha)=(D^2\times S^1)\cup_h M$, where the homeomorphism h is determined by $h(\partial D^2\times\{\mathrm{pt}\})=\alpha$. In particular,

$$\widehat{CF}(M(\alpha))\cong\widehat{CFA}(D^2\times S^1,l,m)\boxtimes\widehat{CFD}(M,\alpha,\beta),$$

where $m=\partial D^2\times\{\mathrm{pt}\}$. In this context, any choice of slopes l dual to m and β dual to α will do, since the family $(D^2\times S^1,l+nm,m)$ are all homeomorphic as bordered manifolds. This is due to the Alexander trick; the Dehn twist along m in $\partial(D^2\times S^1)$ extends to a homeomorphism of $D^2\times S^1$.

Since every bordered manifold is equipped with a preferred choice of slopes, it will be important to distinguish between the two Dehn fillings along these slopes.

Definition 2.3 Let (M,α,β) be a bordered three-manifold. The standard filling is the Dehn filling

$$M(\alpha) = (D^2\times S^1, l, m)\cup(M,\alpha,\beta)$$

(that is, $m\mapsto\alpha$) and the dual filling is the Dehn filling

$$M(\beta) = (D^2\times S^1, m, l)\cup(M,\alpha,\beta)$$

(that is, $m\mapsto\beta$).

It is easy to compute $\widehat{HF}$ of these two fillings from $\widehat{CFD}(M,\alpha,\beta)$. There is a representative of $\widehat{CFA}(D^2\times S^1,\ell,m)$ which has a single generator x with idempotent ι_0 and operations of the form $m_{3+i}(x,\rho_3,\rho_{23},\dots,\rho_{23},\rho_2)=x$, where there are $i\geq 0$ copies of ρ_{23}. It follows that $\widehat{CF}(M(\alpha))$ is generated by the ι_0–generators of $\widehat{CFD}(M,\alpha,\beta)$ with a differential coming from any chain of the form

Similarly, $\widehat{CFA}(D^2\times S^1,m,\ell)$ has a representative with a single generator y with idempotent ι_1 and operations of the form $m_{3+i}(y,\rho_2,\rho_{12},\dots,\rho_{12},\rho_1)=y$, where there are $i\geq 0$ copies of ρ_{12}. It follows that $\widehat{CF}(M(\alpha))$ is generated by the ι_1–generators of $\widehat{CFD}(M,\alpha,\beta)$ with a differential coming from any chain of the form

More generally, given (M,α,β), we would like to compute $\widehat{HF}(M(\gamma))$ for any slope γ expressed in terms of α and β. In particular, we will always make a choice of orientations so that $\alpha\cdot\beta=+1$, resulting in slopes of the form $\gamma=\pm(p\alpha+q\beta)\in H_1(\partial M;\mathbb{Z})/\{\pm1\}$. As is familiar, the fixed choice $\{\alpha,\beta\}$ gives rise to an identification of the set of slopes and the extended rational numbers $\widehat{\mathbb{Q}}:=\mathbb{Q}\cup\{\frac{1}{0}\}$. Our convention is that the slope $p\alpha+q\beta$ is identified with $p/q\in\widehat{\mathbb{Q}}$. We will return to a detailed description of the pairing theorem for an arbitrary Dehn filling in the next section, since the following definition will be of central importance:

Definition 2.4 An L–space is a rational homology sphere Y for which $\dim\widehat{HF}(Y)=|H_1(Y;\mathbb{Z})|$. An L–space slope is a slope γ in ∂M for which the result of Dehn filling $M(\gamma)$ is an L–space. For any M with torus boundary, let $\mathcal{L}_M$ denote the set of L–space slopes in ∂M.

We will need to distinguish certain L–space slopes. To do this, consider the natural inclusion

$$\widehat{\mathbb{Q}} \hookrightarrow \widehat{\mathbb{R}} = \mathbb{R} \cup \left\{\tfrac{1}{0}\right\}$$

arising from orienting the basis slopes so that $\alpha \cdot \beta = +1$, and endow $\widehat{\mathbb{Q}}$ with the subspace topology. With this identification, $\mathcal{L}_M \subset \widehat{\mathbb{Q}}$.

Definition 2.5 The set of strict L–space slopes, denoted by $\mathcal{L}_M^\circ$, is the interior of the subset $\mathcal{L}_M$.

Recall that, if $a, b \in \widehat{\mathbb{Q}}$, then the subsets $(a,b) \cap \widehat{\mathbb{Q}}$ and $[a,b] \cap \widehat{\mathbb{Q}}$ are open and closed, respectively, in $\widehat{\mathbb{Q}}$. By abuse, we will write simply (a,b) and $[a,b]$ with the understanding that these describe subsets of $\widehat{\mathbb{Q}}$.

A key example to consider is that of the exterior of a nontrivial knot K in S^3, with $M = S^3 \smallsetminus \nu K$. In this case it is well known that, if M admits a nontrivial L–space filling, then $\mathcal{L}_M$ is either $[2g-1, \infty]$ or $[\infty, 1-2g]$ relative to the preferred basis consisting of the knot meridian μ (corresponding to $1/0$) and the Seifert longitude λ (corresponding to 0), where g denotes the Seifert genus of K. Notice that μ and $(2g-1)\mu + \lambda$ are nonstrict L–space slopes by definition. On the other hand, if K is the trivial knot then $M \cong D^2 \times S^1$ and $\mathcal{L}_M = \mathcal{L}_M^\circ = \widehat{\mathbb{Q}} \smallsetminus \{0\}$ since these are precisely the fillings that give lens spaces.

Every bordered manifold comes with a preferred identification of the set of slopes with $\widehat{\mathbb{Q}}$; in particular, the notation $p/q \in \mathcal{L}(M, \alpha, \beta)$ should be understood to mean the slope $\pm(p\alpha + q\beta) \in \mathcal{L}_M$. We will adhere to this convention, and use the two interchangeably where there is no potential for confusion.

2.6 Gluing via change of framing

In the interest of determining the set $\mathcal{L}_M$ we will need a means of describing any slope γ in ∂M in terms of a fixed basis of slopes $\{\alpha, \beta\}$. Suppose (M, α, β) is given; we would like to calculate $\widehat{HF}(M(p\alpha + q\beta))$. Then, according to the pairing theorem,

$$\widehat{CF}(M(p\alpha + q\beta)) \cong \widehat{CFA}(D^2 \times S^1, m, l) \boxtimes \widehat{CFD}(M, r\alpha + s\beta, p\alpha + q\beta),$$

where $\left(\begin{smallmatrix} q & p \\ s & r \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. Notice that $p = 0$ recovers a chain complex for the dual filling. Fixing a basis so that $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ represents the standard slope and $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ represents the dual slope, notice that the Dehn twist along α carrying $\beta \mapsto \alpha + \beta$ is encoded by the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Call this the standard Dehn twist and denote it by T_{st}. Similarly, the Dehn twist along β carrying $\alpha \mapsto \alpha + \beta$ is encoded by the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$. Call this the (negative)

dual Dehn twist and denote it by T_{du}^{-1}. Associated with each Dehn twist is a mapping cylinder and to this Lipshitz, Ozsváth and Thurston assign a bimodule. Bimodules are defined similarly to type D and type A structures, except that there are two separate actions which may each be either type D or type A, so that bimodules can have type DD, DA, AD or AA; see [19]. For the Dehn twists $T_{\mathrm{st}}^{\pm 1}$ and $T_{\mathrm{du}}^{\pm 1}$, we use $\widehat{T}_{\mathrm{st}}^{\pm 1}$ and $\widehat{T}_{\mathrm{du}}^{\pm 1}$ to denote the corresponding type DA bimodules.[1]

We can take box tensor products of bimodules or of a module and a bimodule by pairing one type D action with one type A action. The convention is that type D actions are always left actions while type A actions are always right actions, and since the sidedness of the actions is unambiguous we will use the ordering of modules in a product to indicate which actions are involved. That is, for (bi)modules N and M we will understand $N \boxtimes M$ to mean the result of pairing the last action on N with the first action on M, even if the former is a left action and the latter is a right action. In practice this will be clear from the context.

Given an odd-length continued fraction expansion $p/q = [a_1, a_2, \dots, a_n]$, we obtain a decomposition according to Dehn twists,

$$\begin{pmatrix} q & p \\ s & r \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{a_n} \cdots \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{a_2} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{a_1}.$$

The bordered manifold $(M, r\alpha + s\beta, p\alpha + q\beta)$ is obtained from (M, α, β) by attaching the mapping cylinder of a homeomorphism h with representative $h_* = \left(\begin{smallmatrix} q & p \\ s & r \end{smallmatrix}\right)$, and so

$$\widehat{CFD}(M, r\alpha + s\beta, p\alpha + q\beta) \cong \widehat{T}_{\mathrm{st}}^{a_n} \boxtimes \cdots \boxtimes \widehat{T}_{\mathrm{du}}^{-a_2} \boxtimes \widehat{T}_{\mathrm{st}}^{a_1} \boxtimes \widehat{CFD}(M, \alpha, \beta),$$

where $\widehat{T}_{\mathrm{st}}^{n} = \underbrace{\widehat{T}_{\mathrm{st}} \boxtimes \cdots \boxtimes \widehat{T}_{\mathrm{st}}}_{n}$.

More generally, given bordered manifolds (M_1, α_1, β_1) and (M_2, α_2, β_2), we can calculate

$$\widehat{CFA}(M, \alpha_1, \beta_1) \boxtimes \widehat{CFD}(M, r\alpha_2 + s\beta_2, p\alpha_2 + q\beta_2)$$

by considering a homeomorphism h as above. This gives a complex computing $\widehat{HF}(Y)$, where

$$Y \cong M_1 \cup_h M_2 \cong (M_1, \alpha_1, \beta_1) \cup (M_2, r\alpha_2 + s\beta_2, p\alpha_2 + q\beta_2)$$

and $h\colon \partial M_1 \to \partial M_2$ is specified by

$$\alpha_1 \mapsto r\alpha_2 + s\beta_2, \quad \beta_1 \mapsto p\alpha_2 + q\beta_2.$$

[1] In the notation of [19, Section 10], we have $\widehat{T}_{\mathrm{st}} = \widehat{CFDA}(\tau_m)$ and $\widehat{T}_{\mathrm{du}} = \widehat{CFDA}(\tau_\ell^{-1})$.

With these gluing conventions in place, we make an observation that will be of use in the sequel:

Proposition 2.6 *Given type D modules N_1 and N_2, let $N_1' = \widehat{T}_{\mathrm{st}}^n \boxtimes N_1$ and $N_2' = \widehat{T}_{\mathrm{du}}^{-n} \boxtimes N_2$ for some $n \in \mathbb{Z}$. There is a homotopy equivalence*

$$N_1 \boxtimes \widehat{CFAA}(\mathbb{I}) \boxtimes N_2 \cong N_1' \boxtimes \widehat{CFAA}(\mathbb{I}) \boxtimes N_2',$$

where $\widehat{CFAA}(\mathbb{I})$ is the type AA identity bimodule.

Remark 2.7 The type AA identity bimodule is defined in [19] and, in particular, gives rise to $\widehat{CFA}(M,\alpha,\beta) \cong \widehat{CFAA}(\mathbb{I}) \boxtimes \widehat{CFD}(M,\alpha,\beta)$. This observation leads to the algorithm described in Section 2.4.

Proof of Proposition 2.6 The right-hand side of the claimed equivalence can be written as

$$(\widehat{T}_{\mathrm{st}}^n \boxtimes N_1) \boxtimes \widehat{CFAA}(\mathbb{I}) \boxtimes \widehat{T}_{\mathrm{du}}^{-n} \boxtimes N_2.$$

The DA–bimodule $\widehat{T}_{\mathrm{st}}$ is homotopy equivalent to the AD–bimodule

$$\widehat{CFAA}(\mathbb{I}) \boxtimes \widehat{T}_{\mathrm{du}} \boxtimes \widehat{CFDD}(\mathbb{I}).$$

To see this, note that the Heegaard diagram for $T_{\mathrm{st}} = \tau_m$ in [19, Figure 25] can be obtained from the Heegaard diagram for $T_{\mathrm{du}} = \tau_l$ by rotating 180 degrees. Thus the right side above simplifies to

$$\begin{aligned} N_1 \boxtimes (\widehat{CFAA}(\mathbb{I}) \boxtimes \widehat{T}_{\mathrm{du}} \boxtimes \widehat{CFDD}(\mathbb{I}))^n \boxtimes \widehat{CFAA}(\mathbb{I}) \boxtimes \widehat{T}_{\mathrm{du}}^{-n} \boxtimes N_2 & \\ \cong N_1 \boxtimes \widehat{CFAA}(\mathbb{I}) \boxtimes \widehat{T}_{\mathrm{du}}^n \boxtimes \widehat{T}_{\mathrm{du}}^{-n} \boxtimes N_2 & \\ \cong N_1 \boxtimes \widehat{CFAA}(\mathbb{I}) \boxtimes N_2. & \end{aligned}$$

□

Remark 2.8 If $N_1 = \widehat{CFD}(M_1,\alpha_1,\beta_1)$ and $N_2 = \widehat{CFD}(M_2,\alpha_2,\beta_2)$, the homotopy equivalence above corresponds to the fact that, with the gluing conventions described above,

$$(M_1,\alpha_1,\beta_1) \cup (M_2,\alpha_2,\beta_2) \cong (M_1,\alpha_1,\beta_1 + n\alpha_1) \cup (M_2,\alpha_2 + n\beta_2,\beta_2)$$

for any $n \in \mathbb{Z}$.

2.7 Grading

We conclude this discussion with a description of the relative $(\mathbb{Z}/2\mathbb{Z})$–grading on the bordered invariants, summarizing the discussion in [9, Section 2.2]. For more details and developments, see [15; 26]

The torus algebra $\mathcal{A}$ may be promoted to a graded algebra by defining $\mathrm{gr}(\rho_1) = \mathrm{gr}(\rho_3) = 0$ and $\mathrm{gr}(\rho_2) = 1$, and extending according to $\mathrm{gr}(\rho_I \rho_J) \equiv \mathrm{gr}(\rho_I) + \mathrm{gr}(\rho_J)$ reduced modulo 2 (we will always drop this reduction from the notation). In particular, the grading is 1 on all the remaining nonzero products ρ_{12}, ρ_{23} and ρ_{123}.

The relative $(\mathbb{Z}/2\mathbb{Z})$–grading on elements x of a type D structure is determined by

$$\mathrm{gr}(\rho_I \cdot x) \equiv \mathrm{gr}(\rho_I) + \mathrm{gr}(x) \quad \text{and} \quad \mathrm{gr}(\partial x) \equiv \mathrm{gr}(x) + 1.$$

In particular, given a connected $\mathcal{A}$–decorated graph this relative grading is determined by choosing the grading on a given vertex, and then noting that only edges labelled by ρ_1, ρ_3 or the identity element of $\mathcal{A}$ alter the grading.

The relative $(\mathbb{Z}/2\mathbb{Z})$–grading on elements x of a type A structure is determined by

$$\mathrm{gr}(m_{k+1}(x, \rho_{I_1}, \dots, \rho_{I_k})) - k - 1 \equiv \mathrm{gr}(x) + \sum_{j=1}^{k} \mathrm{gr}(\rho_{I_j}).$$

Notice that, given a type D structure with a choice of relative grading, a relative grading on the associated type A structure is obtained by switching the grading of each generator with idempotent ι_0.

If $Y \cong (M_1, \alpha_1, \beta_2) \cup (M_2, \alpha_2, \beta_2)$, choices of relative $(\mathbb{Z}/2\mathbb{Z})$–gradings on each of the objects $\widehat{CFA}(M_1, \alpha_1, \beta_2)$ and $\widehat{CFD}(M_2, \alpha_2, \beta_2)$ give rise to a relative grading on $\widehat{CF}(Y)$ via the pairing theorem and the rule $\mathrm{gr}(x \otimes y) = \mathrm{gr}(x) + \mathrm{gr}(y)$. This agrees with the usual relative $(\mathbb{Z}/2\mathbb{Z})$–grading on $\widehat{CF}(Y)$, so that, in particular, $|\chi(\widehat{CF}(Y))| \geq |H_1(Y; \mathbb{Z})|$ for any rational homology sphere Y. At the level of homology, it follows immediately that Y is an L–space if and only if every generator of $\widehat{HF}(Y)$ has the same grading. More generally, the rule $\mathrm{gr}(x \otimes y) = \mathrm{gr}(x) + \mathrm{gr}(y)$ determines the relative $(\mathbb{Z}/2\mathbb{Z})$–grading on tensor products involving bimodules.

Where required, we will make use of an additional marking on the vertices of an $\mathcal{A}$–decorated graph by $\{\bullet^+, \bullet^-, \circ^+, \circ^-\}$ to indicate a choice of relative $(\mathbb{Z}/2\mathbb{Z})$–grading on the underlying differential module. In addition, given a type D structure N with a choice of relative $(\mathbb{Z}/2\mathbb{Z})$–grading, we define the idempotent Euler characteristics

$$\chi_\bullet(N) = \chi(\iota_0 N) \quad \text{and} \quad \chi_\circ(N) = \chi(\iota_1 N).$$

Note that $\chi_\bullet$ and $\chi_\circ$ are well defined on homotopy classes of type D structures, since changing the basis of N leaves the vertex set of the corresponding graph unchanged and edge reduction removes a pair of $\bullet$–vertices or a pair of $\circ$–generators with opposite gradings.

The following lemma records how $\chi_\bullet$ and $\chi_\circ$ change under reparametrization of the boundary:

Lemma 2.9 *Let* $\left(\begin{smallmatrix} q & p \\ s & r \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. *Then*

$$\pm\chi_\bullet\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta) = r\chi_\bullet\widehat{CFD}(M,\alpha,\beta)+s\chi_\circ\widehat{CFD}(M,\alpha,\beta),$$
$$\pm\chi_\circ\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta) = p\chi_\bullet\widehat{CFD}(M,\alpha,\beta)+q\chi_\circ\widehat{CFD}(M,\alpha,\beta).$$

Here the $\pm$ *choice depends on the choice of absolute* $(\mathbb{Z}/2\mathbb{Z})$*–grading on* $\widehat{CFD}(M,\alpha,\beta)$ *and* $\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta)$.

Proof This is true when $r=q=1$ and $s=p=0$. Assuming the claim is true for given r, s, p and q, we check that it is true when p and q are replaced by $p+r$ and $q+s$ by examining the effect of the standard Dehn twist. The relevant bimodule $\widehat{T}_{\mathrm{st}}$ is pictured in Figure 2. It has three generators: one pairs with ι_0–generators to produce ι_0–generators, one pairs with ι_1–generators to produce ι_1–generators, and one pairs with ι_0–generators to produce ι_1–generators. All three generators have the same relative $(\mathbb{Z}/2\mathbb{Z})$–grading; we will assume they have grading 0. It follows that

$$\begin{aligned}\chi_\bullet\widehat{CFD}(M, r\alpha+s\beta, (p+r)\alpha+(q+s)\beta) &= \chi_\bullet(\widehat{T}_{\mathrm{st}}\boxtimes\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta))\\ &= \chi_\bullet\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta)\\ &= r\chi_\bullet\widehat{CFD}(M,\alpha,\beta)+s\chi_\circ\widehat{CFD}(M,\alpha,\beta),\end{aligned}$$

$$\begin{aligned}&\chi_\circ\widehat{CFD}(M, r\alpha+s\beta, (p+r)\alpha+(q+s)\beta)\\ &\quad= \chi_\circ(\widehat{T}_{\mathrm{st}}\boxtimes\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta))\\ &\quad= \chi_\circ\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta)+\chi_\bullet\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta)\\ &\quad= (p+r)\chi_\bullet\widehat{CFD}(M,\alpha,\beta)+(q+s)\chi_\circ\widehat{CFD}(M,\alpha,\beta).\end{aligned}$$

Similarly, we check the claim for when r and s are replaced with $r-p$ and $s-q$ by considering the dual Dehn twist bimodule $\widehat{T}_{\mathrm{du}}$ (pictured in Figure 4). This bimodule has three generators with relative $(\mathbb{Z}/2\mathbb{Z})$–gradings 0, 0 and 1: the first pairs with ι_0–generators to produce ι_0–generators, the second pairs with ι_1–generators to produce ι_1–generators, and the third pairs with ι_1–generators to produce ι_0–generators. It follows that

$$\begin{aligned}&\chi_\bullet\widehat{CFD}(M, (r-p)\alpha+(s-q)\beta, p\alpha+q\beta)\\ &\quad= \chi_\bullet(\widehat{T}_{\mathrm{du}}\boxtimes\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta))\\ &\quad= \chi_\bullet\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta)-\chi_\circ\widehat{CFD}(M, r\alpha+s\beta, p\alpha+q\beta)\\ &\quad= (r-p)\chi_\bullet\widehat{CFD}(M,\alpha,\beta)+(s-q)\chi_\circ\widehat{CFD}(M,\alpha,\beta),\end{aligned}$$

$$\begin{aligned}\chi_\circ\widehat{CFD}(M,(r-p)\alpha+(s-q)\beta,p\alpha+q\beta) &= \chi_\circ(\widehat{T}_{\mathrm{du}}\boxtimes\widehat{CFD}(M,r\alpha+s\beta,p\alpha+q\beta))\\ &= \chi_\circ\widehat{CFD}(M,r\alpha+s\beta,p\alpha+q\beta)\\ &= p\chi_\circ\widehat{CFD}(M,\alpha,\beta)+q\chi_\circ\widehat{CFD}(M,\alpha,\beta).\end{aligned}$$

The argument extends easily to multiples and inverses of Dehn twists. The claim follows by decomposing an arbitrary element of $\mathrm{SL}_2(\mathbb{Z})$ into Dehn twists. □

For rational homology spheres, we will be interested in a distinguished slope, referred to as the *rational longitude*. Given a bordered structure (α,β) for M, this slope $\gamma = p\alpha+q\beta$ is characterized by the property that some collection of like-oriented copies of γ bounds a surface in M. This slope is recorded by the fraction p/q, and the $(\mathbb{Z}/2\mathbb{Z})$–grading allows us to recover this data from $\widehat{CFD}(M,\alpha,\beta)$ as follows:

Proposition 2.10 *Fix a spin^c–structure $\mathfrak{s}$ on a bordered manifold (M,α,β), and let $\chi_\bullet$ and $\chi_\circ$ denote the Euler characteristics of $\widehat{CFD}(M,\alpha,\beta;\mathfrak{s})$ in the appropriate idempotent. If $\chi_\bullet=\chi_\circ=0$, then M is not a rational homology solid torus. Otherwise, M is a rational homology solid torus, the nullhomologous curves in ∂M are the multiples of $\chi_\circ\alpha-\chi_\bullet\beta$, and, in particular, the rational longitude is $-\chi_\circ/\chi_\bullet$.*

Proof Let $\mathfrak{s}'$ be a spin^c–structure on $M(p\alpha+q\beta)$ that restricts to $\mathfrak{s}$ on M. The Euler characteristic $\chi\widehat{HF}(M(p\alpha+q\beta);\mathfrak{s}')$ is nonzero if and only if $M(p\alpha+q\beta)$ is a rational homology sphere [24, Proposition 5.1]. The closed manifold $M(p\alpha+q\beta)$ is the dual filling of $(M,r\alpha+s\beta,p\alpha+q\beta)$ for any r and s with $rq-ps=1$. Thus, up to sign, $\chi\widehat{HF}(M(p\alpha+q\beta);\mathfrak{s}')$ is

$$\chi_\circ\widehat{CFD}(M,r\alpha+s\beta,p\alpha+q\beta;\mathfrak{s}') = p\chi_\bullet+q\chi_\circ$$

by applying Lemma 2.9. Note that if $\chi_\bullet$ and $\chi_\circ$ are both zero, then $M(p\alpha+q\beta)$ is not a rational homology sphere for any p/q and hence M is not a rational homology solid torus. Otherwise, $M(p\alpha+q\beta)$ is a rational homology sphere unless $p/q=-\chi_\circ/\chi_\bullet\in\widehat{\mathbb{Q}}$. It follows that when M is a rational homology solid torus the slope of the rational longitude is $-\chi_\circ/\chi_\bullet$.

For the remainder of the proof, let $p/q=-\chi_\circ/\chi_\bullet$, where p and q are relatively prime. The multicurve $np\alpha+nq\beta$ is nullhomologous in M for some $n>0$. To determine the smallest such n consider the standard filling of $(M,r\alpha+s\beta,p\alpha+q\beta)$ for some r and s with $rq-ps=1$. This manifold is obtained from M by adding a solid torus identifying the longitude l with $p\alpha+q\beta$; the resulting first homology group is a direct sum of $H_1(M,\partial M)$ and $H_1(D^2\times S^1)$ (which is generated by the longitude l) with the

relation $nl = np\alpha + nq\beta = 0$. Since $H_1(M, \partial M)$ indexes the spin^c–structures on M, for each spin^c structure $\mathfrak{s}$ on M, there are n elements of $H_1(M(r\alpha + s\beta))$. It follows that $\chi\widehat{HF}(M(r\alpha + s\beta); \mathfrak{s}') = n$ for each $\mathfrak{s}'$. By Lemma 2.9, we have

$$n = \chi_\bullet \widehat{CFD}(M, r\alpha + s\beta, p\alpha + q\beta; \mathfrak{s}') = r\chi_\bullet + s\chi_\circ,$$

which may be combined with the fact that $n = nqr - nps$. So, up to sign, $(\chi_\bullet, \chi_\circ) = (nq, -np)$ and hence the minimal nullhomologous multicurve in ∂M is $\chi_\circ\alpha - \chi_\bullet\beta$, as claimed. □

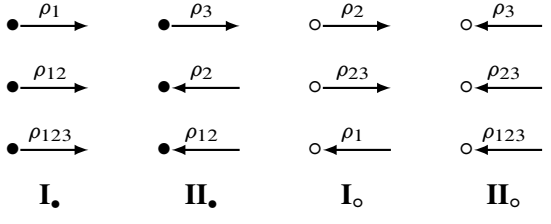
Remark 2.11 The minimal integer n appearing in this proof, multiplied by the number of spin^c structures, corresponds to the constant c_M associated with M described in [29, Section 3.1].

3 Loop calculus

A focus of this work is the development of a calculus for studying bordered invariants. This will be achieved by restricting to a class of manifolds whose bordered invariants can be represented by certain valence two $\mathcal{A}$–decorated graphs, which we will call loops.

3.1 Loops and loop-type manifolds

Towards defining loops, consider the following arrows which may appear in an $\mathcal{A}$–decorated graph:



We will be interested in valence two $\mathcal{A}$–decorated graphs subject to the following restriction:

($\star$) Each ι_0–vertex is adjacent to an edge of type $\mathbf{I}_\bullet$ and an edge of type $\mathbf{II}_\bullet$, and each ι_1–vertex is adjacent to an edge of type $\mathbf{I}_\circ$ and an edge of type $\mathbf{II}_\circ$.

Definition 3.1 A loop is a connected valence two $\mathcal{A}$–decorated graph satisfying ($\star$).

Since any loop describes a differential module over $\mathcal{A}$, a loop may be promoted to a graded loop via the $(\mathbb{Z}/2\mathbb{Z})$–grading described in Section 2.7. In particular, where needed, the vertex set will be extended to $\{\bullet^+, \bullet^-, \circ^+, \circ^-\}$.

As abstract combinatorial objects loops provide a tractable structure to work with; this section develops a calculus for doing so. The results derived from this calculus apply to the following class of bordered invariants:

Definition 3.2 The bordered invariant $\widehat{CFD}(M,\alpha,\beta)$ is said to be of loop-type if, up to homotopy, it may be represented by a collection of loops, that is, by a (possibly disconnected) $\mathcal{A}$–decorated valence two graph satisfying $(\star)$. For simplicity, in this paper we will make the additional assumption that the number of connected components of the valence two graph describing $\widehat{CFD}(M,\alpha,\beta)$ coincides with $|\mathrm{Spin}^c(M)|$.

We will refer to the bordered manifold (M,α,β) being of loop-type when the associated bordered invariant has this property. Some motivation for this definition is provided by the following:

Proposition 3.3 *Let $\mathcal{H}$ be a bordered Heegaard diagram describing (M,α,β) and suppose $\widehat{CFD}(M,\alpha,\beta)=\widehat{CFD}(\mathcal{H})$ is reduced and represented by a valence two $\mathcal{A}$–decorated graph having a single connected component per spin^c–structure. Then $\widehat{CFD}(M,\alpha,\beta)$ is loop-type.*

Proof The hypothesis $\widehat{CFD}(M,\alpha,\beta)=\widehat{CFD}(\mathcal{H})$ allows us to use the notion of generalized coefficient maps developed in [20, Section 11.6], which force restrictions on the type D modules that can occur as invariants of manifolds with torus boundary. Briefly, generalized coefficient maps are extra differentials obtained by counting holomorphic curves that run over the basepoint. The torus algebra is extended with additional Reeb chords: ρ_0, ρ_{01}, ρ_{30}, ρ_{012}, ρ_{301}, ρ_{230}, ρ_{0123}, ρ_{3012}, ρ_{2301} and ρ_{1230}. With these additional differentials, $\partial^2=0$ is no longer satisfied. However, we have instead that

$$\partial^2(x)=\rho_{1230}x+\rho_{3012}x \quad \text{for any } x \text{ in } \iota_0\widehat{CFD}(M,\alpha,\beta),$$
$$\partial^2(x)=\rho_{2301}x+\rho_{1230}x \quad \text{for any } x \text{ in } \iota_1\widehat{CFD}(M,\alpha,\beta).$$

Let x be a generator with idempotent ι_0. Since $\partial^2(x)$ contains the term $\rho_{1230}x$, $\partial(x)$ contains the term $\rho_{I_1}y$ and $\partial(y)$ contains the term $\rho_{I_2}x$ for some $y\in\widehat{CFD}(M,\alpha,\beta)$ and some (generalized) Reeb chords ρ_{I_1} and ρ_{I_2} such that $\rho_{I_1}\rho_{I_2}=\rho_{1230}$. Since we assumed that $\widehat{CFD}(M,\alpha,\beta)$ is reduced, I_1 and I_2 are not $\varnothing$. It follows that $I_1\in\{1,12,123\}$ and, in the graphical representation, the vertex corresponding to x has an incident edge of type $\mathbf{I}_\bullet$.

Since $\partial^2(x)$ contains the term $\rho_{3012}x$, $\partial(x)$ contains the term $\rho_{I_1}y$ and $\partial(y)$ contains the term $\rho_{I_2}x$ for some $y\in\widehat{CFD}(M,\alpha,\beta)$ and some (generalized) Reeb chords ρ_{I_1} and ρ_{I_2} such that $\rho_{I_1}\rho_{I_2}=\rho_{3012}$. It follows that either $I_1=3$ or $I_2\in\{2,12\}$ and,

in the graphical representation, the vertex corresponding to x has an incident edge of type $\mathbf{II}_\bullet$. Since any vertex has valence two by assumption, the vertex corresponding to x must have exactly one edge from each of $\mathbf{I}_\bullet$ and $\mathbf{II}_\bullet$.

The argument when x has idempotent ι_1 is similar. Since $\partial^2(x)$ contains the term $\rho_{2301}x$, it follows that the vertex corresponding to x is adjacent an edge of type $\mathbf{I}_\circ$; and since $\partial^2(x)$ contains the term $\rho_{0123}x$, it follows that the corresponding vertex is adjacent to an edge of type $\mathbf{II}_\circ$. □

Remark 3.4 While this argument required a particular choice of Heegaard diagram, it seems likely that this hypothesis is not a necessary. However, for the purposes of this work the more general statement will not be required — we simply restrict to the graphs satisfying $(\star)$ by definition — and leave developing the necessary algebra to future work.

It is important to note that loops in the abstract need not be related to three-manifolds: it is not necessarily true that every loop (or disjoint union of loops) arises as the bordered invariant of some three-manifold with torus boundary. This section concludes with an explicit example for illustration.

3.2 Standard and dual notation

It is natural to decompose a loop into pieces by breaking along vertices corresponding to one of the two idempotents; the constraint $(\star)$ suggests that the pieces resulting from such a decomposition are quite limited. Indeed, breaking along generators with idempotent ι_0, five essentially different chains are possible and these possibilities are listed in Figure 1, left. Since $(\star)$ also puts restrictions on how these segments can be concatenated, each piece is depicted with puzzle-piece ends. For instance, a type a piece can be followed by a type b piece, but not by a type c piece. Any given piece may also appear backwards; we denote this with a *bar*.

Segments of type a and b behave differently from the other segment types in several ways. We will call type a and b segments *stable chains* and type c, d and e segments *unstable chains*. This terminology comes from [20, Theorem 11.26], where the notion of unstable chains was introduced in order to describe a procedure for extracting the bordered invariants from knot Floer homology. Three types of unstable chains were described, which correspond precisely to our type c, d and e segments. Type a and b segments also appeared as the segments coming from horizontal arrows and vertical arrows, respectively, in knot Floer homology. The motivation for this terminology is

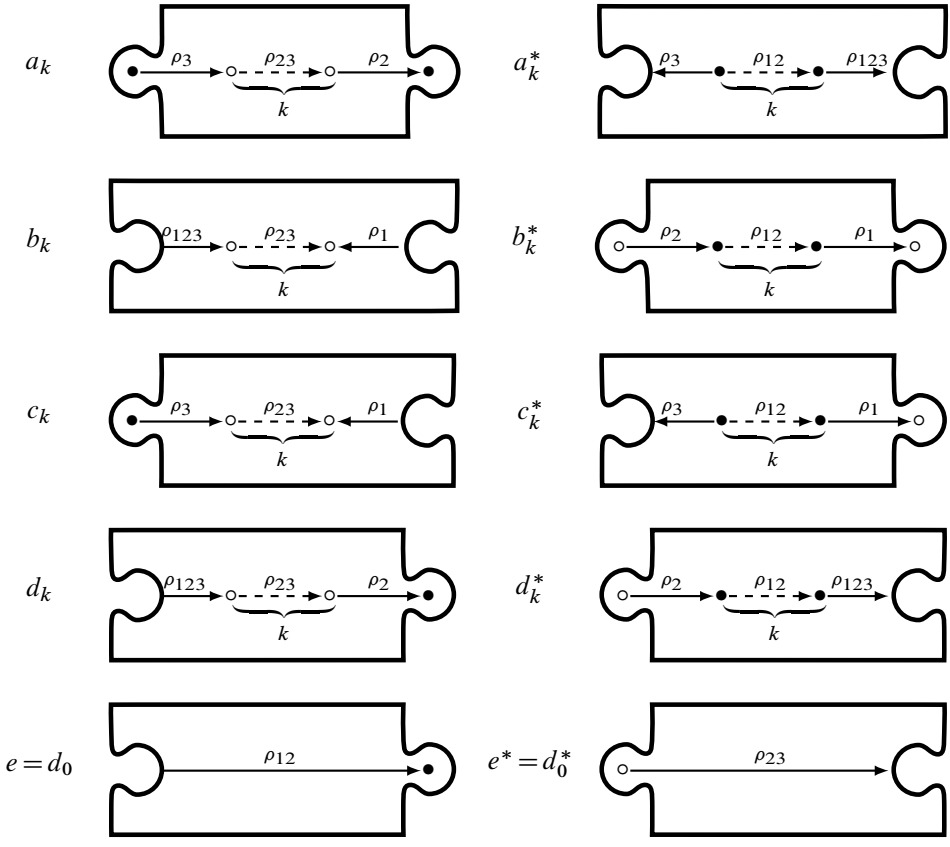


Figure 1: Possible segments: Standard notation (left) is obtained by breaking a loop along ι_0 idempotents (•–vertices) and dual notation (right) is obtained by breaking a loop along ι_1 idempotents (∘–vertices). Note that the integer k records the number of interior vertices, so that the ρ_{23} (for standard notation) or ρ_{12} (for dual notation) appears $k-1$ times.

that, as we will see in Section 3.4, stable chains are preserved by certain Dehn twists while unstable chains are not.

Denote the standard alphabet by

$$\mathfrak{A} = \{a_i, b_i, c_j, d_j\},$$

where $i \in \mathbb{Z} \smallsetminus \{0\}$ and $j \in \mathbb{Z}$. The letters in the standard alphabet correspond directly to the segments depicted in Figure 1, with the relationships

$$a_{-i} = \bar{a}_i \quad \text{and} \quad b_{-i} = \bar{b}_i$$

for $i > 0$ as well as

$$c_{-j} = \bar{d}_j \quad \text{and} \quad d_{-j} = \bar{c}_j$$

for $j \geq 0$ with $d_0 = e = \bar{c}_0$ and $\bar{d}_0 = \bar{e} = c_0$. Throughout this paper we will always assume these shorthand relationships and, by abuse, we will often not distinguish between segments and letters. For instance, the symbols d_0 and e will be used interchangeably and may refer either to a letter in the standard alphabet or the corresponding $\mathcal{A}$–decorated graph segment. As a result of this equivalence, the notation $\bar{s}$ makes sense for any standard letter s, with the understanding that $\bar{\bar{s}} = s$.

We will be interested in the set of cyclic words $W_{\mathfrak{A}}$ on the alphabet $\mathfrak{A}$ that are consistent with the puzzle-piece notation of Figure 1, which encodes the restriction ($\star$). Thus, for example, the cyclic word c_1 is an element of $W_{\mathfrak{A}}$ while the cyclic word a_1 is not. Another immediate restriction is that an element of $W_{\mathfrak{A}}$ must contain an equal number of a–type letters and b–type letters.

Proposition 3.5 *A loop $\boldsymbol{\ell}$ with at least one ι_0–generator may be represented as a cyclic word $\boldsymbol{w_\ell}$ in $W_{\mathfrak{A}}$. Moreover, this representation is unique up to overall reversal of the word $\boldsymbol{w_\ell}$, that is, writing $\boldsymbol{w_\ell}$ with the opposite cyclic ordering and replacing each letter s with $\bar{s}$.*

Proof This is immediate from the definitions. □

As a result, we will be interested in the equivalence class that identifies $\boldsymbol{w_\ell}$ and $\overline{\boldsymbol{w}}_{\boldsymbol{\ell}}$, where $\overline{\boldsymbol{w}}_{\boldsymbol{\ell}}$ denotes the reversal of $\boldsymbol{w_\ell}$. Denote this equivalence class of cyclic words by $(\boldsymbol{w_\ell})$; by abuse we will continue referring to $(\boldsymbol{w_\ell})$ as a cyclic word. By Proposition 3.5, this sets up a one-to-one correspondence $\boldsymbol{\ell} \leftrightarrow (\boldsymbol{w_\ell})$ between loops with at least one ι_0–decorated vertex and (equivalence classes of) cyclic words in $W_{\mathfrak{A}}$.

Definition 3.6 When $\boldsymbol{\ell}$ is represented by a cyclic word $(\boldsymbol{w_\ell})$ using the standard alphabet, we say $\boldsymbol{\ell}$ is written in standard notation.

A loop cannot be written in standard notation if it does not contain an ι_0–decorated vertex. This suggests it will sometimes be useful to break a loop along ι_1–decorated vertices. There are again five types of chains possible, as listed in Figure 1, right. As before, the first two types will be referred to as stable chains and the rest as unstable chains. Denote the dual alphabet by

$$\mathfrak{A}^* = \{a_i^*, b_i^*, c_j^*, d_j^*\}$$

($i \in \mathbb{Z} \smallsetminus \{0\}$ and $j \in \mathbb{Z}$), where, as before, $a^*_{-i} = \bar{a}^*_i$, $b^*_{-i} = \bar{b}^*_i$ (for $i > 0$), $c^*_{-j} = \bar{d}^*_j$, $d^*_{-j} = \bar{c}^*_j$ (for $j \geq 0$) and $d^*_0 = e^* = \bar{c}^*_0$. Proceeding as before, let $W_{\mathfrak{A}^*}$ denote the set of cyclic words in the alphabet $\mathfrak{A}^*$ which are consistent with the puzzle-piece notation.

Proposition 3.7 *A loop $\boldsymbol{\ell}$ with at least one ι_1–generator may be represented as a cyclic word $\boldsymbol{w}^*_{\boldsymbol{\ell}}$ in $W_{\mathfrak{A}^*}$. Moreover, this representation is unique up to overall reversal of the word $\boldsymbol{w}^*_{\boldsymbol{\ell}}$.*

Proof This is immediate from the definitions. □

Again, we denote by $(\boldsymbol{w}^*_{\boldsymbol{\ell}})$ the equivalence class identifying $\boldsymbol{w}^*_{\boldsymbol{\ell}}$ and $\overline{\boldsymbol{w}}^*_{\boldsymbol{\ell}}$, and remark that this gives rise to a second one-to-one correspondence $\boldsymbol{\ell} \leftrightarrow (\boldsymbol{w}^*_{\boldsymbol{\ell}})$ for loops $\boldsymbol{\ell}$ with at least one ι_1–decorated vertex.

Definition 3.8 When $\boldsymbol{\ell}$ is represented by a cyclic word $(\boldsymbol{w}^*_{\boldsymbol{\ell}})$ using the dual alphabet, we say $\boldsymbol{\ell}$ is written in dual notation.

Collecting these observations, notice that, whenever $\boldsymbol{\ell}$ contains instances of both idempotents, the pair of correspondences $(\boldsymbol{w}^*_{\boldsymbol{\ell}}) \leftrightarrow \boldsymbol{\ell} \leftrightarrow (\boldsymbol{w}_{\boldsymbol{\ell}})$ sets up a natural map between the standard representation and the dual representation. In particular, this allows us to set $(\boldsymbol{w}_{\boldsymbol{\ell}})^* = (\boldsymbol{w}^*_{\boldsymbol{\ell}})$ in a well-defined way; we refer to $(\boldsymbol{w}_{\boldsymbol{\ell}})^*$ as the dual of $(\boldsymbol{w}_{\boldsymbol{\ell}})$ and remark that $(\boldsymbol{w}_{\boldsymbol{\ell}})^{**} = (\boldsymbol{w}_{\boldsymbol{\ell}})$.

This description of loops as cyclic words is the essential starting point for loop calculus. A typical loop will have vertices in both idempotents, so it is expressible in both standard and dual notation; switching between the two will be a key part of the loop calculus. We now make this process explicit.

First notice that, given $(\boldsymbol{w}_{\boldsymbol{\ell}})$ representing a loop $\boldsymbol{\ell}$ in standard notation, there is a natural normal form $(\boldsymbol{w}_{\boldsymbol{\ell}}) = (\boldsymbol{u}_1\boldsymbol{v}_1\boldsymbol{u}_2\boldsymbol{v}_2 \dots \boldsymbol{u}_n\boldsymbol{v}_n)$, where

(N1) the subword $\boldsymbol{u}_i$ is a standard letter with subscript $k_i \neq 0$; and

(N2) the subword $\boldsymbol{v}_i$ is $n_i \geq 0$ consecutive copies of either d_0 or c_0 (here n_i may be zero).

This normal form makes sense for any $\boldsymbol{w}_{\boldsymbol{\ell}}$ that does not consist only of $d_0 = e$ or $c_0 = \bar{e}$ letters; we may safely ignore these sporadic examples since, in these cases, $\boldsymbol{\ell}$ has no ι_1–decorated vertices and cannot be expressed in dual notation.

Now the dual word $\boldsymbol{w}^*_{\boldsymbol{\ell}}$ is obtained as follows:

(D1) Replace each $\boldsymbol{v}_i$ by a dual letter with subscript $n_i + 1$ and type[2] determined by (the types of) the ordered pair $\{\boldsymbol{u}_i, \boldsymbol{u}_{i+1}\}$ (written $\boldsymbol{u}_i\boldsymbol{u}_{i+1}$ for brevity) from the subword $\boldsymbol{u}_i\boldsymbol{v}_i\boldsymbol{u}_{i+1}$ according to the rules

$$\begin{aligned} \bar{a}b, \bar{a}d, \bar{c}b, \bar{c}d &\to a^*, & \bar{b}a, \bar{b}c, \bar{d}a, \bar{d}c &\to \bar{a}^*, \\ a\bar{b}, a\bar{c}, d\bar{b}, d\bar{c} &\to b^*, & b\bar{a}, b\bar{d}, c\bar{a}, c\bar{d} &\to \bar{b}^*, \\ \bar{a}\bar{b}, \bar{a}\bar{c}, \bar{c}\bar{b}, \bar{c}\bar{c} &\to c^*, & ba, bc, ca, cc &\to \bar{c}^*, \\ ab, ad, db, dd &\to d^*, & \bar{b}\bar{a}, \bar{b}\bar{d}, \bar{d}\bar{a}, \bar{d}\bar{d} &\to \bar{d}^*. \end{aligned}$$

Note that indices are taken mod n, so $\boldsymbol{u}_{n+1}$ is identified with $\boldsymbol{u}_1$.

(D2) Replace each $\boldsymbol{u}_i$ with the subword consisting of $k_i - 1$ consecutive d_0^* letters if $k_i > 0$ or $1 - k_i$ consecutive c_0^* letters if $k_i < 0$.

To convert from dual notation to standard notation, we use exactly the same procedure (interchanging the words *standard* and *dual* and adding/removing stars where appropriate in the discussion above). Note that letters in dual (resp. standard) notation correspond to consecutive pairs of letters in standard (resp. dual) notation after we ignore letters with subscript 0. We make the following observation about this correspondence:

Observation 3.9 *Stable chains in dual (resp. standard) notation correspond precisely to consecutive pairs of letters in standard (resp. dual) notation, ignoring letters with subscript 0, whose subscripts have opposite signs.*

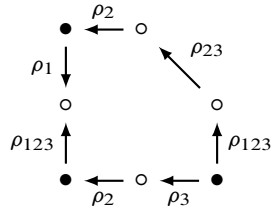
With the forgoing in place we will not distinguish between loops in standard or in dual notation; for a given loop $\boldsymbol{\ell}$ containing (vertices decorated by) both idempotents, it is always possible to choose a representative for $\boldsymbol{\ell}$ in either of the alphabets $\mathfrak{A}$ or $\mathfrak{A}^*$. In summary: We will regard a loop as both a graph-theoretic object and as an equivalence class of words $(\boldsymbol{w}_{\boldsymbol{\ell}})$ and $(\boldsymbol{w}_{\boldsymbol{\ell}}^*)$ modulo dualizing. In particular, we will adopt the abuse of notation that $\boldsymbol{\ell}$ is such an equivalence class of words, where convenient, and think of loops as graph-theoretic objects and word-theoretic objects interchangeably. In particular, we will let the notation $\boldsymbol{\ell}$ stand in for a choice of cyclic word representative (in either alphabet).

We will often refer to a subword $\boldsymbol{w}$ of a given loop $\boldsymbol{\ell}$, so the notation $(\boldsymbol{w})$ indicating the cyclic closure of a word (when it exists) will be used to distinguish subwords from

[2]Recall that a letter of type a, b, c or d with negative subscript can also be written as a letter of type $\bar{a}$, $\bar{b}$, $\bar{d}$ or $\bar{c}$ with positive subscript. Here the *type* of a standard letter (other than d_0 and c_0) refers to the element of $\{a, \bar{a}, b, \bar{b}, c, \bar{c}, d, \bar{d}\}$ corresponding to the representation with positive subscript. Similarly, a dual letter other than d_0^* and c_0^* has a well-defined type in $\{a^*, \bar{a}^*, b^*, \bar{b}^*, c^*, \bar{c}^*, d^*, \bar{d}^*\}$.

(equivalence classes of) cyclic words. The length of a subword is the number of pieces (or letters) in the word.

To conclude with a particular example, let M denote the complement of the left-hand trefoil and consider the bordered manifold (M, μ, λ), where μ is the knot meridian and λ is the Seifert longitude. Following [20, Chapter 11], $\widehat{CFD}(M, \mu, \lambda)$ is described by a loop, as shown in



We may express this in standard notation as $(a_1 b_1 \bar{d}_2)$ or in dual notation as $(a_1^* e^* b_1^* \bar{d}_1^*)$. In particular, according to the discussion above, this pair of words is a representative of the same equivalence specifying the loop $\boldsymbol{\ell}$ shown. That is, we regard $\boldsymbol{\ell} \sim (a_1 b_1 \bar{d}_2) \sim (a_1^* e^* b_1^* \bar{d}_1^*)$. More generally, the homotopy type of the invariant $\widehat{CFD}(M, \mu, \lambda - (2+n)\mu)$ is represented by the loop $(a_1 b_1 c_n)$ for $n \in \mathbb{Z}$, following the convention that $c_0 = \bar{e}$ and $c_n = \bar{d}_{-n}$ when $n < 0$.

3.3 Operations on loops

We now define two abstract operations on loops: $\mathtt{T}$ and $\mathtt{H}$. These operations are easy to describe for loops in terms of standard and dual notation, respectively, and we will see in the next subsection that they correspond to important bordered Floer operations.

If a loop $\boldsymbol{\ell}$ cannot be written in standard notation (that is, it is a collection of only e^* segments), then $\mathtt{T}(\boldsymbol{\ell}) = \boldsymbol{\ell}$. Otherwise, express $\boldsymbol{\ell}$ in standard notation and consider the operation $\mathtt{T}$ determined on individual letters via

$$\mathtt{T}(a_i) = a_i, \quad \mathtt{T}(b_i) = b_i, \quad \mathtt{T}(c_j) = c_{j-1}, \quad \mathtt{T}(d_j) = d_{j+1}$$

for any $i \in \mathbb{Z} \smallsetminus \{0\}$ and $j \in \mathbb{Z}$. For collections of loops, we also define $\mathtt{T}(\{\boldsymbol{\ell}_i\}_{i=1}^n) = \{\mathtt{T}(\boldsymbol{\ell}_i)\}_{i=1}^n$. The operation $\mathtt{T}$ is invertible; denote the inverse by $\mathtt{T}^{-1}$.

Note that the letters a_i and b_i are fixed by $\mathtt{T}$. As a result, these are sometimes referred to as stable chains (or standard stable chains). The (standard) unstable chains are the letters c_j and d_j. This is consistent with the notion of stable and unstable chains from [20].

The operator H is defined similarly, but with respect to dual notation. If a loop ℓ cannot be written in dual notation, then $\mathrm{H}(\ell) = \ell$. Otherwise, express ℓ in dual notation and consider the operation H on ℓ defined on individual dual letters via

$$\mathrm{H}(a_i^*) = a_i^*, \quad \mathrm{H}(b_i^*) = b_i^*, \quad \mathrm{H}(c_j^*) = c_{j-1}^*, \quad \mathrm{H}(d_j^*) = d_{j+1}^*$$

for any $i \in \mathbb{Z} \smallsetminus \{0\}$ and $j \in \mathbb{Z}$. For collections of loops, we also define $\mathrm{H}(\{\ell_i\}_{i=1}^n) = \{\mathrm{H}(\ell_i)\}_{i=1}^n$. As with T, H is an invertible operation with inverse H^{-1}.

Note that a_i^* and b_i^* are fixed by H. When we have need for it, a^*– and b^*–type letters will be referred to as dual stable chains, while c^*– and d^*–type letters will be referred to as dual unstable chains.

The operation T is easy to define for loops in standard notation, but it would be difficult to describe purely in terms of dual letters. Similarly, the operation H is simple to define in dual notation, but would be complicated in terms of standard letters. This suggests the need to comfortably switch between the two; in particular, given a loop ℓ expressed in standard notation, finding $\mathrm{H}(\ell)$ in standard notation is a three-step process: dualize, apply H and dualize again.

We will often need to apply combinations of the operations T and H to a loop. The composition $\mathrm{T} \circ \mathrm{H}^{-1} \circ \mathrm{T}$, in particular, appears often; it will be convenient to regard this composition as another loop operation, which we call E. The following lemma describes the effect of E on a cyclic word in standard notation; this provides a convenient shortcut compared with computing the operations T, H^{-1} and T individually. We state the lemma for general loops, but we will only prove it in a special case.

Lemma 3.10 *If ℓ is written in standard notation, then $\mathrm{E}(\ell)$ is determined by the following action of* E *on standard letters*:

$$\mathrm{E}(a_k) = a_{-k}^*, \quad \mathrm{E}(b_k) = b_{-k}^*, \quad \mathrm{E}(c_k) = c_{-k}^*, \quad \mathrm{E}(d_k) = d_{-k}^*.$$

Proof We give the proof in the special case that ℓ consists only of d_k segments with $k \geq 0$. The general proof is left to the reader.

Let $\ell = (d_{k_1} d_{k_2} \dots d_{k_n})$, with $k_i \geq 0$. Then $\mathrm{T}(\ell) = (d_{k_1+1} \dots d_{k_n+1})$. Writing this in dual notation and applying H^{-1}, we have

$$\mathrm{T}(\ell) = (d_1^* \underbrace{d_0^* \dots d_0^*}_{k_1} d_1^* \underbrace{d_0^* \dots d_0^*}_{k_2} \dots d_1^* \underbrace{d_0^* \dots d_0^*}_{k_n}),$$

$$\mathrm{H}^{-1} \circ \mathrm{T}(\ell) = (d_0^* \underbrace{d_{-1}^* \dots d_{-1}^*}_{k_1} d_0^* \underbrace{d_{-1}^* \dots d_{-1}^*}_{k_2} \dots d_0^* \underbrace{d_{-1}^* \dots d_{-1}^*}_{k_n}).$$

Dualizing and twisting again gives

$$\begin{aligned}
\mathtt{H}^{-1}\circ\mathtt{T}(\boldsymbol{\ell}) &= (c_2\underbrace{c_1\dots c_1}_{k_1-1}c_2\underbrace{c_1\dots c_1}_{k_2-1}\dots c_2\underbrace{c_1\dots c_1}_{k_n-1}),\\
\mathtt{T}\circ\mathtt{H}^{-1}\circ\mathtt{T}(\boldsymbol{\ell}) &= (c_1\underbrace{c_0\dots c_0}_{k_1-1}c_1\underbrace{c_0\dots c_0}_{k_2-1}\dots c_1\underbrace{c_0\dots c_0}_{k_n-1}),\\
\mathtt{T}\circ\mathtt{H}^{-1}\circ\mathtt{T}(\boldsymbol{\ell}) &= (d^*_{-k_1}d^*_{-k_2}\dots d^*_{-k_n}).
\end{aligned}$$

□

As suggested by the example of the left-hand trefoil exterior given at the end of the last subsection, expressing loops as cyclic words gives rise to clean way of describing a type D structure. This extends to reparametrization: consulting [20, Chapter 11], when M is the exterior of the left-hand trefoil, $\widehat{CFD}(M,\mu,\lambda-(2+n)\mu)$ is represented by the loop $(a_1b_1c_n)$ for $n\in\mathbb{Z}$. Motivating the next subsection, this loop may be expressed as $\mathtt{T}^{-n}(\boldsymbol{\ell})$, where $\boldsymbol{\ell}=(a_1b_1c_0)$.

3.4 Dehn twists

The operators T and H naturally encode the effect of a Dehn twist on a loop representing the type D structure of a bordered manifold. Recall that

$$\widehat{T}_{\mathrm{st}}^{\pm1}\boxtimes\widehat{CFD}(M,\alpha,\beta)\cong\widehat{CFD}(M,\alpha,\beta\pm\alpha)$$

and, more generally, for any type D structure N over $\mathcal{A}$ the pair of type D structures $\widehat{T}_{\mathrm{st}}^{\pm1}\boxtimes N$ are well defined.

Proposition 3.11 *If $\boldsymbol{\ell}$ is a loop with corresponding type D structure $N_{\boldsymbol{\ell}}$, then $N_{\boldsymbol{\ell}}^{\pm}=\widehat{T}_{\mathrm{st}}^{\pm1}\boxtimes N_{\boldsymbol{\ell}}$ is a loop-type module represented by the loop $\mathtt{T}^{\pm}(\boldsymbol{\ell})$.*

Proof We compute the box tensor product $\widehat{T}_{\mathrm{st}}\boxtimes N_{\boldsymbol{\ell}}$ by considering the effect on one segment at a time in a standard representative for $\boldsymbol{\ell}$.

Recall from [19, Section 10] (see Figure 2) that the type DA bimodule $\widehat{T}_{\mathrm{st}}$ has three generators p, q and r (denoted by $\bullet$, $\circ$ and $*$, respectively, in Figure 2) with idempotents determined by $p=\iota_0p\iota_0$, $q=\iota_1q\iota_1$ and $r=\iota_1r\iota_0$. Thus each generator $x\in\iota_1N_{\boldsymbol{\ell}}$ gives rise to a single generator $q\otimes x\in\widehat{T}_{\mathrm{st}}\boxtimes N_{\boldsymbol{\ell}}$, and each generator $x\in\iota_0N_{\boldsymbol{\ell}}$ gives rise to two generators $p\otimes x$ and $r\otimes x$ in the box tensor product. There is always a ρ_2 arrow from $r\otimes x$ to $p\otimes x$; that is, $\partial(r\otimes x)$ has a summand $\rho_2\cdot(p\otimes x)$.

If generators $x,y\in\iota_0N_{\boldsymbol{\ell}}$ are connected by a single segment s (in the loop $\boldsymbol{\ell}$), we will consider the portion of a loop representing $N_{\boldsymbol{\ell}}^{+}$ between $p\otimes x$ and $p\otimes y$ and show

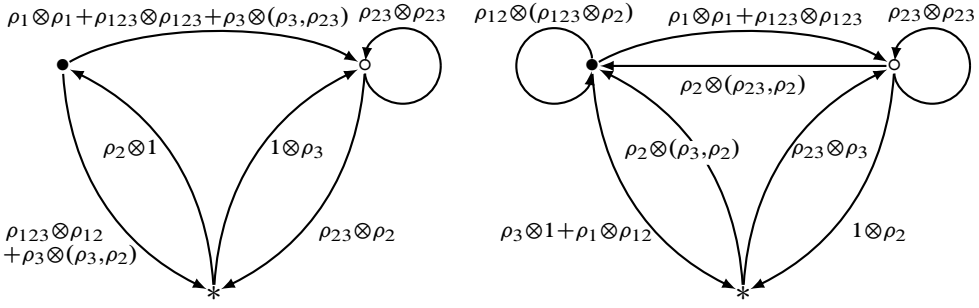


Figure 2: Graphical representations of the Dehn twist bimodules $\widehat{T}_{\mathrm{st}}$ (left) and $\widehat{T}_{\mathrm{st}}^{-1}$ (right), following [19, Section 10].

that (up to homotopy equivalence) it is the segment $\mathtt{T}(s)$. To talk about the portion *between*, we need a (cyclic) ordering on the elements of $N_{\boldsymbol{\ell}}^+$. This is inherited from a choice of cyclic ordering on the elements of $\boldsymbol{\ell}$, together with a specified order of $p\otimes x$ and $r\otimes x$ for each $x\in\iota_0 N_{\boldsymbol{\ell}}$. If there is a segment s from x to y, we say that $r\otimes x$ is between $p\otimes x$ and $p\otimes y$ if the puzzle piece shape of s on the x end agrees with the shape of a_k, that is, if s is a_k, $\bar{a}_k$, c_k, $\bar{d}_k$ or $\bar{e}$.

Consider first a segment a_k from x to y, where x and y are generators of $\iota_0 N_{\boldsymbol{\ell}}$. The cases of $k=1$ and $k\geq 2$ are slightly different; both are pictured in Figure 3. In either case, the effect of tensoring the segment with $\widehat{T}_{\mathrm{st}}$ is pictured. There is a differential starting at $r\otimes x$; after cancelling this differential, the result is a segment of type $a_k=\mathtt{T}(a_k)$ from $p\otimes x$ to $p\otimes y$.

Consider next a segment b_k from x to y. Tensoring the segment with $\widehat{T}_{\mathrm{st}}$, we see that the portion between $p\otimes x$ and $p\otimes y$ is simply a segment of type $b_k=\mathtt{T}(b_k)$. Note that the generators $r\otimes x$ and $r\otimes y$ are not included in this new segment; they must be included in the segments on either side.

If x and y are connected by a segment c_k, then there is a differential starting at $r\otimes x$. If $k>1$, then cancelling this differential leaves a segment from $p\otimes x$ to $p\otimes y$ of type $c_{k-1}=\mathtt{T}(c_k)$. If $k=1$, then cancelling the differential produces a new ρ_{12} arrow, and thus there is a segment $\bar{e}=\mathtt{T}(c_1)$ from $p\otimes x$ to $p\otimes y$. If x and y are connected by segments d_k or e, we see in Figure 3 that the portion of $N_{\boldsymbol{\ell}}^+$ from $x\otimes p$ to $y\otimes p$ is a segment $d_{k+1}=\mathtt{T}(d_k)$ or $d_1=\mathtt{T}(e)$, respectively.

Segments with the opposite orientation behave the same way. A segment of $\bar{s}$ from x to y is the same as a segment s from y to x. In the tensor product, this produces a segment $\mathtt{T}(s)$ from $p\otimes y$ to $p\otimes x$, or a segment $\overline{\mathtt{T}(s)}=\mathtt{T}(\bar{s})$ from $p\otimes x$ to $p\otimes y$.

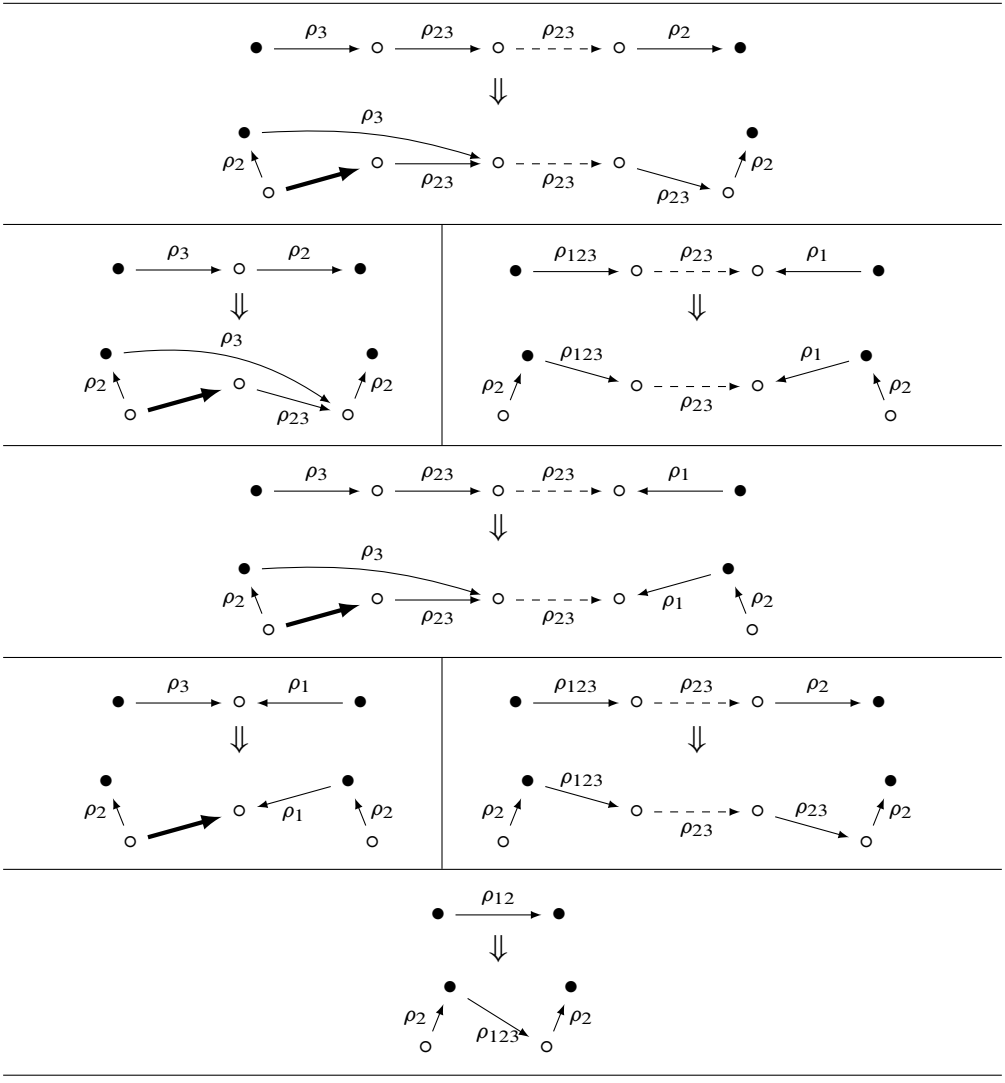


Figure 3: Illustrating the proof of Proposition 3.11: The effect of box tensoring with $\widehat{T}_{\mathrm{st}}$ on each of the possible segments occurring in a loop expressed in standard notation. Unmarked edges, which are eliminated using edge reduction described in Section 2.3, are highlighted.

It remains to check that a loop consisting only of e^* segments represents a type D structure N that is fixed by the action of the standard Dehn twist. This is easy to see, since in this case the only relevant operation in $\widehat{T}_{\mathrm{st}}$ is

$$m_2(q, \rho_{23}) = \rho_{23} \cdot q.$$

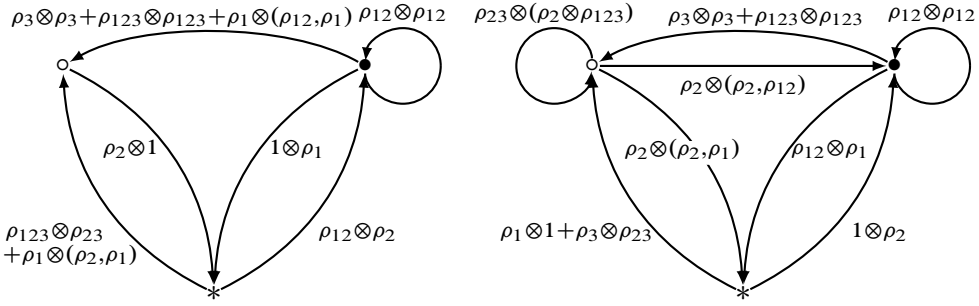


Figure 4: Graphical representations of the Dehn twist bimodules $\widehat{T}_{\mathrm{du}}^{-1}$ (left) and $\widehat{T}_{\mathrm{du}}$ (right), following [19, Section 10].

Each generator of $\iota_1 N = N$ gives rise to one generator of N^+ and each ρ_{23} arrow in N gives a ρ_{23} arrow in N^+.

The case of $\widehat{T}_{\mathrm{st}}^{-1}$ can be deduced from the case of $\widehat{T}_{\mathrm{st}}$. Let N' be a type D module represented by the loop $\mathrm{T}^{-1}(\boldsymbol{\ell})$. We have just shown

$$\widehat{T}_{\mathrm{st}} \boxtimes N' \cong N$$

and it follows that

$$\widehat{T}_{\mathrm{st}}^{-1} \boxtimes N \cong \widehat{T}_{\mathrm{st}}^{-1} \boxtimes \widehat{T}_{\mathrm{st}} \boxtimes N' \cong N'$$

since $\widehat{T}_{\mathrm{st}}^{-1} \boxtimes \widehat{T}_{\mathrm{st}}$ is homotopy equivalent to the identity bimodule. □

Proposition 3.12 *If $\boldsymbol{\ell}$ is a loop with corresponding type D structure $N_{\boldsymbol{\ell}}$, then $N_{\boldsymbol{\ell}}^{\mp} = \widehat{T}_{\mathrm{du}}^{\mp 1} \boxtimes N_{\boldsymbol{\ell}}$ is a loop-type module represented by the loop $\mathrm{H}^{\pm}(\boldsymbol{\ell})$.*

Proof The proof is similar, with the relevant bimodules $\widehat{T}_{\mathrm{du}}$ and $\widehat{T}_{\mathrm{du}}^{-1}$ shown in Figure 4. The result of tensoring $\widehat{T}_{\mathrm{du}}$ with each type of dual segment is shown in Figure 5. We see that a^* and b^* segments are fixed, c_k^* segments become c_{k+1}^* segments, and d_k^* segments become d_{k-1}^* segments. In other words, tensoring $\widehat{T}_{\mathrm{du}}$ with a dual segment $\boldsymbol{s}^*$ gives $\mathrm{H}^{-1}(\boldsymbol{s}^*)$. Thus, for a loop $\boldsymbol{\ell}$, $\widehat{T}_{\mathrm{du}} \boxtimes \boldsymbol{\ell}$ is the loop $\mathrm{H}^{-1}(\boldsymbol{\ell})$. Since $\widehat{T}_{\mathrm{du}}^{-1}$ is the inverse of $\widehat{T}_{\mathrm{du}}$, we conclude that $\widehat{T}_{\mathrm{du}}^{-1} \boxtimes \boldsymbol{\ell}$ is the loop $\mathrm{H}(\boldsymbol{\ell})$. □

We conclude this discussion by observing that the notion of a manifold M (with torus boundary) being of loop-type is now well defined. In particular, since any reparametrization of a loop gives rise to a loop, it follows that the property of loop-type (or having loop-type bordered invariants) is independent of the bordered structure and hence a property of the underlying (unbordered) manifold. Indeed:

Definition 3.13 A compact, orientable, connected three-manifold M with torus boundary is loop-type if $\widehat{CFD}(M, \alpha, \beta)$ is of loop-type for any choice of basis slopes α and β.

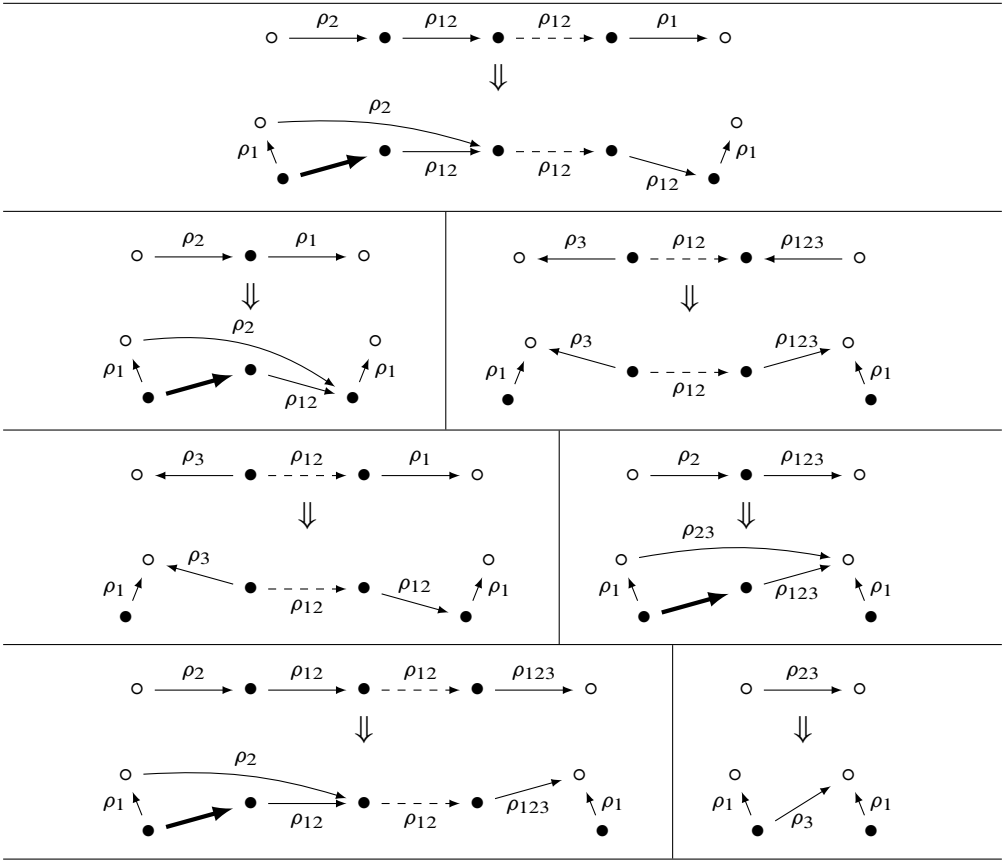


Figure 5: Illustrating the proof of Proposition 3.12: The effect of the Dehn twist $\widehat{T}_{\mathrm{du}}$ on each of the possible segments occurring in a loop expressed in dual notation. Unmarked edges, which are eliminated using edge reduction described in Section 2.3, are highlighted.

3.5 Solid tori

As a simple example of the loop operations described above, we now describe the computation of $\widehat{CFD}$ of a solid torus with arbitrary framing. By a p/q–framed solid torus, we will mean a bordered solid torus $(D^2\times S^1,\alpha,\beta)$ such that the meridian is $p\alpha+q\beta$. Recall that for the standard (0–framed) and dual (∞–framed) solid tori we have

$$\widehat{CFD}(D^2\times S^1,l,m)=(e),\quad \widehat{CFD}(D^2\times S^1,m,l)=(e^*).$$

The bordered invariants for solid tori with other framings can be computed by applying Dehn twists as described in Section 2.6.

Example 3.14 We compute $\widehat{CFD}$ of the $\frac{7}{2}$–framed solid torus using the continued fraction expansion $-\frac{2}{7} = [-1, 2, -2, 3]$:

$$\begin{aligned}
\widehat{CFD}(D^2\times S^1, l, m): &\quad (e)\\
&\Downarrow \mathrm{T}^3\\
\widehat{CFD}(D^2\times S^1, l, m+3l): &\quad (d_3) \quad\sim\quad (d_1^* d_0^* d_0^*)\\
&\qquad\qquad\qquad\qquad\Downarrow \mathrm{H}^{-2}\\
\widehat{CFD}(D^2\times S^1, -2m-5l, m+3l): &\quad (c_1c_1c_0c_1c_0) \quad\sim\quad (d_{-1}^* d_{-2}^* d_{-2}^*)\\
&\Downarrow \mathrm{T}^2\\
\widehat{CFD}(D^2\times S^1, -2m-5l, -3m-7l): &\quad (c_{-1}c_{-1}c_{-2}c_{-1}c_{-2}) \sim (c_{-1}^* c_{-1}^* c_{-1}^* c_0^* c_{-1}^* c_{-1}^* c_0^*)\\
&\qquad\qquad\qquad\qquad\Downarrow \mathrm{H}^{-1}\\
\widehat{CFD}(D^2\times S^1, m+2l, -3m-7l): &\quad (d_{-4}d_{-3}) \quad\sim\quad (c_0^* c_0^* c_0^* c_1^* c_0^* c_0^* c_1^*)
\end{aligned}$$

It is easy to check $m = 7(m+2l)+2(-3m-7l)$, and so $(D^2\times S^1, m+2l, -3m-7l)$ is a $\frac{7}{2}$–framed solid torus. Note that given $\widehat{CFD}$ of a solid torus we can also check the framing by using Proposition 2.10. If we choose the $(\mathbb{Z}/2\mathbb{Z})$–grading so that endpoints of d_k chains in standard notation have grading 0 and endpoints of c_k chains have grading 1, it is not difficult to see that the Euler characteristic $\chi_\bullet$ of a loop in standard notation is the number of d_k segments minus the number of c_k segments and $\chi_\circ$ is given by the sum of the subscripts. Thus $(D^2\times S^1, m+2l, -3m-7l)$ has rational longitude $-\chi_\circ/\chi_\bullet = -\frac{-7}{2}$.

$\widehat{CFD}$ for arbitrarily framed solid tori can be computed by a similar procedure. The result is always a loop of a particularly simple form.

Lemma 3.15 *If $q \neq 0$, then $\widehat{CFD}(D^2\times S^1, pm+ql, rm+sl)$ can be represented by a single loop $\ell = (d_{k_1} d_{k_2} \dots d_{k_m})$. Moreover, the difference between $\max\{k_i\}$ and $\min\{k_i\}$ is at most 1.*

Proof If $p = 0$ (this implies that $q = r = 1$) then

$$\widehat{CFD}(D^2\times S^1, l, m+sl) = \hat{T}_{\mathrm{st}}^s \boxtimes \widehat{CFD}(D^2\times S^1, l, m) \sim \mathrm{T}^s((e)) = (d_s).$$

Otherwise, let $[a_1, \dots, a_{2n}]$ be an even length continued fraction for q/p and choose a_{2n+1} so that $[a_1, \dots, a_{2n+1}]$ is a continued fraction for s/r. Then

$$\begin{aligned}
\widehat{CFD}(D^2\times S^1, pm+ql, rm+sl) &\cong \hat{T}_{\mathrm{st}}^{a_{2n+1}} \boxtimes \cdots \boxtimes \hat{T}_{\mathrm{du}}^{-a_2} \boxtimes \hat{T}_{\mathrm{st}}^{a_1} \boxtimes \widehat{CFD}(M, l, m)\\
&\sim \mathrm{T}^{a_{2n+1}} \circ \mathrm{H}^{a_{2n}} \circ \cdots \circ \mathrm{H}^{a_2} \circ \mathrm{T}^{a_1}((d_0)).
\end{aligned}$$

If p/q is positive, we may assume that $a_i > 0$ for all $1 \le i \le 2n$. We will show by induction on n that $\mathrm{H}^{a_{2n}} \circ \cdots \circ \mathrm{H}^{a_2} \circ \mathrm{T}^{a_1}((d_0))$ is a loop in standard notation consisting only of d_0 and d_1 chains. The base case of $n = 0$ is immediate. Assuming that $\mathrm{H}^{a_{2n-2}} \circ \cdots \circ \mathrm{H}^{a_2} \circ \mathrm{T}^{a_1}((d_0))$ is a loop in standard notation consisting only of d_0 and d_1 chains, applying the twist $\mathrm{T}^{a_{2n-1}}$ produces a loop consisting of d_k chains with only positive subscripts. In dual notation, such a loop involves only d_0^* and d_1^* segments. Applying the twist $\mathrm{H}^{a_{2n}}$ produces a loop consisting of d_k^* segments with positive subscripts. Switching to standard notation, this loop contains only d_0 and d_1 segments. Similarly, if p/q is negative, we assume that $a_i < 0$ for $1 \le i \le n$ and observe by induction that $\mathrm{H}^{a_{2n}} \circ \cdots \circ \mathrm{H}^{a_2} \circ \mathrm{T}^{a_1}((d_0))$ is a loop in standard notation consisting only of d_0 and d_{-1} chains.

Finally, applying the twist $\mathrm{T}^{a_{2n+1}}$ preserves the fact that the loop consists of type d_k unstable chains. It also preserves the relative differences of the subscripts, so the difference between the maximum and minimum subscript remains at most one. □

3.6 Abstract fillings and abstract slopes

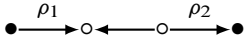
Recall that a loop $\boldsymbol{\ell}$ represents a type D structure, which by slight abuse of notation we denote by $\boldsymbol{\ell}$. Since $\boldsymbol{\ell}$ is reduced, there is an associated type A structure as described in Section 2.4, which we denote by $\boldsymbol{\ell}^A$. As a result, given loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ we can represent the chain complex produced by the box tensor product of the associated modules by $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$. Since loops are connected $\mathcal{A}$–decorated graphs, the type D and type A structures associated to a loop have a well-defined relative $(\mathbb{Z}/2\mathbb{Z})$–grading, as described in Section 2.7. The gradings on $\boldsymbol{\ell}_1^A$ and $\boldsymbol{\ell}_2$ induce a relative $(\mathbb{Z}/2\mathbb{Z})$–grading on $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$.

Consider the loops $\boldsymbol{\ell}_\bullet = (e)$ and $\boldsymbol{\ell}_\circ = (e^*)$. Given any loop $\boldsymbol{\ell}$, we have a pair of chain complexes

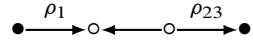
$$C_{\boldsymbol{\ell}}\left(\tfrac{1}{0}\right) = \boldsymbol{\ell}_\bullet^A \boxtimes \boldsymbol{\ell} \quad \text{and} \quad C_{\boldsymbol{\ell}}(0) = \boldsymbol{\ell}_\circ^A \boxtimes \boldsymbol{\ell},$$

which, noting that $\boldsymbol{\ell}_\bullet$ and $\boldsymbol{\ell}_\circ$ represent type D structures of the standard and dual (bordered) solid torus, respectively, might be regarded as abstract standard and dual Dehn filings of $\boldsymbol{\ell}$, respectively.

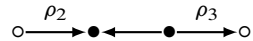
Remark 3.16 We do not need to assume $\boldsymbol{\ell}$ is bounded, since, if it is not, it is homotopy equivalent to a modified loop which is bounded and has the same box tensor product with $\boldsymbol{\ell}_\bullet^A$. Such a modified loop can be obtained by replacing either a ρ_{12} arrow with the homotopy equivalent (but not reduced) sequence



or a ρ_{123} arrow with

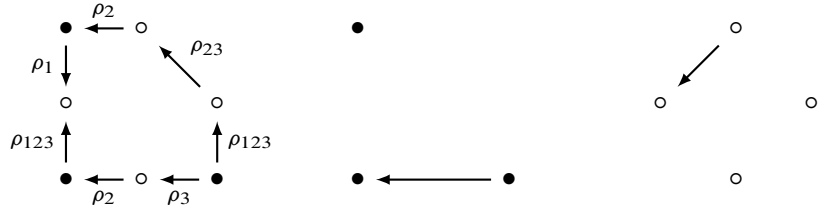


If $\boldsymbol{\ell}$ cannot be written in standard notation — that is, it is a collection of e^* segments — then $H_*(\boldsymbol{\ell}_\bullet^A \boxtimes \boldsymbol{\ell})$ must have two generators of opposite grading. To see this, replace one e^* segment with



to produce a bounded modified loop.

The standard filling picks out the $\bullet$–idempotent, in practice, and adds a differential for each type a_k chain. The dual filling picks out the $\circ$–idempotent and adds a differential for each b_k^* chain. For instance, when M is the exterior of the left-hand trefoil and $\widehat{CFD}(M,\mu,\lambda)$ is represented by $\boldsymbol{\ell} \sim (a_1 b_1 \bar{d}_2) \sim (a_1^* e^* b_1^* \bar{d}_1^*)$, the resulting complexes $\widehat{CF}(S^3) \cong \widehat{CF}(M(\mu)) \cong \boldsymbol{\ell}_\bullet^A \boxtimes \boldsymbol{\ell}$ and $\widehat{CF}(M(\lambda)) \cong \boldsymbol{\ell}_\circ^A \boxtimes \boldsymbol{\ell}$ are



We regard the chain complexes $C_{\boldsymbol{\ell}}\left(\frac{1}{0}\right)$ and $C_{\boldsymbol{\ell}}(0)$ as the result of abstract Dehn fillings along a pair of abstract slopes in $\boldsymbol{\ell}$, which we identify with $\infty = \frac{1}{0}$ (corresponding to the standard filling) and 0 (corresponding to the dual filling). In fact, a given loop $\boldsymbol{\ell}$ gives rise to a natural $\widehat{\mathbb{Q}}$–family of chain complexes: choosing an even-length continued fraction $p/q = [a_1, a_2, \ldots, a_n]$, let $\boldsymbol{\ell}^{p/q} = \mathrm{H}^{a_n} \circ \cdots \circ \mathrm{T}^{a_3} \circ \mathrm{H}^{a_2} \circ \mathrm{T}^{a_1}(\boldsymbol{\ell})$ and define

$$C_{\boldsymbol{\ell}}\left(\frac{p}{q}\right) = \boldsymbol{\ell}_\bullet^A \boxtimes \boldsymbol{\ell}^{p/q}.$$

We will regard the complex $C_{\boldsymbol{\ell}}(p/q)$ as an abstract Dehn filling of the loop $\boldsymbol{\ell}$ along the abstract slope $p/q \in \widehat{\mathbb{Q}}$.

The reason for this definition is illustrated as follows:

Proposition 3.17 *If a given loop $\boldsymbol{\ell}$ represents $\widehat{CFD}(M,\alpha,\beta)$ for some bordered manifold (M,α,β), the chain complex $C_{\boldsymbol{\ell}}(p/q)$ is (homotopy equivalent to) the chain complex $\widehat{CF}(M(p\alpha + q\beta))$. That is, abstract filling along abstract slopes corresponds to Dehn filling along slopes whenever $\boldsymbol{\ell}$ describes the type D structure of a bordered three-manifold.*

Proof This follows immediately from the definitions; note that $\boldsymbol{\ell}_{p/q}$ represents $\widehat{CFD}(M, p\alpha+q\beta, r\alpha+s\beta)$, where $p/q=[a_1,a_2,\dots,a_n]$ is an even length continued fraction and $r/s=[a_1,a_2,\dots,a_{n-1}]$. □

Fixing a relative $(\mathbb{Z}/2\mathbb{Z})$–grading for a loop $\boldsymbol{\ell}$, consider the idempotent Euler characteristics $\chi_\bullet(\boldsymbol{\ell})$ and $\chi_\circ(\boldsymbol{\ell})$. If $\chi_\bullet(\boldsymbol{\ell})$ and $\chi_\circ(\boldsymbol{\ell})$ are not both 0, then $\boldsymbol{\ell}$ has a preferred slope $-\chi_\circ(\boldsymbol{\ell})/\chi_\bullet(\boldsymbol{\ell})$; we call this slope the *abstract rational longitude*. By Proposition 2.10, if $\boldsymbol{\ell}$ represents $\widehat{CFD}(M,\alpha,\beta)$ for a rational homology solid torus (M,α,β), then the abstract rational longitude for $\boldsymbol{\ell}$ is the rational longitude for (M,α,β).

Recall that a slope p/q is an L–space slope for a bordered manifold (M,α,β) if Dehn filling along the curve $p\alpha+q\beta$ yields an L–space. We end this section by defining a similar notion of L–space slopes for loops.

Definition 3.18 Given a loop $\boldsymbol{\ell}$, we say an abstract slope p/q in $\hat{\mathbb{Q}}$ is an L–space slope for $\boldsymbol{\ell}$ if the relatively $(\mathbb{Z}/2\mathbb{Z})$–graded chain complex $C_{\boldsymbol{\ell}}(p/q)$ is an L–space chain complex in the sense that $\dim H_*(C_{\boldsymbol{\ell}}(p/q))=|\chi(C_{\boldsymbol{\ell}}(p/q))|\neq 0$.

Remark 3.19 With these notions in place, we will now drop the modifier *abstract* when treating loops despite the fact that a given loop may or may not describe the type D structure of a three-manifold. In particular, we will not make a distinction between slopes and abstract slopes in the sequel.

By considering loops in the abstract, a particular class of loops is singled out.

Definition 3.20 A loop $\boldsymbol{\ell}$ is solid torus-like if it may be obtained from the loop $(ee\cdots e)$ via applications of $\mathrm{T}^{\pm1}$ and $\mathrm{H}^{\pm1}$.

Note that $\chi_\bullet(ee\cdots e)$ counts the number of e segments appearing and $\chi_\circ(ee\cdots e)=0$ identifies the rational longitude of a solid torus-like loop. In particular, $\boldsymbol{\ell}_\bullet$ (representing $\widehat{CFD}(D^2\times S^1,l,m)$) is solid torus-like. Justifying the chosen terminology, we have the following behaviour:

Proposition 3.21 *If $\boldsymbol{\ell}$ is solid torus-like with $\chi_\circ(\boldsymbol{\ell})=0$, then $\boldsymbol{\ell}^A\boxtimes\boldsymbol{\ell}'\cong\bigoplus_{\chi_\bullet(\boldsymbol{\ell})}\boldsymbol{\ell}_\bullet^A\boxtimes\boldsymbol{\ell}'$ for any loop $\boldsymbol{\ell}'$.*

Proof Recall that $\boldsymbol{\ell}_\bullet^A$ has a single generator x and operations

$$m_{3+i}(x,\rho_3,\underbrace{\rho_{23},\dots,\rho_{23}}_{i\text{ times}},\rho_2)=x,$$

so that, for generators $x \otimes u, x \otimes v \in \boldsymbol{\ell}_\bullet^A \boxtimes \boldsymbol{\ell}'$, with $x \otimes v$ a summand of $\partial(x \otimes u)$, there must be

$$\delta^{i+2}(u) = \rho_3 \otimes \underbrace{\rho_{23} \otimes \cdots \otimes \rho_{23}}_{i \text{ times}} \otimes \rho_2 \otimes v$$

in the type D structure for $\boldsymbol{\ell}'$. Now consider the type A structure described by $\boldsymbol{\ell}^A$. This has $n = \chi_\bullet(\boldsymbol{\ell})$–generators $x_0, \dots, x_{n-1}$ and operations

$$m_{3+i}(x_j, \rho_3, \underbrace{\rho_{23}, \dots, \rho_{23}}_{i \text{ times}}, \rho_2) = x_{i+j+1},$$

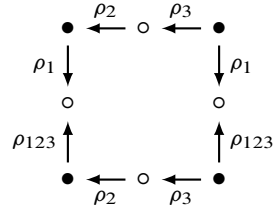
where the subscripts are understood to be reduced modulo n. (Note that the cyclic ordering on the generators is determined by $m_3(x_j, \rho_3, \rho_2) = x_{j+1}$.) So, given generators u and v in the type D structure for $\boldsymbol{\ell}'$ as above, $x_{i+j+1} \otimes v$ is in the image of $\partial(x_j \otimes u)$ for each $j \in \{0, \dots, n-1\}$. This achieves the desired splitting. □

Corollary 3.22 *Suppose $\boldsymbol{\ell}$ is solid torus-like with $\chi_\circ(\boldsymbol{\ell}) = 0$. Then $\boldsymbol{\ell}^A \boxtimes \boldsymbol{\ell}'$ is an L–space complex if and only if ∞ is an L–space slope for the loop $\boldsymbol{\ell}'$.*

As a result, with respect to gluing, solid torus-like loops will need to be treated like solid tori. Consequently, manifolds with solid torus-like invariants (should they exist) will need to be singled out.

Definition 3.23 A loop-type manifold M is solid torus-like if it is a rational homology solid torus and every loop in the representation of $\widehat{CFD}(M, \alpha, \beta)$ is solid torus-like.

We conclude this section with an explicit example, demonstrating that not every abstract loops arises as the type D structure of a bordered three-manifold. Consider the loop $\boldsymbol{\ell}$, described by the cyclic word $(a_1 b_1 \bar{a}_1 \bar{b}_1)$,



Suppose that $\boldsymbol{\ell}$ describes the invariant $\widehat{CFD}(M, \alpha, \beta)$ for some orientable three-manifold with torus boundary M. Then the abstract Dehn filling $\boldsymbol{\ell}_\bullet^A \boxtimes \boldsymbol{\ell}$ yields the chain complex $\widehat{CF}(M(\alpha))$. However, observe that $H_*(\boldsymbol{\ell}_\bullet^A \boxtimes \boldsymbol{\ell}) = 0$ (in particular, this slope is not an L–space slope according to Definition 3.18). This shows that no such M exists: in general, for a closed orientable three-manifold Y, $\widehat{HF}(Y)$ does not vanish [24].

However, we remark that this particular loop arises as a component of the graph describing $\widehat{CFD}(M,\mu,\lambda)$, where $M = S^3 \smallsetminus \nu(K)$ and K is any thin knot in S^3 that does not admit L–space surgeries; see [25].

4 Characterizing slopes

We now turn to an application of the loop calculus developed above. For a given loop $\boldsymbol{\ell}$ we give an explicit description of the L–space and non-L–space slopes. As a consequence, we will prove the following:

Theorem 4.1 *Given a loop $\boldsymbol{\ell}$, the set of non-L–space slopes is an interval in $\widehat{\mathbb{Q}}$, that is, the restriction of a connected subset in $\widehat{\mathbb{R}}$. As a result, if M is a loop-type manifold then there is a decomposition $\widehat{\mathbb{R}} = U \cup V$ into disjoint, connected subsets $U, V \in \widehat{\mathbb{R}}$ such that $\mathcal{L}_M = U \cap \widehat{\mathbb{Q}}$ and $\mathcal{L}_M^c = V \cap \widehat{\mathbb{Q}}$.*

Remark 4.2 The subset U determining $\mathcal{L}_M$ in this statement may be empty.

4.1 L–space slopes

A loop $\boldsymbol{\ell}$ has two preferred slopes, 0 and ∞; we begin by giving explicit conditions on the loop $\boldsymbol{\ell}$ (in terms of its representatives in standard or dual notation) under which these are L–space slopes.

Proposition 4.3 *Given a loop $\boldsymbol{\ell}$, ∞ is an L–space slope if and only if $\boldsymbol{\ell}$ can be written in standard notation with at least one d_k letter and no c_k letters (where k can be any integer). Similarly, 0 is an L–space slope if and only if $\boldsymbol{\ell}$ can be written in dual notation with at least one d_k^* letter and no c_k^* letters. Moreover, by reversing the orientation of $\boldsymbol{\ell}$, the statements hold with the roles of d_k and c_k (or the roles of d_k^* and c_k^*) reversed.*

Proof The slope ∞ is an L–space slope if $C\left(\frac{1}{0}\right) = \boldsymbol{\ell}_\bullet^A \boxtimes \boldsymbol{\ell}$ is an L–space complex, that is, if $H_*\big(C\left(\frac{1}{0}\right)\big)$ is nontrivial and each generator of $H_*\big(C\left(\frac{1}{0}\right)\big)$ has the same $(\mathbb{Z}/2\mathbb{Z})$–grading.

Recall that $\boldsymbol{\ell}_\bullet^A$ has a single generator x with idempotent ι_0 and operations

$$m_{3+i}(x, \rho_3, \underbrace{\rho_{23}, \dots, \rho_{23}}_{i \text{ times}}, \rho_2) = x$$

for each $i \geq 0$. The box tensor product of this module with $\boldsymbol{\ell}$ is easy to describe if $\boldsymbol{\ell}$ is written in standard notation. There is one generator $x \otimes y$ for each $\bullet$–vertex y in $\boldsymbol{\ell}$ (by abuse y is both the vertex in the loop $\boldsymbol{\ell}$ and the corresponding generator

in the type D structure corresponding to ℓ). Given a cyclic word representing ℓ in standard notation, each letter represents a chain between adjacent $\bullet$–vertices; for each type a_k chain from y_1 to y_2 there is a differential from $x \otimes y_1$ to $x \otimes y_2$. Since sequential occurrences of type a chains are impossible in any loop, the differentials on $C\left(\frac{1}{0}\right)$ are isomorphisms mapping single generators to single generators. Therefore, the contribution to $H_*(\ell_\bullet^A \boxtimes \ell)$ (with relative $(\mathbb{Z}/2\mathbb{Z})$–grading) is simply given by the $\bullet$–vertices of ℓ which are not an endpoint of a type a segment.

The $(\mathbb{Z}/2\mathbb{Z})$–grading on $C\left(\frac{1}{0}\right)$ can be recovered as follows: the generators corresponding to two adjacent $\bullet$–vertices in ℓ have the same grading if the vertices are connected by an unstable chain and opposite gradings if the vertices are connected by a stable chain. It follows that endpoints of chains of type d_k all have the same grading and endpoints of chains of type c_k all have the opposite grading. This is because d_k segments must be separated from each other by an even number of stable chains, and from c_k segments by an odd number of stable chains.

Suppose ℓ can be written in standard notation with at least one d_k and no c_k. Every generator of $C\left(\frac{1}{0}\right)$ either comes from an endpoint of a d_k or from the common endpoint of two stable chains. The latter generators vanish in homology, since one of the two stable chains must be type a, and the former generators all have the same $(\mathbb{Z}/2\mathbb{Z})$–grading. Thus, in this case, $C\left(\frac{1}{0}\right)$ is an L–space complex.

Suppose now that ℓ contains both d_k and c_k segments. For any unstable chain, at least one of the two endpoints must correspond to a generator of $C\left(\frac{1}{0}\right)$ that survives in homology, since an unstable chain can be adjacent to a type a chain on at most one side. Since endpoints of type d_k and type c_k unstable chains produce generators of opposite grading, it follows that $C\left(\frac{1}{0}\right)$ has generators of both gradings that survive in homology, and thus $C\left(\frac{1}{0}\right)$ is not an L–space complex.

If ℓ has no unstable chains when written in standard notation, then it has only stable chains, which alternate between types a and b. In this case every $\bullet$–vertex is the endpoint of a type a chain. Thus $H_*\left(C\left(\frac{1}{0}\right)\right)$ is trivial, and $C\left(\frac{1}{0}\right)$ is not an L–space complex. Finally, if ℓ cannot be written in standard notation, then it consists only of e^* segments. $H_*(\ell_\bullet^A \boxtimes \ell)$ has two generators with opposite gradings; it follows that ∞ is not an L–space slope.

The proof for 0–filling is almost identical: $\ell_\circ^A$ has a single generator x with idempotent ι_1 and operations

$$m_{3+i}(x, \rho_2, \underbrace{\rho_{12}, \ldots, \rho_{12}}_{i \text{ times}}, \rho_1) = x$$

for each $i \geq 0$. The box tensor product of this module with $\boldsymbol{\ell}$ has one generator for each generator y of $\boldsymbol{\ell}$ with idempotent ι_1 (that is, each $\circ$–vertex) and a differential for each type a^* segment. Expressing $\boldsymbol{\ell}$ in dual notation, the contribution to $H_*(\boldsymbol{\ell}_\circ^A \boxtimes \boldsymbol{\ell})$ (with relative $(\mathbb{Z}/2\mathbb{Z})$–grading) is given by the ι_0–generators of $\boldsymbol{\ell}$ which do not lie at the end of a type a^* segment. If $\widehat{CFD}(Y, \mathfrak{s})$ is a loop that cannot be written in dual notation — that is, it is a collection of e segments — then the contribution to $H_*(\boldsymbol{\ell}_\circ^A \boxtimes \boldsymbol{\ell}^D)$ is two generators of opposite grading.

Since a^* and b^* segments change the $(\mathbb{Z}/2\mathbb{Z})$–grading while c^* and d^* segments do not, the rest of the proof is completely analogous to the proof for the ∞–filling. □

Combining the two conditions in Proposition 4.3 gives a stronger condition on the loop.

Proposition 4.4 *Given a loop $\boldsymbol{\ell}$, both ∞ and 0 are L–space slopes if and only if the following equivalent conditions hold*:

(i) *$\boldsymbol{\ell}$ can be written in standard notation with at least one d_k letter and no c_k letters (with $k \in \mathbb{Z}$), and $\boldsymbol{\ell}$ contains a subword from exactly one of the sets*

$$A_+ = \{b_i a_j, a_i e^n b_j, a_i e^n d_j, d_i e^n b_j, d_i e^n d_j, d_\ell, a_\ell, b_\ell \mid i, j \geq 1, n \geq 0, \ell \geq 2\},$$
$$A_- = \{b_i a_j, a_i e^n b_j, a_i e^n d_j, d_i e^n b_j, d_i e^n d_j, d_\ell, a_\ell, b_\ell \mid i, j \leq 1, n \geq 0, \ell \leq 2\}.$$

We will say that $\boldsymbol{\ell}$ satisfies condition (i)$_\pm$ *if it satisfies condition* (i) *with a subword in $A_\pm$.*

(ii) *$\boldsymbol{\ell}$ can be written in dual notation with at least one d_k^* letter and no c_k^* letters (with $k \in \mathbb{Z}$), and $\boldsymbol{\ell}$ contains a subword from exactly one of the sets*

$$A_+^* = \{b_i^* a_j^*, a_i^* (e^*)^n b_j^*, a_i^* (e^*)^n d_j^*, d_i^* (e^*)^n b_j^*, d_i^* (e^*)^n d_j^*, d_\ell^*, a_\ell^*, b_\ell^* \mid i, j \geq 1, n \geq 0, \ell \geq 2\},$$
$$A_-^* = \{b_i^* a_j^*, a_i^* (e^*)^n b_j^*, a_i^* (e^*)^n d_j^*, d_i^* (e^*)^n b_j^*, d_i^* (e^*)^n d_j^*, d_\ell^*, a_\ell^*, b_\ell^* \mid i, j \leq 1, n \geq 0, \ell \leq 2\}.$$

We will say that $\boldsymbol{\ell}$ satisfies condition (ii)$_\pm$ *if it satisfies condition* (ii) *with a subword in $A_\pm^*$.*

Moreover, condition (i)$_+$ *is equivalent to condition* (ii)$_+$ *and condition* (i)$_-$ *is equivalent to condition* (ii)$_-$.

Proof By Proposition 4.3, 0 and ∞ are both L–space slopes for $\boldsymbol{\ell}$ if and only if $\boldsymbol{\ell}$ contains either d_k segments or c_k segments in standard notation, but not both, and $\boldsymbol{\ell}$ contains either d_k^* segments or c_k^* segments in dual notation, but not both. Reading

a loop with the opposite orientation takes type d segments to type c segments, so, possibly after switching the orientation on $\boldsymbol{\ell}$, we may assume such a loop contains d_k segments and not c_k segments in standard notation. Similarly, up to switching the orientation on $\boldsymbol{\ell}$, we could arrange that there are d_k^* segments and no c_k^* segments in dual notation. Note however that we cannot ensure both of these simultaneously, as it will be convenient to choose standard notation and dual notation representatives for $\boldsymbol{\ell}$ with respect to the same orientation on the loop.

First assume that $\boldsymbol{\ell}$ contains standard unstable chains of type d_k but not c_k. Under this assumption, the set A_+ is precisely the subwords in standard notation that dualize to give a d_k^* chain. This follows from the discussion on dualizing in Section 3.2. A d_{n+1}^* segment with $n \geq 0$ arises from a subword $a_i e^n b_j$, $a_i e^n d_j$, $d_i e^n b_j$ or $d_i e^n d_j$ with $i, j \geq 1$. A segment d_{-n-1}^* arises from a subword $b_i \bar{e}^n a_j$, $b_i \bar{e}^n c_j$, $c_i \bar{e}^n a_j$ or $c_i \bar{e}^n c_j$ with $i, j \geq 1$; since we assume that $\boldsymbol{\ell}$ contains no c_k segments, including $\bar{e} = c_0$, the only relevant case is $b_i a_j$. Finally, a $d_0^* = e^*$ segment arises from a standard letter of type a_ℓ, b_ℓ, c_ℓ, or d_ℓ with subscript $\ell > 1$; we can ignore the case of c_ℓ by assumption. By similar reasoning, we can check that, under the assumption that $\boldsymbol{\ell}$ contains no c_k letters, the set A_- is precisely the subwords in standard notation that dualize to give c_k^* chains.

If we instead assume that $\boldsymbol{\ell}$ contains dual unstable chains of type d_k^* but not c_k^*, the argument is similar. Note that the sets $A_\pm^*$ are the same as the sets $A_\pm$ with stars added to each letter, and the process for switching from dual to standard notation is the same as the process for switching from standard to dual (up to adding/removing stars from letters). We have immediately that A_+^* is the set of subwords in that dualize to give a d_k letter, and A_-^* is the set of subwords that dualize to give a c_k letter.

We have shown that conditions (i) and (ii) are both equivalent to $\boldsymbol{\ell}$ containing an unstable chain of exactly one of the two types in both standard notation and dual notation, and this is equivalent to both 0 and ∞ being L–space slopes. Finally, if $\boldsymbol{\ell}$ satisfies condition (i)$_+$ then $\boldsymbol{\ell}$ contains both a d_k segment and a d_k^* segment (with respect to the same orientation of the loop). It follows that $\boldsymbol{\ell}$ satisfies condition (ii)$_+$. If $\boldsymbol{\ell}$ satisfies condition (i)$_-$ then it contains both a d_k segment and a c_k^* segment. After reversing the orientation of $\boldsymbol{\ell}$, it contains a c_k segment and a d_k^* segment; thus $\boldsymbol{\ell}$ satisfies condition (ii)$_-$. □

Corollary 4.5 *If 0 and ∞ are L–space slopes for a loop $\boldsymbol{\ell}$, then the set of L–space slopes for $\boldsymbol{\ell}$ contains either all positive slopes (that is, the interval $[0, \infty]$) or all negative slopes (that is, the interval $[-\infty, 0]$).*

Proof Let L_+ be the set of loops satisfying condition (i)$_+$ and (ii)$_+$ in Proposition 4.4, and let L_- be the set of loops satisfying conditions (i)$_-$ and (ii)$_-$. It is not difficult to see that condition (i)$_+$ is preserved by the operation T since, for any $\boldsymbol{w} \in A_+$, $\mathrm{T}(\boldsymbol{w})$ is in A_+ or contains an element of A_+ as a subword. Similarly, condition (ii)$_+$ is preserved by the operation H. Therefore the set L_+ is preserved by T and H. In the same way, the set L_- is preserved by T^{-1} and H^{-1}.

If $p/q > 0$, let $[a_1, \dots, a_n]$ be a continued fraction for p/q of even length with all positive terms. To see if p/q is an L–space slope, we reparametrize by

$$\mathrm{H}^{a_n} \circ \cdots \circ \mathrm{T}^{a_3} \circ \mathrm{H}^{a_2} \circ \mathrm{T}^{a_1}$$

taking the slope p/q to the slope ∞. It follows that, if a loop is in L_+, then any $p/q > 0$ is an L–space slope, since the reparametrized loop is also in L_+.

If $p/q < 0$, let $[-a_1, \dots, -a_n]$ be a continued fraction for p/q of even length with all negative terms. To see if p/q is an L–space slope, we reparametrize by

$$\mathrm{H}^{-a_n} \circ \cdots \circ \mathrm{T}^{-a_3} \circ \mathrm{H}^{-a_2} \circ \mathrm{T}^{-a_1}$$

taking the slope p/q to the slope ∞. It follows that if a loop is in L_-, then any $p/q < 0$ is an L–space slope. □

4.2 Non-L–space slopes

The goal of this section is to establish easily checked conditions certifying that the standard and dual slope of a given loop are non-L–space slopes. Our focus will be on the following result:

Proposition 4.6 *Suppose that ℓ is a loop for which both standard and dual fillings give rise to non-L–spaces. Then the standard and dual slopes bound an interval of non-L–space slopes in $\mathcal{L}^c_M$. That is, one of $[0, \infty]$ or $[-\infty, 0]$ consists entirely of non-L–space slopes.*

The proof of this result is similar to the proof of Corollary 4.5 but is more technical and will therefore be built up in a series of lemmas.

To begin, note that Proposition 4.3 can be restated in terms of non-L–space slopes:

Proposition 4.7 *The slope ∞ is a non-L–space slope for a loop ℓ if and only if either*

(i) *ℓ contains both c_k and d_k unstable chains in standard notation,*

(ii) *ℓ contains no unstable chains in standard notation, or*

(iii) *ℓ cannot be written in standard notation.*

The slope 0 is a non-L–space slope if and only if either

(i) *ℓ contains both c_k^* and d_k^* unstable chains in dual notation,*

(ii) *ℓ contains no unstable chains in dual notation, or*

(iii) *ℓ cannot be written in dual notation.*

It will be helpful to give conditions in standard notation under which 0 is a non-L–space slope. Consider the sets

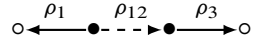
$$A_1 = \{a_k, b_k, c_k, c_i c_0^n c_j, c_i c_0^n a_j, b_i c_0^n c_j, b_i c_0^n a_j\},$$
$$A_2 = \{a_{-k}, b_{-k}, d_{-k}, d_{-i} d_0^n d_{-j}, d_{-i} d_0^n b_{-j}, a_{-i} d_0^n d_{-j}, a_{-i} d_0^n b_{-j}\},$$
$$A_3 = \{a_k, b_k, d_k, d_i d_0^n d_j, d_i d_0^n b_j, a_i d_0^n d_j, a_i d_0^n b_j\},$$
$$A_4 = \{a_{-k}, b_{-k}, c_{-k}, c_{-i} c_0^n c_{-j}, c_{-i} c_0^n a_{-j}, b_{-i} c_0^n c_{-j}, b_{-i} c_0^n a_{-j}\},$$

where the indices and exponents run over all integers satisfying $k > 1$, $i, j > 0$ and $n \geq 0$. These four sets may be interpreted as follows:

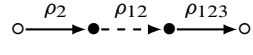
Lemma 4.8 *A loop ℓ written in standard notation contains a word in $A_1 \cup A_3$ if and only if it contains d_n^* in dual notation for some integer n. It contains a word in $A_2 \cup A_4$ if and only if it contains c_n^* in dual notation for some integer n.*

Proof We prove the first statement and leave the second to the reader.

First notice that if $k > 1$ then any letter a_k, b_k, c_k or d_k contains at least one instance of $\circ \xrightarrow{\rho_{23}} \circ$, which, in dual notation, gives an $e^* = d_0^*$. If $i, j > 0$ then each word of the form $c_i c_0^n c_j$, $c_i c_0^n a_j$, $b_i c_0^n c_j$ or $b_i c_0^n a_j$ gives an instance of



which, in dual notation, gives $\bar{c}_{n+1}^* = d_{-n-1}^*$. Similarly, each word of the form $d_i d_0^n d_j$, $d_i d_0^n b_j$, $a_i d_0^n d_j$ or $a_i d_0^n b_j$ gives an instance of



which, in dual notation, gives d_{n+1}^*. For the converse, observe that the segments d_0^*, d_{n+1}^* and d_{-n-1}^* with $n \geq 0$ in dual notation can only arise from the words mentioned above in standard notation. □

Corollary 4.9 *The slope 0 is not an L–space slope for ℓ if and only if either*

(i) *ℓ contains a subword from $A_1 \cup A_3$ and a subword from $A_2 \cup A_4$,*

(ii) *ℓ does not contain a subword from any A_i, or*

(iii) *ℓ cannot be written in dual notation.*

Proof This follows immediately from Proposition 4.7 and Lemma 4.8. □

Based on this observation, the proof of Proposition 4.6 will reduce to several cases, which are treated in the following lemmas.

Lemma 4.10 *If $\boldsymbol{\ell}$ is a loop containing a subword from the set A_1 then $\mathtt{T}^{-1}(\boldsymbol{\ell})$ and $\mathtt{H}^{-1}(\boldsymbol{\ell})$ also contain a subword from A_1. If $\boldsymbol{\ell}$ contains a subword from the set A_2 then $\mathtt{T}^{-1}(\boldsymbol{\ell})$ and $\mathtt{H}^{-1}(\boldsymbol{\ell})$ also contain a subword from A_2.*

Proof First consider $\mathtt{T}^{-1}(\boldsymbol{w})$ for any subword $\boldsymbol{w} \in A_1$. We have $\mathtt{T}^{-1}(a_k) = a_k$, $\mathtt{T}^{-1}(b_k) = b_k$, $\mathtt{T}^{-1}(c_k) = c_{k+1}$, $\mathtt{T}^{-1}(c_i\bar{e}^n c_j) = c_{i+1}c_1^n c_{j+1}$, $\mathtt{T}^{-1}(c_i\bar{e}^n a_j) = c_{i+1}c_1^n a_j$, $\mathtt{T}^{-1}(b_i\bar{e}^n c_j) = b_i c_1^n c_{j+1}$ and $\mathtt{T}^{-1}(b_i\bar{e}^n a_j) = b_i c_1^n a_j$. In each case, $\mathtt{T}^{-1}(\boldsymbol{w})$ is in A_1 or contains a subword in A_1. Thus the operation $\mathtt{T}^{-1}$ preserves the property that $\boldsymbol{\ell}$ contains a word from A_1. On the other hand, revisiting the proof of Lemma 4.8, we see that each of a_k, b_k and c_k produces at least one e^* when dualized, while each of $c_i\bar{e}^n c_j$, $c_i\bar{e}^n a_j$, $b_i\bar{e}^n c_j$ and $b_i\bar{e}^n a_j$ produces a $\bar{c}^*_{n+1}$. Since $\mathtt{H}^{-1}(e^*) = \bar{c}^*_1$ and $\mathtt{H}^{-1}(\bar{c}^*_{n+1}) = \bar{c}^*_{n+2}$, it is enough to observe that

$$\{c_i\bar{e}^n c_j, c_i\bar{e}^n a_j, b_i\bar{e}^n c_j, b_i\bar{e}^n a_j\} \subset A_1$$

is the set of subwords that give rise to a $\bar{c}^*_{n+1}$ under dualizing (where $n \geq 0$ and $i, j > 0$). It follows that the operation $\mathtt{H}^{-1}$ preserves the property that $\boldsymbol{\ell}$ contains a word from A_1. Next consider $\mathtt{T}^{-1}(\boldsymbol{w})$ for any $\boldsymbol{w} \in A_2$. Proceeding as above, observe that $\mathtt{T}^{-1}(\boldsymbol{w})$ is in the set

$$\{\bar{a}_k, \bar{b}_k, \bar{c}_{k+1}, \bar{c}_{i+1}\bar{c}_1^n\bar{c}_{j+1}, \bar{c}_{i+1}\bar{c}_1^n\bar{b}_j, \bar{a}_i\bar{c}_1^n\bar{c}_{j+1}, \bar{a}_i\bar{c}_1^n\bar{b}_j\}.$$

Each of these words is in A_2 or contains a subword in A_2, so the operation $\mathtt{T}^{-1}$ preserves the property that $\boldsymbol{\ell}$ contains a word from A_2. Each of $\bar{a}_k$, $\bar{b}_k$ and $\bar{c}_k$ produces at least one $\bar{e}^*$ when dualized, while each of $\bar{c}_i e^n\bar{c}_j$, $\bar{c}_i e^n\bar{b}_j$, $\bar{a}_i e^n\bar{c}_j$ and $\bar{a}_i e^n\bar{b}_j$ produces a c^*_{n+1}. Since $\mathtt{T}^{-1}(\bar{e}^*) = c_1^*$ and $\mathtt{T}^{-1}(c^*_{n+1}) = c^*_{n+2}$, it is enough to observe that

$$\{\bar{c}_i e^n\bar{c}_j, \bar{c}_i e^n\bar{b}_j, \bar{a}_i e^n\bar{c}_j, \bar{a}_i e^n\bar{b}_j\} \subset A_2$$

is the set of subwords that give rise to a c^*_{n+1} under dualizing (where $n \geq 0$ and $i, j > 0$). It follows that the operation $\mathtt{H}^{-1}$ preserves the property that $\boldsymbol{\ell}$ contains a word from A_2. □

Lemma 4.11 *If $\boldsymbol{\ell}$ is a loop containing a subword from the set A_3 then $\mathtt{T}(\boldsymbol{\ell})$ and $\mathtt{H}(\boldsymbol{\ell})$ also contain a subword from A_3. If $\boldsymbol{\ell}$ contains a subword from the set A_4, then $\mathtt{T}(\boldsymbol{\ell})$ and $\mathtt{H}(\boldsymbol{\ell})$ also contain a subword from A_4.*

Proof This is analogous to the proof of Lemma 4.10 and left to the reader. □

We next consider the case that ℓ contains a subword from each of A_1 and A_4 but does not contain a subword from A_2 or A_3. While this third case is ultimately similar to the two cases we have already treated, it is more technical owing in part to three subcases.

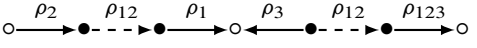
Lemma 4.12 *Suppose ℓ is a loop containing a subword from each of the sets A_1 and A_4 but that ℓ does not contain a subword from either of the sets A_2 or A_3. If ℓ contains a $\bar{c}_1$ then the loop $\mathrm{T}^{-1}(\ell)$ contains a subword from each of A_1 and A_2, while the loop $\mathrm{H}^{-1}(\ell)$ either*

(i) *contains a subword from each of A_1 and A_2, or*

(ii) *contains a $\bar{c}_1$ and a subword from each of the sets A_1 and A_4 but does not contain a subword from A_2 or from A_3.*

Proof First notice that $\mathrm{T}^{-1}(\bar{c}_1) = \bar{c}_2 \in A_2$. Since the set A_1 is closed under T^{-1} (compare Lemma 4.10), it follows that $\mathrm{T}^{-1}(\ell)$ contains a subword from both A_1 and A_2.

Turning now to the behaviour under H^{-1}: Since ℓ does not contain a word from A_2, any occurrence of $\bar{c}_1$ can be preceded only by $a_1 e^i$ or $d_1 e^i$ (for $i \geq 0$), and can be followed only by $e^j b_1$ or $e^j d_1$ (for $j \geq 0$). (Note that instances of a_k or b_k for $k > 1$ are ruled out since ℓ does not contain a word from A_3.) Regardless of which arises, this ensures that a $\bar{c}_1$ is part of a segment in ℓ of the form



and hence a subword $b_{i+1}^* a_{j+1}^*$ when ℓ is expressed in dual notation. It follows that the property that ℓ contains a $\bar{c}_1$ is closed under H^{-1}, under the assumption that ℓ does not contain a subword from A_2 or A_3.

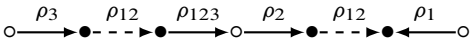
Now observe that the set A_1 is closed under H^{-1} by Lemma 4.10, so it must be the case that $\mathrm{H}^{-1}(\ell)$ contains subwords from A_1 and A_2 or, if this is not the case, from A_1 and A_4. In the latter case, suppose that $\mathrm{H}^{-1}(\ell)$ contains a word from A_3. Then $\mathrm{H}(\mathrm{H}^{-1}(\ell)) = \ell$ contains a word from A_3 also, which is a contradiction. □

Lemma 4.13 *Suppose ℓ is a loop containing a subword from each of the sets A_1 and A_4 but that ℓ does not contain a subword from either of the sets A_2 or A_3. If ℓ contains a d_1 then the loop $\mathrm{T}(\ell)$ contains a subword from each of A_3 and A_4, while the loop $\mathrm{H}(\ell)$ either*

(i) *contains a subword from each of A_3 and A_4, or*

(ii) *contains a d_1 and a subword from each of the sets A_1 and A_4 but does not contain a subword from A_2 or from A_3.*

Proof This is similar to the proof of Lemma 4.12. The key observation is that the instance of d_1 is part of a segment in $\boldsymbol{\ell}$ of the form



which, in dual form, gives $a_{j+1}^* b_{i+1}^*$. Thus the property that $\boldsymbol{\ell}$ contains a d_1 implies that $\mathrm{H}(\boldsymbol{\ell})$ contains a d_1. □

Lemma 4.14 *Suppose $\boldsymbol{\ell}$ is a loop containing a subword from each of the sets A_1 and A_4 but that $\boldsymbol{\ell}$ does not contain a subword from either of the sets A_2 or A_3, and suppose that ∞ is a non-L–space filling for $\boldsymbol{\ell}$. If $\boldsymbol{\ell}$ contains neither a $\bar{c}_1 = d_{-1}$ nor a d_1 then two cases arise: either $\boldsymbol{\ell}$ contains the subword $\boldsymbol{w}_1 = a_1 d_0^n \bar{b}_1$ or $\boldsymbol{\ell}$ contains the subword $\boldsymbol{w}_2 = \bar{a}_1 d_0^n b_1$ for some $n \geq 1$. Regardless of which occurs,*

(i) $\mathrm{T}(\boldsymbol{\ell})$ *contains subwords from both A_3 and A_4,*

(ii) $\mathrm{T}^{-1}(\boldsymbol{\ell})$ *contains subwords from both A_1 and A_2,*

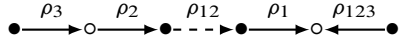
(iii) $\mathrm{H}(\boldsymbol{\ell})$ *either contains subwords A_3 and A_4 or it contains subwords from A_1 and A_4 and contains $\boldsymbol{w}_1$ or $\boldsymbol{w}_2$, and*

(iv) $\mathrm{H}^{-1}(\boldsymbol{\ell})$ *either contains subwords from A_1 and A_2 or it contains subwords from A_1 and A_4 and contains $\boldsymbol{w}_1$ or $\boldsymbol{w}_2$.*

Proof Suppose $\boldsymbol{\ell}$ contains subwords from A_1 and A_4 but not from A_2 or A_3, and $\boldsymbol{\ell}$ contains neither d_1 nor d_{-1}. By Proposition 4.7, if $\boldsymbol{\ell}$ has a non-L–space standard filling, then either $\boldsymbol{\ell}$ contains no standard unstable chains or $\boldsymbol{\ell}$ contains both c_k and d_k unstable chains. In the former case $\boldsymbol{\ell}$ consists of alternating a_k and b_k segments; in the latter case, $\boldsymbol{\ell}$ must contain a sequence of d_0 segments preceded by an $a_{\pm 1}$ segment and followed by a $b_{\pm 1}$ segment (since a_k, b_k and d_k are in A_2 or A_3 for $|k| > 1$, and by assumption $\boldsymbol{\ell}$ also contains no $d_{\pm 1}$ segments). In either case, we see that the subscripts on the type a and b stable chains must have opposite signs (since $a_i d_0^n b_j$ is in A_3 and $a_{-i} d_0^n b_{-j}$ is in A_2), and thus that $\boldsymbol{\ell}$ contains $\boldsymbol{w}_1 = a_1 d_0^n \bar{b}_1$ or $\boldsymbol{w}_2 = \bar{a}_1 d_0^n b_1$.

Note then that $\mathrm{T}(\boldsymbol{w}_1) = a_1 d_1^n \bar{b}_1$, which gives a word in A_3; $\mathrm{T}^{-1}(\boldsymbol{w}_1) = a_1 \bar{c}_1^n \bar{b}_1$, which gives a word in A_2; $\mathrm{T}(\boldsymbol{w}_2) = \bar{a}_1 d_1^n b_1$, which gives a word in A_3; and $\mathrm{T}^{-1}(\boldsymbol{w}_2) = \bar{a}_1 \bar{c}_1^n b_1$, which gives a word in A_2. Since T preserves the property that there is a subword from A_4 (Lemma 4.11) and T^{-1} preserves the property that there is a subword from A_1 (Lemma 4.10), we have shown that conclusions (i) and (ii) hold.

Consider the subword $\boldsymbol{w}_1 = a_1 d_0^n \bar{b}_1$. If a loop $\boldsymbol{\ell}$ contains $\boldsymbol{w}_1$, then the graph for $\boldsymbol{\ell}$ contains the segment



It follows that in dual notation $\boldsymbol{\ell}$ contains a b^*_{n+1}. Moreover, this b^*_{n+1} is preceded by a segment ending with a ρ_3 arrow (that is, a segment of type $\bar{a}^*$ or $\bar{c}^*$) and followed by a segment beginning with a backwards ρ_{123} arrow (that is, a segment of type $\bar{a}^*$ or $\bar{d}^*$). Thus $\boldsymbol{\ell}$ contains $\boldsymbol{w}_1$ in standard notation if and only if it contains one of the following in dual notation: $\bar{a}^*_i b^*_{n+1} \bar{a}^*_j$, $\bar{a}^*_i b^*_{n+1} \bar{d}^*_j$, $\bar{c}^*_i b^*_{n+1} \bar{a}^*_j$ or $\bar{c}^*_i b^*_{n+1} \bar{d}^*_j$, where $i, j > 0$. Note that the set of such words is closed under the action of H, except for the last two words if the index i is 1 since then H replaces the $\bar{c}_1$ with $\bar{c}_0 = e = d_0$. Thus, if $\boldsymbol{\ell}$ contains one of these subwords, then $\mathrm{H}(\boldsymbol{\ell})$ also contains one of these subwords (and thus the subword $\boldsymbol{w}_1$ in standard notation) or it contains $d^*_0 b^*_{n+1}$. In the latter case, $\boldsymbol{\ell}$ contains a ρ_{23} arrow followed by a ρ_2 arrow. This means that $\boldsymbol{\ell}$ contains an a_k or d_k segment with $k > 1$, each of which are in A_3. Similarly, $\mathrm{H}^{-1}(\boldsymbol{\ell})$ either contains $\boldsymbol{w}_1$ in standard notation or it contains $b^*_{n+1} \bar{d}^*_0$, which implies that $\boldsymbol{\ell}$ contains a subword in A_2.

The details of a similar argument for the subword $\boldsymbol{w}_2 = \bar{a}_1 d_0^n \bar{b}_1$ are left to the reader. We similarly find that, if $\boldsymbol{\ell}$ contains $\boldsymbol{w}_2$, then $\mathrm{H}(\boldsymbol{\ell})$ contains either $\boldsymbol{w}_2$ or a subword in A_3 and $\mathrm{H}^{-1}(\boldsymbol{\ell})$ contains either $\boldsymbol{w}_2$ or a subword in A_2.

Since by assumption $\boldsymbol{\ell}$ contains subwords from A_1 and A_4, $\mathrm{H}(\boldsymbol{\ell})$ contains a subword from A_4 and from either A_1 or A_3 (Lemma 4.13). Since $\boldsymbol{\ell}$ also contains a subword $\boldsymbol{w}_1$ or $\boldsymbol{w}_2$, the discussion above implies that $\mathrm{H}(\boldsymbol{\ell})$ contains $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ or a subword in A_3; that is, (iii) is satisfied. Similarly, (iv) is satisfied since $\mathrm{H}^{-1}(\boldsymbol{\ell})$ contains a subword in A_2 and a subword in either A_1 or A_3 (Lemma 4.12), and a subword $\boldsymbol{w}_1$ or $\boldsymbol{w}_2$ if there is no subword from A_1. □

Lemma 4.15 *Suppose $\boldsymbol{\ell}$ contains no subwords in any A_i and that $\boldsymbol{\ell}$ contains both a d_k segment and a c_k segment. Then $\mathrm{T}^{-1}(\boldsymbol{\ell})$ contains a subword from each of A_1 and A_2 and $\mathrm{T}(\boldsymbol{\ell})$ contains a subword from each of A_3 and A_4.*

Proof The assumption that $\boldsymbol{\ell}$ contains no words from any of the sets A_i implies that in standard notation $\boldsymbol{\ell}$ consists of the segments $a_{\pm 1}$, $b_{\pm 1}$, $c_{\pm 1}$, $d_{\pm 1}$, c_0 and d_0, and the nonzero subscripts alternate between $+1$ and -1. By assumption, $\boldsymbol{\ell}$ must contain at least one d_1, d_0 or d_{-1} segment. If $\boldsymbol{\ell}$ contains d_1 then $\mathrm{T}(\boldsymbol{\ell})$ contains $d_2 \in A_3$. Moreover, d_1 must be preceded by d_0, d_{-1} or a_{-1} and followed by d_0, d_{-1} or b_{-1}. It follows that $\mathrm{T}^{-1}(\boldsymbol{\ell})$ contains a d_0 preceded by a d_{-1}, d_{-2} or a_{-1} and followed by a d_{-1}, d_{-2} or b_{-1}; thus $\mathrm{T}^{-1}(\boldsymbol{\ell})$ contains a subword in A_2. Similarly if $\boldsymbol{\ell}$ contains d_{-1} then $\mathrm{T}^{-1}(\boldsymbol{\ell})$ contains $d_{-2} \in A_2$, and, since d_{-1} must be preceded by d_0, d_1 or a_1 and followed by d_0, d_1 or b_1, $\mathrm{T}(\boldsymbol{\ell})$ contains a subword in A_3. Finally, suppose $\boldsymbol{\ell}$ contains a d_0 segment but no d_1 or d_{-1}. If the d_0 segment is adjacent to another d_0 segment,

then $\mathtt{T}(\boldsymbol{\ell})$ contains $d_1d_1 \in A_3$ and $\mathtt{T}^{-1}(\boldsymbol{\ell})$ contains $d_{-1}d_{-1} \in A_2$. Otherwise, $\boldsymbol{\ell}$ must contain either $a_{-1}d_0b_1$ or $a_1d_0b_{-1}$. It is easy to check that $\mathtt{T}(\boldsymbol{\ell})$ contains either d_1b_1 or a_1d_1, both elements of A_3, and that $\mathtt{T}^{-1}(\boldsymbol{\ell})$ contains either $a_{-1}d_{-1}$ or $d_{-1}b_{-1}$ in A_2.

A similar argument shows that, if $\boldsymbol{\ell}$ contains c_1, c_0 or c_{-1}, then $\mathtt{T}(\boldsymbol{\ell})$ contains a subword in A_4 and $\mathtt{T}^{-1}(\boldsymbol{\ell})$ contains a subword in A_1. This can also be deduced by reversing the orientation on $\boldsymbol{\ell}$, which takes d_k segments to c_{-k} segments and takes the sets A_2 and A_3 to A_1 and A_4. □

Proof of Proposition 4.6 Let $\boldsymbol{\ell}$ be a loop for which both standard and dual filling gives a non-L–space. According to Corollary 4.9, the fact that dual filling is a non-L–space implies that either $\boldsymbol{\ell}$ contains a word from A_1 or A_3 and from A_2 or A_4, $\boldsymbol{\ell}$ contains no words from any A_i, or $\boldsymbol{\ell}$ cannot be written in dual notation. In the last case, $\boldsymbol{\ell}$ consists only of type e segments and it is clear that standard filling is an L–space. Thus we have the following cases to consider:

(1) $\boldsymbol{\ell}$ contains a subword from A_1 and from A_2.

(2) $\boldsymbol{\ell}$ contains a subword from A_3 and from A_4.

(3) $\boldsymbol{\ell}$ contains a subword from A_1 and from A_4, but not A_2 or A_3.

(4) $\boldsymbol{\ell}$ does not contain a subword from any A_i.

Note that, a priori, there is another case similar to (3) in which $\boldsymbol{\ell}$ contains subwords in A_2 and A_3; however, this is equivalent to (3) after reversing orientation on $\boldsymbol{\ell}$.

In case (1), Lemma 4.10 implies that applying any combination of $\mathtt{T}^{-1}$ and $\mathtt{H}^{-1}$ to $\boldsymbol{\ell}$ gives rise to a loop $\boldsymbol{\ell}'$ which contains a subword in each of A_1 and A_2, and thus has non-L–space dual filling. The p/q filling on $\boldsymbol{\ell}$ is given by dual filling $\mathtt{T}^{a_n}\circ\cdots\circ\mathtt{H}^{a_2}\circ\mathtt{T}^{a_1}(\boldsymbol{\ell})$, where $[a_1,\dots,a_n]$ is an odd length continued fraction for p/q. For any $-\infty < p/q < 0$, we can choose a continued fraction with each $a_i \le 0$, and so p/q is not an L–space slope for $\boldsymbol{\ell}$. Similarly, in case (2) we choose a continued fraction with positive terms and use Lemma 4.11 to see that any $0 < p/q < \infty$ is a non-L–space.

Case (3) has three subcases, depending on whether $\boldsymbol{\ell}$ contains a $\bar{c}_1$, a d_1 or neither. If $\boldsymbol{\ell}$ contains a $\bar{c}_1$, then, by Lemmas 4.12 and 4.10, applying any combination of $\mathtt{T}^{-1}$ and $\mathtt{H}^{-1}$ to $\boldsymbol{\ell}$ with at least one application of $\mathtt{T}^{-1}$ produces a loop $\boldsymbol{\ell}'$ with a subword in A_1 and A_2. Given any $-\infty < p/q < 0$ we can choose an odd length continued fraction $[a_1,\dots,a_n]$ for which each a_i is strictly negative except that a_1 may be 0, and if $a_1 = 0$ then $n \ge 3$. It follows that p/q is a non-L–space slope for $\boldsymbol{\ell}$. Similarly, if $\boldsymbol{\ell}$ contains a d_1, then Lemmas 4.13 and 4.11 imply that any $0 < p/q < \infty$ is a

non-L–space slope. If ℓ contains neither a d_1 nor a $\bar{c}_1$, then any p/q is a non-L–space slope by Lemma 4.14.

In case (4), ℓ has no dual unstable chains. Since dual stable chains are fixed by $\mathrm{H}^{\pm 1}$, we have immediately that $\mathrm{H}(\ell) = \mathrm{H}^{-1}(\ell) = \ell$. Since ∞ is a non-L–space slope, ℓ must have either no standard unstable chains or unstable chains of both type c_k and type d_k. If ℓ has no standard unstable chains then $\mathrm{T}(\ell) = \mathrm{T}^{-1}(\ell) = \ell$; it follows that any combination of Dehn twists preserves ℓ, and so any slope p/q is a non-L–space slope (in fact in this case any filling of ℓ has trivial homology). If ℓ has standard unstable chains of both types, then Lemma 4.15 implies that any slope p/q is a non-L–space slope. □

4.3 Proof of Theorem 4.1

We are now ready to prove that the set of L–space slopes for a loop ℓ is a (possibly empty) interval in $\widehat{\mathbb{Q}}$. We will use the fact that applying T and H to a loop ℓ changes the set of L–space slopes in a controlled way: the slope p/q for ℓ is equivalent to the slope $(p-nq)/q$ for $\mathrm{T}^n(\ell)$ and the slope $p/(q-np)$ for $\mathrm{H}^n(\ell)$. These transformations preserve the cyclic ordering on $\widehat{\mathbb{Q}}$, and thus preserve the connectedness of the set of L–space slopes.

Remark 4.16 In particular, while Corollary 4.5 and Proposition 4.6 are stated in terms of the 0 and ∞ slopes, analogous statements hold for any two slopes of distance 1 since such pair of slopes can be taken to 0 and ∞ by a combination of $\mathrm{T}^{\pm 1}$ and $\mathrm{H}^{\pm 1}$. More precisely, for any two slopes of distance 1, if both are L–space slopes then one of the two intervals between them consists entirely of L–space slopes, and if both are non-L–space slopes then one of the two intervals between them consists entirely of non-L–space slopes.

In addition to Corollary 4.5 and Proposition 4.6, we will need the following lemma:

Lemma 4.17 *For any loop ℓ, there do not exist integers $n_1 < n_2 < n_3 < n_4$ such that n_1 and n_3 are L–space slopes and n_2 and n_4 are not L–space slopes, or vice versa.*

Proof An integer n is an L–space slope for ℓ if and only if 0 is an L–space slope for $\mathrm{T}^n(\ell)$. By Lemma 4.8, this happens when $\mathrm{T}^n(\ell)$ contains a word from the set $A_1 \cup A_3$ or from the set $A_2 \cup A_4$, but not both. Recall from the proof of Lemmas 4.10 and 4.11 that A_1 and A_2 are closed under T^{-1} and A_3 and A_4 are closed under T.

Suppose that $n_1 < n_2 < n_3$, where n_2 is an L–space slope for ℓ while n_1 and n_3 are non-L–space slopes; we will show that all $n < n_1$ and all $n > n_3$ are non-L–space

slopes. Up to reversing the loop, we can assume $\mathtt{T}^{n_2}(\boldsymbol{\ell})$ contains a word in $A_1 \cup A_3$ and not in $A_2 \cup A_4$. We consider the case that $\mathtt{T}^{n_2}(\boldsymbol{\ell})$ contains a word in A_1; the case that it contains a word in A_3 is similar and left to the reader.

By the closure property mentioned above (Lemma 4.10), $\mathtt{T}^{n_1}(\boldsymbol{\ell})$ contains a word in A_1. Since n_1 is a non-L–space slope, $\mathtt{T}^{n_1}(\boldsymbol{\ell})$ must contain a word from $A_2 \cup A_4$. It does not contain a word in A_4, since $\mathtt{T}^{n_2}(\boldsymbol{\ell})$ does not and the property of containing a word in A_4 is preserved by $\mathtt{T}$ (Lemma 4.11), and so it must contain a word in A_2. It follows that, for any $n < n_1$, $\mathtt{T}^{n}(\boldsymbol{\ell})$ contains words in A_1 and A_2 and thus n is a non-L–space slope for $\boldsymbol{\ell}$.

Since the property of containing a word in A_2 is preserved by $\mathtt{T}^{-1}$ (Lemma 4.10), we know that $\mathtt{T}^{n_3}(\boldsymbol{\ell})$ does not contain a word in A_2. We would like to show that $\mathtt{T}^{n_3}(\boldsymbol{\ell})$ contains a word in A_3, since this would imply it also contains a word in A_4; it would follow that for all $n > n_3$ the loop $\mathtt{T}^{n}(\boldsymbol{\ell})$ contains words from both A_3 and A_4 and thus n is a non-L–space slope for $\boldsymbol{\ell}$. Suppose, to the contrary, that $\mathtt{T}^{n_3}(\boldsymbol{\ell})$ does not contain a word in A_3. Then we see immediately that $\mathtt{T}^{n_2}$ does not contain a word in A_3. It also does not contain a word in A_2, so it must consist only of the segments

$$\{a_1, a_{-1}, b_1, b_{-1}, d_1, d_0, d_{-1}, c_k\},$$

where k can be any integer. Moreover, since $n_2 + 1 \leq n_3$, $\mathtt{T}^{n_2+1}(\boldsymbol{\ell})$ does not contain a word in A_3. In particular this implies that $\mathtt{T}^{n_2}(\boldsymbol{\ell})$ does not contain d_1. Any d_{-1} cannot be preceded by d_{-1} or a_{-1} or followed by a d_{-1} or b_{-1}, since this would give a word in A_2. The alternative, that d_{-1} is preceded by a d_0 or a_1 and followed by a d_0 or b_1, is also ruled out since this would produce a word in A_3 in $\mathtt{T}^{n_2+1}(\boldsymbol{\ell})$; thus $\mathtt{T}^{n_2}(\boldsymbol{\ell})$ does not contain d_{-1}. If $\mathtt{T}^{n_2}(\boldsymbol{\ell})$ contains d_0, then it contains either $a_1 d_0^m b_1$, $a_{-1} d_0^m b_1$, $a_1 d_0^m b_{-1}$ or $a_{-1} d_0^m b_1$. None of these are possible; the first two are words in A_3 and A_2, respectively, and the last two give rise to words in A_3 in $\mathtt{T}^{n_2+1}(\boldsymbol{\ell})$. We have now shown that $\mathtt{T}^{n_2}(\boldsymbol{\ell})$ does not contain any d_k segments, but, since it also does not contain $a_{-1}b_{-1} \in A_2$ this contradicts the fact that $\mathtt{T}^{n_1}(\boldsymbol{\ell})$ has a word in A_2. Therefore, $\mathtt{T}^{n_3}(\boldsymbol{\ell})$ contains a word in A_3 and a word in A_4, and n is a non-L–space slope for all $n > n_3$. □

Proof of Theorem 4.1 Any two distinct slopes r/s and p/q determine two intervals in $\widehat{\mathbb{Q}}$; we need to show that if r/s and p/q are L–space slopes then one of these intervals lies entirely in the set of L–space slopes for $\boldsymbol{\ell}$. By reparametrizing with $\mathtt{T}$ and $\mathtt{H}$ as discussed above, we may assume that $r/s = 0$. We will assume $p/q > 0$ below;

similar arguments apply when $p/q < 0$. If $p/q = \infty$, the result holds by Corollary 4.5, so we will assume $p/q < \infty$.

Suppose that 0 and p/q are L–space slopes and there exist non-L–space slopes $u \in (0, p/q)$ and $v \in [0, p/q]^c$. To reach a contradiction, we choose a continued fraction $[a_1, \dots, a_n]$ for p/q with $a_1 \geq 0$ and $a_i > 0$ for $i > 1$ and proceed by induction on the length n. In the base case of $n = 1$, $p/q = a_1$ is an integer. We claim that there is an integer $u' \in (0, a_1)$ that is a non-L–space slope. To see this, if u is not an integer, consider the slopes $\lfloor u \rfloor$ and $\lceil u \rceil$. These slopes are distance 1 and both intervals between them contain non-L–space slopes ($u \in (\lfloor u \rfloor, \lceil u \rceil)$ and $v \in [\lfloor u \rfloor, \lceil u \rceil]^c$). Thus, by Corollary 4.5 (taking into account Remark 4.16), at least one of them is a non-L–space slope. Similarly, if $v \neq \infty$ then either $\lfloor v \rfloor$ or $\lceil v \rceil$ is a non-L–space slope. If $v = \infty$ is a non-L–space slope, note that by Proposition 4.3 either $\boldsymbol{\ell}$ contains no unstable chains or both types of unstable chains in standard notation or $\boldsymbol{\ell}$ cannot be written in standard notation. If $\boldsymbol{\ell}$ cannot be written in standard notation (in which case it is a collection of e^* segments) or if $\boldsymbol{\ell}$ contains no unstable chains in standard notation, then $\boldsymbol{\ell}$ is fixed by $\mathrm{T}^{\pm 1}$. It follows that all integer slopes behave the same, but this contradicts the fact that 0 is an L–space slope and u' is not. Thus $\boldsymbol{\ell}$ must contain both types of unstable chains in standard notation. Applying T^m for m sufficiently large produces a loop where all unstable chains are of type d or $\bar{d}$ (with large index), and there is at least one of each type. It follows from Corollary 4.9 that m is a non-L–space slope, since $d_k \in A_3$ and $\bar{d}_k = c_{-k} \in A_4$. Thus we have integers $0 < u' < a_1$ and $v' < 0$ or $v' > a_1$ such that 0 and a_1 are L–space slopes and u' and v' are not; this contradicts Lemma 4.17.

Suppose the continued fraction $[a_1, \dots, a_n]$ for p/q has length $n > 1$ and $a_1 = 0$. We consider the loop $\boldsymbol{\ell}' = \mathrm{H}^{a_2}(\boldsymbol{\ell})$. Note that the slope p/q for $\boldsymbol{\ell}$ corresponds to the slope $p'/q' = p/(q - a_2 p)$ for $\boldsymbol{\ell}'$, which has continued fraction $[a_3, \dots, a_n]$. The slope 0 for $\boldsymbol{\ell}$ corresponds to 0 for $\boldsymbol{\ell}'$, and the slopes u and v for $\boldsymbol{\ell}$ correspond to slopes u' and v' for $\boldsymbol{\ell}'$. Thus $\boldsymbol{\ell}'$ has L–space slopes 0 and p'/q' and non-L–space slopes $u' \in (0, p'/q')$ and $v' \in (-\infty, 0) \cup (p'/q', \infty]$. Since p'/q' has a continued fraction of length $n - 2$, this produces a contradiction by induction.

Suppose p/q has continued fraction $[a_1, \dots, a_n]$ of length $n > 1$ with $a_1 > 0$. Consider the distance 1 slopes $a_1 = \lfloor p/q \rfloor$ and $a_1 + 1 = \lceil p/q \rceil$; both intervals between these slopes contain L–space slopes (either 0 or p/q), so by Proposition 4.6 one of them must be an L–space slope. First suppose that a_1 is an L–space slope. By the base case of induction, we cannot have both $u \in (0, a_1)$ and $v \in [0, p/q]^c \subset [0, a_1]^c$, and so we must have $u \in (a_1, p/q)$. Consider the loop $\boldsymbol{\ell}' = \mathrm{T}^{a_1}(\boldsymbol{\ell})$. The slopes 0 and

$p'/q' = p/q - a_1$ are L–space slopes for $\boldsymbol{\ell}'$ while the slopes $u' = u - a_2 \in (0, p'/q')$ and $v' = v - a_2 \in [0, p'/q']^c$ are non-L–space slopes. A continued fraction for p'/q' is $[0, a_2, \dots, a_n]$; as shown above, this produces a contradiction.

Finally, suppose instead that a_1 is a non-L–space slope for $\boldsymbol{\ell}$ and $a_1 + 1$ is an L–space slope. The base case of induction rules out the possibility that $v \in [0, a_1 + 1]^c$, so we must have $v \in (p/q, a_1 + 1)$. Consider the loop $\boldsymbol{\ell}' = \mathrm{H}^{a_2 - 1} \circ \mathrm{T}^{a_1+1}(\boldsymbol{\ell})$. The slope a_1 for $\boldsymbol{\ell}$ corresponds to the slope 0 for $\boldsymbol{\ell}'$, and the slope p/q corresponds to a slope p'/q' with continued fraction $[-1, 1, a_3, \dots, a_n] \sim [0, -a_3, \dots, -a_n]$. There are non-$L$–space slopes for $\boldsymbol{\ell}'$ in both intervals between 0 and p'/q'. By induction (using the analogue of the above cases when $p/q < 0$), this is a contradiction.

This proves that the set of L–space slopes for a loop $\boldsymbol{\ell}$ is an interval. To prove the statement for bordered manifolds (M, α, β) of loop-type, we simply observe that if M is a loop-type manifold, $M(p\alpha + q\beta)$ is an L–space if and only if p/q is an L–space slope for each loop $\boldsymbol{\ell}$ in $\widehat{CFD}(M, \alpha, \beta)$. The set of L–space slopes for M is the intersection of the intervals of L–space slopes for each loop, and hence an interval. □

4.4 Strict L–space slopes

The notion of L–space slope is fairly natural, however, in the context of the theorems in this paper, it is not quite the right condition. We will be interested primarily in slopes that are not only L–space slopes but are also surrounded by a neighbourhood of L–space slopes.

Definition 4.18 A slope in the boundary of a three-manifold M with torus boundary is a strict L–space slope if it is an L–space slope in the interior of $\mathcal{L}_M$. Denote the set of strict L–space slopes by $\mathcal{L}^\circ_M$.

We will give a geometric interpretation of strict L–space slopes in Section 7. For now, we will use loop notation and the results in this section about L–space slopes to determine when an L–space slope is strict.

Proposition 4.19 *Given a loop* $\boldsymbol{\ell}$:

(1) *The slope* ∞ *is a strict* L*–space slope for* $\boldsymbol{\ell}$ *if and only if* $\boldsymbol{\ell}$ *can be written in standard notation using only the subwords*

$$d_k, \quad b_1 a_{-1} \quad \text{and} \quad b_{-1} a_1,$$

where k *can be any integer, with at least one* d_k *and such that* $b_1 a_{-1}$ *is never adjacent to* $b_{-1} a_1$.

(2) *The slope 0 is a strict L–space slope for $\boldsymbol{\ell}$ if and only if $\boldsymbol{\ell}$ can be written in dual notation using only the subwords*

$$d_k^*, \quad b_1^* a_{-1}^* \quad \text{and} \quad b_{-1}^* a_1^*,$$

where k can be any integer, with at least one d_k^* and such that $b_1^* a_{-1}^*$ is never adjacent to $b_{-1}^* a_1^*$.

Proof Suppose ∞ is a strict L–space slope. In particular, it is an L–space slope, so, by Proposition 4.3, $\boldsymbol{\ell}$ can be written in standard notation with no c_k and at least one d_k. Since it does not contain c_k, $\boldsymbol{\ell}$ can be broken into pieces of the form d_k or $b_i a_j$ for any integer k and nonzero integers i and j, with at least one d_k. We also have that n and $-n$ are L–space slopes for sufficiently large n. Equivalently, 0 is an L–space slope after we apply T^n or T^{-n} for sufficiently large n. The twist T^n has the effect of replacing all unstable chains with chains of type d_k for $k \gg 0$, while the twist T^{-n} replaces all unstable chains with d_k for $k \ll 0$.

Consider a loop consisting of the pieces d_k with $k \gg 0$ and $b_i a_j$ with any nonzero integers i and j, with at least one d_k. Referring to the discussion of dualizing in Section 3.2, note that the loop certainly contains $d_0^* = e^*$, since it contains d_k with $k > 1$. Thus, by Proposition 4.3, 0 is an L–space slope if and only if the loop contains no c_k^* segments. The loop contains c_k^* with $k > 0$ if and only if it contains $a_i b_j$ with $i, j < 0$. It contains c_k^* with $k < 0$ if and only if it contains $b_i a_j$ with $i, j < 0$. It contains $c_0^* = \bar{e}^*$ segments if and only if it contains a_ℓ or b_ℓ with $\ell < -1$.

Similarly, consider a loop consisting of the pieces d_k with $k \ll 0$ and $b_i a_j$ with any nonzero integers i and j, with at least one d_k. The loop contains $c_0^* = \bar{e}^*$, so 0 is an L–space slope if and only if the loop contains no d_k^* segments. Moreover, the loop contains d_k^* if and only if it contains $b_i a_j$ or $a_i b_j$ with $i, j > 0$ or a_ℓ or b_ℓ with $\ell > 1$.

Therefore, given a loop for which ∞ is an L–space slope, n and $-n$ are also L–space slopes for all sufficiently large n if and only if the loop contains no chains of type a_k or b_k with $k > 1$, a_1 segments are never adjacent to b_1 segments, and a_{-1} segments are never adjacent to b_{-1} segments.

The proof of part (2) is completely analogous. Given that 0 is an L–space slope, we must check that $1/n$ and $-1/n$ are L–space slopes for sufficiently large n, which is equivalent to checking that ∞ is an L–space slope after applying H^n or H^{-n}. These twists have the effect of replacing unstable chains in dual notation with d_k^* segments with either $k \gg 0$ or $k \ll 0$. The rest of the proof is identical to the proof above after adding/removing stars on each segment. □

4.5 Simple loops

We will often restrict to a special class of loops, which we call simple.

Definition 4.20 A loop $\boldsymbol{\ell}$ is *simple* if there is a loop $\boldsymbol{\ell}'$ consisting only of unstable chains such that $\boldsymbol{\ell}$ can be obtained from $\boldsymbol{\ell}'$ using the operations $\mathrm{T}^{\pm 1}$ and $\mathrm{H}^{\pm 1}$. A collection of loops $\{\boldsymbol{\ell}_i\}_{i=1}^n$ is said to be simple if, possibly after applying a sequence of the operations $\mathrm{T}^{\pm 1}$ and $\mathrm{H}^{\pm 1}$, every loop consists only of unstable chains.

Remark 4.21 In the above definition, we do not specify whether the unstable chains are in standard or dual notation. In fact, it does not matter: if $\boldsymbol{\ell}$ contains only unstable chains in standard notation then $\mathrm{T}^n(\boldsymbol{\ell})$ contains only unstable chains in dual notation for sufficiently large n, and if $\boldsymbol{\ell}$ contains only dual unstable chains then $\mathrm{H}^n(\boldsymbol{\ell})$ contains only standard unstable chains. Definition 4.20 can be stated in terms of standard notation only (compare [10, Definition 4]), but it is sometimes convenient to check the condition in dual notation.

The notion of simple loops gives rise to a refinement of loop-type manifolds: M is said to be of *simple loop-type* if it is loop-type and, for some choice of α and β, the loops representing $\widehat{CFD}(M,\alpha,\beta)$ consist only of unstable chains. Equivalently, M is of simple loop-type if, for any choice of α and β, $\widehat{CFD}(M,\alpha,\beta)$ is represented by a simple collection of loops.

Note that solid torus-like loops (Definition 3.20) are examples of simple loops. Although we do not give an explicit description of all simple loops, we prove the following useful property:

Proposition 4.22 *If $\boldsymbol{\ell}$ is simple then, up to reorienting the loop, $\boldsymbol{\ell}$ has no a_k or b_k segments with $k<0$.*

The proof makes use of the following two observations.

Lemma 4.23 *If $\boldsymbol{\ell}$ contains no a_k, b_k or d_k segments with $k<0$, then the same is true for $\mathrm{T}(\boldsymbol{\ell})$ and $\mathrm{H}(\boldsymbol{\ell})$.*

Proof Recall that an a_k segment with $k<0$ corresponds to a segment of type $\bar{a}$, a b_k with $k<0$ corresponds to type $\bar{b}$, and a d_k with $k<0$ corresponds to type $\bar{c}$. The statement for $\mathrm{T}(\boldsymbol{\ell})$ is obvious, since T fixes a_k and b_k segments and increases the subscript on d_k segments by one. For the second, we observe that $\boldsymbol{\ell}$ contains no a_i^*, b_i^* or c_i^* segments in dual notation with $i>0$. Indeed, an a_i^* or a c_i^* contains a backwards

ρ_3 arrow, which in standard notation implies the presence of a segment of type $\bar{a}$ or $\bar{c}$ (equivalently, a segment of type a_k or d_k with $k<0$), and a b^* segment contains a forward ρ_1 segment, which implies the presence of type $\bar{b}$ or $\bar{c}$ segment (equivalently, of a b_k or d_k with $k<0$). It is clear that this property is preserved by H, which fixes a^* and b^* segments and decreases the index on c^* segments. Since $\mathrm{H}(\boldsymbol{\ell})$ has no a_i^*, b_i^* or c_i^* segments with $i>0$, it is straightforward to check that $\mathrm{H}(\boldsymbol{\ell})$ has no forward ρ_1 arrows or backward ρ_3 arrows, and thus it does not contain any a_k, b_k or d_k segments with $k<0$. □

Lemma 4.24 *If $\boldsymbol{\ell}$ contains no a_k, b_k or c_k segments with $k<0$, then the same is true for $\mathrm{T}^{-1}(\boldsymbol{\ell})$ and $\mathrm{H}^{-1}(\boldsymbol{\ell})$.*

Proof The proof is completely analogous to the previous lemma. □

Proof of Proposition 4.22 Since $\boldsymbol{\ell}$ is simple, there is some $\boldsymbol{\ell}'$ consisting of only unstable chains such that $\boldsymbol{\ell}$ is obtained from $\boldsymbol{\ell}'$ by a sequence of Dehn twists. That sequence of Dehn twists determines an element of the mapping class group, which can be represented by the matrix $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$. Let $[k_1,\dots,k_{2n}]$ be a continued fraction for p/q of even length; the element of the mapping class group above can be decomposed as the sequence of Dehn twists

$$\mathrm{T}^m \circ \mathrm{H}^{k_{2n}} \circ \mathrm{T}^{k_{2n-1}} \circ \cdots \circ \mathrm{H}^{k_2} \circ \mathrm{T}^{k_1},$$

where m is an integer. Let

$$\boldsymbol{\ell}'' = \mathrm{H}^{k_{2n}} \circ \mathrm{T}^{k_{2n-1}} \circ \cdots \circ \mathrm{H}^{k_2} \circ \mathrm{T}^{k_1}(\boldsymbol{\ell}).$$

First suppose that p/q is positive. We may assume that each k_i is positive for $1 \le i \le 2n$. Then $\boldsymbol{\ell}'$ can be written with no a_k, b_k or d_k segments with $k<0$, and, by Lemma 4.23, this property is closed under all positive twists, so $\boldsymbol{\ell}''$ can also be written with no a_k, b_k or d_k segments with $k<0$. It follows that $\boldsymbol{\ell} = \mathrm{T}^m(\boldsymbol{\ell}'')$ can be written with no a_k or b_k segments with $k<0$.

If instead p/q is negative, we may chose a continued fraction with each k_i negative. Then $\boldsymbol{\ell}'$ can be written with no a_k, b_k or c_k segments with $k<0$, and, by Lemma 4.24, the same is true for $\boldsymbol{\ell}''$. It follows that $\boldsymbol{\ell} = \mathrm{T}^m(\boldsymbol{\ell}'')$ can be written with no a_k or b_k segments with $k<0$. □

The argument above can be repeated in dual notation, taking a continued fraction for r/s instead of p/q, to prove the following:

Proposition 4.25 *If ℓ is simple then, up to reorienting the loop, ℓ has no a_k^* or b_k^* segments with $k < 0$.*

The main advantage of restricting to simple loops is that it greatly simplifies the conditions under which a given slope is a strict L–space slope. Recall that, by Proposition 4.19, ∞ is a strict L–space slope for ℓ if and only if ℓ can be decomposed into words of the form d_k, b_1a_{-1} or $b_{-1}a_1$. If ℓ is a simple loop, then the last two words cannot appear by Proposition 4.22, and ∞ is a strict L–space slope if and only if ℓ consists only of unstable chains in standard notation. Equivalently (using Observation 3.9), ∞ is a strict L–space slope if and only if ℓ does not contain both positive and negative subscripts in dual notation. Similarly, 0 is a strict L–space slope if and only if ℓ consists only of unstable chains in dual notation, which is equivalent to ℓ having only nonnegative or only nonpositive subscripts in standard notation.

We conclude this section by refining Theorem 4.1 in the case of simple loops. Theorem 4.1 states that set of L–space slopes for a loop ℓ is a (possibly empty) interval in $\widehat{\mathbb{Q}}$; the endpoints of this interval could be irrational, a priori, but we find that only rational endpoints are possible.

Proposition 4.26 *Let ℓ be a simple loop. The set of L–space slopes for ℓ is either*

(i) *identified with $\mathbb{Q}$ (in practice, every slope other than the rational longitude); or*

(ii) *the restriction to $\widehat{\mathbb{Q}}$ of a closed interval U in $\widehat{\mathbb{R}}$ with rational endpoints.*

Proof Note that applying the twist operations T and H to a loop preserves the cyclic ordering on abstract slopes, and so properties (i) and (ii) are preserved under these operations. Given an arbitrary simple loop ℓ, we will describe an algorithm for applying twists to produce a loop ℓ' with one of the following properties:

(1) ℓ' contains only d_0 segments;

(2) ℓ' contains only d_1, d_0 and d_{-1} segments, with d_1 and d_{-1} segments alternating (ignoring d_0 segments); or

(3) ℓ' contains only d_k segments with $k \geq -1$, including at least one d_{-1} segment, such that each d_{-1} is followed by $(d_0)^j d_k$ for some $j \geq 0$ and $k > 0$, and ℓ' does not satisfy (2).

By definition, some sequence of twists produces a loop ℓ_1 which can be written in standard notation using only type d_k segments. We may assume that the minimum subscript for the d_k segments in ℓ_1 is 0. Given ℓ_i (starting with $i = 1$), the algorithm

proceeds as follows: If the subscripts in $\boldsymbol{\ell}_i$ are all 0, then $\boldsymbol{\ell}_i$ satisfies (1) and the algorithm stops. Otherwise, consider the loop $\mathtt{T}^{-1}(\boldsymbol{\ell}_i)$; this loop consists of d_k segments with $k \geq -1$ including at least one d_{-1}. If $\mathtt{T}^{-1}(\boldsymbol{\ell}_1)$ satisfies (2) or (3), the algorithm stops. Otherwise, consider the loop $\mathtt{E}(\boldsymbol{\ell}_i)$, which by Lemma 3.10 contains only d_k^* segments with $k \leq 0$ in dual notation. It follows that $\mathtt{E}(\boldsymbol{\ell}_i)$ contains only c_k segments with $k \geq 0$ in standard notation; reversing the loop, we have that $\mathtt{E}(\boldsymbol{\ell}_i)$ can be written with only d_k segments with $k \leq 0$ in standard notation. Let m_i denote the minimum subscript for a d_k segment in $\mathtt{E}(\boldsymbol{\ell}_i)$, and define $\boldsymbol{\ell}_{i+1}$ to be $\mathtt{T}^{-m_i} \circ \mathtt{E}(\boldsymbol{\ell}_i)$. Note that $\boldsymbol{\ell}_{i+1}$ consists of d_k segments with $k \geq 0$, including at least one d_0. We now repeat the algorithm using $\boldsymbol{\ell}_{i+1}$.

To see that this algorithm terminates, let κ_i denote the number of d_0 segments in the loop $\boldsymbol{\ell}_i$. Observe that d_0 segments in $\boldsymbol{\ell}_{i+1}$ come from minimal subscripts in $\mathtt{E}(\boldsymbol{\ell}_i)$, which come from maximal sequences of d_0's in $\boldsymbol{\ell}_i$; in particular, there is at least one d_0 in $\boldsymbol{\ell}_i$ for each d_0 in $\boldsymbol{\ell}_{i+1}$, and so $\kappa_{i+1} \leq \kappa_i$. Moreover, the inequality is strict unless $\boldsymbol{\ell}_i$ has no consecutive d_0 segments. If $\mathtt{T}^{-1}(\boldsymbol{\ell}_i)$ satisfies (2) or (3) then the algorithm terminates; otherwise, $\boldsymbol{\ell}_i$ must contain $d_0(d_1)^j d_0$ for some $j \geq 0$. Let λ_i denote the minimal such j. As observed above, if $\lambda_i = 0$ then $\kappa_{i+1} < \kappa_i$. If $\lambda_i > 0$, then the subword $d_0(d_1)^{\lambda_i} d_0$ in $\boldsymbol{\ell}_i$ gives rise to the subword $d_{-2}(d_{-1})^{\lambda_i - 1} d_{-2}$ in $\mathtt{E}(\boldsymbol{\ell}_i)$ and the subword $d_0(d_1)^{\lambda_i - 1} d_0$ in $\boldsymbol{\ell}_{i+1}$. Thus, at each step in the algorithm, either $\kappa_{i+1} < \kappa_i$ or $\kappa_{i+1} = \kappa_i$ and $\lambda_{i+1} < \lambda_i$. Since κ_i and λ_i are nonnegative integers, the algorithm must terminate after finitely many steps.

Let $\boldsymbol{\ell}'$ be the result of the algorithm above. In cases (1) and (2), we can easily check that condition (i) is satisfied. First note that ∞ is an L–space slope for $\boldsymbol{\ell}'$ since it contains only unstable chains in standard notation. Moreover, for any $n \in \mathbb{Z}$, $1/n$ is an L–space slope for $\boldsymbol{\ell}'$ if and only if ∞ is an L–space slope for $\mathtt{H}^n(\boldsymbol{\ell}')$. But $\mathtt{H}^n(\boldsymbol{\ell}') = \boldsymbol{\ell}'$ since $\boldsymbol{\ell}'$ either cannot be written in dual notation (case (1)) or contains only stable chains in dual notation (case (2)). Since the set of L–space slopes is an interval and it contains slopes arbitrarily close to 0 on both sides, it must be all of $\widehat{\mathbb{Q}} \setminus \{0\}$. In case (3), note that writing $\boldsymbol{\ell}'$ in dual notation produces stable chains, but each b^* segment is immediately followed by an a^* segment. Moreover, since $\boldsymbol{\ell}'$ does not satisfy (2), it has at least one unstable chain in dual notation. It follows from Propositions 4.3 and 4.19 that 0 is an L–space slope for $\boldsymbol{\ell}'$ but not a strict L–space slope. Thus 0 must be a boundary of the interval of L–space slopes. Since $\boldsymbol{\ell}'$ was obtained from $\boldsymbol{\ell}$ by applying a finite number of twists, the slope 0 for $\boldsymbol{\ell}'$ can be expressed as a rational slope for $\boldsymbol{\ell}$.

In case (3), we have found one rational boundary of the interval of L–space slopes. In fact, we can check that it is the left boundary. In this case $\chi_\circ(\boldsymbol{\ell}')$ and $\chi_\bullet(\boldsymbol{\ell}')$ have the

same sign, and so the slope of the rational longitude $-\chi_\circ/\chi_\bullet$ is negative. By Corollary 4.5 it follows that the set of L–space slopes for $\boldsymbol{\ell}'$ contains $[0,\infty]$, and so 0 must be the left boundary of the interval of L–space slopes for $\boldsymbol{\ell}'$. A similar algorithm, with opposite signs for subscripts in property (3), shows that the right endpoint is also rational. □

The proof of Theorem 1.2 is now complete: it follows from (and is made more precise by) Theorem 4.1 in combination with Proposition 4.26

5 Gluing results

This section is devoted to proving Theorem 1.3. We first prove the analogous result on the level of abstract loops, and then deduce the gluing theorem for simple loop-type manifolds. We end the section with an application to generalized splicing of L–space knot complements and give the proof of Theorem 1.6.

5.1 A gluing result for abstract loops

We will say that two loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are *L–space aligned* if, for every slope $r/s \in \widehat{\mathbb{Q}}$, either r/s is a strict L–space slope for $\boldsymbol{\ell}_1$ or s/r is a strict L–space slope for $\boldsymbol{\ell}_2$. This section is devoted to proving the following proposition:

Proposition 5.1 *If $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are simple loops which are not solid-torus-like, then $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ is an L–space chain complex if and only if $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are L–space aligned.*

An essential observation is that $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are L–space aligned if and only if $\mathrm{T}(\boldsymbol{\ell}_1)$ and $\mathrm{H}(\boldsymbol{\ell}_2)$ are L–space aligned, since T takes the slope r/s to $(r+s)/s$ while H takes the slope s/r to $s/(r+s)$. More generally, we can apply a sequence of twists $\{\mathrm{T}^{k_1}, \mathrm{H}^{k_2}, \dots, \mathrm{T}^{k_{2n-1}}, \mathrm{H}^{k_{2n}}\}$ to $\boldsymbol{\ell}_1$ and a corresponding sequence of twists $\{\mathrm{H}^{k_1}, \mathrm{T}^{k_2}, \dots, \mathrm{H}^{k_{2n-1}}, \mathrm{T}^{k_{2n}}\}$ to $\boldsymbol{\ell}_2$ without changing whether or not the pair of loops is L–space aligned. By Proposition 2.6, the quasi-isomorphism type of $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ is also unchanged. Thus, in proving Proposition 5.1, we may first reparametrize the pair of loops to get a more convenient form.

Proof of Proposition 5.1, "only if" direction Suppose $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are not L–space aligned; that is, there exists a slope p/q that is not a strict L–space slope for $\boldsymbol{\ell}_1$ and q/p is not a strict L–space slope for $\boldsymbol{\ell}_2$. In fact, we may assume that $p/q = \infty$; if not, we reparametrize as described above, replacing $\boldsymbol{\ell}_1$ with

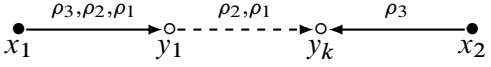
$$\boldsymbol{\ell}_1' = \mathrm{H}^{k_{2n}} \circ \mathrm{T}^{k_{2n-1}} \circ \cdots \circ \mathrm{H}^{k_2} \circ \mathrm{T}^{k_1}(\boldsymbol{\ell}_1)$$

so that the slope p/q for $\boldsymbol{\ell}_1$ becomes the slope ∞ for $\boldsymbol{\ell}_1'$, and replacing $\boldsymbol{\ell}_2$ with

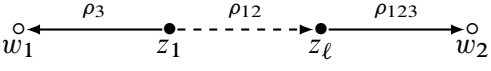
$$\boldsymbol{\ell}_2' = \mathrm{T}^{k_{2n}} \circ \mathrm{H}^{k_{2n-1}} \circ \cdots \circ \mathrm{T}^{k_2} \circ \mathrm{H}^{k_1}(\boldsymbol{\ell}_2).$$

Furthermore, we may assume that $\boldsymbol{\ell}_1$ contains no segments of type d_k, $\bar{d}_k$, e or $\bar{e}$, and that $\boldsymbol{\ell}_2$ contains no segments of type d_k^*, $\bar{d}_k^*$, e^* or $\bar{e}^*$; if necessary, we replace $\boldsymbol{\ell}_1$ with $\mathrm{T}^{-n}(\boldsymbol{\ell}_1)$ and $\boldsymbol{\ell}_2$ with $\mathrm{H}^{-n}(\boldsymbol{\ell}_2)$ for sufficiently large n.

Since ∞ is not a strict L–space slope for $\boldsymbol{\ell}_1$, the loop $\boldsymbol{\ell}_1$ must contain stable chains in standard notation (note that, since $\boldsymbol{\ell}_1$ is not solid torus-like, it can be written in standard notation). In particular, after possibly reversing the loop, $\boldsymbol{\ell}_1$ contains a b_k segment. The corresponding segment in $\boldsymbol{\ell}_1^A$ is



Since 0 is not a strict L–space slope for $\boldsymbol{\ell}_2$, this loop must contain stable chains in dual notation. In particular it contains an a_ℓ^* segment; we label the corresponding generators



Consider the generator y_k in $\boldsymbol{\ell}_1^A$, which has no outgoing $\mathcal{A}_\infty$ operations. To determine the possible incoming operations, note that the segment b_k must be followed by either a type c segment or a type a segment. This is because we assumed that $\boldsymbol{\ell}_1$ contains no $\bar{e}$ or $\bar{d}_j$ segments, and $\boldsymbol{\ell}_1$ cannot contain both b_k and $\bar{a}_j$ segments since it is simple. In either case, x_2 has an outgoing ρ_3 labelled arrow in $\boldsymbol{\ell}_1^A$ and no incoming arrows. It follows that y_k has only the incoming $\mathcal{A}_\infty$ operations

$$\begin{aligned} m_2(x_2, \rho_3) &= y_k, \\ m_{2+i}(y_{k-i}, \rho_2, \rho_{12}, \ldots, \rho_{12}, \rho_1) &= y_k \quad \text{for } 1 \le i < k, \\ m_{3+k}(x_1, \rho_3, \rho_2, \rho_{12}, \ldots, \rho_{12}, \rho_1) &= y_k, \end{aligned}$$

and possibly more operations whose inputs end with $\rho_2, \rho_{12}, \ldots, \rho_{12}, \rho_1$.

Consider the generator w_2 in $\boldsymbol{\ell}_2$. Note that a_ℓ^* must be followed by either a b^* segment or a $\bar{c}^*$ segment. It follows that the only incoming sequences of arrows consist of a ρ_{123} or ρ_1 arrow preceded by some number of ρ_{12} arrows. Comparing this with the $\mathcal{A}_\infty$ operations terminating at y_k described above, it is clear that the generator $y_k \otimes w_2$ in $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ has no incoming differentials. It also has no outgoing differentials, since y_k has no outgoing $\mathcal{A}_\infty$ operations. Thus $y_k \otimes w_2$ survives in homology.

Similarly, consider the generator x_1 in $\boldsymbol{\ell}_1^A$ and z_1 in $\boldsymbol{\ell}_2$. The generator z_1 has no incoming sequences of arrows, and the only outgoing sequences consist of a single ρ_3 arrow or begin with some number of ρ_{12} arrows followed by a ρ_{123}. Here we use the fact that the segment a_ℓ^* in the simple loop $\boldsymbol{\ell}_2$ can only be preceded by a b^* segment or a c^* segment, so the outgoing ρ_3 arrow cannot be followed by another outgoing arrow. In $\boldsymbol{\ell}_1^A$, the segment b_k must be preceded by an a segment or a $\bar{c}$ segment. It follows that, for any nontrivial operation $m_{n+1}(x_1, \rho_{I_1}, \dots, \rho_{I_n})$, we have

- $\rho_{I_1} \neq \rho_{123}$;
- if $\rho_{I_1} = \rho_{12}$, then $\rho_{I_i} = \rho_{12}$ for $1 \leq i \leq n-1$ and $\rho_{I_n} = \rho_1$;
- if $\rho_{I_1} = \rho_3$, then $n > 1$.

We see that no $\mathcal{A}_\infty$ operations starting at x_1 match with the δ^n maps starting at z_1. Thus the generator $x_1 \otimes z_1$ in $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ has no incoming or outgoing differentials and survives in homology.

Finally, we observe that $\mathrm{gr}(z_1) = \mathrm{gr}(w_2)$ since z_1 and w_2 are connected by only ρ_{12} and ρ_{123} arrows, but $\mathrm{gr}(y_k) = \mathrm{gr}(y_1) = -\mathrm{gr}(x_1)$, since arrows labelled (ρ_2, ρ_1) preserve grading but arrows labelled (ρ_3, ρ_2, ρ_1) flip grading. It follows that $y_k \otimes w_2$ and $x_1 \otimes z_1$ have opposite $(\mathbb{Z}/2\mathbb{Z})$–grading. Since both survive in homology, $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ is not an L–space complex. □

To prove the converse we will use the fact that $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are L–space aligned to put strong restrictions on the segments that may appear in the loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$. Once again, we can apply twists to $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ to obtain $\boldsymbol{\ell}_1'$ and $\boldsymbol{\ell}_2'$ with convenient parametrizations, such that $\boldsymbol{\ell}_1'$ and $\boldsymbol{\ell}_2'$ are still L–space aligned and $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ is homotopy equivalent to $(\boldsymbol{\ell}_1')^A \boxtimes \boldsymbol{\ell}_2'$. The set of strict L–space slopes for $\boldsymbol{\ell}_2$ is some nonempty open interval in $\widehat{\mathbb{Q}}$. This interval is either all of $\widehat{\mathbb{Q}}$ except the rational longitude or it has distinct rational endpoints; see Proposition 4.26. In the latter case, we can reparametrize so that for $\boldsymbol{\ell}_2'$ these boundaries have slope 0 and p/q for some $1 < p/q \leq \infty$. To see this, take $p > 0$ and choose n so that $0 \leq q + np < p$ and $1 < p/(q+np) \leq \infty$; we can replace p/q with $p/(q+np)$ by applying dual twists, in particular, leaving the slope 0 fixed. Now the set of strict L–space slopes for $\boldsymbol{\ell}_2'$ is exactly $(0, p/q)$. The fact that $\boldsymbol{\ell}_1'$ and $\boldsymbol{\ell}_2'$ are L–space aligned then implies that the set of strict L–space slopes for $\boldsymbol{\ell}_1'$ contains $[-\infty, q/p]$. If the set of strict L–space slopes for $\boldsymbol{\ell}_2$ is all of $\widehat{\mathbb{Q}}$ except the rational longitude, then we can choose a parametrization such that the rational longitude of $\boldsymbol{\ell}_2'$ is 0 and such that the set of strict L–space slopes for $\boldsymbol{\ell}_1'$ contains $[-\infty, 0]$.

Lemma 5.2 *If $q/p \in [0,1)$ and ℓ is a simple loop for which the interval of strict L–space slopes contains $[-\infty, q/p]$, then ℓ consists only of segments c_k^* with $0 \leq k \leq \lceil p/q \rceil$.*

Proof Since 0 is a strict L–space slope, ℓ can be written with only dual unstable chains. Up to reading the loop in reverse order, we can assume the unstable chains are c_k^* segments. Moreover, the fact that ∞ is a strict L–space slope implies that ℓ cannot contain c_k^* segments with both positive and negative subscripts. Since the rational longitude is given by $-\chi_\circ(\ell)/\chi_\bullet(\ell)$ and falls in the interval $(q/p, \infty)$, we must have that $\chi_\circ(\ell)$ and $\chi_\bullet(\ell)$ have opposite signs. This only happens if ℓ is composed of c_k^* segments with $k \geq 0$. Let $n = \lceil p/q \rceil$. Observe that ℓ must contain at least one c_k^* with $0 \leq k < n$ (recall that $c_0 = \bar{e}^*$), since otherwise the rational longitude is less than $1/n$. Finally, the fact that $1/n$ is a strict L–space slope implies that ∞ is a strict L–space slope for the loop $\mathrm{H}^n(\ell)$. Since ℓ contains $\bar{e}^*$ or c_k^* with $k < n$, $\mathrm{H}^n(\ell)$ contains at least one c_k^* with $k < 0$ and therefore does not contain any c_k^* with $k > 0$. Therefore ℓ does not contain c_k^* with $k > n$. □

Lemma 5.3 *If $p/q \in (1, \infty]$ and ℓ is a simple loop that is not solid torus-like for which the interval of strict L–space slopes contains $(0, p/q)$, then ℓ consists only of a_k, b_k, c_k and d_k segments (for $k > 0$) and e and $\bar{c}_1$ segments. Moreover,*

- *ℓ contains no two $\bar{c}_1 = d_{-1}$ segments separated only by $e = d_0$ segments,*
- *ℓ contains no c_k segments with $k < \lceil p/q \rceil - 1$, and*
- *if 0 is not a strict L–space slope for ℓ then there is at least one $\bar{c}_1$ segment.*

Proof Since 1 is a strict L–space slope, $\mathrm{T}(\ell)$ can be written with no $\bar{a}$, $\bar{b}$, $\bar{c}$ or $\bar{d}$ segments. It follows that ℓ can be written with no $\bar{a}$, $\bar{b}$, $\bar{d}$ or $\bar{e}$ segments and no $\bar{c}_k$ segments with $k > 1$.

Let $n = \lceil p/q \rceil$. Since $n - 1 < p/q$ is a strict L–space slope, $\mathrm{T}^{n-1}(\ell)$ does not contain both barred and unbarred segments (ignoring e's). Suppose ℓ contains c_k with $k < n - 1$. Then $\mathrm{T}^{n-1}(\ell)$ contains at least one $\bar{d}$ segment and cannot contain any unbarred segments. Any a, b or c_k segments with $k > n - 1$ in ℓ produces an unbarred segment in $\mathrm{T}^{n-1}(\ell)$, so we must have that ℓ consists only of c_k segments with $k \leq n - 1$. However, in this case it is easy to see that $\chi_\bullet(\ell)$ and $\chi_\circ(\ell)$ have opposite signs and $|\chi_\circ(\ell)| \leq (n-1)|\chi_\bullet(\ell)|$, which contradicts the fact that the rational longitude $-\chi_\circ/\chi_\bullet$ does not fall in the interval $(0, p/q)$. Thus ℓ does not contain c_k with $k < n - 1$.

Now $\boldsymbol{\ell}$ must contain an a_k, b_k, c_k or d_k segment with $k > 1$ or two segments of type a_1, b_1, c_1 or d_1 separated only by e's. Otherwise, $\boldsymbol{\ell}$ would consist of only $\bar{c}_1$, e and d_1 segments with at least one $\bar{c}_1$ for each d_1; in this case, $\chi_\bullet(\boldsymbol{\ell})$ and $\chi_\circ(\boldsymbol{\ell})$ have opposite signs and $|\chi_\circ(\boldsymbol{\ell})| \leq |\chi_\bullet(\boldsymbol{\ell})|$, so the rational longitude falls in $(0, 1]$. It follows that in dual notation $\boldsymbol{\ell}$ contains $\bar{c}_k^*$, e^* or d_k^*, and thus $\mathrm{H}^m(\boldsymbol{\ell})$ contains a d^* segment for sufficiently large m. Since $1/m$ is a strict L–space slope for sufficiently large m, we must have that $\mathrm{H}^m(\boldsymbol{\ell})$ does not contain a $\bar{d}^*$ segment. Therefore $\boldsymbol{\ell}$ does not contain any c_k^* segments, and thus in standard notation it does not have two $\bar{c}_1$ segments separated only by e's.

Finally, if 0 is not a strict L–space slope for $\boldsymbol{\ell}$, then $\boldsymbol{\ell}$ must contain both barred and unbarred segments in standard notation; it follows that $\boldsymbol{\ell}$ must contain at least one $\bar{c}_1$. □

The two previous lemmas only depend on $\lceil p/q \rceil$. If p/q is not an integer, it is possible to give further restrictions on subwords that can appear in the loop. We will prove one such restriction using two properties for a pair of loops. For an integer $r \geq 0$, we will say that two loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ satisfy property λ (or property λ for $r \geq 0$) if:

- (λ1) $\boldsymbol{\ell}_1$ consists only of c_k^* with $0 \leq k \leq n$, for some n, with at least one c_n^*.
- (λ2) $\boldsymbol{\ell}_2$ consists only of a_k, b_k ($k > 0$), d_k ($k \geq -1$) and c_l ($l \geq m$) segments for some $m > 0$, with at least one c_m and at least one $\bar{c}_1 = d_{-1}$.
- (λ3) There is an integer $N > 0$ and subscripts $k_i \in \{N, N-1\}$ for $1 \leq i \leq r$ such that $\boldsymbol{\ell}_1$ contains the subword $c_N^* c_{k_1}^* \dots c_{k_r}^* c_N^*$ and $\boldsymbol{\ell}_2$ contains the subword $c_{N-1} c_{k_1} \dots c_{k_r} c_{N-1}$.

The integer $r \geq 0$ appearing in condition (λ3) is the complexity of the pair $(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ satisfying property λ; when we appeal to pairs satisfying property λ the aim will be to decrease this complexity. Similarly, two loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ satisfy property property λ^* (or property λ^* for $r \geq 0$) if:

- (λ^*1) $\boldsymbol{\ell}_1$ consists only of $\bar{d}_k$ with $0 \leq k \leq n$, for some n, with at least one $\bar{d}_n$.
- (λ^*2) $\boldsymbol{\ell}_2$ consists only of $\bar{a}_k^*$, $\bar{b}_k^*$ (for $k > 0$), $\bar{c}_k^*$ (for $k > -1$) and $\bar{d}_l^*$ (for $l \geq m$) segments for some m, with at least one $\bar{d}_m^*$ and at least one $d_1^* = \bar{c}_{-1}^*$.
- (λ^*3) There is an integer $N > 0$ and subscripts $k_i \in \{N, N-1\}$ for $1 \leq k \leq r$ such that $\boldsymbol{\ell}_1$ contains the subword $\bar{d}_N \bar{d}_{k_1} \dots \bar{d}_{k_r} \bar{d}_N$ and $\boldsymbol{\ell}_2$ contains the subword $\bar{d}_{N-1}^* \bar{d}_{k_1}^* \dots \bar{d}_{k_r}^* \bar{d}_{N-1}^*$.

Lemma 5.4 *If two simple loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ satisfy either property λ or property λ^*, then $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are not L–space aligned.*

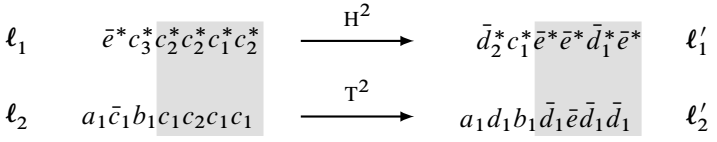


Figure 6: Loops ℓ_1 and ℓ_2 satisfying property λ for $n = 3$ and $m = 1$. The relevant subwords, with $N = r = 2$, have been highlighted. Note that this illustrates the key step in the proof of Lemma 5.4 when $n > m+1$: one checks that 0 is not a strict L–space slope for ℓ_1' (this loop contains a $\bar{d}^*$ segment) while the set of strict L–space slopes for ℓ_2' does not intersect $[0, \infty]$ thus ℓ_1' and ℓ_2' (and hence ℓ_1 and ℓ_2) are not L–space aligned.

Proof Suppose first that property λ is satisfied, where n is the maximum subscript for c_k^* segments in ℓ_1 and m is the minimum subscript of c_k segments in ℓ_2. Since ℓ_1 contains c_N^* and ℓ_2 contains c_{N-1} for some N, we have that $n \geq m+1$. Consider the reparametrized loops $\ell_1' = H^{m+1}(\ell_1)$ and $\ell_2' = T^{m+1}(\ell_2)$; ℓ_1' and ℓ_2' are L–space aligned if and only if ℓ_1 and ℓ_2 are. We now have:

- ℓ_1' consists only of c_k^* with $-m \leq k \leq n-m-1$ with at least one c_{n-m-1}^* (where, as usual, $c_0^* = \bar{e}^*$).
- ℓ_2' consists only of a_k, b_k (for $k > 0$), c_k (for $k \geq -1$) and d_l segments with $l \geq m$, with at least one $c_{-1} = \bar{d}_1$ and at least one d_m.

Note that, for ℓ_2', 0 is not a strict L–space slope because the loop contains at least one barred segment, $\bar{d}_1$, and at least one unbarred segment, d_m. Note also that ∞ is not a strict L–space slope, since ℓ_2' contains unstable chains with both orientations. The slope -1 is a strict L–space slope, since $T^{-1}(\ell_2')$ has no barred segments (ignoring e's). It follows that ℓ_2' has no strict L–space slopes in $[0, \infty]$.

First consider the case that $n > m+1$. In this case, ℓ_1' contains at least one c_k^* segment with $k > 0$. If ℓ_1' also contains a c_k^* segment with $k < 0$, then 0 is not a strict L–space slope for ℓ_1'. Since ∞ is not a strict L–space slope for ℓ_2', ℓ_1' and ℓ_2' are not L–space aligned. If ℓ_1' does not contain a c_k^* segment with $k < 0$, then it consists only of c_k^* segments with $k \geq 0$. It follows that the rational longitude $-\chi_\circ(\ell_1')/\chi_\bullet(\ell_1')$ is positive. Since the rational longitude is not a strict L–space slope, and all positive slopes for ℓ_2' are not strict L–space slopes, ℓ_1' and ℓ_2' are not L–space aligned.

In the case that $n = m+1$, we have $N = n$ in the statement of property λ (with complexity r) for ℓ_1 and ℓ_2. The shifted loops ℓ_1' and ℓ_2' satisfy an additional condition:

- There are subscripts $k_i' \in \{0,1\}$ for $1 \le i \le r$ such that $\boldsymbol{\ell}_1'$ contains the subword $\bar{e}^* \bar{d}_{k_1'}^* \dots \bar{d}_{k_r'}^* \bar{e}^*$ and $\boldsymbol{\ell}_2'$ contains the subword $\bar{d}_1 \bar{d}_{k_1'} \dots \bar{d}_{k_r'} \bar{d}_1$, following (as usual) the convention that $\bar{d}_0 = \bar{e}$ and $\bar{d}_0^* = \bar{e}^*$.

The next step is to write $\boldsymbol{\ell}_1'$ in standard notation and $\boldsymbol{\ell}_2'$ in dual notation. First, $\boldsymbol{\ell}_1'$ consists only of $\bar{e}^*$ and $\bar{d}^*$ segments, so in standard notation it consists only of $\bar{e}$ and $\bar{d}$ segments. There is some maximum subscript on the $\bar{d}$ segments; call it n'. Note that this says that $\boldsymbol{\ell}_1'$ and $\boldsymbol{\ell}_2'$ satisfy condition $(\lambda^* 1)$ for property λ^*. Dualizing $\boldsymbol{\ell}_2'$ is slightly harder. Types of segments in dual notation are determined by adjacent pairs of standard segments (ignoring e's and $\bar{e}$'s); see Section 3.2. The possible pairs of segments in $\boldsymbol{\ell}_2'$ are ab, ad, dd, db, ba, bc, $b\bar{d}$, ca, cc, $c\bar{d}$, $\bar{d}a$, $\bar{d}c$ and $\bar{d}\bar{d}$. These correspond to dual segments d^*, d^*, d^*, d^*, $\bar{c}^*$, $\bar{c}^*$, $\bar{b}^*$, $\bar{c}^*$, $\bar{c}^*$, $\bar{b}^*$, $\bar{a}^*$, $\bar{a}^*$ and $\bar{d}^*$, respectively. Since $\boldsymbol{\ell}_2'$ has no e segments, the subscripts on the d^* segments are at most 1. Since the only bar segments in $\boldsymbol{\ell}_2'$ in standard notation have subscript 1, there are no $\bar{e}^*$ segments in $\boldsymbol{\ell}_2'$. Thus $\boldsymbol{\ell}_2'$ consists only of $\bar{a}_k^*$, $\bar{b}_k^*$, $\bar{c}_k^*$, $\bar{d}_k^*$ (for $k > 0$), e^* and d_1^* segments. There is some minimum subscript on the $\bar{d}$ segments; call it m'. Moreover, since $\boldsymbol{\ell}_2'$ contains at least one d_k segment, it also contains at least one d_1^* segment. Note that this says that $\boldsymbol{\ell}_1'$ and $\boldsymbol{\ell}_2'$ satisfy condition $(\lambda^* 2)$ of property λ^*.

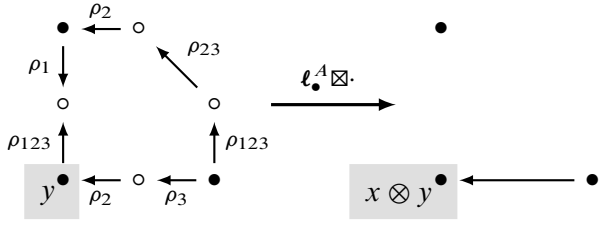
If $n' > m' + 1$, then we proceed in a similar fashion to the $n > m + 1$ case for property λ. We replace $\boldsymbol{\ell}_1'$ with $\boldsymbol{\ell}_1'' = \mathrm{T}^{-m'-1}(\boldsymbol{\ell}_1')$ and $\boldsymbol{\ell}_2'$ with $\boldsymbol{\ell}_2'' = \mathrm{H}^{-m'-1}(\boldsymbol{\ell}_2')$. We can observe that $\boldsymbol{\ell}_2''$ contains at least one c_1^* segment and at least one $\bar{c}_{m'}$ segment; thus 0 and ∞ are not strict L–space slopes for $\boldsymbol{\ell}_2''$, and in fact no slope in $[-\infty, 0]$ is a strict L–space slope. We can also observe that $\boldsymbol{\ell}_1''$ consists of c, $\bar{e}$ and $\bar{d}$ segments with at least one $\bar{d}$. Thus either 0 is not a strict L–space slope or the rational longitude is negative. In either case, $\boldsymbol{\ell}_1''$ and $\boldsymbol{\ell}_2''$ are not L–space aligned.

Now assume that $n' = m' + 1$. Consider the sequence $k_1', \dots, k_r'$. This sequence consists of some number of (possibly empty) strings of 0's each separated by a single 1; let $l_1, \dots, l_s$ be the sequence of the lengths of these strings of 0's. Note that s is at most $r + 1$. Since $\boldsymbol{\ell}_2'$ contains the word $\bar{d}_1 \bar{d}_{k_1'} \dots \bar{d}_{k_r'} \bar{d}_1$, it follows that it contains the dual word $\bar{d}_{l_1}^* \bar{d}_{l_2}^* \dots \bar{d}_{l_{s-1}}^* \bar{d}_{l_s}^*$. Similarly, $\boldsymbol{\ell}_1'$ contains the dual word $\bar{e}^* \bar{d}_{k_1'}^* \dots \bar{d}_{k_r'}^* \bar{e}^*$, which must be followed and proceeded by more $\bar{d}^*$ segments, possibly with additional $\bar{e}^*$ segments in between. It follows that $\boldsymbol{\ell}_1'$ contains the word $\bar{d}_{l_1'} \bar{d}_{l_2} \dots \bar{d}_{l_{s-1}} \bar{d}_{l_s'}$, where $l_1' > l_1$ and $l_s' > l_s$. Since we have assumed that $n' = m' + 1$, we must have that $l_1' = l_1 + 1$ and $l_s' = l_s + 1$, and moreover that $l_1 = l_s$ and $l_i \in \{l_1, l_1 + 1\}$ for every other i. Let $N' = l_1 + 1$, so that $\boldsymbol{\ell}_1'$ contains the subword $\bar{d}_{N'} \bar{d}_{l_2} \dots \bar{d}_{l_{s-1}} \bar{d}_{N'}$ and $\boldsymbol{\ell}_2'$

contains the subword $\bar{d}^*_{N'-1}\bar{d}^*_{l_1}\dots\bar{d}^*_{l_{s-1}}\bar{d}^*_{N'-1}$. In other words, the property (λ^*3) of property λ^* is satisfied with complexity $r'=s-2$. Note that $r'<r$, since $s\le r+1$. We also have that $r'\ge 0$, since if $s=1$ then each k'_i is 0 for each $1\le i\le r$; it would follow that $\boldsymbol{\ell}'_2$ contains $\bar{d}^*_{r+1}$ and $\boldsymbol{\ell}'_1$ contains a $\bar{d}$ segment with subscript at least $r+3$, and so $n'>m'+1$.

We have shown that if $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ satisfy property λ (for some integer $r\ge 0$), then either they are not L–space aligned or they can be modified to $\boldsymbol{\ell}'_1$ and $\boldsymbol{\ell}'_2$ which satisfy property λ^* (for some integer $r'\ge 0$), where $0\le r'<r$. A similar proof shows that if $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ satisfy property λ^* then either they are not L–space aligned or they can be modified to $\boldsymbol{\ell}'_1$ and $\boldsymbol{\ell}'_2$ which satisfy property λ with $0\le r'<r$. In both cases, the complexity is reduced, so, by induction on r, we have that $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are not L–space aligned if they satisfy either property. □

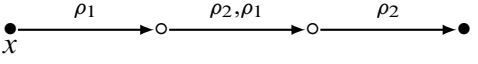
Now that we have placed restrictions on loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ that are L–space aligned, we can complete the proof of Proposition 5.1 by analyzing the box tensor product segment by segment. In the proof, we will determine when certain generators in $\boldsymbol{\ell}_1^A\boxtimes\boldsymbol{\ell}_2$ cancel in homology. To aid in this, we introduce the following terminology: we refer to differentials starting or ending at $x\otimes y$ in $\boldsymbol{\ell}_1^A\boxtimes\boldsymbol{\ell}_2$ as being "on the right" or "on the left" depending on whether the type D operations in $\boldsymbol{\ell}_2$ that give rise to the differential in the tensor product are to the right or left of y, relative to the cyclic ordering on $\boldsymbol{\ell}_2$. In



having fixed the loop $(d_2\bar{b}_1\bar{a}_1)$ (that is, reading the loop counterclockwise) the generator $x\otimes y$ cancels on the right when paired with the standard solid torus. (Recall that this example may be identified with the trivial surgery on the right-hand trefoil.) This terminology is motivated by picturing the tensor product on a grid with rows indexed by generators of $\boldsymbol{\ell}_1^A$ and columns indexed by generators of $\boldsymbol{\ell}_2$; compare Figure 7. With this terminology in place, we observe:

Lemma 5.5 *For any loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ and any generator $x\otimes y\in\boldsymbol{\ell}_1^A\boxtimes\boldsymbol{\ell}_2$, there is at most one differential into or out of $x\otimes y$ on the right, and at most one on the left.*

Proof This follows from examining the arrows of type $\mathbf{I}_\bullet$, $\mathbf{II}_\bullet$, $\mathbf{I}_\circ$ and $\mathbf{II}_\circ$ introduced in Section 3.1 and the corresponding type A arrows. For instance, suppose y has a type $\mathbf{I}_\bullet$ arrow on the right. A differential on the right of $x \otimes y$ must arise from an $\mathcal{A}_\infty$ operation starting at x with inputs starting with ρ_1, ρ_{12} or ρ_{123}. Such $\mathcal{A}_\infty$ operations correspond to sequences of arrows in $\boldsymbol{\ell}_1^A$ starting at x in which the first arrow is labelled by ρ_1. There is at most one such arrow from x in $\boldsymbol{\ell}_1^A$, since the corresponding vertex in $\boldsymbol{\ell}_1$ is adjacent to only one arrow of type $\mathbf{II}_\bullet$. A directed sequence of arrows in $\boldsymbol{\ell}_1^A$ of length k beginning with this ρ_1 labelled arrow gives rise to a collection of k $\mathcal{A}_\infty$ operations with first input ρ_1, ρ_{12} or ρ_{123}, each corresponding to the first i arrows in the sequence for some $i \leq k$. We observe that none of these operations have inputs that are the first n inputs of another operation in the collection. The reason for this is, when reading off the inputs for an $\mathcal{A}_\infty$ operation from a sequence of arrows following the conventions in Section 2.4, the effect of adding one more arrow is to leave all but the last input in the list unchanged, multiply the last input by something nontrivial and possibly add more inputs at the end of the list. For example, for a sequence of arrows, in $\boldsymbol{\ell}_1^A$,



the first arrow yields a contribution to $m_2(x, \rho_1)$, the first two arrows yield a contribution to $m_3(x, \rho_{12}, \rho_1)$ and all three arrows yields a contribution to $m_3(x, \rho_{12}, \rho_{12})$. Thus at most one operation in the collection can pair with a sequence of arrows on the right of y in $\boldsymbol{\ell}_2$. It follows that $x \otimes y$ has at most one differential on the right.

Similar arguments show that there is at most one differential on the right if the arrow on the right of y is type $\mathbf{II}_\bullet$, $\mathbf{I}_\circ$ or $\mathbf{II}_\circ$, and the same is true on the left. □

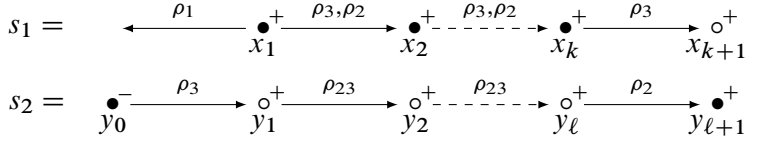
Thus, all differentials in $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ appear in linear chains and $x \otimes y$ will cancel in homology if it has an odd length chain of differentials on either side; we say that $x \otimes y$ cancels on the right (resp. left) if there is an odd length chain of differentials on the right (resp. left). Note that if there is an odd length chain of differentials on one side of $x \otimes y$, successively applying edge reduction to cancel the outermost differential on that side eventually results in cancelling $x \otimes y$ without using the differentials on the other side. If there is an even length chain, this chain can be removed using edge reduction leaving no differentials on that side of $x \otimes y$. If $x \otimes y$ does not cancel on the right or the left, we say $x \otimes y$ does not cancel in homology, since it, or potentially a linear combination of it with other generators of the same $(\mathbb{Z}/2\mathbb{Z})$–grading, survives in homology.

Proof of Proposition 5.1, "if" direction Suppose $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are L–space aligned. Up to changing parametrization, we can assume that the interval of strict L–space slopes for $\boldsymbol{\ell}_2$ contains $(0, p/q)$ for some $1 < p/q \leq \infty$ and does not contain 0 and the interval of strict L–space slopes for $\boldsymbol{\ell}_1$ contains $[-\infty, q/p]$. By Lemmas 5.2 and 5.3, $\boldsymbol{\ell}_1$ can be written with only e^* and c_k^* segments and $\boldsymbol{\ell}_2$ consists of a_k, b_k, c_k, d_k, e and $\bar{c}_1$ segments. We will fix the $(\mathbb{Z}/2\mathbb{Z})$–grading on each loop so that every generator of $\boldsymbol{\ell}_1^A$ has grading 0 and all ι_1–generators in $\boldsymbol{\ell}_2$ have grading 0 except those coming from $\bar{c}_1$ segments.

Consider a generator $x \otimes y$ in $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$. Then x belongs to a segment s_1 in $\boldsymbol{\ell}_1^A$ and y belongs to a segment s_2 in $\boldsymbol{\ell}_2$. We will consider cases depending on the type of the segments s_1 and s_2 and in each case show that either $x \otimes y$ has grading 0 or it cancels in homology. Therefore $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ is an L–space complex.

First note that $\mathrm{gr}(x)$ is always 0 by assumption. If s_2 is an e, d_k or b_k segment, then y also has grading 0 and the grading of $x \otimes y$ is 0. If s_1 is $\bar{e}^*$ then x is in idempotent ι_1 and so y must also have idempotent ι_1. All ι_1–generators of $\boldsymbol{\ell}_2$ have grading 0 except those in $\bar{c}_1$ segments, so $x \otimes y$ has grading 0 if s_1 is $\bar{e}^*$ and s_2 is not $\bar{c}_1$.

Suppose that s_1 is c_k^* and s_2 is a_ℓ with generators labelled by



Every generator in each segment has grading 0 except for y_0, so $x \otimes y$ has grading 0 unless $y = y_0$ and $x = x_i$ for $i \in \{1, \dots, k\}$. For each i, $\boldsymbol{\ell}_1^A$ has the operation

$$m_{2+k-i}(x_i, \rho_3, \rho_{23}, \dots, \rho_{23}) = x_{k+1}.$$

If $k - i < \ell$, it follows that in the tensor product there is a differential from $x_i \otimes y_0$ to $x_{k+1} \otimes y_{k-i+1}$. It is not difficult to check that $x_{k+1} \otimes y_{k-i+1}$ does not cancel from the right; the arrow in $\boldsymbol{\ell}_2$ to the right of y_{k-i+1} is outgoing and s_1 is followed by a c^* or $\bar{e}^*$ segment, so x_{k+1} has only incoming $\mathcal{A}_\infty$ operations. Therefore, in this case $x_i \otimes y_0$ cancels in homology. If instead $k - i \geq \ell$ then $\boldsymbol{\ell}_1^A$ has the operation

$$m_{2+\ell}(x_i, \rho_3, \rho_{23}, \dots, \rho_{23}, \rho_2) = x_{i+\ell}$$

and there is a differential in the box tensor product from $x_i \otimes y_0$ to $x_{i+\ell} \otimes y_{\ell+1}$. Again it is not difficult to see that $x_{i+\ell} \otimes y_{\ell+1}$ does not cancel from the right, since the arrow in $\boldsymbol{\ell}_2$ to the right of $y_{\ell+1}$ must be an outgoing ρ_1, ρ_{12} or ρ_{123} arrow and $x_{i+\ell}$ has

no outgoing $\mathcal{A}_\infty$ operations with first input ρ_1, ρ_{12} or ρ_{123}. Thus $x_i \otimes y_0$ cancels in homology.

Suppose that s_1 is c_k^*, with generators labelled as above, and s_2 is $\bar{c}_1$ with generators labelled as follows (note that the generator y_0 is not actually part of the segment s_2):

$$s_2 = \quad \overset{+}{\bullet}_{y_0} \xrightarrow{\rho_1} \overset{-}{\circ}_{y_1} \xleftarrow{\rho_3} \overset{+}{\bullet}_{y_2}$$

The generator y_2 has grading 0 and the generator y_1 has grading 1, so $x \otimes y$ only has grading 1 if $x = x_{k+1}$ and $y = y_1$. In this setting, $x_{k+1} \otimes y_1$ has an incoming differential on the right which starts from $x_k \otimes y_2$. To ensure that $x_{k+1} \otimes y_1$ cancels in homology we need to check that $x_k \otimes y_2$ does not cancel on the right. To the right of y_2 in $\boldsymbol{\ell}_2$ is an outgoing sequence of arrows that starts with some number of ρ_{12} arrows followed by a ρ_{123} arrow (here we use that a $\bar{c}_1$ segment in $\boldsymbol{\ell}_2$ is not followed by another $\bar{c}_1$ segment with only e's in between). If $k > 1$ then x_k has no outgoing $\mathcal{A}_\infty$ operations except $m_2(x_k, \rho_3) = x_{k+1}$. If $k = 1$, then x_k has additional operations, but, since s_1 is preceded by some number of $\bar{e}^*$ segments and then a c^* segment, the inputs for these operations can only be some number of ρ_{12}'s followed by a ρ_1. In either case, it is clear that $x_k \otimes y_2$ does not cancel on the right.

Suppose that s_1 is $\bar{e}^*$ and s_2 is $\bar{c}_1$. In this case, x is the only generator of $\bar{e}^*$, y is the only generator of $\bar{c}_1$ with idempotent ι_1, and $x \otimes y$ has grading 1. In $\boldsymbol{\ell}_2$, s_2 is preceded by some number i of e segments, which are preceded by a b_k or d_k segment. Thus to the left of y in $\boldsymbol{\ell}_2$ there is an incoming sequence of arrows that ends with ρ_2, i ρ_{12}'s and ρ_1. In $\boldsymbol{\ell}_1^A$, s_1 is followed by some number j of $\bar{e}^*$ segments followed by a c_k^*. If $j > i$ then x has an incoming operation with inputs $(\rho_2, \rho_{12}, \dots, \rho_{12}, \rho_1)$ with i ρ_{12}'s, and if $j \leq i$ then x has an incoming operation with inputs $(\rho_{12}, \dots, \rho_{12}, \rho_1)$ with j ρ_{12}'s. In either case, these operations give rise to a differential in the box tensor product ending at $x \otimes y$. In both cases it is also easy to check that the initial generator of this differential has no other differentials, and so $x \otimes y$ cancels in homology.

The only case remaining is the case that s_1 is c_k^* and s_2 is c_ℓ. This case is depicted in Figure 7 for $k = 3$ and $\ell = 2$. We label the generators of s_1 sequentially as $x_1, \dots, x_{k+1}$, and we label the generators of s_2 as $y_0, \dots, y_\ell$. For the remainder of this proof, assume that s_1 is c_k^*, s_2 is c_ℓ, $x \otimes y$ has grading 1 and $x \otimes y$ does not cancel in homology; we will produce a contradiction, proving the proposition.

Since $x \otimes y$ has grading 1, y is y_0 and x is x_i with $i \in \{1, \dots, k\}$. For each $i > k - \ell$, $x_i \otimes y_0$ has an outgoing differential which ends at $x_{k+1} \otimes y_{k-i+1}$. If $i > k - \ell + 1$, then

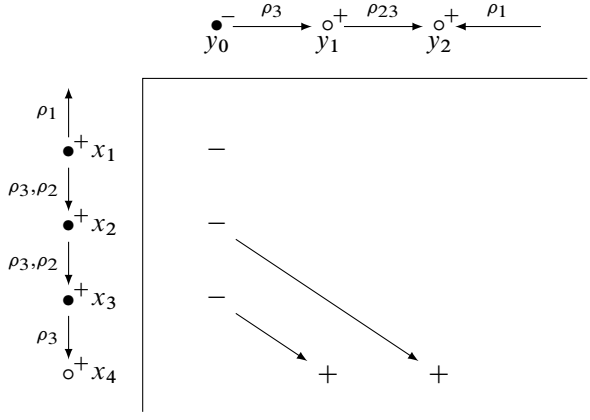


Figure 7: A portion of the chain complex $\boldsymbol{\ell}_1^A \boxtimes \boldsymbol{\ell}_2$ coming from a segment c_3^* in $\boldsymbol{\ell}_1^A$ (left edge) and a segment c_2 in $\boldsymbol{\ell}_2$ (top edge). Signs $+$ and $-$ indicate generators in the box tensor product with grading 0 and 1, respectively.

$x_{k+1} \otimes y_{k-i+1}$ does not cancel on the right, since the arrow to the right of y_{k-i+1} is an outgoing ρ_{23} arrow but x_{k+1} has only incoming $\mathcal{A}_\infty$ operations. Thus, if $i > k-\ell+1$, $x_i \otimes y_0$ cancels on the right. By Lemmas 5.2 and 5.3, $k \le n$ and $\ell \ge n-1$; it follows that $k-\ell$ is at most 1 and $x_i \otimes y_0$ cancels on the right for any $i > 2$. Thus x must be either x_1 or x_2.

If $k \le \ell$ then $x_2 \otimes y_0$ cancels on the right. If $k = n$ and $\ell = n-1$, $x_2 \otimes y_0$ potentially cancels from the right. It has an outgoing differential on the right ending at $x_{k+1} \otimes y_\ell$, so it cancels from the right if and only if $x_{k+1} \otimes y_\ell$ does not cancel on the right. Now $x_{k+1} \otimes y_\ell$ has a differential on the right only if $s_1 = c_k^*$ is followed by a segment $c_{k'}^*$, in which case the differential starts with the generator $x_1' \otimes y_0'$, where x_1' is the first generator of the $c_{k'}^*$ segment following s_1 and y_0' is the first generator of the segment following s_2. If $x_1' \otimes y_0'$ cancels on the right then so does $x_2 \otimes y_0$, and $x_1' \otimes y_0'$ automatically cancels on the right if s_2 is followed by a type a segment. Thus, if $x = x_2$ we must have that s_1 is followed by $c_{k'}^*$, s_2 is followed by c_ℓ', and $x_1' \otimes y_0'$ does not cancel from the right.

If $k < \ell$ then $x_1 \otimes y_0$ cancels on the right, and if $k > \ell$ then $x_1 \otimes y_0$ does not cancel on the right. If $k = \ell = n$ or $k = \ell = n-1$, then $x_1 \otimes y_0$ potentially cancels on the right. By the same reasoning as above, it does not cancel on the right if and only if s_1 is followed by $c_{k'}^*$, s_2 is followed by c_ℓ', and $x_1' \otimes y_0'$ does not cancel from the right. Now $x_1 \otimes y_0$ may also cancel from the left. It has an outgoing differential on the left ending at $x_0 \otimes y_{\ell'}'$, where x_0 is the last generator in the segment immediately preceding s_1

and $y'_{\ell'}$ is the last generator in the segment preceding s_2. If s_2 is preceded by $b_{\ell'}$ then it is easy to see that $x_0 \otimes y'_{\ell'}$ has no differentials on the left, so $x_1 \otimes y_0$ cancels from the left. If instead s_2 is preceded by $c_{\ell'}$, $x_0 \otimes y'_{\ell'}$ cancels from the left unless s_1 is preceded by $c^*_{k'}$ and $k' < \ell'$ or $k' = \ell'$ and $x'_1 \otimes y'_0$ cancels from the left.

Suppose that $x \otimes y = x_2 \otimes y_0$ does not cancel in homology (in particular it does not cancel from the right). Then we have shown that $k = n$, $\ell = n-1$, $s_1 = c^*_k$ is followed by $c^*_{k'}$, $s_2 = c_\ell$ is followed by $c_{\ell'}$, and the generator $x'_1 \otimes y'_0$ does not cancel from the right. Furthermore, this last fact implies that either $k' = n$ and $\ell' = n-1$ or that $k' = \ell'$, $c^*_{k'}$ is followed by $c^*_{k''}$, $c_{\ell'}$ is followed by $c_{\ell''}$, and $x''_1 \otimes y''_0$ does not cancel from the right, where x''_1 and y''_0 are the appropriate generators of $c^*_{k''}$ and $c_{\ell'}$. Repeating this argument, we see that $s_1 = c^*_n$ is followed by a sequence of c^*_n and c^*_{n-1} segments ending with a c^*_n and $s_2 = c_{n-1}$ is followed by a sequence of c_n and c_{n-1} segments ending with c_{n-1} but with indices otherwise the same as the indices in the sequence of c^* segments following s_1.

Suppose that $x \otimes y = x_1 \otimes y_0$ does not cancel in homology. The fact that it does not cancel from the right implies that $k = \ell = n$ or $k = \ell = n-1$, $s_1 = c^*_k$ is followed by $c^*_{k'}$, $s_2 = c_\ell$ is followed by $c_{\ell'}$, and the generator $x'_1 \otimes y'_0$ does not cancel from the right. As in the preceding paragraph, this implies that $s_1 = c^*_n$ is followed by a sequence of c^*_n and c^*_{n-1} segments ending with a c^*_n and $s_2 = c_{n-1}$ is followed by a corresponding sequence of c_n and c_{n-1} segments ending with c_{n-1}. The fact that $x \otimes y = x_1 \otimes y_0$ does not cancel from the left implies that $s_1 = c^*_k$ is preceded by $c^*_{k'}$, $s_2 = c_\ell$ is preceded by $c_{\ell'}$, and either $k' = n$ and $\ell' = n-1$ or $k' = \ell'$ and the generator $x'_1 \otimes y'_0$ does not cancel from the left. Repeating the argument, we see that s_1 is preceded by a sequence of c^*_n and c^*_{n-1} segments starting with c^*n and s_2 is preceded by a sequence of c_n and c_{n-1} segments with matching sequence of indices except that the initial segment is c_{n-1}.

Regardless of whether $x = x_1$ or $x = x_2$, we find that there is a sequence of indices $k_1, k_2, \dots, k_r$ with $k_i \in \{n, n-1\}$ such that $\boldsymbol{\ell}_1^A$ contains $c^*_n c^*_{k_1} c^*_{k_2} \dots c^*_{k_r} c^*_n$ and $\boldsymbol{\ell}_2$ contains $c_{n-1} c_{k_1} c_{k_2} \dots c_{k_r} c_{n-1}$. It follows that $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ satisfy property λ. By Lemma 5.4, this implies that $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are not L–space aligned, a contradiction. □

As noted previously, the gluing statement in Proposition 5.1 needs to be modified if either loop is solid torus-like. Note that for a solid torus-like loop, all slopes are L–space slopes except the rational longitude. If $\boldsymbol{\ell}_1$ is solid torus-like with rational longitude r/s and $\boldsymbol{\ell}_2$ is simple, then $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are L–space aligned if and only if s/r

is a strict L–space slope for ℓ_2. This is a sufficient, but not a necessary, condition for $\ell_1^A \boxtimes \ell_2$ to be an L–space complex.

Proposition 5.6 *If ℓ_1 and ℓ_2 are simple loops and ℓ_1 is solid torus-like with rational longitude r/s, then $\ell_1^A \boxtimes \ell_2$ is an L–space chain complex if and only if s/r is an L–space slope for ℓ_2.*

Proof We may choose a framing so that ℓ_1 is a collection of e segments, that is, so that the rational longitude is represented by 0. Correspondingly, the slope s/r for ℓ_2 is represented by ∞. The result now follows from Corollary 3.22. □

5.2 A gluing result for loop-type manifolds

Returning to loop-type manifolds, we are now in a position to collect the material proved in this section and, in particular, apply Proposition 5.1 to establish a gluing theorem.

Theorem 5.7 *Let (M_1, α_1, β_1) and (M_2, α_2, β_2) be simple loop-type bordered manifolds with torus boundary which are not solid torus-like, and let Y be the closed manifold $(M_1, \alpha_1, \beta_1) \cup (M_2, \alpha_2, \beta_2)$ (with the gluing map $\alpha_1 \mapsto \beta_2$, $\beta_1 \mapsto \alpha_2$, as in Section 2.5). Then Y is an L–space if and only if every essential simple closed curve on $\partial M_1 = \partial M_2 \subset Y$ determines a strict L–space slope for either M_1 or M_2.*

Remark 5.8 There is an alternative statement of the conclusion on Theorem 5.7 using the notation laid out in this paper: the closed manifold $M_1 \cup M_2$ is an L–space if and only if, for each rational p/q, either $p/q \in \mathcal{L}^\circ(M_1, \alpha_1, \beta_1)$ or $q/p \in \mathcal{L}^\circ(M_2, \alpha_2, \beta_2)$. Recall that, following the conventions in Section 2.5, $p/q \in \mathcal{L}^\circ(M, \alpha, \beta)$ if and only if $\pm(p\alpha + q\beta) \in \mathcal{L}^\circ_M$.

Note that Theorem 5.7 implies Theorem 1.3.

Remark 5.9 If either bordered manifold in the statement of Theorem 5.7 is solid torus-like, then Y is an L–space if and only if every essential simple closed curve on $\partial M_1 = \partial M_2 \subset Y$ determines an L–space slope for either M_1 or M_2. This is an immediate consequence of Proposition 5.6, and amounts to replacing $\mathcal{L}^\circ(M_i, \alpha, \beta)$ by $\mathcal{L}(M_i, \alpha, \beta)$. Notice that Dehn surgery — by definition of an L–space slope — is a special case of this version of the gluing result, since the solid torus is solid torus-like, in the sense of Definition 3.23.

Proof of Theorem 5.7 Let $\widehat{CFD}(M_1,\alpha_1,\beta_1)$ be represented by simple loops $\boldsymbol{\ell}_1^1,\dots,\boldsymbol{\ell}_1^n$ and let $\widehat{CFD}(M_2,\alpha_2,\beta_2)$ be represented by simple loops $\boldsymbol{\ell}_2^1,\dots,\boldsymbol{\ell}_2^m$. A given slope is a (strict) L–space slope for (M_i,α_i,β_i) if and only if it is a (strict) L–space slope (abstractly) for each loop $\boldsymbol{\ell}_i^k$. Y is an L–space if and only if $\boldsymbol{\ell}_1^k\boxtimes\boldsymbol{\ell}_2^j$ is and L–space complex (again, abstractly) for each $1\le k\le n$ and $1\le j\le m$.

Suppose there is a slope p/q such that $p/q\notin\mathcal{L}^\circ(M_1,\alpha_1,\beta_1)$ and $q/p\notin\mathcal{L}^\circ(M_2,\alpha_2,\beta_2)$. Then p/q is not a strict L–space slope for $\boldsymbol{\ell}_1^k$ for some k and q/p is not a strict L–space slope for $\boldsymbol{\ell}_2^j$ for some j. We may assume that $\boldsymbol{\ell}_1^k$ is not solid torus-like; if it is solid torus-like, then p/q must be the rational longitude for (M_1,α_1,β_1) and thus not a strict L–space slope for any of the loops $\boldsymbol{\ell}_1^1,\dots,\boldsymbol{\ell}_1^n$. Similarly, we may assume that $\boldsymbol{\ell}_2^j$ is solid torus-like. Since $\boldsymbol{\ell}_1^k$ and $\boldsymbol{\ell}_2^j$ are not L–space aligned, $\boldsymbol{\ell}_1^k\boxtimes\boldsymbol{\ell}_2^j$ is not an L–space complex, and Y is not an L–space.

If every slope is a strict L–space slope for either M_1 or M_2, then every slope is a strict L–space slope either for each $\boldsymbol{\ell}_1^k$ or for each $\boldsymbol{\ell}_2^j$. It follows that, for every pair (k,j), $\boldsymbol{\ell}_1^k$ and $\boldsymbol{\ell}_2^j$ are L–space aligned, and thus Y is an L–space. □

5.3 L–space knot complements

In light of Theorem 5.7, it is natural to ask which 3–manifolds M with torus boundary have simple loop-type bordered invariants. We first observe that complements of L–space knots in S^3, that is, those knots admitting an L–space surgery, have this property.

Proposition 5.10 *If K is an L–space knot in S^3 and $M=S^3\smallsetminus\nu(K)$, then for any choice of parametrizing curves α and β, $\widehat{CFD}(M,\alpha,\beta)$ can be represented by a single simple loop.*

Proof We need only show that, for some choice of parametrizing curves, $\widehat{CFD}(M,\alpha,\beta)$ is a loop which consists of only unstable chains in either standard notation or dual notation. Consider then $\widehat{CFD}(M,\mu,n\mu+\lambda)$, where μ is the meridian of the knot K and λ is the Seifert longitude of K. Suppose that $|n|$ is chosen sufficiently large that the result of Dehn surgery $S^3_n(K)=M(n\mu+\lambda)$ is an L–space.

The knot Floer homology $CFK^-(K)$ has a staircase shape of alternating horizontal and vertical arrows (see [25], for example). By the construction in [20, Section 11.5], $\widehat{CFD}(M,\mu,n\mu+\lambda)$ consists of alternating horizontal and vertical chains, with the ends connected by a single unstable chain. Thus $\widehat{CFD}(M,\mu,n\mu+\lambda)$ is a loop. In standard loop notation, the horizontal chains are type a segments, the vertical chains are type b segments, and the unstable chain is a type d segment if $n>0$ and a type c

segment if $n < 0$. More precisely,

$$\widehat{CFD}(M,\mu,n\mu+\lambda) = \begin{cases} b_{k_1}a_{k_2}b_{k_3}a_{k_4}\cdots b_{k_{2r-1}}a_{k_{2r}}d_j & \text{if } n>0,\\ a_{k_1}b_{k_2}a_{k_3}b_{k_4}\cdots a_{k_{2r-1}}b_{k_{2r}}c_j & \text{if } n<0. \end{cases}$$

In either case, the loop has no bar segments in standard notation, so when we switch to dual notation it has no stable chains. Thus, $\widehat{CFD}(M,\mu,n\mu+\lambda)$ is a simple loop. □

Note that the behaviour established in this application of Theorem 5.7 is expected in general and, in particular, should not require the hypothesis that the knot complement be a loop-type manifold; see [9, Conjecture 1; 5, Conjecture 4.3].

Conjecture 5.11 *For knots K_1 and K_2 in S^3, with $M_i = S^3 \smallsetminus \nu(K)$ for $i = 1,2$, the generalized splicing $M_1 \cup_h M_2$ is an L–space if and only if either $\gamma \in \mathcal{L}^\circ_{M_1}$ or $h(\gamma) \in \mathcal{L}^\circ_{M_2}$ for every slope γ.*

On the other hand, it is clear that the case where two L–space knot complements are identified in the generalized splice is the interesting case, and therefore the loop restriction in this setting is quite natural. In this case, we remark that the question is now settled.

Corollary 5.12 *Conjecture 5.11 holds when the K_i are L–space knots.*

Proof This is immediate on combining Theorem 5.7 and Proposition 5.10. □

Note that Corollary 5.12 implies Theorem 1.6.

Some authors define L–space knots to be the class of knots for which some *positive* surgery yields an L–space. We will not follow this convention because, while this is a natural definition for knots in S^3 in that certain statements become simpler, our interest is in the more general setting of manifolds with torus boundary admitting L–space fillings, wherein the distinction seems to be less meaningful. In particular, the next section is concerned with this more general setting.

6 Graph manifolds

Note that knots in the three-sphere admitting L–space surgeries contain torus knots — those knots admitting a Seifert structure on their complement — as a strict subset. Our goal now is to establish another class of loop-type manifolds: Seifert-fibred rational homology solid tori. This should be viewed as the natural geometric enlargement of the

class of torus knot exteriors. To do so, we study these as a subset of graph manifolds, and establish sufficient conditions for a rational homology solid torus admitting a graph manifold structure to be of loop-type.

6.1 Preliminaries on graph manifolds

We will represent a graph manifold rational homology sphere Y by a plumbing tree, following the notation developed in [22]. A plumbing tree is an acyclic graph with integer weights associated to each vertex. Such a graph specifies a graph manifold as follows: To each vertex v_i of weight e_i and valence d_i we assign the Euler number e_i circle bundle over the sphere minus d_i disks, and for each edge connecting vertices v_i and v_j glue the corresponding bundles along a torus boundary component by a gluing map that takes the fibre of one bundle to a curve in the base surface of the other bundle, and vice versa.

To allow for graph manifolds with boundary, we associate an additional integer $b_i \geq 0$ to each vertex. To construct the corresponding manifold, we associate to each vertex a bundle over S^2 minus $(d_i + b_i)$; d_i boundary components of this bundle will glue to bundles corresponding to other vertices, but the remaining b_i boundary components remain unglued. The resulting graph manifold has $\sum_i b_i$ toroidal boundary components. In diagrams of plumbing trees, we will indicate the presence of boundary tori by drawing b_i half-edges (dotted lines which do not connect to another vertex) at each vertex v_i.

Given a plumbing tree Γ with a single boundary half-edge, let M_Γ denote the corresponding graph manifold. The torus ∂M_Γ has a natural choice of parametrizing curves: one corresponds to a fibre in the S^1 bundle containing the boundary, and one corresponds to a curve in the base surface of that bundle (note that these are precisely the curves used above to specify the gluing maps in the construction based on a given graph). Call these two slopes α and β, respectively. These slopes do not have a preferred orientation, but reversing the orientation on the bundle associated to every vertex in the construction of M_Γ gives a diffeomorphism between $(M_\Gamma, \alpha, \beta)$ and $(M_\Gamma, -\alpha, -\beta)$ as bordered manifolds.

Thus $(M_\Gamma, \alpha, \beta)$ is a canonical bordered manifold associated to Γ. Since we will be interested in the bordered invariants of such graph manifolds, for ease of notation we will refer to $\widehat{CFD}(M_\Gamma, \alpha, \beta)$ simply as $\widehat{CFD}(\Gamma)$.

We will make use of three important operations on single boundary plumbing trees. These are summarized in Figure 8, and described as follows:

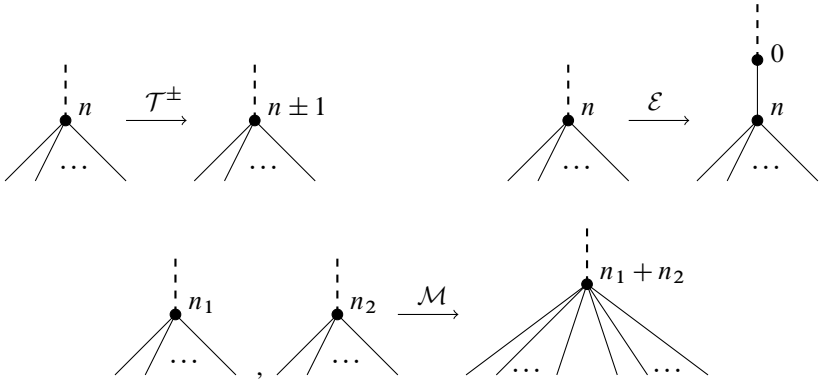


Figure 8: Three operations on graphs for constructing and graph manifold with torus boundary.

Twist The operation $\mathcal{T}^{\pm}$ adds ± 1 to the weight of the vertex containing the boundary edge.

Extend The operation $\mathcal{E}$ inserts a new 0–framed vertex between the boundary vertex and the boundary edge.

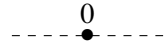
Merge The operation $\mathcal{M}$ takes two single boundary plumbing trees and identifies their boundary vertices, removing one of the two boundary edges to produce a new single boundary plumbing tree; the weight of the new boundary vertex is the sum of the weights of the original boundary vertices.

With these operations we can, in particular, construct any single-boundary plumbing tree.

6.2 Operations on bordered manifolds

The graph operations described above may be thought of as operations on the corresponding (bordered) graph manifolds. In particular, we will abuse notation and write $M_{\mathcal{T}(\Gamma)} = \mathcal{T}(M_\Gamma)$. $\mathcal{T}(M_\Gamma)$ is obtained from M_Γ by attaching the mapping cylinder of a positive (standard) Dehn twist to the boundary of M_Γ. More precisely, by considering the gluing conventions prescribed by the plumbing tree, the effect on the bordered manifold is $\mathcal{T}(M_\Gamma, \alpha, \beta) = (M_\Gamma, \alpha, \beta + \alpha)$.

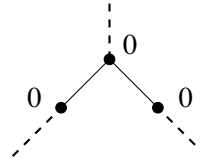
Similarly, $\mathcal{E}(M_\Gamma) = M_{\mathcal{E}(\Gamma)}$ is obtained from M_Γ by attaching the bimodule corresponding to the two-boundary plumbing tree



The graph manifold assigned to this plumbing tree is the trivial S^1–bundle over the annulus, or $T^2 \times [0,1]$. Recall that the bordered structure on each boundary component of this manifold is given by the convention that α is an S^1 fibre and β lies in the base surface. One checks that the resulting bordered manifold can be realised as the mapping cylinder for the diffeomorphism represented by the matrix $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. The effect on the level of bordered manifolds is $\mathcal{E}(M_\Gamma, \alpha, \beta) = (M_\Gamma, -\beta, \alpha)$.

Finally, $\mathcal{M}(M_{\Gamma_1}, M_{\Gamma_2})$ is the result of attaching the bundle $S^1 \times \mathcal{P}$, where $\mathcal{P}$ is a pair of pants. The bordered structure on $S^1 \times \mathcal{P}$ (which we suppress from the notation) is determined as follows: the two "input" boundary components which glue to M_{Γ_1} and M_{Γ_2} are parametrized by pairs (α, β) with β a fibre and α a curve in the base surface $\mathcal{P}$, while the third "output" boundary component is parametrized by (α, β) with α a fibre and β a curve in $\mathcal{P}$. As usual, the bordered structure on $S^1 \times \mathcal{P}$ determines the gluing as well as the bordered structure on the resulting three-manifold.

To see that $\mathcal{M}(M_{\Gamma_1}, M_{\Gamma_2})$ is indeed the manifold corresponding to $\mathcal{M}(\Gamma_1, \Gamma_2)$, note that $S^1 \times \mathcal{P}$ with the specified bordered structure is the manifold associated with the three-boundary plumbing tree $\Gamma_{\mathcal{M}}$,

$\Gamma_{\mathcal{M}} =$ 

On the other hand, attaching Γ_1 and Γ_2 to the two lower boundary edges of $\Gamma_{\mathcal{M}}$ produces a single-boundary plumbing tree that is equivalent to $\mathcal{M}(\Gamma_1, \Gamma_2)$ (by equivalent, we mean that the corresponding manifolds are diffeomorphic; to see this equivalence, use rule R3 of [22] to contract the 0–framed valence two vertices).

From this discussion, it should be clear that the operations $\mathcal{T}$, $\mathcal{E}$ and $\mathcal{M}$ may be extended to natural operations on arbitrary bordered manifolds. We are interested in the effect of these operations on bordered invariants. Following the conventions laid out, we have that $\mathcal{T}^{\pm}(M, \alpha, \beta) = (M, \alpha, \beta \pm \alpha)$, so that, at the level of bordered invariants,

$$\widehat{CFD}(\mathcal{T}^{\pm}(M, \alpha, \beta)) \cong \widehat{T}_{\mathrm{st}}^{\pm 1} \boxtimes \widehat{CFD}(M, \alpha, \beta). \tag{1}$$

Similarly, the action of $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, giving $\mathcal{E}(M, \alpha, \beta) = (M, -\beta, \alpha)$, is realized by

$$\widehat{CFD}(\mathcal{E}(M, \alpha, \beta)) \cong \widehat{T}_{\mathrm{st}} \boxtimes \widehat{T}_{\mathrm{du}} \boxtimes \widehat{T}_{\mathrm{st}} \boxtimes \widehat{CFD}(M, \alpha, \beta) \tag{2}$$

on type D structures.

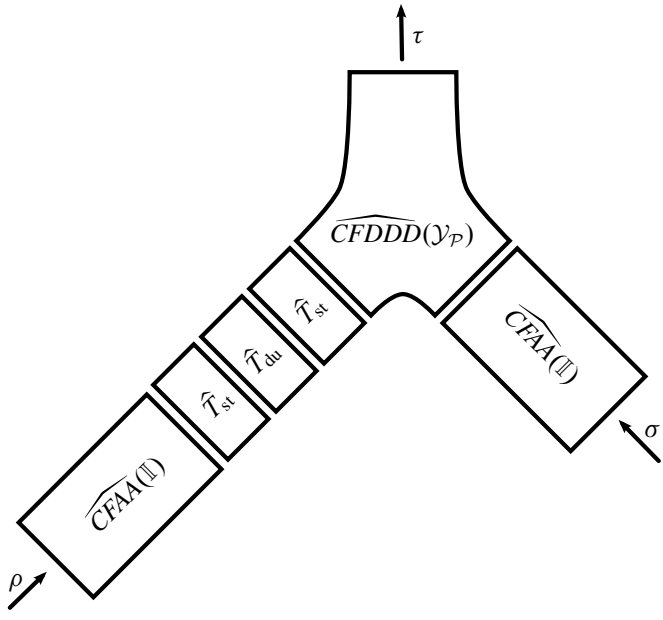


Figure 9: A schematic description of the trimodule $\widehat{CFDAA}(S^1\times\mathcal{P})$ extracted from the calculation and conventions in [8].

The most involved is the merge operation: for a pair of bordered manifolds (M_i,α_i,β_i) for $i=1,2$, this produces a bordered manifold

$$\mathcal{M}_{1,2}=\mathcal{M}((M_1,\alpha_1\beta_1),(M_2,\alpha_2,\beta_2))$$

using $S^1\times\mathcal{P}$. As a bordered 3–manifold with three boundary components, $S^1\times\mathcal{P}$ gives rise to a trimodule in the bordered theory (the object of study in work of the first author [8]). More precisely, we consider the type DAA trimodule $\widehat{CFDAA}(S^1\times\mathcal{P})$.

We will extract $\widehat{CFDAA}(S^1\times\mathcal{P})$ from the trimodule $\widehat{CFDDD}(\mathcal{Y}_\mathcal{P})$ computed in [8]. Note that $S^1\times\mathcal{P}$ and $\mathcal{Y}_\mathcal{P}$ agree as manifolds, but have different bordered structure. The trimodule $\widehat{CFDDD}(\mathcal{Y}_\mathcal{P})$ has five generators, v, w, x, y and z, and the differential is given by (cf [8, Figure 10])

$$\begin{aligned}
\partial(v)&=\rho_3\otimes x+\rho_1\sigma_3\tau_{123}\otimes y+\rho_{123}\sigma_{123}\tau_{123}\otimes y+\tau_3\otimes z,\\
\partial(w)&=\rho_3\sigma_{12}\otimes x+\rho_1\sigma_3\tau_1\otimes y+\rho_{123}\sigma_{123}\tau_1\otimes y,\\
\partial(x)&=\rho_2\sigma_{12}\otimes v+\rho_2\otimes w+\sigma_1\tau_3\otimes y,\\
\partial(y)&=\sigma_2\tau_2\otimes x+\rho_2\sigma_2\otimes z,\\
\partial(z)&=\rho_3\sigma_1\otimes y+\tau_2\otimes w.
\end{aligned}$$

$m_2(x,\sigma_3)=z$	$m_3(v,\rho_1,\sigma_1)=\tau_1\otimes y$
$m_2(y,\sigma_2)=x$	$m_3(v,\rho_1,\sigma_{12})=\tau_1\otimes x$
$m_2(y,\sigma_{23})=z$	$m_3(v,\rho_1,\sigma_{123})=\tau_1\otimes z$
$m_2(w,\rho_3)=z$	$m_3(v,\rho_1,\sigma_{123})=\tau_{123}\otimes y$
$m_2(y,\rho_2)=w$	$m_3(v,\rho_{12},\sigma_1)=\tau_1\otimes w$
$m_2(y,\rho_{23})=z$	$m_3(v,\rho_{12},\sigma_{12})=\tau_{12}\otimes v$
$m_2(w,\sigma_{23})=\tau_{23}\otimes w$	$m_3(v,\rho_{12},\sigma_{123})=\tau_{123}\otimes w$
$m_2(x,\sigma_3)=\tau_{23}\otimes y$	$m_3(v,\rho_{123},\sigma_1)=\tau_1\otimes z$
$m_2(x,\rho_2)=\tau_2\otimes v$	$m_3(v,\rho_{123},\sigma_{12})=\tau_{123}\otimes x$
$m_2(x,\rho_{23})=\tau_{23}\otimes x$	$m_3(v,\rho_{123},\sigma_{123})=\tau_{123}\otimes z$
$m_2(v,\sigma_3)=\tau_3\otimes w$	$m_5(v,\rho_3,\rho_2,\rho_1,\sigma_1)=\tau_1\otimes z$
$m_2(y,\sigma_{23})=\tau_{23}\otimes y$	$m_5(v,\rho_3,\rho_2,\rho_1,\sigma_1)=\tau_{123}\otimes y$
$m_2(v,\rho_3)=\tau_3\otimes x$	$m_5(v,\rho_3,\rho_2,\rho_{12},\sigma_1)=\tau_{123}\otimes w$
$m_2(w,\sigma_2)=\tau_2\otimes v$	$m_5(v,\rho_3,\rho_2,\rho_{123},\sigma_1)=\tau_{123}\otimes z$
$m_2(z,\sigma_2)=\tau_{23}\otimes x$	$m_7(v,\rho_3,\rho_2,\rho_3,\rho_2,\rho_1,\sigma_1)=\tau_{123}\otimes z$
$m_2(z,\sigma_{23})=\tau_{23}\otimes z$	

Table 1: Operations for $\widehat{CFDAA}(\Gamma_{\mathcal{M}})$.

Since the trimodule has three commuting actions by three copies of the torus algebra (one corresponding to each boundary component), we use ρ, σ and τ to distinguish between algebra elements in each copy of $\mathcal{A}$. The σ–boundary of $\mathcal{Y}_{\mathcal{P}}$ is parametrized by a pair (α,β) such that β is a fibre of $\mathcal{P}\times S^1$ and α lies in the base $\mathcal{P}$, while the ρ– and τ–boundaries have the opposite parametrization. Thus the bordered manifold $S^1\times\mathcal{P}$ can be obtained from $\mathcal{Y}_{\mathcal{P}}$ by switching the role of α and β on (say) the ρ–boundary.

It is now clear how to compute $\widehat{CFDAA}(S^1\times\mathcal{P})$ from $\widehat{CFDDD}(\mathcal{Y}_{\mathcal{P}})$: we change the parametrization of the ρ–boundary by applying the bimodule $\widehat{T}_{\mathrm{st}}\boxtimes\widehat{T}_{\mathrm{du}}\boxtimes\widehat{T}_{\mathrm{st}}$ and then change the ρ– and σ–boundaries to type A by tensoring with $\widehat{CFAA}(\mathbb{I})$ (see Figure 9 for a schematic description). These are both straightforward computations; the resulting trimodule has five generators and the operations are listed in Table 1.

With this trimodule in hand, we have

$$\widehat{CFD}(\mathcal{M}_{1,2})\cong\widehat{CFDAA}(S^1\times\mathcal{P})\boxtimes((M_1,\alpha_1,\beta_1),(M_2,\alpha_2,\beta_2)),$$

where $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\cdot, \cdot)$, by convention, tensors against the ρ–boundary in the first factor and against the σ–boundary in the second factor; compare Figure 9.

6.3 The effect of twist, extend and merge on loops

Restricting to the case of loop-type bordered invariants, the effect of the operations $\mathcal{T}^{\pm 1}$, $\mathcal{E}$ and $\mathcal{M}$ can be given simpler descriptions. Recall that $\mathrm{E} = \mathrm{T} \circ \mathrm{H}^{-1} \circ \mathrm{T}$, and that E can be easily calculated using Lemma 3.10.

Proposition 6.1 *If $\widehat{CFD}(M, \alpha, \beta)$ is of loop-type, and represented by the collection $\{\boldsymbol{\ell}_i\}_{i=1}^n$, then $\widehat{CFD}(\mathcal{T}^{\pm}(M, \alpha, \beta))$ is represented by $\{\mathrm{T}^{\pm 1}(\boldsymbol{\ell}_i)\}_{i=1}^n$ and $\widehat{CFD}(\mathcal{E}(M, \alpha, \beta))$ is represented by $\{\mathrm{E}(\boldsymbol{\ell}_i)\}_{i=1}^n$.*

Proof The proof is immediate from Proposition 3.11 with (1) and Proposition 3.12 with (2). □

It follows easily that the operations $\mathcal{T}^{\pm}$ and $\mathcal{E}$ preserve the simple loop-type property.

Lemma 6.2 *Given a (simple) loop-type bordered three-manifold, the operations $\mathcal{T}^{\pm}$ and $\mathcal{E}$ produce (simple) loop-type manifolds.*

Proof It is an immediate consequence of Proposition 6.1 that if $\widehat{CFD}(M, \alpha, \beta)$ is represented by a collection of loops then the same is true for $\widehat{CFD}(\mathcal{T}(M, \alpha, \beta))$ and $\widehat{CFD}(\mathcal{E}(M, \alpha, \beta))$, since the operations $\mathrm{T}^{\pm 1}$ and E take loops to loops. Moreover, if the loops defining $\widehat{CFD}(M, \alpha, \beta)$ are simple then the loops resulting from $\mathcal{T}^{\pm}$ and $\mathcal{E}$ are simple, since changing framing by Dehn twists does not, by definition, change whether or not a loop is simple. It only remains to check that $\widehat{CFD}(\mathcal{T}(M, \alpha, \beta))$ and $\widehat{CFD}(\mathcal{E}(M, \alpha, \beta))$ have exactly one loop for each spin^c–structure on the corresponding manifolds. This is again clear since it is true for $\widehat{CFD}(M, \alpha, \beta)$ and the operations $\mathcal{T}^{\pm 1}$ and $\mathcal{E}$ amount to changing the parametrization on the boundary of M; changing the parametrization does not change the number of spin^c–structures, and the loop operations $\mathrm{T}^{\pm 1}$ and E do not change the number of loops. □

For the merge operation, we will restrict further to the case that the loop(s) representing $\widehat{CFD}(M_1, \alpha_1, \beta_1)$ can all be written in standard notation with no stable chains. In this case, $\widehat{CFD}(\mathcal{M}_{1,2})$ is also a collection of loops.

Remark 6.3 By contrast, if $\widehat{CFD}(M_1, \alpha_1, \beta_1)$ and $\widehat{CFD}(M_2, \alpha_2, \beta_2)$ both contain stable chains in standard notation, $\widehat{CFD}(\mathcal{M}_{1,2})$ is not obviously of loop-type. However, in many cases it can be realized as a collection of loops after a homotopy equivalence.

It is enough to describe $\mathcal{M}$ on individual loops; we use M to denote the corresponding operation on abstract loops. If $\widehat{CFD}(M_1,\alpha_1,\beta_1)$ is represented by a collection of loops $\{\boldsymbol{\ell}_i\}_{i=1}^n$ and $\widehat{CFD}(M_2,\alpha_2,\beta_2)$ is represented by a collection of loops $\{\boldsymbol{\ell}_j\}_{j=1}^m$, then $\widehat{CFD}(\mathcal{M}_{1,2})$ is given by $\bigcup_{1\le i\le n,1\le j\le m}\mathrm{M}(\boldsymbol{\ell}_i,\boldsymbol{\ell}_j)$. Determining $\mathrm{M}(\boldsymbol{\ell}_i,\boldsymbol{\ell}_j)$ is a direct calculation using the trimodule and a key application of loop calculus.

Proposition 6.4 *Let $\boldsymbol{\ell}_1$ be a loop which can be written in standard notation with only type d_k unstable chains and let $\boldsymbol{\ell}_2$ be any loop. Then $\mathrm{M}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ is a collection of loops. If $\boldsymbol{\ell}_2$ cannot be written in standard notation, then $\mathrm{M}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ is one copy of $\boldsymbol{\ell}_2$ for each segment in $\boldsymbol{\ell}_1$. Otherwise, $\mathrm{M}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ is determined as follows: The ι_0–vertices of the $\mathcal{A}$–decorated graph correspond to pairs (u,v), where u is an ι_0–vertex of $\boldsymbol{\ell}_1$ and v is an ι_0–vertex of $\boldsymbol{\ell}_2$, and for each d_k segment from u_1 to u_2 in $\boldsymbol{\ell}_1$ we have:*

(1) *For each a_l segment from v_1 to v_2 in $\boldsymbol{\ell}_2$, there is an a_l segment from (u_2,v_1) to (u_2,v_2).*

(2) *For each b_l segment from v_1 to v_2 in $\boldsymbol{\ell}_2$, there is n b_l segment from (u_1,v_1) to (u_1,v_2).*

(3) *For each c_l segment from v_1 to v_2 in $\boldsymbol{\ell}_2$, there is a c_{l-k} segment from (u_2,v_1) to (u_1,v_2).*

(4) *For each d_l segment from v_1 to v_2 in $\boldsymbol{\ell}_2$, there is a d_{k+l} segment from (u_1,v_1) to (u_2,v_2).*

Proof The type D module represented by $\mathrm{M}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ is obtained by tensoring the type D modules corresponding to $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ with the ρ– and σ–boundaries, respectively, of the trimodule $\widehat{CFDAA}(\mathcal{Y}_{\mathcal{P}})$. We will denote this tensor by $\widehat{CFDAA}(\mathcal{Y}_{\mathcal{P}})\boxtimes(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$. This trimodule has five generators; the idempotents associated with each generator on the ρ–, τ– and σ–boundaries are as follows:

generator	v	w	x	y	z
idempotents	$(\iota_0^\rho,\iota_0^\sigma,\iota_0^\tau)$	$(\iota_0^\rho,\iota_1^\sigma,\iota_1^\tau)$	$(\iota_1^\rho,\iota_0^\sigma,\iota_1^\tau)$	$(\iota_1^\rho,\iota_1^\sigma,\iota_1^\tau)$	$(\iota_1^\rho,\iota_1^\sigma,\iota_1^\tau)$

Note that the ι_0–generators in $\widehat{CFDAA}(S^1\times\mathcal{P})\boxtimes(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ arise precisely from the generator v tensored with ι_0–generators in $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$.

First suppose that $\boldsymbol{\ell}_2$ is written in standard notation. Note also that, since $\boldsymbol{\ell}_1$ is assumed to have no a_k segments, we may ignore the m_5 and m_7 operations in Table 1. As a result, there are no operations in $\mathrm{M}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ that arise from more than one segment

in either loop. This means that to compute $\mathrm{M}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ we can feed $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ into the trimodule one segment at a time. For each combination of segment in $\boldsymbol{\ell}_1$ and segment in $\boldsymbol{\ell}_2$, the resulting portion of $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ is homotopy equivalent to the segment determined by (1)–(4) in the statement of the proposition. The proof is essentially contained in Figures 10, 11 and 12; we will describe one case in detail and leave the details of the other cases to the reader, with the figures as a guide.

Consider a segment d_k in $\boldsymbol{\ell}_1$ and a segment a_l in $\boldsymbol{\ell}_2$ with $k, l > 0$ (see Figure 10, top left, for the case of $k = l = 2$). Let the generators in these two segments be labelled as

$$d_k = \xrightarrow{\rho_{123}} \underset{x_1}{\circ} \xrightarrow{\rho_{23}} \cdots \xrightarrow{\rho_{23}} \underset{x_k}{\circ} \xrightarrow{\rho_2} \underset{u_2}{\bullet} \rightarrow \qquad a_l = \leftarrow \underset{v_1}{\bullet} \xrightarrow{\sigma_3} \underset{y_1}{\circ} \xrightarrow{\sigma_{23}} \cdots \xrightarrow{\sigma_{23}} \underset{y_l}{\circ} \xrightarrow{\sigma_2} \underset{v_2}{\bullet} \rightarrow$$

Note that the arrow adjacent to u_2 on the right is determined by the segment following d_k in $\boldsymbol{\ell}_1$, but, following Section 3.1 it must be an outgoing ρ_1, ρ_{12} or ρ_{123} arrow. Similarly, the arrows to the left of v_1 and the right of v_2 are outgoing σ_1, σ_{12} or σ_{123} arrows. The portion of $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ coming from d_k and a_l has generators

$$\begin{aligned}
x \otimes x_i \otimes v_j &\quad \text{for } 1 \le i \le k,\ j \in \{1, 2\}, \\
y \otimes x_i \otimes y_j &\quad \text{for } 1 \le i \le k,\ 1 \le j \le l, \\
z \otimes x_i \otimes y_j &\quad \text{for } 1 \le i \le k,\ 1 \le j \le l, \\
v \otimes u_2 \otimes v_j &\quad \text{for } j \in \{1, 2\}, \\
w \otimes u_2 \otimes y_j &\quad \text{for } 1 \le j \le l.
\end{aligned}$$

For each i in $\{1, \dots, k\}$, the trimodule operations $m_2(x, \sigma_3)$, $m_2(y, \sigma_{23})$ and $m_2(y, \sigma_2)$ give rise to the unlabelled edges

$$\begin{aligned}
x \otimes x_i \otimes v_1 &\to z \otimes x_i \otimes y_1 &&\text{for } 1 \le i \le k, \\
y \otimes x_i \otimes y_j &\to z \otimes x_i \otimes y_{j+1} &&\text{for } 1 \le i \le k, 1 \le j \le l-1, \\
y \otimes x_i \otimes y_l &\to x \otimes x_i \otimes v_2 &&\text{for } 1 \le i \le k.
\end{aligned}$$

We can cancel these unlabelled edges using the edge reduction algorithm described in Section 2.3. We cancel them in order of increasing i and, for fixed i, in the order above. It is not difficult to check that the only additional incoming arrows at $z \otimes x_i \otimes y_j$ and $z \otimes x_i \otimes v_2$ are given by

$$\begin{aligned}
z \otimes x_i \otimes y_j &\xrightarrow{\tau_{23}} z \otimes x_i \otimes y_{j+1} &&\text{for } 1 \le i \le k, 1 \le j \le l-1, \\
z \otimes x_i \otimes y_l &\xrightarrow{\tau_{23}} x \otimes x_i \otimes v_2 &&\text{for } 1 \le i \le k, \\
y \otimes x_{i-1} \otimes y_j &\to z \otimes x_i \otimes y_j &&\text{for } 1 \le i \le k, 1 \le j \le l-1, \\
x \otimes x_{i-1} \otimes y_l &\to x \otimes x_i \otimes v_2 &&\text{for } 1 \le i \le k.
\end{aligned}$$

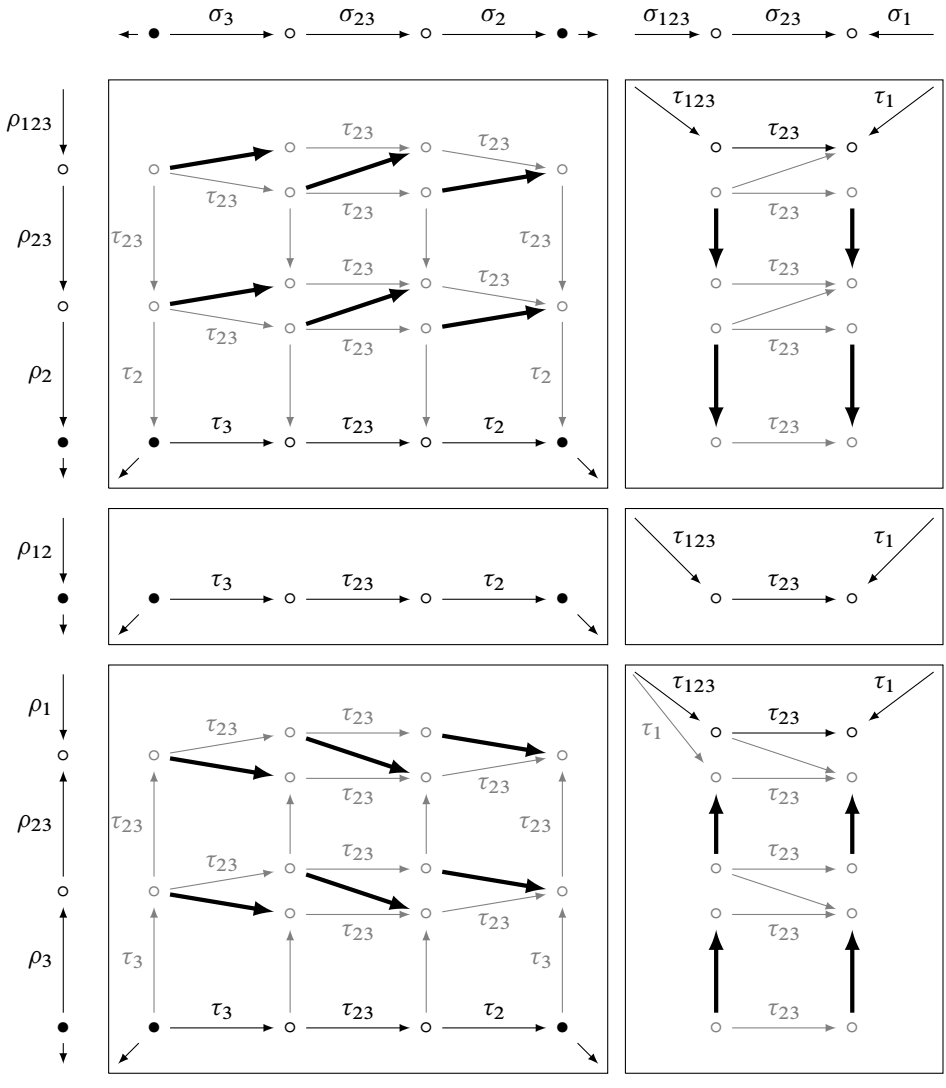


Figure 10: Each box contains the portion of $\widehat{CFDAA}(S^1\times\mathcal{P})\boxtimes(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ coming from a d_2, d_0 or d_{-2} segment in $\boldsymbol{\ell}_1$ (top, middle and bottom, respectively) and an a_2 or b_2 segment in $\boldsymbol{\ell}_2$ (left and right, respectively). Thick unlabelled arrows can be removed with the edge reduction algorithm; gray indicates generators and arrows that are eliminated when the differentials are cancelled.

It follows that each time we use the edge reduction algorithm on one of the unlabelled arrows mentioned above, there are no other incoming arrows at the terminal vertex that have not already been cancelled, and so cancelling the arrow produces no new

arrows. After cancelling all of the unlabelled arrows, the only remaining generators are $v \otimes u_2 \otimes v_1$, $v \otimes u_2 \otimes v_2$, and $v \otimes u_2 \otimes y_j$ for $1 \leq j \leq l$. Arrows between these generators arise from trimodule operations involving only the generators v and w and with no ρ inputs; there are only three:

$$m_2(v, \sigma_3) = \tau_3 \otimes w, \quad m_2(w, \sigma_{23}) = \tau_{23} \otimes w \quad \text{and} \quad m_2(w, \sigma_2) = \tau_2 \otimes v.$$

It follows that the only arrows in the portion of $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ coming from the segments d_k and a_l are

$$\begin{aligned} v \otimes u_2 \otimes v_1 &\xrightarrow{\tau_3} w \otimes u_2 \otimes y_1, \\ w \otimes u_2 \otimes y_j &\xrightarrow{\tau_{23}} w \otimes u_2 \otimes y_{j+1} \quad \text{for } 1 \leq j \leq l-1, \\ w \otimes u_2 \otimes y_l &\xrightarrow{\tau_2} v \otimes u_2 \otimes v_2. \end{aligned}$$

That is, the portion of $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ coming from the segments d_k and a_l is an a_l segment from $v \otimes u_2 \otimes v_1$ to $v \otimes u_2 \otimes v_2$.

Finally, note that there can be no arrows connecting $w \otimes u_2 \otimes y_j$ to any other portions of $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ arising from different segments, since the only arrow connecting u_2 to a generator in a different segment of $\boldsymbol{\ell}_1$ is an outgoing ρ_1, ρ_{12} or ρ_{123} arrow, and the only trimodule operations involving these inputs also have σ_1, σ_{12} or σ_{123} as an input. The outgoing arrows from u_2 and from v_1 and v_2 do give rise to additional arrows out of $v \otimes u_2 \otimes v_1$ and $v \otimes u_2 \otimes v_2$; these show up in the portion of $\widehat{CFDAA}(\mathcal{Y}_{\mathcal{P}}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ coming from the segment following d_k in $\boldsymbol{\ell}_1$ and the segment following or preceding a_l in $\boldsymbol{\ell}_2$.

For other pairs of segments in $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$, the proof is similar. Figure 10 depicts the relevant portion of $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ for d_2, d_0 or d_{-2} paired with a_2 or b_2. Any d_k paired with a_l or b_l behaves like one of these cases, depending on the sign of k. Note that, if $l < 0$, we simply take the mirror image of these diagrams, since $a_{-l} = \bar{a}_l$. This proves (1) and (2).

(3) can be deduced from (4) by observing that a c_l segment from v_1 to v_2 is the same as a d_{-l} segment from v_2 to v_1. To prove (4), consider pairing d_k in $\boldsymbol{\ell}_1$ with d_l in $\boldsymbol{\ell}_2$. The behaviour depends on the sign of k and l; Figures 11 and 12 depict the cases with k in $\{2, 0, -2\}$ and l in $\{3, 0, -3\}$. If $k = 0$, it is clear that the result is a segment of type d_l, and similarly, if $l = 0$, the result is a segment of type d_k. The case of k and l positive behaves like the top left box in Figure 11, and the case of k and l negative behaves like the bottom box in Figure 12. In each case all generators cancel except for those along the top and right edges, resulting in a segment of type d_{k+l}.

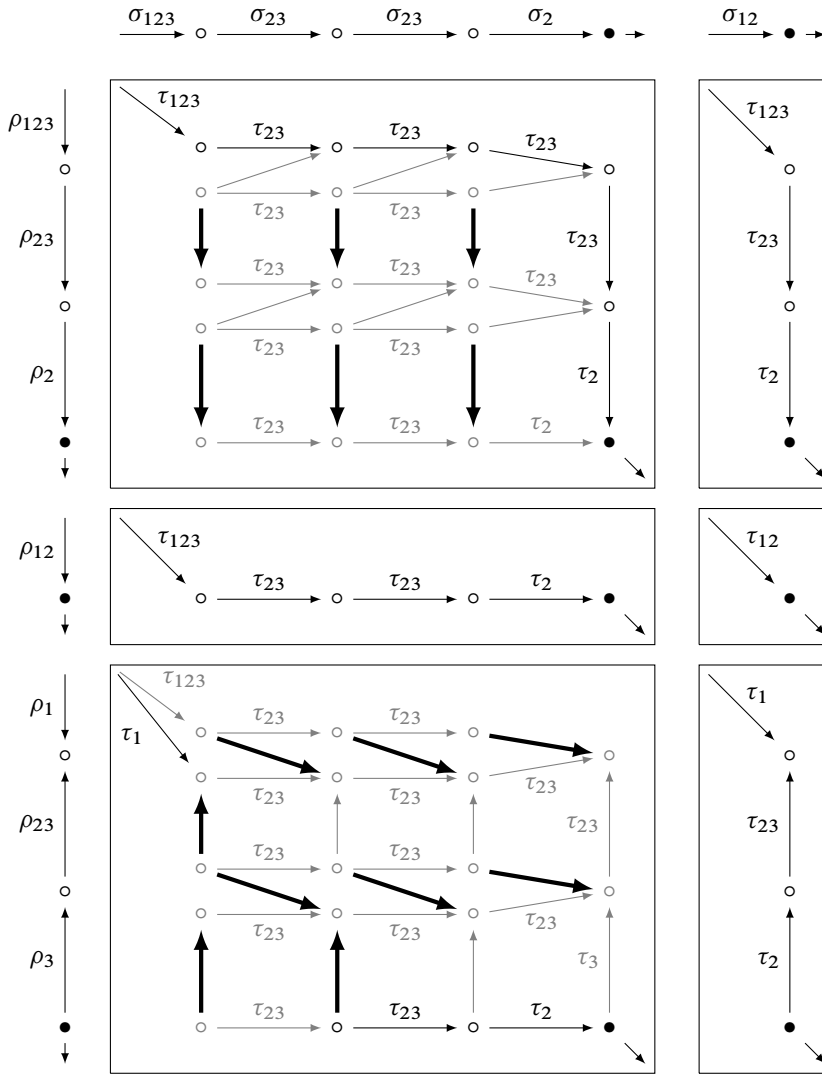


Figure 11: The portions of $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ coming from a d_2, d_0 or d_{-2} segment in $\boldsymbol{\ell}_1$ and a d_3 or d_0 segment in $\boldsymbol{\ell}_2$. Thick unlabelled arrows can be removed with the edge reduction algorithm; gray indicates generators and arrows that are eliminated when the differentials are cancelled.

The case that k and l have opposite signs is slightly more complicated. Assume first that k is negative and l is positive. If $k \geq -l$, the resulting complex looks like the bottom left box in Figure 11. Note that starting from the top left corner, there is a path to the bottom right corner consisting of a τ_1 arrow, an odd length "zigzag" sequence

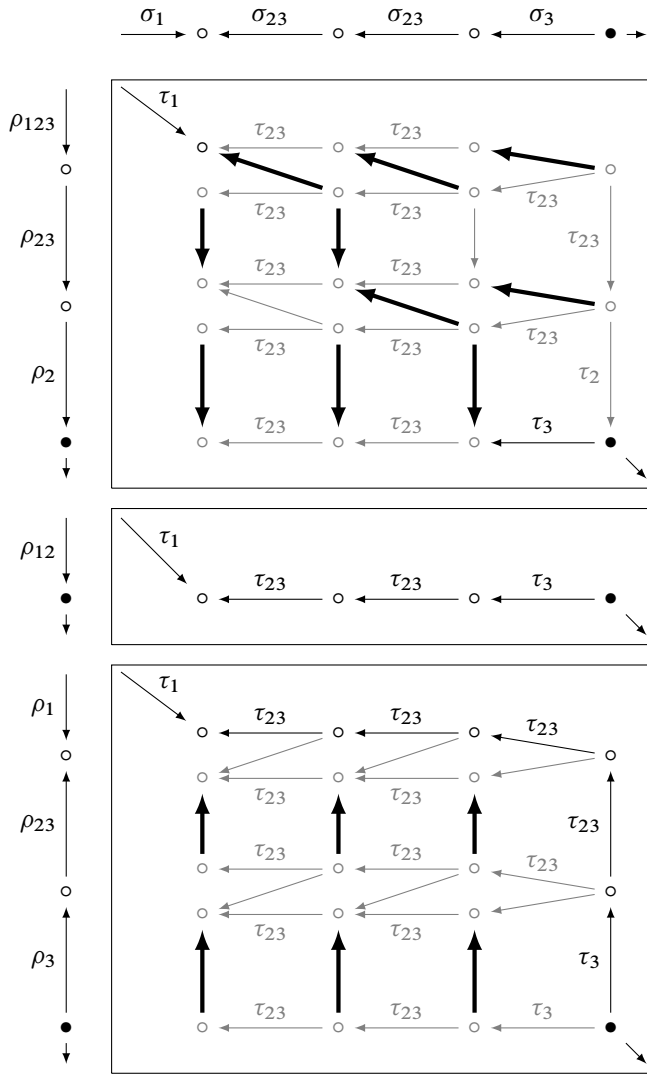


Figure 12: Each box contains the portion of $\widehat{CFDAA}(S^1\times\mathcal{P})\boxtimes(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ coming from a d_2, d_0 or d_{-2} segment in $\boldsymbol{\ell}_1$ and a d_{-3} segment in $\boldsymbol{\ell}_2$. Thick unlabelled arrows can be removed with the edge reduction algorithm; gray indicates generators and arrows that are eliminated when the differentials are cancelled.

of unlabelled arrows, $k+l$ τ_{23} arrows and a τ_2 arrow. Everything in the diagram not involved in this sequence can be cancelled without adding new arrows. Cancelling the remaining unlabelled arrows turns the τ_1 arrow and the first τ_{23} arrow into a τ_{123} arrow if $k+l>0$, or it turns the τ_1 arrow and τ_2 arrow into a τ_{12} arrow if $k+l=0$. The

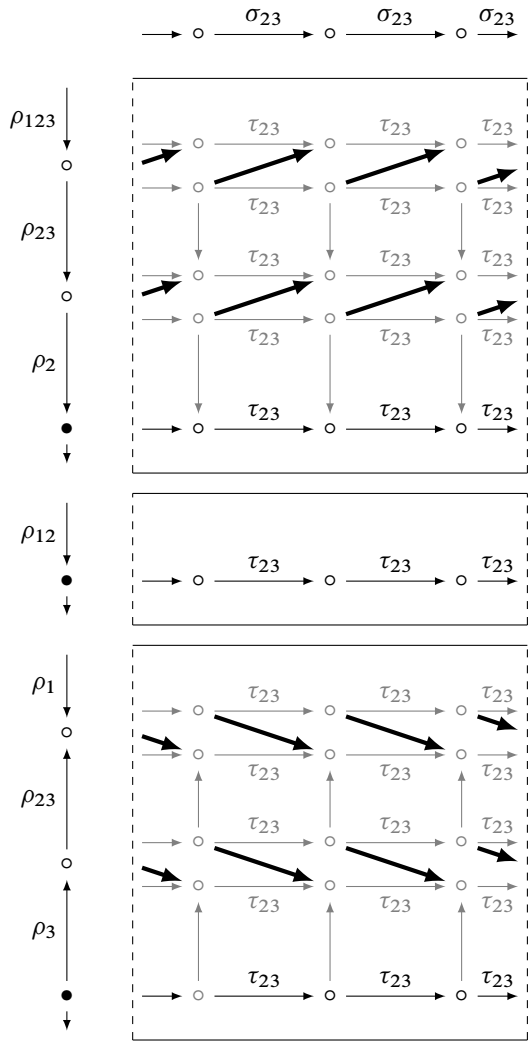


Figure 13: The portions of $\widehat{CFDAA}(S^1 \times \mathcal{P}) \boxtimes (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ coming from a d_2, d_0 or d_{-2} segment in $\boldsymbol{\ell}_1$ when $\boldsymbol{\ell}_2$ is a collection of e^* segments. The left and right edges of each box should be identified. Thick unlabelled arrows can be removed with the edge reduction algorithm; gray indicates generators and arrows that are eliminated when the differentials are cancelled. The result is a copy of $\boldsymbol{\ell}_2$ for each ι_0–generator in $\boldsymbol{\ell}_1$.

result is a segment of type d_{k+l}. The case that $k < -l$ is slightly different. It is not pictured separately, but the main difference is that the zigzag sequence of unlabelled arrows starting at the end of the τ_1 arrow has even length and ends on the right side of

the diagram instead of the bottom. It is then followed by $-k-l-1$ backwards τ_{23} arrows and a backwards τ_3 arrow. Everything not involved in this sequence cancels, and removing the unlabelled arrows produces a segment of type d_{k+l} from the top left corner to the bottom right corner.

To complete the proof of (4), consider the case that k is positive and l is negative. If $l<-k$, the resulting complex looks like the top box in Figure 12. The complex reduces to a τ_1 arrow followed by a chain of unlabelled arrows, $-k-l-1$ backwards τ_{23} arrows and a backwards τ_3 arrow. This further reduces to a chain of type d_{k+l}. If instead $l\geq -k$, the chain of unlabelled arrows following the τ_1 arrow ends on the right side of the diagram instead of the bottom. Again, the complex reduces to a chain of type d_{k+l}.

Finally, we must consider the case the case that $\boldsymbol{\ell}_2$ cannot be written in standard notation. In this case, $\boldsymbol{\ell}_2$ is a collection of e^* segments. Note that in this case we can ignore all operations in Table 1 that involve v, x, or σ inputs other than σ_{23} (this only leaves seven operations). It is easy to see that the complex $\widehat{CFDAA}(S^1\times\mathcal{P})\boxtimes(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ collapses to a copy of $\boldsymbol{\ell}_2$ for each d_k segment in $\boldsymbol{\ell}_1$ (see Figure 13). □

In practice, it is helpful to compute $\mathrm{M}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ by creating an i by j grid, where i is the length of $\boldsymbol{\ell}_1$ and j is the length of $\boldsymbol{\ell}_2$. The (i,j) entry of this grid is a square containing a single segment connecting two of its corners, as dictated by Proposition 6.4 (see, for example, Figure 14). The collection of loops $\mathrm{M}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ can now be read off the grid by identifying the top and bottom edges and the left and right edges. Note that the number of disjoint loops in $\mathrm{M}(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)$ is given by $\gcd(i,j')$, where i is the number of d_k segments in $\boldsymbol{\ell}_1$ (by assumption, this is the length of $\boldsymbol{\ell}_1$) and j' is the number of d_k segments in $\boldsymbol{\ell}_2$ minus the number of c_k segments in $\boldsymbol{\ell}_2$ (if $j'=0$, we use the convention that $\gcd(i,0)=i$; our assumptions rule out the case that $i=0$). Note that i and j' can be given in terms of the $(\mathbb{Z}/2\mathbb{Z})$–grading on $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$: up to sign, we have $i=\chi_\bullet(\boldsymbol{\ell}_1)$ and $j'=\chi_\bullet(\boldsymbol{\ell}_2)$. To see this, observe that, for an appropriate choice of relative grading, each d_k segment contributes 1 to $\chi_\bullet$, each c_k segment contributes -1 and each a_k or b_k segment contributes 0.

Lemma 6.5 *Consider a pair of loop-type, bordered, rational homology solid tori (M_1,α_1,β_1) and (M_2,α_2,β_2). If the loops representing $\widehat{CFD}(M_1,\alpha_1,\beta_1)$ contain only standard unstable chains then*

$$\mathcal{M}_{1,2}=\mathcal{M}((M_1,\alpha_1,\beta_1),(M_2,\alpha_2,\beta_2))$$

is of loop-type. If, in addition, the loops representing $\widehat{CFD}(M_2,\alpha_2,\beta_2)$ contain only standard unstable chains, then $\mathcal{M}_{1,2}$ is of simple loop-type.

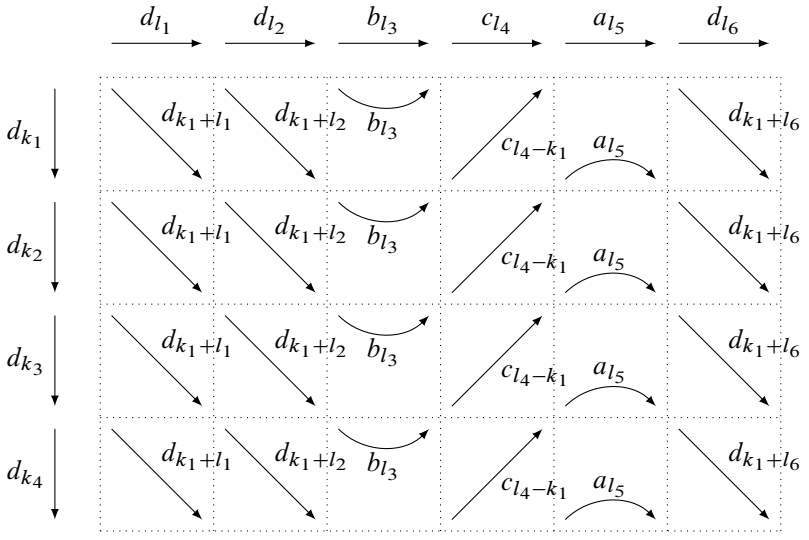


Figure 14: Computing $\mathrm{M}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ for two loops $\boldsymbol{\ell}_1 = (d_{k_1} d_{k_2} d_{k_3} d_{k_4})$ and $\boldsymbol{\ell}_2 = (d_{l_1} d_{l_2} b_{l_3} c_{l_4} a_{l_5} d_{l_6})$.

Proof If $\widehat{CFD}(M_1, \alpha_1, \beta_1)$ and $\widehat{CFD}(M_2, \alpha_2, \beta_2)$ are collections of loops and the loops in $\widehat{CFD}(M_1, \alpha_1, \beta_1)$ contain only standard unstable chains, then, by Proposition 6.4, $\widehat{CFD}(\mathcal{M}_{1,2})$ is a collection of loops. Moreover, if the loops in $\widehat{CFD}(M_2, \alpha_2, \beta_2)$ also contain only standard unstable chains, Proposition 6.4 implies that the resulting loops in $\widehat{CFD}(\mathcal{M}_{1,2})$ contain only standard unstable chains, and in particular are simple. It only remains to check that $\widehat{CFD}(\mathcal{M}_{1,2})$ has exactly one loop for each spin^c–structure on $\mathcal{M}_{1,2}$.

Recall that the operation $\mathcal{M}$ corresponds to gluing two bordered manifolds to two of the three boundary components of $\mathcal{P} \times S^1$, where $\mathcal{P}$ is S^2 with three disks removed. Thus $\partial(S^1 \times \mathcal{P})$ has three connected components; denote the i^{th} connected component by $\partial(S^1 \times \mathcal{P})_i$. For each $i \in \{1, 2, 3\}$ let f_i denote a curve in $\partial(S^1 \times \mathcal{P})_i$ which is a fibre $\{\text{pt}\} \times S^1$ and let b_i denote the relevant component of $\partial\mathcal{P} \times \{\text{pt}\}$. According to the conventions introduced in Section 6.1, applying the operation $\mathcal{M}$ to two bordered manifolds (M_1, α_1, β_1) and (M_2, α_2, β_2) corresponds to gluing M_1 and M_2 to $\mathcal{P} \times S^1$ by identifying α_1 with f_1, β_1 with b_1, α_2 with f_2 and β_2 with b_2. The result is the manifold $M_3 = \mathcal{M}_{1,2}$.

We consider homology groups with coefficients in $\mathbb{Z}$. The spin^c–structures on a 3–manifold M with boundary are indexed by $H^2(M) \cong H_1(M, \partial M)$. Using the

appropriate Mayer–Vietoris sequences, we have that $H_1(M_3, \partial M_3)$ is homomorphic to the quotient

$$H_1(S^1\times\mathcal{P}, \partial(S^1\times\mathcal{P})_3)\oplus H_1(M_1)\oplus H_1(M_2)/\{\alpha_1\sim f_1,\, \beta_1\sim b_1,\, \alpha_2\sim f_2,\, \beta_2\sim b_2\}.$$

Note that $H_1(S^1\times\mathcal{P}, \partial(S^1\times\mathcal{P})_3)$ is generated by f_i and b_i for $i\in\{1,2,3\}$ with the relations $f_1 = f_2 = f_3$ and $b_1 + b_2 = b_3 = 0$. For $i = 1, 2$, since M_i is a rational solid torus there is a unique (possibly disconnected) curve in ∂M_i that bounds a surface in M_i, to which we associate p_i/q_i (where p_i and q_i may not be relatively prime), so that $p_i\alpha_i + q_i\beta_i$ generates the kernel of the inclusion of $H_1(\partial M_i)$ into $H_1(M_i)$. The long exact sequence for relative homology gives

$$H_1(M_i) \cong H_1(M_i, \partial M_i)\oplus H_1(\partial M_i)/\langle p_i\alpha_i + q_i\beta_i\rangle.$$

In $H_1(M_3, \partial M_3)$, the relation $p_i\alpha_i + q_i\beta_i = 0$ translates to $p_i f_i + q_i b_i = q_i b_i = 0$. It follows that

$$H_1(M_3, \partial M_3) \cong H_1(M_1, \partial M_1)\oplus H_1(M_2, \partial M_2)\oplus\langle b_1\rangle/(q_1b_1 = q_2b_1 = 0).$$

Thus, for each spin^c–structure on M_1 and each spin^c–structure on M_2, there are $\gcd(q_1, q_2)$ spin^c–structures on M_3. Note that q_2 may be 0; in this case we use the convention that $\gcd(q, 0) = q$. The assumption that $\widehat{CFD}(M_1, \alpha_1, \beta_1)$ contains only standard unstable chains implies that $q_1 \neq 0$.

Recall that the rational longitude of a rational homology solid torus can be read off of the bordered invariants. More precisely, by Proposition 2.10, the curves in ∂M_i which are nullhomologous in M_i are integer multiples of

$$\chi_\circ(\widehat{CFD}(M_i, \alpha_i, \beta_i; \mathfrak{s}_i))\alpha_i + \chi_\bullet(\widehat{CFD}(M_i, \alpha_i, \beta_i; \mathfrak{s}_i))\beta_i,$$

where $\mathfrak{s}_i$ is any spin^c–structure on M_i. Thus $q_i = \chi_\bullet(\widehat{CFD}(M_i, \alpha_i, \beta_i; \mathfrak{s}_i))$. By assumption, $\widehat{CFD}(M_i, \alpha_i, \beta_i; \mathfrak{s}_i)$ is a loop for each $\mathfrak{s}_i$. It was observed above that, given two loops $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ with $\boldsymbol{\ell}_1$ consisting only of standard unstable chains, $\mathrm{M}(\boldsymbol{\ell}_1, \boldsymbol{\ell}_2)$ is a collection of $\gcd(\chi_\bullet(\boldsymbol{\ell}_1), \chi_\bullet(\boldsymbol{\ell}_2))$ loops. Thus, for each loop in $\widehat{CFD}(M_1, \alpha_1, \beta_1)$ and for each loop in $\widehat{CFD}(M_2, \alpha_2, \beta_2)$, there are $\gcd(q_1, q_2)$ loops in $\widehat{CFD}(\mathcal{M}_{1,2})$. It follows that $\widehat{CFD}(\mathcal{M}_{1,2})$ has exactly one loop for each spin^c–structure on M_3 (where $\mathcal{M}_{1,2} \cong (M_3, \alpha_3, \beta_3)$ as bordered manifolds). □

6.4 Bordered invariants of graph manifolds

With this description of $\mathcal{T}$, $\mathcal{E}$ and $\mathcal{M}$ on bordered invariants in hand, we may now return our focus to graph manifolds. The first author has described (and implemented)

an algorithm for computing the (bordered) Heegaard Floer invariants of graph manifolds [8]; we will now outline a version of this algorithm for graph manifolds with a single boundary component and adapt it to the loops setup. Recall that, given a graph Γ with associated bordered manifold $(M_\Gamma, \alpha, \beta)$ (as described in Section 6.1), we write $\widehat{CFD}(\Gamma)$ for $\widehat{CFD}(M_\Gamma, \alpha, \beta)$.

In order to compute $\widehat{CFD}(\Gamma)$, we inductively build up the plumbing tree Γ using the three plumbing tree operations $\mathcal{T}^{\pm 1}$, $\mathcal{E}$ and $\mathcal{M}$ depicted in Figure 8, starting from the plumbing tree

$$\Gamma_0 = \overset{0}{\bullet}\text{- - - -}$$

Note that M_{Γ_0} is a solid torus, with bordered structure $(M_{\Gamma_0}, \alpha, \beta)$, where α is a fibre of the S^1–bundle over D^2 (ie a longitude of the solid torus) and β is a curve in the base surface (ie a meridian of the solid torus, identified by $\partial D^2 \times \{\text{pt}\}$). Thus $\widehat{CFD}(\Gamma_0)$ is represented by the loop $\boldsymbol{\ell}_\bullet \sim (d_0)$. As we apply the operations $\mathcal{T}$, $\mathcal{E}$ and $\mathcal{M}$, we keep track of the bordered invariants of the relevant manifolds. Let Γ_1 and Γ_2 be single-boundary plumbing trees. As shown in the previous section,

$$\begin{aligned}
\widehat{CFD}(\mathcal{T}(\Gamma_1)) &\cong \hat{T}_{\text{st}} \boxtimes \widehat{CFD}(\Gamma_1),\\
\widehat{CFD}(\mathcal{E}(\Gamma_1)) &\cong \hat{T}_{\text{st}} \boxtimes \hat{T}_{\text{du}} \boxtimes \hat{T}_{\text{st}} \boxtimes \widehat{CFD}(\Gamma_1),\\
\widehat{CFD}(\mathcal{M}(\Gamma_1, \Gamma_2)) &\cong \widehat{CFDAA}(\Gamma_{\mathcal{M}}) \boxtimes (\widehat{CFD}(\Gamma_1), \widehat{CFD}(\Gamma_2)).
\end{aligned}$$

Specializing Lemmas 6.2 and 6.5 to graph manifold leads to the following observation:

Lemma 6.6 *Let Γ be a single boundary plumbing tree constructed as described above from copies of Γ_0 using the operations $\mathcal{T}^{\pm}$, $\mathcal{E}$, and $\mathcal{M}$. $\widehat{CFD}(\Gamma)$ has simple loop-type as long as, each time the operation $\mathcal{M}$ is applied, the two input plumbing trees have simple loop-type bordered invariants with only unstable chains in standard notation.*

Proof This follows from Lemma 6.2, which says that $\mathcal{T}^{\pm}$ and $\mathcal{E}$ take simple loop-type manifolds to simple loop-type manifolds, and Lemma 6.5, which says that $\mathcal{M}$ takes two simple loop-type manifolds to a simple loop-type manifold provided the simple loops corresponding to both inputs consist only of unstable chains. □

With this observation, we can describe a large family of simple loop-type manifolds.

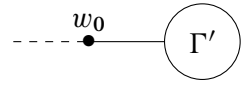
For each vertex v of a plumbing tree Γ, let $w(v)$ denote the Euler weight associated to v and let $n_+(v)$ and $n_-(v)$ denote the number of neighbouring vertices v' for which $w(v') \geq 0$ and $w(v') \leq 0$, respectively. We will say that v is a *bad vertex* if

$-n_-(v) < w(v) < n_+(v)$, and otherwise v is a *good vertex* (this should be viewed as a generalization of the notion of bad vertices defined for negative definite plumbing trees in [23]).

Proposition 6.7 *Let Γ be a plumbing tree with a single boundary edge at the vertex v_0, and suppose that every vertex other than v_0 is good. Then $\widehat{CFD}(\Gamma)$ is a collection of loops consisting only of standard unstable chains; up to reversal we can assume these unstable chains are of type d_k. Moreover, if $w(v_0)$ is (strictly) greater than $n_+(v_0)$, then these unstable chains all have subscripts (strictly) greater than 0. If $w(v_0)$ is (strictly) less than $n_-(v_0)$ then the unstable chains all have subscript (strictly) less than 0.*

Proof We proceed by induction on the number of vertices of Γ. The base case, where Γ has only one vertex v_0, is trivial; $\widehat{CFD}(\Gamma)$ in this case is given by the loop $(d_{w(v_0)})$.

For the inductive step, first suppose the boundary vertex v_0 of Γ has valence two (including the boundary edge). That is, Γ has the form

$\Gamma =$ 

where $w_0 = w(v_0)$. Let v_1 be the boundary vertex of Γ', and let w_1 be the corresponding weight. Let $n_\pm(v_1)$ denote the counts of neighbouring vertices defined above for v_1 as a vertex in Γ, and let $n'_\pm(v_1)$ denote these counts for v_1 as a vertex in Γ' (ie ignoring the vertex v_0). Since v_1 is a good vertex, we have one of the following two cases:

(1) $w_1 \le -n_-(v_1) \le -n'_-(v_1) \le 0$.

(2) $w_1 \ge n_+(v_1) \ge n'_+(v_1) \ge 0$.

Note that $\Gamma = \mathcal{T}^{w_0}(\mathcal{E}(\Gamma'))$, and so $\widehat{CFD}(\Gamma)$ is given by $\mathrm{T}^{w_0}\big(\mathrm{E}(\widehat{CFD}(\Gamma'))\big)$. Furthermore, we assume by induction that the proposition holds for Γ'.

In case (1), we have that $\widehat{CFD}(\Gamma')$ consists of d_k segments with nonpositive subscripts. Thus in dual notation $\mathrm{E}(\widehat{CFD}(\Gamma'))$ is a loop consisting of d_k^* segments with nonnegative subscripts, and in standard notation $\mathrm{E}(\widehat{CFD}(\Gamma'))$ consists of d_k segments with nonnegative subscripts. Moreover, if $w_0 \le 0$, then w_1 is strictly less than $-n'_-(v_1) = -n_-(v_1) + 1$, so the d_k segments in $\widehat{CFD}(\Gamma)$ have strictly negative subscripts, the d_k^* segments in $\mathrm{E}(\widehat{CFD}(\Gamma'))$ have strictly positive subscripts, and in standard notation $\mathrm{E}(\widehat{CFD}(\Gamma'))$ consists only of d_0 and d_1 segments. It follows that

$\widehat{CFD}(\Gamma)$ is a collection of loops consisting of d_k segments with subscripts satisfying $k<0$ if $w_0<-n_-(v_0)=-1$, $k\le 0$ if $w_0=n_-(v_0)=-1$, $k\ge 0$ if $w_0=n_+(v_0)\ge 0$ and $k>0$ if $w_0>n_+(v_0)\ge 0$. Thus the proposition holds for Γ.

In case (2), $\widehat{CFD}(\Gamma')$ consists of d_k segments with nonnegative subscripts. Thus $\mathrm{E}(\widehat{CFD}(\Gamma'))$ consists of d_k^* with nonpositive subscripts in dual notation and of d_k with nonpositive subscripts in standard notation. Moreover, if $w_0\ge 0$ then $n_+(v_1)>n'_+(v_1)$, so the d_k segments in $\widehat{CFD}(\Gamma')$ have strictly positive subscripts, the d_k^* segments in $\mathrm{E}(\widehat{CFD}(\Gamma'))$ have strictly negative subscripts and the d_k segments in $\mathrm{E}(\widehat{CFD}(\Gamma'))$ have subscripts in $\{0,-1\}$. It follows that $\widehat{CFD}(\Gamma)$ is a collection of loops consisting of d_k segments with subscripts satisfying $k<0$ if $w_0<-n_-(v_0)\le 0$, $k\le 0$ if $w_0=n_-(v_0)\le 0$, $k\ge 0$ if $w_0=n_+(v_0)=1$ and $k>0$ if $w_0>n_+(v_0)=1$. Thus the proposition holds for Γ.

Now suppose that v_0 has valence higher than two. This means that Γ can be obtained as $\mathcal{M}(\Gamma_1,\Gamma_2)$, where Γ_1 and Γ_2 have fewer vertices than Γ. By induction, the proposition holds for Γ_1 and Γ_2, and so $\widehat{CFD}(\Gamma_1)$ and $\widehat{CFD}(\Gamma_2)$ may be represented by a collection of loops consisting only of standard type d_k chains. By Proposition 6.4, merging two (collections of) loops with only d_k chains produces a new collection of loops with only d_k chains. Moreover, each chain in the $\widehat{CFD}(\Gamma)$ is of the form d_{k+l} for some chains d_k in $\widehat{CFD}(\Gamma_1)$ and d_l in $\widehat{CFD}(\Gamma_2)$, so the maximum (resp. minimum) subscript in $\widehat{CFD}(\Gamma)$ is the sum of the maximum (resp. minimum) subscripts in $\widehat{CFD}(\Gamma_1)$ and $\widehat{CFD}(\Gamma_2)$.

For $i\in\{0,1\}$, let v_i be the boundary vertex of Γ_i and let $w_i=w(v_i)$ be the corresponding weight. We can choose any values for w_i provided $w_1+w_2=w_0$. If $w_0\ge n_+(v_0)=n_+(v_1)+n_+(v_2)$, then we can choose $w_1=n_+(v_1)$ and $w_2\ge n_+(v_2)$. By the inductive assumption, $\widehat{CFD}(\Gamma_1)$ and $\widehat{CFD}(\Gamma_2)$ both have only nonnegative subscripts, so the same is true of $\widehat{CFD}(\Gamma_2)$. Furthermore, if $w_0>0$ then $w_2>n_+(v_2)$, so $\widehat{CFD}(\Gamma_2)$ has only strictly positive subscripts and the same follows for $\widehat{CFD}(\Gamma)$. Similarly, if w_0 is (strictly) less than $n_-(v_0)=n_-(v_1)+n_-(v_2)$, we can choose $w_1=n_-(v_1)$ and w_2 (strictly) less than $n_-(v_2)$ and conclude that $\widehat{CFD}(\Gamma)$ only has subscripts (strictly) less than 0. □

Note that if Γ is a plumbing tree with at most one bad vertex, we can compute $\widehat{HF}(M_\Gamma)$ by taking the dual filling of $\widehat{CFD}(\Gamma')$, where Γ' is obtained from Γ by adding a boundary edge to the bad vertex, or to any vertex if there are no bad vertices. $\widehat{CFD}(\Gamma')$, which has simple loop-type by Proposition 6.7, can be computed using the operations $\mathrm{T}^{\pm 1}$, E and M. In particular, we have the following generalization of [23, Lemma 2.6]:

Corollary 6.8 *If Γ is a closed plumbing tree with no bad vertices, then the manifold M_Γ is an L–space.*

Proof Add a boundary edge to any vertex of Γ to produce a single boundary plumbing tree Γ'. By Proposition 6.7, $\widehat{CFD}(\Gamma')$ is a collection of loops consisting only of type d_k segments with either all subscripts greater than or equal to zero or all subscripts less than or equal to zero. By Observation 3.9, it follows that in dual notation the loops representing $\widehat{CFD}(\Gamma')$ have no stable chains. M_Γ is obtained from $M_{\Gamma'}$ by dual filling, and dual filling is an L–space if there are no stable chains in dual notation. □

Recall that in the context of Theorem 1.3 it is important to distinguish solid torus-like manifolds from other manifolds of simple loop-type. Toward that end, we check the following:

Proposition 6.9 *Let Γ be a plumbing tree with a single boundary edge at the vertex v_0, and suppose that every vertex other than v_0 is good. Then M_Γ is solid torus-like if and only if it is a solid torus.*

Proof We proceed by induction on the size of Γ as in the proof of Proposition 6.7. Applying Dehn twists does not change whether or not a manifold is a solid torus, nor does it change whether or not the corresponding bordered invariants are solid torus-like. It follows that the proposition holds for $\mathcal{T}^{\pm}(\Gamma)$ and $\mathcal{E}(\Gamma)$ if it holds for Γ. We only need to check that it holds for $\mathcal{M}(\Gamma_1, \Gamma_2)$, assuming by induction that it holds for Γ_1 and Γ_2. By Proposition 6.7, we may assume that the loops representing $\widehat{CFD}(\Gamma_1)$ and $\widehat{CFD}(\Gamma_2)$ consist only of d_k chains.

Suppose $\widehat{CFD}(\mathcal{M}(\Gamma_1, \Gamma_2))$ is solid torus-like. Recall that, by the appropriate generalization of Lemma 3.15, the loops representing $\widehat{CFD}(\mathcal{M}(\Gamma_1, \Gamma_2))$ consist only of d_k segments such that the maximum and minimum subscripts appearing differ by at most one. Note also that the difference between maximum and minimum subscripts is additive under $\mathcal{M}$. It follows that either Γ_1 or Γ_2 must have bordered invariant consisting only of d_n segments for a fixed $n \in \mathbb{Z}$; say Γ_2 has this property. M_{Γ_2} is solid torus-like; by induction, M_{Γ_2} is a solid torus, and $\widehat{CFD}(\Gamma_2)$ is given by the single loop (d_n). Proposition 6.4 then implies that $\widehat{CFD}(\mathcal{M}(\Gamma_1, \Gamma_2))$ is obtained from $\widehat{CFD}(\Gamma_1)$ by adding n to the subscript of each segment. In other words, $\widehat{CFD}(\mathcal{M}(\Gamma_1, \Gamma_2)) \cong \widehat{CFD}(\mathcal{T}^n(\Gamma_1))$.

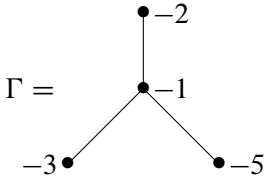
This equivalence does not only hold on the level of bordered invariants. Indeed, $\mathcal{M}(\Gamma_1, \Gamma_2)$ is equivalent to $\mathcal{T}^n(\Gamma_1)$ up to the graph moves in [22]; in particular, the

corresponding graph manifolds are diffeomorphic. To see this, note that Γ_2 must be equivalent to the plumbing tree

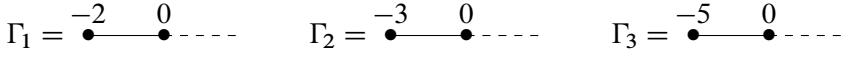
since $\widehat{CFD}$ of this tree is (d_n), and it is clear from Figure 8 that merging with this tree has the same effect as applying $\mathcal{T}^n$. Since $\widehat{CFD}(\mathcal{T}^n(\Gamma_1))$ is solid torus-like, $\widehat{CFD}(\Gamma_1)$ is solid torus-like and, by the inductive hypothesis, M_{Γ_1} is a solid torus. It follows that $M_{\mathcal{M}(\Gamma_1,\Gamma_2)} = M_{\mathcal{T}^n(\Gamma_1)} = \mathcal{T}^n(M_{\Gamma_1})$ is a solid torus. □

6.5 An explicit example: the Poincaré homology sphere

As an example of the algorithm and loop operations described above, we will compute $\widehat{HF}$ of the Poincaré homology sphere using the plumbing tree



We start with the loop (d_0) representing $\widehat{CFD}(\Gamma_0)$. For this example, by abuse of notation, we will equate the bordered invariants with their loop representatives; thus, $\widehat{CFD}(\Gamma_0) = (d_0)$. We use the twist and extend operations to compute invariants for the plumbing trees



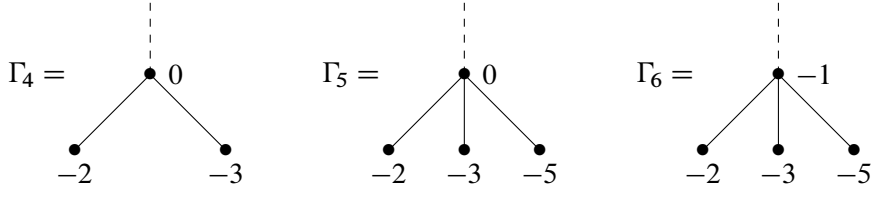
Note that $\Gamma_1 = \mathcal{E}(\mathcal{T}^{-2}(\Gamma_0))$, so

$$\widehat{CFD}(\Gamma_1) = \mathrm{E}\big(\mathrm{T}^{-2}((d_0))\big) = \mathrm{E}((d_{-2})) = (d_2^*) \sim (d_1 d_0).$$

Similarly, we find that

$$\widehat{CFD}(\Gamma_2) = \mathrm{E}\big(\mathrm{T}^{-3}((d_0))\big) = \mathrm{E}((d_{-3})) = (d_3^*) \sim (d_1 d_0 d_0),$$
$$\widehat{CFD}(\Gamma_3) = \mathrm{E}\big(\mathrm{T}^{-5}((d_0))\big) = \mathrm{E}((d_{-5})) = (d_5^*) \sim (d_1 d_0 d_0 d_0 d_0).$$

Now let



We have that $\Gamma_4 = \mathcal{M}(\Gamma_1, \Gamma_2)$, $\Gamma_5 = \mathcal{M}(\Gamma_4, \Gamma_3)$ and $\Gamma_6 = \mathcal{T}^{-1}(\Gamma_5)$. Reading diagonally from the first grid below, we see that $\widehat{CFD}(\Gamma_4) = \mathrm{M}((d_1 d_0), (d_1 d_0 d_0)) = (d_2 d_0 d_1 d_1 d_1 d_0)$:

	d_1	d_0	d_0
d_1	d_2	d_1	d_1
d_0	d_1	d_0	d_0

	d_1	d_0	d_0	d_0	d_0
d_2	d_3	d_2	d_2	d_2	d_2
d_0	d_1	d_0	d_0	d_0	d_0
d_1	d_2	d_1	d_1	d_1	d_1
d_1	d_2	d_1	d_1	d_1	d_1
d_1	d_2	d_1	d_1	d_1	d_1
d_0	d_1	d_0	d_0	d_0	d_0

The second grid tells us that

$$\widehat{CFD}(\Gamma_5) = (d_3 d_0 d_1 d_1 d_1 d_1 d_2 d_0 d_1 d_1 d_2 d_0 d_2 d_0 d_1 d_2 d_1 d_0 d_2 d_0 d_2 d_1 d_1 d_0 d_2 d_1 d_1 d_1 d_1 d_0).$$

Applying the operation T^{-1}, we find that

$$\widehat{CFD}(\Gamma_6) = \big(d_2 d_{-1} d_0 d_0 d_0 d_0 d_1 d_{-1} d_0 d_0 d_1 d_{-1} d_1 d_{-1} d_0 d_1 d_0 d_{-1} d_1 d_{-1} d_1 d_0 d_0 \\ d_{-1} d_1 d_0 d_0 d_0 d_0 d_{-1}\big).$$

Finally, to compute $\widehat{CF}$ of the closed manifold M_Γ we must fill in the boundary of $(M_{\Gamma_6}, \alpha, \beta)$ with a $D^2 \times S^1$ such that the meridian $\partial D^2 \times \{\mathrm{pt}\}$ glues to β and the longitude $\{\mathrm{pt}\} \times S^1$ glues to α. In other words, M_Γ is the 0–filling of $(M_{\Gamma_6}, \alpha, \beta)$, so $\widehat{CF}(M_\Gamma)$ is obtained from $\widehat{CFD}(\Gamma_6)$ by tensoring with $\boldsymbol{\ell}_\circ^A$. To do this, we first write $\widehat{CFD}(\Gamma_6)$ in dual notation (using the procedure described in Section 3.2),

$$\widehat{CFD}(\Gamma_6) = (d_0^* b_1^* a_5^* b_1^* a_3^* b_1^* a_1^* b_1^* a_2^* b_2^* a_1^* b_1^* a_1^* b_3^* a_1^* b_5^* a_1^*).$$

Tensoring with $\boldsymbol{\ell}_\circ$ produces one generator for each segment in dual notation, but also one differential for each type a^* segment. Since $\widehat{CFD}(\Gamma_6)$ has 17 dual segments and 8 a^* segments, all but one generator in $\boldsymbol{\ell}_\circ^A \boxtimes \widehat{CFD}(\Gamma_6) \cong \widehat{CF}(M_\Gamma)$ cancels in homology. Thus $\widehat{HF}(M_\Gamma)$ has dimension 1, as is now well known (this was first calculated in [24, Section 3.2]).

Remark 6.10 While computer computation is not our primary motivation, it is worth pointing out that using loop calculus as described in this section instead of taking box tensor products of modules, bimodules and trimodules greatly improves the efficiency of the algorithm in [8] for rational homology sphere graph manifolds. This is illustrated by the fact that the example above can easily be done by hand, while computing the relevant tensor products would be tedious without a computer. The largest computation

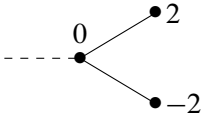
in [8] (for which the dimension of $\widehat{HF}$ is 213 312) took roughly 12 hours; when the computer implementation is adapted to use loop calculus, the same computation runs in 30 seconds.[3] The caveat is that the purely loop calculus algorithm may not work for all rational homology sphere graph manifolds, since there may be graph manifold rational homology solid tori that are not of loop-type, but in practice it works for most examples.

7 L–spaces and non-left-orderability

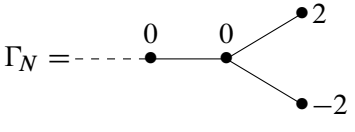
We conclude by proving the remaining results quoted in the introduction: Theorem 1.4, Theorem 1.5 and, finally, Theorem 1.1.

To begin, we observe that Seifert-fibred rational homology tori have simple loop-type bordered invariants. Any Seifert-fibred space over S^2 can be given a star-shaped plumbing tree in which all vertices of valence one or two have weight at most -2; in particular, there are no bad vertices except for the central vertex. Adding a boundary edge to the central vertex corresponds to removing a neighbourhood of a regular fibre, creating a Seifert-fibred space over D^2. By Proposition 6.7, such a plumbing tree has simple loop-type $\widehat{CFD}$. Moreover, by Proposition 6.9, such a plumbing tree is nonsolid torus-like unless the corresponding manifold is a solid torus.

The only other option to consider is a Seifert-fibred space over the Möbius band, since a Seifert-fibred space over any other base orbifold has $b_1 > 0$. Such a manifold can be obtained from a Seifert-fibred space over D^2 by removing a neighbourhood of a regular fibre and gluing in the Euler number 0 bundle over the Möbius band, fibre to fibre and base to base. On the plumbing tree, this corresponds to adding



to the central vertex, or equivalently to merging with the plumbing tree



Proposition 7.1 *$\widehat{CFD}$ for the plumbing tree Γ_N above consists of two loops, one $(a_1 b_1) \sim (d_1^* d_{-1}^*)$ and the other $(e^* e^*)$.*

[3] An implementation is available from the authors upon request.

Proof Using the loop operations described in the preceding section, this is a simple computation. Note that $\Gamma_N = \mathcal{E}\big(\mathcal{M}\big(\mathcal{E}(\mathcal{T}^2(\Gamma_0)), \mathcal{E}(\mathcal{T}^{-2}(\Gamma_0))\big)\big)$. Thus $\widehat{CFD}(\Gamma_N)$ is given by

$$\begin{aligned}
\mathrm{E}\big(\mathrm{M}\big(\mathrm{E}\big(\mathrm{T}^2((d_0))\big), \mathrm{E}\big(\mathrm{T}^{-2}((d_0))\big)\big)\big) &= \mathrm{E}\big(\mathrm{M}\big(\mathrm{E}((d_2)), \mathrm{E}((d_{-2}))\big)\big) \\
&= \mathrm{E}\big(\mathrm{M}((d_{-1}d_0), (d_1d_0))\big) \\
&= \mathrm{E}((d_0d_0) \amalg (d_1d_{-1})) \\
&= (d_0^*d_0^*) \amalg (d_1^*d_{-1}^*). \qquad \square
\end{aligned}$$

Remark 7.2 This result was first established by a direct calculation by Boyer, Gordon and Watson [3]; this calculation is greatly simplified by appealing to loop calculus.

We now complete the proof that Seifert-fibred rational homology tori have simple loop-type.

Proof of Theorem 1.4 As observed above, the case of Seifert-fibred manifolds over D^2 is a special case of Proposition 6.7. For a Seifert-fibred manifold over the Möbius band, a plumbing tree Γ is given by $\mathcal{M}(\Gamma', \Gamma_N)$, where Γ' is a star shaped plumbing tree for a Seifert-fibred manifold over D^2. By Proposition 6.7, the loops in $\widehat{CFD}(\Gamma')$ contain only unstable chains in standard notation. By Propositions 6.4 and 7.1, we find that $\widehat{CFD}(\Gamma)$ is a collection of disjoint copies of $\widehat{CFD}(\Gamma_N)$, and in particular is a collection of simple loops. Moreover, by Lemma 6.5 there is one loop for each spin^c–structure, and so $\widehat{CFD}(\Gamma)$ is of simple loop-type. Finally, with the foregoing in place (in particular Proposition 6.9) one checks that the only solid torus-like manifold in this class is the solid torus itself. $\square$

We are now in a position to assemble the pieces and give the proof of Theorem 1.1. A key observation is the following consequence of our gluing theorem, which provides some alternative characterizations of the set of strict L–space slopes of a given simple loop-type manifold:

Theorem 7.3 *Let M be a simple loop-type manifold that is not solid torus-like. The following are equivalent*:

(i) *γ is a strict L–space slope for M, that is, $\gamma \in \mathcal{L}^\circ_M$.*

(ii) *$M \cup_h M'$ is an L–space, where $h(\gamma) = \lambda$ and M' is a nonsolid torus-like, simple, loop-type manifold with rational longitude λ for which $\mathcal{L}_{M'}$ includes every slope other than λ.*

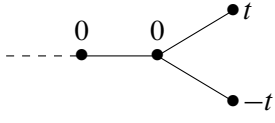
(iii) *$M \cup_h N$ is an L–space, where $h(\gamma) = \lambda$ and N is the twisted I–bundle over the Klein bottle with rational longitude λ.*

Proof Given $\gamma \in \mathcal{L}^\circ_M$ and a simple, loop-type manifold M' for which $\mathcal{L}_{M'}$ includes every slope other than λ, Theorem 5.7 ensures that $M \cup_h M'$ is an L–space. Indeed, for any $\gamma' \neq \gamma$ we have that $\lambda \neq h(\gamma') \in \mathcal{L}^\circ_{M'}$. This proves that (i) implies (ii).

To see that (ii) implies (iii), is suffices to observe that $N(\gamma)$ is an L–space for all γ other than the rational longitude; this is indeed the case, as observed, for example, in [3, Proposition 5] (alternatively, this fact is an exercise in loop calculus).

Finally, suppose that $M \cup_h N$ is an L–space and consider the slope γ in ∂M determined by $h^{-1}(\lambda)$. Since $N(\lambda)$ is not an L–space, by Theorem 5.7 it must be that $\gamma \in \mathcal{L}^\circ_M$, as required, so that (iii) implies (i). □

Boyer and Clay consider a collection of Seifert-fibred rational homology solid tori $\{N_t\}$ for integers $t > 1$. In this collection, $N_2 = N$, the twisted I–bundle over the Klein bottle. More generally, the N_t are examples of Heegaard Floer homology solid tori (see [12, Section 1.5] for a expanded discussion on this class of manifolds). These manifolds are easily described by the plumbing tree



Translated into loop notation, the invariant $\widehat{CFD}(N_t, \varphi, \lambda)$ is simple, described by

$$(d_0^*)^t \amalg \left[\coprod_{i=1}^{t-1} (d_i^* d_{i-t}^*) \right].$$

This calculation is similar to that of Proposition 7.1; these are members of a much larger class of manifolds that are interesting in their own right. Note, in particular, that $N_t(\gamma)$ is an L–space for all γ other than the rational longitude; see Section 4 but compare also [3]. Therefore, the N_t satisfy the conditions of (ii) in Theorem 7.3 and we have the following:

Corollary 7.4 *Let M be a simple loop-type manifold that is not solid torus-like. The following are equivalent*:

(i) *γ is a strict L–space slope for M, that is, $\gamma \in \mathcal{L}^\circ_M$.*

(ii) *$M \cup_h N_t$ is an L–space, where $h(\gamma) = \lambda$ is the rational longitude, for any integer $t > 1$.*

(iii) *$M \cup_h N_2$ is an L–space, where $h(\gamma) = \lambda$ is the rational longitude.*

Note that Theorem 1.5 follows from the equivalence between (i) and (iii) in Corollary 7.4. This answers [2, Question 1.8] and considerably simplifies [2, Theorem 1.6] when restricting to Seifert-fibred rational homology solid tori. Indeed, we have shown:

Theorem 7.5 *Suppose $M \ncong D^2 \times S^1$ is a Seifert-fibred rational homology solid torus. The following are equivalent*:

(i) $\gamma \in (\mathcal{L}^\circ_M)^c$.

(ii) *γ is detected by a left-order* (*in the sense of Boyer and Clay* [2]).

(iii) *γ is detected by a taut foliation* (*in the sense of Boyer and Clay* [2]).

Proof This follows immediately from [2, Theorem 1.6] combined with Corollary 7.4. □

Proof of Theorem 1.1 The equivalence between (ii) and (iii) is due to Boyer and Clay [2]. To see that (i) is equivalent to either of these we first note that, if M is one of the two Seifert-fibred pieces in Y, then, according to Theorem 4.1, the set of all slopes $\widehat{\mathbb{Q}}$ is divided into (the restriction to $\widehat{\mathbb{Q}}$ of) two disconnected intervals $\mathcal{L}^\circ_M$ and $(\mathcal{L}^\circ_M)^c$. The latter is precisely the set of *NLS detected slopes* in the sense of Boyer and Clay [2, Definition 7.16], according to Theorem 7.5. (In particular, this observation should be compared with [2, Theorem 8.1].) Thus the desired equivalence follows from Theorem 1.3, on comparison with [2, Theorem 1.7] restricted to rational homology solid tori. □

References

[1] **M Boileau**, **S Boyer**, *Graph manifold $\mathbb{Z}$–homology 3–spheres and taut foliations*, J. Topol. 8 (2015) 571–585 MR Zbl

[2] **S Boyer**, **A Clay**, *Foliations, orders, representations, L-spaces and graph manifolds*, Adv. Math. 310 (2017) 159–234 MR Zbl

[3] **S Boyer**, **C M Gordon**, **L Watson**, *On L-spaces and left-orderable fundamental groups*, Math. Ann. 356 (2013) 1213–1245 MR Zbl

[4] **S Boyer**, **D Rolfsen**, **B Wiest**, *Orderable 3–manifold groups*, Ann. Inst. Fourier (Grenoble) 55 (2005) 243–288 MR Zbl

[5] **A Clay**, **T Lidman**, **L Watson**, *Graph manifolds, left-orderability and amalgamation*, Algebr. Geom. Topol. 13 (2013) 2347–2368 MR Zbl

[6] **N M Dunfield**, *Floer homology, group orderability, and taut foliations of hyperbolic 3–manifolds*, Geom. Topol. 24 (2020) 2075–2125 MR Zbl

[7] **F Haiden**, **L Katzarkov**, **M Kontsevich**, *Flat surfaces and stability structures*, Publ. Math. Inst. Hautes Études Sci. 126 (2017) 247–318 MR Zbl

[8] **J Hanselman**, *Bordered Heegaard Floer homology and graph manifolds*, Algebr. Geom. Topol. 16 (2016) 3103–3166 MR Zbl

[9] **J Hanselman**, *Splicing integer framed knot complements and bordered Heegaard Floer homology*, Quantum Topol. 8 (2017) 715–748 MR Zbl

[10] **J Hanselman**, **J Rasmussen**, **S D Rasmussen**, **L Watson**, *L-spaces, taut foliations, and graph manifolds*, Compos. Math. 156 (2020) 604–612 MR Zbl

[11] **J Hanselman**, **J Rasmussen**, **L Watson**, *Bordered Floer homology for manifolds with torus boundary via immersed curves*, preprint (2016) arXiv 1604.03466 To appear in J. Amer. Math. Soc.

[12] **J Hanselman**, **J Rasmussen**, **L Watson**, *Heegaard Floer homology for manifolds with torus boundary: properties and examples*, Proc. Lond. Math. Soc. (3) 125 (2022) 879–967 MR

[13] **J Hanselman**, **L Watson**, *Cabling in terms of immersed curves*, Geom. Topol. 27 (2023) 925–952

[14] **M Hedden**, **A S Levine**, *Splicing knot complements and bordered Floer homology*, J. Reine Angew. Math. 720 (2016) 129–154 MR Zbl

[15] **J Hom**, **T Lidman**, **L Watson**, *The Alexander module, Seifert forms, and categorification*, J. Topol. 10 (2017) 22–100 MR Zbl

[16] **A Kotelskiy**, **L Watson**, **C Zibrowius**, *Immersed curves in Khovanov homology*, preprint (2019) arXiv 1910.14584

[17] **A S Levine**, *Knot doubling operators and bordered Heegaard Floer homology*, J. Topol. 5 (2012) 651–712 MR Zbl

[18] **T Li**, **R Roberts**, *Taut foliations in knot complements*, Pacific J. Math. 269 (2014) 149–168 MR Zbl

[19] **R Lipshitz**, **P S Ozsváth**, **D P Thurston**, *Bimodules in bordered Heegaard Floer homology*, Geom. Topol. 19 (2015) 525–724 MR Zbl

[20] **R Lipshitz**, **P S Ozsvath**, **D P Thurston**, *Bordered Heegaard Floer homology*, Mem. Amer. Math. Soc. 1216, Amer. Math. Soc., Providence, RI (2018) MR Zbl

[21] **M Mauricio**, *On lattice cohomology and left-orderability*, preprint (2013) arXiv 1308.1890

[22] **W D Neumann**, *A calculus for plumbing applied to the topology of complex surface singularities and degenerating complex curves*, Trans. Amer. Math. Soc. 268 (1981) 299–344 MR Zbl

[23] **P Ozsváth**, **Z Szabó**, *On the Floer homology of plumbed three-manifolds*, Geom. Topol. 7 (2003) 185–224 MR Zbl

[24] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and three-manifold invariants: properties and applications*, Ann. of Math. 159 (2004) 1159–1245 MR Zbl

[25] **I Petkova**, *Cables of thin knots and bordered Heegaard Floer homology*, Quantum Topol. 4 (2013) 377–409 MR Zbl

[26] **I Petkova**, *The decategorification of bordered Heegaard Floer homology*, J. Symplectic Geom. 16 (2018) 227–277 MR Zbl

[27] **J Rasmussen**, **S D Rasmussen**, *Floer simple manifolds and L-space intervals*, Adv. Math. 322 (2017) 738–805 MR Zbl

[28] **S D Rasmussen**, *L-space intervals for graph manifolds and cables*, Compos. Math. 153 (2017) 1008–1049 MR Zbl

[29] **L Watson**, *Surgery obstructions from Khovanov homology*, Selecta Math. 18 (2012) 417–472 MR Zbl

Department of Mathematics, University of Texas at Austin
Austin, TX, United States

Current address: *Department of Mathematics, Princeton University*
Princeton, NJ, United States

School of Mathematics and Statistics, University of Glasgow
Glasgow, United Kingdom

Current address: *Department of Mathematics, University of British Columbia*
Vancouver BC, Canada

jh66@math.princeton.edu, liam@math.ubc.ca

Proposed: Peter Ozsváth Received: 24 September 2015
Seconded: Ciprian Manolescu, Cameron Gordon Revised: 10 July 2020

msp

Geometry & Topology **27**:3 (2023) 925–952
DOI: 10.2140/gt.2023.27.925
Published: 8 June 2023

Cabling in terms of immersed curves

JONATHAN HANSELMAN
LIAM WATSON

In joint work with J Rasmussen (Proc. Lond. Math. Soc. (3) 125 (2022) 879–967), we gave an interpretation of Heegaard Floer homology for manifolds with torus boundary in terms of immersed curves in a punctured torus. In particular, knot Floer homology is captured by this invariant (arXiv 1810.10355). Appealing to earlier work of the authors on bordered Floer homology (Geom. Topol. 27 (2023) 823–924), we give a formula for the behaviour of these immersed curves under cabling.

57M25, 57M27

Knot Floer homology, as introduced by Ozsváth and Szabó [18] and Rasmussen [20], provides a categorification: Given a knot K in the three-sphere, this invariant is a bigraded vector space $\bigoplus_{a,m\in\mathbb{Z}} \widehat{HFK}_m(K,a)$ with the property that

$$\sum_{a,m}(-1)^m \dim(\widehat{HFK}_m(K,a))t^a$$

recovers the (symmetrized) Alexander polynomial. This polynomial knot invariant satisfies natural properties associated with operations on knots; for instance, it is well behaved under cabling. Understanding how this particular property manifests at the categorified level drove some of the early calculations of knot Floer homology; see in particular work of Hedden [7; 8; 9].

Bordered Floer homology provides an essential tool for studying decompositions of three-manifolds along essential tori; see Lipshitz, Ozsváth and Thurston [15]. They laid out a framework of bimodules, of relevance to satellite operations, in [14]. The work of Levine [13], Hom [11], and Petkova [19], for example, puts this to use in an essential way. In the setting of manifolds with a single toroidal boundary component, the relevant bordered invariants have been recast in terms of immersed curves in the once-punctured torus; see Hanselman, Rasmussen and Watson [4; 5]. For the purpose of this note, the examples of interest will be provided by the complement of a knot in the three-sphere; our aim is to establish formulas for how these invariants behave under cabling. Namely,

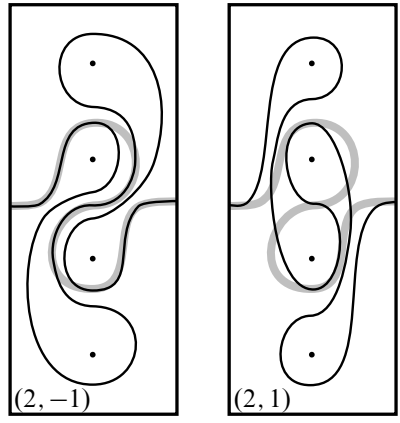


Figure 1: The Heegaard Floer homology for the $(2,-1)$– and $(2,1)$–cables of the right-hand trefoil; the invariant for the trefoil complement is shown in grey.

for a knot K in S^3, let $K_{p,q}$ denote the (p,q)–cable of K, and denote the respective knot complements by $M = S^3 \setminus \nu(K)$ and $M_{p,q} = S^3 \setminus \nu(K_{p,q})$; given the immersed multicurve $\widehat{HF}(M)$, we wish to describe $\widehat{HF}(M_{p,q})$ explicitly. For example, immersed curves for two cables of the right-hand trefoil are illustrated in Figure 1; the expert reader already familiar with the passage from $\widehat{HFK}(K)$ to $\widehat{HFK}(K_{p,q})$ should compare these pictures with the detailed calculations of Hedden [7] or Ozsváth, Stipsicz and Szabó [17]. Our calculation makes an explicit appeal to a bordered trimodule calculated by the first author [3], which was reinterpreted combinatorially in work of the authors predating the immersed curves invariant [6]. Indeed, central to this article is the work of translating our merge operation (described in terms of loop calculus) into the language of immersed curves (Section 1); cabling is then seen as a special case of the merge operation (Section 2).

Recall from [4; 5] that, for a (connected, orientable) three-manifold M with torus boundary, the invariant $\widehat{HF}(M)$ takes the form of a collection of immersed curves, possibly decorated with local systems, in the punctured torus $T_\bullet = \partial M \setminus z$, where z is some fixed basepoint in ∂M. If we choose a pair of parametrizing curves (α, β) on ∂M, then $T_\bullet$ can be identified with the square $[0,1]\times[0,1]$ with opposite sides identified such that α runs in the positive vertical direction, β runs in the positive horizontal direction, and the puncture z is identified with $(0,0)$. For a knot complement, there is a preferred choice of parametrizing curves, (μ, λ), where μ is the meridian and λ is the Seifert longitude. The invariant $\widehat{HF}(M)$ comes equipped with grading data which, among other things, specifies a lift of these curves to the punctured cylinder

$\overline{T}_\bullet = (\mathbb{R}^2 \setminus \mathbb{Z}^2)/\langle\lambda\rangle$; in the standard framing, this can be identified with $(\mathbb{R}/\mathbb{Z})\times\mathbb{R}$ with punctures at each lattice point $(0,n)$. (Note that this is the point of view taken in the presentation of the invariants in Figure 1: in each rectangle, the sides are identified to form a cylinder.) Thus as a graded object it makes sense to view $\widehat{HF}(M)$ as a collection of closed immersed curves $\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_n)$ in $\overline{T}_\bullet$, defined up to homotopy in $\overline{T}_\bullet$, possibly decorated with local systems. For knots in S^3, these curves have the property that, possibly after a homotopy, the curve set intersects the vertical line $\{\frac{1}{2}\}\times\mathbb{R}$ exactly once; we will always assume that γ_0 is the curve component containing this intersection. In other words, γ_0 wraps around the cylinder exactly once, while the remaining $\gamma_{i>0}$ can be confined to a neighbourhood of the vertical line through the punctures. We remark also that, while $\boldsymbol{\gamma}$ may carry nontrivial local systems, γ_0 always carries the trivial 1–dimensional local system (otherwise the rank of $\widehat{HF}$ of the meridional filling of K would be greater than one). Finally, it is sometimes convenient to work in the plane $\widetilde{T}_\bullet = \mathbb{R}^2 \setminus \mathbb{Z}^2$ rather than the cylinder, with the multicurve $\boldsymbol{\gamma}$ lifting to one that is invariant under translation by λ; note that in this cover γ_0 lifts to a single periodic curve while each $\gamma_{i>0}$ lifts to infinitely many copies of the same curve.

We will show that the (p,q)–cable operation acts on $\widehat{HF}(M)$ by applying a particular diffeomorphism to the plane. Let $\boldsymbol{g}_{p,q}$ be a diffeomorphism of $\mathbb{R}^2$ defined on the lattice $\mathbb{Z}^2$ by sliding each lattice point leftward along lines of slope $\frac{q}{p}$ until they first meet a vertical line $x = np$ for some integer n. Note that $\boldsymbol{g}_{p,q}$ does not fix the lattice $\mathbb{Z}^2$ but rather takes it to $p\mathbb{Z}\times\frac{1}{p}\mathbb{Z}$; let $\boldsymbol{f}_{p,q}$ be the composition of this map with vertical stretching by a factor of p and horizontal compression by a factor of p, so that $\boldsymbol{f}_{p,q}$ takes $\mathbb{Z}^2$ to $\mathbb{Z}^2$, followed by a vertical shift of $\frac{1}{2}(p-1)(q-1)$. We remark that the vertical shift is forced by the symmetry of the curves $\widehat{HF}(M)$ for any M and our convention that these curves are centred at height $\frac{1}{2}$; with this convention understood we will generally ignore the vertical positioning of the curves, but it is sometimes helpful to keep track of this vertical translation explicitly. The map $\boldsymbol{f}_{p,q}$ is not linear, though in some sense it is as close to being linear as possible: it is the composition of linear transformations, which can each be realized as a sequence of plane shears, with a single *fractional plane shear* (defined in Section 2).

Theorem 1 *If $\boldsymbol{\gamma}$ is the immersed multicurve associated with K, $\boldsymbol{\gamma}_{p,q}$ is the immersed multicurve associated with $K_{p,q}$, and $\widetilde{\boldsymbol{\gamma}}$ and $\widetilde{\boldsymbol{\gamma}}_{p,q}$ are the corresponding lifts to $\widetilde{T}_\bullet = \mathbb{R}^2 \setminus \mathbb{Z}^2$, then $\widetilde{\boldsymbol{\gamma}}_{p,q}$ is homotopic to $\boldsymbol{f}_{p,q}(\widetilde{\boldsymbol{\gamma}})$.*

Note that $\boldsymbol{f}_{p,q}$ is periodic with period p in the horizontal direction, so it makes sense to view $\boldsymbol{f}_{p,q}$ as a map from the cylinder $p\overline{T} := (\mathbb{R}/p\mathbb{Z})\times\mathbb{R}$ to $\overline{T} := (\mathbb{R}/\mathbb{Z})\times\mathbb{R}$ taking

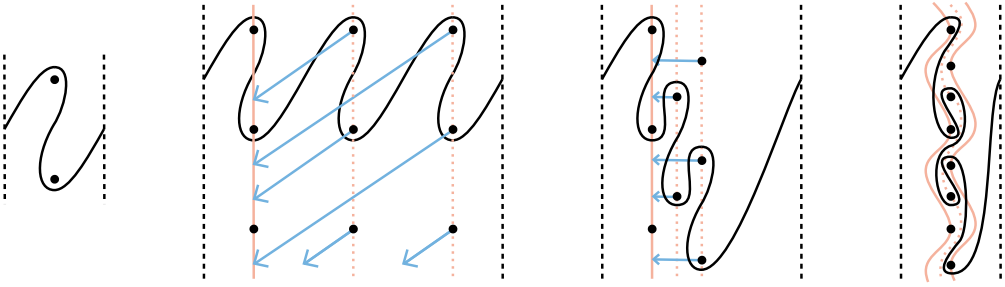


Figure 2: Computation of the immersed curve associated with the $(3,2)$–cable of the right-hand trefoil, starting from the trefoil curve pictured on the left. The two middle diagrams are two ways of thinking about the construction starting from three copies of the trefoil curve: we either slide lattice points along lines of slope $\frac{2}{3}$ or we stagger the heights of the three copies of the trefoil curve and then slide lattice points horizontally. Either way the result is the curve on the right.

lattice points to lattice points. With this view, the process of computing $\widehat{HF}(M_{p,q})$ from $\widehat{HF}(M)$ is to lift from $\overline{T}$ to $p\overline{T}$ and then apply $\boldsymbol{f}_{p,q}$. In practice, this amounts to drawing p copies of $\widehat{HF}(M)$ in sequence, perturbing the curve by pushing lattice points along lines of slope $\frac{q}{p}$ until they all lie on the same vertical line, and then scaling vertically by a factor of p. This procedure is depicted in Figure 2 for the case of the $(3,2)$–cable of the right-hand trefoil. It is helpful to note that the procedure of pushing lattice points along lines of slope $\frac{q}{p}$ can equivalently be viewed as drawing p copies of the input curve with staggered heights and then translating punctures horizontally. In practice, the process for computing $\widehat{HF}(M_{p,q})$ from $\widehat{HF}(M)$ amounts to a three-step process:

(1) draw p copies of $\widehat{HF}(M)$ next to each other, each scaled vertically by a factor of p, staggered in height such that each copy of the curve is a height of q units lower than the previous copy;

(2) connect the loose ends of the successive copies of the curve; and

(3) translate the pegs horizontally so that they lie in the same vertical line, carrying the curve along with them.

Numerical concordance invariants extracted from curves

As an illustration of Theorem 1 at work, we can revisit the work of Hedden [8; 9] and Van Cott [21], culminating in a result of Hom [11], which establishes the behaviour of the τ–invariant under cabling. Since $\tau(K)$ can be easily extracted from the immersed

multicurve $\widehat{HF}(M)$, we can recover this cabling behaviour from Theorem 1. The same is true for some other numerical invariants. We begin by making an observation that is implicit in earlier work. Let $\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_n)$ denote the underlying set of immersed curves for $\widehat{HF}(M)$, with γ_0 the unique component which wraps around the cylinder. This component is itself an invariant of K, so it will sometimes be convenient to express it as $\gamma_0(K)$.

Proposition 2 *The curve $\gamma_0(K)$ is an invariant of the concordance class of K.*

Proof This follows from Hom [12] and the recipe for deriving $\widehat{HF}(M)$ from $CFK^-(K)$ described in [5, Section 4]. The concordance invariant described in [12] is the smallest direct summand of $CFK^-(K)$, up to homotopy equivalence, which supports the homology of S^3. The set of immersed curves derived from this summand is a subset of the immersed curves $\widehat{HF}(M)$ which necessarily contains $\gamma_0(K)$. This subset of curves, and in particular $\gamma_0(K)$, is thus a concordance invariant. □

Note that the concordance invariant described in [12] is slightly stronger than γ_0 since some information may be lost when passing from complexes to immersed curves (namely, diagonal arrows are ignored). In fact, $\gamma_0(K)$ carries exactly the same information as the ϵ–equivalence class of K defined in [12]. Any number that can be extracted from γ_0 is automatically a concordance invariant, and several familiar concordance invariants can be defined in this way. The two most common are τ and Hom's ϵ–invariant, which are extracted from γ_0 as follows: Starting on the section of γ_0 which wraps around the back of the cylinder — say, at the unique intersection of γ_0 with the line $x = \frac{1}{2}$ — and moving rightward along γ_0, let a denote the first intersection of γ_0 with the vertical axis $x = 0$. Then the integer τ records the height of the intersection point a (here we use a discrete notion of height given by the greatest integer less than the y–coordinate of a). Continuing along γ_0 from a, one of three things can happen: γ_0 can turn downwards, it can turn upwards, or it can continue straight to wrap around the cylinder. This is recorded by ϵ, which takes the values $+1$, -1 or 0 in these three cases, respectively. (Both of these observations are made in [5].) Note that if $\epsilon = 0$ then there is only one intersection of γ_0 with the vertical axis, so γ_0 is simply a horizontal curve, which is the immersed curve associated with the complement of the unknot. Now consider the effect of cabling on each of these invariants. Throughout, let $\gamma_0 = \gamma_0(K)$ and let $\gamma_0' = \gamma_0(K_{p,q})$.

Theorem 3 (Hom [11, Theorem 2]) *If $\epsilon(K) = \pm 1$ then $\epsilon(K_{p,q}) = \epsilon(K)$; and if $\epsilon(K) = 0$ then $\epsilon(K_{p,q}) = \epsilon(T_{p,q})$.*

A quick reproof of Theorem 3 By Theorem 1, γ_0' is obtained from γ_0 by placing p copies of γ_0 next to each other, with appropriate vertical shifts, and compressing them into one vertical line. The first intersection of γ_0' with the vertical axis thus comes from the first intersection of the first copy of γ_0 with the vertical axis, and clearly if γ_0 turns upward or downward at this point then γ_0' does also. On the other hand, if $\epsilon(K)=0$ then γ_0 is simply a horizontal line, the same as the curve associated with the unknot. It follows that γ_0' agrees with $\gamma_0(T_{p,q})$, since $T_{p,q}$ is the (p,q)–cable of the unknot, and thus $\epsilon(K_{p,q})=\epsilon(T_{p,q})$. □

The value $\epsilon(T_{p,q})$ was also computed in [11, Theorem 2]; we can recover this computation by viewing $T_{p,q}$ as the (p,q)–cable of the unknot. In this case γ_0 is horizontal in $\overline{T}$ and lifts to a horizontal line in $p\overline{T}$. To compute γ_0' from this we shift the i^{th} column downwards by $\frac{iq}{p}$ and then compress horizontally. If $|q|=1$ then $T_{p,q}$ is unknotted and we must have $\epsilon(T_{p,q})=0$; indeed, in this case every column of lattice points shifts by less than one unit, so it is possible for γ_0' to remain horizontal despite the shift. On the other hand, if $q>1$ then the shift causes γ_0' to turn downwards, so $\epsilon(T_{p,q})=+1$; similarly, if $q<-1$ then γ_0' turns upward and $\epsilon(T_{p,q})=-1$.

Theorem 4 [11, Theorem 1] *If $\epsilon(K)=\pm1$ then $\tau(K_{p,q})=p\tau(K)+\frac{1}{2}(p-1)(q\mp1)$; and if $\epsilon(K)=0$ then $\tau(K_{p,q})=\tau(T_{p,q})=(-1)^{\operatorname{sign}(q)}\frac{1}{2}(p-1)(|q|-1)$.*

A quick reproof of Theorem 4 The first intersection of γ_0' with the vertical axis clearly comes from the first intersection of the first copy of γ_0 with the vertical axis. This intersection occurs between the lattice points at height $\tau(K)$ and $\tau(K)+1$; after applying $\boldsymbol{f}_{p,q}$ and the appropriate vertical shift, these lattice points map to heights $h_1=p\tau(K)+\frac{1}{2}(p-1)(q-1)$ and $h_2=p\tau(K)+p+\frac{1}{2}(p-1)(q-1)$. Note that there are $p-1$ lattice points between these two heights; whether or not γ_0' first intersects the vertical axis above or below these points depends on the behaviour of γ_0 just after it crosses the vertical axis, as pictured in Figure 3. If γ_0 turns downward (ie if

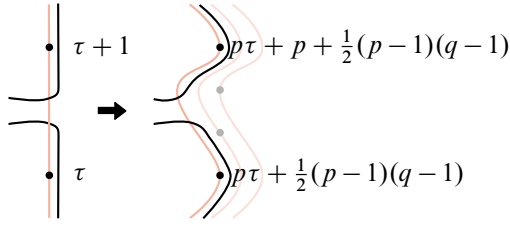


Figure 3: Calculating $\tau(K_{p,q})$.

$\epsilon(K) = +1$), then γ_0' will also turn downward and meet the vertical axis just above height h_1; thus $\tau(K_{p,q}) = h_1$. If γ_0 turns upward (ie if $\epsilon(K) = -1$), then γ_0' will also turn upwards and meet the vertical axis just below height h_2; thus $\tau(K_{p,q}) = h_2 - 1 = p\tau(K) + \frac{1}{2}(p-1)(q+1)$. Finally, if $\epsilon(K) = 0$ then γ_0 agrees with the curve invariant of the unknot, so γ_0' is the curve associated with $T_{p,q}$, and thus $\tau(K_{p,q}) = \tau(T_{p,q})$. In particular, if $q > 1$ then γ_0' bends down after its first intersection with the vertical axis, and as above $\tau(K_{p,q}) = h_1 = \frac{1}{2}(p-1)(q-1)$. If $q < -1$ then γ_0' bends upward and $\tau(K_{p,q}) = h_2 - 1 = -\frac{1}{2}(p-1)(-q-1)$, while if $|q| = 1$ then γ_0' is horizontal and $\tau(K_{p,q}) = 0$. □

Other concordance invariants can be extracted from $\gamma_0(K)$. For instance, for any positive integer i, the invariant $\phi_i(K)$ introduced recently by Dai, Hom, Stoffregen and Truong [1] counts the number of left arcs of γ_0 of length i, where a left arc of length i refers to a segment of γ_0 connecting successive intersections with the vertical axis whose height differ by i which does not wrap around the cylinder and which lies to the left of the vertical axis. These arcs are counted with sign coming from the orientation of γ_0, with downward oriented arcs counting positively.[1] Like τ, the integers ϕ_i are of particular interest in the study of knot concordance because they are additive under connected sum; that is, they define concordance homomorphisms.

Returning to cabling, the behaviour of the invariants ϕ_i is more complicated. In particular, $\phi_i(K_{p,q})$ does not depend only on $\phi_i(K)$, or even on the collection of invariants $\tau(K)$, $\epsilon(K)$ and $\phi_j(K)$ for all j. In order to express the effect of cabling we need to keep track of how each left arc in $\gamma_0(K)$ behaves at each end. For example, we can define refined invariants ϕ_i^{++}, ϕ_i^{+-}, ϕ_i^{-+} and ϕ_i^{--} encoding the signed count of four different types of length i left arcs in γ_0. The type is determined by the direction γ_0 turns at each end of the segment; $+$ indicates that γ_0 turns upward and $-$ indicates that γ_0 turns downward, with the first sign indicating the behaviour at the top of the arc and the second sign indicating the behaviour at the bottom of the arc, as in Figure 4. Note that $\phi_i = \phi_i^{++} + \phi_i^{+-} + \phi_i^{-+} + \phi_i^{--}$. With these extra quantities defined, it is possible to derive explicit formulas for $\phi_i(K_{2,1})$. More generally, we could derive explicit formulas for $\phi_i(K_{p,q})$ in a similar way; the key difference is that the notions of turning up or turning down used in defining the invariants $\phi_i^{\pm\pm}$ are dependent on p

[1] This is a straightforward translation of the definition of ϕ_i given in [1] to the language of immersed curves. The *standard complex* described in [1] corresponds precisely to the component γ_0 of $\widehat{HF}(M)$. The integers ϕ_i count horizontal arrows of length i in the standard complex, which correspond to length i right arcs in γ_0. By symmetry, we can equivalently count length i left arcs in γ_0.

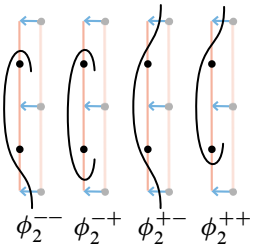


Figure 4: Four cases complete the proof.

and q (here turning up means moving upward vertically or rightward with slope greater than $\frac{q}{p}$). The formulas are cumbersome so, rather that derive the general case, we focus instead on the special case of $(2,1)$–cabling.

Proposition 5 *For $i>1$ all the variants of $\phi_i(K_{2,1})$ are either determined by*

$$\phi_{2n}^{\pm\pm}(K_{2,1})=\phi_n^{\pm\pm}(K), \quad \phi_{2n\pm1}^{\pm\mp}(K_{2,1})=\phi_n^{\pm\mp}(K)$$

or they are trivial. In particular, for $n\geq 1$,

$$\phi_{2n}(K_{2,1})=\phi_n^{++}(K)+\phi_n^{--}(K), \quad \phi_{2n+1}(K_{2,1})=\phi_n^{+-}(K)+\phi_{n+1}^{-+}(K).$$

Proof The curve $\gamma_0'=\gamma_0(K_{2,1})$ is constructed in three steps: take two consecutive copies of $\gamma_0=\gamma_0(K)$; scale vertically by a factor of two and shift the second copy of γ_0 down one unit; and compress horizontally (compare Figure 2). Before compressing horizontally, we can divide this curve into two (nonconnected) subcurves which lie to the left and right of vertical line through the first column of lattice points; let γ_L and γ_R denote the images of these subcurves after horizontal compression, so that $\gamma_0'=\gamma_L\cup\gamma_R$.

The key observation is that every component of γ_R lies to the right of every even height lattice point, and therefore any left arc on γ_0' which lies in γ_R must have length 1. It is also clear that any left arc of γ_0' intersects at most one component of γ_L, since otherwise it contains a full component of γ_R which must lie to the right of some lattice point. Thus each left arc of length greater than 1 in γ_0' comes from a component of γ_L, which in turn comes from a left arc of the first copy of γ_0. Conversely, every left arc of length i in the first copy of γ_0 gives rise to exactly one left arc in γ_0', which has the same end behaviour. The length of this new arc depends on the end behaviour: it is $2i-1$ for $-+$ arcs, $2i$ for $++$ or $--$ arcs, and $2i+1$ for $+-$ arcs (see Figure 4). □

We will say that γ_0 has a unique maximal-length left arc of type $++$ and length N if $\phi_N^{++}(K)=1$, $\phi_i^{++}(K)=0$ for all $i>N$, and $\phi_i^{+-}(K)=\phi_i^{-+}(K)=\phi_i^{--}(K)=0$ for all $i\geq N$. The following is an immediate consequence of the formulas above:

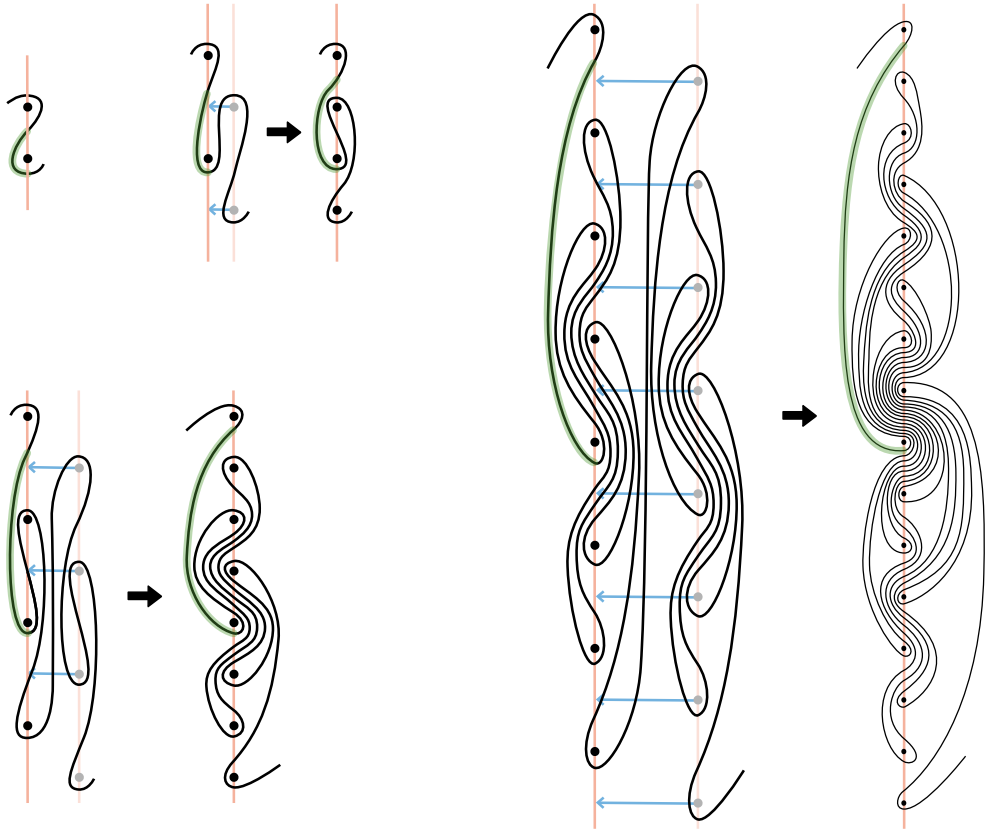


Figure 5: Immersed curves for the first few iterated $(2,1)$–cables of the right-hand trefoil. These are also the distinguished curve γ_0 for the knots K_0, K_1, K_2 and K_3 from Corollary 7. The longest left arc (highlighted) is stretched by a factor of two with each cabling iteration; thus, the length of the longest left arc for K_n is 2^n.

Proposition 6 *If $\gamma_0(K)$ has a unique maximal-length left arc of type $++$ and length N, then $\gamma_0(K_{2,1})$ has a unique maximal-length left arc of type $++$ and length $2N$.*

Consider for example iterated $(2,1)$–cables of the right-hand trefoil $T_{2,3}$; the immersed curves for the first few of these knots are shown in Figure 5. The immersed curve $\gamma_0(T_{2,3})$ has only one left arc, which has type $++$ and length 1. If we repeatedly $(2,1)$–cable this knot, there is always a single left arc of maximal length, which always has type $++$, and the length of this arc doubles in length with each iteration.

In [1], the concordance invariants ϕ_i were used to identify a $\mathbb{Z}^\infty$ direct summand in the topologically slice smooth concordance group $\mathcal{C}_{\mathrm{TS}}$; see also [17]. The relevant

infinite family of knots is built from cables of a certain knot D, the untwisted positively clasped Whitehead double of $T_{2,3}$. More precisely, the family of knots is given by $D_{n,n+1} \# -T_{n,n+1}$. Using Proposition 6, we can construct another $\mathbb{Z}^\infty$ summand from D by instead taking iterated $(2,1)$–cables. The key properties of D are that

(i) the Alexander polynomial of D is trivial, and

(ii) the distinguished component γ_0 associated to D agrees with $\gamma_0(T_{2,3})$.

The knot D can be replaced with any other knot which shares these two properties — an example of a hyperbolic knot with this property is $15n113775$.

Corollary 7 *Let $K = K_0$ be a knot for which $\Delta_K(t) = 1$ and $\gamma_0(K) = \gamma_0(T_{2,3})$. For $n \geq 1$ let K_n be the $(2,1)$–cable of K_{n-1}. The knots $\{K_n\}_{n=0}^\infty$ span a $\mathbb{Z}^\infty$ summand of $\mathcal{C}_{\mathrm{TS}}$.*

Proof According to a result of Freedman, $\Delta_K(t) = 1$ implies that K is topologically slice [2]. The $(2,1)$–cable of a topologically slice knot is topologically concordant to the $(2,1)$–cable of the unknot, which is the unknot; thus, by induction, K_n is topologically slice for all n. On the other hand, K_0 has a unique maximal-length left arc of type $++$ and length 1, so Proposition 6 and induction implies that K_n has a unique maximal-length left arc of type $++$ and length 2^n. In particular, for each n we have $\phi_{2^n}(K_n) = 1$ and $\phi_i(K_n) = 0$ for all $i > 2^n$. Since each ϕ_i is a concordance homomorphism, it follows that the knots are linearly independent in the smooth concordance group. Moreover, it is straightforward to see that the homomorphism

$$\bigoplus_{i=0}^{\infty} \phi_{2^i} : \mathcal{C}_{\mathrm{TS}} \to \bigoplus_{i=0}^{\infty} \mathbb{Z}$$

is an isomorphism when restricted to the span of the K_n. Indeed, it follows from the information above that the images of the K_n form a basis for $\mathbb{Z}^\infty$ (see for instance [17, Proposition 6.4]). In fact, by the remark below we also have that $\phi_i(K_n) = 0$ for all $i < 2^n$, so the image of K_n is the standard i^{th} basis vector of $\mathbb{Z}^\infty$. □

Remark 8 We leave the behaviour of ϕ_1 under $(2,1)$–cabling, which was not needed in the above application, as an exercise to the motivated reader, who will find that

$$\phi_1(K_{2,1}) = -\sum_{j \geq 1} \phi_j(K) + \begin{cases} 1 & \text{if } \tau(K) > 0, \\ 0 & \text{if } \tau(K) \leq 0. \end{cases}$$

Using corresponding formulas for variants of ϕ_1 along with those from Proposition 5 and induction on n, it can be shown that the knots K_n in Corollary 7 in fact satisfy

$\phi_i(K_n) = 1$ if $i = 2^n$ and $\phi_i(K_n) = 0$ otherwise. Thus the map

$$\bigoplus_{i=0}^{\infty} \phi_{2^i} : \mathrm{Span}\{K_n\} \to \bigoplus_{i=0}^{\infty} \mathbb{Z}$$

defines an explicit isomorphism from the span of the K_n in the topologically slice concordance group to $\mathbb{Z}^\infty$.

Curves that do not come from cables

In many cases Theorem 1 provides a simple obstruction to a knot being a nontrivial cable. For example, if the curve set corresponding to a knot contains a figure-eight component enclosing two adjacent lattice points (as is the case, for example, for any nontorus alternating knot; see [19]), then the knot is not a cable of a knot in S^3. This is because any closed component of $\widehat{HF}(M_{p,q})$ comes from one copy of $\widehat{HF}(M)$ before the transformation of the plane and thus encloses only lattice points with the same height modulo p. If we restrict our attention to γ_0, we can find knots which are not concordant to a nontrivial cable. An example of this is the knot $12n242$, whose γ_0 is pictured in Figure 6. Because the first left arc has length 2, if this curve comes from a (p,q)–cable then p must be at most 3. If $p = 1$, the cable operation is trivial. If $p = 2$, then the curve should not pass to the left of an odd-height lattice point (grey in Figure 6) after passing to the right of an even-height lattice point (black in the figure), but this clearly happens. Similarly, if $p = 3$ then the curve should not pass to the left of a point whose height is congruent to 2 mod 3, after passing to the right of a lattice point with a different height modulo 3, but this also happens. We thank Tye Lidman for asking us about the existence of such an example.

Figure 6

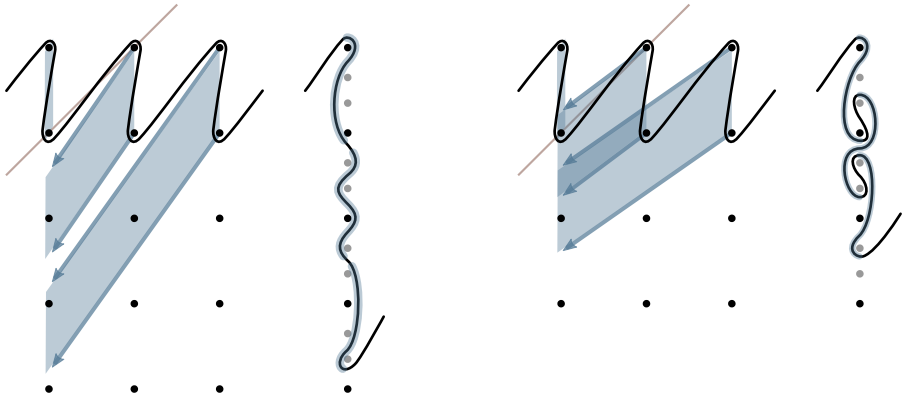


Figure 7: The $(3,4)$–cable (left) and the $(3,2)$–cable (right) of the right-hand trefoil.

L–space surgeries on cable knots

Other properties of the knot Floer homology of cable knots are made relatively transparent by Theorem 1. The following is a well-known property that was established by Hom [10] (building on work of Hedden [9]):

Theorem 9 [10, Theorem] *For any knot K in S^3, $K_{p,q}$ admits a positive L–space surgery if and only if K admits a positive L–space surgery and $\frac{q}{p}$ is at least $2g-1$, where g denotes the Seifert genus of K.*

Note that K admits a positive L–space surgery if and only if $\widehat{HF}(M)$ is a single curve which, apart from the segment that wraps around the cylinder, moves monotonically downward in the neighbourhood of the vertical axis (see [4, Section 7.5]). When this curve is pulled tight in the cylinder $\overline{T}_\bullet$ (or in the plane $\widetilde{T}_\bullet$), the slope of the nonvertical segment is $2g-1$. Following Theorem 1, we construct $\widehat{HF}(M_{p,q})$ from p columns of the lift of $\widehat{HF}(M)$ to $\widetilde{T}_\bullet$ by translating lattice points along lines of slope $\frac{q}{p}$.

A quick reproof of Theorem 9 If $\widehat{HF}(M)$ is oriented upward at any point apart from the nonvertical segment, it is clear the same will be true at the image of this point on $\widehat{HF}(M_{p,q})$; thus K having a positive L–space surgery is a necessary condition for $K_{p,q}$ to have one. Supposing K has a positive L–space surgery, it is clear that if $\frac{q}{p} > 2g-1$ then the p copies of the downward-oriented portion of $\widehat{HF}(M)$ miss each other, so the resulting curve moves monotonically downward and $K_{p,q}$ has a positive L–space surgery. On the other hand, if $\frac{q}{p} < 2g-1$ then these sections of curves overlap, forcing some backtracking in the resulting curve, implying that $K_{p,q}$ has no L–space surgeries. An example is given in Figure 7. □

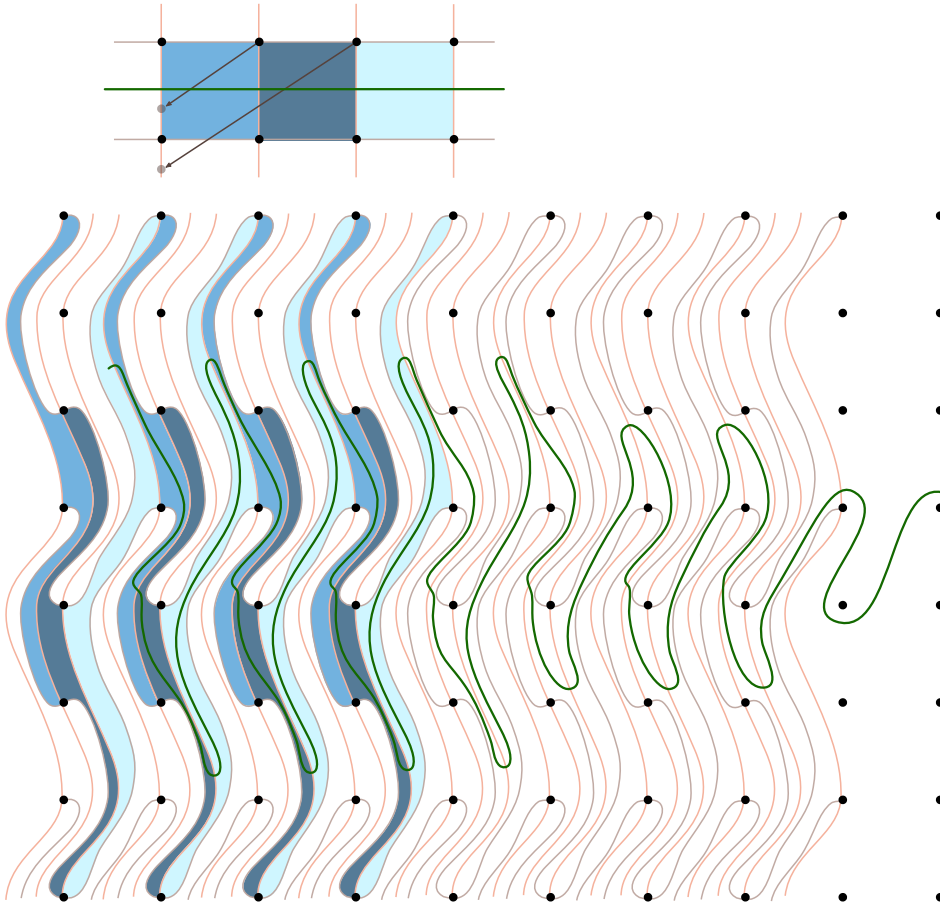


Figure 8: The $(3,2)$–cabling operation interpreted as a plane tiling: three copies of the standard square tile (above) are carried to a new regular tile in $\mathcal{T}_{3,2}$ (below) under the operation $\boldsymbol{f}_{p,q}$ appearing in Theorem 1. To illustrate, the image of the longitude has been included (gradually homotoped to a simpler form moving rightward), which recovers the invariant associated with the right-hand trefoil as expected.

Cabling via tiling

From Theorem 1 it is possible to interpret cabling in terms of plane tilings. That is, in a visual summary of the above discussion, we record the following:

Corollary 10 *For every relatively prime pair (p,q) there is a periodic tiling $\mathcal{T}_{p,q}$ of the plane, unique up to lattice-fixing planar isotopy, such that $\boldsymbol{\gamma}(K_{p,q})$ is the image of $\boldsymbol{\gamma}(K)$ under the transformation taking the lattice $\widetilde{T}_\bullet$ to $\mathcal{T}_{p,q}$.*

Proof This is a simple reformulation of Theorem 1: Consider the standard square tiling of the plane $\widetilde{T}_\bullet$ defined by the preferred (μ,λ)–framing. The image of p square tiles aligned horizontally, under the application of $\boldsymbol{f}_{p,q}$, gives a tile in a periodic tiling of the plane. □

This is best illustrated in an example, and we have shown the tiling associated with $(3,2)$–cabling in Figure 8. Note that this point of view comes with a built-in sanity check: one can check that the image of a longitudinal curve under the transformation to $\mathcal{T}_{p,q}$ is the immersed curve $\boldsymbol{\gamma}(T_{p,q})$. Recalling that, as a polynomial in t, the Alexander polynomial satisfies

$$\Delta(K_{p,q}) = \Delta(K)|_{t^p} \cdot \Delta(T_{p,q}),$$

our formula has $\mathcal{T}_{p,q}$ playing the role of $\Delta(T_{p,q})$ in this formula while replacing t with t^p corresponds to the p repeated copies of $\boldsymbol{\gamma}(K)$.

In general, one expects bimodules in bordered Floer homology (for manifolds with two boundary tori) to be associated with Lagrangians in $T_\bullet \times T_\bullet$. A simple first example of this is the bimodule associated with a diffeomorphism of the torus, where the (embedded) Lagrangian surface is the graph of the diffeomorphism. In that case we can interpret the action of the bimodule as follows: to compute the image of an immersed curve γ, we consider $\gamma \times T_\bullet$, intersect with the Lagrangian surface, and project to the second coordinate. Cabling bimodules provide a first glimpse at how this construction might be generalized to arbitrary bimodules. The diffeomorphism of the plane $\boldsymbol{f}_{p,q}$ does not descend to a diffeomorphism of the torus, but, since $\boldsymbol{f}_{p,q}$ is periodic and is determined by its effect on p consecutive tiles of the plane, it can be viewed as a p–valued function on $T_\bullet$; that is, to each point in $T_\bullet$ it associates an unordered tuple of p points in $T_\bullet$. The graph of this multivalued function is an (immersed) Lagrangian surface in $T_\bullet \times T_\bullet$, and the action of the bimodule on curves can be interpreted geometrically as before.

1 Immersed curves and the merge operation

For any orientable manifold M with torus boundary, the Heegaard Floer homology $\widehat{HF}(M)$ is an immersed multicurve in the marked torus ∂M [4], as introduced above. This view of the Heegaard Floer invariants of M arises from an interpretation of bordered Floer homology [15], and is closely related to the *loop calculus* introduced in [6]. This section builds a glossary between loop calculus [6] and immersed curves [4];

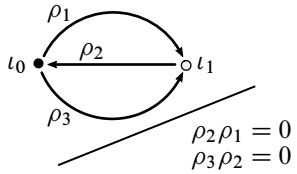


Figure 9: The torus algebra $\mathcal{A}$ as the path algebra of a quiver with relations.

in the former we developed the machinery for understanding gluing pairs of manifolds along essential annuli in their boundaries, which we aim to interpret in terms of the immersed curves in the case of cabling knots in the three-sphere.

1.1 From puzzle pieces to curve segments

Assuming familiarity with some subset of [4; 6; 15], we give a very terse summary of the bordered invariants in order to set up the desired glossary.

The torus algebra $\mathcal{A}$ is obtained as the path algebra of the quiver described in Figure 9. Let $\mathcal{I} \subset \mathcal{A}$ denote the subring of idempotents generated by ι_0 and ι_1. Working over the two-element field $\mathbb{F}$, a type D structure over $\mathcal{A}$ is a finite-dimensional left $\mathcal{I}$–module V together with a map $\delta\colon V \to \mathcal{A} \otimes_{\mathcal{I}} V$. This map must satisfy a compatibility condition equivalent to ensuring that $\partial(a \otimes x) = a \cdot \delta(x)$ is a differential on the $\mathcal{A}$–module $\mathcal{A} \otimes_{\mathcal{I}} V$.

There is a simple interpretation of the above data in terms of decorated graphs: the vertices encode the generating set (these come in two types $\bullet$ and $\circ$, depending on the idempotents ι_0 and ι_1, respectively) and, by passing to type D structures that are reduced, the directed edges are labelled by the set $\{1, 2, 3, 12, 23, 123\}$ in order to encode the coefficient maps; see Figure 11. These graphs can be naturally immersed in the marked torus or, more precisely, in the once-punctured torus with a fixed choice of 1–handle cocores cutting the surface into a disk. In our case, these cocores will always coincide with the preferred (μ, λ)–pair, since we are focussed on knots in S^3. With this data in hand, we can decompose the torus into the familiar square patch with opposite edges identified. The type D structures of interest then are immersed train tracks (in the sense of Thurston [16]) where all of the vertices/switches lie on the horizontal or vertical edges; when such a train track comes from a three-manifold, the classification theorem proved in [4] tells us it is equivalent to an immersed multicurve, possibly decorated with local systems, which we denote by $\widehat{HF}(M)$ [4].

In the case where the local systems are trivial, we recover the class of loop-type manifolds considered in our earlier work [6] (see also [5, Section 1]). Central to

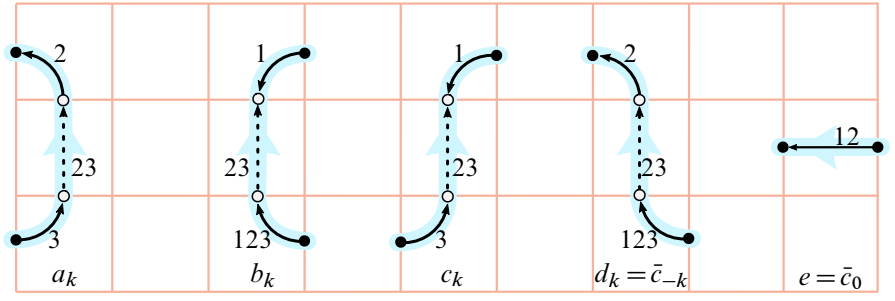


Figure 10: Segments of immersed curve in the cover of the marked torus, labelled to be consistent with the puzzle pieces given in [6]. The integer subscript $k > 0$ indicates the number of $\circ$ generators in the segment. These letters can appear forwards or backwards in a cyclic word, so that $\bar{a}_k$ runs against the direction indicated by the blue arrow. We can also extend our notation by setting $a_{-k} = \bar{a}_k$, $b_{-k} = \bar{b}_k$, $c_{-k} = \bar{d}_k$, $d_{-k} = \bar{c}_k$ and $d_0 = \bar{c}_0 = e$; note that then a segment with subscript k moves upward k units in the plane.

this is the observation that, when the type D structure in question can be represented by a valence 2 graph, it is possible to decompose along $\bullet$ vertices into segments, each of which takes one of five possible forms as described in Figure 10 (compare [6, Figure 1]).[2] As a result, studying these type D structures amounts to a calculus for manipulating cyclic words in the infinite alphabet $\mathfrak{A} = \{a_k, b_k, c_k, d_k, e\}$ for all positive integers k. The segments corresponding to these letters may appear backwards as we traverse a loop; this is indicated by a *bar*. There are rules governed by the algebra restricting the letters that can be concatenated, which are most easily described by noting that each segment also corresponds to a segment of immersed curve as in Figure 10: if two curve segments share an endpoint, they must lie on opposite sides of the vertical near that point. (In [6], a puzzle piece convention is used to describe these rules.) Note that the a_k and the b_k correspond to the two types of stable chains introduced in [15], while c_k, d_k and e correspond to the three types of unstable chains. In fact, it makes sense to view the three types of unstable chain as part of a single family, and with this in mind we set $c_0 = \bar{e}$ and $c_{-k} = \bar{d}_k$. The example in Figure 11 explains this for the right-hand trefoil exterior.

Now consider a component of $\widehat{HF}(M)$, that is, an immersed curve γ decorated with a local system (V, Φ) of dimension n. Following [4], we can interpret this as a *curve-like*

[2]There are certain exceptional type D structures that cannot be decomposed in this way; however, these examples are not particularly important in this setting. The interested reader can consult [6] for a *dual* notation that decomposes along $\circ$ vertices.

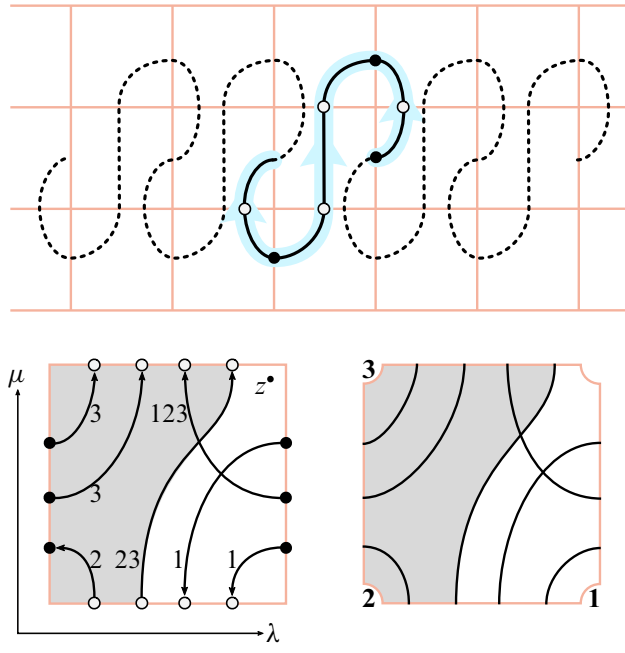


Figure 11: Three different views of the invariant associated with the exterior of the right-hand trefoil. In all three cases, we have fixed the preferred (μ, λ)–framing in order to present the torus boundary. On the lower left, the decorated graph describing the type D structure has been immersed in the marked torus as a train track. This description exhibits the redundancy in the edge labels: as shown in the lower right figure, the idempotents can be recovered from the horizontal and vertical edges while the coefficient maps are determined by which of the labelled corners are traversed by the curve segments (a region indicating the 23 edge is shaded). Finally, lifting the curve to the cover $\overline{T}_\bullet$ (or, as pictured, $\widetilde{T}_\bullet$) makes obvious the cyclic word $a_1\bar{c}_2b_1$, which in [6] is referred to as a *loop*.

train track, which consists of n parallel copies of γ along with some additional edges that we may assume all lie on a portion of γ corresponding to a single segment (that is, along one letter of $\mathfrak{A}$ as described above). When M is the complement of a knot K in S^3, we may in fact assume that these edges lie on a segment of type a_k; this is because the curve $\gamma_0(K)$ does not carry a nontrivial local system, and all other curves are closed in the lift to the plane $\widetilde{T}_\bullet$ and thus must contain a type a_k segment. The portion of the train track containing the extra edges is precisely a type a_k segment with local system (V, Φ). These extra edges determine an $n \times n$ matrix over $\mathbb{F}$, where the (i, j) entry is nonzero if the curve-like train track contains a copy of the a_k segment

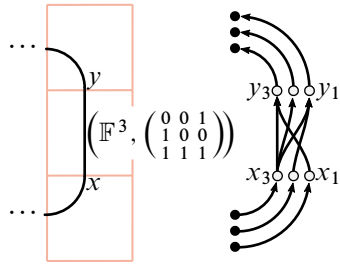


Figure 12: A 3–dimensional local system, expanded at an a_2 to give a train track.

from the i^{th} copy of the initial generator in the segment to the j^{th} copy of the final generator of the segment. By construction, this matrix represents the local system Φ.

It is relatively straightforward to extend this language to cases admitting a nontrivial local system. Recall that each letter in $\mathfrak{A}$ corresponds to a (portion of a) type D structure that is a linear chain of arrows. We allow a letter in $\mathfrak{A}$ to be decorated by a local system, as follows. Let V be a vector space over $\mathbb{F}$ of dimension n, and let $\Phi\colon V \to V$ be an endomorphism. Decorating a segment with (V, Φ) amounts to taking n parallel copies of the appropriate chain, with the n parallel copies of any one arrow in the chain replaced with a collection of arrows determined by Φ; see Figure 12 for an example. Suppose the relevant arrow in the chain connects generators (ie vertices) x and y, with x occurring first in the chain (ie there is an edge connecting x to y). Fix bases $\langle x_1, \dots, x_n\rangle$ and $\langle y_1, \dots, y_n\rangle$ for V and consider an $n \times n$ matrix over $\mathbb{F}$ representing Φ; these arrows connect the i^{th} copy of x (ie x_i) to the j^{th} copy of y (ie y_j) if and only if the (i, j) entry of the matrix is nonzero (again, see Figure 12). If all letters in a cyclic word carry a local system (each having the same dimension), then the local system on the cyclic word is determined by composing the endomorphisms. Note that a letter decorated by the trivial local system of dimension n corresponds to n parallel copies of the relevant curve segment.

To summarize, given a knot K in S^3, the invariant $\widehat{HF}(M)$ is an immersed multicurve $\boldsymbol{\gamma}(K) = (\gamma_0, \gamma_1, \dots, \gamma_n)$ where each $\gamma_{i>0}$ caries a (possibly trivial) local system. By the above discussion, we can assume that each component of $\boldsymbol{\gamma}(K)$ is represented by a cyclic word in $\mathfrak{A}$, possibly with a nontrivial local system on a single a_k segment.

1.2 The merge operation

Given type D structures $\boldsymbol{\vartheta}$ and $\boldsymbol{\gamma}$, we describe a new type D structure $\mathrm{M}(\boldsymbol{\gamma}, \boldsymbol{\vartheta})$. This follows the notation set out in [6], where we showed that this type D structure agrees

with $\widehat{HF}(\mathcal{M}(M_1, M_2))$ in the case where $\boldsymbol{\gamma}$ and $\boldsymbol{\vartheta}$ correspond to $\widehat{HF}(M_1)$ and $\widehat{HF}(M_2)$, respectively.[3] The operations M and $\mathcal{M}$ are referred to as merges; the latter glues two manifolds along essential annuli in their torus boundaries. We will first describe the operator M algebraically, and then explain the gluing conventions for $\mathcal{M}$ in the next section in the context of cabling.

Some simplifications are possible in the present setting. First, we assume that $\boldsymbol{\gamma}$ is a loop consisting only of some c_k for integers k. This assumption holds in particular when M_1 is a solid torus, in which case $\mathcal{M}(M_1, -)$ will give rise to a cabled knot.[4] Further, as described above, we assume that $\boldsymbol{\vartheta}$ is represented as a curve-like train track. This may consist of several disjoint components, but we can restrict to connected train tracks without loss of generality: if $\boldsymbol{\vartheta} = (\vartheta_0, \dots, \vartheta_n)$ then $\mathrm{M}(\boldsymbol{\gamma}, \boldsymbol{\vartheta}) = (\mathrm{M}(\boldsymbol{\gamma}, \vartheta_0), \dots, \mathrm{M}(\boldsymbol{\gamma}, \vartheta_n))$.

The main tool used in this paper is a distilled version of [6, Proposition 6.4]:

Proposition 11 *Let $\boldsymbol{\vartheta}$ be a type D structure represented by a single cyclic word in $\mathfrak{A}$ and let $\boldsymbol{\gamma}$ be a word containing only the c_k. If the local system on $\boldsymbol{\vartheta}$ is trivial then the type D structure $\mathrm{M}(\boldsymbol{\gamma}, \boldsymbol{\vartheta})$ is obtained by applying the rules*

$$\mathrm{M}(c_k, a_j) = a_j, \quad \mathrm{M}(c_k, b_j) = b_j, \quad \mathrm{M}(c_k, c_j) = c_{j+k}$$

to every letter in $\boldsymbol{\vartheta}$, ranging over all letters for $\boldsymbol{\gamma}$, and assembling the result together using a toroidal grid to match up the endpoints, as described in Figure 13.

The proof of this result is contained in [6]; however, because nontrivial local systems are not handled there, we want to be precise about how to extend the result based on the material in our earlier work.

Proposition 12 *Proposition 11 holds when $\boldsymbol{\vartheta}$ carries a local system, where, for each letter u in the word representing $\boldsymbol{\vartheta}$ and each c_k in the word representing $\boldsymbol{\gamma}$, $\mathrm{M}(c_k, u)$ carries the same local system as u.*

[3] More specifically, treating $\boldsymbol{\gamma}$ and $\boldsymbol{\vartheta}$ as type D structures, in [6] we use $\mathrm{M}(\boldsymbol{\gamma}, \boldsymbol{\vartheta})$ as a shorthand for the type D structure $\widehat{CFDAA}(\mathcal{P} \times S^1) \boxtimes (\boldsymbol{\gamma}, \boldsymbol{\vartheta})$, where $\widehat{CFDAA}(\mathcal{P} \times S^1)$ is the bordered trimodule calculated in [3] and the three-manifold $\mathcal{P} \times S^1$ is a circle bundle over a pair of pants (this plays a key role in the next section).

[4] In fact, everything we do works in a much more general setting: Any manifold admitting L–space surgeries has a type D structure that, relative to a slope corresponding to an L–space filling, can be expressed in terms of only letters c_k. We have opted to simplify matters and focus on a well-known construction with well-established conventions in order to illustrate the key principle. More general cases follow the same lines, and can be extracted from [4; 6].

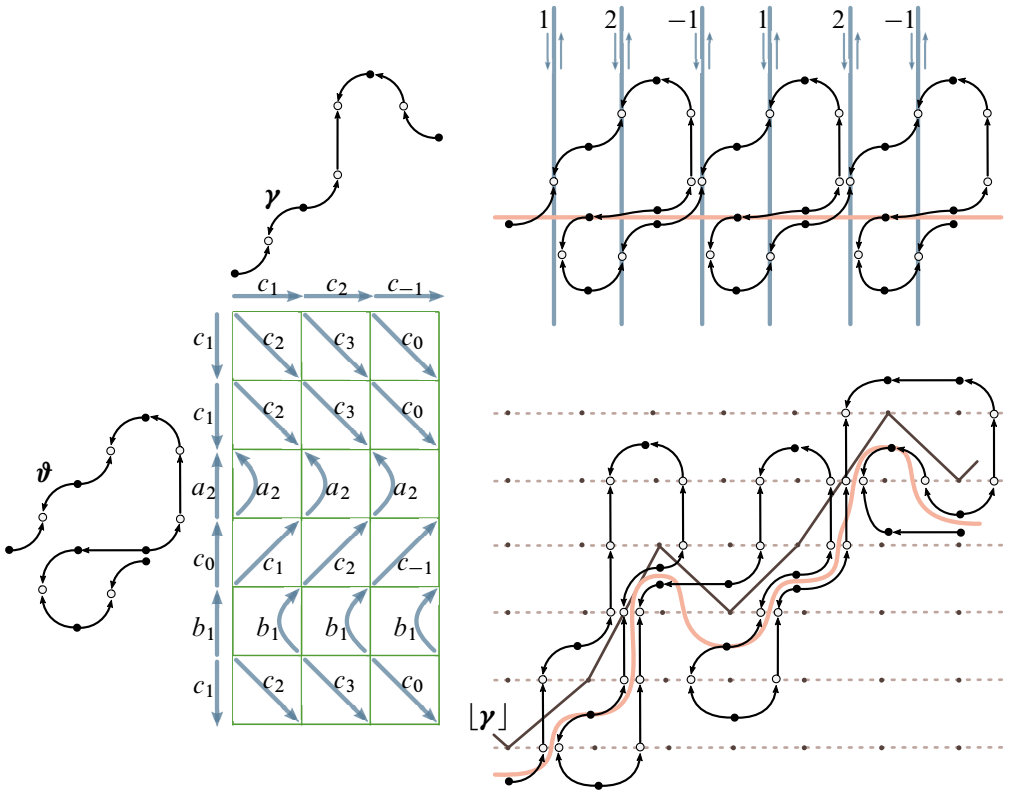


Figure 13: Merging a pair of curves, as in Proposition 11, described graphically: On the left-hand side of the diagram, the output curve is interpreted on a toroidal grid, where the c_k from $\boldsymbol{\gamma}$ (written on the horizontal) act on the letters in $\boldsymbol{\vartheta}$ (written on the vertical). On the right, this process in interpreted in terms of curves, where the top right figure gives a section of the (periodic) curve in $\widetilde{T}_\bullet$ while the bottom right figure is the result of the merge. Note that the horizontal is moved to the *key curve* $\lfloor\boldsymbol{\gamma}\rfloor$.

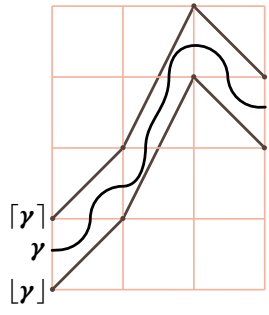


Figure 14: The PL *key curves* approximating $\boldsymbol{\gamma}$.

Proof There is nothing to check for trivial local systems, as these are just disjoint copies of some curve and Proposition 11 applies. For a nontrivial local system, we need to carry out the computation in [6, Figures 10, 11 and 12], replacing the simple segments for the ℓ_2 input with segments carrying an arbitrary local system as in, for example, Figure 12. This is a straightforward computation. Note that we do not need to check this computation for type b_k pieces, since we may assume a local system on a loop is concentrated on any one letter and a loop containing a b_k must also contain an a_j. (Furthermore if $\boldsymbol{\vartheta}$ is the multicurve corresponding to a knot in S^3, as in this paper, it is enough to check the computation for a_k pieces since any component with a nontrivial local system must contain an a_k piece.) □

The takeaway from Proposition 11 (and its extension to nontrivial local systems in Proposition 12) is a graphical calculus used to determine the merge of two curves when one contains only c_k segments; this is the content of Figure 13. Consider a word (in the c_k) representing $\boldsymbol{\gamma}$, and write this along the top of a rectangular grid; consider a word in $\mathfrak{A}$ representing $\boldsymbol{\vartheta}$, and write this along the side of the rectangular grid. Then, following the letter-by-letter instructions in Proposition 11, the new word $\mathrm{M}(\boldsymbol{\gamma}, \boldsymbol{\vartheta})$ can be obtained by running through the grid, starting at the top left. As shown in Figure 13, the $\mathrm{M}(c_j, c_k) = c_{j+k}$ run diagonally, while the $\mathrm{M}(c_j, a_k) = a_k$ and $\mathrm{M}(c_j, b_k) = b_k$ change direction. The sides of this grid are identified to form a toroidal grid, and this connects up the endpoints of the segments to form the new loop $\mathrm{M}(\boldsymbol{\gamma}, \boldsymbol{\vartheta})$.

Note that, as the resulting loop is traversed, horizontal motion in the grid corresponds exactly to horizontal motion of the corresponding curve in the plane. In particular, each vertical line of lattice points in the plane corresponds to a column of the grid containing some c_k in $\boldsymbol{\gamma}$, and the effect of merging on the curve for $\boldsymbol{\vartheta}$ is to shear the plane along that vertical line by k.

The graphical shorthand from Figure 13 suggests an interpretation of the merge operation in terms of immersed curves, which we can think of as $\boldsymbol{\gamma}$ acting on $\boldsymbol{\vartheta}$. To describe this, it is useful to have a piecewise-linear representative of the curve $\boldsymbol{\gamma}$. Let $\boldsymbol{\gamma}$ be expressible as a word in only the c_k, so that, viewed in the plane $\widetilde{T}_\bullet$, $\boldsymbol{\gamma}$ is a graph. Let $\lfloor \boldsymbol{\gamma} \rfloor$ be the curve consisting of linear segments that, at each integer in the horizontal direction, intersects the lattice point immediately below $\boldsymbol{\gamma}$. The curve $\lceil \boldsymbol{\gamma} \rceil$ is defined similarly, by instead pushing up to the lattice points immediately above $\boldsymbol{\gamma}$; see Figure 14. Note that this is closely related to the *pegboard diagrams* introduced in [4].

Recall that the immersed multicurve $\widehat{HF}(M)$ coming from a bordered 3–manifold M lives in the cylinder $(\mathbb{R}^2 \setminus \mathbb{Z}^2)/\langle\lambda\rangle$, where λ corresponds to the homological longitude of M. Equivalently, we think of this as a multicurve in $\mathbb{R}^2 \setminus \mathbb{Z}^2$ which is invariant under the action of λ. We will say that such a curve has *horizontal period* p if translation by λ moves p units in the horizontal direction.

Corollary 13 *Let $\boldsymbol{\vartheta}$ be an immersed multicurve with local systems in $\widetilde{T}_\bullet$ with horizontal period q, and let $\boldsymbol{\gamma}$ be a curve in $\widetilde{T}_\bullet$ with no vertical tangent lines (ie the graph of a function) with horizontal period p, with p and q relatively prime. Then the immersed multicurve for $\mathrm{M}(\boldsymbol{\gamma}, \boldsymbol{\vartheta})$ is obtained by adding $\lfloor\boldsymbol{\gamma}\rfloor$ to $\boldsymbol{\vartheta}$ vertically. That is, we find the image $\boldsymbol{\vartheta}$ under the transformation of $\widetilde{T}_\bullet$ which translates along each vertical line to take the horizontal axis to $\lfloor\boldsymbol{\gamma}\rfloor$.*

Proof This is the main thrust of Figure 13: For a component homologous to λ, the new cyclic word moves q columns to the right in the grid each time it traverses the grid vertically. Since p and q are relatively prime, the new word makes p vertical passes, tracing out the entire grid, before returning to the starting point. The new word is p copies of the word representing $\boldsymbol{\vartheta}$, with the indices on type c letters shifted according to the column in the grid; this corresponds to p copies of the fundamental region in $\widetilde{\boldsymbol{\vartheta}}$, each of which moves q units to the right, with a plane shear applied along each column of lattice points. The magnitude of each shear is determined by the index of the corresponding letter in $\boldsymbol{\gamma}$, which amounts to shifting each column upwards by the height of $\lfloor\widetilde{\boldsymbol{\gamma}}\rfloor$ in that column. The resulting curve has horizontal period pq. For a nullhomologous component, the grid gives rise to p separate cyclic words, each traversing the grid vertically once starting in a different column. Each word is a copy of $\boldsymbol{\vartheta}$ with shifted indices on type c letters. The nullhomologous component of $\boldsymbol{\vartheta}$ lifts to infinitely many copies of the same closed curve in $\widetilde{\boldsymbol{\vartheta}}$, which are translations of each other by multiples of λ. Taking p consecutive copies corresponds to the p cyclic words in the grid, and adding $\lfloor\widetilde{\boldsymbol{\gamma}}\rfloor$ corresponds to the required shifts in indices. □

As mentioned previously, the setup of Corollary 13 is more general than we need for cabling; we will only need the case that $\boldsymbol{\gamma}$ is in fact a straight line of some rational slope. Note that, when $\boldsymbol{\gamma}$ is a line of slope 1, the transformation taking the horizontal axis to $\lfloor\boldsymbol{\gamma}\rfloor$ is a lift to $\widetilde{T}_\bullet$ of a Dehn twist in $T_\bullet$. This is a linear transformation of the plane, which we refer to as a *plane shear* in the vertical direction. The case that $\boldsymbol{\gamma}$ is a line of rational slope is a mild generalization of this, which we call a *fractional plane shear*.

2 The proof of Theorem 1

In order to complete the proof of our theorem, we need to connect the operation described in Corollary 13 to the specific context of cabling. To do this we first set our conventions.

2.1 Cabling conventions

Recall that, fixing a knot K, we let M denote the complement $S^3 \setminus \nu(K)$ and $M_{p,q}$ denote the complement of the cable $C_{p,q}(K)$. Let $\mathcal{P}$ denote a two-sphere with three disks removed (so that $\mathcal{P}$ is homeomorphic to a pair of pants). The manifold $M_{p,q}$ can be obtained by gluing M into one boundary component of $\mathcal{P} \times S^1$ and an appropriately framed solid torus $D^2 \times S^1$ into another boundary component. We will briefly review this construction, paying particular attention to framing conventions.

Each torus boundary in this construction has a natural choice of parametrizing curves. For ∂M, we use a meridian μ and the Seifert longitude λ, fixing orientations on these curves with the convention that $\lambda \cdot \mu = +1$. For $\partial(D^2 \times S^1)$, we let m be a meridian $\partial D^2 \times \{\text{pt}\}$ and let ℓ be the longitude $\{\text{pt}\} \times S^1$, with the orientation convention that $m \cdot \ell = +1$. For $i \in \{1, 2, 3\}$, the i^{th} boundary component of the S^1–bundle $\mathcal{P} \times S^1$ is parametrized by a fibre $f_i = \{\text{pt}\} \times S^1$ and $b_i = \partial_i(\mathcal{P}) \times \{\text{pt}\}$, where $\partial_i(\mathcal{P})$ denotes the i^{th} boundary component of the base surface $\mathcal{P}$. We set orientations on these curves so that $b_i \cdot f_i = +1$.

The third boundary component of $\mathcal{P} \times S^1$ will ultimately become the boundary of $M_{p,q}$; however, it is helpful for the moment to fill this third boundary component in with a solid torus in a trivial way so that $\mathcal{P} \times S^1$ becomes $A \times S^1$, where A is an annulus. This solid torus can be removed later by deleting a neighbourhood of a fibre of $A \times S^1$. We glue $D^2 \times S^1$ to the first boundary component of $\mathcal{P} \times S^1$ (now $A \times S^1$) such that f_1 is identified with $p\ell + qm$; this means that b_1 is identified with $r\ell + sm$ for some integers r and s with $ps - qr = -1$ (we can choose r and s arbitrarily subject to this condition, but the choice affects the framings on the resulting boundary components). The result of this gluing is a solid torus, equipped with a Seifert fibration in which the core of the solid torus is a singular fibre and the regular fibres wind p times longitudinally and q times meridionally. This solid torus is glued to the knot complement M such that the result is S^3 and the core of the solid torus is identified with K. As a result, a regular fibre of $D^2 \times S^1 \cup A \times S^1$ is the cable $C_{p,q}(K)$, and removing a neighbourhood of

one of these (or, equivalently, not filling in the third boundary of $\mathcal{P}\times S^1$) yields the complement $M_{p,q}$.

Note that inserting $A\times S^1\cong T^2\times[0,1]$ between M and $D^2\times S^1$ amounts to a change of framing and, in particular, f_1 can be identified with f_2 and b_1 can be identified with $-b_2$. To recover S^3, we want μ to be identified with m and λ to be identified with ℓ. It follows that f_2 glues to $p\lambda+q\mu$ and b_2 glues to $-r\lambda-s\mu$. To summarize, we have

$$\begin{bmatrix} f_1\\ b_1\end{bmatrix}=\begin{bmatrix} p & q\\ r & s\end{bmatrix}\begin{bmatrix}\ell\\ m\end{bmatrix},\quad \begin{bmatrix} f_2\\ b_2\end{bmatrix}=\begin{bmatrix} p & q\\ -r & -s\end{bmatrix}\begin{bmatrix}\lambda\\ \mu\end{bmatrix}$$

with $ps-qr=-1$. Inverting these matrices,

$$\begin{bmatrix}\ell\\ m\end{bmatrix}=\begin{bmatrix} -s & q\\ r & -p\end{bmatrix}\begin{bmatrix} f_1\\ b_1\end{bmatrix},\quad \begin{bmatrix}\lambda\\ \mu\end{bmatrix}=\begin{bmatrix} -s & -q\\ r & p\end{bmatrix}\begin{bmatrix} f_2\\ b_2\end{bmatrix}.$$

If we do not fill in the third boundary of $\mathcal{P}\times S^1$ in the construction above, the resulting manifold with torus boundary is $M_{p,q}$; it is clear that the meridian μ_C of the cable knot is given by b_3. While not required, as will follow from the computation below, one can check that the Seifert longitude λ_C of the cable knot is given by $-f_3+q^2b_3$.

2.2 Applying the merge operation

We are interested in obtaining the immersed curve set $\widehat{HF}(M_{p,q})$ from the immersed curve set $\widehat{HF}(M)$. We can do this by applying the merge operation to $\widehat{HF}(M)$ and $\widehat{HF}(D^2\times S^1)$, keeping in mind the framings discussed above. Following Section 1 and the conventions in [6], the first step is to draw lifts of both curve sets in the plane with respect to the parametrization by f_i and b_i (or, more precisely, by the curves in $\partial(D^2\times S^1)$ or ∂M which are identified with f_i and b_i) such that b_i is the horizontal direction and f_i is the vertical direction. Recall our convention is that $b_i\cdot f_i=1$, but we are now considering the plane as a lift of $\partial(D^2\times S^1)$ or of ∂M, which are identified with boundary tori of $\mathcal{P}\times S^1$ by an orientation-reversing diffeomorphism, so if we take b_i to be the positive horizontal axis then f_i is the negative vertical axis. Note that we could instead choose the opposite orientation for both axes, but this ambiguity can be ignored since immersed curves for bordered invariants are symmetric under the elliptic involution of the torus by [5, Theorem 7]. Since $\widehat{HF}(D^2\times S^1)$ is the meridian $m=rf_1-pb_1$, this curve is simple to describe in the relevant basis: it is a line of slope $\frac{r}{p}$ [4]. The second step is to apply Corollary 13 by taking the vertical sum of $\widehat{HF}(M)$ and $\lfloor m\rfloor$; note that the conditions of the corollary are satisfied because

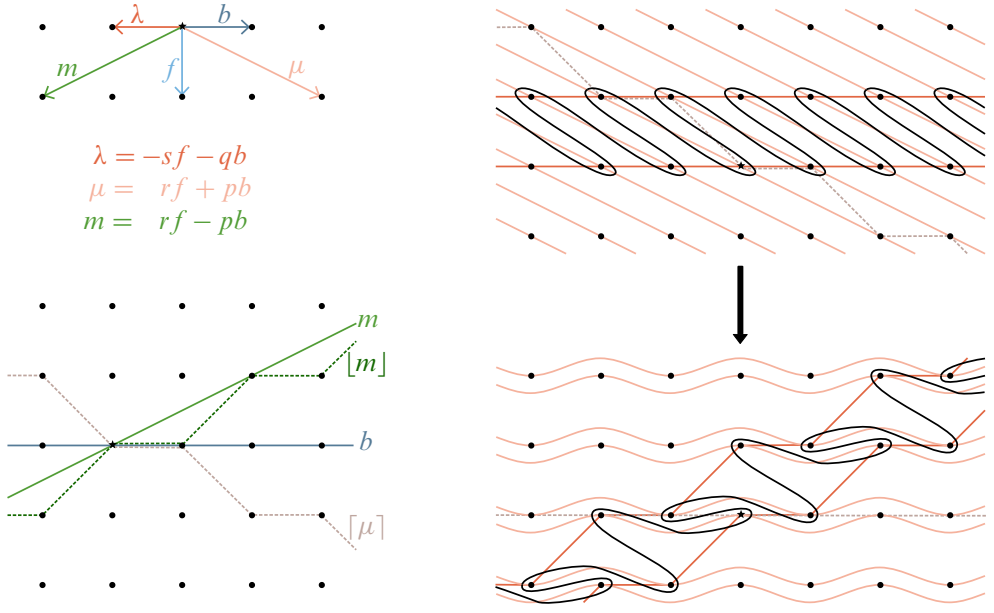


Figure 15: The fractional plane shear in the vertical direction associated with computing a (p,q)–cable, viewed with respect to the $(-f,b)$–framing. Other relevant curves are shown, with respect to this framing, in the top left. The bottom left shows a copy of m through the origin and the corresponding curve $\lfloor m \rfloor$ obtained by dropping down to the highest peg below m in each column. This curve serves as a "key" for the plane shear — that is, we shift each column of pegs upward by the height of $\lfloor m \rfloor$ in that column. Thus the plane shear is determined by the fact that it takes b to $\lfloor m \rfloor$, or equivalently that it takes $\lceil \mu \rceil$ to b. The right shows the effect of this shear on the curve for the right-hand trefoil. For the concrete example in the figure, $(p,q) = (2,1)$ and $(r,s) = (1,0)$.

components $\widehat{HF}(M)$ are homologous to zero or to the rational longitude λ, which moves horizontally by q units, while m moves horizontally by p units. The result is (a lift to the plane of) $\widehat{HF}(M_{p,q})$, though given with respect to the framing (f_3, b_3) rather than the usual (μ_C, λ_C); see Figure 15.

While the previous paragraph gives a complete procedure for computing $\widehat{HF}(M_{p,q})$, performing the change of basis to draw the curve set $\widehat{HF}(M)$ with respect to the (f,b)–framing can be cumbersome. Instead, we can follow the same operation but view the plane with respect to (μ,λ), the preferred framing for ∂M, throughout the process. Now, instead of shifting pegs in each vertical column, we shift along lines parallel to the fibre direction; since $f_2 = p\lambda + q\mu$, this is a line of slope $\frac{q}{p}$. To keep

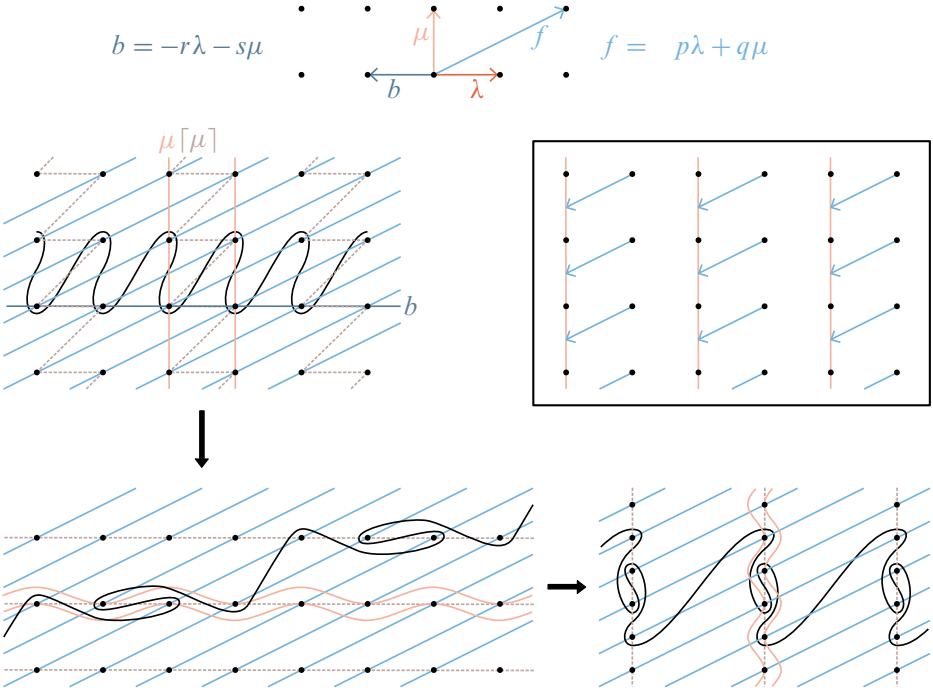


Figure 16: Starting with the curve $\widehat{HF}(M)$ drawn in the plane with respect to the standard (μ, λ)–framing, the fractional plane shear in the f direction which takes $\lceil\mu\rceil$ to b produces the curve $\widehat{HF}(M_{p,q})$, though not in terms of a convenient parametrization. shearing back partially along f gives the curve with the standard parametrization, up to rescaling the lattice. These two steps can be combined into one, as shown in the box: each lattice point is translated leftward along lines of slope $\frac{q}{p}$ until its x–coordinate is a multiple of p. Note that p copies of the curve $\widehat{HF}(M)$ are involved in each copy of $\widehat{HF}(M_{p,q})$. The figure shows the case of the $(2, 1)$–cable of the right-hand trefoil.

track of how much to shift along each line of slope $\frac{q}{p}$, we can draw a copy of the piecewise linear curve $\lceil\mu\rceil$; note that this is obtained from a vertical line μ through the origin by pushing each point $\left(0, \frac{n}{p}\right)$ on μ leftward along a line of slope $\frac{q}{p}$ to the first lattice point it encounters (see Figure 16, top left). To perform the cable operation, we shear along lines of slope $\frac{q}{p}$ to bring this curve $\lceil\mu\rceil$ to b (see Figure 16, bottom left).

We can now start with the curve $\widehat{HF}(M)$ represented in terms of its standard framing (μ, λ) and produce the immersed curve $\widehat{HF}(M_{p,q})$ in one simple step. However, as before, the output is not given with respect to the standard framing by (μ_C, λ_C). Of course, it is straightforward to determine the slopes of μ_C and λ_C in the output picture

and then we simply need to change basis applying a linear map to the plane which takes these to the vertical and horizontal directions, respectively. This can always be accomplished by a sequence of (integral) plane shears in the horizontal and vertical directions. However, this too is cumbersome, so we will describe a shortcut to this reparametrization making use of a linear transformation of the plane which does not preserve the lattice. More precisely, consider the linear transformation which fixes f and takes b to μ; this can be understood as translating each lattice point on b along a line of slope $\frac{q}{p}$ until it reaches the vertical line μ (see Figure 16, bottom right). Note that the lattice $\mathbb{Z}^2$ is not mapped to itself under this transformation, but rather its image is $p\mathbb{Z} \times \frac{1}{p}\mathbb{Z}$. Even so, in this new deformed lattice the directions corresponding to μ_C and λ_C are vertical and horizontal, as desired, and we can recover the usual lattice by ending with another linear transformation which scales and compresses by a factor of p in the vertical direction and horizontal direction, respectively.

Finally, we mention that there are now two steps which involve shearing along the lines of slope $\frac{q}{p}$: the fractional plane shear taking $\lceil\mu\rceil$ to b (this transformation is not linear), and the linear transformation taking b to μ. These steps can be combined in one by shearing along lines of slope $\frac{q}{p}$ to push $\lceil\mu\rceil$ onto μ. In other words, every p^{th} vertical column of lattice points is fixed, while all other points are pushed leftward along lines of slope $\frac{q}{p}$ until they reach a vertical line containing one of the fixed columns (see the boxed portion of Figure 16). This proves Theorem 1.

Acknowledgements Hanselman was partially supported by NSF grant DMS-1812527; Watson was partially supported by an NSERC discovery/accelerator grant.

References

[1] **I Dai**, **J Hom**, **M Stoffregen**, **L Truong**, *More concordance homomorphisms from knot Floer homology*, Geom. Topol. 25 (2021) 275–338 MR Zbl

[2] **M H Freedman**, *The topology of four-dimensional manifolds*, J. Differential Geometry 17 (1982) 357–453 MR Zbl

[3] **J Hanselman**, *Bordered Heegaard Floer homology and graph manifolds*, Algebr. Geom. Topol. 16 (2016) 3103–3166 MR Zbl

[4] **J Hanselman**, **J Rasmussen**, **L Watson**, *Bordered Floer homology for manifolds with torus boundary via immersed curves*, preprint (2016) arXiv 1604.03466

[5] **J Hanselman**, **J Rasmussen**, **L Watson**, *Heegaard Floer homology for manifolds with torus boundary: properties and examples*, Proc. Lond. Math. Soc. (3) 125 (2022) 879–967 MR

[6] **J Hanselman**, **L Watson**, *A calculus for bordered Floer homology*, Geom. Topol. 27 (2023) 823–924

[7] **M Hedden**, *On knot Floer homology and cabling*, PhD thesis, Columbia University (2005) MR Available at https://www.proquest.com/docview/305015665

[8] **M Hedden**, *On knot Floer homology and cabling*, Algebr. Geom. Topol. 5 (2005) 1197–1222 MR Zbl

[9] **M Hedden**, *On knot Floer homology and cabling, II*, Int. Math. Res. Not. 2009 (2009) 2248–2274 MR Zbl

[10] **J Hom**, *A note on cabling and L–space surgeries*, Algebr. Geom. Topol. 11 (2011) 219–223 MR Zbl

[11] **J Hom**, *Bordered Heegaard Floer homology and the tau-invariant of cable knots*, J. Topol. 7 (2014) 287–326 MR Zbl

[12] **J Hom**, *A survey on Heegaard Floer homology and concordance*, J. Knot Theory Ramifications 26 (2017) art. id. 1740015 MR Zbl

[13] **A S Levine**, *Knot doubling operators and bordered Heegaard Floer homology*, J. Topol. 5 (2012) 651–712 MR Zbl

[14] **R Lipshitz**, **P S Ozsváth**, **D P Thurston**, *Bimodules in bordered Heegaard Floer homology*, Geom. Topol. 19 (2015) 525–724 MR Zbl

[15] **R Lipshitz**, **P S Ozsvath**, **D P Thurston**, *Bordered Heegaard Floer homology*, Mem. Amer. Math. Soc. 1216, Amer. Math. Soc., Providence, RI (2018) MR Zbl

[16] **L Mosher**, *What is a train track?*, Notices Amer. Math. Soc. 50 (2003) 354–355

[17] **P S Ozsváth**, **A I Stipsicz**, **Z Szabó**, *Concordance homomorphisms from knot Floer homology*, Adv. Math. 315 (2017) 366–426 MR Zbl

[18] **P Ozsváth**, **Z Szabó**, *Holomorphic disks and knot invariants*, Adv. Math. 186 (2004) 58–116 MR Zbl

[19] **I Petkova**, *Cables of thin knots and bordered Heegaard Floer homology*, Quantum Topol. 4 (2013) 377–409 MR Zbl

[20] **J A Rasmussen**, *Floer homology and knot complements*, PhD thesis, Harvard University (2003) MR arXiv math/0306378

[21] **C A Van Cott**, *Ozsváth–Szabó and Rasmussen invariants of cable knots*, Algebr. Geom. Topol. 10 (2010) 825–836 MR Zbl

Department of Mathematics, Princeton University
Princeton, NJ, United States

Department of Mathematics, University of British Columbia
Vancouver BC, Canada

jh66@math.princeton.edu, liam@math.ubc.ca

Proposed: Ciprian Manolescu Received: 12 August 2019
Seconded: Peter Ozsváth, András I Stipsicz Revised: 22 November 2019

msp

Geometry & Topology **27**:3 (2023) 953–1082
DOI: 10.2140/gt.2023.27.953
Published: 8 June 2023

Combinatorial Reeb dynamics on punctured contact 3–manifolds

RUSSELL AVDEK

Let $\Lambda^{\pm} = \Lambda^{+} \cup \Lambda^{-} \subset (\mathbb{R}^3, \xi_{\text{std}})$ be a contact surgery diagram determining a closed, connected contact 3–manifold $(S^3_{\Lambda^{\pm}}, \xi_{\Lambda^{\pm}})$ and an open contact manifold $(\mathbb{R}^3_{\Lambda^{\pm}}, \xi_{\Lambda^{\pm}})$. Following work of Bourgeois, Ekholm and Eliashberg, we demonstrate how $\Lambda^{\pm}$ determines a family α_ϵ of contact forms for $(\mathbb{R}^3_{\Lambda^{\pm}}, \xi_{\Lambda^{\pm}})$ whose closed Reeb orbits are in one-to-one correspondence with cyclic words of composable Reeb chords on $\Lambda^{\pm}$. We compute the homology classes and integral Conley–Zehnder indices of these orbits diagrammatically and develop algebraic tools for studying holomorphic curves in surgery cobordisms between the $(\mathbb{R}^3_{\Lambda^{\pm}}, \xi_{\Lambda^{\pm}})$.

These new techniques are used to describe the first known examples of closed, tight contact manifolds with vanishing contact homology: they are contact $1/k$ surgeries along the right-handed, $\text{tb} = 1$ trefoil for $k > 0$, which are known to have nonzero Heegaard Floer contact classes by work of Lisca and Stipsicz.

53D42; 57K33

1 Introduction

The main objects of interest in this paper are *contact 3–manifolds* and their *Legendrian submanifolds*. A *contact form* on an oriented 3–manifold M is a 1–form $\alpha \in \Omega^1(M)$ for which $\alpha \wedge d\alpha > 0$ with respect to the orientation of M. A contact 3–manifold is a pair (M, ξ) consisting of an oriented 3–manifold M together with an oriented 2–dimensional distribution $\xi \subset TM$ which is the kernel of a contact form α satisfying $d\alpha|_\xi > 0$ with respect to the orientation on ξ. We say that α *is a contact form for* (M, ξ). A Legendrian submanifold of (M, ξ) is a link which is tangent to ξ. We'll typically denote Legendrian submanifolds by Λ or Λ^0.

Given a contact 1–form α for some (M, ξ) its *Reeb vector field*, R, is determined by the equations

$$\alpha(R) = 1, \quad d\alpha(R, *) = 0.$$

For the purposes of studying invariants of (M, ξ) and its Legendrian submanifolds defined by counting holomorphic curves — see Eliashberg, Givental and Hofer [23], Etnyre and Ng [26], Hutchings [40] and Seidel [59] — we are interested in finding contact forms on a given (M, ξ) for which R is easy to analyze. Specifically we want to have visibility into the closed orbits of R as well the *chords* of Legendrians $\Lambda^0 \subset (M, \xi)$, that is, the orbits of R parametrized by compact intervals which both begin and end on Λ^0.

Let $(\mathbb{R}^3, \xi_{\text{std}})$ denote the standard contact structure on Euclidean 3–space, where

$$\xi_{\text{std}} = \ker(\alpha_{\text{std}}), \quad \alpha_{\text{std}} = dz - y\, dx,$$

and let (S^3, ξ_{std}) denote the standard contact structure on the unit 3–sphere S^3, where

$$\xi_{\text{std}} = \ker\left(\sum_1^2 x_i\, dy_i - y_i\, dx_i\right).$$

A *contact surgery diagram* is a Legendrian link

$$\Lambda^\pm = \Lambda^+ \cup \Lambda^- \subset (\mathbb{R}^3, \xi_{\text{std}}).$$

Performing contact ± 1 surgery on the components of the $\Lambda^\pm$ as defined by Ding and Geiges [16] produces a contact 3–manifold, which we will denote by $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$. By considering $(\mathbb{R}^3, \xi_{\text{std}})$ as being contained in (S^3, ξ_{std}), we can view the surgery diagram $\Lambda^\pm$ as determining a closed contact 3–manifold $(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$, with $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ obtained by removing a point from $(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$. As proved by Ding and Geiges

in [16] — see also Avdek [2] — every closed, connected contact 3–manifold (M, ξ) can be described as $(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ for some choice of $\Lambda^\pm$.

For the remainder of this introduction we assume basic familiar with contact surgery, Weinstein handle attachment and symplectic field theory (SFT). Further background and references will be provided in Section 2.

1.1 Combinatorial Reeb dynamics on punctured contact 3–manifolds

The primary purpose of this article is to describe a family of particularly well-behaved contact forms α_ϵ for $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ which are determined by the surgery diagram $\Lambda^\pm$. Our intention is to extend the analysis of Reeb dynamics appearing in work of Bourgeois, Ekholm and Eliashberg [7; 18] to allow for contact $+1$ surgeries. In particular, the following theorem states that their "chord-to-orbit correspondence" is applicable to any closed contact 3–manifold:[1]

Theorem 1.1 *Let $\Lambda^\pm$ be a contact surgery diagram presented in the front projection, where each component is equipped with an orientation. Possibly after a Legendrian isotopy of $\Lambda^\pm$ which preserves the front projection up to isotopy, there is*

(1) *a constant ϵ_0,*

(2) *a neighborhood N_{ϵ_0} of $\Lambda^\pm$ in $\mathbb{R}^3$, and*

(3) *a family of contact forms α_ϵ with Reeb vector fields R_ϵ parametrized by $\epsilon < \epsilon_0$ on $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$*

such that the following conditions hold:

(1) *For any $\epsilon < \epsilon_0$, performing contact surgery along a neighborhood $N_\epsilon \subset N_{\epsilon_0}$ produces $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ equipped with the contact form α_ϵ.*

(2) *$\alpha_\epsilon = \alpha_{\text{std}}$ on the complement of N_ϵ.*

(3) *For any $\epsilon < \epsilon_0$, there is a one-to-one correspondence between cyclic words of composable ∂_z chords of $\Lambda^\pm$ and closed orbits of R_ϵ (Theorem 5.1).*

(4) *For a given cyclic word of chords w, there exists $\epsilon_w < \epsilon_0$ such that the orbits of R_ϵ corresponding to w are hyperbolic for $\epsilon < \epsilon_w$ (Theorem 5.3).*

(5) *There is pair of sections (X, Y) of $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ determined by $\Lambda^\pm$ and its orientation, providing a symplectic trivialization of the restriction of $(\xi_{\Lambda^\pm}, d\alpha_\epsilon)$ to all closed orbits of R_ϵ. The zero locus $X^{-1}(0) = Y^{-1}(0)$ is a link contained in*

[1] Contact -1 surgery — also known as Legendrian surgery — describes how the convex boundaries of Liouville domains are modified by critical-index Weinstein handle attachment.

$(\mathbb{R}^3 \setminus N_\epsilon) \subset \mathbb{R}^3_{\Lambda^\pm}$ *whose connected components are given by transverse push-offs of the components of* $\Lambda^\pm$ *with nonzero rotation number* (Theorem 6.1).

(6) *The integral Conley–Zehnder indices* $\mathrm{CZ}_{X,Y}$ (Theorem 7.1) *and homology classes* (Theorem 9.1) *of the closed orbits of* R_ϵ *can be computed combinatorially from the surgery diagram.*

By "computed combinatorially", we mean computed via extensions of methods typically used to set up chain complexes for the Legendrian contact homology (LCH) [26] or the Legendrian rational symplectic field theory (*LRSFT*) of Ng [50] of $\Lambda^\pm$. Analogous results are stated for chords of Legendrian links $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ throughout the paper, providing a "chord-to-chord" correspondence with diagrammatically computable Maslov indices. The content of Theorem 1.1 is sufficient to compute some algebraic invariants of tight contact structures on the lens space $L(2,1)$ and $S^1 \times S^2$, as shown in Section 12.1.

The dynamics analysis of Theorem 1.1 can be supplemented with a direct limit argument as in Ekholm and Ng [20, Section 4] to obtain a description of the Reeb dynamics on the closed contact manifolds $(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ associated to a contact surgery diagram, which introduces a pair of embedded elliptic orbits.[2] We will not pursue analysis of closed contact manifolds in this paper, as the open manifolds $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ have particularly friendly geometries, which we'll leverage in applications.

1.2 Constrained topology of holomorphic curves and applications

The secondary purpose of this article is to develop tools for studying holomorphic curves in symplectizations of the $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ and in surgery cobordisms between them. Our intention is to make "hat versions" of holomorphic curve invariants of $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ — as defined by Colin, Ghiggini, Honda and Hutchings [14, Section 7.1] — more computationally accessible. Theorem 1.1 already provides us with rather complete descriptions of the chain complexes underlying such invariants.[3] In particular, we'll be interested in the hat version of contact homology (CH),

$$\widehat{\mathrm{CH}}(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm}) = \mathrm{CH}(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm}).$$

[2] See for example Bourgeois [6, Section 4.1] and Hutchings [40, Example 1.8].

[3] There is some subtlety for $\widehat{\mathrm{ECH}}$: In order to compute relative ECH indices, the links underlying collections of simple Reeb orbits should be known, whereas we will describe the homotopy classes of closed Reeb orbits. Such link embeddings can be computed as solutions to matrix arithmetic problems described in Section 5.6.

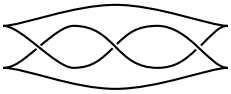


Figure 1: The Legendrian trefoil of Theorem 1.2 shown in the front projection.

Hat versions of other holomorphic curve invariants of $(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ such as embedded contact homology ($\widehat{\mathrm{ECH}}$) and the *SFT* algebra ($\widehat{SFT}$) are defined analogously.[4]

We demonstrate the utility of our tools in two applications: First we provide a (slightly) new proof of the vanishing of CH of overtwisted contact manifolds (see Eliashberg and Yau [65]) using surgery-theoretic methods (Section 12.3). Second, we prove the following theorem (Section 12.5):

Theorem 1.2 *If $\Lambda^- = \varnothing$ and Λ^+ has a component which is a right-handed trefoil, then*

$$\mathrm{CH}(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm}) = \widehat{\mathrm{CH}}(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm}) = 0$$

In particular, contact $1/k$ surgery on the right-handed, $\mathrm{tb} = 1$ trefoil for $k > 0$ produces a closed, tight contact manifold $(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ with vanishing contact homology. (See Figure 1.)

The development of our tools (Section 11) starts with a variation of the construction of transverse knot filtrations of holomorphic curve invariants from [14, Section 7.2]: Lines in $\mathbb{R}^3$ directed by ∂_z over points $(x, y) \in \mathbb{R}^2 \setminus \pi_{x,y}(N_\epsilon)$ determine infinite-energy holomorphic planes $\mathbb{C}_{x,y}$ in $\mathbb{R} \times \mathbb{R}^3_{\Lambda^\pm}$. The $\mathbb{C}_{x,y}$ form a holomorphic foliation whose existence constrains the topology of curves à la the proofs of uniqueness-of-symplectic-manifold theorems of Eliashberg [22], Gromov [32], Geiges and Zehmisch [30], Hind [34], McDuff [44; 45] and Wendl [61]. Counting intersections $\mathbb{C}_{x,y} \cdot U$ of these planes with finite-energy curves U asymptotic to collections $\gamma^\pm$ of closed R_ϵ orbits yields locally constant, $\mathbb{Z}_{\geq 0}$–valued functions on SFT moduli spaces — topological invariants determined by the relative homology classes

$$[\pi_{\mathbb{R}^3_{\Lambda^\pm}} \circ U] \in H_2(\mathbb{R}^3_{\Lambda^\pm}, \gamma^\pm)$$

of holomorphic curves. Surgery cobordisms may be similarly considered when equipped with special almost-complex structures described in Section 11.2. By tracking these intersections, we can:

(1) Show that certain disks appearing in Ng's combinatorially defined Legendrian *RSFT* [50] determine rigid holomorphic planes in $\mathbb{R} \times \mathbb{R}^3_{\Lambda^\pm}$ (Section 12.2). This

[4] We use *SFT* to denote the *SFT* algebra, while SFT — without italics — refers to Eliashberg, Givental and Hofer's framework for defining holomorphic curve invariants of contact and symplectic manifolds of [23].

follows a Lagrangian-boundary version of Hofer's bubbling argument [35], in which case the $\mathbb{C}_{x,y} \cdot U$ completely dictate the ways in which certain families of holomorphic disks can degenerate into multilevel SFT buildings.

(2) Equip the $\widehat{\mathrm{CH}}$ chain complexes with a new grading, denoted by $\mathcal{I}_\Lambda$, which depends on the surgery diagram (Section 12.4). Variants of this grading may similarly be applied to any holomorphic curve invariant of $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$.

In the proof of Theorem 1.2, we show that $+1$ surgery on the $\mathrm{tb} = 1$ trefoil provides a $\mathrm{CZ}_{X,Y} = 2$ closed orbit γ of R_ϵ with $\partial_{\mathrm{CH}}\gamma = \pm 1 \in \mathbb{Q}$. Computations of Conley–Zehnder indices, homology classes, and $\mathcal{I}_\Lambda$ shows that any $\mathrm{ind} = 1$ rational holomorphic curves positively asymptotic to γ must be a plane, and that such planes may be counted using our bubbling argument.

Theorem 1.2 provides the first examples of closed, tight contact manifolds with $\mathrm{CH} = 0$.[5] The tightness of $1/k$ surgeries on the $\mathrm{tb} = 1$ trefoil is provided by computations of Heegaard Floer (HF) contact classes — see Honda, Kazez and Matić [38] and Ozsváth and Szabó [53] — by Lisca and Stipsicz in [43, Section 3]. As the HF contact class contains the same information as the ECH contact class — see Colin, Ghiggini and Honda [13] and Kutluhan, Lee and Taubes [42] — and both ECH and SFT count holomorphic curves of arbitrary topological type — in particular, arbitrary genus — it would be interesting to know if there is some SFT invariant of this contact manifold which is nonvanishing. Broadening the scope of this inquiry, we ask the following:

Question 1.3 *For 3–dimensional contact manifolds, does* $\mathrm{CH}(M, \xi) \neq 0$ *imply that the* $\mathrm{HF} = \mathrm{ECH}$ *contact class of* (M, ξ) *is nonzero? Do there exist tight contact manifolds of dimension greater than three with vanishing contact homology?*

We note that, using the algebraic formalism of [23], the vanishing of contact homology is equivalent to the vanishing of SFT according to Bourgeois and Niederkrüger [9].

Outline of this paper

In Section 2 we outline notation and background information which will be used throughout the rest of the paper. Section 3 is also primarily concerned with notation, associating algebraic data to chords of Legendrian links in $(\mathbb{R}^3, \xi_{\mathrm{std}})$, which will be used to package invariants of chords and closed orbits in the surgered contact manifolds $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$.

[5] Due to CH functoriality under Liouville cobordism, Honda's tight contact manifold which becomes overtwisted after contact -1 surgery [37] already provides an example of a contact manifold with convex boundary whose *sutured contact homology* [14] is zero.

Sections 4–9 carry out the computational details of Theorem 1.1 and analogous results for chords of Legendrian links $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$. In Section 10 we describe handle-attachment cobordisms between the $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ associated to surgeries along their Legendrian knots. The construction of these cobordisms — slight modifications of Ekholm [18] and Weinstein [60] — provides us with model geometry facilitating analysis of holomorphic curves.

Section 11 describes holomorphic curves in symplectizations of and surgery cobordisms between the $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$. The algebraic tools described in that section are prerequisite for the applications appearing in Section 12, culminating in the proof of Theorem 1.2.

Content pertaining to Legendrian links $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ may be skipped by readers only interested in the applications of Section 12. This material is included to provide a complete picture of relative SFT chain complexes in anticipation of their use in future applications.

Acknowledgments

The author is partly supported by the grant KAW 2016.0198 from the Knut and Alice Wallenberg Foundation. We send our gratitude to Erkao Bao, Guillaume Dreyer, Tobias Ekholm, Ko Honda, Yang Huang and Vera Vértesi for interesting conversations. Special thanks go to Fabio Gironella and András Stipsicz, as well as Georgios Dimitroglou Rizell, for their interest in this project and invitations to give talks in their seminars. Finally, we thank our referees for their thoughtful and detailed commentary.

2 Prerequisites

2.1 General notation

Throughout this paper $\delta_{*,*}$ — with a double subscript — will denote the Kronecker delta and $\lfloor * \rfloor$ will be the floor function $\mathbb{R} \to \mathbb{Z}$. A *collection* will be a set in which elements are allowed to have nontrivial multiplicity. We use set notation for collections. For example $\{1,1,2\}$ is a collection with $\{1,1,2\} \setminus \{1\} = \{1,2\}$ and $\{1,1,2\} \cup \{2,3\} = \{1,1,2,2,3\}$. We'll use often use collections and ordered collection to organize chords and orbits as they may appear in CH, ECH, LCH, etc.

Unless otherwise specified, we use I to denote a connected 1–manifold and, for a positive number ϵ, we write $I_\epsilon = [-\epsilon, \epsilon]$. For $a > 0$, the circle $\mathbb{R}/a\mathbb{Z}$ will be denoted

by S^1_a and, without the subscript, $S^1 = S^1_1$. The unit disk of dimension n and radius C centered about $x \in \mathbb{R}^n$ will be denoted by $\mathbb{D}^n_C(x)$. We'll typically use the simplified notation $\mathbb{D}^n = \mathbb{D}^n_1(0)$ and $\mathbb{D}$ for $\mathbb{D}^2$. The complex projective space will be written $\mathbb{P}^n$.

For a closed manifold M, $\widehat{M}$ will denote the open manifold obtained from M by removing a point or closed disk. When (M,ξ) is a closed contact manifold, $(\widehat{M,\xi})$ will denote (M,ξ) with a point or standard Darboux disk removed. We say that $(\widehat{M,\xi})$ is a *punctured contact manifold*.[6]

For a space M, we denote homology and cohomology groups by $H_*(M)$ and $H^*(M)$, respectively. Integral coefficients will be assumed unless otherwise explicitly stated. When M is a closed manifold, PD will be used to denote the Poincaré duality isomorphism in either direction, $H_i \leftrightarrow H^{\dim(M)-i}$. Abusing notation, we also use PD to denote the associated isomorphisms for punctured manifolds $\widehat{M}$ in degrees $i \neq 0$. By a $\mathbb{Q}$–*homology sphere*, we mean a closed or punctured 3–manifold with finite H_1 (implying that $H_2 = 0$ by the universal coefficients theorem; see [33, Corollary 3.3]).

For a vector bundle E over a manifold M, the space of C^∞ sections will denoted by $\Gamma(E)$. The space of nowhere-zero sections — which may be empty — will be denoted by $\Gamma_{\neq 0}(E)$. Provided that E has finite rank n and trivializations (V_i) and (W_i) of E over some set $U \subset M$, transformations of the form $\sum_{i,j} a_{i,j} W_i \otimes V_j^*$ can be written as matrices, with respect to which we say that (V_i) is the *incoming basis* and (W_i) is the *outgoing basis*. In such situations, provided $a_1, \dots, a_n \in C^\infty(U)$, $\mathrm{Diag}(a_1, \dots, a_n)$ will be the diagonal matrix with a_1 in the top-left corner and J_0 will denote standard complex multiplication where applicable. The Euler class of a finite-dimensional bundle will be written $e(E)$ and Chern classes will be written $c_k(E)$ when the bundle is equipped with a (homotopy class of) complex structure. We will be predominantly interested in the case $E = \xi$ for a 3–dimensional contact manifold (M,ξ), in which case the Euler and first Chern classes coincide: $e(\xi) = c_1(\xi)$.

2.2 Vector fields and almost-complex structures

In this section we review vector fields and almost-complex structures typically encountered in symplectic and contact geometry, primarily for the purpose of establishing conventions which often vary in the literature. We'll use Option 1 of [63]. See that article or [46, Remark 3.3] for further discussion.[7]

[6] In [14], the notation $M(1)$ is used for what we call $\widehat{M}$.

[7] Regarding work we'll be frequently referencing, our signs for symplectic forms on cotangent bundles will be opposite that of [18] and our signs for Hamiltonian vector fields are opposite that of [4; 3; 14].

Let (W,β) be a $2n$–dimensional exact symplectic manifold. That is, W is an oriented $2n$–manifold on which $d\beta$ is symplectic. We call such β a *Liouville form* or *symplectic potential*. If $H\in\mathcal{C}^\infty(W)$ is a smooth function with values in $\mathbb{R}$ or S^1, the associated *Hamiltonian vector field*, denoted by X_H, is the unique solution to the equation

$$d\beta(*,X_H)=dH.$$

Clearly H is constant along the flow-lines of X_H and X_H depends only on $d\beta$ (rather than β). If J is an almost-complex structure for which g_J, defined by

$$g_J(u,v)=d\beta(u,Jv),\quad u,v\in T_p\Sigma,$$

is a J–invariant Riemannian metric, then

$$X_H=J\nabla H,$$

where ∇H is the gradient of H with respect to g_J solving $g_J(\nabla H,*)=dH$. We say that such J is a *compatible* almost-complex structure.

The *Liouville vector field*, denoted by X_β, on W is the unique solution to the equation

$$d\beta(X_\beta,*)=\beta.$$

If W is compact and X_β points outward along the boundary of W, we say that the pair (W,β) is a *Liouville domain*. Given a function $H\in\mathcal{C}^\infty(W)$, the 1–form $\beta_H=\beta+dH$ is also a primitive for $d\beta$ such that

$$X_{\beta_H}=X_\beta+X_H.$$

By our choice of convention, Hamiltonian and Liouville vector fields interact with $d\beta$ via

$$\beta(X_H)=d\beta(X_\beta,X_H)=dH(X_\beta).$$

Given a contact manifold (M,ξ) equipped with a contact form α, action of the chords and closed orbits of its Reeb vector field may be computed as

$$\mathcal{A}(\gamma)=\int_\gamma\alpha.$$

2.3 Contact and symplectic manifolds

Here we review some contact and symplectic manifolds which will appear frequently.

2.3.1 Cotangent bundles Our convention for Liouville forms on the cotangent bundle T^*L of a smooth manifold L will be to use the form $(T^*L,\lambda_{\text{can}})$ with $\lambda_{\text{can}}=p_i\,dq_i$ in

a local coordinate system (q_i) on L. Provided such coordinates on L, we use (p_i, q_i) as local coordinates on T^*L, so that $d\lambda_{\text{can}}$ is symplectic with respect to the induced orientation.

2.3.2 Contactizations Provided an exact symplectic manifold (W, β), we have a contact form $dz+\beta$ on $I\times W$. We will refer to the contact manifold $(I\times W, \ker(dz+\beta))$ and the pair $(\mathbb{R}\times W, dz+\beta)$ both as the *contactization of* (W, β).

It's easy to see that deformations of an exact symplectic manifold give rise to contactomorphic contactizations. For, if $H\in C^\infty(W, \mathbb{R})$, then the contactization of $(W, \beta+dH)$ is equivalent to the contactization of (W, β) by the transformation

$$(t, w)\mapsto (t+H, w).$$

We'll further analyze geometry of contactizations in Sections 10.1 and 11.1. The quintessential example of a contactization is the 1–*jet space* of a closed manifold, which is the contactization of its cotangent bundle.

2.3.3 Symplectizations Provided (M, ξ) and α as above, $(\mathbb{R}\times M, e^t\alpha)$ is an exact symplectic manifold, called the *symplectization* of the pair (M, α). By considering diffeomorphisms of the form $(t, x)\mapsto (t+f(x), x)$ on $\mathbb{R}\times M$ for $f\in C^\infty(M, (0, \infty))$, it is clear that the symplectization is independent of the choice of α for ξ, up to symplectomorphism.

For any constant C, we will likewise refer to $\big([C, \infty)\times M, e^t\alpha\big)$ as the *positive half-infinite symplectization* and $\big((-\infty, C]\times M, e^t, \alpha\big)$ as the *negative half-infinite symplectization* of the pair (M, α). For constants $C<C'$, we will call $([C, C']\times M, e^t\alpha)$ a *finite symplectization* of the pair (M, α).

Here we can compute

$$X_\beta = \partial_t, \quad X_t = e^{-t}R.$$

Hence, there is a one-to-one correspondence between periodic orbits of R and 1–periodic orbits of X_t by associating to each γ in M the loop $(\log \mathcal{A}(\gamma), \gamma)$ in the symplectization.

2.3.4 Liouville cobordisms between closed and punctured contact manifolds Here we review some standard vocabulary regarding symplectic cobordisms, modified to deal with punctured contact manifolds. What are sometimes called "strong symplectic cobordisms" we will simply refer to as *symplectic cobordisms* for notational simplicity.

What are sometimes called "exact symplectic cobordisms" we will refer to as *Liouville cobordisms*. Our reasoning is that there exist symplectic cobordisms which are exact symplectic manifolds but which are not "exact symplectic cobordisms"; see Section 2.4 of [62]. See that paper or [52, Chapter 12] for a review of various notions of fillings and cobordisms with emphasis on low dimensions. We will only be concerned with Liouville cobordisms here.

Let (M, ξ) be a closed contact manifold of dimension $2n+1$ and $p \in M$ a point. We say that a contact form α for ξ defined on $M \setminus \{p\}$ is *standard at infinity* if there exists a ball B_p about $p \in M$, a positive constant C and a diffeomorphism

$$\Phi\colon (B_p \setminus \{p\}) \to (\mathbb{R}^{2n+1} \setminus \mathbb{D}^{2n+1}_C(0))$$

such that $\Phi^*(dz - y_i\, dx_i) = \alpha$ and $|\Phi(\gamma(t))| \to \infty$ for paths $\gamma(t)$ in $B_p \setminus \{p\}$ tending towards p.

A *Liouville cobordism between contact manifolds* (M^+, ξ^+) *and* (M^-, ξ^-) is a compact exact symplectic manifold (W, λ) for which

(1) $\partial W = M^+ - M^-$,

(2) the Liouville vector field X_λ points into W along M^- and out of W along M^+, and

(3) $\lambda|_{TM^\pm}$ is a contact form for $\xi^\pm$.

We call M^+ the *convex boundary of* (W, λ) and M^- the *concave boundary of* (W, λ). We may think of a Liouville domain as cobordism whose concave boundary is the empty set.

A *Liouville cobordism between punctured contact manifolds* $\widehat{(M^+, \xi^+)}$ *and* $\widehat{(M^-, \xi^-)}$ is defined analogously as in the case where the $(M^\pm, \xi^\pm)$ are closed. However, we require that there exists a region

$$I_C \times (\mathbb{R}^{2n+1} \setminus \mathbb{D}^{2n+1}_C(0)) \subset W, \quad \{\pm C\} \times (\mathbb{R}^{2n+1} \setminus \mathbb{D}^{2n+1}_C(0)) \subset M^\pm$$

along which $\lambda = e^t(dz - y_i\, dx_i)$ and such that the $t = \pm C$ slices provide standard at infinity neighborhoods of the punctures of the $M^\pm$.

We won't bother to specify that a Liouville cobordism is between closed or punctured contact manifolds, as it should be clear from the context. In either case, we may define the *completion* of a Liouville cobordism to be the noncompact exact symplectic manifold obtained from a Liouville cobordism by appending a positive half-infinite

symplectization to a collar of its convex boundary and a negative half-infinite symplectization to a collar of its concave boundary. We denote the completion of such a cobordism (W, λ) by $(\overline{W}, \overline{\lambda})$.

2.4 Remarks on SL(2, ℝ)

We briefly review some properties of $\mathrm{SL}(2,\mathbb{R})$ which will be useful for analyzing Reeb dynamics on contact 3–manifolds. By definition, $\mathrm{SL}(2,\mathbb{R})$ coincides with $\mathrm{Symp}(2,\mathbb{R})$ — the space of matrices preserving the standard symplectic form $dx \wedge dy$.

An element $A \in \mathrm{SL}(2,\mathbb{R})$ has characteristic polynomial

$$\det(A - \lambda\,\mathrm{Id}) = \lambda^2 - \mathrm{tr}(A)\lambda + 1, \tag{1}$$

using which eigenvalues of A can be found using the quadratic formula. The *nondegenerate* elements are those for which 1 is not an eigenvalue. A nondegenerate element A falls into one of two classes:

(1) A is called *elliptic* if its eigenvalues lie on the unit circle or, equivalently, $|\mathrm{tr}(A)| < 2$.

(2) A is called *hyperbolic* if its eigenvalues are elements of $\mathbb{R}$ or, equivalently, $|\mathrm{tr}(A)| > 2$.

Hyperbolic elements are further classified as *positive* (*resp. negative*) *hyperbolic* if the eigenvalues are positive (resp. negative) real numbers. The classification of $A \in \mathrm{SL}(2,\mathbb{R})$ as elliptic, positive hyperbolic or negative hyperbolic depends only on the conjugacy class of A.

2.5 Conley–Zehnder indices of Reeb orbits in contact 3–manifolds

Throughout the remaining subsections covering Reeb dynamics and contact homology, we follow the expositions [6] of Bourgeois (which covers all dimensions) and [40, Section 3.2] of Hutchings (which specifically focuses on the 3–manifolds). Let γ be a closed Reeb orbit of a contact manifold (M, ξ) equipped with a contact form α for ξ, whose Reeb vector field will be denoted by R. We assume γ is embedded and comes with a parametrization $\gamma(t)$; write γ^k for its k–fold iterate with $k > 0$.

As the Reeb flow preserves ξ, the Poincaré return map for time $t = \mathcal{A}(\gamma)$ sends $\xi|_{\gamma(0)}$ to itself and — provided a symplectic basis of $(\xi|_{\gamma(0)}, d\alpha)$ — determines a matrix $\mathrm{Ret}_\gamma \in \mathrm{SL}(2,\mathbb{R})$. The orbit γ will be called nondegenerate, elliptic, positive (negative)

hyperbolic if the matrix Ret_γ has the associated property. We say that the contact form α is *nondegenerate* if all of its Reeb orbits are nondegenerate.[8]

Remark 2.1 Having a nondegenerate contact form for which all closed orbits are hyperbolic — as is the case with the contact forms α_ϵ of Theorem 1.1 — is generally desirable as branched covers of trivial cylinders over elliptic orbits can have negative index [41, Section 1]. Likewise, in ECH chain complexes only simple covers of hyperbolic orbits are considered, whereas multiple covers of elliptic orbits cannot be avoided [40]. See also [3; 56], where analysis of holomorphic maps is simplified by considering only hyperbolic orbits.

Suppose that γ is a nondegenerate orbit equipped with a framing $s \in \Gamma_{\neq 0}(\xi|_\gamma)$. By extending s to a symplectic trivialization of the normal bundle $(\xi|_\gamma, d\alpha)$ to γ, we can write the restriction of the linearized flow to $\xi|_\gamma$ as a path $\phi = \phi(t)$ in $\mathrm{SL}(2,\mathbb{R})$. Then we define the *Conley–Zehnder index of the orbit γ with framing s*, denoted by $\mathrm{CZ}_s(\gamma)$, to be the Conley–Zehnder index $\mathrm{CZ}(\phi)$ of the path ϕ.

If γ is hyperbolic, ϕ rotates the eigenspaces of Ret_γ by an angle πn for some $n \in \mathbb{Z}$, in which case

$$\mathrm{CZ}_s(\gamma^k) = kn.$$

Negative hyperbolic orbits have n odd and positive hyperbolic orbits have n even. If γ is elliptic, ϕ rotates the eigenspaces of Ret_γ by some angle $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$ in which case the Conley–Zehnder index is

$$\mathrm{CZ}_s(\gamma^k) = 2\left\lfloor \frac{k\theta}{2\pi} \right\rfloor + 1.$$

Note that CZ_s depends only on the isotopy class of the framing s. If we write $s+n$ for a framing whose isotopy class is given by twisting s by n meridians, then

$$\mathrm{CZ}_{s+n}(\gamma^k) = \mathrm{CZ}_s(\gamma^k) - 2nk. \tag{2}$$

An orbit γ^k is *bad* if the parity of its Conley–Zehnder index disagrees with that of the underlying embedded orbit γ. Orbits which are not bad are *good*. Hence (when $\dim(M) = 3$), the only bad orbits are even covers of negative hyperbolic orbits. See [23, Remarks 1.9.2 and 1.9.6].

[8] In practice, one is typically interested in studying sequences of contact forms α_n with "nice" limiting behavior, namely there exists a sequence $C_n \to \infty$ such that the orbits of α_n of action $\leq C_n$ are nondegenerate. See for example [4; 3; 5; 18]. We take a similar approach in this article.

Note that, as $\mathrm{CZ}_s(\gamma)$ mod 2 is independent of s, so is the property that an orbit is good or bad. We write $\mathrm{CZ}_2(\gamma) \in \mathbb{Z}/2\mathbb{Z}$ for the index modulo 2 which satisfies

$$\operatorname{sgn} \circ \det(\mathrm{Ret}_\gamma - \mathrm{Id}) = (-1)^{\mathrm{CZ}_2 + 1}. \tag{3}$$

The following method of computing the Conley–Zehnder index of a path $\phi(t)$ for $t \in [0,1]$ of symplectic matrices is due to Robbin and Salamon [55]. For a path $\phi\colon [0,1] \to \mathrm{SL}(2,\mathbb{R})$, a point $t \in [0,1]$ is *crossing* if 1 is an eigenvalue of $\phi(t)$. Writing

$$\frac{\partial \phi}{\partial t}(t) = J_0 S(t) \phi(t) \tag{4}$$

for symmetric matrices $S(t)$, we say that a crossing t is *regular* if the quadratic form $\Gamma(t)$ defined as the restriction of $S(t)$ to $\ker(\phi(t) - \mathrm{Id})$ is nondegenerate. For a path ϕ with only regular crossings, we can compute $\mathrm{CZ}(\phi)$ as

$$\mathrm{CZ}(\phi) = \tfrac{1}{2} \operatorname{sgn}(\Gamma(0)) + \sum_{t>0 \text{ crossing}} \operatorname{sgn}(\Gamma(t)). \tag{5}$$

Also of utility for computation is the *loop property* of CZ, which states that, given $k \in \mathbb{Z}$ and a nondegenerate path ϕ, the path $\widetilde{\phi}(t) = e^{i2\pi k t}\phi(t)$ has

$$\mathrm{CZ}(\widetilde{\phi}) = 2k + \mathrm{CZ}(\phi). \tag{6}$$

2.6 Holomorphic curves in symplectizations and the index formula

Now suppose that α is a nondegenerate contact form for some contact 3–manifold (M,ξ) and that J is an almost-complex structure which is *adapted to the symplectization* $(\mathbb{R}\times M, e^t\alpha)$. That is,

(1) J is compatible with $d(e^t\alpha)$,

(2) it is t–invariant and preserves ξ, and

(3) $J\partial_t = R$.

Let

$$\gamma^+ = \{\gamma_1^+, \dots, \gamma_{m^+}^+\} \quad \text{and} \quad \gamma^- = \{\gamma_1^-, \dots, \gamma_{m^-}^-\}$$

be collections of Reeb orbits with γ^+ nonempty and let (Σ, j) be a Riemann surface with marked points $\{p_1^+, \dots, p_{m^+}^+, p_1^-, \dots, p_{m^-}^-\}$. We write Σ' for Σ with its marked

points removed. We say that $(t, U)\colon \Sigma' \to \mathbb{R} \times M$ is *holomorphic* if

$$\overline{\partial}(t, U) = \tfrac{1}{2}(T(t, U) + J \circ T(t, U) \circ j)$$

vanishes. If we wish to specify J and j, we'll say that the map is (J, j)–*holomorphic*. This is equivalent to the conditions

$$dt = U^*\alpha \circ j, \quad J\pi_\alpha \circ TU = \pi_\alpha \circ TU \circ j, \tag{7}$$

where $\pi_\alpha\colon TM \to \xi$ is the projection $V \mapsto V - \alpha(V)R$. We provide a few simple examples.

Example 2.2 (trivial strips, planes and cylinders) Provided a map $\gamma\colon I \to M$ parametrizing a Reeb trajectory for a connected 1–manifold I, $\mathbb{R} \times \operatorname{im}(\gamma) \subset \mathbb{R} \times M$ is an immersion with J–complex tangent planes. Some examples of particular interest:

(1) If I is compact with nonempty boundary parametrizing a chord of R with endpoints on a Legendrian submanifold Λ, we'll call $\mathbb{R} \times \operatorname{im}(\gamma)$ a *trivial strip*.

(2) If $I = \mathbb{R}$ and the map γ is an embedding, we'll say that $\mathbb{R} \times \operatorname{im}(\gamma)$ is a *trivial plane*.

(3) If $I = S^1_a$ parametrizing a Reeb orbit of action a, then we'll say that $\mathbb{R} \times \operatorname{im}(\gamma)$ is a *trivial cylinder*.

Given a holomorphic map $(t, U)\colon (\Sigma, j) \to (\mathbb{R} \times M, J)$, we say that the puncture p_i^+ is *positively asymptotic* to the orbit γ_i^+ if there exists a neighborhood $[C, \infty) \times S^1$ of p_i^+ in Σ with coordinates r and θ for which j is the standard cylindrical complex structure such that $t(r, \theta) \to \infty$ and $U(r, \theta)$ tends to a parametrization of γ_i^+ as $r \to \infty$. Likewise, we say that the puncture p_i^- is *negatively asymptotic* to the orbit γ_i^- if $t(r, \theta) \to -\infty$ and $U(r, \theta)$ tends to a parametrization of $-\gamma_i^+$ as $r \to \infty$. Allowing j and the location of the marked points to vary and then modding out by reparametrization in the domain, we write $\mathcal{M}_{(t,U)}$ for the *moduli space of holomorphic maps* asymptotic to the $\gamma^\pm$ containing the map (t, U).

The *index* of a holomorphic map as above is defined by

$$\operatorname{ind}((t, U)) = \operatorname{CZ}_s(\gamma^+) - \operatorname{CZ}_s(\gamma^-) - \chi(\Sigma') + 2c_s(U) \in \mathbb{Z}, \tag{8}$$

where

$$\operatorname{CZ}_s(\gamma^\pm) = \sum_{i=1}^{m^\pm} \operatorname{CZ}_s(\gamma_i^\pm).$$

The *relative first Chern class* $c_s(U)$ is the signed count of zeros of $U^*\xi$ over Σ' using a section which coincides with s near the punctures. Note that ind is independent of s.

In ideal geometric settings, $\mathcal{M}_{(t,U)}$ is a manifold near the point (t,U) of dimension $\mathrm{ind}((t,U))$.

Remark 2.3 Here we are disregarding asymptotic markers for orbits which are required for a rigorous functional-analytic setup for moduli spaces and curve counts. We refer to [4; 54] for details.

The *energy* of a holomorphic curve is defined by

$$\mathcal{E}(t,U)=\int_{\Sigma'} d\alpha=\sum_1^{m^+}\mathcal{A}(\gamma_i^+)-\sum_1^{m^-}\mathcal{A}(\gamma_i^-).$$

The energy is nonnegative and is zero if and only if (t,U) is a branched cover of a trivial cylinder. Energies of curves will be presumed finite unless otherwise explicitly stated.

2.7 Contact homology and its variants

We now give a brief overview of contact homology and symplectic field theory. As in previous subsections, we focus specifically on the case of contact 3–manifolds.

For each closed Reeb orbit γ with framing s, we define its degree $|\gamma|_s=\mathrm{CZ}_s(\gamma)-1\in\mathbb{Z}$. This degree modulo 2 will be denoted by $|\gamma|$. We write $\mathrm{CC}(\alpha)$ for the supercommutative algebra with unit 1 generated by the good Reeb orbits of α over $\mathbb{Q}$. Here supercommutativity means $\gamma_1\gamma_2=(-1)^{|\gamma_1||\gamma_2|}\gamma_2\gamma_1$. We note that $\mathrm{CC}(\alpha)$ has two canonical gradings:

(1) The *degree grading* given by $|\gamma_1\cdots\gamma_n|:=\sum_1^n|\gamma_i|\in\mathbb{Z}/2\mathbb{Z}$.

(2) The H_1 *grading* given by $[\gamma_1\cdots\gamma_n]:=\sum_1^n[\gamma_i]\in H_1(M)$.

For $i\in\mathbb{Z}/2\mathbb{Z}$ and $h\in H_1(M)$, we will use the notation $\mathrm{CC}_{i,h}$ to denote the relevant graded $\mathbb{Q}$–subspaces. The contact homology differential

$$\partial_{\mathrm{CH}}\colon \mathrm{CC}_{i,h}\to\mathrm{CC}_{i-1,h}$$

is defined by counting $\mathrm{ind}=1$ (possibly perturbed) solutions to $\overline{\partial}(t,U)=0$ with one positive puncture, any number of negative punctures and genus 0. For such curves (t,U) positively asymptotic to some γ^+ and negatively asymptotic to γ_j^- simultaneously framed with some choice of s, equation (8) becomes

$$\mathrm{ind}((t,U))=|\gamma^+|_s-\sum_j|\gamma_j^-|_s+2c_s(U). \tag{9}$$

Each such solution contributes a term to $\partial\gamma^+$ of the form $m(\gamma^+;\gamma_i^-)\gamma_1^-\cdots\gamma_n^-$ with $m(\gamma^+;\gamma_i^-)\in\mathbb{Q}$. If there are no negative punctures, we get a term of the form $m(\gamma^+)1$

and we set $\partial_{\mathrm{CH}} 1 = 0$. The differential is then extended to products of orbits using the graded Leibniz rule

$$\partial_{\mathrm{CH}}(\gamma_1\gamma_2) = (\partial\gamma_1)\gamma_2 + (-1)^{|\gamma_1|}\gamma_1(\partial\gamma_2)$$

and to sums of products linearly.

Definition 2.4 The resulting differential graded algebra $\ker(\partial_{\mathrm{CH}})/\mathrm{im}(\partial_{\mathrm{CH}})$ is defined to be the *contact homology of* (M,ξ), denoted by $\mathrm{CH}(M,\xi)$. As in the case of $\mathrm{CC}(\alpha)$, $\mathrm{CH}(M,\xi)$ also has degree and H_1 gradings. We write $\mathrm{CH}_{i,h}(M,\xi)$ for the subspace of $\mathrm{CH}(M,\xi)$ with degree i and H_1 grading h.

This theory, first proposed by Eliashberg, Givental and Hofer [23], has been proven to be rigorously defined and independent of all choice involved by Bao and Honda [4] and Pardon [54]. We defer to these citations for the specifics of how the coefficients $m(\gamma^+;\gamma_i^-) \in \mathbb{Q}$ are computed and details around any required perturbations of $\overline{\partial}$. For the purposes of this paper, it suffices to know that, for generic J adapted to the symplectization of a contact manifold,

(1) curves which are somewhere injective may be assumed regular,

(2) regularity for these curves may be achieved by perturbations of J in arbitrarily small neighborhoods of the closed orbits of R, and

(3) assuming such regularity, the moduli space of holomorphic planes positively asymptotic to a closed, embedded orbit will be a manifold (rather than an orbifold), so that such ind $= 1$ planes can be counted over $\mathbb{Z}$.

Additional algebraic structures — which require more sophisticated underlying chain complexes — may be constructed as follows:

(1) By counting ind $= 1$, genus 0 holomorphic curves with arbitrary numbers of positive and negative punctures via a differential ∂_{RSFT}, the *rational SFT algebra* (*RSFT*) may be defined.

(2) By counting ind $= 1$ holomorphic curves with arbitrary genus and numbers of positive and negative punctures via a differential ∂_{SFT}, the *SFT algebra* (*SFT*) may be defined.

See [23] for a more complete picture or the exposition [64, Lecture 12] for further details regarding these invariants.[9] For other *RSFT*-like algebraic structures associated

[9] At the time of writing, rigorous definitions of *RSFT* and *SFT* are under construction using a variety of frameworks. We refer to [4; 48; 54] for accounts of the current state of the development of SFT.

to counts of rational curves with multiple positive punctures, see [48], which constructs such invariants and provides an overview of recent additions to the literature.

The lecture notes [6] and Section 1.8 of [54] also contain a rather exhaustive list of additional structures such as grading refinements and twisted coefficient systems for contact homology. We won't address such additional structures in this article, except in the following simple situations:

Proposition 2.5 (canonical $\mathbb{Z}$ gradings) *The $0 \in H_1(M)$ part of $\mathrm{CH}(M,\xi)$ is a subalgebra of CH. Suppose that $\Gamma_{\neq 0}(\xi)$ is nonempty (equivalently, $c_1(\xi) = 0$).*

(1) *The $\mathbb{Z}$–valued degree gradings $|\cdot|_s$ on $\mathrm{CC}(\alpha)$ determine $\mathbb{Z}$–valued gradings on $\mathrm{CH}_{*,0}(M,\xi)$ and are independent of the choice of $s \in \Gamma_{\neq 0}(\xi)$.*

(2) *Moreover, if $H^1(M) = H_2(M) = 0$, then the $\mathbb{Z}$–valued degree gradings $|\cdot|_s$ on $\mathrm{CC}(\alpha)$ determine $\mathbb{Z}$–valued gradings $\mathrm{CH}(M,\xi)$ which are independent of the choice of $s \in \Gamma_{\neq 0}(\xi)$.*

We get canonical $\mathbb{Z}$ gradings on CH when we have a nondegenerate Reeb vector field with only homologically trivial Reeb orbits or when M is a 3–dimensional $\mathbb{Q}$–homology sphere.

Proof The fact that $\mathrm{CH}_{*,0}(M,\xi)$ is a subalgebra of $\mathrm{CH}(M,\xi)$ is clear from the fact that ∂_{CH} preserves H_1 and that $\mathrm{CC}_{*,0}$ is closed under products.

Provided $s \in \Gamma_{\neq 0}(\xi)$, extend s to a trivialization $\xi \to \mathbb{C}$. For our extension, we may use Js for an almost-complex structure J on ξ. In this way, we see that any other nonvanishing section s' defines a map $M \to \mathbb{C}^* \simeq S^1$ and recall that homotopy classes of maps to S^1 are in bijective correspondence with elements of H^1 [33, Theorem 4.57]. Write $[s'-s] \in H^1$ for the cohomological element provided by this correspondence. If γ is a closed orbit of some α for (M,ξ), then $[s'-s]\cdot[\gamma] \in \mathbb{Z}$ equals the difference in meridians between the framings of $\xi|_\gamma$ determined by s and s'. Then $\mathrm{CZ}_s(\gamma) - \mathrm{CZ}_{s'}(\gamma)$ will be determined by this framing difference according to (2).

If $[\gamma] = 0 \in H_1$, then the above tells us $\mathrm{CZ}_s(\gamma) = \mathrm{CZ}_{s'}(\gamma)$, so that the gradings $|\gamma|_s$ on $\mathrm{CC}_{*,0}$ are independent of choice of nonvanishing s. If $H^1(M) = 0$ then s' is necessarily homotopic to s, so that all of the gradings $|\cdot|_s$ are equivalent on $\mathrm{CC}_{*,*}$. As s is nonvanishing, the c_s term in (9) is always 0, meaning that ∂ always lowers the degree $|\cdot|_s$ by exactly 1 and so the $\mathbb{Z}$–valued degree gradings on CC determines a $\mathbb{Z}$ grading on homology.

To complete the proof, we must show that the $\mathbb{Z}$ grading is independent of the choices used to compute CH. Proofs of invariance of CH (see as they appear in [4; 54]) are obtained by considering the symplectization of (M,α) — for some α — equipped with almost-complex structures which are adapted to α at the negative end $(-\infty,-C]\times M$ of the symplectization and adapted to $H\alpha$ at the positive end $[C,\infty)\times M$ for some $C>0$ and $H\in\mathcal{C}^\infty(M,(0,\infty))$. In such a scenario, $T(\mathbb{R}\times M)$ can be split as the direct sum $\operatorname{span}_{\mathbb{R}}(\partial_t, J\partial_t)\oplus\xi$, and we can extend s over $\mathbb{R}\times M$ in the obvious way to frame Reeb orbits at both ends of $\mathbb{R}\times M$. The isomorphism between the contact homologies of the ends of the cobordism is defined by counting $\operatorname{ind}=0$ holomorphic curves in $\mathbb{R}\times M$, which, by the index formula of (9), must preserve the $\mathbb{Z}$ grading. □

The variant of contact homology which will be of the most interest to us is the *hat version*, denoted by $\widehat{\mathrm{CH}}(M,\xi)$ and defined in [14]. To define this theory for (M,ξ), we can equip $\widehat{M}$ with a standard-at-infinity α for $\hat{\xi}=\xi|_{\widehat{M}}$, choose an appropriately convex J on $\hat{\xi}$, and compute CH as above. We describe such J for the $(\mathbb{R}^3_{\Lambda^\pm},\xi_{\Lambda^\pm})$ in Section 11.2.[10]

The following theorem summarizes some properties of $\widehat{\mathrm{CH}}$ laid out in the introduction of [14] (coupled with some well-known results):

Theorem 2.6 *The invariant $\widehat{\mathrm{CH}}(M,\xi)$ satisfies the following properties*:

(1) *For the standard contact 3–sphere (S^3,ξ_{std}), $\widehat{\mathrm{CH}}(S^3,\xi_{\mathrm{std}})=\mathbb{Q}1$.*

(2) *If (M,ξ) is overtwisted, then $\widehat{\mathrm{CH}}(M,\xi)=0$.*

(3) *For a contact-connected sum $(M_1,\xi_1)\#(M_2,\xi_2)$,*

$$\widehat{\mathrm{CH}}((M_1,\xi_1)\#(M_2,\xi_2))\simeq\widehat{\mathrm{CH}}(M_1,\xi_1)\otimes\widehat{\mathrm{CH}}(M_2,\xi_2).$$

(4) *The inclusion $\widehat{(M,\xi)}\to(M,\xi)$ induces an algebra homomorphism*

$$\widehat{\mathrm{CH}}(M,\xi)\to\mathrm{CH}(M,\xi).$$

Consequently, $\mathrm{CH}(M,\xi)\neq0$ implies $\widehat{\mathrm{CH}}(M,\xi)\neq0$.

(5) *A Liouville cobordism (W,λ) with convex boundary $\widehat{(M^+,\xi^+)}$ and concave boundary $\widehat{(M^-,\xi^-)}$ determines an algebra homomorphism*

$$\Phi_{(W,\lambda)}\colon\widehat{\mathrm{CH}}(M^+,\xi^+)\to\widehat{\mathrm{CH}}(M^-,\xi^-).$$

(6) *Consequently, if (M,ξ) admits a Liouville filling, then both $\mathrm{CH}(M,\xi)$ and $\widehat{\mathrm{CH}}(M,\xi)$ are nonzero.*

[10] In [14], less restrictive conditions are placed on α and J to define $\widehat{\mathrm{CH}}$ within the framework of the more general sutured contact homology. We choose more restrictive conditions so as to simplify our exposition and avoid general discussion of sutured contact manifolds and their completions as well as to simplify J–convexity arguments.

Item (5), which will refer to as *Liouville functoriality*, does not explicitly appear in the literature for $\widehat{\mathrm{CH}}$, though it follows from a simple combination of existing arguments and constructions. Liouville functoriality is established for closed contact manifolds in [4; 54]. To extend the results to punctured contact manifolds, one needs to establish SFT compactness [8] of (possibly perturbed) moduli spaces of holomorphic curves positively asymptotic to closed orbits of a standard-at-infinity contact form on $\widehat{(M^+,\xi^+)}$ and negatively asymptotic to closed orbits of a standard-at-infinity form on $\widehat{(M^-,\xi^-)}$ in the completion of (W,λ). To obtain compactness, we may restrict to almost-complex structures J which are t–invariant over the neighborhood of the puncture of the $M^{\pm}$ to ensure that sequences of curves cannot escape the completed cobordism through the horizontal boundary of the symplectization of the puncture. Our definition of Liouville cobordism between punctured contact manifolds and the J of Section 11.2 ensure that these desired hypotheses are in place. Perturbations of $\overline{\partial}$ required to achieve transversality for the counting of curves and gluing of multilevel SFT buildings can be implemented in arbitrarily small neighborhoods of closed Reeb orbits [4, Section 5], so that such perturbations do not interfere with convexity. In this way, the compactness results of [14, Section 5] carry over without issue.

For Theorem 2.6(6), Liouville functoriality tells us that a Liouville filling of a closed contact manifold induces an algebra homomorphism from $\mathrm{CH}(M,\xi)$ to $\mathbb{Q}$ (also known as an augmentation). Therefore, $\mathrm{CH}(M,\xi)\neq 0$, implying $\widehat{\mathrm{CH}}(M,\xi)\neq 0$ by (4).

2.7.1 Relative contact homology We now briefly review SFT invariants of a Legendrian link $\Lambda\subset(M,\xi)$. For the case $(M,\xi)=(\mathbb{R}^3,\xi_{\mathrm{std}})$, we recommend [26], with the general theory laid out in [23, Section 2.8].

Provided a Legendrian link $\Lambda\subset(M,\xi)$ and a contact form α for ξ, consider the space of chords of R which begin and end on Λ. A chord $r=r(t)$ is *nondegenerate* if it satisfies the transversality condition

$$\mathrm{Flow}_R^{\mathcal{A}(r)}(T_{r(0)}\Lambda)\pitchfork T_{r(\mathcal{A}(r))}\Lambda\subset\xi_{r(\mathcal{A}(r))}.$$

We then say that the pair (α,Λ) is *nondegenerate* if all chords for the pair and all closed orbits of R are nondegenerate. Provided nondegeneracy, we consider a $\mathbb{Z}/2\mathbb{Z}$–graded supercommutative algebra $\mathrm{CC}(\alpha,\Lambda)$ generated by the chords of Λ and the good closed orbits of R.[11] As in the nonrelative case, $\mathrm{CC}(\alpha,\Lambda)$ comes with an additional homological grading, given by the relative homology classes of chords and orbits in $H_*(M,\Lambda)$.

[11] We are skipping definition of the gradings of chords in the general case. See [26; 50] for gradings in the case of Legendrians in $(\mathbb{R}^3,\xi_{\mathrm{std}})$.

We may then define a differential

$$\partial_{\mathrm{LCH}}\colon \mathrm{CC}_{i,h}(\alpha,\Lambda)\to \mathrm{CC}_{i-1,h}(\alpha,\Lambda)$$

for $i\in\mathbb{Z}/2\mathbb{Z}$ and $h\in H_1(M,\Lambda)$ as follows: for a chord r, ∂_{LCH} counts $\mathrm{ind}=1$ holomorphic disks in the symplectization of (M,ξ) with

(1) a single boundary puncture positively asymptotic to r,

(2) any number m of boundary punctures negatively asymptotic to chords r_i^- of Λ,

(3) $\partial\mathbb{D}$ with its punctures removed mapped to the Lagrangian cylinder over Λ, and

(4) n interior punctures negatively asymptotic to closed orbits γ_j^-

Each such disk contributes a term of the form $m(r^+;r^-,\gamma^-)r_1^-\cdots r_m^-\gamma_1^-\cdots\gamma_n^-$ to $\partial_{\mathrm{LCH}}r^+$. For a closed orbit γ^+, the differential $\partial_{\mathrm{LCH}}\gamma^+$ coincides with the contact homology differential of γ^+. The differential is then extended to products and sums of products using the Leibniz rule and linearity as in the case of nonrelative contact homology.

Definition 2.7 The resulting differential graded algebra $\ker(\partial_{\mathrm{LCH}})/\mathrm{im}(\partial_{\mathrm{LCH}})$ is defined to be the *Legendrian contact homology* of the triple (M,ξ,Λ), denoted by $\mathrm{LCH}(M,\xi,\Lambda)$. As in the case of CC_Λ, LCH has degree and relative H_1 gradings.

The computation $\partial_{\mathrm{LCH}}^2=0$ and proof of invariance for links in $(\mathbb{R}^3,\xi_{\mathrm{std}})$ — in which case there are no closed Reeb orbits — is carried out in [19], with a proof of the general case sketched in [23]. In the case $\Lambda\subset(\mathbb{R}^3,\xi_{\mathrm{std}})$, a combinatorial version of LCH — originally due to Chekanov [10] — may be computed by counting immersions of disks into the xy–plane with boundary mapped to the Lagrangian projection of Λ, in which case $\partial_{\mathrm{LCH}}^2=0$ may be proved diagrammatically. Additional algebraic structures may derived from the triple (M,ξ,Λ) by considering disks with multiple positive punctures as in [17; 50]. Again, we point to [26] for further references regarding proofs that the combinatorially and analytically defined invariants coincide for $(\mathbb{R}^3,\xi_{\mathrm{std}})$ as well as extensions and generalizations of LCH in both algebraic and geometric directions.

2.8 Legendrian knots and links in $(\mathbb{R}^3,\xi_{\mathrm{std}})$

Legendrian knots and links will be denoted by Λ with sub- and superscripts. Throughout this article, we assume that each component of Λ is equipped with a predetermined orientation. For a Legendrian link Λ in a contact manifold (M,ξ) with contact form α and Reeb vector field R,

$$\mathrm{Flow}_R^\delta(\Lambda)$$

for $\delta > 0$ arbitrarily small will be called the *push-off of* Λ. The Legendrian isotopy class of the pair $(\Lambda, \mathrm{Flow}_R^{\delta}(\Lambda))$ is independent of R and δ. We write λ_ξ for the Legendrian isotopy class of the push-off.

For a Legendrian link Λ in $(\mathbb{R}^3, \xi_{\mathrm{std}})$, the front and Lagrangian projections will be denoted by π_{xz} and π_{xy}, respectively. We will use front projections as our default starting point for analysis of Λ, from which we will transition to the Lagrangian projection — see Section 4.4.

Assuming that Λ has a single connected component, its *classical invariants* are

(1) the Thurston–Bennequin number $\mathrm{tb}(\Lambda)$;

(2) the rotation number, $\mathrm{rot}(\Lambda)$, which depends on an orientation of Λ; and

(3) the smooth topological knot underlying Λ.

In the Lagrangian projection, we may compute $\mathrm{tb}(\Lambda)$ as the writhe and $\mathrm{rot}(\Lambda)$ as the winding number. Geometrically, the Thurston–Bennequin number is defined as the linking number

$$\mathrm{tb}(\Lambda) = \mathrm{lk}(\Lambda, \lambda_\xi),$$

whereas rot is defined as the degree of the Gauss map of $T\Lambda$ in ξ_{std} with respect to a nowhere-vanishing trivialization. If we replace $\mathbb{R}^3$ with any contact $\mathbb{Q}$–homology sphere, then tb is defined for null-homologous Legendrian knots and rot is defined for all Legendrian knots using the framings of Proposition 2.5. See also Definition 6.3.

Classical invariants of a Legendrian knot $\Lambda \subset (\mathbb{R}^3, \xi_{\mathrm{std}})$ are constrained by the *slice-Bennequin bound* of [57]:

$$\tfrac{1}{2}\big(\mathrm{tb}(\Lambda) + |\mathrm{rot}(\Lambda)| + 1\big) \leq g_s(\Lambda) \leq g(\Lambda). \tag{10}$$

Here $g_s(\Lambda)$ is the smooth slice genus of the topological knot underlying Λ and $g(\Lambda)$ is the Seifert genus. See [24, Section 3] for an overview of related results.

Let Λ be a Legendrian knot in a contact manifold (M, ξ). Take a cube $I^3_\epsilon \subset M$ with $\epsilon > 0$ and coordinates x, y and z such that

$$\xi = \ker(\alpha_{\mathrm{std}}), \quad \Lambda \cap I^3_\epsilon = \{y = z = 0\}$$

and ∂_x orients Λ. Then Λ is locally described by Figure 2, left. The *positive and negative stabilizations* of Λ, denoted by $S_+(\Lambda)$ and $S_-(\Lambda)$, are defined as the Legendrian knots determined by modifying Λ in the front projection of I^3_ϵ as described in Figure 2, right. We say that a Legendrian knot Λ is *stabilized* if it is a positive or negative stabilization of some $\Lambda' \subset (M, \xi)$.

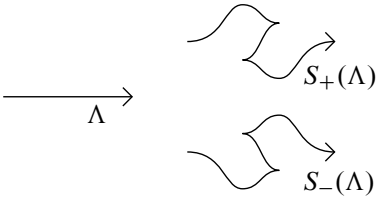


Figure 2: Positive and negative stabilizations of Λ described in the front projection as in [24, Figure 19].

2.9 Contact surgery

Contact surgery — first defined in [16] — provides a way of performing Dehn surgery on a Legendrian link $\Lambda^{\pm}$ so that the surgered manifold carries a contact structure uniquely determined by $\Lambda^{\pm}$ and (M, ξ). We recommend Ozbagci and Stipsicz [52] as a general reference.

We take the coefficients of the components of the sublinks Λ^+ (resp. Λ^-) to be $+1$ (resp. -1). Intuitively speaking, contact -1 (resp. $+1$) surgery removes a neighborhood of a Legendrian knot of the form $I_\epsilon \times I_\epsilon \times S^1$ — the first coordinate being directed by ∂_z — and then glues it back in using a positive (resp. negative) Dehn twist along $\{\epsilon\} \times I_\epsilon \times S^1$. The construction may be formalized using the gluing theory of convex surfaces. A rigorous account of the construction will be carried out in Section 4. For $k \in \mathbb{Z} \setminus \{0\}$ one may analogously perform *contact* $1/k$ *surgery* on a Legendrian knot Λ by applying $-k$ Dehn twists as above. We will take as definition that contact $1/k$ surgery for $k \neq 0$ is given by performing contact $\operatorname{sgn}(k)$ surgery on $|k|$ parallel push-offs of Λ.

We write

$$\Lambda = \Lambda^+ \cup \Lambda^0 \cup \Lambda^- \subset (\mathbb{R}^3, \xi_{\text{std}})$$

to specify a Legendrian link Λ^0 sitting inside of the contact manifold $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$. Since the neighborhoods of the components of Λ defining surgery many be chosen to be disjoint from Λ^0, we may consider it to be a Legendrian link in $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ post surgery. The superscript 0 on Λ^0 may be thought of as indicating a trivial $\frac{1}{0} = \infty$ surgery in the usual notation of Kirby calculus.

In Section 10 we will review how contact surgeries may be viewed as the result of handle attachments. We refer the reader to [52] for a review in the low-dimensional case and to [11] for the general case.

Theorem 2.8 *We summarize some known results about contact surgery relevant to this paper*:

(1) *The contact manifold obtained by contact $+1$ surgery on the Legendrian unknot with $\mathrm{tb} = -1$ yields the standard fillable contact structure ξ_{std} on $S^1 \times S^2$.*

(2) *Applying contact -1 surgery on a Legendrian knot in (M, ξ) produces the same contact manifold as is obtained by attaching a Weinstein handle to the convex boundary of the symplectization of (M, ξ).*

(3) *Then performing ± 1 surgery on a Legendrian knot $\Lambda \subset (M, \xi)$ followed by ∓ 1 surgery on a push-off λ_ξ leaves (M, ξ) unchanged.*

(4) *A contact 3–manifold is overtwisted if and only if it can be described as the result of a contact $+1$ surgery along a stabilized Legendrian knot Λ in some (M, ξ).*[12]

(5) *If $\Lambda \subset (\mathbb{R}^3, \xi_{\mathrm{std}})$ satisfies $\mathrm{tb}(\Lambda) = 2g_s(\Lambda) - 1$, then contact $1/k$ surgery on Λ produces a tight contact manifold [43] for any $k \in \mathbb{Z}$.*

(6) *For a Legendrian knot $\Lambda \subset (\mathbb{R}^3, \xi_{\mathrm{std}})$ and an integer $k > 0$, contact $1/k$ surgery on Λ produces a symplectically fillable contact 3–manifold if and only if both $k = 1$ and Λ bounds a Lagrangian disk in the standard symplectic 4–disk [15].*

3 Notation and algebraic data associated to chords

In this section we describe notation and algebraic data associated to chords of Legendrian links which will be used throughout the remainder of the paper. We take $\Lambda \subset (\mathbb{R}^3, \xi_{\mathrm{std}})$ to be a nonempty link with sublinks Λ^+, Λ^- and Λ^0 — any of which may be empty. We write $\Lambda^\pm = \Lambda^+ \cup \Lambda^-$.

Assumptions 3.1 It is assumed throughout that Λ is *chord generic*, meaning that all chords of Λ are nondegenerate and that distinct chords are disjoint as subsets of $\mathbb{R}^3$.

3.1 Surgery coefficients and chords of Λ

It will be convenient to write $\Lambda = \cup \Lambda_i$ with the subscript i indexing the connected components of Λ. Using this notation, we use $c_i \in \{-1, 0, +1\}$ to indicate that $\Lambda_i \subset \Lambda^{c_i}$.

Denote by r_j the Reeb chords of Λ with the contact form $\alpha_{\mathrm{std}} = dz - y\,dx$, which are in one-to-one correspondence with the double points of the Lagrangian projection π_{xy}.

[12] One proof is obtained by proving the "if" statement using [51] and proving "only if" by following the proof of Theorem 12.3. Alternatively, one can apply a handle-slide to [2, Theorem 5.5(2)].

We write $\operatorname{sgn}_j \in \{\pm 1\}$ for the sign of the crossing of Λ in the Lagrangian projection associated with the chord r_j in accordance with the orientation of Λ.

We define l_j^- to be the index i of the Λ_i on which r_j begins and l_j^+ to be the index of the component of Λ on which r_j ends. The *tip* of a chord r_j is the point $q_j^+ \in \Lambda_{l_j^+}$ where the chord ends. The *tail* of r_j is the point $q_j^- \in \Lambda_{l_j^-}$ at which the chord r_j begins. We write the surgery coefficient of the components of Λ corresponding to $l_j^\pm$ as $c_j^\pm$. That is,

$$c_j^\pm = c_{l_j^\pm}.$$

3.2 Words of chords

An ordered pair of chords (r_{j_1}, r_{j_2}) is *composable* if $l_{j_1}^+ = l_{j_2}^-$. A *word of Reeb chords for* Λ is a formal product of chords $w = r_{j_1} \cdots r_{j_n}$ in which each pair $(r_{j_k}, r_{j_{k+1}})$ is composable for $k = 1, \dots, n-1$.

We say that a word of Reeb chords $r_{j_1} \cdots r_{j_n}$ is a *word of chords with boundary on* Λ^0 if r_{j_1} begins on Λ^0 and r_{j_n} ends on Λ^0 and all other endpoints of chords touch components of $\Lambda^+ \cup \Lambda^-$.

A *cyclic word of Reeb chords for* Λ, denoted by $r_{j_1} \cdots r_{j_n}$, is a word of Reeb chords for which (r_{j_n}, r_{j_1}) is composable. Cyclic permutations of cyclic words are considered to be equivalent:

$$r_{j_1} r_{j_2} \cdots r_{j_n} = r_{j_2} \cdots r_{j_n} r_{j_1}.$$

When speaking of cyclic words of Reeb chords on Λ, we will implicitly assume that it is a cyclic word of Reeb chords on $\Lambda^+ \cup \Lambda^-$.

The *word length* of a word w of Reeb chords is the number of individual chords it contains and will be denoted by $\operatorname{wl}(w)$. The actions of each r_j will be denoted by $\mathcal{A}_j$ and the *action* of a word is defined by

$$\mathcal{A}(r_{j_1} \cdots r_{j_n}) = \sum_{k=1}^{n} \mathcal{A}_{j_k}.$$

3.3 Capping paths

Provided a composable pair of chords (r_{j_1}, r_{j_2}), their *capping path* is the unique embedded, oriented segment of $\Lambda_{l_{j_1}^+} = \Lambda_{l_{j_2}^-}$, traveling in the direction of the orientation of Λ from the tip of r_{j_1} to the tail of r_{j_2}. The capping path will be denoted by η_{j_1, j_2}.

The analogously defined path, which travels opposite the orientation of $\Lambda_{l_{j_1}^+}$ will be denoted by $\overline{\eta}_{j_1,j_2}$ and called the *opposite capping path*. We will use ζ_{j_1,j_2} to denote one of either η_{j_1,j_2} or $\overline{\eta}_{j_1,j_2}$. By setting $\overline{\overline{\eta}}_{j_1,j_2} = \eta_{j_1,j_2}$, we can define $\overline{\zeta}_{j_1,j_2}$ in the obvious way.

3.3.1 Rotation angles and numbers Denote by G the Gauss map sending the unit tangent bundle of $\mathbb{R}^2$ to $S^1_{2\pi}$ with

$$G(\cos(t)\partial_x + \sin(t)\partial_y) = t.$$

This determines a map $G_\Lambda : \Lambda \to S^1_{2\pi}$ assigning to each point in Λ the unit tangent vector at that point determined by the orientation on Λ.

For any path $\zeta : [0,1] \to \Lambda$, we can associate an angle $\theta(\zeta) \in \mathbb{R}$ as follows: Composing ζ with G_Λ determines a map

$$\phi = G_\Lambda \circ \zeta : [0,1] \to S^1_{2\pi}.$$

Denoting by $\widetilde{\phi}$ the lift of this map to $\mathbb{R}$, the *rotation angle* of ζ, denoted by $\theta(\zeta)$, is defined by

$$\theta(\zeta) = \widetilde{\phi}(1) - \widetilde{\phi}(0).$$

If $q : S^1 \to \mathbb{R}^3$ is a parametrization of a component Λ_i of Λ, then the rotation angle of the associated path $[0,1] \to \Lambda_i$ is $2\pi \operatorname{rot}(\Lambda_i)$.

The *rotation angle* of a composable pair (r_{j_1}, r_{j_2}), denoted by $\theta_{j_1,j_2} \in \mathbb{R}$, will later help us to compute Conley–Zehnder indices of closed Reeb orbits. It is defined as $\theta_{j_1,j_2} = \theta(\eta_{j_1,j_2})$. We write $\overline{\theta}_{j_1,j_2}$ for the rotation angle computed with the opposite capping path $\overline{\eta}_{j_1,j_2}$, whence

$$\theta_{j_1,j_2} - \overline{\theta}_{j_1,j_2} = 2\pi \operatorname{rot}(\Lambda_{l_{j_1}^+}). \tag{11}$$

The *rotation number of a composable pair of chords* (r_{j_1}, r_{j_2}), denoted by $\operatorname{rot}_{j_1,j_2}$, is defined as

$$\operatorname{rot}_{j_1,j_2} = \left\lfloor \frac{\theta_{j_1,j_2}}{\pi} \right\rfloor \in \mathbb{Z}.$$

3.3.2 Crossing monomials Now we define the *crossing monomials*, which will later facilitate our computations of the homology classes of Reeb orbits of the R_ϵ. Consider a collection of variables μ_i indexed by the connected components Λ_i of Λ.

The *crossing monomial of a chord* r_j, denoted by cr_j, is defined by

$$\operatorname{cr}_j = \tfrac{1}{2}((c_j^- + \operatorname{sgn}_j)\mu_{l_j^-} + (c_j^+ + \operatorname{sgn}_j)\mu_{l_j^+}) \in \bigoplus \mathbb{Z}\mu_i. \tag{12}$$

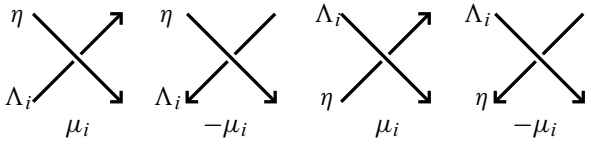


Figure 3: Each subfigure gives a local picture of a crossing of a capping path η with a component of Λ in the Lagrangian projection. Labelings of the strands appears at the left of each subfigure, with the local contribution to the crossing number appearing below. Each subfigure may be rotated by π.

The *crossing monomial of a composable pair of chords* (r_{j_1}, r_{j_2}), denoted by cr_{j_1,j_2}, is defined by

$$\mathrm{cr}_{j_1,j_2} = \sum_{q_{j-}^- \in \mathrm{int}(\eta_{j_1,j_2})} \mathrm{sgn}_{j-} \mu_{l_{j-}^+} + \sum_{q_{j+}^+ \in \mathrm{int}(\eta_{j_1,j_2})} \mathrm{sgn}_{j+} \mu_{l_{j+}^-} \in \bigoplus \mathbb{Z}\mu_i .$$

The contributions are as described in Figure 3.

Remark 3.2 (crossing monomials for connected Λ) When Λ consists of a single connected component, we get a single surgery coefficient c and a single μ. In this case, $\mathrm{cr}_j = (c + \mathrm{sgn}_j)\mu \in 2\mu\mathbb{Z}$ and $\mathrm{cr}_{j_1,j_2} = m\mu$, where m is the number of times the interior capping path η_{j_1,j_2} touches the tips and tails of chords, counted with signs given by the sgn_j.

3.4 Broken closed strings

We temporarily work with an arbitrary contact 3–manifold (M, ξ) containing a Legendrian submanifold Λ. Equip (M, ξ) with a contact form α and write κ_j for the chords of Λ, which will be assumed nondegenerate. Words of chords with boundary on Λ and cyclic words of chords on Λ are defined as above in the obvious fashion.

Let κ_k for $k = 1, \dots, n$ be a sequence of chords on Λ and let $a_k \in \{\pm 1\}$. Let ζ_k be a collection of oriented arcs $\zeta_k : [0, 1] \to \Lambda$ starting at the endpoint (starting point) of κ_k if a_k is positive (negative) and ending at the starting point (endpoint) of κ_{k+1} if a_{k+1} is positive (negative). Assume that the a_k and ζ_{j_k} are such that

$$b = (a_1\kappa_1) * \zeta_1 * \cdots * (a_k\kappa_n) * \zeta_n \tag{13}$$

forms a closed, oriented loop, where $*$ denotes concatenation and $\pm\kappa_k$ is κ_k parametrized with positive (negative) orientation.

Definition 3.3 We call a map b as in (13) a *broken closed string on* Λ. We call the a_k *asymptotic indicators*. We consider broken closed strings which differ by cyclic rotation of indices involved to be equivalent and say that a broken closed string is *parametrized* if a fixed ordering of the indices is in use. We also consider broken closed strings which differ by homotopy of the ζ_k (relative to their endpoints) to be equivalent.

Example 3.4 Let (t, U) be a holomorphic map from a disk with boundary punctures $\{p_j\}$ removed, $\mathbb{D} \setminus \{p_j\}$, to the symplectization of (M, ξ), with boundary punctures asymptotic to chords of Λ. Suppose that

(1) the p_j are indexed according to their counterclockwise ordering along $\partial\mathbb{D}$,

(2) the p_j are a_j–asymptotic to chords κ_j ($a_j = 1$ for positively asymptotic and $a_j = -1$ for negatively asymptotic), and

(3) $U(\partial\mathbb{D} \setminus \{p_j\}) \subset \Lambda$ with ζ_j denoting the restriction of U to the component of $\partial\mathbb{D} \setminus \{p_j\}) \subset \Lambda$ whose oriented boundary is $p_{j+1} - p_j$.

With the data a_j, κ_j and ζ_j specified by (t, U) as above, equation (13) is a broken closed string on Λ. Of particular interest are broken closed strings determined by disks appearing in the *LRSFT* differential for Legendrian links in $(\mathbb{R}^3, \xi_{\text{std}})$ [50].

This may be generalized in the obvious way to holomorphic maps (t, U) whose domain is a compact Riemann surface (Σ, j) decorated with interior punctures (asymptotic to closed Reeb orbits) and boundary punctures (asymptotic to chords). Then any boundary component of Σ determines a broken closed string on Λ.

Definition 3.5 A broken closed string determined by a holomorphic map as in Example 3.4 will be called a *holomorphic boundary component*.

We note that the ζ_k in the definition of a holomorphic boundary component may be constant: for example, if (t, U) is a trivial strip with domain $\mathbb{R} \times I_\epsilon$ for some chord κ, consider

$$b = \kappa * \zeta_1 * (-\kappa) * \zeta_2$$

with ζ_1 being a constant path with value the tip of κ and ζ_2 a constant path with value the tail of κ.

Example 3.6 Suppose that $\Lambda^\pm$ is a contact surgery diagram and let $w = r_{j_1} \cdots r_{j_n}$ be a cyclic word of composable Reeb chords on Λ. There are 2^n parametrized broken closed strings associated to this cyclic word, given by all of the ways that we may

choose orientations for the capping path starting at the tip of each r_{j_k} and ending at the tail of each $r_{j_{k+1}}$:

$$\begin{gathered} r_{j_1} * \eta_{j_1,j_2} * \cdots * r_{j_n} * \eta_{j_n,j_1}, \\ r_{j_1} * \eta_{j_1,j_2} * \cdots * r_{j_n} * \overline{\eta}_{j_n,j_1}, \\ \vdots \\ r_{j_1} * \overline{\eta}_{j_1,j_2} * \cdots * r_{j_n} * \eta_{j_n,j_1}, \\ r_{j_1} * \overline{\eta}_{j_1,j_2} * \cdots * r_{j_n} * \overline{\eta}_{j_n,j_1}. \end{gathered}$$

Definition 3.7 We call each of the broken closed strings described in Example 3.6 an *orbit string* associated to w.

When dealing with orbit strings, the r_j are determined by the indices of the capping paths involved, and so will be omitted from our notation.

Note that a broken closed string on a Legendrian submanifold of dimension n in a contact manifold of dimension $2n+1$ for $n>1$ is uniquely determined by its chords up to homotopy through broken closed strings. We will see in Section 9.3 that a parametrized capping string provides instructions for homotoping a Reeb orbit of $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ into the complement of a neighborhood of $\Lambda^\pm$ in $\mathbb{R}^3$.

3.5 Maslov indices of broken closed strings

Here we define Maslov indices on broken closed strings on Legendrians in contact 3–manifolds, which are relevant to index computation of holomorphic curves. Essentially, we are packaging terminology appearing in the above subsection so as to be cleanly plugged into index computations appearing in [19; 17]. See Section 8.

We assume that $\dim(M)=3$ and that we are working with κ_k, a_k and ζ_k for $k=1,\dots,n$, as described in the previous subsection, determining a broken closed string b whose domain we take to be $\mathrm{dom}(b)=S^1$. We remark on the case $\dim(M)>3$ later in this subsection. Our discussion follows [17, Section 3]. We write $q_k^- \in \Lambda$ for the starting point of each κ_k and q_k^+ for its endpoint.

We assume that ξ is equipped with an adapted almost-complex structure J and suppose that we have a trivialization $s\colon \xi|_{\mathrm{im}(b)} \to \mathbb{C}$ of ξ over the image of a broken closed string b in M which identifies the symplectic structure $d\alpha$ and complex structure J on the target with the standard symplectic and complex structures on $\mathbb{C}$. The trivialization s provides us with an identification

$$b^*\xi \simeq \mathbb{C}\times S^1$$

Denote by $\mathcal{L}(\xi) \to M$ the bundle whose fiber $\mathcal{L}(\xi|_x) \simeq S^1_\pi$ over a point $x \in M$ is the space of unoriented Lagrangian subspaces — that is, unoriented real lines — in $(\xi_x, d\alpha)$.[13] Then s likewise determines an identification

$$b^*\mathcal{L}(\xi) \simeq S^1_\pi \times S^1.$$

Over the subset of S^1 parametrizing the ζ_k, we have a section of this bundle determined by the unoriented Gauss map,

$$t \mapsto T_{b(t)}\Lambda \subset \xi_{b(t)}.$$

Using s, this section determines a map ϕ^G over this subset to S^1_π. We now describe how to extend this section over the subset of S^1 parametrizing the $a_k\kappa_k$.

For each chord κ_k, the time $t \in [0, \mathcal{A}(\kappa_k)]$ flow of R determines a path in $\mathrm{SL}(2,\mathbb{R})$ by writing $\mathrm{Flow}_R^t(\xi_{q_k^-})$ in the standard basis of $\mathbb{R}^2$ determined by s. This likewise determines a section of $\mathcal{L}(\xi)$ over the chord by $\mathrm{Flow}_R^t(T_{q_k^-}\Lambda)$. As we've assumed that κ_k is nondegenerate,

$$\mathrm{Flow}_R^{\mathcal{A}(\kappa_k)}(T_{q_k^-}\Lambda) \neq T_{q_k^+}\Lambda$$

as Lagrangian subspaces of $\xi_{q_k^+}$. In order to assign a Maslov number to b, we must make a correction to obtain a closed loop of Lagrangian subspaces:

(1) If $a_k = 1$, then the orientation of b and the chord coincide. To form a closed loop we join $\mathrm{Flow}_R^{\mathcal{A}(\kappa_k)}(T_{q_k^-}\Lambda)$ to $T_{q_k^+}\Lambda$ by making the smallest possible clockwise rotation to $\mathrm{Flow}_R^{\mathcal{A}(\kappa_k)}(T_{q_k^-}\Lambda)$.

(2) If $a_k = -1$, then the orientation of b and the chord disagree. To form a closed loop of Lagrangian subspaces along b, we start at the endpoint of the chord, follow the negative flow of R, and then join $\mathrm{Flow}_R^{-\mathcal{A}(\kappa_k)}(T_{q_k^+}\Lambda)$ to $T_{q_k^-}\Lambda$ by making the smallest clockwise rotation possible.

Denote by $\phi_{b,s}\colon S^1 \to S^1_\pi$ the map so obtained.

Definition 3.8 We call the degree of the map $\phi_{b,s}$ described above the *Maslov index of the broken closed string b with respect to the framing s*, denoted by $\mathrm{M}_s(b) \in \mathbb{Z}$. It is easy to see that $\mathrm{M}_s(b)$ does not depend on the cyclic ordering of its indices involved, so that it is well defined.

[13] We use the circle of radius π, S^1_π, rather than $S^1_{2\pi}$ due to our ignoring the orientations of the lines involved.

The following easily follows from the construction of M_s:

Proposition 3.9 *Let b be a broken closed string on $\Lambda \subset (M,\xi)$ with a trivialization s of $\xi|_{\mathrm{im}(b)}$. Smooth homotopies of such trivializations s leave $\mathrm{M}_s(b)$ unchanged. The* mod 2 *reduction of $\mathrm{M}_s(b)$ is independent of s, so that we may define $\mathrm{M}_2(b) \in \mathbb{Z}/2\mathbb{Z}$ as an invariant of b.*

Now suppose that $\Gamma_{\neq 0}(\xi)$ is nonempty as in Proposition 2.5, which clearly applies to any $\Lambda \subset (\mathbb{R}^3, \xi_{\mathrm{std}})$:

(1) *If b is homologically trivial in M, then $\mathrm{M}_s(b)$ is independent of $s \in \Gamma_{\neq 0}(\xi)$.*

(2) *If $H_2(M) = H^1(M) = 0$, then $\mathrm{M}_s(b)$ is independent of $s \in \Gamma_{\neq 0}(\xi)$, regardless of the homotopy class of b in M.*

3.6 Generalizations and comparison with existing conventions

3.6.1 Generalized crossing signs and Maslov indices Crossing signs generalize to n–dimensional Legendrian submanifolds inside contact manifolds of dimension $2n+1$ as follows. As above, consider a generic chord κ on an oriented Legendrian submanifold $\Lambda \subset (M,\xi)$ parametrized by an interval $[0,a]$ given by the flow of some R. Then we may define $\mathrm{sgn}(\kappa)$ by

$$\big(\textstyle\bigwedge^n T_{\kappa(a)}\Lambda\big) \wedge \big(\bigwedge^n \mathrm{Flow}_R^a(T_{\kappa(0)}\Lambda)\big) = \mathrm{sgn}(\kappa)\big(\bigwedge^{2n} \xi_{\kappa(a)}\big)$$

as an orientation on $\xi_{\kappa(a)}$. Note that $\mathrm{sgn}(\kappa)$ is independent of the orientation of Λ if and only if Λ is connected. However, the product of sgn over the chords appearing in a broken closed string is always independent of choice of orientation.

We also briefly address generalizations of the Maslov index to higher dimensions. Provided a contact manifold (M,ξ) of dimension $2n+1$, we write $\mathcal{L}(2n) = U(n)/O(n)$ for the space of (unoriented) Lagrangian planes in the standard symplectic vector space and define the bundle

$$\mathcal{L}(2n) \hookrightarrow \mathcal{L}(\xi) \twoheadrightarrow M$$

as above without modification. Provided a trivialization

$$s\colon b^*(\xi) \to \mathbb{C}^n \times S^1,$$

we can view the sections of $b^*\mathcal{L}(\xi)$ as maps from S^1 to $U(n)/O(n)$, in which case $\mathrm{M}_s(b)$ may be defined and computed as the usual Maslov index of loops in the Lagrangian Grassmannian. See for example [46, Theorem 2.35]. The required "clockwise rotation" correction in arbitrary dimensions is described by the paths $\mathrm{f}_j(s)$ appearing in Section 5.9 of [19].

3.6.2 Conventions for capping paths We briefly address how our conventions for capping paths and rotation angles differ from those used to construct gradings in Legendrian contact homology. See for example the exposition [26, Section 3.1]. Assume that $\Lambda \subset (\mathbb{R}^3, \xi_{\text{std}})$ consists of a single component and has a designated basepoint $*$ not coinciding with the tip or tail of any chord.

For a chord r_j, exactly one of $\eta_{j,j}$ or $\overline{\eta}_{j,j}$ will pass through $*$. Denoting by ϕ_j the rotation angle of the path not passing through $*$, the LCH grading is defined — by a slight manipulation of conventional notation — as

$$|r_j| = \left\lfloor \frac{\phi_j}{\pi} \right\rfloor - 1.$$

This is very similar to our computation of rotation numbers except that

(1) knots along which we are performing surgery do not have basepoints,

(2) our capping paths do not necessarily begin and end at endpoints of the same chord,

(3) our capping paths follow the orientation of Λ by default.

We will see that our conventions for computation arise naturally when computing Conley–Zender indices of Reeb orbits of the R_ϵ using the framing construction of Section 6. This convention is also convenient as it will simplify the statements of homology classes of closed Reeb orbits in Section 9.

Our framing construction can be modified so as to naturally lead to computations of rotation angles using basepoints as in LCH. See Remark 6.2. By (11), if $\text{rot}(\Lambda) = 0$ then our computation of rotation angles coincide when the endpoints of a capping path lie over the same chord:

$$\theta_{j,j} = \overline{\theta}_{j,j} = \phi_j.$$

3.6.3 Conventions for broken closed strings In [50, Definition 3.1], broken closed strings have discontinuities at Reeb chords, whereas our broken closed strings are continuous maps. We have chosen to define broken closed strings to include the data of the chords in question, so as reduce ambiguity when discussing chords on Legendrians contained in surgered contact manifolds $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$.

4 Model geometry for Legendrian links and contact surgery

In this section we construct neighborhoods of Legendrian links and then perform contact surgery on $\Lambda^\pm$ using these neighborhoods to describe the contact manifolds $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ and the contact forms α_ϵ.

Our strategy is to develop highly specialized models for the objects involved in contact surgery, determining Reeb vector fields on surgered contact manifolds which are linear in a way which will be made precise in Section 5.1. The main benefits of this approach are that the proofs of the following will be considerably simplified:

(1) The chord-to-orbit (Theorem 5.1) and chord-to-chord (Theorem 5.10) correspondences.

(2) The Conley-Zehnder index (Theorem 7.1) and Maslov index (Theorem 7.2) computations.

We will also be able to determine the embeddings of simple closed orbits in surgered manifolds as fixed points of explicitly defined affine endomorphisms of $\mathbb{R}^2$ (Section 5.6). While we don't pursue computation in this paper, we anticipate this being of utility in future applications.

The primary disadvantage to our contact forms being so specialized is that surgery cobordisms between the $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ will be less explicitly defined and will require greater effort in their construction (Section 10). Furthermore, we will be imposing restrictions on the Lagrangian projections of Legendrian links in the style of [49], so that our analysis — which is applicable to all Legendrian *isotopy classes* $\Lambda^\pm$ — will not be applicable to all chord-generic Legendrian links in $(\mathbb{R}^3, \xi_{\text{std}})$.

Remark 4.1 Our approach to contact surgery is quite similar to that of Foulon and Hasselblatt [29], who defined surgery using a model Dehn twist as in our Section 4.6.

In [7; 18], Bourgeois, Ekholm and Eliashberg describe surgeries as the result of critical-index Weinstein handle attachments and then study the resulting Reeb dynamics. This contrasts with our approach in that we will first describe our contact forms α_ϵ and then build specialized Weinstein handles that have the α_ϵ as the restriction of their Liouville form to their contact boundaries.

The approaches to contact surgery here and [29; 7; 18] all have at least one feature in common: shrinking the size of the surgery locus is used to control Reeb dynamics.

4.1 Almost-complex structures, metrics and the Gauss map

We will want our Legendrians and their neighborhoods to interact nicely with an almost-complex structure J_0 and a metric $g_{\mathbb{R}^3}$, which we now describe.

Define vector fields $X, Y \in \Gamma(\xi_{\text{std}})$ by lifting the derivatives of the usual coordinates:

$$X = \partial_x + y\partial_z, \quad Y = \partial_y.$$

We define a complex structure J_0 on ξ_{std} as the lift of the usual complex structure on $\mathbb{R}^2 = \mathbb{C}$:

$$J_0 X = Y, \quad J_0 Y = -X. \tag{14}$$

This determines an almost-complex structure, which we'll also call J_0, adapted to the symplectization $(\mathbb{R} \times \mathbb{R}^3, e^t \alpha_{\text{std}})$, defined by

$$J_0 \partial_t = \partial_z, \quad J_0 \partial_z = -\partial_t.$$

This almost-complex structure determines a J_0–invariant metric $g_{\mathbb{R}^3}$ on $\mathbb{R}^3$, defined by

$$g_{\mathbb{R}^3}(u, v) = \alpha(u)\alpha(v) + d\alpha(\pi_\alpha u, J_0 \pi_\alpha v), \quad \pi_\alpha(u) = u - \alpha(u)\partial_z \in \xi_{\text{std}}.$$

The metric yields a simple formula for the lengths of vectors in ξ_{std}:

$$Z = aX + bY \in \xi_{\text{std}} \implies |Z| = \sqrt{a^2 + b^2}. \tag{15}$$

4.2 Good position and Lagrangian resolution

Definition 4.2 We say that a Legendrian link $\Lambda \subset \mathbb{R}^3$ is in *good position* if it is chord generic and, for each double point $(x_0, y_0) \in \mathbb{R}^2$ of its Lagrangian projection $\pi_{xy}(\Lambda)$, there exists a neighborhood within which

(1) the over-crossing arc admits a parametrization satisfying

$$(x, y)(q) = (x_0 + q, y_0 - q),$$

and

(2) the under-crossing arc admits a parametrization satisfying

$$(x, y)(q) = (x_0 + q, y_0 + q).$$

Good position guarantees that the Gauss map of a parametrization of Λ evaluates to $\frac{3}{4}\pi$ or $\frac{7}{4}\pi$ near an over-crossing and to $\frac{1}{4}\pi$ or $\frac{5}{4}\pi$ near an under-crossing.[14] Likewise, the condition ensures that capping paths of composable pairs of chords satisfy

$$\theta_{j_1, j_2} \bmod 2\pi \in \left\{\tfrac{1}{2}\pi, \tfrac{3}{2}\pi\right\}.$$

[14] In [7], it is presumed that the tangent map of Reeb flow along a chord r sends $T_{r(0)}\Lambda \subset \xi$ to the subspace $J T_{r(a)}\Lambda$, which is achieved by an appropriate choice of almost-complex structure on the contact hyperplane of the manifold containing Λ. In our case, this is achieved by assuming that Λ is in good position. We will see in the proof of Theorem 7.1 that our analysis is contingent upon this assumption. Similarly precise perturbations of Legendrian submanifolds near endpoints of chords appear in [19] for the purpose of guaranteeing transversality of moduli spaces used to compute differentials for the contact homology of Legendrians in $(\mathbb{R}^{2n+1}, \xi_{\text{std}})$.

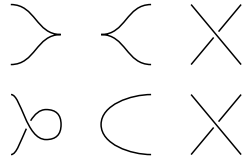


Figure 4: The first row of subfigures shows segments of a Legendrian link appearing in the front projection. Directly below each subfigure is how it appears in the Lagrangian resolution.

Proposition 4.3 *Provided a front projection of a Legendrian link Λ, we may perform a Legendrian isotopy so that the following properties are satisfied*:

(1) *Λ is in good position.*

(2) *The Lagrangian diagram is obtained by resolving singularities of the front as depicted in Figure 4.*

(3) *The arc length of each connected component of Λ with respect to $g_{\mathbb{R}^3}$ is 1.*

Proof The proof proceeds in three steps. The first step establishes the first two desired properties of Λ. This step is essentially the proof of Proposition 2.2 from [49] and so we will omit the details. The only modification required to ensure a link is in good position after Legendrian isotopy is to control $\partial z/\partial x$ of a parametrization of Λ near the right-pointing cusps and what are called "exceptional segments" in that proof. In particular, $\partial z/\partial x$ can be made quadratic with highest-order coefficient $\frac{1}{2}$ $\left(\text{resp. } -\frac{1}{2}\right)$ on neighborhoods of the positive (resp. negative) endpoints of chords.

In our second step, we modify Λ so that the arc length of each component is arbitrarily small while maintaining our desired conditions on the Lagrangian projection. For $\rho > 0$, consider the linear transformation ϕ_ρ of $\mathbb{R}^3$, defined by $\phi_\rho(x, y, z) = (\rho x, \rho y, \rho^2 z)$. Then $\phi_\rho^* \alpha_{\text{std}} = \rho^2 \alpha_{\text{std}}$, so that each ϕ_ρ is a contact transformation. The map ϕ_ρ also has the following useful properties:

(1) It preserves the angles of vectors in ξ_{std}.

(2) If Λ_i is a Legendrian curve with arc length ℓ, then $\phi_\rho(\Lambda_i)$ has arc length $\rho\ell$.

Take the family of Legendrians $\phi_{e^{-T}}(\Lambda)$ for $T \in [0, T_0]$ with T_0 large enough that each connected component of $\phi_{e^{-T_0}}(\Lambda)$ will have arc length ≤ 1. This interpolation between Λ and $\phi_{e^{-T_0}}(\Lambda)$ determines a 1–parameter family of Legendrian submanifolds and so may be realized by a Legendrian isotopy.

In the case that Λ connected, we choose T_0 so that the arc length is exactly equal to 1 after the isotopy and conclude the proof. When Λ is disconnected, a final, third step is

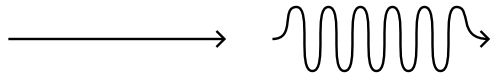


Figure 5: Locally modifying a Legendrian in the Lagrangian projection by a rapidly oscillating function $\tilde{y}$ to increase its arc length.

required. In this step we increase the arc lengths of the connected components of Λ so that they are all 1 while preserving good position and the smooth isotopy type of the Lagrangian projection.

We demonstrate how to increase arc lengths so as to achieve the desired result. Consider a segment of Λ_i along which the x–derivative is nonzero, parametrized via the x variable by $x \mapsto (x, y(x), z(x))$ with $x \in [-\delta, \delta]$ for an arbitrarily small positive constant δ. We assume that the Lagrangian projection of the segment does not touch any double points. Let $\tilde{y} \in \mathcal{C}^\infty([-\delta, \delta])$ be a function with compact support contained in $(-\delta, \delta)$ and for which $\int_{-\delta}^{\delta} \tilde{y}\, dx = 0$. Consider perturbations $\Lambda_{i,T}$ of Λ_i parametrized by $T \in [0, 1]$ which modify Λ_i along our segment to take the form

$$x \mapsto \left(x, y + T\tilde{y}, z + T\int_{-\delta}^{x} \tilde{y}\, dx\right).$$

The vanishing of the integral of $\tilde{y}$ ensures that the z–values at the endpoints of the segment are unaffected by the perturbation. By making $\sup \tilde{y}$ small and $\int |\partial\tilde{y}/\partial x|\, dx$ very large, we can ensure that, for $T \in [0, 1]$, our perturbations introduce no new double points in the Lagrangian projection and that $\Lambda_{i,T}$ has arc length as large as we like, say 2 when $T = 1$. See Figure 5. Hence, for some $T_0 \in [0, 1]$, the arc length of Λ_{i,T_0} will be exactly 1.

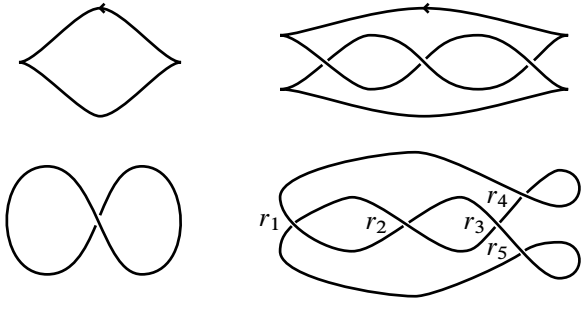


Figure 6: The left column shows Legendrian $\mathrm{tb} = -1$ unknot in the front and Lagrangian projections. A right-handed trefoil knot with $\mathrm{tb} = 1$ and $\mathrm{rot} = 0$ is shown in the front and Legendrian projections on the right. The Reeb chords of the Lagrangian projection of the trefoil are labeled r_i.

Apply such perturbations to each connected component of Λ so that no new double points are created and neighborhoods of double points are unaffected. Each perturbation is realizable by a Legendrian isotopy. Thus, we have obtained a Legendrian isotopy of Λ having all of the desired properties. □

Provided Λ as a front projection diagram, we call the Lagrangian projection of (an isotopic copy of) Λ obtained as in the above proposition the *Lagrangian resolution* of the front diagram. Figure 6 displays Lagrangian resolutions of an unknot and a trefoil. Following [49], we say that a front projection of a Legendrian link Λ is *nice* if there exists some $x_0 \in \mathbb{R}$ for which all right-pointing cusps have x–value x_0. It's not difficult to see that any Λ can be isotoped to have a nice front projection.

4.3 Conventions for link diagrams

We will not concern ourselves with specific requirements of good position or arc length when drawing Legendrian links in the Lagrangian projection and consider such a diagram to be valid if it recovers the Lagrangian projection of a Legendrian link after an isotopy of the xy–plane. In particular, we will not take care to ensure that angles at crossings are precise or that the components of $\mathbb{R}^2 \setminus \pi_{xy}(\Lambda)$ satisfy the area requirements of [24, Section 2].

Throughout, Legendrian knots with surgery coefficient $+1$ will be colored blue and knots with surgery coefficient -1 will be colored red. If the coefficient of a knot is not already determined or the knot corresponds to a component of Λ^0, it will be colored black.

4.4 Standard neighborhoods

Before stating the properties we will want our neighborhoods of Λ to have, we will create model neighborhoods near under- and over-crossings of chords. The neighborhood construction is completed in Proposition 4.5.

4.4.1 Model neighborhoods near endpoints of chords Here we describe a construction of a neighborhood of Λ along the arcs described in Definition 4.2. We can reparametrize the arcs to have unit speed, so that they take the form

$$q \mapsto \left(x_0 + \tfrac{1}{\sqrt{2}}q,\, y_0 - \tfrac{1}{\sqrt{2}}q,\, z_0 + \tfrac{1}{\sqrt{2}}y_0 q - \tfrac{1}{4}q^2\right)$$

near an over-crossing and

$$q \mapsto \left(x_0 + \tfrac{1}{\sqrt{2}}q,\, y_0 + \tfrac{1}{\sqrt{2}}q,\, z_0 + \tfrac{1}{\sqrt{2}}y_0 q + \tfrac{1}{4}q^2\right)$$

along an under-crossing. For $\epsilon > 0$ sufficiently small, we extend these embeddings to embeddings of $I_\epsilon \times I_\epsilon \times I_{2\epsilon}$ into $\mathbb{R}^3$ using coordinates (z, p, q). Near an over-crossing, this embedding takes the form

$$\begin{aligned}(16)\quad &\Phi^+_{x_0,y_0,z_0}(z,p,q)\\ &= \big(x_0 - \tfrac{1}{\sqrt{2}}p + \tfrac{1}{\sqrt{2}}q,\, y_0 - \tfrac{1}{\sqrt{2}}p - \tfrac{1}{\sqrt{2}}q,\, z_0 + z + y_0\tfrac{1}{\sqrt{2}}q - p + \tfrac{1}{4}p^2 + \tfrac{1}{2}pq - \tfrac{1}{4}q^2\big).\end{aligned}$$

Near an under-crossing arc, this takes the form

$$\begin{aligned}(17)\quad &\Phi^-_{x_0,y_0,z_0}(z,p,q)\\ &= \big(x_0 + \tfrac{1}{\sqrt{2}}p + \tfrac{1}{\sqrt{2}}q,\, y_0 - \tfrac{1}{\sqrt{2}}p + \tfrac{1}{\sqrt{2}}q,\, z_0 + z + y_0\tfrac{1}{\sqrt{2}}p + q - \tfrac{1}{4}p^2 + \tfrac{1}{2}pq + \tfrac{1}{4}q^2\big).\end{aligned}$$

Properties 4.4 *The following properties are satisfied by the* $\Phi^\pm_{x_0,y_0,z_0}$:

(1) $\Phi^\pm_{x_0,y_0,z_0}(0,0,q)$ *provides a parametrization of* Λ *with unit speed.*

(2) $(\Phi^\pm_{x_0,y_0,z_0})^*\alpha_{\text{std}} = dz + p\,dq$.

(3) *With respect to the basis* $P = \partial_p$ *and* $Q = \partial_q - p\partial_z$, *we have* $J_0 = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

(4) $\pi_{xy} \circ \Phi^\pm_{x_0,y_0,z_0}$ *is an affine map.*

(5) *The images of* $\pi_{xy} \circ \Phi^\pm_{x_0,y_0,z_0}$ *overlap in squares of the form* $I_\epsilon \times I_\epsilon$ *near a crossing* (see Figure 9).

4.4.2 Neighborhood construction We now assume that Λ satisfies the conclusions of Proposition 4.3.

Proposition 4.5 *For* ϵ_0 *sufficiently small, there exists a neighborhood* $N_{\epsilon_0,i}$ *of each* Λ_i *parametrized by an embedding*

$$\Phi_i : I_{\epsilon_0} \times I_{\epsilon_0} \times S^1 \to \mathbb{R}^3$$

with coordinates (z, p, q) *such that the following conditions are satisfied*:

(1) $\Phi_i^*\alpha_{\text{std}} = dz + p\,dq$.

(2) *The* $N_{\epsilon_0,i}$ *are disjoint.*

(3) $\Phi_i(0,0,q)$ *provides a unit-speed parametrization of* Λ_i.

(4) J_0 *is* z*–invariant in* $N_{\epsilon_0,i}$ *and, with respect to the basis* $P = \partial_p$ *and* $Q = \partial_q - p\partial_z$, *it satisfies*

$$\Phi_i^* J_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \mathcal{O}(p).$$

(5) *Near the endpoints* $(x_j, y_j, z_j^{\pm})$ *with* $z_j^+ > z_j^-$ *of each chord* r_j *of touching* Λ, *we can find a matrix of the form* $M = \mathrm{Diag}(1,1,1)$ *or* $\mathrm{Diag}(1,-1,-1)$ *such that*

$$\Phi_i(z,p,q) = \Phi^{\pm}_{x_j,y_j,z_j^{\pm}} \circ M(z,p,q-q_j^{\pm}),$$

where the $\Phi_{x_j,y_j,z_j^{\pm}}$ *are as in Properties 4.4.*

Proof Presuming that Λ_i is parametrized with a variable q with respect to which it has unit speed, we pick an arbitrarily small positive constant ϵ_1 and define a map $I_{\epsilon_1} \times S^1 \to \mathbb{R}^3$ as

$$\phi_1 \colon (p,q) \mapsto \exp_{\Lambda_i(q)}\left(-pJ_0\frac{\partial \Lambda_i}{\partial q}(q) + h_1(p,q)\right),$$

where $h_1 \in C^\infty(I_{\epsilon_1} \times S^1, \mathbb{R}^3)$ vanishes up to second order in p and is chosen so that it produces the map

$$(p,q) \mapsto \Phi^{\pm}_{x_j,y_j,z_j^{\pm}} \circ M(0,p,q-q_j^{\pm})$$

near the endpoints of the chords of Λ as in the statement of the proposition. Here exp is the exponential map with respect to the metric $g_{\mathbb{R}^3}$ and the matrix M is as in the statement of the proposition.

Since the tangent map of the exponential map is the identity along the zero section,

$$T\exp_{\Lambda_i(q)} = \mathrm{Id} \colon T_{\Lambda_i(q)}\mathbb{R}^3 \to T_{\Lambda_i(q)}\mathbb{R}^3$$

and h_1 is $\mathcal{O}(p^2)$, we have

$$\left.\frac{\partial \phi_1}{\partial p}\right|_{p=0} = -J_0\frac{\partial \Lambda_i}{\partial q} + \left.\frac{\partial h_1}{\partial p}\right|_{p=0} = -J_0\frac{\partial \Lambda_i}{\partial q} = -J_0 e^{J_0 G_i} X = -e^{J_0 G_i(q)} Y,$$

$$\left.\frac{\partial \phi_1}{\partial q}\right|_{p=0} = \frac{\partial}{\partial q}\exp_{\Lambda_i(q)}(0) = \frac{\partial \Lambda_i}{\partial q} = e^{J_0 G_i} X.$$

Therefore, the tangent map for ϕ_1 can be expressed along $\{p=0\}$ as a matrix

$$T\phi_1|_{p=0} = -J_0 e^{J_0 G_i(q)} \tag{18}$$

with incoming basis (P, Q) and outgoing basis (X, Y). This map will be an embedding when restricted to some $I_{\epsilon_1} \times S^1$ for ϵ_1 sufficiently small.

From (18), we compute $\phi_1^* \, d\alpha_{\mathrm{std}} = dp \wedge dq$ along $\{p=0\}$. More generally, we can write $\phi_1^* \, d\alpha_{\mathrm{std}} = F \, dp \wedge dq$ for some smooth function F satisfying $F|_{\{p=0\}} = 1$. Hence, F will be strictly positive on some tubular neighborhood of $\{p=0\} \subset I_{\epsilon_1} \times S^1$, so that $\phi_1^* \, d\alpha_{\mathrm{std}}$ will be symplectic on some $I_{\epsilon_2} \times S^1$ for ϵ_2 sufficiently small.

Applying a fiberwise Taylor expansion to $\phi_1^*\alpha_{\text{std}}$ along the annulus $I_{\epsilon_2}\times S^1$, we write

$$\phi_1^*\alpha_{\text{std}} = (f_0 + pf_1 + p^2 f_2 + f_{\text{hot}})\,dp + (g_0 + pg_1 + p^2 g_2 + g_{\text{hot}})\,dq,$$

where

(1) f_{hot} and g_{hot} are functions of p and q which are $\mathcal{O}(p^3)$, and

(2) $f_0,\dots,g_2$ are functions of q.

As Λ_i is Legendrian and J_0 preserves the contact structure, we must have $f_0 = g_0 = 0$. Then, computing

$$\phi_1^*\,d\alpha_{\text{std}} = d\phi_1^*\alpha_{\text{std}} = \left(g_1 + p\left(2g_2 - \frac{\partial f_1}{\partial q}\right) - p^2\frac{\partial f_2}{\partial q} + \frac{\partial g_{\text{hot}}}{\partial p} - \frac{\partial f_{\text{hot}}}{\partial q}\right) dp\wedge dp,$$

we must have $g_1 = 1$, so that

$$\phi_1^*\alpha_{\text{std}} = (pf_1 + p^2 f_2 + f_{\text{hot}})\,dp + (p + p^2 g_2 + g_{\text{hot}})\,dq.$$

We can eliminate the f_1 term in this equation with a perturbation in the z direction. With $h_2 = \frac{1}{2}p^2 f_1$, we have $dh_2 = pf_1\,dp + \frac{1}{2}p^2\partial f_1/\partial q\,dq$. Hence,

$$\phi_2(p,q) = \phi_1(p,q) - (0,0,h_2(p,q)),$$

admits an expansion of the form

$$\phi_2^*\alpha_{\text{std}} = (p^2 f_2 + f_{\text{hot}})\,dp + (p + p^2 g_2 + g_{\text{hot}})\,dq.$$

To ensure that this map is an embedding, we restrict its domain to $I_{\epsilon_3}\times S^1$ for some sufficiently small $\epsilon_3 \le \epsilon_2$. We note that the f_2 and g_2 here may differ from those in the Taylor expansion of $\phi_1^*\alpha_{\text{std}}$.

Now we'll apply a Moser argument as in [46, Section 3.2] to modify ϕ_2 by precomposing it with an isotopy to produce a map ϕ_3 so that $\phi_3^*\,d\alpha_{\text{std}} = dp\wedge dq$. Due to the facts that the annulus is not closed and that we'll require the result to be an codimension 1 embedding, we cannot simply quote [46, Section 3.2].

Writing $\phi_2^*\alpha_{\text{std}} = p\,dq + \sigma$ and, solving for a vector field X_σ satisfying

$$dp\wedge dq(*, X_\sigma) = \sigma,$$

we see that σ and X_σ have coefficient functions, vanishing up to second order,

$$\sigma = \mathcal{O}(p^2)\,dp + \mathcal{O}(p^2)\,dq, \quad X_\sigma = \mathcal{O}(p^2)\partial_p + \mathcal{O}(p^2)\partial_q.$$

Writing $\text{Flow}^t_{X_\sigma}$ for the time t flow of X_σ, choose $\epsilon_4 \le \epsilon_3$ so that $\text{Flow}^t_{X_\sigma}(I_{\epsilon_4}\times S^1) \subset I_{\epsilon_3}\times S^1$ for all $t\in[0,1]$ and define

$$\phi_3(p,q) = \phi_2 \circ \text{Flow}^1_{X_\sigma}(p,q)\colon I_{\epsilon_4}\times S^1 \to \mathbb{R}^3.$$

The Moser argument shows that $\phi_3^* d\alpha_{\text{std}} = dp \wedge dq$, as desired. Moreover our conditions on X_σ imply that $\text{Flow}^1_{X_\sigma}$ must agree with the identity mapping up to third order along $\{p = 0\}$. Hence, we can continue to write $\phi_3^*\alpha_{\text{std}} = p\,dq + \sigma$ for some σ which vanishes up to second order in p. Using $\phi_3^* d\alpha_{\text{std}} = dp \wedge dq$, we know that σ is closed, and, since it mush vanish along $\{p = 0\}$, we conclude that it is exact. Hence, $\phi_3^*\alpha_{\text{std}} = p\,dq + dh_4$ for some $h_4 \in C^\infty(I_{\epsilon_3} \times S^1, \mathbb{R})$. Possibly restricting to some $I_{\epsilon_4} \times S^1$, we define

$$\phi_4(p, q) = \phi_3(p, q) - (0, 0, h_4),$$

so that ϕ_4 is an embedding, whence $\phi_4^*\alpha_{\text{std}} = p\,dq$. Now define

$$\Phi_i(z, p, q) = \phi_4(p, q) + (0, 0, z).$$

Restricting to some $I_{\epsilon_0} \times I_{\epsilon_0} \times S^1$ for ϵ_0 sufficiently small, we can ensure that $\bigsqcup \Phi_i$ is an embedding. By construction of the Φ_i, we have

$$\Phi_i^*\alpha_{\text{std}} = dz + p\,dq.$$

Regarding the formula for J_0 in the basis (P, Q), note that this is satisfied for the map ϕ_1 and that subsequent perturbations — ϕ_2, ϕ_3 and ϕ_4 — preserve (P, Q) up to second order in p. The z–invariance of J_0 is clear from the definition of the Φ_i and z–invariance of the almost-complex structure on ξ_{std}.

For the last condition stated in the proposition, we note that ϕ_1 produces the desired result by definition of the function h_1. As all other required conditions are satisfied by ϕ_1, where the last condition is required to be satisfied as per Properties 4.4. The perturbations of ϕ_1 carried out in the remainder of the proof are trivial where this condition is required to be satisfied. Indeed, near the endpoints of chords h_2 (used to define ϕ_2), σ (used to define ϕ_3) and h_4 (used to define ϕ_4) all vanish. □

Assumptions 4.6 We assume throughout the remainder of this article that the Legendrian link Λ is in good position and has unit arc length with respect to $g_{\mathbb{R}^3}$, and write

$$N_\epsilon = \bigcup_i N_{\epsilon,i}$$

for a neighborhood of Λ as described in the above proposition with $\epsilon \le \epsilon_0$. We call the set $\{z = 0\} \subset N_{\epsilon,i}$ the *ribbon* of Λ_i. From the above proof, we may assume that the image of the projection of the ribbon of Λ_i to the xy–plane coincides with the image of the projection of $N_{\epsilon,i}$.

4.5 Transverse push-offs

The boundary of the ribbon of a component Λ_i of Λ consists of two knots

$$T^+_{i,\epsilon} = \{z = 0, p = \epsilon\} \quad T^-_{i,\epsilon} = \{z = 0, p = -\epsilon\}. \tag{19}$$

Definition 4.7 With ϵ fixed, the knots $T^+_{i,\epsilon}$ and $T^-_{i,\epsilon}$ will be called the *positive and negative transverse push-offs* of Λ_i. We orient both of these knots so that $\partial_q > 0$ in the coordinate system on $N_{\epsilon,i}$.

The positive (negative) transverse push-off is positively (negatively) transverse to ξ_Λ. Because these knots live on the boundary of $N_{\epsilon,i}$, we may consider them as living within either $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ or $(\mathbb{R}^3, \xi_{\text{std}})$.

4.6 Model Dehn twists

In this and the following subsection we describe contact forms on $\mathbb{R}^3_{\Lambda^\pm}$ which will facilitate analysis on Reeb orbits after contact ± 1 surgery. We begin by providing an explicit model for a Dehn twist and then describe the gluing map used to define contact ± 1 surgery explicitly.

Provided a smooth function $f : \mathbb{R} \to S^1$, we define $\tau_f \in \text{Diff}^+(\mathbb{R} \times S^1)$ by

$$\tau_f(p, q) = (p, q + f(p))$$

and note that $\tau_{-f} = \tau_f^{-1}$. We'll call the map τ_f a *positive* (*resp. negative*) *Dehn twist by* f if:

(1) The derivative of f has compact support in $\mathbb{R}$.

(2) $\int_{\mathbb{R}} \partial f/\partial p \, dp = -1$ (resp. $+1$).

A positive (resp. negative) Dehn twist by f is a positive (resp. negative) Dehn twist in the usual sense of the expression. We compute

$$\tau_f^* p \, dq = p \, dq + p \frac{\partial f}{\partial p} \, dp, \quad \tau_f^*(dp \wedge dq) = dp \wedge dq, \tag{20}$$

so that τ_f is always a symplectomorphism with respect to $dp \wedge dq$ but does not preserve $p \, dq$ unless f is constant. For any f and $\epsilon > 0$, we write

$$f_\epsilon(p) = f\left(\frac{p}{\epsilon}\right).$$

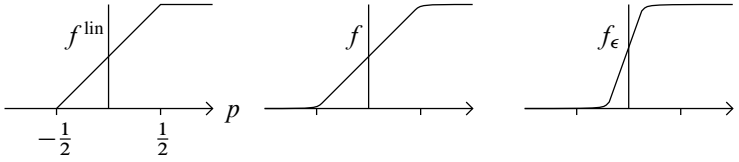


Figure 7: The functions f^{lin}, f and f_ϵ.

Assumptions 4.8 Throughout the remainder of this paper, f will denote a function for which τ_f is a negative Dehn twist whose derivative $\partial f/\partial p$ is

(1) nonnegative,

(2) an even function of p,

(3) supported on $I_1 = [-1, 1]$, and

(4) bounded in absolute value pointwise by 1.

We think of f as being a smooth approximation to a piecewise-linear function

$$f^{\mathrm{lin}}(p) = \begin{cases} 0 & \text{if } p \in \left(-\infty, -\frac{1}{2}\right], \\ p + \frac{1}{2} & \text{if } p \in I_{1/2}, \\ 1 & \text{if } p \in \left[\frac{1}{2}, \infty\right). \end{cases} \tag{21}$$

See Figure 7.

The following proposition gathers some properties of the deviation of twists by f_ϵ from preserving $p\,dq$, as described in (20):

Proposition 4.9 *Suppose that f satisfies Assumptions 4.8 and, for $\epsilon \in (0, 1)$, define*

$$H_\epsilon(p) = \int_{-\infty}^{p} P \frac{\partial f_\epsilon}{\partial p}(P)\, dP.$$

Then H is well defined, zero on the complement of I_ϵ, symmetric and satisfies $-\epsilon \le H_\epsilon \le 0$ pointwise.

Proof The first two statements are clear from the compact support and symmetry of the derivative of f. Then, using the fact that $\partial f_\epsilon/\partial p$ is supported on I_ϵ, we have

$$|H_\epsilon(p)| = \frac{1}{\epsilon}\left|\int_{-\epsilon}^{p} P \frac{\partial f}{\partial p}\left(\frac{P}{\epsilon}\right) dP\right| \le \frac{1}{\epsilon} \sup_p \left|\frac{\partial f}{\partial p}\right| \int_{-\epsilon}^{\epsilon} |P|\, dP = \epsilon \sup_p \left|\frac{\partial f}{\partial p}\right| \le \epsilon.$$

For $p \le 0$, $H_\epsilon(p)$ is an integral of a nonpositive function and so must be nonpositive. Then, for $p \ge 0$,

$$H_\epsilon(p) = H_\epsilon(0) + \int_0^p P \frac{\partial f_\epsilon}{\partial p}(P)\, dP = H_\epsilon(0) - \int_{-p}^0 P \frac{\partial f_\epsilon}{\partial p}(P)\, dP = H_\epsilon(-p)$$

by the symmetry of the derivative of f. □

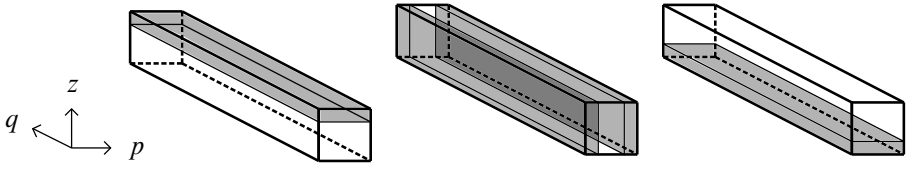


Figure 8: From left to right: the top, side and bottom pieces of our neighborhood are shaded.

4.7 Gluing maps

Now we define the gluing maps to define contact surgery on Λ and contact forms α_ϵ on the surgered manifold $\mathbb{R}^3_{\Lambda\pm}$.

Let ϵ_0 be a sufficiently small as described in Proposition 4.5 and choose $\epsilon \in (0, \epsilon_0)$. We decompose a neighborhood of each $\partial N_{\epsilon,i}$ into top, side and bottom pieces as shown in Figure 8:

(1) $T_{\delta,\epsilon} = \{z \geq \epsilon - \delta\}$.

(2) $S_{\delta,\epsilon} = \{|p| \geq \epsilon - \delta\}$.

(3) $B_{\delta,\epsilon} = \{z \leq -\epsilon + \delta\}$.

To perform contact surgery along Λ_i with surgery coefficient c_i, we define a map $\phi_{c_i,f,\epsilon,\delta}$ in coordinates (z, p, q) by

$$\phi_{c,f,\epsilon,\delta}(z,p,q) \sim \begin{cases} (z - c_i H_\epsilon(p), p, q + c f_\epsilon(p)) & \text{along } T_{\delta,\epsilon}, \\ (z,p,q) & \text{along } S_{\delta,\epsilon} \cup B_{\delta,\epsilon}, \end{cases} \tag{22}$$

where H_ϵ is as described in Proposition 4.9. Due to the properties of f_ϵ and H_ϵ described in the previous section, we have that $\phi_{c,f,\epsilon,\delta}$ agrees on the overlaps of the top, bottom and sides of $N_{\epsilon,i}$ for δ sufficiently small. Therefore, the map determines a smooth gluing.

The tangent map of the gluing map is given by

$$T\phi_{c_i,f,\epsilon,\delta} = \partial_z \otimes \left(dz - c_i\, p \frac{\partial f_\epsilon}{\partial p}\, dp \right) + \partial_p \otimes dp + \partial_q \otimes \left(dq + c_i \frac{\partial f_\epsilon}{\partial p}\, dp \right) \tag{23}$$

along $T_{\delta,\epsilon}$ and $T\phi_{c_i,f,\epsilon,\delta} = \mathrm{Id}$ along $S_{\delta,\epsilon} \cup B_{\delta,\epsilon}$, so that

$$\phi^*_{c_i,f,\epsilon,\delta}(dz + p\, dq) = dz + p\, dq.$$

The gluing map therefore determines a contact form $\alpha_{c_i,f,\epsilon,\delta}$ on the manifold $\mathbb{R}^3_{\Lambda_i}$ obtained by performing the surgery and hence a contact structure $\xi_{\Lambda\pm} = \ker(\alpha_{c_i,f,\epsilon,\delta})$

on this manifold. Shrinking δ amounts to a restriction of the domain of the map and so does not affect the associated contact manifold.

Definition 4.10 For $\epsilon \in (0, \epsilon_0)$, we write α_ϵ for the contact form on $\mathbb{R}^3_{\Lambda^\pm}$ determined by performing surgery using the gluings $\phi_{c_i, f, \epsilon, \delta}$ as described in (22) to each connected component $N_{\epsilon,i}$ of N_ϵ. The Reeb vector field of α_ϵ will be denoted by R_ϵ.

5 Chord-to-orbit and chord-to-chord correspondences

In this section we study the dynamics of the Reeb vector fields R_ϵ for the contact forms α_ϵ for $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ as described in Definition 4.10. Our results are summarized by the following:

Theorem 5.1 *There exist one-to-one correspondences between*:

(1) *Closed orbits of R_ϵ in $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ and cyclic words of chords on $\Lambda^\pm \subset (\mathbb{R}^3, \xi_{\text{std}})$.*

(2) *Chords of R_ϵ with boundary on $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ and words of chords with boundary on $\Lambda^0 \subset (\mathbb{R}^3, \xi_{\text{std}})$.*

A description of the correspondences will be given below.

Definition 5.2 Via the above theorem, we use the notation

$$(r_{j_1} \cdots r_{j_n})$$

to denote either a closed orbit of R_ϵ or a chord of $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ whose underlying word is $r_{j_1} \cdots r_{j_n}$.

After establishing Theorem 5.1, we estimate the actions of chords and closed orbits in $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ in Section 5.5. Then, in Section 5.6, we describe equations whose solutions determine the embeddings of closed Reeb orbits and allow exact calculation of their actions. While we do not provide a closed form solutions to these equations, their analysis provides the following:

Theorem 5.3 *For each $n > 0$, there exists ϵ_n such that, for all $\epsilon \le \epsilon_n$, all orbits γ of word length $\le n$ are hyperbolic with*

$$\mathrm{CZ}_2(\gamma) = \sum_{k=1}^{n} (\mathrm{rot}_{j_k, j_{k+1}} + \delta_{1, c^+_{j_k}}) \in \mathbb{Z}/2\mathbb{Z}.$$

Moreover, if either $\Lambda^+ = \varnothing$ or $\Lambda^- = \varnothing$, then all closed orbits of R_ϵ are hyperbolic for all $\epsilon < \min\{\frac{1}{2}, \epsilon_0\}$.

Throughout this section, γ will denote a closed orbit of R_ϵ and κ will denote a chord of $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$.

5.1 Overlapping rectangles

In order to state our chord-to-orbit and chord-to-chord correspondences, we need to introduce the objects which will define them: embedded squares in $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ which record the positions of Reeb orbits as they propagate through the manifold. Along the way, we slightly refine the specifications of the function f in our surgery construction so as to reduce our analysis of dynamics of R_ϵ to analysis of affine linear transformations.

With ϵ — the constant which governs the size of N_ϵ — sufficiently small, the projection of N_ϵ to the xy–plane will have overlaps only at rectangles centered about double points of the Lagrangian projection of Λ. There is a unique rectangle $\mathcal{D}_j \subset \mathbb{R}^2$ for each chord r_j. As per Assumptions 4.6 and Properties 4.4, each $\mathcal{D}_j$ is the image of a map of the form

$$(p, q) \mapsto (x_0 + p - q, y_0 + p + q)$$

for $(p, q) \in I_{\epsilon_1} \times I_{\epsilon_2}$ for some $\epsilon_i \in (0, \infty)$ with $(x_0, y_0) \in \mathbb{R}^2$ being the coordinates of the double point of Λ in the xy–plane corresponding to r_j.

We write $\mathcal{D}_j^{\text{ex}}$ for the lift of this disk to the top of the N_{ϵ, l_j^-} and $\mathcal{D}_j^{\text{en}}$ for the lift to the bottom of N_{ϵ, l_j^+}. The superscripts are indicative of the fact that closed orbits of R_ϵ enter N_ϵ through the $\mathcal{D}_j^{\text{en}}$ and exit N_ϵ through the $\mathcal{D}_j^{\text{ex}}$. See Lemma 5.6.

Again using Assumptions 4.6 and Properties 4.4, we have that each $\mathcal{D}_j^{\text{en}}$ and $\mathcal{D}_j^{\text{ex}}$ can be described as

$$\{z = \pm\epsilon, q \in [q_j^\pm - \delta, q_j^\pm + \delta]\} \tag{24}$$

for some $\delta \in (0, \infty)$ with respect to the coordinates (z, p, q) provided by Proposition 4.5 on the "outside" of the surgery handle.

If we flow $\mathcal{D}_j^{\text{en}}$ through the surgery handle in which it is contained, we will see it pass through the top $\{z = \epsilon\}$ in a set $\mathcal{D}_j^\tau$, which, when projected onto the (p, q) coordinates, is of the form

$$\mathcal{D}_j^\tau = \tau_{f_\epsilon}^{c_j^+}(\{q \in [q_0 - \delta, q_0 + \delta]\})$$

for some $q_0 \in S^1$ and $\delta > 0$. This set will intersect the each $\mathcal{D}_{j'}^{\text{ex}}$ for $j' \neq j$ in a connected set diffeomorphic to a square. These intersections are depicted as the dark gray regions in Figure 9, right.

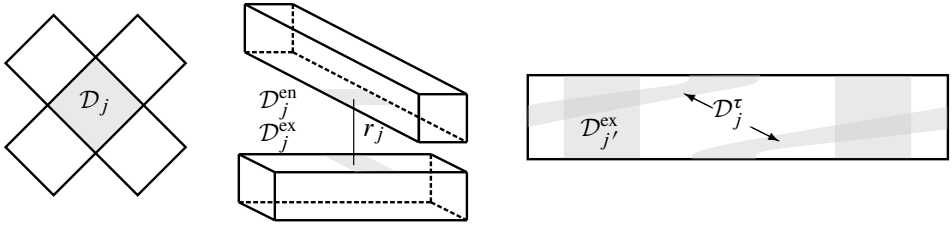


Figure 9: On the left, we see the xy–projection of the ribbon of Λ overlapping at a rectangle $\mathcal{D}_j$. In the middle — with a slightly offset point of view — we see the $\mathcal{D}_j^*$ touching the endpoints of a chord r_j. Here the boxes represent portions of N_ϵ. On the right, we see τ_{f_ϵ} applied to one rectangle intersecting other rectangles. In this portion of the diagram, ∂_p points upward and ∂_q points to the left.

Assumptions 5.4 For a fixed ϵ, we refine our choice of f in Assumptions 4.8 so that it is affine with derivative equal to 1 on some $I_{1-\delta} \subset I_1$, with δ chosen sufficiently small that each $\mathcal{D}_j^\tau \cap \mathcal{D}_{j'}^{\text{ex}}$ with $j \neq j'$ is determined by a pair of linear inequalities

$$\mathcal{D}_j^\tau \cap \mathcal{D}_{j'}^{\text{ex}} = \{q \in [q_0 - \delta_1, q_0 + \delta_1],\, a + bq \in [\delta_2, \delta_3]\}$$

for constants a, b, δ_1, δ_2 and δ_3.

Properties 5.5 *Under Assumptions 5.4, we have that, at any point $(p,q) \in \mathcal{D}_j^{\text{en}}$ for which $\tau_{f_\epsilon}(p,q) \in \mathcal{D}_{j'}^{\text{ex}}$,*

$$\frac{\partial f_\epsilon}{\partial p}(p) = \frac{1}{\epsilon}, \quad H_\epsilon(p) = H_\epsilon(0) + \frac{p^2}{2\epsilon},$$

where $i = l_j^+$. At such points we can write $\tau_{f_\epsilon}^{c_i}$ as

$$(p,q) \mapsto \left(p, q + \frac{1}{2} + \frac{c_i\, p}{\epsilon}\right).$$

5.2 Cyclic words from Reeb orbits

Here we prove the easy part of the of the (cyclic words) $\leftrightarrow$ (closed orbits) correspondence, showing that each γ uniquely determines a cyclic word of chords on $\Lambda^+ \cup \Lambda^-$.

Lemma 5.6 *Any closed orbit γ of R_ϵ must pass through N_ϵ. Every time γ enters N_ϵ, it must pass through some $\mathcal{D}_j^{\text{en}}$, and, every time it exits N_ϵ, it must pass through some $\mathcal{D}_j^{\text{ex}}$.*

Proof The Reeb vector field R_ϵ agrees with ∂_z on the complement of N_ϵ and flows $\mathcal{D}_j^{\text{ex}}$ into $\mathcal{D}_j^{\text{en}}$. The orbit γ must pass through N_ϵ as otherwise $z(\gamma)$ would take on

arbitrarily large values, implying that γ is not closed. If when passing through some component $N_{\epsilon,i}$ of the surgery handles γ exits the top of $N_{\epsilon,i}$ in the complement of the $\mathcal{D}_j^{\mathrm{ex}}$, then, again, $z(\gamma)$ would tend to ∞ as we follow the trajectory of the orbit. Likewise, if γ enters some $N_{\epsilon,i}$ in the complement of the $\mathcal{D}_j^{\mathrm{en}}$, then, following the orbit backwards in time, we see that $z(\gamma)$ is unbounded from below. □

Then γ must intersect some nonempty finite collection of the $\mathcal{D}_j^{\mathrm{ex}}$. Let $j_1, \dots, j_n$ be the indices of the $\mathcal{D}_j^{\mathrm{ex}}$ through which γ passes, ordered in accordance with a parametrization of γ.

Definition 5.7 We define the *cyclic word map* as

$$\mathrm{cw}(\gamma) = r_{j_1} \cdots r_{j_n}$$

and write $\mathrm{wl}(\gamma)$ for the word length of $\mathrm{cw}(\gamma)$.

5.3 Reeb orbits from cyclic words

In this section we describe how a cyclic word of composable Reeb chords uniquely determines an closed orbit of R_ϵ. Let $r_{j_1} \cdots r_{j_n}$ be a cyclic word and consider the squares $\mathcal{D}_{j_k}^*$, $* = \mathrm{en}, \mathrm{ex}, \tau$ as described in the previous subsection.

Theorem 5.8 *For $\epsilon \leq \epsilon_0$ and each word $w = r_{j_1} \cdots r_{j_n}$, there is a unique closed Reeb orbit γ_w of R_ϵ for which $\mathrm{cw}(\gamma_w) = w$.*

Our logic follows directly from arguments in [7, Section 6.1] — carried out in detail in [18] — which are simplified by our reduction of dynamics to that of affine transformations in Section 5.1.

Proof The proof follows from an analysis of $\mathrm{Flow}_{R_\epsilon}^t$ applied to the disk $\mathcal{D}_{j_1}^{\mathrm{ex}}$. Recall that this disk is contained in the "top" of a surgery handle $N_{\epsilon,l_{j_1}^-}$.

Write $S_1 = \mathcal{D}_{j_1}^{\mathrm{ex}}$ and let $G_1 = \mathrm{Id}_{S_1}$. Consider the following iterative process, for which Figure 10 serves as a visual aid:

(1) **Flow through the handle complement** There is a function $t(p,q)$ solving for the minimal $t > 0$ such that $\mathrm{Flow}_{R_\epsilon}^{t(p,q)}$ applied to $(p,q) \in S_1$ is an element of the square $\mathcal{D}_{j_1}^{\mathrm{en}}$ directly above S_1. Write $F_1^{\mathrm{co}}(p,q) = \mathrm{Flow}_{R_\epsilon}^{t(p,q)}(p,q)$, whose image is the square S_1', which is contained in the bottom of $N_{\epsilon,l_{j_1}^+}$. By the results of Section 5.1, $S_1' = \mathcal{D}_{j_1}^{\mathrm{en}}$. Briefly, F_1^{co} is the flow of our square $\mathcal{D}_{j_1}^{\mathrm{en}}$ through the handle complement.

(2) **Flow through the handle** Similarly define a function F_1^h which flows $S_1' \subset \{z = -\epsilon\} \subset N_{\epsilon,l_{j_1}^+}$ up to the top, $\{z = \epsilon\}$, of the surgery handle. The square $F^h(S_1')$

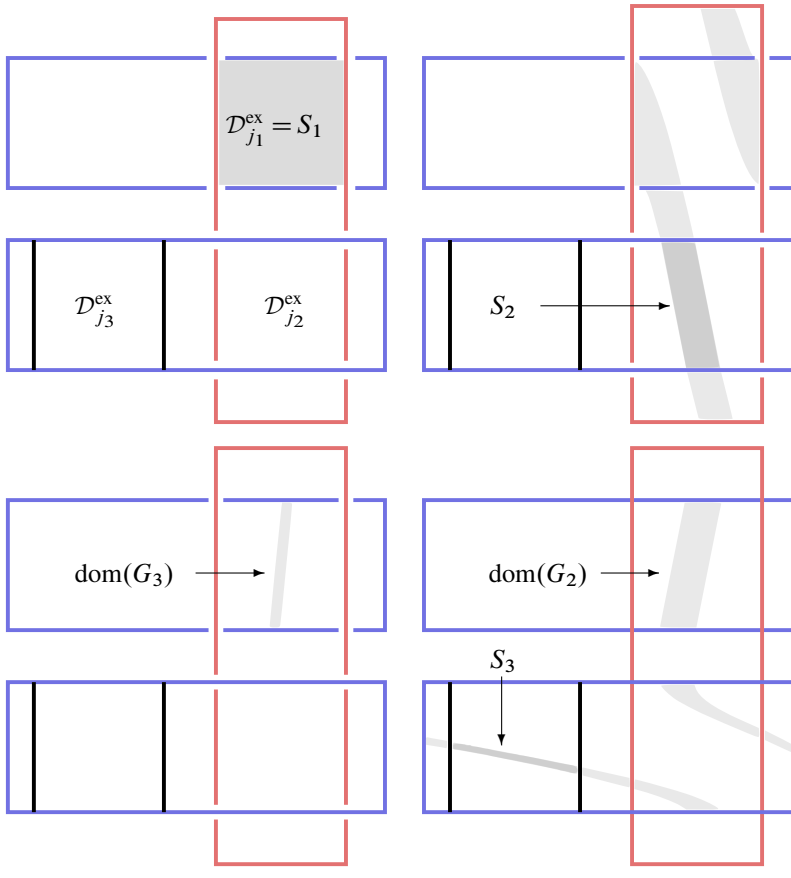


Figure 10: Following the subfigures clockwise we see the sets S_k and $\mathrm{dom}(G_k)$ drawn schematically. The S_k are shaded dark gray. Each rectangle represents the ribbon of some component of Λ cut at some value of q, layered over each other as indicated by the crossings so that the value of z increases as we traverse each subfigure clockwise. Within each rectangle, the sides of shorter (resp. longer) length are directed by p (resp. q). In the top-right we see $F_1^h \circ F_1^{\mathrm{co}}(S_1)$ as a subset of the top of $N_{\epsilon,l^+_{j_1}} = N_{\epsilon,l^-_{j_2}}$. Taking the intersection of this set with $\mathcal{D}^{\mathrm{ex}}_{j_2}$ determines S_2.

will appear in the coordinates (z, p, q) on the "outside" of the surgery handle as the application of a (positive or negative) Dehn twist to S_1'. That is, in the notation of Section 5.1, $F_1^h(S_1') = \mathcal{D}^{\tau}_{j_2}$ is the flow through the handle.

(3) **Trim** We write $S_2 = F_1^h(S_1') \cap \mathcal{D}^{\mathrm{ex}}_{j_2}$ for the intersection of $F_1^h(S_1')$ with the next square in the sequence $\mathcal{D}^{\mathrm{ex}}_{j_k}$ determined by w. Then S_2 is contained in the top of $N_{\epsilon,l^-_{j_2}}$.

We get a diffeomorphism G_2 from $\mathrm{dom}(G_2) = (F_1^h \circ F_1^{\mathrm{co}})^{-1}(S_2) \subset S_1$ to $\mathrm{im}(G_2) = S_2$ by $G_2 = F_1^h \circ F_1^{\mathrm{co}}$.

(4) **Repeat** We now inductively repeat the process by applying it to $S_k \subset \mathcal{D}^{\mathrm{ex}}_{j_k}$. We analogously define F_k^{co} and F_k^h with domain S_k then apply $F_k^h \circ F_k^{\mathrm{co}}$ to flow S_k up through the next handle in the sequence $N_{\epsilon,l^+_{j_k}}$ whose image we trim to define S_{k+1}. This determines a diffeomorphism

$$G_{k+1} = F_k^h \circ F_k^{\mathrm{co}} \circ \cdots \circ F_1^h \circ F_1^{\mathrm{co}} \colon (\mathrm{dom}(G_{k+1}) \subset S_1) \to (S_{k+1} \subset \mathcal{D}^{\mathrm{ex}}_{j_{k+1}}).$$

Making use of the results in Section 5.1, we have the following observations:

(1) Each F_k^{co} — considered with the domain $\mathcal{D}^{\mathrm{ex}}_{j_k}$ in which S_k is contained — is an affine transformation with respect to the (p,q) coordinates of the components of N_ϵ. Each F_k^{co} sends $\mathcal{D}^{\mathrm{ex}}_{j_k}$ diffeomorphically to $\mathcal{D}^{\mathrm{en}}_{j_k}$ and is a symplectomorphism with respect to $d\alpha_\epsilon$.

(2) Each F_k^h — considered with the domain $\mathcal{D}^{\mathrm{en}}_{j_k}$ — is nonlinear, as can by seen by looking at where p is extremal. It is also a symplectomorphism with respect to $d\alpha_{\mathrm{std}}$. The restriction of F_k^h to $(F_k^h)^{-1}(\mathcal{D}^{\mathrm{ex}}_j)$ for each j is an affine transformation by Properties 4.4.

(3) We see by induction that S_k is a connected, nonempty quadrilateral determined by a nondegenerate pair of linear inequalities, one of which is of the form $q \in [q_0 - \delta, q_0 + \delta]$.

(4) Combining the above with the fact that a composition of affine transformations is an affine transformation, $\mathrm{dom}(G_k)$ is a quadrilateral determined by a pair of linear inequalities, one of which is the trivial $p \in [-\epsilon, \epsilon]$.

(5) Each trimming step monotonically decreases the area with respect to $d\alpha_\epsilon$ and, for each k, we have $\mathrm{dom}(G_{k+1}) \subsetneq \mathrm{dom}(G_k)$:

$$0 < \int_{S_{k+1}} d\alpha_\epsilon < \int_{S_k} d\alpha_\epsilon, \quad 0 < \int_{\mathrm{dom}(G_{k+1})} d\alpha_\epsilon < \int_{\mathrm{dom}(G_k)} d\alpha_\epsilon.$$

Now observe that $\mathrm{dom}(G_n)$ stretches across $\mathcal{D}^{\mathrm{ex}}_{j_1}$ in the p direction and that S_n stretches across $\mathcal{D}^{\mathrm{ex}}_{j_1}$ in the q direction. Since both sets are convex, $U_1 = \mathrm{dom}(G_n) \cap S_n$ must be nonempty and convex. We likewise define U_k as the intersection of S_{nk} with $\mathrm{dom}(G_{nk})$ for all $k > 0$. See Figure 11.

The U_k satisfy $U_{k+1} \subsetneq U_k$ and we claim that $\mathrm{area}(U_k) \to 0$ as $k \to \infty$. To see this, recall that f_ϵ is linear on $p \in [-\epsilon(1-\delta), \epsilon(1-\delta)]$, where $\delta \in (0,1)$ is as described in Assumptions 5.4. By the conditions which characterize δ, for each $k \geq 0$ the set of points in S_k which reach S_{k+1} via the map $F_k^h \circ F_k^{\mathrm{co}}$ must be contained in the set

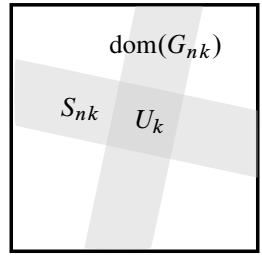


Figure 11: Overlaps of the sets S_{nk} and $\mathrm{dom}(G_{nk})$ within in the set $S_1 = \mathcal{D}^{\mathrm{ex}}_{j_1}$.

$S^{\mathrm{lin}}_k = \{p \in [-\epsilon(1-\delta), \epsilon(1-\delta)]\} \cap S_k$. By the fact that S^k is a rectangle stretching across the p coordinate of the annulus, $\mathrm{area}(S^{\mathrm{lin}}_k) = (1-\delta)\,\mathrm{area}(S_k)$. By the definition of S^{lin}_k and the fact that $F^h_k \circ F^{\mathrm{co}}_k$ is symplectic, $\mathrm{area}(S_{k+1}) \leq \mathrm{area}(S^{\mathrm{lin}}_k)$. Inductively, we conclude $\mathrm{area}(S_k) \leq (1-\delta)^{k-1}\,\mathrm{area}(S_1)$. Since U_k is contained in S_{nk}, our claim is established.

By our construction, any Reeb orbit with word w^k must intersect $\mathcal{D}^{\mathrm{ex}}_{j_1}$ at a point in $\mathrm{dom}(G_{nk})$ which is sent to itself via G_{nk}. Hence, such a point of intersection must lie in U_k. By considering multiple covers of the orbit γ_w — whose existence we seek to establish — we see that, if such a point of intersection lies in U_1, then it must lie in U_k for all $k > 0$. We therefore define

$$U_\infty = \bigcap_1^\infty U_k \subset \mathcal{D}^{\mathrm{ex}}_{j_1},$$

which, by our previous analysis, consists of a single point.

To complete our proof, it suffices to show that $G_{nk}(U_\infty) = U_\infty$ for all $k > 0$. This amounts to unwinding the definitions established in the proof so far. If we write $k = k_1 + k_2$ for any pair of natural numbers k_1 and k_2, then we must have

$$G_{nk_1}(\mathrm{dom}(G_{nk})) \subset \mathrm{dom}(G_{nk_2})$$

as otherwise $G_{nk_2} \circ G_{nk_1}(\mathrm{dom}(G_{nk}))$ would not be contained in $\mathcal{D}^{\mathrm{ex}}_{j_1}$. On the other hand, $\mathrm{dom}(G_{nk}) \subsetneq \mathrm{dom}(G_{nk_1})$ implies that

$$G_{nk_1}(\mathrm{dom}(G_{nk})) \subsetneq S_{nk_1}.$$

Combining the above two equations, we conclude that

$$U_\infty = \bigcap_1^\infty (S_{nk} \cap \mathrm{dom}(G_{nk})) = \left(\bigcap_1^\infty S_{nk}\right) \cap \left(\bigcap_1^\infty \mathrm{dom}(G_{nk})\right)$$

satisfies $G_{nk}(U_\infty) = U_\infty$. □

5.4 Reeb chords of Λ^0 after surgery

In this section, we describe open-string versions of our results for closed Reeb orbits, establishing the chord-to-chord correspondence of Theorem 5.10.

Definition 5.9 Suppose that a chord κ of $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ passes through a sequence of the $\mathcal{D}^*_j$ of the form

$$\mathcal{D}^{\text{en}}_{j_1}, \mathcal{D}^{\text{ex}}_{j_2}, \mathcal{D}^{\text{en}}_{j_2}, \dots, \mathcal{D}^{\text{ex}}_{j_{n-1}}, \mathcal{D}^{\text{en}}_{j_{n-1}}, \mathcal{D}^{\text{ex}}_{j_n}.$$

Then we write $\mathrm{w}(\kappa) = r_{j_1} \cdots r_{j_n}$. We call the association $\kappa \mapsto \mathrm{w}(\kappa)$ the *word map*.

Theorem 5.10 *For each $\epsilon < \epsilon_0$, the word map* w *determines a one-to-one correspondence between words of chords with boundary on $\Lambda^0 \subset (\mathbb{R}^3, \xi_{\text{std}})$ and Reeb chords of $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ determined by the contact form α_ϵ. For each such word w, the associated chord κ_w is nondegenerate for all $\epsilon < \epsilon_0$.*

The chords with word length 1 are those which exist for $\Lambda^0 \subset \mathbb{R}^3$ prior to surgery, while the rest of the chords in Theorem 5.10 are created after the performance of surgery along $\Lambda^\pm$.

Proof The proof is analogous to the proof of Theorem 5.8, although considerably simpler.

Let $w = r_{j_1} \cdots r_{j_n}$ be a word of chords on Λ^0 with word length $n > 1$. By (16), flowing Λ^0 up to N_ϵ along a chord sends Λ^0 to a strand in N_ϵ of the form $q = q_0$, which we call A_1'. Flow this arc up to the top of N_ϵ and take its intersection with $\mathcal{D}^{\text{ex}}_{j_2}$ to obtain an arc we'll call A_1. Define arcs A_k for $k > 1$ as follows:

(1) **Flow through the handle complement** Flow $A_{k-1} \subset \mathcal{D}^{\text{ex}}_{j_k}$ up to $\mathcal{D}^{\text{en}}_{j_k}$ using the map F^{co}_k as in the proof of Theorem 5.8.

(2) **Flow through the handle** Now we apply the map F^h_k as defined in Theorem 5.8 to flow $F^{\text{co}}_k(A_{k-1})$ up to the top of N_ϵ.

(3) **Trim** Define $A_k = F^h_k \circ F^{\text{co}}_k(A_{k-1}) \cap \mathcal{D}^{\text{ex}}_{j_k}$.

(4) **Repeat** Repeat the above steps until we obtain an arc $A_n \subset \mathcal{D}^{\text{ex}}_{j_n}$.

Again, following the logic of the proof of Theorem 5.8 using the linearity conditions of Section 5.1, each $A_k \subset \mathcal{D}^{\text{ex}}_{j_k}$ is a line segment which wraps across $\mathcal{D}^{\text{ex}}_{j_k}$ in the q direction. In other words, each admits a parametrization of the form

$$A_k = \{(aq + b, q) : q \in [q_0 - \delta, q_0 + \delta]\}$$

for some constants $a \neq 0, b, q_0, \delta$. Since flowing Λ^0 downward to $\mathcal{D}^{\mathrm{ex}}_{j_n}$ along the chord r_{j_n} is a set of the form $q = q_0$, the intersection of this set with A_k consists of a single point. Since A_k wraps across $\mathcal{D}^{\mathrm{ex}}_{j_n}$ in the q direction, this intersection is transverse. By construction, the collection of such intersections are in one-to-one correspondence with the collection of chords of $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$.

For words of length 1, the restriction of R_ϵ to the complement of the surgery handles is ∂_z, so that words of length 1 correspond exactly to the chords of Λ^0 present prior to surgery. □

5.5 Action estimates

To obtain refined estimates of the actions of the chords and orbits of R_ϵ we'll need the following lemmas. The first lemma tells us how much time it takes to flow from the top of N_ϵ to the bottom in a neighborhood of a chord r_j.

Lemma 5.11 *Let r_j be some chord of $\Lambda \subset (\mathbb{R}^3, \xi_{\mathrm{std}})$ with action $\mathcal{A}(r_j)$ and parametrize the disk $\mathcal{D}^{\mathrm{ex}}_j \subset \partial N_\epsilon$ with coordinates (p, q) as in (17). Then, for each $(p, q) \in \mathcal{D}^{\mathrm{en}}_j$, there exists a minimal-length chord from $\mathcal{D}^{\mathrm{ex}}_j$ to $\mathcal{D}^{\mathrm{en}}_j$ starting at (P, Q) with action*

$$t = \mathcal{A}(r_j) - 2\epsilon - pq.$$

Proof This is a straightforward calculation, so we omit the details. For a given j, write $(p^{\mathrm{ex}}, q^{\mathrm{ex}})$ and $(p^{\mathrm{en}}, q^{\mathrm{en}})$ for the coordinates on $\mathcal{D}^{\mathrm{ex}}_j$ and $\mathcal{D}^{\mathrm{en}}_j$ provided by (17) and (16), respectively. Then $p^{\mathrm{en}} = -q^{\mathrm{ex}}$ and $q^{\mathrm{en}} = p^{\mathrm{ex}}$. Plug these into the equations provided to compute the differences in the z coordinates and consider the fact that $R_\epsilon = \partial_z$ on $\mathbb{R}^3 \setminus N_\epsilon$. □

Our second lemma tells how much time it takes for an orbit to flow through one of the surgery handles.

Lemma 5.12 *For some j, again consider coordinates (p, q) on $\mathcal{D}^{\mathrm{ex}}_j \subset \partial N_\epsilon$ as provided by (17). Then the time it takes a point in $\mathcal{D}^{\mathrm{en}}_j$ to reach this point via the flow of R_ϵ is*

$$t = 2\epsilon + c_i H_\epsilon(p).$$

This becomes obvious if we look at the graph of the "top" part of the gluing map of (22). See Figure 12. By comparing Proposition 4.9 with the definition of the gluing map in (22), actions increase slightly as we pass through a surgery handle with coefficient -1 and decrease slightly as we pass through a surgery handle with coefficient $+1$.

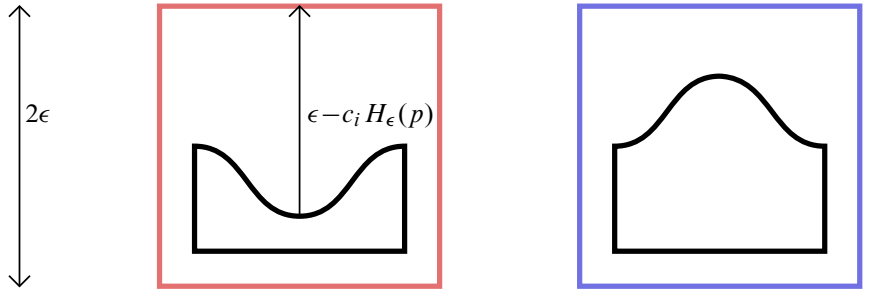


Figure 12: The squares represent $\{q = q_0\} \subset N_{\epsilon,i}$ slices of the N_ϵ at components $N_{\epsilon,i}$ with surgery coefficient $c_i = -1$ (left) and $c_i = 1$ (right). The black arcs represent the boundaries of the gluing region, as it intersects each slice.

Proposition 5.13 *For all closed Reeb orbits γ of R_ϵ, we have*

$$|\mathcal{A}(\gamma) - \mathcal{A}(\mathrm{cw}(\gamma))| < 3\epsilon\, \mathrm{wl}(\gamma).$$

For each chord r of R_ϵ with boundary on $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$, we have

$$|\mathcal{A}(r) - \mathcal{A}(\mathrm{w}(\gamma))| < 3\epsilon\, \mathrm{wl}(r).$$

This is obvious from Lemmas 5.11 and 5.12 together with Proposition 4.9.

5.6 Calculating orbit embeddings

Let $\gamma = (r_{j_1} \cdots r_{j_n})$ be a closed orbit of some R_ϵ. Let (p_k, q_k) be coordinates on the squares $\mathcal{D}^{\mathrm{ex}}_{j_k}$ described by (16). Suppose that, in these coordinates, γ passes through the points (P_k, Q_k). If γ is simply covered and we compute the exact values of the (P_k, Q_k), then we can see the knot formed by γ inside of $\mathbb{R}^3_{\Lambda^\pm}$ and be able to compute the action $\mathcal{A}(\gamma)$ exactly. In this section, we describe how these (P_k, Q_k) can be calculated. The analysis here will be the starting point for the computation of Conley–Zehnder indices.

In the above notation, we can describe (P_1, Q_1) as a fixed point of an affine transformation

$$A + b\colon \mathbb{R}^2 \to \mathbb{R}^2, \quad A \in \mathrm{SL}(2, \mathbb{R}),\ b \in \mathbb{R}^2,$$

as follows.

Starting at (a subset of) $\mathcal{D}^{\mathrm{ex}}_{j_k}$, apply $\mathrm{Flow}_{R_\epsilon}$ to pass through the handle complement to $\mathcal{D}^{\mathrm{en}}_{j_k}$ and then through the surgery handle $N_{\epsilon,l^+_{j_k}}$ to $\mathcal{D}^{\mathrm{ex}}_{j_{k+1}}$. As we are only interested in the set of points in $\mathcal{D}^{\mathrm{ex}}_{j_k}$ along which the $\tau_{\pm f_\epsilon}$ are linear, we can write this as a map

$A_k + b_k$ with $A_k \in \mathrm{SL}(2,\mathbb{R})$. The $b_k \in \mathbb{R}^2$ term is required by (16) centering the q coordinate about the endpoint of the Reeb chord r_{k+1}.

Hence, we may write

$$\begin{aligned} A + b &= (A_n + b_n) \circ \cdots \circ (A_1 + b_1) \\ &= (A_n \cdots A_1) + (A_n \cdots A_2) b_1 + \cdots + A_n b_{n-1} + b_n \end{aligned} \tag{25}$$

with (P_1, Q_1) being the fixed point of this map. By linearity of the equations involved and our prior knowledge (Theorem 5.8) that there exists a unique fixed point, we may as well consider the $A_k + b_k$ to be transformations of $\mathbb{R}^2$. We can then solve for $u_1 = (P_1, Q_1)$ as

$$u_1 = (\mathrm{Id} - A)^{-1} b = (\mathrm{Id} - A_n \cdots A_1)^{-1} (A_n \cdots A_2 b_1 + \cdots + A_n b_{n-1} + b_n).$$

Provided u_1, we can then find the $u_k = (P_k, Q_k)$ by applying the $(A_k + b_k)$:

$$u_{k+1} = A_k u_k + b_k = (A_k + b_k) \cdots (A_1 + b_1) u_1.$$

Proposition 5.14 *In the above notation,*

$$\begin{aligned} (-1)^{\mathrm{rot}_{j_k, j_{k+1}}} A_k &= \begin{pmatrix} 0 & -1 \\ 1 & -c_{j_k}/\epsilon \end{pmatrix} = J_0 \begin{pmatrix} 1 & -c_{j_k}/\epsilon \\ 0 & 1 \end{pmatrix}, \\ (-1)^{\mathrm{rot}_{j_k, j_{k+1}}} b_k &= \begin{pmatrix} 0 \\ \frac{1}{2} - d_{j_k, j_{k+1}} \end{pmatrix} = J_0 \begin{pmatrix} \frac{1}{2} - d_{j_k, j_{k+1}} \\ 0 \end{pmatrix}, \end{aligned}$$

where $d_{j_k, j_{k+1}}$ is the minimal length of a capping path for the pair $(r_j, r_{j_{k+1}})$ projected to the xy–plane using the standard Euclidean metric on $\mathbb{R}^2$.

Proof We can determine $A_k + b_k$ as a composition of the following elementary mappings:

(1) The change of coordinates from $\mathcal{D}^{\mathrm{ex}}_{j_k}$ to $\mathcal{D}^{\mathrm{en}}_{j_k}$, which we see when flowing points (p, q) through the handle complement,

$$(p, q) \mapsto (-q, p).$$

(2) The flow from $\mathcal{D}^{\mathrm{en}}_{j_k}$ to the top of $N_{\epsilon, l^+_{j_k}}$, which, according to Properties 5.5, is given by

$$(p, q) \mapsto \left(p, q + \frac{1}{2} + \frac{c^+_{j_k}}{\epsilon} p \right).$$

(3) A shift in the q coordinate such that $(0, 0)$ is identified with the tail of $r_{j_{k+1}}$. Here $d^{\mathrm{in}}_{k,k+1}$ is the magnitude of this shift when Λ_{j_k} is parametrized with Φ_{i+},

as in Proposition 4.5,

$$(p,q)\mapsto(p,q-d^{\text{in}}_{k,k+1}).$$

(4) A mapping of the coordinates on the top of $N_{\epsilon,l^+_{j_k}}$ to $\mathcal{D}^{\text{ex}}_{k+1}$,

$$(p,q)\mapsto(-1)^{\text{rot}_{j_k,j_{k+1}}}(p,q).$$

The result of composing the above maps produces

$$(p,q)\mapsto(-1)^{\text{rot}_{j_k,j_{k+1}}}\left(-q,\,p+\frac{1}{2}-d^{\text{in}}_{k,k+1}-\frac{c^+_{j_k}}{\epsilon}q\right). \qquad \square$$

5.7 Hyperbolicity and the $\mathbb{Z}/2\mathbb{Z}$ index

For a given closed orbit $\gamma=(r_{j_1}\cdots r_{j_n})$, we can use the above formula to write its Poincaré return map as $\text{Ret}_\gamma=A_n\cdots A_1$, where the A_k are given by (25). By using the calculation of the A_k in Proposition 5.14, we have an explicit representation of Ret_γ as

$$\begin{aligned}(-1)^{\text{rot}}\,\text{Ret}_\gamma&=\prod_{K=1}^{n}J_0\begin{pmatrix}1&-c^+_{j_{n+1-K}}\epsilon^{-1}\\0&1\end{pmatrix}\\&=J_0\begin{pmatrix}1&-c^+_{j_n}\epsilon^{-1}\\0&1\end{pmatrix}\cdots J_0\begin{pmatrix}1&-c^+_{j_1}\epsilon^{-1}\\0&1\end{pmatrix}\\&=J_0^n+\sum_{K=1}^{n}\left(\sum_{k\in I_K}\left(\prod_{i=1}^{K}-c^+_{j_{k_i}}\right)M_k\right)\epsilon^{-K},\qquad(26)\\ \text{rot}&=\sum_{K=1}^{n}\text{rot}_{j_K,j_{K+1}},\\M_k&=J_0^{n-k_K}\,\text{Diag}(0,1)J_0^{k_K-k_{K-1}-1}\cdots J_0^{k_2-k_1-1}\,\text{Diag}(0,1)J_0^{k_1-1},\\I_K&=\{k=(k_1,\dots,k_K):1\le k_1<\cdots<k_K\le n\}.\end{aligned}$$

The equality in the third line involving the M_k easily follows from an induction on n.

Observe that I_n consists of a single element $(1,\dots,n)$, so that the $K=n$ term in the above formula is

$$\epsilon^{-n}\left(\prod_{k=1}^{n}-c^+_{j_k}\right)\text{Diag}(0,1)=\epsilon^{-n}(-1)^{\#(c^+_{j_k}=1)}\,\text{Diag}(0,1). \qquad (27)$$

Thus, for a fixed word, $\text{tr}(\text{Ret}_\gamma)$ can be expressed as a polynomial in ϵ^{-1} whose highest-order term is given by the above expression.

Proof of Theorem 5.3 For ϵ_w sufficiently small, the ϵ^{-n} term in the polynomials for $\mathrm{tr}(\mathrm{Ret}_\gamma)$ determines their sign for all $\epsilon < \epsilon_w$ and words of length $\leq n$ as there are only finitely many cyclic words less than a given length. Possibly making ϵ_w smaller, we can guarantee that the absolute values of the traces are bounded below by 2. To compute CZ_2, we apply (1) and (3), noting that $\det(\mathrm{Ret} - \mathrm{Id}) = 2 - \mathrm{tr}(\mathrm{Ret})$.

If one of Λ^+ or Λ^- is empty, then each orbit of word length n has return map

$$\mathrm{Ret}_\gamma = \pm M_a^n, \quad M_a = \begin{pmatrix} 0 & -1 \\ 1 & a \end{pmatrix}, \quad a = \pm\epsilon^{-1}.$$

If $\epsilon < \frac{1}{2}$, then M_a is hyperbolic and so is conjugate to $\mathrm{Diag}(\lambda, \lambda^{-1})$ with

$$\lambda = \tfrac{1}{2}(a + \sqrt{a^2 - 4}), \quad \lambda^{-1} = \tfrac{1}{2}(a - \sqrt{a^2 - 4}),$$

implying that Ret_γ is conjugate to $\pm\,\mathrm{Diag}(\lambda^n, \lambda^{-n})$. In this case, it's clear that $|\mathrm{tr}(\mathrm{Ret}_\gamma)| > 2$ independent of n, implying that all closed orbits of R_ϵ are hyperbolic for $\epsilon < \frac{1}{2}$. □

6 The semiglobal framing (X, Y)

Having computed the $\mathbb{Z}/2\mathbb{Z}$ Conley–Zehnder indices of the closed Reeb orbits of R_ϵ, we now seek to compute $\mathbb{Z}$–valued indices with respect to a framing as well as Maslov indices of broken closed strings on $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$.

In this section we describe sections of $\xi_{\Lambda^\pm}$, which we will later use to compute these indices. This will allow us to draw a cycle representing $\mathrm{PD}(c_1(\xi_{\Lambda^\pm})) = \mathrm{PD}(e(\xi_{\Lambda^\pm}))$ as a link in the Lagrangian projection: See Figure 13 for an example. The results of this section are summarized as follows:

Theorem 6.1 *For each $\epsilon < \epsilon_0$ there are sections $X, Y \in \Gamma(\xi_\Lambda)$ such that the following conditions hold*:

(1) $(X, Y) = (\partial_x + y\partial_z, \partial_y)$ *on* $\mathbb{R}^3_{\Lambda^\pm} \setminus N_\epsilon \simeq \mathbb{R}^3 \setminus N_\epsilon$.

(2) (X, Y) *is a symplectic basis of* $(\xi_{\Lambda^\pm}, d\alpha_\epsilon)$ *at each point contained in a closed Reeb orbit of* R_ϵ.

(3) $X^{-1}(0) = Y^{-1}(0)$ *is a union of connected components of* $\bigcup_i T_i^{c_i}$, *where the* $T_i^\pm$ *are the transverse push-offs of the* Λ_i *as described in Definition 4.7.*

Using the (X, Y), the first Chern class of $\xi_{\Lambda^\pm}$ may be computed as

$$\mathrm{PD}(c_1(\xi_{\Lambda^\pm})) = \sum_{\Lambda_i \subset \Lambda^\pm} -c_i \,\mathrm{rot}(\Lambda_i)[T_i^{c_i}] = \sum_{\Lambda_i \subset \Lambda^\pm} \mathrm{rot}(\Lambda_i)\mu_i \in H_1(\mathbb{R}^3_{\Lambda^\pm}).$$

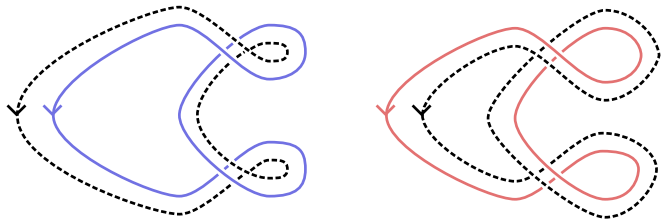


Figure 13: Here we consider contact ± 1 surgery on the Legendrian unknot with $\mathrm{rot}(\Lambda) = 1$. In each case, Theorem 6.1 provides a framing of ξ_Λ on the complement of a transverse push-off of Λ which travels along the right-hand side of Λ when the surgery coefficient is $+1$ and along the left side of Λ when the coefficient is -1. These push-offs are depicted as the dashed, black circles.

The classes μ_i are given by a standard presentation of $H_1(\mathbb{R}^3_{\Lambda^\pm})$ determined by the surgery diagram, which we will describe in Section 9.

Theorem 6.1 may be compared with [31, Proposition 2.3], where a similar result is stated for Chern classes integrated over 2–cycles in Stein surfaces, and with [27, Section 3], where Chern classes are computed when performing surgery along Legendrians lying in pages of open book decompositions.

For notational simplicity, we assume throughout this section that $\Lambda^\pm$ has a single connected component unless otherwise stated. Accordingly, we temporarily drop the indices i appearing in the notation of Section 3. The surgery coefficient of this knot will be denoted by c.

Our framing is constructed in three steps:

(1) We start with a framing of ξ_{std} over the complement of N_ϵ and express it in terms of our local coordinate system (z, p, q) along the boundary of our surgery handles.

(2) Next, we describe an explicit extension of this framing throughout most of the handle. We will need this explicit description to compute Conley–Zehnder and Maslov indices in Section 7.

(3) Finally, we describe the zero locus of this extension.

6.1 Change of bases between trivializations

Consider the following pairs of sections of ξ_{std} and $\xi_{\Lambda^\pm}$, which form symplectic bases:

$$\{X = \partial_x + y\partial_x,\ Y = \partial_y\},\quad \{P_{\text{in}} = \partial_p,\ Q_{\text{in}} = \partial_q - p\partial_z\},\quad \{P_{\text{out}} = \partial_p,\ Q_{\text{out}} = \partial_q - p\partial_z\}.$$

These come from

(1) the coordinate systems (x, y, z) on $\mathbb{R}^3$,

(2) the coordinates (z, p, q) on N_ϵ viewed "from the outside" of the surgery handle prior to surgery, and

(3) (z, p, q) on N_ϵ viewed "from the inside" of the surgery handle after surgery,

respectively. After performing surgery, the pairs (X, Y) and $(P_{\text{in}}, Q_{\text{in}})$ are well defined on the complement of a neighborhood of the form $N_{\epsilon'}$ for some $\epsilon' < \epsilon$. Our strategy will be to apply a series of change-of-basis transformations to extend the framing (X, Y) of $\xi_{\Lambda^\pm}$ throughout the surgery handle in so far as cohomological obstruction — $c_1(\xi_{\Lambda^\pm})$ — will allow.

First we describe change of bases from $(P_{\text{in}}, Q_{\text{in}})$ to $(P_{\text{out}}, Q_{\text{out}})$. Following (23), the restriction of the tangent map of the gluing map $\phi_{c,f,\epsilon,\delta}$ — defined in (22) — to $\xi_{\Lambda^\pm}$ can be written as

$$T\phi_{c,f,\epsilon,\delta}(z, p, q)|_\xi = \begin{pmatrix} 1 & 0 \\ c\,\partial f_\epsilon/\partial p(p) & 1 \end{pmatrix} \tag{28}$$

along $T_{\delta,\epsilon}$ and as $\mathrm{Diag}(1, 1)$ along $B_{\delta,\epsilon} \cup S_{\delta,\epsilon}$. Here the incoming basis is $(P_{\text{in}}, Q_{\text{in}})$, the outgoing basis is $(P_{\text{out}}, Q_{\text{out}})$, and coordinates (z, p, q) correspond to the coordinate system inside of the surgery handle.

Now we describe change of bases from $(P_{\text{out}}, Q_{\text{out}})$ to (X, Y). To this end, let G be the Gauss map for a parametrization of Λ as described in Section 4.4. Using the construction of N_ϵ in Proposition 4.5, we can write the change of basis at a point (p, q, z) as

$$E(p, q)e^{J_0(G(q)-\pi/2)}, \tag{29}$$

where $E = \mathrm{Diag}(1, 1) + \mathcal{O}(p)$. Here the incoming basis is $(P_{\text{out}}, Q_{\text{out}})$, the outcoming basis is (X, Y), and coordinates (z, p, q) correspond to the coordinate system on "the outside" — the complement of the surgery handle in N_ϵ.

By composing the changes of bases described above in (28) and (29) and then inverting, we can write (X, Y) in the basis $(P_{\text{in}}, Q_{\text{in}})$ on a neighborhood of ∂N_ϵ as follows: Along $B_{\delta,\epsilon} \cup S_{\delta,\epsilon}$ the change of basis is given by

$$e^{J_0(\pi/2-G(q))}E^{-1}(p, q). \tag{30}$$

Along $T_{\delta,\epsilon}$ the transition map is

$$\begin{pmatrix} 1 & 0 \\ -c\,\partial f_\epsilon/\partial p(p) & 1 \end{pmatrix} e^{J_0(\pi/2-G(q+cf_\epsilon(p)))}E^{-1}(p, q + cf_\epsilon(p)). \tag{31}$$

Here the incoming basis is (X, Y), the outcoming basis is $(P_{\text{in}}, Q_{\text{in}})$, and coordinates (z, p, q) correspond to the coordinate system inside of the surgery handle. Then, where they are defined, equations (30) and (31) provide X and Y as a linear combination of P_{in} and Q_{in} by multiplying the above expressions on the left by $\binom{1}{0}$ and $\binom{0}{1}$, respectively.

6.2 Framing extension up to obstruction

We use the above equations to extend the framing (X, Y) of $\xi_{\Lambda^\pm}$ inside of the surgery handle. To this end, let $\delta > 0$ be an arbitrarily small constant and consider a smooth function $\nu\colon I_\epsilon \to [0, 1]$ with the following properties:

(1) $\nu(-\epsilon) = 0$ and $\nu(\epsilon) = 1$.

(2) All of its derivatives vanish outside of $I_{\epsilon-\delta}$.

When $c = 1$, we use (30) and (31) to extend the definitions of (X, Y) over the set $\{p < \epsilon - \delta\} \subset N_\epsilon$ using the family of matrices

$$\begin{pmatrix} 1 & 0 \\ -\partial f_\epsilon/\partial p(\zeta^+) & 1 \end{pmatrix} e^{J_0(\pi/2 - G(\eta^+))} E^{-1}(\zeta^+, \eta^+), \tag{32}$$

where

$$\zeta^+(z, p) = p\nu(z) - \epsilon(1 - \nu(z)), \quad \eta^+(z, p, q) = q + f_\epsilon(\zeta^+(z, p)).$$

Note that $\zeta^+ = -\epsilon$ along $\{z = -\epsilon\} \cup \{p = -\epsilon\}$ and that $\zeta^+ = p$ along $z = \epsilon$. By these properties and the properties of f_ϵ and its derivatives in Section 4.7, we have that this family of matrices agrees with (30) along $B_{\delta,\epsilon}$ and with (31) along $T_{\delta,\epsilon}$.

Likewise, when $c = -1$, we extend the definitions of (X, Y) over $\{p > -\epsilon + \delta\} \subset N_\epsilon$, using the family of matrices which provide (X, Y) in the basis $(P_{\text{in}}, Q_{\text{in}})$,

$$\begin{pmatrix} 1 & 0 \\ \partial f_\epsilon/\partial p(\zeta^-) & 1 \end{pmatrix} e^{J_0(\pi/2 - iG(\eta^-))} E^{-1}(\zeta^-, \eta^-), \tag{33}$$

where

$$\zeta^-(z, p) = p\nu(z) + \epsilon(1 - \nu(z)), \quad \eta^-(z, p, q) = q - f_\epsilon(\zeta^-(z, p)).$$

Note that $\zeta^- = \epsilon$ along $\{z = -\epsilon\} \cup \{p = \epsilon\}$ and that $\zeta^- = p$ along $z = \epsilon$. As in the $c = 1$ case, this family of matrices agrees with (30) along $B_{\delta,\epsilon} \cup S_{\delta,\epsilon}$ and with (31) along $T_{\delta,\epsilon}$.

The extension of the fields (X, Y) through the surgery handle N_ϵ is illustrated in Figure 14.

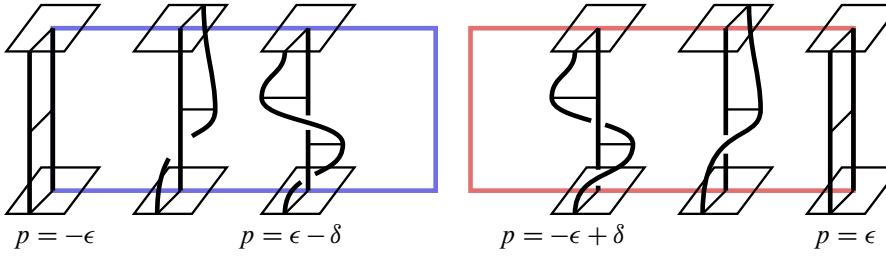


Figure 14: On the left we have the extension of the framing (X, Y) through the surgery handle over a square of the form $\{q = q_0,\, p < \epsilon - \delta\}$ in N_ϵ when $c = 1$ and $\mathrm{rot}(\Lambda) = 1$. On the right is the case $c = -1$, $\mathrm{rot}(\Lambda) = 1$. Here ∂_p points to the right, ∂_q points in to the page, and ∂_z points upwards. Along the bottom of the square, the framing is constant. At each p, the framing twists with respect to the trivialization $(\partial_p, \partial_q - p\partial_z)$ according to the twisting of Gauss map along the path in Λ from q_0 to $q_0 + f_\epsilon(p)$. For $c = 1$, moving from left to right, we eventually get to $p_0 = \epsilon - \delta$ such that $f_\epsilon(p) = 1$ for $p > p_0$.

6.3 Obstruction to global definition of (X, Y)

The Chern class $c_1(\xi_{\Lambda^\pm})$ agrees with the Euler class of $\xi_{\Lambda^\pm}$ and so can be represented as the zero locus of a generic section $s \in \Gamma(\xi_{\Lambda^\pm})$. In attempting to extend the definition of (X, Y) over the squares $S^+_{q_0,\epsilon,\delta} := \{q = q_0, \epsilon - \delta \le p \le \epsilon\}$ when $c = 1$ and $S^+_{q_0,\epsilon,\delta} := \{q = q_0, -\epsilon \le p \le -\epsilon + \delta\}$ when $c = -1$, we may complete the proof of Theorem 6.1.

Proof of Theorem 6.1 We attempt to extend X throughout the entirety of the handle, assuming that δ is small enough so that f_ϵ is constant on each component of $[-\epsilon, -\epsilon + \delta) \cup (\epsilon - \delta, \epsilon]$. For the case of $+1$ contact surgery we study (32). We orient $S^+_{q_0,\epsilon,\delta}$ so that ∂_q points positively through it. Parametrize the oriented boundary of each $S^+_{q_0,\epsilon,\delta}$ with a piecewise-smooth curve $\gamma = \gamma(t)$ so that

$$\frac{\partial \gamma}{\partial t} = \begin{cases} \partial_z & \text{if } p = \epsilon - \delta, \\ \partial_p & \text{if } z = \epsilon, \\ -\partial_z & \text{if } p = \epsilon, \\ -\partial_p & \text{if } z = \epsilon. \end{cases}$$

Applying the vector $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ to the left of (31) gives us the section X as a linear combination of P_{in} and Q_{in} along γ. By throwing away the shearing and rescaling terms in (32), this section is homotopic through nonvanishing sections of $\xi_{\Lambda^\pm}$ to a section of the form

$$t \mapsto \begin{cases} e^{-J_0 G(q_0 + f_\epsilon(t\epsilon))} & \text{for } p = \epsilon - \delta, t \in [-1, 1], \\ e^{-J_0 G(q_0 + f_\epsilon(\epsilon))} & \text{on } \{z = -\epsilon\} \cup \{p = \epsilon\} \cup \{z = \epsilon\}. \end{cases} \tag{34}$$

This is homotopic to

$$t \mapsto e^{\text{const}-2\pi i t\, \text{rot}(\Lambda)}, \quad t \in [0,1].$$

Therefore, a generic extension of X over each $S^+_{q_0,\epsilon,\delta}$ will have $-\text{rot}(\Lambda)$ zeros counted with multiplicity. Taking a generic extension of X over $\{p > \epsilon - \delta\}$ will then be an oriented link which transversely intersects each square with multiplicity $\text{rot}(\Lambda)$. Pushing this zero locus through the side $p = \epsilon$ of the surgery handle provides $\text{PD}(c_1(\xi_{\Lambda^\pm})) = -\text{rot}(\Lambda)\lambda_\xi$.

The case $c = -1$ is similar; we only check signs. Consider a parametrization of the boundary of the square $S^-_{q_0,\epsilon,\delta}$ with a loop γ satisfying

$$\frac{\partial \gamma}{\partial t} = \begin{cases} -\partial_z & \text{if } p = -\epsilon + \delta, \\ -\partial_p & \text{if } z = -\epsilon, \\ \partial_z & \text{if } p = -\epsilon, \\ \partial_p & \text{if } z = \epsilon. \end{cases}$$

Then, following (33), the analog of (34) for the $c = -1$ case is

$$t \mapsto \begin{cases} e^{-J_0 G(q_0 - f_\epsilon(t\epsilon))} & \text{for } p = -\epsilon + \delta, t \in [-1,1], \\ e^{-J_0 G(q_0 - f_\epsilon(\epsilon))} & \text{on } \{z = \epsilon\} \cup \{p = -\epsilon\} \cup \{z = -\epsilon\}, \end{cases}$$

so that the zero locus of the extension of the vector field X throughout the handle is homologous to $\text{rot}(\Lambda)\lambda_\xi$. □

Remark 6.2 We sketch how the framing (X, Y) can be modified so that its zero locus is contained in a union of meridians of the Λ_i. Take a meridian μ_i of Λ_i and handle-slide it through N_ϵ to obtain a longitude $-c_i \lambda_i$, which we may take to be $-c_i T^{c_i}$.

This homotopy, say parametrized by $[0,1]$ may be chosen so that the surface S it sweeps out is an embedded cylindrical cobordism parametrized by an annulus $[0,1] \times S^1$. Then we can find a family (X_t, Y_t) of sections of ξ_Λ whose zero loci are contained in $\{t\} \subset S^1$, so that (X_1, Y_1) will vanish along some union of the Λ_i, as desired.

If γ is a Reeb orbit of R_ϵ, then, according to (2), we can compute CZ_{X_1,Y_1} from $\text{CZ}_{X,Y}$ by counting the number of intersections of γ with S, which measures the meridional framing difference.

6.4 Rotation numbers and Chern classes in arbitrary contact 3–manifolds

We briefly state how the above can be generalized to understand how c_1 changes after contact surgery on an arbitrary contact manifold (M, ξ). A section $s \in \Gamma(\xi)$ determines a homotopy class of oriented trivialization of ξ on the complement of $s^{-1}(0)$ by considering $\xi_x = \text{span}_{\mathbb{R}}(s_x, Js_x)$ for an almost-complex structure J on ξ

compatible with $d\alpha$ for a contact form α for ξ and $x \in M \setminus s^{-1}(0)$. Suppose that s is transverse to the zero section and nonvanishing along a neighborhood N_ϵ of Λ. Write η_s for the oriented link

$$T_s = s^{-1}(0) \subset (M \setminus N_\epsilon) = (M_\Lambda \setminus N_\epsilon).$$

Definition 6.3 The *rotation number* $\mathrm{rot}_s(\Lambda)$ is the winding number of ∂_q in ξ_Λ determined by the trivialization of $\xi|_\Lambda$ provided by s.

Note that changing the orientation of Λ multiplies the rotation number by -1 and that rot_s agrees with the standard definition of the rotation number for oriented Legendrians in $(\mathbb{R}^3, \xi_{\mathrm{std}})$ if we take $s \in \Gamma_{\neq 0}(\xi_{\mathrm{std}})$. More generally, the rotation number of Legendrian knot Λ is canonically defined whenever ξ admits a nonvanishing section and at least one of $H^1(M) = 0$ or $[\Lambda^0] = 0 \in H_1(M)$ holds as in Proposition 2.5.

We note that Definition 6.3 may be applied to Legendrian knots in $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ even when these hypotheses are not satisfied: if such a knot Λ^0 is contained in $\mathbb{R}^3 \setminus N_\epsilon = \mathbb{R}^3_{\Lambda^\pm} \setminus N_\epsilon$, then $\mathrm{rot}_{X,Y}(\Lambda^0)$ may be computed using the typical methods for Legendrian knots in $(\mathbb{R}^3, \xi_{\mathrm{std}})$ as described in Section 2.8. This follows immediately from the first condition listed in Theorem 6.1.

Proposition 6.4 *Following the notation in the preceding discussion and writing λ_ξ for a longitude of Λ determined by ξ, and μ for a meridian of Λ, the Chern class $c_1(\xi_\Lambda)$ for the contact manifold (M_Λ, ξ_Λ) obtained by performing contact ± 1 surgery on $\Lambda \subset M$ is determined by the formula*

$$\mathrm{PD}(c_1(\xi_\Lambda)) = [T_s] \mp \mathrm{rot}_s(\Lambda)\lambda_\xi = [T_s] \pm \mathrm{rot}_s(\Lambda)\mu \in H_1(M_\Lambda).$$

This can be proved using the same strategy as Theorem 6.1, replacing X with s.

7 Conley–Zehnder and Maslov index computations

The goal of this section is to compute the integral Conley–Zehnder indices of closed orbits of the R_ϵ and the Maslov indices of broken closed strings on $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ using the framing (X, Y) defined in Section 6.

Theorem 7.1 *For each $n > 0$, there exists ϵ_0 such that, for all $\epsilon \leq \epsilon_0$, all orbits γ of word length $\leq n$ are hyperbolic with*

$$\mathrm{CZ}_{X,Y}(\gamma) = \sum_{k=1}^{n} (\mathrm{rot}_{j_k, j_{k+1}} + \delta_{1, c^+_{j_k}}) \in \mathbb{Z}.$$

Theorem 7.2 *Let b be a broken closed string on $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ of the form*

$$b = \zeta_1 * (a_1\kappa_1) * \cdots \zeta_n * (a_n\kappa_n),$$

where each ζ_k is a path in Λ^0 and each κ_k is a chord of Λ^0 with respect to R_ϵ. By Theorem 5.10, we can write

$$\kappa_k = (r_{k_1} \cdots r_{k_{n_k}})$$

for some word of chords with boundary on $\Lambda^0 \subset (\mathbb{R}^3, \xi_{\text{std}})$. In this notation,

$$\mathrm{M}_{X,Y}(b) = \sum_1^n \left(\frac{\theta(\zeta_k)}{\pi} - \frac{1}{2} + a_k m_{X,Y}(\kappa_k) \right),$$
$$m_{X,Y}(\kappa_k) = \sum_{l=1}^{n_k - 1} (\mathrm{rot}_{j_{k_l}, j_{k_l+1}} + \delta^+_{1,c_{k_l}}).$$

Proving the above theorems requires further analysis of (32) and (33). The analysis will provide an expression of the linearized flow of R_ϵ as a path of matrices in $\mathrm{SL}(2,\mathbb{R})$ with entries in $\mathbb{R}[\epsilon^{-1}]$ determined by $\mathrm{cw}(\gamma)$. Analysis of the highest-order terms of these polynomials gave us the proof of Theorem 5.3. Analysis of the second-highest-order terms of these polynomials will yield a formula for integral Conley–Zehnder indices, $\mathrm{CZ}_{X,Y}$.

7.1 Matrix model for the linearized flow

With respect to the coordinate system (z, p, q) inside of the surgery handles $R_\epsilon = \partial_z$. Hence computing the restriction of the linearized to flow to $\xi_{\Lambda^\pm}$ from the bottom ($z = -\epsilon$) to a point above it ($z > -\epsilon$) in the surgery handle with respect to (X, Y) amounts to writing $(X, Y)_{-\epsilon,p,q}$ in the basis $(X, Y)_{z,p,q}$. We write $F_i(z, p, q) \in \mathrm{SL}(2,\mathbb{R})$ for this path of matrices associated to points (z, p, q) in the component of N_ϵ associated to Λ_i.

By composing (30) with equations (32) — in the case of $+1$ surgery — and (33) — in the case of -1 surgery — we have

$$F_i(z,p,q) = E(\zeta^{c_i}, \eta^{c_i}) e^{J_0 G_i(\eta^{c_i})} \begin{pmatrix} 1 & -c_i\, \partial f_\epsilon / \partial p(\zeta^{c_i}) \\ 0 & 1 \end{pmatrix} e^{-J_0 G_i(q)} E^{-1}(p,q) \tag{35}$$
$$= e^{J_0 G_i(\eta^{c_i})} \begin{pmatrix} 1 & -c_i\, \partial f_\epsilon / \partial p(\zeta^{c_i}) \\ 0 & 1 \end{pmatrix} e^{-J_0 G(q)} (\mathrm{Id} + \mathcal{O}(p)).$$

We have preemptively simplified the expression with some basic arithmetic. Here G_i is the Gauss map associated to the component Λ_i of Λ. The following collection of assumptions will allow us to further simplify the above expression:

Assumptions 7.3 We refine our previous constructions as follows: At any point through which a closed Reeb orbit passes, the sections X and Y of $\xi_{\Lambda^\pm}$ described in Section 6.2 are defined according to the formula contained within that section. This can be achieved by setting the constant δ to be sufficiently small.

Some consequences of the above assumptions coupled with Assumptions 5.4 are:

(1) Equation (35) is valid for any point contained in a closed Reeb orbit.

(2) The expression $-c_i\,\partial f_\epsilon/\partial p(\zeta^{c_i})$ in that formula simplifies to $-c_i\epsilon^{-1}$ for any point lying in a closed Reeb orbit.

Combining these consequences with a conjugation, we have that F_i in a neighborhood of a Reeb segment which exits $N_{\epsilon,i}$ near $l^+_{j_1}$ and exits near $l^-_{j_2}$ for composable Reeb chords r_{j_i} and r_{j_2} is homotopic — relative endpoints — to a path of the form

$$\mathcal{F}_{j_1,j_2}(t) = e^{J_0 t\theta_{j_1,j_2}} \begin{pmatrix} 1 & -tc^+_{j_1}\epsilon^{-1} \\ 0 & 1 \end{pmatrix} (\mathrm{Id} + \mathcal{O}(\epsilon)) \in \mathrm{SL}(2,\mathbb{R}), \quad t \in [0,1], \tag{36}$$

where we use the basis $e^{i\pi/4}(X, Y)$.

Using (36), we can write the restriction of Poincaré return map to ξ of a closed Reeb orbit γ of α_ϵ with $\mathrm{cw}(\gamma) = r_{j_1}\cdots r_{j_n}$ as

$$\mathrm{Ret}_\gamma = \mathcal{F}_{j_n,j_1}(1)\mathcal{F}_{j_{n-1},j_n}(1)\cdots\mathcal{F}_{j_1,j_2}(1) \tag{37}$$

by composing the flow maps as an orbit passes through the various surgery handles. If the word consists of a single chord, then we have $\mathrm{Ret} = \mathcal{F}_{j_1,j_1}(1)$. Note that, while our expression for Ret depends on a particular representation of the associated cyclic word, its conjugacy class in $\mathrm{SL}(2,\mathbb{R})$ does not.

7.2 Integral Conley–Zehnder indices

In this subsection we prove Theorem 7.1 via induction on the word length n of γ. The proof is computational, making use of the Robbin–Salamon characterization of the Conley–Zehnder index described in (5).

7.2.1 The case $n = 1$ We begin with the case $n = 1$, analyzing (36). A slight modification of the proof of the following lemma along with further analysis of (26) will provide the general case. For the sake of notational simplicity, we temporarily drop the subscripts required to describe words of length greater than 1.

Lemma 7.4 *Theorem 7.1 is valid for Reeb orbits of word length* 1.

Proof We homotop the path $\mathcal{F}$ so that it is parametrized with the interval $[0, 3]$, taking the form

$$\mathcal{F}(t) = e^{J_0 t_1 \theta} \begin{pmatrix} 1 & -t_2 c\epsilon^{-1} \\ 0 & 1 \end{pmatrix} E(t_3), \quad E(0) = \mathrm{Id}, \quad E(1) = \mathrm{Id} + \mathcal{O}(\epsilon),$$

$$t_k = \begin{cases} 0 & \text{if } t \leq k-1, \\ t-k+1 & \text{if } t \in (k-1, k), \\ 1 & \text{if } t \geq k. \end{cases} \tag{38}$$

With this parametrization we are performing the rotation first so that the path is nondegenerate at $\mathcal{F}(0) = \mathrm{Id}$. A standard computation shows that, along the interval $[0, 1]$, the contributions to $\mathrm{CZ}_{X,Y}$ are given by $2\lfloor \theta/2\pi \rfloor + 1$. Then, along $t \in [1, 2]$, we have

$$\mathcal{F}(t) = (-1)^{\mathrm{rot}} \begin{pmatrix} 0 & -1 \\ 1 & -t_2 c\epsilon^{-1} \end{pmatrix}$$

By the $\mathrm{SL}(2, \mathbb{R})$ trace formula of (1), $t \in [1, 2]$ will be crossing exactly when

$$\mathrm{tr}(\mathcal{F}(t)) = (-1)^{1+\mathrm{rot}} t_2 c = 2.$$

Therefore, we find a crossing in the interval — and a single one at that — if and only if $c = (-1)^{1+\mathrm{rot}}$. At such a crossing, if it exists, the matrix $S(t)$ of (4) is $S(t) = \mathrm{Diag}(t_2 c\epsilon^{-1}, 0)$. So the contribution to $\mathrm{CZ}_{X,Y}$ can be computed as $\frac{1}{2}(c - (-1)^{\mathrm{rot}})$.

Adding up the contributions along $t \in [0, 2]$, we have

$$\begin{aligned} \mathrm{CZ} &= 2\left\lfloor \frac{\theta}{2\pi} \right\rfloor + 1 + \tfrac{1}{2}(c - (-1)^{\delta}) = 2\left\lfloor \frac{\theta}{2\pi} \right\rfloor + \tfrac{1}{2}(1 - (-1)^{\mathrm{rot}}) + \tfrac{1}{2}(c+1) \\ &= \mathrm{rot} + \delta_{1,c}. \end{aligned}$$

Along the interval $[2, 3]$, the addition of the E term to the formula contributes a term to the trace which is bounded by a constant which is independent of ϵ. Thus, this interval is devoid of crossings for ϵ small. □

7.2.2 The case $n > 1$ Now we prove the induction step in our index computation. We suppose that the Reeb orbit in question has word length $n + 1 > 1$ and is parametrized with an interval $[0, n + 1]$. Then we can compute the Conley–Zehnder index using the

path of symplectic matrices

$$\phi(t) = \mathcal{F}_{j_n,j_{n+1}}(t_{n+1}) \cdots \mathcal{F}_{j_{n+1},j_1}(t_1), \quad t_k = \begin{cases} 0 & \text{if } t \le k-1, \\ t-k+1 & \text{if } t \in (k-1,k), \\ 1 & \text{if } t \ge k, \end{cases}$$

by combining (36) and (37). As in the proof of Lemma 7.4, we can drop the E terms in the equations, count the contributions to CZ coming from the rotation and shearing matrices, then reintroduce the E terms, noting that they do not contribute to CZ due to the large absolute values of traces. Consequently, we ignore these E terms during computation. With this simplification, ϕ takes the form, when $t \in [n, n+1]$,

$$\phi(t) = \widetilde{\phi}(t_{n+1}) \operatorname{Ret}_n, \tag{39}$$

where

$$\begin{aligned} \widetilde{\phi}(t_{n+1}) &= e^{i t_n \theta_{j_{n+1},j_1}} \begin{pmatrix} 1 & -t_{n+1} c^+_{j_{n+1}} \epsilon^{-1} \\ 0 & 1 \end{pmatrix}, \\ \operatorname{Ret}_n &= \mathcal{F}_{j_{n-1},j_n}(1) \cdots \mathcal{F}_{j_{n+1},j_1}(1) \\ &= (-1)^{\mathrm{rot}_n} \Bigg(J^n + \sum_{K=1}^{n} \Bigg(\sum_{k \in I_K} \Bigg(\prod_{i=1}^{K} -c^+_{j_{k_i}} \Bigg) M_k \Bigg) \epsilon^{-K} \Bigg), \\ \mathrm{rot}_n &= \sum_{0}^{n-1} \mathrm{rot}_{j_k,j_{k+1}}, \end{aligned}$$

over the subinterval $[n, n+1]$. Here indices are cyclic, so that $\mathrm{rot}_{j_0,j_1} = \mathrm{rot}_{j_{n+1},j_1}$. The M_k are as in (26).

By Theorem 5.3, the trace of Ret_n has absolute value of order ϵ^{-n}. Therefore, $n, n+1 \in [0, n+1]$ are not crossing for ϵ small. The ϵ^{-n} term in Ret_n is given by (27). The ϵ^{1-n} term is also easily computable. Noting that $\operatorname{Diag}(0,a) J \operatorname{Diag}(0,b) = 0$ for $a, b \in \mathbb{R}$, the only k for which M_k is nonzero with $k \in I_n$ are $(1, \dots, n-1)$ and $(2, \dots, n)$. Thus, the ϵ^{1-n} terms in Ret_n are

$$\Bigg(\prod_{1}^{n-1} -c^+_{j_{k_i}} \Bigg) \operatorname{Diag}(0,1) J + \Bigg(\prod_{2}^{n} -c^+_{j_{k_i}} \Bigg) J \operatorname{Diag}(0,1) = \begin{pmatrix} 0 & -\prod_2^n -c^+_{j_{k_i}} \\ \prod_1^{n-1} -c^+_{j_{k_i}} & 0 \end{pmatrix}.$$

Combining this with (27), we have

$$\begin{aligned} \operatorname{Ret}_n &= (-1)^{\mathrm{rot}_n} \begin{pmatrix} 0 & -\epsilon^{-n+1} \prod_2^n -c^+_{j_{k_i}} \\ \epsilon^{-n+1} \prod_1^{n-1} -c^+_{j_{k_i}} & \epsilon^{-n} \prod_1^n -c^+_{j_{k_i}} \end{pmatrix} + \mathcal{O}(\epsilon^{2-n}) \\ &= (-1)^{\mathrm{rot}_n} \epsilon^{-n+1} \Bigg(\prod_{1}^{n} -c^+_{j_{k_i}} \Bigg) \begin{pmatrix} 0 & c^+_{j_{k_1}} \\ -c^+_{j_{k_n}} & \epsilon^{-1} \end{pmatrix} + \mathcal{O}(\epsilon^{2-n}). \end{aligned} \tag{40}$$

Lemma 7.5 *For ϵ sufficiently small, the contribution to $\mathrm{CZ}_{X,Y}$ along the interval $[n, n+1]$ in (39) is*

$$\mathrm{rot}_{j_n, j_{n+1}} + \delta_{1, c^+_{j_{n+1}}}.$$

Proof We begin by making some temporary notational simplifications and further subdivide the interval $[n, n+1]$ along which the map $\widetilde{\phi}$ is changing and Ret_n is constant. We are referring here to (39) and use notation from that equation throughout the proof. We write

$$\sigma = (-1)^{\mathrm{rot}_n}\left(\prod_1^n -c^+_{j_{k_i}}\right) \in \{\pm 1\},$$

$$c_1 = c^+_{j_1}, \quad c_n = c^+_{j_n}, \quad c_{n+1} = c^+_{j_{n+1}},$$

$$\mathrm{rot} = \mathrm{rot}_{j_n, j_{n+1}}, \quad \mathrm{rot}_2 = \mathrm{rot} \mod 2 \in \mathbb{Z}/2\mathbb{Z},$$

$$\theta_{j_n, j_{n+1}} = \pi\big(2k + \delta_{1,\mathrm{rot}_2} + \tfrac{1}{2}\big), \quad k = \left\lfloor \frac{\theta_{j_n, j_{n+1}}}{2\pi} \right\rfloor \in \mathbb{Z}.$$

By combining the above notation with (40), we can write $\phi(t) = \widetilde{\phi}(t_n)\,\mathrm{Ret}_n$, where

$$\widetilde{\phi}(t_{n+1}) = e^{J_0 t_{n+1} \theta_{j_n, j_{n+1}}} \begin{pmatrix} 1 & -t_{n+1} c_{n+1} \epsilon^{-1} \\ 0 & 1 \end{pmatrix},$$

$$\mathrm{Ret}_n = \sigma \epsilon^{1-n} \begin{pmatrix} 0 & c_1 \\ -c_n & \epsilon^{-1} \end{pmatrix} + \mathcal{O}(\epsilon^{2-n}).$$

Along the subset $t \in [n, n+1]$, we homotop $\widetilde{\phi}$ to take the form

$$\widetilde{\phi}(t_{n+1}) = e^{J_0 \theta} \begin{pmatrix} 1 & -s_2 c_{n+1} \epsilon^{-1} \\ 0 & 1 \end{pmatrix}, \quad \theta = \pi\big(s_1 \cdot \tfrac{1}{4} + s_3\big(\delta_{1,\mathrm{rot}_2} + \tfrac{1}{4}\big) + s_4 \cdot 2k\big),$$

where s_i are functions of t taking values in $[0, 1]$ as described in Figure 15, so that $\theta = \theta_{j_n, j_{n+1}}$ when $s_1 = \cdots = s_4 = 1$. In words, we will be applying a rotation by $\frac{1}{4}\pi$, a shear, a rotation by $\pi\big(\delta_{1,\mathrm{rot}_2} + \frac{1}{4}\big)$, and then finally a rotation by $2\pi k$. Taking the arguments of all trigonometric functions to be θ,

$$\begin{aligned} \phi &= \begin{pmatrix} \cos & -\sin \\ \sin & \cos \end{pmatrix} \begin{pmatrix} 1 & -s_2 c_{n+1} \epsilon^{-1} \\ 0 & 1 \end{pmatrix} \left(\sigma \epsilon^{1-n} \begin{pmatrix} 0 & c_1 \\ -c_n & \epsilon^{-1} \end{pmatrix} + \mathcal{O}(\epsilon^{2-n}) \right) \\ &= \sigma \epsilon^{-n} \begin{pmatrix} s_2 c_n c_{n+1} \cos & -s_2 c_{n+1} \epsilon^{-1} \cos - \sin \\ s_2 c_n c_{n+1} \sin & -s_2 c_{n+1} \epsilon^{-1} \sin + \cos \end{pmatrix} + \mathcal{O}(\epsilon^{1-n}), \\ \mathrm{tr}(\phi) &= \sigma \epsilon^{-n} (s_2 c_n c_{n+1} \cos - s_2 c_{n+1} \epsilon^{-1} \sin + \cos) + \mathcal{O}(\epsilon^{1-n}). \end{aligned}$$

Along our first subinterval parametrized by s_1, we have $s_2 = s_3 = s_4 = 0$ with θ increasing from 0 to $\frac{1}{4}\pi$. Here $\mathrm{tr}(\phi) = \sigma \epsilon^{-n} \cos + \mathcal{O}(\epsilon^{1-n})$ has large absolute value

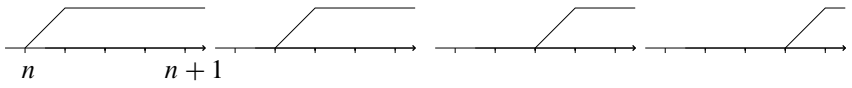


Figure 15: From left to right are graphs of the functions $s_1(t), \dots, s_4(t)$.

when ϵ is small. Hence, for ϵ small, there are no crossings over the s_1 subinterval and so there are no contributions to the Conley–Zehnder index.

Along the subinterval parametrized by s_2, we have $\theta = \frac{1}{4}\pi$, so that $\cos(\theta) = \sin(\theta) = \frac{1}{\sqrt{2}}$. Hence,

$$\operatorname{tr}(\phi) = \tfrac{1}{\sqrt{2}}\sigma\epsilon^{-n}(s_2c_nc_{n+1} - s_2c_{n+1}\epsilon^{-1} + 1) + \mathcal{O}(\epsilon^{1-n}). \tag{41}$$

At both the $s_2 = 0$ and $s_2 = 1$ endpoints of the interval, $\operatorname{tr}(\phi)$ will have large absolute value of orders ϵ^{-n} and ϵ^{-1-n}, respectively, so that ϕ cannot be crossing at either of these endpoints. Over the interior of the subinterval, we see that there is a single crossing if $c_{n+1} = 1$ and no crossings if $c_{n+1} = -1$.

If $c_{n+1} = 1$, then at the unique crossing we compute the crossing form

$$S = -J_0\frac{\partial\phi}{\partial s_2}\phi^{-1} = \frac{1}{2\epsilon}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} a & b \end{pmatrix} S \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{2\epsilon}(a-b)^2.$$

The quadratic form determined by S vanishes exactly along the line $\mathbb{R}\binom{1}{1}$. Therefore, to see that the unique crossing along the s_2 subinterval is nondegenerate, we only need to check that $\phi\binom{1}{1} \neq 0$ at the crossing. Plugging $c_{n+1} = 1$ into (41), at the crossing we must have

$$s_2(c_n - \epsilon^{-1}) + 1 = \mathcal{O}(\epsilon)$$

in order to eliminate the ϵ^{-n} and ϵ^{-n-1} terms. Therefore, at the crossing,

$$\phi\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \tfrac{1}{\sqrt{2}}\sigma\epsilon^{-n}\begin{pmatrix} s_2(c_n - \epsilon^{-1}) - 1 \\ s_2(c_n - \epsilon^{-1}) + 1 \end{pmatrix} + \mathcal{O}(\epsilon^{1-n}) = \tfrac{1}{\sqrt{2}}\sigma\epsilon^{-n}\begin{pmatrix} -2 \\ 0 \end{pmatrix} + \mathcal{O}(\epsilon^{1-n})$$

is nonzero for ϵ small. We conclude that, at the crossing, $\ker(\phi - \mathrm{Id})$ must be 1–dimensional and the restriction of S to $\ker(\phi - \mathrm{Id})$ must be positive. Hence, the s_2 subinterval contributes $\delta_{c_{n+1},1}$ to the Conley–Zehnder index.

Now we study the s_3 subinterval along which $s_2 = 1$ and $\theta \in \left[\frac{1}{4}\pi, \pi\left(\delta_{1,\mathrm{rot}_2} + \frac{1}{2}\right)\right]$. Along this subinterval,

$$\operatorname{tr}(\phi) = \sigma\epsilon^{-n}(c_nc_{n+1}\cos - c_{n+1}\epsilon^{-1}\sin + \cos) + \mathcal{O}(\epsilon^{1-n}).$$

If $\delta_{1,\mathrm{rot}_2} = 0$ and ϵ is very small, then, as θ increases from $\frac{1}{4}\pi$ to $\frac{1}{2}\pi$, the sin term dominates, $\mathrm{tr}(\phi)$ maintains a large absolute value, and there are no crossings.

If $\delta_{1,\mathrm{rot}_2} = 1$, then we have a single crossing at which the crossing form is determined by the matrix

$$S = -J_0 \frac{\partial \phi}{\partial s_2} \phi^{-1} = \tfrac{3}{4}\pi \,\mathrm{Id}.$$

Because the crossing form is positive definite, the contribution to the Conley–Zehnder index is the dimension of $\ker(\phi - \mathrm{Id})$ — either 1 or 2. At the crossing, sin is $\mathcal{O}(\epsilon)$, implying that $\theta \to \pi$ (and so $\cos(\theta) \to -1$) as $\epsilon \to 0$. Therefore, at the crossing,

$$\begin{aligned}(\phi - \mathrm{Id}) \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= \sigma \epsilon^{-n} \begin{pmatrix} -c_{n+1}\epsilon^{-1} \cos - \sin \\ -c_{n+1}\epsilon^{-1} \sin + \cos \end{pmatrix} + \mathcal{O}(\epsilon^{1-n}) \\ &= \sigma \epsilon^{-1-n} \begin{pmatrix} -c_{n+1} \cos \\ 0 \end{pmatrix} + \mathcal{O}(\epsilon^{-n}) \neq 0,\end{aligned}$$

implying that $\dim \ker(\phi - \mathrm{Id}) = 1$. We conclude that the contribution to CZ along the s_3 interval is $\delta_{1,\mathrm{rot}_2}$.

For the s_4 subinterval, we appeal to the loop property of CZ described in (6) to see a contribution of $2k$. Combining the contributions over the four s_i subintervals, we get

$$2k + \delta_{\mathrm{rot}_2,1} + \delta_{c_{n+1},1} = \mathrm{rot}_{j_n, j_{n+1}} + \delta_{c^+_{j_{n+1}},1},$$

by reverting to our original notation, thereby completing the proof. □

The combination of the above lemmas completes our induction, thereby proving Theorem 7.1.

7.3 Integral Maslov indices

The proof of Theorem 7.2 follows from the same methods of calculation as Theorem 7.1.

Proof of Theorem 7.2 According to Definition 3.8, we need to measure the rotation of $\mathrm{Flow}^t_{R_\epsilon}(T_{q^-_{k_1}} \Lambda^0)$ along each chord $(r_{k_1} \cdots r_{k_{n_k}})$ with respect to the framing (X, Y). For chords κ_k of word length 1, the flow is trivial, so we restrict attention to chords of word length > 1. The required analysis can be carried out via analysis of (36); we recall that this describes the restriction of the linearized flow of R_ϵ to $\xi_{\Lambda^\pm}$ through a component N_{ϵ,j_1} of N_ϵ starting at a point near the tip of one chord r_{j_1} up to a point near the tail of another chord r_{j_2}.

We study the rotation along a single $\kappa = (r_{j_1} \cdots r_{j_n})$: The matrix expression $\mathcal{F}_{j_1,j_2}(t)$ in (36) applies to the basis

$$e^{i\pi/4}(X, Y) \simeq (\partial_q - p\partial_z, -\partial_p),$$

beginning on the bottom, $\{z = -\epsilon\}$, of the surgery handle N_{ϵ,j_1}. For $j_1 = k_1$ and $j_2 = k_2$, the strand of Λ^0 touching the starting point of the chord r_{j_1} is such that $T\Lambda^0 = \mathbb{R}\binom{0}{1}$. Therefore, we need to see how $\mathcal{F}_{j_1,j_2}(t)$ rotates this subspace for $t \in [0, 1]$. As in the proof of Theorem 7.1, we can modify the path so as to apply the shearing first, and then the rotation.

For the shearing, we study the family of real lines in $\mathbb{R}^2$ given by

$$\mathbb{R}\begin{pmatrix} 1 & -tc^+_{j_1}\epsilon^{-1} \\ 0 & 1 \end{pmatrix}(\mathrm{Id} + \mathcal{O}(\epsilon))\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad t \in [0, 1].$$

The end result is a line of the form $\mathbb{R}\left(\binom{1}{0} + \mathcal{O}(\epsilon)\right)$ obtained by rotating $\mathbb{R}\binom{0}{1}$ by an angle of

$$c^+_{j_1} \cdot \tfrac{1}{2}\pi + \mathcal{O}(\epsilon). \tag{42}$$

Then, applying the rotation through angles $t\theta_{j_1,j_2}$ as in (36), we rotate this subspace by

$$\theta_{j_1,j_2} = \pi \,\mathrm{rot}_{j_1,j_2} + \tfrac{1}{2}\pi, \tag{43}$$

which we recall from Section 3 is the rotation angle of the capping path $\eta_{i,j}$ associated to the pair of composable chords (r_{j_1}, r_{j_2}).

Continuing the flow by applying the remaining $\mathcal{F}_{k_l,k_{l+1}}$ for $l = 2, \ldots, n_k - 1$ provides us a total rotation angle of

$$\pi\left(\sum_{l=1}^{n_k-1} \mathrm{rot}_{k_l,k_{l+1}} + \tfrac{1}{2} + \tfrac{1}{2}c^+_{k_l}\right) + \mathcal{O}(\epsilon) = \pi\left(\sum_{l=1}^{n_{k-1}} \mathrm{rot}_{k_l,k_{l+1}} + \delta_{c^+_{k_l},1}\right) + \mathcal{O}(\epsilon),$$

leaving us on a neighborhood of the strand of Λ lying at the starting point of the chord $r_{k_{n_k}}$ which ends on Λ^0. Each summand in the above formula is the result of adding the contributions of (42) and (43). As in the case $\mathrm{wl}(\kappa) = 1$, the linearized flow up to Λ^0 along $r_{k_{n_k}}$ is trivial in the basis (X, Y). This nearly completes the construction of the section ϕ^G along the chord κ: we must apply one more rotation to ensure that the unoriented Lagrangian line closes up as described in the discussion preceding Definition 3.8.

In the case that our asymptotic indicator is $a_k = 1$, the total rotation along the chord will be

$$\pi\left(-\tfrac{1}{2} + \sum_{l=1}^{n_{k-1}} (\mathrm{rot}_{k_l,k_{l+1}} + \delta_{c^+_{1,k_l}})\right),$$

where the $\frac{1}{2}$ is the contribution of the clockwise correction rotation at the end of the chord. From the above analysis we know that the angle for this rotation is $-\frac{1}{2}\pi + \mathcal{O}(\epsilon)$.

If the asymptotic indicator is -1, we must travel in the opposite direction, from the tip to the tail of the chord, and then apply small the clockwise rotation to obtain a total rotation angle of

$$\pi\left(-\tfrac{1}{2} - \sum_{l=1}^{n_{k-1}} (\mathrm{rot}_{k_l,k_{l+1}} + \delta_{c^+_{1,k_l}})\right).$$

The section ϕ^G appearing in the definition of M_s is determined by the $\theta(\zeta_k)$ as the framing (X, Y) coincides with $(\partial_x - y\partial_z, \partial_y)$ — from which the $\theta(\zeta_k)$ are computed — on the complement of N_ϵ, in which Λ^0 is presumed to be contained. □

8 Diagrammatic index formulas

In this section we compute indices of holomorphic curves in $(\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}, d(e^t\alpha_\epsilon))$. We begin by covering the case of curves whose domain is a closed surface with punctures, which is a simple application of (8) to our existing computations of Conley–Zehnder indices and Chern classes. Next we cover the case of a holomorphic disk which is asymptotic to a broken closed string on $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ in the sense of Example 3.4. The case $\Lambda^\pm = \varnothing$ recovers a classic index formula appearing in combinatorial versions of LCH and Legendrian *RSFT*. These index formulas are then combined to describe indices associated to holomorphic curves with arbitrary configurations of interior and boundary punctures in Theorem 8.3.

All indices computed will depend only on topological data, so mention of any specific almost-complex structures are ignored.

8.1 Index formulas for closed orbits

Let $\{w_1^+, \dots, w_{m^+}^+\}$ and $\{w_1^+, \dots, w_{m^-}^-\}$ be collections of cyclic words of chords on Λ. By Theorem 7.1, we may choose some $\epsilon_\gamma > 0$ such that for all $\epsilon < \epsilon_\gamma$, the closed orbits $\gamma_j^\pm$ of R_ϵ corresponding to these cyclic words via Theorem 5.8 are all nondegenerate hyperbolic à la Theorem 5.3. Write $\gamma^+ = \{\gamma_1^+, \dots, \gamma_{m^+}^+\}$ and $\gamma^- = \{\gamma_1^-, \dots, \gamma_{m^-}^-\}$ for the corresponding collections of orbits.

Suppose that (Σ, j) is a closed Riemann surface containing a nonempty collection of punctures and that $(t, U)\colon \Sigma' \to \mathbb{R} \times \mathbb{R}^3_{\Lambda^\pm}$ is a holomorphic curve (as in Section 2.7) which is positively asymptotic to the punctures γ^+ and negatively asymptotic to the γ^-.

Theorem 8.1 *Using the framing (X, Y) described in Section 6.2, we can write the expected dimension of the moduli space of curves near (t, U) as*

$$\text{(44)}\quad \operatorname{ind}((t,U)) = \mathrm{CZ}_{X,Y}(\gamma^+) - \mathrm{CZ}_{X,Y}(\gamma^-) - \chi(\Sigma') - 2\sum_1^n c_i \operatorname{rot}(\Lambda_i)(U \cdot T_i^{c_i})$$

for all $\epsilon < \epsilon_\gamma$, where the sum runs over the connected components of $\Lambda^\pm$ and the $\mathrm{CZ}_{X,Y}$ are computed as in Theorem 7.1.

Proof By comparing with (8), we only need to check that

$$c_{X,Y}(U) = -\sum_1^n c_i \operatorname{rot}(\Lambda_i)(U \cdot T_i^{c_i}),$$

where $c_{X,Y}(U)$ is the relative Chern class of the framing (X, Y). Letting $X_U \in U^*(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ be a section for which $T_z U(X_U) = X(U(z))$ for $z \in \Sigma'$ to compute $c_{X,Y}(U)$ provides the desired result, as $X^{-1}(0)$ is a union of connected components of $\bigcup T_i^{c_i}$ and the coefficients $-c_i \operatorname{rot}(\Lambda_i)$ account for the multiplicities of the zeros of X_U by the construction of (X, Y) in Section 6. □

8.2 Index formulas for disks with boundary punctures

Now suppose that $\{p_j\}_1^m \subset \partial\mathbb{D}$ is a collection of distinct points on the boundary of a disk. Write $\mathbb{D}' = \mathbb{D} \setminus \{p_j\}$ for the complement of the boundary punctures in $\mathbb{D}$ and write j for the standard complex structure on $\mathbb{D}$. Suppose that $(t, U)\colon \mathbb{D}' \to \mathbb{R} \times \mathbb{R}^3_{\Lambda^\pm}$ is a (j, J)–holomorphic map satisfying

(1) $(t, U)(\partial\mathbb{D}') \subset \mathbb{R} \times \Lambda^0$, and

(2) the punctures $\{p_j\}$ are asymptotic to chords of R_ϵ with boundary on $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$.

As described in Example 3.4, such a map determines a broken closed string which we will denote by $\operatorname{bcs}(U)$. As in the case of (8), we use $\operatorname{ind}((t, U))$ to denote the expected dimension of the space of holomorphic maps with the same $\operatorname{bcs}(U)$ boundary conditions as (t, U) and in the same relative homotopy class obtained by allowing the locations to vary and then modding out by holomorphic reparametrization in the domain (when $m < 3$).

Theorem 8.2 *The moduli space of holomorphic disks with boundary condition* $\mathrm{bcs}(U)$ *in the homotopy class of* U *has expected dimension*

$$\mathrm{ind}((t,U)) = \mathrm{M}_{X,Y}(\mathrm{bcs}(U)) + m - 1 - 2\sum_{1}^{n} c_i\,\mathrm{rot}(\Lambda_i)(U \cdot T_i^{c_i}) \tag{45}$$

near the point (t,U)*. The sum appearing in the above formula is indexed over the components* Λ_i *of* Λ*.*

Proof We are simply plugging our definition of broken closed strings into formulas appearing in [19; 17].

Assume first that X is nonvanishing over $\mathrm{im}(U)$, so that ∂_t, R_ϵ, X and Y determine a trivialization of $U^*T(\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm})$ which splits as a pair of complex lines. Using framing deformation-invariance of $\mathrm{M}_{X,Y}$, we may perturb (X,Y) so that it is invariant under the flow of R_ϵ, in which case the geometric setup described in [17, Section 3.1] applies. Our choices of "clockwise rotations" along positive punctures and "counterclockwise rotations" along negative punctures in the definition of the path of symplectic matrices defining M_s coincide with those used to define the Maslov numbers (which are denoted by $\mu(\gamma)$) in that text. The tangent space of our Lagrangian — $\mathbb{R}\times\Lambda^0$ — splits as $\mathbb{R}\partial_t \oplus T\Lambda^0$, so the only contribution to the Maslov number in question comes from the rotation of $T\Lambda^0$ along the boundary of the disk by the direct sum formula for Maslov numbers. Then the moduli space dimension formula of [17, Section 3.1] completes our proof.

Now suppose that X is nonvanishing along $\mathrm{im}(U)$. By the construction of the framing (X,Y), we have that this section must be nonvanishing along Λ^0 and all of its Reeb chords, and so is nonvanishing along $\mathrm{im}(\mathrm{bcs}(U))$. Therefore, the Maslov index can corrected by a relative Chern class term as in (8), which may be computed as signed count of intersections of U with the transverse push-offs of the Λ_i as in the statement of that theorem. □

8.3 Index formulas for curves with interior and boundary punctures

Now we state an index formula for holomorphic curves of general topological type. The geometric setup is as follows.

Let (Σ, j) be a compact, connected Riemann surface with boundary components

$$(\partial\Sigma)_k, \quad k = 1,\dots,\#(\partial\Sigma),$$

marked points $p_i^{\mathrm{int},\pm}$ contained in $\mathrm{int}(\Sigma)$, and marked points $p_i^{\partial,\pm}$ contained in $\partial\Sigma$. We write Σ' for Σ with all of its marked points removed. Consider a holomorphic map $(t,U)\colon \Sigma' \to \mathbb{R}\times(\mathbb{R}^3_{\Lambda^\pm},\xi_{\Lambda^\pm})$ subject to the following conditions:

(1) the $p_i^{\mathrm{int},+}$ are positively asymptotic to some collection γ^+ of closed orbits of R_ϵ,

(2) the $p_i^{\mathrm{int},-}$ are negatively asymptotic to some collection γ^- of closed orbits of R_ϵ,

(3) the $p_i^{\partial,+}$ are positively asymptotic to some collections κ^+ of chords of $\Lambda^0\subset(\mathbb{R}^3_{\Lambda^\pm},\xi_{\Lambda^\pm})$,

(4) the $p_i^{\partial,-}$ are negatively asymptotic to some collections κ^- of chords of $\Lambda^0\subset(\mathbb{R}^3_{\Lambda^\pm},\xi_{\Lambda^\pm})$, and

(5) $(t,U)(\partial\Sigma')\subset\mathbb{R}\times\Lambda^0$.

In this setup, we have a broken closed string bcs_k associated to each component $(\partial\Sigma)_k$ of Σ. We may consider the moduli space of curves subject to the same asymptotics — $\gamma^\pm$ and bcs_k — allowing the complex structure on Σ to vary and taking a quotient by j–holomorphic symmetries on the domain.

Theorem 8.3 *In the above notation, the expected dimension of the moduli space of holomorphic maps is*

$$\begin{aligned}\mathrm{ind}((t,U)) = \mathrm{CZ}_{X,Y}(\gamma^+) - \mathrm{CZ}_{X,Y}(\gamma^-) + \sum_k \mathrm{M}_{X,Y}(\mathrm{bcs}_k)\\ -\chi(\Sigma) + \#(p^{\mathrm{int}}) + \#(p^{\partial}) - 2\sum_{\Lambda_i\subset\Lambda^\pm} c_i\,\mathrm{rot}(\Lambda_i)(U\cdot T_i^{c_i}).\end{aligned}$$

The proof is a simple combination of Theorems 8.1 and 8.2 using index additivity; see [58, Section 3].

9 H_1 computations and push-outs of closed orbits

Here we compute the first homology $H_1(\mathbb{R}^3_{\Lambda^\pm})$ of $\mathbb{R}^3_{\Lambda^\pm}$ and the homology classes of the closed orbits of R_ϵ.

Theorem 9.1 *The first homology $H_1(\mathbb{R}^3_{\Lambda^\pm})$ is presented with generators μ_i and relations*

$$(\mathrm{tb}(\Lambda_i)+c_i)\mu_i + \sum_{j\neq i}\mathrm{lk}(\Lambda_i,\Lambda_j)\mu_j = 0.$$

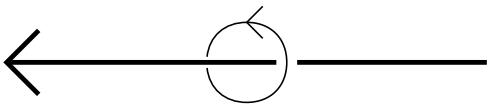


Figure 16: Default orientations for meridians.

Let γ be a Reeb orbit of α_ϵ with $\mathrm{cw}(\gamma) = r_{j_1} \cdots r_{j_n}$. Then its homology class in $H_1(\mathbb{R}^3_{\Lambda\pm})$ with respect to the above basis is

$$[\gamma] = \frac{1}{2} \sum_{k=1}^{n} (\mathrm{cr}_{j_k} + \mathrm{cr}_{j_k, j_{k+1}}),$$

where the k are considered modulo n.

Relative homology classes $[\kappa] \in H_1(\mathbb{R}^3_{\Lambda\pm}, \Lambda^0)$ of chords κ with boundary on $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda\pm}, \xi_{\Lambda\pm})$ can similarly be computed using the technique of the proof of Theorem 9.1 which is carried out in Section 9.3. It will be clear that the method of proof allows the reader to compute $[\gamma]$ as an element of the H_0 of the free loop space of $\mathbb{R}^3_{\Lambda\pm}$. In Section 9.4, we show how the proof can be generalized to provide a general means of homotoping closed orbits of R_ϵ into $\mathbb{R}^3 \setminus N$, a technique we will need for the proof of Theorem 1.2.

9.1 Conventions for meridians and longitudes

Before proving Theorem 9.1, we quickly review some standard notation. Let μ_j denote a meridian of Λ and λ_i a longitude of Λ provided by the Seifert framing and orientation of Λ_i. We note that, with respect to the Seifert framing of Λ_i, the longitude provided by ξ, denoted by $\lambda_{\xi,i}$ is

$$\lambda_{\xi,i} = \lambda_i + \mathrm{tb}(\Lambda_i)\mu_i.$$

Each μ_i is oriented so that

$$(\text{meridian, longitude, outward-pointing normal})$$

is a basis for $T\mathbb{R}^3$ agreeing with the usual orientation over ∂N (after rounding the edges of ∂N in the obvious fashion). See Figure 16.

9.2 First homology of the ambient space

The computation of $H_1(\mathbb{R}^3_{\Lambda\pm})$ easily follows from the fact that contact ± 1 surgery is a form of Dehn surgery. Suppose that $\mathbb{R}^3_L$ is a 3–manifold obtained by Dehn surgery on a smooth link $L = \bigcup L_i$ for which the surgery coefficients with respect to the Seifert

framing are p_i/q_i for relatively prime integers p_i and q_i. Writing μ_j for the oriented meridians of the L_i, we have the following theorem from Kirby calculus — see eg [52, Theorem 2.2.11]:

Theorem 9.2 *Denote by $\mathbb{R}^3_L$ a 3–manifold determined by a surgery diagram where each component L_i of L has Dehn surgery coefficient p_i/q_i for relative prime integers p_i and q_i. Then $H_1(\mathbb{R}^3_L)$ is presented with generators μ_i and relations*

$$p_i\mu_i + q_i \sum_{j\neq i} \operatorname{lk}(L_i, L_j)\mu_j = 0,$$

where $\operatorname{lk}(L_i, L_j)$ is the linking number.

When performing contact surgery on the component Λ_i of Λ, the meridian μ_i bounding a core disk of the surgery handle is sent to

$$\mu_i + c_i\lambda_{\xi,i} = (1 + c_i \operatorname{tb}(L_i))\mu_i + c_i\lambda_i.$$

Thus, for Legendrian knots in $\mathbb{R}^3$, contact ±1 surgery on Λ_i is topologically a $(\operatorname{tb}(\Lambda_i)\pm 1)$ surgery. From this computation, the calculation of $H_1(M)$ in Theorem 9.1 is then immediate.

9.3 Homology classes of Reeb orbits

In this section we describe how to compute homology classes of the Reeb orbits of α_ϵ. Our strategy will be to homotop orbits to the complement of N_ϵ in $\mathbb{R}^3_{\Lambda^\pm}$, after which the following computational tool may be applied:

Theorem 9.3 *Let γ be an oriented link in $\mathbb{R}^3 \setminus L$. Then the homology classes of γ in $H_1(\mathbb{R}^3 \setminus L)$ and $H_1(\mathbb{R}^3_L)$ is given by*

$$[\gamma] = \sum_i \operatorname{lk}(\gamma, L_i)\mu_i.$$

Proof Assume that γ is embedded and let $S \subset \mathbb{R}^3$ be a Seifert surface which transversely intersects the L_i. Punch holes in S near its intersections with the L_i, producing a surface S' which is disjoint from L and whose oriented boundary is a union of γ and a linear combination $\sum a_i\mu_i$. Then S' provides a cobordism from γ to these μ_i, providing an equivalence $[\gamma] = \sum a_i\mu_i$ in homology. By the definition of lk, we have $a_i = \operatorname{lk}(\gamma, L_i)$. □

Warning 9.4 The homotopies which we apply to closed Reeb orbits γ are not guaranteed to preserve the isotopy class of their embedding in $\mathbb{R}^3_{\Lambda^\pm}$ (assuming γ is embedded).

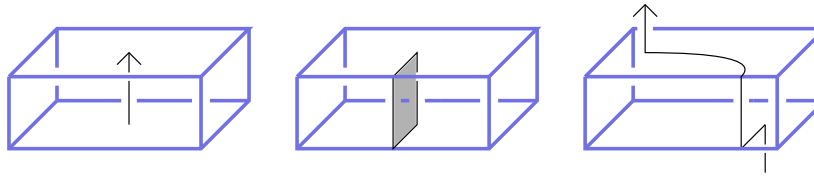


Figure 17: Homotoping a Reeb orbit into $\mathbb{R}^3 \setminus N_\epsilon$ as it passes through a $c = +1$ surgery handle.

Figure 17 demonstrates how to homotop a segment of a Reeb orbit γ into the exterior of the surgery handle N_ϵ as it passes through a component $N_{\epsilon,i}$ for which $c_i = 1$. The boxes represent the surgery handles with ∂_p pointing into the page, ∂_q pointing to the left, and ∂_z pointing up. On the left we have an arc parallel to the Reeb vector field entering the handle as seen from the inside of $N_{\epsilon,i}$. The arc extends in the ∂_z direction through the handle, along which it can be realized as being contained in the boundary of a square of the form $\{p \leq p_0, q = q_0\}$, depicted in gray. On the right, we see intersection of the boundary of this square with ∂N_ϵ as seen from the outside of the surgery handle $\mathbb{R}^3 \setminus N_\epsilon$. By homotoping γ across the gray disk, we obtain this arc shown on the right.

Figure 18 demonstrates the same procedure for orbits as they pass through surgery handles with surgery coefficient -1. In this case we consider squares of the form $\{p \geq p_0, q = q_0\}$ through which we homotop our arcs. Note that our choice of homotopy for both surgery coefficients is such that the homotoped arcs traverse ∂N in the ∂_q direction in which the components of Λ are oriented.

For a Reeb orbit γ, we can perform homotopies as described above at the tips of all chords in $\mathrm{cw}(\gamma)$ to push it to the exterior of N_ϵ. Away from the chords, we may arrange that the homotoped orbit traverses the $p = \mp\epsilon$ side of $N_{\epsilon,i}$ when the surgery coefficient of Λ_i is ± 1. The image of γ after homotopy is shown in the Lagrangian projection in Figure 19.

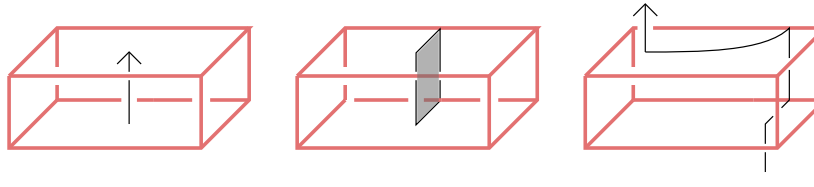


Figure 18: Homotoping a Reeb orbit into $\mathbb{R}^3 \setminus N_\epsilon$ as it passes through a $c = -1$ surgery handle.

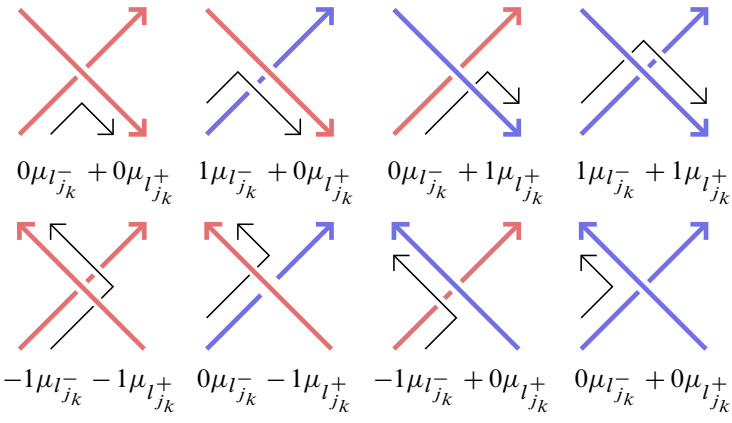


Figure 19: Here the homotopy described in Figures 17 and 18 are depicted in the Lagrangian projection. The top (bottom) row shows positive (negative) crossings of $\Lambda^{\pm}$. Each subfigure may be rotated by π. Local contributions to linking numbers are indicated below each subfigure.

The computation of homology classes of orbits in Theorem 9.1 then amounts to packaging the above observations algebraically:

Proof of Theorem 9.1 We homotop γ to $\mathbb{R}^3 \setminus N_\epsilon$ as described above and then apply Theorem 9.3. We write γ' for the image of γ under the homotopy. The linking number of two knots in $\mathbb{R}^3$ may be computed from a diagram as half of the signed count of crossings in the diagram. Therefore, in order to compute $[\gamma]$, it suffices to show that the signed count of crossings between γ' and each Λ_i is given by the μ_i coefficients in $\sum(\mathrm{cr}_{j_k} + \mathrm{cr}_{j_k, j_{k+1}})$.

In a neighborhood of a crossing, γ' will be as depicted in Figure 19 in the Lagrangian projection, where the contribution to the signed count of crossings between γ and the Λ_i is given by the terms

$$\tfrac{1}{2}((c^-_{j_k} + \mathrm{sgn}_{j_k})\mu_{l^-_{j_k}} + (c^+_{j_k} + \mathrm{sgn}_{j_k})\mu_{l^+_{j_k}}).$$

The formula may be verified on a case-by-case basis for each of the eight components of the figure. This is exactly the definition of cr_{j_k} given in (12).

Away from a crossing, γ' will continue following alongside arc components of the Λ_i, to the right (in the $p > 0$ direction) of Λ when the component of Λ has coefficient -1 and to the left otherwise as it travels from a crossing j_k to j_{k+1}. The contributions to the signed count of crossings with each of the Λ_i are given by the coefficients of μ_i in $\mathrm{cr}_{j_k, j_{k+1}}$ in the formula, as is clear from the definition of the crossing monomial. □

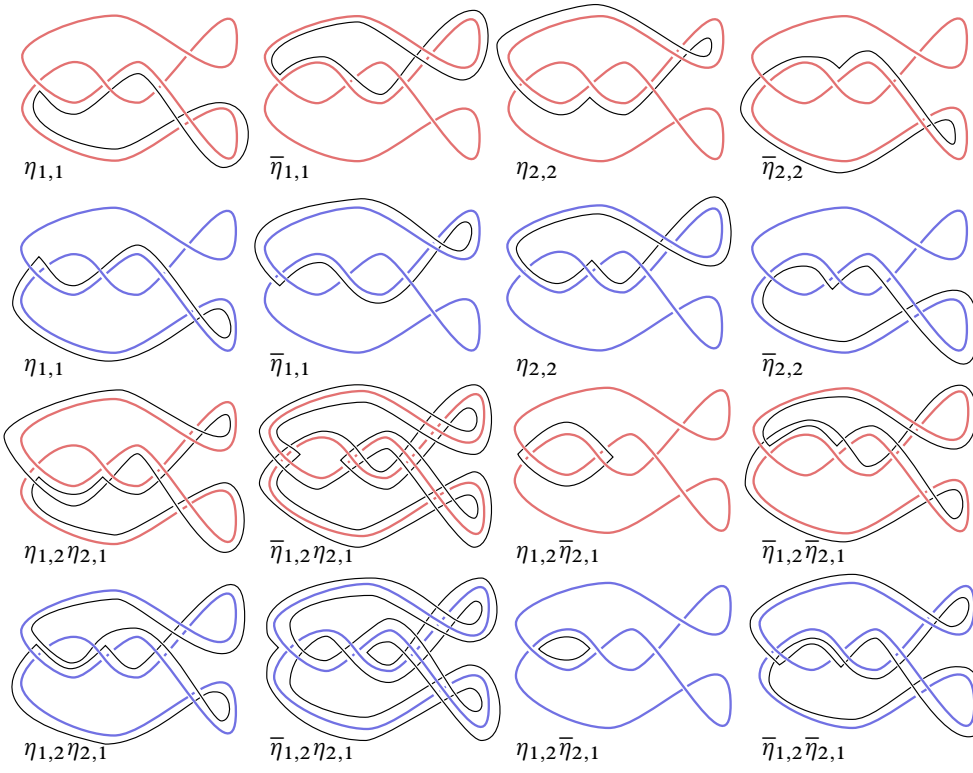


Figure 20: Push-outs of Reeb orbits in $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$, where Λ is the trefoil of Figure 6. Default orientations for Λ and hence for capping paths are determined by the arrow on Λ appearing in that figure. Each subfigure is labeled (to its lower-left) with the capping paths which determine the homotopy shown with homotoped Reeb orbits appearing in black.

9.4 Push-outs of Reeb orbits

We've demonstrated how squares of the form $\{p \leq p_0, q = q_0\} \subset N_\epsilon$ in the case of $+1$ surgery and of the form $\{p \geq p_0, q = q_0\}$ in the case of -1 surgery are used to homotop Reeb orbits into $\mathbb{R}^3_{\Lambda^\pm} \setminus N_\epsilon = \mathbb{R}^3 \setminus \epsilon$ so that the homotoped circles ride along some $\eta_{j_1,j_2} \subset \Lambda$ according to its prescribed orientation.

Squares of the form $\{p \geq p_0, q = q_0\}$ inside of a $c_i = +1$ component of Λ and of the form $\{p \leq p_0, q = q_0\}$ inside of a $c_i = -1$ component could also be used. As may be checked with the same local model — Figures 17 and 18 — but with opposite the prescribed orientation for Λ, we may use these squares to homotop an orbit γ to $\mathbb{R}^3 \setminus \epsilon$. Using these squares will result in the homotoped arcs riding along some $\overline{\eta}_{j_1,j_2} \subset \Lambda$.

We then have two choices of homotoping square each time our orbit γ passes through N_ϵ, with each choice corresponding to a choice of either a η_{j_1,j_2} or a $\overline{\eta}_{j_1,j_2}$. Hence, for a Reeb orbit $\gamma = (r_{j_1} \cdots r_{j_n})$, a choice of $\zeta_1, \dots, \zeta_n$ with each $\zeta_j \in \{\eta_{j_k,j_{k+1}}, \overline{\eta}_{j_k,j_{k+1}}\}$ determines a means of homotoping γ into $\mathbb{R}^3 \setminus N_\epsilon$.

Definition 9.5 Provided $\zeta_1, \dots, \zeta_n$ as above, we say that the homotopy class of the map of the circle in $\mathbb{R}^3 \setminus N_\epsilon$ determined by homotoping γ as described above is the *push-out of* $\zeta_1, \dots, \zeta_n$.

In other words, each orbit string — recall Definition 3.7 — determines instructions for homotoping γ into the complement of the surgery locus. Various examples are depicted in Figure 20, displaying all push-outs for orbits (r_1), (r_2) and $(r_1 r_2)$ for $\mathbb{R}^3_{\Lambda^\pm}$, where Λ is the trefoil of Figure 6 for both choices of surgery coefficient.

10 Surgery cobordisms and Lagrangian disks

The purpose of this section is to build symplectic cobordisms between the $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ with specialized properties. We consider the following setup: Take $\Lambda \subset (\mathbb{R}^3, \xi_{\text{std}})$ in good position with $\Lambda^0 \subset \Lambda$ nonempty. After performing surgery on $\Lambda^\pm \subset \Lambda$, we have a contact form α_ϵ on $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ and consider Λ^0 as a Legendrian link in $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$. We choose a constant $c = \pm 1$ and denote the contact manifold obtained by performing contact c surgery along $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ by $(\mathbb{R}^3_\Lambda, \xi_\Lambda)$; we also denote the contact form on $(\mathbb{R}^3_\Lambda, \xi_\Lambda)$ so obtained by α_ϵ. We write N^0_ϵ for a standard neighborhood of $\Lambda^0 \subset \mathbb{R}^3_{\Lambda^\pm}$, as described in Section 4.4, of size ϵ.

Theorem 10.1 *For any $\epsilon > 0$, there exists a positive constant $C > 0$ and a Liouville cobordism (W_c, λ_c) with the following properties*:

(1) *If $c = +1$, the convex end of the cobordism is $(\mathbb{R}^3_{\Lambda^\pm}, e^C \alpha_\epsilon)$ and the concave end is $(\mathbb{R}^3_\Lambda, e^{-C} \alpha_\epsilon)$.*

(2) *If $c = -1$, the convex end of the cobordism is $(\mathbb{R}^3_\Lambda, e^C \alpha_\epsilon)$ and the concave end is $(\mathbb{R}^3_{\Lambda^\pm}, e^{-C} \alpha_\epsilon)$.*

(3) *(W_c, λ_c) contains a disjoint collection of disks $\mathbb{D}_{c,i}$ along which $\lambda_c = 0$, bounding Λ^0 in the convex end of the cobordism when $c = +1$ and bounding Λ^0 in the concave end of the cobordism when $c = -1$.*

(4) *A finite symplectization*

$$([-C, C] \times (\mathbb{R}^3_{\Lambda^\pm} \setminus N^0_\epsilon), e^t \alpha_\epsilon)$$

of $(\mathbb{R}^3_{\Lambda^\pm}\setminus N^0_\epsilon)$ is contained in (W_c,λ_c), so that the restriction of its inclusion map to $(\partial[-C,C])\times(\mathbb{R}^3_{\Lambda^\pm}\setminus N^0_\epsilon)$ provides the obvious inclusions into $(\mathbb{R}^3_{\Lambda^\pm}, e^{\pm C}\alpha_\epsilon)$ and $(\mathbb{R}^3_{\Lambda}, e^{\pm C}\alpha_\epsilon)$.

We will construct (W_c,λ_c) by attaching 4–dimensional surgery handles to $\mathbb{R}^3_{\Lambda^\pm}$. As mentioned in the above theorem, the key properties of our cobordism are that

(1) we get exactly the contact forms α_ϵ on its boundaries, and

(2) all of the perturbations required to achieve this end happen within a standard neighborhood of Λ^0 whose size shrinks as ϵ tends to zero.

Then the analysis of Sections 4–7 applies to contact forms on the ends of our cobordisms without modification.

We are only slightly modifying known handle-attachment constructions — corresponding to the case $c=-1$ above — such as appear in Weinstein's original work [60] and Ekholm [18].

An outline of this section is as follows:

(1) In Section 10.1 we collect lemmas required to perturb contact forms on contactizations, being particularly interested in standard neighborhoods of Legendrian knots.

(2) In Section 10.2 we describe a square surgery handle sitting inside of $\mathbb{R}^4$ and outline the properties of its ambient geometry.

(3) In Section 10.4 we flatten the corners of the handle to prepare for later attachment.

(4) In Section 10.5 we describe Reeb dynamics on the convex end of this handle, showing that its flow is described as a Dehn twist.

(5) In Section 10.6 we modify the handle so that the Dehn twist determined by the Reeb flow is a linear Dehn twist as described in the gluing construction of Section 4.7.

(6) In Section 10.7 we finalize our construction by attaching our handle to finite symplectizations of $(\mathbb{R}^3_{\Lambda^\pm},\alpha_\epsilon)$.

10.1 Geometry of 1–forms on contactizations and their symplectizations

Let $(I\times W,\alpha=dz+\beta)$ be a contactization of an exact symplectic manifold (W,β) as in Section 2.3.2.

10.1.1 ξ–preserving perturbations We first look at how the Reeb vector field changes if we multiply α by a positive function, thereby preserving the contact structure.

Lemma 10.2 *Given $H \in \mathcal{C}^\infty(I \times W)$, the Reeb vector field R_H of the contact form*

$$\alpha_H = e^H(dz + \beta)$$

on $I \times W$ is

$$R_H = e^{-H}\Big((1 + \beta(X_H))\partial_z - X_H - \frac{\partial H}{\partial z} X_\beta\Big),$$

where X_H is computed with respect to $d\beta$.

This is a straightforward computation. We'll be interested in the following special case:

Lemma 10.3 *Suppose that $H = H(z, p)$ is a smooth function on $I \times I \times S^1$. Then the Reeb vector field of $\alpha_H = e^H(dz + p\,dq)$ is*

$$R_H = e^{-H}\Big(\Big(1 + p\frac{\partial H}{\partial p}\Big)\partial_z - \frac{\partial H}{\partial p}\partial_q - p\frac{\partial H}{\partial z}\partial_p\Big)$$

and the function pe^H is invariant under $\mathrm{Flow}^t_{R_H}$.

For the last item, we see that the projection of R_H onto the (z, p) coordinates is the Hamiltonian vector field associated to $dp \wedge dz$ and the function pe^{-H}.

10.1.2 ξ–modifying perturbations Now we study perturbations of α which modify ξ. Similar modifications of contact forms appear in [4, Definition 3.1.1; 14, Corollary 2.5].

Lemma 10.4 *Given a smooth function $h \in \mathcal{C}^\infty(I \times W, (0, \infty))$, the contact form*

$$\alpha_h = h\,dz + \beta$$

is contact if and only if

$$h\,d\beta + \beta \wedge dh$$

is a symplectic form on each $\{z\} \times W$. If this form is contact, its Reeb vector field R_h is

$$R_h = (h - \beta(X_h))^{-1}(\partial_z - X_h),$$

where X_h is computed with respect to $d\beta$. The contact structure $\xi_h = \ker(\alpha_h)$ is given by

$$\xi_h = \{hV - \beta(V)\partial_z : V \in TW\}.$$

The following technical result will allow us to modify the Reeb vector field on standard neighborhoods of Legendrians so that the flow map from the bottom to the top of the neighborhood realizes a Dehn twist τ_g associated to a function g. For applications to surgery later in this section, it will be important to keep track of the size of our neighborhood.

Proposition 10.5 *For positive constants $\epsilon_p, \epsilon_g > 0$, let $g = g(p)\colon I_{\epsilon_p} \to \mathbb{R}$ be a smooth function which vanishes for all orders on ∂I_{ϵ_p} and satisfies the pointwise bound $|g(p)| \leq \epsilon_g$. Then, for constants ϵ_z and ϵ_t satisfying*

$$\epsilon_p \epsilon_g \leq \tfrac{1}{2}\epsilon_z, \frac{\epsilon_t \epsilon_z}{2(1+\epsilon_t)},$$

there exists a function $h = h(z, p)$ on $I_{\epsilon_z} \times I_{\epsilon_p}$ and an exact symplectic manifold

$$([-\epsilon_t, 0] \times I_{\epsilon_z} \times I_{\epsilon_p} \times S^1, \lambda)$$

such that the following conditions hold:

(1) $\lambda|_{\{-\epsilon_t\}\times I_{\epsilon_z}\times I_{\epsilon_p}\times S^1} = e^{-\epsilon_t}(dz + p\,dq)$.

(2) $\lambda|_{\{0\}\times I_{\epsilon_z}\times I_{\epsilon_p}\times S^1} = \alpha_h$, *where α_h is as in Lemma 10.4 for a positive function h.*

(3) $s\alpha_h + (1-s)(dz + p\,dq)$ *is contact for all* $s \in [0, 1]$.

(4) $\alpha_h - (dz + p\,dq)$ *and all of its derivatives vanish along* $\partial(I_{\epsilon_z} \times I_{\epsilon_p} \times S^1)$.

(5) *The Reeb vector field R_h of α_h satisfies $dz(R_h) > 0$ everywhere.*

(6) *For each point $(p, q) \in I_{\epsilon_p} \times S^1$ a flow-line of R_h passing through $(-\epsilon_z, p, q)$ will pass through $(\epsilon_z, p, q + g(p))$.*

(7) *The Liouville vector field of λ agrees with ∂_t on a collar neighborhood of the boundary of its domain.*

Proof We first outline the contact forms we'll need. Consider functions of the form $h = 1 + F(z)G(p)$ on $I_{\epsilon_z} \times I_{\epsilon_p}$ and 1–forms

$$\alpha_h = h\,dz + p\,dq$$

as studied in Lemma 10.4. We assume $F \geq 0$ and that both F and G and all of their derivatives vanish on collar neighborhoods of the boundary of their domains. By Lemma 10.4, α_h is contact if and only if

$$0 < 1 + FG - pF\frac{\partial G}{\partial p}. \tag{46}$$

Second we outline the construction of Liouville forms which interpolate between $\alpha = dz + p\,dq$ and α_h. Consider functions E on an interval $[-\epsilon_t, 0]$ satisfying $E(-\epsilon_t) = 0$ and $E(0) = 1$ with $\partial^k E/\partial t^k = 0$ for all $k > 0$ at the endpoints of its domain and $\partial E/\partial t \geq 0$ everywhere. Define a 1–form

$$\lambda_{EFG} \in \Omega^1([0, \epsilon_t] \times I_{\epsilon_z} \times I_{\epsilon_p} \times S^1)$$

determined by

$$\lambda_{EFG} = e^t\big((1 + E(t)F(z)G(p))\,dz + p\,dq\big). \tag{47}$$

Then we compute

$$d\lambda_{EFG} \wedge d\lambda_{EFG} = e^{2t}\left(1 + EFG - pEF\frac{\partial G}{\partial p} + \frac{\partial E}{\partial t}FG\right) dt \wedge dz \wedge dp \wedge dq. \tag{48}$$

We seek to specify the E, F and G so that

(1) α_h is contact and its flow determines a Dehn twist by g,

(2) $d\lambda_{EFG}$ is symplectic, and

(3) the sizes of our neighborhood and symplectic cobordism — governed by the constants ϵ_z and ϵ_t — are reasonably small.

First we show that G is determined by g. If α_h is contact, its Reeb vector field is

$$R_h = \left(1 + FG - pF\frac{\partial G}{\partial p}\right)^{-1}\left(\partial_z - F\frac{\partial G}{\partial p}\partial_q\right).$$

This Reeb vector field is particularly friendly in that it preserves p and provides us with a separable ODE. For, provided an initial condition (z_0, p_0, q_0) and some $z > z_0$, we see that, after some time $t > 0$, $\mathrm{Flow}^t_{R_h}$ will pass through the point (z, p_0, q) with

$$q = q_0 - \frac{\partial G}{\partial p}\int_{z_0}^{z} F(Z)\, dZ.$$

In order to realize the flow from $\{-\epsilon_z\}\times\mathbb{R}\times S^1$ to $\{\epsilon_z\}\times\mathbb{R}\times S^1$ as a Dehn twist by g, we need

$$G(p) = -\left(\int_{-\epsilon_z}^{\epsilon_z} F(z)\, dz\right)^{-1}\int_{-\infty}^{p} g(P)\, dP.$$

This quantity is well defined by our presumption that g is compactly supported.

With this choice of G, the contact condition provided by (46) is equivalent to

$$F\cdot\left(\int_{-\infty}^{p} g(P)\, dP - pg(p)\right) \le \int_{-\epsilon_z}^{\epsilon_z} F(z)\, dz \tag{49}$$

for all (z, p, q). The condition that $d\lambda_{EFG}$ is symplectic provided by (48) is equivalent to

$$EF\cdot\left(\int_{-\infty}^{p} g(P)\, dP - pg(p)\right) + \frac{\partial E}{\partial t}F\cdot\left(\int_{-\infty}^{p} g(P)\, dP\right) \le \int_{-\epsilon_z}^{\epsilon_z} F(z)\, dz. \tag{50}$$

Now choose F and a constant ϵ_F so that

$$\epsilon_F = \sup|F(z)|, \quad \epsilon_F\epsilon_z = \int_{-\epsilon_z}^{\epsilon_z} F(z)\, dz.$$

It's easy to see by drawing pictures of bump functions that these choices can be made. Then (49) is satisfied so long as

$$\epsilon_p\epsilon_g \le \tfrac{1}{2}\epsilon_z$$

and, since $0 \leq E \leq 1$, we have that (50) is satisfied so long as

$$\Big(2 + \frac{\partial E}{\partial t}\Big)\epsilon_p \epsilon_g \leq \epsilon_z.$$

Choose E so that $\sup \partial E/\partial t = 2/\epsilon_t$. Then this last inequality, which we seek to satisfy, becomes

$$2(1+\epsilon_t^{-1})\epsilon_p\epsilon_g \leq \epsilon_z \iff \epsilon_p\epsilon_g \leq \frac{\epsilon_t\epsilon_z}{2(1+\epsilon_t)}. \qquad \square$$

10.2 The square handle

Having established the above lemmas, we proceed with the construction of our symplectic handle. Here we construct a square Weinstein handle sitting in $\mathbb{R}^4$.

Consider the Liouville form on $\mathbb{R}^4 = \mathbb{C}^2$,

$$\lambda_0 = \sum_1^2 2x_i\, dy_i + y_i\, dx_i.$$

This is a potential for the standard symplectic form $d\lambda_0 = dx_i \wedge dy_i$ with Liouville vector field

$$X_{\lambda_0} = 2x_i\,\partial_{x_i} - y_i\,\partial_{y_i},$$

whose time t flow is given by

$$\mathrm{Flow}^t_{X_{\lambda_0}}(x,y) = (e^{2t}x, e^{-t}y). \tag{51}$$

For $\rho_0 > 0$, consider also the convex set with corners

$$\mathbb{D}_{\rho_0} \times \mathbb{D} = \{|x| \leq \rho_0,\, |y| \leq 1\} \subset \mathbb{R}^4,$$

whose smooth boundary strata we denote by

$$M^+_{\rho_0} = \partial\mathbb{D}_{\rho_0} \times \mathbb{D}, \quad M^-_{\rho_0} = \mathbb{D}_{\rho_0} \times \partial\mathbb{D}.$$

Then X_{λ_0} is positively transverse to the $M^\pm_{\rho_0}$ if we equip $M^+_{\rho_0}$ with the outward-pointing orientation and equip $M^-_{\rho_0}$ with its inward-pointing orientation. Therefore, $\lambda_0|_{M^\pm}$ is contact. Applying $\mathrm{Flow}^t_{X_{\lambda_0}}$ for $t \in (-\infty, 0]$, we have embeddings of the negative half-infinite symplectizations of the $(M^\pm_{\rho_0}, \lambda_0|_{TS^\pm})$ into $\mathbb{R}^4$,

$$\mathrm{Flow}^t_{X_{\lambda_0}} \circ i^\pm \colon (-\infty, 0] \times M^\pm_{\rho_0} \to \mathbb{R}^4, \tag{52}$$

where $i^\pm \colon M^\pm_{\rho_0} \to \mathbb{R}^4$ denote the inclusion mappings.

10.2.1 Reeb trajectories across the square handle The Reeb vector field R_{ρ_0} along $M^+_{\rho_0}$ is

$$R_{\rho_0} = \frac{1}{2\rho_0^2} x_i\,\partial_{y_i} \implies \mathrm{Flow}^t_{R_{\rho_0}}(x,y) = \Big(x, y + \frac{t}{2\rho_0^2}x\Big).$$

Starting at points $(x_\theta, y_0) = (\rho_0 \cos(\theta), \rho_0 \sin(\theta), 1, 0)$, Reeb trajectories are

$$\mathrm{Flow}^t_{R_{\rho_0}}(x_\theta, y_0) = \left(\rho_0 \cos(\theta), \rho_0 \sin(\theta), 1 + \frac{t}{2\rho_0}\cos(\theta), \frac{t}{2\rho_0}\sin(\theta)\right).$$

In order that such a trajectory does not immediately exit $M^+_{\rho_0}$ (maintaining the condition $|y| \le 1$ for small $t \ge 0$), we must have $\theta \in \left[\frac{1}{2}\pi, \frac{3}{2}\pi\right]$. These trajectories touch $\partial M^+_{\rho_0}$ when

$$1 = \left(1 + \frac{t}{2\rho_0}\cos(\theta)\right)^2 + \left(\frac{t}{2\rho_0}\sin(\theta)\right)^2 \iff -4\rho_0 \cos(\theta) = t,$$

at which point the y coordinate will be

$$\begin{aligned} y_\theta &= (1 - 2\cos^2(\theta), -2\cos(\theta)\sin(\theta)) = (-\cos(2\theta), -\sin(2\theta)) \\ &= (\cos(2\theta - \pi), \sin(2\theta - \pi)). \end{aligned}$$

We can then measure the angle from y_0 to y_θ as $2\theta - \pi \in [0, 2\pi]$.

10.3 Identification of the concave end of the handle as a 1–jet space

We define an embedding of a standard neighborhood of a Legendrian into M_1^- as

$$\Phi_-(z, p, q) = \left(z\cos - \frac{p}{2\pi}\sin, z\sin + \frac{p}{2\pi}\cos, \cos, \sin\right),$$

where the arguments of cos and sin are both $2\pi q$. The map parametrizes M_1^- so that

(1) $2\pi q$ is the angle in the y–plane,

(2) $z = x \cdot y$,

(3) $p = x \cdot \partial y/\partial q$, and

(4) $|x|^2 = z^2 + (p/2\pi)^2$.

The tangent map of Φ_- is

$$T\Phi_- = \begin{pmatrix} \cos & -(1/2\pi)\sin & -2\pi z \sin - p\cos \\ \sin & (1/2\pi)\cos & 2\pi z \cos - p \sin \\ 0 & 0 & -2\pi \sin \\ 0 & 0 & 2\pi \cos \end{pmatrix}$$

with incoming basis $\{\partial_z, \partial_p, \partial_q\}$ and outgoing basis $\{\partial_{x_1}, \partial_{x_2}, \partial_{y_1}, \partial_{y_2}\}$, from which it follows that

$$\Phi_-^* \lambda_0 = dz + p\, dq.$$

We can extend Φ_- to an embedding of the symplectization of the 1–jet space into $\mathbb{R}^4$ by

$$\begin{aligned}(53)\quad \overline{\Phi}_-(t,z,p,q) &= \mathrm{Flow}^t_{X_{\lambda_0}} \circ \Phi_-(z,p,q)\\ &= \Big(e^{2t}\Big(z\cos -\frac{p}{2\pi}\sin\Big), e^{2t}\Big(z\sin +\frac{p}{2\pi}\cos\Big), e^{-t}\cos, e^{-t}\sin\Big).\end{aligned}$$

By (51) and $\mathcal{L}_{X_{\lambda_0}}\lambda_0 = \lambda_0$, we have

$$(54)\qquad \overline{\Phi}_-^*\lambda_0 = e^t(dz + p\,dq).$$

10.4 Shaping the handle

Here we shape our handle so that the manifold obtained by the handle attachment will be smooth. Moreover, we will choose a specific shape which allows us to control Reeb dynamics on the surgered contact manifold.

Pick a positive constant $\rho_1 < \rho_0$ and a smooth function $B = B(\rho)\colon (0,\infty) \to [0,\infty)$ satisfying the conditions

(1) $B(\rho) = \log\sqrt{\rho_0/\rho} = -\frac{1}{2}(\log\rho - \log\rho_0)$ for $\rho \in (0,\rho_1)$,

(2) $B(\rho) = 0$ for $\rho > \rho_0$, and

(3) $0 \le -\partial B/\partial\rho < \rho^{-1}$ everywhere.

Along $\rho \in (0,\rho_1)$, we have $\partial B/\partial\rho = -1/2\rho$, so that our last condition is satisfied. To find such a function B, we can take a smoothing of the piecewise-smooth function

$$(55)\qquad B^{\mathrm{PW}}(\rho) = \begin{cases} \log\sqrt{\rho_0/\rho} & \text{if } \rho \le \rho_0,\\ 0 & \text{if } \rho \ge \rho_0.\end{cases}$$

Let $N = I \times I \times S^1$ be a standard neighborhood of a Legendrian Λ with $\Lambda = \{0\}\times\{0\}\times S^1$. Using the function B, we define an embedding

$$\Phi_H\colon (N\setminus\Lambda) \to \mathbb{R}^4, \quad \Phi_H = \mathrm{Flow}^H_{X_{\lambda_0}} \circ \Phi_-,$$

where

$$H(p,z) = B(\rho(p,z)), \quad \rho(p,z) = \sqrt{z^2 + \Big(\frac{p}{2\pi}\Big)^2}.$$

We outline some important properties of the map Φ_H:

(1) From (54), $\Phi_H^*\lambda_0 = e^H(dz + p\,dq)$.

(2) Along the set $\{z^2 + (p/2\pi)^2 \ge \rho_0\}$, Φ_H is the same as Φ_-.

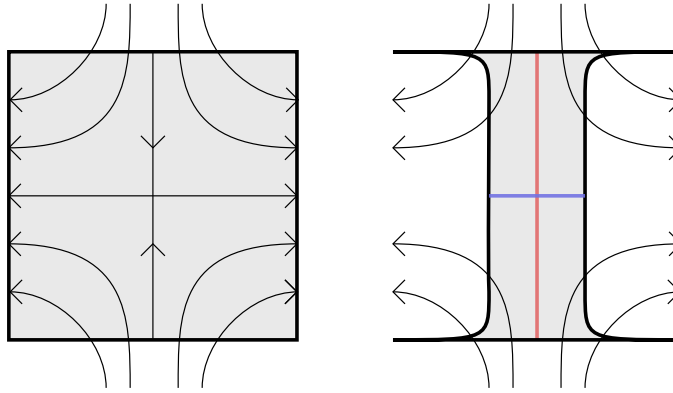


Figure 21: On the left is the square handle $\mathbb{D}_{\rho_0} \times \mathbb{D}$ and on the right is the handle W_H. Flow-lines of X_{λ_0} transversely pass through $M^{\pm}_{\rho_0}$ and M_H. The Lagrangian disks $\{x = 0\}$ and $\{y = 0\}$ are shown in red and blue, respectively.

(3) From the first property characterizing B and (53), we see that, on the set $\{\rho \le \rho_1\}$, the x and y coordinates of the embedding satisfy

$$\begin{aligned} |x \circ \Phi_H(z, p, q)| &= e^{2H}\sqrt{z^2 + \left(\frac{p}{2\pi}\right)^2} = \rho_0, \\ |y \circ \Phi_H(z, p, q)| &= e^{-H} = \sqrt{\frac{\rho}{\rho_0}}. \end{aligned} \tag{56}$$

From the last equation, we have the equivalences

$$\begin{aligned} \Phi_H(\{\rho \le \rho_1\}) &= \left\{|x| = \rho_0, |y| \le \sqrt{\frac{\rho_1}{\rho_0}}\right\} \setminus \{y = 0\}, \\ \overline{\Phi_H(\{\rho \le \rho_1\})} &= \left\{|x| = \rho_0, |y| \le \sqrt{\frac{\rho_1}{\rho_0}}\right\}. \end{aligned}$$

The closure of the image of Φ_H in $\mathbb{R}^4$ is a smooth hypersurface M_H which is positively transverse to X_{λ_0}. We write M_H for this hypersurface and define $W_H \subset \mathbb{R}^4$ to be the set enclosed by M_1^- and M_H. The handle W_H is depicted in Figure 21, right.

10.5 Analysis of R_H over M_H

Here we analyze dynamics on M_H of the Reeb vector field R_H for the contact form $\alpha_H = \lambda_0|_{M_H}$. Because of our use of the imprecisely defined function B, we won't be able to solve for $\mathrm{Flow}^t_{R_H}$ explicitly. However, we'll be able to capture enough information about this flow for the applications to handle attachment.

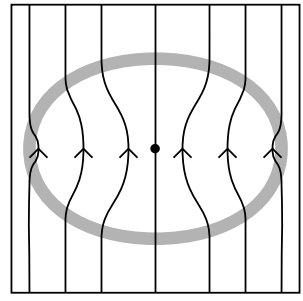


Figure 22: Projections of flow-lines of R_H to the (p,z) coordinates. By Lemma 10.3, these are the level sets of pH. Here p points to the right and z points upward. The dot represents the circle $|x|=0$ along which our flow is not defined in the (z,p,q) coordinate system. The region $\rho\in[\rho_1,\rho_0]$ where the function B^{PW} is smoothed to obtain B is shaded.

On the complement of the set $\{z=p=0\}$, our contact form is $\alpha_H=e^H(dz+p\,dq)$, whose Reeb field will be denoted by R_H. Writing $\rho=\rho(p,z)$, we compute

$$dH=\frac{\partial B}{\partial\rho}\,d\rho,\quad d\rho=\rho^{-1}\Big(z\,dz+\frac{p}{(2\pi)^2}\,dp\Big).$$

Now apply Lemma 10.3 to compute R_H using the coordinates (z,p,q) on $M_H\setminus\{y=0\}$ as

$$R_H=e^{-H}\Big(\Big(1+\frac{\partial B}{\partial\rho}\rho^{-1}\Big(\frac{p}{2\pi}\Big)^2\Big)\partial_z-\frac{p}{\rho}\frac{\partial B}{\partial\rho}\Big(\frac{1}{(2\pi)^2}\partial_q+z\partial_p\Big)\Big). \tag{57}$$

Here is a collection of observations regarding R_H and its flow:

(1) The ∂_z part of R_H is always strictly positive. This is a consequence of the inequalities

$$-\frac{\partial B}{\partial\rho}\rho^{-1}\Big(\frac{p}{2\pi}\Big)^2\le-\frac{\partial B}{\partial\rho}\rho^{-1}\rho^2=-\frac{\partial B}{\partial\rho}\rho<1\implies 1+\frac{\partial B}{\partial\rho}\rho^{-1}\Big(\frac{p}{2\pi}\Big)^2>0$$

following from the definition of ρ and the third defining property of the function B.

(2) For each p and $\epsilon>0$, a flow-line starting at the point $(-\epsilon,p,q)$ will pass through some (ϵ,p,q'). This follows from the facts that $(pe^{-H})(-z,p)=(pe^{-H})(z,p)$ and that the projection of R_H onto the (z,p)–plane is Hamiltonian with respect to $dp\wedge dz$ as per Lemma 10.3. See Figure 22.

(3) The flow-line passing through $(-\epsilon,0,q)$ will pass through the point $\big(\epsilon,0,q+\frac{1}{2}\big)$. To see this, observe that such a flow-line with such an initial condition must flow up into the circle $\{z=p=0\}$ along the line $\{p=0\}$ and compare with the definition of the map Φ_H.

(4) A twist map $f_{H,\rho_0}\colon I_{2\pi\rho_0} \to S^1$ is defined by following the flow-line of R_H passing through $(-\epsilon, p, q)$ to a point $(\epsilon, p, q + f_{H,\rho_0}(p))$. By the properties we've used to specify B, the derivatives of f_{H,ρ_0} are supported on $I_{2\pi\rho_0}$ as R_H coincides with ∂_z outside of this region. Likewise, $R_H = \partial_z$ on $\{|\rho| \geq \rho_0\}$.

(5) The twist map satisfies

$$f_{H,\rho_0}(-p) = -f_{H,\rho_0}(p), \quad f_{H,\rho_0}(-2\pi\rho_0) = f_{H,\rho_0}(2\pi\rho_0) = 0.$$

The first equality follows from the fact that the ∂_z factor of R_H is a function of p^2 while the ∂_p and ∂_q factors are antisymmetric in p. The second equality follows from the previous item.

(6) As $\partial B/\partial p \leq 0$, the ∂_q coefficient of R_H from (57) has sign equal to $\operatorname{sgn}(p)$ where is it nonzero. Hence, f_{H,ρ_0} always twists to the right for $p < 0$ and to the left along $p > 0$. This is the expected behavior of a positive Dehn twist.

Proposition 10.6 *Write $\tilde{f}_{H,\rho_0}\colon I_{2\pi\rho_0} \to \mathbb{R}$ for the lift of the twist map $f_{H,\rho_0}\colon I_{2\pi\rho_0} \to \mathbb{R}$ with initial condition*

$$\tilde{f}_{H,\rho_0}(-2\pi\rho_0) = 0 \quad \Longrightarrow \quad \tilde{f}_{H,\rho_0}(0) = -\tfrac{1}{2}, \quad \tilde{f}_{H,\rho_0}(2\pi_{\rho_0}) = -1$$

by the preceding analysis. Suppose that $\tilde{f}\colon I_{2\pi_{\rho_0}} \to [-1, 0]$ is a decreasing function also satisfying the above equalities. Then, for H constructed using a function B which is sufficiently C^0–close to the function B^{PW}, the estimate

$$|\tilde{f}_{H,\rho_0}(p) - \tilde{f}(p)| \leq \tfrac{1}{2} \tag{58}$$

is satisfied for all $p \in I_{2\pi\rho_0}$.

Proof The analysis of Section 10.2.1 provides a very explicit approximation of the function $\tilde{f}_{H,\rho_0}$. Let's consider the degenerate case when $B = B^{\mathrm{PW}}$ as described in (55), writing H^{PW} for the associated piecewise-smooth function. Then $M_{H^{\mathrm{PW}}}$ will be piecewise smooth as a submanifold of $\mathbb{R}^4$. We have a C^0 flow on $M_{H^{\mathrm{PW}}}$ given by following ∂_z on $M_{H^{\mathrm{PW}}} \setminus M^+_{\rho_0}$ and by following R_{ρ_0} on $M^+_{\rho_0}$. We can view $M_{H^{\mathrm{PW}}}$ as a smooth manifold by viewing its nonsmooth part to be the graph of a C^0 function, and observe that M_H and $M_{H^{\mathrm{PW}}}$ coincide along the sets $|x| \geq \rho_0$.

We look at flow trajectories passing over the $q = 0$ slice of our neighborhood, which corresponds to the $y = y_0$ subset of $M_{H^{\mathrm{PW}}}$. Then $\Phi_{H^{\mathrm{PW}}}$ maps the arc

$$A = \left\{ z = -\sqrt{\rho_0^2 - \left(\frac{p}{2\pi}\right)^2}, q = 0 \;\middle|\; p \in [-2\pi\rho_0, 2\pi\rho_0] \right\}$$

to the semicircle

$$\left\{\left(-\sqrt{\rho_0^2-\left(\frac{p}{2\pi}\right)^2},\frac{p}{2\pi},1,0\right)\;\middle|\;p\in[-2\pi\rho_0,2\pi\rho_0]\right\}\subset M_{H^{\mathrm{PW}}}. \tag{59}$$

In the language of Section 10.2.1, this semicircle is the set

$$\left\{(x_\theta,y_0)\mid\theta\in\left[\tfrac{1}{2}\pi,\tfrac{1}{2}3\pi\right]\right\}\subset\mathbb{R}^4$$

equipped with a clockwise parametrization (determined by the variable p). When a trajectory passes through the handle entering at angle $\theta=\theta(p)\in\left[\frac{1}{2}\pi,\frac{3}{2}\pi\right]$ determined by p in the x–plane, it will come out on the top of our neighborhood at angle $2\theta-\pi$, as described in Section 10.2.1. Therefore, when using the piecewise-smooth handle, the lift $\tilde{f}_{H,\rho_0}$ of our continuous flow map $f_{H^{\mathrm{PW}},\rho_0}$ can be written as

$$\tilde{f}_{H^{\mathrm{PW}},\rho_0}(p)=\begin{cases}0 & \text{if } p<-2\pi\rho_0,\\ (1/2\pi)(2\theta(p)-\pi) & \text{if } p\in[-2\pi\rho_0,2\pi\rho_0],\\ -1 & \text{if } p>2\pi\rho_0,\end{cases}$$

where $\theta(p)$ is the angle in the x–plane given by (59).

As the p coordinate wraps around the semicircle of (59) in a clockwise fashion, we conclude that $\tilde{f}_{H^{\mathrm{PW}},\rho_0}$ is a decreasing function. Moreover, $\tilde{f}_{H,\rho_0}(-2\pi\rho)=0$, $\tilde{f}_{H^{\mathrm{PW}},\rho_0}(0)=-\frac{1}{2}$ and $\tilde{f}_{H^{\mathrm{PW}},\rho_0}(-2\pi\rho)=-1$, just like our test function $\tilde{f}$. Therefore, both $\tilde{g}=\tilde{f}$, $\tilde{f}_{H^{\mathrm{PW}},\rho_0}$ must satisfy

$$\tilde{g}([-2\pi\rho_0,0])=\left[-\tfrac{1}{2},0\right],\quad \tilde{g}([0,2\pi\rho_0])=\left[-\tfrac{1}{2},-1\right]. \tag{60}$$

From this we conclude that (58) holds for $\tilde{f}_{H^{\mathrm{PW}},\rho_0}$ given any function $\tilde{f}$ satisfying the required properties.

Now we suppose that B is smooth and $\mathcal{C}^0$–close to B^{PW}. Then the twist map $\tilde{f}_{H,\rho_0}$ will be $\mathcal{C}^0$–close to $\tilde{f}_{H^{\mathrm{PW}},\rho_0}$. This is because the twist map is entirely determined by the flow of the arc A across the surgery handle. All Reeb trajectories starting at points in A which pass through the $|y|<1$ portion of the handle will intercept the region $\rho\in[\rho_1,\rho_0]$ where B^{PW} is smoothed to obtain B. See Figure 22.

Because of item (6) in the observations preceding this proof, $\tilde{f}_{H,\rho_0}(p)<0$ for $p<0$, we can guarantee to that $\tilde{f}_{H,\rho_0}$ satisfies $\tilde{g}([-2\pi\rho_0,-\delta])\subset\left[-\frac{1}{2},0\right]$ for some arbitrarily small $\delta>0$. Therefore, we have $|\tilde{f}_{H,\rho_0}(p)-\tilde{f}(p)|\leq\frac{1}{2}$ for $p\in[-2\pi\rho_0,-\delta]$. By continuity we can also ensure that the desired inequality holds for $p\in[-\delta,0]$ by making B $\mathcal{C}^0$–close enough to B^{PW}. The same arguments with modified notation apply to ensure that (58) holds over $[0,2\pi\rho_0]$ for B close enough to B^{PW}. □

10.6 Perturbing λ_0

Now that we've shown that the flow from the set $\{z = -\rho_0\}$ to the set $\{z = \rho_0\}$ defined by R_H is determined by a Dehn twist by f_{H,ρ_0}, which is supported on $I_{2\pi\rho_0} \times S^1$. Moreover, equation (58) tells us that we can use Proposition 10.5 to correct λ_0 so that the flow over our handle will be an "approximately linear twist" satisfying Assumptions 5.4. We now carry out the details of this correction.

We construct a new coordinate system (z, p, q) on M_H as follows: On the set

$$\{|y| = 1, |x| > \rho_0\},$$

we have coordinates (p, q, z) on M_H coming from the embedding Φ_- as M_1^- and M_H overlap on this region. To get a standard coordinate system on M_H, apply the map

$$(z, p, q) \mapsto \mathrm{Flow}_{R_H}^{z+\rho_0} \circ \Phi_-(-\rho_0, p, q), \quad z \in I_{\rho_0}. \tag{61}$$

With respect to this coordinate system,

$$\lambda_0|_{M_H} = dz + p\, dq.$$

Due to our identification of the flow from the top to bottom of this region — with respect to the (z, p, q) coordinates on M_- — as being determined by a Dehn twist by $f_{H,\delta}$, the change of coordinates on the overlap

$$(M_H \setminus \{\rho < \rho_0\}) \to M_1^-$$

is given exactly as the gluing map of Section 4.7 with the "height perturbation function" — denoted in that section by $H_{f,\epsilon}$ — uniquely determined by f_{H,ρ_0}.

We seek to modify $f_{H,\epsilon}$ using Proposition 10.5 so that the flow over the convex boundary of our handle satisfies the linear dynamics assumptions described in Assumptions 5.4. To this end, let $g\colon I_{2\pi\rho_0} \to \mathbb{R}$ be a function satisfying the following properties:

(1) A Dehn twist by $f_{\rho_0}(0) = f_{H,\rho_0}(p) + g(p)$ satisfies Assumptions 5.4 with $\partial f_{\rho_0}/\partial p(0) = (2\pi\rho_0)^{-1}$.

(2) $|g(p)| \le \frac{1}{2}$.

(3) g and all of its derivatives vanish along $\partial I_{2\pi\rho_0}$.

Such a choice of g is possible by (58). According to Proposition 10.5, using

$$\epsilon_p = \epsilon_z = 2\pi\rho_0, \quad \epsilon_g = \tfrac{1}{2} \tag{62}$$

and ϵ_t arbitrarily large, we can modify the contact form within the coordinate system on M_H by

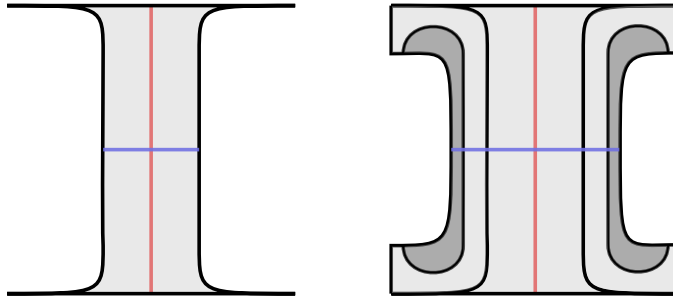


Figure 23: On the left, the rounded handle W_H of Figure 21. On the right, the perturbed handle $W \subset \mathbb{R}^4$. The region along which λ_0 is modified — as in Proposition 10.5 — is shaded in dark gray. The lightly shaded extension of W_H indicates extension by the Liouville flow.

(1) adding a finite symplectization $([0, e^{\epsilon_t}] \times M_H, \lambda_0 = e^t(dz + p\,dq))$ to obtain a handle $W \subset \mathbb{R}^4$ containing W_H, and

(2) perturbing λ_0 within a proper subset of this region to obtain a contact form λ on W,

so that the flow over M_H in the coordinates (z, p, q) is given by a Dehn twist by f_{ρ_0}. A schematic for this extension and perturbation is depicted in Figure 23, right.

Now we rework through (61) and its consequences this time using the new Reeb vector field R. The map

$$(z, p, q) \mapsto \mathrm{Flow}_R^{z+\rho_0} \circ \Phi_-(-\rho_0, p, q) \tag{63}$$

will provide us with a coordinate system (z, p, q) on the convex boundary of W. Now our attaching map is determined by the composition of the Dehn twists

$$(p, q) \mapsto (p, q + g(p)), \quad (p, q) \mapsto (p, q + f_{H,\rho_0}),$$

yielding a Dehn twist by f_{ρ_0}, as desired.

10.7 Attaching the handle to finite symplectizations

To finish our construction, we attach the handle (W, λ) to a finite symplectization of $(\mathbb{R}^3_{\Lambda^\pm}, \alpha_\epsilon)$. In doing so, we will omit the specific choices of ρ_0 required provided that they are determined by ϵ as described in Definition 4.10. Likewise, we assume that Λ^0 consists of a single connected component to simplify notation.

We first consider the case $c = -1$; the map Φ_- provides us with an identification of standard neighborhood N^0_ϵ of Λ^0. The map Φ_- provides us with an identification

of this neighborhood with the convex end of the handle W. By considering $\mathbb{R}^3_{\Lambda^\pm}$ as being contained in the top of a finite symplectization $[-C,0]\times\mathbb{R}^3_{\Lambda^\pm}$, we may attach the handle W via this identification to obtain a 4–manifold along which we set

$$\lambda_{-1}|_W=\lambda,\quad \lambda_{-1}|_{[-C,0]\times\mathbb{R}^3_{\Lambda^\pm}}=e^t\alpha_\epsilon.$$

Outside of a neighborhood of the form $\{\rho(p,z)<\text{const}\}\subsetneq\{0\}\times N^0_\epsilon$, we may extend by some $[0,C]\times\mathbb{R}^3_{\Lambda^\pm}\setminus\{|z|+|p|<\text{const}\}$, over which we take

$$\lambda_{-1}|_{[0,C]\times\mathbb{R}^3_{\Lambda^\pm}\setminus\{\rho(p,z)<\text{const}\}}=e^t\alpha_\epsilon.$$

The constant C may be chosen so that the top of this region coincides with the convex end of the handle W. By the fact that the perturbation of λ described in the previous subsection occurs away from the attaching locus, W_{-1} is smooth with λ_{-1} determining a smooth form, as desired. The disk $\mathbb{D}_{-1}$ is obtained by taking the intersection of the plane $\{|x|=0\}\subset\mathbb{R}^4$ with the handle $W_H\subset W$, depicted as the red line in Figures 21 and 23, and then extending through $[-C,0]\times\mathbb{R}^3_{\Lambda^\pm}$ by a Lagrangian cylinder $[-C,0]\times\Lambda^0$.

Now set $c=+1$. In this case our disk $\mathbb{D}_{+1}$ is taken to be the intersection of the plane $\{|y|=0\}$ with the handle W. According to (47), $\lambda|_{\mathbb{D}_{+1}}=0$. Using the coordinates (p,q,z) on (63), we may identify a neighborhood of the boundary of this disk with a standard neighborhood of Λ^0, which we may consider as being contained in the bottom of a finite symplectization $[0,C]\times\mathbb{R}^3_{\Lambda^\pm}$. We extend the disk by a Lagrangian cylinder over Λ^0 within $[0,C]\times\mathbb{R}^3_{\Lambda^\pm}$ so that its boundary lies in $\{C\}\times\mathbb{R}^3_{\Lambda^\pm}$. To complete the construction of our Liouville cobordism (W_{+1},λ_{+1}), we layer on $[-C,0]\times\mathbb{R}^3_{\Lambda^\pm}\setminus\{|z|+|p|<\text{const}\}$ so that the concave end of the cobordism is smooth and coincides with $(\mathbb{R}^3_\Lambda,\alpha_\epsilon)$.

11 Holomorphic foliations, intersection numbers and the Λ quiver

In this section we describe some tools which allow us to frame geometric questions regarding holomorphic curves in the 4–manifolds relevant to this article — symplectizations and surgery cobordisms — as algebraic problems. We will largely be relying on intersection positivity for holomorphic curves in 4–manifolds [47, Appendix E] and basic algebraic topology.

These tools serve to establish some properties of holomorphic curves in $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$ and surgery cobordisms which we believe to be true intuitively but which are more difficult to articulate precisely: curves with "high energy" should look like Legendrian *RSFT*

disks as they pass through the complement of the surgery locus N_ϵ while "low energy" curves should be trapped inside of the union of N_ϵ with a neighborhood of the chords of Λ and have constrained asymptotics. This will be formalized in Section 11.7 as the *exposed/hidden alternative*.

The first three subsections deal with geometry: In Section 11.1, we describe special almost-complex structures on contactizations and how combinatorial *LRSFT* disks can be "lifted" to holomorphic disks. Section 11.2 described how these complex structures J can be used on large open subsets of symplectizations and surgery cobordisms. Next, in Section 11.3, we show that such J endow open subsets of our 4–manifolds with a foliation by J–holomorphic planes. This is another area of analysis which is considerably simplified by working with $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ rather than $(S^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$.

The remainder of the section is concerned with algebra: Section 11.4 describes some properties of intersections between these planes and finite-energy holomorphic curves asymptotic to chords and orbits of the R_ϵ. These intersection numbers are essentially homological invariants of curves. In the event that the intersection numbers all vanish, an alternative bookkeeping device can be used to keep track of holomorphic curves — an object we call the Λ *quiver*, Q_Λ. This quiver can be used as an algebraic tool to encode $\mathrm{LCH}^{\mathrm{cyc}}$ chain complexes — see Remark 4.1 of [7] — but we will be most interested in the fact that it is a quotient of a space homotopy equivalent to the complement of the $\mathbb{C}$–foliated region of our 4–manifold.

11.1 Model almost-complex structures on symplectizations of contactizations of Stein manifolds

Here we review some generalities regarding holomorphic curves in symplectizations of contactizations of Stein manifolds. For the purposes of this paper, we're really only interested in the cotangent bundles of the real line — for $(\mathbb{R}^3, \xi_{\mathrm{std}})$ is the 1–jet space of $\mathbb{R}$ — though the results are no harder to state or prove in fuller generality. The results here are known; for example, they are implicit in the convexity arguments of [14] and definitions of LCH moduli spaces in [19].

Let W be a manifold of dimension $2n$ with complex structure J and suppose that $F \in C^\infty(\Sigma)$ is such that

$$\beta = -dF \circ J$$

is a Liouville form on W. In other words, (W, J, F) is a *Stein manifold* except that we have omitted any requirements regarding transversality between X_β and ∂W. Define a

contact 1–form $\alpha = dz + \beta$ on $\mathbb{R} \times W$ so that

$$\xi = \{V - \beta(V)\partial_z : V \in TW\}$$

for $V \in TW$. We can define a J' adapted to the symplectization of $(\mathbb{R} \times W, \alpha)$ by

$$J'\partial_t = \partial_z, \quad J'(V - \beta(V)\partial_z) = JV - \beta(JV)\partial_z.$$

As previously mentioned, we're primarily concerned with the cases $W = \mathbb{R} \times I$ for a 1–manifold I with $\beta = p\,dq = -\frac{1}{2}d(p^2) \circ j$. We get $(\mathbb{R}^2, -y\,dx)$ by a change of coordinates.

Lemma 11.1 *If a map $(t, z, u)\colon \Sigma' \to \mathbb{R} \times \mathbb{R} \times W$ is (J', j)–holomorphic, then*

(1) *z is harmonic, and*

(2) *u is (J, j)–holomorphic.*

Moreover, if Σ' is simply connected and we have (z, u) for which z is harmonic and u is (J, j)–holomorphic, then there exists $t\colon \Sigma' \to \mathbb{R}$ for which (t, z, u) is (J', j)–holomorphic. Such t is unique up to addition by a constant.

Proof This is a local calculation: Take coordinates x and y on $\mathbb{D}$, which we may consider being contained in Σ' with j denoting the standard complex structure on $T\mathbb{D}$. We will be studying (7).

Write $\overline{\partial}_{J',j}(t, z, u) = \frac{1}{2}(T(t, z, u) + J'T(t, z, u) \circ j)$ for the usual Cauchy–Riemann operator. For $V \in TW$, we calculate

$$\pi_\alpha(a\partial_z + V) = V - \beta(V)\partial_z,$$

so that the ξ–valued part $\frac{1}{2}(\pi_\alpha + J' \circ \pi_\alpha \circ j)$ of $\overline{\partial}_{J',j}(s, t, u)$ depends only on u. Then

$$(\pi_\alpha + J' \circ \pi_\alpha \circ j)(s, t, u) = \overline{\partial}_{J,j}u - \beta \circ (\overline{\partial}_{J,j}u)\partial_z,$$

where $\overline{\partial}_{J,j}$ is the Cauchy–Riemann operator for u. The TW part of this expression vanishes if and only if u is (J, j)–holomorphic, which would imply that the ∂_z part of the expression vanishes as well.

Assuming that (t, z, u) is (J', j)–holomorphic, then u is (J, j)–holomorphic and $dt = ((z, u)^*\alpha) \circ j$, implying

$$\begin{aligned} u^*\beta = -u^*(dF \circ J) &= -d(F \circ u) \circ j, \\ (z, u)^*\alpha \circ j &= dz \circ j + d(F \circ u), \\ d^2t = d((z, u)^*\alpha \circ j) &= d(dz \circ j) = -\Delta(z) = 0, \end{aligned}$$

where Δ is the Laplacian. Therefore, z is harmonic.

Now, provided harmonic z and (J, j)–holomorphic u for simply connected Σ', the above expression tells us that $(z,u)^*\alpha \circ j$ is closed, and so is exact. Therefore, we have a function t — determined uniquely up to addition by scalars — satisfying $dt = (z,u)^*\alpha \circ j$. Then, by the above formula and (7), (t,z,u) is (J', j)–holomorphic. □

Corollary 11.2 (drawing-to-disk correspondence) *Suppose that (W, J, F) is a Stein manifold of complex dimension 1 and that Λ is a chord-generic Legendrian link in $(I \times W, dz - dF \circ J)$. Suppose that*

$$u : \mathbb{D} \setminus \{p_j\} \to W$$

is an orientation-preserving immersion of the disk with a finite set of boundary punctures $\{p_k\}$ removed such that $u(\partial\mathbb{D} \setminus \{p_k\}) \subset \pi_W(\Lambda)$. Then there exists a set $\{p'_k\}$ of boundary punctures on the disk, a diffeomorphism $\phi : \mathbb{D} \setminus \{p'_k\} \to \mathbb{D} \setminus \{p_k\}$ and functions $t, z : \to \mathbb{R}$ such that

$$(t, z, u \circ \phi) : \mathbb{D} \setminus \{p'_k\} \to \mathbb{R} \times \mathbb{R} \times W$$

is (J', j)–holomorphic with $(z, u \circ \phi)(\partial\mathbb{D} \setminus \{p_k\}) \subset \Lambda$. Provided ϕ, z is uniquely determined and t is uniquely determined up to addition by a positive constant.

Proof Because u is an immersion, we can force it to be (J, j')–holomorphic for some almost-complex structure j' on $\mathbb{D}$ by defining $j'\partial_x = (Tu)^{-1} J(Tu)\partial_x$. We can then find a diffeomorphism ϕ which is (j', j)–holomorphic by the uniformization theorem.

By the chord-genericity and smoothness of Λ, there exists a unique, bounded, smooth function z_∂ on $\partial\mathbb{D} \setminus \{p'_k\}$ for which

$$(z_\partial, u \circ \phi) \in \Lambda.$$

Applying [1, Chapter 6, Section 4.2], there is a function $z : \mathbb{D} \setminus \{p'_k\} \to \mathbb{R}$ solving the Dirichlet problem

$$\Delta(f) = 0, \quad z|_{\partial\mathbb{D}\setminus\{p'_k\}} = z_\partial,$$

which is unique by the maximum principle. By Lemma 11.1, we can find t for which $(t, z, u \circ \phi)$ is (J', j)–holomorphic, as desired. □

11.2 N–standard almost-complex structures

As always, let N_ϵ be a tubular neighborhood of Λ, whose complement we may consider to be a codimension-0 submanifold of either $\mathbb{R}^3$ or $\mathbb{R}^3_{\Lambda^\pm}$. Define

$$\widetilde{N}_\epsilon = \pi_{xy}^{-1}(\pi_{xy}(N_\epsilon)),$$

which we may view as an open set in either $\mathbb{R}^3$ or $\mathbb{R}^3_{\Lambda^\pm}$. Denote its complement by $\widetilde{N}_\epsilon^{\complement}$.

Definition 11.3 We say that an almost-complex structure J on $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$ is *N–standard* if its restriction to $\xi_{\Lambda^\pm}$ agrees with the standard almost-complex structure J_0 on $\xi_{\Lambda^\pm}$ described by (14) on $\widetilde{N}^{\complement}_{\epsilon}$ as well as on a neighborhood

$$N_{C,\infty} = \{x^2 + y^2 + z^2 > C\}$$

of the puncture of our 3–manifold for some $\epsilon, C > 0$. In order that J be adapted to the symplectization, we require $J\partial_t = \partial_z$ on $\mathbb{R}\times\widetilde{N}^{\complement}_{\epsilon}$.

We may define N–standard for almost-complex structures on completions of surgery cobordisms (W_c, λ_c) of Section 10 analogously as the cobordisms contain the symplectizations of $(\mathbb{R}^3_{\Lambda^\pm}\setminus N_\epsilon, \alpha_{\text{std}})$.

For an N–standard almost-complex structure J and a (J, j)–holomorphic curve

$$U : \Sigma' \to \mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm},$$

along $U^{-1}(\mathbb{R}\times N_{C,\infty})$ we can write $U = (t, z, u)$. By Lemma 11.1, z is harmonic and u is holomorphic, so that $x\circ u$ and $y\circ u$ are harmonic as well. It follows that $-d(d(z^2+|u|^2)\circ j)$ is nonnegative as an area form on Σ'. Hence, for $C' > C$, finite-energy curves with punctures asymptotic to chords and orbits of R_ϵ cannot touch spheres of radius C' by the maximum principle.

11.2.1 Compatibility with perturbation schemes and adaption to symplectizations

Note that perturbations of almost-complex structures required to achieve the transversality required to define SFT curve counts in $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$ or $(\overline{W}_c, \overline{\lambda}_c)$ may be defined in arbitrarily small neighborhoods of the orbits of R_ϵ [4, Section 5] and these orbits are properly contained in open sets unconstrained by the N–standard condition. Hence, these perturbations may be carried out for N–standard almost-complex structures while maintaining their defining properties.

Similarly, the cobordisms (W_c, λ_c) of Section 10 are designed to support N–standard almost-complex structures which are adapted to their cylindrical ends. For such cobordisms, we'll be additionally interested in studying somewhere-injective curves positively asymptotic to chords of the Legendrian boundaries of the disks $\mathbb{D}_{c,i}\subset W_c$ with Lagrangian boundary. See Section 12.2. In this context, the perturbation scheme of [18, Section 2] may be applied, which likewise only deforms Cauchy–Riemann equations in arbitrarily small neighborhoods of chords and orbits. Again, there is no lack of compatibility with the N–standard condition.

Assumptions 11.4 Throughout the remainder of this section, we assume that any almost-complex structure J on a symplectization or surgery cobordisms is N–standard and that all somewhere-injective curves under consideration are regular. When discussing surgery cobordisms, we assume that J is adapted to the cylindrical ends of its completion and that almost-complex structures on symplectizations are adapted.

11.3 Semiglobal foliation by holomorphic planes

Here we describe holomorphic foliations by infinite-energy planes in symplectizations and surgery cobordisms.

11.3.1 $\mathbb{C}$ foliations in symplectizations Observe that $\widetilde{N}_\epsilon^{\mathbb{C}}$ is foliated by embedded, $\mathbb{R}$–parametrized Reeb orbits of the form $t \to (t, x_0, y_0)$. Then $\mathbb{R} \times \widetilde{N}_\epsilon^{\mathbb{C}}$ is foliated by holomorphic planes parametrized

$$(s,t) \mapsto (s,t,x,y)$$

for $(x,y) \in \mathbb{R}^2 \setminus \pi_{x,y}(N_\epsilon)$. We denote each such unparametrized plane by $\mathbb{C}_{x,y}$.

11.3.2 $\mathbb{C}$ foliations in surgery cobordisms For the following, we require that Λ^0 be nonempty. The link $\Lambda^\pm$ is allowed to be empty, in which case we would have $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm}) = (\mathbb{R}^3, \xi_{\text{std}})$ and set $\alpha_\epsilon = dz - y\,dx$. Let (W_c, λ_c) be a surgery cobordism associated to the pair

$$\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm}), \quad c \in \{\pm 1\}$$

as described in the introduction of Section 10, with completion $(\overline{W}_c, \overline{\lambda}_c)$. Because the handles are attached along a neighborhood of Λ^0, we can view $\mathbb{R} \times \widetilde{N}^{\mathbb{C}}$ as a subset of $\overline{W}_c$ which is also foliated by infinite-energy planes $\mathbb{C}_{x,y}$.

11.4 Intersection numbers

For the following, let (Σ, j) be a compact Riemann surface, possibly with boundary, with fixed collections of interior points p_k^{int} and boundary points p_k^{∂}. As usual, we write Σ' for Σ with all of its marked points removed. When discussing completions $(\overline{W}_c, \overline{\lambda}_c)$, we write

$$\overline{\mathbb{D}}_{c,i} \subset \overline{W}_c$$

for the Lagrangian planes obtained by extending the disks $\mathbb{D}_{c,i}$ of Theorem 10.1 by the positive (resp. negative) half-infinite Lagrangian cylinders over their Legendrian boundaries when $c = 1$ (resp. $c = -1$).

Definition 11.5 We say that a holomorphic map $U\colon \Sigma' \to \overline{W}_c$ is a $\overline{W}_c$ *curve* if it its boundary is mapped to the $\overline{\mathbb{D}}_{c,i}$, its boundary punctures are asymptotic to chords of their Legendrian boundaries, and all interior punctures are asymptotic to closed Reeb orbits at the convex and concave ends of $\overline{W}_c$.

We say that a holomorphic map $U\colon \Sigma' \to \mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$ is an $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$ *curve* if the boundary of Σ' is mapped to the Lagrangian cylinder over Λ^0, its boundary punctures are asymptotic to chords of Λ^0, and its interior punctures are asymptotic to closed orbits of R_ϵ.

We recall — see [47, Definition E.2.1] — that, provided a pair of maps $u_i\colon \Sigma_i' \to W$ from surfaces Σ_i' for $i=1,2$ into a 4–manifold W whose images are disjoint outside of some open sets $S_i \subset \Sigma_i$ with compact closures outside of which the Σ_i are disjoint, then we can define a *intersection number* $u_1\cdot u_2 \in \mathbb{Z}$ by perturbing the u_i along the S_i so that the maps are transverse and counting their intersections with signs.[15]

Theorem 11.6 *Suppose that U is a $\overline{W}_c$ curve or an $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$ curve. Then, for $(x,y)\in\mathbb{R}^2\setminus\pi_{x,y}(N_\epsilon)$, the intersection number $\mathbb{C}_{x,y}\cdot U_s \in \mathbb{Z}$ is well defined and nonnegative. Furthermore, they are homological invariants in the following sense*:

(1) ***Boundaryless curves in symplectizations*** *Suppose that U is an $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$ curve positively asymptotic to a collection γ^+ of Reeb orbits and negatively asymptotic to some γ^-. Then the intersection number $\mathbb{C}_{x,y}\cdot U$ depends only on the relative homology class*

$$[\pi_{\mathbb{R}^3_{\Lambda^\pm}}(U)] \in H_2(\mathbb{R}^3_{\Lambda^\pm}, \gamma^+\cup\gamma^-).$$

(2) ***Curves in symplectizations with Lagrangian boundary*** *Suppose that U is an $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$ curve asymptotic to collections $\gamma^\pm$ of Reeb orbits and collections of chords $\kappa^\pm$ of Λ^0. Then the intersection number $\mathbb{C}_{x,y}\cdot U$ depends only on the relative homology class*

$$[\pi_{\mathbb{R}^3_{\Lambda^\pm}}(U)] \in H_2(\mathbb{R}^3_{\Lambda^\pm}, \gamma^+\cup\gamma^-\cup\kappa^+\cup\kappa^-\cup\Lambda^0).$$

(3) ***Boundaryless curves in surgery cobordisms*** *Suppose that U is a $\overline{W}_c$ curve positively asymptotic to a collection of closed Reeb orbits γ^+ in ∂^+W and*

[15] We're taking a slight modification of [47, Definition E.2.1] by defining the intersection number to be the sum of the *local intersection numbers* over all points of intersection. This is feasible for holomorphic curves in 4–manifolds with our hypotheses as distinct curves have isolated intersections [47, Proposition E.2.2].

negatively asymptotic to some collection of closed Reeb orbits γ^- in $\partial^- W$. Then the intersection number $\mathbb{C}_{x,y} \cdot U$ depends only on the relative homology class

$$[U] \in H_2(W_c, \gamma^+ \cup \gamma^-).$$

Here we view U as a cobordism in the compact manifold W_c bounding the orbit collections $\gamma^\pm$ in its boundary.

(4) ***Curves in surgery cobordisms with Lagrangian boundary*** *Suppose that U is a $\overline{W}_c$ curve positively asymptotic collection of closed Reeb orbits $\gamma^\pm$ in $\partial^+ W$ with boundary punctures asymptotic to some collection $\kappa^\pm$ of chords of the Legendrian boundaries of disks $\mathbb{D}_k$. Then the intersection number $\mathbb{C}_{x,y} \cdot U$ depends only on the relative homology class*

$$[U] \in H_2\Big(W_c, \gamma^+ \cup \gamma^- \cup \kappa^+ \cup \kappa^- \cup \bigcup \overline{\mathbb{D}}_{c,i}\Big).$$

Proof To check well-definedness, we need to ensure that any intersections between $\mathbb{C}_{x,y}$ and $U(\Sigma')$ occur away from the boundary and punctures of Σ', so that intersection numbers are independent of the perturbation required in their definition. By our boundary conditions, U must be such that there exists some open neighborhood $S \subset \Sigma'$ of the punctures and boundary of Σ which maps into the complement of $\mathbb{R} \times (\mathbb{R}^3_{\Lambda^\pm} \setminus N_\epsilon)$. The images of the complements of the S_s must be contained in some compact set of the form $[-C_1, C_1] \times (\mathbb{R}^3_{\Lambda^\pm} \setminus N_\epsilon)$. Likewise, the images of the complements of the S_s must be bounded in the z coordinate on $\mathbb{R}^3$. Hence, all intersections occur within a subset of the form $[-C_1, C_1] \times [-C_2, C_2] \times \{(x, y)\} \subset \mathbb{C}_{x,y}$, implying that the $\mathbb{C}_{x,y} \cdot U_s \in \mathbb{Z}$ are well defined.

Intersection nonnegativity follows from positivity of intersections of holomorphic curves in 4–manifolds. See for example [47, Section E.2]. For homological invariance, we will work out the details in the case of boundaryless curves in symplectizations. The other cases follow similar reasoning.

As in the statement of the theorem, we can slightly perturb U near its asymptotic ends to obtain a 2–cycle in $[-C, C] \times \mathbb{R}^3_{\Lambda^\pm}$ bounding $\{C\} \times \gamma^+ - \{-C\} \times \gamma^-$ for some large $C > 0$. Using the coordinates on $\mathbb{R} \times \mathbb{R}^3$ in which we may consider $\mathbb{C}_{x,y}$ to be contained, each intersection between U and $\mathbb{C}_{x,y}$ occurs at some (t, z, x, y). Possibly perturbing U near each such intersection to achieve transversality and isolation of intersections, the sign of each intersection is given by the sign of $T\mathbb{C}_{x,y} \wedge TU$ considered as an oriented ray in the orientation line bundle $\mathbb{R}\partial_t \wedge \partial_z \wedge \partial_x \wedge \partial_y$ for

$T_{(t,z,x,y)}W$. As $T\mathbb{C}_{x,y} = \mathrm{span}_{\mathbb{R}}(\partial_t, \partial_z)$, this sign only depends on the ∂_x and ∂_y parts of the tangent map TU of U. Hence, the intersection number $\mathbb{C}_{x,y} \cdot U$ only depends on (x, y) and $\pi_{\mathbb{R}^3_{\Lambda\pm}} \circ U$. □

The ending of the above proof also immediately implies the following:

Lemma 11.7 *Suppose that $U : \mathbb{D} \setminus \{p_k\} \to \mathbb{R} \times \mathbb{R}^3$ is a holomorphic disk determined by an immersion $u : \mathbb{D} \setminus \{p_k\} \to \mathbb{R}^2$ as in Corollary 11.2. Given a point $(x, y) \in \mathbb{R}^2 \setminus \pi_{x,y}(N_\epsilon)$, the intersection number is*

$$\mathbb{C}_{x,y} \cdot U = \#u^{-1}((x, y)).$$

11.5 Bases and energy bounds

Here we'll reduce the information of the $\mathbb{C}_{x,y}$ down to that of a finite collection of planes. Write $\mathcal{R}_k$ for the connected components of $\mathbb{R}^2 \setminus \pi_{x,y}(N_\epsilon)$ of finite area and write

$$\mathcal{E}_k = \int_{\mathcal{R}_k} dx \wedge dy$$

for their areas. There is also a single connected component $\mathbb{R}^2 \setminus \pi_{x,y}(N_\epsilon)$ of infinite area, which we will denote by $\mathcal{R}_\infty$.

Pick a point (x_k, y_k) within the interior of each $\mathcal{R}_k$ as well as a point $(x_\infty, y_\infty) \in \mathcal{R}_\infty$. We'll call such a choice of indices and points a *point basis* for Λ. Provided a point basis, we may abbreviate

$$\mathbb{C}_k = \mathbb{C}_{(x_k,y_k)}.$$

Such a choice allows us to package a simple-to-state energy estimate:

Proposition 11.8 *Let U be a finite-energy $\mathbb{R} \times \mathbb{R}^3_{\Lambda\pm}$ curve with interior punctures asymptotic to some collections of orbits of R_ϵ and boundary punctures asymptotic to chords of $\Lambda^0 \subset \mathbb{R}^3_{\Lambda\pm}$. Then*

$$\mathcal{E}(U) > \sum_k \mathcal{E}_k \mathbb{C}_k \cdot U.$$

Proof For each $k \neq \infty$ for which

$$\Sigma'_k = U^{-1}(\mathbb{R} \times \mathbb{R} \times \mathcal{R}_k)$$

is not empty,

$$\pi_{x,y} \circ \pi_{\mathbb{R}^3_{\Lambda\pm}} \circ U : \Sigma'_k \to \mathcal{R}_k$$

is a nonconstant holomorphic map. By our boundary conditions, each Σ'_k is disjoint from some neighborhood of the boundary and punctures of Σ' and so must be a branched

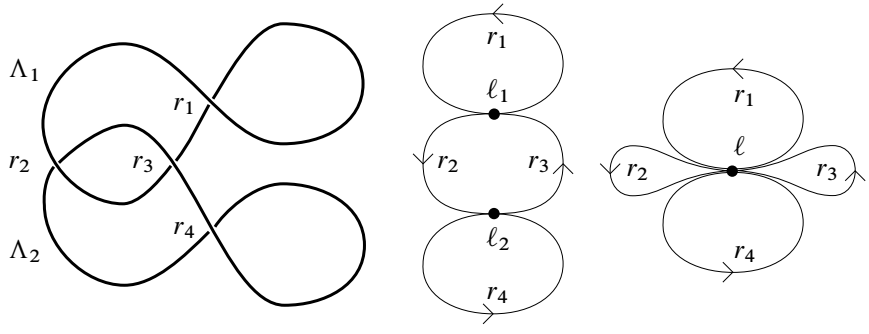


Figure 24: From left to right: a Legendrian Hopf link Λ in the Lagrangian projection, the associated quiver Q_Λ and the quiver Q_Λ/ℓ.

covering. The degree of the associated map

$$(\overline{\Sigma}'_k, \partial\overline{\Sigma}'_k) \to (\overline{\mathcal{R}}_k, \partial\overline{\mathcal{R}}_k)$$

is equal to $\mathbb{C}_k \cdot U$, so that our requirement that α_ϵ coincides with $\alpha_{\text{std}} = dz - y\,dx$ on the compliment of $N^\pm$ implies

$$\mathcal{E}(U) > \sum_k \int_{\Sigma'_k} d\alpha_\epsilon = \sum_k \int_{\Sigma'_k} dx \wedge dy = \sum_k \mathcal{E}_k \mathbb{C}_k \cdot U. \qquad \square$$

11.6 The Λ quiver

In the event that all intersection numbers $\mathbb{C}_k \cdot U$ are zero for a given curve U, we can employ another device to keep track of holomorphic curves and their boundary conditions.

Definition 11.9 The Λ *quiver*, denoted by Q_Λ, is the directed graph with

(1) one vertex ℓ_i for each connected component Λ_i of Λ, and

(2) one directed edge for each chord r_j of $\Lambda \subset \mathbb{R}^3$ starting at the vertex $\ell_{l_j^-}$ and ending at $\ell_{l_j^+}$.[16]

Also define a graph Q_Λ/ℓ which is the quotient of Q_Λ obtained by identifying all of its vertices. We write

$$\pi_\ell \colon Q_\Lambda \to Q_\Lambda/\ell$$

for the quotient map.

An example is provided in Figure 24.

[16] We recall the $l_j^\pm$ are defined in Section 3.

11.6.1 Algebraic aspects of Q_Λ and Q_Λ/ℓ The primary utility of the space Q_Λ/ℓ is that its homology has a particularly nice presentation, with H_1 freely generated by the chords of $\Lambda \subset (\mathbb{R}^3, \xi_{\text{std}})$,

$$H_1(Q_\Lambda/\ell) = \oplus \mathbb{Z} r_j,$$

while its fundamental group — based at its unique vertex, ℓ — is a free group on the chords of Λ,

$$\pi_1(Q_\Lambda/\ell, \ell) = \langle r_j \rangle.$$

In applications, we'll make use of the following definitions and lemma:

Definition 11.10 For an edge e of a directed graph G we define the *collapse map at* e, denoted by $\pi_e \colon G \to S^1$, as the map which takes the quotient by $G \setminus \text{int}(e)$. The target is naturally pointed and oriented by the direction of e. A continuous map $\Phi \colon S^1 \to G$ from an oriented circle is *nonnegative* if, for every edge e of g, the composition

$$S^1 \xrightarrow{\Phi} G \xrightarrow{\pi_e} S^1$$

with the collapse map has nonnegative degree. We say that the map is *positive* if it is nonnegative and there exists at least one $e \subset G$ for which $\pi_e \circ \Phi$ has positive degree.

Definition 11.11 Let $\mathcal{S}$ be a set with associated free group $\langle \mathcal{S} \rangle$. We say that an element $x \in \langle \mathcal{S} \rangle$ is *positive* if it can be described as a word

$$x = x_1 \cdots x_n, \quad x_k \in \mathcal{S}.$$

Alternatively, the set of positive elements in $\langle \mathcal{S} \rangle$ is equivalent to the image of the natural monoid homomorphism from the free monoid on $\mathcal{S}$ into $\langle \mathcal{S} \rangle$.

If x is positive then the above factorization is necessarily unique. We say that two positive elements x and y of $\langle \mathcal{S} \rangle$ are *cyclically equivalent* if their positive factorizations differ by a cyclic rotation. That is, provided a factorization of x as above, there exists k for which

$$y = x_k \cdots x_n x_1 \cdots x_{k-1}.$$

We say that $x \in \langle \mathcal{S} \rangle$ is *negative* if x^{-1} is positive. Two negative elements x and y are *cyclically equivalent* if x^{-1} and y^{-1} are cyclically equivalent.

Cyclic equivalence is no stronger than conjugacy equivalence.

Lemma 11.12 *Suppose that $x, y \in \langle \mathcal{S} \rangle$ are positive and conjugate in $\langle \mathcal{S} \rangle$. Then they are cyclically equivalent.*

Proof Suppose there exists some z for which $zx = yz$ and write $z = z_1 \cdots z_n$ with the z_k being elements of $\mathcal{S}$ or inverses of such letters. We can assume that at least one of z_1 or z_n is positive. Otherwise we can write $z^{-1}y = xz^{-1}$ to obtain the desired hypothesis by a change of notation.

Suppose that z_1 is positive. Then the positive factorization of y must start with z_1. Then $y' = z_1^{-1} y z_1$ is positive, so we can write $z'x = y'z'$ with $z' = z_2 \cdots z_n$. We have reduced the problem to finding a cyclic equivalence between two positive elements x and y' which are conjugate by a word z' of length $n-1$. A similar argument may be applying in the case that z_n is positive.

To complete the proof, loop through this argument n times. □

11.6.2 Geometric aspects of Q_Λ and Q_Λ/ℓ The primary utility of the space Q_Λ in relation to the present discussion is given by the following result:

Proposition 11.13 *There exist surjective maps*

$$\mathbb{R}^3_{\Lambda^\pm} \setminus \widetilde{N}^{\complement}_\epsilon \to Q_\Lambda, \quad \overline{W}_c \setminus \mathbb{R} \times \widetilde{N}^{\complement}_\epsilon \to Q_\Lambda,$$

both of which we will denote by π_Q, such that, for each chord r_j of Λ and each line segment I directed by ∂_z connecting $\mathcal{D}^{\mathrm{ex}}_j$ to $\mathcal{D}^{\mathrm{en}}_j$, the submanifold $\mathbb{R} \times I$ is mapped onto the edge r_j of Q_Λ in a way such that, for each t in $\mathbb{R}$, $\{t\} \times I \to e_j$ is a homeomorphism.[17]

Proof We start with the case in which the domain of π_Q is $\mathbb{R}^3_{\Lambda^\pm} \setminus \widetilde{N}^{\complement}_\epsilon$. We have that $\mathbb{R}^3_{\Lambda^\pm} \setminus \widetilde{N}^{\complement}_\epsilon$ is homotopy equivalent the union of N_ϵ with all of the chords r_j of Λ. We can perform this homotopy so that the intervals connecting the $\mathcal{D}^{\mathrm{en}}_j$ and $\mathcal{D}^{\mathrm{ex}}_j$ (forming a neighborhood of r_j) collapse onto r_j as a fibration. Note that N_ϵ is a collection of solid tori, so that $N_\epsilon \cup \{r_j\}$ is homotopy equivalent to a 1–dimensional CW complex. If we collapse each connected component $N_{\epsilon,i}$ of N_ϵ to a point ℓ_i, the graph Q_Λ is obtained.

The proof for $\overline{W}_c \setminus \mathbb{R} \times \widetilde{N}^{\complement}_\epsilon$ is nearly identical except at the last step; the addition of the surgery handles already provides the effect of attaching 2–cells along the circles in our 1–complex corresponding to components of $\Lambda^\pm$. We then collapse these 2 cells to points, which has the same effect — in the homotopy category — as collapsing the circles corresponding to the components of Λ to points. □

[17] We recall that the $\mathcal{D}^*_j$ are defined in Section 5.1.

Proposition 11.14 *Suppose that $\gamma(t)$ parametrizes a Reeb orbit in $\mathbb{R}^3_{\Lambda\pm}$ or ∂W_c. Then $\pi_Q \circ \gamma$ is positive in the sense of Definition 11.10.*

The open string version of this assertion is as follows: Let U be a $\overline{W}_c$ or $\mathbb{R}\times\mathbb{R}^3_{\Lambda\pm}$ curve with domain Σ' having a boundary component $\partial_i\Sigma \subset \partial\Sigma$ for which all punctures along $\partial_i\Sigma$ have positive asymptotics. Then $\pi_Q \circ U|_{\partial_i\Sigma}$ is a positive loop. If all punctures along $\partial_i\Sigma$ have negative asymptotics, then this loop is negative.

This is clear from the construction of the map π_Q. For a parametrization γ of a Reeb orbit with cyclic word $r_{j_1}\cdots r_{j_n}$, we have

$$[\pi_\ell \circ \pi_Q \circ \gamma] = \sum_1^n [r_{j_k}] \in H_1(Q_\Lambda/\ell).$$

Intuitively, the map $\pi_\ell \circ \pi_Q$ induces a map on homology which abelianizes boundary conditions for holomorphic curves. We can also view $[\pi_\ell \circ \pi_Q \circ \gamma]$ as an element of the H_0 of the free loop space of Q_Λ/ℓ which records the word map of γ.

For a single chord κ with boundary on some $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda\pm}, \xi_{\Lambda\pm})$, we can view $\pi_\ell\circ\pi_Q\circ\kappa$ as a pointed map

$$(\kappa, \partial\kappa) \to (Q_\Lambda/\ell, \ell)$$

as Λ is mapped to ℓ by $\pi_\ell \circ \pi_Q$. In this way, κ determines a positive element of $\pi_1(Q_\Lambda/\ell)$ as well as a relative homology class

$$[\pi_\ell \circ \pi_Q \circ \kappa] \in H_1(Q_\Lambda/\ell, \ell).$$

Both the π_1 and H_1 classes record the word map of κ.

11.7 The exposed/hidden alternative

Assume that Λ is equipped with a basis of points $(x_k, y_k) \in \mathbb{R}^2 \setminus \pi_{x,y}(N_\epsilon)$ as described in Section 11.5.

Definition 11.15 (exposed/hidden alternative) We say that an $\mathbb{R}\times\mathbb{R}^3_{\Lambda\pm}$ or $\overline{W}_c$ curve U is *exposed* if there exists at least one k for which $\mathbb{C}_k \cdot U > 0$. Otherwise we say that U is *hidden*.

If a curve U is exposed, then we can use the intersection numbers to keep track of the location of its image within the target manifold. If the curve is hidden, then, by intersection positivity, its image must be entirely contained in the complement of $\mathbb{R}\times\widetilde{N}^{\mathbb{C}}_\epsilon$, whence we can apply the map $\pi_\ell \circ \pi_Q$. We state some simple applications, the first few of which tell us that the homology of Q_Λ/ℓ dictates whether a curve is exposed or hidden.

Proposition 11.16 (homological mismatches are exposed) *Suppose that U is an $\mathbb{R}\times\mathbb{R}^3_{\Lambda\pm}$ or $\overline{W}_c$ curve without boundary components positively asymptotic to some collection $\gamma^+=\{\gamma_k^+\}$ of closed orbits and negatively asymptotic to some collection $\gamma^-=\{\gamma_k^-\}$ of Reeb orbits. If the 1–cycle*

$$\sum[\pi_\ell\circ\pi_Q\circ\gamma_k^+]-\sum[\pi_\ell\circ\pi_Q\circ\gamma_k^-]\neq 0\in H_1(Q_\Lambda/\ell),$$

then U is exposed.

Proof If the curve was hidden, then we could apply the map $\pi_\ell\circ\pi_Q$ to the image of U. Our hypotheses on asymptotics imply that we would get a 2–cycle in $\mathbb{R}^3_{\Lambda\pm}\setminus\widetilde{N}_\epsilon^{\complement}$ or $W_c\setminus\mathbb{R}\times\widetilde{N}_\epsilon^{\complement}$ bounding a homologically nontrivial 1–cycle, providing a contradiction. □

A slight modification applies to chords as well.

Proposition 11.17 (exposure of filling curves) *Suppose that U is an $\mathbb{R}\times\mathbb{R}^3_{\Lambda\pm}$ or $\overline{W}_c$ curve for which all asymptotic chords and orbits are positive. Then U must be exposed.*

Proposition 11.18 (homological matches are hidden) *Let $h\in H_1(Q_\Lambda/\ell)$ be a positive homology class.*[18] *Then there exists ϵ_h such that, for each $\epsilon<\epsilon_h$, given a holomorphic curve in $\mathbb{R}\times\mathbb{R}^3_{\Lambda\pm}$ positively asymptotic to a collection of orbits γ^+ and negatively asymptotic to a collection γ^- of R_ϵ orbits with*

$$[\pi_\ell\circ\pi_Q\circ\gamma^+]=[\pi_\ell\circ\pi_Q\circ\gamma^-]=h\in H_1(Q_\Lambda/\ell),$$

then U is hidden.

Proof By the action estimates of Section 5.5, we have

$$\mathcal{E}(U)=\mathcal{O}\Big(3\epsilon\sum\mathrm{wl}(\gamma_k^+)\Big).$$

For ϵ sufficiently small, we could guarantee that this quantity is less than the energies $\mathcal{E}_k$ of the regions $\mathcal{R}_k$ (which grow slightly as ϵ tends to 0 with N_ϵ shrinking). Therefore, the energy bound of Proposition 11.8 would imply that U must be hidden. □

Proposition 11.19 (cyclic order preservation of open–closed interpolations) *Suppose that U is a hidden $\overline{W}_c$ curve whose domain is a disk with a single interior puncture and any number of boundary punctures. We require that*:

(1) *If $c=+1$, the boundary punctures are positively asymptotic to chords of Λ^0 with words $w_1,\dots,w_n$.*

[18] That is, h may be represented as a sum of positive cycles.

(2) *If $c = -1$, the boundary punctures are negatively asymptotic to chords of Λ^0 with words $w_1, \dots, w_n$.*

Here indices follow the counterclockwise cyclic ordering of the punctures around $\partial\mathbb{D}$. Then interior puncture of U asymptotic to the orbit $(w_1 \cdots w_n)$.

The $c = -1$ curves described are those used to determine homomorphisms from linearized contact homology to a cyclic version of Legendrian contact homology when performing a contact -1 surgery in [7; 18; 20].[19] We'll see some of the $c = 1$ curves shortly in Theorem 12.2.

Proof Consider the map $\pi_\ell \circ \pi_Q \circ U$ from the punctured disk to the graph Q_Λ/ℓ. Then $\partial\mathbb{D}$ — compactified appropriately — will give us an element of the free loop space of Q_Λ/ℓ. It is clear from the construction of the map π_Q that the connected component of the free loop space of Q_Λ/ℓ containing this loop is indexed by $w_1 \cdots w_n$. Looking at circles of varying radii in $\mathbb{D}$ provides a homotopy between this loop and the one provided by the interior puncture. Again by the construction of π_Q, observe that, if the orbit to which the puncture is asymptotic has cyclic word $r_{j_1} \cdots r_{j_n}$, then this word must also index the component of the free loop space of Q_Λ/ℓ to which the puncture is associated. The connected components of the free loop space of Q_Λ/ℓ are in bijective correspondence with conjugacy classes on $\langle r_j \rangle$, so that the expressions $r_{j_1} \cdots r_{j_n}$ and $w_1 \cdots w_n$ are conjugate by the existence of the aforementioned homotopy. They are also both positive in the sense of Definition 11.11 and so differ by a cyclic permutation of their letters by Lemma 11.12. □

Proposition 11.20 (triviality of hidden cylinders and strips) *Suppose that U is a hidden holomorphic cylinder in $\mathbb{R} \times \mathbb{R}^3_{\Lambda^\pm}$. Then U is a trivial cylinder.*

If U has domain $\mathbb{R} \times I_C$ for some $C > 0$, is hidden, with boundary on the Lagrangian cylinder over $\Lambda^0 \subset (\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$, and with punctures asymptotic to chords of Λ^0, then U is a trivial strip.

Proof If U is positively asymptotic to some orbit $(r_{j_1} \cdots r_{j_n})$, then we can follow the proof of Proposition 11.19 verbatim to conclude that U is negatively asymptotic to $(r_{j_1} \cdots r_{j_n})$. Hence, the energy of U is zero and it must be a trivial cylinder.

The case of a holomorphic strip is even easier. Suppose the strip is parametrized by $s \in \mathbb{R}$ and $t \in I_C$, and consider the family of paths $\gamma_s(t) = \pi_{\mathbb{R}^3_{\Lambda^\pm}} \circ U(s,t)$ with boundary on $\Lambda^0 \subset \mathbb{R}^3_{\Lambda^\pm}$. Then we may consider the $\pi_\ell \circ \pi_Q \circ \gamma_s$ as an $\mathbb{R}$ family of based

[19] We're ignoring anchors, which can be avoided in some settings, such as [20].

loops in Q_Λ/ℓ. As $s \to \infty$, the $\pi_1(Q_\Lambda/\ell)$ element recorded by this based loop is the word map of the chord to which U is positively asymptotic. As $s \to -\infty$, the element recorded is the word map of the chord to which U is negatively asymptotic. Hence, the asymptotics are equivalent by our chord-to-chord correspondence (Theorem 5.10), the energy of U is zero, and U is a trivial strip. □

12 Applications

In this section we apply our computational tools to study the contact homology of various contact manifolds. A summary of the results are as follows:

(1) In Section 12.1, we compute the contact homology of contact ± 1 surgeries on the $\mathrm{tb} = -1$, $\mathrm{rot} = 0$ unknot in $\mathbb{R}^3$.

(2) In Section 12.2, we use the results of Section 11 to prove a general existence result for holomorphic planes in $\mathbb{R} \times \mathbb{R}^3_{\Lambda^\pm}$ when $\Lambda^+ \neq \varnothing$.

(3) In Section 12.3, we use the existence of these holomorphic planes to provide a new proof of the vanishing of CH for overtwisted contact structures.

(4) In Section 12.4, we state how the intersection numbers of Section 11 can be used to define a grading $\mathcal{I}_\Lambda$ on the CH chain complex for α_ϵ.

(5) In Section 12.5, we compute the homology classes and Conley–Zehnder indices of R_ϵ orbits appearing after application of contact surgeries to the $\mathrm{tb} = 1$, right-handed trefoil.

(6) In Section 12.6, we combine computations of Section 12.5 with the results of Sections 12.2 and 12.4 to prove Theorem 1.2.

For notational simplicity, we will ignore mention of specific contact forms α_ϵ, assuming that each contact manifold $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ is equipped with such a contact form with ϵ small enough to guarantee that all orbits under consideration are hyperbolic and that Theorem 7.1 may be applied. Assumptions 11.4 are also in effect. When working with symplectizations of $(\mathbb{R}^3, \xi_{\mathrm{std}})$, we assume that we're using the standard almost-complex structure J_0.

12.1 Surgeries on the standard unknot

Let Λ be the Legendrian unknot with $\mathrm{tb} = -1$ and $\mathrm{rot} = 0$, depicted as a figure 8 in the Lagrangian projection in Figure 25. Performing contact -1 surgery will produce the standard contact lens space $L(2,1)$ — the unit cotangent bundle of S^2, or alternatively

Figure 25: Contact surgeries on the $\mathrm{tb} = -1$ unknot with push-outs of their unique embedded Reeb orbits. A -1 $(+1)$ surgery is applied on the left (right) subfigure.

the unit circle bundle associated to the line bundle $\mathcal{O}(-2) \to \mathbb{P}^1$. We'll denote this contact lens space by $(L(2,1), \xi_{\mathrm{std}})$. Performing contact $+1$ surgery produced the standard contact $S^1 \times S^2$ — see Theorem 2.8 — denoted by $(S^1 \times S^2, \xi_{\mathrm{std}})$.

We can arrange that the Lagrangian projection of Λ has a single crossing corresponding to a Reeb chord we denote by r, so that after performing a contact ± 1 surgery there is only a single embedded orbit (r) with cyclic word r. Push-outs of (r) using a choice of capping path are shown in Figure 25. As $\mathrm{rot}(\Lambda) = 0$, the framing (X, Y) described in Section 6 is nowhere-vanishing. For either choice of surgery coefficient, the first homology H_1 is generated by a meridian μ of Λ with

$$H_1(L(2,1)) = (\mathbb{Z}/2\mathbb{Z})\mu, \quad H_1(S^1 \times S^2) = \mathbb{Z}\mu.$$

Theorem 12.1 *The Conley–Zehnder gradings $|*|_{X,Y}$ on*

$$\widehat{\mathrm{CH}}(L(2,1), \xi_{\mathrm{std}}) \quad \text{and} \quad \widehat{\mathrm{CH}}(S^1 \times S^2, \xi_{\mathrm{std}})$$

are canonical in the sense of Proposition 2.5. We compute

$$\widehat{\mathrm{CH}}(L(2,1), \xi_{\mathrm{std}}) = \mathbb{Q}[z_0, z_2, \dots, z_{2k}, \dots], \quad |z_{2k}|_{X,Y} = 2k, \quad [z_{2k}] = \mu \in H_1,$$

for the lens space and

$$\widehat{\mathrm{CH}}(S^1 \times S^2, \xi_{\mathrm{std}}) = \textstyle\bigwedge_{k=1}^{\infty} \mathbb{Q} z_{2k-1}, \quad |z_{2k-1}|_{X,Y} = 2k-1, \quad [z_{2k-1}] = 0 \in H_1,$$

for $S^1 \times S^2$.

Proof For either choice of surgery coefficient $c = \pm 1$, we may compute Conley–Zehnder indices of (r) using a capping path η. We see that the rotation angle of η is $\frac{3}{2}\pi$, so that its rotation number 1. We conclude that

$$\mathrm{CZ}_{X,Y}((r^k)) = \begin{cases} k & \text{if } c = -1, \\ 2k & \text{if } c = +1. \end{cases}$$

Here and throughout the remainder of the proof, $(r^k) = (r \cdots r)$ is the k–fold cover of the embedded orbit (r) for $k > 0$. To sanity check our index computations against known results, we may:

(1) Compare the case $c = -1$ with [7, Section 7.1], in which contact -1 surgery is applied to Λ.

(2) Compare the case $c = +1$ with [20, Lemma 4.2], in which a contact 1–handle is attached to $(\mathbb{R}^3, \xi_{\text{std}})$ to obtain $\widehat{(S^1 \times S^2, \xi_{\text{std}})}$.

In each case a single closed, embedded orbit is produced with Conley–Zehnder index as described in the present scenario.

For the homology classes of orbits, we may apply Theorem 9.1, or simply look at the push-outs depicted in Figure 25 to compute

$$[(r)] = \begin{cases} \mu & \text{if } c = -1, \\ 0 & \text{if } c = +1. \end{cases}$$

As the framing (X, Y) is nonvanishing, we conclude that $\widehat{\text{CH}}$ is canonically $\mathbb{Z}$–graded for either choice of surgery coefficient, for, when $c = -1$, we have a $\mathbb{Q}$–homology sphere and, when $c = +1$, all orbits are homologically trivial.

When $c = -1$, an orbit (r^k) is bad exactly when $k \mod 2 = 0$. Write z_{2k} for the orbit (r^{2k-1}). Then the $\widehat{\text{CH}}$ chain algebra is freely generated by the z_{2k} with gradings as described in the statement of the theorem. As the $\text{CZ}_{X,Y}$ grading is even, ∂_{CH} must vanish. The theorem is now complete in the case $c = -1$.

When $c = +1$, all of the (r^k) are good orbits, which we will denote by z_{2k-1}. These are graded as described in the statement of the theorem. As (r) is the unique orbit of index 1, $\partial_{\text{CH}}(r)$ must be a count of holomorphic planes. If this count was nonzero, then the unit in $\widehat{\text{CH}}$ would be exact. This is impossible, as $(S^1 \times S^2, \xi_{\text{std}})$ bounds the Liouville domain

$$(S^1 \times \mathbb{D}^3, x\, dy - y\, dx + z\, d\theta),$$

implying that $\text{CH}(S^1 \times S^2, \xi_{\text{std}}) \neq 0$ and so $\widehat{\text{CH}}(S^1 \times S^2, \xi_{\text{std}}) \neq 0$ by Theorem 2.6. We conclude $\partial_{\text{CH}}(r) = 0$.

For $c = +1$ and $k > 1$, the contact homology differential of (r^k) is determined by counts of pairs of pants $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ with

(1) ∞ positively asymptotic to (r^k),

(2) 0 negatively asymptotic to some (r^{k_0}),

(3) 1 negatively asymptotic to some (r^{k_1}), and

(4) $k = k_0 + k_1$, as required by the index formula, equation (8).

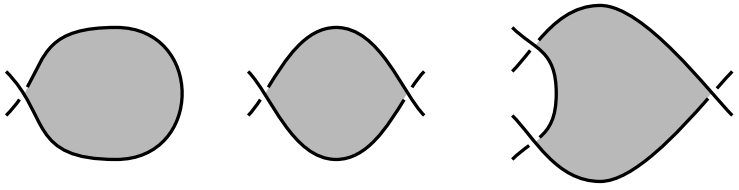


Figure 26: Some *RSFT* disks with only positive punctures.

The energies of any such curves must be 0, indicating that these curves must be branched covers of the trivial cylinder over (r). According to calculations of Fabert [28], the contact homology differential must be strictly action-decreasing, implying that the counts of such curves are 0. We conclude $\partial_{\mathrm{CH}}(r^k) = 0$, completing the proof. □

12.2 Bubbling planes in surgery diagrams

In this section we use the results of Section 11 to count holomorphic curves in completed surgery cobordisms $(\overline{W}_{+1}, \overline{\lambda}_{+1})$ determined by certain *LRSFT* disks on Legendrian links in $(\mathbb{R}^3, \xi_{\mathrm{std}})$ with only positive punctures. The arguments can be generalized to Legendrians Λ^0 in arbitrary punctured contact manifolds $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$, with additional notation and hypothesis. We consider *LRSFT* disks with arbitrary numbers of positive punctures, although in the applications of Sections 12.3 and 12.5 we'll only need to look at disks with a single positive puncture.

As mentioned in the introduction, the inspiration for our construction is Hofer's bubbling argument [35], used to prove the Weinstein conjecture — that every Reeb vector field on a given contact manifold has a closed orbit — for certain contact 3–manifolds. We also have in mind the holomorphic curves in contact -1 surgery cobordisms of [7; 18] positively asymptotic to closed orbits and negatively asymptotic to chords of a Legendrian link. In the case of $+1$ surgery, we will see some curves for which these boundary conditions have been flipped upside-down, allowing us to interpolate between chords of Legendrian links and Reeb orbits appearing after contact $+1$ surgery.

Suppose that $\Lambda^0 \subset (\mathbb{R}^3, \xi_{\mathrm{std}})$ has an immersed *LRSFT* disk $u\colon \mathbb{D} \setminus \{p_k\} \to \mathbb{R}^2$ for some boundary punctures $\{p_k\}$ as in Figure 26. Specifically, we assume that u is an embedding with only positive punctures, completely covering a connected component of $\mathbb{R}^2 \setminus \pi_{x,y}(\Lambda_0)$. Write $r_{j_1}, \dots, r_{j_n}$ for the chords associated to the punctures of the disk indexed in a counterclockwise fashion along its boundary and write

$$U : \mathbb{D} \setminus \{p_k\} \to \mathbb{R} \times (\mathbb{R}^3, \xi_{\mathrm{std}})$$

for the associated holomorphic curve with boundary mapping to $\mathbb{R}\times\Lambda^0$ determined by the drawing-to-disk correspondence, Corollary 11.2.

Let (x_k, y_k) be a basis of points for Λ^0, indexed so that (x_1, y_1) lies in the interior of the image of u. Then, by our hypothesis on u,

$$\mathbb{C}_k \cdot U = \begin{cases} 1 & \text{if } k = 1, \\ 0 & \text{if } k \neq 1. \end{cases} \tag{64}$$

Consider the completed cobordism $(\overline{W}_{+1}, \overline{\lambda}_{+1})$ obtained by performing contact $+1$ surgery on Λ^0 as described by Theorem 10.1. Then we may consider U as having $\overline{W}_{+1}$ as its target with boundary on an embedded union of Lagrangian planes $\overline{\mathbb{D}}_{+1,i}$ — as described in Section 10 — whose intersection with the positive end of $\overline{W}_{+1}$ is $[0,\infty)\times\Lambda^0$. We simply write $\overline{\mathbb{D}}_{+1}$ for this union of planes. We may consider the planes $\mathbb{C}_k$ as being contained in any of $\mathbb{R}\times\mathbb{R}^3$, $\overline{W}_{+1}$ or $\mathbb{R}\times\mathbb{R}^3_\Lambda$.

We consider the following moduli spaces:

(1) $\mathcal{M}_{\mathbb{R}^3}$ is the moduli space of holomorphic disks in $\mathbb{R}\times\mathbb{R}^3$ with positive punctures asymptotic to the $r_1,\dots,r_n$ and boundary on $\mathbb{R}\times\Lambda^0$ satisfying (64).

(2) $\mathcal{M}_{\overline{W}_{+1}}$ is the moduli space of holomorphic disks in $\overline{W}_{+1}$ with positive punctures asymptotic to the $r_1,\dots,r_n$ and boundary on $\overline{\mathbb{D}}_{+1}$ satisfying (64).

(3) $\mathcal{M}_{\mathbb{R}^3_\Lambda}$ is the moduli space of holomorphic planes in $\mathbb{R}\times\mathbb{R}^3_\Lambda$ positively asymptotic to the closed orbit $(r_{j_1}\cdots r_{j_n})$ and satisfying (64).

Within the positive end of the completed cobordism, we can translate U positively in the $\mathbb{R}$ direction, determining a half-infinite ray $[0,\infty)\subset\mathcal{M}_{\overline{W}_{+1}}$. The index of U is equal to 1, so that these curves are regular. Following the analogy with [35], these disks will serve as our Bishop family.

Theorem 12.2 *The boundary of the* SFT *compactification $\overline{\mathcal{M}}_{\overline{W}_{+1}}$ of the moduli space $\mathcal{M}_{\overline{W}_{+1}}$ consists of two points (when curves in symplectizations are considered equivalent modulo $\mathbb{R}$–translation). One point is given by $\mathbb{R}$–translations of the curve U, considered as living in $\mathbb{R}\times\mathbb{R}^3$. The other point is given by a height* 3 SFT *building consisting of*:

(1) *A collection of trivial strips over the r_{j_k} in $\mathbb{R}\times\mathbb{R}^3$.*

(2) *A hidden curve U_c^o in $\overline{W}_{+1}$ from a disk with n boundary punctures positively asymptotic to the r_{j_k} — preserving the cyclic ordering of the r_{j_k} — and a single interior puncture negatively asymptotic to the closed Reeb orbit $(r_{j_1}\cdots r_{j_n})$.*

(3) *A curve $U_\varnothing^c\in\mathcal{M}_{\mathbb{R}^3_\Lambda}$.*

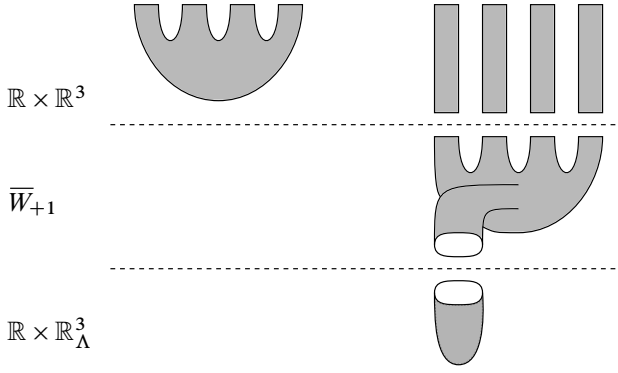


Figure 27: Elements of $\partial\overline{\mathcal{M}}_{\overline{W}_{+1}}$.

The algebraic count of such U_c^o is ±1 and the algebraic count of points in $\mathcal{M}_{\mathbb{R}^3_\Lambda}$ is also ±1.

The two buildings in $\partial\overline{\mathcal{M}}_{\overline{W}_{+1}}$ are shown in Figure 27. The notation U_c^o indicates that the curves interpolates between open and closed strings — that is, between chords and orbits — and this curve is shown in the center-right of Figure 27. The curve $U_\varnothing^c$ is shown in the bottom-right of the figure.

Proof The space $\partial\overline{\mathcal{M}}_{\overline{W}_{+1}}$ consists of multilevel SFT buildings such that, when their levels are glued together, an index 1 curve obeying the topological hypotheses on $\mathcal{M}_{\overline{W}_{+1}}$ is obtained. Subject to these conditions, such buildings may be of any of the following configurations:

(1) **Case $(1,\varnothing,\varnothing)$** A 3–level building consisting of an index 1 curve in $\mathbb{R}\times\mathbb{R}^3$, an empty curve in $\overline{W}_{+1}$, and an empty curve in the symplectization of the surgered manifold $\mathbb{R}\times\mathbb{R}^3_\Lambda$.

(2) **Case $(1,0,\varnothing)$** A 3–level building consisting of an index 1 curve in $\mathbb{R}\times(\mathbb{R}^3,\xi_{\text{std}})$, a collection of index 0 curves in $\overline{W}_{+1}$, and an empty curve in $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$.

(3) **Case $(0,0,1)$** A 3–level building consisting of a collection of index 0 curves in $\mathbb{R}\times\mathbb{R}^3$, a collection of index 0 curves in $\overline{W}_{+1}$, and an index 1 curve in $\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}$.

The buildings are required to recover the boundary conditions of U when glued in the obvious way. Buildings of height greater than 3 are ruled out by presumption of transversality for somewhere-injective curves in Assumptions 11.4, index additivity and the fact that all closed orbits of R_ϵ at the negative end of $\overline{W}_{+1}$ are assumed hyperbolic,

so that there cannot be levels consisting of branched covers of trivial cylinders with ind ≤ 0 as described in [41, Section 1].

We will show, using the intersections with the $\mathbb{C}_k$, that

(1) U is the only possibility for the case $(1, 0, \varnothing)$,

(2) there are no curves in the case $(1, 0, \varnothing)$, and

(3) the second configuration described in the statement of the proposition — appearing in Figure 27, right — is the only possibility for the case $(0, 0, 1)$.

Case (1, ∅, ∅) For the case $(1, \varnothing, \varnothing)$, our assumptions on the immersion u indicate that U is the only disk in $\mathbb{R} \times \mathbb{R}^3$ satisfying (64). We conclude that U is then the only possibility in this case.

Case (1, 0, ∅) Next, suppose we have a holomorphic building satisfying the conditions of the case $(1, 0, \varnothing)$ and note that the middle level — a union of curves in $\overline{W}_{+1}$ we'll denote by $U_{\overline{W}_{+1}}$ — must be positively asymptotic to some number of chords and have no negative asymptotics. Hence, each connected component of $U_{\overline{W}_{+1}}$ must be exposed by Proposition 11.17.[20] The conditions on intersection numbers of (64) then indicate that $U_{\overline{W}_{+1}}$ must consist of a single component and that the upper level of this building $U_{\mathbb{R}\times\mathbb{R}^3}$ must be hidden.

For each component of $U_{\mathbb{R}\times\mathbb{R}^3}$, the number of positive punctures must match the number of negative punctures, as otherwise Proposition 11.16 would indicate that this component is exposed. If any component had more than a single negative puncture, then $U_{\overline{W}_{+1}}$ would have more than a single connected component in violation of the above arguments. We conclude that $U_{\mathbb{R}\times\mathbb{R}^3}$ must be a union of hidden strips, which are then trivial by Proposition 11.20.

Since $U_{\mathbb{R}\times\mathbb{R}^3}$ is a collection of trivial strips, it must then have ind $= 0$, in violation of our hypothesis. We conclude that no buildings of type $(1, 0, \varnothing)$ can exist.

Case (0, 0, 1) Finally, we address configurations of type $(0, 0, 1)$. Suppose that we have such a height 3 building whose levels — going from top to bottom — will be denoted by $U_{\mathbb{R}\times\mathbb{R}^3}$, $U_{\overline{W}_{+1}}$ and $U_{\mathbb{R}\times\mathbb{R}^3_\Lambda}$. By Proposition 11.17, $U_{\mathbb{R}\times\mathbb{R}^3_\Lambda}$ must be exposed and so, by (64), both $U_{\mathbb{R}\times\mathbb{R}^3}$ and $U_{\overline{W}_{+1}}$ must be hidden. Then $U_{\mathbb{R}\times\mathbb{R}^3_\Lambda}$ must consist of a holomorphic plane positively asymptotic to some orbit γ. The curve $U_{\overline{W}_{+1}}$

[20] By connected component we intend that nodal configurations, such as those appearing in the appendix of [12], are broken up into their irreducible pieces, with any removable boundary singularities filled in. We maintain this convention throughout the remainder of the proof.

must then consist of a single connected component negatively asymptotic to γ, as any additional components would necessarily have trivial negative asymptotics and therefore be exposed by Proposition 11.17. As its index is zero, $U_{\mathbb{R}\times\mathbb{R}^3}$ must be a collection of trivial strips. We conclude that $U_{\overline{W}_{+1}}$ must consist of a punctured disk exactly as described in the statement of the proposition. We know that the negative puncture of $U_{\overline{W}_{+1}}$ must be asymptotic to $(r_{j_1}\cdots r_{j_n})$ by Proposition 11.19.

Apart from the statement regarding algebraic counts, our proof is complete. To prove this last statement, observe that $\partial\overline{\mathcal{M}}_{\overline{W}_{+1}}$ has a count of 0 points when taking into account some choice of orientation as it is the boundary of a 1–manifold. We can also write

$$\#\partial\overline{\mathcal{M}}_{x_0,y_0} = \#((1,\varnothing,\varnothing)\text{ buildings}) + \#((1,0,\varnothing)\text{ buildings}) + \#((0,0,1)\text{ buildings}),$$

where the $\#(\cdots)$ are counted with signs. We know that the set of $(1,\varnothing,\varnothing)$ buildings consists of a single element yielding a count of ±1 and that the set of $(1,0,\varnothing)$ buildings must be empty by our previous arguments providing a count of 0. Hence, the number of $(0,0,1)$ buildings must be ∓1. But this number is equal to $\#(U_{\overline{W}_{+1}})\cdot\#(U_{\mathbb{R}\times\mathbb{R}^3_\Lambda})$, so that both numbers must have absolute value 1. Observing that $\#(U_{\mathbb{R}\times\mathbb{R}^3_\Lambda})$ coincides with a count of points in the moduli space $\mathcal{M}_{\mathbb{R}\times\mathbb{R}^3_\Lambda}$, the proof is complete. □

12.3 Vanishing invariants of overtwisted contact manifolds

Here we use the holomorphic planes of Section 12.2 to prove that the contact homologies of overtwisted contact 3–manifolds are 0. Throughout, we write $(M_{\mathrm{OT}},\xi_{\mathrm{OT}})$ for a closed, overtwisted contact 3–manifold.

Theorem 12.3 [65] $\widehat{\mathrm{CH}}(M_{\mathrm{OT}},\xi_{\mathrm{OT}}) = \mathrm{CH}(M_{\mathrm{OT}},\xi_{\mathrm{OT}}) = 0$.

Proof Applying Eliashberg's theorem [21; 39], which asserts that isotopy classes of overtwisted contact structures on a given contact 3–manifold are classified by the homotopy classes of their underlying oriented 2–plane fields, we know that, for each $n\in\mathbb{Z}$, there exists a unique overtwisted contact structure ξ_n on S^3 whose d_3 invariant is $n-\frac{1}{2}$. For the tight contact structure (S^3,ξ_{std}) on S^3, we have $d_3(\xi_{\mathrm{std}}) = -\frac{1}{2}$.[21] Denoting contact-connected sum by $\#$ and isotopic contact structures by $\simeq$,

$$(M_{\mathrm{OT}},\xi_{\mathrm{OT}}) \simeq (M_{\mathrm{OT}},\xi_{\mathrm{OT}})\#(S^3,\xi_{\mathrm{std}}) \simeq (M_{\mathrm{OT}},\xi_{\mathrm{OT}})\#(S^3,\xi_{-1})\#(S^3,\xi_1).$$

By the connected-sum formula of Theorem 2.6, then, we only need to show that $\widehat{\mathrm{CH}}(S^3,\xi_1)=0$.

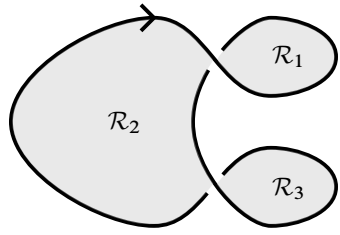


Figure 28: A basis for the $\mathrm{tb} = -2$, $\mathrm{rot} = 1$ unknot.

A contact surgery diagram for (S^3, ξ_1) is provided by a contact $+1$ surgery on a $\mathrm{tb} = -1$, $\mathrm{rot} = 1$ unknot. See [52, Lemma 11.3.10]. A Lagrangian resolution of this knot Λ — shown in Figure 28 — has two chords, say r_1 and r_2. Perturbing Λ as necessary, we may assume that the actions of the chords are distinct and that r_1 has the least action of the two chords with r_1 corresponding to the positive puncture of the disk determined by the region $\mathcal{R}_1$ of Figure 28. Applying the Conley–Zehnder index calculations of Theorem 7.1 to the figure, we see that the Reeb orbit (r_1) has $\mathrm{CZ}_{X,Y} = 2$. Moreover, the orbit is contractible as can be seen by considering a push-out by the orbit string $\eta_{1,1}$.

As the action of (r_1) is the least among all orbits R_ϵ according to our chord-to-orbit correspondence (Theorem 5.1) and the action estimates of Proposition 5.13, $\partial_{\mathrm{CH}}(r_1)$ and $\partial_{\mathrm{SFT}}(r_1)$ are counts of planes bounding (r_1). Using the notation of Section 11.5, write $\mathcal{E}_i$ for the areas of the regions $\mathcal{R}_i \subset \mathbb{R}^2 \setminus \pi_{xy}(N_\epsilon)$ shown in Figure 28. By taking the ϵ parameter in α_ϵ to be sufficiently small, we may assume that $\mathcal{E}_2, \mathcal{E}_3 > \mathcal{A}((r_1))$. Likewise, by Stokes' theorem, $\mathcal{A}((r_1)) - \mathcal{E}_1$ is positive. and may be assumed arbitrarily small by taking ϵ to be arbitrarily small. Then, by the action-energy bound of Proposition 11.8 and the exposure of filling curves (Proposition 11.17), any plane $U : \mathbb{C} \to \mathbb{R} \times \mathbb{R}^3_\Lambda$ bounding (r_1) must satisfy

$$\mathbb{C}_k \cdot U = \begin{cases} 1 & \text{if } k = 1, \\ 0 & \text{if } k \neq 1. \end{cases}$$

We can view $\mathcal{R}_1$ as determining a disk with a positive puncture at the chord r_1, apply Theorem 12.2 to obtain a holomorphic plane bounding (r_1), and conclude that the count of such planes is ± 1. Hence,

$$\partial_{\mathrm{CH}}(r_1) = \pm 1 \in \mathbb{Q},$$

so that the unit in $\widehat{\mathrm{CH}}$ is zero. This implies that $\mathrm{CH}(M_{\mathrm{OT}}, \xi_{\mathrm{OT}})$ must also be zero by Theorem 2.6. □

[21] See [52, Section 11.3] for an overview of d_3 invariants (which we will be following in this proof) as defined by Gompf [31, Section 4].

12.4 Intersection gradings on $\widehat{\mathrm{CH}}$ chain complexes

Here we describe how the intersections of finite-energy curves with the planes $\mathbb{C}_k$ of Section 11.4 can define gradings on the $\mathrm{CC}_{*,0}(\alpha_\epsilon)$ chain complexes of punctured $\mathbb{Q}$–homology spheres which take values in a free $\mathbb{Z}$–module. As described in the introduction, this is simply a variation of the transverse knot filtrations of [14, Section 7.2].

It will be clear from their construction that analogous gradings — which depend on a surgery presentation of our punctures contact manifold — can be constructed for holomorphic curve invariants of $\mathbb{Q}$–homology spheres $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ such as $\widehat{\mathrm{ECH}}$ and the $\widehat{SFT}$. It will also be clear that the assumption that $H_2(M) = 0$ may be dropped by considering $\mathbb{Q}[H_2(M)]$ coefficient systems as described in [6]. Likewise, such gradings can be extended to all of $\mathrm{CC}_{*,*}$ using $\mathbb{Q}[H_2(M)]$ coefficients and spanning surfaces bounding unions of closed orbits and fixed representatives of homology classes as in [6]. In Section 12.6 we will use this grading to prove Theorem 1.2, in which case we will only need the $\mathrm{CC}_{*,0}$ version of this construction for $\mathbb{Q}$–homology spheres.

Let $(\mathbb{R}^3_{\Lambda^\pm}, \xi_{\Lambda^\pm})$ be a contact manifold determined by a contact surgery diagram $\Lambda^\pm$ with $\mathbb{R}^3_{\Lambda^\pm}$ a $\mathbb{Q}$–homology sphere. Let (x_k, y_k) for $k = 1, \dots, K$ be a point basis for the surgery diagram determining a finite collection of infinite-energy holomorphic planes $\mathbb{C}_k$ as described in Section 11.5.

Suppose $\gamma = \{\gamma_k\}$ is a collection of Reeb orbits for which $[\gamma] = 0 \in H_1(\mathbb{R}^3_{\Lambda^\pm})$ and let S_γ be a surface in $\mathbb{R}^3_{\Lambda^\pm}$ with $\partial S_\gamma = \gamma$. To the surface S_γ and each point (x_k, y_k), we define

$$I_k(\gamma) = \big(\{(x_k, y_k)\} \times \mathbb{R}\big) \cdot S_\gamma \in \mathbb{Z}.$$

By Theorem 11.6 and the fact that $H_2(\mathbb{R}^3_{\Lambda^\pm}) = 0$, the numbers $I_k(\gamma)$ are independent of choice of spanning surface S_γ for γ. We collect all of these numbers as monomials

$$\mathcal{I}_\Lambda(\gamma) = \sum_1^K I_k(\gamma)\iota_k \in \mathbb{Z}^K$$

for formal variables ι_k for $k = 1, \dots, K$. It follows from this definition that, provided two homologically trivial collections γ_1, γ_2 of closed Reeb orbits, we have

$$\mathcal{I}_\Lambda(\gamma_1 \cup \gamma_2) = \mathcal{I}_\Lambda(\gamma_1) + \mathcal{I}_\Lambda(\gamma_2).$$

We set $\mathcal{I}_\Lambda(\varnothing) = 0 \in \mathbb{Z}^K$. Then $\mathcal{I}_\Lambda$ determines a $\mathbb{Z}^K$–valued grading on the $H_1 = 0$ subalgebra $\mathrm{CC}_{*,0}$ of the chain algebra CC for the contact homology associated to the contact form α_ϵ of $\mathbb{R}^3_{\Lambda^\pm}$.

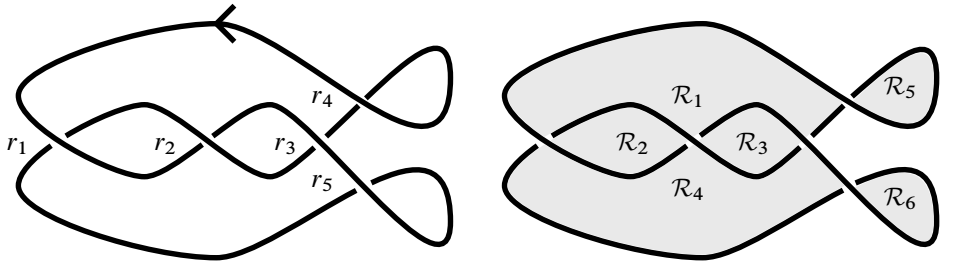


Figure 29: A Legendrian trefoil with $\mathrm{tb} = -1$ and $\mathrm{rot} = 0$ in the Lagrangian projection together with a basis for $\mathbb{R}^2 \setminus N$.

Now suppose that γ^+ and γ^- are two homologically trivial collections of closed orbits and that U is a map from a surface with boundary into $\mathbb{R}^3_{\Lambda^\pm}$ for which $\partial U = \gamma^+ - \gamma^-$. Then, relative to its boundary, we have

$$\big(\{(x_k, y_k)\} \times \mathbb{R}\big) \cdot U = I_k(\gamma^+) - I_k(\gamma^-) \in \mathbb{Z}.$$

In particular, if $(t, U) \colon \Sigma' \to \mathbb{R} \times \mathbb{R}^3_{\Lambda^\pm}$ is a holomorphic curve positively asymptotic to the γ^+ and negatively asymptotic to the γ^-, then

$$\mathcal{I}_\Lambda(\gamma^+) - \mathcal{I}_\Lambda(\gamma^-) = \sum \big(\big(\{(x_k, y_k)\} \times \mathbb{R}\big) \cdot U\big)\iota_k = \sum (\mathbb{C}_k \cdot (t, U))\iota_k \in \mathbb{Z}^K_{\geq 0}. \tag{65}$$

In summary, the $\mathcal{I}_\Lambda$ allows us to make a priori computations of intersection numbers between holomorphic curves asymptotic to orbits with leaves of the foliation as described in Section 11. In particular, if

$$\mathcal{I}_\Lambda(\gamma^+) - \mathcal{I}_\Lambda(\gamma^-) \notin \mathbb{Z}^K_{\geq 0}, \tag{66}$$

then the coefficient of γ^- in $\partial_{\mathrm{CH}}(\gamma^+)$ must be zero.[22]

12.5 Surgery on a trefoil

Take Λ to be the trefoil depicted in Figure 29 with chords $r_1, \ldots, r_5$. This is a reproduction of Figure 6 with a point basis shown in the right-hand side of the figure. This trefoil is the unique nondestabilizable $m(3_1)$ by [25].

[22] Here it is implicit that, if the collection γ^+ contains more than a single orbit, then a holomorphic map (t, U) as above contributing to ∂_{CH} will consist of a connected index 1 holomorphic curve positively asymptotic to some orbit in γ^+ together with a union of trivial cylinders over the remaining orbits in the collection. This deviation from convention allows us to associate cobordisms to differentials of monomials consisting of γ^+ containing more than one orbit.

12.5.1 Ambient geometry According to Theorem 9.1, the first homology of $\mathbb{R}^3_{\Lambda^\pm}$ is generated by the meridian μ with

$$H_1(\mathbb{R}^3_{\Lambda^\pm}) = \begin{cases} \mathbb{Z}/2\mathbb{Z}\mu & \text{if } c = 1, \\ \mathbb{Z}\mu & \text{if } c = -1. \end{cases}$$

Since Λ is smoothly fibered, with fiber a punctured torus, the closed manifold obtained by contact -1 surgery — a topological 0 surgery with respect to the Seifert framing — is a torus bundle over S^1. This manifold is Liouville fillable, and hence tight, and so is a torus bundle covered by the classification in [36, Section 2].

Performing $+1$ contact surgery produces a tight but nonfillable contact manifold studied in [43] — see also [52, Theorem 1.3.4] — which is a Brieskorn sphere with reversed orientation, $-\Sigma(2,3,4)$. Nonfillability may also be viewed as a consequence of the fact that the trefoil is not slice by [15], as mentioned in Theorem 2.8.

12.5.2 Rotation numbers and crossing monomials Here we compute rotation numbers and crossing monomials for the trefoil, which will allow us to compute Conley–Zehnder indices and homology classes of the orbits in the surgered manifolds by applying Theorems 7.1 and 9.1, respectively.

To compute the rotation numbers, we first find the rotation angles θ_{j_1,j_2}, which we see are all either $\frac{1}{2}\pi$, $\frac{3}{2}\pi$ or $\frac{5}{2}\pi$, producing the following table:

chord	$\mathrm{rot}_{j,1}$	$\mathrm{rot}_{j,2}$	$\mathrm{rot}_{j,3}$	$\mathrm{rot}_{j,4}$	$\mathrm{rot}_{j,5}$
r_1	0	0	0	0	1
r_2	0	0	0	0	1
r_3	0	0	0	0	1
r_4	1	1	1	1	2
r_5	0	0	0	0	1

For the computation of the crossing monomials, there is only a single μ_i, so that Remark 3.2 is applicable. The μ coefficients of the relevant crossing monomials are:

chord	sgn	$\mathrm{cr}_j : c = 1$	$\mathrm{cr}_j : c = -1$	$\mathrm{cr}_{j,1}$	$\mathrm{cr}_{j,2}$	$\mathrm{cr}_{j,3}$	$\mathrm{cr}_{j,4}$	$\mathrm{cr}_{j,5}$
r_1	1	2	0	0	0	2	3	1
r_2	1	2	0	0	0	0	1	1
r_3	1	2	0	-2	0	0	1	-1
r_4	-1	0	-2	1	1	3	4	2
r_5	-1	0	-2	-1	1	1	2	0

$\mathrm{cw}(\gamma)$	$\mu : c = 1$	$\mathrm{CZ}_{X,Y} : c = 1$	$\mu : c = -1$	$\mathrm{CZ}_{X,Y} : c = -1$
r_1	1	1	0	0
r_2	1	1	0	0
r_3	1	1	0	0
r_4	0	2	1	1
r_5	0	2	-1	1
r_1r_2	0	2	0	0
r_1r_3	0	2	0	0
r_1r_4	1	3	1	1
r_1r_5	1	3	-1	1
r_2r_3	0	2	0	0
r_2r_4	0	3	0	1
r_2r_5	0	3	0	1
r_3r_4	1	3	1	1
r_3r_5	1	3	-1	1
r_4r_5	0	4	0	2

Table 1

12.5.3 Homology classes and indices of orbits after surgery Using the above computations, we can produce in Table 1 the homology classes and Conley–Zehnder indices of Reeb orbits with word length ≤ 2 using Theorems 9.3 and 7.1. Multiply covered orbits have been omitted. Coefficients for μ in the case $c = 1$ are taken modulo 2.

12.6 Proof of Theorem 1.2

In this section, we prove Theorem 1.2 by computing $\partial_{\mathrm{CH}}(r_4)$.

12.6.1 The subalgebra $C_{0,0}$ and intersection gradings As the rotation numbers of capping paths on Λ are bounded below by 0, Theorem 7.1 tells us that the Conley–Zehnder indices of all orbits of are bounded below by their word lengths. We conclude

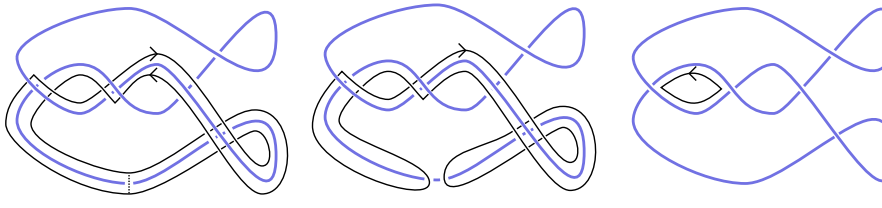


Figure 30: An annulus bounding $(r_1) \cup (r_2)$.

that $\partial_{\mathrm{CH}}(r_4)$ must be an element of $\mathrm{CC}_{0,0}$ which is a commutative algebra on generators

$$1, \quad (r_1)^2, \quad (r_2)^2, \quad (r_3)^2, (r_1)(r_2), \quad (r_1)(r_3), \quad (r_2)(r_3).$$

We'll compute the $\mathcal{I}_\Lambda$ gradings on $\mathrm{CC}_{0,0}$ using points (x_k, y_k) appearing in the centers of the regions $\mathcal{R}_k$ of Figure 29:

$\mathrm{CC}_{*,0}$ monomial	I_1	I_2	I_3	I_4	I_5	I_6
(r_4)	0	0	0	0	1	0
$(r_1)^2$	−1	−1	−2	−1	1	1
$(r_2)^2$	1	2	2	1	−1	−1
$(r_3)^2$	−1	−2	−1	−1	1	1
$(r_1)(r_2)$	0	1	0	0	0	0
$(r_1)(r_3)$	−1	−1	−1	−1	1	1
$(r_2)(r_3)$	0	0	1	0	0	0

To establish the calculations appearing in the above table, we construct surfaces bounding $(r_1)(r_2)$, $(r_2)(r_3)$ and $(r_2)(r_2)$, filling in the remainder of the table using arithmetic. Such surfaces will be constructed out of simple cobordisms built out of homotopies and skein operations. For (r_4), we have an obvious disk bounding a push-out along $\overline{\eta}_4$, obtained by perturbing $\mathcal{R}_5$.

In Figure 30, we construct a spanning surface for the union of the orbits $(r_1) \cup (r_2)$. We begin by homotoping the union of orbits into the complement of N_ϵ, as described in Section 9.4. The result — associated to capping paths η_1 and $\overline{\eta}_2$ — is shown on the left-most subfigure. To get from the left column of the figure to the center, we apply a skein cobordism along the dashed arc, resulting in a pair-of-pants cobordism. The resulting knot can be homotoped to the Reeb orbit $(r_1 r_2)$ as shown in the right-hand side of the figure. So far our surface has avoided passing through any of the lines $\{(x, y) = (x_k, y_k)\} \subset \mathbb{R}^3_{\Lambda^\pm}$. To complete our cobordism, we fill in the knot shown in the right-most subfigure using the obvious disk which is a perturbation of the disk $\mathcal{R}_2$. The union of our pair of pants with this disk provides us with an annular filling of $(r_1) \cup (r_2)$ which intersects the link $\{(x, y) = (x_2, y_2)\}$ exactly once with positive sign. We conclude that

$$\mathcal{I}_\Lambda((r_1)(r_2)) = \iota_2.$$

A similar construction can be carried out to find an annular filling of $(r_2) \cup (r_3)$: we start with a push-out corresponding to capping paths $\overline{\eta}_2$ and η_3, apply a skein cobordism giving us a pair of pants with boundary $(r_2) \cup (r_3) - (r_2 r_3)$, and then fill in $(r_2 r_3)$ with

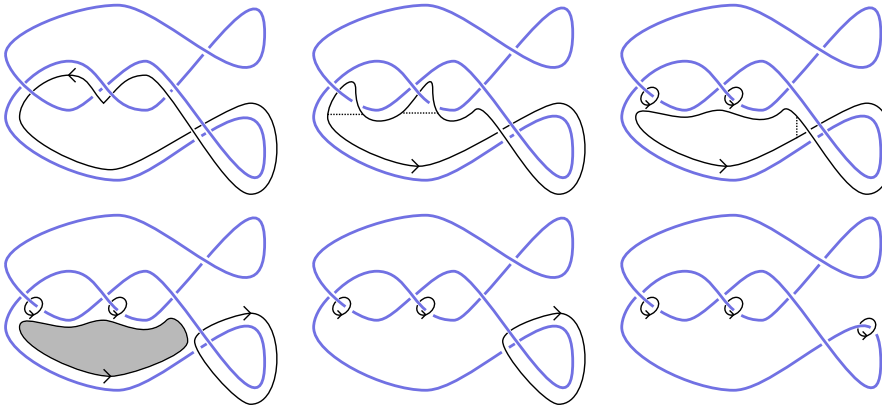


Figure 31: A cobordism with boundary $(r_2)+\mu$.

a perturbation of the disk $\mathbb{D}_3$. We conclude that

$$\mathcal{I}_\Lambda((r_2)(r_3)) = \iota_3.$$

Now we construct a spanning surface for $(r_2)\cup(r_2)$. The construction is more complicated in this case: we construct two cobordisms from (r_2) from a positive and negative meridian of Λ, which can then be patched together to give us a surface with boundary $(r_2)\cup(r_2)$.

We break down the construction of one such cobordism whose boundary is $(r_2)+\mu$ into a sequence of elementary cobordisms, as shown in Figure 31:

(1) We start with a push-out of (r_2) using the capping path $\overline{\eta}_2$, as shown in the top-left subfigure.

(2) Going from the top-left to top-center, we homotop our knot across the disks $\mathcal{R}_2$ and $\mathcal{R}_3$. Along the way, we pick up two intersections with the lines associated to the points (x_2, y_2) and (x_3, y_3) with positive signs.

(3) Going from the top-center to the top-right, we apply skein cobordisms along the dashed arcs appearing in the top-center.

(4) Going from the top-right to the bottom-left, we apply another skein cobordism along the dashed arc appearing in the top-right yielding a 4–component link.

Figure 32: A tube bounding $\mu-\mu$.

(5) Going from the bottom-left to the bottom-center, we fill in one of the components of our link with a disk which is a perturbation of the disk $\mathcal{R}_4$. In doing so, we pick up a positive intersection with the line over the point (x_4, y_4).

(6) Going from the bottom-center to the bottom-right, we homotop one component of our knot over $-\mathcal{R}_6$ to a $-\mu$

Combining all of the above steps, we've constructed a homotopy from (r_2) to a collection of meridians. We can cancel a pair of them with a tube as shown in Figure 32. The end result is a cobordism with boundary $(r_2)+\mu$ passing through the lines associated to the points (x_2, y_2), (x_3, y_3) and (x_4, y_4) once each with positive intersection number and passing through the line over (x_6, y_6) with negative intersection number.

We can also construct a cobordism with boundary $(r_2)-\mu$ by flipping Figure 31 about a horizontal line, starting with a push-out of η_2. The resulting cobordism passes through the lines associated to the points (x_1, y_1), (x_2, y_2) and (x_3, y_3) once each with positive intersection number and passing through the line over (x_5, y_5) with negative intersection number.

We can connect the two cobordisms with another tube bounding $\mu-\mu$ to obtain a spanning surface for $(r_2)\cup(r_2)$. By the above counts of intersections, we have

$$\mathcal{I}_\Lambda((r_2)^2) = \iota_1 + 2\iota_2 + 2\iota_3 + \iota_4 - \iota_5 - \iota_6.$$

Using our calculations of $\mathcal{I}_\Lambda((r_1)(r_2))$, $\mathcal{I}_\Lambda((r_2)(r_3))$ and $\mathcal{I}_\Lambda((r_2)^2)$, we can fill out the remainder of the above table by computing

$$\begin{aligned}
\mathcal{I}_\Lambda((r_1)^2) &= 2\mathcal{I}_\Lambda((r_1)(r_2)) - \mathcal{I}_\Lambda((r_2)^2),\\
\mathcal{I}_\Lambda((r_3)^2) &= 2\mathcal{I}_\Lambda((r_2)(r_3)) - \mathcal{I}_\Lambda((r_2)^2),\\
\mathcal{I}_\Lambda((r_1)(r_3)) &= \mathcal{I}_\Lambda((r_1)(r_2)) + \mathcal{I}_\Lambda((r_2)(r_3)) - \mathcal{I}_\Lambda((r_2)^2).
\end{aligned}$$

12.6.2 Intersection numbers of curves positively asymptotic to (r_4) Now suppose that we have a holomorphic curve U positively asymptotic to (r_4) and negatively asymptotic to a collection of generators γ^- from $C_{0,0}$. Writing γ^- as a monomial in $C_{0,0}$, there are nonnegative constants $C_{i,j}$ for which

$$\gamma^- = (r_1)^{2C_{1,1}}(r_2)^{2C_{2,2}}(r_3)^{2C_{3,3}}((r_1)(r_2))^{C_{1,2}}((r_1)(r_3))^{C_{1,3}}((r_2)(r_3))^{C_{2,3}}.$$

We'll use the intersection grading to show that all of the $C_{i,j}$ must be zero, so that U cannot have any negative asymptotics. We can count the intersection of U with the planes $\mathbb{C}_k$ as the coefficients of the ι_k in the expression $\mathcal{I}_\Lambda((r_4)) - \mathcal{I}_\Lambda(\gamma^-)$ as

described in (65). Using the table above, we compute

$$\begin{aligned}\mathcal{I}_\Lambda((r_4))-\mathcal{I}_\Lambda(\gamma^-) = {}& (C_{1,1}-C_{2,2}+C_{3,3}+C_{1,3})\iota_1\\ &+(C_{1,1}-2C_{2,2}+2C_{3,3}-C_{1,2}+C_{1,3})\iota_2\\ &+(2C_{1,1}-2C_{2,2}+C_{3,3}+C_{1,3}-C_{2,3})\iota_3\\ &+(C_{1,1}-C_{2,2}+C_{3,3}+C_{1,3})\iota_4\\ &+(1-C_{1,1}+C_{2,2}-C_{3,3}-C_{1,3})\iota_5\\ &+(-C_{1,1}+C_{2,2}-C_{3,3}-C_{1,3})\iota_6.\end{aligned}$$

All of the ι_k coefficients above must be nonnegative by intersection positivity.

As the ι_4 and ι_6 coefficients are the same with opposite sign, both must be zero so that

$$C_{2,2}=C_{1,1}+C_{3,3}+C_{1,3}.$$

Therefore, we must have

$$\mathcal{I}_\Lambda((r_4))-\mathcal{I}_\Lambda(\gamma^-)=(-C_{1,1}-C_{1,2}-C_{1,3})\iota_2+(-C_{3,3}-C_{1,3}-C_{2,3})\iota_3+\iota_5,$$

implying that the remaining $C_{i,j}$ are all zero.

12.6.3 Completion of the proof The above analysis implies that, if U is an index 1 holomorphic curve contributing to $\partial_{\mathrm{CH}}(r_4)$, then it cannot have any negative asymptotics and must satisfy

$$\mathbb{C}_k\cdot U=\begin{cases}1 & \text{if } k=5,\\ 0 & \text{if } k\neq 5.\end{cases} \tag{67}$$

Such a curve must be parametrized by $\mathbb{C}$ as per the definition of ∂_{CH}. To complete our proof, we analyze the moduli space of finite-energy curves

$$\mathcal{M}_{4,5}=\{\mathbb{C}\xrightarrow{U}\mathbb{R}\times\mathbb{R}^3_{\Lambda^\pm}: U \text{ asymptotic to } (r_4), \text{ satisfying (67)}\}/\text{reparametrization}.$$

By the above analysis, $\partial_{\mathrm{CH}}((r_4))=\#(\mathcal{M}_{4,5})1$, counting points algebraically. This moduli space exactly describes the lowest levels $U^c_\varnothing$ of the height 3 SFT buildings studied in Theorem 12.2, which when applied to the disk $\mathcal{R}_5$ tell us that $\#(\mathcal{M}_{4,5})=\pm 1$.

The proof of Theorem 1.2 is then complete in the case of the tb $=1$ trefoil shown in Figure 29. By the classification torus knots in $(\mathbb{R}^3,\xi_{\mathrm{std}})$ [25], all other right-handed trefoils are stabilizations of this one — contact $+1$ surgeries on these stabilized knots will be overtwisted and so will have $\mathrm{CH}=0$. The proof is now complete in the case that Λ^+ consists of a single component. In the case that $\Lambda^+=\bigcup_i^n\Lambda_i^+$ has multiple components, we have — as described in Section 10 — a Liouville cobordism (W,λ) whose convex end $(M^+,\xi^+)=(\partial^+W,\ker(\lambda)|_{\partial^+W})$ is given by contact $+1$ surgery

on Λ_1^+ and whose concave end $(M^-, \xi^+) = (\partial^- W, \ker(\lambda)|_{\partial^- W})$ is given by contact surgery on Λ^+. If we index the components of Λ^+ so that Λ_1^+ is a right-handed trefoil, then $\mathrm{CH}(M^+, \xi^+) = 0$, and so, by Liouville functoriality, $\mathrm{CH}(M^-, \xi^-) = 0$ as well. The proof is now complete for all right-handed trefoils and all contact surgery coefficients $1/k$ with $k > 0$.

References

[1] **L V Ahlfors**, *Complex analysis: an introduction to the theory of analytic functions of one complex variable*, 3rd edition, McGraw-Hill, New York (1978) MR Zbl

[2] **R Avdek**, *Contact surgery and supporting open books*, Algebr. Geom. Topol. 13 (2013) 1613–1660 MR Zbl

[3] **E Bao**, **K Honda**, *Definition of cylindrical contact homology in dimension three*, J. Topol. 11 (2018) 1002–1053 MR Zbl

[4] **E Bao**, **K Honda**, *Semi-global Kuranishi charts and the definition of contact homology*, Adv. Math. 414 (2023) art. id. 108864 MR Zbl

[5] **F Bourgeois**, *A Morse–Bott approach to contact homology*, PhD thesis, Stanford University (2002) MR https://www.proquest.com/docview/305591502

[6] **F Bourgeois**, *Introduction to contact homology*, lecture notes (2003) https://www.imo.universite-paris-saclay.fr/~bourgeois/papers/Berder.pdf

[7] **F Bourgeois**, **T Ekholm**, **Y Eliashberg**, *Effect of Legendrian surgery*, Geom. Topol. 16 (2012) 301–389 MR Zbl

[8] **F Bourgeois**, **Y Eliashberg**, **H Hofer**, **K Wysocki**, **E Zehnder**, *Compactness results in symplectic field theory*, Geom. Topol. 7 (2003) 799–888 MR Zbl

[9] **F Bourgeois**, **K Niederkrüger**, *Towards a good definition of algebraically overtwisted*, Expo. Math. 28 (2010) 85–100 MR Zbl

[10] **Y Chekanov**, *Differential algebra of Legendrian links*, Invent. Math. 150 (2002) 441–483 MR Zbl

[11] **K Cieliebak**, **Y Eliashberg**, *From Stein to Weinstein and back: symplectic geometry of affine complex manifolds*, American Mathematical Society Colloquium Publications 59, Amer. Math. Soc., Providence, RI (2012) MR Zbl

[12] **K Cieliebak**, **J Latschev**, *The role of string topology in symplectic field theory*, from "New perspectives and challenges in symplectic field theory" (M Abreu, F Lalonde, L Polterovich, editors), CRM Proc. Lecture Notes 49, Amer. Math. Soc., Providence, RI (2009) 113–146 MR Zbl

[13] **V Colin**, **P Ghiggini**, **K Honda**, *Equivalence of Heegaard Floer homology and embedded contact homology via open book decompositions*, Proc. Natl. Acad. Sci. USA 108 (2011) 8100–8105 MR Zbl

[14] **V Colin**, **P Ghiggini**, **K Honda**, **M Hutchings**, *Sutures and contact homology I*, Geom. Topol. 15 (2011) 1749–1842 MR Zbl

[15] **J Conway**, **J B Etnyre**, **B Tosun**, *Symplectic fillings, contact surgeries, and Lagrangian disks*, Int. Math. Res. Not. 2021 (2021) 6020–6050 MR Zbl

[16] **F Ding**, **H Geiges**, *A Legendrian surgery presentation of contact 3–manifolds*, Math. Proc. Cambridge Philos. Soc. 136 (2004) 583–598 MR Zbl

[17] **T Ekholm**, *Rational symplectic field theory over $\mathbb{Z}_2$ for exact Lagrangian cobordisms*, J. Eur. Math. Soc. 10 (2008) 641–704 MR Zbl

[18] **T Ekholm**, *Holomorphic curves for Legendrian surgery*, preprint (2019) arXiv

[19] **T Ekholm**, **J Etnyre**, **M Sullivan**, *The contact homology of Legendrian submanifolds in $\mathbb{R}^{2n+1}$*, J. Differential Geom. 71 (2005) 177–305 MR Zbl

[20] **T Ekholm**, **L Ng**, *Legendrian contact homology in the boundary of a subcritical Weinstein 4–manifold*, J. Differential Geom. 101 (2015) 67–157 MR Zbl

[21] **Y Eliashberg**, *Classification of overtwisted contact structures on 3–manifolds*, Invent. Math. 98 (1989) 623–637 MR Zbl

[22] **Y Eliashberg**, *On symplectic manifolds with some contact properties*, J. Differential Geom. 33 (1991) 233–238 MR Zbl

[23] **Y Eliashberg**, **A Givental**, **H Hofer**, *Introduction to symplectic field theory*, from "Visions in mathematics" (N Alon, J Bourgain, A Connes, M Gromov, V Milman, editors), volume 2, Birkhäuser (= GAFA special volume), Boston, MA (2000) 560–673 MR Zbl

[24] **J B Etnyre**, *Legendrian and transversal knots*, from "Handbook of knot theory" (W Menasco, M Thistlethwaite, editors), Elsevier, Amsterdam (2005) 105–185 MR Zbl

[25] **J B Etnyre**, **K Honda**, *Knots and contact geometry, I: Torus knots and the figure eight knot*, J. Symplectic Geom. 1 (2001) 63–120 MR Zbl

[26] **J B Etnyre**, **L L Ng**, *Legendrian contact homology in $\mathbb{R}^3$*, from "Surveys in differential geometry 2020: surveys in 3–manifold topology and geometry" (I Agol, D Gabai, editors), Surv. Differ. Geom. 25, International, Boston, MA (2022) 103–161 MR Zbl

[27] **J B Etnyre**, **B Ozbagci**, *Invariants of contact structures from open books*, Trans. Amer. Math. Soc. 360 (2008) 3133–3151 MR Zbl

[28] **O Fabert**, *Obstruction bundles over moduli spaces with boundary and the action filtration in symplectic field theory*, Math. Z. 269 (2011) 325–372 MR Zbl

[29] **P Foulon**, **B Hasselblatt**, *Contact Anosov flows on hyperbolic 3–manifolds*, Geom. Topol. 17 (2013) 1225–1252 MR Zbl

[30] **H Geiges**, **K Zehmisch**, *How to recognize a 4–ball when you see one*, Münster J. Math. 6 (2013) 525–554 MR Zbl

[31] **R E Gompf**, *Handlebody construction of Stein surfaces*, Ann. of Math. 148 (1998) 619–693 MR Zbl

[32] **M Gromov**, *Pseudo holomorphic curves in symplectic manifolds*, Invent. Math. 82 (1985) 307–347 MR Zbl

[33] **A Hatcher**, *Algebraic topology*, Cambridge Univ. Press (2002) MR Zbl

[34] **R Hind**, *Stein fillings of lens spaces*, Commun. Contemp. Math. 5 (2003) 967–982 MR Zbl

[35] **H Hofer**, *Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three*, Invent. Math. 114 (1993) 515–563 MR Zbl

[36] **K Honda**, *On the classification of tight contact structures, II*, J. Differential Geom. 55 (2000) 83–143 MR Zbl

[37] **K Honda**, *Gluing tight contact structures*, Duke Math. J. 115 (2002) 435–478 MR Zbl

[38] **K Honda**, **W H Kazez**, **G Matić**, *On the contact class in Heegaard Floer homology*, J. Differential Geom. 83 (2009) 289–311 MR Zbl

[39] **Y Huang**, *A proof of the classification theorem of overtwisted contact structures via convex surface theory*, J. Symplectic Geom. 11 (2013) 563–601 MR Zbl

[40] **M Hutchings**, *Lecture notes on embedded contact homology*, from "Contact and symplectic topology" (F Bourgeois, V Colin, A Stipsicz, editors), Bolyai Soc. Math. Stud. 26, János Bolyai Math. Soc., Budapest (2014) 389–484 MR Zbl

[41] **M Hutchings**, **C H Taubes**, *Gluing pseudoholomorphic curves along branched covered cylinders, I*, J. Symplectic Geom. 5 (2007) 43–137 MR Zbl

[42] **Ç Kutluhan**, **Y-J Lee**, **C H Taubes**, HF = HM*, I: Heegaard Floer homology and Seiberg–Witten Floer homology*, Geom. Topol. 24 (2020) 2829–2854 MR Zbl

[43] **P Lisca**, **A I Stipsicz**, *Ozsváth–Szabó invariants and tight contact three-manifolds, I*, Geom. Topol. 8 (2004) 925–945 MR Zbl

[44] **D McDuff**, *The structure of rational and ruled symplectic 4–manifolds*, J. Amer. Math. Soc. 3 (1990) 679–712 MR Zbl

[45] **D McDuff**, *Symplectic manifolds with contact type boundaries*, Invent. Math. 103 (1991) 651–671 MR Zbl

[46] **D McDuff**, **D Salamon**, *Introduction to symplectic topology*, 2nd edition, Oxford Univ. Press (1998) MR Zbl

[47] **D McDuff**, **D Salamon**, *J–holomorphic curves and symplectic topology*, American Mathematical Society Colloquium Publications 52, Amer. Math. Soc., Providence, RI (2004) MR Zbl

[48] **A Moreno**, **Z Zhou**, *A landscape of contact manifolds via rational SFT*, preprint (2020) arXiv

[49] **L L Ng**, *Computable Legendrian invariants*, Topology 42 (2003) 55–82 MR Zbl

[50] **L Ng**, *Rational symplectic field theory for Legendrian knots*, Invent. Math. 182 (2010) 451–512 MR Zbl

[51] **B Ozbagci**, *A note on contact surgery diagrams*, Internat. J. Math. 16 (2005) 87–99 MR Zbl

[52] **B Ozbagci**, **A I Stipsicz**, *Surgery on contact 3–manifolds and Stein surfaces*, Bolyai Society Mathematical Studies 13, Springer (2004) MR Zbl

[53] **P Ozsváth**, **Z Szabó**, *Heegaard Floer homology and contact structures*, Duke Math. J. 129 (2005) 39–61 MR Zbl

[54] **J Pardon**, *Contact homology and virtual fundamental cycles*, J. Amer. Math. Soc. 32 (2019) 825–919 MR Zbl

[55] **J Robbin**, **D Salamon**, *The Maslov index for paths*, Topology 32 (1993) 827–844 MR Zbl

[56] **J Rooney**, *Cobordism maps in embedded contact homology*, preprint (2019) arXiv

[57] **L Rudolph**, *The slice genus and the Thurston–Bennequin invariant of a knot*, Proc. Amer. Math. Soc. 125 (1997) 3049–3050 MR Zbl

[58] **M Schwarz**, *Cohomology operations from S^1 cobordisms in Floer homology*, PhD thesis, ETH Zurich (1995) https://doi.org/10.3929/ethz-a-001475365

[59] **P Seidel**, *A biased view of symplectic cohomology*, from "Current developments in mathematics, 2006" (B Mazur, T Mrowka, W Schmid, R Stanley, S-T Yau, editors), International, Somerville, MA (2008) 211–253 MR Zbl

[60] **A Weinstein**, *Contact surgery and symplectic handlebodies*, Hokkaido Math. J. 20 (1991) 241–251 MR Zbl

[61] **C Wendl**, *Strongly fillable contact manifolds and J–holomorphic foliations*, Duke Math. J. 151 (2010) 337–384 MR Zbl

[62] **C Wendl**, *Non-exact symplectic cobordisms between contact 3–manifolds*, J. Differential Geom. 95 (2013) 121–182 MR Zbl

[63] **C Wendl**, *Signs (or how to annoy a symplectic topologist)*, blog post (2015) https://tinyurl.com/Wendl-annoy

[64] **C Wendl**, *Lectures on symplectic field theory*, preprint (2016) arXiv

[65] **M-L Yau**, *Vanishing of the contact homology of overtwisted contact 3–manifolds*, Bull. Inst. Math. Acad. Sin. 1 (2006) 211–229 MR Zbl With an appendix by Yakov Eliashberg

Department of Mathematics, Uppsala University
Uppsala, Sweden

russell.avdek@math.uu.se

https://russellavdek.com

Proposed: András I Stipsicz
Seconded: Leonid Polterovich, Yakov Eliashberg

Received: 8 July 2020
Revised: 11 October 2021

msp

Geometry & Topology **27**:3 (2023) 1083–1202
DOI: 10.2140/gt.2023.27.1083
Published: 8 June 2023

Unexpected Stein fillings, rational surface singularities and plane curve arrangements

OLGA PLAMENEVSKAYA
LAURA STARKSTON

We compare Stein fillings and Milnor fibers for rational surface singularities with reduced fundamental cycle. Deformation theory for this class of singularities was studied by de Jong and van Straten (1998); they associated a germ of a singular plane curve to each singularity and described Milnor fibers via deformations of this singular curve.

We consider links of surface singularities, equipped with their canonical contact structures, and develop a symplectic analog of de Jong and van Straten's construction. Using planar open books and Lefschetz fibrations, we describe all Stein fillings of the links via certain arrangements of symplectic disks, related by a homotopy to the plane curve germ of the singularity.

As a consequence, we show that many rational singularities in this class admit Stein fillings that are not strongly diffeomorphic to any Milnor fibers. This contrasts with previously known cases, such as simple and quotient surface singularities, where Milnor fibers are known to give rise to all Stein fillings. On the other hand, we show that if for a singularity with reduced fundamental cycle, the self-intersection of each exceptional curve is at most -5 in the minimal resolution, then the link has a unique Stein filling (given by a Milnor fiber).

14J17, 32S30, 32S50, 57K33, 57K43; 14H50, 32S25

1 Introduction

The goal of this paper is to compare and contrast deformation theory and symplectic topology of certain rational surface singularities. Using topological tools, we examine symplectic fillings for links of rational surface singularities with reduced fundamental cycle and compare these fillings to Milnor fibers of the singularities. Each Milnor fiber carries a Stein structure and thus gives a Stein filling of the link; however, we show that there is a plethora of Stein fillings that do not arise from Milnor fibers. Milnor fibers and deformation theory are studied in the work of de Jong and van Straten [27] for sandwiched surface singularities (this class includes rational singularities with reduced fundamental cycle). The main feature of their construction is a reduction from *surfaces* to *curves*: deformations of a surface singularity in the given class can be understood via deformations of the germ of a reducible plane curve associated to the singularity. To describe Stein fillings, we develop a symplectic analog of de Jong and van Straten's constructions, representing the fillings via arrangements of smooth (or symplectic) disks in $\mathbb{C}^2$. Our approach is purely topological and thus different from de Jong and van Straten's; their algebrogeometric techniques do not apply in our more general symplectic setting. We work with Lefschetz fibrations and open books, referring to algebraic geometry only for motivation and for the description of smoothings from [27].

Let $X \subset \mathbb{C}^N$ be a singular complex surface with an isolated singularity at the origin. For small $r > 0$, the intersection $Y = X \cap S^{2N-1}_r$ with the sphere $S^{2N-1}_r = \{|z_1|^2 + |z_2|^2 + \cdots + |z_N|^2 = r\}$ is a smooth 3–manifold called the *link of the singularity* $(X, 0)$. The induced contact structure ξ on Y is the distribution of complex tangencies to Y, and is referred to as the *canonical* or *Milnor fillable* contact structure on the link. The contact manifold (Y, ξ), which we will call the *contact link*, is independent of the choice of r, up to contactomorphism.

An important problem concerning the topology of a surface singularity is to compare the Milnor fibers of smoothings of $(X, 0)$ to symplectic or Stein fillings of the link (Y, ξ). A *smoothing* is given by a deformation of X to a surface (the *Milnor fiber*) that is no longer singular. (We discuss smoothings in more detail in Section 2.) Milnor fibers themselves are Stein fillings of (Y, ξ), called *Milnor fillings*. An additional Stein filling can be produced by deforming the symplectic structure on the minimal resolution of $(X, 0)$; see Bogomolov and de Oliveira [10]. For rational singularities, this filling agrees with the Milnor fiber of the *Artin smoothing* component and need not be considered separately; see Section 4. An interesting question is whether the collection

of these expected fillings, taken for all singularities with the same link (Y,ξ), gives all possible Stein fillings of the link. In this article, we will use the term *unexpected Stein filling* to refer to any Stein filling which does not arise as a Milnor fiber or the minimal resolution.

There are very few examples of unexpected Stein fillings in the previously existing literature, none of which are simply connected. In this article, we show that, in fact, unexpected Stein fillings are abundant, and in many cases simply connected, even for the simple class of *rational singularities with reduced fundamental cycle*. These singularities, also known as *minimal singularities* (see Kollár [29]), can be characterized by the conditions that the dual resolution graph is a tree, where each vertex v corresponds to a curve of genus 0, and its self-intersection $v\cdot v$ and valency $a(v)$ satisfy the inequality $-v\cdot v \geq a(v)$. (See Section 2 for more details.) In low-dimensional topology, such graphs are often referred to as trees with no bad vertices. The corresponding plumbed 3–manifolds are L–spaces, ie they have the simplest possible Heegaard Floer homology; see Ozsváth and Szabó [51]. In a sense, links of rational singularities with reduced fundamental cycle are just slightly more complicated than lens spaces. As another measure of low complexity, these contact structures admit planar open book decompositions. In the planar case, the set of Stein fillings satisfies a number of finiteness properties (see for instance Kaloti [28], Lisi and Wendl [35], Plamenevskaya [54] and Stipsicz [60]), which makes it rather surprising that these singularities diverge from the expected.

We construct many specific examples of unexpected Stein fillings for rational singularities with reduced fundamental cycle. Then we show that our examples can be broadly generalized to apply to a large class of singularities with reduced fundamental cycle: we only require that the resolution graph of the singularity contain a certain subgraph to ensure that the link has many unexpected Stein fillings.

Theorem 1.1 *For any $N>0$, there is a rational singularity with reduced fundamental cycle whose contact link (Y_N,ξ_N) admits at least N pairwise nonhomeomorphic simply connected Stein fillings, none of which is diffeomorphic to a Milnor filling (**rel** certain boundary data). Examples of such (Y_N,ξ_N) include Seifert fibered spaces over S^2 corresponding to certain star-shaped resolution graphs.*

The statement also holds for any rational singularity with reduced fundamental cycle whose resolution graph has a star-shaped subgraph as above.

More precise statements are given in Section 7. Our first example which admits simply connected unexpected Stein fillings corresponds to the singularity with resolution

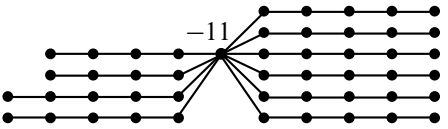


Figure 1: A resolution graph for a singularity whose link admits simply connected unexpected fillings. (Unlabeled vertices have self-intersection -2.) Any graph containing this as a subgraph corresponds to a singularity which also admits simply connected unexpected fillings.

graph in Figure 1. More generally, for every $N > 4$ we can find N distinct unexpected Stein fillings for singularities whose dual resolution graph is star-shaped with at least $2N + 5$ sufficiently long legs, the self-intersection of the central vertex is a large negative number, and the self-intersection of any other vertex is -2.

By contrast, previous results have indicated that for simple classes of singularities, all Stein fillings come from Milnor fibers or the minimal resolution (there are no unexpected fillings). This is true for (S^3, ξ_{std}) by Eliashberg [13], for links of simple and simple elliptic singularities by Ohta and Ono [49; 50], for lens spaces (links of cyclic quotient singularities) by Lisca [34] and Némethi and Popescu-Pampu [45], and in general for quotient singularities by Bhupal and Ono [8] and Park, Park, Shin and Urzúa [52]. Theorem 1.1 breaks this pattern and provides many unexpected fillings. However, we are also able to show that certain classes of rational singularities with reduced fundamental cycle do not admit any unexpected fillings:

Theorem 1.2 *Let $(X, 0)$ be a rational singularity with reduced fundamental cycle with link (Y, ξ), and suppose that each exceptional curve in its minimal resolution has self-intersection at most -5. Then the resolution of $(X, 0)$ is the unique weak symplectic filling of (Y, ξ), up to blow-up, symplectomorphism and symplectic deformation.*

This theorem proves a symplectic analogue of [27, Theorem 6.21], which establishes a special case of a conjecture of Kollár, showing that for singularities as in Theorem 1.2, the base space of a semiuniversal deformation has one component. Thus, they show there is a unique smoothing, whereas we generalize this to show there is a unique minimal symplectic filling. To prove Theorem 1.2, we build on the combinatorial argument of [27] and use mapping class group arguments to establish the symplectic case.

The bound of -5 on the self-intersection of the exceptional curves in Theorem 1.2 cannot generally be improved. Indeed, any singularity whose minimal resolution

contains a sphere of self-intersection -4 has at least two distinct Stein fillings, because a neighborhood of the (-4) sphere can be rationally blown down to produce another filling with smaller Euler characteristic; see Symington [61]. This corresponds to the fact that the singularity has at least two smoothing components if a (-4) sphere is present; see Kollár [30]. While our Theorem 1.1 shows there are unexpected fillings in many examples, we do not cover all examples which fail the hypotheses of Theorem 1.2; there are many cases where we cannot determine whether or not the link has unexpected fillings.

Theorem 1.2 extends the list of singularities with no unexpected Stein fillings. However, when complexity of the singularity increases, one should expect the unexpected: as predicted in Némethi [43], more complicated singularities are likely to have Stein fillings that do not arise from Milnor fibers. To our knowledge, the only previous examples of unexpected Stein fillings in the literature are detected by their first Betti number. By Greuel and Steenbrink [22], Milnor fibers for normal surface singularities always have $b_1 = 0$. An infinite family of Stein fillings with $b_1 \neq 0$ was given in Akhmedov and Ozbagci [3; 4] for links of certain nonrational singularities; these links are Seifert fibered spaces over higher genus surfaces. It follows from [3; 4] that most of these fillings are different from both the Milnor fibers and the resolution of the singularity. The constructions in these papers use surgeries and produce infinite families of exotic fillings (which are all homeomorphic but pairwise nondiffeomorphic). Note that for rational singularities, the first Betti number cannot detect unexpected fillings: the link is a rational homology sphere, and a homology exact sequence argument shows that $b_1 = 0$ for any Stein filling; see Remark 6.5.

Note that, in general, known results allow us to find many nonrational singularities whose links have infinitely many Stein fillings. As an example, consider a normal surface singularity whose resolution has a unique exceptional curve of genus $g \geq 2$ with self-intersection $-d$, for $d > 0$. The resolution is the total space of the complex line bundle of degree d over the corresponding Riemann surface, and the singularity can be thought of as cone point. If $g = \frac{1}{2}(d-1)(d-2)$, one of the analytic singularities in this topological type is the hypersurface $(X_d, 0)$ in $\mathbb{C}^3$, given by $x^d + y^d + z^d = 0$. For each $d \geq 5$, the results of Baykur, Monden and Van Horn-Morris [7] produce arbitrarily long positive factorizations of the corresponding open book monodromy, which in turn yields infinitely many Stein fillings for the link (Y_d, ξ_d); in particular, there are Stein fillings with arbitrarily large b_2. One might hope that most of these Stein fillings are unexpected: indeed, a hypersurface singularity has a unique Milnor fiber, and its topology is well understood; see Milnor [39] and Tyurina [64]. However,

the question is more subtle: because $(X_d, 0)$ is not (*pseudo*)*taut* (see Laufer [32]), there are infinitely many singularities with the same link (Y_d, ξ_d). Milnor fibers of these singularities may yield additional Stein fillings. Describing all such Milnor fibers seems to be out of reach; conceivably, they may produce all the Stein fillings given by the arbitrarily long factorizations of [7]. We will discuss related questions in more detail in Section 4, although we do not have any answers for this case.

Our present work gives the first examples of unexpected Stein fillings for rational singularities, and for the case where the link Y is a rational homology sphere. In the case of rational singularities, the fillings must be differentiated from Milnor fibers by more subtle means than b_1, as all Stein fillings have $b_1 = 0$ in this case. For singularities with reduced fundamental cycle, the contact link admits a planar open book decomposition; see Némethi and Tosun [46] and Schönenberger [58]. By Kaloti [28], Plamenevskaya [54] and Stipsicz [60], it follows that the number of Dehn twists in any positive monodromy factorization, and thus b_2 of Stein fillings, is bounded above. This means that we cannot generate unexpected fillings by arbitrarily long positive factorizations. On the other hand, even though there is typically an infinite collection of singularities with the given link, the reduced fundamental cycle hypothesis, together with the de Jong–van Straten theory, gives us certain control over the topology of all possible Milnor fibers.

In general, comparing Stein fillings to Milnor fillings is a two-fold challenge: classification is typically out of reach, both on the deformation theory side (smoothings and Milnor fibers) and on the symplectic side (Stein fillings). In the particular case of rational singularities with reduced fundamental cycle, two important tools facilitate the study of fillings. On the algebraic geometry side, de Jong and van Straten reduce the study of deformations of the surface to certain deformations of a decorated germ of a reducible singular complex curve $\mathcal{C} \subset \mathbb{C}^2$. (The germ $\mathcal{C}$ is associated to the surface as explained in Section 2. For now, we omit the decoration from notation.) The construction of [27] works for a more general class of *sandwiched* rational singularities; in the case of reduced fundamental cycle, the associated plane curve germ has smooth irreducible components. Thus in this case, $\mathcal{C}$ is simply the union of smooth complex disks $C_1, C_2, \dots, C_m$, all passing through 0. The decoration of the germ is given by marked points, initially concentrated at the origin. To encode deformations of the surface singularity, one considers 1–parameter δ–constant deformations of $\mathcal{C}$, where the marked points are redistributed so that all singularities of the deformed curve $\mathcal{C}^s$ are marked (additional "free" marked points are also allowed). Smoothings of the corresponding singularities are given by *picture deformations*, where the only singularities of the

deformed curve are transverse multiple points. While picture deformations are still hard to classify directly and thus rarely give explicit classification of smoothings, they do provide a lot of useful information. In certain examples, they allow us to understand the topology of Milnor fibers and compute their basic invariants.

The following theorem summarizes the results of de Jong–van Straten that we use. Detailed definitions and precise statements will be given in Section 2.

Theorem 1.3 [27, Theorem 4.4, Lemma 4.7] *Let $(X,0)$ be a rational singularity with reduced fundamental cycle, and $\mathcal{C}\subset\mathbb{C}^2$ its decorated germ of a reducible complex curve such that all the branches $C_1,\dots,C_m$ of $\mathcal{C}$ are smooth complex disks. Then smoothings of $(X,0)$ are in one-to-one correspondence with **picture deformations** of $\mathcal{C}$. A picture deformation gives an arrangement $\mathcal{C}^s$ of the deformed branches $C_1^s,\dots,C_m^s$, $s\neq 0$, with marked points that include all the intersections of the branches. The Milnor fiber $W_{\mathcal{C}^s}$ of the corresponding smoothing can be constructed by blowing up at all marked points and taking the complement of the proper transforms of $C_1^s,\dots,C_m^s$.*

The Milnor fibers described in Theorem 1.3 are noncompact, but a slight modification yields compact Milnor fillings of the contact link (Y,ξ) of $(X,0)$. We consider the germ $\mathcal{C}$ in a small closed ball $B\subset\mathbb{C}^2$ centered at 0, such that all the branches of $\mathcal{C}$, and thus all the deformed branches for small s, intersect ∂B transversely, and B contains all marked points. To obtain a smooth compact 4–manifold whose boundary is the link Y, we blow up B at the marked points, take the complement of disjoint tubular neighborhoods of the proper transforms of $C_1^s,\dots,C_m^s$, and smooth the corners.

In turn, on the symplectic side, contact links of singularities with reduced fundamental cycle are more accessible because they are supported by planar open books; see Némethi and Tosun [46] and Schönenberger [58]. By a theorem of Wendl [65], all Stein fillings of a planar contact manifold are given by Lefschetz fibrations whose fiber is the page of the open book. In other words, all these Lefschetz fibrations arise from factorizations of the monodromy of the *given* open book into a product of positive Dehn twists. In most cases, such positive factorizations cannot be explicitly classified, but they give a combinatorial approach to Stein fillings.

To relate the two sides of the story, we generalize the notion of picture deformation and consider *smooth graphical homotopies* of the decorated germ $\mathcal{C}$ with smooth branches. A smooth graphical homotopy of $\mathcal{C}$ is a real 1–parameter family of embedded disks $C_1^t,\dots,C_m^t$ such that for $t=0$ the disks $C_1^0,\dots,C_m^0$ are the branches of $\mathcal{C}$, and

for $t = 1$, the intersections between C_i^1 and C_j^1 are transverse and positive for all i, j. There is a collection of marked points on $C_1^1, \dots, C_m^1$, coming from a redistribution of the decoration on $\mathcal{C}$, such that all intersection points are marked. (See Definition 3.1.)

We prove that just as picture deformations yield smoothings in [27], every smooth graphical homotopy gives rise to a Stein filling naturally supported by a Lefschetz fibration.

Theorem 1.4 *Let (Y, ξ) be the contact link of a singularity $(X, 0)$ with reduced fundamental cycle, and let $\mathcal{C}$ be a decorated plane curve germ representing $(X, 0)$, with m smooth components $C_1^0, \dots, C_m^0$. For any smooth graphical homotopy, let W be the smooth 4–manifold obtained by blowing up at all marked points and taking the complement of the proper transforms of $C_1^1, \dots, C_m^1$. (In the case of a picture deformation $\mathcal{C}^s$, W is the Milnor fiber $W_{\mathcal{C}^s}$ from Theorem 1.3).*

Then W carries a planar Lefschetz fibration that supports a Stein filling of (Y, ξ). When $W = W_{\mathcal{C}^s}$, the Lefschetz fibration is compatible with the Stein structure on the Milnor fiber.

The fiber of the Lefschetz fibration on W is a disk with m holes, and the vanishing cycles can be computed directly from the decorated curve configuration $C_1^1, \dots, C_m^1$. On (Y, ξ), the Lefschetz fibration induces a planar open book decomposition, which is independent of the smooth graphical homotopy of the given decorated germ $\mathcal{C}$.

Each rational singularity with reduced fundamental cycle has a distinguished *Artin* smoothing component, which corresponds to a picture deformation called the *Scott deformation*; see Section 4. Applying Theorem 1.4 to the Scott deformation yields a planar Lefschetz fibration filling (Y, ξ) where the vanishing cycles are disjoint; see Proposition 4.1. This gives a natural model for the planar open book decomposition on (Y, ξ). This open book is closely related to the braid monodromy of the singularity of $\mathcal{C}$. Note that we need to consider all singularities topologically equivalent to $(X, 0)$ to describe all Milnor fillings for (Y, ξ), since all such singularities have the same contact link. However, topologically equivalent singularities can be represented by topologically equivalent decorated germs and produce the same open book decompositions.

The process of computing the monodromy factorization resembles a known strategy for monodromy calculation for a plane algebraic curve; see Moishezon and Teicher [40; 41]. The necessary information can be encoded by a *braided wiring diagram* given by the intersection of $\mathcal{C}^s$ with a suitably chosen copy of $\mathbb{C} \times \mathbb{R} \subset \mathbb{C}^2$.

A reversal of the above constructions allows us to represent Stein fillings of (Y,ξ) via arrangements of symplectic curves, as follows. Let W be an arbitrary Stein filling of the link (Y,ξ). We fix an open book for (Y,ξ) defined by the germ $\mathcal{C}$ as above. By Wendl's theorem, W can be represented by a Lefschetz fibration with the planar fiber given by the page. The Lefschetz fibration corresponds to a factorization of the open book monodromy into a product of positive Dehn twists. We reverse-engineer a braided wiring diagram producing this factorization, and then use the diagram to construct an arrangement Γ of symplectic disks. (In fact, an arrangement of smooth graphical disks is sufficient for our constructions, but the symplectic condition can be satisfied at no extra cost.) We require that the disks intersect transversally (multiple intersections are allowed), and equip Γ with a collection of marked points that include all intersections and possibly additional "free" points. We also show that the resulting arrangement of disks and points is related to the decorated germ $\mathcal{C}$ by a smooth homotopy, which is graphical in suitable coordinates. (The homotopy moves the disks and the marked points.) This yields a symplectic analog of Theorem 1.3.

Theorem 1.5 *Let (Y,ξ) be the contact link of a singularity $(X,0)$ with reduced fundamental cycle that corresponds to a decorated plane curve germ $\mathcal{C}$. Then any Stein filling of (Y,ξ) arises from an arrangement Γ of symplectic graphical disks with marked points, as in Theorem 1.4. The arrangement Γ is related to the decorated germ $\mathcal{C}$ by a smooth graphical homotopy.*

Theorems 1.3 and 1.5 mean that both Milnor fibers and arbitrary Stein fillings of a given link of rational singularity with reduced fundamental cycle can be constructed in a similar way, starting with the decorated plane curve germ $\mathcal{C}$ representing the singularity. Milnor fibers arise from algebraic picture deformations of the branches of $\mathcal{C}$, while Stein fillings come from smooth graphical homotopies of the branches.

Once the comparison of Milnor fibers and Stein fillings is reduced to comparison of arrangements of complex curves or smooth disks with certain properties, we can construct examples of arrangements that generate Stein fillings not diffeomorphic to Milnor fibers. We need arrangements that are related to a particular plane curve germ by a smooth graphical homotopy but not by an algebraic picture deformation. We build *unexpected line arrangements* satisfying this property in Section 7, using classical projective geometry and a study of analytic deformations. We use these to construct unexpected Stein fillings; then we verify that they are not diffeomorphic (relative to the boundary open book data) to Milnor fillings by an argument based on Némethi

and Popescu-Pampu [44]. This leads to the proof of Theorem 1.1 and other similar examples.

At first glance, the difference between algebraic and smooth plane curve arrangements seems rather obvious. However, because we are in an open situation, working with germs of curves and smooth disks with boundary as opposed to closed algebraic surfaces, the question is quite subtle. In particular, we cannot simply use known examples of topological or symplectic line arrangements in $\mathbb{C}\mathrm{P}^2$ not realizable by complex lines. Indeed, in many cases the smooth surfaces can be closely approximated by high-degree polynomials, so that a Lefschetz fibration on the corresponding Stein filling can be realized by a Milnor fiber. We discuss the relevant features of the picture deformations and smooth (or symplectic) graphical homotopies in detail in Section 8, and explain what makes our examples work.

It is worth stating that while Stein fillings and Milnor fillings are the same for certain small families of singularities, the two notions are in fact fundamentally different. A Milnor filling is given by a smoothing of a singular complex surface, so there is a family of Stein homotopic fillings of (Y,ξ) that degenerate to the singular surface. A Stein filling of the link has no *a priori* relation to the singular surface and is not part of any such family. This distinction becomes apparent in our present work, by the following heuristic reasoning. A picture deformation $\mathcal{C}^s$ of the decorated germ $\mathcal{C}$ gives, for any $s \neq 0$, a Milnor filling $W_{\mathcal{C}^s}$, so that all these fillings are diffeomorphic and even Stein homotopic. The Milnor fillings look the same for all $s \neq 0$ because the arrangements of deformed branches $\{C_1^s, \dots, C_m^s\}$ have the same topology. By contrast, if the germ $\mathcal{C}$ is homotoped via a family of smooth disk arrangements Γ^t, the topology of the arrangement $\{\Gamma_1^t, \dots, \Gamma_m^t\}$ may change during the homotopy. Under certain conditions we can construct a family of Lefschetz fibrations W_t that includes the given Stein filling and changes its diffeomorphism type at finitely many discrete times as it connects to the minimal resolution. In other cases, at some time t the homotopy gives an arrangement Γ^t which produces an *achiral* Lefschetz fibration, so the 4–manifolds in the corresponding family do not necessarily carry a Stein structure. We return to this discussion in Section 8.

One can also ask whether unexpected fillings exist for rational singularities with reduced fundamental cycle that are not covered by Theorem 1.1 or Theorem 1.2. For certain additional simple examples, we can use Theorem 1.5 and pseudoholomorphic curve arguments to verify that there are no unexpected fillings, even though the smoothing

may not be unique. This approach only works when the germ of the singularity is a pencil of lines satisfying certain restrictive constraints. Namely, we can consider

(1) arrangements of 6 or fewer symplectic lines, or

(2) arrangements of symplectic lines where one of the lines has at most two marked points where it meets all the other lines in the arrangement.

Since the boundary behavior of symplectic lines is controlled, we can cap off symplectic lines in a ball to symplectic projective lines in $\mathbb{C}\mathrm{P}^2$, together with the line at infinity. The corresponding arrangements in $\mathbb{C}\mathrm{P}^2$ are shown to have a unique symplectic isotopy class and are symplectically isotopic to an actual complex algebraic line arrangement in $\mathbb{C}\mathrm{P}^2$; see Starkston [59, Lemma 3.4.5]. It follows that every symplectic arrangement as above can be obtained as picture deformation of a pencil of complex lines, and therefore, the corresponding Stein fillings are given by Milnor fibers. The links of the corresponding singularities are Seifert fibered spaces, for which Stein fillings were completely classified and presented as planar Lefschetz fibrations in [59, Chapter 4]. The line arrangements appearing in that classification precisely coincide with the symplectic disk arrangements from the perspective of this article. (Here, gluing in the deleted neighborhood of the disk provides an embedding of the Stein filling into a blow-up of $\mathbb{C}^2$. In [59], gluing on the cap, which augments the configuration of lines by the additional line at infinity, provides an embedding of the Stein filling in a blow-up of $\mathbb{C}\mathrm{P}^2$.) In general, Theorem 1.5 seems to have limited applications to classification of fillings, due to complexity of arrangements of curves.

It is interesting to note that while de Jong and van Straten describe deformations of sandwiched singularities, our constructions only work for the subclass of rational singularities with reduced fundamental cycle. Indeed, a planar open book decomposition of the contact link plays a key role in our work because we need Wendl's theorem to describe Stein fillings. By Ghiggini, Golla and Plamenevskaya [19] the Milnor fillable contact structure on the link of a normal surface singularity is planar *only if* the singularity is rational and has reduced fundamental cycle. This means that our methods in the present paper cannot be used for classification for any other surface singularities. However, for future work, we are investigating extensions of these methods to produce examples of unexpected fillings for more general surface singularities. Finally, recall that all weak symplectic fillings of a planar contact 3–manifold are in fact given by planar Lefschetz fibrations, up to blow-ups and symplectic deformation; see Niederkrüger and Wendl [48]. It follows that Theorem 1.5 and related results apply to describe all

minimal weak symplectic fillings. However, we focus on Stein fillings and will give all statements, with the exception of Theorem 1.2, only for the Stein case.

Organization of the paper

In Section 2 we review the definitions of rational singularities with reduced fundamental cycle as well as their deformation theory from [27], and prove some of their properties from the topological perspective. In Section 3 we prove the first direction of the symplectic correspondence, namely Theorem 1.4. In Section 4 we explain the smoothing in the Artin component from the perspective of symplectic topology, discuss the corresponding open books, and also raise some questions related to open book factorizations and nonrational singularities. In Section 5 we prove the other half of the correspondence, establishing Theorem 1.5 using braided wiring diagrams and Wendl's theorem [65]. In Section 6 we prove Theorem 1.2 and explain how to calculate algebraic topological invariants of the fillings, which we will use to distinguish our examples of unexpected Stein fillings from Milnor fillings. In Section 7 we prove that there are many examples of unexpected Stein fillings for links of rational surface singularities with reduced fundamental cycle, establishing Theorem 1.1. Finally, in Section 8 we explain what key differences between picture deformations and smooth graphical homotopies contributed to the distinction between expected and unexpected Stein fillings.

Acknowledgements We are grateful to Stepan Orevkov for suggesting Example 7.14 to us. This example played a crucial role in our understanding of arrangements that produce unexpected fillings. We thank Roger Casals, Eugene Gorsky and Marco Golla for their interest in this project and numerous motivating and illuminating discussions in the early stages of this work. In particular, Eugene helped us understand some of the results of [27]. We are grateful to Eugene and Marco for their comments on the preliminary version of this article, and to Jonathan Wahl, Jeremy Van Horn-Morris and Patrick Popescu-Pampu for interesting correspondence. Many thanks to İnanc Baykur for illuminating correspondence and discussion on the higher genus case. Starkston would also like to thank Sari Ogami, who learned and explained to her a great deal about the monodromy of braided wiring diagrams. Plamenevskaya is also grateful to John Etnyre, Jonny Evans, Mark McLean and Oleg Viro for a few helpful discussions. We thank the referees for their thoughtful comments and suggestions. Plamenevskaya was supported by NSF grants DMS-1510091 and DMS-1906260 and a Simons Fellowship. Starkston was supported by NSF grant DMS-1904074.

2 Rational singularities with reduced fundamental cycle, their decorated curve germs, and relation to deformations

In this section, we collect some facts about rational singularities with reduced fundamental cycle and state de Jong and van Straten's results on their smoothings [27]. De Jong and van Straten's results are in fact more general: they fully describe deformation theory for a wider class of *sandwiched* singularities. We state only the results we need. Some of our statements are slightly different from [27]: we describe their constructions from the topological perspective and set the stage for our work. Although we aim for a mostly self-contained discussion, the reader may find it useful to consult [42] for a general survey on topology of surface singularities. The survey [56] focuses on the interplay between singularity theory and contact topology and provides very helpful background. Additionally, a brief survey of the key results of [27] from the topological perspective can be found in [44].

2.1 Resolutions and smoothings.

We begin with some general facts about surface singularities. Let $(X,0)$ be a normal surface singularity. Its resolution $\pi\colon \widetilde{X} \to X$ is a proper birational morphism such that $\widetilde{X}$ is smooth. The *exceptional divisor* $\pi^{-1}(0)$ is the inverse image of the singular point. For a given singularity $(X,0)$, the resolution is not unique, as one can always make additional blow-ups; however, for a surface singularity, there is a unique *minimal* resolution [31]. The minimal resolution is characterized by the fact that $\widetilde{X}$ contains no embedded smooth complex curves of genus 0 and self-intersection -1 (thus it does not admit a blow-down).

After performing additional blow-ups if necessary, we can assume that the exceptional divisor $\pi^{-1}(0)$ has normal crossings. This means that $\pi^{-1}(0) = \bigcup_{v\in G} E_v$, where the irreducible components E_v are smooth complex curves that intersect transversally at double points only. A resolution with this property is called a *good* resolution. For a surface singularity, a minimal good resolution is also unique [31].

The topology of a good resolution is encoded by the (dual) resolution graph G. The vertices $v \in G$ correspond to the exceptional curves E_v and are weighted by the genus and self-intersection $E_v \cdot E_v$ of the corresponding curve. We will often refer to $E_v \cdot E_v$ as the self-intersection of the vertex v, and use the notation $v \cdot v$ for brevity. The edges of G record intersections of different irreducible components. Note that the link of the

singularity is the boundary of the plumbing of disk bundles over surfaces according to G. In this paper, we focus on *rational* singularities; in this case G is always a tree, and each exceptional curve E_v has genus 0. (Genus 0 curves are also called *rational curves*.) Therefore we will typically omit the genus from the markings on the vertices and only record the self-intersection numbers.

It is well known that the dual resolution graph of every normal surface singularity is negative definite, and conversely, every negative definite connected graph corresponds to some normal surface singularity; see eg [42]. The link of the singularity determines the dual graph of the minimal good resolution, and vice versa. By a result of W Neumann [47], the links of two normal surface singularities have the same oriented diffeomorphism type if and only if their dual resolution graphs are related by a finite sequence of blow-ups/blow-downs along rational (-1) curves. Moreover, the links of two normal surface singularities are orientation-preserving diffeomorphic if and only if their minimal good resolutions have the same dual graphs. Minimal good resolutions are easy to recognize: if a good resolution is not minimal, its graph will have a vertex representing a genus 0 curve with self-intersection -1. (This follows from [47]; see also [19, Lemma 5.2] for a direct proof that any possible blow-downs can be seen directly from the graph.)

The local topological type of the singularity $(X,0)$ can be understood from its link Y, as a cone on the corresponding 3–manifold. We will say that two singularities are *topologically equivalent* if they have the same link. It is important to note that the analytic type of the singularity is not uniquely determined by the link; typically, many analytically different singularities have diffeomorphic links. It is known that the canonical contact structures are all isomorphic for different singularities of the same topological type [11]; thus, the dual resolution graph encodes the canonical contact structure. Indeed, this contact structure can be recovered as the convex boundary of the plumbing, according to the graph, of the standard neighborhoods of the corresponding symplectic surfaces.

We now turn our attention to deformations and Milnor fibers. A deformation of a surface singularity $(X,0)$ is any flat map $\lambda\colon(\mathcal{X},0)\to(\mathcal{T},0)$ such that $\lambda^{-1}(0)=(X,0)$. A versal (or semiuniversal) deformation $f\colon(\mathcal{X},0)\to(B,0)$ parametrizes all possible deformations of $(X,0)$. The base space $(B,0)$ generally has multiple irreducible components, which may have different dimensions. It is generally difficult to understand the space B, its irreducible components, and the dimensions of these components.

A deformation $\lambda:(\mathcal{X},0)\to(D,0)$ over the disk in $\mathbb{C}$ is called a (1–parameter) *smoothing* of $(X,0)$ if $X_s:=\lambda^{-1}(s)$ is smooth for all $s\neq 0$. For any smoothing all such X_s are diffeomorphic, and we call X_s the *Milnor fiber* of the smoothing. For example, for a hypersurface $X=\{f(x,y,z)=0\}\subset\mathbb{C}^3$ with $f(0)=0$ and $df(0)=0$, a smoothing of the singularity at 0 can be given by $f:\mathbb{C}^3\to\mathbb{C}$, with Milnor fiber $X_\epsilon=\{f(x,y,z)=\epsilon\}$ for a small $\epsilon\neq 0$. Each Milnor fiber is endowed with a Stein structure, and for different $t_0,t_1\in D\setminus 0$, X_{s_0} and X_{s_1} are Stein homotopic (the Stein homotopy is obtained by choosing a path from s_0 to s_1 in D which avoids 0).

We need to work with a compact version of the Milnor fiber. For a sufficiently small radius $r>0$, the surface $X\subset\mathbb{C}^N$ is transverse to the sphere S_r^{N-1}. We fix a ball $B_r^N\subset\mathbb{C}^N$ centered at 0, sometimes called *a Milnor ball*, and consider $X\cap B_r^N$ as the *Milnor representative* of X. The boundary $\partial(X\cap B_r^N)$ is the link Y of $(X,0)$, and the complex structure on X induces the canonical contact structure ξ on Y. For sufficiently small $s\neq 0$, we can similarly find a compact version of X_s whose boundary is contactomorphic to the link (Y,ξ), which provides a Stein filling of (Y,ξ).

For a semiuniversal deformation $f:(\mathcal{X},0)\to(B,0)$ of the surface singularity $(X,0)$, an irreducible component B_i of B is called a *smoothing component* of $(X,0)$ if the general fiber over B_i is smooth. We note that B_i may have lower (complex) dimensional strata where the fibers over these strata are not smooth. For example, these nongeneral strata could arise from singularities in the component B_i or intersections of B_i with other irreducible components of B. Nevertheless, these nongeneral strata have positive complex codimension, so the subset of B_i over which the fiber is smooth will be connected. Any 1–parameter smoothing of $(X,0)$ lies in a unique smoothing component B_i.

In general, not every surface singularity admits a smoothing. However, for rational singularities every irreducible component of B is a smoothing component; see [5] and also [56, Theorem 4.24]. Moreover, there is one distinguished component, called the *Artin component*. This component is associated to the minimal resolution $\widetilde{X}$ of $(X,0)$; see [5] and also [56, Theorem 4.25]. (For rational singularities, deformations of $\widetilde{X}$ come from deformations of $(X,0)$, and these deformations of $(X,0)$ form the Artin component.) We discuss Milnor fibers in this component in greater detail in Section 4.

In this paper, we study Stein fillings for the contact link (Y,ξ) of a surface singularity, and compare them to Milnor fillings. As explained above, in general the link determines only the topological, but not the analytic, type of the singularity. Normal surface singularities whose topological type admits a unique analytic type are called *taut*; if there are only finitely many analytic types, the singularity is *pseudotaut*. Taut and

pseudotaut singularities were classified by Laufer [32]: there are several very restrictive lists for the dual resolution graphs, in particular, the graphs cannot have any vertices of valency greater than 3. Thus, most singularities are not (pseudo)taut, even if we restrict to a very special kind that we consider in this paper, rational singularities with reduced fundamental cycle. If we are to compare Stein fillings and Milnor fillings of the link, we need to consider Milnor fibers for *all* possible singularities of the given topological type. In principle, it is quite possible that topologically equivalent singularities have nondiffeomorphic Milnor fibers: for example, the hypersurface singularities $x^2+y^7+z^{14}=0$ and $x^3+y^4+z^{12}=0$ have the same topological type, but their (unique) Milnor fibers have different b_2; see [33] and also the discussion in [56, Section 6.2]. Fortunately, in the case of reduced fundamental cycle we will have some control over the topology of Milnor fibers for different analytic types, thanks to the de Jong–van Straten construction.

2.2 Sandwiched singularities, extended graphs, and decorated germs

Definition 2.1 $(X,0)$ is a *rational singularity with reduced fundamental cycle* if it admits a normal crossing resolution such that all exceptional curves have genus 0, the dual resolution graph G is a tree, and for each vertex $v\in G$, the valency $a(v)$ of v and the self-intersection $v\cdot v$ satisfy the inequality

$$a(v)\le -v\cdot v. \tag{2-1}$$

It follows from (2-1) that the graph as above can only have vertices with self-intersection -1 as the leaves of the tree. Blowing down all such vertices, we obtain a graph that still satisfies (2-1) and represents the minimal resolution of $(X,0)$.

To explain the terminology of Definition 2.1, we recall the definition of a fundamental cycle. For a given resolution, consider the set of divisors

$$\left\{ Z=\sum_{v\in G} m_v E_v \;\middle|\; Z>0, \text{ and } Z\cdot E_v\le 0 \text{ for all } E_v \right\}.$$

This set has a partial order, defined by $\sum m_vE_v\ge\sum n_vE_v$ if $m_v\ge n_v$ for all v. There is a minimal element with respect to this partial order, denoted by $Z_{\min}$ and called Artin's fundamental cycle. The resolution dual graph is connected, different components E_v intersect positively, and $Z>0$, so any element in the set has $m_v>0$. Therefore, $Z_{\min}\ge\sum_{v\in G}E_v$. It is easy to see that $\big(\sum_{v\in G}E_v\big)\cdot E_v\le 0$ for all E_v if and only if condition (2-1) is satisfied. In this case $Z_{\min}=\sum_{v\in G}E_v$, and since each exceptional curve enters with multiplicity 1, we say that the fundamental cycle $Z_{\min}$ is reduced.

In [27], de Jong and van Straten work with *sandwiched* singularities. By definition, a sandwiched singularity $(X,0)$ is analytically isomorphic to the germ of an algebraic singular surface which admits a birational morphism to $(\mathbb{C}^2,0)$. For a resolution $\pi\colon \widetilde{X}\to X$, we get a diagram $(\widetilde{X},\pi^{-1}(0)) \dashrightarrow (X,0) \dashrightarrow (\mathbb{C}^2,0)$. In particular, X is sandwiched between two smooth spaces via birational maps. Sandwiched singularities are rational and can be characterized by their resolution graphs as follows, by translating the sandwiched condition. The graph G is *sandwiched* if we can add to it a number of edges and their end vertices with self-intersections (-1), so that the resulting graph G' gives a plumbing whose boundary represents S^3. In other words, G' gives a configuration of rational curves that can be blown down to a smooth point. The choice of the graph G' is not unique. It is not hard to see that every rational singularity with reduced fundamental cycle is sandwiched. In Proposition 2.2 below, we discuss in detail the construction of the possible graphs G' for this case.

Any sandwiched singularity can be associated to a (germ of a) complex plane curve singularity, constructed as follows. The choice of the graph G' corresponds to an embedding of the tubular neighborhood of the exceptional set of the resolution $\widetilde{X}$ into some blow-up of $\mathbb{C}^2$. This blown-up surface also has a distinguished collection of (-1) curves, so that the configuration of these (-1) together with the exceptional set can be completely blown down. For each distinguished (-1) curve, choose a transverse complex disk (called a *curvetta*) through a generic point. Now, contract the curve configuration corresponding to G'. The union of the curvettas becomes a germ of a reducible curve $\mathcal{C}$ in $\mathbb{C}^2$, with components passing through 0. Let C_i, $i=1,2,\dots,m$, be the irreducible components of $\mathcal{C}$; following [27], we also refer to C_i as *curvettas*. We emphasize that only the germ of $\mathcal{C}$ at the origin is defined; when we use the notation $\mathcal{C}\subset\mathbb{C}^2$, we only consider a small neighborhood of $0\in\mathbb{C}^2$. In particular, we are only interested in the singularity of the reducible curve $\mathcal{C}$ at 0. In this paper, we will focus on the case where the components C_i are smooth at 0, so that locally C_i is a smooth disk. This suffices to study rational singularities with reduced fundamental cycle, as we will soon see. This disk may be locally parametrized by a high-degree algebraic curve in $\mathbb{C}^2$, but the global topology of this curve is unimportant to us, because we only use the part of the curve in a neighborhood of the origin.

Each curvetta C_i comes with a weight $w_i = w(C_i)$, given by the number of exceptional spheres that intersect the corresponding curve in the blow-down process from G' to the empty graph. In other words, w_i is the number of blow-down steps that affect the corresponding curvetta before it becomes C_i. The weighted curve $(\mathcal{C},w)$ is called a

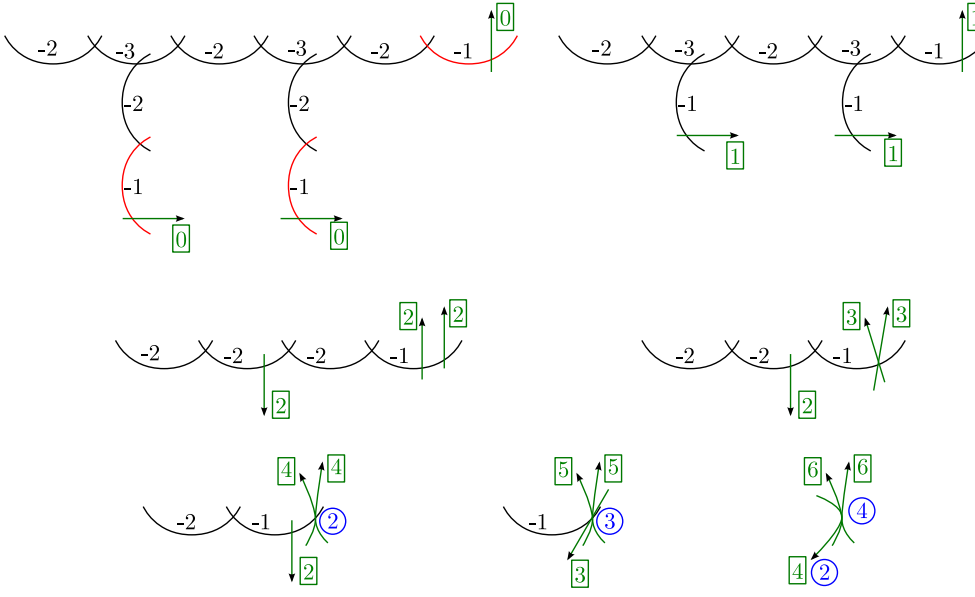


Figure 2: An example of a sandwiched singularity and a choice of corresponding curvettas (green arrows). The first diagram shows the resolution curves together with extra (red) (-1) exceptional curves attached. Then there is a sequence of blow-downs. We keep track of the weights $w(C_i)$ in rectangular boxes next to each green curvetta arrow. The multiplicities of tangencies between bunches are recorded in blue circled numbers.

decorated germ corresponding to $(X, 0)$. An example of this process, and the resulting decorated germ for the given singularity, is shown in Figure 2.

It is convenient to start the process with the minimal normal crossings resolution of $(X, 0)$. For rational singularities with reduced fundamental cycle, it is easy to see that the graph of the minimal normal crossings resolution has no (-1) vertices. (From (2-1), only vertices of valency 1 can have self-intersection -1 in any resolution graph, and these can be blown down to get the minimal graph.) If G has no (-1) vertices, then *all* the (-1) vertices of G' are those that come from the extension: each (-1) vertex is a leaf of G', connected by an edge to a unique vertex of G. The transverse curvetta slices are added to all these (-1) vertices.

In what follows, we will only consider decorated germs that arise from the above construction. (These are called *standard* decorated germs in [44]. Some statements in [27] allow for more general decorated germs.)

The singularity $(X,0)$ can be reconstructed from $(\mathcal{C},w)$. We iteratively blow up points infinitely near 0 on proper transforms of curvettas $C_1,\dots,C_k$ until we obtain a minimal embedded resolution of $\mathcal{C}$. Then we perform additional blow-ups at the intersection of C_i with the corresponding exceptional curve, so that the sum of multiplicities of proper transforms of C_i at the blow-up points is exactly w_i. The union of the exceptional curves that *do not* meet the proper transforms of the curvettas is then contracted to form $(X,0)$.

We emphasize that $\mathcal{C}$ depends on the choice of the graph G', ie on the particular extension of the resolution graph of $(X,0)$ by (-1) curves. Any of these choices can be used to classify Milnor fillings as in [27]. In general, the branches of $\mathcal{C}$ are singular curves. However, if $(X,0)$ is a rational singularity with reduced fundamental cycle, an appropriate choice of G' ensures that $\mathcal{C}$ has smooth branches. We will always work in this setting and only consider decorated germs with smooth components. In the following proposition, we establish a necessary and sufficient condition for smoothness purely in terms of the graph G'. Although similar questions were studied in [27; 26], we formulate the condition here in a way that seems simplest from the topological point of view. In the next section, we will reinterpret the statement for open book decompositions.

Proposition 2.2 *Let the graph G' be a negative definite plumbing tree, and P' the corresponding plumbing of disk bundles over rational curves. Suppose that the boundary of the plumbing P' is S^3; equivalently, G' encodes a configuration of rational curves that can be blown down to a smooth point. For each (-1) vertex, let $\widetilde{C}_j$ be a complex disk intersecting the corresponding (-1) sphere in P' transversally once. Let $C_1,\dots,C_m$ be the images of $\widetilde{C}_1,\dots,\widetilde{C}_m$ under blowing down the configuration G'. Then the following are equivalent*:

(1) *Each C_j is smooth.*

(2) *There exists exactly one $v_0'\in G'$ such that*

$$\begin{aligned} v_0'\cdot v_0' + a(v_0') &= -1,\\ v'\cdot v' + a(v') &= 0 \quad \text{for all } v'\neq v_0'. \end{aligned}$$

(We will often refer to v_0' as the *root*.)

As before, $v'\cdot v'$ denotes the self-intersection of a vertex $v'\in G'$, and $a(v')$ its valence.

Proof Consider $\mathcal{C} = C_1 \cup \cdots \cup C_m$ with smooth branches C_j. We obtain G' as described above, by blowing up repeatedly at intersections of the C_j with each other and with the exceptional divisors. We stop when the resulting configuration of curves has the following property: if an exceptional divisor intersects a proper transform $\widetilde{C}_j$ then it is disjoint from all other proper transforms $\widetilde{C}_{j'}$, $j' \neq j$ (in particular, different $\widetilde{C}_j$ are disjoint from each other), and the total number of blow-ups performed on (proper transforms of) C_j is exactly w_j, the weight on C_j.

We will show that G' has the structure of a rooted tree by repeatedly applying the following procedure. For the root v_0', we will have $v_0' \cdot v_0' + a(v_0') = -1$, and for all other vertices $v' \neq v_0'$, $v' \cdot v' + a(v') = 0$. We show that this condition is satisfied at every stage of the process.

Blow up at the common intersection point of all C_j. The resulting exceptional divisor (and its future proper transforms) gives the root of the tree. If proper transforms of all C_j still have a common point, we repeatedly blow up at the same point until some of the proper transforms $\widetilde{C}_j$ become disjoint from each other. (With a slight abuse of notation, $\widetilde{C}_j$ will denote the proper transform of C_j at any stage of the process.) Additional blow-ups create a chain of exceptional (-2) spheres with the root at one end and the most recent exceptional (-1) sphere at the other end. Up to relabeling, we can assume there are distinct intersection points $\widetilde{C}_1 \cap \cdots \cap \widetilde{C}_{a_1} = p_1^1$, $\widetilde{C}_{a_1+1} \cap \cdots \cap \widetilde{C}_{a_2} = p_2^1, \ldots, \widetilde{C}_{a_{r_1}^1} \cap \cdots \cap \widetilde{C}_m = p_{r_1}^1$ lying on the most recently introduced exceptional divisor B_1.

Assuming $m > 1$, since all the $\widetilde{C}_j$ intersect B_1, we must blow up exactly once at each p_i^1 to make them all disjoint from B_1. Here we use smoothness of the curvettas C_j (and thus of their proper transforms) to ensure that they become disjoint from B_1 after a single blow-up: every point on C_j has multiplicity 1, thus $\widetilde{C}_j$ intersects each exceptional divisor with multiplicity at most 1. Note that once $\widetilde{C}_1, \ldots, \widetilde{C}_m$ are all disjoint from B_1, we will not blow up at any point on B_1 again, therefore at this stage we can already compute the self-intersection and valency of the corresponding vertex in G'. The self-intersection of the proper transform of B_1 in G' (which we will also denote by B_1) is $-r_1 - 1$. If B_1 is not the root, it has valency $r_1 + 1$, and if it is the root it has valency r_1. Thus, condition (2) is satisfied for the vertex of G' given by B_1. All the other vertices in the graph at this stage are either (-2) spheres in a chain of valency 2 (if not the root) or valency 1 (if the root), or newly introduced (-1) vertices of valency 1, so condition (2) is satisfied at this stage.

In order to obtain G' we repeat this process iteratively, replacing the first exceptional sphere with the exceptional sphere obtained by blowing up at some p_i^s. (The points $p_1^1, \dots, p_{r_1}^1$ were introduced above; after blowing up at each of these new points, the new exceptional curves intersect the proper transforms of the curvettas at points $p_1^2, \dots, p_{r_2}^2$; similarly, points $p_i^s, \dots, p_{r_s}^s$ are the intersections that appear at step s.) Each time, condition (2) is preserved, since each curve $\widetilde{C}_j$ intersects each exceptional divisor with multiplicity at most 1. Repeating sufficiently many times, eventually all of the $\widetilde{C}_j$ will intersect only disjoint exceptional spheres. After potentially blowing up more times at the intersection of $\widetilde{C}_j$ with its intersecting exceptional sphere until the number of blow-ups is w_j, we obtain G'. (The additional blow-ups create a chain of (-2) vertices connecting to the last (-1) vertex.) Since condition (2) is preserved at each step of this procedure, G' satisfies condition (2).

Conversely, if G' satisfies condition (2), the only (-1) vertices are leaves of the rooted tree (valency 1). Blowing down a leaf preserves condition (2) because it decreases the valency of the adjacent vertex by 1 and increases the self-intersection by 1. The $\widetilde{C}_j$ are disks which transversally intersect the (-1) leaves of G' with multiplicity 1. Therefore each $\widetilde{C}_j$ intersects each exceptional divisor with multiplicity at most 1. This property is preserved under blowing down a (-1) leaf, because a multiplicity 1 intersection of $\widetilde{C}_j$ on a (-1) leaf becomes a multiplicity 1 intersection on the adjacent exceptional divisor after blowing down. Blowing down an exceptional divisor which intersects $\widetilde{C}_j$ with multiplicity 1 preserves smoothness of $\widetilde{C}_j$. Therefore after blowing down all leaves of G' and finally the root, the resulting proper transforms C_j are still smooth. □

Remark 2.3 Another way to see that G' must satisfy condition (2) is to consider what happens if G' has a vertex with $a(v') > -v' \cdot v'$. After blowing down, eventually the vertex v' will correspond to a (-1) sphere with valency ≥ 2, with at least one $\widetilde{C}_j$ intersecting it with multiplicity at least 1. (The existence of the intersecting $\widetilde{C}_j$ comes from the fact that intersections are transferred under blow-down to the adjacent vertices. Initially, every (-1) sphere in G' has an intersecting curvetta. Each time that a (-1) sphere is blown down, the curvetta intersection is transferred to the adjacent vertices, whose self-intersections are correspondingly increased. For v' to reach self-intersection -1, one must have blown down (-1) vertices adjacent to it. Throughout the process of blowing down, we maintain the condition that (-1) vertices always have at least one intersecting curvetta.) After blowing down the (-1) sphere of valency ≥ 2, we obtain a point where at least two exceptional divisors intersect at the same point

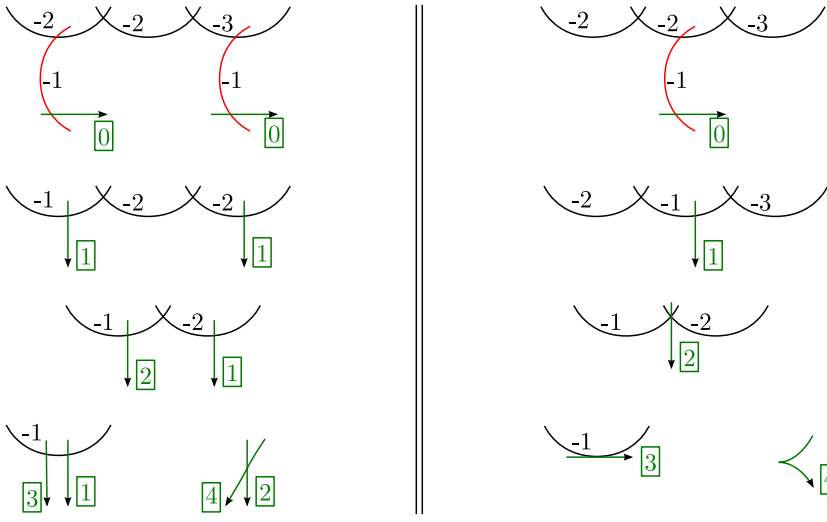


Figure 3: Two possible choices to add -1 curves to the same resolution graph, resulting in different curvettas, one with smooth components and another with a singular (cuspidal) component.

with a $\widetilde{C}_j$. Eventually one of these exceptional divisors will be blown down, forcing $\widetilde{C}_j$ to intersect the other exceptional divisor with multiplicity ≥ 2. Once this other exceptional divisor is blown down, the proper transform of $\widetilde{C}_j$ becomes singular.

Note that it is possible to have different choices of extension for G, such that one choice yields smooth curvettas and another yields singular curvettas; see Figure 3 for an example. In other words, some sandwiched resolution graphs G have extensions both to a graph which does satisfy condition (2) of Proposition 2.2 and to a graph which does not. For our classifications, we will always work with a choice of extension of G which does satisfy condition (2) and the corresponding smooth curvettas.

We can also deduce some basic numerical properties from Proposition 2.2. It turns out that for a rational singularity $(X, 0)$ with reduced fundamental cycle, the multiplicity of the singular point determines the number of choices for the defining plane curve germ $\mathcal{C}$ with smooth branches, as well as the number of curvetta branches in each such germ. Assuming that $(X, 0) \subset (\mathbb{C}^N, 0)$ for some large N, recall that the multiplicity mult X can be defined geometrically as the number of intersections $\#X \cap L$ of X with a generic complex $(N-2)$–dimensional affine subspace $L \subset \mathbb{C}$, passing close to the origin. For rational singularities, multiplicity is a topological invariant, which can be

computed from the resolution graph by the formula $\operatorname{mult} X = -Z_{\min}^2$; see eg [42]. The two statements below are also discussed in [27] from the algebrogeometric perspective, but they follow easily from the combinatorics of the resolution graph.

Proposition 2.4 *Let $(X,0)$ be a rational singularity with reduced fundamental cycle, and $\mathcal{C}$ a plane curve germ corresponding to $(X,0)$. If $\mathcal{C}$ has smooth branches, the number of branches is given by* $\operatorname{mult} X - 1$.

Proof The minimal normal crossings resolution graph G for $(X,0)$ has no (-1) vertices. Then G is obtained from (any choice of) the graph G' by deleting all vertices $v' \in G'$ with $v' \cdot v' = -1$. The curvetta branches are obtained by putting transverse slices on each (-1) sphere $v' \in G'$, thus the number m of curvetta branches is given by the number of the (-1) vertices in G'. By condition (2) of Proposition 2.2,

$$\sum_{v' \in G'} (v' \cdot v' + a(v')) = -1.$$

Again by condition (2), each (-1) vertex has valency 1 in G', so each addition of a (-1) vertex to G increases the sum $\sum_{v \in G}(v \cdot v + a(v))$ by 1, thus we have

$$\sum_{v \in G}(v \cdot v + a(v)) = \sum_{v' \in G'} (v' \cdot v' + a(v')) - m = -1 - m.$$

Finally, we relate this quantity to the fundamental cycle $Z_{\min}$, which is the sum of homology classes of the exceptional divisors, $Z_{\min} = \sum_{v \in G} E_v$:

$$\sum_{v \in G}(v \cdot v + a(v)) = \sum_{v} E_v^2 + \sum_{v \neq u} E_v \cdot E_u = Z_{\min}^2.$$

So $m = -1 - Z_{\min}^2 = \operatorname{mult} X - 1$. □

Decorated germs representing a given $(X,0)$ are obtained from extensions G' of the resolution graph G as above. These can be thought of as combinatorial choices for the decorated germ; in the next lemma, we compute the number of such extensions. Then, we show that the combinatorial choice, namely the choice of vertices of G on which the additional (-1) vertices are placed to form G', determines the *topological type* of the resulting decorated germ. By definition, the topological type of a germ of a singular curve $\mathcal{C} \subset \mathbb{C}^2$ is given by its link, which is the intersection of $\mathcal{C}$ with a sufficiently small 3–sphere $S^3 \subset \mathbb{C}^2$ centered at the origin. For a decorated germ, we additionally record the weights of the curvetta components. Later on, we will see that the different choices of G' correspond to natural different choices of data on the open book decomposition we construct in Section 4.

Lemma 2.5 *Up to topological equivalence, there are at most* mult X *choices of plane curve germs with smooth branches representing* $(X, 0)$.

Proof We first show that there are at most mult $X = -Z_{\min}^2$ possible *combinatorial choices* for germs with smooth components representing $(X, 0)$. These correspond to choices of extensions of G to G' by adding (-1) vertices. If we have a minimal graph G with an extension G' satisfying condition (2) of Proposition 2.2, then we can add another (-1) sphere leaf adjacent to the root to get a new graph G'' so that the valency of each vertex of G'' equals its negative self-intersection. All the other possible extensions of G to a graph satisfying condition (2) can be obtained by deleting one of the (-1) vertices of G''. (Indeed, adding a (-1) vertex to any other position in G would violate condition (2).) Since G' has (mult $X - 1$) vertices of self-intersection -1, we know that G'' has exactly mult X vertices with this property, one of which must be deleted. Note that because of potential symmetries in the graph G'', some of the choices of G' will result in isomorphic germs C, but mult X gives an upper bound on the number of combinatorially different curvetta configurations.

Once the choice of the extension G' of the graph G is made, the topological type of the decorated germ $\mathcal{C}$ can be read off directly from G'. In particular, we can compute the relevant numerical invariants, such as linking numbers between the components of $\mathcal{C} \subset \mathbb{C}^2$. As before, we assume that G' satisfies condition (2) of Proposition 2.2, so that $\mathcal{C}$ has smooth branches.

Following [27, Definition 4.14], we define the *length* and *overlap* functions on the vertices of the graph G. For $v_0, v_i \in G$, let the length $l(v_0, v_i)$ be the number of vertices in the path from v_i to v_0 in the tree G (including endpoints). For $v_0, v_i, v_j \in G$, let the overlap $\rho(v_i, v_j; v_0)$ be the number of common vertices in the paths from v_i to v_0 and v_j to v_0.

Let $v_0 \in G \subset G'$ be the root. Now, if the curvetta C_i comes from the transverse slice on a (-1) sphere corresponding to a leaf of G', and this leaf is attached to the vertex $v_i \in G$, then the blow-down process gives $w(C_i) = 1 + l(v_0, v_i)$. If C_i and C_j are the curvettas at the (-1) vertices attached to v_i and v_j, the order of tangency $\operatorname{tang}(C_i, C_j)$ between the corresponding branches of $\mathcal{C}$ is given by $\operatorname{tang}(C_i, C_j) = \rho(v_i, v_j; v_0)$.

The topological type of $\mathcal{C} \subset \mathbb{C}^2$ is described via its link, given by the intersection $\mathcal{C} \cap S^3$, where S^3 is a small sphere centered at the origin. As each of the curvettas $C_1, \dots, C_m$ is a smooth disk, the intersection of C_i with S^3 is an unknot; $\mathcal{C} \cap S^3$ is a link with m components $C_1 \cap S^3, \dots, C_m \cap S^3$, each of them unknotted. The components of

$C_i \cap S^3$ are oriented as boundaries of $C_i \cap B^4$. Then, the linking number between two link components equals the order of tangency between the corresponding curvettas,

$$\mathrm{lk}(C_i \cap S^3, C_j \cap S^3) = \mathrm{tang}(C_i, C_j).$$

The topological equivalence of germs follows from the above calculations, by construction of the links of the germs that we consider; it can also be seen more directly. Any decorated germ for $(X, 0)$ comes, after a blow-down, from a particular placement of the transverse curvetta slices on the (-1) curves corresponding to vertices that we added to G to from the graph G'. This gives a configuration of curvetta slices together with the curve configuration corresponding to the graph G', embedded in a blow-up of $\mathbb{C}^2$. Clearly, for two different choices of the generic curvetta slices for the same graph G', the two configurations of curvettas+curves can be identified by an ambient homeomorphism (in the blown-up $\mathbb{C}^2$). After the blow-down, the induced ambient homeomorphism will identify the links of the resulting germs, showing that the germs are topologically equivalent. We already know that the weights will be same, so the decorated germs have the same topological type. □

The following observation will also be useful later. Let $t(C_i) = \max_j \mathrm{tang}(C_i, C_j)$ be the maximal order or tangency between C_i and another branch of $\mathcal{C}$. Then it follows that

$$t(C_i) < w(C_i) \tag{2-2}$$

for all curvettas C_i.

Remark 2.6 De Jong and van Straten [27] study deformation theory of the surface singularity $(X, 0)$; in particular, they are interested in the analytic type of the singularity and its deformations. To encode the analytic type of $(X, 0)$, one needs the analytic type of the corresponding decorated germ $\mathcal{C}$. By contrast, our focus is on the contact link (Y, ξ) of $(X, 0)$ and its Stein fillings. A priori there may be another surface singularity $(X', 0)$ whose link is Y, and by [11], the singularities $(X, 0)$ and $(X', 0)$ have contactomorphic links. By Neumann's results [47], all singularities with the same link have the same dual graph of minimal resolution, so both $(X, 0)$ and $(X', 0)$ correspond to the same minimal graph G. (Note that by [32], if G has any vertices of valency greater than 3, the analytic type of the singularity is not uniquely determined, so indeed $(X, 0)$ and $(X', 0)$ may be analytically different in the above scenario.) We can compare the decorated germs that describe singularities $(X, 0)$ and $(X', 0)$: any choice of the decorated germ for $(X, 0)$ arises from an extension G' of the graph G and

the corresponding placement of the curvettas. Although analytically the exceptional divisors of resolutions of $(X,0)$ and $(X',0)$ may be different, topologically they look the same, and we can choose the same extension G' and the corresponding placement of curvettas for $(X',0)$. By the argument above, the resulting germ for $(X',0)$ will be topologically equivalent to the germ for $(X,0)$, even if the two germs may be analytically different. This fact will play an important role in the proof of Theorem 7.8. In particular, the two germs will have the same number of branches, the same weights and the same pairwise orders of tangency for the branches.

Of course, if we only know the combinatorics of the graph G, we lose analytic information on the plane curve germ $\mathcal{C}$ (such as, for example, the angles between its transverse branches), but we will never need the analytic information. The contact 3–manifold (Y,ξ) is fully determined by the weights and pairwise orders of tangency of the branches of the decorated germ $\mathcal{C}$.

2.3 De Jong–van Straten theory: Milnor fibers from germ deformations

The main result of [27] says that deformations of the sandwiched singularity can be encoded via deformations of the germ $(C,0)$ satisfying certain hypotheses. We will state a special case of their theorem that will be relevant to us, but first we introduce some notation.

We have defined the weights as positive integers w_i associated to the irreducible components (curvettas) C_i of $\mathcal{C}$. It will be convenient to interpret the weight w_i as a collection of w_i marked points concentrated at $0 \in C_i$. More formally, we consider a subscheme $w(i)$ of length w_i at 0 in C_i. The normalization $\widetilde{\mathcal{C}}$ of the reducible curve $\mathcal{C}$ with smooth components is given by the disjoint union of the components C_i; thus we can think of the decoration $w = (w_1, w_2, \dots, w_m)$ as a subscheme of $\widetilde{\mathcal{C}}$, with components $w(i) \subset C_i$ as above. (We use the notation $\widetilde{\mathcal{C}}$ for normalization here and in the discussion below. Similar notation $\widetilde{C}_j$ had a different meaning in Proposition 2.2, though in a sense, both uses refer to resolutions of the curve $C_j \subset \mathcal{C}$. This should not lead to confusion as normalization is only mentioned in the next few paragraphs.)

De Jong and van Straten prove that for sandwiched singularities, 1–parameter smoothings correspond to *picture deformations*, which are 1–parameter deformations of the germ $\mathcal{C}$ together with the subscheme w. In fact, de Jong and van Straten describe all deformations of $(X,0)$, but in this paper we are only interested in smoothings. Since we do not use their results in full generality, we omit some technical points and give simpler versions of the definitions and statements from [27].

Informally, picture deformations look as follows. The deformation $\mathcal{C}^s$ is given by individual deformations C_i^s of the curvetta components, so that the deformed germ $\mathcal{C}^s$ is reduced and has irreducible smooth components C_i^s corresponding to the original curvettas. (In the case of plane curves, any deformation is given by *unfolding*, ie by deforming the defining equation of the curve.) The deformation is required to eliminate tangencies between the curvettas, so that for $s \neq 0$ all deformed curvettas C_i^s intersect transversally. Thus, the only singularities of the deformed germ $\mathcal{C}^s = \bigcup_i C_i^s$ for $s \neq 0$ are transverse multiple points. For $s = 0$, the decoration w consists of w_i marked points on the curvetta C_i for each $i = 1, \dots, m$, concentrated at 0. During the deformation, these marked points move along the curvettas, so that for $s \neq 0$, the deformed curvetta C_i^s contains exactly w_i *distinct* marked points, and all intersection points $C_i^s \cap C_j^s$ for $j \neq i$ are marked.

More formally, deforming the curvettas C_i individually means that we consider δ–constant deformations of the reducible germ $\mathcal{C} = \bigcup_i C_i$. Intersection points between deformed curvettas define the total multiplicity scheme m^s on the normalization $\widetilde{\mathcal{C}}^s$ for $s \neq 0$; if all intersections are transverse, the corresponding divisor is reduced, ie each point enters with multiplicity one. The requirement that all intersection points are marked means that the deformation $w^S \subset \widetilde{\mathcal{C}} \times S$ of the decoration w must satisfy the condition $m^s \subset w^s$. The requirement that all marked points be distinct on each C_i^s for $s \neq 0$ is the same as saying that the divisor given by w_i^s is reduced for $s \neq 0$. The condition $m^s \subset w^s$ then implies automatically that all singularities of the deformed germ $\mathcal{C}^s$ are ordinary multiple points, ie the deformed curvettas intersect transversally.

Definition 2.7 A picture deformation $\mathcal{C}^S$ of the decorated germ $(\mathcal{C}, w)$ with smooth components $C_1, \dots, C_m$ over a germ of a smooth curve $(S, 0)$ is given by a δ–constant deformation $\mathcal{C}^S \to S$ of $\mathcal{C}$ and a flat deformation $w^S \subset \widetilde{\mathcal{C}^S} = \widetilde{\mathcal{C}} \times S$ of the scheme w such that for $s \neq 0$, the divisor w^s is reduced, the only singularities of $\mathcal{C}^s$ are ordinary multiple points, and $m^s \subset w^s$.

Strictly speaking, w^S lives in the normalization, but for $s \neq 0$ we can think of w^s as the set of marked points $\{p_1, p_2, \dots, p_n\} \subset \bigcup_{i=1}^m C_i^s$ such that *all* intersection points $C_i^s \cap C_j^s$ are marked. We say that p_i is a *free* marked point if it lies on a single C_i^s (away from the intersections). (Note that these points, and the number of such points n, can generally be different for different picture deformations.)

With these definitions in place, de Jong and van Straten's results on smoothings are as stated in Theorem 1.3: every picture deformation of $(\mathcal{C}, w)$ gives rise to a smoothing

of the corresponding surface singularity $(X,0)$, and every smoothing arises in this way. Specifically, the Milnor fiber of the smoothing that corresponds to the picture deformation $\mathcal{C}^s = \bigcup_{i=1}^m C_i^s \subset \mathbb{C}^2$ with marked points $\{p_1, p_2, \dots, p_n\}$ is obtained by blowing up $\mathbb{C}^2$ at all points $p_1, p_2, \dots, p_n$ and taking the complement of the proper transforms of $C_1^s, \dots, C_m^s$ in $\mathbb{C}\#\#_{j=1}^n \overline{\mathbb{C}\mathrm{P}}^2$. Picture deformations of $\mathcal{C}$ generate all Milnor fibers, that is, each Milnor fiber of $(X,0)$ arises from some picture deformation of $(X,0)$ via this construction. Note that Theorem 1.3 makes no claim of a precise one-to-one correspondence between picture deformations and smoothings: one expects that isomorphic smoothings only come from isomorphic picture deformations (in the appropriate sense), but this has not been established. In certain cases, one can distinguish Milnor fibers by their topological invariants, or by comparing incidence matrices of the corresponding curvetta arrangements; see [27, Section 5] or [44]. We discuss this in Section 6 and use a similar technique to distinguish Stein fillings.

Remark 2.8 To be more precise, we need to consider the compact version of the construction of Milnor fibers, as follows. Fix a closed Milnor ball $B \subset \mathbb{C}^2$ for the germ $\mathcal{C}$. For sufficiently small $s \neq 0$, the deformed arrangement $\mathcal{C}^s$ will have a representative in B which meets $\partial B = S^3$ transversally, and all marked points $p_1, \dots, p_n$ are contained in the interior of B. Let $\widetilde{B}$ be the blow-up of B at $p_1, \dots, p_n$. Because in the picture deformation all the intersections between deformed curvettas are transverse, the proper transforms of $C_1^s, \dots, C_m^s$ in $\widetilde{B}$ will be disjoint smooth disks. Let $T_1, \dots, T_m$ be pairwise disjoint tubular neighborhoods of these proper transforms. As a compact 4–manifold with boundary, the Milnor fiber that corresponds to $\mathcal{C}^s$ is given by $W = \widetilde{B} \setminus \bigcup_{i=1}^m T_i$, after corners are smoothed, and the Stein structure is homotopic to the complex structure induced from the blow-up.

3 Graphical deformations of curvettas yield fillings

Let $(X,0)$ be a rational surface singularity with reduced fundamental cycle, and consider the associated decorated germ $(\mathcal{C}, w)$ of a reducible plane curve as in the previous section, with smooth branches $C_1, C_2, \dots, C_m$ equipped with weights. Our goal is to build an analog of [27] in the symplectic category: it turns out that Stein fillings of the link of $(X,0)$ can be obtained from certain *smooth* homotopies of the branches of the decorated germ $\mathcal{C}$. We will restrict to graphical homotopies to streamline our definition and constructions. (In our setting, one can always choose an appropriate coordinate system, so the graphical hypothesis leads to no loss of generality.)

Fix a closed Milnor ball B for $\mathcal{C}$ as in Remark 2.8, so that each branch C_i intersects ∂B transversally. If B is small enough, the complex coordinates (x, y) in $\mathbb{C}^2$ can be chosen so that all branches $C_1, C_2, \dots, C_m$ are graphical in B: $C_i = \{y = f_i(x)\}$. We will consider smooth graphical arrangements $\Gamma = \{\Gamma_1, \Gamma_2, \dots, \Gamma_m\}$ such that each Γ_i is a smooth graphical disk, so that $\Gamma_i = \{y = g_i(x)\}$ for a smooth function g_i, and Γ_i intersects ∂B transversally.

The following definition is given for homotopies of the branches defined for a real parameter $t \in [0, 1]$. Sometimes we will use the same notion for homotopies defined in a parameter interval $t \in [0, \tau]$, with obvious notational changes. We assume that coordinates (x, y) are chosen as above.

Definition 3.1 Let $(\mathcal{C}, w)$ be a decorated plane curve germ, with weights $w_i = w(C_i)$ of its smooth graphical branches $C_1, C_2, \dots, C_m$. A *smooth graphical homotopy* of $(\mathcal{C}, w)$ is a smooth homotopy C_i^t of the branches of $\mathcal{C}$, so that $\mathcal{C} = \bigcup_{i=1}^m C_i^0$, together with distinct marked points p_k, $k = 1, \dots, n$ (for some n), on $\bigcup_{i=1}^m C_i^1$. We assume that in a Milnor ball B the following conditions are satisfied:

(1) Each branch is given by $C_i^t = \{y = f_i^t(x)\}$ for a function $f_i^t(x) = f_i(x, t)$ smooth in (x, t), and C_i^t intersects ∂B transversally for all t.

(2) Intersections between the branches remain in the interior of B during the homotopy.

(3) At $t = 1$, all intersections of any two branches C_i^1 and C_j^1 are positive and transverse.

(4) At $t = 1$, all intersection points on each branch C_i^1 are marked, and there may be additional free marked points. Each free point lies in the interior of B on a unique branch C_i^1. The total number of marked points on C_i^1 is w_i.

The choice of Milnor ball B is unimportant as all our considerations are local. For brevity, we will often omit B from notation and talk about decorated germs and their homotopies in $\mathbb{C}^2$. In that case, we implicitly work in a fixed neighborhood of the origin, and assume that all intersections between branches which begin in this neighborhood remain in this neighborhood during the homotopy, and thus the components of the arrangement have controlled behavior near the boundary of the neighborhood.

Conditions (1) and (2) are automatically satisfied for "small" homotopies. Indeed, if t is close to 0, C_i^t is C^1–close to C_i. The reducible curve $\mathcal{C}$ with smooth branches has a

finite set of tangent directions at the origin, and the branches C_i^t will have tangent spaces lying in a small neighborhood of these directions in the Grassmannian of symplectic planes in $\mathbb{C}^2$. Therefore we can choose coordinates so that the fiber of the projection avoids these directions. We only include the intersection of the branches of $\mathcal{C}$ at 0 in the Milnor ball B, so for small t, intersections will remain in B. For larger homotopies, we require these conditions nontrivially.

Picture deformations satisfy all of the conditions (1)–(4), so a picture deformation is a special case of a smooth graphical homotopy of the germ (in appropriate coordinates). In contrast to picture deformations of [27], condition (4) on the marked points and the weight restrictions is only required at $t = 1$ for homotopies. For a closer analogy with Definition 2.7, we can consider marked points $\{p_i^j(t)\}_{j\in\{1,\dots,w_i\}}$ on C_i^t for all $0 \le t \le 1$. For $t = 0$, the marked points are concentrated at the origin on each branch, giving the decoration of $(\mathcal{C}, w)$. Suppose that $p_i^j(t)$, $0 \le t \le 1$, are smooth functions describing the motion of marked points during homotopy, so that $p_i^j(t) \in C_i^t$ for all t. For $t = 1$, the points $p_i^j(1) = p_i^j$ satisfy condition (4) above. This implies, in particular, that at $t = 1$, the branch C_i^t has no more than w_i intersection points with other branches. However, for $0 < t < 1$, the marked points $p_i^j(t)$ are not subject to any restrictions and have little significance. The homotoped curvettas C_i^t can have an arbitrary number of intersections, and intersections may be positive or negative. By contrast, for picture deformations, the weights control the number of intersection points between deformed curvettas at all times, the intersections between branches are always marked during deformation, and all intersections are positive because curvettas are deformed through complex curves.

Let (Y, ξ) be the link of the singularity $(X, 0)$ with the decorated germ $(\mathcal{C}, w)$. We will show that every smooth graphical homotopy of the germ $\mathcal{C}$ gives rise to a Stein filling of (Y, ξ).

First, we focus on the curvetta arrangement $\{C_1^1, C_2^1, \dots, C_m^1\}$ with marked points, produced at the end of homotopy at the time $t = 1$. Lemma 3.2 below produces a certain Lefschetz fibration from this input. The lemma applies to any arrangement of smooth graphical disks $\{\Gamma_1, \Gamma_2, \dots, \Gamma_m\}$ satisfying the stated hypotheses; the homotopy is not used at this stage. We use different notation to emphasize that $\{\Gamma_i\}$ need not be related to $\mathcal{C}$. Then, Lemma 3.4 uses the homotopy between the decorated germ $(\mathcal{C}, w)$ and the curvetta arrangement $\{C_1^1, C_2^1, \dots, C_m^1\}$ with its marked points $p_1, p_2, \dots, p_n$ to show that the open book on the boundary of the Lefschetz fibration supports (Y, ξ). It follows that our construction produces a Stein filling of (Y, ξ).

As a smooth 4–manifold, the filling produced by Lemma 3.2 is constructed similarly to the Milnor fibers in Theorem 1.3. Namely, we blow up at each of the intersection points of the homotoped curvettas, as well as at the free marked points, and then take the complement of the proper transforms of the curvettas. Even though C_i^1 are smooth disks (rather than complex curves), we will assume that they are locally modeled on complex curves near the intersection point, so the blow-up and the proper transforms can be understood in the usual sense. Alternatively, one could also think about the proper transform in the smooth sense, as the closure of the complement of the blown-up point; see [20, Definitions 2.2.7 and 2.2.9]. To obtain a 4–manifold with given boundary, we consider a compact version of the construction in a Milnor ball, as explained in Remark 2.8. It is convenient to consider the Milnor ball of the form $B = D_x \times D_y \subset \mathbb{C}^2$, with corners smoothed, where D_x and D_y are disks in the coordinate planes $\mathbb{C}_x$ and $\mathbb{C}_y$. For every $x_0 \in D_x$, the graphical disks Γ_i intersect $\{x_0\} \times D_y$ transversally, and the intersection with $\partial(D_x \times D_y)$ lies as a braid in $\partial D_x \times D_y$. To simplify notation, we do not mention the Milnor ball B explicitly in the first part of the lemma.

Lemma 3.2 *Let $\Gamma_1, \dots, \Gamma_m$ be smooth disks in $\mathbb{C}^2$ which are graphical with respect to the projection π_x, that is, $\Gamma_i = \{y = f_i(x)\}$. Assume that at each intersection point of two or more Γ_i, there exists a neighborhood U of the intersection such that $\bigcup_i \Gamma_i$ is cut out by complex linear equations inside U. (Up to graphical isotopy, this only requires the Γ_i to intersect transversally and positively with respect to the orientation on the graph Γ_i induced from the natural orientation on $\mathbb{C}$.) Let $p_1, \dots, p_n$ be points on the disks Γ_i which include all intersection points, and let $\alpha\colon \mathbb{C}^2 \# n\overline{\mathbb{C}\mathrm{P}}^2 \to \mathbb{C}^2$ be the blow-up at the points $p_1, \dots, p_n$. Let $\widetilde{\Gamma}_1, \dots, \widetilde{\Gamma}_m$ denote the proper transforms of $\Gamma_1, \dots, \Gamma_m$. Then $\pi_x \circ \alpha\colon (\mathbb{C}^2 \# n\overline{\mathbb{C}\mathrm{P}}^2) \setminus (\widetilde{\Gamma}_1 \cup \dots \cup \widetilde{\Gamma}_m) \to \mathbb{C}$ is a Lefschetz fibration whose regular fibers are punctured planes, where each puncture corresponds to a component $\widetilde{\Gamma}_i$. There is one vanishing cycle for each point p_j, which is a curve in the fiber enclosing the punctures that correspond to the components Γ_i passing through p_j.*

Similarly, if $B = D_x \times D_y$ is a Milnor ball that contains all the points $p_i, \dots, p_n$ and contains $(D_x \times \mathbb{C}) \cap \left(\bigcup_i \Gamma_i\right)$, and T_i is a small tubular neighborhood of $\widetilde{\Gamma}_i$, then $\pi_x \circ \alpha\colon (\alpha^{-1}(D_x \times D_y)) \setminus (T_1 \cup \dots \cup T_m) \to D_x$ is a Lefschetz fibration with compact fiber. The fiber is a disk with holes corresponding to the components Γ_i. The vanishing cycles correspond to the points p_j in the same way.

If the curvettas $C_1^s, \dots, C_m^s$ with marked points are the result of picture deformation of the germ (C, w) associated to a surface singularity, then the Lefschetz fibration

constructed as above is compatible with the complex structure on the Milnor fiber of the corresponding smoothing.

Proof Before blowing up, the projection $\pi_x\colon \mathbb{C}^2 \to \mathbb{C}$ is clearly a fibration, and the smooth disks Γ_i are sections of this fibration. If they were disjoint sections, then their complement would be a fibration whose fiber is $\mathbb{C}$ with m punctures. Since the sections intersect, we blow up at each of the intersection points, along with blow-ups at other chosen points on the curves. For each fiber containing one of the p_j where we blow up, the corresponding fiber in the blow-up is the total transform, which is a nodal curve containing the exceptional sphere and the proper transform of the fiber. More specifically, translating the coordinates (x, y) on $\mathbb{C}^2$ to be centered at p_j, the coordinates on the blow-up are

$$\mathbb{C}^2 \# \overline{\mathbb{C}\mathrm{P}}^2_{p_i} = \{((x, y), [u : v]) \mid xv = yu\}.$$

The singular fiber is the total transform of $F = \{x = 0\}$, which has two irreducible components:

$$\big(E = \{((0, 0), [u : v])\}\big) \cup \big(\widetilde{F} = \{((0, y), [0 : 1])\}\big).$$

The node occurs at the intersection of these two components at $((0, 0), [0 : 1])$. Therefore in a neighborhood of the node we can take $v = 1$, so we have local coordinates on the blow-up given by $(y, u) \in \mathbb{C}^2$ where $x = yu$. The projection $\pi_x \circ \alpha$ is given in these coordinates by

$$\pi_x \circ \alpha(y, u) = yu,$$

which is exactly the model for a Lefschetz singularity at $(y, u) = (0, 0)$.

In the coordinate chart on $\mathbb{C}^2$ centered at p_j, let $\Gamma_i = \{(x, f_i(x))\}$. The total transforms of the curves Γ_i which pass through p_j — ie which have $f_i(0) = 0$ — are given by

$$\big(E = \{((0, 0), [u : v])\}\big) \cup \left(\widetilde{\Gamma}_i = \left\{\left((x, f_i(x)), \left[1 : \lim_{a \to x} \frac{f_i(a)}{a}\right]\right)\right\}\right),$$

and those which do not pass through p_j — ie which have $f_i(0) \neq 0$ — lift isomorphically to the blow-up

$$\big\{\big((x, f_i(x)), [x : f_i(x)]\big)\big\}.$$

Note that the proper transforms do not pass through the node $((0, 0), [0 : 1])$. Moreover, since the intersections between the Γ_i were assumed to be transverse, $\lim_{a \to 0} f_i(a)/a$ have different values for different values of i where $f_i(0) = 0$. Therefore, the $\widetilde{\Gamma}_i$

are disjoint sections of the Lefschetz fibration from the blow-up of $\mathbb{C}^2$ to $\mathbb{C}$, so their complement gives a Lefschetz fibration with punctured fibers. Moreover, in the singular fibers, the sections which intersect the exceptional sphere part of the fiber are precisely the proper transforms $\widetilde{\Gamma}_i$ such that Γ_i passed through p_j.

Regular neighborhoods T_i of the $\widetilde{\Gamma}_i$ can be chosen sufficiently small to be disjoint from each other and the Lefschetz singular points, thus yielding the compact Lefschetz fibration. This changes the fiber (converting the punctures into holes) but does not change the fibration structure and the vanishing cycles. The total space of a Lefschetz fibration over a disk is a compact 4–manifold with boundary; the fibration induces a planar open book decomposition on the boundary.

In the case of a picture deformation of the germ (C, w), the deformed curvettas $C_1^s, C_2^s, \dots, C_m^s$ are smooth complex disks with marked points satisfying the hypotheses of the lemma. The Stein structure induced by the Lefschetz fibration is compatible with the complex structure on the Milnor fiber, because $\pi_x \circ \alpha$ is holomorphic. □

Consider a smooth graphical arrangement $\Gamma = \{\Gamma_1, \dots, \Gamma_m\}$ in a Milnor ball $B = D_x \times D_y$, such that each Γ_i transversally intersects the vertical part $\partial D_x \times D_y$ of ∂B and is disjoint from $D_x \times \partial D_y$. Taking the boundaries of the graphical disks, we have an m–braid $\partial\Gamma = \partial\Gamma_1 \cup \partial\Gamma_2 \cup \cdots \cup \partial\Gamma_m \subset \partial B = S^3$. (Each component $\partial\Gamma_i$ is an unknot, but the components are linked.) The monodromy of this braid is called the monodromy of the arrangement Γ. We can interpret the braid group on m strands as the mapping class group $\mathrm{MCG}(\mathbb{C}_m)$ of the m–punctured plane. Then the braid $\partial\Gamma$ is identified with the monodromy ϕ_Γ of the $\mathbb{C}_m$–bundle over S^1, given by the projection $\pi_x : \mathbb{C}^2 \setminus \bigcup_{i=1}^m \Gamma_i \to \mathbb{C}$ restricted to the preimage $\pi_x^{-1}(\partial D_x)$ of the circle $\partial D_x \subset \mathbb{C}$.

To construct the Lefschetz fibration corresponding to Γ in Lemma 3.2, we perform blow-ups at points p_i that project to the interior of D_x. These blow-ups do not affect the bundle over ∂D_x. Therefore, the noncompact version of the Lefschetz fibration (with fiber $\mathbb{C}_m$) has the monodromy ϕ_Γ given by the braid $\partial\Gamma$.

For the compact version of the Lefschetz fibration from Lemma 3.2, the general fiber P_m is the disk D_y with m holes. The fibration induces an open book on its boundary, with page P_m. The boundary of the total space of the fibration $\mathcal{L}^s$ is the union of two parts: the horizontal boundary $\partial P_m \times D$, which forms the binding of the open book, and the vertical boundary, a fiber bundle over $S^1 = \partial D_x$ with fiber P_m, which forms

the mapping torus for the open book. This fiber bundle is given by the projection $(\pi_x \circ \alpha)^{-1}(\partial D_x) \to D_x$, which is the same as the projection $\pi_x \colon B \setminus \bigcup_{i=1}^m \Gamma_i \to D_x$ restricted to $\pi_x^{-1}(\partial D_x)$, because the blow-up map α is the identity over ∂D_x. Let $\phi \colon P_m \to P_m$ denote the monodromy of this fiber bundle, ie the monodromy of the open book. We then have a commutative diagram

$$\begin{array}{ccc} P_m & \xrightarrow{\phi} & P_m \\ \downarrow & & \downarrow \\ \mathbb{C}_m & \xrightarrow{\phi_\Gamma} & \mathbb{C}_m \end{array}$$

where the vertical maps are inclusions. This proves:

Lemma 3.3 *Let $\Gamma = \{\Gamma_1, \dots, \Gamma_m\}$ be a smooth graphical arrangement with marked points $\{p_j\}$, and $B = D_x \times D_y$ a Milnor ball whose interior contains all marked points, such that $\Gamma_i \cap (D_x \times \mathbb{C}) \subset B$ and Γ_i is transverse to ∂B for all $i = 1, \dots, m$. Let ϕ_Γ be the monodromy of the braid $\partial\Gamma = \partial\Gamma_1 \cup \cdots \cup \partial\Gamma_m \subset \partial B = S^3$.*

Let $\phi \colon P_m \to P_m$ be the monodromy of the open book induced by the Lefschetz fibration constructed for $(\Gamma, \{p_j\})$ in Lemma 3.2. Then ϕ_Γ is the image of ϕ under the projection

$$\eta \colon \mathrm{MCG}(P_m) \to \mathrm{MCG}(\mathbb{C}_m)$$

induced by the inclusion $P_m \hookrightarrow \mathbb{C}_m$ of the compact disk with m holes into the m–punctured plane.

When the arrangement $(\Gamma, \{p_j\})$ is related to the decorated germ $(\mathcal{C}, w)$ by a smooth graphical homotopy, the monodromy ϕ_Γ of the braid $\partial\Gamma$ is the same as the monodromy of the braid $\partial\mathcal{C} = \partial C_1 \cup \cdots \cup \partial C_m$, because the homotopy between disks gives an isotopy of the two boundary braids. By definition, the braid monodromy of $\partial\mathcal{C}$ is the monodromy φ of the singular point of $\mathcal{C}$.

We next examine the monodromy of the open book corresponding to Γ in the case of the compact fiber, and find its relation to the monodromy of the singular curve $\mathcal{C}$.

Lemma 3.4 *Let $\{\Gamma_1, \dots, \Gamma_m\}$ and $\{\Gamma_1', \dots, \Gamma_m'\}$ be two smooth graphical arrangements, such that the boundary braid of are braid-isotopic (respecting labels) and the weights on the corresponding components agree. Let $\mathcal{L}$ and $\mathcal{L}'$ be the corresponding Lefschetz fibrations constructed in Lemma 3.2. Then the induced open book decompositions on the boundary have the same page and same monodromy.*

We will prove the lemma after pointing out its consequences. Since the plane curve arrangements at either end of a smooth graphical homotopy have braid-isotopic boundaries, and the weights on the components remain constant during the smooth graphical homotopy, it follows that the open book decomposition induced on the boundary for any Lefschetz fibration arising in this way is independent of the choice of smooth graphical homotopy.

For the case where (Γ, w) is the end point of a *picture deformation* of a plane curve germ $(\mathcal{C}, w)$, $\mathcal{L}'$ is a Lefschetz fibration on the (compactified) Milnor fiber of the associated smoothing of the surface singularity $(X, 0)$, as in Theorem 1.3. In this case the boundary of the Milnor fiber is the link Y of the singularity $(X, 0)$, and the Milnor fiber gives a Stein filling of the canonical contact structure ξ on the link, so the open book supports ξ. Because every rational singularity has a picture deformation yielding a Milnor fiber arising in such a manner (see Section 4 in our case), the open book on the boundary of any Lefschetz fibration arising from the endpoint of a smooth graphical homotopy of the same germ must support the canonical contact structure on the link of the singularity.

Combining Lemmas 3.2 and 3.4 with this discussion completes the proof of Theorem 1.4, which we summarize in the following corollary.

Corollary 3.5 *A smooth graphical homotopy of the decorated germ $(\mathcal{C}, w)$ gives rise to a Stein filling of the link (Y, ξ) of the corresponding singularity.*

Proof of Lemma 3.4 Applying the previous discussion and Lemma 3.3 to the arrangement $\Gamma = \{\Gamma_1, \Gamma_2, \dots, \Gamma_m\}$, we see that the homomorphism $\eta\colon \mathrm{MCG}(P_m) \to \mathrm{MCG}(\mathbb{C}_m)$ sends the open book monodromies ϕ and ϕ' to the same braid monodromy $\varphi \in \mathrm{MCG}(\mathbb{C}_m)$. The kernel of the map $\eta\colon \mathrm{MCG}(P_m) \to \mathrm{MCG}(\mathbb{C}_m)$ is generated by the boundary-parallel Dehn twists around the holes in the fiber P_m. (Recall that the monodromy of an open book is considered *rel* boundary of the page, so while the twists around individual strands are trivial in the braid case, the boundary twists become nontrivial for open books.) It follows that the monodromies ϕ and ϕ' of the open books on the boundaries of $\mathcal{L}$ and $\mathcal{L}'$ can differ *only* by boundary twists, since $\eta(\phi) = \eta(\phi') = \varphi$.

Let T_i denote a positive Dehn twist around the i^{th} hole. Then we have

$$\phi' = \phi \circ T_1^{\alpha_1} \circ T_2^{\alpha_2} \circ \cdots \circ T_m^{\alpha_m} \tag{3-1}$$

for some integers $\alpha_1, \alpha_2, \dots, \alpha_m$. The order is unimportant since the boundary twists are in the center of $\mathrm{MCG}(P_m)$.

It remains to pin down the boundary twists around each hole, ie to show that $\alpha_i = 0$ for every $i = 1, \dots, m$. To do so, we need to take into account the blow-ups at the free marked points p_i (the marked points that lie on the branches away from the intersections). These correspond to boundary twists. Recall a basic fact about diffeomorphisms of a planar surface *rel* boundary: for any two factorizations Ψ and Ψ' of $\psi : P_m \to P_m$, the number of Dehn twists that enclose a given hole h is the same for Ψ and Ψ'. (Here, we count all twists, not only the boundary ones.) The above statement easily follows from the fact that lantern relations generate all relations in the mapping class group of a planar surface [36], and the number of Dehn twists enclosing a given hole is unchanged under a lantern relation. This implies that the number of Dehn twists enclosing the i^{th} hole is well defined for a monodromy $\psi : P_m \to P_m$; let $n_i = n_i(\psi)$ denote this number. If two monodromies ϕ and ϕ' are related by (3-1), we have

$$n_i(\phi') = n_i(\phi) + \alpha_i. \tag{3-2}$$

On the other hand, the number n_i is determined by the vanishing cycles of the Lefschetz fibration. By construction of the fibration $\mathcal{L}^1$ associated to the homotopy $\mathcal{C}^t$, the number of Dehn twists enclosing the i^{th} hole is given by the number of blow-ups at the marked points on C_i^1, which in turn equals the weight w_i of the component C_i of the original germ $\mathcal{C}$. So $n_i(\phi) = w_i = n_i(\phi')$, and $\alpha_i = 0$ from (3-2). □

Remark 3.6 Our description of the open book monodromy for an arrangement is somewhat similar to E Hironaka's results [24] on the monodromy of complexified real line arrangements in $\mathbb{C}^2$. An important difference is that we consider Lefschetz fibrations on the complement of the proper transform of the curves in a blow-up of $\mathbb{C}^2$, while Hironaka computes the monodromy of the fiber bundle over S^1 obtained by projecting the complement of the complex lines in $\mathbb{C}^2$ to a circle of large radius; compare with the proof of Lemma 3.4. She also considers the setting with compactified fibers, by taking the complement of tubular neighborhoods of the lines, and computes the monodromy of line arrangements as an element of $\mathrm{MCG}(P_m)$. It is important to note that even in the compactified setting, her answers are different from the monodromy of the corresponding Lefschetz fibrations that we consider. (The difference is given by some boundary twists.) The discrepancy appears because when the tubular neighborhoods of the C_i^1 are removed from $\mathbb{C}^2$, their parametrization is induced from $\mathbb{C}^2$. When we blow up and take proper transforms of C_i^1, the parametrization of tubular neighborhoods is induced by the Lefschetz fibration structure on the blown-up manifold. These two

parametrizations are different in the two settings, affecting the choice of the meridian of the tubular neighborhood of a line and the framing of the boundary of the corresponding hole.

4 The Lefschetz fibration for the Artin smoothing

4.1 The Scott deformation

We can now use a specific deformation to describe the monodromy of the open book decomposition of (Y,ξ). We will use a canonical deformation, called the *Scott deformation* in [27], which yields a smoothing in the Artin component. This deformation yields a particularly nice arrangement of curvettas where the associated Lefschetz vanishing cycles are disjoint. This in turn yields a model factorization for the monodromy of the boundary open book decomposition. In Proposition 6.7, we will show that the corresponding Stein filling is uniquely recognizable from its combinatorics. Recall that $\operatorname{tang}(C_i,C_j)$ stands for the order of tangency between branches C_i and C_j of $\mathcal{C}$, and $t(C_i)=\max_j \operatorname{tang}(C_i,C_j)$.

Proposition 4.1 *Let $(X,0)$ be a rational surface singularity with reduced fundamental cycle, and $(\mathcal{C},w)$ one of its decorated reducible plane curve germs with m smooth irreducible components. Let (Y,ξ) be the contact link of $(X,0)$. Then (Y,ξ) has a planar open book decomposition whose page is a disk with m holes $h_1,\dots,h_m$, corresponding to the branches of $\mathcal{C}$. The open book monodromy admits a factorization into **disjoint** positive Dehn twists with the following properties*:

(1) *For any two branches C_i and C_j, the corresponding holes h_i and h_j are enclosed by exactly $\operatorname{tang}(C_i,C_j)$ of these Dehn twists.*

(2) *There are $w(C_i)-t(C_i)>0$ boundary Dehn twists around the hole h_i.*

(3) *There is at least one positive Dehn twist about the outer boundary component of the page.*

Proof We use the picture deformation of $(\mathcal{C},w)$ referred to as the *Scott deformation* in [27, Proposition 1.10]. This deformation arises from iteratively applying the following procedure. (We refer the reader to [27; 1] for details, including the explanation why the procedure below can be actually realized by a 1–parameter deformation.)

The input of the procedure is an isolated singular point p of a plane curve C with multiplicity m. In our case C is a union of smooth components, and the multiplicity m

is the number of components through the point p. The output of the procedure is a deformation C' whose singularities are

(I) one m–fold point where m branches intersect transversally, and

(II) the collection of singularities occurring on the proper transform of C in the blow-up of $\mathbb{C}^2$ at p.

The idea of the deformation is to blow up at p, perform a small deformation of the curves so that the singularities of the proper transform become disjoint from the exceptional divisor, and then blow down the exceptional divisor to return to the plane and obtain the curve C'.

We demonstrate this process in an example in Figure 4. The initial configuration in the bottom left consists of five curves. The curves C_1 and C_2 are tangent with multiplicity 3, and these two curves are tangent to C_3 with multiplicity 2. The curves C_4 and C_5 are transverse to C_1, C_2 and C_3 but tangent to each other with multiplicity 4. After blowing up at the common intersection point, we obtain the proper transforms together with an exceptional divisor as shown in the top left of Figure 4. Now C_1 and C_2 are tangent with multiplicity 2 and transversally intersect C_3 at the same point on the exceptional divisor. The curves C_4 and C_5 become disjoint from C_1, C_2 and C_3, and they are tangent to each other with multiplicity 3 at another point on the exceptional divisor. Next we perform the deformation of the curves, fixing the exceptional divisor, but translating the proper transforms $\widetilde{C}_1, \widetilde{C}_2, \dots, \widetilde{C}_5$ of the curvettas slightly so that the intersection of the exceptional divisor with the proper transforms now occurs away from the intersections of the proper transforms with each other, as shown in the top right of Figure 4. Finally, we blow down the exceptional divisor, which results in a transverse intersection of the resulting curvettas $C_1^s, C_2^s, \dots, C_5^s$ together with the singularities (intersections) of the proper transforms, as required.

Since the multiplicity of the orders of tangency between components decreases each time we take the proper transform, applying this procedure iteratively to the singularities of type (II) eventually yields a deformation to a plane curve with only transverse intersections. See Figure 5 for the iterations of the Scott deformation in our example, until all of the singularities are transverse intersection points. When working with a decorated germ $(\mathcal{C}, w)$, with the marked points of w initially concentrated at 0, the same blow-up procedure will separate the marked points. Indeed, if there are additional marked points which increase the weight, they can be separated by additional iterations

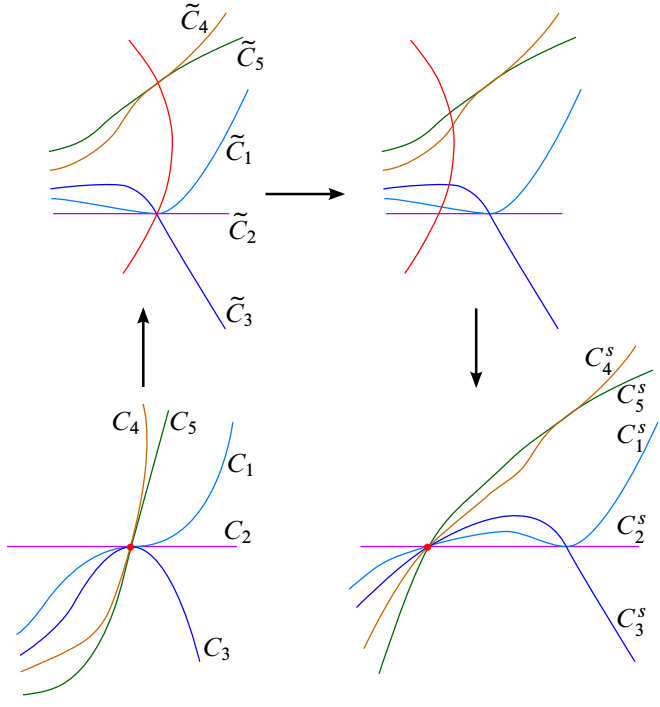


Figure 4: One iteration of the Scott deformation in an example.

of the blow-ups and translations, so that at the end all marked points are disjoint. (In this sense the scheme w^s is reduced.) Note that the total weight $w(C_i)$ of each component is equal to the total number of marked points on that component (including the intersection points).

When the components of $\mathcal{C}$ are smooth, the result of this deformation is as follows. If some components of $\mathcal{C}$ were tangent to order r_1 before the deformation, they will all

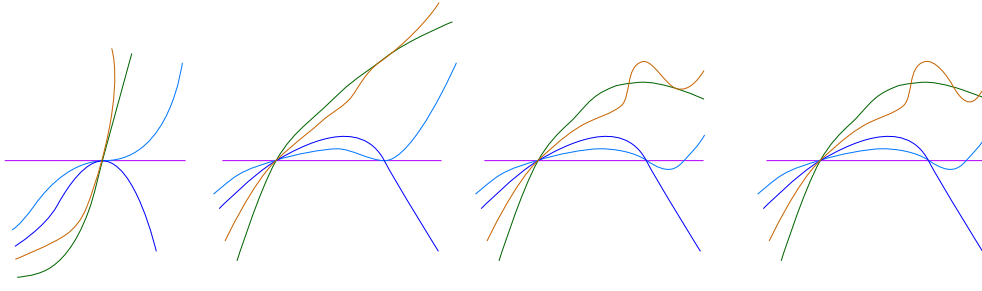


Figure 5: A Scott deformation applied iteratively until all intersections are transverse.

pass through the same r_1 transverse multipoints $p_{i_1}, \dots, p_{i_{r_1}}$. If another component of $\mathcal{C}$ intersects these components with multiplicity $r_2 < r_1$ before the deformation, this component will pass through r_2 of these points afterwards. The total number of intersection points appearing on the Scott deformation of a component C_i is precisely $t(C_i)$, the highest possible order of tangency between C_i and another branch in the original germ $\mathcal{C}$. In this sense, the intersection points are used as efficiently as possible. The number of additional marked points on C_i is $w(C_i) - t(C_i)$.

Now, consider the Lefschetz fibration constructed from the Scott deformation via Lemma 3.2. We claim that up to a curve isotopy, the vanishing cycles of this fibration are disjoint curves on the planar page. The reason for this is built into the iterative nature of the Scott deformation, which results in a nesting of the vanishing cycles as follows.

Consider the equivalence relations on the components $C_1, \dots, C_m$ of the germ $\mathcal{C}$ defined by $C_i \sim_l C_j$ if C_i and C_j intersect at 0 with multiplicity at least l. The transitivity of this relation comes from the fact that if C_1 intersects C_2 with multiplicity r at 0 and C_2 intersects C_3 with multiplicity s at 0, then C_1 must intersect C_3 with multiplicity at least $\min\{r, s\}$. These equivalence relations induce partitions of the components of $\mathcal{C}$, and $\sim_l$ refines $\sim_{l'}$ for $l > l'$.

If we apply the Scott deformation procedure iteratively, on the first iteration, we obtain one transverse intersection of all of the branches (the singularity of type (I)), which groups the components of $\mathcal{C}$ according to the (unique) block of the partition induced from $\sim_1$. Applying the Scott deformation procedure to all the singularities of type (II) generates a transverse multipoint of type (I) for every block in the partition induced by $\sim_2$. Iterating this procedure, for $l \geq 1$ we obtain a transverse intersection for every block of each partition P_l induced by $\sim_l$. For sufficiently large l, each block will consist of a single smooth component, and thus no new transverse intersections of type (I) will result from the procedure. When a block contains a single element, there may or may not be additional marked points placed. Instead of using the partition and Scott deformation to place additional marked points, we can simply use the formula that C_i must have $w(C_i) - t(C_i)$ total additional marked points.

Recall that there is one vanishing cycle in the Lefschetz fibration for each marked point of the Scott deformed curve, and this vanishing cycle encircles the punctures/holes corresponding to the components of curves which pass through the given marked point. Because the equivalence relations $\sim_l$ refine each other as l increases, the subsets

of C_i which intersect at the $(l+1)^{\text{st}}$ iteration are nested within the subsets of C_i which intersect at the l^{th} iteration. Moreover, because the isotopy in the blow-up procedure can be made arbitrarily small, we can assume that there is no braiding of the components C_i between the l^{th} and $(l+1)^{\text{st}}$ iterations; see Section 5 for more details on how braiding of the curves can occur and be understood in general. More specifically, observe that in the Scott deformation procedure, as in Figure 4, the deformation from right to left in the blow-up (at the top of the figure) can be performed by an arbitrarily small translation of the exceptional divisor. By making the translation sufficiently small, we can ensure that in each subset intersecting at the $(l+1)^{\text{st}}$ iteration, the curves stay close together and do not interact with another such subset. (In the language of Section 5, nontrivial braiding would correspond to a crossing of the wires, and a small translation ensures that the wires cannot cross in between the singularities produced iteratively by the Scott deformation.) Then, the vanishing cycles corresponding to the intersections of type (I) which are introduced at the $(l+1)^{\text{st}}$ iteration will be nested inside (and thus disjoint from) the vanishing cycles corresponding to the intersections of type (I) introduced at the l^{th} iteration. We can also assume that any two vanishing cycles introduced in this way at the l^{th} iteration are disjoint, because the application of Lemma 3.2 to the Scott deformation actually realizes these Lefschetz singularities simultaneously in the same fiber (we can later perturb so they arise in different fibers if desired). Finally, the additional marked points at smooth points of the C_i correspond to vanishing cycles which are boundary parallel to the i^{th} hole, and thus can be realized disjointly from each other and all other vanishing cycles. Thus we conclude that the Scott deformation yields a Lefschetz fibration with disjoint vanishing cycles. This means that the compatible planar open book for the link (Y,ξ) has monodromy which is a product of positive Dehn twists about the *disjoint* curves described above. Because at the first step we get a transverse intersection of *all* deformed curvettas, the corresponding vanishing cycle encloses *all* holes, ie we have a Dehn twist about the outer boundary component of the page. □

4.2 Symplectic resolution and Lefschetz fibrations

It is noted in [27] that the Scott deformation corresponds to the Artin smoothing, which in this situation is diffeomorphic to the resolution of the singularity. In fact, we can see more directly, through symplectic topological means, that the Lefschetz fibration corresponding to this Scott deformation gives a plumbing which necessarily corresponds to the resolution of the singularity.

We recall the procedure of [18, Theorem 1.1]. Starting with the plumbing graph G, this procedure produces a planar Lefschetz fibration compatible with the symplectic resolution of a rational singularity with reduced fundamental cycle. (The symplectic structure on the plumbing can be deformed to the corresponding Stein structure.) In fact, [18, Theorem 1.1] applies to a wider class of singularities (see Section 4.3 below), but we first describe it for this particular case. To construct the fiber of the Lefschetz fibration, take a sphere S_v for each vertex $v \in G$ and cut $-a(v) - v \cdot v \geq 0$ disks out of this sphere. (As before, $a(v)$ is the valency of the vertex v; the number of disks is nonnegative by (2-1).) Next, make a connected sum of these spheres with holes by adding a connected sum neck for each edge of G. For a sphere S_v corresponding to the vertex v, the number of necks equals the number of edges adjacent to v, ie its valency $a(v)$. The resulting surface S has genus 0 because G is a tree. See the top of Figure 6 for an example.

Proposition 4.2 [18, Theorem 1.1] *The surface S constructed above is the fiber of a Lefschetz fibration on a symplectic neighborhood of symplectic surfaces intersecting ω–orthogonally according to the graph G. The vanishing cycles are given by the curves parallel to the boundaries of the holes (one curve for each hole) and the cores of the necks of the connected sums.*

Let $\widetilde{X}$ be the Milnor fiber of the Artin smoothing component for a rational $(X, 0)$ with reduced fundamental cycle; $\widetilde{X}$ is a Stein filling for the contact link (Y, ξ). We now have several different Lefschetz fibration structures on $\widetilde{X}$. First, because $\widetilde{X}$ is diffeomorphic to the minimal resolution of $(X, 0)$, a Lefschetz fibration is produced by the Gay–Mark construction of Proposition 4.2. Second, for each choice of the decorated germ $(\mathcal{C}, w)$ with smooth branches, the proof of Proposition 4.1 also gives a Lefschetz fibration on $\widetilde{X}$. All these Lefschetz fibrations have planar fibers. In our construction of the Lefschetz fibration from the curvetta arrangement, the general fiber has a distinguished "outer" boundary component coming from the fibration $\pi : B \to \mathbb{C}$ on the Milnor ball $B = D_x \times D_y \subset \mathbb{C}^2$. In the Gay–Mark construction, there is no distinguished boundary component of the fiber. On the other hand, the decorated germ is not uniquely defined: recall from Proposition 2.4 that there are mult X choices of decorated germs with smooth branches representing $(X, 0)$, where some of these germs may coincide due to symmetries in the extension of the resolution graph. Of course, since the link of the singularity is independent of the choice of curvetta germs, the Stein filling arising from the Artin smoothing should not depend on these choices. We now show that the choice of curvettas

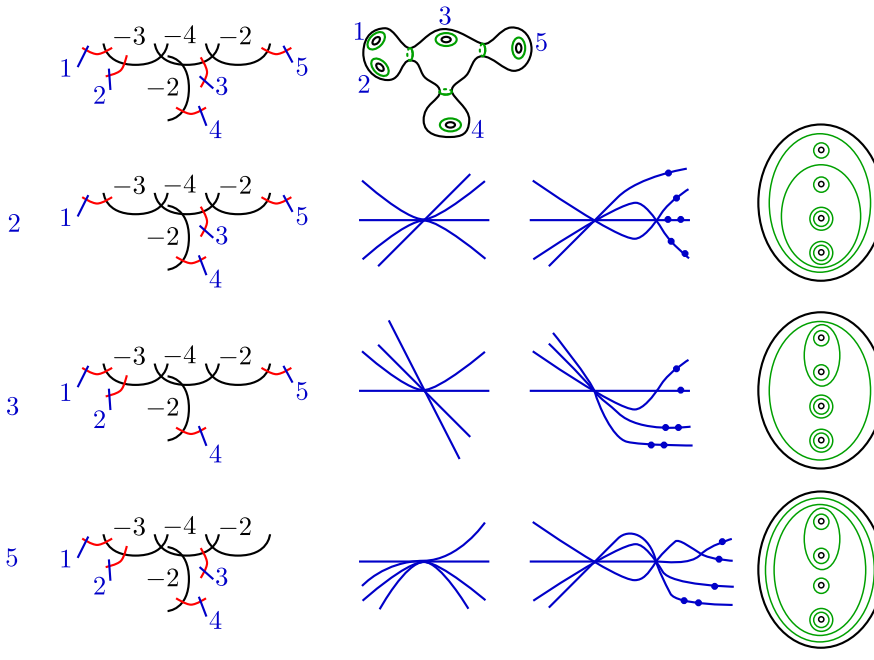


Figure 6: An example demonstrating different choices of curvettas correspond to different choices of outer boundary component for the fiber of the Lefschetz fibration. At the top we have the resolution configuration and the corresponding Gay–Mark Lefschetz fiber with vanishing cycles. The resolution configuration is augmented with red (-1) curves and blue curvettas. For each choice of curvettas we delete exactly one of these red (-1) curves and the corresponding curvetta. We show the resulting curvettas, their Scott deformation, and the corresponding planar Lefschetz fibration obtained from Lemma 3.2 in the cases of excluding the (-1) curves labeled 2, 3, and 5. Note that because of symmetries in the graph, the exclusion of 1 or 2 yield very similar looking cases, and similarly with the exclusion of 4 or 5.

corresponds precisely to the choice of the outer boundary component, so this choice only affects the presentation of the Lefschetz fibration. See Figure 6 for an example.

Lemma 4.3 *Let $\mathcal{L}$ be the planar Lefschetz fibration on $\widetilde{X}$ provided by Proposition 4.2. Then the* mult X *different choices of smooth curvetta germs for $(X,0)$ produce, via the Scott deformation, planar Lefschetz fibrations on $\widetilde{X}$ with a distinguished boundary component of the fiber. The choices of smooth curvetta germs are in one-to-one correspondence with the different choices of outer boundary component of the general fiber of $\mathcal{L}$.*

Proof As before, we associate to each vertex of the resolution graph G for the singularity the quantities $v\cdot v$ for the self-intersection and $a(v)$ for the valency. In the Gay–Mark Lefschetz fibration $\mathcal{L}$, each vertex $v \in G$ contributes $-v\cdot v - a(v)$ boundary components to the fiber. On the other hand, recall from the proof of Propositions 2.2 and 2.4 that the germ $\mathcal{C}$ of smooth curvettas is obtained from an extension of the resolution graph G to a graph G'. We attach $-v\cdot v - a(v)$ vertices with self-intersection -1 and valency 1 to each vertex v to obtain a graph G'', and then delete exactly one of these (-1) vertices to get the graph G'. This shows that the number of choices for the germ matches the number of boundary components of the fiber of $\mathcal{L}$, and this number is exactly $\operatorname{mult} X = -\sum(v\cdot v + a(v))$. The curvetta branches of the germ $\mathcal{C}$ are obtained by taking disks dual to the remaining (-1) vertices and considering their proper transform after blowing down all exceptional divisors; thus the curvettas are in one-to-one correspondence with the (-1) vertices of G'. In turn, in the Lefschetz fibration constructed by Lemma 3.2, the "inner" boundary components of the fiber are in one-to-one correspondence with the curvettas. The deleted (-1) vertex in G'' still corresponds to a boundary component in the fiber of the Gay–Mark Lefschetz fibration $\mathcal{L}$, thus we can say that it corresponds to the outer boundary component of the fiber of the planar Lefschetz fibration produced by Lemma 3.2. Note also that if we enumerate the (-1) vertices of the graph G' by $1, 2, \dots, m = \operatorname{mult} X - 1$, we get an enumeration of the components of $\mathcal{C}$, which in turn gives an enumeration of the holes of the fiber. □

Recall from Remark 2.6 that there may be different analytic types of singularities with the same link (Y,ξ). These singularities are all topologically equivalent and have the same graph G, so that decorated germs for each of these singularities are obtained from extensions of G. A particular choice of extension gives topologically equivalent decorated germs for all singularities with link Y. Topologically equivalent germs yield the same open book decompositions of (Y,ξ) as in Proposition 4.1, since the weights and the orders of tangency between branches are encoded by the topological type. Together with the previous proposition, this gives:

Corollary 4.4 *Let (Y,ξ) be a link of surface singularity with reduced fundamental cycle. Then for any singularity $(X,0)$ whose link is Y, and any choice of the decorated germ $\mathcal{C}$ for $(X,0)$ with smooth branches, the open book decomposition of (Y,ξ) defined by $\mathcal{C}$ is the same; namely, the open book induced by the Gay–Mark Lefschetz fibration. Different extensions G' of the resolution graph G used to construct $\mathcal{C}$ correspond to*

different choices of the outer boundary of the page of the open book. Enumeration of the branches of $\mathcal{C}$ (or equivalently, of the (-1) vertices of G') corresponds to enumeration of the holes in the page.

It is interesting to note that the Milnor fiber of the Artin smoothing is the *only* Stein filling with disjoint vanishing cycles in its Lefschetz fibration.

Proposition 4.5 *Suppose a planar Lefschetz fibration has disjoint vanishing cycles, with at least one boundary parallel vanishing cycle for each boundary component. Then this is a Lefschetz fibration for the Artin smoothing of a rational singularity with reduced fundamental cycle. In particular, the induced open book decomposition on the boundary supports the contact link of a rational singularity with reduced fundamental cycle.*

Proof As in [18], if the vanishing cycles are disjoint, we can realize all Lefschetz singularities simultaneously in the same fiber. The unique singular fiber is thus a configuration of spheres intersecting transversally according to a graph. Note that the boundary parallel twists are important to ensure that the only nonclosed components of the singular fiber are disks which retract to a point. (These disks come from the small annuli around the holes.) The nonsingular fibers provide a regular neighborhood for the configuration, so the entire 4–manifold is a symplectic plumbing. This 4–manifold gives a symplectic filling for a contact structure supported by a planar open book, thus by [15] its intersection form is negative definite, ie the plumbing graph G is negative definite. Thus, the graph can be thought of as the resolution graph of a normal surface singularity $(X, 0)$.

As in [18], $-v \cdot v \geq a(v)$ for each vertex $v \in G$, so $(X, 0)$ is a rational singularity with reduced fundamental cycle. To see this, observe that each vertex $v \in G$ corresponds to a closed component $\widehat{S}_v$ of the singular fiber. Alternatively, $\widehat{S}_v$ can be viewed as the union of a component S_v of the complement of the vanishing cycles in a regular fiber capped off by thimbles for each of its boundary vanishing cycles. Then, $v \cdot v = \widehat{S}_v \cdot \widehat{S}_v$ equals the negative number of thimbles in $\widehat{S}_v$, or equivalently the negative number of vanishing cycles on the boundary of S_v; see [19, Proposition 2.1]. The valency $a(v)$ is the number of other spheres in the singular fiber intersecting $\widehat{S}_v$. Put differently, $a(v)$ is the number of closed surfaces $\widehat{S}_{v'}$, $v' \neq v$, such that S_v and $S_{v'}$ share a vanishing cycle in their boundaries; thus $a(v)$ is the number of the vanishing cycles in ∂S_v that are not adjacent to a boundary component in the fiber. Then $-v \cdot v - a(v)$ is the number of

vanishing cycles adjacent to a boundary component in ∂S_v, so $-v\cdot v - a(v) \geq 0$, as required. Note also that $v\cdot v \leq -2$, as each S_v has at least 2 vanishing cycles on the boundary, so G is the graph of the *minimal* resolution.

The above discussion implies that if we run the construction of Proposition 4.2 for the graph G, we recover the given Lefschetz fibration. It follows that our Lefschetz fibration is compatible with the symplectic structure on the minimal resolution. For a rational singularity, the resolution is diffeomorphic to the Milnor fiber of the Artin smoothing (and the symplectic structure on the symplectic plumbing deforms to the corresponding Stein structure). This shows that the Lefschetz fibration produces the same filling as the Artin smoothing. □

4.3 A digression: some nonrational singularities and potential unexpected fillings

Although we stated Proposition 4.2 for rational singularities, Theorem 1.1 of [18] is more general: the same construction works when the normal crossings resolution has exceptional curves of higher genus, as long as condition (2-1) is satisfied. The fiber of the corresponding Lefschetz fibration is formed by taking the connected sum of surfaces given by the exceptional curves and cutting $-v\cdot v - a(v) \geq 0$ holes in the surface corresponding to $v \in G$. As before, the vanishing cycles are given by the boundary parallel curves around the holes and the curves around the connected sum necks. We can use this construction together with monodromy factorizations of [7] to construct infinite collections of Stein fillings for links of certain nonrational singularities.

Indeed, suppose that a normal surface singularity $(X, 0)$ has a good resolution such that one of the exceptional curves has genus $g \geq 2$ and self-intersection $-d$, with $d \leq 2g-4$. As before, we assume that the resolution graph has no bad vertices, ie satisfies (2-1). Then the fiber of the Lefschetz fibration from [18, Theorem 1.1] has a subsurface of genus g with some necks and holes, and a vanishing cycle around each neck and each hole. (See Figure 7.) The total number of these vanishing cycles is d. We can cut out this subsurface along the curves parallel to the vanishing cycles to get a surface of genus g with d holes, so that the product of the Dehn twists around the vanishing cycles is the boundary multitwist. For $d \leq 2g-4$, [7, Theorem A] establishes that the boundary multitwist has infinitely many positive factorizations as products of Dehn twists about nonseparating curves. These factorizations can consist of arbitrarily many Dehn twists. It follows that the monodromy of the corresponding open book on the

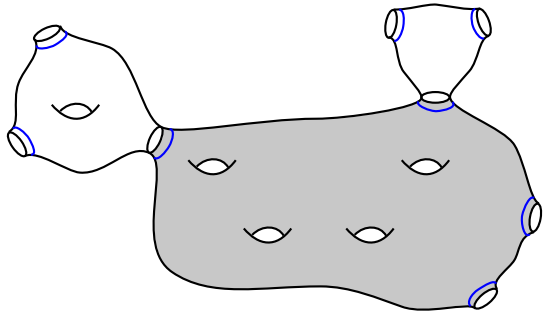


Figure 7: The Gay–Mark Lefschetz fibration for the resolution of a non-rational singularity which admits infinitely many unexpected fillings. The subsurface of genus 4 with $d = 4$ used to produce infinitely many monodromy factorizations is shaded. Vanishing cycles are drawn in blue.

link (Y, ξ) has infinitely many positive factorizations, each of which produces a positive allowable Lefschetz fibration (see [2]) and thus a Stein filling; these Stein fillings can have arbitrarily high Euler characteristic. We ask:

Question 4.6 Does the above construction produce any unexpected Stein fillings?

To answer this question, one would need to contrast these Stein fillings and the Milnor fibers of *all* surface singularities with the given link. Each fixed singularity can only have finitely many Milnor fibers. (Indeed, the Milnor fibers correspond to the components of the base of miniversal deformation; the base is a germ of an analytic space, and as such it can only have finitely many components; see eg [56, Theorem 4.10 and the discussion in Section 7].) However, because of the presence of a higher-genus surface in the resolution, every singularity as above is not (pseudo)taut [32], which means that there exist infinitely many analytic types of singularities with the same dual resolution graph, and thus the same contact link. We are interested in the Stein topology of the Milnor fibers, which is more coarse than the analytic type; in principle, it is possible that the infinite collection of analytic types of the singularity would only give rise to finitely many Stein homotopy types for the Milnor fibers. Thus, we have the following dichotomy: either

(1) there are only finitely many Stein homotopy types (or diffeomorphism types) of the Milnor fibers, which would imply existence of unexpected fillings, or

(2) an infinite collection of possible analytic types gives rise to an infinite collection of pairwise distinct Stein fillings.

Establishing either outcome would be extremely interesting, even for a single example.

It should also be noted that in the nonrational case, one should in principle consider nonnormal singularities as well, as these might generate additional Stein fillings; see [55] for a detailed discussion of this issue (which doesn't arise in the rational case).

Remark 4.7 In a related direction, it is interesting to take a closer look at a family of examples given by cones over curves. Consider a normal surface singularity whose resolution has a unique exceptional curve of genus $g \geq 2$ with self-intersection $-d$ for $d > 0$. The resolution is the total space of the complex line bundle of degree d over the corresponding Riemann surface, and the singularity can be thought of as cone point. The link is a circle bundle over the genus g surface, with Euler number $-d$. The canonical contact structure is the Boothby–Wang structure, which has an open book decomposition as described above: the page is a genus g surface with d boundary components, and the monodromy is the boundary multitwist.

As explained above, for $d \leq 2g-4$ we have an infinite collection of Stein fillings, produced by factorizations of the multitwist. Interestingly, this method no longer applies when $d > 4g+4$: in that range, the boundary multitwist admits no nontrivial positive factorizations, again by [7, Theorem A]. The singularity given by a cone over a projective curve is nonsmoothable when $d > 4g+4$ by [62]; in fact, it is also known that the resolution gives the unique Stein filling in this case [49, Proposition 8.2].

Similarly, for cones over elliptic curves, ie $g = 1$, the singularity is nonsmoothable for $d > 9$ [53], and the only Stein filling is indeed given by the resolution, while for $d \leq 9$, all Stein fillings are given by smoothings and resolutions [50].

5 Every symplectic filling comes from a symplectic deformation of curvettas

5.1 Braided wiring diagrams

A braided wiring diagram is a generalization of a braid in $\mathbb{R} \times \mathbb{C}$ (where the braid condition means that the curves should be transverse to each $\{t\} \times \mathbb{C}$). In a wiring diagram, instead of only looking at smooth braids, we allow the strands to intersect. Let $\pi_{\mathbb{R}} : \mathbb{R} \times \mathbb{C} \to \mathbb{R}$ denote the projection to the first coordinate. We will also use the natural projection from $\mathbb{C}$ to $\mathbb{R}$ sending a complex number to its real part.

Definition 5.1 A *braided wiring diagram* is a union of curves $\gamma_j\colon \mathbb{R} \to \mathbb{R}\times\mathbb{C}$, $j = 1,\dots,n$, each of which is a section of the projection $\pi_{\mathbb{R}}\colon \mathbb{R}\times\mathbb{C} \to \mathbb{R}$, ie each "wire" is given by $\gamma_j(t) = (t, h_j(t) + i\,w_j(t))$. Different wires γ_j may intersect; in this article we will assume that they are not tangent at intersections.

We say a braided wiring diagram is in *standard form* if there are disjoint intervals $I_1,\dots,I_N \subset \mathbb{R}$ such that $I_\ell \times \mathbb{C}$ contains a unique intersection point of some subcollection of the curves γ_j, and in $I_\ell\times\mathbb{C}$, the wires are given as $\gamma_j(t) = (t, k_j t + a_j + i b_j)$. If γ_j does not pass through the intersection point, we require $k_j = 0$.

Note that any braided wiring diagram can be isotoped through braided wiring diagrams to be in standard form.

We can encode a braided wiring diagram by projecting the union of the images of the γ_j to $\mathbb{R}\times\mathbb{R}$ and denoting the crossings of the projection as in a knot diagram.

A braided wiring diagram can be encoded by a sequence

$$(\beta_0, J_1, \beta_1, J_2, \dots, \beta_{m-1}, J_m, \beta_m),$$

where each β_i is a braid and $J_i = \{k_i, k_i+1, \dots, k_i+\ell_i\}$ is a consecutive sequence of integers indicating the local indices of the strands involved in the i^{th} intersection point. For brevity, we will say that J_i is a consecutive set.

Conventions Strands in a wiring diagram are numbered from bottom to top. The convention in [12] is to draw this sequence of braids and intersections from right to left. If one thinks of composing words in the braid group using group notation (left to right) instead of functional notation (right to left), then one will need to read off the braid words from left to right — this is the convention used in [12]. However, in our case since we are always thinking of braids as diffeomorphisms of the punctured plane, we will use functional notation to compose braid words, and thus read everything — the intersections and the braid words — from right to left.

Example 5.2 The braided wiring diagram shown in Figure 8 corresponds to the sequence

$$(\mathrm{id}, \{2,3\}, \mathrm{id}, \{3,4\}, \sigma_1^{-1}\circ\sigma_2^{-1}, \{3,4\}).$$

Braided wiring diagrams were introduced in [12] (inspired by foundational work of [40] and generalized from diagrams of [21]) to study configurations of complex curves,

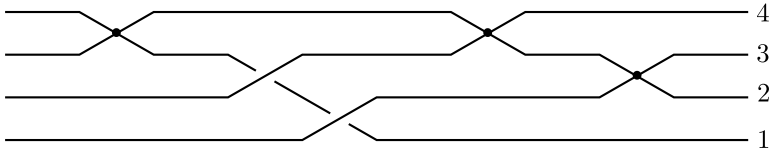


Figure 8: Braided wiring diagram.

particularly line arrangements, and the fundamental groups of their complements. The definition works just as well to study configurations of smooth graphical disks in $\mathbb{C}^2$. As in Section 3, let (x, y) be complex coordinates on $\mathbb{C}^2$, and let π_x be the projection to the first coordinate. Let $\Gamma_1, \dots, \Gamma_m$ be smooth disks in $\mathbb{C}^2$ which are graphical with respect to the projection to x, so $\Gamma_i = \{y = f_i(x)\}$. Assume that all the intersections between the Γ_i are transverse and positive (with respect to the natural orientation on the graphical disks projecting to $\mathbb{C}$).

Definition 5.3 For a graphical configuration $\Gamma = \{\Gamma_1, \Gamma_2, \dots, \Gamma_n\}$ of smooth disks in $\mathbb{C}^2$, a braided wiring diagram is obtained as follows. Choose a (real) embedded curve $\eta\colon [0,1] \to \mathbb{C}$ which passes once through the projection of each singular point of the configuration and whose real part $\operatorname{Re}\eta$ is nonincreasing. The preimage of the curve η under π_x in $\mathbb{C}^2$ is diffeomorphic to $[0,1]\times\mathbb{C}$, and the intersection of this copy of $[0,1]\times\mathbb{C}$ with the configuration Γ is the braided wiring diagram.

The transversality of each smooth disk Γ_j to the projection π_x ensures that the wiring diagram curves are transverse to the projection $\pi_{\mathbb{R}}\colon \mathbb{R}\times\mathbb{C} \to \mathbb{R}$. Note that different choices of η may result in different braided wiring diagrams, which are related by certain generalized Markov moves. See for example [12] for more details. We will show in Section 5.3 that one can always construct a configuration Γ with a given braided wiring diagram; moreover, the components Γ_j of Γ can be chosen to be symplectic.

5.2 Braided wiring diagrams to vanishing cycles

Given a configuration $\Gamma = \{\Gamma_1, \Gamma_2, \dots, \Gamma_m\}$ in $\mathbb{C}^2$ as above, Lemma 3.2 produces an associated Lefschetz fibration. Recall that a Lefschetz fibration is completely determined by its fiber and an ordered list of vanishing cycles. (Critical points are assumed to have distinct critical values.) The fiber in this situation is planar with m boundary components, where m is the number of curves in the configuration. If we are given a braided wiring diagram of Γ, we can explicitly determine the vanishing cycles, as follows.

To describe the vanishing cycles of a Lefschetz fibration $L\colon M \to \mathbb{C}$, we first need to fix certain data. Choose a regular fiber $F_0 := L^{-1}(p_0)$ as the reference fiber. Let $p_1, \dots, p_n$ denote the critical values of L. Choose paths η_j connecting p_0 to p_j in the complement of the p_j, such that the paths η_j are ordered counterclockwise from 1 to n locally around p_0. Then the j^{th} vanishing cycle V_j is the simple closed curve in F_0 which collapses to a point under parallel transport along the path η_j.

When given a braided wiring diagram, we can construct the paths η_j in a systematic manner and compute the vanishing cycles V_j in terms of the braided wiring data. The wiring diagram lies over a curve $\eta\colon [0,1] \to \mathbb{C}$ whose real part $\operatorname{Re}\eta$ is always decreasing. The Lefschetz fibration from Lemma 3.2 comes from the composition $L := \pi \circ \alpha$ of the blow-down map $\alpha\colon \mathbb{C}^2 \#_n \overline{\mathbb{C}\mathrm{P}}^2 \to \mathbb{C}^2$ with the projection map $\pi_x\colon \mathbb{C}^2 \to \mathbb{C}$. One then takes the complement of the sections given by proper transforms of the curves $\Gamma_1, \Gamma_2, \dots, \Gamma_m$ in $\mathbb{C}^2 \#_n \overline{\mathbb{C}\mathrm{P}}^2$, so that each Γ_j corresponds to a hole in the planar fiber. Thus the j^{th} hole corresponds to the wire γ_j in the diagram, and in the standard form the holes are arranged vertically in the fiber, labeled $1, \dots, m$, consecutively. Each consecutive set J_i corresponds to a subcollection of holes contained in a convex subset of $\mathbb{C}$. The Lefschetz critical points occur in $\mathbb{C}^2 \#_n \overline{\mathbb{C}\mathrm{P}}^2$ above the intersection points of the braided wiring diagram. Let $0 < t_1 < \cdots < t_n < 1$ denote the times at which the j^{th} intersection point of the wiring diagram lies over $\eta(t_j)$. We will choose our reference fiber to lie over the right endpoint $p_0 = \eta(0)$ of the curve η in $\mathbb{C}$. Strictly speaking, we need a compact version of this construction, which is obtained by working in a closed Milnor ball and taking complements of tubular neighborhoods of the $\overline{\Gamma}_i$, but for simplicity we omit the Milnor ball from the notation.

We will choose paths $\eta_j : [0, t_j] \to \mathbb{C}$ given by $\eta_j(t) = \eta(t) - \varepsilon_j \rho_j(t) i$, where $\rho_j : [0, t_j] \to [0,1]$ is a bump function which is 0 near $t = 0$ and $t = t_j$, and 1 outside a small neighborhood of 0 and t_j, and $0 < \varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_n < \varepsilon$. See Figure 9.

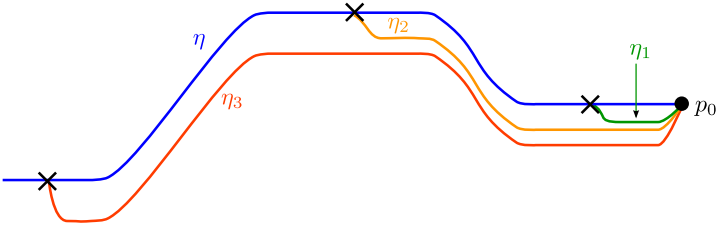


Figure 9: The vanishing paths η_j chosen to identify the vanishing cycles in the fiber over p_0 relative to the wiring diagram path η.

Our local model for the Lefschetz fibration in Lemma 3.2 shows that the curve which collapses to a point in the fiber $L^{-1}(\eta_j(t_j-\delta))$ (for small $\delta>0$) is a convex curve enclosing the holes in the set J_j. To determine the vanishing cycle in our reference fiber $F_0=L^{-1}(p_0)$, we need to track the monodromy over the path η_j for $t\in[0,t_j-\delta]$. This is the monodromy of the braid given by the intersection of the configuration with the slice of $\mathbb{C}^2$ which projects to η_j. (Note that this intersection is indeed a braid over the interior of η_j, because each curve η_j is disjoint from the critical points away from its endpoints.) By assuming ε to be sufficiently small, we see that this braid agrees with corresponding portion of the braided wiring diagram, except when passing near an intersection point. When η_j passes an interval near t_k for $k<j$, the braid resolves the intersection by separating the strands. The strands are ordered from bottom to top in decreasing order by slope in the projection $\mathbb{R}\times\mathbb{C}\to\mathbb{R}\times\mathbb{R}$ (the most positive slope is the lowest strand in the crossing). This can be verified by checking the local model for the complexification of real lines because all of our intersections are positive and transverse; see [40]. After resolving an intersection of the strands in the set $J_k=\{i_k,i_k+1,\dots,i_k+l_k\}$, the element of the mapping class group which corresponds to this portion of the braid from right to left is Δ^{-1}, where Δ is the positive half-twist of the strands $i_k,i_k+1,\dots,i_k+l_k$. (In terms of the standard generators of the braid group, $\Delta_{J_k}=(\sigma_{i_k}\cdots\sigma_{i_k+l_k-1})(\sigma_{i_k}\cdots\sigma_{i_k+l_k-2})(\sigma_{i_k}\sigma_{i_k+1})(\sigma_{i_k})$.) Therefore, the braid lying above η_j is given by

$$\phi_j=\beta_{j-1}\circ\Delta_{j-1}^{-1}\circ\cdots\circ\beta_1\circ\Delta_1^{-1}\circ\beta_0,$$

where Δ_k denotes the positive half-twist of the strands in the set J_k. Namely, Δ_k is the diffeomorphism supported in a neighborhood of the disk convexly enclosing the holes in the set J_k, which acts by rotating the disk by π counterclockwise. The j^{th} vanishing cycle is the curve which is taken to the convex curve A_j enclosing the holes in the set J_j under the braid lying above η_j. Therefore, $V_j=\phi_j^{-1}(A_j)$.

Remark 5.4 We can encode blow-ups at "free" points (as is allowed by Lemma 3.2) by adding marked points in our braided wiring diagram indicating "intersection points" that involve only a single strand (so the corresponding J will have $|J|=1$).

The total monodromy of the curve configuration around a circle enclosing all of the critical points can now be calculated in two different ways:

(1) Using the total monodromy of the curve configuration encoded by the braided wiring diagram.

(2) Taking the product of positive Dehn twists about the induced vanishing cycles.

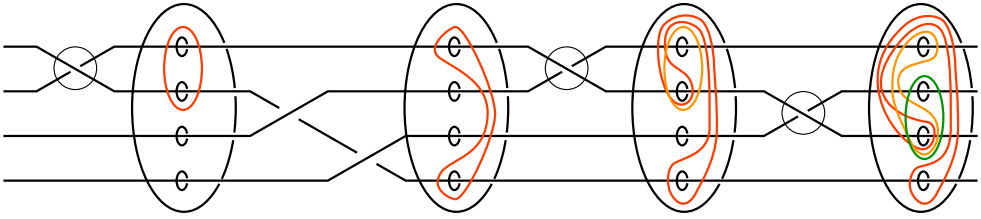


Figure 10: Vanishing cycles corresponding to the braided wiring diagram of Figure 8. The circled crossings correspond to intersections in the wiring diagram. Uncircled crossings come from braiding between intersections.

To reassure the reader that our formulas and conventions are consistent, we verify that these two different ways of calculating the monodromy agree.

The total monodromy encircling a braided wiring diagram

$$(\beta_0, J_1, \beta_1, \dots, \beta_{n-1}, J_n, \beta_n)$$

is given by following the diffeomorphisms induced by a counterclockwise rotation around the wiring interval. Such a counterclockwise circle is obtained by connecting an upward push-off of the wire interval oriented right to left with a downward push-off oriented left to right as in Figure 11. The intersections between the strands of J_j are resolved as the positive half-twist Δ_j in the upward push-off (right to left). In the downward push-off the intersection is resolved as the negative half-twist Δ_j^{-1} right to left, but since we pass through the downward push-off from left to right, each such segment contributes Δ_j to the monodromy. The braids contribute β_j when traversed right to left, and β_j^{-1} when traversed left to right. See Figure 11. The total monodromy is therefore

$$\beta_0^{-1}\circ\Delta_1\circ\beta_1^{-1}\circ\Delta_2\circ\beta_2^{-1}\circ\cdots\circ\beta_{n-2}^{-1}\circ\Delta_{n-1}\circ\beta_{n-1}^{-1}\circ\Delta_n^2\circ\beta_{n-1}\circ\Delta_{n-1}\circ\beta_{n-2}\circ\cdots\circ\beta_2\circ\Delta_2\circ\beta_1\circ\Delta_1\circ\beta_0.$$

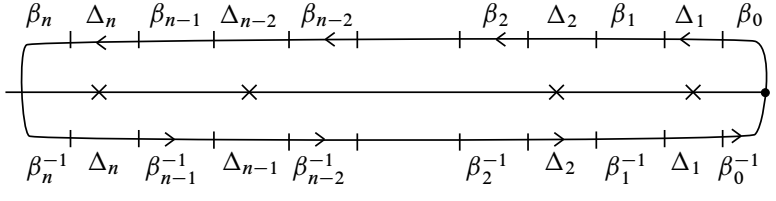


Figure 11: The total monodromy about a braided wiring diagram.

On the other hand, each vanishing cycle is given as

$$V_j = \phi_j^{-1}(A_j) = (\beta_{j-1} \circ \Delta_{j-1}^{-1} \circ \cdots \circ \beta_1 \circ \Delta_1^{-1} \circ \beta_0)^{-1}(A_j).$$

Therefore a Dehn twist τ_{V_j} about V_j is equal to

$$\tau_{V_j} = \phi_j^{-1} \circ \Delta_j^2 \circ \phi_j$$

because $\tau_{A_j} = \Delta_j^2$ and in general $\tau_{\phi(C)} = \phi \circ \tau_C \circ \phi^{-1}$. Thus, the total monodromy of the Lefschetz fibration given by the product of positive Dehn twists about the vanishing cycles is

$$\phi_n^{-1} \circ \Delta_n^2 \circ \phi_n \circ \phi_{n-1}^{-1} \circ \Delta_{n-1}^2 \circ \phi_{n-1} \circ \cdots \circ \phi_1^{-1} \circ \Delta_1^2 \circ \phi_1.$$

We can simplify $\phi_j \circ \phi_{j-1}^{-1}$ as

$$(\beta_{j-1} \circ \Delta_{j-1}^{-1} \circ \cdots \circ \beta_1 \circ \Delta_1^{-1} \circ \beta_0) \circ (\beta_0^{-1} \circ \Delta_1 \circ \beta_1^{-1} \circ \cdots \circ \Delta_{j-2} \circ \beta_{j-2}^{-1}) = \beta_{j-1} \circ \Delta_{j-1}^{-1}.$$

Therefore $\tau_{V_n} \circ \cdots \circ \tau_{V_1}$ is equal to

$$\phi_n^{-1} \circ \Delta_n^2 \circ (\beta_{n-1} \circ \Delta_{n-1}^{-1}) \circ \Delta_{n-1}^2 \circ \cdots \circ (\beta_1 \circ \Delta_1^{-1}) \circ \Delta_1^2 \circ \beta_0,$$

which equals

$$\beta_0^{-1} \circ \Delta_1 \circ \beta_1^{-1} \circ \cdots \circ \Delta_{n-1} \circ \beta_{n-1}^{-1} \circ \Delta_n^2 \circ \beta_{n-1} \circ \Delta_{n-1} \circ \cdots \circ \beta_1 \circ \Delta_1 \circ \beta_0.$$

This coincides with the total monodromy of the braided wiring diagram given above, as required.

5.3 Wiring diagrams to symplectic configurations

Given any braided wiring diagram, we interpret it as a collection of intersecting curves in $\mathbb{R} \times \mathbb{C}$. We will extend each of these curves to a symplectic surface in $\mathbb{C} \times \mathbb{C}$.

Proposition 5.5 *Given a braided wiring diagram $\bigcup_j \gamma_j \subset \mathbb{R} \times \mathbb{C}$ in standard form, there exists a configuration of symplectic surfaces $\bigcup_j \Gamma_j$ in $\mathbb{C} \times \mathbb{C}$ such that Γ_j extends γ_j, that is,*

$$\left(\bigcup_j \Gamma_j \right) \cap (\mathbb{R} \times \{0\} \times \mathbb{C}) = \bigcup_j \gamma_j,$$

and all intersections $\Gamma_j \cap \Gamma_k$ lie in the original wiring diagram in $(\mathbb{R} \times \{0\} \times \mathbb{C})$ and are transverse and positive.

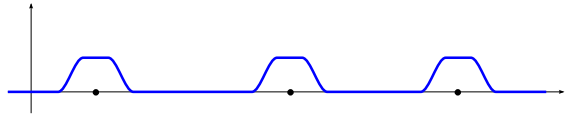


Figure 12: The graph of $\rho\colon \mathbb{R} \to [0,1]$. The marked points on $\mathbb{R}$ are the t_i.

Proof Let $t_1 = \pi_{\mathbb{R}}(p_1), \dots, t_n = \pi_{\mathbb{R}}(p_n)$ denote the $\mathbb{R}$ coordinates of the intersection points $p_1, \dots, p_n$ in the wiring diagram. Braid crossings in the wiring diagram can be viewed as additional intersections that appear in the image of the diagram under the projection $\mathbb{R} \times \mathbb{C} \to \mathbb{R} \times \mathbb{R}$. Choose $\delta > 0$ sufficiently small that there are no crossings in the braided wiring diagram in $\pi_{\mathbb{R}}^{-1}([t_i - 4\delta, t_i + 4\delta])$ (except the intersection at p_n). Let $\rho_i\colon \mathbb{R} \to [0,1]$ be a smooth bump function such that

$$\rho_i(t) = \begin{cases} 1 & \text{for } t \in [t_i - \delta, t_i + \delta], \\ 0 & \text{for } t \notin (t_i - 2\delta, t_i + 2\delta). \end{cases}$$

Let $\rho = \sum_{i=1}^{n} \rho_i$. See Figure 12.

Let $\eta > 0$. Let $\chi\colon \mathbb{R} \to [-\eta, \eta]$ be a smooth function such that

$$\chi(s) = \begin{cases} -\eta & \text{for } s \le -2\eta, \\ s & \text{for } -\frac{1}{2}\eta \le s \le \frac{1}{2}\eta, \\ \eta & \text{for } s \ge 2\eta, \end{cases} \qquad \chi'(s) \ge 0 \quad \text{for all } s \in \mathbb{R}.$$

See Figure 13.

For each wire, we will define its extension to a symplectic surface. Suppose the wire is parametrized as

$$\gamma_j(t) = (t, h_j(t) + i\, w_j(t)) \in \mathbb{R} \times \mathbb{C}.$$

Define $\Gamma_j(t,s)\colon \mathbb{R}^2 \to \mathbb{C}^2$ by

$$\Gamma_j(t,s) = \big(t + i\, s, h_j(t) + i\,(w_j(t) + \rho(t)\chi(s)h_j'(t))\big).$$

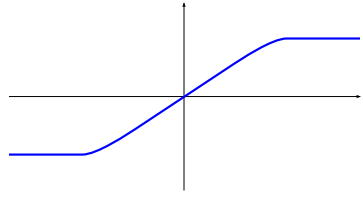


Figure 13: The graph of $\chi\colon \mathbb{R} \to [-\eta, \eta]$.

The tangent space of the image of Γ_j is spanned by

$$\frac{\partial \Gamma_j}{\partial t} = d\Gamma_j\Big(\frac{\partial}{\partial t}\Big) \quad \text{and} \quad \frac{\partial \Gamma_j}{\partial s} = d\Gamma_j\Big(\frac{\partial}{\partial s}\Big).$$

The previous formulas use complex coordinates (x, y) on $\mathbb{C}^2$; now we pass to real coordinates (x_1, x_2, y_1, y_2), so that $x = x_1 + ix_2$, $y = y_1 + iy_2$. In these coordinates, the standard symplectic form is given by $\omega = dx_1 \wedge dx_2 + dy_1 \wedge dy_2$. We have

$$\frac{\partial \Gamma}{\partial t} = \frac{\partial}{\partial x_1} + h_j'(t)\frac{\partial}{\partial y_1} + \big(w'(t) + \rho'(t)\chi(s)h_j'(t) + \rho(t)\chi(s)h_j''(t)\big)\frac{\partial}{\partial y_2},$$
$$\frac{\partial \Gamma_j}{\partial s} = \frac{\partial}{\partial x_2} + \rho(t)\chi'(s)h_j'(t)\frac{\partial}{\partial y_2}.$$

Evaluating the symplectic form gives

$$\omega\left(\frac{\partial \Gamma_j}{\partial t}, \frac{\partial \Gamma_j}{\partial s}\right) = 1 + \rho(t)\chi'(s)(h_j'(t))^2 > 0,$$

so the image of Γ_j is a symplectic surface.

To verify that these extensions do not intersect outside of the original intersections of the wiring diagram, we observe that any intersection between Γ_j and Γ_k would occur at the same parameters (t_0, s_0) and must have

$$h_j(t_0) = h_k(t_0) \quad \text{and} \quad w_j(t_0) + \rho(t_0)\chi(s_0)h_j'(t_0) = w_k(t_0) + \rho(t_0)\chi(s_0)h_k'(t_0).$$

If $h_j(t_0) = h_k(t_0)$, this means that the wires γ_j and γ_k project to the same point under the projection $\mathbb{R} \times \mathbb{C} \to \mathbb{R} \times \mathbb{R}$. This means there is either a crossing or an intersection between wires γ_j and γ_k at t_0.

If t_0 is an intersection point of the wires, $w_j(t_0) = w_k(t_0)$. Additionally, at t_0, the projections of the wires have different slopes, so $h_j'(t_0) \neq h_k'(t_0)$. We also have $\rho(t) \equiv 1$ near t_0. Using this, the intersection assumption that

$$w_j(t_0) + \rho(t_0)\chi(s_0)h_j'(t_0) = w_k(t_0) + \rho(t_0)\chi(s_0)h_k'(t_0)$$

implies that

$$\chi(s_0)(h_k'(t_0) - h_j'(t_0)) = w_j(t_0) - w_k(t_0) = 0.$$

Therefore, $\chi(s_0) = 0$, so $s_0 = 0$ by definition of χ.

If t_0 is a crossing between wires, $w_j(t_0) \neq w_k(t_0)$. Because ρ is supported only near the intersection times, and we assume the crossings occur outside of these intervals, $\rho \equiv 0$. Therefore, the assumption that $w_j(t_0) + \rho(t_0)\chi(s_0)h_j'(t_0) = w_k(t_0) + \rho(t_0)\chi(s_0)h_k'(t_0)$ gives a contradiction.

Finally, we check that Γ_j and Γ_k intersect positively. If we assume that the wiring diagram is in standard form near the intersection points $h_j(t) = k_j t + a_j$ and constant coordinate $w_j(t) \equiv b_j$, then near $(t_i, 0)$ where $\rho(t) \equiv 1$ and $\chi(s) \equiv s$, we have that

$$\Gamma_j(t,s) = (t + is, k_j t + a_j + i(b_j + k_j s)),$$

so the image of Γ_j agrees with the complex line $y = k_j x + a_j + b_j i$, and the intersection of Γ_j and Γ_k locally agrees with an intersection of complex lines. □

5.4 Stein fillings correspond to symplectic configurations

Given a contact structure supported by a planar open book, a theorem of Wendl [65] says that every Stein filling is symplectic deformation equivalent to a Lefschetz fibration with the same planar fiber; Niederkrüger and Wendl [48] extend this result to minimal weak symplectic fillings. Thus, Stein fillings are essentially in one-to-one correspondence with positive factorizations of the monodromy of the given planar open book (and the same is true even for weak symplectic fillings, up to blow-up). The following statement is equivalent to Theorem 1.5.

Proposition 5.6 *Let (Y, ξ) be the link of a rational singularity $(X, 0)$ with reduced fundamental cycle. Fix a decorated germ $(\mathcal{C}, w)$ for $(X, 0)$, with smooth branches $C_1, C_2, \dots, C_m$.*

Then every Stein filling of (Y, ξ) is supported by a Lefschetz fibration built from a configuration of m symplectic disks $\{\Gamma_1, \Gamma_2, \dots, \Gamma_m\}$ in $\mathbb{C}^2$ with marked points, via Lemma 3.2.

Proof Because the contact manifold is planar, any Stein filling is supported by a planar Lefschetz fibration with the same fiber. We will reverse-engineer the required configuration of symplectic disks. Let F_0 be a fixed identification of the planar fiber, where the holes are lined up vertically and labeled by numbers $1, 2, \dots, m$. Let $V_1, \dots, V_n$ be the ordered list of vanishing cycles for the Lefschetz fibration. We begin by producing a collection $(\psi_0, \dots, \psi_{n-1})$ of diffeomorphisms $\psi_i : F_0 \to F_0$ and $(J_1, \dots, J_n)$ of consecutive subsets of $\{1, \dots, m\}$. Here, "consecutive" means that $J_j = \{i, i+1, \dots, i+k\}$ for some i and k.

Choose a diffeomorphism $\beta_0 : F_0 \to F_0$ so that $\beta_0(V_1)$ is isotopic to a curve convexly enclosing a consecutive collection of holes; let J_1 be the corresponding consecutive subset. Let Δ_1 be the counterclockwise half-twist of the convex disk that contains

precisely the holes indexed by J_1. Recursively, choose a diffeomorphism $\beta_j : F_0 \to F_0$ such that $\beta_j \circ \Delta_j^{-1} \circ \cdots \circ \beta_1 \circ \Delta_1^{-1} \circ \beta_0(V_{j+1})$ is isotopic to a curve convexly enclosing a consecutive collection of holes that corresponds to the set J_{j+1}, and let Δ_{j+1} denote the corresponding half-twist.

Consider the braided wiring diagram determined by $(\beta_0, J_1, \beta_1, J_2, \ldots, \beta_{n-1}, J_n)$. By Proposition 5.5, we can construct a configuration of symplectic surfaces $\Gamma_1, \ldots, \Gamma_m$ in $\mathbb{C}^2$ extending this diagram. Using Lemma 3.2, we obtain a planar Lefschetz fibration. We need to use the compact version of the construction to get a fibration whose general fiber is a disk with m holes; for this, we start with a Milnor ball of the form $B = D_x \times D_y$, such that D_x is a neighborhood of η, and D_y is a disk of sufficiently large radius to include the wires above D_x.

As explained in Section 5.2, the vanishing cycles of this Lefschetz fibration will be given by

$$V_j' = (\beta_{j-1} \circ \Delta_{j-1}^{-1} \circ \cdots \circ \beta_1 \circ \Delta_1^{-1} \circ \beta_0)^{-1}(A_j)$$

for $j = 1, \ldots, n$, where A_j is a convex curve enclosing the consecutive holes in the set J_j. The choice of the β_j ensures that these vanishing cycles are identical to our original ones: $V_j' = V_j$.

Along with the symplectic disk configuration $\{\Gamma_1, \ldots, \Gamma_m\}$, we also obtain a collection of marked points on these disks. The marked points include all the intersections as well as additional free marked points, as in Remark 5.4. Each free marked point can be chosen anywhere on the corresponding disk, as long as all marked points are distinct. As in Lemma 3.4, counting multiplicities of pairwise Dehn twists in the monodromy shows that the number of marked points on each disk Γ_j is the same as the weight $w(C_j)$ of the corresponding curvetta C_j of the defining decorated germ (C, w) of the singularity. □

Remark 5.7 The diffeomorphisms β_j are not unique. Any choice will suffice to produce an appropriate braided wiring diagram and corresponding symplectic configuration.

To show that every Stein filling is generated by a symplectic analog of the de Jong–van Straten theorem, it remains to prove that different symplectic configurations with the same monodromy are related by deformations. The role of de Jong and van Straten's picture deformations is played by graphical homotopies.

Proposition 5.8 *Let $(X,0)$ be a rational singularity with reduced fundamental cycle, and (C,w) its decorated plane curve germ with smooth branches $C_1,\dots,C_m$. Let (Y,ξ) be the contact link of $(X,0)$. Suppose that $\Gamma=\{\Gamma_1,\Gamma_2,\dots,\Gamma_m\}$ is a configuration of symplectic disks with marked points $p_1,\dots,p_n$, constructed for a given Stein filling of (Y,ξ) as in Proposition 5.6. Then $(\Gamma,\{p_j\})$ can be connected to (C,w) by a smooth graphical homotopy.*

Lemma 5.9 *Suppose $C_1^0,\dots,C_m^0$ and $C_1^1,\dots,C_m^1$ are two configurations of graphical disks in a Milnor ball $B=D_x\times D_y$, such that $\partial C_j^0=\partial C_j^1$ for $j=1,\dots,m$. Then there is a family of graphical disks $C_1^t,\dots,C_m^t$ (potentially with negative intersections) interpolating between these two configurations with fixed boundary link $\partial C_1^t\cup\cdots\cup\partial C_m^t\subset\partial B$. Here, $\partial C_j^t=C_j^t\cap\partial B=C_j^t\cap(\partial D_x\times D_y)$.*

Proof Because we are not limiting the behavior of the intersections of the components, it suffices to check that there is a family C_j^t interpolating between C_j^0 and C_j^1 for one component. For simplicity of notation we will drop the j. For this, because both C^0 and C^1 are graphical, we can write them as $C^s=\{(x,f^s(x))\}$ for $s=0,1$. Then since $\partial C^0=\partial C^1$, we have that $f^0(x)=f^1(x)$ for $x\in\partial D_x$. Let $C^t=\{(x,tf^1(z)+(1-t)f^0(x)\}$. Then C^t interpolates smoothly between C^0 and C^1, and its boundary is fixed. □

Lemma 5.10 *Suppose $C_1\cup\cdots\cup C_m$ is a configuration of graphical disks, so its boundary $\partial C_1\cup\cdots\cup\partial C_m$ is a braid. Let $L_1,\dots,L_m$ be the components of a braid $L_1\cup\cdots\cup L_m$ which is braid isotopic (with corresponding indices) to $\partial C_1\cup\cdots\cup\partial C_m$. Then there is a homotopy of graphical disks $C_1^t,\dots,C_m^t$ such that $C_j^0=C_j$ and $\partial C_j^1=L_j$.*

Proof If $C_1,\dots,C_m$ are graphical over a disk D_x, choose a larger disk D_x' containing D_x. Then we can extend $C_1,\dots,C_m$ to graphical disks $C_1',\dots,C_m'$ over D_x' so that $\partial C_1',\dots,\partial C_m'$ is the braid L_j, by realizing the trace of the braid isotopy over the annulus $D_x'\setminus D_x$. Next, we can shrink D_x' to D_x continuously via a family of embeddings $\phi_t\colon D_x'\to D_x'$ where $\phi_0=\mathrm{id}$, $\phi_1(D_x')=D_x$, and ϕ_1 identifies points in $\partial D_x'$ with points in ∂D_x according to the same identification used to realize the trace. Then if $C_j'=\{(x,f_j(x))\}$ for $x\in D_x'$, we can let

$$C_j^t=\{(\phi_t(x),f_j(x)\mid x\in D_x'\}\cap(D_x\times\mathbb{C}).$$

Then $C_j^0=C_j$ and $\partial C_j^1=L_j$, as required. □

Proof of Proposition 5.8 When we fix the germ (C, w) and apply the method of Proposition 5.6 to a given Stein filling for (Y, ξ), we first consider the open book on (Y, ξ) induced by the decorated germ as in Proposition 4.1. The Stein filling then carries a Lefschetz fibration that induces the same open book on the boundary, and the arrangement $(\Gamma, \{p_j\})$ is constructed from the monodromy of this Lefschetz fibration. The smooth disks $\Gamma_1, \dots, \Gamma_m$ are contained in the Milnor ball B for $\mathcal{C}$ and are transverse to its boundary S^3, so that $S^3 \cap (\Gamma_1 \cup \cdots \cup \Gamma_m)$ is a braid. By Lemma 3.3, the monodromy of this braid is the image of the monodromy of the open book under the projection $\mathrm{MCG}(P_m) \to \mathrm{MCG}(\mathbb{C}_m)$ of the mapping class group of the compact disk with holes to the mapping class group of the punctured plane, so the two braids are braid-isotopic. Therefore, we can apply Lemma 5.10 to perform a graphical homotopy to $\Gamma_1, \dots, \Gamma_m$ so that its boundary agrees with that of $C_1, \dots, C_m$. Next, apply Lemma 5.9 to continue the graphical homotopy from $C_1, \dots, C_m$ to $\Gamma_1, \dots, \Gamma_m$. □

Remark 5.11 For our construction of a Lefschetz fibration, it is not important that the C_i^t are symplectic disks, we only care that they are graphical. However, by performing a rescaling in the y direction, we can ensure that all of the graphical disks are symplectic if the partial derivatives of the function f are sufficiently small. More specifically, if $C = \{(x, f(x))\}$, where $x = x_1 + ix_2$ and

$$\left|\frac{\partial f}{\partial x_1}\right|, \left|\frac{\partial f}{\partial x_2}\right| < \frac{\sqrt{2}}{2},$$

then C will be symplectic. This bound is sufficient although not necessary; it can be achieved by rescaling f, which itself is a graphical homotopy. Moreover, if f^0 and f^1 both satisfy these bounds, then their convex combination $tf^0 + (1-t)f^1$ also satisfies the bound for all $t \in [0, 1]$, so the interpolation between the two disks will also be symplectic.

6 Incidence matrix and topology of fillings

6.1 Basic topological invariants

It is shown in [27] that the basic topological invariants of the Milnor fibers obtained from the picture deformations can be easily computed from the deformed curvetta arrangement. Moreover, the incidence matrix of the arrangement can be reconstructed from the Milnor fiber [44]. We now review these facts briefly and adapt and generalize

them in our context: the goal is to show that exactly the same results hold for more general Stein fillings, constructed from smooth disk arrangements as in Section 5.

As we have shown in Section 5, every Stein filling W can be described by an arrangement $\Gamma = \{\Gamma_i\}$ of symplectic curvettas with marked points $\{p_j\}_{j=1}^n$, related to the plane curve germ $\mathcal{C} = C_1 \cup \cdots \cup C_m$ by a smooth graphical homotopy. We always assume that curvettas intersect positively. We also treat the components of $\mathcal{C}$ as labeled, so the ordering of components $C_1, \ldots, C_m$ is fixed. The set of marked points $\{p_j\}_{j=1}^n$ contains all intersection points between the Γ_i and possibly a number of free points. The *incidence matrix* $\mathcal{I}(\Gamma, \{p_j\})$ has m rows and n columns, defined so that its entry a_{ij} at the intersection of the i^{th} row and the j^{th} column equals 1 if $p_j \in \Gamma_i$, and 0 otherwise. Note that there is no canonical labeling of the points p_j, so the incidence matrix is defined only up to permutation of columns. We will say that two arrangements $(\Gamma, \{p_j\})$ and $(\Gamma', \{p'_j\})$ are *combinatorially equivalent* if their incidence matrices coincide (up to permutation of columns, ie up to relabeling of the marked points).

Let $\mathcal{L}$ be the Lefschetz fibration constructed for the arrangement $(\Gamma, \{p_j\})$ as in Lemma 3.2. Its general fiber is a disk with m holes that correspond to the curvettas $\Gamma_1, \ldots, \Gamma_m$ of Γ; in particular, the number of holes equals the number of rows in the matrix $\mathcal{I}(\Gamma, \{p_j\})$. The vanishing cycles of $\mathcal{L}$ correspond to the marked points $\{p_j\}_{j=1}^n$ and enclose sets of holes that correspond to curvettas passing through that point: if $\Gamma_{i_1}, \ldots, \Gamma_{i_k}$ are all curvettas that intersect at p_j, the vanishing cycle V_j encloses the holes $h_{i_1}, \ldots, h_{i_k}$. It follows that *homology classes* of the vanishing cycles of $\mathcal{L}$ can be determined from the incidence matrix $\mathcal{I}(\Gamma, \{p_j\})$, and we have:

Proposition 6.1 *Let $\mathcal{L}$ be the Lefschetz fibration for the arrangement $(\Gamma, \{p_j\})$ with incidence matrix $\mathcal{I}(\Gamma, \{p_j\})$. If the j^{th} column of $\mathcal{I}(\Gamma, \{p_j\})$ has 1s in rows $i_1, i_2, \ldots, i_k$, the corresponding vanishing cycle V_j of $\mathcal{L}$ encloses the holes $h_{i_1}, \ldots, h_{i_k}$ in the fiber.*

Corollary 6.2 *Let $(\Gamma, \{p_j\})$ and $(\Gamma', \{p'_j\})$ be two combinatorially equivalent arrangements, and $\mathcal{L}$ and $\mathcal{L}'$ the corresponding Lefschetz fibrations. Then the vanishing cycles of $\mathcal{L}$ and $\mathcal{L}'$ are in one-to-one correspondence, so that the two vanishing cycles that correspond to one another are given by homologous curves in the fiber.*

Because smooth graphical homotopies do not allow intersections to escape through the boundary, the number of pairwise intersections of Γ_i and Γ_j is given by $\operatorname{tang}(C_i, C_j) = \rho(v_i, v_j; v_0)$; see Remark 2.6. The weight of Γ_i (the total number of intersection points and the free marked points on Γ_i) is given by $w(C_i) = 1 + l(v_0, v_i)$. The intersections

between Γ_i and Γ_j correspond to the points among $p_1, p_2, \dots, p_n$ contained in both lines, and each such point gives a 1 in the same column for the i^{th} row and the j^{th} row of the incidence matrix. Therefore we have:

Lemma 6.3 *Let $(\mathcal{C}, w)$ be a decorated germ corresponding to $(X, 0)$, with branches $C_1, C_2, \dots, C_m$. Consider any arrangement $\{\Gamma_i\}_{i=1}^m$ of smooth curvettas encoding a Stein filling of the link of $(X, 0)$. The incidence matrix $\mathcal{I}(\Gamma, \{p_j\})$ has the following properties*:

(i) *The number of 1s in the i^{th} row of $\mathcal{I}(\Gamma, \{p_j\})$ is $w(C_i) = 1 + l(v_0, v_i)$.*

(ii) *The number of 1s which appear in the same columns for the i^{th} row and the j^{th} row is $\operatorname{tang}(C_i, C_j) = \rho(v_i, v_j; v_0)$.*

Here, $l(v_0, v_i)$ and $\rho(v_i, v_j; v_0)$ are the length and overlap functions on the resolution graph G, defined in Remark 2.6, and v_0 is the choice of root.

We now describe how the incidence matrix $\mathcal{I}(\Gamma, \{p_j\})$ determines basic algebraic topology of the filling W, namely $H_1(W)$, $H_2(W)$, the intersection form of W, and the first Chern class $c_1(J)$ of the Stein structure. (Homology is taken with $\mathbb{Z}$ coefficients throughout.) The statements about the homology and the intersection form of W are proved in [27, Section 5] for the algebraic case, but the proofs are entirely topological and apply in the more general settings as well. Alternatively, the same invariants can be computed from the vanishing cycles of the Lefschetz fibration [6, Lemma 16]. For Lefschetz fibrations with planar fiber, detailed proofs for the intersection form and $c_1(J)$ calculations are given in [19]. We write $\mathbb{Z}\langle\{p_j\}\rangle$ for the free abelian group generated by $\{p_j\}_{j=1}^n$, and $\mathbb{Z}\langle\{\Gamma_i\}\rangle$ is defined similarly. The incidence matrix $\mathcal{I}(\Gamma, \{p_j\})$ defines a map between the corresponding lattices.

Proposition 6.4 *There is a short exact sequence*

$$0 \to H_2(W) \to \mathbb{Z}\langle\{p_j\}\rangle \xrightarrow{\mathcal{I}} \mathbb{Z}\langle\{\Gamma_i\}\rangle \to H_1(W) \to 0.$$

Proof Let W be the total space of a Lefschetz fibration over a disk D, with planar fiber P. (We always assume that W, P and D are compatibly oriented.) If $D' \subset D$ is a small disk that contains no critical points, then W is obtained from $P \times D'$ by attaching 2–handles to copies of the vanishing cycles contained in the vertical boundary $P \times \partial D'$, so that distinct handles are attached along knots contained in distinct fibers.

We use the exact sequence of the pair $(W, P \times D')$; since $P \times D'$ retracts onto P, we can replace the former with the latter. Notice also that $H_1(W, P) = 0$, so we get

$$0 \to H_2(W) \xrightarrow{j_*} H_2(W, P) \xrightarrow{\partial_*} H_1(P) \to H_1(W) \to 0.$$

The group $H_2(W, P)$ is freely generated by the cores of the attached 2–handles; we can identify these generators with the vanishing cycles. By construction of the Lefschetz fibration, each vanishing cycle corresponds to a blow-up at some marked point, so we can identify the vanishing cycles with the set $\{p_j\}$. The free abelian group $H_2(W, P)$ is then identified with the lattice $\mathbb{Z}\langle\{p_j\}\rangle$. The generators for the free abelian group $H_1(P)$ can be given by loops around the holes in the planar fiber. The holes correspond to the branches of $\mathcal{C}$, thus $H_1(P)$ can be identified with the lattice $\mathbb{Z}\langle\{\Gamma_i\}\rangle$. The map ∂_* is evaluated as follows: to compute $\partial_*(p_j)$, we take the boundary of the core of the corresponding 2–handle, given by the vanishing cycle associated with p_j, and express this vanishing cycle in terms of the generators of $H_1(P) = \mathbb{Z}\langle\{\Gamma_i\}\rangle$. Since the vanishing cycle is a simple closed curve on the planar page, its first homology class equals the sum of the boundaries of the holes it encloses, which in turn correspond to the branches Γ_i passing through p_j. Therefore, $\partial_*(p_j)$ is given precisely by the j^{th} column of the incidence matrix $\mathcal{I}(\Gamma, \{p_j\})$, as required. □

Remark 6.5 Since the link Y of a rational singularity $(X, 0)$ is always a rational homology 3–sphere, a standard argument shows that $b_1(W) = 0$ for any Stein filling W of Y. Indeed, W has no 3–handles, so $H^3(W; \mathbb{Q}) = 0$; then for the pair $(W, Y) = (W, \partial W)$ we have

$$0 = H_1(\partial W; \mathbb{Q}) \to H_1(W; \mathbb{Q}) \to H_1(W, \partial W; \mathbb{Q}) \cong H^3(W; \mathbb{Q}) = 0.$$

It follows that the matrix $\mathcal{I}(\Gamma, \{p_j\})$ always has full rank.

Note that $H_2(W)$ is isomorphic to $\operatorname{Im} j_*$, which in turn equals $\ker \partial_*$. So $H_2(W)$ can be identified with null-homologous linear combinations of vanishing cycles (thought of as 1–chains in P). One can explicitly describe an oriented embedded surface in W representing a given second homology class, as follows [19, Section 2]. First, one constructs an oriented embedded surface in $P \times D'$ whose boundary is the given null-homologous linear combination of the vanishing cycles, and then the vanishing cycles are capped off in W. A similar construction is given in [27] without Lefschetz fibrations, for Milnor fibers obtained by blowing up the 4–ball at the marked points and taking the complement of the proper transforms of curvettas; exactly the same

argument works for a smooth curvetta arrangement $(\Gamma, \{p_j\})$. After blowing up the 4–ball B at the points $p_1, p_2, \dots, p_n$, we have the 4–manifold $\widetilde{B}$, the blow-up of B, with generators of $H_2(\widetilde{B})$ given by the fundamental classes E_{p_i} of the exceptional divisors. We identify $H_2(\widetilde{B}) = \mathbb{Z}\langle\{p_j\}\rangle$. The intersection form of $\widetilde{B}$ is standard negative definite in the given basis, as $E_p \cdot E_p = -1$. The manifold W is obtained from $\widetilde{B}$ by removing the tubular neighborhoods T_i of the proper transforms $\widetilde{\Gamma}_i$ of the curvettas Γ_i. The inclusion induces a map $H_2(W) \to H_2(\widetilde{B})$, which is in fact the same map as j_* above, under obvious identifications. Every homology class in $H_2(W)$ is represented by an embedded oriented surface which can be constructed by taking the collection of the corresponding exceptional spheres E_{p_i}, punctured at their intersections with $\widetilde{\Gamma}_j$, and connected by tubes running inside the cylinders T_i. The intersection of two such surfaces can be computed by taking the intersections of the corresponding collections of exceptional spheres, as the tubes can be arranged to be disjoint. For the Stein structure J on W associated to the given Lefschetz fibration, we can compute $c_1(J)$ using the same inclusion $H_2(W) \to H_2(\widetilde{B})$. Indeed, J is homotopic to the restriction of the complex structure j on $\widetilde{B}$, and $c_1(j)[E_{p_i}] = 1$ for every E_{p_i}. Therefore we have:

Proposition 6.6 *The intersection form on $H_2(W) \subset \mathbb{Z}\langle\{p_j\}\rangle$ is the restriction of the standard negative definite form given by $p_i \cdot p_j = -\delta_{ij}$ for $i, j = 1, \dots, n$. The first Chern class $c_1(J)$ of the Stein structure is the restriction of the linear form on $\mathbb{Z}\langle\{p_j\}\rangle$ given by $c_1[p_i] = 1$ for $i = 1, \dots, n$.*

See also [19, Propositions 2.1 and 2.4] for a detailed calculation (in terms of the vanishing cycles) of the intersection form and $c_1(J)$ for an arbitrary Lefschetz fibration (W, J) with planar fiber.

6.2 Uniqueness of the Artin filling and proof of Theorem 1.2

In general, the topology of the filling might not be fully determined by the incidence matrix of the corresponding curvettas arrangement; Proposition 6.1 gives the homology classes of the vanishing cycles but not their isotopy classes. However, it turns out that the incidence matrix completely determines the smoothing for picture deformations that are combinatorially equivalent to the Scott deformation, so that one gets the Artin smoothing component [27, Cases 4.13]. We prove that an analogous result holds for Stein fillings as well. Note that the argument in [27] uses simultaneous resolutions and

only works in the algebraic setting, while we work with mapping class groups instead. Our argument works because the Artin filling has a Lefschetz fibration with *disjoint* vanishing cycles in the fiber.

Proposition 6.7 *Let $(X,0)$ be a rational surface singularity with reduced fundamental cycle, with contact link (Y,ξ) and decorated germ $(\mathcal{C},w)$. Let Γ be an arrangement of smooth graphical curves with positive intersections and marked points $\{p_j\}$, related to the germ $(\mathcal{C},w)$ by a smooth graphical homotopy, so that $(\Gamma,\{p_j\})$ gives rise to a Stein filling W of (Y,ξ).*

Suppose that $(\Gamma,\{p_j\})$ is combinatorially equivalent to the Scott deformation $(\mathcal{C}^s,w^s)$ of $(\mathcal{C},w)$. Then the Stein filling given by $(\Gamma,\{p_j\})$ is Stein deformation-equivalent to the Artin filling of (Y,ξ).

Proof Let $\mathcal{L}$ be the Lefschetz fibration for $(\Gamma,\{p_j\})$, constructed as in Lemma 3.2, and let $\mathcal{L}_A$ be the Lefschetz fibration for the Artin smoothing, given in Proposition 4.2. We know that $\mathcal{L}_A$ is given by the monodromy factorization as in Proposition 4.1; let ϕ denote the monodromy of the open book as in the lemma.

Both fibrations $\mathcal{L}$ and $\mathcal{L}_A$ have the same fiber S, and the fibration $\mathcal{L}$ corresponds to some factorization of the same monodromy ϕ. By Corollary 6.2, the vanishing cycles $\{V_j\}$ and $\{V_j^A\}$ of the two fibrations are in one-to-one correspondence, so that the curves V_j and V_j^A are homologous in the fiber. We need to show that V_j and V_j^A are isotopic.

There are two types of vanishing cycle in the fibration $\mathcal{L}_A$: (1) boundary-parallel curves that enclose a single hole each, and (2) the curves that go around the necks connecting the spheres, as shown at the top of Figure 6. The isotopy class of a boundary-parallel curve in the fiber is uniquely determined by its homology class, so if V_j^A is boundary-parallel, then $V_j = V_j^A$. Now, because the total monodromy of $\mathcal{L}$ and $\mathcal{L}_A$ is the same, and the Dehn twists around the boundary-parallel curves are in the center of the mapping class group of the fiber, we see that the products of the Dehn twists around the vanishing cycles homologous to necks are the same for both $\mathcal{L}$ and $\mathcal{L}_A$. In other words, if N denotes the set of vanishing cycles homologous to necks, we have

$$\prod_{V_j\in N}\tau_{V_j}=\prod_{V_j^A\in N}\tau_{V_j^A}. \tag{6-1}$$

Let ψ denote the diffeomorphism of the fiber given by the product (6-1).

To prove that each vanishing cycle V_j is indeed isotopic to the vanishing class V_j^A homologous to V_j, we proceed by induction on the number of necks in the fiber S (this is the same as the number of edges in the dual resolution graph G). Equivalently, we can induct on the number of vertices, since G is a tree. When G has only one vertex, there are no necks, so all the vanishing cycles are boundary-parallel, and $V_j = V_j^A$ for all pairs of vanishing cycles. Assume that the claim is established for all graphs with k vertices or fewer. Consider a graph G with $k+1$ vertices and pick a leaf vertex v of G. We will be able to remove v to reduce the question to a graph G' with k vertices.

In the Lefschetz fibration of Proposition 4.2, the leaf v corresponds to the sphere S_v with holes, connected to the rest of the fiber S by a single neck. The fibration $\mathcal{L}_A$ has a vanishing cycle V^A that goes around this neck, and $\mathcal{L}$ has a vanishing cycle V in the same homology class. Since v is a leaf, S_v is separated from its complement $S \setminus S_v$ by the curve V^A. Observe that all the other nonboundary parallel vanishing cycles of $\mathcal{L}_A$ lie outside S_v. A priori, nonboundary parallel vanishing cycles of $\mathcal{L}$ may belong to different isotopy classes and intersect S_v; we want to show that they can be isotoped to lie outside S_v.

If the self-intersection $v \cdot v = -2$, then in fact V^A encloses only one hole, so it is boundary-parallel, and we can immediately conclude that V and V^A are isotopic, and S_v is a boundary-parallel annulus disjoint from all the other vanishing cycles.

Suppose now that $v \cdot v \leq -3$, so that V^A encloses $r = -1 - v \cdot v > 1$ holes. Connect these holes by $r-1$ disjoint arcs $\alpha_1, \dots, \alpha_{r-1}$ in the sphere S_v, so that if the fiber S is cut along these arcs, the r holes will become a single hole; see Figure 14.

By construction, the arcs $\alpha_1, \dots, \alpha_{r-1}$ are disjoint from all nonboundary parallel vanishing cycles V_j^A of $\mathcal{L}_A$. It follows that each α_i is fixed by the diffeomorphism ψ. As in [7, Proposition 3] and [17, Section 2], we now make the following key observation: after an isotopy removing nonessential intersections, all arcs $\alpha_1, \dots, \alpha_{r-1}$ must be also disjoint from all non-boundary-parallel vanishing cycles V_j of $\mathcal{L}$. To see this, we recall that each right-handed Dehn twist is a right-veering diffeomorphism of the oriented surface S [25]. If α and β are two arcs with the same endpoint $x \in \partial S$, we say that β lies to the right of α if the pair of tangent vectors $(\dot{\beta}, \dot{\alpha})$ at x gives the orientation of S. The right-veering property of a boundary-fixing map $\tau\colon S \to S$ means that for every simple arc α with endpoints on ∂S, the image $\tau(\alpha)$ is either isotopic to α or lies to the right of α at both endpoints, once all nonessential intersections between α and $\tau(\alpha)$ are removed. Now, suppose that $\mathcal{L}$ has a vanishing cycle $V_j \in N$ that

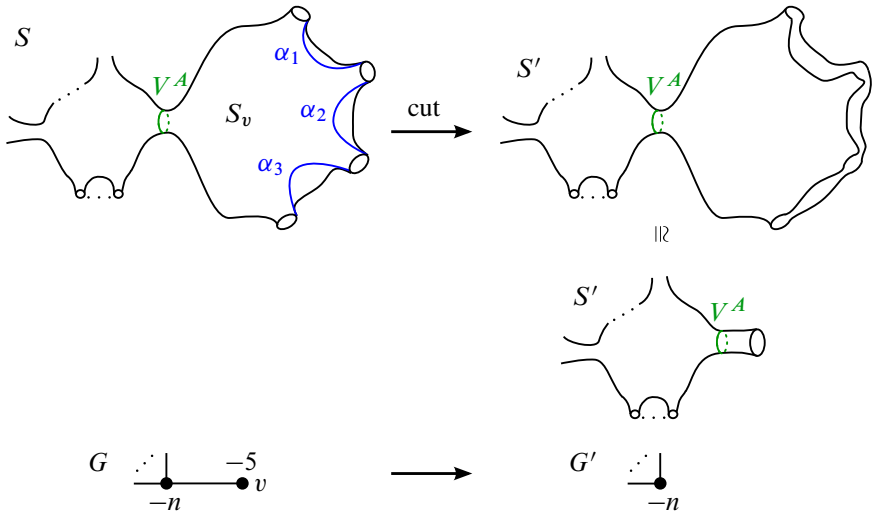


Figure 14: After cutting the fiber S, the vanishing cycle V^A becomes boundary-parallel in S'.

essentially intersects one of the arcs, say α_1. Then the curve $\tau_{V_j}(\alpha_1)$ is not isotopic to α (see eg [16, Proposition 3.2]), so $\tau_{V_j}(\alpha_1)$ lies to the right of α_1. Since the composition of right-veering maps is right-veering, we can only get curves that lie further to the right of α after composing with the other nonboundary parallel vanishing cycles of $\mathcal{L}$. However, the composition $\psi = \prod_{V_j \in N} \tau_{V_j}$ fixes α_1, a contradiction.

Once we know that no vanishing cycles of $\mathcal{L}$ or $\mathcal{L}_A$ intersect any of the arcs $\alpha_1, \dots, \alpha_{r-1}$, we can cut the fiber S along these arcs, and consider the image of the relation (6-1) in the resulting cut-up surface S'. In S', V^A becomes a boundary-parallel curve, and since V lies in the same homology class, we see that V and V^A are isotopic in S' (and therefore in S). We then have

$$\prod_{V_j \in N,\, V_j \neq V} \tau_{V_j} = \prod_{V_j^A \in N,\, V_j^A \neq V^A} \tau_{V_j^A}.$$

Now observe that cutting up S along the arcs as above has the same effect as removing the sphere S_v with its neck from the set of subsurfaces forming the fiber S in Proposition 4.2. Then the cut-up fiber S' with its non-boundary-parallel vanishing cycles $\{V_j\}$ and $\{V_j^A\}$ corresponds to the fibrations for the graph G' obtained by deleting the leaf v and its outgoing edge from the graph G. By the induction hypothesis, we can conclude that all pairs of homologous vanishing cycles V_j, V_j^A are isotopic in S', and

thus in S. It follows that the Lefschetz fibrations $\mathcal{L}$ and $\mathcal{L}_A$ are equivalent, and therefore the Stein filling given by $\mathcal{L}$ is Stein deformation equivalent to the Artin filling. □

The above results have the following interesting application, related to conjectures of Kollár on deformations of rational surface singularities. Suppose that a rational singularity $(X, 0)$ has a dual resolution graph G such that $v \cdot v \leq -5$ for every vertex $v \in G$. In this case, Kollár's conjecture asserts that the base space of a semiuniversal deformation of X has just one component, the Artin component; in particular, there is a unique smoothing, up to diffeomorphism. In the special case of reduced fundamental cycle, this conjecture was proved by de Jong and van Straten via their picture deformations method. We establish the symplectic version of this result, proving Theorem 1.2.

Proof of Theorem 1.2 We can focus on Stein fillings: by [65] and [48], every weak symplectic filling of a planar contact manifold is a blow-up of a Stein filling, up to symplectic deformation. By Section 5, Stein fillings are given by arrangements of symplectic curvettas. The argument in [27, Theorem 6.23] shows that under the given hypotheses on the resolution of $(X, 0)$, there is a unique *combinatorial* solution to the smoothing problem, namely, any arrangement of curvettas must have the same incidence matrix as the Artin incidence matrix given by the Scott deformation. The argument of De Jong and van Straten is somewhat involved, so we will not summarize it here, but we emphasize that the proof of this fact is completely combinatorial and does not use the algebraic nature of arrangements. The same claim holds for an arbitrary smooth arrangement subject to the same hypotheses. The only input used in [27] is the properties of the incidence matrix determined by the resolution graph as in Lemma 6.3, together with the following observation: if all vertices of the resolution graph G have self-intersection -5 or lower, each end vertex of G (except the root) gets at least three (-1) vertices attached in the augmented graph G', so that there are at least three corresponding curvettas. An important step in the inductive proof is that the matrix must have a column where all entries are 1, ie all the Γ_i must have a common point.

Once we know that all arrangements corresponding to possible Stein fillings are combinatorially equivalent to the arrangement given by the Scott deformation, Theorem 1.2 follows from Proposition 6.7. □

In the case where, additionally, the graph G is star-shaped with three legs, uniqueness of minimal symplectic filling (up to symplectomorphism and symplectic deformation) was proved by Bhupal and Stipsicz [9]. (They give a detailed proof under the hypothesis

that the self-intersection of the central vertex is at most -10, but mention that one can go up to -5 with similar techniques.) Their method relies on McDuff's theorem [38] and was previously used by Lisca [34]: one finds a concave symplectic cap which is a plumbing of spheres that completes an arbitrary filling to a rational surface, which must be a blow-up of $\mathbb{C}\mathrm{P}^2$, analyzes possible configurations of (-1) curves, and then verifies that the configurations in the image of the cap plumbing under the blow-down is a pencil of symplectic lines which has a unique symplectic isotopy class. To our knowledge, this strategy has not been applied to non-star-shaped graphs in the existing literature. The difficulty in the non-star-shaped case is that there is not an obvious concave symplectic plumbing which can serve as a cap. Our proof works for completely arbitrary trees.

6.3 Distinguishing Stein fillings

We now turn to constructions that will be needed in the next section, and explain how to use incidence matrices to distinguish Stein fillings, at least relative to certain boundary data. Indeed, as shown by Némethi and Popescu-Pampu [44], the incidence matrix is "remembered" by the Milnor fiber of the corresponding smoothing, which allows us to show that certain Milnor fibers are not diffeomorphic (in the strong sense, ie relative to a boundary marking). The argument in [44] is purely topological, so we can generalize it to arbitrary Stein fillings. While [44] applies more generally to sandwiched singularities, we only consider the case of reduced fundamental cycle.

Instead of the boundary marking used in [44], we will keep track of the boundary data via a choice of a compatible embedded open book for (Y,ξ). As in Section 2, we fix a choice of extension G' of the dual resolution graph G of a singularity with link (Y,ξ), to fix the topological type of the associated decorated germ $(\mathcal{C},w)$ with labeled branches $C_1,\dots,C_m$. Each branch C_j corresponds to a hole h_j of the open book as, explained in Section 4; fixing the embedded open book, up to isotopy, is equivalent to fixing the topological type of the decorated germ. In fact, this open book decomposition provides the data of the "markings" of [44], where the solid tori components of the binding correspond to "pieces" of the marking data which allow one to fix the gluing of the smooth cap of [44] to the filling using the open book instead of the markings.

By Wendl's theorem [65], all Stein fillings of a planar contact 3–manifold are given, up to symplectic deformation, by Lefschetz fibrations with same fiber, so that these fibrations are encoded by monodromy factorizations of the fixed open book as above. Suppose

that Stein fillings W and W' arise from symplectic curvetta arrangements $(\Gamma, \{p_j\})$ and $(\Gamma', \{p'_j\})$ as in Propositions 5.6 and 5.8. On the boundaries ∂W and $\partial W'$, these arrangements induce open books which are isomorphic, because both are isomorphic to the open book induced by the germ $(\mathcal{C}, w)$. Fix these two open books, $\mathcal{OB}$ on ∂W and $\mathcal{OB}'$ on $\partial W'$, defined up to isotopy; as part of the open book data, we also label the binding components (with the exception of the outer boundary of the disk, the boundary components of the page correspond to the branches of the decorated germ).

We will say that W and W' are *strongly diffeomorphic* if there is an orientation-preserving diffeomorphism $W \to W'$ whose restriction to ∂W maps the open book $\mathcal{OB}$ on ∂W to an open book on $\partial W'$ which is isotopic to the given one, $\mathcal{OB}'$. If the open book on $\partial W'$ is isotopic to the image of the open book on ∂W, we can compose the diffeomorphism $W \to W'$ with a self-diffeomorphism of W' which extends the isotopy of $\partial W'$ to obtain a diffeomorphism matching the open books. Therefore, we can equivalently say that W and W' are *strongly diffeomorphic* if there is an orientation-preserving diffeomorphism $W \to W'$ that identifies the open books $\mathcal{OB}$ on ∂W and $\mathcal{OB}'$ on $\partial W'$. This identification is required to preserve the labeling of the binding components. (We will discuss a slightly weaker condition in Remark 6.9.)

Rephrasing the theorem of [44] in our context, we have:

Proposition 6.8 [44, Theorem 4.3.3] *Let (Y, ξ) be the contact link of a rational singularity with reduced fundamental cycle, and fix the isotopy class of an embedded open book as above. Let two strongly diffeomorphic Stein fillings W and W' arise from arrangements $(\Gamma, \{p_j\})$ and $(\Gamma', \{p'_j\})$ of symplectic curvettas with marked points, as in Section 5. Then the incidence matrices $\mathcal{I}(\Gamma, \{p_j\})$ and $\mathcal{I}(\Gamma', \{p'_j\})$ are equal, up to permutation of columns.*

Proof We outline the proof briefly, referring the reader to [44] for details, as we use exactly the same topological argument in a slightly different (in fact, simpler) context.

Let $(\mathcal{C}, w)$ be the decorated germ with labeled smooth branches $C_1, \dots, C_m$, determined up to topological equivalence by the open book data for (Y, ξ). Unlike [44], we only work with the case of smooth components of $\mathcal{C}$; therefore, all δ–invariants of the branches C_i are 0, and the formulas of [44] become simpler.

As in [44], we construct a cap U, which is a smooth manifold with boundary that can be attached to any Stein filling W of (Y, ξ), so that $W \cup U$ is a blow-up of a 4–sphere. To construct U, let $B \subset \mathbb{C}^2$ be a closed Milnor ball as in Section 3, so that

B contains both the branches of the germ $\mathcal{C}$ and the arrangement Γ together with all intersection points between curvettas Γ_i. Let $(B', \mathcal{C}')$ be another copy of this ball with the germ $\mathcal{C}$ inside, with reversed orientation. After an isotopy of the boundaries of the curvettas Γ_i to match ∂C_i, we can glue (B, Γ) and $(B', \mathcal{C}')$ so that the boundary of Γ_i is glued to the boundary of the corresponding germ branch C_i'. Each disk Γ_i is oriented as a graph over $\mathbb{C}$, so the result of gluing is a smooth 4–sphere $B \cup B'$ containing the embedded smooth 2–spheres $\Sigma_i = \Gamma_i \cup C_i'$. Blowing up at the points $p_1, \dots, p_n$, we get $\#_{i=1}^n \overline{\mathbb{C}\mathrm{P}}^2$, represented as the blow-up $\widetilde{B}$ of the ball B glued to B'. Let T_i be a thin tubular neighborhood of the proper transform of Γ_i in $\widetilde{B}$. By Lemmas 3.2 and 3.4, we have $W = \widetilde{B} \setminus \bigcup_{i=1}^m T_i$. Set $U = B' \bigcup_{i=1}^m T_i$, so that we have $U \cup W = \#_{i=1}^n \overline{\mathbb{C}\mathrm{P}}^2$. As in [44, Lemma 4.2.4], the cap U is independent of W and is determined by the boundary data. Indeed, to form U, we attach 2–handles to the 4–ball B'. The attaching circles are given by the boundaries of the Γ_i, and the link $\bigcup_i \partial\Gamma_i$ is isotopic to the link given by the boundaries of the branches of the original decorated germ. The framing for $\partial\Gamma_i$ is $-w_i$, the negative weight on the branch C_i of the decorated germ. The proof of Lemma 3.4 shows that the weight w_i is given by the number of Dehn twists enclosing the i^{th} hole in (any decomposition of) the monodromy of the open book. Thus, the cap U and the way it is glued to W is determined by the decorated germ defining the singularity, together with the fixed open book data of (Y, ξ). Finally, as in [44], we see that there is a unique basis $\{e_j\}_{j=1}^n$ for $H_2\big(\#_{i=1}^n \overline{\mathbb{C}\mathrm{P}}^2\big)$ of classes of square -1 such that the intersection numbers $\Sigma_i \cdot e_j$ are all positive. It follows that these numbers depend only on W and the open book data. On the other hand, the numbers $\Sigma_i \cdot e_j$ form the incidence matrix $\mathcal{I}(\Gamma, \{p_j\})$, as $\Sigma_i \cdot e_j = 1$ if $p_j \in \Gamma_i$, and 0 otherwise. It follows that the incidence matrices $\mathcal{I}(\Gamma, \{p_j\})$ and $\mathcal{I}(\Gamma', \{p_j'\})$ are the same, up to relabeling the marked points, which amounts to permutation of columns. □

Remark 6.9 Our definition of a strong diffeomorphism and the above proof assumes that the binding components of the open book are labeled, and that the diffeomorphism preserves this labeling. In other words, we think of the page of the open book(s) as a disk with holes, where each hole h_i corresponds to the i^{th} branch of the fixed decorated germ; the diffeomorphism matches the i^{th} hole of the page for ∂W to the i^{th} hole for $\partial W'$. It is in fact possible to consider a less restrictive definition of strong diffeomorphism, by allowing permutations of binding components, and to prove a sightly stronger version of Proposition 6.8 and Theorem 7.8. More precisely, the proposition still holds if there is a diffeomorphism $f : W \to W'$ that sends the chosen open book $\mathcal{OB}$ on ∂W to an open book on $\partial W'$ which is isotopic to $\mathcal{OB}'$, in the sense

of isotoping the binding and the pages, but the isotopy matches the binding components in a wrong order. Moreover, it is plausible that the proposition still holds if we only have a diffeomorphism $W \to W'$ whose restriction to ∂W takes the binding of the open book $\mathcal{OB}$ to an oriented link which is isotopic to the binding of $\mathcal{OB}'$ on $\partial W'$ — because $\partial W = Y$ is a link of rational singularity, and thus a rational homology sphere, it seems possible to use [63] to construct an isotopy of pages of the open books if their bindings are isotopic, perhaps under some mild additional hypotheses. We leave most of the details to the motivated reader, only indicating below why the proposition should hold if the identification of the open books permutes the binding components. It should be emphasized that these arguments would yield only a mild generalization of Proposition 6.8: fixing appropriate boundary data is crucial for our proof. Note that by Wendl's theorem, all Stein fillings of a planar contact manifold fill *the same* open book; so in this sense, it is reasonable to think of the boundary open book as fixed.

To consider the case where the diffeomorphism between the fillings permutes the binding components of the open book, assume that there is an orientation-preserving self-diffeomorphism σ of the page of the open book that commutes with the monodromy. We do not assume that σ fixes the boundary of the page; in particular, we are interested in the case where σ permutes the boundary components. It can be shown that if σ acts nontrivially on the set of boundary components, then the decorated germ and/or the resolution graph of the singularity has the corresponding symmetry. For example, if σ exchanges holes h_1 and h_2, these holes must be enclosed by the same number of Dehn twists (in any positive factorizations of the open book); this implies, in particular, the equality of weights for the corresponding curvetta branches,

$$w_1 = w(C_1) = w(C_2) = w_2.$$

Additionally, for any other hole h_i, the number of Dehn twists enclosing the pair h_1, h_i must be the same as the number of Dehn twists enclosing the pair h_2, h_i. Because the Artin factorization is determined by combinatorial data (see Proposition 6.7), it follows that the Artin factorization admits a symmetry interchanging holes h_1 and h_2. Then, we can argue as in Proposition 4.5 to reconstruct the resolution graph of the singularity, and to see that the graph must have a symmetry, and the corresponding curvetta arrangement must admit a symmetry interchanging curvettas C_1 and C_2 (up to a topological equivalence). Similar reasoning would work for a more general self-diffeomorphism σ; we do not give the complete argument to avoid setting up complicated notation. If σ exchanges the boundary of a hole with the outer boundary of

the page (thought of as a disk with holes), there must be a symmetry of the resolution graphs and the corresponding extended graphs; see Section 2.

Since the self-diffeomorphism σ of the page commutes with the monodromy, it induces a self-diffeomorphism of the supporting 3–manifold Y, which is not necessarily isotopic to the identity. We will use the same notation for this self-diffeomorphism of Y, $\sigma : Y \to Y$.

Now, suppose that fillings W and W' are as in Proposition 6.8, and that there is an orientation-preserving diffeomorphism $f : W \to W'$ that maps the open book $\mathcal{OB}$ on W to the open book $f(\mathcal{OB})$ on W' that is isotopic to $\sigma(\mathcal{OB}')$ rather than to $\mathcal{OB}'$. As explained above, the decorated germ admits a symmetry induced by σ; in turn, it follows that the cap U admits a self-diffeomorphism that restricts to the map $\sigma : Y \to Y$ on the boundary, after an orientation reversal. Using this self-diffeomorphism to glue the cap to W', and comparing $W \cup_{\mathrm{id}} U$ and $W' \cup_{\sigma} U$, we can argue as in Proposition 6.8 to conclude that the incidence matrices $\mathcal{I}(\Gamma, \{p_j\})$ and $\mathcal{I}(\Gamma', \{p'_j\})$ are the same.

7 Milnor fibers and unexpected Stein fillings: examples

We now construct examples where the link of a rational singularity with reduced fundamental cycle has Stein fillings that are not realized by Milnor fibers of any smoothing.

Our examples build on results of the previous sections: by [27], Milnor fibers of smoothings correspond to (algebraic) picture deformations of the decorated germ, while Stein fillings of the link can be constructed from arbitrary smooth graphical homotopies of the curvettas. During the picture deformation, the decorated germ $\mathcal{C}$ is *immediately* deformed into an arrangement of curvettas yielding a Milnor fiber, so that the arrangement appears as the deformation $\mathcal{C}^s$ for small s (and for a given deformation, all values of s close to 0 produce diffeomorphic Milnor fibers and equivalent Lefschetz fibrations). Indeed, for an algebrogeometric 1–parameter deformation of the germ $\mathcal{C}$, the general fibers of the deformation all "look the same" (up to diffeomorphism). By contrast, during the course of a smooth graphical homotopy, we are allowed to change the topology of the arrangement of curvettas, and thus will produce Stein fillings whose topology varies during the homotopy. We emphasize that *immediate deformation* vs *long-term homotopy* of the branches of $\mathcal{C}$ makes the key difference between Milnor fillings and Stein fillings of links of rational singularities with reduced

fundamental cycle. In Section 8, we explain why this is the key aspect and compare picture deformations and smooth graphical homotopies in more detail. In this section, we exploit the difference between immediate deformations and long-term homotopies to produce examples of Stein fillings that are not diffeomorphic (rel boundary) to any Milnor fibers.

7.1 Arrangements of symplectic lines and pseudolines

To construct links of singularities that admit unexpected Stein fillings, we first consider decorated germs given by pencils of lines (with weights) and focus on their associated singularities. In this section, we will use the following terminology: several points are *collinear* if they all lie on the same line, and several lines are *concurrent* if they all pass through the same point. Concurrent lines form a *pencil*; we will refer to an arrangement of concurrent lines as a *pencil of lines*. We will also talk about concurrent pseudolines or concurrent smooth disks, with the same meaning.

Note that any two pencils of complex lines in $\mathbb{C}^2$ are isotopic through pencils, therefore the corresponding singularities are topologically equivalent and have contactomorphic links. Let $\mathcal{C} = \{C_1, C_2, \dots, C_m\}$ be a pencil of m complex lines, with each line C_k decorated by a weight $w_k = w(C_k)$. Consider the surface singularity that corresponds to the decorated germ $(\mathcal{C}, w)$, and let $Y(m, w) = Y(m; w_1, \dots, w_m)$ denote its link with the canonical contact structure ξ. Note that $Y(m, w)$ is a Seifert fibered space over S^2 with at most m singular fibers. Indeed, consider the dual resolution graph of the singularity; the graph gives a surgery diagram for the link. This graph has m legs emanating from the central vertex. Legs correspond to the lines of the pencil, so that the k^{th} leg has $w_k - 1$ vertices (including the central vertex).

Note that legs of length 1 consist only of the central vertex and thus will appear invisible. However, in the examples we focus on, every leg will have length greater than 1. The central vertex has self-intersection $-m - 1$, all the other vertices have self-intersection -2. See Figure 16 for an example. The decorated pencil $\mathcal{C}$ can be recovered from the graph as in Section 2: we add (-1) vertices at the end of each leg, take the corresponding collection of curvettas, and blow down the augmented graph.

To construct Stein fillings of $Y(m, w)$, we will use curvetta homotopies taking the pencil of complex lines to a symplectic line arrangement in $\mathbb{C}^2$. We define these arrangements as follows.

Definition 7.1 A *symplectic line arrangement* in $\mathbb{C}^2$ is a collection of m symplectic graphical disks $\Gamma_1, \dots, \Gamma_m$ in $\mathbb{C}^2$ with respect to a projection $\pi : \mathbb{C}^2 \to \mathbb{C}$ such that

(i) for every pair $i, j \in \{1, \dots, m\}$ with $i \neq j$, Γ_i intersects Γ_j positively transversally exactly once, and

(ii) for R sufficiently large, $(\Gamma_1 \cup \cdots \cup \Gamma_m) \cap \pi^{-1}(S_R)$ is isotopic to the braid given by one full twist on m strands in the solid torus $\pi^{-1}(S_R)$, where $S_R \subset \mathbb{C}$ is the circle of radius R.

Equivalently, we can view the symplectic line arrangement in a Milnor ball $B = D_x \times D_y \subset \mathbb{C}^2$ containing all intersections. The intersection of the arrangement with ∂B is then the braid of one full twist in $\partial D_x \times D_y$. A symplectic line arrangement in the closed ball B can always be extended to an arrangement in $\mathbb{C}^2$, so we will give all statements about symplectic line arrangements in $\mathbb{C}^2$.

Example 7.2 A pencil of complex lines intersecting at the origin in $\mathbb{C}^2$ is a symplectic line arrangement. Clearly every pair of lines intersects at a single point (the origin) transversally (and positively because they are complex). That the monodromy in $\pi^{-1}(S_R)$ is one full twist on m strands can be computed directly from a model as in [40].

More generally, any complex line arrangement of m lines in $\mathbb{C}^2$ such that no intersections between lines occur at infinity (ie every complex line has a different complex slope) gives a symplectic line arrangement. This can be seen by compactifying the line arrangement in $\mathbb{C}\mathrm{P}^2$ and looking at the intersection of the lines with the boundary of a regular neighborhood of the $\mathbb{C}\mathrm{P}^1$ at infinity. These intersections form an m component link with one component for each line, such that the link components are isotopic to disjoint fibers of the ε–neighborhood (which can be identified with a subset of the normal bundle) of the $\mathbb{C}\mathrm{P}^1$ at infinity. After changing coordinates from the perspective of the $\mathbb{C}\mathrm{P}^1$ at infinity to the perspective of the complementary ball, the components of the link obtain one full twist. From the Kirby calculus perspective, the boundary of the ε–neighborhood of $\mathbb{C}\mathrm{P}^1$ is presented as $(+1)$ surgery on the unknot, and the link is m parallel meridians of this surgery curve. After reversing orientation to get the boundary of the complementary ball, the surgery coefficient on the unknot becomes a (-1) surgery, and blowing down this surgery curve induces one full twist in the m unknotted meridians.

Since any symplectic line arrangement has the same monodromy as the pencil of complex lines, Lemmas 5.9 and 5.10 imply they are related to the pencil by a smooth graphical homotopy.

Our primary source of examples of noncomplex symplectic line arrangements is given by pseudoline arrangements as described below. However, symplectic line arrangements are more general and can include braiding in the associated wiring diagram.

Example 7.3 A *pseudoline arrangement* is a collection $\ell_1, \dots, \ell_m$ of smooth graphical curves in $\mathbb{R}^2$ where for every pair i, j, the curves ℓ_i and ℓ_j intersect transversally at exactly one point. Such a pseudoline arrangement can be considered a braided wiring diagram as in Definition 5.1, but in the particular case where there is no braiding. In particular, we can apply Proposition 5.5 to extend the pseudoline arrangement to an arrangement of symplectic graphical disks $\Gamma_1, \dots, \Gamma_m$; the extension produces a symplectic line arrangement. Indeed, condition (i) in the definition of a symplectic line arrangement is satisfied because any two pseudolines intersect transversally at one point, and their extensions intersect positively by construction. Condition (ii) follows from the calculation of the total monodromy as in Section 5.2 and a classical theorem of Matsumoto and Tits [37] about uniqueness of reduced factorizations in the braid group.

Alternatively, we can refer to the results of [57, Section 6], where pseudoline arrangements in $\mathbb{R}\mathrm{P}^2$ are extended to symplectic line arrangements in $\mathbb{C}\mathrm{P}^2$ (extensions in $\mathbb{C}\mathrm{P}^2$ are strictly harder to construct than extensions in $\mathbb{C}^2$). Additionally, using the same theorem of Matsumoto and Tits, [57, Proposition 6.4] provides a homotopy of pseudoline arrangements connecting the given arrangement to the pencil. After applying Proposition 5.5, we get a homotopy of the corresponding symplectic line arrangements. Note that by construction, this homotopy of symplectic line arrangements keeps all intersections positive at all times, whereas the smooth graphical homotopy given by Lemmas 5.9 and 5.10 may introduce negative intersections.

We use symplectic line and pseudoline arrangements to construct Stein fillings of Seifert fibered spaces $(Y(m; w), \xi)$ via Lemmas 3.2 and 3.4.

Proposition 7.4 *Let $(\mathcal{C}, w)$ be a decorated pencil of m lines. Suppose that $\Gamma = \{\Gamma_1, \dots, \Gamma_m\}$ is a symplectic line arrangement such that each disk Γ_i has at most w_i distinct intersection points with the other disks of the arrangement. Then, $(\Gamma, \{p_j\})$ yields a Stein filling of $(Y(m; w_1, w_2, \dots, w_m), \xi)$.*

In particular, a pseudoline arrangement $\Lambda = \{\ell_1, \dots, \ell_m\}$ gives a Stein filling of $(Y(m; w_1, w_2, \dots, w_m), \xi)$ via an extension to a symplectic line arrangement, provided that ℓ_i has at most w_i distinct intersection points with the other pseudolines.

7.2 Unexpected line arrangements yield unexpected fillings

Now we will show that some of the Stein fillings as above do not arise as Milnor fibers. In the next lemma, we consider analytic deformations of reducible plane curve germs, associated to a singularity by the de Jong–van Straten theory, and establish a property that will play a key role in our construction of unexpected arrangements.

The term *δ–constant* deformation in the next lemma refers to an algebrogeometric property: the deformation is required to preserve the δ–invariant of a singular plane curve. We keep this terminology since it is used in [27] and [44]; however, under the hypothesis that the germ has smooth branches, the δ–constant condition simply means that the deformation changes the germ componentwise, without merging different components. Intuitively, the δ–invariant counts the number of double points "concentrated" in each singular point [39, Section 10]; for example, an ordinary d–tuple point (where d smooth components meet transversely) contributes $\delta = \frac{1}{2}d(d-1)$, since it can be perturbed to $\frac{1}{2}d(d-1)$ double points. Thus, we can deform a triple point to three double points by a δ–constant deformation, but we are not allowed to deform two transversely intersecting lines into a smooth conic (such a deformation would kill a double point).

Lemma 7.5 *Consider the germ of a reducible plane curve $\mathcal{C}$ in $\mathbb{C}^2$ with m smooth graphical branches $C_1, C_2, \dots, C_m$ passing through 0, and let $\mathcal{C}^s = \bigcup_{k=1}^m C_k^s$ be a δ–constant deformation of $\mathcal{C}$.* (Here, *δ–constant* means that each branch of the germ is deformed individually, ie the deformation is not allowed to merge different branches.) *Suppose that all the branches $C_1, \dots, C_m$ have distinct tangent lines at 0, and that not all deformed branches $C_1^s, \dots, C_m^s$ are concurrent for $s \neq 0$.*

Then there exists a complex line arrangement $\mathcal{A} = \{L_1, \dots, L_m\}$ in $\mathbb{C}^2$ such that not all lines in $\mathcal{A}$ are concurrent, no two lines are equal, and $\mathcal{A}$ satisfies all the incidence relations of $\mathcal{C}^s$. Namely, for any collection of the deformed branches $C_{i_1}^s, C_{i_2}^s, \dots, C_{i_k}^s$ that intersect at one point, the corresponding lines $L_{i_1}, L_{i_2}, \dots, L_{i_k}$ also intersect, ie

$$C_{i_1}^s \cap C_{i_2}^s \cap \cdots \cap C_{i_k}^s \neq \varnothing \implies L_{i_1} \cap L_{i_2} \cap \cdots \cap L_{i_k} \neq \varnothing. \tag{7-1}$$

Note that the incidence pattern for branches of $\mathcal{C}^s$ is the same for all $s \neq 0$, because the definition of a 1–parameter deformation implies that all nearby fibers "look the same". It is important to keep in mind that the complex line arrangement $\mathcal{A}$ may satisfy additional incidences, so that certain intersection points coincide in $\mathcal{A}$ but are distinct

for the arrangement $\{C_1^s, C_2^s, \dots, C_m^s\}$. In particular, a pencil of lines would satisfy incidence relations of any other arrangement, but we postulate that $\mathcal{A}$ cannot be a pencil (the lines in $\mathcal{A}$ are not all concurrent).

Proof of Lemma 7.5 Since any two curvettas intersect positively in the original germ $\mathcal{C}$, any two deformed branches C_i^s, C_j^s intersect for $s \neq 0$. We can make an s–dependent translation to ensure that the first two branches always intersect at the origin, $C_1^s \cap C_2^s = \{0\}$; strictly speaking, this means passing to a slightly different deformation of the germ $\mathcal{C}$.

All components of the reducible curve $\mathcal{C}$ pass through 0 and are graphical analytic disks with respect to the projection to the x–coordinate. Thus we can define the germ of $\mathcal{C}$ near 0 by an equation of the form

$$\prod_{i=1}^{m}(a_i x + c_i(x) - y) = 0,$$

where $c_i(x) = \sum_{k\geq 2} c_{i,k} x^k$ are analytic functions in x with $\operatorname{ord}_x c_i > 1$ at 0. We can also assume that $a_i \neq 0$ for all $i = 1, \dots, m$.

The 1–parameter deformation $\mathcal{C}^s$ is then given, for s close to 0, by an equation of the form

$$\prod_{i=1}^{m}(a_i(s)x + b_i(s) + c_i(x,s) - y) = 0.$$

Here a_i and b_i are analytic functions in s, and at the origin $(0,0)$ we have $\operatorname{ord}_s a_i = 0$ and $\operatorname{ord}_s b_i > 0$; additionally, $c_i(x,s)$ is analytic in x and s, and $\operatorname{ord}_x c_i > 1$. The i^{th} component C_i^s of the deformed curve at time s is given by $a_i(s)x + b_i(s) + c_i(x,s) - y = 0$. Because the branches C_1^s and C_2^s pass through 0 for all s, we have $b_1 \equiv b_2 \equiv 0$. At $s = 0$ all components pass through the origin, so $b_i(0) = 0$ for all i.

Let $r = \min_i(\operatorname{ord}_s b_i)$, where the order is always taken at the origin. Because $b_i(0) = 0$ for all i, we have $r > 0$, and $r = \operatorname{ord}_s b_{i_0}$ for some $3 \leq i_0 \leq m$. Notice also that $r < +\infty$, since otherwise all the components C_i^s would pass through 0 for all $s \neq 0$. We write $b_i(s) = s^r \overline{b}_i(s)$; then $\overline{b}_{i_0}(0) \neq 0$.

Now make a change of variables for $s \neq 0$,

$$x = s^r x' \quad \text{and} \quad y = s^r y'.$$

Since $\operatorname{ord}_x c_i(x,s) \geq 2$, we have $c_i(x,s) = s^{2r}\overline{c}_i(x',s)$ for some analytic function $\overline{c}_i$. Thus, the equation for the deformation becomes

$$\prod_{i=0}^{m}\left(a_i(s)s^r x' + s^r \overline{b}_i(s) + s^{2r}\overline{c}_i(x',s) - s^r y'\right) = 0.$$

Equivalently, for $s \neq 0$ and $i = 1,\dots,m$, the deformed components C_i^s are given by the equations

$$a_i(s)x' + \overline{b}_i(s) + s^r\overline{c}_i(x',s) - y' = 0.$$

When we pass to the limit as $s \to 0$, the equations become

$$a_i(0)x' + \overline{b}_i(0) - y' = 0,$$

so in the limit we obtain an arrangement of straight lines in $\mathbb{C}^2$. Not all of these lines are concurrent, since $\overline{b}_{i_0}(0) \neq 0$ while $\overline{b}_1(0) = \overline{b}_2(0) = 0$.

The curves C_i^s satisfy the same incidence relations for all $s \neq 0$. Since intersection points between curves vary continuously with s, the incidence relations must be preserved in the limit, so (7-1) holds. □

Our examples of unexpected Stein fillings are given by pseudoline arrangements with the following special property.

Definition 7.6 Let $\Lambda = \{\Gamma_1,\dots\Gamma_m\} \subset \mathbb{R}^2$ be a symplectic line arrangement where not all lines are concurrent. We say that Λ is *unexpected* if the only complex line arrangements that satisfy all the incidence relations of Λ are pencils of lines. Namely, whenever a complex line arrangement $\mathcal{A} = \{L_1, L_2, \dots, L_m\} \subset \mathbb{C}^2$ has the property

$$\Gamma_{i_1} \cap \Gamma_{i_2} \cap \cdots \cap \Gamma_{i_k} \neq \varnothing \implies L_{i_1} \cap L_{i_2} \cap \cdots \cap L_{i_k} \neq \varnothing,$$

all the lines $L_1, L_2, \dots, L_m$ of $\mathcal{A}$ must be concurrent.

If an unexpected symplectic line arrangement comes from a pseudoline arrangement, we will say that the pseudoline arrangement is unexpected.

Remark 7.7 It is important to note that unexpected symplectic line arrangements are not the same as symplectic line arrangements not realizable by complex lines. Being an unexpected arrangement is a stronger condition: we want to rule out not only complex line arrangements with the same incidence relations as those of Λ, but also complex line arrangements that satisfy all the incidence relations of Λ and possibly additional incidence relations (without being a pencil). For instance, the pseudo-Pappus

arrangement (Example 8.1 in the next section) is not realizable by complex lines but it is not unexpected, because the classical Pappus arrangement has all of the same incidences and an additional one.

Theorem 7.8 *Suppose that $\Gamma = \{\Gamma_1, \dots, \Gamma_m\}$ is an arrangement of smooth graphical disks with marked points $\{p_j\}$, related by a smooth graphical homotopy to a decorated germ $(\mathcal{C}, w)$. Let (Y, ξ) be the link of the surface singularity that corresponds to $(\mathcal{C}, w)$. Suppose that a subcollection of disks $\{\Gamma_1, \Gamma_2, \dots, \Gamma_r\}$ of Γ forms an unexpected symplectic line arrangement.*

Then the Stein filling W given by $(\Gamma, \{p_j\})$ is not strongly diffeomorphic to any Milnor filling of (Y, ξ). If the weights on $\mathcal{C}$ are large enough, W is simply connected.

By Proposition 7.4, unexpected line arrangements yield unexpected fillings of Seifert fibered spaces of the form $Y(m, w)$.

Corollary 7.9 *Let $\Gamma = \{\Gamma_1, \dots, \Gamma_m\}$ be an unexpected symplectic line arrangement, and for $k = 1, \dots, m$, let $w(\Gamma_k)$ denote the number of intersection points of Γ_k with the disks Γ_i, $i \neq k$. Then for every weight $w = (w_1, w_2, \dots, w_m)$ with $w_k \geq w(\Gamma_k)$ for $k = 1, \dots, m$, the Seifert fibered space $(Y(m, w), \xi)$ has a Stein filling not strongly diffeomorphic to any Milnor filling. This Stein filling is given by a Lefschetz fibration constructed from the arrangement Γ with the appropriate choice of marked points. When strict inequalities $w_k > w(\Gamma_k)$ hold for all k, we get a simply connected unexpected Stein filling.*

Proof of Theorem 7.8 Observe that when the number of intersection points on each Γ_i is smaller than the weight of the corresponding branch of the decorated germ, each Γ_i has a free marked point. Then the Lefschetz fibration constructed from $(\Gamma, \{p_j\})$ has a boundary-parallel vanishing cycle around every hole in the disk fiber, so that the corresponding thimbles kill all generators of $\pi_1(\text{fiber})$, and therefore, in this case $\pi_1(W) = 0$.

Let W_M be a Milnor filling that arises from a smoothing of some surface singularity with the link Y. By Theorem 1.3, W_M corresponds to a picture deformation $\mathcal{C}'^s$ of a decorated germ $\mathcal{C}' = \bigcup_{i=1}^m C_i'$ with weight w, topologically equivalent to $(\mathcal{C}, w)$.

Although the germs $\mathcal{C}$ and $\mathcal{C}'$ may differ analytically, they are topologically equivalent and thus have isotopic boundary braids. Therefore by Lemma 3.4 the open book decomposition naturally induced by the Lefschetz fibration in Lemma 3.2 for W agrees with that for W_M, so comparing them via strong diffeomorphism makes sense.

By Proposition 6.8, if W is strongly diffeomorphic to W_M, the incidence matrix of the deformed curvetta arrangement $\{C_1'^s, \dots, C_m'^s\}$, $s \neq 0$, with its marked points must be the same as the incidence matrix for the arrangement $(\Gamma, \{p_j\})$, up to permutation of columns. In particular, we see that the subarrangement $\{\Gamma_1, \dots, \Gamma_r\}$ of symplectic lines satisfies the same incidence relations as the subarrangement $\{C_1'^s, \dots, C_r'^s\}$ of the deformed curvettas of $\mathcal{C}'$. By assumption, in each of these arrangements not all curvettas are concurrent. Because pairs of curves $\Gamma_1, \dots, \Gamma_r$ intersect algebraically positively once, $C_1', \dots, C_r'$ have distinct tangent lines. Now by Lemma 7.5, there exists a complex line arrangement $\mathcal{A}$ that satisfies all the incidence relations of $\{C_1'^s, \dots, C_r'^s\}$, and thus all the incidence relations of Γ. This is a contradiction because Γ is an unexpected arrangement. □

7.3 Constructing unexpected pseudoline arrangements

We now give examples of unexpected pseudoline arrangements; these will yield concrete examples of unexpected Stein fillings. We start with classical projective geometry constructions.

Example 7.10 Recall that the classical Pappus arrangement in $\mathbb{R}^2$ is constructed as follows. Take two lines, ℓ_1 and ℓ_2, and mark three distinct points a, b, c on L_1 and three distinct points A, B, C on ℓ_2, avoiding the intersection $\ell_1 \cap \ell_2$. Consider the following lines through pairs of marked points:

$$\ell_3 = aB, \quad \ell_4 = aC, \quad \ell_5 = bA, \quad \ell_6 = bC, \quad \ell_7 = cA, \quad \ell_8 = cB.$$

The Pappus theorem asserts that the three intersection points $\ell_3 \cap \ell_5$, $\ell_4 \cap \ell_7$, and $\ell_6 \cap \ell_8$ are collinear; the classical Pappus arrangement consists of the lines $\ell_1, \dots, \ell_8$, together with the line through these three points. We modify this last line to make an unexpected pseudoline arrangement, as follows. Let ℓ_9 be a line through C, distinct from ℓ_4 and ℓ_6. Consider the intersection point $\ell_8 \cap \ell_9$ and let ℓ_{10} be a pseudoline passing through points $\ell_3 \cap \ell_5$, $\ell_4 \cap \ell_7$ and $\ell_8 \cap \ell_9$, as shown in Figure 15. Let $\mathcal{P} = \{\ell_1, \ell_2, \dots, \ell_{10}\}$.

Notice that in this case, it is clear that the pseudoline ℓ_{10} can be homotoped to the classical Pappus line through the points $\ell_3 \cap \ell_5$, $\ell_4 \cap \ell_7$ and $\ell_6 \cap \ell_8$. The resulting arrangement of straight lines in $\mathbb{R}^2$ can be homotoped to a pencil by linear homotopy. (We already know from discussion in Example 7.3 that $\mathcal{P}$ is homotopic to the pencil, but here we have a very simple explicit homotopy.)

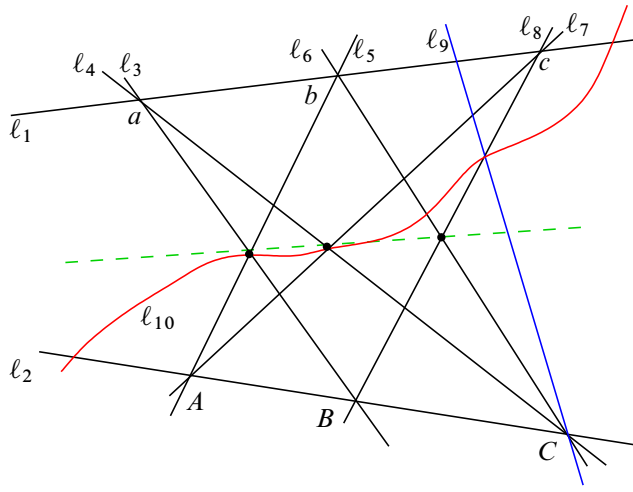


Figure 15: The pseudoline arrangement $\mathcal{P} = \{\ell_1, \ell_2, \dots, \ell_{10}\}$ is given by the black lines, the blue line, and the red line in the figure. The dotted line in the middle is not included. The dotted line and the eight black lines give the classical Pappus arrangement. The intersection points $\ell_1 \cap \ell_2$, $\ell_3 \cap \ell_6$, $\ell_3 \cap \ell_9$, $\ell_5 \cap \ell_8$ and $\ell_5 \cap \ell_9$ are not shown in the figure.

Proposition 7.11 *The arrangement $\mathcal{P}$ is unexpected.*

Proof As already stated, the classical Pappus theorem asserts that for the given arrangement, the intersection points $\ell_3 \cap \ell_5$, $\ell_4 \cap \ell_7$, and $\ell_6 \cap \ell_8$ are collinear. Collinearity holds both in the real and in the complex projective geometry settings, so that if $L_1, L_2, \dots, L_7, L_8 \subset \mathbb{C}^2$ are complex lines with given incidences, then $L_3 \cap L_5$, $L_4 \cap L_7$, and $L_6 \cap L_8$ are collinear. From this, we can immediately see that the arrangement $\mathcal{P}$ is not realizable by complex lines $\{L_1, L_2, \dots, L_{10}\}$: since $L_6 \cap L_8$ and $L_8 \cap L_9$ are distinct points on L_8, the points $L_3 \cap L_5$, $L_4 \cap L_7$ and $L_8 \cap L_9$ cannot be collinear.

To show that $\mathcal{P}$ is unexpected, we need to prove that no complex line arrangement satisfies all the incidence relations of $\mathcal{P}$ even if some (but not all) of the intersection points coincide. Indeed, we show that if a complex line arrangement $\mathcal{A} = \{L_1, L_2, \dots, L_{10}\}$ satisfies the incidence relations of $\mathcal{P}$ and two of the intersection points coincide, then $\mathcal{A}$ must be a pencil. Remember that we always assume that all the lines in the arrangement are distinct.

The following trivial fact, applied systematically, greatly simplifies the analysis of cases:

Observation 7.12 *Let L_1, L_2, L_3, L_4 be four lines in $\mathbb{C}^2$, which are not necessarily distinct. Suppose that two of the pairwise intersection points coincide: $L_1 \cap L_2 = L_3 \cap L_4$. Then L_1, L_2, L_3 and L_4 are concurrent, so that they all intersect at the point $L_1 \cap L_2 = L_1 \cap L_3 = L_1 \cap L_4 = L_2 \cap L_3 = L_2 \cap L_4 = L_3 \cap L_4$.*

In the case of three lines, if $L_1 \cap L_2 = L_3 \cap L_1$, then $L_2 \cap L_3 = L_1 \cap L_2 = L_3 \cap L_1$. Visually, if two vertices of a triangle coincide, the third vertex of the triangle coincides with the first two.

Assuming that some of the intersection points in Figure 15 coincide, we mark these points by "O", and then use Observation 7.12 to chase vertices that coincide: starting with two marked vertices, we look for additional vertices that coincide with the first two, further mark these by "O", and continue. When every line contains a marked intersection point, we know that all lines in the arrangement are concurrent: they form a pencil though O.

We begin this process. First, assume that the intersection points $L_3 \cap L_5 \cap L_{10}$ and $L_4 \cap L_7 \cap L_{10}$ are distinct. By the Pappus theorem, the complex line arrangement $\mathcal{A} = \{L_1, L_2, \dots, L_{10}\}$ can satisfy all the incidence relations of $\mathcal{P}$ only if $L_6 \cap L_8 = L_8 \cap L_9 \cap L_{10}$. Setting $\mathrm{O} = L_6 \cap L_8 = L_8 \cap L_9 \cap L_{10}$, by Observation 7.12 we have $\mathrm{O} = C = L_4 \cap L_6 \cap L_9 \cap L_2$, then $\mathrm{O} = B = L_8 \cap L_2 \cap L_3$, then $\mathrm{O} = a = L_3 \cap L_4 \cap L_1$, then $\mathrm{O} = b = L_5 \cap L_6 \cap L_1$ and $\mathrm{O} = c = L_7 \cap L_8 \cap L_1$. Now, O appears on every line at least once, so the arrangement degenerates to a pencil.

(This can be seen quickly if in the above diagram, you highlight the lines passing through intersection points marked by O, in order. You can mark a new intersection by O if it contains at least two highlighted lines, and then highlight all the lines through that point O. When all the lines are highlighted, you have a pencil.)

For the second case, assume that the intersection points $L_3 \cap L_5 \cap L_{10}$ and $L_4 \cap L_7 \cap L_{10}$ coincide. Set $\mathrm{O} = L_3 \cap L_5 \cap L_{10} = L_4 \cap L_7 \cap L_{10}$. Then $\mathrm{O} = a = L_3 \cap L_4 \cap L_1$ and $\mathrm{O} = A = L_5 \cap L_7 \cap L_2$. Then $\mathrm{O} = c = L_7 \cap L_8 \cap L_1$ and $\mathrm{O} = C = L_4 \cap L_6 \cap L_2 \cap L_9$. Again, every line contains a point marked O, so the arrangement degenerates to a pencil. □

Corollary 7.13 *Let $Y = Y(10; w)$ be a Seifert fibered space given by a star-shaped plumbing graph with 10 legs, as in Figure 16, such that eight of the legs of the graph have at least 5 vertices each, including the central vertex, and two remaining legs have at least 4 vertices each. (Equivalently, two components of w are 5 or greater, and the*

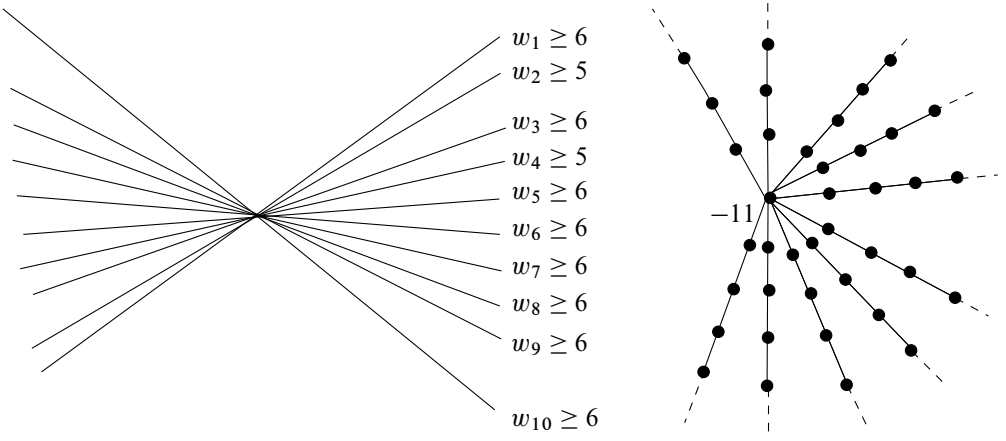


Figure 16: Left, a pencil of 10 lines decorated with weights. Right, the plumbing graph for Y: the central vertex has self-intersection -11, all the rest have self-intersection -2. Eight of the legs have at least 5 vertices each (including the central vertex), and two remaining legs have at least 4 vertices each.

rest are 6 or greater.) Observe that Y is the link of a rational singularity, and let ξ be the Milnor fillable contact structure on Y. Then (Y, ξ) admits a Stein filling which is not strongly diffeomorphic to any Milnor filling.

Proof We count the intersection points on each line in the arrangement $\mathcal{P}$: $w(\ell_2) = w(\ell_4) = 5$, $w(\ell_k) = 6$ for $k \neq 2, 4$. Then for any collection of integers $w_1, w_2, \ldots, w_{10}$ such that $w_2 \geq 5$, $w_4 \geq 5$ and $w_k \geq 6$ for $k \neq 2, 4$, we can mark the lines of the arrangement $\mathcal{P}$ as required in Corollary 7.9. The corresponding singularity has the dual resolution graph as shown in Figure 16, with one leg of length $w_k - 1$ for each line L_k in the arrangement, so the link is the Seifert fibered space $Y(10, w)$. The result now follows from Corollary 7.9 and Proposition 7.11. □

A different example comes from a version of the Desargues theorem; we use complete quadrangles and harmonic conjugates. The example in Figure 17 was pointed out to us by Stepan Orevkov. He suggested an approach to proving that this arrangement cannot appear as an algebraic deformation of a pencil. We are grateful for his input, which inspired us to define unexpected line arrangements and prove Lemma 7.5.

Example 7.14 In the standard $\mathbb{R}^2 \subset \mathbb{R}\mathrm{P}^2$, we take four vertical lines $\ell_1, \ell_2, \ell_3, \ell_4$, three horizontal lines ℓ_5, ℓ_6, ℓ_7, the two parallel diagonal lines ℓ_8, ℓ_9, and a "bent"

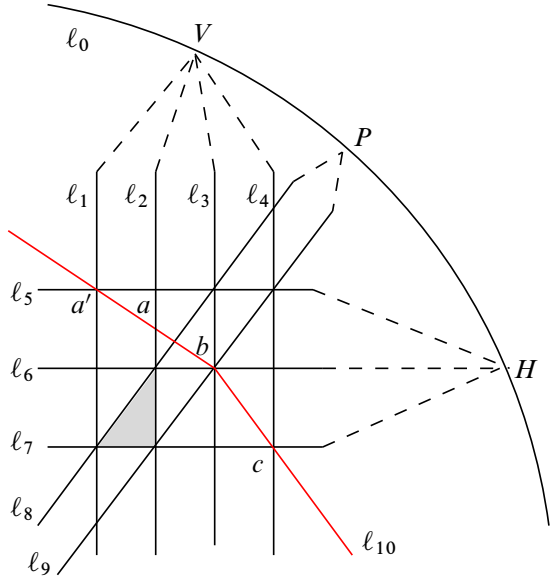


Figure 17: An arrangement of real pseudolines. The intersection of ℓ_0 and ℓ_{10} is not shown.

pseudoline ℓ_{10}, as shown in Figure 17. Let ℓ_0 be the line at infinity. Note that because $\ell_1, \ell_2, \ell_3, \ell_4$ are all parallel in $\mathbb{R}^2$, they intersect at a point V on ℓ_0. Similarly, the lines ℓ_5, ℓ_6, ℓ_7 have a common intersection with ℓ_0 at a point H, and the lines ℓ_8 and ℓ_9 intersect on ℓ_0 at a point P. Removing from $\mathbb{R}\mathrm{P}^2$ a line which is different from all the ℓ_i and intersects them generically, we can consider $\mathcal{Q} = \{\ell_i\}_{i=0}^{10}$ as a pseudoline arrangement in $\mathbb{R}^2$. (See Figure 18 for a version where ℓ_0 is no longer the line at infinity.)

Proposition 7.15 *The pseudoline arrangement $\mathcal{Q}$ is unexpected.*

Proof Suppose that a complex line arrangement $\mathcal{A} = L_0, L_1, \dots, L_{10}$ satisfies all the incidence relations of $\mathcal{Q}$. This means that for all intersections between the pseudolines in Figure 17, the corresponding lines of $\mathcal{A}$ intersect. We claim that unless $\mathcal{A}$ is a pencil, all of these intersection points must be distinct — that is, no two distinct intersection points in Figure 17 can coincide for the arrangement $\mathcal{A}$. To see this, we use Observation 7.12 repeatedly, as in Proposition 7.11. Recall that $V = L_1 \cap L_2 \cap L_3 \cap L_4 \cap L_0$ and $H = L_5 \cap L_6 \cap L_7 \cap L_0$.

If $H = V = \mathrm{O}$, then we have $L_i \cap L_j = \mathrm{O}$ for all $1 \le i \le 4$ and $5 \le j \le 7$, so $\mathcal{A}$ is a pencil.

If one of the intersection points $L_i \cap L_j$ with $1 \le i \le 4$ and $5 \le j \le 7$ coincides with V or H, then we have two intersection points marked with O on a vertical or horizontal line in Figure 17; then $O = V = H$, and all lines are concurrent.

If any two intersection points $L_i \cap L_j$ with $1 \le i \le 4$ and $5 \le j \le 7$ coincide, Observation 7.12 implies that they will both coincide with at least one of V or H, so we revert to the previous case.

Finally, if all the points V, H and $L_i \cap L_j$ with $1 \le i \le 4$ and $5 \le j \le 7$ are distinct, all remaining intersection points which do not coincide with one of these are necessarily generic double points (otherwise we would have a pair of lines intersecting more than once).

Once we know that all the distinct intersections for $\mathcal{Q}$ are distinct for $\mathcal{A}$, it remains to show that $\mathcal{Q}$ cannot be realized as a complex line arrangement $\mathcal{A} = \{L_i\}_{i=0}^{10}$. Suppose that it is, for the sake of contradiction.

We will show that the intersection points

$$a = L_2 \cap L_5, \quad b = L_3 \cap L_6 \quad \text{and} \quad c = L_4 \cap L_7$$

are collinear. (See Figure 18.) Then we can conclude that the points $a' = L_1 \cap L_5$, b and c cannot be collinear. Indeed, $a \neq a'$, since all intersection points in the diagram are distinct. If all four points a, a', b and c were collinear, then the line L_5 through a and a' would coincide with the line L_{10} through a', b and c, but we assume that L_5 and L_{10} are distinct.

To see that the points a, b and c are collinear, we will use some notions of classical projective geometry, namely complete quadrangles and harmonic conjugates. (In Remark 7.16 below, we also indicate an alternative proof, in the more familiar Euclidean terms.) Observe that the lines L_5, L_6, L_2, L_3, L_8 and the line L through a and b form the four sides and the two diagonals of a complete quadrangle. Then the point $Q = L \cap L_0$ is the harmonic conjugate of the point $P = L_8 \cap L_0$ with respect to the points $V = L_2 \cap L_3$ and $H = L_5 \cap L_6$. Now, consider the lines L_2, L_4, L_5, L_7, L_9 and the line L' through a and c. Again these form a complete quadrangle, so that the point $Q' = L' \cap L_0$ is the harmonic conjugate of the point $P = L_9 \cap L_0$ with respect to $V = L_2 \cap L_4$ and $H = L_5 \cap L_7$. Since the harmonic conjugate of P with respect to V and H is unique, it follows that $Q = Q'$. Since the lines L and L' both pass through $Q = Q'$ and a, we must have $L = L'$, and so all three points a, b and c lie on this line. □

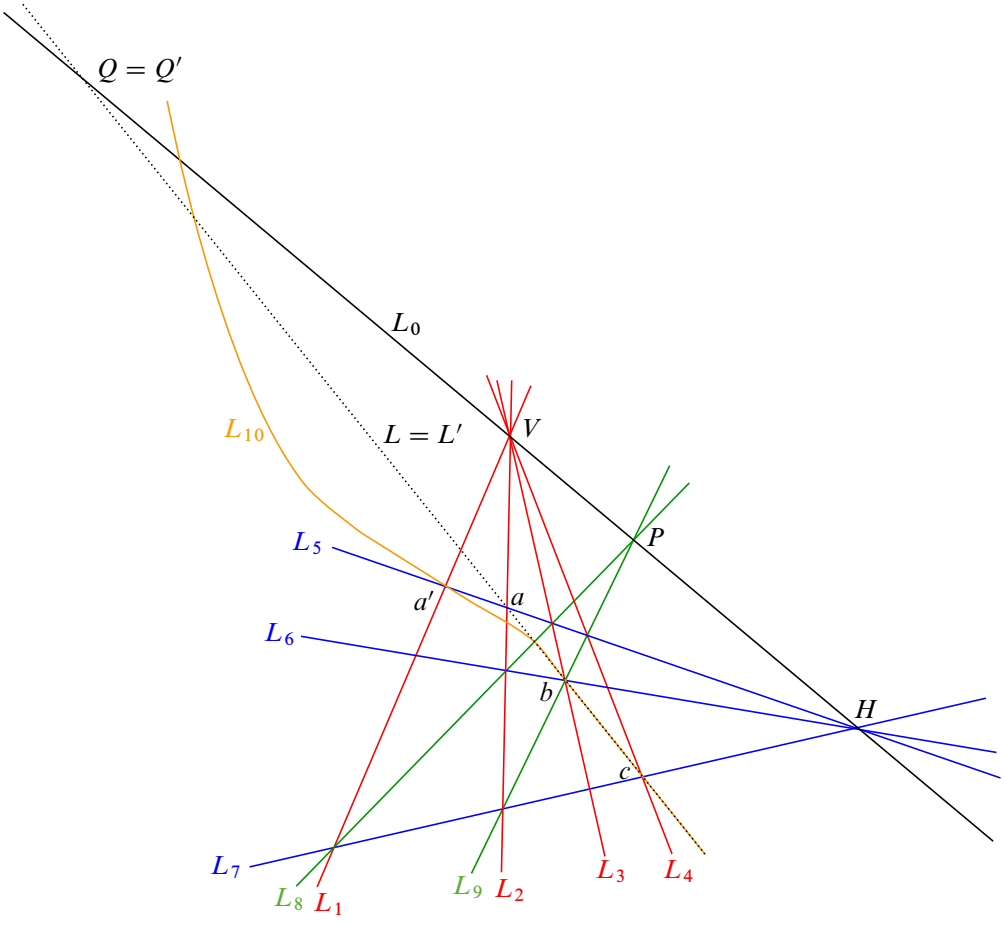


Figure 18: An arrangement of lines $L_0, L_1, \dots, L_9$ and pseudoline L_{10} with incidences as in Figure 17. We show that the line L through a and b and the line L' through a and c coincide (with the dotted line shown), so the points a, b and c are collinear. Therefore a', b and c cannot be collinear.

Remark 7.16 The above statement also has an easy Euclidean geometry proof, after some projective transformations. Indeed, we can find an automorphism of $\mathbb{C}\mathrm{P}^2$ such that

$$L_1 \cap L_5 \mapsto (0:0:1), \quad L_1 \cap L_6 \mapsto (1:0:1),$$
$$L_2 \cap L_5 \mapsto (0:1:1), \quad L_2 \cap L_6 \mapsto (1:1:1).$$

Then $H \mapsto (1:0:0)$ and $V \mapsto (0:1:0)$, and it is not hard to see that all the lines in the figure must be complexifications of real lines. The line L_0 is the line at infinity; the remaining lines are (complexifications of) the corresponding real lines in $\mathbb{R}^2$. We use the same notation for the real lines. Now we see that L_1, L_2, L_3, L_4 are parallel

vertical lines, L_5, L_6, L_7 are parallel horizontal lines, etc. So the arrangement looks like Figure 17. The lines in the figure form a number of triangles that are similar to the shaded triangle; it then follows that the points a, b, c are collinear, so a', b, c are not.

Note, however, that the above proof is somewhat incomplete: Figure 17 assumes a particular position of the lines L_3, L_4, L_7 relative to L_1, L_2, L_5, L_6. For a complete proof, an additional analysis of cases is required, with slightly different figures for other possible relative positions of the lines. Our projective argument with harmonic conjugates allows us to avoid this analysis, and also to emphasize the projective nature of the statement and the proof.

Corollary 7.17 *Let $Y = Y(11; w)$ be the Seifert fibered space given by a star-shaped plumbing graph with 11 legs such that two legs have at least 5 vertices each, two legs have at least 3 vertices, and the remaining 7 legs have at least 4 vertices each (including the central vertex). In other words, two components of the multiweight w are 4 or greater, two are 6 or greater, and the remaining seven are 5 or greater. Let ξ be the Milnor fillable contact structure on Y. Then (Y, ξ) admits a Stein filling which is not strongly diffeomorphic to any Milnor filling.*

Proof Exactly as in Corollary 7.13, this follows from Corollary 7.9 and Proposition 7.15. The picture is similar to Figure 16, with the obvious minor changes. Indeed, the pseudoline arrangement of Proposition 7.15 has two lines ℓ_0 and ℓ_3 with weight 4, two lines ℓ_9 and ℓ_{10} with weight 6, and seven remaining lines with weight 5. Note that a permutation of the components of w does not change the contact manifold, so we avoided labeling the components of w in the statement of the corollary. □

It is easy to generalize the above examples to star-shaped graphs with higher negative self-intersection values of the central vertex. Indeed, by Theorem 7.8, we can construct unexpected Stein fillings from an arbitrary arrangement of smooth graphical disks that contains an unexpected symplectic line arrangement. We turn to the general case later in this section; for now, we create more unexpected pseudoline arrangements simply by adding extra lines.

Lemma 7.18 *Suppose that Λ is an unexpected symplectic line arrangement. Let ℓ be a symplectic line that passes through at least one intersection point of two or more lines in Λ. Then the pseudoline arrangement $\Lambda \cup \{\ell\}$ is also unexpected.*

Proof If there exists a complex line arrangement $\mathcal{A}\cup\{L\}$ that satisfies all the incidence relations of $\Lambda\cup\{\ell\}$, and L corresponds to ℓ, then $\mathcal{A}$ satisfies all incidences of Λ, and so $\mathcal{A}$ is a pencil. The line L must pass through the intersection of two or more lines of $\mathcal{A}$, so $\mathcal{A}\cup\{L\}$ is also a pencil. □

Theorem 7.19 *For any $m\geq 10$, consider the Seifert fibered space $Y_m = Y(m,w)$ with $m\geq 10$, with weights $w=(w_1,\dots,w_m)$ such that $w_i\geq m-1$ for all $i=1,\dots,m$. The space Y_m is given by a star-shaped graph with $m\geq 10$ legs, such that the length of each leg is at least $m-1$. The central vertex has self-intersection $-m-1$, and all other vertices have self-intersection -2. Let ξ be the Milnor fillable contact structure on Y. Then (Y,ξ) admits a simply connected Stein filling not strongly diffeomorphic to any Milnor fiber.*

Proof We can add lines to the arrangement $\mathcal{P}$ to form unexpected arrangements of $m\geq 10$ pseudolines. Since any pseudolines intersect at most once, each pseudoline has at most $m-1$ intersections with other lines. By Corollary 7.9, $Y=Y(n,w)$ is an unexpected Stein filling if $w_i\geq m-1$ for all $i=1,\dots,m$, which is simply connected if all inequalities are strict. □

Varying the positions of the additional lines and/or applying a similar procedure to different arrangements such as $\mathcal{P}$ and $\mathcal{Q}$, it is possible to construct a variety of pairwise nonhomeomorphic Stein fillings of the same link, so that none of the Stein fillings is strongly diffeomorphic to a Milnor filling. We give one such construction below to prove the first part of Theorem 1.1. The second part of Theorem 1.1 follows from the discussion at the end of this section, where we extend star-shaped graphs that correspond to unexpected arrangements to a much wider collection of graphs of rational singularities with reduced fundamental cycle.

Theorem 7.20 *For every $N>0$ there exists a rational singularity with reduced fundamental cycle whose link (Y,ξ) admits at least N pairwise nonhomeomorphic simply connected Stein fillings, none of which is strongly diffeomorphic to any Milnor fiber. The link Y is given by a Seifert fibered space $Y=Y(2N+5,w)$ with sufficiently large weights w.*

Proof We will start with the arrangement $\mathcal{Q}$ of Figure 17 and augment it to other unexpected arrangements, using Lemma 7.18. First, we add more "vertical" and "horizontal" lines to the arrangement, so that it has N vertical and N horizontal lines, creating a

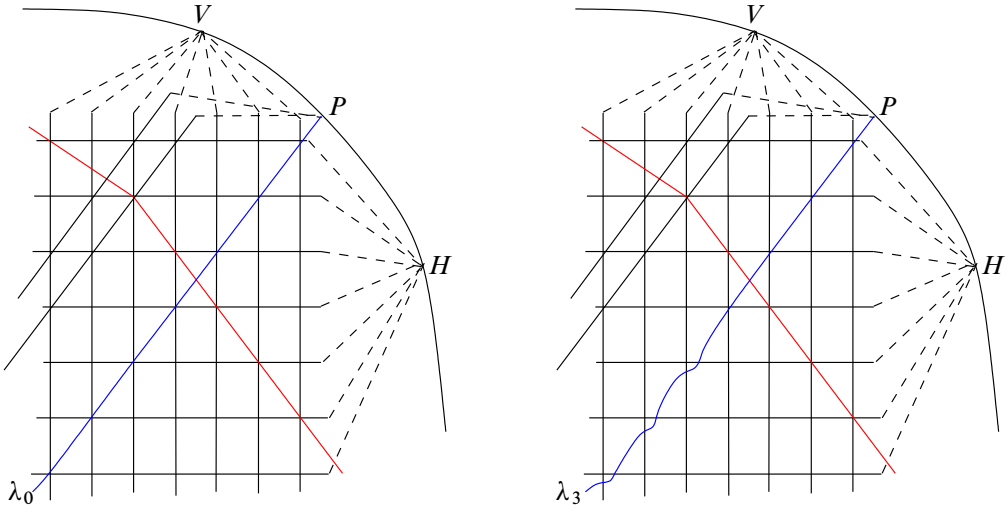


Figure 19: Pseudoline arrangements and fillings with different topology.

grid as shown in Figure 19. (We assume $N \geq 4$ as the $N = 4$ case fulfills the statement for lower values of N.) All "vertical" lines intersect at the point V, all horizontal lines intersect at the point H. The two diagonal lines ℓ_8 and ℓ_9 intersecting at P, the bent pseudoline ℓ_{10}, and the line at infinity ℓ_0 are present as in the arrangement $\mathcal{Q}$. Let $\mathcal{Q}'$ denote this arrangement. We will now produce $N+1$ unexpected arrangements $\mathcal{Q}'_k = \mathcal{Q}' \cup \lambda_k$ for $k = 0, 1, \ldots, N$, by adding to $\mathcal{Q}'$ different additional "diagonal" pseudolines $\lambda_0, \lambda_1, \ldots, \lambda_N$ passing through P; see Figure 19. Each arrangement $\mathcal{Q}'_k$ consists of $2N+5$ pseudolines. The pseudoline λ_0 is taken to be the main diagonal of the grid formed by the vertical and horizontal lines; it is a straight line in $\mathbb{R}\mathrm{P}^2$ passing through the point P. The pseudoline λ_1 differs from λ_0 in a small neighborhood of a single grid intersection: while λ_0 passes through the chosen intersection point of a vertical and a horizontal line, λ_1 intersects these two lines at distinct points. Similarly, λ_k differs from λ_0 in neighborhoods of k grid intersections and meets the corresponding vertical and horizontal lines at distinct points. Figure 19 shows the arrangements $\mathcal{Q}'_0 = \mathcal{Q}' \cup \lambda_0$ and $\mathcal{Q}'_3 = \mathcal{Q}' \cup \lambda_3$.

Now, consider the decorated germ given by a pencil of $2N+5$ lines, each with a weight greater than $2N+4$. We choose the weights to be greater than the number of intersection points on each line in any of the arrangements $\mathcal{Q}'_k$; obviously, taking weights greater than $2N+4$ suffices because each line intersects the other $2N+4$ lines once (in fact, $w \geq 2N+2$ suffices for this arrangement). Let (Y_N, ξ) be the contact link of the corresponding singularity. Similarly to the previous examples, Y_N is the

Seifert fibered space given by a star-shaped plumbing graph with $2N+5$ sufficiently long legs, with the central vertex having the self-intersection $-2N-6$ and all the other vertices self-intersection -2. By Corollary 7.9, each arrangement $\mathcal{Q}'_k$ yields a Stein filling W_k of (Y_N, ξ) which is not strongly diffeomorphic to any Milnor filling.

Finally, we argue that all fillings $W_0, W_1, \dots, W_N$ have different Euler characteristic. Each W_k has the structure of a Lefschetz fibration with the same planar fiber (a disk with $2N+5$ holes), but these Lefschetz fibrations have different numbers of vanishing cycles. Every time we replace a triple intersection of pseudolines in the arrangement by three double points (and arrange the marked points on the lines accordingly), the number of vanishing cycles decreases by 1. Indeed, three double points correspond to three vanishing cycles in the Lefschetz fibration (each enclosing two holes), while a triple intersection together with an additional free marked point on each of three lines corresponds to four vanishing cycles (one vanishing cycle enclosing three holes, the remaining three enclosing a single hole each). Thus, replacing a triple point by three double points corresponds to a lantern relation monodromy substitution, which in turn corresponds to a rational blow-down of a (-4) sphere. Therefore, $\chi(W_0) > \chi(W_1) > \cdots > \chi(W_N)$, as required. □

7.4 Generalizations

All our previous examples were given by singularities with star-shaped graphs where most vertices have self-intersection -2. It is not hard to obtain examples with much more general graphs, using the full power of Theorem 7.8: we add more smooth disks to an unexpected symplectic line arrangement.

Example 7.21 In the arrangement $\mathcal{Q}$ of Figure 17, replace the line ℓ_3 by several pseudolines that all pass through the same four intersection points. Note that because of multiple intersections, the result is no longer a pseudoline arrangement, but we still have a braided wiring diagram and can apply Proposition 5.5 to extend it to an arrangement of symplectic disks. In Figure 20, we take three curves replacing ℓ_3. In the decorated germ, the complex line corresponding to ℓ_3 will be replaced by 3 curvettas that are tangent to order 4 (and transverse to the other 10 branches of the germ). By (2-2), the weight of each new curvetta must be 5 or greater. We take the weights to be exactly 5 for the three new curvettas. Consider the symplectic curve arrangement given by the extension of the diagram in Figure 20, with marked points at all intersections and one additional free marked point on each of the three new curves (to account for higher weights). The resolution graph for $\mathcal{Q}$ is star-shaped with 11 legs. The

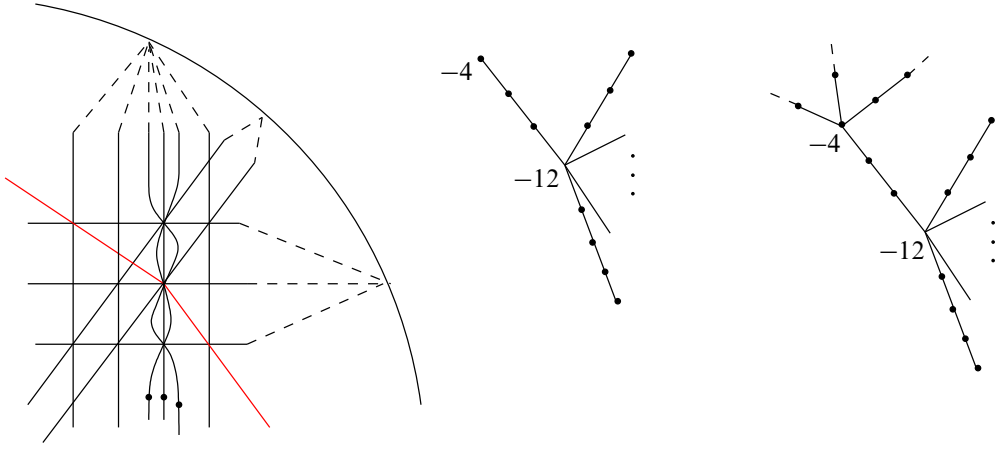


Figure 20: The pseudoline arrangement $\mathcal{Q}$ of Figure 17 is modified: the pseudoline ℓ_3 is replaced by three smooth curves with 4 intersections, as shown. There are 3 free marked points, one on each of the new curves; the rest of the marked points are the intersections in the diagram. The germ of the corresponding singularity has three curvettas tangent to order 4, each of weight 5, replacing one of the lines. The resolution graph of the corresponding singularity is shown in the middle of the figure. If the weights of the three tangent curvettas are taken to be higher, the graph will have additional branching as shown on the right. All unlabeled vertices have self-intersection -2.

self-intersection of the central vertex is -12 and all other self-intersections are -2. The legs of the resolution graph for $\mathcal{Q}$ with minimal weights had two legs of length 3, two of length 5, and seven of length 4. For this revised arrangement, the corresponding singularity has an augmented graph. Specifically, one of the legs of length 3 (which corresponded to ℓ_3) gains an additional vertex of self-intersection -4. If the three tangent curvettas have higher weights, so they have additional free marked points in the deformed arrangement, the -4 vertex becomes a branching point with 3 additional legs (each vertex on these legs has self-intersection -2). See Figure 20. By Theorem 7.8, the links of the corresponding singularities have unexpected Stein fillings.

In general, if we replace ℓ_3 with k curves commonly intersecting at the four points where ℓ_3 intersected other pseudolines as above, the additional vertex will have self-intersection $-k-1$ and increased weights will yield k additional legs with (-2) vertices.

Further, we can replace each of the k pseudolines by a bundle of curves that go through the same intersections.

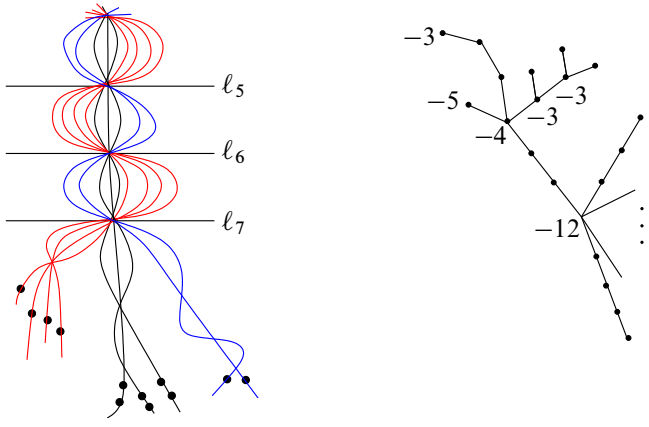


Figure 21: In the pseudoline arrangement $\mathcal{Q}$ of Figure 17, we replace ℓ_3 with a bundle of curves passing through the existing intersections of ℓ_3 with ℓ_5, ℓ_6, ℓ_7 and ℓ_0. (Only part of the arrangement is shown.) The additional curves create no extra intersections with the pseudolines of $\mathcal{Q}$. All the intersection points are marked, and there are additional free marked points that correspond to higher weights. In the resolution graph of the singularity, the leg corresponding to ℓ_3 is replaced by a tree with additional branching, as shown. All unlabeled vertices have self-intersection -2.

Example 7.22 Figure 21 shows a possible bundle replacing ℓ_3, instead of the bundle of three curves in the previous arrangement of Figure 20. All the new curves run C^1–close to and are isotopic to the original pseudoline, and they pass through the same intersection points with the other pseudolines. Within each bundle, the curves may have additional intersections, which lead to higher-order tangencies between the corresponding curvettas in the decorated germ. In particular, for the arrangement in Figure 21, the bundle of curves replacing ℓ_3 will have three subbundles of curves intersecting each other 4 times, and intersecting each of the other pseudolines once. One of these subbundles has four curves which intersect each other a total of 5 times, another has two curves which intersect a total of 7 times, and the third has two curves intersecting each other a total of 6 times, with an additional curve intersecting these two 5 times.

The corresponding decorated germ (with the weights given by the number of intersection points in the disk arrangement) encodes the singularity whose graph has more branching and some vertices with higher negative self-intersections, as shown in Figure 21. If we vary the incidence pattern of the additional curves (subject to the weight restrictions), we can obtain a number of unexpected Stein fillings with different topology.

Example 7.22 demonstrates how, once we have an unexpected symplectic line arrangement $\Gamma = \{\Gamma_i\}$, the star-shaped graph G of the corresponding singularity can be extended to arbitrarily complicated graphs of rational singularities with reduced fundamental cycle. The following proposition explains how to form these bundles in general from a given extension of the graph, completing the proof of Theorem 1.1. It is not hard to see that under the hypotheses of the proposition, the extended graph H corresponds to a singularity with reduced fundamental cycle.

Proposition 7.23 *Let G be the star-shaped resolution graph corresponding to the surface singularity associated to an unexpected symplectic line arrangement with minimal possible weights. Let I be the set of leaves of G, and let $\{G_i\}_{i\in I}$ be a collection of (possibly empty) negative definite rooted trees; assume that G and G_i have no (-1) vertices.*

Consider a graph H constructed by attaching to G the rooted trees G_i, $i \in I$, so that the root of G_i is connected to the leaf u_i by a single edge. Assume that the resulting graph H satisfies condition (2-1). Let (Y,ξ) be the link of a rational surface singularity with reduced fundamental cycle whose dual resolution graph is H.

Then (Y,ξ) admits a Stein filling which is not strongly diffeomorphic to any Milnor filling.

Remark 7.24 Proposition 7.23 provides a fairly general class of rational surface singularities with reduced fundamental cycle which admit unexpected fillings. The construction can be further generalized to include variations in the bundling structure and to apply to more general graphs G as the input. Despite all variations, getting rid of the (-2) vertices in the resolution graph seems difficult. Indeed, we could add a curve intersecting ℓ_3 only twice in Example 7.22, which would lower the self-intersection to (-3) for one of the vertices on the leg of the star-shaped graph G. However, such a curve would intersect the other pseudolines in the arrangement $\mathcal{Q}$ at new points. This would increase the weights on the curvettas corresponding to these other pseudolines, producing free marked points and yielding additional (-2) vertices elsewhere in the graph. In fact, we already know from Theorem 1.2 that our strategy must have limitations, as there are no unexpected fillings when each vertex of the resolution graph has self-intersection -5 or lower.

Proof of Proposition 7.23 The initial unexpected symplectic line arrangement $\{L_i\}$ consists of symplectic lines associated to the legs of the star-shaped graph G. As above,

let u_i denote the valency 1 vertex of the leg that corresponds to L_i. Choose a braided wiring diagram for the symplectic line arrangement such that a symplectic line L_i corresponds to the wire γ_i. The braided wiring diagram should be chosen so that γ_i contains all the marked points of L_i (including free points). We will replace each wire γ_i with a bundle of curves (with intersections but no braiding between the components of the bundle) constructed according to the tree G_i, as follows.

All curves in the i^{th} bundle must intersect at all marked points on γ_i. We will specify the additional intersections and explain how to determine the number of curves and free marked points in the bundle. The bundle will be described recursively, via its subbundles and iterative (sub)^k–bundles, which we determine by moving through the graph G_i. We start at the root and move upward in the graph G_i with respect to the partial order induced by the root, stopping when we either reach either a vertex v_0 of self-intersection number $-s_0$ for $s_0 \geq 3$, or exhaust the graph G_i.

By condition (2-1), (-2) vertices can only occur in a linear chain. Thus, if we never reach a vertex with self-intersection $-s_0$ for $s_0 \geq 3$, then all vertices of G_i have self-intersection -2 (and G_i is a linear chain). Suppose there are $r_0 \geq 0$ such (-2) vertices. In that case, the bundle for G_i should consist of only a single curve, but with $r_0 \geq 0$ additional free points. (The weights of the decorated germ increase accordingly.)

If there exists a vertex v_0 of self-intersection $-s_0$ for $s_0 \geq 3$ after passing through a linear chain of r_0 vertices of self-intersection -2, then the bundle will consist of exactly $s_0 - 1$ nonempty subbundles. The subbundles will be described as we travel further along G_i. We require that all curves in the bundle intersect exactly r_0 additional times (where each of these r_0 intersection points gets marked) and increase the weight of each curve by $r_0 + 1$, yielding one additional free marked point on each curve. Two curves in different subbundles will not intersect at any additional points beyond those specified so far.

Note that v_0 can have at most $s_0 - 1$ vertices directly above it in G_i, since its valency is at most s_0. In particular, G_i itself is built by attaching $s_0 - 1$ (potentially empty) trees onto the subgraph $\{v \leq v_0\} \subset G_i$. We associate the $s_0 - 1$ subbundles to these $s_0 - 1$ rooted trees $G_1^1, \ldots, G_{s_0-1}^1$, which may be empty or nonempty. (The partial order on G induced by its root induces a partial order and root on each G_j^1.)

Now we will create subbundles and their subsubbundles by iteratively repeating a slight modification of the process above. For each tree G_j^1, we construct a subbundle as follows. Starting at the root of G_j^1, we again have a linear chain of $r_1 \geq 0$ vertices

with self-intersection -2, which either exhausts the graph G_j^1 or ends in a vertex v_1 of self-intersection number $-s_1$ for $s_1 \geq 3$. (Note that r_1 and s_1 depend on j, but we drop this index to avoid further notational clutter.) If we are in the first case, where there is no such vertex v_1, the subbundle associated to G_j^1 will consist of a single curve with r_1 additional free marked points. If we are in the second case, where the chain of length r_1 of (-2) vertices ends at a vertex v_1 with self-intersection $-s_1$ with $s_1 \geq 3$, the subbundle itself will be a union of $s_1 - 1$ nonempty subsubbundles, intersecting at $r_1 + 1$ additional points. (Accordingly, the weights increase by $r_1 + 1$, but no new free marked points are added.) Two curves in different subsubbundles will not intersect at any additional points beyond those previously specified.

The $s_1 - 1$ subsubbundles correspond to the $s_1 - 1$ potentially empty trees $G_1^2, \dots, G_{s_1-1}^2$ attached above v_1. We determine these subsubbundles by iteratively repeating this process, where G_l^2 takes the role of G_j^1 and the subsubbundle takes the role of the subbundle. The $(\text{sub})^k$–bundles will generally have $(\text{sub})^{k+1}$–bundles, leading to additional iterations of the procedure. The situation where a $(\text{sub})^k$–bundle does not have a $(\text{sub})^{k+1}$–bundle is when the $(\text{sub})^k$–bundle consists of a single component (as in the first case of the procedure). Since the graph is finite, there will be a finite number of iterations, so this process will eventually describe the bundle completely.

Having constructed such bundles individually for each G_i, we now superimpose them onto the wires γ_i as satellites to get a new braided wiring diagram by inserting them into a small neighborhood of γ_i so that each wire of the bundle is C^1–close to the original wire γ_i. Recall that all intersection points between wires are marked in the original diagram, and all curves from the i^{th}–bundle are required to intersect at all marked points. It follows that curves from the different bundles are allowed to intersect *only* at the marked points of the original diagram.

We can apply Proposition 5.5 to extend the new braided wiring diagram to an arrangement Γ of symplectic disks. We claim that via Lemma 3.2, the resulting arrangement Γ provides a Stein filling for the link of the singularity with the resolution graph H. To check the claim, we need to show that the open book decomposition on the boundary of the Lefschetz fibration constructed from Γ supports the canonical contact structure for the link associated to H. Recall that H is associated to a decorated germ $\mathcal{C}^H$ with smooth branches, by attaching (-1) vertices and curvettas and blowing down. We will show that Γ is related by a smooth graphical homotopy to another decorated germ $\mathcal{C}$, which is topologically equivalent to $\mathcal{C}^H$. The topological type of $\mathcal{C}$ will be determined by the intersections and marked points in Γ: the order of tangency between

two components in $\mathcal{C}$ is equal to the number of intersections between the corresponding components of Γ. The weight on each curve is the total number of marked points on the corresponding disk of Γ, including intersections and free marked points. After showing that Γ and $\mathcal{C}$ are related by a smooth graphical homotopy, we will verify that $\mathcal{C}$ and $\mathcal{C}^H$ are topologically equivalent (with corresponding weights), to conclude that the open book decompositions are equivalent.

To relate Γ and $\mathcal{C}$, we first construct a smooth graphical homotopy from Γ to a "pencil of the bundles". In the pencil of the bundles, all curves will intersect at one point, and curves from different bundles do not intersect anywhere else, but curves from the same bundle may intersect at other points along the corresponding line. We can use a smooth graphical homotopy of the original symplectic line arrangement $\{L_i\}$ to a pencil as a guide to build the required homotopy of Γ, because each bundle is C^1–close to the corresponding symplectic line inside the chosen Milnor ball. Essentially, at this step we treat each bundle as a whole, bringing different bundles together without perturbing curves inside each bundle. More precisely, we satellite the bundle onto the family of wiring diagrams corresponding to the smooth graphical homotopy of the symplectic lines to the pencil. The intersection points within a bundle will remain distinct in this smooth graphical homotopy. At intermediate times during the homotopy, we allow many additional intersection points in the arrangement, as curves from different bundles will intersect outside the common marked intersections.

Next, we show that each bundle can be homotoped so that all the intersections come together to high-order tangencies. Let Γ^i denote the i^{th} bundle constructed above, and let $\mathcal{C}^i$ denote the curves in the germ $\mathcal{C}$ corresponding to those in Γ^i. To show that Γ^i and $\mathcal{C}^i$ are related by a smooth graphical homotopy, it suffices to check that they have the same boundary braid. To verify this, we observe that the subbundling structure looks like the nested structure produced by the Scott deformation of $\mathcal{C}^i$ as in the proof of Proposition 4.1. The bundle, as drawn in $\mathbb{R}^2$, provides a wiring diagram which is planar isotopic to the wiring diagram of the Scott deformation, and thus their braid monodromy is the same. As a consequence, each bundle Γ^i is related by a smooth graphical homotopy to $\mathcal{C}^i$. Applying these homotopies to all bundles, we see that Γ is related to $\mathcal{C}$ by a smooth graphical homotopy, and their induced open books agree.

Now, we need to check that $\mathcal{C}$ and $\mathcal{C}^H$ are topologically equivalent. To this end, we will compare the weights and the pairwise orders of tangency between curvettas in the two germs. For $\mathcal{C}$, these quantities are computed from the intersections and marked points in Γ, while Remark 2.6 shows how to compute them from the graph H.

First, we make a few observations to relate the curvettas on the graph H to the bundling construction above. Before the star-shaped graph G is extended, the lines L_i correspond to the legs of the graph. For each i, the i^{th} leg is a chain of (-2) vertices, with an end vertex u_i. We attach a single (-1) vertex to u_i and put a curvetta on this vertex; this curvetta gives rise to the line L_i. By Remark 2.6, the weight of L_i is $1 + l(u_0, u_i)$, where u_0 is the root of G. In this case, the root has been chosen to be the center of the star-shaped graph.

When G_i is nonempty, the symplectic line L_i is replaced by a collection of m_i curves (we compute m_i below) in the germ associated to H. These new curves come from curvettas on the additional (-1) vertices attached to G_i. For each $v \in G_i$, $(v \cdot v + a(v))$ additional (-1) vertices are attached to v, and each (-1) vertex has a curvetta attached, thus

$$m_i = -\sum_{v \in G_i} (v \cdot v + a(v)),$$

as in Proposition 2.4. Note that m_i agrees with the number of curves in the bundle Γ_i constructed above for the graph G_i. This is because the subbundling process terminates when you reach a $(\text{sub})^k$–bundle which is a single component. This occurs when the $(\text{sub})^k$–bundle corresponds to a $(\text{sub})^k$–tree consisting of only $r \geq 0$ vertices of self-intersection -2. When $r > 0$, this means that there is a (-2) vertex leaf which contributes one to m_i, and when $r = 0$, this means there is a $(-s)$ vertex v with fewer than $(s-1)$ branches above it, and there are correspondingly $-(v \cdot v + a(v)) = s - a(v)$ such $(\text{sub})^k$–bundles, each consisting of a single curve.

Now, let C_x be one of the curvettas for the graph H, and let $\widetilde{v}_x$ be a vertex of G such that C_x intersects a (-1) vertex attached to $\widetilde{v}_x$. According to Remark 2.6, the weight of C_x according to the graph H is $1 + l(\widetilde{v}_x, u_0)$, where $l(\widetilde{v}_x, u_0)$ counts the number of vertices in the path from the root u_0 of G to the vertex $\widetilde{v}_x$. This path consists of several parts. From the original graph G, the path contains the $l(u_i, u_0)$ vertices connecting the root u_0 to the vertex u_i where G_i is attached. Next, there are vertices from G_i, which can be organized into $(K+1)$ chains as shown in Figure 22. For $0 \leq k \leq K-1$, the k^{th} chain consists of $r_k \geq 0$ vertices of self-intersection (-2), followed by a vertex of self-intersection $-s_k < -2$. Finally, there may be a last chain of (-2) vertices, of length $r_K \geq 0$, such that $\widetilde{v}_x$ is its last vertex. (If $\widetilde{v}_x \cdot \widetilde{v}_x < -2$, then $r_K = 0$.) Therefore,

$$1 + l(\widetilde{v}_x, u_0) = 1 + l(u_i, u_0) + r_K + \sum_{k=0}^{K-1} (r_k + 1).$$

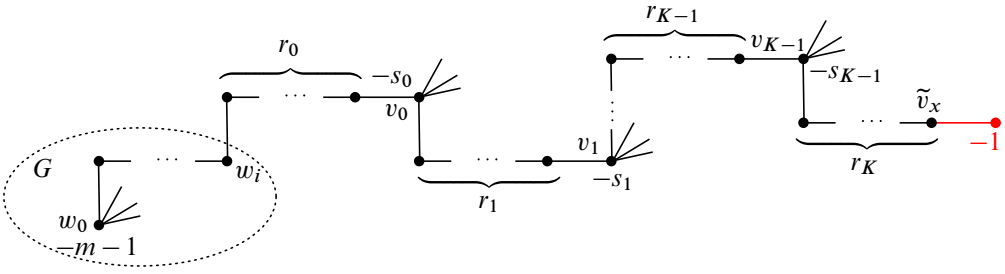


Figure 22: How to compute the weights from the graph G following the proof of Proposition 7.23.

On the other hand, in the construction of the bundle, the initial weight on each curve begins at $1+l(u_i,u_0)$. For each iterative $(\text{sub})^k$–bundle it is included in, the weight is increased by r_k+1, until we reach a stage K where the $(\text{sub})^K$ graph consists of $r_K \geq 0$ vertices, all of self-intersection -2. For this K^{th} stage, the weight is increased by r_K (the increase is associated to free marked points). Therefore, the total weight on C_x will be

$$w(C_x) = 1 + l(u_i,u_0) + r_K + \sum_{k=0}^{K-1}(r_k+1),$$

which agrees with $1+l(\widetilde{v}_x,u_0)$, as required.

Next, we compare the orders of tangency between the curves. According to Remark 2.6, the order of tangency between two components C_x and C_y is $\rho(\widetilde{v}_x,\widetilde{v}_y;u_0)$, the number of common vertices in the path from $\widetilde{v}_x$ to u_0 with the path from $\widetilde{v}_y$ to u_0. By condition (2-1), the vertex v_L where these two paths diverge has self-intersection $-s_L$ for $s \geq 3$. See Figure 23. The path from u_0 to v_L includes the path from u_0 to u_i in G. This contributes $l(u_i,u_0)$ vertices. The path continues into G_i, with sequential chains of r_k vertices of self-intersection (-2), each ending in a vertex v_k of self-intersection $-s_k < -2$, for $0 \leq k \leq L$. Therefore,

$$\rho(\widetilde{v}_x,\widetilde{v}_y;v_0) = l(u_i,u_0) + \sum_{k=0}^{L}(r_k+1).$$

On the other hand, in the bundle construction, the curves C_x and C_y lie in two distinct $(\text{sub})^{L+1}$–bundles created for two of the distinct trees lying above vertex v_L. No intersections between C_x and C_y will be created after the L^{th} stage. At the beginning of the bundle construction, all curves are required to intersect $1+l(u_i,u_0)$ times. All other intersections between C_x and C_y are created in the procedure above at some

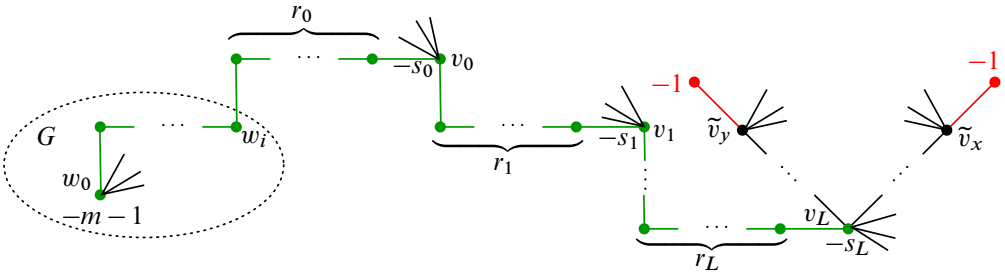


Figure 23: How to compute the tangencies from the graph G following the proof of Proposition 7.23.

iteration k, $0 \le k \le L$. At the $k = 0$ stage, we add r_0 intersections between C_x and C_y. At stage k for $1 \le k \le L$, we add additional $r_k + 1$ intersections between C_x and C_y. Therefore the total number of intersections between C_x and C_y is

$$1 + l(u_0, u_i) + r_0 + \sum_{k=1}^{L} (r_k + 1),$$

which agrees with $\rho(\widetilde{v}_x, \widetilde{v}_y; v_r)$.

To complete the proof, observe that the arrangement Γ contains the original unexpected symplectic line arrangement as a subarrangement (choose a single component of each bundle). By Theorem 7.8, we obtain unexpected Stein fillings of the link of the singularity corresponding to the graph H. □

8 Further comments and questions on curvetta homotopies

In the previous section we showed that Stein fillings of the link of a singularity do not always arise from the Milnor fibers, even for the simple class of rational singularities with reduced fundamental cycle. Our examples of unexpected Stein fillings come from curvetta arrangements that do not arise as picture deformations of the decorated germ representing the singularity, although these arrangements are still related to the decorated germ through a smooth graphical homotopy. In this section, we make a detailed comparison of de Jong and van Straten's picture deformations (Definition 2.7) with smooth graphical homotopies (Definition 3.1). Observe that the two notions differ in several essential ways. Indeed, the curvetta branches are required to be algebraic in the former, and just smooth in the latter; positivity of all intersections and the weight restrictions must hold at all times during a picture deformation but only at the end of a graphical homotopy; the topology of the arrangement may change at nonzero times

	smooth graphical homotopy	picture deformation
type of curvetta branch C_j^t	smooth graphical disk	disk given by (germ of) algebraic curve
topology of curvetta arrangement	may change with time	remains the same
weight restrictions: C_j^t has at most w_j intersections	only hold for final arrangement, may be violated during homotopy	hold at all times
positivity of intersection points: $C_i^t \cdot C_j^t > 0$	only hold for the final arrangement, may be violated during homotopy	hold at all times

Table 1

during graphical homotopy but not during a picture deformation. This is summarized in Table 1. We will explore each of these aspects and their role in differentiating Stein fillings from Milnor fibers. The most important aspect seems to be the topology of the curvetta arrangement, and whether it is allowed to vary during the homotopy.

8.1 Algebraic versus smooth

The first difference between picture deformations and homotopies is that a smooth graphical homotopy includes curvettas which need not be complex algebraic curves, either during the course of the homotopy or at the end of the homotopy. It turns out that this is not the key aspect contributing to the difference between Milnor fillings and Stein fillings in our examples. Indeed, adding higher-order terms, one can produce some surprising curvetta arrangements. Because the curvettas are open algebraic disks, possibly given by high-degree algebraic equations, curvetta arrangements can be more general than arrangements of complex lines or global algebraic curves. To illustrate, we recall the example of the pseudo-Pappus arrangement from [27]; see Figure 24.

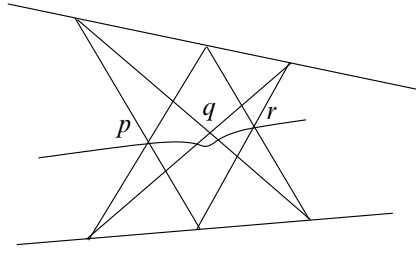


Figure 24: The pseudo-Pappus arrangement.

Example 8.1 [23; 27] The classical Pappus arrangement consists of 9 lines; we have already discussed this arrangement in Example 7.10. By the Pappus theorem, the points p, q, r in the middle of Figure 24 are collinear. In the pseudo-Pappus arrangement, the line through these three points is replaced by a bent pseudoline that passes through two points but not through the third. The pseudo-Pappus arrangement cannot be realized by complex lines. However, the bent pseudoline can be given by a graph of a high-degree polynomial whose additional intersections with the other lines occur sufficiently far outside the ball we restrict to. Thus, the pseudo-Pappus arrangement can be realized by higher-degree open algebraic curves. In fact, as mentioned in [27], the pseudo-Pappus arrangement arises as a picture deformation of the pencil of 9 lines, with the weights of each line given by the number of intersection points on the corresponding line in the arrangement. The picture deformation can be obtained by adding small higher-order terms to the linear deformation of the pencil to the classical Pappus arrangement. Thus, the pseudo-Pappus arrangement gives rise to Milnor fibers of smoothings of the singularities given by the corresponding decorated pencil of 9 lines.

In fact, all of the fillings produced via arrangements of real pseudolines can be obtained from an algebraic curvetta arrangement which can be deformed by a polynomial homotopy (through algebraic curves) to a pencil of lines. (However, this family does not constitute a picture deformation because the topology may vary at different $t \neq 0$, and the weight constraints may fail at intermediate times.) Note that we only consider a portion of the algebraic curves in a chosen ball surrounding the origin. In particular, the algebraic curves may intersect additional times outside of this ball, but we do not need to count such intersections in the incidence data of our arrangement.

Proposition 8.2 *Let $\Lambda = \{\ell_1, \dots, \ell_m\}$ be an arrangement of real pseudolines in $\mathbb{R}^2$. Then there exists a family of complex algebraic curves $\{\Gamma_1^t, \dots, \Gamma_m^t\}$, given by polynomial equations*

$$\Gamma_i^t = \{y = p(x,t)\}$$

and a smoothly embedded closed 4–ball $B \subset \mathbb{C}^2$, such that $\{\Gamma_1^t, \dots, \Gamma_m^t\}$ is a symplectic line arrangement in B (with intersections in the interior of B) for every $t \in [0,1]$, where

- *$B \cap (\Gamma_1^0 \cup \cdots \cup \Gamma_m^0)$ has the incidences of a pencil of lines, and*
- *$B \cap (\Gamma_1^1 \cup \cdots \cup \Gamma_m^1)$ is isotopic in B to the symplectic extension of the pseudoline arrangement $\ell_1 \cup \cdots \cup \ell_m$ given by Proposition 5.5.*

Before proving the proposition, we discuss its consequences.

Remark 8.3 Consider an arbitrary pseudoline arrangement $\ell_1, \dots, \ell_m$ and the corresponding symplectic line arrangement $\{\Gamma_1, \dots, \Gamma_m\}$. By Proposition 7.4, this arrangement gives Stein fillings of the spaces $(Y(m; w_1, \dots, w_k), \xi)$ whenever the weights satisfy inequalities $w_k \geq w(\Gamma_k)$ for $k = 1, \dots, m$. Let $\Gamma^t = \{\Gamma_1^t, \dots, \Gamma_m^t\}$ be a polynomial homotopy between a pencil of lines and the arrangement $\{\Gamma_1, \dots, \Gamma_m\}$; such a homotopy always exists by Proposition 8.2. A priori, the homotopy may violate the weight constraints: at some moment t, the number of intersections may increase, so that $w(\Gamma_k^t) > w_k$. (In fact, the homotopy constructed in Proposition 8.2 converts all multiple intersections into double points and thus creates a lot of additional intersections.) However, since Γ_k^t intersects each of the other $m-1$ components exactly once, $w(\Gamma_k^t)$ will never exceed $m-1$. Thus, if $w_k \geq m-1$ for all k, any homotopy as above will satisfy the weight constraints. By construction, intersections between any two components Γ_i^t and Γ_j^t remain positive for all t. Thus, the homotopy Γ^t satisfies the requirements of the first, third and fourth lines in Table 1, sharing these properties with picture deformations, but it changes the topology of the arrangement. Accordingly, the arrangement $\{\Gamma_1^t, \dots, \Gamma_m^t\}$ gives a Stein filling W_t of $(Y(m; w_1, \dots, w_k), \xi)$ for every t, and W_t carries a Lefschetz fibration as in Lemma 3.2, but the topology of the fillings W_t changes with t. Note also that for small $t > 0$, the defining polynomials for Γ_k^t give an unfolding, and thus a 1–parameter deformation of $\mathcal{C}$. Equipped with marked points, this gives a picture deformation. Therefore, for small $t > 0$ the Stein filling W_t is given by a Milnor fiber. As t increases and the topology of the arrangement changes, we obtain new fillings W_t, which may not be realizable by Milnor fibers. We will consider a specific example of such a topology change in Section 8.3.

The conclusion we wish to draw here is that the difference between algebraic curves and smooth curves is not essential to our counterexamples, as we can realize the corresponding symplectic line arrangements by complex algebraic curves and construct polynomial homotopies. The positivity of intersections and the weight constraints can often be trivially satisfied, although we further discuss the role of weights in Section 8.4. In fact, the important difference comes from the second aspect in Table 1, namely smooth graphical homotopies can vary their topology and singularities in various different ways during the homotopy, whereas picture deformations must maintain the same topology for all nonzero parameters t.

We now turn to the proof of Proposition 8.2. Given any pseudoline arrangement, it can be isotoped in $\mathbb{R}^2$ to be in a standard wiring diagram form, with the following properties. Each pseudoline is graphical, $\ell_i = \{y = f_i(x)\}$. Away from intersection

points, each pseudoline is horizontal with $f_i(x) = 2\delta n$ for some integer $1 \le n \le m$ and a fixed constant $\delta > 0$. There are disjoint intervals $(a_1, b_1), \dots, (a_r, b_r)$ at which $f_i(x)$ is nonconstant, such that there is a unique point in each interval (a_k, b_k) at which ℓ_i intersects other pseudolines. Furthermore, we ask that f_i and f_j are linear whenever $|f_i(x) - f_j(x)| < \delta$, and each $f_i(x)$ is monotonic in each interval (a_k, b_k). We will assume after a planar isotopy of Λ that our pseudoline arrangement is initially given in this form. To construct our algebraic family, we first require a smooth family of pseudolines connecting this given pseudoline arrangement in standard wiring diagram form to a pencil, and satisfying a quantitative transversality property, as follows.

Lemma 8.4 *Let $\Lambda = \{\ell_1, \dots, \ell_m\}$ be an arrangement of real pseudolines in $\mathbb{R}^2$ in standard wiring diagram form with constant δ, such that all intersections occur in $[-M, M] \times \mathbb{R}$. Then there exist smooth functions $f_i : [-M, M] \times [0, 1] \to \mathbb{R}$ with the following properties*:

(1) *$\ell_i = \{y = f_i(x, 1)\}$, ie at time 1 the graphs of the functions give the chosen pseudoline arrangement.*

(2) *$f_i(x, 0) = c_i x$, ie at time 0 the graphs of the functions give a linear pencil.*

(3) *For any $t_0 \in [0, 1]$ and any $i \neq j$, there is a unique point $\overline{x} \in [-M, M]$ such that $f_i(\overline{x}, t_0) = f_j(\overline{x}, t_0)$ and an interval $(a, b) \subset [-M, M]$ containing $\overline{x}$ such that $|f_i(x, t_0) - f_j(x, t_0)| < \delta$ if and only if $x \in (a, b)$, ie the pseudolines remain at least distance δ apart except in a neighborhood of their unique intersection.*

(4) *For any $t_0 \in [0, 1]$ and any $x_0 \in [-M, M]$ such that $|f_i(x_0, t_0) - f_j(x_0, t_0)| < \delta$, we have that*

$$\left| \frac{\partial f_i}{\partial x}(x_0, t_0) - \frac{\partial f_j}{\partial x}(x_0, t_0) \right| > \eta := \frac{\delta}{2M},$$

ie whenever the pseudolines become close enough to intersect, their slopes are quantitatively far enough from each other to ensure isolated transverse intersections.

Proof Note that when the original pseudoline arrangement $\{\ell_i\}$ is in standard wiring diagram form, it does satisfy property (4) of the lemma when $t_0 = 1$. This is because whenever $|f_i(x, 1) - f_j(x, 1)| < \delta$, the function $f_i - f_j$ is linear, and it interpolates a height difference greater than δ over an interval smaller than $2M$, so its slope is greater than η.

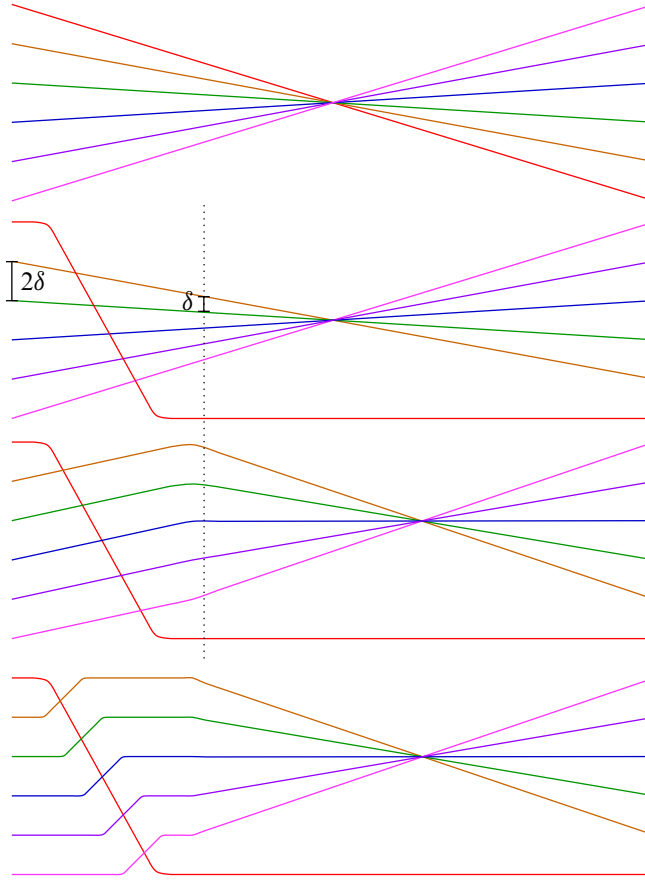


Figure 25: Key move used to construct a family of pseudolines, slightly modified from [57].

It was proven in [57, Proposition 6.4] that any arrangement of pseudolines in standard wiring diagram form can be related through a family of pseudolines to a pencil. In that paper, what is needed is that the pseudolines maintain transverse intersections throughout the family, whereas we need a quantitative measure of this transversality. We demonstrate here that this stronger condition is in fact satisfied by the family in [57].

We briefly recall the key aspects in the construction of the family and refer the reader to [57, Proposition 6.4] for further details. This family is graphical and thus can be written as $\ell_i^t = \{y = f_i(x,t)\}$ for $i = 1,\dots,m$, where $\ell_i^1 = \ell_i$. The key move to modify the pseudoline arrangement into a pencil through a family is shown in Figure 25; this figure is a slight modification of that appearing in [57, Figure 8]. This move is used

Figure 26: First reordering move.

iteratively to break up k–tuple points into a sequence of double points in a particular order. This procedure can be reversed to form an m–tuple point from a collection of appropriately ordered double points at the end to obtain a pencil. The order of the double points can be modified through the moves shown in Figures 26 and 27, by a classical theorem of Matsumoto and Tits [37].

If a pseudoline arrangement satisfies the transversality property (4) before the move in Figure 27, then it will continue to satisfy the same property throughout the move, because the relative slopes remain the same; only the interval where they occur is translated.

For the move from Figure 26, this can be realized using Figure 25 once in reverse to form a triple point, and then again in the forwards time direction, but mirrored to break up the triple point in the opposite manner; see [57, Figure 10]. Therefore it suffices to ensure that property (4) is satisfied throughout the move shown in Figure 25. Indeed, throughout this move, whenever a pair of pseudolines have height difference less than δ (recall that the spacing between the heights of the strands at the left and right ends of the figure is 2δ), both pseudolines are linear in this interval. The difference of pairwise slopes whenever $|f_i(x,t)-f_j(x,t)|<\delta$ is always greater than η throughout this family, because each crossing changes the difference in f_i-f_j by at least 2δ across the interval, whereas the interval has length at most $2M$. Moreover, this move preserves the property that there is a unique interval at which a given pair satisfies $|f_i(x,t)-f_j(x,t)|<\delta$. □

Proof of Proposition 8.2 We use the functions $\{f_i(x,t)\}$, representing a family of pseudolines through their graphs at a fixed time t, and approximate these by real polynomials intersecting in somewhat controlled ways. We assume that $x\in[-M,M]$ and that $M\geq 1$. Our final pseudoline arrangement is given by $\ell_i=\{y=f_i(x,1)\}$. Let $x_1,\dots,x_n$ be the points at which $f_i(x_k,1)=f_j(x_k,1)$ for some $i\neq j$.

Figure 27: Second reordering move.

Let $\varepsilon > 0$. Let $\zeta = \min\{1, \min_{i\neq j}\{|x_i - x_j|\}\}$. In particular, $\zeta \leq 1$.

Using the Stone–Weierstrass approximation theorem, choose polynomials $\widetilde{p}_i(x,t)$ such that

$$\left|\frac{\partial f_i}{\partial x}(x,t) - \widetilde{p}_i(x,t)\right| < \frac{\varepsilon\zeta^{n-1}}{4n^2(2M)^n}.$$

Then by integrating $\widetilde{p}_i(x,t)$ and shifting by a constant, we can find $\overline{p}_i(x,t)$ such that $(\partial\overline{p}_i/\partial x)(x,t) = \widetilde{p}_i(x,t)$ and

$$|\overline{p}_i(x,t) - f_i(x,t)| < \frac{\varepsilon\zeta^{n-1}}{4n^2(2M)^{n-1}}.$$

Now for $k = 1,\dots,n$ let

$$a_k^i = \frac{(f_i(x_k,1) - \overline{p}_i(x_k,1))}{(x_k - x_1)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)}.$$

Let $a_0^i = \overline{p}_i(0,0)$. Define

$$\begin{aligned} p_i(x,t) = \overline{p}_i(x,t) &+ a_0^i(t-1) + a_1^i t(x-x_2)\cdots(x-x_n) \\ &+ a_2^i t(x-x_1)(x-x_3)\cdots(x-x_n) + \cdots + a_n^i t(x-x_1)\cdots(x-x_{n-1}). \end{aligned}$$

Then for every $k = 1,\dots,n$, we have that $p_i(x_k,1) = f_i(x_k,1)$ and $p_i(0,0) = p_j(0,0) = 0$ for all i and j. In particular, for every multi-intersection point of the pseudolines $\ell_1,\dots,\ell_m$, there is a multi-intersection point of the corresponding $\{p_1(x,1)=0\},\dots,\{p_m(x,1)=0\}$. We will show that the curves $\gamma_1^{t_0} := \{p_1(x,t_0)=0\}$, $\dots,\gamma_m^{t_0} := \{p_m(x,t_0)=0\}$ form a pseudoline arrangement at each time t_0 (namely every pair of components intersects exactly once). In particular, this suffices to show that at $t_0 = 1$, the algebraic arrangement has the same intersections as the smooth pseudoline arrangement. For this, we use the bounds

$$\begin{aligned} |p_i(x,t) - f_i(x,t)| &\leq |p_i(x,t) - \overline{p}_i(x,t)| + |\overline{p}_i(x,t) - f_i(x,t)| \\ &\leq a_0^i + \sum_{k=1}^n a_k^i (2M)^{n-1} + \frac{\varepsilon\zeta^{n-1}}{4n^2(2M)^{n-1}} \\ &\leq \frac{\varepsilon\zeta^{n-1}}{4n^2(2M)^{n-1}} + \sum_{k=1}^n \frac{\varepsilon}{4n^2(2M)^{n-1}}\cdot(2M)^{n-1} + \frac{\varepsilon\zeta^{n-1}}{4n^2(2M)^{n-1}} \\ &< \varepsilon. \end{aligned}$$

We can similarly bound the difference of the derivatives with respect to x,

$$\left|\frac{\partial p_i}{\partial x}(x,t) - \frac{\partial f_i}{\partial x}(x,t)\right| \leq a_0^i + \sum_{k=1}^n a_k^i n(2M)^{n-2} + \frac{\varepsilon\zeta^{n-1}}{4n^2(2M)^{n-1}} < \varepsilon.$$

Now we want to show that the graphs $\lambda_i^t := \{y = p_i(x,t) \mid x \in [-M,M]\}$ provide a family of algebraic pseudoline arrangements whose incidences agree with those of $\{\ell_i\}$ at $t=1$, and agree with the incidences of a pencil at $t=0$. We will use the intersection and quantitative transversality properties of Lemma 8.4 to verify that for each time $t_0 \in [0,1]$, there is a unique transverse intersection between $\lambda_i^{t_0}$ and $\lambda_j^{t_0}$ where $p_i(x,t_0) = p_j(x,t_0)$ for $x \in [-M,M]$.

Since we could choose $\varepsilon > 0$ arbitrarily in the argument above, we now set $\varepsilon = \min\{\frac{1}{3}\delta, \frac{1}{3}\eta\}$. For each $t_0 \in [0,1]$ and each pair $i \neq j$, there is an interval (a,b) such that for $x \in [-M,M] \setminus (a,b)$, we have $|f_i(x,t_0) - f_j(x,t_0)| \geq \delta$. By the triangle inequality, for $x \in [-M,M] \setminus (a,b)$,

$$|p_i(x,t_0) - p_j(x,t_0)| \geq |f_i - f_j| - |f_i - p_i| - |p_j - f_j| > \delta - 2\varepsilon \geq \tfrac{1}{3}\delta > 0.$$

Therefore $p_i(x,t_0) \neq p_j(x,t_0)$ for $x \in [-M,M] \setminus (a,b)$. Now for $x \in (a,b)$, we have that $|f_i(x,t_0) - f_j(x,t_0)| < \delta$, so by the last property of Lemma 8.4,

$$\left| \frac{\partial f_i}{\partial x}(x,t_0) - \frac{\partial f_j}{\partial x}(x,t_0) \right| > \eta.$$

Again by the triangle inequality and the bounds above we get that

$$\left| \frac{\partial p_i}{\partial x}(x,t_0) - \frac{\partial p_j}{\partial x}(x,t_0) \right| > \tfrac{1}{3}\eta.$$

Since the difference of the derivatives is bounded away from zero, this implies that there can be *at most* one value $x \in (a,b)$ such that $p_i(x,t_0) = p_j(x,t_0)$.

Because $f_i(x,t_0)$ and $f_j(x,t_0)$ intersect once in the interval (a,b) and their distance is δ at the endpoints a and b, up to switching i and j, we have $f_i(a,t_0) - f_j(a,t_0) = \delta = f_j(b,t_0) - f_i(b,t_0)$. Since $|p_i(x,t) - f_i(x,t)| < \frac{1}{3}\delta$ and $|p_j(x,t) - f_j(x,t)| < \frac{1}{3}\delta$, this implies that $p_i(a,t_0) > p_j(a,t_0)$ and $p_j(b,t_0) > p_i(b,t_0)$. Therefore there must exist *at least* one value $x \in (a,b)$ such that $p_i(x,t_0) = p_j(x,t_0)$. Therefore the arrangement $\{\lambda_i^{t_0}\}_{i=1}^m$ is a pseudoline arrangement for all $t_0 \in [0,1]$.

Finally, view x as a complex variable. Let $B = [-M,M] \times i[-\alpha,\alpha] \times D_R \subset \mathbb{C}^2$, where D_R is a disk of sufficiently large radius R so that all $|p_i(x,t)| < R$ for $x \in [-M,M] \times i[-\alpha,\alpha]$. We consider the locus $\{\prod_{i=1}^m (y - p_i(x,t)) = 0\} \subset B$ for each $t \in [0,1]$, and label its irreducible components as $\Gamma_i^t = \{y - p_i(x,t) = 0 \mid (x,y) \in B\}$. If $\alpha > 0$ is chosen sufficiently small, then all of the intersections where $p_i(x,t) = p_j(x,t)$ with $x \in [-M,M] \times i[-\alpha,\alpha]$ occur at real values of x. Therefore this complexification of the $\lambda_i^{t_0}$ restricted to B gives an algebraic family of curves, which for any $t_0 \in [0,1]$

is a symplectic line arrangement, at $t_0 = 0$ has the incidences of a pencil, and at $t_0 = 1$ has the incidences of the original pseudoline arrangement $\{\ell_i\}$. □

Remark 8.5 To prove Proposition 8.2, we started with a particular smooth homotopy between the given pseudoline arrangement and the pencil; this homotopy was provided by Lemma 8.4. The same argument applies to an *arbitrary* smooth graphical homotopy that has the properties stated in Lemma 8.4. In many examples such as those in Section 7, a homotopy with the required properties can be easily constructed directly, thus we can find its polynomial approximation without resolving all multiple intersections into double points as required by the algorithm of Lemma 8.4. However, we are unable to do the polynomial approximation while preserving all the incidence relations during the homotopy (we only guarantee the required incidences agree with those of the homotopy for $t = 0$ and $t = 1$ but not for $0 < t < 1$).

8.2 Smooth graphical homotopies imitating picture deformations

Even without the algebraic condition, we can define a subclass of smooth graphical homotopies which produce Stein fillings constrained in a similar way as Milnor fibers. We now isolate these key properties of a picture deformation needed to detect the examples of unexpected Stein fillings in Section 7.

We can describe a smooth graphical homotopy with branches $C_k^t \subset \mathbb{C}^2$ via equations

$$f_k(x_1, x_2, t) - y = 0, \tag{8-1}$$

where (x, y) are the complex coordinates on $\mathbb{C}^2$, $x = x_1 + ix_2$, and t is the real homotopy parameter. At $t = 0$, we assume that $\bigcup_{i=1}^k C_k^0 = \mathcal{C}$ is the germ of a complex algebraic curve where each branch passes through the origin. In particular, $f_k(0, 0, 0) = 0$ for all k. Additionally, any two branches of $\mathcal{C}$ have positive total algebraic intersection number, so any two deformed branches C_i^t and C_j^t intersect for small $t > 0$. Composing the homotopy with a t–dependent translation, we can also assume that the first two branches always intersect at the origin, $C_1^t \cap C_2^t = 0$.

As before, we will assume that the deformed branches C_k^t are not all concurrent for $t > 0$. This means that for $t > 0$, at least one of the functions $f_k(0, 0, t)$, with $k > 2$, is nonzero. We need a nondegenerate version of nonconcurrence:

$$\frac{\partial^r f_k}{\partial t^r}(0, 0, 0) \neq 0 \quad \text{for some } k \in \{3, \dots, m\} \text{ and } r > 0. \tag{8-2}$$

In other words, if we set

$$\operatorname{ord}_t f_k = \min\Big\{r : \frac{\partial^r f_k}{\partial t^r}(0,0,0) \neq 0\Big\},$$

then $\operatorname{ord}_t f_k$ is finite for at least some values $k = 3, \dots, m$. Intuitively, this condition says that the branches move away from being concurrent at the infinitesimal level.

In addition to the above nondegeneracy hypothesis, assume that for all $t > 0$ the arrangements $\{C_1^t, C_2^t, \dots, C_m^t\}$ are topologically equivalent. It follows that each curvetta C_i^t has a finite number of intersections with the other curvettas C_j^t, $i \neq j$; the incidence pattern, and the number of intersections, remain constant during the homotopy. We can add decorations so that all intersection points on $\bigcup_{i=1}^m C_i^t$ are marked; as for picture deformations, we allow free marked points as well. Let w_k be the total number of marked points on the branch C_k^t for any $t > 0$, and set $w = (w_1, w_2, \dots, w_m)$. We will use the term *small smooth deformation* to refer to a smooth graphical homotopy of the decorated germ $(\mathcal{C}, w)$ with special properties as above. Small smooth deformations mimic picture deformations in the smooth category, using smooth graphical instead of algebraic curvettas: they preserve the topology of the curvetta arrangement and satisfy the same weight restrictions and positivity of intersection properties.

Proposition 8.6 *Lemma 7.5 holds for small smooth deformations of plane curve germ $\mathcal{C}$ with smooth branches.*

Proof The proof remains almost the same, but we have to use Taylor approximations of smooth functions instead of power series for analytic functions.

In complex coordinates (x, y) on $\mathbb{C}^2$, the complex tangent line to C_k at 0 has the form $a_k x - y = 0$ for $a_k \in \mathbb{C}$. Setting $x = x_1 + i x_2$ and identifying $\mathbb{C}^2$ with $\mathbb{R}^2 \times \mathbb{C}$, the complex tangent line becomes the 2–plane $a_k x_1 + i a_k x_2 - y = 0$. Set $b_k(t) = f_k(0,0,t)$ and $g_k(x,y,t) = f_k(x,y,t) - a_k x_1 - i a_k x_2 - b_k(t)$. Since $g_k(0,0,t) = 0$ for all t, we have

$$\frac{\partial^\gamma g_k}{\partial t^\gamma}(0,0,0) = 0$$

for all γ; additionally,

$$\frac{\partial g_k}{\partial x}(0,0,0) = 0 \quad \text{and} \quad \frac{\partial g_k}{\partial y}(0,0,0) = 0.$$

Equation (8-1) for the deformed branch C_k^t becomes

$$a_k x_1 + i a_k x_2 + b_k(t) + g_k(x_1, x_2, t) - y = 0. \tag{8-3}$$

Using (8-2), we have $r = \min_k \operatorname{ord}_t b_k(t) = \operatorname{ord}_t b_{k_0}(t) < +\infty$, and write $b_k(t) = t^r \bar{b}_k(t)$ for all k.

We now use the Taylor formula for each function $g_k(x_1, x_2, t)$ at $(0, 0, 0)$, writing out the terms up to r^{th} order, followed by the remainder. This gives

$$a_k x_1 + i a_k x_2 + t^r \bar{b}_k(t) + \sum_{\substack{1<\alpha+\beta+\gamma\le r \\ \alpha>0 \text{ or } \beta>0}} \frac{\partial^{\alpha+\beta+\gamma} g_k}{\partial x_1^\alpha \partial x_2^\beta \partial t^\gamma}(0,0,0) x_1^\alpha x_2^\beta t^\gamma + \sum_{\substack{\alpha+\beta+\gamma=r \\ \alpha>0 \text{ or } \beta>0}} h_{k;\alpha,\beta,\gamma}(x_1, x_2, t) x_1^\alpha x_2^\beta t^\gamma + h_{k;0,0,r}(0,0,t) t^r - y = 0. \tag{8-4}$$

The remainder function $h_{k;\alpha,\beta,\gamma}$ is continuous for each $(k;\alpha,\beta,\gamma)$, and we have that $h_{k;\alpha,\beta,\gamma}(x_1, x_2, t) \to 0$ when $(x_1, x_2, t) \to (0, 0, 0)$. Now make a change of variables

$$x_1 = t^r x_1', \quad x_2 = t^r x_2', \quad y = t^r y'.$$

It is not hard to see that, as in Lemma 7.5, after the change of variables we can divide equation (8-4) by t^r for $t \neq 0$ and take the limit as $t \to 0$. The result is an arrangement of nonconcurrent complex lines given by equations $a_k x' + \bar{b}_k(0) - y' = 0$. Since we have assumed that the incidence relations for $C_1^t, \dots, C_m^t$ remain the same for all $t \neq 0$, the same relations must hold for the lines. □

As a consequence, small smooth deformations cannot produce the unexpected symplectic line arrangements that gave unexpected Stein fillings in Section 7. In such examples, to obtain deformations which produce only Milnor fibers, the algebraic condition on the curves and deformation is less important than keeping the topology of the curves constant for $t \neq 0$. For rational singularities with reduced fundamental cycle, small smooth deformations give a symplectic analogue of smoothings, picking out the Stein fillings which are "closest" to the singularity and its resolution.

8.3 Smooth graphical homotopies changing topology

The key difference between picture deformations and smooth graphical homotopies in Table 1 is that the topology of the union of the curves is allowed to change multiple times during a smooth graphical homotopy — for picture deformations, the only change happens at time 0. In other words, the types of singularities where the curves intersect can vary during the homotopy.

Here we provide an explicit example to illustrate the topology change in the family of Lefschetz fibrations. Our example is related to the configuration $\mathcal{Q}$ from Example 7.14, but with a careful choice of weights.

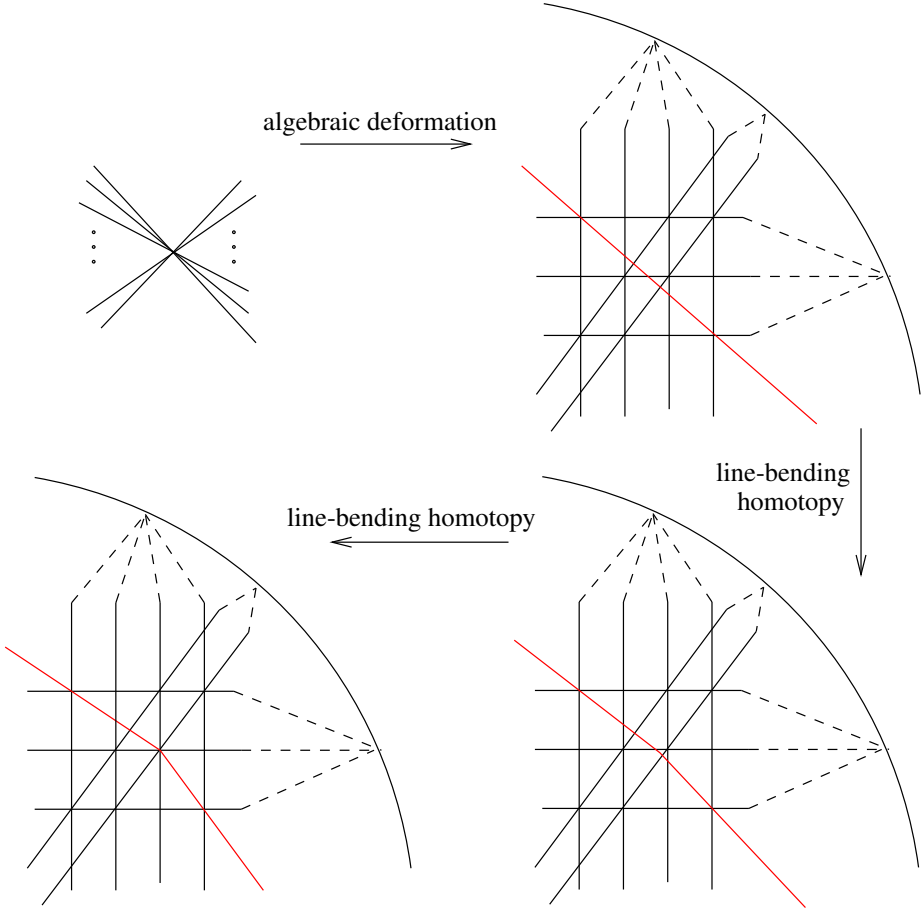


Figure 28: A long-term homotopy from a pencil of lines to $\mathcal{Q}$.

Example 8.7 Consider the pencil of 11 lines indexed from 0 to 10, with weights $w_0 = 4$, $w_1 = w_2 = w_3 = w_4 = w_5 = w_7 = 5$, $w_6 = w_8 = w_9 = 6$ and $w_{10} = 8$. Observe that any arrangement of straight lines is related to the pencil by linear deformation (scaling the constant terms of the linear equations to 0). Using such a deformation, let $\mathcal{Q}_{t_0}$ be the arrangement shown in Figure 28, where ℓ_{10} is a straight line. Unlike the arrangement $\mathcal{Q}$, ℓ_{10} does not pass through the intersection point b of ℓ_3, ℓ_6 and ℓ_9. The corresponding picture deformation of the weighted pencil gives a deformation of the surface singularity. We can extend the picture deformation to a smooth graphical homotopy which for $t_0 < t < 1$ bends the pseudoline ℓ_{10} towards the intersection $\ell_3 \cap \ell_6 \cap \ell_9$, and at $t = 1$ realizes the configuration $\mathcal{Q}$. (We implicitly use Proposition 5.5 to symplectify the family of pseudolines to a smooth graphical homotopy of symplectic line arrangements.)

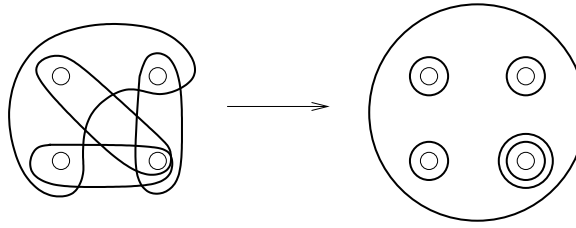


Figure 29: The Stein filling W is related to the Milnor fibers W_t by the monodromy substitution as shown.

Now, consider the Stein fillings W_t correspond to the arrangements $\mathcal{Q}_t$, $0 \le t \le 1$. For $0 < t < 1$, the Stein fillings are diffeomorphic to Milnor fibers of the corresponding smoothings of the singular complex surface. Indeed, the Lefschetz fibrations given by Lemma 3.2 are all equivalent, and for t close to 0 the smooth graphical homotopy is a picture deformation. When $t = 1$, Corollary 7.17 says that the Stein filling W arising from $\mathcal{Q}$ is not strongly diffeomorphic to any Milnor fiber. The topology of W is different from that of W_t: as a smooth manifold, W_t for $t < 1$ is obtained from W by rational blow-down. The corresponding Lefschetz fibrations are related via the positive monodromy substitution given by the daisy relation [14]; see Figure 29.

8.4 Violating positivity of intersections and weight constraints

Although we have seen that we can produce many examples of unexpected Stein fillings using smooth graphical deformations which satisfy positivity of intersections and the weight constraints, we also can construct examples where a Stein filling arises from a configuration of curves such that every smooth graphical homotopy from the germ curvetta violates the weight constraints.

Example 8.8 Consider again the configuration $\mathcal{Q}$, from Example 7.14, of 11 symplectic lines $\{L_k\}_{k=1}^{11}$. We compare this to a pencil of lines with weights

$$w_0 = w_3 = 4, \quad w_1 = w_2 = w_4 = w_5 = w_6 = w_7 = w_9 = 5, \quad w_8 = w_{10} = 6. \tag{8-5}$$

These are chosen such that $w_k = w(L_k)$, so they are the minimal possible weights satisfying the hypotheses of Corollary 7.9. We can show that there is no smooth graphical homotopy from this pencil to $\mathcal{Q}$ satisfying these weight constraints.

Proposition 8.9 *The arrangement $\mathcal{Q}$ cannot be obtained from the pencil of lines by a smooth graphical homotopy satisfying the weight constraints as above if we consider homotopies that are analytic in t or satisfy a nondegeneracy condition such as (8-2).*

This statement follows from the following lemma, which shows that for combinatorial reasons, there are no "intermediate" arrangements between the pencil and $\mathcal{Q}$, so if a homotopy existed, it would have to deform the pencil immediately into an arrangement with the same incidence relations as $\mathcal{Q}$.

Lemma 8.10 *Let $\mathcal{Q}_t = \bigcup_{k=0}^{10} L_k^t$ be a smooth graphical homotopy such that $\mathcal{Q}_0$ is a pencil of 11 lines, and $\mathcal{Q}_1 = \mathcal{Q}$ (after an appropriate choice of coordinates). Suppose that all intersections $L_i^t \cdot L_j^t$ are positive, and each L_k^t has no more than w_k intersection points at all times $t \in [0,1]$. Then, the homotopy $\mathcal{Q}_t$ immediately deforms the pencil of lines into an arrangement combinatorially equivalent to $\mathcal{Q}$, perhaps after restricting to a smaller time interval: there exists $\tau \geq 0$ such that $\mathcal{Q}_\tau$ is a pencil, and $\mathcal{Q}_t$ is combinatorially equivalent to $\mathcal{Q}$ for all $t \in (\tau, 1]$.*

Proof Any two lines in the pencil have algebraic intersection number 1. Since intersections remain inside the Milnor ball during the homotopy and remain positive at all times, throughout the homotopy any two components L_i^t and L_j^t of $\mathcal{Q}_t$ intersect exactly once. This allows us to work with $\mathcal{Q}_t$ as with pseudoline arrangements in Proposition 7.15.

We examine possible combinatorics of an arrangement with the weight restrictions as above. The analysis below works at any time t. For each individual line L_k, we write L_k^t for its image under the homotopy at time t. For $t = 0$, the lines L_k^0 form a pencil; for $t = 1$, we have $\mathcal{Q} = \bigcup L_k^1$.

In the arrangement $\mathcal{Q}$, the line L_0 contains 4 intersection points. These are points where L_0 meets the pencil L_1, L_2, L_3, L_4 of vertical lines, the pencil L_5, L_6, L_7 of horizontal lines, the two diagonal lines L_8, L_9, and the bent line L_{10}. The weight condition then implies that L_0^t can never have more than 4 intersection points. Note that L_3 also has only 4 intersection points, so the same is true for L_3^t. It follows that at most one intersection point on L_0^t can have multiplicity 5 or greater: if there were two such points, there would be two pencils of 5 or more lines. Even if L_3^t is in one of these pencils, it would intersect the lines of the other pencil in 5 or more distinct points, a contradiction. Next, observe that no line has more than 6 intersection points, so no pencil can contain more than 6 lines unless all the lines are concurrent. We conclude that L_0^t must have at least 3 intersection points for all t, because it is not possible to distribute the 10 other lines into two intersection points on L_0 subject to these conditions.

Observe that $\mathcal{Q}_t$ must be combinatorially equivalent to $\mathcal{Q}$ for t close to 1. Indeed, for t sufficiently close to 1, the four distinct intersection points on L_0 remain distinct on L_0^t. Similarly, for t close to 1, each of L_5^t, L_6^t and L_7^t have *at least* 5 distinct intersection points with the other curves in the arrangement $\mathcal{Q}_t$. On the other hand, due to weight restrictions, each of these curves has *at most* 5 intersection points. It follows that L_5^t, L_6^t and L_7^t have exactly 5 intersection points each, and the curves of $\mathcal{Q}_t$ meeting at each intersection have the same incidence relations as the corresponding lines in $\mathcal{Q}$. Thus, the incidences involving L_0^t, as well as the incidences for the "grid" intersections between $L_1^t, L_2^t, L_3^t, L_4^t$ and L_5^t, L_6^t, L_7^t, are the same as in $\mathcal{Q}$ for t close to 1. All the remaining intersections in $\mathcal{Q}_t$ are double points, and they cannot merge with other intersections if t is sufficiently close to 1.

The above argument shows that $\{t \in [0,1] : \mathcal{Q}_t$ is combinatorially equivalent to $\mathcal{Q}\}$ is open. Now, suppose that $\mathcal{Q}_t$ is equivalent to $\mathcal{Q}$ for $t > t_0$. We examine the combinatorial possibilities for $\mathcal{Q}_{t_0}$, assuming that this arrangement is not a pencil. Consider two cases:

(1) $L_0^{t_0}$ has 4 distinct intersection points.

(2) $L_0^{t_0}$ has 3 distinct intersection points.

In the first case, it follows that $\mathcal{Q}_{t_0}$ must be combinatorially equivalent to $\mathcal{Q}$. This is because all the incidence relations valid for $t > t_0$ still hold by taking a limit as $t \to t_0$. As in the proof of Proposition 7.15, we see that no two intersection points can collapse (if they do, all the curves must be concurrent). It follows that in this case, all the incidence relations in $\mathcal{Q}_{t_0}$ are the same as in $\mathcal{Q}$.

In the second case, there are 3 intersection points on L_0. Again, because all incidences hold after taking limits as $t \to t_0$, the arrangement $\mathcal{Q}_{t_0}$ satisfies all the incidence relations of $\mathcal{Q}$. Additionally, two of the intersection points on L_0 collapse. It follows from the proof of Proposition 7.15 that in this case $\mathcal{Q}_{t_0}$ must be a pencil, contradicting the assumption that $L_0^{t_0}$ has 3 distinct intersection points.

We conclude that if $\mathcal{Q}_t$ is combinatorially equivalent to $\mathcal{Q}$ for all $1 \geq t > t_0$, and $\mathcal{Q}_{t_0}$ is different, then $\mathcal{Q}_{t_0}$ must be a pencil. □

We have just seen that there are examples of Stein fillings arising from graphical smooth homotopies which do not satisfy the weight constraint (and such that there is no possible graphical smooth homotopy which does satisfy the weight constraint). On the other hand, we do not have examples of Stein fillings associated to a configuration of graphical curves which cannot be related to the curvetta germ by a smooth graphical

homotopy satisfying positivity of intersections between the curve components. We suspect that in fact, there may always be a smooth graphical homotopy maintaining positivity of intersections.

Question 8.11 Suppose $C^0 = \{C_1^0, C_2^0, \dots, C_m^0\}$ and $C^1 = \{C_1^1, C_2^0, \dots, C_m^1\}$ are two collections of symplectic disks in B_r^4 such that C_i^t intersects C_j^t positively transversally or with a local holomorphic model. Further assume that the boundaries of C^0 and C^1 are isotopic braids in S_r^3. Does there exist a continuous family $\{C_1^t, C_2^t, \dots, C_m^t\}$ of symplectic disks, all with isotopic boundary braid for $t \in [0,1]$, extending this pair of arrangements, such that for each t, C_i^t and $C_{i'}^t$ have positive intersections?

To prove existence of such a homotopy, one could realize C^0 and C^1 as J_0- and J_1–holomorphic curves, respectively, for almost complex structures J_0 and J_1 which are compatible with the standard symplectic structure, with appropriate convexity conditions at the boundary of the ball. One could connect J_0 and J_1 through a family J_t of almost complex structures with the same properties, and then try to find a family C_i^t of J_t–holomorphic disks interpolating between C_i^0 and C_i^1 for each i. The difficulty arises in analyzing the moduli spaces of J–holomorphic curves with appropriately chosen boundary conditions (either using an SFT set-up or a totally real boundary condition). Compactness issues in the moduli space must be overcome to obtain a positive answer to Question 8.11. Because such techniques are far beyond the scope of this article, and the answer to the question is not central to our investigations, we leave this open.

Remark 8.12 If a smooth graphical homotopy fails to satisfy the weight constraints or positivity of intersections, we cannot construct a sequence of Stein fillings using Lemma 3.2. However, we can "connect" the singular complex surface $(X, 0)$ to the Stein filling W via a family of *achiral* Lefschetz fibrations; see [20, Section 8.4].

Consider Example 8.8. We will use the homotopy of pseudoline arrangements given in Example 8.7. For $0 < t < 1$, the pseudolines ℓ_3, ℓ_6, ℓ_9 and ℓ_{10} have more intersection points than the weights (8-5) allow. We need to compensate for the higher weights to obtain the required open book monodromy, so we place *negative* free marked points on these lines: ℓ_3, ℓ_6, ℓ_9 need one negative marked point each to compensate for one extra positive intersection, and ℓ_{10} needs 2 negative points. In the open book monodromy, every negative marked point contributes a negative Dehn twist around the corresponding hole. It follows from the proof of Lemma 3.4 that with these additional

negative twists, the resulting open book supports (Y,ξ). The corresponding vanishing cycles determine an achiral Lefschetz fibration. The negative Dehn twists correspond to a "negative" blow-up in the smooth category (the 4–manifold changes by taking a connected sum with $\mathbb{C}\mathrm{P}^2$).

References

[1] **N A'Campo**, *Le groupe de monodromie du déploiement des singularités isolées de courbes planes, I*, Math. Ann. 213 (1975) 1–32 MR Zbl

[2] **S Akbulut**, **B Ozbagci**, *Lefschetz fibrations on compact Stein surfaces*, Geom. Topol. 5 (2001) 319–334 MR Zbl

[3] **A Akhmedov**, **B Ozbagci**, *Singularity links with exotic Stein fillings*, J. Singul. 8 (2014) 39–49 MR Zbl

[4] **A Akhmedov**, **B Ozbagci**, *Exotic Stein fillings with arbitrary fundamental group*, Geom. Dedicata 195 (2018) 265–281 MR Zbl

[5] **M Artin**, *Algebraic construction of Brieskorn's resolutions*, J. Algebra 29 (1974) 330–348 MR Zbl

[6] **D Auroux**, *A stable classification of Lefschetz fibrations*, Geom. Topol. 9 (2005) 203–217 MR Zbl

[7] **R İ Baykur**, **N Monden**, **J Van Horn-Morris**, *Positive factorizations of mapping classes*, Algebr. Geom. Topol. 17 (2017) 1527–1555 MR Zbl

[8] **M Bhupal**, **K Ono**, *Symplectic fillings of links of quotient surface singularities*, Nagoya Math. J. 207 (2012) 1–45 MR Zbl

[9] **M Bhupal**, **A I Stipsicz**, *Smoothings of singularities and symplectic topology*, from "Deformations of surface singularities" (A Némethi, Á Szilárd, editors), Bolyai Soc. Math. Stud. 23, János Bolyai Math. Soc., Budapest (2013) 57–97 MR Zbl

[10] **F A Bogomolov**, **B de Oliveira**, *Stein small deformations of strictly pseudoconvex surfaces*, from "Birational algebraic geometry" (Y Kawamata, V V Shokurov, editors), Contemp. Math. 207, Amer. Math. Soc., Providence, RI (1997) 25–41 MR Zbl

[11] **C Caubel**, **A Némethi**, **P Popescu-Pampu**, *Milnor open books and Milnor fillable contact 3–manifolds*, Topology 45 (2006) 673–689 MR Zbl

[12] **D C Cohen**, **A I Suciu**, *The braid monodromy of plane algebraic curves and hyperplane arrangements*, Comment. Math. Helv. 72 (1997) 285–315 MR Zbl

[13] **Y Eliashberg**, *On symplectic manifolds with some contact properties*, J. Differential Geom. 33 (1991) 233–238 MR Zbl

[14] **H Endo**, **T E Mark**, **J Van Horn-Morris**, *Monodromy substitutions and rational blowdowns*, J. Topol. 4 (2011) 227–253 MR Zbl

[15] **J B Etnyre**, *Planar open book decompositions and contact structures*, Int. Math. Res. Not. 2004 (2004) 4255–4267 MR Zbl

[16] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Math. Ser. 49, Princeton Univ. Press (2012) MR Zbl

[17] **E Fossati**, *Contact surgery on the Hopf link: classification of fillings*, preprint (2019) arXiv 1905.13026

[18] **D Gay**, **T E Mark**, *Convex plumbings and Lefschetz fibrations*, J. Symplectic Geom. 11 (2013) 363–375 MR Zbl

[19] **P Ghiggini**, **M Golla**, **O Plamenevskaya**, *Surface singularities and planar contact structures*, Ann. Inst. Fourier (Grenoble) 70 (2020) 1791–1823 MR Zbl

[20] **R E Gompf**, **A I Stipsicz**, *4–Manifolds and Kirby calculus*, Graduate Studies in Math. 20, Amer. Math. Soc., Providence, RI (1999) MR Zbl

[21] **J E Goodman**, *Proof of a conjecture of Burr, Grünbaum, and Sloane*, Discrete Math. 32 (1980) 27–35 MR Zbl

[22] **G-M Greuel**, **J Steenbrink**, *On the topology of smoothable singularities*, from "Singularities, I" (P Orlik, editor), Proc. Sympos. Pure Math. 40, Amer. Math. Soc., Providence, RI (1983) 535–545 MR Zbl

[23] **B Grünbaum**, *Configurations of points and lines*, Graduate Studies in Math. 103, Amer. Math. Soc., Providence, RI (2009) MR Zbl

[24] **E Hironaka**, *Generalized lantern relations and planar line arrangements*, from "Computational algebraic and analytic geometry" (M Seppälä, E Volcheck, editors), Contemp. Math. 572, Amer. Math. Soc., Providence, RI (2012) 113–125 MR Zbl

[25] **K Honda**, **W H Kazez**, **G Matić**, *Right-veering diffeomorphisms of compact surfaces with boundary*, Invent. Math. 169 (2007) 427–449 MR Zbl

[26] **T de Jong**, **D van Straten**, *On the deformation theory of rational surface singularities with reduced fundamental cycle*, J. Algebraic Geom. 3 (1994) 117–172 MR Zbl

[27] **T de Jong**, **D van Straten**, *Deformation theory of sandwiched singularities*, Duke Math. J. 95 (1998) 451–522 MR Zbl

[28] **A Kaloti**, *Stein fillings of planar open books*, preprint (2013) arXiv 1311.0208

[29] **J Kollár**, *Toward moduli of singular varieties*, Compos. Math. 56 (1985) 369–398 MR Zbl

[30] **J Kollár**, *Flips, flops, minimal models, etc.*, from "Surveys in differential geometry" (C-C Hsiung, S-T Yau, editors), Lehigh Univ., Bethlehem, PA (1991) 113–199 MR Zbl

[31] **H B Laufer**, *Normal two-dimensional singularities*, Ann. of Math. Stud. 71, Princeton Univ. Press (1971) MR Zbl

[32] **H B Laufer**, *Taut two-dimensional singularities*, Math. Ann. 205 (1973) 131–164 MR Zbl

[33] **H B Laufer**, *On μ for surface singularities*, from "Several complex variables, I" (R O Wells, Jr, editor), Proc. Sympos. Pure Math. 30, Amer. Math. Soc., Providence, RI (1977) 45–49 MR Zbl

[34] **P Lisca**, *On symplectic fillings of lens spaces*, Trans. Amer. Math. Soc. 360 (2008) 765–799 MR Zbl

[35] **S Lisi**, **C Wendl**, *Spine removal surgery and the geography of symplectic fillings*, Michigan Math. J. 70 (2021) 403–422 MR Zbl

[36] **D Margalit**, **J McCammond**, *Geometric presentations for the pure braid group*, J. Knot Theory Ramifications 18 (2009) 1–20 MR Zbl

[37] **H Matsumoto**, *Générateurs et relations des groupes de Weyl généralisés*, C. R. Acad. Sci. Paris 258 (1964) 3419–3422 MR Zbl

[38] **D McDuff**, *The structure of rational and ruled symplectic 4–manifolds*, J. Amer. Math. Soc. 3 (1990) 679–712 MR Zbl

[39] **J Milnor**, *Singular points of complex hypersurfaces*, Ann. of Math. Stud. 61, Princeton Univ. Press (1968) MR Zbl

[40] **B Moishezon**, **M Teicher**, *Braid group technique in complex geometry, I: Line arrangements in $\mathbb{C}P^2$*, from "Braids" (J S Birman, A Libgober, editors), Contemp. Math. 78, Amer. Math. Soc., Providence, RI (1988) 425–555 MR Zbl

[41] **B Moishezon**, **M Teicher**, *Braid group technique in complex geometry, II: From arrangements of lines and conics to cuspidal curves*, from "Algebraic geometry" (S Bloch, I Dolgachev, W Fulton, editors), Lecture Notes in Math. 1479, Springer (1991) 131–180 MR Zbl

[42] **A Némethi**, *Five lectures on normal surface singularities*, from "Low dimensional topology" (K Böröczky, Jr, W Neumann, A Stipsicz, editors), Bolyai Soc. Math. Stud. 8, János Bolyai Math. Soc., Budapest (1999) 269–351 MR Zbl

[43] **A Némethi**, *Some meeting points of singularity theory and low dimensional topology*, from "Deformations of surface singularities" (A Némethi, A Szilárd, editors), Bolyai Soc. Math. Stud. 23, János Bolyai Math. Soc., Budapest (2013) 109–162 MR Zbl

[44] **A Némethi**, **P Popescu-Pampu**, *On the Milnor fibers of sandwiched singularities*, Int. Math. Res. Not. 2010 (2010) 1041–1061 MR Zbl

[45] **A Némethi**, **P Popescu-Pampu**, *On the Milnor fibres of cyclic quotient singularities*, Proc. Lond. Math. Soc. 101 (2010) 554–588 MR Zbl

[46] **A Némethi**, **M Tosun**, *Invariants of open books of links of surface singularities*, Studia Sci. Math. Hungar. 48 (2011) 135–144 MR Zbl

[47] **W D Neumann**, *A calculus for plumbing applied to the topology of complex surface singularities and degenerating complex curves*, Trans. Amer. Math. Soc. 268 (1981) 299–344 MR Zbl

[48] **K Niederkrüger**, **C Wendl**, *Weak symplectic fillings and holomorphic curves*, Ann. Sci. École Norm. Sup. 44 (2011) 801–853 MR Zbl

[49] **H Ohta**, **K Ono**, *Symplectic fillings of the link of simple elliptic singularities*, J. Reine Angew. Math. 565 (2003) 183–205 MR Zbl

[50] **H Ohta**, **K Ono**, *Simple singularities and symplectic fillings*, J. Differential Geom. 69 (2005) 1–42 MR Zbl

[51] **P Ozsváth**, **Z Szabó**, *On the Floer homology of plumbed three-manifolds*, Geom. Topol. 7 (2003) 185–224 MR Zbl

[52] **H Park**, **J Park**, **D Shin**, **G Urzúa**, *Milnor fibers and symplectic fillings of quotient surface singularities*, Adv. Math. 329 (2018) 1156–1230 MR Zbl

[53] **H C Pinkham**, *Deformations of algebraic varieties with G_m action*, Astérisque 20, Soc. Math. France, Paris (1974) MR Zbl

[54] **O Plamenevskaya**, *On Legendrian surgeries between lens spaces*, J. Symplectic Geom. 10 (2012) 165–181 MR Zbl

[55] **P Popescu-Pampu**, *On the smoothings of non-normal isolated surface singularities*, J. Singul. 12 (2015) 164–179 MR Zbl

[56] **P Popescu-Pampu**, *Complex singularities and contact topology*, Winter Braids Lect. Notes 3 (2016) art. id. 3 MR Zbl

[57] **D Ruberman**, **L Starkston**, *Topological realizations of line arrangements*, Int. Math. Res. Not. 2019 (2019) 2295–2331 MR Zbl

[58] **S Schönenberger**, *Determining symplectic fillings from planar open books*, J. Symplectic Geom. 5 (2007) 19–41 MR Zbl

[59] **L Starkston**, *Classifications and applications of symplectic fillings of Seifert fibered spaces over S^2*, PhD thesis, University of Texas at Austin (2015) Available at https://www.math.ucdavis.edu/~lstarkston/Dissertation

[60] **A I Stipsicz**, *On the geography of Stein fillings of certain 3–manifolds*, Michigan Math. J. 51 (2003) 327–337 MR Zbl

[61] **M Symington**, *Symplectic rational blowdowns*, J. Differential Geom. 50 (1998) 505–518 MR Zbl

[62] **S Tendian**, *Surfaces of degree d with sectional genus g in $\mathbb{P}^{d+1-g}$ and deformations of cones*, Duke Math. J. 65 (1992) 157–185 MR Zbl

[63] **W P Thurston**, *A norm for the homology of 3–manifolds*, Mem. Amer. Math. Soc. 339, Amer. Math. Soc., Providence, RI (1986) 99–130 MR Zbl

[64] **G N Tyurina**, *Locally semi-universal flat deformations of isolated singularities of complex spaces*, Izv. Akad. Nauk SSSR Ser. Mat. 33 (1969) 1026–1058 MR In Russian; translated in Math. USSR-Izv. 33 (1969) 967–999

[65] **C Wendl**, *Strongly fillable contact manifolds and J–holomorphic foliations*, Duke Math. J. 151 (2010) 337–384 MR Zbl

Department of Mathematics, Stony Brook University
Stony Brook, NY, United States
Mathematics Department, University of California, Davis
Davis, CA, United States
olga@math.stonybrook.edu, lstarkston@math.ucdavis.edu

Proposed: András I Stipsicz Received: 21 July 2020
Seconded: Dan Abramovich, Paul Seidel Revised: 14 June 2021

msp

Geometry & Topology **27**:3 (2023) 1203–1272
DOI: 10.2140/gt.2023.27.1203
Published: 8 June 2023

A smooth compactification of the space of genus two curves in projective space: via logarithmic geometry and Gorenstein curves

LUCA BATTISTELLA
FRANCESCA CAROCCI

We construct a modular desingularisation of $\overline{\mathcal{M}}_{2,n}(\mathbb{P}^r, d)^{\text{main}}$. The geometry of Gorenstein singularities of genus two leads us to consider maps from prestable admissible covers; with this enhanced logarithmic structure, it is possible to desingularise the main component by means of a logarithmic modification. Both isolated and nonreduced singularities appear naturally. Our construction gives rise to a notion of reduced Gromov–Witten invariants in genus two.

14H10; 14H45, 14N35, 14T90

1 Introduction

Modern enumerative geometry is based on a series of compactifications of the moduli space of smooth embedded curves of genus g and curve class $\beta \in H_2^+(X, \mathbb{Z})$ in a smooth projective variety X, and on their virtual intersection theory. The boundary of M Kontsevich's space of stable maps represents maps from nodal curves, including multiple covers and contracted components. The genus zero theory of projective space provides a smooth compactification with normal crossing boundary. In higher genus, instead, contracted subcurves and finite covers may give rise to boundary components of excess dimension. Even though this moduli space satisfies R Vakil's Murphy's law,

a desingularisation of the main component certainly exists after H Hironaka's work on resolution of singularities. The main application of the methods developed in this paper is an *explicit modular desingularisation* of the main component of the moduli space of stable maps to projective space in genus two and degree $d \geq 3$.

Theorem *There exists a logarithmically smooth and proper DM stack $\mathcal{VZ}_{2,n}(\mathbb{P}^r, d)$ over $\mathbb{C}$, with locally free logarithmic structure (and therefore smooth) and a birational forgetful morphism to $\overline{\mathcal{M}}_{2,n}(\mathbb{P}^r, d)^{\mathrm{main}}$, parametrising the following data*:

- *A pointed admissible hyperelliptic cover*

$$\psi : (C, D_R, x_1, \dots, x_n) \to (T, D_B, \psi(x_1), \dots, \psi(x_n)),$$

  where C is a prestable curve of arithmetic genus two, D_R and D_B are length six (ramification and branch) divisors, and $(T, D_B, \psi(\boldsymbol{x}))$ is a stable rational tree.
- *A map $f : C \to \mathbb{P}^r$.*
- *A bubbling destabilisation $C \leftarrow \widetilde{C}$ and a contraction $\widetilde{C} \to \overline{C}$ to a curve with Gorenstein singularities such that f factors through $\bar{f} : \overline{C} \to \mathbb{P}^r$, and the latter is not special on any subcurve of $\overline{C}$.*

More generally, for any smooth projective variety X, we construct a proper moduli space $\mathcal{VZ}_{2,n}(X, \beta)$ admitting a perfect obstruction theory and defining *reduced Gromov–Witten invariants*. When X is a projective complete intersection, they satisfy the *quantum Lefschetz* hyperplane principle. In general, we expect them to have a simpler enumerative content compared to standard Gromov–Witten invariants.

1.1 Main and boundary components

The moduli space of stable maps to X (see Kontsevich [47]) represents $f : C \to X$, where C is a *prestable* (nodal and reduced) curve such that every rational component of C contracted by f contains at least three *special points* (markings and nodes).

When $X = \mathbb{P}^r$, f is equivalently determined by the data of a line bundle L on C and $r+1$ sections of L that do not vanish simultaneously on C. Forgetting the sections, we obtain a morphism to the universal Picard stack, parametrising pairs (C, L) of a prestable curve and a line bundle on it,

$$\overline{\mathcal{M}}_{g,n}(\mathbb{P}^r, d) \to \mathfrak{Pic}_{\mathscr{C}_{g,n}/\mathfrak{M}_{g,n}}.$$

Obstructions lie in $H^1(C, L)$; see Wang [78]. This implies that $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^r, d)$ is smooth. On the other hand, for $g \geq 1$, boundary components may arise where L restricts to

a special line bundle on a subcurve of C. When $d > 2g-2$, we can identify a *main component*: the closure of the locus of maps from a smooth source curve.

In genus one, the generic point of a boundary component has a contracted subcurve of genus one, together with $k \leq r$ tails of positive degree; see Vakil [74].

In genus two, two types of boundary phenomena occur:

- a subcurve of positive genus is contracted, or
- the restriction of f to the minimal subcurve of genus two (*core*) is the hyperelliptic cover of a line in $\mathbb{P}^r$.

As an example of the second phenomenon, the main component of $\overline{\mathcal{M}}_2(\mathbb{P}^3, 5)$ has dimension 20; but the locus of maps from a reducible curve $C = Z \cup_q R$, with Z of genus two covering a line two-to-one and $R \simeq \mathbb{P}^1$ parametrising a twisted cubic, has dimension 21.

We stress that $\overline{\mathcal{M}}_{2,n}(\mathbb{P}^r, d)^{\text{main}}$ has no natural modular interpretation, since it is obtained by taking the closure of the nice locus. Moreover, its singularities along the boundary can be nasty; see Vakil [75]. In [12], we draw the consequences of the present construction, analysing the locus of smoothable maps in $\overline{\mathcal{M}}_2(\mathbb{P}^2, 4)$, a moduli space with more than twenty irreducible components.

1.2 Good factorisation through a Gorenstein curve

Our approach to the desingularisation of the main component takes off from a simple observation: a line bundle of degree at least $2g-1$ on a *minimal Gorenstein* curve of genus g has vanishing h^1; we have already used this implicitly for nodal curves in the discussion of the irreducible components of $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^r, d)$. The Gorenstein assumption makes this fact a straightforward consequence of Serre duality. Minimality is a weaker notion than irreducibility, but it is needed to ensure that the line bundle has sufficiently positive degree on every subcurve.

This observation raises a natural question: is it possible to replace every $f : C \to \mathbb{P}^r$ with a "more positive"/"less obstructed" $\bar{f} : \overline{C} \to \mathbb{P}^r$ by "contracting" the higher-genus subcurves on which $f^*\mathcal{O}_{\mathbb{P}^r}(1)$ is special? The answer is "no", since every such map is *smoothable*, ie it can be deformed into a map from a smooth curve.[1]

[1]The answer could also be "yes": it has been shown by M Viscardi [77] in genus one that maps from Gorenstein curves satisfying certain stability conditions give rise to irreducible compactifications of $\mathcal{M}_{1,n}(\mathbb{P}^r, d)$; yet, their deformation theory is hard to grasp, because such is the deformation theory of the singularities that are involved [68].

This, on the other hand, gives us a strategy to study the main component:

(1) Construct a universal contraction $\mathscr{C} \to \bar{\mathscr{C}}$ to a Gorenstein curve, by collapsing the subcurves of C on which f has low degree. This Gorenstein curve will not be ready available on $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^r, d)$, because the contraction map $\mathscr{C} \to \bar{\mathscr{C}}$ has moduli (called *crimping spaces* by van der Wyck [80] and *moduli of attaching data* by D I Smyth [68]); we first need to introduce a modification of $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^r, d)$ along the boundary.

(2) Take only those maps that admit a factorisation through $\bar{\mathscr{C}}$: as we have discussed, these are all smoothable, so the moduli space is at the very least irreducible. It provides a birational model of $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^r, d)^{\mathrm{main}}$, with the advantage of admitting a modular interpretation.

Moreover, this space is unobstructed over a base that can be assumed to be smooth in the low-genus examples that we have at our disposal: those of D Ranganathan, K Santos-Parker and J Wise [63], where the base is a logarithmic modification of the moduli space of prestable curves; and those of this paper, where the base is a logarithmic modification of the moduli space of hyperelliptic admissible covers. Once such a moduli space is constructed, the proof of smoothness for target $\mathbb{P}^r$ is entirely conceptual; furthermore, the same methods can be employed to approach the study of different targets (eg products of projective spaces, toric varieties, flag varieties) and stability conditions (eg quasimaps).

While point (2) was essentially established for us by Ranganathan, Santos-Parker and Wise, point (1) is at all open for $g \geq 2$. Here we make a hopefully meaningful step in this direction.

1.3 Logarithmic geometry and singularities

The moduli space of prestable curves has a natural logarithmic structure induced by its boundary divisor, thus keeping track of the nodes and their smoothing parameters; see F Kato [43]. This induces a logarithmic structure on the moduli space of maps as well. It is a natural question whether the desingularisation of the main component can be achieved by means of a logarithmic modification; it is indeed the case in genus one [63], but not in our construction. We think that the following are truly high-genus phenomena.

1.3.1 Augmenting the logarithmic structure: admissible covers Our first most relevant finding is that, instead, in genus two it is necessary to enrich the logarithmic

structure of the base by passing to a moduli space of admissible covers; see J Harris and D Mumford [37].

Every smooth curve of genus two is hyperelliptic, ie the canonical class induces a two-to-one cover of a line branched along six points. The hyperelliptic cover is essentially unique, but, when the curve becomes nodal and in the presence of markings, the uniqueness is lost. It can be restored by making it part of the moduli functor. The resulting space of admissible covers is as nice as the moduli space of curves, but it has the advantage of encoding the Brill–Noether theory of the curve in the logarithmic structure; see S Mochizuki [55]. The necessity of such enrichment can be understood already by looking at the isolated Gorenstein singularities of genus two (see Battistella [11]): there are two families of these, basically corresponding to the choice of either a Weierstrass point or a conjugate pair in the semistable model. In order to tell these two cases apart logarithmically, in the construction of $\mathscr{C} \to \bar{\mathscr{C}}$, we start by considering stable maps from the source of an admissible cover.

1.3.2 Realisable tropical canonical divisors This points to our second finding. The line bundle giving the contraction $\mathscr{C} \to \mathscr{C}'$ (the first step towards $\bar{\mathscr{C}}$) is a vertical twist of the relative dualising line bundle. The twist is indeed logarithmic, and the piecewise-linear function on $\mathrm{trop}(\mathscr{C})$ determining it is nothing but a *tropical canonical divisor* satisfying certain requirements of compatibility with the admissible cover, and in particular realisable; see Möller, Ulirsch and Werner [57]. Pulling back from the target of the admissible cover, which is a rational curve, allows us to simplify several computations. The subdivision of the tropical moduli space according to the domain of linearity of such function should therefore be thought of as parametrising all possible Gorenstein contractions of $\mathscr{C}$ compatible with the degree of f.

1.3.3 Nonisolated singularities Finally, we underline that nonreduced curves appear as fibres of $\bar{\mathscr{C}}$, which requires a careful study of what we call *tailed ribbons*, and further distinguishes our work from its genus one ancestor. This should not come as a surprise: first, nonreduced curves can still be Gorenstein if the nilpotent structure is supported along one-dimensional components, rather than isolated points; second, *ribbons* were introduced in the 1990s as limits of smooth canonical curves when the latter tend to the hyperelliptic locus in moduli; see Fong [31]. They appear naturally in our work at the intersection of the main component with the hyperelliptic ones. It is possibly less expected that they show up as well when the core is contracted, as a way of interpolating between isolated singularities whose special branches differ. The construction of $\bar{\mathscr{C}}$ is

concluded by gluing in $\mathscr{C}'$ the hyperelliptic cover of a genus two subcurve supported along a boundary divisor.

For future investigation, we remark that we perform both the contraction $\mathscr{C} \to \mathscr{C}'$ and the pushout $\mathscr{C}' \to \bar{\mathscr{C}}$ in sufficiently general families. We wonder whether a pointwise construction might be possible, realising both steps as special instances of a more general pushout of logarithmic subcurves, in the spirit of S Bozlee [18].

1.4 Relation to other work and further directions

1.4.1 Local equations and resolution by blowing up It is always possible to find a local embedding of $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^r, d)$ in a smooth ambient space by looking at

$$\mathrm{Tot}(\pi_* L)^{\oplus r+1} \subseteq \mathrm{Tot}(\pi_* L \otimes \mathscr{O}_C(n))^{\oplus r+1}$$

over the Picard stack, where $\mathscr{O}_C(1)$ is a relative polarisation for the universal curve $\pi : \mathscr{C} \to \mathfrak{Pic}_{g,n}$ [25, Section 3.2]. If the polarisation is chosen carefully, the embedding is given by $\approx g$ local equations, repeated $r+1$ times. An approach to the desingularisation of the main component is by blowing up according to these local equations; see Y Hu, J Li and J Niu [38; 39; 40]. Possibly the hardest task is to provide a compatible global description of the blow-up centres and their sequence. This method is close in spirit to the original construction of Vakil and A Zinger [84; 76] (in particular, it involves an iterated blow-up procedure and a good deal of bookkeeping), and it has the advantage of simultaneously "desingularising" the sheaves $\pi_* \mathscr{L}^{\otimes k}$, $k \geq 1$, on the main component, making them into vector bundles. This in turn makes the theory of projective complete intersections accessible via torus localisation; see Zinger [83] and Popa [62].

Hu, Li and Niu [40] carry out this strategy in genus two. For the reader's benefit, we provide a brief conversion chart: in round 3, phase 3 of their routine, they blow up loci in the Picard stack, where the line bundle is required to restrict to the canonical bundle on the core; this part is subsumed in our work by starting from a birational model of the moduli space of weighted curves, namely the space of admissible covers. In round 1 they blow up loci in $\mathfrak{M}_2^{\mathrm{wt}}$ that correspond explicitly to the boundary phenomena that we have summarised in Section 1.1. In rounds 2 and 3 they further blow up loci contained in the exceptional divisor of the previous rounds; these loci often depend on the attaching points of some rational tails being Weierstrass, or conjugate. In the language we have adopted, all of the blow-up centres are encoded in *the domain of linearity of a certain piecewise-linear function* — a tropical canonical divisor — while the Brill–Noether theory of the curve is built in the moduli space at the level of admissible covers

already. Our construction is thus more intrinsic. We remark, for example, that the genus one phenomena (corresponding to round 1, phases 2–4 in [40]) are addressed uniformly within this language, with the tropical canonical divisor resembling the function "distance from the core" that is at the base of [63]. The desingularisation of Hu, Li and Niu seems more efficient than ours, in the (vague) sense that the information encoded by the admissible cover *far from the special branch* is useless for the rest of our construction. Their procedure appears to select a "special branch" to start with in round 1, and then continue blowing up accordingly only "around that branch"; the price to be paid seems to be a certain noncanonicity of the blow-up procedure, with subdivisions appearing that are not quite natural from our perspective (when the core is reducible; see Example 3.39).

In [41], Hu and Niu reconstruct the blow-up by gluing projective bundle strata indexed by *treelike structures* and *level trees*; these data are reminiscent of piecewise-linear functions on tropical curves, though missing both the slope and the metric data. These authors have already pointed out the similarities between their indexing set and the combinatorial data appearing in the moduli space of multiscale differentials due to Bainbridge, Chen, Gendron, Grushevsky and Möller [9]. This relation certainly deserves further attention; we think that canonical divisors on tropical curves could provide the right language to talk about it, and the geometry of Gorenstein curves could be the informing principle of further investigations. We expect a crucial ingredient for an all-genus desingularisation will be a good compactification of canonical curves.

1.4.2 Computations in Gromov–Witten theory Naive computations of what we may now expect to coincide with our reduced invariants have made seldom appearances in the literature, for example, with Zinger's enumeration of genus two curves with a fixed complex structure in $\mathbb{P}^2$ and $\mathbb{P}^3$ [81], and the computation of characteristic numbers of plane curves due to T Graber, J Kock and R Pandharipande [33]. To make the relation with the latter work precise, we should first extend our methods to the analysis of relative and logarithmic stable maps (compare with Battistella, Nabijou and Ranganathan [13] and Ranganathan, Santos-Parker and Wise [64] in genus one).

$\mathcal{VZ}_{2,n}(X,\beta)$ is only the main component of a moduli space of aligned admissible maps $\mathscr{A}_{2,n}(X,\beta)$, which dominates $\overline{\mathcal{M}}_{2,n}(X,\beta)$ and is virtually birational to it. The virtual class of $\mathscr{A}_{2,n}(X,\beta)$ is expected to split as the sum of its main and boundary components; the contribution of the latter should be expressible in terms of genus zero and reduced genus one invariants via virtual pushforward; see Manolache [52].

This would deliver an extension of Li and Zinger's formula to genus two (see Li and Zinger [51], Zinger [82] and Coates and Manolache [26]). Together with a localisation computation of reduced invariants, as mentioned in the previous section, this would provide an alternative proof of the genus two mirror theorem for the quintic threefold; see Guo, Janda and Ruan [35] and Chang, Guo and Li [24]. Along the same lines, it would ease the computation of genus two Gromov–Witten invariants of Fano and Calabi–Yau complete intersections in projective space.

1.4.3 Logarithmic maps to toric varieties and tropical realisability In [64], Ranganathan, Santos-Parker and Wise apply their techniques to desingularising the space of genus one logarithmic maps to a toric variety with respect to its boundary. As an application, they are able to solve the *realisability problem* for tropical maps of genus one (see Speyer [69]): which of them arise as the tropicalisation of a map from a smooth genus one curve to an algebraic torus? With Ranganathan, we are working towards a similar result in genus two — where, as far as we know, there is at the moment no reasonable guess as to what the full answer should be, although a clear understanding of the moduli space of tropical curves has been obtained by Cueto and Markwig [27].

1.4.4 Birational geometry of moduli spaces of curves In [11], the first author produces a sequence of alternative compactifications of $\mathcal{M}_{2,n}$ based on replacing genus one and two subcurves with few special points by isolated Gorenstein singularities. Although we do not discuss it here, the techniques we develop also provide a resolution of the rational maps among these spaces. Moreover, the universal Gorenstein curve constructed here unveils the possibility of defining new birational models of $\overline{\mathcal{M}}_{2,n}$ by including nonreduced curves as well — contrary to $\overline{\mathcal{M}}_{2,n}(m)$, these models could respect the $\mathfrak{S}_n$–symmetry in the markings. It would be interesting to compare them with the work of Johnson and Polishchuk [42], and to establish their position in the Hassett–Keel program of Alper, Fedorchuk, Smyth and van der Wyck [68; 7]. Recently, Bozlee, B Kuo and A Neff [19] have classified all the compactifications of $\mathcal{M}_{1,n}$ in the stack of Gorenstein curves with distinct markings — it turns out that there are many more than envisioned by Smyth, although the numerosity arises more from combinatorial than geometric complications. It would be interesting if all compactifications of $\mathcal{M}_{2,n}$ could be classified by a mixture of our techniques.

Plan of the paper

In Section 2, we establish some language and background material concerning *logarithmic curves*, their tropicalisation and the use of piecewise-linear functions; *admissible*

covers and their logarithmic structure; and *Gorenstein curves*, with both isolated singularities and nonreduced structures, including a number of useful properties and formulas, the classification of isolated singularities of genus one (due to Smyth) and two (due to the first author), and a novel study of the nonisolated singularities of genus two with nonnegative canonical class.

In Section 3, we introduce the key player: a subdivision of the tropical moduli space of weighted admissible covers based on aligning (ordering) the vertices of the tropical curve with respect to a piecewise-linear function constructed from tropical canonical divisors. This subdivision induces a logarithmically étale model of the moduli space of weighted admissible covers. It is on this model that we are able to construct a universal family of Gorenstein curves. This process consists of two steps: first, a birational contraction $\mathscr{C} \to \mathscr{C}'$; then, a pushout/normalisation $\mathscr{C}' \to \bar{\mathscr{C}}$.

In Section 4, we apply these methods to desingularise $\overline{\mathcal{M}}_{2,n}(\mathbb{P}^r, d)^{\mathrm{main}}$. To be more precise, we repeat the contraction/pushout process twice: first, to ensure that any curve of positive genus has weight at least one; then, to exclude the possibility that the map is hyperelliptic on a genus two subcurve. The intermediate space, although not being smooth, is already interesting in that its invariants satisfy quantum Lefschetz; see Lee [50].

Notation and conventions

We work throughout over $\mathbb{C}$.

By *trait* we mean the spectrum Υ of a discrete valuation ring. It only has two points; the closed one is often denoted by s (or 0), and the generic one by η.

Curves will always be projective and $S1$, ie without embedded points, but they may be nonreduced. Subcurves are not supposed to be irreducible, but they are usually connected. We call *core* the minimal subcurve of genus two. We may refer to subcurves $C' \subseteq C$ that intersect the rest of C and the markings in only one (resp. two) at worst nodal point(s) — or markings — as *tails* (resp. *bridges*).

The *dual graph* Γ of a nodal curve C has a vertex for every irreducible component, an edge for every node, and a leg for every marking (labelled or not); it is endowed with a genus function $g\colon V(\Gamma) \to \mathbb{Z}$, and the genus of the graph is given by

$$g(\Gamma) = h^1(\Gamma) + \sum_{v \in V(\Gamma)} g(v).$$

The graph Γ may be further weighted with a function $w\colon V(\Gamma)\to\mathbb{Z}$, which should be thought of as recording the degree of a line bundle on C.

A graph with no loops is called a *tree*; its valence one vertices are called *leaves* (sometimes one of them plays a different role and it is therefore named the *root*). A tree with only two leaves is a *chain*.

A *tropical curve* is a graph metrised in a monoid $\overline{M}$ (most basically, $\mathbb{R}_{\geq 0}$), ie a graph Γ as above together with a length function $\ell\colon E(\Gamma)\to\overline{M}$ on the set of edges (legs are considered to be infinite, instead).

A *logarithmic space* is denoted by $X=(\underline{X},M_X)$ with logarithmic structure $\alpha\colon M_X\to(\mathscr{O}_X,\cdot)$ (also simply indicated by exponential/logarithmic notation). M_X is a sheaf in the étale topology of $\underline{X}$. We assume that logarithmic structures are *FS*, ie fine (they admit an atlas of finitely generated and integral charts) and saturated. Let $\overline{M}_X$ denote the characteristic sheaf $M_X/\alpha^{-1}(\mathscr{O}_X^*)$. This is a constructible sheaf of abelian groups.

A logarithmic space X gives rise to a cone stack $\mathrm{trop}(X)$ via *tropicalisation*. In the case of logarithmically smooth curves, the tropicalisation is a tropical curve in the above sense. We abide to the rule that the tropicalisation of C should be denoted by the corresponding piecewise-linear character $\sqsubset$.

We have quite a few families of curves; usually we adopt the following notation:

- $\pi\colon C\to S$ will denote a prestable curve, $\widetilde{C}$ (with $\tilde{\pi}$) a partial destabilisation of C, C' (with π') a (not necessarily Gorenstein) contraction of $\widetilde{C}$, and $\overline{C}$ (with $\overline{\pi}$) a (not necessarily reduced) Gorenstein curve dominated by C'.
- $f\colon C\to X$ will denote a (stable) map to (a smooth and projective) X.
- $\psi\colon C\to T$ will denote an admissible cover from a genus two to a rational curve, $\psi'\colon C'\to T'$ a double cover (where the curves are not necessarily prestable).

Acknowledgements

We are grateful to Dhruv Ranganathan and Jonathan Wise for nurturing this project. We would also like to thank Daniele Agostini, Fabio Bernasconi, Sebastian Bozlee, David Holmes, Yi Hu, Navid Nabijou and Jingchen Niu for helpful conversations, and Zak Turčinović for his support with the pictures. We thank Cristina Manolache and Richard Thomas for pushing us in this direction during our doctoral studies. We thank the referee for their quick and detailed response, that was of great help in improving the overall

quality of the paper. Battistella is supported by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy EXC-2181/1 – 390900948 (the Heidelberg STRUCTURES Cluster of Excellence). Carocci is supported by the starter grant 759967, *Categorified Donaldson–Thomas theory*, of the European Research Council. This work was initiated at the Max-Planck-Institut für Mathematik in Bonn; it was partly carried out during the workshop on *Enumerative logarithmic geometry and mirror symmetry* at the Mathematisches Forschungsinstitut Oberwolfach and during a visit of Battistella to the University of Edinburgh; we gratefully acknowledge all these institutions for their hospitality.

2 Preliminaries and background material

2.1 Logarithmically smooth curves and their tropicalisations

Let (S, M_S) be a logarithmic scheme. A family of *logarithmically smooth curves* over S is a proper and logarithmically smooth morphism $\pi : (C, M_C) \to (S, M_S)$ with connected one-dimensional (geometric) fibres such that π is integral and saturated. These hypotheses guarantee flatness and that the fibres are reduced. Logarithmically smooth curves naturally provide a compactification of the moduli space of smooth curves.

2.1.1 Characterisation of logarithmically smooth curves Kato proved in [43] that logarithmically smooth curves have at worst nodal singularities; moreover, he provided the following local description. For every geometric point $p \in C$, there exists an étale local neighbourhood of p in C with a strict étale morphism to:

- **Smooth point** $\mathbb{A}^1_S$ with the log structure pulled back form the base.
- **Marking** $\mathbb{A}^1_S$ with the log structure generated by the zero section and $\pi^* M_S$.
- **Node** $\mathscr{O}_S[x, y]/(xy = t)$ for some $t \in \mathscr{O}_S$, with semistable log structure induced by the multiplication map $\mathbb{A}^2_S \to \mathbb{A}^1_S$ and $t : S \to \mathbb{A}^1$.

In the last case, the class of $\log(t)$ in $\overline{M}_S$ is called a *smoothing parameter* for the node.

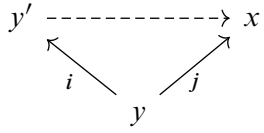
At times, we may have to consider more general logarithmic orbicurves [60].

2.1.2 Minimal logarithmic structures For every prestable curve $\underline{\pi} : \underline{C} \to \underline{S}$, there is a *minimal* logarithmic structure M_S^{can} on $\underline{S}$ together with a logarithmically smooth enhancement $\pi : (\underline{C}, M_C^{\mathrm{can}}) \to (\underline{S}, M_S^{\mathrm{can}})$ such that any other logarithmically smooth

enhancement of $\underline{\pi}$ is pulled back from the minimal one. If $S = \{s\}$ is a geometric point, the characteristic sheaf of the minimal structure is freely generated by the smoothing parameters of the nodes of C_s:

$$\overline{M}_S^{\mathrm{can}} = \mathbb{N}^{\#E(\Gamma(C_s))}.$$

More generally, the concept of minimality — introduced by W D Gillam [32] — serves the purpose of describing which stacks $\mathscr{X}$ over (LogSch) are induced by ordinary stacks $\mathcal{X}$ over (Sch) endowed with a logarithmic structure $\alpha\colon M_{\mathcal{X}} \to (\mathscr{O}_X, \cdot)$. Gillam defines a *minimal object* x of $\mathscr{X}$ to be one such that, for every solid diagram



with i and j over the identity of underlying schemes, there exists a unique dashed arrow in $\mathscr{X}$ making the diagram commutative. He then shows that $\mathscr{X}$ is induced by an $(\mathcal{X}, M_{\mathcal{X}})$ if and only if

(1) for every object z of $\mathscr{X}$, there exist a minimal x, and $z \to x$ covering the identity of underlying schemes;

(2) for every $w \to x$ with x minimal, the corresponding morphism of logarithmic schemes is strict if and only if w is minimal as well.

From this discussion, it follows that the stack of logarithmically smooth curves over (LogSch) is induced by a logarithmic stack over (Sch). It can be identified as the Artin stack of prestable (marked) curves endowed with the divisorial logarithmic structure corresponding to its normal crossings boundary.

2.1.3 Tropicalisation The combinatorial structure of an FS logarithmic space X is encoded by its *Artin fan* $\mathcal{A}_X$ [2; 5]. An Artin fan is a logarithmic stack that looks like the quotient of a toric variety by its dense torus $[V/T]$ locally in the strict étale topology. The morphism $\underline{X} \to \mathrm{Log}$ classifying the logarithmic structure on X [59] factors through $X \to \mathcal{A}_X$. The 2–category of Artin fans is equivalent to the 2–category of *cone stacks*, ie stacks over the category of rational polyhedral cone complexes [21, Theorem 6.11]. The latter should be thought of as a generalisation of the category of rational polyhedral cone complexes, where cones are allowed to be glued to themselves via automorphisms. In this way, we can associate to the logarithmic stack X a cone stack $\mathrm{trop}(X)$ known as its *tropicalisation*, which is nothing but another incarnation

of the Artin fan. When X is a scheme with Zariski logarithmic structure (without monodromy), it is the generalised (standard) cone complex

$$\operatorname{trop}(X) = \Big(\coprod_{x \in X} \operatorname{Hom}(\overline{M}_{X,x}, \mathbb{R}_{\geq 0}) \Big) \Big/ \sim,$$

where x runs through the schematic points of X and gluing takes place along the face inclusions induced by the specialisation morphisms [34, Appendix B]. The language developed in [21] allows us to talk about *the tropicalisation map* within the category of logarithmic stacks. See also [1; 72; 73].

For a logarithmically smooth curve (C, M_C) over a geometric point (S, M_S), the tropicalisation $\sqsubset$ consists of its dual graph $\Gamma(C)$ metrised in $\overline{M}_S$; the length function $\ell\colon E(\Gamma(C)) \to \overline{M}_S$ associates to each edge the smoothing parameter of the corresponding node. More precisely, vertices correspond to irreducible components weighted by their geometric genus, and there are (labelled) infinite legs corresponding to the markings. The tropicalisation of C is thus a family of classical (ie metrised in $\mathbb{R}$) tropical curves over the cone $\operatorname{Hom}(\overline{M}_S, \mathbb{R}_{\geq 0})$. The construction can be generalised to more general base logarithmic schemes [21, Section 7]. It induces a (strict, smooth, surjective) morphism

$$\mathfrak{M}_{g,n}^{\log} \to \widetilde{\mathfrak{M}}_{g,n}^{\operatorname{trop}},$$

where the latter is the lift of the stack of tropical curves to the category of logarithmic schemes through the tropicalisation map.

2.1.4 Piecewise-linear functions and line bundles The characteristic monoid at a node $q \in C$ is the amalgamated sum

$$\overline{M}_{C,q} = \overline{M}_{S,\pi(q)} \oplus_{\mathbb{N}} \mathbb{N}^{\oplus 2},$$

where the map $\mathbb{N} \to \overline{M}_{S,\pi(q)}$ is the smoothing parameter δ_q and $\mathbb{N} \to \mathbb{N}^{\oplus 2}$ is the diagonal. It has been noticed in [34] that

$$\overline{M}_{C,q}^{\operatorname{gp}} \simeq \{(\lambda_1, \lambda_2) \in \overline{M}_{S,\pi(q)}^{\operatorname{gp},\oplus 2} \mid \lambda_2 - \lambda_1 \in \mathbb{Z}\delta_q\}.$$

If $\sqsubset$ is a tropical curve metrised in $\overline{M}$, a *piecewise-linear* (PL) *function* on $\sqsubset$ with values in $\overline{M}^{\operatorname{gp}}$ is a function λ from the vertex set of $\sqsubset$ to $\overline{M}^{\operatorname{gp}}$ such that, for any two adjacent vertices v_1 and v_2 connected by an edge e_q, we can write

$$\lambda(v_2) - \lambda(v_1) = s(\lambda, e_q)\delta_q$$

for some $s(\lambda, e_q) \in \mathbb{Z}$ (called the *slope* of λ along e_q, outgoing from v_1), where δ_q is the smoothing parameter of the node q in $\overline{M}$.

Over a geometric point (or, more generally, if $\overline{M}_S$ and $\sqsubset$ are constant over S — in particular, if no nodes are smoothed out) the previous observation, together with a similar analysis at the markings, shows (see [21, Remark 7.3]) that

$$H^0(C, \overline{M}_C^{\mathrm{gp}}) = \{\text{PL functions on } \sqsubset \text{ with values in } \overline{M}_S^{\mathrm{gp}}\}.$$

Similarly, every section of $\pi_*\overline{M}_C/\overline{M}_S$ is described by the collection of slopes of a rational function on $\sqsubset$.

The exact sequence

$$0 \to \mathscr{O}_C^\times \to M_C^{\mathrm{gp}} \to \overline{M}_C^{\mathrm{gp}} \to 0$$

shows that to every section $\lambda \in \Gamma(C, \overline{M}_C^{\mathrm{gp}})$ there is an associated $\mathscr{O}_C^\times$–torsor (or, equivalently, a line bundle) of lifts of λ to M_C^{gp}. Under the assumption that $\overline{M}_S$ and $\sqsubset$ are constant over S, every vertex v of $\sqsubset$ determines an irreducible component of C, and the restriction of $\mathscr{O}_C(-\lambda)$ to C_v is given explicitly by [63, Proposition 2.4.1]

$$\mathscr{O}_{C_v}(\lambda) \simeq \mathscr{O}_{C_v}\Big(\sum s(\lambda, e_q)q\Big) \otimes \pi^*\mathscr{O}_S(\lambda(v)),$$

where $s(\lambda, e_q)$ denotes the outgoing slope of λ along the edge corresponding to q (either a marking or a node of C).

If we started from $\lambda \in \Gamma(C, \overline{M}_C)$, the associated line bundle $\mathscr{O}_C(-\lambda)$ would come with a cosection $\mathscr{O}_C(-\lambda) \to \mathscr{O}_C$ (induced by the logarithmic structure). Such a cosection is not always injective, but when it is it defines an effective Cartier divisor on C; when it is not, it will vanish along components of C. Nonetheless, the association of this *generalised effective Cartier divisor* to λ behaves well under pullbacks, and in fact a useful point of view on logarithmic structures is to consider them as a functorial system of generalised effective Cartier divisors indexed by $\overline{M}_C$ (see [17] for the details). See also [18, page 9] for a description in local coordinates.

Finally, we briefly recall the theory of divisors on tropical curves [36]. For this, let $\sqsubset$ be a tropical curve over $\mathbb{R}_{\geq 0}$, thought of as its metric realisation. A *divisor* on $\sqsubset$ is a finite $\mathbb{Z}$–linear combination of points on $\sqsubset$; it is said to be *effective* if all the coefficients are nonnegative. The *tropical canonical divisor* is

$$K_\sqsubset = \sum_{v \in \sqsubset} (2g(v) - 2 + \mathrm{val}(v))v.$$

It is effective unless the curve has rational tails. To a piecewise-linear function λ on (a subdivision of) $\sqsubset$, we can associate a divisor supported on its nonlinearity locus,

$$\mathrm{div}(\lambda) = \sum_{v \in \sqsubset} \Big(\sum_{e \in \mathrm{Star}(v)} s(\lambda, e) \Big) v.$$

Notice that the dependence on λ is only up to a global translation by $\mathbb{R}$. Divisors of the form $\operatorname{div}(\lambda)$ are called *principal*. Two divisors are *linearly equivalent* if their difference is principal. We can define the *divisor class group* of $\sqsubset$ by taking equivalence classes of divisors on $\sqsubset$ modulo linear equivalence. The *linear system* of a divisor is given by

$$|D| = \{D' \sim_{\text{lin}} D \mid D' \geq 0\}.$$

We will loosely refer to a member of $|K_\sqsubset|$ as a tropical canonical divisor.

Notice that a tropical linear system is *tropically convex*, ie it is closed under taking max and real translations (see [36, Lemma 4]).

Everything here can be restated for tropical curves over a more general base. Subdivisions, though, may only make sense after enlarging the base monoid, often corresponding to a subdivision of its dual cone.

If $\psi : \sqsubset \to \mathsf{T}$ is a harmonic morphism of tropical curves [8], both divisors and piecewise-linear functions can be pulled back from T to $\sqsubset$ — in both cases, a coefficient accounting for the expansion factor of ψ along edges must be included.

2.1.5 Alignments and blow-ups The monoid $\overline{M}$ induces a partial order on $\overline{M}^{\text{gp}}$,

$$m_1 \leq m_2 \iff m_2 - m_1 \in \overline{M}.$$

Given a logarithmic scheme $(\underline{S}, M_S)$ with a logarithmic ideal $K = (m_1, \dots, m_h)$, the logarithmic subfunctor of S defined by requiring that there always be a minimum among the m_i for $i = 1, \dots, h$ — ie that the ideal be locally principal — is represented by the blow-up of S in the ideal $\alpha(K)$ [61, Section 3.4]. Tropically, it corresponds to a subdivision of the cone $\operatorname{Hom}(\overline{M}_S, \mathbb{R}_{\geq 0})$, and vice versa.

This simple observation has had many fruitful applications in moduli theory, starting from [53].

2.2 Admissible covers and their logarithmic structure

Admissible covers have been introduced by Harris and Mumford [37] as a compactification of Hurwitz spaces. A fully fledged moduli theory for them has been developed only later by Mochizuki with the introduction of logarithmic techniques [55], and by D Abramovich, A Corti and A Vistoli with the introduction of twisted curves [3]. They have also been generalised by B Kim [44, Sections 5.2 and 7.2] (he calls them log-stable μ–ramified maps). We will only be concerned with the case of double covers of $\mathbb{P}^1$, ie hyperelliptic curves.

In order to motivate the construction, we observe that every smooth curve of genus two is hyperelliptic (under the canonical map) with six Weierstrass points (ie ramification points for the hyperelliptic cover; this follows from the Riemann–Hurwitz formula). When the curve is allowed to become nodal, there still exists a degree two map to the projective line, but it is no longer finite. In this case, it is appropriate to define Weierstrass points to be limit of Weierstrass points in a smoothing family; the drawback is that whether a point on a (contracted) rational subcurve is Weierstrass or not may depend on the choice of the smoothing family. To resolve this ambiguity within the definition of the moduli problem, a solution is to expand the target, so as to consider the space of *finite* morphisms from nodal curves of genus two to rational *trees*. Here is a formal definition:

Definition 2.1 A family of *admissible hyperelliptic covers* over S is a finite morphism $\psi\colon (C, D_R) \to (T, D_B)$ over S such that

(1) (C, D_R) and (T, D_B) are prestable curves with (unlabelled) smooth disjoint multisections D_R and D_B of length $2g+2$, C has arithmetic genus g, and (T, D_B) is a stable rational tree;

(2) ψ is a double cover on an open $U \subseteq T$ dense over S;

(3) ψ is étale on $C^{\mathrm{sm}} \setminus D_R$, it maps D_R to D_B with simple ramification, and it maps nodes of C to nodes of T so that, in local[2] coordinates,

$$\psi^{\#}\colon \mathscr{O}_S[u,v]/(uv-s) \to \mathscr{O}_S[x,y]/(xy-t)$$

maps $u \mapsto x^i$, $v \mapsto y^i$ and $s \mapsto t^i$ for $i = 1$ or 2.

Mochizuki shows that condition (3) can be replaced by requiring that ψ lift to a *logarithmically étale* morphism of logarithmic schemes $(C, M_C) \to (T, M_T)$ over (S, M_S), so that the image of a smoothing parameter of T is either a smoothing parameter of C or its double. Moreover there is a minimal logarithmic structure over S that makes this possible: if p is a node of C that ψ maps to the node q of T,

$$M_S^{\psi-\mathrm{can}} = M_S^{C-\mathrm{can}} \oplus_{\mathscr{O}_S^*} M_S^{T-\mathrm{can}}/\sim, \quad (0, \delta_q) \sim (i\delta_p, 0),$$

where i is the local multiplicity defined at the end of (3). In the case of double covers, this simply means that the minimal logarithmic structure for the admissible cover is the same as that of the source curve, except that the smoothing parameters of two nodes

[2] We refer the reader to [55], and in particular Remark 2 of Section 3.9, for a detailed discussion.

have been identified if ψ matches them both with the same node of the target. We thus obtain a logarithmic Deligne–Mumford stack.

Theorem 2.2 [55, Section 3.22] *The moduli stack of admissible hyperelliptic covers $\mathscr{A}_{g,0,2}$ is a logarithmically smooth with locally free logarithmic structure* (*and therefore smooth*), *proper Deligne–Mumford stack*, *with a logarithmically étale morphism to $\overline{\mathcal{M}}_{0,2g+2}/\mathfrak{S}_{2g+2}$.*

To an admissible hyperelliptic cover we can associate a harmonic morphism of degree two to a metric tree $\psi : \sqsubset \to \mathsf{T}$. The tropical geometry of hyperelliptic and admissible covers has been analysed, for instance, in [23; 22].

2.3 Gorenstein curves

2.3.1 Tools and formulas Curves shall always be assumed Cohen–Macaulay, ie $S1$; they might still be nonreduced along some subcurve, but they have no embedded points.

Definition 2.3 A curve X is Gorenstein if its dualising sheaf ω_X is a line bundle.

A fundamental role in the study of singularities is played by the conductor ideal.

Definition 2.4 Let $\nu : \widetilde{X} \to X$ be a finite and birational morphism. The *conductor ideal* of ν is

$$\mathfrak{c} = \mathrm{Ann}_{\mathscr{O}_X}(\nu_*\mathscr{O}_{\widetilde{X}}/\mathscr{O}_X) = \mathscr{H}om_X(\nu_*\mathscr{O}_{\widetilde{X}}, \mathscr{O}_X).$$

The conductor is the largest $\mathscr{O}_X$–ideal sheaf that is also a $\nu_*\mathscr{O}_{\widetilde{X}}$–ideal sheaf. When X is a reduced curve with finite normalisation — which is always the case over a field — blowing up the conductor ideal recovers the normalisation [79]. The conductor ideal admits a further characterisation in terms of duality theory as

$$\mathfrak{c} = \nu^!\mathscr{O}_X = \omega_\nu.$$

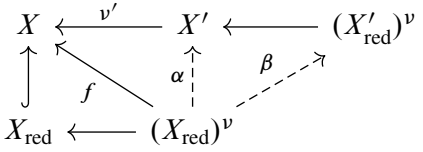
Remark 2.5 Let us restrict to schemes of finite type over a field. For such an X, the normalisation $X^\nu \to X$ has two universal properties (see eg [16, Tag 035Q]):

(1) It is final in the category of dominant morphisms $Y \to X$ with Y normal.

(2) For every finite and birational morphism $Y \to X$, the normalisation of X factors uniquely through a morphism $X^\nu \to Y$.

Now suppose that X is a Cohen–Macaulay curve. The smooth (possibly disconnected) curve $(X_{\mathrm{red}})^\nu$ is what is normally referred to as the normalisation of X. Notice that

$(X_{\mathrm{red}})^\nu$ can be obtained as a (sequence of) blow-up(s) along a subscheme (points) δ of X_{red}. Set $X' = \mathrm{Bl}_\delta X$ and consider the diagram



The arrow α exists by the universal property of blowing up (because $f^{-1}\mathcal{I}_{\delta/X}$ is principal in $\mathcal{O}_{(X_{\mathrm{red}})^\nu}$). Since $(X_{\mathrm{red}})^\nu$ is reduced and normal, α factors through $(X'_{\mathrm{red}})^\nu$; hence, the arrow β exists. Since it is a finite and birational morphism of normal curves, β is an isomorphism. Abusing notation, we are going to call X' the *normalisation* of X. Note that the underlying reduced curve of every irreducible component is regular. Moreover, the sheaf $\nu'_*\mathcal{O}_{X'}/\mathcal{O}_X$ is supported along the singular locus of X_{red}. Suppose that X' contains a ribbon (see Definition 2.17) or, more generally, a *multiple curve* in the sense of [29]; the reader should be aware that, if we only ask of a morphism $g\colon Y \to X$ that the sheaf $g_*\mathcal{O}_Y/\mathcal{O}_X$ be supported at points, we can construct infinitely many more Y by blowing up X' successively at a number of smooth points of X_{red}. Indeed, for ribbons, these are essentially all the morphisms that restrict to the identity on the underlying $\mathbb{P}^1$ yet alter the square-zero ideal that defines the double structure (see [14, Theorem 1.10]).

It is known that, for a nodal curve X, the conductor ideal relates the dualising line bundle of X to that of its normalisation X^ν; more generally [20, Proposition 1.2]:

Proposition 2.6 (Noether's formula) *Let $\nu\colon \widetilde{X} \to X$ be a finite birational morphism of Gorenstein curves. Then, viewing $\mathfrak{c}$ as an ideal sheaf on $\widetilde{X}$,*

$$\omega_{\widetilde{X}} = \nu^*\omega_X \otimes \mathfrak{c}.$$

Corollary 2.7 *Let X be a Gorenstein curve with Gorenstein normalisation $\nu\colon \widetilde{X} \to X$. The conductor is a principal ideal on $\widetilde{X}$.*

Definition 2.8 For a finite birational morphism $\nu\colon \widetilde{X} \to X$ of curves, the coherent $\mathcal{O}_X$–module $\nu_*\mathcal{O}_{\widetilde{X}}/\mathcal{O}_X$ has finite support; its length is called the *δ–invariant* of ν. When ν is the normalisation, we simply call it the δ–invariant of X. When X is reduced, it is the sum of the local contributions of all the isolated singularities of X.

We review a result of J-P Serre [66, Section 4, Proposition 7] for possibly nonreduced curves.

Lemma 2.9 *If X is a Gorenstein curve and $\widetilde{X}$ its normalisation, then*

$$\text{length}(\mathscr{O}_{\widetilde{X}}/\mathfrak{c}) = 2\delta. \tag{1}$$

Proof Since the normalisation is finite and birational, we may work locally around the support of $\nu_*\mathscr{O}_{\widetilde{X}}/\mathscr{O}_X$, which we may therefore assume is a single point. Let $X = \text{Spec}(R)$ and $\widetilde{X} = \text{Spec}(\widetilde{R})$. Using $\mathfrak{c} \subseteq R \subseteq \widetilde{R}$, we see that (1) is equivalent to

$$\dim_{\mathbb{C}}(R/\mathfrak{c}) = \delta.$$

Let ω be a local generator of Ω_R; given $f \in R$, Noether's formula implies that $f\omega \in \Omega_{\widetilde{R}}$ if and only if $f \in \mathfrak{c}$. We therefore get

$$R/\mathfrak{c} \hookrightarrow \Omega_R/\Omega_{\widetilde{R}}.$$

Since Ω_R is free of rank one and ω a generator, the above map is surjective as well. Finally, applying $\text{Hom}_R(-, \Omega_R)$ to the normalisation exact sequence, we obtain $\Omega_R/\Omega_{\widetilde{R}} = \text{Ext}^1_R(\widetilde{R}/R, \Omega_R)$, and therefore $\dim_{\mathbb{C}} \Omega_R/\Omega_{\widetilde{R}} = \delta$. □

Remark 2.10 The converse is true when X is reduced, but may fail otherwise, an example of which is provided by the subalgebra of $\mathbb{C}[\![s,\epsilon]\!] \times \mathbb{C}[\![t]\!]$ generated by $\langle s, \epsilon, t\rangle$, ie the transverse union of a ribbon with a line, which, though satisfying (1), is not Gorenstein as a consequence of the following lemma. What seems to be lacking in the nonreduced case is a good theory of dualising sheaves as rational differential forms on the normalisation satisfying some residue condition.

Definition 2.11 A curve is *decomposable* if it is obtained by gluing two curves along a (reduced closed) point.

Remark 2.12 This is equivalent to [70, Definition 2.1]. Indeed, the fibre product over $\mathbb{C}$ of $\mathbb{C}[x_1,\dots,x_m]/I_1$ and $\mathbb{C}[y_1,\dots,y_n]/I_2$ is isomorphic to

$$\underbrace{\mathbb{C}[x_1,\dots,x_m,y_1,\dots,y_n]}_{S} \Big/ \underbrace{(I_1(\boldsymbol{x}), y_1,\dots,y_n)}_{J_1} \cap \underbrace{(x_1,\dots,x_m, I_2(\boldsymbol{y}))}_{J_2}. \tag{2}$$

On the other hand, with notation as in (2), there is a short exact sequence

$$0 \to S/J_1 \cap J_2 \to S/J_1 \times S/J_2 \to S/J_1 + J_2 \simeq \mathbb{C} \to 0,$$

which is exact in the middle because every element there can be represented as a pair of polynomials in $\boldsymbol{x}$ only (resp. in $\boldsymbol{y}$ only), which is easy to lift.

We review [6, Proposition 2.1] for not necessarily reduced curves.

Lemma 2.13 *A decomposable curve may be Gorenstein only if it is a node.*

Proof Let us assume that X is the decomposable union of X_1 and X_2.

Notice first that, if X is Gorenstein, then so are X_1 and X_2. Indeed, say we can find $(a_1,\dots,a_n)\subseteq \mathfrak{m}_X$ a regular sequence such that $\mathscr{O}_{X^0}=\mathscr{O}_X/(a_1,\dots,a_n)$ is Gorenstein of dimension zero, ie $\operatorname{Hom}_{X^0}(\mathbb{C},\mathscr{O}_{X^0})\cong\mathbb{C}$ (and all higher Ext groups vanish, which is automatic by [54, Theorem 18.1]). Since $\mathfrak{m}_X=\mathfrak{m}_{X_1}\oplus\mathfrak{m}_{X_2}$, this defines a regular sequence in the latter as well, and $\mathscr{O}_{X_1}\oplus\mathscr{O}_{X_2}/(a_1,\dots,a_n)\cong\mathscr{O}_{X_1^0}\oplus\mathscr{O}_{X_2^0}$ is an extension of k with $\mathscr{O}_{X^0}$. From this we see that $\dim_{\mathbb{C}}\operatorname{Hom}_{X_i^0}(\mathbb{C},\mathscr{O}_{X_i^0})=1$ for $i=1,2$, which is to say that X_1 and X_2 are Gorenstein as well.

Now, X and $X_1\sqcup X_2$ have the same normalisation $\widetilde{X}$, and

$$\delta_X=\delta_{X_1}+\delta_{X_2}+1. \tag{3}$$

On the other hand, $\mathscr{O}_{\widetilde{X}}/\mathscr{O}_X\twoheadrightarrow\mathscr{O}_{\widetilde{X}}/\mathscr{O}_{X_1\sqcup X_2}$ implies that $\mathfrak{c}_X\subseteq\mathfrak{c}_{X_1\sqcup X_2}$ (as ideals of $\mathscr{O}_{\widetilde{X}}$). We claim that, if X is Gorenstein, the reduced curves underlying X_1 and X_2 are both smooth:

- If they are both singular, then $\mathfrak{c}_{X_1\sqcup X_2}\subseteq\mathfrak{m}_{X_1}\oplus\mathfrak{m}_{X_2}=\mathfrak{m}_X$, so $\mathfrak{c}_X=\mathfrak{c}_{X_1\sqcup X_2}$. By Lemma 2.9, we obtain $2\delta_X=\dim_{\mathbb{C}}(\mathscr{O}_{\widetilde{X}}/\mathfrak{c}_X)=\dim_{\mathbb{C}}(\mathscr{O}_{\widetilde{X}}/\mathfrak{c}_{X_1\sqcup X_2})=2\delta_{X_1}+2\delta_{X_2}$, which contradicts (3).
- If X_1 is singular and X_2 smooth, (3) reduces to $\delta_X=\delta_{X_1}+1$. On the other hand, $\mathfrak{c}_{X_2}=\mathscr{O}_{X_2}$, and the contradiction comes from $2\delta_X=\dim_{\mathbb{C}}(\mathscr{O}_{\widetilde{X}}/\mathfrak{c}_X)=\dim_{\mathbb{C}}(\mathscr{O}_{\widetilde{X}}/\mathfrak{c}_{X_1})=2\delta_{X_1}$.

Finally, since both underlying curves are smooth, we see from the definition of decomposability that

$$\widehat{\mathscr{O}}_X=\mathbb{C}[\![x,\epsilon_i,y,\eta_j]\!]/(xy,x\eta_j,\epsilon_i y,\epsilon_i\eta_j,\epsilon_i^{m_i},\eta_j^{n_j})_{i=1,\dots,h,\ j=1,\dots,k}.$$

If X_1 and X_2 are reduced, we recover the node. In all other cases, one can verify that $\dim_{\mathbb{C}}\operatorname{Hom}(\mathbb{C},\mathscr{O}_X)\geq 2$ (it is generated by $xy, x\epsilon_1^{m_1-1},\dots$). □

2.3.2 Isolated singularities Let (X,x) be (the germ of) a reduced curve with a unique singular point x, with normalisation $\nu\colon\widetilde{X}\to X$. The following is a measure of how much of the arithmetic genus of a projective curve is hiding in its singularities:

Definition 2.14 [67] If X has m branches at x, the *genus* of (X,x) is

$$g=\delta-m+1.$$

The classification of isolated Gorenstein singularities of genus one has been carried out by Smyth [67, Proposition A.3].

Proposition 2.15 *An (X, x) of genus one with m branches is locally isomorphic to*:

- $\boldsymbol{m = 1}$ *The cusp $V(y^2 - x^3) \subseteq \mathbb{A}^2_{x,y}$.*
- $\boldsymbol{m = 2}$ *The tacnode $V(y^2 - yx^2) \subseteq \mathbb{A}^2_{x,y}$.*
- $\boldsymbol{m \geq 3}$ *The union of m general lines through the origin of $\mathbb{A}^{m-1}$.*

All of these singularities are smoothable. Choosing a one-parameter smoothing and passing to a regular semistable model, the *semistable tail*, ie the subcurve contracted to the singularity, admits a simple description: if we mark the semistable tail by the intersection with the rest of the central fibre, then it is a *balanced* nodal curve of genus one. This means that it consists of a genus one core — which can be either smooth, or a circle of rational curves — together with some rational trees supporting the markings, and the distance between a marking and the core — ie the length of the corresponding rational chain — is independent of the chosen marking [67, Proposition 2.12].

The classification of isolated Gorenstein singularities of genus two has been carried out by the first author [11, Section 2].

Proposition 2.16 *The unique unibranch Gorenstein singularity of genus two is the* ***ramphoidal cusp*** *or A_4–singularity $V(y^2 - x^5) \subseteq \mathbb{A}^2_{x,y}$. For every $m \geq 2$, there are exactly two isomorphism classes of germs of isolated Gorenstein singularities of genus two*; *see Table 1.*

m	type I	type II
	$x_1 = t_1 \oplus 0 \oplus \cdots \oplus t_m^3$ $x_2 = 0 \oplus t_2 \oplus \cdots \oplus t_m^3$ $\vdots$ $x_{m-1} = 0 \oplus \cdots \oplus t_{m-1} \oplus t_m^3$ $x_m = 0 \oplus \cdots \oplus 0 \oplus t_m^2$	$x_1 = t_1 \oplus 0 \oplus \cdots \oplus t_m$ $x_2 = 0 \oplus t_2 \oplus \cdots \oplus t_m^2$ $\vdots$ $x_{m-1} = 0 \oplus \cdots \oplus t_{m-1} \oplus t_m^2$ $(y = 0 \oplus t_2^3$ if $m = 2)$
1	$x^5 - y^2$ (A_4)	
2	$x_2(x_2^3 - x_1^2)$ (D_5)	$y(y - x_1^3)$ (A_5)
3	$\langle x_3(x_1 - x_2), x_3^3 - x_1x_2 \rangle$	$x_1x_2(x_2 - x_1^2)$ (D_6)
≥ 4	$\langle x_i(x_j - x_k), x_m(x_i - x_j), x_m^3 - x_1x_2 \rangle$ $(i \neq j \neq k \in \{1, \ldots, m-1\})$	$\langle x_3(x_1^2 - x_2), x_i(x_j - x_k) \rangle$ $(1 \leq i < j < k \leq m-1$ or $1 < j < k < i \leq m-1)$

Table 1: The isolated Gorenstein singularities of genus two. The first row gives the parametrisation; the second gives the equation(s) for different values of m.

It is not hard to see that, for $m \geq 3$, every type I singularity is the union of a cusp with $m-1$ lines, and every type II singularity is the union of a tacnode with $m-2$ lines; in each case, we refer to the components of the genus one subcurve as the *special branches*.

The description of semistable tails is a bit more cumbersome, but enlightening: assuming that the genus two core is smooth, the special branches are closer to the core and attached to a Weierstrass point (resp. two conjugate points) in type I (resp. II); all other branches are equidistant and further away from the core. In fact, the ratio between the length of the special rational chain and the others is fixed to $\frac{1}{3}$ $\left(\text{resp. } \frac{1}{2}\right)$ in type I (resp. II). For a more detailed statement when the core is not smooth, see [11, Propositions 4.3–4.6].

2.3.3 Nonreduced structures Multiple curves were investigated in the 1990s in connection to Green's conjecture [14]. Ribbons, in particular, were understood to arise as limits (in the Hilbert scheme of $\mathbb{P}^{g-1}$) of canonical curves, as the curve becomes hyperelliptic [31].

Definition 2.17 A *ribbon* is a double structure on $\mathbb{P}^1$, ie a nonreduced curve R with $R_{\mathrm{red}} = \mathbb{P}^1$ defined by a square-zero ideal $\mathcal{I}_{R_{\mathrm{red}}/R}$ that is a line bundle on R_{red}.

Example 2.18 There is only one ribbon of genus two, R_2, up to isomorphism: it is the first infinitesimal neighbourhood of the zero section in $\mathrm{Tot}_{\mathbb{P}^1}(\mathcal{O}(3))$. The short exact sequence

$$0 \to \mathcal{I}_{\mathbb{P}^1/R_2} \simeq \mathcal{O}_{\mathbb{P}^1}(-3) \to \mathcal{O}_{R_2} \to \mathcal{O}_{\mathbb{P}^1} \to 0$$

is split by restricting the projection $\mathrm{Tot}_{\mathbb{P}^1}(\mathcal{O}(3)) \to \mathbb{P}^1$ to R_2. R_2 is therefore called a *split ribbon*. In fact, all ribbons of genus at most two are split.

Automorphisms and moduli of multiple curves have been studied in [29]. We are going to encounter nonreduced structures along singular curves as well.

Example 2.19 The simplest example of a Gorenstein nonreduced structure with singular underlying curve is given by the union of a ribbon and a line *along a double point*. Local equations are $\mathbb{C}[\![x, y]\!]/(x^2 y)$. It is easy to see that, for such a curve to have genus two, the ribbon needs to have ideal $\mathcal{I}_{\mathbb{P}^1/R} \simeq \mathcal{O}_{\mathbb{P}^1}(-2)$. Such a curve can be realised in the linear system $|2D_+ + F|$ on the Hirzebruch surface $\mathbb{F}_2$, where D_+ denotes the class of the positive section and F the class of a fibre.

Example 2.20 More generally, the union of a ribbon, R, with a rational k–fold point along a double point (representing a generic tangent vector to the k–fold point) is Gorenstein. Local equations are given by

$$\mathbb{C}[\![x_1,\dots,x_k,y]\!]/(x_ix_j,(x_i-x_j)y)_{1\le i<j\le k}.$$

To see that this is Gorenstein, it is enough to find a regular element ξ such that the quotient is Gorenstein of dimension zero. Let $\xi=\sum_{i=1}^k x_i-y$. The quotient A is a graded finite-dimensional algebra, with $A_0=\mathbb{C}$, $A_1=\mathbb{C}\langle x_1,\dots,x_k\rangle$ and $A_2=\mathbb{C}\langle x_1^2=\cdots=x_k^2\rangle$, having one-dimensional socle A_2.

Now we can obtain a Gorenstein projective curve of genus two, C, by gluing a ribbon together with some k_i–fold points ($i=1,\dots,r$) at distinct (closed) points of the ribbon, by iterating the local construction above. From the short exact sequence

$$0\to\mathscr{O}_C\to\mathscr{O}_R\oplus\bigoplus_{i=1}^r\mathscr{O}_{\mathbb{P}^1}^{\oplus k_i}\to\bigoplus_{i=1}^r(\mathbb{C}^{k_i-1}\oplus\mathbb{C}[\epsilon])\to 0,$$

we can compute the Euler characteristic of R, whose structure sheaf thus satisfies

$$0\to\mathscr{O}_{\mathbb{P}^1}(r-3)\to\mathscr{O}_R\to\mathscr{O}_{\mathbb{P}^1}\to 0,$$

depending only on the total number of "noded" points r and not on the number of branches of each k_i–fold point.

Definition 2.21 We call C as in Example 2.20 a $(k_1,\dots,k_r)$–*tailed ribbon* of genus two.

Remark 2.22 We can employ Noether's formula (Proposition 2.6) and adjunction (on a surface containing the ribbon, ie $\mathrm{Tot}(\mathscr{O}_{\mathbb{P}^1}(3-r))$) to compute the restriction of the dualising sheaf to every component of a $(k_1,\dots,k_r)$–tailed ribbon C. First, the "normalisation" is given by

$$\begin{aligned}\mathbb{C}[\![x_1,\dots,x_k,y]\!]/(x_ix_j,(x_i-x_j)y)_{1\le i<j\le k}&\to\mathbb{C}[\![s]\!][\epsilon]/(\epsilon^2)\times\mathbb{C}[\![t_1]\!]\times\cdots\times\mathbb{C}[\![t_k]\!]\\ x_i&\mapsto(\epsilon,0,\dots,t_i,\dots,0),\quad i=1,\dots,k,\\ y&\mapsto(s,0,\dots,0).\end{aligned}$$

From this we compute the conductor $\mathfrak{c}=\langle x_1^2,\dots,x_k^2,y\rangle$. The restriction to a tail is

$$\omega_C|_T\simeq\omega_T\otimes\mathscr{O}(2q)\simeq\mathscr{O}_T.$$

The restriction to the $\mathbb{P}^1$ underlying the ribbon is

$$\omega_C|_{R_{\mathrm{red}}}\simeq\omega_{R_{\mathrm{red}}}\otimes\mathscr{I}^\vee_{R_{\mathrm{red}}/R}\otimes\mathfrak{c}^\vee_{R_{\mathrm{red}}}\simeq\mathscr{O}_{R_{\mathrm{red}}}(-2+(3-r)+r)\simeq\mathscr{O}_{R_{\mathrm{red}}}(1).$$

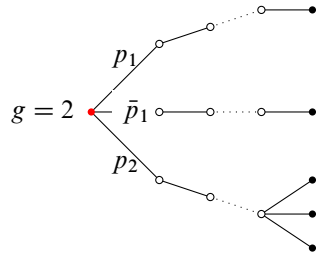


Figure 1: Dual graph of the central fibre of a regular semistable model for a $(2,3)$–tailed ribbon with p_1 and $\bar{p}_1$ conjugate.

For a nodal curve C with core Z, we say that a component D of C *cleaves* to $q \in Z$ if D is joined to Z at q through a chain of rational curves.

Lemma 2.23 *Regular semistable models of a $(k_1, \dots, k_r)$–tailed ribbon can be classified. If the core is a smooth curve of genus two, Z, the sets of tails $\{T_1^i, \dots, T_{k_i}^i\}_{i=1,\dots,r}$ cleave to the same (or to conjugate) point(s) of Z, and are all equidistant from the core (independently of i and j in T_j^i). The configuration of attaching points on Z is such that the hyperelliptic cover maps it to the corresponding configuration of noded points on R_{red}, up to reparametrisation.*

Proof This is a word-by-word repetition of the argument of [11, Proposition 4.3]. □

See Figure 1 for an example of a semistable tail of a tailed ribbon.

Remark 2.24 The description of the semistable models of a $(k_1, \dots, k_r)$–tailed ribbon (Lemma 2.23) in case the former has reducible core is more cumbersome. Indeed, the length of a tree depends on which component of the core it is attached to. See Proposition 2.26 and Figure 2 for a precise description in terms of piecewise-linear functions on the tropicalisation. A ribbon arises when the special component (which we denote pictorially by a red dot) belongs to the core. The piecewise-linear function then has constant slope 1 on any tree emanating from the core, and in this sense it is reminiscent of the genus one situation.

Example 2.25 A semistable model of the 1–tailed ribbon C (Example 2.19) can be computed by taking the pencil

$$u_+^2 f_1 - t u_-^2 p_5(f_1, f_2) = 0$$

on $\mathbb{F}_2 \times \mathbb{A}^1_t$, with p_5 a generic homogeneous polynomial of degree 5 in two variables. The resulting smoothing family is singular at six points along the central fibre, including

the node of C. After blowing them up and normalising, we obtain a genus two curve covering the ribbon two-to-one; the tail of C is attached to a Weierstrass point. This appears to be an accident due to the restriction that the smoothing comes from a pencil on $\mathbb{F}_2$.

In order to describe semistable tails in complete generality, there is no better way than expressing the problem in terms of existence of a certain piecewise-linear function on the dual graph. We refer the reader to Sections 2.1.4 and 2.2 for background and notation.

Proposition 2.26 *Let C be a Gorenstein curve of arithmetic genus two. Let $\mathscr{C} \to \Upsilon$ be a one-parameter smoothing of C. Let $\phi\colon \mathscr{C}^{\mathrm{ss}} \to \mathscr{C}$ be a semistable model with exceptional locus Z. Then, up to destabilising $\mathscr{C}^{\mathrm{ss}}$, there exists a hyperelliptic admissible cover $\psi : \mathscr{C}^{\mathrm{ss}} \to \mathscr{T}$ and a tropical canonical divisor $\lambda : \sqsubset \to \mathbb{R}$ pulled back along* $\mathrm{trop}(\psi)$ *such that Z is cut out by $\mathscr{O}_{\mathscr{C}^{\mathrm{ss}}}(-\lambda)$, and we have $\omega_{\mathscr{C}^{\mathrm{ss}}} = \phi^*\omega_{\mathscr{C}}(\lambda)$.*

Proof The proof in the isolated case has appeared in [11, Section 4.4]. The argument is insensitive to rational tails away from the core, so we may as well assume that C is minimal (see Definition 2.29 below). In this case, the canonical linear series gives a two-to-one map $\bar{\psi}\colon \mathscr{C} \to \mathbb{P}_{\Upsilon}(\pi_*\omega_{\mathscr{C}}) =: \mathbb{P}$. The general fibre is the hyperelliptic cover of a smooth curve of genus two. Applying semistable reduction to the target and branch divisor, we may lift $\bar{\psi}$ to a map of nodal curves, thanks to the properness of the moduli space of admissible covers

$$\begin{array}{ccc} \mathscr{C}^{\mathrm{ss}} & \xrightarrow{\psi} & \mathscr{T} \\ \downarrow{\scriptstyle \phi} & & \downarrow{\scriptstyle \phi_T} \\ \mathscr{C} & \xrightarrow{\bar{\psi}} & \mathbb{P} \end{array}$$

The line bundles $\phi_T^*\mathscr{O}_{\mathbb{P}}(1)$ and $\omega_{\mathscr{T}}\left(\frac{1}{2}D_B\right)$ have the same total degree on $\mathscr{T}$, so, since the latter is a tree, their difference is associated to a piecewise-linear function λ_T on T. Pulling back via $\mathscr{C}$, on the other hand, we find

$$\phi^*\omega_{\mathscr{C}} = \omega_{\mathscr{C}^{\mathrm{ss}}}(-\mathrm{trop}(\psi)^*\lambda_T). \qquad \square$$

For the reader's benefit, Figure 2 provides a pictorial description of such piecewise-linear functions in case the core is a genus two configuration of rational curves, including the *slopes* along the nodes (or rational chains). This is where the *admissible functions* of Section 3.2 originate from. The two columns correspond to the two maximally

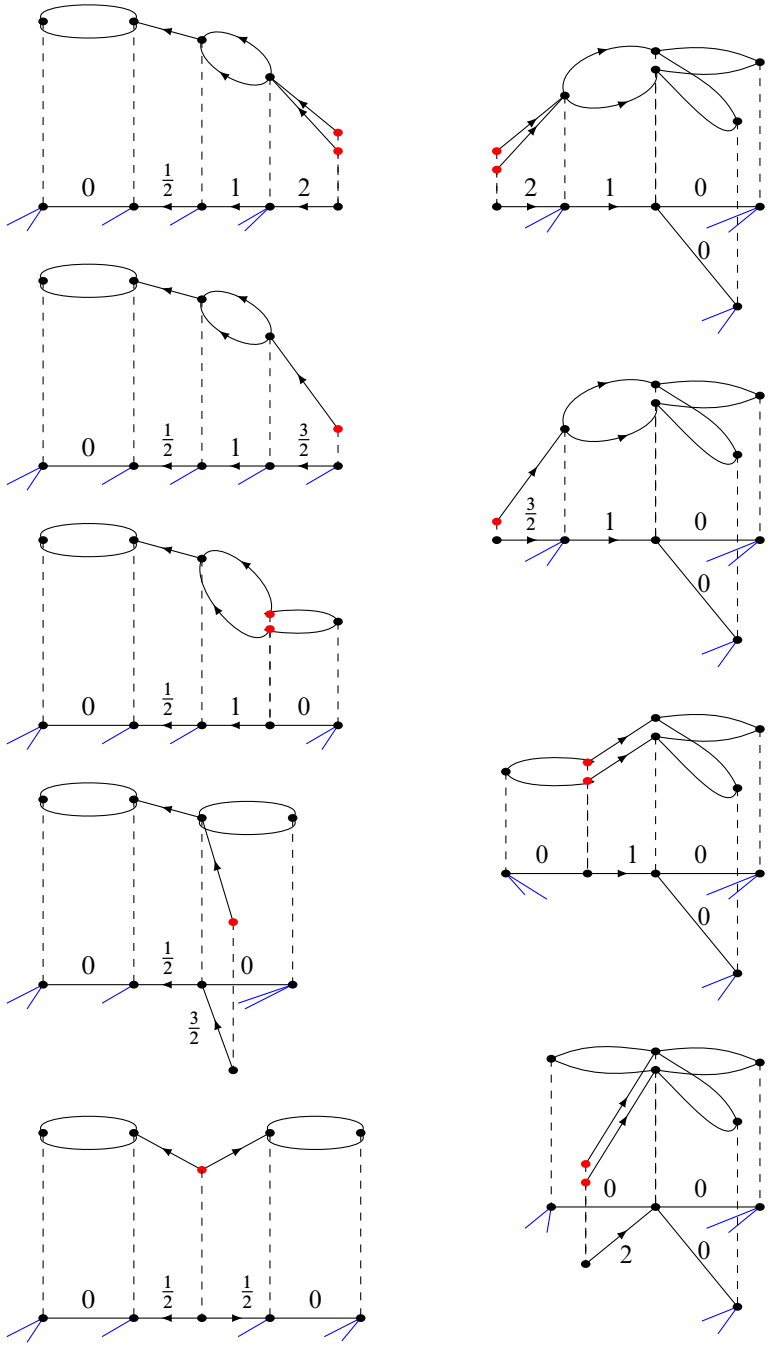


Figure 2: Admissible functions on maximally degenerated tropical hyperelliptic curves: the dumbbell (left) and the theta graph (right).

degenerate (stable) tropical curves of genus two. Here, each red vertex corresponds to a component mapping to a special component of the singularity/ribbon. Blue legs represent branch points of the admissible cover. Tails are not drawn, but can be understood by a tropical modification: the corresponding vertices would lie on trees of slope 1 towards the core, all at the same height, and lower than each red vertex. See also the conventions set out in Section 3.2.

Remark 2.27 The nonreduced structures of genus two are numerous; indeed, as opposed to the case of isolated singularities, where the δ–invariant is always related to the number of branches via the genus, in the nonreduced case a very large δ–invariant can always be compensated by lowering the genus of the ribbon (which can be negative).

Yet, if we require $\omega_C \geq 0$, we find a characterisation of tailed ribbons.

Proposition 2.28 *Let C be a Gorenstein curve of genus two, consisting of a ribbon glued **in some way** to a number of (possibly singular) rational curves. Assume that the dualising sheaf is nonnegative, and positive on the ribbon. Then C is a $(k_1, \dots, k_r)$–tailed ribbon.*

Proof We may assume that the normalisation sequence is

$$0 \to \mathscr{O}_C \to \mathscr{O}_R \oplus \bigoplus_{i=1}^{r} \mathscr{O}_{\mathbb{P}^1}^{\oplus k_i} \to \bigoplus_{i=1}^{r} \mathscr{F}_i \to 0,$$

where the $\mathscr{F}_i$ are sheaves supported at the gluing points; call δ_i their lengths. We claim that $\delta_i \geq k_i$; indeed, this is the minimum number of conditions that we need to impose to ensure that the values of the functions on different branches agree at the preimage of the singular point. But in fact we may even assume $\delta_i \geq k_i + 1$, by the Gorenstein assumption and Lemma 2.13.

Now it is easy to compute that, for the genus of C to be two, we need

$$\mathscr{I}_{R_{\mathrm{red}}/R} \simeq \mathscr{O}_{\mathbb{P}^1}\left(\sum_{i=1}^{r}(\delta_i - k_i) - 3\right).$$

So, by applying Noether's formula and adjunction as in Remark 2.22, we find

$$\omega_C|_{R_{\mathrm{red}}} \simeq \omega_{\mathbb{P}^1} \otimes \mathscr{I}_{R_{\mathrm{red}}/R}^{\vee} \otimes (\mathfrak{c}|_{R_{\mathrm{red}}})^{\vee} \simeq \mathscr{O}_{\mathbb{P}^1}\left(1 + \sum_{i=1}^{r}(k_i - \deg(\mathfrak{c}_i|_{R_{\mathrm{red}}}) - \delta_i)\right).$$

We study the local contributions $k_i - \deg(\mathfrak{c}_i|_{R_{\mathrm{red}}}) - \delta_i$ around each singular point. We are then looking for Gorenstein subalgebras A of $\widetilde{A} = \mathbb{C}[\![s]\!][\epsilon](\epsilon^2) \times \mathbb{C}[\![t_1]\!] \times \cdots \times \mathbb{C}[\![t_k]\!]$. By Corollary 2.7, we know that $\mathfrak{c}$ is principal as an ideal of $\widetilde{A}$; up to rescaling the generators, we may assume that $\mathfrak{c}$ is generated by the element

$$(s^c + \epsilon s^d, t_1^{c_1}, \dots, t_k^{c_k}).$$

Indeed, under our assumptions, the degree of ω_C has to be one on R_{red} and zero on every tail; thus, it is easy to see that $c_1 = \cdots = c_k = 2$. Now, if $d \geq c$, then $s^c \in \mathfrak{c}$; if $d < c$, then $s^{2c-d} \in \mathfrak{c}$. In the second case,

$$\widetilde{A}/\mathfrak{c} = \langle 1, s, \epsilon, \dots, \epsilon s^d (= -s^c), s^{d+1}, \dots, s^{c-1}, \epsilon s^{c-1} (= s^{2c-d-1}); 1, t_1; \dots; 1, t_k \rangle.$$

In any case we find $2\delta = \dim_{\mathbb{C}} \widetilde{A}/\mathfrak{c} = 2c + 2k$ (by Lemma 2.9), and $\delta_i = k_i - \deg(\mathfrak{c}_i|_{R_{\mathrm{red}}})$ for all i.

By assumption (that the reduced subcurve underlying R is smooth), we must have a generator y of A whose linear part contains $s + \epsilon p(s)$. By Lemma 2.13 and $t_i^2 \in \mathfrak{c}$, there must be generators with nontrivial linear part in t_i. Up to taking polynomial combinations of these generators, we may assume they take the form

$$y = s + \epsilon p(s), \qquad x_i = \epsilon q_i(s) \oplus t_i, \quad i = 1, \dots, k,$$

where p and q_i are monomial. Since $1 + \epsilon \tilde{p}(s)$ is invertible in $\widetilde{A}$, we may assume that $y = s$. Furthermore, by looking at the conductor ideal, we see that for x_i we may assume $q_i(s) = q_i s^{c-1}$, with $q_i \in \mathbb{C}^\times$. Up to blowing $\widetilde{A}$ down (see Remark 2.5), we may reduce to the case $c = 1$. Thus, we find the singularities of Example 2.20, with local equations

$$\mathbb{C}[\![x_1, \dots, x_k, y]\!]/(x_i x_j, x_i y - x_j y)_{1 \leq i < j \leq k}$$

and ideal sheaf $\mathscr{I}_{R_{\mathrm{red}}/R} = \mathscr{O}_{R_{\mathrm{red}}}(r-3)$. □

2.4 h^1–vanishing

First we describe all the possible cores. Recall the following:

Definition 2.29 A curve is *minimal* if there is no proper subcurve of the same arithmetic genus.

Remark 2.30 Up to replacing nodes with rational bridges, minimal curves of genus one are either irreducible or elliptic m–fold points with rational branches ($m \geq 2$).

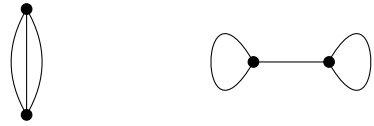


Figure 3: Minimal theta and dumbbell configurations (dual graphs).

Proposition 2.31 *Up to replacing nodes with rational bridges, a* ***minimal*** *Gorenstein curve of genus two with isolated singularities is one of the following*:

(1) *Irreducible.*

(2) *The nodal union of two minimal curves of genus one* (including a dumbbell configuration of $\mathbb{P}^1$s; see Figure 3).

(3) *A theta configuration of* $\mathbb{P}^1$*s* (see Figure 3).

(4) *A genus one singularity with a minimal genus one branch.*

(5) *A type I singularity with rational branches.*

(6) *A type II singularity with rational branches.*

Lemma 2.32 *With numbering as in the previous proposition, the dualising line bundle of a minimal Gorenstein curve of genus two has*

(4) *degree* 2 *on the genus one branch, or on the common branch of the two elliptic* m*–fold points*;

(5) *degree* 2 *on the special branch*;

(6) *multidegree* $(1,1)$ *on the special branches.*

Proof This follows from Noether's formula (Proposition 2.6) and an explicit calculation. □

Now we describe sufficient conditions for the vanishing of higher cohomology of a line bundle on a minimal curve of genus two. This will allow us to prove the unobstructedness of the λ–aligned admissible maps satisfying the factorisation property. The proof of the following lemmas boils down to a simple albeit tedious application of the normalisation exact sequence.

Lemma 2.33 *Let C be a minimal Gorenstein curve of genus two with isolated singularities. A line bundle L on C having nonnegative multidegree, positive degree on every subcurve of genus one, degree at least two and $L \neq \omega_C$ has vanishing h^1.*

Lemma 2.34 *Let C be a $(k_1, \dots, k_r)$–tailed ribbon and L a nonnegative line bundle on it. Then $h^1(C, L) = 0$ if*

- *L has positive degree on at least two k–fold points; or*
- *L restricts to $\mathcal{O}_{\mathbb{P}^1}(1)$ on R_{red} (where R denotes the component of multiplicity two), and it has positive degree on at least one k–fold point.*

3 Gorenstein curves over aligned admissible covers

Definition 3.1 A weighted admissible cover consists of a hyperelliptic admissible cover

$$\psi : (C, D_R) \to (T, D_B)$$

with prestable target, together with a weight function $w : V(\sqsubset) \to \mathbb{N}$, such that, with the induced weight function $w_T : V(\mathsf{T}) \to \mathbb{N}$ given by

$$w_T(v) = \sum_{v' \in \psi^{-1}(v)} w(v'),$$

the rational tree T is *weighted-stable* (every weight zero component has at least three nodes or branch points). We denote by $\mathscr{A}_2^{\mathrm{wt}}$ the smooth Artin stack of weighted admissible covers where C has arithmetic genus two.

3.1 The intuition and strategy

Our goal is to produce a morphism to a Gorenstein curve $\mathscr{C} \to \bar{\mathscr{C}}$ such that a line bundle on $\bar{\mathscr{C}}$ of degree as prescribed by the weight function w would have vanishing higher cohomology. This would ensure the unobstructedness of the space of maps to projective space.

Classically, we would look for a vertical divisor $\mathcal{Z}$, supported on the exceptional locus of $\mathscr{C} \to \bar{\mathscr{C}}$, such that the contraction is associated to a line bundle of the form $\omega_{\mathscr{C}}(\mathcal{Z})$ — this guarantees that $\bar{\mathscr{C}}$ is Gorenstein. Tropically, $\mathcal{Z}$ is replaced by a PL function λ on $\sqsubset$. The relevant information encoded by λ is a collection of slopes along the edges of $\sqsubset$. These are the objects of the stack **Div**, introduced in [53, Section 4.2].

Finding λ such that $\omega_{\mathscr{C}}(\lambda)$ is trivial on the exceptional locus of $\mathscr{C} \to \bar{\mathscr{C}}$ reduces to a simple calculation of degrees — as opposed to a more complicated equality in the

Picard group of a genus two curve — because we impose that λ is pulled back from the target of the admissible cover.

When the weight of the core is zero, basically any of the positive-weight tails can be elected as the special branch of the Gorenstein singularity $\overline{C}$, be it isolated or a ribbon. The choice of λ can be thought of as the choice of a degree one divisor (ie a point) on the target of the admissible cover. The support of this divisor corresponds to the special component of the singularity.

Given a standard tropical weighted admissible cover over $\mathbb{R}_{\geq 0}$, the function λ is uniquely determined. It is a member of the tropical canonical linear series on the maximal subcurve $\Delta \subseteq \sqsubset$ (making it into a *level graph* in the sense of [9]) such that

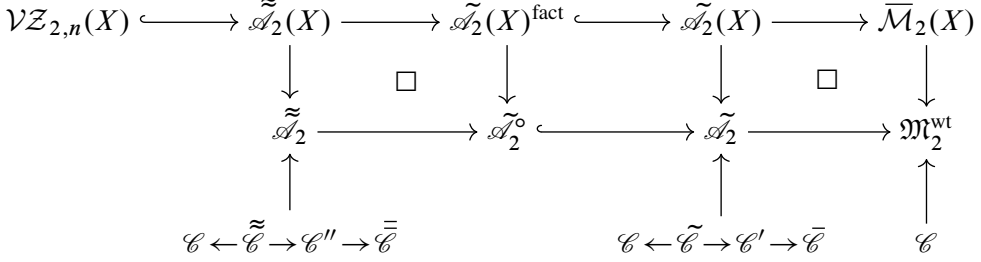
- the interior Δ° contains every subcurve of arithmetic genus g and $w \leq 2g-2$;
- every connected component Ω of Δ° has weight at most $2p_a(\Omega)-2$, and $\overline{\Omega}$ has weight at least $2p_a(\Omega)-1$.

For a more general family of tropical curves, this determines a (polyhedral and simplicial) subdivision of the base cone. In order to prove this, we actually define λ by interpolating among finitely many piecewise-linear functions, which we call *admissible*. Our definition of admissible functions has been inspired by that of Bozlee's *mesa curves* [18, Section 3]. It could be said that it constitutes a Copernican revolution with respect to [63].

When the special component has weight 2, if this represents the degree of a map, the only possibility is that the map restrict on the core to the hyperelliptic cover of a line. In this case, factorisation through the Gorenstein singularity is not enough to ensure smoothability. If we again replace the core by a ribbon, imposing a second factorisation requires the map to satisfy some ramification condition on the nearby branches, which turns out to be sufficient for the obstructions to vanish.

In order to avoid complications (and nasty singularities) arising from nonfactoring situations — ie when the core has weight 2, but it cannot possibly be the weight of the dualising bundle of a Gorenstein singularity — we proceed in two steps: First, in Section 3.2, we make sure that every subcurve of positive genus has positive weight. Then, in Section 3.4, we discard the locus where the weight of the core is 1 or $1+1$. After doing this, we can safely replace any weight two core with a ribbon. We sum up

the construction of this and the next sections in the diagram of moduli spaces



3.2 Admissible functions and aligned admissible covers

Let $\psi\colon (C, D_R) \to (T, D_B)$ be a weighted admissible cover over S, endowed with the minimal logarithmic structure. Let T denote the tropicalisation of the target. T is therefore a weighted *tree*, metrised in the monoid $\overline{M}_S$. The metric structure can be disregarded for the minute.

It follows from the Riemann–Hurwitz formula that the $\frac{1}{2}\mathbb{Z}$–divisor, supported on the vertices of T,

$$D'(v) = \mathrm{val}(v) - 2 + \tfrac{1}{2}\deg(D_B|_v), \tag{4}$$

where $\mathrm{val}(v)$ denotes the valence (number of adjacent edges) of the vertex v and D_B is the branch divisor of the cover, pulls back to the canonical divisor of $\sqsubset$ (see Section 2.1.4). The degree of D' on T is 1.

Let D be another $\frac{1}{2}\mathbb{Z}$–divisor of degree 1 on T. Since T is a tree, its Jacobian is trivial, so any two divisors of the same degree are linearly equivalent (this carries over to $\frac{1}{2}\mathbb{Z}$–divisors). Therefore, there exists a unique collection $\overline{\lambda}_T$ of half-integral slopes on the edges of T such that

$$D = D' + \mathrm{div}(\overline{\lambda}_T).$$

If D has integral coefficients, we observe that half-integral slopes may only appear along edges over which $\mathrm{trop}(\psi)$ has expansion factor two, so it is possible to think of $\overline{\lambda}_T$ as a piecewise-linear function on T with values in $\overline{M}_S^{\mathrm{gp}}$, up to a global translation of $\overline{M}_S^{\mathrm{gp}}$. In fact, it may be more accurate to think of $\overline{\lambda}_T$ as a PL function on the tropicalisation of the orbicurve $[C/\mathfrak{S}_2]$.

If D is effective, the pullback $\overline{\lambda}$ of $\overline{\lambda}_T$ along $\mathrm{trop}(\psi)$, or, rather, the divisor $K_{\sqsubset} + \mathrm{div}(\overline{\lambda})$, is an element of the tropical canonical linear system.

Definition 3.2 Let S be a geometric point. An *admissible function* $\overline{\lambda}$ on $\sqsubset$ is a piecewise-linear function with integral slopes (with values in $\overline{M}_S^{\mathrm{gp}}$, but defined only up to a global translation by $\overline{M}_S^{\mathrm{gp}}$) such that

- $\overline{\lambda}$ is the pullback along $\mathrm{trop}(\psi)$ of a (possibly half-integral) $\overline{\lambda}_T$ on T;
- $K_\sqsubset + \mathrm{div}(\overline{\lambda}) \geq 0$ defines an element of the tropical canonical system, which is the pullback of an effective, degree 1 divisor D supported on *exactly one* vertex of T.

We are going to call *special* the vertex of T supporting D, or its preimage(s) in $\sqsubset$. With an eye to the future, it will correspond to the special branch(es) of the Gorenstein curve of genus two.

Remark 3.3 There are only *finitely many* admissible functions compatible with a given ψ. Indeed, they are in bijection with the vertices of T.

When the core is maximally degenerate, ie a configuration of rational curves, all the possible admissible functions are depicted in Figure 2. The red vertices represent the ones supporting $K_\sqsubset + \mathrm{div}(\overline{\lambda})$. Missing from the picture is what happens outside the core, but this is easily explained: there may be any number of rational trees, on which $\overline{\lambda}$ has constant slope 1 towards the core.

For the sake of concreteness, we now look at some examples, adopting the:

Convention 3.4 With the application to stable maps in mind, we think of the source of an admissible cover as the destabilisation of a weighted-stable curve. In the following, we represent a component of positive weight by a black circle, and a component of weight zero by a white one, unless it is unstable (either a rational tail introduced as the conjugate of a rational tail of positive weight, or a rational bridge introduced by slicing $\sqsubset$), in which case it is represented by a cross and the corresponding edge is dotted. A red vertex represents the component supporting the divisor D, or its preimages in $\sqsubset$. The blue legs (*B–legs*) represent the branching divisor D_B of ψ (see Definition 2.1): the number of B–legs attached to a vertex v of T is the degree of $D_B \cap T_v$, where T_v is the irreducible component of T corresponding to v.

Example 3.5 Assume the core is irreducible of weight zero, and there are two tails of (large) positive weight: one of them (R) is attached to a Weierstrass point of the

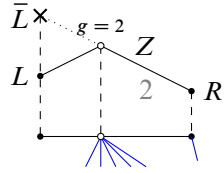


Figure 4: A weighted admissible cover with weight zero core.

core; the other one (L) is attached to a general point, so that a weight zero tail ($\bar{L}$) must be sprouted from the conjugate point in order for the admissible cover to exist (Figure 4). The tropical cover has expansion factor 2 (in grey) along the rightmost edge, corresponding to the ramification order of the algebrogeometric map at each branch of the node.

There are three admissible functions $\bar{\lambda}_i$ compatible with the given weighted admissible cover; see Figure 5.

Note that the pullback of D always has degree 2 — although the tropical cover is injective on the rightmost edge, the expansion factor 2 provides the correct multiplicity. It may be instructive to compute the multidegree of various tropical divisors on $\sqsubset$ — we represent them as vectors in $\mathbb{Z}^4$ by ordering the components from left to right, $(\bar{L}, L, Z, R)$:

	$\operatorname{div}(\bar{\lambda}_i)$	$\operatorname{trop}(\psi)^* D = K_\sqsubset + \operatorname{div}(\bar{\lambda}_i)$
(i)	$(1, 1, -3, 1)$	$(0, 0, 2, 0)$
(ii)	$(1, 1, -5, 3)$	$(0, 0, 0, 2)$
(iii)	$(2, 2, -3, 1)$	$(1, 1, 0, 0)$

The canonical $K_\sqsubset$ gives $(-1, -1, 5, -1)$.

Definition 3.6 A *prealigned* admissible cover over a logarithmic scheme (S, M_S) is one for which the values

$$\{\bar{\lambda}(v) - \bar{\lambda}(v') \mid v, v' \in V(\sqsubset)\} \subseteq (\overline{M}_S, \geq) \tag{5}$$

are comparable for every admissible function $\bar{\lambda}$ compatible with ψ (Definition 3.2).

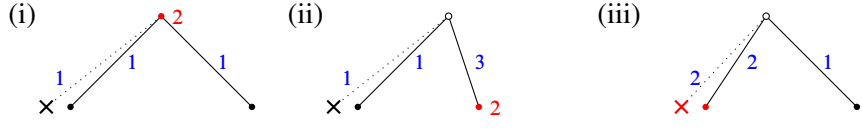


Figure 5: Admissible functions compatible with Figure 4.

Prealigned admissible covers form a subfunctor of weighted admissible covers over (LogSch). The next lemma follows from the definitions and Section 2.1.2.

Lemma 3.7 *The minimal logarithmic structure of a prealigned admissible cover is obtained, starting from the minimal logarithmic structure of the admissible cover, by adding in the elements of $\overline{M}_S^{\mathrm{gp}}$ indicated in (5) and sharpifying.*

Remark 3.8 Aligning determines a subdivision Σ of the tropical moduli space $\sigma = \mathrm{Hom}(\overline{M}_S, \mathbb{R}_{\geq 0})$ or, equivalently, a logarithmic modification of $\mathscr{A}_2^{\mathrm{wt}}$, which will be denoted by $\mathscr{A}_2^{\mathrm{pre}}$. See Section 2.1.5.

Let ψ be a prealigned admissible cover, and $\{\overline{\lambda}_1, \dots, \overline{\lambda}_s\}$ the set of admissible functions compatible with ψ. We are going to choose genuine PL functions λ_i on $\sqsubset$ with values in $\overline{M}_S^{\mathrm{gp}}$ lifting the $\overline{\lambda}_i$. A lift determines and is determined by the set of vertices mapping to 0 in $\overline{M}_S^{\mathrm{gp}}$; let us denote by $\sqsubset_{\geq 0}$ (resp. $\sqsubset_{>0}$) the set of vertices with values in $\overline{M}_S \subseteq \overline{M}_S^{\mathrm{gp}}$ (resp. $\overline{M}_S \setminus \overline{M}_S^*$; note that we may assume $\overline{M}_S^* = \{0\}$). In the end, we are going to define the object of interest as an interpolation/truncation of these lifts.

Definition 3.9 Let S be a geometric point. Define a lift $\lambda_i \in \Gamma(S, \pi_* \overline{M}_C^{\mathrm{gp}})$ of $\overline{\lambda}_i$ by requiring that $w(\sqsubset_{>0}) \leq 0$ and $w(\sqsubset_{\geq 0}) \geq 1$.

Definition 3.10 Define a PL function on a subdivision $\widetilde{\sqsubset}$ of $\sqsubset$ by

$$\lambda = \max\{0, \lambda_1, \dots, \lambda_s\}.$$

We denote by Δ° the support of λ (appeared as $\sqsubset_{>0}$ before), and by Δ the minimal subcurve of arithmetic genus two in $\widetilde{\sqsubset}$ containing the closure of Δ°.

Remark 3.11 Parallel to the above definition, we denote by Δ_T° the support of λ_T, and by Δ_T the image of Δ under $\mathrm{trop}(\psi)$.

Since tropical linear systems are tropically convex (see Section 2.1.4), it is still true that λ is the pullback of a PL function λ_T (with slopes in $\frac{1}{2}\mathbb{Z}$) on T, such that $D' + \mathrm{div}(\lambda_T)$ is an effective divisor of degree 1 on Δ_T (notice that here, in the definition of D', the valence of a vertex is the one in Δ_T and not the one in T). This is clear away from 0, where λ coincides with one of the admissible functions above; at 0, on the other hand, all the slopes of λ_T are nonnegative, so the only doubt is for a vertex of valence 1 supporting exactly one B–leg; but such a vertex corresponds to a rational tail in Δ,

so it can be included in Δ_T only if the slope of λ_T along the unique adjacent edge is strictly positive.

The divisor $D = D' + \operatorname{div}(\lambda_T)$ need not be supported on a vertex of T, and could be the sum of two half-points in the locus over which $\operatorname{trop}(\psi)$ has expansion factor 2.

Since by Riemann–Hurwitz we have $\operatorname{trop}(\psi)^*(D') = K_\Delta$,

$$\operatorname{trop}(\psi)^*(D) = K_\Delta + \operatorname{div}(\lambda). \tag{6}$$

Lemma 3.12 *On every cone of* $\operatorname{trop}(\mathscr{A}_2^{\mathrm{pre}})$, λ *is well defined as a PL function on a subdivision of* $\sqsubset$.

Proof For λ to be well defined, we need to argue that, for every vertex v of $\sqsubset$, the values $\{\lambda_i(v) \mid i = 1, \dots, s\}$ are comparable, and there is a unique way of subdividing $\sqsubset$ in order to make λ piecewise linear. Notice that, a priori, we only know that the values of a single λ_i at the vertices of $\sqsubset$ are ordered.

It is convenient to work on T. The leaves of T have positive weight due to stability, so λ_T necessarily takes the value 0 on them.

In order to prove that the maximum is well defined at every vertex, we can proceed inductively from the leaves and run through all of T. Assume that $\lambda_T(v_1)$ has been determined. We need to establish the behaviour of λ_T on the edge e between v_1 and v_2, and its value on v_2. Upon relabelling, we may suppose that λ_1 is the function with maximal slope along e among the ones with $\lambda_i(v_1) = \lambda_T(v_1)$; so λ_T coincides with λ_1 in a neighbourhood of v_1. If λ_1 is also the function of maximal slope along e among all the admissible functions, then $\lambda_T(v_2) = \lambda_1(v_2)$, and we are done.

If not, there must be a function λ_2 with slope along e greater than that of λ_1 but $\lambda_2(v_1) \le \lambda_1(v_1)$. First assume that $\overline{\lambda}_1$ and $\overline{\lambda}_2$ differ only along e, ie $D(\lambda_1, v_2) = 1$ and $D(\lambda_2, v_1) = 1$. Then there is a unique way to interpolate between λ_1 and λ_2 along e, namely by solving the equations

$$\begin{cases} \ell(e) = \ell(e') + \ell(e''), \\ \lambda_2(v_2) - \lambda_1(v_1) = s(\lambda_1, e)\ell(e') + s(\lambda_2, e)\ell(e''). \end{cases}$$

This uniquely determines the subdivision of e on which λ_T is piecewise linear, and the values of λ_T on both v_2 and the newly introduced vertex $v_{1,2}$.

More generally, λ_1 and λ_2 will differ along the chain R connecting the support of D_1 to that of D_2, and containing the edge e. Then $\lambda_2 - \lambda_1 = D_2 - D_1$ is a rational function

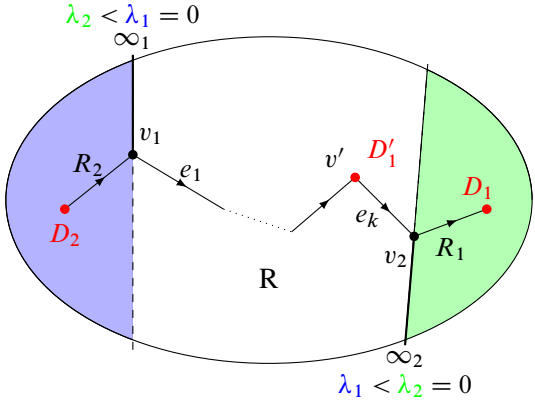


Figure 6: Comparison of λ_1 and λ_2.

with constant slope 1 along R. In order to define λ_T, we need to interpolate between λ_1 and λ_2 along R; see Figure 6.

We will argue that this is possible because we can always inductively reduce to the above case, for which we have an explicit formula for the interpolation.

We may assume that R_2 has zero length, because changing λ_2 on R_2 increases it without affecting $\lambda_2 \le \lambda_1$ (in the blue-shaded region), and similarly for R_1.

Now let $\overline{\lambda}_1'$ be the function that differs from $\overline{\lambda}_1$ only on the last edge e_k of R. If $\lambda_1'(\infty_1) = 0$, then $\lambda_1' \ge \lambda_1$, so that we can instead compare λ_1' and λ_2, which is possible by the inductive assumption. If instead $\lambda_1'(\infty_1) < 0$, then $\lambda_1'(\infty_1') = 0$ for some ∞_1' in the green-shaded region (since $\lambda_1' < \lambda_1$ on the complement). In fact, it is easy to see that $\lambda_1'(\infty_2) = 0$, so $\lambda_2(v') = \lambda_1'(v') < \lambda_1(v')$, and we conclude that λ_T interpolates between λ_1 and λ_2 (equivalently λ_1') only along e_k — which we have already solved. □

Remark 3.13 Over a geometric point, there is no obstruction to defining the lifts λ_i of $\overline{\lambda}_i$. On the other hand, they are stable under generisation (edge contraction), because the weight of $\sqsubset_{>0}$ stays the same, and the weight of $\sqsubset_{\ge 0}$ can only go up. So both the λ_i and λ define sections of $\pi_* \overline{M}_C$ over any base S.

Definition 3.14 Let Σ' be the subdivision of $\sigma = \mathrm{trop}(\mathscr{A}_2^{\mathrm{wt}})$ determined by applying universal semistable reduction [56, Theorem 2.4.2] to the morphism $\widetilde{\sqsubset} \to \sigma$.

Let $a\colon \widetilde{\mathscr{A}_2} \to \mathscr{A}_2^{\mathrm{wt}}$ be the corresponding logarithmically étale model.

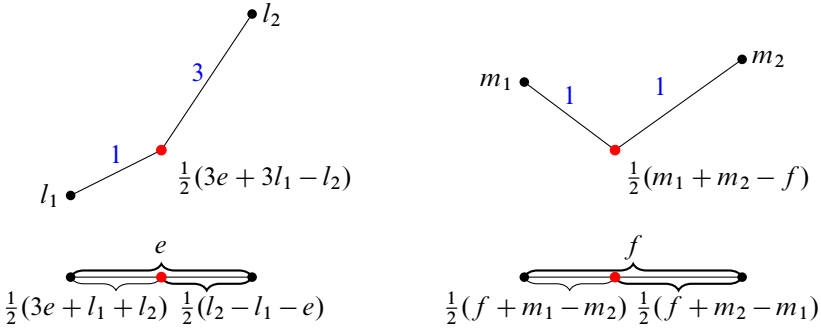


Figure 7: The two situations in which half-integral lengths occur.

It follows from the construction and from [56; 4] that there is a logarithmically smooth curve $\widetilde{\mathscr{C}} \to \widetilde{\mathscr{A}_2}$, which is a partial destabilisation of $a^*\mathscr{C}$, such that $\widetilde{\Sigma}$ is its tropicalisation, $\lambda \in \Gamma(\widetilde{\mathscr{C}}, \overline{M}_{\widetilde{\mathscr{C}}})$ and all of its values at vertices of $\widetilde{\Sigma}$ are comparable.

Remark 3.15 Σ' is neither finer nor coarser than the subdivision Σ previously defined in Definition 3.6. It is not finer, because λ is insensitive to the comparison of values of λ_i below 0. It is not coarser, because cones of $\widetilde{\Sigma}$ do not map surjectively to cones of Σ; see Example 3.39.

Notice also that a is not simply a logarithmic blow-up in general. Indeed, due to the existence of some simplicial but not smooth cones in the subdivision, in order for the lattice map to be surjective it may be necessary to perform a Kummer extension of the base logarithmic structure [17, Section 4]:

$$M_{\mathscr{A}} \subseteq M_{\widetilde{\mathscr{A}}} \subseteq \tfrac{1}{2} M_{\mathscr{A}}^{\mathrm{gp}},$$

as prescribed by the subdivision $\widetilde{\Sigma}$ and the slopes of λ; see Figure 7. This operation should be thought of as a generalised root stack construction, or generalising the presentation of a simplicial affine toric variety as the coarse moduli of $[\mathbb{A}^n/G]$, where G is the finite abelian group encoding the difference between the two lattices at stake. This is different from the genus one case (it has to do with the appearance of slopes other than 1 in λ), but it is nothing new with respect to [56].

Theorem 3.16 *The moduli space $\widetilde{\mathscr{A}_2}$ of aligned admissible covers is a logarithmically smooth stack with locally free logarithmic structure* (*and therefore smooth*) *admitting a log étale morphism to $\mathscr{A}_2^{\mathrm{wt}}$.*

Proof The construction of the minimal logarithmic structure can be traced back to Definition 3.14. The fact that it is locally free can be justified as follows: λ provides us

with a PL map from $\widetilde{\sqsubset}$ to a polyhedral subdivision of $\mathbb{R}_{\geq 0}$. The rank of the logarithmic structure is the number of finite edges in the polyhedral subdivision of $\mathbb{R}_{\geq 0}$. For every level, there is at most one edge of $\widetilde{\sqsubset}$ mapping to it with expansion factor strictly greater than 1; the smoothing parameter corresponding to the said edge of $\widetilde{\sqsubset}$ can be taken as a generator of the minimal logarithmic structure.

Logarithmic smoothness follows from that of $\mathscr{A}_2^{\mathrm{wt}}$, since $a\colon \widetilde{\mathscr{A}_2} \to \mathscr{A}_2$ is logarithmically étale.

Finally, a scheme logarithmically smooth over the trivial log point is smooth if and only if the stalks of its characteristic sheaf are locally free [58, Lemma 5.2]. □

3.3 Combinatorial properties of λ

In the following, we work over the standard log point, so $\sqsubset$ is a graph metrised in $\mathbb{R}$. To an open subcurve $\sqsubset_* \subseteq \sqsubset$ there corresponds a proper subcurve $C_* \subseteq C$; for a closed tropical subcurve, we take its algebrogeometric counterpart to be the closure of the preimage under tropicalisation. By abuse of notation, we will refer to $\sqsubset_*$ or to C_* interchangeably.

Lemma 3.17 *Every component Ω of Δ° has positive genus.*

Proof Restricting (6) to Ω, we find

$$0 \leq \deg(\mathrm{trop}(\psi)^*(D)|_\Omega) = 2p_a(\Omega) - 2 + \sum_{i=1}^{h}(1 + s(\lambda, e_i)),$$

where $e_1, \dots, e_h$ are the edges outgoing from Ω. Since by assumption $s(\lambda, e_i) \leq -1$ for every i, we see that $p_a(\Omega)$ cannot be 0. □

Lemma 3.18 *If $D = \frac{1}{2}(D_1 + D_2)$, then $\lambda(D_1) = \lambda(D_2) = 0$.*

Proof Since by definition every λ_i has D supported on a single vertex, this situation may only occur after interpolating with the zero function.

We first explain how to reduce to the path between D_1 and D_2. Consider an edge e of Δ_T outside this path, and denote by Δ_- and Δ_+ the connected components of $\Delta_T \setminus \{e\}$. Assume that D is entirely supported on Δ_+. Restricting (6) to Δ_-, we find

$$0 = \mathrm{val}(\Delta_-) - 2 + \tfrac{1}{2}\,\#D_B|_{\Delta_-} + \mathrm{div}(\lambda_T|_{\Delta_-}).$$

We notice that $\mathrm{val}(\Delta_-) = 1$ and that $\mathrm{div}(\lambda_T|_{\Delta_-})$ is simply the slope s of λ_T along e, oriented away from Δ_-. Then $s = 1 - \frac{1}{2}\,\#D_B|_{\Delta_-}$. Examining the restriction of (6) to Δ_+, we notice that it is unaltered by contracting e and all the edges in Δ_-.

Assume by contradiction that $\lambda_T(D_1) < \lambda_T(D_2)$. By the previous reduction, we may assume that Δ_T consists of the path between D_1 and D_2 only. By assumption, there must be an edge e on this path along which λ_T has nonzero slope. Contract all the rest. Once again by (6) it is easy to see that the slope can be $\frac{1}{2}$ (resp. 1) from D_1 to D_2 if and only if the distribution of B–legs is $(2,4)$ (resp. $(1,5)$). Now, in order to see the half-points, the slope should be $\frac{1}{2}$; on the other hand, it has to be the slope of one of the λ_i, and we notice that they have integral slope along e if the distribution is $(2,4)$. This is a contradiction, so we conclude that $\lambda_T(D_1) = \lambda_T(D_2)$. The value has to be 0 because half-points can only appear after interpolating with the zero function. □

Corollary 3.19 *In the situation of the previous lemma, let e be any edge of Δ_T between D_1 and D_2. Then each of the connected components of $\mathrm{trop}(\psi)^{-1}(\Delta_T \setminus e^\circ)$ has genus one.*

Proof Once we know that λ_T has slope 0 between D_1 and D_2, this follows from a straightforward application of (6). □

We can now make the following:

Definition 3.20 Let ρ_1 denote the value of λ on the special vertex/vertices.

Lemma 3.21 *The subcurve $\sqsubset_{\geq \rho_1}$ is connected.*

Proof Suppose that there are two connected components Δ_1 and Δ_2. Since $\sqsubset$ is path-connected, we can find an oriented path P from Δ_1 to Δ_2. The slope of λ must be negative at the beginning and positive at the end of P, so $\mathrm{div}(\lambda)|_P \geq 2$. Restricting (6) to P, we find

$$0 = \mathrm{trop}(\psi)^*(D)|_P = 2p_a(P) - 2 + (\mathrm{div}(\lambda)|_P + 2),$$

or $p_a(P) \leq -1$, which is a contradiction. □

Lemma 3.22 *The subcurve $\sqsubset_{\geq \rho_1}$ contains the core.*

Proof Equivalently, $p_a(\sqsubset_{\geq \rho_1}) = 2$. This follows from inspection of Figure 2, but we give a more formal argument; the following has been suggested by the referee. If $\rho_1 = 0$, there is nothing to prove, because $\sqsubset_{\geq 0} = \sqsubset$, so let us assume that $\rho_1 > 0$. Let $\sqsubset_{<\rho_1,\bullet}$

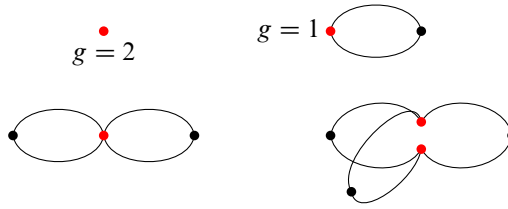


Figure 8: The top level of Δ over $\mathcal{D}_1 \setminus \mathcal{D}_2$.

be one of the h connected components of $\sqsubset \setminus \sqsubset_{\geq \rho_1}$, let $e_1, \dots, e_k$ be the edges between $\sqsubset_{<\rho_1,\bullet}$ and $\sqsubset_{\geq \rho_1}$, with positive slope $s_1, \dots, s_k$. Restricting (6) to $\sqsubset_{<\rho_1,\bullet}$, we find

$$0 = 2p_a(\sqsubset_{<\rho_1,\bullet}) - 2 + k + \sum_{i=1}^{k} s_i,$$

which holds if and only if $p_a(\sqsubset_{<\rho_1,\bullet}) = 0$, $k = 1$ and $s_1 = 1$.

Restricting (6) to $\sqsubset_{\geq \rho_1}$ we find

$$2 = 2p_a(\sqsubset_{\geq \rho_1}) - 2 + h - \sum_{j=i}^{h} 1,$$

which gives the desired result. □

Corollary 3.23 *When λ is constant on $\sqsubset_{\geq \rho_1}$, every positive genus subcurve of $\sqsubset_{\geq \rho_1}$ supports D. The possible shapes of $\sqsubset_{\geq \rho_1}$ are depicted in Figure 8.*

Proof Let F be any positive genus subcurve on which $\lambda|_F \equiv \rho_1$. Restricting (6) to F, we find

$$\deg(\mathrm{trop}(\psi)^*(D)|_F) = 2p_a(F) - 2 + \sum_{i=1}^{k}(1 + s(\lambda, e_i)),$$

where $e_1, \dots, e_k$ are the edges from $\sqsubset \setminus F$ to F. It follows from the assumption and the proof above that the slope of λ along any e_i can only be 0 or -1. The only possibility for $\deg(\mathrm{trop}(\psi)^*(D)|_F)$ to be zero remains that of a genus one curve with all outgoing slopes -1. But then λ would not be constant on $\sqsubset_{\geq \rho_1}$. □

By construction, there is always at least one vertex of positive weight on the boundary of Δ°_T. We make some stronger statements that we will need below.

Lemma 3.24 *Suppose that Δ° is connected of genus two. If there is only one vertex of positive weight on the boundary of Δ°_T, then it is the special one.*

Proof Let us call v_1 the vertex in the statement. Observe that, by Lemma 3.18 and Corollary 3.19, there has to be a single point of T supporting D. We argue by contradiction, showing that if v_1 did not support D we could find a different (interpolation of) admissible function(s) greater than λ; but the latter is defined as the maximum. Extend λ with slope 1 outside Δ. Let v' denote the first vertex of T (not supporting D) encountered on the path from D to v_1, and let ℓ_1 denote the distance from D to v'. Consider the function μ_1 that is 0 from v_1 to v', has slope 1 from v' to D, and has value ℓ_1 on D and further away from v_1; extend it to T by making it locally constant outside the path from v_1 to D. The function $\lambda+\mu_1 > \lambda$ should still be cut at $\lambda+\mu_1(v_1)=0$, unless one (call it v_2) of the vertices of positive weight beyond D has distance $\ell_2 < \ell_1$ from $\partial\Delta_T$. In this case, let ℓ_2 denote the minimal such length. Then consider the function μ_2 that has value ℓ_2 on D and beyond, and decreases with slope 1 towards v_1, until it reaches 0 (which happens before v' since $\ell_2 < \ell_1$). We observe that $\lambda+\mu_2 > \lambda$ should still be cut at $\lambda+\mu_2(v_1) = \lambda+\mu_2(v_2) = 0$, and it has two vertices of positive weight on $\partial\Delta_T$ (on two different sides of D). This is a contradiction unless $l_1 = 0$. □

Example 3.25 Figure 9 illustrates the proof of Lemma 3.24 in one example. In these pictures, the circle delimits the locus where λ is positive. Outside the circle, λ can be extended with slope 1. Conjugate branches are not drawn; we leave it to the reader to complete the picture as necessary.

In the left-most picture — which is the one we start our argument by contradiction with — the values taken by λ are

$$\lambda(v_1)=0, \quad \lambda(\text{core})=\ell, \quad \lambda(v')=\ell-2\ell'=m+2\ell_1, \quad \lambda(w)=m, \quad \lambda(v_2)=-\ell_2,$$

where we have denoted by w the vertex supporting D (it might be arising from the subdivision of an edge) and by m the distance from w to the circle.

Now we start "moving" D towards v_1, as shown in the second picture from the left; if ℓ_2 is long enough, D may reach v' before v_2 reaches $\partial\Delta_T$. The value of λ stays the same on the core, v_1, v' and v_3, but it increases otherwise:

$$\lambda(w)=m+\ell_1, \quad \lambda(v_2)=\ell_1-\ell_2.$$

We can keep pushing D towards v_1 until either D reaches v_1, or v_2 reaches $\partial\Delta_T$, in which case we are no longer in the hypotheses of the lemma. These are the two possibilities illustrated in the right-most pictures of Figure 9.

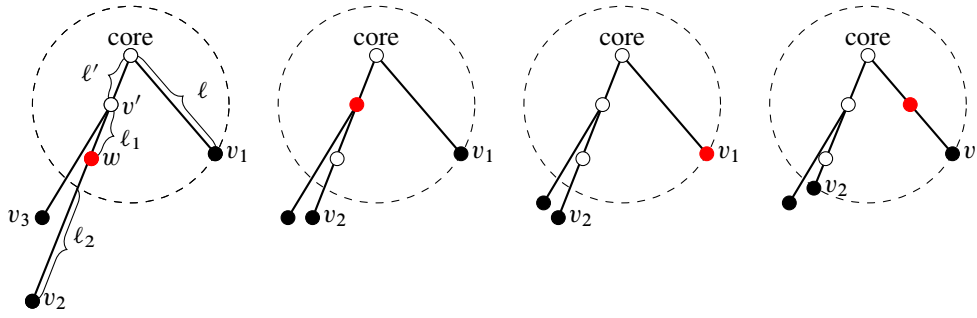


Figure 9: A graphical proof of Lemma 3.24.

The following statements are byproducts of the proof given above:

Corollary 3.26 *Suppose that Δ° is connected of genus two. If D is supported on the boundary of Δ°_T, the corresponding vertex has positive weight.*

Corollary 3.27 *Suppose that Δ° is connected of genus two. If D is supported in Δ°_T, then there is at least one vertex of positive weight on the boundary of Δ°_T on either side of D.*

Finally, we deal with a genus one situation.

Lemma 3.28 *Let Δ°_1 be a connected component of Δ° of genus one. Then there is at least one vertex of positive weight on its boundary.*

Proof Let E_1 denote the genus one core of Δ°_1. There is an admissible function $\overline{\lambda}_i$ (at least one) behaving like the distance from E_1: it is the one with D supported on (any vertex of) the other genus one subcurve E_2 of $\sqsubset$. Then the vertex of positive weight closest to E_1 can be at most as far as E_2, because otherwise Δ° would be connected of genus two. Now λ looks like λ_i in a neighbourhood of E_1. □

For the sake of concreteness, we notice that, in the situation of the above lemma, either E_2 is on the boundary of Δ°_1, supports D and has positive weight (essentially by Corollary 3.26), or the path from E_1 to E_2 crosses the boundary of Δ°_1 at a rational vertex supporting $\frac{1}{2}D$, and proceeds with slope 0 in a neighbourhood.

3.4 The secondary alignment

Lemma 3.29 *The locus V where $w(\Delta)$ is 1, or it is 2 but supported on two non-conjugate vertices of $\sqsubset$, is closed in $\widetilde{\mathscr{A}}_2$.*

Proof The statement is of a topological nature. The loci where the combinatorial type of $\sqsubset$ is constant form a constructible stratification of $\widetilde{\mathscr{A}}_2$. By Noetherianity, it is enough to check that the described locus is closed under specialisation for any trait, ie if we have an edge contraction $\sqsubset_s \to \sqsubset_\eta$ such that the latter is in V, then the former must be as well. This is easily checked. □

Definition 3.30 Let $\widetilde{\mathscr{A}}_2^\circ$ be the open substack of $\widetilde{\mathscr{A}}_2$ defined as the complement of the locus V described in the previous lemma.

Definition 3.31 Let Σ'' be the subdivision of $\mathrm{trop}(\widetilde{\mathscr{A}}_2^\circ)$ defined by

- formally contracting Δ to a point z;
- aligning the rest of the curve with respect to the distance from z.

Let $(\widetilde{\mathscr{A}}_2^\circ)^{\mathrm{pre}}$ denote the corresponding logarithmic blow-up.

On $\widetilde{\mathscr{A}}_2^\circ$ the weight of Δ is at least 2. If it is 2, it is supported on a single vertex v_1 of Δ_T. In this case, by Lemma 3.24, λ has D supported on v_1. By relabelling, assume that $\overline{\lambda}_1$ was the admissible function (Definition 3.2) determined by having v_1 as the special vertex. By Lemma 3.22, the slope of $\overline{\lambda}_1$ is always 1 below the level of v_1; hence, the alignment of the previous definition makes sense. We extend λ to reach the next vertex of positive weight.

Definition 3.32 Let S be a geometric point of $(\widetilde{\mathscr{A}}_2^\circ)^{\mathrm{pre}}$, and suppose that $w(\Delta) = 2$. Let $\lambda_1' \in \Gamma(S, \pi_* \overline{M}_C)$ be the lift of $\overline{\lambda}_1$ determined by $w(\sqsubset_{>0}) \le 2$ and $w(\sqsubset_{\ge 0}) \ge 3$. Let

$$\tilde{\lambda} = \max\{0, \lambda_1'\}.$$

Lemma 3.33 *If a specialisation $\eta \rightsquigarrow s$ in $(\widetilde{\mathscr{A}}_2^\circ)^{\mathrm{pre}}$ induces an edge contraction $\sqsubset_s \to \sqsubset_\eta$ such that $w(\sqsubset_{s,\ge 0}) = 2$ but $w(\sqsubset_{\eta,\ge 0}) \ge 3$, then $\tilde{\lambda}$ generises to λ.*

Proof By Lemma 3.24, on the cone of $\mathrm{trop}(\widetilde{\mathscr{A}}_2^\circ)$ corresponding to the point s, the function λ was a lift of $\overline{\lambda}_1$, with a cutoff at the first vertex v_1 of positive weight. Extending it to the second vertex v_2 of positive weight makes sense thanks to Definition 3.31. When an edge contraction makes $w(\Delta)$ grow, it means in particular that v_2 and v_1 become at the same height, so $\tilde{\lambda}$ coincides with the original λ. □

This shows that $\tilde{\lambda}$ extends λ as a well-defined (continuous) PL function on a subdivision $\tilde{\tilde{\Sigma}}$ of $\tilde{\Sigma}$ for any S–point of $(\tilde{\mathscr{A}}_2^\circ)^{\mathrm{pre}}$, which we think of as an extension of λ on certain cones of $\mathrm{trop}((\tilde{\mathscr{A}}_2^\circ)^{\mathrm{pre}})$.

Definition 3.34 Let $\tilde{\tilde{\mathscr{A}}}_2$ denote the logarithmic blow-up of $\tilde{\mathscr{A}}_2^\circ$ induced by applying universal semistable reduction to $\tilde{\tilde{\Sigma}} \to \mathrm{trop}((\tilde{\mathscr{A}}_2^\circ)^{\mathrm{pre}})$.

Remark 3.35 The spaces $(\tilde{\mathscr{A}}_2^\circ)^{\mathrm{pre}}$ and $\tilde{\tilde{\mathscr{A}}}_2$ are analogous to, respectively, the space of radially aligned and the space of centrally aligned logarithmic curves in [63].

Similarly to the previous section, we have the following result:

Theorem 3.36 *The moduli space $\tilde{\tilde{\mathscr{A}}}_2$ is a logarithmically smooth stack with locally free logarithmic structure* (*and therefore smooth*). *There is a logarithmically smooth curve $\tilde{\tilde{\mathscr{C}}} \to \tilde{\tilde{\mathscr{A}}}_2$, which is a partial destabilisation of $\tilde{a}^*\mathscr{C}$, such that $\tilde{\tilde{\Sigma}}$ is its tropicalisation, $\tilde{\lambda} \in \Gamma(\tilde{\tilde{\mathscr{C}}}, \overline{M}_{\tilde{\tilde{\mathscr{C}}}})$ and all of its values at vertices of $\tilde{\tilde{\Sigma}}$ are comparable.*

3.5 Examples of subdivisions

In this section we collect a few examples to show what the subdivisions Σ and Σ' of a cone $\sigma \in \mathrm{trop}(\mathscr{A}_2^{\mathrm{wt}})$, the combinatorics of λ on the various cones, and the associated singularities in $\bar{\mathscr{C}}$ look like. The construction of $\bar{\mathscr{C}}$ will be carried out in the next section.

Example 3.37 See Figure 10. The stabilisation of C has smooth core of weight zero, and two rational tails of high weight, one of which is attached to a Weierstrass point (compare with Example 3.5). In this case, the subdivisions Σ and Σ' coincide. Notice that there is a simplicial nonsmooth cone σ_3; correspondingly, half edge-lengths occur in a Kummer extension of $M_{\tilde{A}}$.

Example 3.38 See Figure 11. The stabilisation of C has a smooth core of weight zero, and three rational tails attached to general points. Aligning with respect to all the admissible functions produces the nonsimplicial subdivision (a); the subdivisions (b) and (c) are the coarsening in case of high-weight, respectively weight two, tails. We also represent λ_T and the associated singularity on some cones of the subdivision; the other ones can be derived by symmetry. In case (c) the singularity over σ_2 is replaced by a (sprouted) ribbon.

Example 3.39 See Figure 12. The core of C consists of two elliptic curves of weight zero, meeting in a node; each elliptic curve is attached to a high-degree rational tail.

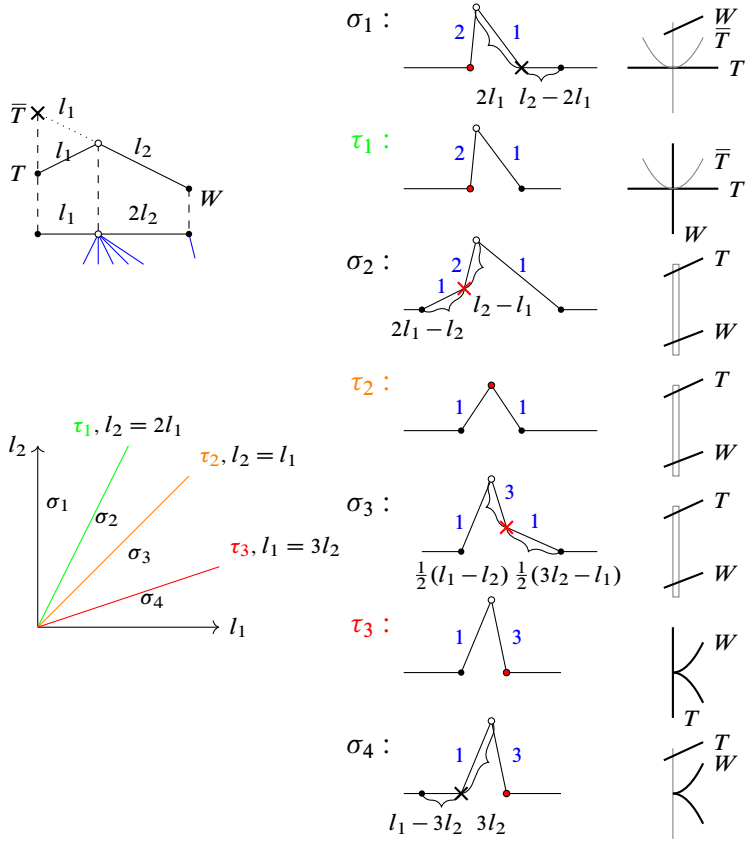


Figure 10: Tropical admissible cover (top left); the subdivision of $\mathrm{trop}(\mathscr{A}_2^{\mathrm{wt}})$ (bottom left); λ_T and the Gorenstein singularities (right).

Notice that the central cone σ_2 is not smooth; again, we need a Kummer extension (half-lengths). We also remark that the hyperplane $l_1 + l_2 = m$ does not come from the alignment, but from the procedure of Definition 3.14.

3.6 The primary construction

We construct a universal morphism $\widetilde{\mathscr{C}} \to \overline{\mathscr{C}}$ over $\widetilde{\mathscr{A}_2}$, where $\overline{\mathscr{C}}$ is a family of Gorenstein singularities (both isolated and ribbons) with core of positive weight. We do so in two steps: first, a contraction informed by λ, producing a possibly non-Gorenstein curve. The image in $\widetilde{\mathscr{A}_2}$ of the non-Gorenstein locus is contained in a divisor, which we name $\mathcal{D}_1$ below. We complete the construction of $\overline{\mathscr{C}}$ by gluing in a portion of ψ over this locus, thus producing a nonreduced structure along the fibres.

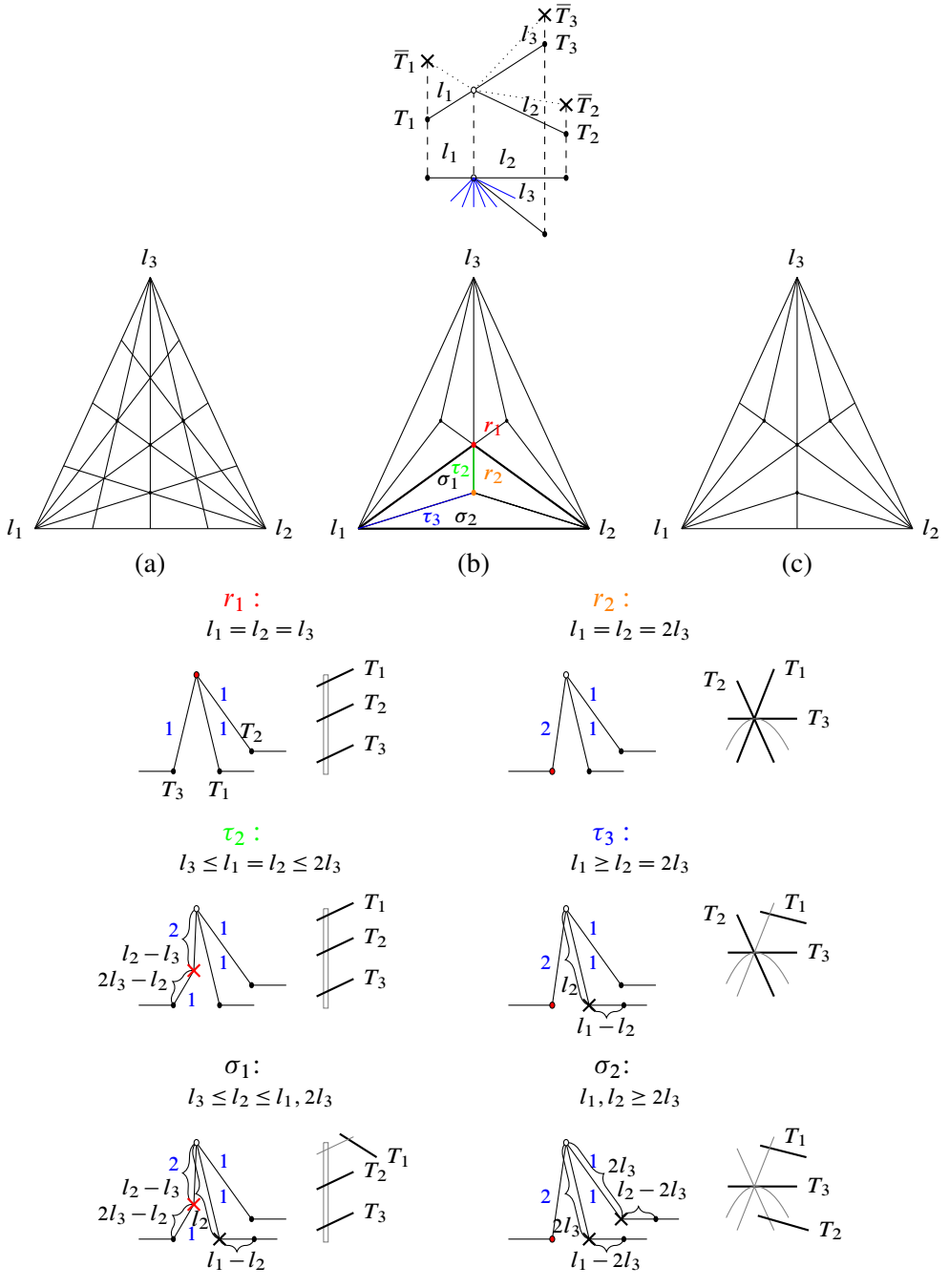


Figure 11: (a) Subdivision Σ (nonsimplicial). (b) Subdivision Σ' for high-degree tails. (c) Subdivision Σ' for degree two tails.

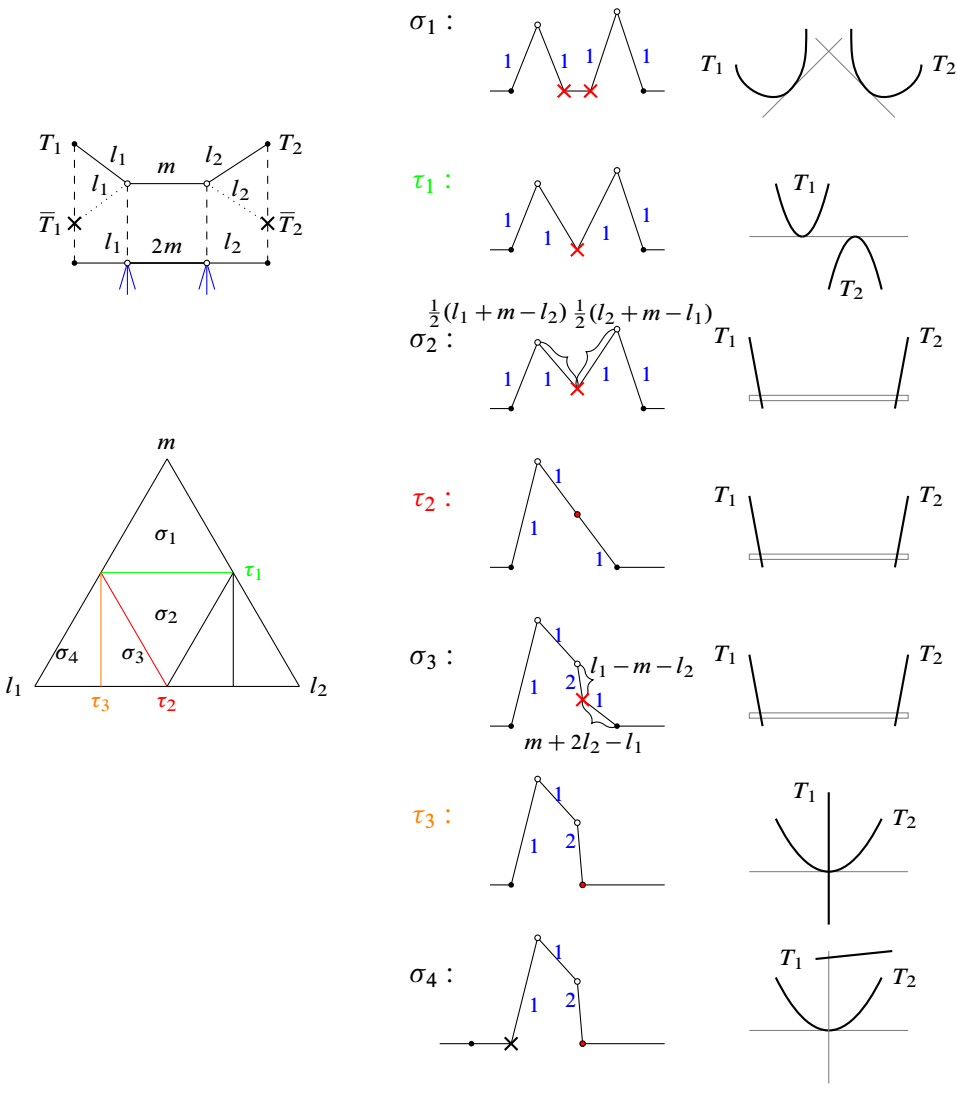


Figure 12: Tropical admissible cover (top left); the subdivision of $\mathrm{trop}(\mathscr{A}_2^{\mathrm{wt}})$ (bottom left); λ_T and the Gorenstein singularities (right).

Definition 3.40 Let $\rho_{\max}$ denote the maximum value attained by λ on $\sqsubset$, and let $\mathcal{D} \subseteq \widetilde{\mathscr{A}_2}$ be the Cartier divisor determined by the ideal sheaf $\mathscr{O}_{\widetilde{\mathscr{A}}}(-\rho_{\max}) \hookrightarrow \mathscr{O}_{\widetilde{\mathscr{A}}}$.

Definition 3.41 Recall from Definition 3.20 that ρ_1 denotes the value attained by λ at the vertex(ices) supporting $K_\Delta + \mathrm{div}(\lambda)$. Let $\mathcal{D}_1 \subseteq \widetilde{\mathscr{A}_2}$ be the Cartier divisor cut by the ideal sheaf $\mathscr{O}_{\widetilde{\mathscr{A}}}(-\rho_1) \hookrightarrow \mathscr{O}_{\widetilde{\mathscr{A}}}$.

Note that $\rho_1 \le \rho_{\max}$ implies $\mathcal{D}_1 \subseteq \mathcal{D}$. By Lemma 3.22, the subcurve $\sqsubset_{\ge\rho_1}$ always contains the core, so, in particular, $p_a(\Delta^\circ) = 2$ over $\mathcal{D}_1$. Ribbons will appear in $\bar{\mathscr{C}}$ precisely over $\mathcal{D}_1$.

Definition 3.42 Let $\mathcal{D}_2$ be cut by $\rho_{\max} - \rho_1$, so that we have an exact sequence

$$0 \to \mathscr{O}_{\mathcal{D}_2}(-\rho_1) \to \mathscr{O}_{\mathcal{D}} \to \mathscr{O}_{\mathcal{D}_1} \to 0.$$

The Gorenstein curve $\bar{\mathscr{C}}$ will have only isolated singularities over $\mathcal{D}_2$. They may have genus one or two according to $p_a(\Delta^\circ)$.

Definition 3.43 Let $\mathcal{Z}$ be the Cartier divisor on $\widetilde{\mathscr{C}}$, supported over $\mathcal{D}$, which is determined by the inclusion $\mathscr{O}_{\widetilde{\mathscr{C}}}(-\lambda) \hookrightarrow \mathscr{O}_{\widetilde{\mathscr{C}}}$.

Remark 3.44 Locally, we construct a line bundle on $\widetilde{\mathscr{C}}$ that is trivial on $\mathcal{Z}$ except when it contains the special component, and relatively ample elsewhere, as follows. We pick smooth disjoint sections $p_1, \dots, p_{2d}$ of $\widetilde{\mathscr{C}}$ according to the weight function, namely so that $\#\{i \mid p_i \in \widetilde{C}_v\} = 2w(v)$. These sections exist only locally, but we will show that our construction does not depend on this choice; therefore, it glues on the whole of $\widetilde{\mathscr{A}}_2$.

Notation Let $\sigma : \widetilde{\mathscr{C}} \to \widetilde{\mathscr{C}}$ be the hyperelliptic involution. Let $\boldsymbol{p}_i$ denote the multisection $p_i + \sigma(p_i)$, and $\boldsymbol{p} = \sum \boldsymbol{p}_i$. We denote by $\mathscr{L}$ the line bundle $\omega_{\widetilde{\mathscr{C}}/\widetilde{\mathscr{A}}}(\boldsymbol{p})(\lambda)$ on $\widetilde{\mathscr{C}}$.

Theorem 3.45 *The line bundle $\mathscr{L}$ is π–semiample. In the diagram*

$$\begin{array}{ccc} \widetilde{\mathscr{C}} & \xrightarrow{\ \phi\ } & \mathscr{C}' := \underline{\mathrm{Proj}}_{\widetilde{\mathscr{A}}_2}\big(\pi_* \bigoplus_{k\ge 0} \mathscr{L}^{\otimes k}\big) \\ & {\scriptstyle\pi}\searrow \quad \swarrow{\scriptstyle\pi'} & \\ & \widetilde{\mathscr{A}}_2 & \end{array}$$

the morphism π' is a flat family of reduced, projective, Cohen–Macaulay curves of arithmetic genus two, with Gorenstein fibres outside $\mathcal{D}_1$. Moreover, we can perform a parallel contraction $\mathscr{T} \to \mathscr{T}'$ so that $\psi' : \mathscr{C}' \to \mathscr{T}'$ remains finite. Neither $\mathscr{C}'$ nor $\mathscr{T}'$ depend on the choice of sections respecting the weight function, as per Remark 3.44.

Remark 3.46 $\mathscr{L}$ is symmetric under the hyperelliptic involution, ie it is pulled back from the target of the admissible cover. Any weight zero branch appearing as the conjugate of a rational tail (or bridge) of positive weight is therefore not contracted under ϕ. This ensures that $\mathscr{C}'$ remains "hyperelliptic", ie it admits a double cover to a rational curve, possibly with ordinary m–tuple points. As far as the factorisation of

the stable map is concerned, all those unstable components of weight zero (different from the special branches) do not interfere. We could have just as well carried out the construction without symmetrising the line bundle.

Theorem 3.47 *Let $\mathcal{Z}'$ denote the image of $\mathcal{Z}$ under ϕ. Over $\mathcal{D}_1$, it is a flat family of Gorenstein curves of arithmetic genus two, and the image of $\mathcal{Z}'$ under ψ' is a rational curve $\mathscr{T}'_{\mathcal{Z}}$. Let $\bar{\mathscr{C}}$ be obtained as the pushout of*

$$\begin{array}{ccc} \mathcal{Z}'_{\mathcal{D}_1} & \hookrightarrow & \mathscr{C}' \\ \downarrow{\scriptstyle \psi'} & & \\ \mathscr{T}'_{\mathcal{Z}} & & \end{array}$$

Then $\bar{\mathscr{C}} \to \widetilde{\mathscr{A}_2}$ is a flat family of projective, Gorenstein curves of arithmetic genus two. Moreover, the weight of any subcurve of positive genus is at least one.

The proof of the two theorems above occupies the next two sections.

3.7 First step: the contraction

Lemma 3.48 *For $k \geq 2$, $R^1\pi_*\mathscr{L}^{\otimes k}$ is supported along $\mathcal{D}_2$. Moreover, it admits a two-term resolution that remains such after pullback to a sufficiently generic base; in particular, if $f : S \to \widetilde{\mathscr{A}_2}$ is a morphism such that $\mathscr{O}_S(-\rho_{\max}) \to \mathscr{O}_S$ remains injective, then*

$$f^*\pi_*(\mathscr{L}^{\otimes k}) \to (\pi_S)_*(\mathscr{L}_S^{\otimes k})$$

is an isomorphism.

Proof There are three things to show:

Support Since $\mathscr{L}^{\otimes k}$ is flat on the base, by cohomology and base change it is enough to show the vanishing of $H^1(\widetilde{\mathscr{C}}_s, \mathscr{L}^{\otimes k})$ for $s \in \widetilde{\mathscr{A}_2} \setminus \mathcal{D}_2$.

By weighted stability, $\mathscr{L}$ is relatively ample outside of $\mathcal{Z}$, and in particular over $\widetilde{\mathscr{A}_2} \setminus \mathcal{D}$.

For $s \in \mathcal{D} \setminus \mathcal{D}_2$ we have $0 \neq \rho_1 = \rho_{\max}$. We note that $\mathscr{L} \geq 0$ and, by Corollary 3.23, it has degree two on every positive genus subcurve of $\widetilde{\mathscr{C}}_s$. Then it is clear for degree reasons that $h^1(\mathscr{L}^{\otimes k}) = 0$ for $k \geq 2$.

Resolution Note first that the rank of $R^1\pi_*\mathscr{L}^{\otimes k}$ is not constant along $\mathcal{D}_2$.

Indeed, $\mathcal{D}_2$ has two types of irreducible components:

- $\mathcal{D}_{2,1}$, where *generically* Δ° has genus one, and $R^1\pi_*\mathscr{L}^{\otimes k}$ has rank one;
- $\mathcal{D}_{2,2}$, where Δ° has genus two, and $R^1\pi_*\mathscr{L}^{\otimes k}$ has rank two.

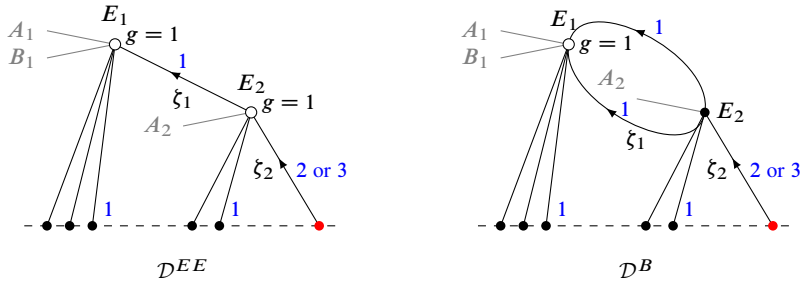


Figure 13: The generic points of $\mathcal{D}_{2,1} \cap \mathcal{D}_{2,2}$.

Away from their intersection, it is easy to find the desired *local* resolution

$$0 \to \mathscr{O}_U^{\oplus 2} \xrightarrow{\begin{pmatrix} e^{\rho_{\max}} & 0 \\ 0 & e^{\rho_{\max}} \end{pmatrix}} \mathscr{O}_U^{\oplus 2} \to \mathscr{O}_{\mathcal{D}_{2,2}\cap U}^{\oplus 2} \to 0$$

on $U \subseteq \widetilde{\mathscr{A}_2}$ a neighbourhood of a generic point of $\mathcal{D}_{2,2}$ (assuming $\rho_1 = 0$), where $e^{\rho_{\max}}$ denotes a local equation for $\mathcal{D}$. Similarly,

$$0 \to \mathscr{O}_U \xrightarrow{e^{\rho_{\max}}} \mathscr{O}_U \to \mathscr{O}_{\mathcal{D}_{2,1}\cap U} \to 0$$

around a generic point in $\mathcal{D}_{2,1}$.

The intersection of $\mathcal{D}_{2,1}$ with $\mathcal{D}_{2,2}$ has two types of irreducible components; see Figure 13. In order to obtain a local resolution, we adapt an argument of [40]. We sketch it here for the reader's benefit.

Locally on the base, there is a section p_{2d+1} such that $\omega_C(\lambda) = \mathscr{O}_C(\boldsymbol{p}_{2d+1})$, where $\boldsymbol{p}_{2d+1} = p_{2d+1} + \sigma(p_{2d+1})$, and therefore $\mathscr{L} = \mathscr{O}_C\big(\sum_{i=1}^{2d+1} \boldsymbol{p}_i\big)$.

Again locally on the base, we may find disjoint generic sections A_1, A_2 and B such that A_1 and B pass through E_1, and A_2 passes through E_2. Then $\mathscr{L}(A_1 + A_2 - B)$ has vanishing h^1 on fibres, and therefore, by cohomology and base change, $\pi_*\mathscr{L}(A_1 + A_2 - B)$ is a vector bundle. Locally, we can write $\pi_*\mathscr{L} \simeq \mathscr{O}_U \oplus \pi_*\mathscr{L}(-B)$, and the second factor is the kernel of the evaluation map

$$\pi_*\mathscr{L}(A_1 + A_2 - B) \to \pi_*(\mathscr{O}_{A_1}(A_1) \oplus \mathscr{O}_{A_2}(A_2)).$$

The evaluation map can be studied fibrewise, since both sheaves in question have vanishing h^1; moreover, the source can be decomposed into

$$\pi_*\mathscr{L}(A_1 + A_2 - B) \cong \bigoplus_{i=1}^{2d+2} \pi_*\mathscr{O}(\boldsymbol{p}_i + A_1 + A_2 - B),$$

and the evaluation map can be studied componentwise:

$$\mathrm{ev}_{i,j}\colon \pi_*\mathscr{O}(\boldsymbol{p}_i + A_1 + A_2 - B) \to \pi_*(\mathscr{O}_{A_j}(A_j)), \quad i = 1,\dots,2d+2,\ j = 1,2.$$

The cokernel of the latter is $H^1(C, \mathscr{O}_C(\boldsymbol{p}_i + A_{2-j} - B))$. See [38, Section 4.2].

It follows from an argument similar to that of [38, Proposition 4.13] that the latter is nonzero precisely when p_i stays away from E_j; therefore, in some local trivialisation of the line bundles involved,

$$\mathrm{ev}_{i,j} = c_{i,j} \prod_{q\in[p_i,A_j]} \zeta_q,$$

where $c_{i,j} \in \mathscr{O}_U^*$, we denote by $[p_i, A_j]$ the set of nodes separating p_i from A_j, and $\zeta_q \in \mathscr{O}_U$ is the smoothing parameter of the node q. Thanks to the alignment,

- the smoothing parameter ζ_2 of the node separating E_2 from the component supporting D divides all the expressions of the form $\mathrm{ev}_{i,2}$ for $i = 1,\dots,2d+2$;
- if ζ_1 denotes the smoothing parameter of the node separating E_1 from E_2, the product $\zeta_1\zeta_2$ divides all the expressions of the form $\mathrm{ev}_{i,1}$ for $i = 1,\dots,2d+2$.

(Notice that ζ_1 and ζ_2 should be replaced by some products of smoothing parameters when the curve degenerates.) We can therefore use the column of the evaluation matrix associated to a marking p_i (up to relabelling, $i = 1$) on the component supporting D in order to put the matrix in triangular form. In order to diagonalise it, we need more refined information, which we borrow from [40, Sections 2.6–2.7] and we restate here in streamlined form:

Proposition (Hu, Li and Niu) *The determinant of the matrix*

$$\begin{pmatrix} c_{1,1} & c_{j,1} \\ c_{1,2} & c_{j,2} \end{pmatrix}$$

is invertible when the markings p_1 and p_j are not conjugate under ψ.

Since the component supporting D contains at least three markings (p_1, p_{2d+1} and p_{2d+2}), we can find two nonconjugate ones.

Summing up, the evaluation matrix can be put in the form

$$\begin{pmatrix} \zeta_1 & 0 \\ 0 & \zeta_1\zeta_2 \end{pmatrix}$$

(the remaining columns are zero). Noticing that $\mathcal{D}_{2,i} = \{\zeta_i = 0\}$ for $i = 1, 2$, around the given point, we have thus found the desired local resolution of $R^1\pi_*\mathscr{L}^k$. In particular, it follows that $\pi_*\mathscr{L}^{\otimes\geq 2}$ is a vector bundle on $\widetilde{\mathscr{A}_2}$.

Base change See the proof of [63, Lemma 3.7.5.3]. Note that the evaluation matrix above is an explicit instance of the Grothendieck–Mumford complex for cohomology and base change, and $\zeta_1\zeta_2 = e^{\rho_{\max}}$ is a local equation for $\mathcal{D}$. □

Proof of Theorem 3.45 This is analogous to [67, Lemma 2.13; 63, Proposition 3.7.6.1]. We recap for the reader's convenience.

Flatness This is equivalent to $\pi_*\mathscr{L}^{\otimes k}$ being a vector bundle for $k \gg 0$ [16, Tag 0D4D].

Basepoint-freeness That is, existence of the morphism ϕ. This is clear outside $\mathcal{D}$, where $\mathscr{L}^{\otimes k}$ is π–ample. Even on $\mathcal{D}$, we have that $\mathscr{L}^{\otimes k}$ is π–ample outside of $\mathcal{Z}$. Moreover, from the short exact sequence

$$0 \to \mathscr{L}(-\lambda)^{\otimes k} \to \mathscr{L}^{\otimes k} \to \mathscr{L}|_{k\mathcal{Z}}^{\otimes k} \to 0$$

and the vanishing of $R^1\pi_*\mathscr{L}(-\lambda)^{\otimes k}$ (stability), it is enough to show that, for any $x \in \mathcal{Z}$, there exists a section of $\pi_*\mathscr{L}|_{k\mathcal{Z}}^{\otimes k}$ around $\pi(x)$ which does not vanish in x. By definition of λ, the line bundle $\mathscr{L}|_{\mathcal{Z}}$ is the pullback along ψ of a line bundle of degree one and nonnegative multidegree on $\mathscr{T}_{\mathcal{Z}}$; the latter has enough sections.

Properties of the fibres This can be studied after base change to a generic trait Υ. We may assume that the generic point corresponds to a smooth curve and the closed point maps to $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$. Thus,

$$\phi_\Upsilon : C := \mathscr{C}_\Upsilon \to \mathscr{C}'_\Upsilon =: C'$$

is a birational contraction satisfying $\phi_*\mathscr{O}_C = \mathscr{O}_{C'}$. It follows that C' is a normal surface. In particular, the central fibre is $S1$; it is also generically reduced, being birational to that of C; thus, it is reduced (and Cohen–Macaulay).

Finally, we want to argue that the fibres are Gorenstein outside $\mathcal{D}_1$. We may assume that the special point of Υ maps to $\mathcal{D}_2 \setminus \mathcal{D}_1$. Then $\mathscr{L}$ is trivial along Z, which is therefore contracted to a codimension two locus Z' of C'. Outside of Z, ϕ restricts to an isomorphism. The equality of line bundles

$$\mathscr{O}_{C'}(1)\Big(-\sum \boldsymbol{p}_i\Big)\Big|_{C'\setminus Z'} = \omega_{C/\Upsilon}|_{C\setminus Z} = \omega_{C'/\Upsilon}|_{C'\setminus Z'},$$

together with the fact that $\omega_{C'/\Upsilon}$ is an $S2$ sheaf, shows that the latter coincides with the line bundle $\mathscr{O}_{C'}(1)\big(-\sum \boldsymbol{p}_i\big)$ on the whole of C' (Hartogs' theorem). Thus, C' is Gorenstein over Υ.

Compatible contraction of $\mathscr{T}$ This can be performed using the line bundle

$$\omega_{\mathscr{T}}(D_{\boldsymbol{B}})(\lambda_T)(\psi(\boldsymbol{p})),$$

whose pullback to $\widetilde{\mathscr{C}}$ is $\mathscr{L}$.

Well-posedness The construction of $\mathscr{C}'$ and $\mathscr{T}'$ is independent of the choice of markings respecting the weight function by birational rigidity [28, Lemma 1.15]. □

3.8 Second step: the pushout

Ribbons can be used to interpolate between isolated singularities over $\mathcal{D}_1$ — this is completely natural from the point of view of piecewise-linear functions on the tropical side, and serves as a correction of the failure of $\mathscr{C}'$ at being Gorenstein.

Recall that $\mathcal{Z}'$ was defined as the image of $\mathcal{Z}$ under ϕ. First, we prove that the definition of $\mathcal{Z}'$ commutes with base change to a generic trait; then, we will show that the pushout construction commutes with such a base change, and thus we may reduce to the case of surfaces in order to study the singularities of the fibres.

Proposition 3.49 *Let $\mathcal{Z}'$ be the subscheme of $\mathscr{C}'$ defined by the ideal sheaf $\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)$ and supported on $\mathcal{D}$. Then*:

(1) $R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda) = 0$; *in particular*, $\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda) = \mathrm{Fitt}(\phi_*\mathscr{O}_{\mathcal{Z}})$.

(2) *For every generic trait $\Upsilon \xrightarrow{\iota} \widetilde{\mathcal{A}}$ with generic point mapping to the smooth locus, the definition of $\mathcal{Z}'$ commutes with base change, ie $\iota^*\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda) = \phi_{\Upsilon *}\mathscr{O}_{\mathscr{C}_\Upsilon}(-\lambda)$.*

The analogous statements about $\mathscr{T}'_{\mathcal{Z}}$ hold as well.

Proof The discussion has been somewhat inspired by [71, Section 1].

(1) Let C_s be a fibre on which ϕ_s is not an isomorphism; in particular, $\lambda_s \neq 0$. Working locally on the base, we can choose smooth and disjoint sections $p_1, \dots, p_d$ of $\mathscr{C}$ respecting the weight function. It is enough to prove that $R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\,\boldsymbol{p}) = 0$ for $k \gg 0$. Indeed, once we know the latter is vanishing, since ϕ is an isomorphism around $\boldsymbol{p}$, we have that $\mathscr{O}_{\mathscr{C}}(k\,\boldsymbol{p}) = \phi^*\mathscr{O}_{\mathscr{C}'}(k\,\boldsymbol{p}')$ and thus

$$\begin{aligned} 0 &= R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\,\boldsymbol{p}) = H^1(R\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\,\boldsymbol{p})) \\ &= H^1(R\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda) \otimes \phi^*\mathscr{O}_{\mathscr{C}'}(k\,\boldsymbol{p}')) = R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda) \otimes \mathscr{O}_{\mathscr{C}'}(k\,\boldsymbol{p}'), \end{aligned}$$

implying the desired vanishing.

From the spectral sequence computing $R\pi'_* \circ R\phi_*$, the five-term exact sequence

$$0 \to R^1\pi'_*\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p}) \to R^1\pi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p}) \to \pi'_*R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p})$$

ends there, because the next term would involve an $R^2\pi'_*(-)$, which vanishes as the fibre dimension is bounded by one.

Notice that if $R^1\pi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p})\otimes k(s)=0$ for those s where ϕ_s is not an isomorphism, so is $\pi'_*R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p})\otimes k(s)$.

On the other hand, $R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p})$ is a coherent sheaf supported on the closed locus V' in $\mathscr{C}'$ over which the fibre of ϕ has positive dimension. The latter is finite over $\tilde{\mathcal{A}}$, and thus $\pi'_*R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p})\otimes k(s)=0$ implies $R^1\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p})\otimes k(x)=0$ for each $x\in V'_s$.

Since $R^1\pi_*\mathscr{O}_{\mathscr{C}}(-\lambda)(k\boldsymbol{p})$ satisfies cohomology and base change, the vanishing can be checked after restricting to a fibre. Let Z_s denote the support of λ_s, C_i the trees rooted at the vertices of $\partial\Delta$, and e_i the first edges encountered in Δ°. Taking the normalisation of C_s at the nodes corresponding to the edges e_i, we get

$$0 \to \mathscr{O}_{C_s}(-\lambda)(k\boldsymbol{p}) \to \mathscr{O}_{Z_s}(-\lambda)\oplus\bigoplus_i \mathscr{O}_{C_i}(-\lambda)(k\boldsymbol{p}) \to \bigoplus_i \mathbb{C}_{e_i} \to 0.$$

The evaluation map on the nodes e_i is clearly surjective at the level of H^0, as the line bundle restricted to C_i is very ample for k big enough, and the desired vanishing follows from that of $H^1(\mathscr{O}_{Z_s}(-\lambda_s))$, which can be argued by the definition of λ and Serre duality.

(2) We know by Lemma 3.48 that the construction of $\mathscr{C}'$ commutes with the given base change, so the following diagram is Cartesian:

$$\begin{array}{ccc} \mathscr{C}_\Upsilon & \xrightarrow{\phi_\Upsilon} & \mathscr{C}'_\Upsilon \\ \downarrow{\scriptstyle\iota} & \square & \downarrow{\scriptstyle\iota'} \\ \mathscr{C} & \xrightarrow{\phi} & \mathscr{C}' \end{array}$$

Furthermore, since the source and target are smooth, $\Upsilon \hookrightarrow \tilde{\mathcal{A}}$ is an LCI morphism, and so are ι and ι'; it thus follows from [49, Corollary 2.27] that

$$L\iota'^*R\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda) = R\phi_{\Upsilon *}L\iota^*\mathscr{O}_{\mathscr{C}}(-\lambda).$$

On the other hand, the higher pushforward vanishes by the previous point, so

$$R\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda) = \phi_*\mathscr{O}_{\mathscr{C}}(-\lambda),$$

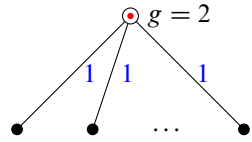


Figure 14: λ at the generic point of $\mathcal{D}_1$.

and $L\iota^*\mathscr{O}_{\mathscr{C}}(-\lambda) = \mathscr{O}_{\mathscr{C}_\Upsilon}(-\lambda)$ because it is a line bundle; hence, the derived statement is equivalent to what we want. □

Proposition 3.50 *The restriction of $\mathcal{Z}'$ to $\mathcal{D}_1$ is a flat family of Gorenstein curves of genus two. Similarly, $\mathcal{T}'_{\mathcal{Z}}$ is a flat family of Gorenstein curves of genus zero (ie at worst nodal).*

Notice that $\mathscr{C}'$ is not always Gorenstein over $\mathcal{D}_1$; in particular, it can be the decomposable union of $\mathcal{Z}'$ with some lines. For the proof we need the following:

Lemma 3.51 *$\mathcal{D}_1$ is a reduced divisor.*

Proof $\mathcal{D}_1$ is a Cartier divisor in a smooth ambient space, so it is enough to check that it is generically reduced.

The generic point of $\mathcal{D}_1$ looks like in Figure 14.

Therefore, a generic trait with uniformiser t will intersect $\mathcal{D}_1$ in (t). □

Proof of Proposition 3.50 We may change the base to a generic trait Υ with closed point mapping to the given point of $\mathcal{D}_1$, ie

$$\begin{array}{ccc} C & \xrightarrow{\ \phi\ } & C' \\ & {\scriptstyle\pi}\searrow \quad \swarrow{\scriptstyle\pi'} & \\ & \Upsilon & \end{array}$$

so C is a smooth surface, ϕ is a birational contraction, which is an isomorphism outside the divisor Z defined by $\mathscr{I}_Z = \mathscr{O}_C(-\lambda)$. Notice that there is a subcurve of Z on which the line bundle defining the contraction is ample; therefore, $Z' \subseteq C'$ has pure codimension one. We want to show that $\omega_{Z'}$ is a line bundle and that $\chi(\omega_{Z'}) = 1$ (or equivalently $p_a(Z') = 2$).

Recall that, by Grothendieck duality,

$$\omega_{Z'} = \mathscr{E}xt^1_{C'}(\mathscr{O}_{Z'}, \omega_{C'}) = \mathscr{H}om_C(\mathscr{I}_{Z'}, \omega_{C'})|_{Z'}.$$

By adjunction for the Cartier divisor $Z \subseteq C$, there is a short exact sequence

$$0 \to \omega_C \to \omega_C(\lambda) \to \omega_Z \to 0, \tag{7}$$

which stays exact after pushforward along ϕ by Grauert–Riemenschneider vanishing [46, Corollary 2.68]. By applying $\mathscr{H}om_{\mathscr{O}_{C'}}(-, \omega_{C'})$ to the exact sequence

$$0 \to \phi_* \mathscr{O}_C(-\lambda) \to \mathscr{O}_{C'} \to \mathscr{O}_{Z'} \to 0,$$

we obtain the bottom row of the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \phi_*\omega_C & \longrightarrow & \phi_*\omega_C(\lambda) & \longrightarrow & \phi_*\omega_Z & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \omega_{C'} & \longrightarrow & \mathscr{H}om_C(\phi_*\mathscr{O}_C(-\lambda), \omega_{C'}) & \longrightarrow & \omega_{Z'} & \longrightarrow & 0 \end{array}$$

By Grothendieck duality ($\phi_* \dashv \phi^!$ and $\phi^! \omega_{C'} = \omega_C$) and the snake lemma, we conclude that the vertical arrows are isomorphisms. This implies that $\omega_{Z'}$ is a line bundle if $\phi_*\omega_C(\lambda)$ is. Since the sections $\boldsymbol{p}$ are away from $\mathrm{Exc}(\phi)$, by construction

$$\phi_*\omega_C(\lambda) = \mathscr{O}_{C'}(1)(-\phi(\boldsymbol{p}))$$

is a line bundle. We have thus proved that Z' is Gorenstein.

Moreover, $\phi_*\omega_Z = \omega_{Z'}$. Now, to prove that $\chi(\omega_{Z'}) = 1$, it is enough to prove that $\chi(\omega_{Z'}) = \chi(\omega_Z)$. Indeed, over $\mathcal{D}_1$, the definition of λ and adjunction (7) show that $\omega_C(\lambda)$ restricts to a line bundle of degree two on Z, which is therefore a curve of genus two. Since $\phi_*\mathscr{O}_Z = \mathscr{O}_{Z'}$, it is enough to prove the vanishing of $R^1\phi_*\mathscr{O}_Z$. By Proposition 3.49, $R^1\phi_*\mathscr{O}_Z \simeq R^1\phi_*\mathscr{O}_C$; therefore, the desired statement is equivalent to the fact the C' has rational singularities. This follows from $\phi_*\omega_C = \omega_{C'}$; see for example [48]. We have thus proved that Z' has genus two.

In order to prove the flatness of $\mathcal{Z}'_{\mathcal{D}_1} \to \mathcal{D}_1$, it is sufficient to show that

$$\pi'_*(\mathscr{O}_{\mathcal{Z}'_{\mathcal{D}_1}} \otimes \mathscr{O}_{\mathscr{C}'}(n))$$

is a vector bundle on $\mathcal{D}_1$ for n large enough. Since $\mathcal{D}_1$ is a reduced divisor (Lemma 3.51), we only have to show that its rank is constant along $\mathcal{D}_1$. It is easy to see that

$$\mathscr{O}_{\mathscr{C}'}(1)|_{\mathcal{Z}'_{\mathcal{D}_1}} \cong \omega_{\mathcal{Z}'_{\mathcal{D}_1}}.$$

Since $\mathcal{Z}'_{\mathcal{D}_1}$ is a curve of genus two, it follows from Riemann–Roch that

$$H^0(Z'_s, \omega_{Z'_s}^{\otimes n}) = 2n - 1,$$

so it is enough to show that $\mathscr{O}_{\mathscr{C}'}(1)|_{\mathcal{Z}'_{\mathcal{D}_1}}$ satisfies cohomology and base change. For this, we observe the short exact sequence

$$\begin{array}{ccccccc} 0 & \longrightarrow & \pi'_*(\phi_*\mathscr{O}_{\mathscr{C}}(-\lambda)\otimes\mathscr{O}_{\mathscr{C}'}(n)) & \longrightarrow & \pi'_*\mathscr{O}_{\mathscr{C}'}(n) & \to & \pi'_*\mathscr{O}_{\mathscr{C}'}(n)|_{\mathcal{Z}'} \to 0 \\ & & \downarrow\cong & & & & \\ & & \pi_*\big(\omega_\pi^{\otimes n}(n\boldsymbol{p})\otimes\mathscr{O}_{\mathscr{C}}((n-1)\lambda)\big) & & & & \end{array}$$

As $\mathscr{C}$ and $\mathscr{C}'$ are flat over the base, and both $\omega_\pi(\boldsymbol{p})$ and $\mathscr{O}_{\mathscr{C}'}(1)$ are relatively ample, the first two bundles satisfy cohomology and base change. It follows from a diagram chase that so does the third. We have thus proved that $\mathcal{Z}'$ is flat on $\mathcal{D}_1$.

The statement about $\mathscr{T}'_{\mathcal{Z}}$ can be proven in an analogous (but easier) fashion. □

Proof of Theorem 3.47 Noticing that $\mathcal{Z}' \to \mathscr{T}'_{\mathcal{Z}}$ is finite, the existence of the pushout as a scheme over $\widetilde{\mathscr{A}_2}$ follows from results of D Ferrand [30]. We have already proved that the construction of $\mathscr{C}'$, $\mathcal{Z}'$ and $\mathscr{T}'_{\mathcal{Z}}$ commutes with pullback to a generic trait. The pushout does as well in virtue of [16, Tag 0ECK]. So, in order to prove that the fibres of $\overline{\mathscr{C}}$ are Gorenstein, we may work with fibred surfaces, in which case we may apply some results of M Reid [65].

Following [65, Section 2.1], $\bar{\phi}\colon C' \to \overline{C}$ is the normalisation, with conductor

$$\operatorname{Ann}(\bar{\phi}_*\mathscr{O}_{C'}/\mathscr{O}_{\overline{C}}) = \operatorname{Ann}(\psi_*\mathscr{O}_{Z'}/\mathscr{O}_{T'_Z}).$$

Since ψ is a double cover, it is in particular flat, and $\psi_*\mathscr{O}_{Z'}/\mathscr{O}_{T'_Z}$ is a line bundle; it follows from [65, Proposition 2.2] that $\overline{C}$ is $S2$.

Moreover, we have seen in the proof of Proposition 3.50 that

$$\omega_{C'}(Z') = \mathscr{H}om_{\mathscr{C}'}(\mathscr{I}_{\mathcal{Z}'}, \omega_{\mathscr{C}'})$$

is a line bundle on C'. Since $\psi\colon Z' \to T'_Z$ is a double cover of a rational curve, it follows that the kernel of the canonical map $\psi_*\omega_{Z'} \to \omega_{T'_Z}$ is a line bundle as well. The criterion of [65, Corollary 2.8(iv)] allows us to conclude that $\overline{C}$ is Gorenstein.

Finally, the statement about weights is obvious from the construction of λ if its support is connected. If there are two components of Δ°, each of them necessarily of genus one by Lemma 3.17, the statement follows from Lemma 3.28. □

3.9 The secondary construction

We restrict now to $\tilde{\tilde{\mathscr{A}}}_2$, defined in Definition 3.34. Thanks to the alignment, we have:

Definition 3.52 Let $\mathcal{D}_1^2 \subseteq \tilde{\tilde{\mathscr{A}}}_2$ be the logarithmic divisor where the weight of the subcurve Δ determined by λ is two.

The curve $\bar{\mathscr{C}}|_{\tilde{\tilde{\mathscr{A}}}_2}$ will be manipulated further to insert ribbons over $\mathcal{D}_1^2$. We could descend the line bundle $\mathscr{O}_{\mathscr{C}'}(1)$ to $\bar{\mathscr{C}}$ using the results of [30, Theorem 2.2] and proceed from there. For simplicity we chose instead to work on $\tilde{\tilde{\mathscr{C}}}$ from Theorem 3.36, so that the phrasing of the following two theorems is analogous to that of the main theorems in Section 3.6. The proof goes along the same lines too, so we leave it to the reader to figure out the details.

Let $\tilde{\mathscr{L}}$ denote the line bundle $\omega_{\tilde{\tilde{\mathscr{C}}}/\tilde{\tilde{\mathscr{A}}}_2}(\boldsymbol{p})(\tilde{\lambda})$ on $\tilde{\tilde{\mathscr{C}}}$, where $\boldsymbol{p}$ is a local multisection compatible with the weight function; see the notation at the beginning of Section 3.6.

Theorem 3.53 *The line bundle $\tilde{\mathscr{L}}$ is $\tilde{\pi}$–semiample. In the diagram*

$$\begin{array}{ccc} \tilde{\tilde{\mathscr{C}}} & \xrightarrow{\tilde{\phi}} & \mathscr{C}'' := \underline{\mathrm{Proj}}_{\tilde{\tilde{\mathscr{A}}}_2}\big(\tilde{\pi}_* \bigoplus_{k\geq 0} \tilde{\mathscr{L}}^{\otimes k}\big) \\ & {\scriptstyle\tilde{\pi}}\searrow \quad \swarrow{\scriptstyle\pi''} & \\ & \tilde{\tilde{\mathscr{A}}}_2 & \end{array}$$

the morphism π'' is a flat family of reduced, projective, Cohen–Macaulay curves of arithmetic genus two, with Gorenstein fibres outside $\mathcal{D}_1 \cup \mathcal{D}_1^2$. Moreover, we can perform a parallel contraction $\mathscr{T} \to \mathscr{T}''$ so that $\psi'' : \mathscr{C}'' \to \mathscr{T}''$ remains finite. Neither $\mathscr{C}''$ nor $\mathscr{T}''$ depend on the choice of local sections $\boldsymbol{p}$.

Let $\tilde{\tilde{\mathcal{Z}}}$ be the logarithmic divisor on $\tilde{\tilde{\mathscr{C}}}$ defined by $\mathscr{O}_{\tilde{\tilde{\mathscr{C}}}}(-\tilde{\lambda}) \hookrightarrow \mathscr{O}_{\tilde{\tilde{\mathscr{C}}}}$.

Theorem 3.54 *Let $\mathcal{Z}''$ denote the image of $\tilde{\tilde{\mathcal{Z}}}$ under $\tilde{\phi}$. Over $\mathcal{D}_1 \cup \mathcal{D}_1^2$, it is a flat family of Gorenstein curves of arithmetic genus two, and the image of $\mathcal{Z}''$ under ψ'' is a rational curve $\mathscr{T}''_{\mathcal{Z}}$. Let $\bar{\bar{\mathscr{C}}}$ be obtained as the pushout of*

$$\begin{array}{ccc} \mathcal{Z}''_{\mathcal{D}_1\cup\mathcal{D}_1^2} & \hookrightarrow & \mathscr{C}'' \\ \downarrow{\scriptstyle\psi''} & & \\ \mathscr{T}''_{\mathcal{Z}} & & \end{array}$$

Then $\bar{\bar{\mathscr{C}}} \to \tilde{\tilde{\mathscr{A}}}_2$ is a flat family of projective, Gorenstein curves of arithmetic genus two. Moreover, the weight of the minimal subcurve of genus two is at least three.

Remark 3.55 $\mathcal{D}_1$ and $\mathcal{D}_1^2$ do not intersect; thus, the flatness of $\mathcal{Z}''$ can be checked independently on the two components. To this goal, we notice that $\mathscr{O}_{\mathscr{C}''}(1)$ restricts to $\omega_{\mathcal{Z}''_{\mathcal{D}_1}/\mathcal{D}_1}$ on $\mathcal{Z}''_{\mathcal{D}_1}$ (as seen in the previous section) and to $\omega^{\otimes 3}_{\mathcal{Z}''_{\mathcal{D}_1^2}/\mathcal{D}_1^2}$ on $\mathcal{Z}''_{\mathcal{D}_1^2}$. The pushout construction can be carried out independently on these two loci.

Remark 3.56 We could have extended λ to $\tilde{\lambda}$ over the whole $\widetilde{\mathscr{A}_2}$ by choosing a different cutoff level whenever $w(\Delta) \leq 2$. In this case, though, we would have had to deal with the cases that the weight is 1, or $1+1$, or 2 but supported on a different vertex than v. In these cases, the construction above produces singularities worse than tailed ribbons. When there are two vertices in Δ° on which $\mathscr{L}$ has positive degree, the singularity looks like a chain of two ribbons on an underlying node, with local equations $\mathbb{C}[\![x,y]\!]/(x^2y^2)$. When there are three, if the contraction ϕ acts nontrivially, there may even be three ribbons with underlying curve an ordinary 3–fold point. We chose to keep the singularities under control by discarding the bad locus in Definition 3.30. We shall see below that this forces an intermediate step on us (which can actually have some interest of its own), but it does not affect the end result.

3.10 Markings

To avoid overloading the notation and the exposition, we have so far considered only weighted admissible covers without markings. However, with the application to stable maps in mind, markings are necessary to impose cohomological constraints using the evaluation maps. Our construction extends to the marked version essentially unchanged. We can consider:

Definition 3.57 A weighted admissible cover with markings consists of

$$\big(\psi : (C, D_R, \boldsymbol{x}) \to (T, D_B, \boldsymbol{y} = \psi(\boldsymbol{x})), w : V(\sqsubset) \to \mathbb{N}\big)$$

such that D_R and $\boldsymbol{x}$ are separately disjoint (multi)sections of C, and T is weighted-stable, ie every weight zero component has at least three special (marked, branch or nodal) points. We denote the moduli space of weighted admissible covers with n markings by $\mathscr{A}^{\mathrm{wt}}_{2,n}$.

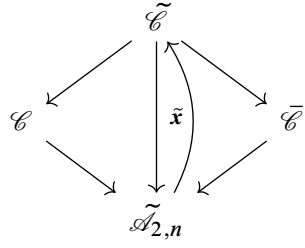
Markings are represented by infinitely long legs on $\sqsubset$. They play no role in the alignment: admissible functions will have slope 1 along them, and the infinite legs will be subdivided accordingly. In particular, if a marking is supported on a vertex of Δ°, then we may *sprout* (blow up) the marking as many times as is necessary for its strict

transform to be supported on a vertex of $\partial\Delta$. (In fact, only one blow-up is necessary if it is allowed to be weighted.)

Theorem 3.58 *There exists a logarithmically étale modification $\widetilde{\mathscr{A}}_{2,n} \to \mathscr{A}^{\mathrm{wt}}_{2,n}$ (resp. $\widetilde{\widetilde{\mathscr{A}}}_{2,n} \to \mathscr{A}^{\mathrm{wt}}_{2,n}$) parametrising weighted admissible covers with markings and a primary (resp. secondary) alignment. The moduli space $\widetilde{\mathscr{A}}_{2,n}$ (resp. $\widetilde{\widetilde{\mathscr{A}}}_{2,n}$) is a logarithmically smooth stack with locally free logarithmic structure (and therefore smooth).*

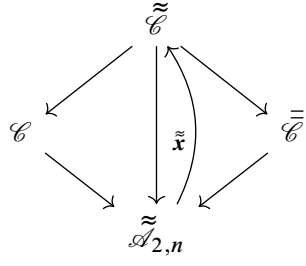
In particular, the strict transform of the markings never touches the singularity.

Theorem 3.59 *There is a diagram of flat families of projective, Gorenstein curves of arithmetic genus two, with n corresponding disjoint sections of the smooth locus:*



Moreover, the weight of every subcurve of positive genus in $\bar{\mathscr{C}}$ is at least one.

Theorem 3.60 *There is a diagram of flat families of projective, Gorenstein curves of arithmetic genus two, with n corresponding disjoint sections of the smooth locus:*



Moreover, the weight of the core of $\bar{\bar{\mathscr{C}}}$ is at least three.

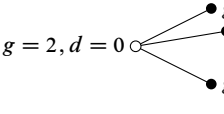
4 A modular desingularisation of $\overline{\mathcal{M}}_{2,n}(\mathbb{P}^r, d)^{\mathbf{main}}$

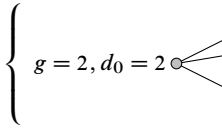
4.1 Irreducible components

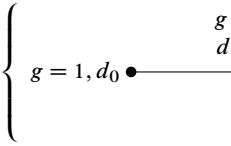
We draw the weighted dual graph of the general member of all possible irreducible components of $\overline{\mathcal{M}}_2(\mathbb{P}^r, d)$. Our running convention is that a white vertex corresponds

to a contracted component, a grey one to a genus two subcurve covering a line two-to-one, and a black vertex to a nonspecial subcurve. Vertices are labelled with their genus and weight.

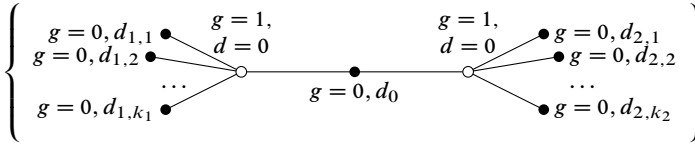
(1) *Main* is the closure of the locus of maps from a smooth curve of genus two.

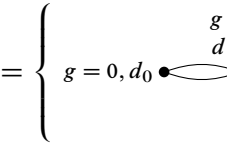
(2) $\mathcal{D}^{(d_1,\dots,d_k)} = \left\{ \right.$  $\left. \right\}$.

(3) ${}^{\mathrm{hyp}}\mathcal{D}^{(d_1,\dots,d_k)} = \left\{ \right.$  $\left. \right\}$.

(4) ${}_{d_0}\mathcal{E}^{(d_1,\dots,d_k)} = \left\{ \right.$  $\left. \right\}$.

(5) ${}_{(d_{1,1},\dots,d_{1,k_1})}\mathcal{E}_{d_0}\mathcal{E}^{(d_{2,1},\dots,d_{2,k_2})}$

$= \left\{ \right.$  $\left. \right\}$.

(6) ${}^{\mathrm{br}=d_0}\mathcal{E}^{(d_1,\dots,d_k)} = \left\{ \right.$  $\left. \right\}$.

This is taken from the first author's PhD thesis [10], and is implicit in [40].

4.2 Factorisation through a Gorenstein curve

Definition 4.1 Let $(X, \mathscr{O}_X(1))$ be a polarised variety, and $\beta \in H_2^+(X, \mathbb{Z})$ an effective curve class. The moduli space $\mathscr{A}_{2,n}(X, \beta)$ of *admissible maps* to X is defined by the fibre diagram

$$\begin{array}{ccc} \mathscr{A}_{2,n}(X,\beta) & \longrightarrow & \overline{\mathcal{M}}_{2,n}(X,\beta) \\ \downarrow & \square & \downarrow \\ \mathscr{A}^{\mathrm{wt}}_{2,n} & \longrightarrow & \mathfrak{M}^{\mathrm{wt}}_{2,n} \end{array}$$

It can be described as a space of maps and admissible covers with the same source

$$\begin{array}{ccc}(C, D_R, \boldsymbol{x}) & \xrightarrow{f} & \mathbb{P}^r \\ \downarrow{\scriptstyle \psi} & & \\ (T, D_B, \psi(\boldsymbol{x})) & & \end{array}$$

subject to the stability condition that, if $\sigma : C \to C$ is the hyperelliptic involution,

$$\omega_C(\boldsymbol{x}) \otimes f^*\mathscr{O}_X(2) \otimes \sigma^* f^*\mathscr{O}_X(2)$$

is relatively ample.

If we set $d = \mathscr{O}_X(1) \cdot \beta$, we can restrict to the component of the base where the weight is d.

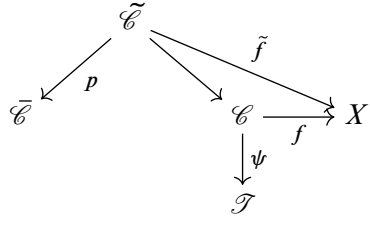
Remark 4.2 On the locus of maps from a smooth curve, $\mathscr{A}_{2,n}(X, \beta) \to \overline{\mathcal{M}}_{2,n}(X, \beta)$ is an isomorphism; therefore, for $X = \mathbb{P}^r$, the main components are birational.

Definition 4.3 Let $\widetilde{\mathscr{A}}_{2,n}(X, \beta)$ denote the fibre product

$$\begin{array}{ccc}\widetilde{\mathscr{A}}_{2,n}(X, \beta) & \longrightarrow & \mathscr{A}_{2,n}(X, \beta) \\ \downarrow & \square & \downarrow \\ \widetilde{\mathscr{A}}_{2,n} & \longrightarrow & \mathscr{A}^{\mathrm{wt}}_{2,n}\end{array}$$

We call it the moduli space of *aligned admissible maps*.

Remark 4.4 $\widetilde{\mathscr{A}}_{2,n}(X, \beta)$ is logarithmically étale over $\mathscr{A}_{2,n}(X, \beta)$. It comes with universal structures



Definition 4.5 Let $\widetilde{\mathscr{A}}_{2,n}(X, \beta)^{\mathrm{fact}} \subseteq \widetilde{\mathscr{A}}_{2,n}(X, \beta)$ be the locus of maps such that $\tilde{f} : \widetilde{\mathscr{C}} \to X$ factors through a map $\bar{f} : \bar{\mathscr{C}} \to X$. We call it the moduli space of aligned admissible maps satisfying the *first factorisation property*.

Remark 4.6 The map f need *not* factor through the admissible cover ψ.

Proposition 4.7 *$\widetilde{\mathscr{A}}_{2,n}(X,\beta)^{\mathrm{fact}} \subseteq \widetilde{\mathscr{A}}_{2,n}(X,\beta)$ is a closed substack. If X is smooth, there is a perfect obstruction theory*

$$(R^\bullet \bar{\pi}_* \bar{f}^* T_X)^\vee \to \mathbb{L}^\bullet_{\widetilde{\mathscr{A}}_{2,n}(X,\beta)^{\mathrm{fact}}/\widetilde{\mathscr{A}}_{2,n}}$$

endowing $\widetilde{\mathscr{A}}_{2,n}(X,\beta)^{\mathrm{fact}}$ with a virtual fundamental class in $A_{\mathrm{vdim}}(\widetilde{\mathscr{A}}_{2,n}(X,\beta)^{\mathrm{fact}})$, where

$$\mathrm{vdim}(\mathcal{VZ}_{2,n}(X,\beta)) = 3 - \dim(X) + n - K_X \cdot \beta.$$

Proof For the first claim, we refer the reader to [63, Theorem 4.3]. The second claim goes back to K Behrend and B Fantechi [15, Proposition 6.3]. □

Although $\widetilde{\mathscr{A}}_{2,n}(\mathbb{P}^r,d)^{\mathrm{fact}}$ is not necessarily smooth, as it may still have a hyperelliptic component, it can already be useful to the end of computing the invariants of a projective complete intersection.

Lemma 4.8 *If $\iota\colon X \hookrightarrow \mathbb{P}^r$ is a complete intersection of degree $(l_1,\dots,l_h)$ ($l_i \geq 2$ for all i),*

$$\widetilde{\mathscr{A}}_{2,n}(X,\beta)^{\mathrm{fact}} \subseteq \widetilde{\mathscr{A}}_{2,n}(\mathbb{P}^r,\iota_*\beta)^{\mathrm{fact}}$$

is cut out by a section of the vector bundle

$$\mathscr{E}_{(l_1,\dots,l_k)} = \bar{\pi}_* \bar{f}^* \Big(\bigoplus_{i=1}^{h} \mathscr{O}_{\mathbb{P}^r}(l_i) \Big).$$

*In particular, the invariants satisfy the **quantum Lefschetz** hyperplane principle.*

Proof Thanks to the last statement of Theorem 3.47, the pullback $f^*\mathscr{O}_{\mathbb{P}^r}(1)$ has degree at least one on any subcurve of positive genus. Then the degree has to be at least two on the minimal subcurve of genus two in order for a nonconstant map to exist. This implies that $\bar{f}^*\mathscr{O}_{\mathbb{P}^r}(l_i)$ has vanishing h^1 along the fibres of $\bar{\mathscr{C}}$ for $l_i \geq 2$. Therefore, $\mathscr{E}_{(l_1,\dots,l_k)}$ is a vector bundle by "cohomology and base change". The virtual statement follows as in genus zero from [45]. □

Lemma 4.9 *The projection $\widetilde{\mathscr{A}}_{2,n}(\mathbb{P}^r,d)^{\mathrm{fact}} \to \widetilde{\mathscr{A}}_{2,n}$ factors through $\widetilde{\mathscr{A}}^\circ_{2,n}$.*

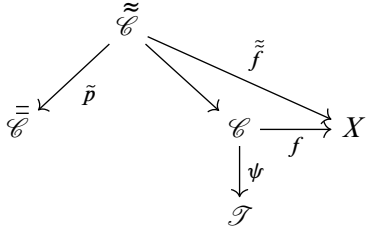
Proof See Definition 3.30 for the notation. The key observation is that the only (nonconstant) maps of degree two from a minimal Gorenstein curve of genus two are those which factor through the hyperelliptic cover of a line. Hence, we may discard the locus where the weight of Δ is two but (partly) supported away from D. □

Thanks to this lemma we can make the following:

Definition 4.10 Let $\widetilde{\widetilde{\mathscr{A}}}_{2,n}(X,\beta)$ be the moduli space of *admissible maps with a secondary alignment* defined by the fibre diagram

$$\begin{array}{ccc} \widetilde{\widetilde{\mathscr{A}}}_{2,n}(X,\beta) & \longrightarrow & \widetilde{\mathscr{A}}_{2,n}(X,\beta)^{\mathrm{fact}} \\ \downarrow & \square & \downarrow \\ \widetilde{\widetilde{\mathscr{A}}}_{2,n} & \longrightarrow & \widetilde{\mathscr{A}}^{\circ}_{2,n} \end{array}$$

This space comes with universal structures



Let $\mathcal{VZ}_{2,n}(X,\beta) \subseteq \widetilde{\widetilde{\mathscr{A}}}_{2,n}(X,\beta)$ be the locus of maps such that $\tilde{\tilde{f}}\colon \tilde{\tilde{\mathscr{C}}} \to X$ factors through a map $\bar{f}\colon \bar{\bar{\mathscr{C}}} \to X$. We call it the moduli space of aligned admissible maps satisfying the *second factorisation property*.[3]

Definition 4.11 Let $\mathrm{ev}_i\colon \mathcal{VZ}_{2,n}(X,\beta) \to X$ denote the evaluation map at the i^{th} marked point for $i = 1,\dots,n$. Let $\alpha_1,\dots,\alpha_n \in H^*(X)$ be cohomology classes on the target manifold. The *reduced genus two Gromov–Witten invariants* are defined as

$$\langle \alpha_1,\dots,\alpha_n\rangle^{X,\mathrm{red}}_{2,\beta,n} = \int_{[\mathcal{VZ}_{2,n}(X,\beta)]^{\mathrm{vir}}} \mathrm{ev}_1^*\,\alpha_1 \cup \cdots \cup \mathrm{ev}_n^*\,\alpha_n.$$

Theorem 4.12 *For $d \geq 3$, $\mathcal{VZ}_{2,n}(\mathbb{P}^r,d)$ is a desingularisation of $\overline{\mathcal{M}}_{2,n}(\mathbb{P}^r,d)^{\mathrm{main}}$.*

Proof Consider the factorisation

$$\mathcal{VZ}_{2,n}(\mathbb{P}^r,d) \to \mathfrak{Pic}_{\widetilde{\mathscr{A}}_{2,n}} \to \widetilde{\mathscr{A}}_{2,n}.$$

Obstructions to the first map can be found in $H^1(\bar{C},L)$, where $L = \bar{f}^*\mathscr{O}_{\mathbb{P}^r}(1)$. If there are two disconnected subcurves of genus one, the degree must be positive on each of them by Theorem 3.59. On the other hand, if the core is a minimal subcurve of genus two, it must have degree at least three by Theorem 3.60, and either the special component has positive degree (at least two) by Corollary 3.26, or the special component is a collapsed ribbon with at least two tails of positive degree by Corollary 3.27. It follows from Lemmas 2.33 and 2.34 that the obstructions vanish.

[3]The notation is reminiscent of the celebrated desingularisation of $\mathcal{M}_{1,n}(\mathbb{P}^r,d)^{\mathrm{main}}$ due to Vakil and Zinger [76].

Obstructions to the second map lie in $H^2(\overline{C}, \mathscr{O})$, which vanishes for dimension reasons. The map is therefore unobstructed. The base is smooth by Theorem 3.36. We conclude that $\mathcal{VZ}_{2,n}(\mathbb{P}^r, d)$ is smooth as well. Since it is proper and it contains the locus of maps from a smooth curve as an open dense, $\mathcal{VZ}_{2,n}(\mathbb{P}^r, d) \to \overline{\mathcal{M}}_{2,n}(\mathbb{P}^r, d)^{\mathrm{main}}$ is birational (see Remarks 4.2 and 4.4). □

Remark 4.13 A posteriori, we note that aligning and the factorisation property do not alter the main component of $\overline{\mathcal{M}}_2(\mathbb{P}^r, 2)$, which therefore is already smooth.

References

[1] **D Abramovich**, **Q Chen**, **S Marcus**, **M Ulirsch**, **J Wise**, *Skeletons and fans of logarithmic structures*, from "Nonarchimedean and tropical geometry" (M Baker, S Payne, editors), Springer (2016) 287–336 MR Zbl

[2] **D Abramovich**, **Q Chen**, **S Marcus**, **J Wise**, *Boundedness of the space of stable logarithmic maps*, J. Eur. Math. Soc. 19 (2017) 2783–2809 MR Zbl

[3] **D Abramovich**, **A Corti**, **A Vistoli**, *Twisted bundles and admissible covers*, Comm. Algebra 31 (2003) 3547–3618 MR Zbl

[4] **D Abramovich**, **K Karu**, *Weak semistable reduction in characteristic 0*, Invent. Math. 139 (2000) 241–273 MR Zbl

[5] **D Abramovich**, **J Wise**, *Birational invariance in logarithmic Gromov–Witten theory*, Compos. Math. 154 (2018) 595–620 MR Zbl

[6] **J Alper**, **M Fedorchuk**, **D I Smyth**, *Singularities with $\mathbb{G}_m$–action and the log minimal model program for $\overline{\mathcal{M}}_g$*, J. Reine Angew. Math. 721 (2016) 1–41 MR Zbl

[7] **J Alper**, **M Fedorchuk**, **D I Smyth**, **F van der Wyck**, *Second flip in the Hassett–Keel program: a local description*, Compos. Math. 153 (2017) 1547–1583 MR Zbl

[8] **O Amini**, **M Baker**, **E Brugallé**, **J Rabinoff**, *Lifting harmonic morphisms, II: Tropical curves and metrized complexes*, Algebra Number Theory 9 (2015) 267–315 MR Zbl

[9] **M Bainbridge**, **D Chen**, **Q Gendron**, **S Grushevsky**, **M Möller**, *The moduli space of multi-scale differentials*, preprint (2019) arXiv 1910.13492

[10] **L Battistella**, *Alternative compactifications in low genus Gromov–Witten theory*, PhD thesis, Imperial College London (2018) `https://spiral.imperial.ac.uk/handle/10044/1/64777`

[11] **L Battistella**, *Modular compactifications of $\mathcal{M}_{2,n}$ with Gorenstein curves*, Algebra Number Theory 16 (2022) 1547–1587 MR Zbl

[12] **L Battistella**, **F Carocci**, *A geographical study of $\overline{\mathcal{M}}_2(\mathbb{P}^2, 4)^{\mathrm{main}}$*, Adv. Geom. 22 (2022) 463–480 MR Zbl

[13] **L Battistella**, **N Nabijou**, **D Ranganathan**, *Curve counting in genus one: elliptic singularities and relative geometry*, Algebr. Geom. 8 (2021) 637–679 MR Zbl

[14] **D Bayer**, **D Eisenbud**, *Ribbons and their canonical embeddings*, Trans. Amer. Math. Soc. 347 (1995) 719–756 MR Zbl

[15] **K Behrend**, **B Fantechi**, *The intrinsic normal cone*, Invent. Math. 128 (1997) 45–88 MR Zbl

[16] **P Belmans**, **A J de Jong**, et al., *The Stacks project*, electronic reference (2005–) http://stacks.math.columbia.edu

[17] **N Borne**, **A Vistoli**, *Parabolic sheaves on logarithmic schemes*, Adv. Math. 231 (2012) 1327–1363 MR Zbl

[18] **S Bozlee**, *Contractions of subcurves of families of log curves*, Comm. Algebra 49 (2021) 4616–4660 MR Zbl

[19] **S Bozlee**, **B Kuo**, **A Neff**, *A classification of modular compactifications of the space of pointed elliptic curves by Gorenstein curves*, Algebra Number Theory 17 (2023) 127–163 MR Zbl

[20] **F Catanese**, *Pluricanonical Gorenstein curves*, from "Enumerative geometry and classical algebraic geometry" (P Le Barz, Y Hervier, editors), Progr. Math. 24, Birkhäuser, Boston, MA (1982) 51–95 MR Zbl

[21] **R Cavalieri**, **M Chan**, **M Ulirsch**, **J Wise**, *A moduli stack of tropical curves*, Forum Math. Sigma 8 (2020) art. id. e23 MR Zbl

[22] **R Cavalieri**, **H Markwig**, **D Ranganathan**, *Tropicalizing the space of admissible covers*, Math. Ann. 364 (2016) 1275–1313 MR Zbl

[23] **M Chan**, *Tropical hyperelliptic curves*, J. Algebraic Combin. 37 (2013) 331–359 MR Zbl

[24] **H-L Chang**, **S Guo**, **J Li**, *BCOV's Feynman rule of quintic* 3*–folds*, preprint (2018) arXiv 1810.00394

[25] **I Ciocan-Fontanine**, **B Kim**, *Moduli stacks of stable toric quasimaps*, Adv. Math. 225 (2010) 3022–3051 MR Zbl

[26] **T Coates**, **C Manolache**, *A splitting of the virtual class for genus one stable maps*, preprint (2018) arXiv 1809.04162

[27] **M A Cueto**, **H Markwig**, *Tropical geometry of genus two curves*, J. Algebra 517 (2019) 457–512 MR Zbl

[28] **O Debarre**, *Higher-dimensional algebraic geometry*, Springer (2001) MR Zbl

[29] **J-M Drézet**, *Paramétrisation des courbes multiples primitives*, Adv. Geom. 7 (2007) 559–612 MR Zbl

[30] **D Ferrand**, *Conducteur, descente et pincement*, Bull. Soc. Math. France 131 (2003) 553–585 MR Zbl

[31] **L-Y Fong**, *Rational ribbons and deformation of hyperelliptic curves*, J. Algebraic Geom. 2 (1993) 295–307 MR Zbl

[32] **W D Gillam**, *Logarithmic stacks and minimality*, Internat. J. Math. 23 (2012) art. id. 1250069 MR Zbl

[33] **T Graber**, **J Kock**, **R Pandharipande**, *Descendant invariants and characteristic numbers*, Amer. J. Math. 124 (2002) 611–647 MR Zbl

[34] **M Gross**, **B Siebert**, *Logarithmic Gromov–Witten invariants*, J. Amer. Math. Soc. 26 (2013) 451–510 MR Zbl

[35] **S Guo**, **F Janda**, **Y Ruan**, *A mirror theorem for genus two Gromov–Witten invariants of quintic threefolds*, preprint (2017) arXiv 1709.07392

[36] **C Haase**, **G Musiker**, **J Yu**, *Linear systems on tropical curves*, Math. Z. 270 (2012) 1111–1140 MR Zbl

[37] **J Harris**, **D Mumford**, *On the Kodaira dimension of the moduli space of curves*, Invent. Math. 67 (1982) 23–88 MR Zbl

[38] **Y Hu**, **J Li**, *Genus-one stable maps, local equations, and Vakil–Zinger's desingularization*, Math. Ann. 348 (2010) 929–963 MR Zbl

[39] **Y Hu**, **J Li**, *Derived resolution property for stacks, Euler classes and applications*, Math. Res. Lett. 18 (2011) 677–690 MR Zbl

[40] **Y Hu**, **J Li**, **J Niu**, *Genus two stable maps, local equations and modular resolutions*, preprint (2018) arXiv 1201.2427v3

[41] **Y Hu**, **J Niu**, *A theory of stacks with twisted fields and resolution of moduli of genus two stable maps*, preprint (2020) arXiv 2005.03384

[42] **D Johnson**, **A Polishchuk**, *Birational models of $\mathcal{M}_{2,2}$ arising as moduli of curves with nonspecial divisors*, Adv. Geom. 21 (2021) 23–43 MR Zbl

[43] **F Kato**, *Log smooth deformation and moduli of log smooth curves*, Internat. J. Math. 11 (2000) 215–232 MR Zbl

[44] **B Kim**, *Logarithmic stable maps*, from "New developments in algebraic geometry, integrable systems and mirror symmetry" (M-H Saito, S Hosono, K Yoshioka, editors), Adv. Stud. Pure Math. 59, Math. Soc. Japan, Tokyo (2010) 167–200 MR Zbl

[45] **B Kim**, **A Kresch**, **T Pantev**, *Functoriality in intersection theory and a conjecture of Cox, Katz, and Lee*, J. Pure Appl. Algebra 179 (2003) 127–136 MR Zbl

[46] **J Kollár**, **S Mori**, *Birational geometry of algebraic varieties*, Cambridge Tracts in Mathematics 134, Cambridge Univ. Press (1998) MR Zbl

[47] **M Kontsevich**, *Enumeration of rational curves via torus actions*, from "The moduli space of curves" (R Dijkgraaf, C Faber, G van der Geer, editors), Progr. Math. 129, Birkhäuser, Boston, MA (1995) 335–368 MR Zbl

[48] **S J Kovács**, *Rational singularities*, preprint (2017) arXiv 1703.02269

[49] **A G Kuznetsov**, *Hyperplane sections and derived categories*, Izv. Ross. Akad. Nauk Ser. Mat. 70 (2006) 23–128 MR Zbl In Russian; translated in Izv. Math. 70 (2006) 447–547

[50] **Y-P Lee**, *Quantum Lefschetz hyperplane theorem*, Invent. Math. 145 (2001) 121–149 MR Zbl

[51] **J Li**, **A Zinger**, *On the genus-one Gromov–Witten invariants of complete intersections*, J. Differential Geom. 82 (2009) 641–690 MR Zbl

[52] **C Manolache**, *Virtual push-forwards*, Geom. Topol. 16 (2012) 2003–2036 MR Zbl

[53] **S Marcus**, **J Wise**, *Logarithmic compactification of the Abel–Jacobi section*, Proc. Lond. Math. Soc. 121 (2020) 1207–1250 MR Zbl

[54] **H Matsumura**, *Commutative ring theory*, 2nd edition, Cambridge Studies in Advanced Mathematics 8, Cambridge Univ. Press (1989) MR Zbl

[55] **S Mochizuki**, *The geometry of the compactification of the Hurwitz scheme*, Publ. Res. Inst. Math. Sci. 31 (1995) 355–441 MR Zbl

[56] **S Molcho**, *Universal stacky semistable reduction*, Israel J. Math. 242 (2021) 55–82 MR Zbl

[57] **M Möller**, **M Ulirsch**, **A Werner**, *Realizability of tropical canonical divisors*, J. Eur. Math. Soc. 23 (2021) 185–217 MR Zbl

[58] **W Nizioł**, *Toric singularities: log-blow-ups and global resolutions*, J. Algebraic Geom. 15 (2006) 1–29 MR Zbl

[59] **M C Olsson**, *Logarithmic geometry and algebraic stacks*, Ann. Sci. École Norm. Sup. 36 (2003) 747–791 MR Zbl

[60] **M C Olsson**, *(Log) twisted curves*, Compos. Math. 143 (2007) 476–494 MR Zbl

[61] **K S Parker**, *Semistable modular compactifications of moduli spaces of genus one curves*, PhD thesis, University of Colorado at Boulder (2017) MR `https://www.proquest.com/docview/1904507277`

[62] **A Popa**, *The genus one Gromov–Witten invariants of Calabi–Yau complete intersections*, Trans. Amer. Math. Soc. 365 (2013) 1149–1181 MR Zbl

[63] **D Ranganathan**, **K Santos-Parker**, **J Wise**, *Moduli of stable maps in genus one and logarithmic geometry, I*, Geom. Topol. 23 (2019) 3315–3366 MR Zbl

[64] **D Ranganathan**, **K Santos-Parker**, **J Wise**, *Moduli of stable maps in genus one and logarithmic geometry, II*, Algebra Number Theory 13 (2019) 1765–1805 MR Zbl

[65] **M Reid**, *Nonnormal del Pezzo surfaces*, Publ. Res. Inst. Math. Sci. 30 (1994) 695–727 MR Zbl

[66] **J-P Serre**, *Algebraic groups and class fields*, Graduate Texts in Math. 117, Springer (1988) MR Zbl

[67] **D I Smyth**, *Modular compactifications of the space of pointed elliptic curves, I*, Compos. Math. 147 (2011) 877–913 MR Zbl

[68] **D I Smyth**, *Modular compactifications of the space of pointed elliptic curves, II*, Compos. Math. 147 (2011) 1843–1884 MR Zbl

[69] **D E Speyer**, *Tropical geometry*, PhD thesis, University of California, Berkeley (2005) MR `https://www.proquest.com/docview/305031446`

[70] **J Stevens**, *On the classification of reducible curve singularities*, from "Algebraic geometry and singularities" (A Campillo López, L Narváez Macarro, editors), Progr. Math. 134, Birkhäuser, Basel (1996) 383–407 MR Zbl

[71] **B Teissier**, *The hunting of invariants in the geometry of discriminants*, from "Real and complex singularities" (P Holm, editor), Sijthoff and Noordhoff, Alphen aan den Rijn (1977) 565–678 MR Zbl

[72] **M Ulirsch**, *Functorial tropicalization of logarithmic schemes: the case of constant coefficients*, Proc. Lond. Math. Soc. 114 (2017) 1081–1113 MR Zbl

[73] **M Ulirsch**, *Non-Archimedean geometry of Artin fans*, Adv. Math. 345 (2019) 346–381 MR Zbl

[74] **R Vakil**, *The enumerative geometry of rational and elliptic curves in projective space*, J. Reine Angew. Math. 529 (2000) 101–153 MR Zbl

[75] **R Vakil**, *Murphy's law in algebraic geometry: badly-behaved deformation spaces*, Invent. Math. 164 (2006) 569–590 MR Zbl

[76] **R Vakil**, **A Zinger**, *A desingularization of the main component of the moduli space of genus-one stable maps into $\mathbb{P}^n$*, Geom. Topol. 12 (2008) 1–95 MR Zbl

[77] **M Viscardi**, *Alternate compactifications of the moduli space of genus one maps*, Manuscripta Math. 139 (2012) 201–236 MR Zbl

[78] **J Wang**, *Deformations of pairs (X, L) when X is singular*, Proc. Amer. Math. Soc. 140 (2012) 2953–2966 MR Zbl

[79] **P M H Wilson**, *On blowing up conductor ideals*, Math. Proc. Cambridge Philos. Soc. 83 (1978) 445–450 MR Zbl

[80] **F D W van der Wyck**, *Moduli of singular curves and crimping*, PhD thesis, Harvard University (2010) MR https://www.proquest.com/docview/613377231

[81] **A Zinger**, *Enumeration of genus-two curves with a fixed complex structure in $\mathbb{P}^2$ and $\mathbb{P}^3$*, J. Differential Geom. 65 (2003) 341–467 MR Zbl

[82] **A Zinger**, *Standard versus reduced genus-one Gromov–Witten invariants*, Geom. Topol. 12 (2008) 1203–1241 MR Zbl

[83] **A Zinger**, *The reduced genus 1 Gromov–Witten invariants of Calabi–Yau hypersurfaces*, J. Amer. Math. Soc. 22 (2009) 691–737 MR Zbl

[84] **A Zinger**, *A sharp compactness theorem for genus-one pseudo-holomorphic maps*, Geom. Topol. 13 (2009) 2427–2522 MR Zbl

Humboldt-Universität zu Berlin, Institut für Mathematik
Berlin, Germany

Institute of Mathematics, École Polytechnique Fédérale de Lausanne
Lausanne, Switzerland

luca.battistella@hu-berlin.de, francesca.carocci@epfl.ch

Proposed: Dan Abramovich
Seconded: Mark Gross, Richard P Thomas
Received: 10 January 2021
Revised: 12 August 2021

Guidelines for Authors

Submitting a paper to Geometry & Topology

Papers must be submitted using the upload page at the GT website. You will need to choose a suitable editor from the list of editors' interests and to supply MSC codes.

The normal language used by the journal is English. Articles written in other languages are acceptable, provided your chosen editor is comfortable with the language and you supply an additional English version of the abstract.

Preparing your article for Geometry & Topology

At the time of submission you need only supply a PDF file. Once accepted for publication, the paper must be supplied in LaTeX, preferably using the journal's class file. More information on preparing articles in LaTeX for publication in GT is available on the GT website.

`arXiv` papers

If your paper has previously been deposited on the `arXiv`, we will need its `arXiv` number at acceptance time. This allows us to deposit the DOI of the published version on the paper's `arXiv` page.

References

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited at least once in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used, but will be converted to the house style (see a current issue for examples).

Figures

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Fuzzy or sloppily drawn figures will not be accepted. For labeling figure elements consider the pinlabel LaTeX package, but other methods are fine if the result is editable. If you're not sure whether your figures are acceptable, check with production by sending an email to graphics@msp.org.

Proofs

Page proofs will be made available to authors (or to the designated corresponding author) in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

GEOMETRY & TOPOLOGY
Volume 27 Issue 3 (pages 823–1272) 2023