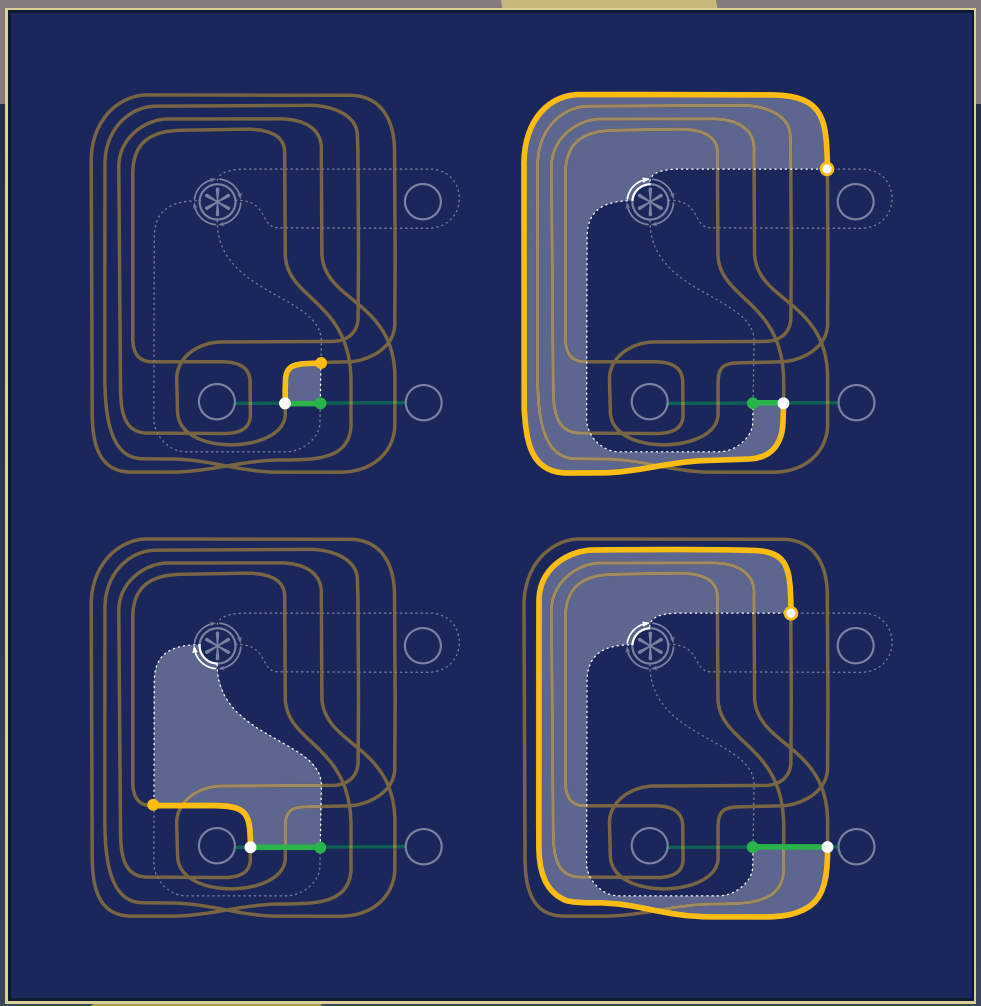
THE OPEN BOOK SERIES 5

Gauge Theory and Low-Dimensional Topology: Progress and Interaction

Banff International Research Station, 2020

edited by
John A. Baldwin, Hans U. Boden, John B. Etnyre, and Liam Watson



msp

Gauge Theory and
Low-Dimensional Topology:
Progress and Interaction

THE OPEN BOOK SERIES 5

Gauge Theory and Low-Dimensional Topology: Progress and Interaction

Banff International Research Station, 2020

Edited by
John A. Baldwin
Hans U. Boden
John B. Etnyre
Liam Watson

msp

Volume Editors:

John A. Baldwin
Boston College
Boston, MA
United States

Hans U. Boden
McMaster University
Hamilton, ON
Canada

John B. Etnyre
Georgia Institute of Technology
Atlanta, GA
United States

Liam Watson
University of British Columbia
Vancouver, BC
Canada

The cover image is based on an illustration from the article "Khovanov homology and strong inversions", by Artem Kotelskiy, Liam Watson and Claudius Zibrowius (see p. 230).

Electronic copies can be obtained free of charge from http://msp.org/obs/5 and printed copies can be ordered from MSP (contact@msp.org).

The Open Book Series is a trademark of Mathematical Sciences Publishers.

ISSN: 2329-9061 (print), 2329-907X (electronic)

ISBN: 978-1-935107-09-5 (print), 978-1-935107-10-1 (electronic)

First published 2022.

MATHEMATICAL SCIENCES PUBLISHERS
798 Evans Hall #3840, c/o University of California, Berkeley CA 94720-3840
contact@msp.org **https://msp.org**

This book is a proceedings of the sixth Banff International Research Station workshop on "Interactions of gauge theory with contact and symplectic topology in dimensions 3 and 4", held in 2020 over the internet. The volume contains sixteen refereed papers, with an emphasis on new developments and interconnections of gauge theory, low-dimensional topology, contact and symplectic topology. It is representative of the high level of talks and results presented at the Interactions workshops over the years since its inception in 2007.

Contents

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.9

Preface

This volume is a proceedings of the 2020 Banff International Research Station (BIRS) workshop "Interactions of gauge theory with contact and symplectic topology in dimensions 3 and 4". This was the sixth iteration of a recurring workshop held in Banff. Regrettably, the workshop was not held onsite but was instead an online gathering over Zoom, as a result of the Covid-19 pandemic. However, one benefit of the online format was that the participant list could be expanded beyond the usual strict limit of 42 individuals. It seemed to be also fitting, given the altered circumstances and larger than usual list of participants, to take the opportunity to put together a conference proceedings.

The result is this volume, which features papers showcasing research from participants at the sixth Interactions workshop (or earlier ones). As the title suggests, the emphasis is on research in gauge theory, contact and symplectic topology, and low-dimensional topology. The volume contains sixteen refereed papers, and it is representative of the many excellent talks and fascinating results presented at the Interactions workshops over the years since its inception in 2007.

We take this opportunity to acknowledge the contributions of Dennis Auroux and Olivier Collin, who were instrumental in founding and organizing the BIRS Interactions workshops. Partial funding was provided by NSF grant DMS1454865 and the Georgia Institute of Technology's Elaine M. Hubbard Distinguished Faculty Award. We thank Conall Hegarty and Fintan Hegarty for their rapid and professional copy-editing services. We also thank Alex Scorpan and Silvio Levy, our contacts at Mathematical Sciences Publishers, for all their help and support throughout the publication process.

JOHN A. BALDWIN, Boston College
john.baldwin@bc.edu

HANS U. BODEN, McMaster University
boden@mcmaster.ca

JOHN B. ETNYRE, Georgia Institute of Technology
etnyre@math.gatech.edu

LIAM WATSON, University of British Columbia
liam@math.ubc.ca

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.1

A friendly introduction to the bordered contact invariant

Akram Alishahi, Joan E. Licata, Ina Petkova and Vera Vértesi

We give a short introduction to the contact invariant in bordered Floer homology defined by Földvári, Hendricks, and the authors. We survey the contact geometry required to understand the new invariant but assume some familiarity with bordered Heegaard Floer invariants. The input for the construction is a special class of foliated open books, which are introduced carefully and with multiple examples. We discuss how a foliated open book may be constructed from an open book for a closed manifold, and how it may be modified to ensure compatibility with the contact bordered invariant. As an application of these techniques, we give a "local proof" of the vanishing of the contact invariant for overtwisted structures in the form of an explicit bordered computation.

1. Introduction

Contact geometry, often pitched as the odd-dimensional complement to symplectic geometry, considers a $(2k+1)$-dimensional manifold equipped with some additional structure. In dimension three — where we reside henceforth — this extra data is a nowhere-integrable plane field called a contact structure. Adding this extra data prompts interesting new questions, but one of the most intriguing features of the subject is that this "extra" data also offers insight into topological structure apparently unrelated to plane fields at all. Two notable examples are the role of contact geometry in the proof of the property P conjecture [11] and the proofs that knot Floer homology detects knot genus [17] and fiberedness [3; 16].

Contact structures themselves split into two mutually exclusive types, known as *tight* and *overtwisted*. Overtwisted structures are determined by homotopical

Alishahi was supported by NSF Grant DMS-2019396. Vértesi was supported by the ANR grant "Quantum topology and contact geometry".
MSC2020: 57K33, 57R58.
Keywords: Heegaard Floer homology, open book, TQFT, contact topology.

data, and so are easy to understand. In contrast, tight contact structures are more mysterious: some, but not all, tight contact structures arise naturally as the boundary of symplectic manifolds, and tight contact structures do not satisfy an h-principle. Many existence and classification questions for tight contact structures remain open, but significant progress has been made since the advent of Heegaard Floer homology in the early 2000s and the subsequent development of Floer-theoretic contact invariants.

Like other Heegaard Floer invariants, the input data for these constructions is a Heegaard diagram for the three-manifold, but in this setting, the Heegaard diagram is induced from an *open book decomposition*, a topological decomposition of a three-manifold that captures the additional data of an equivalence class of contact structure. Ozsváth and Szabó defined the first Heegaard Floer invariant of closed contact three-manifolds in [18]. Given a closed, contact manifold (M, ξ), this invariant is a class $c(\xi)$ in the Heegaard Floer homology $\widehat{HF}(-M)$. In [8], Honda, Kazez, and Matić gave an alternative description of $c(\xi)$, again using open books. This "contact class" gives information about overtwistedness: if ξ is overtwisted, then $c(\xi) = 0$, whereas if ξ is Stein fillable, then $c(\xi) \neq 0$ [18]. The contact class was used in the knot Floer homology proofs noted above, and also to distinguish notions of fillability: Ghiggini used it to construct examples of strongly symplectically fillable contact three-manifolds which do not have Stein fillings [2].

In this paper, we discuss a recent extension of the contact class to three-manifolds with boundary. Namely, in [1], a contact invariant was defined in the bordered sutured Floer homology of a foliated contact three-manifold $(M, \xi, \mathcal{F})$, which is a contact manifold with a certain type of singular foliation on the boundary. We associate to a foliated contact three-manifold a bordered sutured manifold $(M, \Gamma, \mathcal{Z})$. The resulting sutures are particularly simple, so one can think of $(M, \Gamma, \mathcal{Z})$ as a bordered manifold $(M, \mathcal{Z})$ of a type slightly more general than in [14]. Below, we rephrase the main results of [1], translating from "bordered sutured" to "multipointed" language. Section 4 explores the correspondence between these two viewpoints in more detail.

Using a special decomposition of $(M, \xi, \mathcal{F})$ called a *sorted foliated open book*, one can construct an admissible multipointed bordered Heegaard diagram for the manifold $(M, \mathcal{Z})$ and identify a preferred generator. This preferred generator is an invariant of the contact structure.

Theorem 1.1 [1, Theorem 1]. *Let $(M, \xi, \mathcal{F})$ be a foliated contact three-manifold with associated bordered manifold $(M, \mathcal{Z})$. Then there are invariants $c_D(M, \xi, \mathcal{F})$ and $c_A(M, \xi, \mathcal{F})$ of the contact structure which are well defined homotopy equivalence classes in the multipointed bordered Floer homologies $\widetilde{CFD}(-M, \overline{\mathcal{Z}})$ and $\widetilde{CFA}(-M, \overline{\mathcal{Z}})$, respectively.*

Furthermore, this generator vanishes for overtwisted manifolds, in the following sense.

Theorem 1.2 [1, Corollary 4]. *If $(M, \xi, \mathcal{F})$ is overtwisted, the classes $c_D(M, \xi, \mathcal{F})$ and $c_A(M, \xi, \mathcal{F})$ are zero in $H_*(\widehat{CFD}(-M, \overline{\mathcal{Z}}))$ and $H_*(\widehat{CFA}(-M, \overline{\mathcal{Z}}))$, respectively.*

Given a pair of foliated contact three-manifolds $(M^L, \xi^L, \mathcal{F}^L)$ and $(M^R, \xi^R, \mathcal{F}^R)$ whose boundaries agree in an appropriate sense, there is a natural way to glue them to obtain a closed contact three-manifold (M, ξ). The contact invariants of the two foliated contact three-manifolds pair to recover the contact invariant of (M, ξ).

Theorem 1.3 [1, Theorem 2]. *The tensor product*

$$c_A(M^L, \xi^L, \mathcal{F}^L) \boxtimes c_D(M^R, \xi^R, \mathcal{F}^R)$$

recovers the contact invariant $c(M, \xi)$.

This paper offers a hands-on introduction to the bordered contact invariant, favoring geometric intuition over the formal proofs that may be found in [12] and [1]. We assume minimal background in contact geometry, so Section 2 focuses on understanding contact structures via characteristic foliations. Section 4 introduces multipointed bordered Floer homology as a special case of bordered sutured Floer homology, laying the groundwork for a simplified description of the construction of the bordered contact invariant. Section 3 discusses open books, reviewing the classical case for closed manifolds before introducing foliated open books for manifolds with boundary. After exploring some topological examples we define the contact structure supported by a foliated open book. We also define the technical condition "sorted" for a foliated open book and explain how it may be achieved by stabilization preserving the supported contact structure. We illustrate this in a carefully chosen example of a foliated open book for a neighborhood of an overtwisted disk. In Section 5 we describe how to construct a Heegaard diagram from a sorted foliated open book and define an associated generator that represents the contact invariant. Finally, in Section 6 we extend the earlier example to construct a Heegaard diagram for an overtwisted ball. A local computation, in conjunction with Theorem 1.3, then recovers the following vanishing result:

Corollary 1.4 [18]. *Let (M, ξ) be a closed contact three-manifold. If ξ is overtwisted, then $c(\xi) = 0$.*

Note that [7] establishes the vanishing of the sutured contact class for a neighborhood of an overtwisted disk. The TQFT gluing map from [6] then yields a sutured argument that $c(\xi)$ vanishes for overtwisted closed manifolds. Our local construction explicitly constructs the "contact compatible" layer needed in the sutured setting, giving a bordered counterpart to the argument.

2. Contact manifolds and surfaces

A key aim of this paper is to render more accessible a new invariant in bordered sutured Floer homology, but we'd like to start with a discussion of what this is an invariant *of*. Since its inception in the early 2000s, Heegaard Floer theory has given rise to invariants for a large range of mathematical objects; this one is distinguished not simply by its input, but also by the fact that the algebraic invariant behaves well under a natural topological operation.

2A. ***Contact structures.*** Recall from the introduction that a contact structure is a nowhere-integrable two-plane field. We will consider contact structures only on orientable three-manifolds, and we further require that contact structures be *cooriented*. That is, each contact plane is oriented, so there is a consistent choice of positive normal vector. It will be useful to reference a coordinate model, so we introduce the *standard contact structure* on $\mathbb{R}^3$, where the contact plane at each point is the kernel of the one-form $dz - ydx$. (A cooriented contact structure may always be described as the kernel of such a *contact form*.) In this case, the vector field ∂_z coorients the contact planes. We are primarily interested in studying contact manifolds up to *contactomorphism*, that is, up to diffeomorphism preserving the plane fields.

Like topological manifolds, contact manifolds are locally simple but globally complicated. The contact Darboux theorem states that every point in a contact three-manifold has a neighborhood contactomorphic to a neighborhood of the origin in the standard contact $\mathbb{R}^3$. In fact, some higher-dimensional substructures also have well-behaved neighborhoods. For example, a curve segment everywhere transverse to the contact planes has a neighborhood contactomorphic to a neighborhood of the z-axis in the standard $\mathbb{R}^3$. In this paper, we will focus on the kind of two-dimensional submanifolds with particularly nice neighborhoods: convex surfaces. We will characterize convex surfaces by considering certain foliations they carry, but since codimension-one foliations are so central to the rest of the paper, we briefly detour into some general discussion before returning specifically to foliations on convex surfaces.

2B. ***Foliations on surfaces.*** Throughout this article we will consider only oriented singular foliations whose singularities are isolated and are either elliptic (see bottom-right picture of Figure 2) or hyperbolic (see bottom-left picture of Figure 2). We denote the set of elliptic singularities by E, and the set of hyperbolic singularities by H. Unless explicitly noted, we assume that all regular leaves of the foliation compactify to oriented intervals. Elliptic points are either *sources*, in which case they are also called *positive* elliptic points, or *sinks*, which are also called *negative* elliptic points. At a four-pronged hyperbolic singularity, the two opposite prongs

oriented towards the hyperbolic point form the *stable separatrix*, while the two prongs oriented away from the hyperbolic point form the *unstable separatrix*. The topological type of these foliations can be described combinatorially by the embedded graph formed by the stable and unstable separatrices of the hyperbolic points.

The foliations appearing in this paper will have additional structure given by assigning a sign to each singular point. The signs of elliptic points have already been introduced, but the signs of hyperbolic points are not visible from the combinatorics of the foliation. (The sign of a hyperbolic point comes from the orientation of the surface and additional local data that depends on the source of the foliation; see Sections 2C and 3A.) A foliation with the properties above, together with the extra partition of $H = H_+ \sqcup H_-$, is called a *signed singular foliation*; in the following we refer to oriented signed singular foliations simply as foliations.

Given a foliation $\mathcal{F}$ on a surface F satisfying the hypotheses imposed above, we say that a multicurve $\Gamma \subset F$ is a *dividing set* if Γ is everywhere transverse to the leaves of $\mathcal{F}$ and separates F into two subsurfaces, each of which contains all the singularities of a fixed sign. With this structure in hand, we are ready to introduce the characteristic foliation on a surface in a contact manifold, which is the key to the local neighborhood theorem mentioned above. We introduce the aspects of this theory that we will need, and we recommend [15] for further reading on the topic.

2C. ***Convex surfaces.*** An oriented surface F embedded in a contact three-manifold (M, ξ) inherits a *characteristic foliation* from ξ. Intersecting the contact plane with the tangent plane at each point in the surface defines a line field, and the leaves of the characteristic foliation are the integral curves of these intersections. Characteristic foliations may be more general than the foliations described above, admitting leaves that are circles or even nonmanifolds. However, we will not consider any cases where these phenomena arise. The orientation of the leaves follows from the coorientation of the contact structure, while the signs of the singular points depend on whether the coorientation of the contact structure is a positive or negative normal for F.

A surface in a contact structure is *convex* if the contact structure is I-invariant in some product neighborhood; a key result states that a surface is convex if and only if its characteristic foliation admits a dividing curve Γ [4]. Remarkably, the local neighborhood of a convex surface is determined by the dividing set alone. The property of admitting a dividing curve (and hence, convexity) can be checked combinatorially, and in fact, convex surfaces are C^∞-generic [4]. Another important aspect of this equivalence is *Giroux's flexibility*, it describes the sense in which Γ captures the essential data of the contact structure in a neighborhood of F. Specifically, if Γ is a dividing set on $F \subset (M, \xi)$, then any foliation divided by Γ

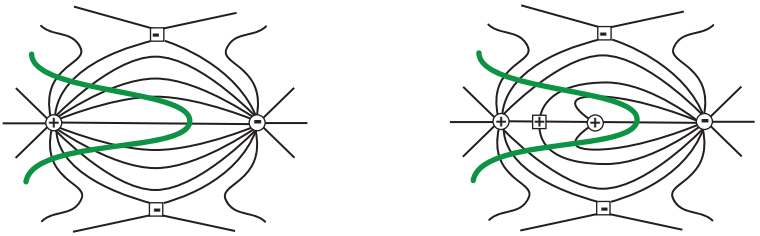


Figure 1. Local pictures of two characteristic foliations divided by the same curve Γ, shown in green. Circles are elliptic points and the squares are hyperbolic points.

can be realized as the characteristic foliation of some isotopic surface F' in a neighborhood of F. Thus, given a separating multicurve Γ on a surface F, one may choose any foliation divided by Γ and construct a compatible contact structure on $F \times I$. If we choose another foliation, then Giroux's flexibility implies this is the characteristic foliation on some surface inside this neighborhood, so our original neighborhood in fact contains the contact structure determined by this new foliation.

Characteristic foliations may exhibit many leaf types, but we will restrict attention to the cases where the hypotheses of Section 2B are satisfied; this is also a generic property. In addition, we will require that each signed singular foliation has no closed leaves or leaves connecting two hyperbolic points, and any such foliation will admit a dividing set, thus ensuring convexity. To see this, we introduce two graphs $G_\pm$ embedded into F and associated to $\mathcal{F}$. The vertices of G_+ are the positive elliptic points, and the edges between them are the stable separatrices of positive hyperbolic points. The graph G_- is analogously defined using the negative elliptic points and unstable separatrices. Observe that G_+ and G_- are disjoint and that the complement of their neighborhoods $N(G_+)$ and $N(G_-)$ has no singularities and is thus foliated by intervals. The *dividing curve* Γ of such a foliation is given by the oriented boundary of $N(G_+)$, which is isotopic through curves transverse to the foliation to $-\partial N(G_-)$. When a foliation admits a dividing set, it is unique up to isotopy, so we will often refer to "the" dividing set.

With the given restrictions on leaf types (i.e., only intervals and leaves containing a single hyperbolic point), the complement of the union of the separatrices is a collection of disks, each with one positive and one negative elliptic point on its boundary. The interior of each disk is foliated by an I-family of leaves from the positive to the negative elliptic point; this can be seen in Figure 1.

3. Foliated open books

We saw in Section 2 that the dividing set on a convex surface suffices to determine the contact structure in a neighborhood of the surface. Although the precise

information about the characteristic foliation is lost, enough data is retained to identify the relevant equivalence class. This theme is pervasive throughout contact geometry, with open books being one of the most notable illustrations. An open book decomposition of a contact manifold loses information about the specific contact structure, but with the benefit that the isotopy class of the contact structure is determined by a minimal amount of data. This economical encoding of the isotopy class was first studied in the contact setting by Thurston and Winkelnkemper and rose in prominence with the work of Giroux [5; 20]. After recalling the classical construction, we will describe the *foliated open books* first introduced in [12] as a new version of open books for contact manifolds with convex boundary. Although the definition of a foliated open book will require us to keep track of more data on the boundary than simply the dividing set, the payoff will be a more user-friendly set of gluing theorems than seen with previous types of open books.

An *abstract open book* for a closed three-manifold is a pair (S, h), where S is a surface with boundary and h an element of its mapping class group. This data suffices to construct an S-bundle over S^1, and after collapsing the boundary in a controlled way, yields a closed three-manifold. A foliated open book adapts this approach to the setting of a manifold with boundary. This time, the data consist of a sequence of $2k$ topologically distinct surfaces and the maps identifying one surface with the next. Analogously, this determines a manifold with foliated boundary.

3A. *Classical open books.* This section reviews the definition of an open book decomposition of a closed three-manifold, along with the notion of an *open book foliation* developed in [9].

Definition 3.1. An *abstract open book* is a pair (S, h) where S is a surface with boundary $\partial S = B$ and $h : S \to S$ is a diffeomorphism that preserves B pointwise.

An abstract open book determines a closed three-manifold M as follows. First, consider the product $S \times I$ and identify the points $(h(x), 0) \sim (x, 1)$ to form the mapping torus of h. Then collapse each component of the boundary $\partial S \times S^1$ to a circle via $(x, t) \sim (x, t')$ whenever $x \in \partial S$. The image of $\partial S \times S^1$ is an oriented link called the *binding* and again denoted by B, while the surfaces $S \times \{t\}$ become the *pages*. We will also make use of the function $\pi : M \setminus B \to S^1$ that sends each point on $S \times \{t\}$ to t.

The simplest example of an open book is given by setting $S = D^2$, so that h is necessarily isotopic to the identity. The pair (D^2, id) determines S^3; to see this, observe that $N(B)$ and $M \setminus N(B)$ give a genus-one Heegaard splitting with meridional curves on the two solid tori intersecting once. In fact, an open book determines not only a topological three-manifold, but actually a contact three-manifold, but this will be explored in the next section. For now, we consider further topological structure associated to an open book.

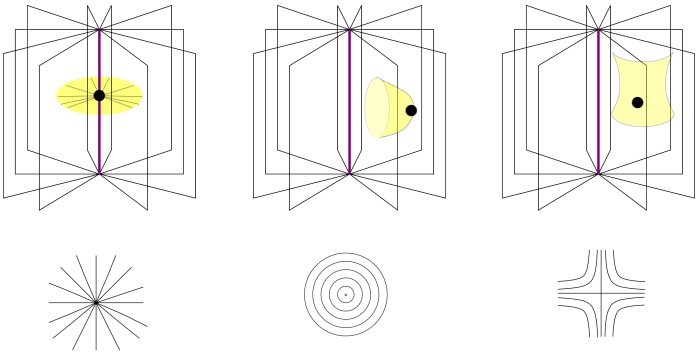


Figure 2. The intersection of F with the pages and binding (above) induces the singularity of $\mathcal{F}_\pi$ (below). Left: the foliation on a disk transverse to the binding has an elliptic point. Center: the foliation on a cup with one point tangent to a page has a center. Right: the foliation on a saddle with one point tangent to a page has a hyperbolic point.

Suppose that M is the closed three-manifold built from (S, h). Then the pages of S induce a foliation on a generic surface embedded in M. Assume that a surface F is transverse to the binding B, so that $E = B \cap F$ is finite. Additionally, assume that the restriction $\widetilde{\pi} = \pi|_F : F \setminus E \to S^1$ is an S^1-valued Morse function with only one critical point for any critical value. Then the *open book foliation* $\mathcal{F}_\pi$ on F is the foliation induced by the level sets of $\widetilde{\pi}$ together with the elliptic points E. Equivalently, the leaves of the foliation are the intersections of F with the pages of the open book. As seen in Figure 2, such a foliation may have three types of singularities: the points in E are elliptic points; the index 0 and 2 critical points of $\widetilde{\pi}$ are centers; and the index 1 critical points are hyperbolic points. Each level set of $\widetilde{\pi}$ has at most one critical point, and there are no leaves connecting hyperbolic points. Although closed leaves may arise, one may eliminate them via a bigger isotopy of the surface [9].

As above, we can associate signs to the elliptic points depending on whether the binding coorients F or not, whereas the sign of a critical point of $\widetilde{\pi}$ is given by the sign of $d\pi$ evaluated on the normal to F. Just as characteristic foliations on convex surfaces determine the nearby contact structure, open book foliations determine the open book decomposition near the surface.

3B. ***Foliated open books.*** Intuitively, a foliated open book is the structure on a manifold with boundary formed by cutting a classical open book along a surface with an open book foliation. We consider two examples of this sort before carefully stating a definition in parallel to Definition 3.1.

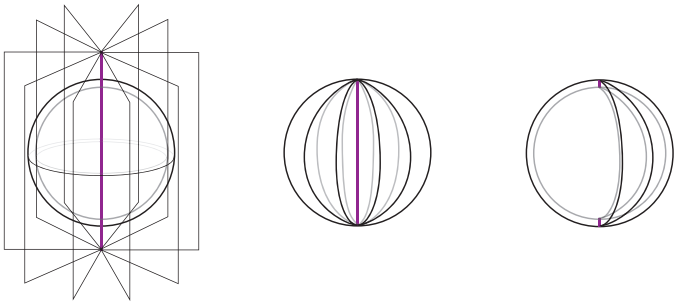


Figure 3. Cutting the open book (D^2, id) for S^3 (left) along a pair of parallel spheres yields a (pair of) foliated book(s) for the three-ball(s) (center) and a foliated open book for the thickened sphere (right, selected pages shown). On the boundary spheres of the resulting foliated open books, each leaf of the open book foliation is a line of longitude, and the only singularities are the two elliptic points at the poles.

Example 3.2. Consider the open book for S^3 described above with connected binding and disk pages. Choose a neighborhood of a point on the binding and cut S^3 along the boundary of this ball as shown in the center of Figure 3. Discarding the complement of this ball, one sees that it inherits a binding and pages from the original open book, and that the new boundary is naturally equipped with the foliation whose leaves are boundary intervals of the pages. This is the simplest possible foliated open book.

For an example that is one step more interesting, cut S^3 along a pair of parallel spheres to get a thickened sphere that intersects the binding in two intervals. The complement of these binding intervals is a union of rectangles.

We will see more interesting examples after the formal definition.

Definition 3.3 [12, Definition 3.12]. An *abstract foliated open book* is a tuple $(\{S_i\}_{i=0}^{2k}, h)$ where S_i is a surface with boundary $\partial S_i = B \cup A_i$[1] and corners at $E = B \cap A_i$ such that

(1) for all i, A_i is a finite union of intervals and B is a union of intervals or circles;

(2) the surface S_i is obtained from S_{i-1} by either

- attaching a 1-handle along two points on A_{i-1}, or
- cutting S_{i-1} along a properly embedded arc γ_i with endpoints in A_{i-1} and then smoothing.[2]

[1] By a slight abuse of notation we denote the "constant" part of the boundary of S_i by B for all i.

[2] The indices of γ_i in this paper are shifted compared to [12], where the cutting arcs were denoted by γ_{i-1}.

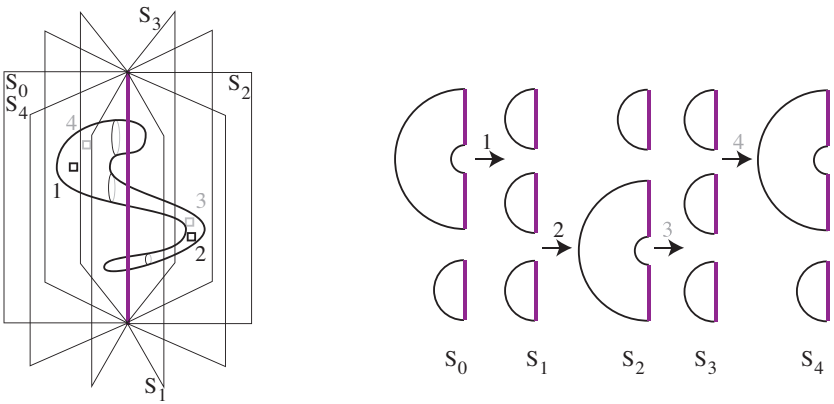


Figure 4. A different foliated open book for a ball cut from S^3.

Furthermore, $h : S_{2k} \to S_0$ is a diffeomorphism between cornered surfaces that preserves B pointwise.

We invite the reader to pause and compare Definitions 3.1 and 3.3. The latter has two levels of complexity not seen in the classical definition: first, the definition replaces a single surface S with a family of surfaces S_i of distinct topological type, and second, the boundary of each surface is partitioned into A_i-intervals and B-intervals or -circles. This second feature was seen already in Example 3.2: cutting each page in the open book for S^3 along its intersection with the sphere resulted in two new bigon pages each bounded by an A_i-interval and a B-interval.

Example 3.4. To illustrate the differences between classical and foliated open books, we consider a further example built by cutting the standard open book for S^3 along a separating S^2; see Figure 4. Here, the intersections between the indicated ball and the pages of the original open book are not all homeomorphic. The points on the embedded S^2 where the changes in topological type occur are labeled by squares on the figure; the right-hand side of the figure shows the distinct subsurfaces (the pages of the resulting abstract open book), labeled to match the (embedded) pages in the original open book.

We now take on the full complexity of Definition 3.3 and describe how to build a manifold from a sequence of pages of distinct topological types. Throughout, we will use subscript indices to distinguish topologically distinct page types, referring to these as "abstract pages" for convenience.

Any pair of consecutively indexed abstract pages S_i and S_{i+1} defines an elementary cobordism. We build an analogue of the mapping torus by concatenating these elementary cobordisms and gluing S_{2k} to S_0 by the map h. More precisely, each abstract page S_i yields a product $S_i \times \left(\frac{i}{2k}, \frac{i+1}{2k}\right)$, for $0 \leq i \leq 2k-1$, and consecutive products join smoothly along a singular page which is a surface with two points

on its boundary identified. (Since $h : S_{2k} \to S_0$ is a diffeomorphism, we need not assign a separate product to S_{2k}). After collapsing $B \times S^1$ to a multicurve called the *binding* and still denoted by B, the remaining boundary is decorated by the nonbinding boundaries of the pages. With the above decomposition in mind, define a function $\pi : M \setminus B \to S^1$ so that on each piece $S_i \times \left(\frac{i}{2k}, \frac{i+1}{2k}\right)$, the function is projection to the second coordinate; here S^1 is identified with $[0, 1]/(0 \sim 1)$. Below, we abuse notation a couple of times and write S_t for $\pi^{-1}(t)$.

This construction induces a singular foliation $\mathcal{F}$ on ∂M whose regular leaves are copies of A_i, oriented as the boundary of the pages, and whose singular leaves each contain a single four-pronged hyperbolic point. Equivalently, leaves are level sets of the restriction of the function π to ∂S. The elliptic points E and the hyperbolic points H each come with signs: each interval component of A_i is oriented from a positive elliptic point towards a negative one. Hyperbolic points associated to attaching a one-handle are negative, while hyperbolic points associated to cutting along an arc are positive; for an illustration of the latter, refer to Figure 10.

We denote the resulting smooth object by the triple $(M, \partial M, \mathcal{F})$ and call it a *manifold with foliated boundary*. We remember the identification of leaves with intervals on the boundary of abstract pages, and, in particular, the foliation has a distinguished union of 0-leaves, which are always regular. Because there are no closed leaves or saddle-saddle connections, we may use the signs of the singular points to associate a dividing set to the foliation: as seen in Section 2C, the boundary of a neighborhood of the positive separatrices of positive hyperbolic points is a dividing set, and this is unique up to isotopy transverse to the leaves. Note that a manifold with foliated boundary does not have an associated foliated open book structure; rather, it has a boundary foliation that is compatible with the existence of a foliated open book.

We conclude with one more topological example before turning attention to the relationship between open books and contact structures.

Example 3.5. For a final example in this section, we describe a process for promoting a nicely foliated surface F to a foliated open book $F \times I$ with the property that the open book foliation on each $F \times \{s\}$ is isotopic to the original foliation. This procedure is described in detail in [12, Section 4.2], but we summarize it here for later use in this article.

The open book decomposition near a surface F is completely determined by the open book foliation $\mathcal{F}_\pi$ on F [12, Corollary 4.6]. In the following, we describe this local structure by constructing a foliated open book for $F \times [-1, 1]$ that embeds into any other (foliated) open book inducing $\mathcal{F}_\pi$. Naively, one might try to cross the original surface with $[-1, 1]$ and take the pages to be the products of leaves with the interval. This works in the case of a foliation with only elliptic singularities, as in Example 3.2, but the process is more subtle in the case that the original foliation has hyperbolic points.

We first briefly describe the open book determined by $\mathcal{F}_\pi$ near F before using the foliation to construct its abstract pages. The binding of the open book is transverse to F, so we can assume it embeds as $E \times I$ in $F \times I$, oriented by $\partial/\partial s$ (respectively, $-\partial/\partial s$) for positive (negative) elliptic points. Recall that Γ denotes the dividing set for a signed foliation. Away from a neighborhood of $\Gamma \times I$, each page S_t is the union of the leaves $\widetilde{\pi}^{-1}(t \pm \epsilon s) \times \{s\} \subset F \times \{s\}$, where the sign depends on whether we are in F_+ or F_-, and ϵ is sufficiently small so that no page contains more than one hyperbolic point. We connect these across $\Gamma \times I$ by bands which twist to compensate for the shearing of leaves in opposite directions as $|s|$ increases. (Figure 8 in [12] provides local models for this construction near Γ and E.)

As noted above, when t is not near a singular point, this yields pages which are simply thickened copies of the original interval leaves; when F is closed, these are rectangular pages with two binding intervals separating a pair of leaves, one on each of $F \times \{\pm 1\}$. This is illustrated by the thickened sphere in Example 3.2.

To see what happens near a singular value t_0 for the original foliation, consider the page which contains the corresponding hyperbolic point on $F \times \{0\}$. The boundary of this page on $F \times \{-1\}$ is a copy of the $\widetilde{\pi}^{-1}(t_0 - \epsilon)$ leaf in which the saddle resolution has not yet happened, while the boundary of this page on $F \times \{1\}$ is a copy of the $\widetilde{\pi}^{-1}(t_0 + \epsilon)$ leaf where the saddle resolution has already occurred. This gives a recipe for writing down abstract pages: starting from the regular value 0, set $S_0 = \widetilde{\pi}^{-1}(0) \times I$. To form S_1, perform the first cut/add operation on the corresponding $F \times \{1\}$ edges of S_0; to form S_2, perform the corresponding add/cut operation on the $F \times \{-1\}$ edges of S_1. Note that S_2 can be thought of as $\widetilde{\pi}^{-1}(t_0 + \epsilon) \times I$, where t_0 is the first singular value encountered after 0. We can continue to obtain a pair of pages for each hyperbolic point in the same way. If the original foliation had n hyperbolic points, the new foliated open book will therefore have $2n + 1$ pages. Each even-indexed page is a thickened regular leaf, while odd-indexed pages interpolate between these. Finally, note that the monodromy h will always be trivial, as the first and last pages are simply unions of disks.

3C. ***Foliated open books and contact structures.*** With the topological constructions well in hand, we are ready to recall the compatibility between foliated open books and contact structures.

Definition 3.6. [12, Definition 3.7] The abstract foliated open book $(\{S_i\}, h)$ *supports* the contact structure ξ on $(M, \partial M, \mathcal{F})$ if

(1) TB is positively transverse to ξ;

(2) there exists a nowhere zero vector field everywhere transverse to the interior of each page and to ξ whose flow preserves ξ;

(3) there is a topological isotopy of ∂M taking $\mathcal{F}$ to the characteristic foliation $\mathcal{F}_\xi$ such that some Γ is a dividing set for each foliation throughout the isotopy.

We will often want to consider a manifold with foliated boundary $(M, \partial M, \mathcal{F})$ together with a contact structure ξ supported by a foliated open book inducing the boundary foliation; we call this a *foliated contact three-manifold* and denote it by the triple $(M, \xi, \mathcal{F})$.

As above, we may ignore the third condition to recover the classical definition of a contact structure supported by an open book decomposition of a closed manifold. If a three-manifold M has both an open book decomposition (B, π) and a contact structure ξ supported by this open book, then a sufficiently generic surface will carry both a characteristic foliation $\mathcal{F}_\pi$ and an open book foliation $\mathcal{F}_\xi$. A priori these foliations are unrelated, but if the open book foliation has no circle leaves, then the contact structure can be isotoped so that the characteristic foliation and the open book foliation have the same combinatorics and further, that the singular points agree [9]. This is the key observation that gives the boundary criteria for foliated open books.

In the examples produced by cutting an honest open book along a separating surface, observe that the open book foliations on the two new boundaries match, but with the signs of all singular points reversed. Conversely, any two foliated open books with matching, sign-reversed boundary foliations may be glued to produce a closed manifold with an open book structure. In fact, these cutting and gluing results respect the contact structures supported by each of the open books in the sense of Definition 3.6 [12, Theorems 6.1 and 6.2].

In the remainder of this section, we construct several additional foliated open books for specific contact manifolds. Example 3.7 constructs foliated open books for a pair of distinct contact structures on the three-ball. In this case, as in the examples above, the foliated open books are identified as submanifolds of an open book for a closed three-manifold. Finally, Example 3.8 is a specific instance of the procedure described in Example 3.5 above; we endeavor to provide a plausible construction here while referring the reader to [12] for the technical details.

Example 3.7. Different open book decompositions of a fixed topological manifold may determine different contact structures, and the same holds true in the case of foliated open books. In this example we consider a pair of open books for S^3, one of which supports the unique tight contact structure and the other of which supports an overtwisted contact structure. Cutting each of these along a separating S^2 yields foliated open books for tight and overtwisted balls.

Let $(A, h^{\pm})$ denote the open book for S^3 with annular pages and monodromy a single Dehn twist of the indicated sign. The binding of the associated open book decomposition is a positive (resp. negative) Hopf link, denoted by H^+ (resp. H^-). To picture this, consider the genus one Heegaard splitting of $S^3 = H_1 \cup_\partial (-H_2)$ into two solid tori where H^+ (resp. H^-) is embedded on the Heegaard torus as in Figure 5. Here $\pi^{-1}([0, \frac{1}{2}]) = H_1$ and $\pi^{-1}([\frac{1}{2}, 1]) = H_2$. The positive twist

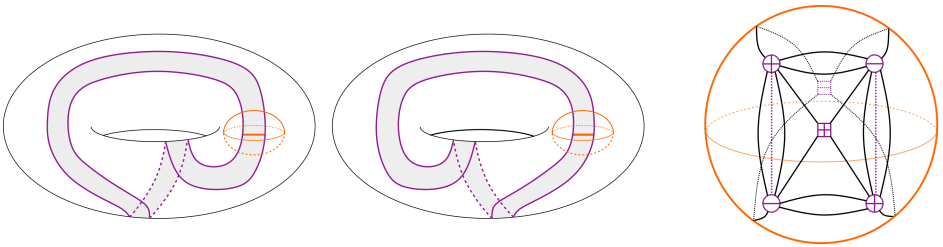


Figure 5. Left: H^+ embedded in ∂H_1 in a tight S^3. Center: H^- embedded in ∂H_1 in an overtwisted S^3. Right: the open book foliation on the boundary of a neighborhood of the spanning arc in the shaded annulus; as in Figure 1, elliptic points are drawn as circles, and hyperbolic ones as squares.

monodromy induces the tight contact structure on S^3, while the negative twist monodromy induces an overtwisted structure.

In each of these open books, consider the embedded S^2 bounding a neighborhood of the orange arc in $\pi^{-1}\left(\frac{1}{2}\right)$ shown in Figure 5. Discard this ball, leaving a pair of foliated open books for the complementary tight and overtwisted balls. The open book foliation on the boundary sphere has four elliptic points and two hyperbolic points as in the right-hand picture of Figure 5. The pages of these foliated open books are shown as the shaded subsurfaces in Figure 6. The Dehn twists from the original open book restrict to Dehn twists on the annular pages of the foliated open books.

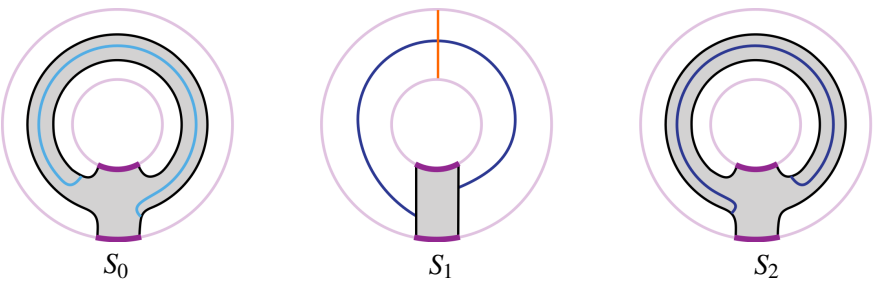


Figure 6. Removing a neighborhood of the orange arc from S^3 yields the shaded pages for a foliated open book for the three-ball. Each abstract page is shown embedded into an annular page for the open book $(A, h^{\pm})$, where $h^{\pm}$ is a positive (resp. negative) Dehn twist around the core of the annulus. These twists restrict to the cornered annulus S_2 as the monodromy for the foliated open books $(\{S_0, S_1, S_2\}, h^+)$ for a tight three-ball and $(\{S_0, S_1, S_2\}, h^-)$ for an overtwisted three-ball. The light and dark blue curves are sorting arcs, which are introduced in Section 3D.

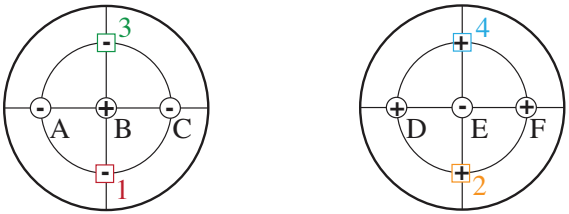


Figure 7. The left-hand picture shows the top of an overtwisted disk with transverse boundary; elliptic and hyperbolic points are again drawn as circles and squares, respectively. The right-hand picture labels the singularities of the characteristic foliation on the underside of the disk; the two points in the pairs (A, D), (B, E), and (C, F) coincide, but the sign of each singular point is reversed when viewed from the opposite side.

Example 3.8. In this example we construct a foliated open book for a ball supporting an overtwisted contact structure. This example is borrowed from [13], following the procedure summarized in Example 3.5. The motivation for including this initially opaque construction is that it will allow us to characterise any foliated open book for an overtwisted contact manifold in terms of a particular embedded foliated open book. To begin this process, we introduce a nonstandard definition of overtwistedness:

Definition 3.9. A contact manifold (M, ξ) is *overtwisted* if it contains an embedded disk whose boundary is everywhere transverse to ξ and whose characteristic foliation is as shown in Figure 7.

Overtwistedness is more commonly characterized in terms of the existence of an embedded disk with a different characteristic foliation, but it's a consequence of Giroux flexibility that the existence of a disk with this foliation is equivalent to the existence of disks with related characteristic foliations. We choose Definition 3.9 with a later application in mind. We now apply the construction sketched in Example 3.5 to build a foliated open book for a neighborhood of this disk; it follows that inside any overtwisted contact manifold, we may find an overtwisted ball that admits this foliated open book.

The existence of a transverse boundary requires us to slightly modify the construction, smoothing the boundary of pages associated to leaves that terminate on ∂F. Thus a regular leaf connecting an elliptic point e to ∂F will give rise to a bigon with one $e \times [-1, 1]$ component and one A_i component connecting $e \times \{\pm 1\}$.

We now apply this construction to the overtwisted disk shown in Figure 7, yielding an abstract foliated open book with five abstract pages. We set $t = 0$ to consist of the leaves where intervals connect (1) elliptic points A and B, and (2) the elliptic point C to the boundary. The first leaf becomes a rectangular page with two

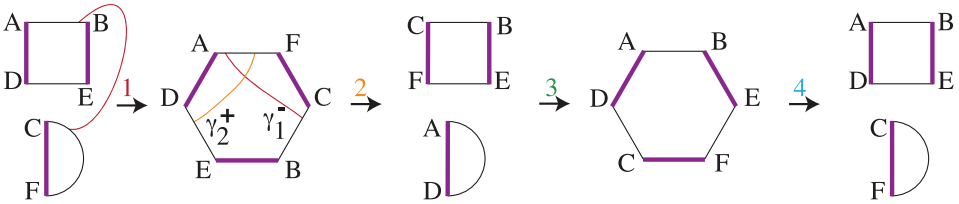


Figure 8. The pages of a foliated open book for a neighborhood of the disk in Figure 7. Each hyperbolic point in the original foliation induces a pair of hyperbolic points of opposite sign in the foliated open book. (The labeled arcs are explained in Example 3.11.)

boundary intervals, one connecting $A \times \{1\}$ and $B \times \{1\}$ and the other connecting $D \times \{-1\}$ with $E \times \{-1\}$. The leaf connecting C to the boundary becomes a bigon whose boundary interval connects $C \times \{1\}$ with $F \times \{-1\}$. See Figure 8. Around a positive elliptic point, t increases in the positive direction; following the procedure outlined in Example 3.5, the first hyperbolic singularity corresponds to adding a handle to connect these two pages. Figure 8 shows all the abstract pages of the resulting foliated open book.

Since each component of each page is a disk, there is a unique (up to isotopy) way to identify successive pages, and the foliated open book is completely determined by this data. One may also reconstruct the dividing set on the ball. One component encircles B on the "top" of the ball, while two further components bound an annulus containing D, F, and the two positive hyperbolic points on the "bottom" of the ball. In contrast, the foliated open book for the overtwisted ball constructed in Example 3.8 has a connected dividing set.

To illustrate how this ball might embed in an overtwisted contact manifold, we consider the open book for an overtwisted S^3 from Example 3.7. Recall that the pages are annuli and the monodromy is a left-handed Dehn twist. The top half of Figure 9 shows a ball intersecting two representative pages of this open book. The elliptic points are labeled to identify these subsurfaces with the first and third abstract pages from Figure 8; although we find it difficult to visualize further pages embedded in S^3, it is not difficult to embed the foliated open book pages in abstract pages, as shown below.

3D. ***Sorted foliated open books.*** Foliated open books will be our means to associating a Floer-theoretic invariant to a three-manifold with foliated boundary. However, in order to generate a multipointed Heegaard diagram, we will need to require the further technical condition that our foliated open book is *sorted*. Since the notation to verify this condition is somewhat involved, we pause to motivate it first.

The definition of a foliated open book requires successive pages to evolve by cutting or by gluing, but we may equivalently think of this as the condition that

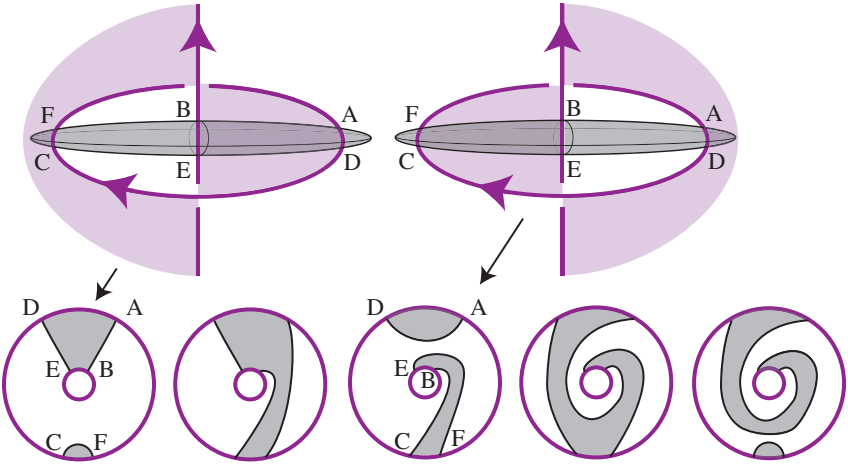


Figure 9. The foliated open book for the minimal overtwisted ball embeds in the simplest open book for an overtwisted S^3. The monodromy is a left-handed Dehn twist.

evolution is always by addition, but in either direction: either S_i is obtained from S_{i-1} by a one-handle addition or else S_i is obtained from S_{i+1} by a one-handle addition. One-handles associated to negative hyperbolic points are those already described in Definition 3.3 as "adds", while positive hyperbolic points correspond to adding a handle as the page index decreases. We will call a foliated open book *sorted* if a one-handle, after being added with respect to some direction (i.e., increasing or decreasing indices), persists for all subsequent pages with respect to that direction. See Figure 10.

Recall that the elliptic points $E = A_i \cap B$ partition as $E = E_+ \sqcup E_-$, where each interval is oriented from a point $e_+ \in E_+$ to a point $e_- \in E_-$. We impose the

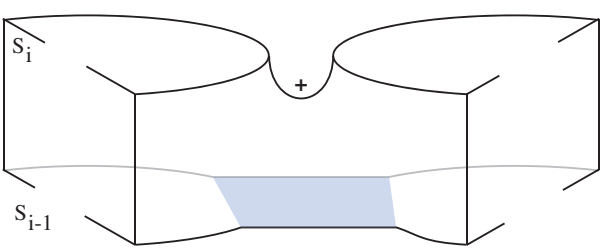


Figure 10. Here S_{i-1} is obtained from S_i by adding the shaded one-handle, inducing a positive hyperbolic singularity at the saddle point. The sorted condition requires that this handle persist for all S_j with $0 \leq j < i$. Note that the binding has been blown up as $B \times I$ for ease of viewing.

following conventions on the cutting and gluing arcs that govern how the pages evolve:

- If $S_i \to S_{i+1}$ cuts S_i along a properly embedded arc, the endpoints of the arc lie near the e_+ end of the intervals of A_i. We decorate S_i and all prior pages with a copy of the cutting arc and label these arcs as γ_i^+. If S_j is decorated with multiple cutting arcs near the same point e_+, the indices decrease with the orientation of A_j.
- If $S_i \to S_{i+1}$ adds a one-handle to S_i, the points of the attaching sphere separate any γ^+ endpoints from the e^- on the intervals of A_i. We decorate S_{i+1} and all subsequent pages with a copy of the cocore of the attached one-handle and label these arcs as γ_i^-. If S_j is decorated with multiple cocores near the same point e_-, the indices decrease with the orientation of A_j.

If we take the perspective that gluing is simply cutting in with the order of the indices reversed, then the second bullet point can be phrased in identical language to the first. Figure 11 illustrates these conventions in an example.

Definition 3.10. A foliated open book is *sorted* if the arcs $\gamma^- \cup \gamma^+$ are mutually disjoint on all the pages where they appear. We denote a sorted foliated open book by $(\{S_i\}_{i=0}^{2k}, h, \{\gamma_i^{\pm}\})$.

A foliated open book which is sorted has a ghost page: a minimal surface formed by cutting along all of the arcs. Although this surface may not actually coincide with any S_i in the foliated open book, it embeds as a subsurface in each abstract page. Remembering this may help in understanding the following notation-heavy definition.

Suppose $(\{S_i\}_{i=0}^{2k}, h, \{\gamma_i^{\pm}\})$ is a sorted foliated open book for foliated contact three-manifold $(M, \xi, \mathcal{F})$. On each page S_i, let P_i be the complement of a "cornered" neighborhood of $A_i \cup \left(\bigcup_{i<j} \gamma_j^+\right) \cup \left(\bigcup_{i \geq j} \gamma_j^-\right)$, with corners at E. This P_i is the ghost page and exists as a subsurface of each S_i. The copies of P_i may be identified via the flow of a vector field transverse to the pages, and we denote the composition of these identifications from $P_0 \subset S_0$ onto $P_{2k} \subset S_{2k}$ by ι.

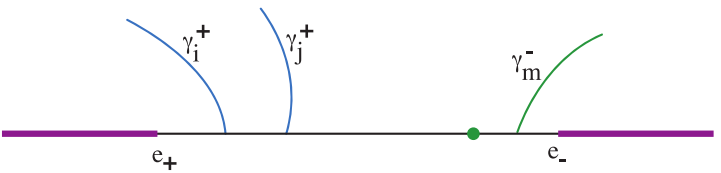


Figure 11. An indicative interval of A_n. Here $i > j \geq n > m$. The arcs γ_i^+ and γ_j^+ show arcs that will be cut along on higher-index pages. The bold dot indicates where a one-handle could be attached on some later page, while the arc γ_m^- is the cocore of a handle already attached.

3E. ***Sorting by stabilization.*** In this subsection we examine the operation of positive stabilization on a foliated open book and show how it can be used to render a nonsorted foliated open book sorted. The idea is straightforward: each stabilization adds a one-handle to every page of the foliated open book by taking a connect sum with a foliated open book for the standard tight S^3. For a simple example, we note that the foliated open book for a tight three-ball constructed in Example 3.7 is a positive stabilization of the foliated open book from Example 3.2.

The number of sorting arcs $\gamma^{\pm}$ is controlled by the foliation, and hence unchanged by stabilization. Repeating the process sufficiently many times gives the sorting arcs more space in the enlarged page to avoid each other. Of course, the arcs that guide the stabilization must be chosen carefully, and we explain how to do this below. The formal proof that this is always possible may be found in [12].

As shown in [12], stabilization may be understood as a concrete example of gluing two foliated open books. Choose an arc $(\gamma, \partial\gamma)$ embedded in a fixed page (S_t, B) of a foliated open book. A regular neighborhood of this arc may be chosen so that its boundary is a sphere whose signed singular foliation has two hyperbolic singularities. Choosing such neighborhoods in two separate foliated open books yields manifolds with matching foliated boundaries. Since we can only glue foliations where the singularities match, but with opposite signs, shifting the t coordinate by $\frac{1}{2}$ allows us to glue the two spheres to construct a foliated open book for the connect sum of the two original manifolds; the new pages are the Murasugi sum of the pages of the original foliated open books. If one of the manifolds was an open book with annular pages supporting the tight contact structure on S^3, then the contactomorphism type of the manifold is unchanged and we say that the foliated open book has been *positively stabilized*. The open book in Example 3.7 with positive Hopf twist binding is a stabilization of the elementary open book for S^3 from Example 3.2.

The description above applies with minor modification to all versions of open books, but a distinguishing feature of foliated open books is the nonhomogeneity of the pages. An arc on S_t may not persist to some later page $S_{t'}$, or $S_{t'}$ may have a nontrivial mapping class group even though S_t was a collection of disks. This highlights that there are two choices to be made when defining a stabilization of a foliated open book: which page, and which arc?

With a goal of removing intersections of the form $\gamma_i^+ \cap \gamma_j^-$, choose a regular page between the hyperbolic points h_i^+ and h_j^-. We will stabilize along an arc in this page so that as γ_i^+ rises up through the manifold, the subinterval that would collide with the descending γ_j^- picks of the monodromy of the foliated open book for S^3 and instead undergoes a Dehn twist around the core of the annular Murasugi summand of the page.

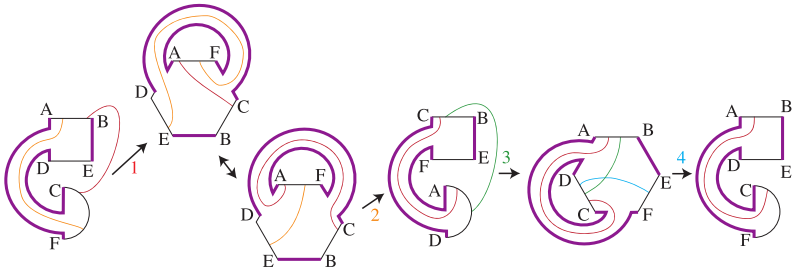


Figure 12. The stabilization of the foliated open book from Figure 8. Note that γ_3^+ and γ_4^- intersect on the new page S_3, so the foliated open book remains unsorted.

Example 3.11. Example 3.8 introduced a foliated open book for an overtwisted ball which embeds into any overtwisted contact manifold. Examining Figure 8 will show that it is not sorted, and this example will perform the sorting stabilizations.

The first hyperbolic singularity is negative and corresponds to adding a one-handle to S_0 as shown; on the second page S_1, the cocore of the one-handle is recorded as an arc γ_1^-. However, the second hyperbolic singularity is positive and corresponds to cutting the second page along the arc labeled γ_2^+ to get the third page. As shown in the figure, γ_1^- and γ_2^+ intersect.

To remove this obstruction to sortedness, choose a copy of S_1 and stabilize along an arc that crosses γ_2^+ and γ_1^- once. The result is shown in Figure 12. One can think of γ_1^- as undergoing a right-handed twist as it ascends or γ_2^+ as undergoing a left-handed twist as it descends, and since the two curves now avoid each other, we may proceed with increasing t until γ_3^- and γ_4^+ intersect on the new S_3 page.

To remove the intersection $\gamma_3^- \cap \gamma_4^+$, we analogously stabilize along an arc intersecting each of these curves once. Finally, a sorted foliated open book is seen in Figure 13. Since the gluing map is inherited from the original foliated open book, it remains translation in the page as drawn.

For any i, cutting along all the sorting arcs on S_i yields a pair of disks, the "ghost page" described in the previous section.

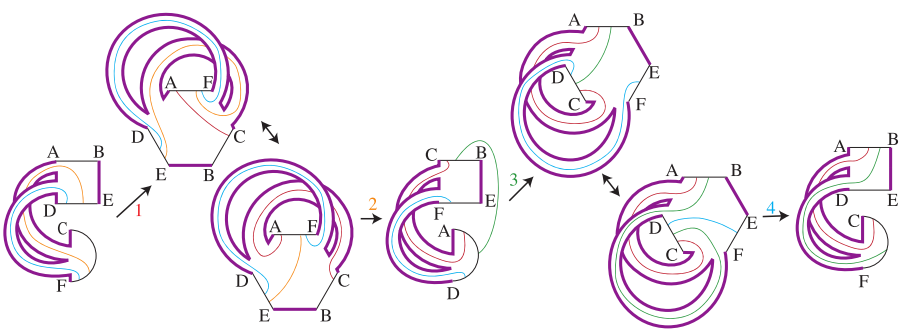


Figure 13. A sorted foliated open book for a neighborhood of an overtwisted disk, obtained from the foliated open book in Figure 8 by a sequence of two stabilizations.

4. Multipointed bordered Floer homology

As a stepping stone for defining link Floer homology, Ozsváth and Szabó defined a *multipointed* version of $\widehat{HF}$ denoted by $\widetilde{HF}$ [19]. This version is defined using Heegaard diagrams with multiple basepoints, and, given a closed, oriented three-manifold M, it is related to $\widehat{HF}(M)$ by the isomorphism

$$\widetilde{HF}(M,n) \cong \widehat{HF}(M) \otimes V^{n-1}.$$

Here, n is the number of basepoints and V is a 2-dimensional graded $\mathbb{Z}/2\mathbb{Z}$-vector space with generators in gradings 1 and 0; i.e., $V \cong H_*(S^1)$.

In [10], Juhász defines an extension of $\widehat{HF}$ for nonclosed three-manifolds whose boundary is *sutured*, called sutured Floer homology. Note that both $\widehat{HF}(M)$ and $\widetilde{HF}(M,n)$ are sutured Floer homologies of specific sutured manifolds corresponding to M. Specifically, let $M(n)$ be the sutured manifold obtained from M by removing n pairwise disjoint balls and adding as a suture one oriented simple closed curve on each resulting sphere boundary component. Then, we have

$$\widehat{HF}(M) \cong SFH(M(1))$$

while $\widetilde{HF}(M,n) \cong SFH(M(n))$.

Lipshitz, Ozsváth, and Thurston define bordered Floer homology as an extension of $\widehat{HF}$ for three-manifolds with parametrized boundary [14]. First, they associate a differential graded algebra $\mathcal{A}(\partial M)$ to the parametrization. Then, they define an $\mathcal{A}_\infty$-module, or *type A structure*, $\widehat{CFA}(M)$ over $\mathcal{A}(\partial M)$, or equivalently, a *type D structure* (roughly, a dg module) $\widehat{CFD}(M)$ over $\mathcal{A}(-\partial M)$. These invariants are constructed to satisfy a nice gluing formula which recovers $\widehat{HF}$. Specifically, if M is a closed three-manifold obtained by a gluing $M_1 \cup_\partial M_2$, then the derived tensor product $\widehat{CFA}(M_1) \widetilde{\otimes}_{\mathcal{A}(\partial M_1)} \widehat{CFD}(M_2)$ (which often has a smaller model denoted $\boxtimes$) is homotopy equivalent to $\widehat{CF}(M)$.

A generalization of bordered Floer homology, called bordered sutured Floer homology, was defined by Zarev [21]. It is an invariant of three-manifolds whose boundary is "part sutured, part parametrized". This invariant satisfies a gluing formula which recovers sutured Floer homology.

In this section, we introduce a *multipointed* theory for bordered Floer homology as a special case of bordered sutured Floer homology. First, we recall the definition of the boundary parametrization in bordered Floer homology. Let M be a three-manifold with boundary of genus k. A parametrization for ∂M consists of a disk $D \subset \partial M$; a basepoint $z \in \partial D$; and $2k$ pairwise disjoint properly embedded arcs $\amalg_{i=1}^{2k} \delta_i$ in $\partial M \setminus \mathrm{Int}(D)$ such that $M \setminus \left(D \cup \amalg_{i=1}^{2k} \delta_i\right)$ is an open disk. The parametrization data is recorded by a pointed matched circle $\mathcal{Z} = (Z, a, m)$, where $Z = \partial D$ with $z \in Z$, $a = \partial(\amalg_{i=1}^{2k} \delta_i)$ is a union of $4k$ points on Z, and M is a matching on a that pairs endpoints of the same arc δ_i.

Definition 4.1. A *pointed matched multicircle* is a triple $\mathcal{Z} = (Z, a, m)$ where $Z = \amalg_{i=1}^{n} Z_i$ is a union of n circles with a basepoint z_i on each Z_i, $a \subset Z$ is a set of an even number of points, and $m : a \to a$ is a matching. Given a three-manifold M with boundary of genus k, a (*multipointed*) *parametrization* of ∂M is a pointed matched multicircle $\mathcal{Z}$ with $|a| = 4n + 4k - 4$, along with an embedding of Z and of pairwise disjoint arcs $\delta = \amalg_{i=1}^{2n+2k-2} \delta_i$ into M, satisfying the following:

(1) the image of each Z_i bounds a disk D_i in ∂M whose interior is disjoint from the arcs δ_i for all i;

(2) $\partial \delta = a$ and each $\partial \delta_i$ is a pair of points matched by m;

(3) $\partial M \setminus \left(\left(\amalg_{i=1}^{n} D_i \right) \cup \left(\amalg_{i=1}^{2n+2k-2} \delta_i \right) \right)$ is the union of n open disks such that each disk contains exactly one of the marked points z_i for $i = 1, \ldots, n$.

We call the three-manifold with multipointed parametrized boundary a *bordered manifold*, as in [14], and denote it by $(M, \mathcal{Z})$, omitting from the notation the implicit data of how the arcs δ_i are embedded on ∂M.

A three-manifold with multipointed parametrized boundary $(M, \mathcal{Z})$ can be reinterpreted as a bordered sutured manifold $(M, \Gamma, \mathcal{Z}^\circ)$ where $\amalg_{i=1}^{n} D_i$ is the sutured part while its complement is the parametrized part, and $\mathcal{Z}^\circ$ is the arc diagram obtained from $\mathcal{Z}$ by removing neighborhoods of the basepoints. Thus, Zarev's construction associates a type A structure $\widehat{BSA}(M, \Gamma, \mathcal{Z}^\circ)$ over $\mathcal{A}(\mathcal{Z}) := \mathcal{A}(\mathcal{Z}^\circ)$, or equivalently a type D structure $\widehat{BSD}(M, \Gamma, \mathcal{Z}^\circ)$ over $\mathcal{A}(-\mathcal{Z})$. The construction uses a Heegaard diagram presentation $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathcal{Z}^\circ)$ for the bordered sutured manifold. The arc diagram $\mathcal{Z}^\circ$ is embedded on $\partial \mathcal{H}$ so that there is one interval on each component of $\partial \mathcal{H}$. The structures $\widehat{BSA}$ and $\widehat{BSD}$ are generated by certain sets of intersection points in $\boldsymbol{\alpha} \cap \boldsymbol{\beta}$ on the Heegaard diagram and they have structure maps defined by counting certain holomorphic curves in $\Sigma \times I \times \mathbb{R}$ whose projection onto Σ avoids the regions of $\Sigma \setminus (\boldsymbol{\alpha} \cup \boldsymbol{\beta})$ containing $\partial \mathcal{H} \setminus \mathcal{Z}^\circ$.

The embedding of $\mathcal{Z}^\circ$ on $\partial \mathcal{H}$ can be extended to an identification of $\mathcal{Z}$ with $\partial \mathcal{H}$, by reinserting the basepoints, one in each component of $\partial \mathcal{H} \setminus \mathcal{Z}^\circ$. The result is a *multipointed bordered Heegaard diagram* for $(M, \mathcal{Z})$. Since there is no loss of information when moving from one perspective to the other, we denote $\mathcal{Z}^\circ$ simply by $\mathcal{Z}$ in this paper. We will denote the structures $\widehat{BSA}(M, \Gamma, \mathcal{Z}^\circ)$ and $\widehat{BSD}(M, \Gamma, \mathcal{Z}^\circ)$ by $\widetilde{CFA}(M, \mathcal{Z})$ and $\widetilde{CFD}(M, \mathcal{Z})$, respectively. Explicitly, given a multipointed bordered Heegaard diagram, these structures are defined by counting the "usual" holomorphic curves; the condition of "avoiding the basepoints" is equivalent to "avoiding $\partial \mathcal{H} \setminus \mathcal{Z}^\circ$". The gluing formula for bordered sutured Floer homology implies that if the closed three-manifold M with multiple basepoints is obtained by gluing multipointed bordered three-manifolds $M_1 \cup_\partial M_2$, with M_1 parametrized by $\mathcal{Z}$ and M_2 by $-\mathcal{Z}$, then $\widetilde{CFA}(M_1) \boxtimes_{\mathcal{A}(\mathcal{Z})} \widetilde{CFD}(M_2)$ is homotopy equivalent to $\widetilde{CF}(M, n)$.

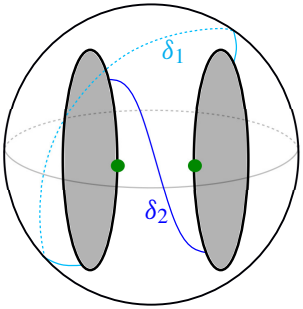


Figure 14. The bordered three-manifold associated to the foliated ball that is the neighborhood of the orange arc from Example 3.7. The two grey disks make up D, their boundary is Z, the two arcs δ_i are drawn in blue and labeled on the figure, and the basepoints are drawn in green.

5. The bordered contact invariant

Let $(M, \xi, \mathcal{F})$ be a foliated contact three-manifold. In [1], a sorted foliated open book for $(M, \xi, \mathcal{F})$ was used to construct a Heegaard diagram for an associated bordered sutured manifold $(M, \Gamma, \mathcal{Z})$, along with a preferred generator of the diagram. The homotopy equivalence class of this generator in the resulting bordered sutured Floer homology is an invariant of the foliated contact three-manifold [1, Theorem 1]. In particular, the class is independent of the choice of open book. We recall the construction next, slightly rephrasing to use multipointed bordered Floer homology, and we work out a small example.

As explained in Section 4, we can convert the data of a bordered sutured manifold $(M, \Gamma, \mathcal{Z})$ to multipointed bordered data for a simpler perspective. We describe the parametrization on the boundary of the resulting bordered manifold $(M, \mathcal{Z})$ directly below.

We use the foliation to define a natural parametrization of ∂M via a pointed matched multicircle $\mathcal{Z} = (Z, a, m)$. Recall that the data of the foliation remembers the page index associated to each leaf, and in particular, that there is a distinguished union of regular leaves denoted by A_0. Let $D \subset \partial M$ be a closed neighborhood of A_0, and let $Z = \partial D$. Note that D is a union of n disks, where $2n$ is the number of elliptic points in the foliation. Let δ_i be a subarc of the positive (resp. negative) separatrix for h_i^+ (resp. h_i^-) that lies in $\partial M \setminus (\text{int} D)$. Define $a \subset Z$ to be the set of points that are the boundaries of δ_i and let m be the matching induced on the points in a by δ_i. For each component of Z, mark a basepoint with a smallest possible $(0, 2\pi)$-coordinate. See Figure 14 for an example. It is easy to check that $\mathcal{Z} = (Z, a, m)$ together with the embedding of the arcs δ_i parametrizes ∂M.

Now, fix an abstract sorted foliated open book $(\{S_i\}_{i=0}^{2k}, h, \{\gamma_i^{\pm}\})$ for the foliated contact three-manifold $(M, \xi, \mathcal{F})$. The sortedness condition ensures that the first

page of the open book together with its (indexed) γ_i^+ arcs, the last page together with its (indexed) γ_j^- arcs, and the monodromy h fully describe the manifold. In fact, the union of the first and last page naturally describes a (cornered) handlebody decomposition for M. Using the data of $(\{S_i\}, h, \{\gamma_i^\pm\})$, we describe a multipointed bordered Heegaard diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathcal{Z})$ for this handlebody decomposition, along with a preferred generator. We outline the construction below; see [1, Section 3].

Let g_i be the genus of S_i and let n_i be the number of boundary components of S_i. Recall that the boundary of the cornered surface S_i is $B \cup A_i$, where B is a union of circles and arcs, and A_i is a union of intervals only.

We let $\Sigma = S_0 \cup_B -S_0$. In order to distinguish the two copies, we will write

$$\Sigma = S_\epsilon \cup_B -S_0,$$

but we emphasize that S_ϵ can be identified with S_0. The surface Σ has genus $2g_0 + n_0 - 1$ and $|A_0|$ boundary components.

For $i \in H_-$, consider the S_{2k} copies of the sorting arcs γ_i^-, and let β_i^- equal $-h(\gamma_i^-)$ on $-S_0$. For $i \in H_+$, consider the S_ϵ copies of the sorting arcs γ_i^+. The endpoints of γ_i^+ lie near the E_+ end of intervals of A_ϵ. Isotope the arcs $\{\gamma_i^+\}$ (simultaneously, to preserve disjointness) near the endpoints along $-\partial\Sigma$ until they all lie in $I_+ \subset A_0$; the isotopy stops after crossing E_+ and before encountering $\cup_{j\in H_-} -h(\gamma_j^-) \subset -S_0$. Call the resulting arcs β_i^+. Define a set of arcs $\boldsymbol{\beta}^a = \{\beta_1^a, \dots, \beta_{2k}^a\}$ by

$$\beta_i^a = \begin{cases} \beta_i^+ & \text{if } i \in H_+, \\ \beta_i^- & \text{if } i \in H_-. \end{cases}$$

As in [1], we use the notation β_i^a or $\beta_i^\pm$ if $i \in H_\pm$ interchangeably.

Let $\boldsymbol{b} = \{b_1, \dots, b_{2g_0+n_0+|A_0|-k-2}\}$ be a set of cutting arcs for $P_\epsilon \subset S_\epsilon$ disjoint from $\boldsymbol{\beta}^a$ and with endpoints on B, so that each connected component of $S_\epsilon \setminus (\boldsymbol{b} \cup \boldsymbol{\beta}^a)$ is a disk with exactly one interval of A_ϵ on its boundary. (In [1], we show this can always be achieved.) In other words, $\boldsymbol{b}$ is a basis for $H_1(P_\epsilon, B)$. Recalling the identification $S_\epsilon = S_0$, we may push $b_i \subset S_0$ through M to lie on S_0 again and define

$$\beta_i = b_i \cup -h \circ \iota(b_i) \subset S_\epsilon \cup_B -S_0,$$

where ι is the identification of P_0 with P_{2k} from Section 3D. Write

$$\boldsymbol{\beta}^c = \{\beta_1, \dots, \beta_{2g_0+n_0+|A_0|-k-2}\}.$$

For each cutting arc $b_i \in \boldsymbol{b}$ on S_ϵ, let a_i be an isotopic curve formed by pushing the endpoints negatively along the boundary so that a_i and b_i intersect once transversely. Similarly, for each arc $b_i^+ := S_\epsilon \cap \beta_j^+$, let $\tilde{a}_j$ be an isotopic curve formed by pushing

the endpoints negatively along the boundary so that $\tilde{a}_j$ and b_i^+ (and equivalently $\tilde{a}_j$ and β_j^+) intersect once transversely. We "double" each of these arcs to form the α-circles which define the handlebody $S_0 \times [0, \epsilon]$. Namely, define

$$\alpha_i = a_i \cup -a_i \subset S_\epsilon \cup_B -S_0, \qquad \widetilde{\alpha}_j = \tilde{a}_j \cup -\tilde{a}_j \subset S_\epsilon \cup_B -S_0,$$

and write $\boldsymbol{\alpha}^c = \{\widetilde{\alpha}_i\}_{i \in H_+} \cup \{\alpha_1, \ldots, \alpha_{2g_0+n_0+|A_0|-k-2}\}$. Place a basepoint on each interval of $A_\epsilon \subset S_\epsilon \subset \Sigma$. Write

$$\mathbf{z} = \{z_1, \ldots, z_{|A_\epsilon|}\}$$

for the set of basepoints.

We say that a multipointed bordered Heegaard diagram $\mathcal{H} = (\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathcal{Z})$ constructed as above is *adapted* to the sorted abstract foliated open book $(\{S_i\}, h, \{\gamma_i^\pm\})$ and to the corresponding foliated contact three-manifold $(M, \xi, \mathcal{F})$.

Let $\mathcal{H}$ be a multipointed bordered Heegaard diagram adapted to $(\{S_i\}, h, \{\gamma_i^\pm\})$. In [1], we show that any such diagram is admissible. (In fact, in [1], neighborhoods of basepoints are drilled out to obtain a bordered sutured diagram for a certain bordered sutured manifold naturally associated to $(M, \xi, \mathcal{F})$, but we suppress this discussion here.) Using the notation introduced above, define

$$\mathbf{x} = \{x_1, \ldots, x_{2g_0+n_0+|A_0|-k-2}\} \cup \{x_i^+ \mid i \in H_+\}$$

to be the set of unique intersection points

$$x_i = a_i \cap b_i \in S_\epsilon \subset \Sigma,$$
$$x_i^+ = \tilde{a}_i \cap b_i^+ \in S_\epsilon \quad \text{if } i \in H_+.$$

We will use $\mathbf{x}$ to define two contact invariants in multipointed bordered Floer homology.

By [1, Proposition 3.4] and [22, Section 3.4], the diagram $\overline{\mathcal{H}} = (\Sigma, \boldsymbol{\beta}, \boldsymbol{\alpha}, \overline{\mathcal{Z}})$ obtained by exchanging the roles of the two sets of curves and formally replacing the arc diagram $\mathcal{Z}$ of β-type (which is to say, parametrized by arcs which are part of the second set of curves) with the identical arc diagram $\overline{\mathcal{Z}}$ of α-type (parametrized by arcs which are part of the first set of curves) is a multipointed bordered diagram for $(-M, \overline{\mathcal{Z}})$. Write $\overline{\mathcal{Z}} = (Z, a, m)$. We have the following proposition.

Proposition 5.1 [1, Proposition 3.5]. *The above* $\mathbf{x}$ *gives a well-defined generator*

$$\mathbf{x}_D := \mathbf{x} \in \widetilde{CFD}(\overline{\mathcal{H}})$$

with $I_D(\mathbf{x}) = I(H_-)$ *and* $\delta^1(\mathbf{x}_D) = 0$, *and a well-defined generator*

$$\mathbf{x}_A := \mathbf{x} \in \widetilde{CFA}(\overline{\mathcal{H}})$$

with $I_A(\mathbf{x}_A) = I(H_+)$ *and* $m_{i+1}(\mathbf{x}_A, a(\boldsymbol{\rho}_1), \ldots, a(\boldsymbol{\rho}_i)) = 0$ *for all* $i \geq 0$ *and all sets of Reeb chords* $\boldsymbol{\rho}_j$ *in* (Z, a).

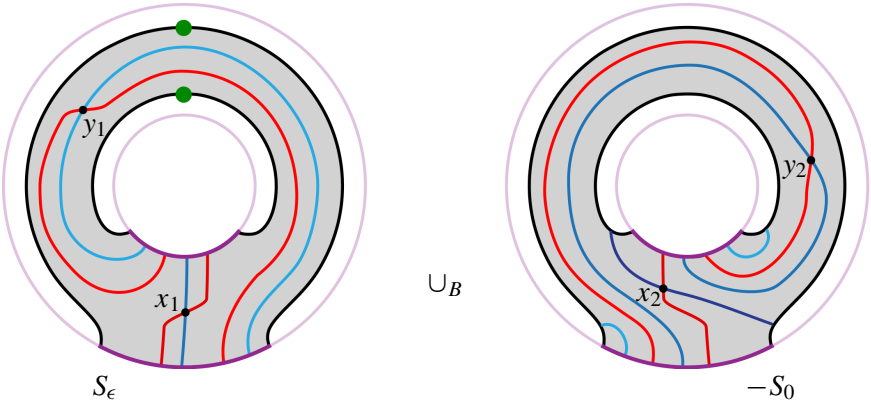


Figure 15. The Heegaard diagram for the sorted foliated open book $(\{S_0, S_1, S_2\}, h^+)$ from Figure 6. The monodromy h^+ is a positive Dehn twist, so the images $\beta_2^- = -h^+(\gamma_2^-)$ and $-h^+ \circ \iota(b_1)$ are the dark and medium-dark blue curves on $-S_0$, respectively. Intersection points are labeled differently from the above definition, for convenience. The contact generator $\mathbf{x}$ is the pair $\{x_1, y_1\}$, or x_1y_1 for short.

Example 5.2. We illustrate the construction outlined above using the (sorted) foliated open book in Figure 6. Recall that the three pages depicted in Figure 6 in fact can be used to construct different foliated open books, depending on the choice of monodromy τ^n, for $n \in \mathbb{Z}$, where τ is a positive Dehn twist along the core of the annular page S_2.

First, consider the foliated open book with pages depicted in Figure 6 and monodromy τ (this was denoted by h^+ in Figure 6). Figure 15 shows the associated Heegaard diagram $\mathcal{H}^+$. We label the intersection points in the Heegaard diagram $\mathcal{H}^+$ by x_1, x_2, y_1, and y_2 as in Figure 15. The diagram has two generators, x_1y_1 and x_2y_2, where x_1y_1 is the special generator $\mathbf{x}$ defined above. Let ρ_1 and ρ_2 be the algebra elements in $\mathcal{A}(\partial\overline{\mathcal{H}^+})$ corresponding to the Reeb chords on the inside and outside boundary components of the Heegaard diagram, respectively, as seen on Figure 6. The type D structure $\widetilde{CFD}(\overline{\mathcal{H}^+})$ is generated by x_1y_1 and x_2y_2, and has structure maps

$$\delta^1(x_1y_1) = 0,$$
$$\delta^1(x_2y_2) = (\rho_1 + \rho_2) \otimes x_1y_1.$$

The contact class $c_D(B^3, \xi, \mathcal{F})$ is the homotopy equivalence class of x_1y_1.

Next, consider the foliated open book with pages depicted in Figure 6 and monodromy τ^{-1} (which was denoted h^- in Figure 6). Figure 16 shows the associated Heegaard diagram $\mathcal{H}^-$. We label the intersection points in $\mathcal{H}^-$ by x_1', x_2', x_3', x_4', y_1', y_2', y_3', and y_4' as in Figure 16. Let $\rho_1, \rho_2 \in \mathcal{A}(\partial\overline{\mathcal{H}^-})$ be as in the previous example.

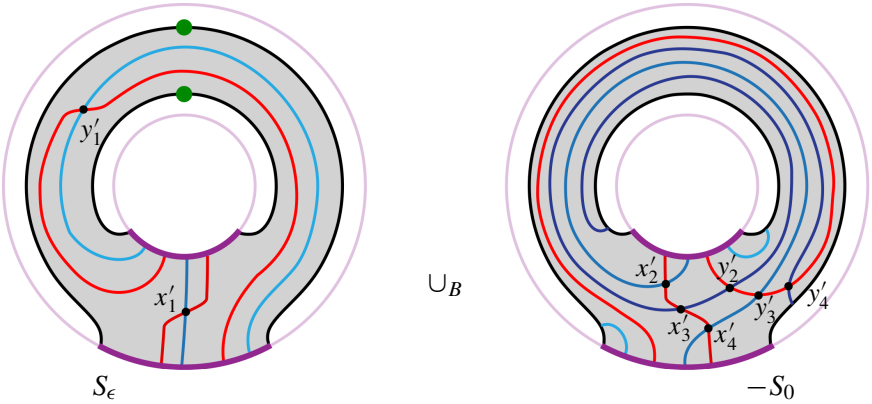


Figure 16. The Heegaard diagram for the sorted foliated open book $(\{S_0, S_1, S_2\}, h^-)$ from Figure 6. The monodromy h^- is a negative Dehn twist, so the images $\beta_2^- = -h^-(\gamma_2^-)$ and $-h^- \circ \iota(b_1)$ are the dark and medium-dark blue curves on $-S_0$, respectively. The contact generator $\mathbf{x}$ is $x_1' y_1'$.

The type D structure $\widetilde{CFD}(\overline{\mathcal{H}^-})$ is generated by $x_1'y_1'$, $x_1'y_2'$, $x_1'y_4'$, $x_2'y_1'$, $x_2'y_2'$, $x_2'y_4'$, $x_3'y_3'$, $x_4'y_1'$, $x_4'y_2'$, and $x_4'y_4'$, and has structure maps

$$\begin{aligned}
\delta^1(x_1'y_1') &= 0,\\
\delta^1(x_1'y_2') &= \rho_1 \otimes x_1'y_1',\\
\delta^1(x_1'y_4') &= \rho_2 \otimes x_1'y_1',\\
\delta^1(x_2'y_1') &= I \otimes x_1'y_1',\\
\delta^1(x_2'y_2') &= \rho_1 \otimes x_2'y_1' + I \otimes x_1'y_2',\\
\delta^1(x_2'y_4') &= I \otimes x_1'y_4' + \rho_2 \otimes x_2'y_1',\\
\delta^1(x_3'y_3') &= 0,\\
\delta^1(x_4'y_1') &= I \otimes x_1'y_1',\\
\delta^1(x_4'y_2') &= I \otimes x_1'y_2' + I \otimes x_3'y_3',\\
\delta^1(x_4'y_4') &= I \otimes x_1'y_4' + \rho_2 \otimes x_4'y_1'.
\end{aligned}$$

In particular, $\delta^1(x_2'y_1') = I \otimes x_1'y_1'$ implies that there is a type D homotopy equivalence from $\widetilde{CFD}(\overline{\mathcal{H}^-})$ to an equivalent structure, carrying $x_1'y_1'$ to zero.

6. Vanishing of the contact class for overtwisted structures: a local argument

In this section, we illustrate the power of invariants compatible with cut-and-paste constructions by providing a local argument that the contact class $c(\xi)$ for closed contact manifolds vanishes if the contact structure is overtwisted.

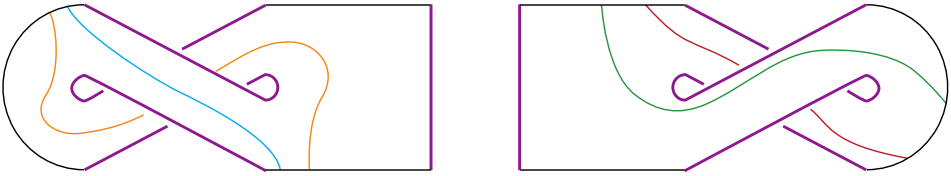


Figure 17. The first page (to the left) and the mirror of the last page (to the right) of the sorted foliated open book in Figure 13.

We begin by showing that the bordered contact invariant vanishes for a neighborhood of an overtwisted disk. Specifically, we consider the foliated open book constructed in [13] for a three-ball neighborhood $(B^3, \xi_{\mathrm{OT}}, \mathcal{F}_{\mathrm{OT}})$. In Example 3.8, we stabilized the foliated open book from [13] to a sorted one. We now construct the Heegaard diagram $\mathcal{H}$ associated to the resulting sorted foliated open book from Figure 13. For convenience, in Figure 17 we display again the pages S_0 and $-S_4$, along with the sorting arcs decorations.

Figure 18 shows the associated Heegaard diagram $\mathcal{H}$.

The generator $x_1y_1w_1 \in \widetilde{CFD}(\overline{\mathcal{H}})$ represents the contact class. We claim that there is a unique holomorphic curve that avoids the basepoints and is asymptotic to $x_1y_1w_4$ at $-\infty$, and this curve ends at $x_1y_1w_1$.

Indeed, x_1 and y_1 cannot be starting moving coordinates for a holomorphic curve; the only nonbasepointed regions at these intersection points are the thin strips supported on the S_ϵ part of the diagram, but the orientation on these strips is *into* x_1 and y_1. So any holomorphic curve starting from $x_1y_1w_4$ must only have w_4 as a moving coordinate. A curve that hits the boundary of the Heegaard diagram would need to have a moving coordinate on a β-arc. Since w_4 is on a β-circle, all holomorphic curves starting from $x_1y_1w_4$ project to the interior of the diagram. Thus, any such curve with a single moving coordinate projects to an immersed bigon. By counting local coefficients, the yellow bigon from $x_1y_1w_4$ to $x_1y_1w_1$ in Figure 18 represents the unique such curve.

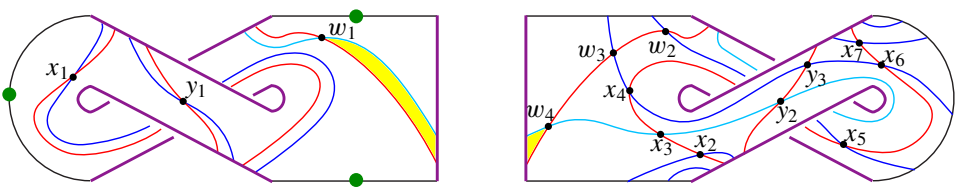


Figure 18. The Heegaard diagram for the sorted foliated open book in Figure 13. The monodromy h is the identity, so the images $\beta_i^- = -h(\gamma_i^-)$ are simply the sorting arcs γ_i^- on $-S_0$. Intersection points are labeled differently from the above definition, for convenience. The contact generator $\mathbf{x}$ is the triple $\{x_1, y_1, w_1\}$, or $x_1y_1w_1$ for short.

Thus, considering $\widetilde{CFD}(\overline{\mathcal{H}})$, we have $\delta^1(x_1y_1w_4) = I \otimes x_1y_1w_1$. Or, if one prefers to consider $\widetilde{CFA}(\overline{\mathcal{H}})$, we have $m_1(x_1y_1w_4) = x_1y_1w_1$, whereas higher products m_i vanish on $x_1y_1w_4$. It follows that there is a type D (resp. type A) homotopy equivalence from $\widetilde{CFD}(\overline{\mathcal{H}})$ (resp. $\widetilde{CFA}(\overline{\mathcal{H}})$) to an equivalent structure, carrying $x_1y_1w_1$ to zero.

Recall from the introduction that we claimed the Ozsváth–Szabó vanishing result for overtwisted contact manifolds can be recovered from gluing properties of the bordered contact invariant. In fact, the necessary technical results have already been established, and we conclude by assembling them into the promised proof.

Proof of Corollary 1.4. Suppose (M, ξ) is a closed overtwisted three-manifold. As discussed in Section 3C, (M, ξ) contains an overtwisted disk whose neighborhood is contactomorphic to the contact three-ball $(B^3, \xi_{\mathrm{OT}}, \mathcal{F}_{\mathrm{OT}})$ studied in Example 3.8. Thus, (M, ξ) decomposes as the union of two foliated contact three-manifolds, one of which is $(B^3, \xi_{\mathrm{OT}}, \mathcal{F}_{\mathrm{OT}})$.The computation, above, together with Theorem 1.3 and functoriality for $\boxtimes$, implies that $c(\xi) = 0$. □

Acknowledgements

We are grateful to BIRS for hosting the workshop *Interactions of gauge theory with contact and symplectic topology in dimensions* 3 *and* 4. We also appreciate the helpful feedback provided by the referee on an earlier version of this paper.

References

[1] A. Alishahi, V. Földvári, K. Hendricks, J. Licata, I. Petkova, and V. Vértesi, "Bordered Floer homology and contact structures", preprint, 2020. arXiv 2011.08672

[2] P. Ghiggini, "Strongly fillable contact 3-manifolds without Stein fillings", *Geom. Topol.* **9** (2005), 1677–1687. MR Zbl

[3] P. Ghiggini, "Knot Floer homology detects genus-one fibred knots", *Amer. J. Math.* **130**:5 (2008), 1151–1169. MR Zbl

[4] E. Giroux, "Convexité en topologie de contact", *Comment. Math. Helv.* **66**:4 (1991), 637–677. MR Zbl

[5] E. Giroux, "Géométrie de contact: de la dimension trois vers les dimensions supérieures", pp. 405–414 in *Proceedings of the International Congress of Mathematicians, Vol. II* (Beijing, 2002), Higher Ed. Press, Beijing, 2002. MR Zbl

[6] K. Honda, W. H. Kazez, and G. Matic, "Contact structures, sutured Floer homology and TQFT", 2008. arXiv 0807.2431

[7] K. Honda, W. H. Kazez, and G. Matić, "The contact invariant in sutured Floer homology", *Invent. Math.* **176**:3 (2009), 637–676. MR Zbl

[8] K. Honda, W. H. Kazez, and G. Matić, "On the contact class in Heegaard Floer homology", *J. Differential Geom.* **83**:2 (2009), 289–311. MR Zbl

[9] T. Ito and K. Kawamuro, "Open book foliation", *Geom. Topol.* **18**:3 (2014), 1581–1634. MR Zbl

[10] A. Juhász, "Holomorphic discs and sutured manifolds", *Algebr. Geom. Topol.* **6** (2006), 1429–1457. MR Zbl

[11] P. B. Kronheimer and T. S. Mrowka, "Witten's conjecture and property P", *Geom. Topol.* **8** (2004), 295–310. MR Zbl

[12] J. E. Licata and V. Vértesi, "Foliated open books", preprint, 2020.

[13] J. E. Licata and V. Vértesi, "A User's Guide to Morse Foliated Open Books", preprint, 2021. arXiv 2104.06705

[14] R. Lipshitz, P. S. Ozsvath, and D. P. Thurston, *Bordered Heegaard Floer homology*, vol. 254, 2018. MR Zbl

[15] P. Massot, "Topological methods in 3-dimensional contact geometry", pp. 27–83 in *Contact and symplectic topology*, Bolyai Soc. Math. Stud. **26**, János Bolyai Math. Soc., Budapest, 2014. MR Zbl

[16] Y. Ni, "Knot Floer homology detects fibred knots", *Invent. Math.* **170**:3 (2007), 577–608. MR Zbl

[17] P. Ozsváth and Z. Szabó, "Holomorphic disks and genus bounds", *Geom. Topol.* **8** (2004), 311–334. MR

[18] P. Ozsváth and Z. Szabó, "Heegaard Floer homology and contact structures", *Duke Math. J.* **129**:1 (2005), 39–61. MR

[19] P. Ozsváth and Z. Szabó, "Holomorphic disks, link invariants and the multi-variable Alexander polynomial", *Algebr. Geom. Topol.* **8**:2 (2008), 615–692. MR

[20] W. P. Thurston and H. E. Winkelnkemper, "On the existence of contact forms", *Proc. Amer. Math. Soc.* **52** (1975), 345–347. MR Zbl

[21] R. Zarev, "Bordered Floer homology for sutured manifolds", preprint, 2009. arXiv 0908.1106

[22] R. Zarev, "Joining and gluing sutured Floer homology", preprint, 2010. arXiv 1010.3496

Received 1 Mar 2021. Revised 22 Dec 2021.

AKRAM ALISHAHI: akram.alishahi@uga.edu
Department of Mathematics, University of Georgia, Boyd Graduate Studies Research Center, Athens, GA, United States

JOAN E. LICATA: joan.licata@anu.edu.au
Australian National University, Ainslie ACT, Australia

and

Mathematical Sciences Institute, The Australian National University, Canberra, Australia

INA PETKOVA: ina.petkova@dartmouth.edu
Department of Mathematics, Dartmouth College, Hanover, NH, United States

VERA VÉRTESI: vera.vertesi@univie.ac.at
University of Vienna, Vienna, Austria

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.31

Branched covering simply connected 4-manifolds

David Auckly, R. İnanç Baykur, Roger Casals,
Sudipta Kolay, Tye Lidman and Daniele Zuddas

We prove that any closed simply connected smooth 4-manifold is 16-fold branched covered by a product of an orientable surface with the 2-torus, where the construction is natural with respect to spin structures. In particular this solves Problem 4.113(C) in Kirby's list. We also discuss analogous results for other families of 4-manifolds with infinite fundamental groups.

1. Introduction

The work in this note was prompted by the following natural question:

> Does every closed 4-manifold admit a branched covering by a symplectic 4-manifold?

Problem 4.113(C) in Kirby's list [15] is the instance of this question for simply connected, irreducible manifolds. This question was studied by the authors at the 2018 American Institute of Mathematics Workshop on "Symplectic four-manifolds through branched coverings", motivated by a conjecture of Eliashberg [9, Conjecture 6.2]. We provide the following fairly strong answer to the above question in the case of simply connected 4-manifolds.

Theorem 1. *Let X be a closed oriented simply connected smooth 4-manifold. Then there exist $g \in \mathbb{N}$ and a degree 16 branched covering $f : X' \to X$ such that X' is the smooth 4-manifold $T^2 \times \Sigma_g$. In addition, if the 4-manifold X is spin, the branched covering f is natural with respect to a spin structure on $T^2 \times \Sigma_g$.*

Note that the smooth 4-manifold $T^2 \times \Sigma_g$ admits a symplectic structure. It follows from Theorem 1 that if instead X is a closed (possibly nonorientable) connected smooth 4-manifold with finite $\pi_1(X)$, then there is a branched covering $T^2 \times \Sigma_g \to X$ of degree $16|\pi_1(X)|$, which factors through the universal covering $\widetilde{X} \to X$.

MSC2020: 57K40, 57M12.
Keywords: branched cover, 4-manifold, symplectic.

The above do not generalize to 4-manifolds with arbitrary fundamental groups; for instance, no Σ_g-bundle over Σ_h with $g, h \geq 2$ and infinite monodromy group (e.g., any surface bundle with nonzero signature) can be dominated by a product 4-manifold by [16, Theorem 1.4]. Nonetheless, there are comparable results for many other 4-manifolds with infinite fundamental groups. For example, $X = \#_g(S^1 \times S^3)$, with $\pi_1(X) \cong *\,\mathbb{Z}^g$, is degree 4 branched covered by $X' = S^2 \times \Sigma_g$ by [25, Theorem 1.2]. In addition, the branched virtual fibering theorem of Sakuma [29, Addendum 1] implies the following:

Proposition 2. *Let $X = S^1 \times Y$ be a smooth 4-manifold which is the product of S^1 and a closed connected oriented 3-manifold Y. Then there exist $g \in \mathbb{N}$ and a double branched covering $X' \to X$, where X' is a symplectic 4-manifold which is a Σ_g-bundle over T^2.*

Indeed, [29] shows that any closed oriented 3-manifold Y is double branched covered by a surface bundle over a circle (see also [21] for a different proof), from which Proposition 2 is immediately deduced; this provides yet another class of 4-manifolds with infinite fundamental group for which a (symplectic) branched cover can be readily described. Here, we recall that the product of a fibered 3-manifold and the circle is symplectic [30].

It is worth noting that with a little more information on the smooth topology of X, one can easily determine the topology of the branched coverings $X' \to X$ in Theorem 1 and Proposition 2. For the former, one only needs to know the number of stabilizations by taking the connected sum with $S^2 \times S^2$ that are required before the simply connected 4-manifold X completely decomposes into a connected sum of copies of $\mathbb{CP}^2$, $S^2 \times S^2$ and the K3 surface, taken with either orientation. This of course can always be achieved by a classical result of Wall [31], and for vast families of simply connected 4-manifolds, one stabilization is known to be enough [2]. Similarly, for Proposition 2, one just needs to know a Heegaard decomposition of the 3-manifold factor Y [29], or any open book on it [21]. See Remark 6 for some explicit examples.

In all the results we have discussed above, the covering symplectic 4-manifold X' is not of general type, in contrast with the symplectic domination results of Fine and Panov [10]; see Remark 7 below. It would be interesting to find more general families of nonsymplectic 4-manifolds branched covered (with universally fixed degree) by specific families of symplectic 4-manifolds like ours, say by Σ_g-bundles over Σ_h, for arbitrary h.

2. Proof of Theorem 1

Henceforth all the manifolds and maps we consider are assumed to be smooth. We denote by $\overline{X}$ the oriented 4-manifold X with the reversed orientation, and by $\#_a X \#_b Y$ the smooth connected sum of a copies of X and b copies of Y. We

denote by Σ_g^b a closed connected oriented surface of genus g with b boundary components, and we drop b from the notation when there is no boundary.

2.1. ***Preliminaries.*** Let us briefly recall the definition of a branched covering.

Definition 3. Let X and X' be compact connected smooth manifolds (possibly with boundary) of the same dimension, and let $f : X' \to X$ be a smooth proper surjective map. We say that f is a branched covering if it is finite-to-one and open, and moreover the (open) subset of X' where f is locally injective coincides with the subset of X' where f is a local diffeomorphism. □

The subset $B'_f \subset X'$ where f fails to be locally injective is called the *branch set* of f, and its image $B_f = f(B'_f) \subset X$ is called the *branch locus* of f. By a result of Church [7, Corollary 2.3], either $B'_f = \varnothing$ or $\dim B'_f = \dim B_f = \dim X - 2$, and then the restriction of f over the complement of B_f is an ordinary connected covering space $X' \setminus f^{-1}(B_f) \to X \setminus B_f$.

Moreover, for every *smooth* point of B'_f at which $f|_{B'_f} : B'_f \to X$ is a local smooth embedding, the map f is *topologically* locally equivalent to the map $p_d : \mathbb{C} \times \mathbb{R}^{n-2} \to \mathbb{C} \times \mathbb{R}^{n-2}$ defined by $p_d(z, x) = (z^d, x)$, for some $d \geq 2$, where $n = \dim X' = \dim X$. However, the branched coverings f_i that we consider below turn out to be *smoothly* locally equivalent to p_2, while their composition, which will be indicated by f, has this property away from the singular points of B'_f. Notice that every finite composition of branched coverings is a branched covering, and the restriction to the boundary of a branched covering is a branched covering as well. Throughout, we assume that branched coverings between oriented manifolds are orientation-preserving.

2.2. ***The argument.*** Let X be a closed oriented simply connected smooth 4-manifold. We will describe the branched covering in the statement of Theorem 1, that is $f : T^2 \times \Sigma_g \to X$, as a composition of four simpler double branched coverings f_1, f_2, f_3, f_4. While all the latter will be branched over embedded orientable surfaces, the branch locus of the composition will typically be singular.

For clarity of exposition, we will not explicitly keep track of how the topology is growing at each step, but instead, we will illustrate with some examples in Remark 6 how one can deduce this information.

Step 1: By Wall [31], the connected sum of X with a certain number m of copies of $S^2 \times S^2$ is diffeomorphic to a connected sum of copies of the standard 4-manifolds $\mathbb{CP}^2$, $S^2 \times S^2$ and the K3 surface, taken with either orientation. Note that when X is spin, the decomposition has only spin connected summands, and also that the resulting 4-manifold does satisfy $11/8$ when m is large enough.

Moreover, since we have $\mathrm{K3} \# \overline{\mathrm{K3}} \cong \#_{22}(S^2 \times S^2)$ and $\mathbb{CP}^2 \# (S^2 \times S^2) \cong \#_2 \mathbb{CP}^2 \# \overline{\mathbb{CP}}^2$ [12, page 344], the complete decomposition as above can be written

as $\#_a \mathrm{K3} \#_b (S^2 \times S^2)$ or $\#_a \overline{\mathrm{K3}} \#_b (S^2 \times S^2)$ when X is spin (depending on the sign of the signature $\sigma(X)$), and as $\#_a \mathbb{CP}^2 \#_b \overline{\mathbb{CP}}^2$ when X is nonspin, for some nonnegative integers a and b which are not both zero. In the spin case, we can guarantee that $b \geq 2a$ by taking sufficiently many stabilizations.

The conjugation map $(z_1, z_2) \mapsto (\bar{z}_1, \bar{z}_2)$, which is an antiholomorphic involution on $\mathbb{CP}^1 \times \mathbb{CP}^1 \cong S^2 \times S^2$, induces a double branched covering $S^2 \times S^2 \to S^4$, where the branch locus is the unknotted $T^2 \subset S^4$ (bounding a handlebody). Taking the equivariant connected sum of m copies of it, we get an involution on $\#_m(S^2 \times S^2)$, which induces a double covering $\#_m(S^2 \times S^2) \to S^4$ branched along an unknotted Σ_m, for every $m \geq 1$.

We can now take a double covering of X branched along an unknotted Σ_m in X (viewing $X \cong X \# S^4$, take an unknotted Σ_m in S^4), which we denote by $f_1 : X_1 \to X$, where clearly $X_1 \cong X \# X \#_m (S^2 \times S^2)$. We choose $m \geq 1$ such that $X \#_m (S^2 \times S^2)$ completely decomposes, and so does X_1 (as one gets at least m copies of $S^2 \times S^2$ after decomposing $X \#_m (S^2 \times S^2)$). Then X_1 is diffeomorphic to one of the standard connected sums we listed above.

Step 2: We would like to obtain a double branched covering of X_1 by some $\#_g(S^2 \times S^2)$. We will describe this covering in essentially two different ways, depending on whether X (and thus X_1) is spin or not.

The K3 surface can be obtained as a holomorphic double covering of $S^2 \times S^2$ branched along a curve of bidegree $(4, 4)$ in $\mathbb{CP}^1 \times \mathbb{CP}^1 \cong S^2 \times S^2$ [12, page 262]. Reversing the orientations, we see that $\overline{\mathrm{K3}}$ is also a double branched covering of $S^2 \times S^2$ (recall that $S^2 \times S^2$ admits an orientation-reversing diffeomorphism). By taking equivariant connected sums, we can then express both $\#_n \mathrm{K3}$ and $\#_n \overline{\mathrm{K3}}$ as branched double coverings of $\#_n(S^2 \times S^2)$. Taking $n = 2a$, we then conclude that $\#_a \mathrm{K3} \#_b (S^2 \times S^2)$ admits a double branched covering by $\#_{2a} \mathrm{K3} \#_{2a} \overline{\mathrm{K3}} \#_{2(b-2a)} (S^2 \times S^2)$. Since $\mathrm{K3} \# \overline{\mathrm{K3}} \cong \#_{22}(S^2 \times S^2)$, we have obtained the desired double branched cover $\#_g(S^2 \times S^2)$, for $g = 40a + 2b$. Mirroring the same argument, we see that $\#_a \overline{\mathrm{K3}} \#_b (S^2 \times S^2)$ is also double branched covered by some $\#_g(S^2 \times S^2)$. This concludes the construction in the spin case.

The following variation can be run for both spin and nonspin manifolds. Switching the two factors $(z_1, z_2) \mapsto (z_2, z_1)$, which is a holomorphic involution on $\mathbb{CP}^1 \times \mathbb{CP}^1 \cong S^2 \times S^2$, induces a double branched covering $S^2 \times S^2 \to \mathbb{CP}^2$, where the branch locus is the quadric (this may be interpreted as the map taking a pair of numbers to the quadratic equation having those roots). Reversing the orientations, we obtain a double branched covering over $\overline{\mathbb{CP}}^2$. Taking equivariant connected sums once again, we then deduce that $\#_a(S^2 \times S^2) \#_b (S^2 \times S^2)$ is a double branched covering of $\#_a \mathbb{CP}^2 \#_b \overline{\mathbb{CP}}^2$. So in the nonspin case, we arrive at the desired double covering $\#_g(S^2 \times S^2)$ as well.

We let $f_2 : X_2 \to X_1$ denote the double branched covering we described in either case.

Step 3: We next show that $S^1 \times \#_g(S^1 \times S^2)$ is a double branched covering of $\#_g(S^2 \times S^2)$, which will prescribe our next covering $f_3 : X_3 \to X_2$. A similar double branched covering over $\#_g(\mathbb{CP}^2 \# \overline{\mathbb{CP}}^2)$ was described by Neofytidis in [23, Theorem 1]. (Also see [24] for similar constructions in other dimensions.)

The hyperelliptic involution on T^2 induces a double branched covering $p : T^2 \to S^2$ with four simple branch points $x_1, x_2, x_3, x_4 \in S^2$. Taking its product with the identity map on S^2 yields a double branched covering $p \times \mathrm{id}_{S^2} : T^2 \times S^2 \to S^2 \times S^2$ with branch locus $\{x_1, x_2, x_3, x_4\} \times S^2$. Note that if $g = 1$, we can stop here and skip Step 4.

Let $D \subset S^2$ be a 2-disk containing exactly two branch points of p. So $A = p^{-1}(D) \subset T^2$ is an equivariant annulus that contains two fixed points of the hyperelliptic involution. Moreover, D can be chosen such that A is a union of fibers of the trivial S^1-bundle $T^2 = S^1 \times S^1 \to S^1$ given by the canonical projection onto the second factor.

Let $S^2 \cong S^2 \times \{y\} \subset S^2 \times S^2$ be a fiber sphere, for a certain $y \in S^2$. Let $D' \subset S^2$ be a disk centered at y. Then, $U = D \times D' \subset S^2 \times S^2$ is a fibered bidisk, whose preimage $V = (p \times \mathrm{id}_{S^2})^{-1}(U) \cong A \times D'$ is a fibered neighborhood of a fiber of the trivial S^1-bundle

$$T^2 \times S^2 = S^1 \times (S^1 \times S^2) \to S^1 \times S^2.$$

By taking two copies of the branched covering $T^2 \times S^2 \to S^2 \times S^2$, and performing an equivariant fiber sum upstairs along V and connected sum downstairs along $U \cong D^4$, and repeating the construction for every $g \geq 2$, we finally get a branched double covering $S^1 \times \#_g(S^1 \times S^2) \to \#_g(S^2 \times S^2)$.

We can also describe this branched covering as follows: start with a double covering $q : S^1 \times D^1 \to D^2$ branched over two points in $\mathrm{Int}\, D^2$ (this is the above branched covering $A \to D$), so the product $q \times \mathrm{id}_{D^1} : S^1 \times D^1 \times D^1 \to D^2 \times D^1$ yields a double covering $q' : S^1 \times D^2 \to D^3$ branched over the union of two parallel proper trivial arcs in D^3 (this fills the above branched covering $p : T^2 \to S^2$), up to the identifications $S^1 \times D^1 \times D^1 \cong S^1 \times D^2$ and $D^2 \times D^1 \cong D^3$. Then, we get a double branched covering $q'' = q' \times \mathrm{id}_{S^2} : S^1 \times D^2 \times S^2 \to D^3 \times S^2$. Let $D \subset S^2$ be a 2-disk. Up to the identification $D^2 \times D^1 \times S^2 \cong D^3 \times S^2$, we consider the bidisks $C^- = D^2 \times \{-1\} \times D$ and $C^+ = D^2 \times \{1\} \times D \subset \partial(D^3 \times S^2)$, each of which intersects the branch locus of q'' along the union of two parallel proper trivial 2-disks. Consider g copies of q'', say $q''_i : (S^1 \times D^2 \times S^2)_i \to (D^3 \times S^2)_i$, and let $C_i^-, C_i^+ \subset \partial(D^3 \times S^2)_i$ be the corresponding bidisks. Thus, we obtain a double branched covering

$$q''' = q''_1 \cup \cdots \cup q''_g : \cup_i (S^1 \times D^2 \times S^2)_i \to \cup_i (D^3 \times S^2)_i,$$

where $(D^3 \times S^2)_i$ is attached to $(D^3 \times S^2)_{i+1}$ by identifying C_i^+ with C_{i+1}^- and $(S^1 \times D^2 \times S^2)_i$ is attached to $(S^1 \times D^2 \times S^2)_{i+1}$ by identifying $(q''_i)^{-1}(C_i^+)$ with

$(q''_{i+1})^{-1}(C^-_{i+1})$ in the obvious way, for all $i = 1, \dots, g-1$. This in turn is a double branched covering

$$q''' : S^1 \times \sharp_g(D^2 \times S^2) \to \sharp_g(D^3 \times S^2),$$

as it can be easily realized by looking at the attaching maps, where $\sharp$ denotes the boundary connected sum. Finally, the desired branched covering $S^1 \times \#_g(S^1 \times S^2) \to \#_g(S^2 \times S^2)$ can be obtained by restricting q''' to the boundary.

Step 4: Our final double branched covering $f_4 : T^2 \times \Sigma_g \to S^1 \times \#_g(S^1 \times S^2)$ is a special case of Proposition 2 and can be obtained by taking the product of the identity map on the S^1 factor with a double branched covering $S^1 \times \Sigma_g \to \#_g(S^1 \times S^2)$. The latter can be derived from the work of Sakuma [29] we mentioned in the introduction, or from Montesinos' alternative construction [21], which is quicker to describe here: the involution $(z, t) \to (\bar{z}, -t)$ on the annulus $A = S^1 \times [-1, 1] \subset \mathbb{C} \times \mathbb{R}$ induces a double covering $q : A \to D^2$ branched at two points (this is same as the double branched cover described in Step 3), so we get a double branched covering $q \times \mathrm{id}_{S^1} : A \times S^1 \to D^2 \times S^1$. Then, for any open book decomposition of a closed connected oriented 3-manifold Y with pages Σ^m_k and monodromy ϕ, we can get a double covering $h : Y' \to Y$ branched over two parallel copies of the binding, where Y' is now a surface bundle whose fiber and the monodromy are the doubles of Σ^m_k and ϕ. Indeed, by lifting the usual splitting $Y = (D^2 \times \partial\Sigma^m_k) \cup_\partial T(\phi)$ that gives the open book decomposition of Y, with the branch link contained in $D^2 \times \partial\Sigma^m_k$, and where $T(\phi)$ denotes the mapping torus of ϕ, one obtains a splitting $Y' = (A \times \partial\Sigma^m_k) \cup_\partial (T(\phi)_1 \cup T(\phi)_2)$, with the annulus A instead of D^2, where $T(\phi)_1$ and $T(\phi)_2$ are two disjoint copies of $T(\phi)$ (the branched covering $h : Y' \to Y$ is trivial over $T(\phi)$). By looking at the attaching maps, it is immediate to get the bundle structure on Y' as above.

In our case, since $\#_g(S^1 \times S^2)$ admits a planar open book with pages Σ^{g+1}_0 and $\phi = \mathrm{id}$, we obtain the desired covering. (The covering produced by the arguments of both Sakuma and Montesinos in this simple setting is equivalent to the one given in [17, Proposition 4].)

The composition $f = f_1 \circ f_2 \circ f_3 \circ f_4 : T^2 \times \Sigma_g \to X$ gives the desired covering.

The spin case: Let us conclude by observing that our construction is natural with respect to the spin structures, when X is spin, and then briefly discuss the topology of the branch locus of f.

Recall that a spin structure on a 4-manifold is the same as a trivialization of the tangent bundle over the 1-skeleton that extends over the 2-skeleton [20; 14]. We may use a handlebody decomposition in this definition. Given an unramified cover over a spin 4-manifold, the trivialization will lift to the tangent bundle of the cover restricted to the 1-skeleton and any extension to the 2-skeleton, so there is a natural lift of a spin structure to a covering space.

Now consider a 2-fold branched covering with branch locus B. We may build a handle decomposition of the base in the following way. Start with a handle decomposition of B. This extends to a handle decomposition of a tubular neighborhood of B with only zero, one and two handles. Now extend this to a handle decomposition of the rest of the base X. Finally turn the entire handle decomposition over. Notice that all of the 1-handles of this new handle decomposition are in the exterior of B. Each of these handles lifts to the cover of the exterior and the restriction of the spin structure to the exterior lifts to the cover. We now complete the handle decomposition of the total space of the branched cover as follows. Use the identification of the inverse image $\widetilde{B}$ with B to construct a decomposition of $\widetilde{B}$ which is then extended to a decomposition of the normal bundle of $\widetilde{B}$. Turn this upside down and add it the to decomposition of the inverse image of the exterior. This only adds 2-, 3- and 4-handles to the decomposition. It is not necessarily true that the trivialization of the tangent bundle over the 1-skeleton will extend over the 2-skeleton. It will extend precisely when the mod two reduction of the integral homology class $[B]/2$ is zero in the second homology of the base with $\mathbb{Z}_2$ coefficients [4; 22]. Note that the class of $[B]$ is necessarily divisible by 2 due to the existence of the double branched cover. So, a spin structure does not have to lift to the total space of a 2-fold branched covering, but if it does, there is a natural lift.

It is now straightforward to check that each double cover f_i that we employed in our construction when X is spin satisfies the above criterion, so for the initial spin structure $\mathfrak{s}$ on X, there is a spin structure $\mathfrak{s}'$ on $X' \cong T^2 \times \Sigma_g$ constructed this way. (Note that there are $2^{2(g+1)}$ different spin structures on X'.) Thus the branched covering $X' \to X$ is compatible with the spin structures $\mathfrak{s}$ on X and $\mathfrak{s}'$ on X'.

The branch locus: The branch locus $B_f \subset X$ of f is given by

$$B_f = B_{f_1} \cup f_1(B_{f_2} \cup f_2(B_{f_3} \cup f_3(B_{f_4}))),$$

where $B_{f_i} \subset X_{i-1}$ denotes the branch locus of f_i, for $i = 1, 2, 3, 4$, with $X_0 = X$. Each B_{f_i} is a smooth embedded closed orientable surface in X_{i-1}. By taking into account that each covering f_i is two-to-one and its tangent map has a 2-dimensional kernel along the branch set, an easy transversality argument based on perturbing the f_i's up to isotopy, shows that the branch locus $B_f \subset X$ can be assumed to be a smooth orientable surface away from at most finitely many singular points, which are transversal or tangential double points (at the latter the local link has two trivial components with linking number ± 2). □

3. Ancillary remarks

Let us list a few comments in relation to Theorem 1, its proof and related works.

Remark 4 (variations). In Step 1 above we could have stabilized by taking connected sums with copies of $\mathbb{CP}^2$ and $\overline{\mathbb{CP}}^2$ so that we got a double covering $g_1 : \#_a \mathbb{CP}^2 \#_b \overline{\mathbb{CP}}^2 \to X$ branched over a genus m *nonorientable* surface, which is trivially embedded in X, for certain integers a, b and m (once again by Wall [31]). The complex conjugation on $\mathbb{CP}^2$ induces this double covering $\mathbb{CP}^2 \to S^4$ branched over the standard smooth $\mathbb{RP}^2 \subset S^4$ [18; 19]. Now, we can invoke Theorem 1.2 in [25] to conclude that there exists a 4-fold simple branched covering

$$g_2 : \Sigma_h \times \Sigma_g \to \#_a \mathbb{CP}^2 \#_b \overline{\mathbb{CP}}^2$$

for *every* given $a, b \geq 0$ and $h \geq 1$, and for some g large enough. Thus, the composition $g_1 \circ g_2 : \Sigma_g \times \Sigma_h \to X$ is a degree 8 branched covering.

Again by Theorem 1.2 in [25] (see also Remark 2 therein), there exist degree 4 branched coverings $T^4 = T^2 \times T^2 \to X$, with $X = \#_m \mathbb{CP}^2 \#_n \overline{\mathbb{CP}}^2$ and $X = \#_n (S^2 \times S^2)$, for every $m, n \leq 3$. Note that the case $X = S^2 \times S^2$ is straightforward by taking the product $p \times p : T^2 \times T^2 \to S^2 \times S^2$, and the case $X = \#_2 (S^2 \times S^2)$ was previously obtained by Rickman [28]. Branched coverings from the n-dimensional torus are relevant in connection with the theory of *quasiregularly elliptic* manifolds; see Bonk and Heinonen [3]. In this direction, a result by Prywes [26, Theorem 1.1] implies that if there is a branched covering $T^4 \to X$, then $b_1(X) \leq 4$ and $b_2(X) \leq 6$, so in Theorem 1 we cannot take $g \leq 1$ if $b_2(X) \geq 7$.

However, unlike in our construction above, the results in [25] do not give explicit branched coverings, and there is not much control on the topology of the branch locus.

Remark 5 (branched cover geometries). Theorem 1 and our subsequent remark in the introduction imply that any X with finite $\pi_1(X)$ is branched covered by $T^2 \times \Sigma_g$, where it is easy to see from our proof that we can always assume $g \geq 2$. In terms of 4-dimensional geometries [13], this shows that all such X can be branched covered by a 4-manifold with $\mathbb{E}^2 \times \mathbb{H}^2$ geometry. However, if we replace the double branched covering $h : Y' \to \#_g (S^1 \times S^2)$ we used in the construction of $f_4 = \mathrm{id}_{S^1} \times h$ with the one built by Brooks in [5], we can also get Y' to be a Σ_g-bundle over S^1 with hyperbolic total space. Therefore, any X with finite $\pi_1(X)$ can also be branched covered by a 4-manifold with $\mathbb{E} \times \mathbb{H}^3$ geometry. Similarly, one can modify the construction in Proposition 2 to get a double branched cover of any product 4-manifold $S^1 \times Y$ by a 4-manifold with $\mathbb{E} \times \mathbb{H}^3$ geometry.

Remark 6 (topology of the branched coverings). Here we will try to demonstrate by way of example how one can control the topology of the branched coverings in Theorem 1. For some variety, we will run our construction for two infinite families of irreducible 4-manifolds which are not completely decomposable: Dolgachev surfaces, which are nonspin complex surfaces of general type, and knot surgered K3

surfaces of Fintushel and Stern, which include spin 4-manifolds that do not admit any symplectic structures [11]. The members of either one of these two families of simply connected 4-manifolds completely decompose after a single stabilization by $S^2 \times S^2$; see, e.g., [1]. Now, if X is a Dolgachev surface, we can take the branch locus of f_1 as an unknotted T^2, and get $X_1 = \#_3 \mathbb{CP}^2 \#_{19} \overline{\mathbb{CP}}^2$. The branched covering f_2 performed along the connected sum of 22 quadrics (each coming from distinct copies of $\mathbb{CP}^2$ and $\overline{\mathbb{CP}}^2$) then gives $X_2 = \#_{22}(S^2 \times S^2)$, and the last two coverings yield $X' = T^2 \times \Sigma_{22}$. If we take X to be a knot surgered K3 surface instead, thinking ahead of the second step, we take the branch locus of f_1 this time as Σ_4, so $X_1 = \#_2 \mathrm{K3} \#_4 (S^2 \times S^2)$. The next double cover f_2 is taken along a connected sum of four bidegree $(4,4)$ curves (each coming from distinct copies of $S^2 \times S^2$, with noncomplex orientation), and we get $X_2 = \#_4 \mathrm{K3} \#_4 \overline{\mathrm{K3}} \cong \#_{88}(S^2 \times S^2)$. The last two coverings this time yield $X' = T^2 \times \Sigma_{88}$.

Remark 7 (symplectic domination). A recent article of Fine and Panov provides a symplectic domination result [10, Theorem 1] which is worth mentioning here. Their beautiful construction is very general: for any closed oriented even-dimensional smooth manifold M, they build a closed symplectic manifold S of the same dimension with a positive degree map $f : S \to M$. In dimension 4, where we can compare their result with ours in Theorem 1, their symplectic manifold S is constructed as a Donaldson hypersurface in the 6-dimensional symplectic twistor space Z of a negatively pinched manifold N, where the latter admits a degree one map $g : N \to M$. The construction of N, with sectional curvature arbitrarily close to -1, is implicit, and relies on the recent works of Ontaneda involving rather intricate new techniques in Riemannian geometry. (The condition on the sectional curvature is to guarantee that the twistor space Z of N is a symplectic 6-manifold.) Secondly, the construction of a symplectic hypersurface S in N, which is built through asymptotically holomorphic techniques of [8], is also implicit and the smooth topology of S is effectively impossible to control. Hence, one does not have any information on the smooth topology of the dominating symplectic 4-manifold S, other than that it is of general type, i.e., of Kodaira dimension 2 [10]. Besides the very implicit nature of this construction, since the map $f : S \to M$ factors through the degree one map g above, Fine and Panov's domination is essentially never a branched covering. Moreover, because the symplectic twistor space Z is in fact known to be non-Kähler [27], the dominating symplectic 4-manifold S has a priori no reason to be a Kähler surface. On the other hand, the dominating symplectic 4-manifold $X' = T^2 \times \Sigma_g$ of Theorem 1 is obviously a Kähler surface, and X' in both Theorem 1 and Proposition 2 is of Kodaira dimension $-\infty$, 0 or 1, depending on whether this (possibly trivial) Σ_g-bundle over T^2, has fiber genus $g = 0$, 1 or ≥ 2, respectively.

Domination is certainly distinct from branched covering as the following example shows. There is a degree one map from $\Sigma_4 \times \Sigma_2$ to $\Sigma_3 \times \Sigma_2$ given by the extension

of the natural collapse of a copy of $\Sigma_1^1 \times \Sigma_2$ to $\Sigma_0^1 \times \Sigma_2$. However there can be no branched covering from $\Sigma_4 \times \Sigma_2$ to $\Sigma_3 \times \Sigma_2$ since the Gromov norm of the former is $24(4-1)(2-1) = 72$, the Gromov norm of the latter is $24(3-1)(2-1) = 48$ and the Gromov norm is super multiplicative with respect to degree [6].

Acknowledgments

This project was started at the 2018 AIM workshop on *"Symplectic four-manifolds through branched coverings"*, and was resumed following the 2020 BIRS Workshop on *"Interactions of gauge theory with contact and symplectic topology in dimensions* 3 *and* 4*."* The authors would like to thank the American Institute of Mathematics and the Banff International Research Station, and the other organizers of these workshops. Auckly was partially supported by the Simons Foundation grant 585139 and NSF grant DMS 1952755. İnanç Baykur was partially supported by the NSF grants DMS-200532 and DMS-1510395. Casals is supported by the NSF grant DMS-1841913, the NSF CAREER grant DMS-1942363 and the Alfred P. Sloan Foundation. Lidman was partially supported by the NSF grant DMS-1709702 and a Sloan Fellowship. Zuddas was partially supported by the 2013 ERC Advanced Research Grant 340258 TADMICAMT; he is member of GNSAGA, Istituto Nazionale di Alta Matematica "Francesco Severi", Italy. We would like to thank the anonymous referee for helpful comments.

References

[1] R. I. Baykur, "Dissolving knot surgered 4-manifolds by classical cobordism arguments", *J. Knot Theory Ramifications* **27**:5 (2018), art. id. 1871001. MR Zbl

[2] R. I. Baykur and N. Sunukjian, "Round handles, logarithmic transforms and smooth 4-manifolds", *J. Topol.* **6**:1 (2013), 49–63. MR Zbl

[3] M. Bonk and J. Heinonen, "Quasiregular mappings and cohomology", *Acta Math.* **186**:2 (2001), 219–238. MR Zbl

[4] N. Brand, "Necessary conditions for the existence of branched coverings", *Invent. Math.* **54**:1 (1979), 1–10. MR

[5] R. Brooks, "On branched coverings of 3-manifolds which fiber over the circle", *J. Reine Angew. Math.* **362** (1985), 87–101. MR Zbl

[6] M. Bucher-Karlsson, "The simplicial volume of closed manifolds covered by $\mathbb{H}^2 \times \mathbb{H}^2$", *J. Topol.* **1**:3 (2008), 584–602. MR Zbl

[7] P. T. Church, "Differentiable open maps on manifolds", *Trans. Amer. Math. Soc.* **109** (1963), 87–100. MR Zbl

[8] S. K. Donaldson, "Symplectic submanifolds and almost-complex geometry", *J. Differential Geom.* **44**:4 (1996), 666–705. MR Zbl

[9] Y. Eliashberg, "Recent advances in symplectic flexibility", *Bull. Amer. Math. Soc.* (*N.S.*) **52**:1 (2015), 1–26. MR Zbl

[10] J. Fine and D. Panov, "Symplectic domination", *Bull. Lond. Math. Soc.* **53**:1 (2021), 100–103. MR Zbl

[11] R. Fintushel and R. J. Stern, "Knots, links, and 4-manifolds", *Invent. Math.* **134**:2 (1998), 363–400. MR Zbl

[12] R. E. Gompf and A. I. Stipsicz, *4-manifolds and Kirby calculus*, Graduate Studies in Mathematics **20**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[13] J. A. Hillman, *Four-manifolds, geometries and knots*, Geometry & Topology Monographs **5**, Geometry & Topology Publications, Coventry, 2002. MR Zbl

[14] R. C. Kirby, *The topology of 4-manifolds*, Lecture Notes in Mathematics **1374**, Springer, 1989. MR Zbl

[15] R. Kirby, "Problems in low-dimensional topology", pp. 35–473 in *Geometric topology* (Athens, GA, 1993), edited by R. Kirby, AMS/IP Stud. Adv. Math. **2**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[16] D. Kotschick and C. Löh, "Fundamental classes not representable by products", *J. Lond. Math. Soc.* (2) **79**:3 (2009), 545–561. MR Zbl

[17] D. Kotschick and C. Neofytidis, "On three-manifolds dominated by circle bundles", *Math. Z.* **274**:1-2 (2013), 21–32. MR Zbl

[18] N. H. Kuiper, "The quotient space of $\mathbf{C}P(2)$ by complex conjugation is the 4-sphere", *Math. Ann.* **208** (1974), 175–177. MR Zbl

[19] W. S. Massey, "The quotient space of the complex projective plane under conjugation is a 4-sphere", *Geometriae Dedicata* **2** (1973), 371–374. MR Zbl

[20] J. Milnor, "Spin structures on manifolds", *Enseign. Math.* (2) **9** (1963), 198–203. MR Zbl

[21] J. M. Montesinos, "On 3-manifolds having surface bundles as branched coverings", *Proc. Amer. Math. Soc.* **101**:3 (1987), 555–558. MR Zbl

[22] S. Nagami, "On spin structures of double branched covering spaces", *JP J. Geom. Topol.* **14**:2 (2013), 119–147. MR Zbl

[23] C. Neofytidis, "Branched coverings of simply connected manifolds", *Topology Appl.* **178** (2014), 360–371. MR Zbl

[24] C. Neofytidis, *Non-zero degree maps between manifolds and groups presentable by products*, Ph.D. thesis, LMU München, 2014.

[25] R. Piergallini and D. Zuddas, "Branched coverings of $\mathbb{CP}^2$ and other basic 4-manifolds", *Bull. Lond. Math. Soc.* **53**:3 (2021), 825–842. Zbl

[26] E. Prywes, "A bound on the cohomology of quasiregularly elliptic manifolds", *Ann. of Math.* (2) **189**:3 (2019), 863–883. MR Zbl

[27] A. G. Reznikov, "Symplectic twistor spaces", *Ann. Global Anal. Geom.* **11**:2 (1993), 109–118. MR Zbl

[28] S. Rickman, "Simply connected quasiregularly elliptic 4-manifolds", *Ann. Acad. Sci. Fenn. Math.* **31**:1 (2006), 97–110. MR Zbl

[29] M. Sakuma, "Surface bundles over S^1 which are 2-fold branched cyclic coverings of S^3", *Math. Sem. Notes Kobe Univ.* **9**:1 (1981), 159–180. MR Zbl

[30] W. P. Thurston, "Some simple examples of symplectic manifolds", *Proc. Amer. Math. Soc.* **55**:2 (1976), 467–468. MR Zbl

[31] C. T. C. Wall, "On simply-connected 4-manifolds", *J. London Math. Soc.* **39** (1964), 141–149. MR Zbl

Received 26 Jan 2021. Revised 27 Feb 2021.

DAVID AUCKLY: dav@ksu.edu
Department of Mathematics, Kansas State University, Manhattan, KS, United States

R. İNANÇ BAYKUR: baykur@math.umass.edu
Department of Mathematics and Statistics, University of Massachusetts, Lederle Graduate Research Tower, Amherst, MA, United States

ROGER CASALS: casals@math.ucdavis.edu
Department of Mathematics, UC Davis, Davis, CA, United States

SUDIPTA KOLAY: skolay3@math.gatech.edu
School of Mathematics, Georgia Institute of Technology, Atlanta, GA, United States

TYE LIDMAN: tlid@math.ncsu.edu
Department of Mathematics, North Carolina State University, Raleigh, NC, United States

DANIELE ZUDDAS: dzuddas@units.it
Dipartimento di Matematica e Geoscienze, Università di Trieste, Trieste, Italy

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.43

Lifting Lagrangian immersions in $\mathbb{C}P^{n-1}$ to Lagrangian cones in $\mathbb{C}^n$

Scott Baldridge, Ben McCarty and David Vela-Vick

We show how to lift Lagrangian immersions in $\mathbb{C}P^{n-1}$ to produce Lagrangian cones in $\mathbb{C}^n$, and use this process to produce several families of examples of Lagrangian cones and special Lagrangian cones. As an application of this theorem, for $n = 3$ we show how to produce Lagrangian cones that are isotopic to the Harvey–Lawson special Lagrangian cone and the trivial cone. The projections of the Legendrian links of both of these cones to $\mathbb{C}P^2$ are immersions with four and seven transverse double points. We expect that these double points represent the chord generators of the 0-filtration level of a suitably defined version of Legendrian contact homology of the links.

1. Introduction

This paper focuses on creating models for Lagrangian cones. The motivation for this paper arises from the string theory model in physics. According to the theory, our universe consists of the standard Minkowski space-time, $\mathbb{R}^4$, together with a complex Calabi–Yau 3-fold, X. Based upon physical grounds, the SYZ-conjecture of Strominger, Yau, and Zaslow [32] expects that this Calabi–Yau manifold can be viewed as a fibration by 3-tori with some singular fibers. These singular fibers are not well understood. The standard approach is to model them locally as special Lagrangian cones $C \subset \mathbb{C}^3$ (by cone, we mean a subset $C \subset \mathbb{C}^3$ such that $r \cdot C = C$ for any real number $r > 0$). Such a cone can be characterized by its link, $C \cap S^5$, which is a Legendrian surface.

Special Lagrangian cones in $\mathbb{C}^3$ are solutions to nonlinear degree 2 and 3 partial differential equations. Many papers on the subject to date have used this perspective, often by using examples from algebraic geometry. However, given that the cone

MSC2020: 53D17, 53D35, 57R17.
Keywords: lagrangian, cone, SYZ, Calabi–Yau, knot, Harvey–Lawson, hypercube diagram, grid diagram, contact homology.

can be characterized by the Legendrian link, this topic is very closely related to the study of knotted Legendrian submanifolds. This relationship connects it to a great deal of work in the area of contact topology. In this area much progress has been made, at least in part, due to the fact that there are topological and combinatorial representations of such submanifolds. In dimension 3, where the problem of understanding Legendrian submanifolds amounts to classifying Legendrian knots up to isotopy, such diagrammatic representations are easy to generate. For instance, grid diagrams can be used to obtain combinatorial representations of both front and Lagrangian projections of Legendrian knots (see [4; 5; 7; 19; 28]). In higher dimensions, there are fewer such constructions. In [9], Ekholm, Etnyre, and Sullivan present front spinning as a way of constructing one class of knotted Legendrian tori, showing that the theory of Legendrian submanifolds of $\mathbb{R}^{2n+1}$ is at least as rich in higher dimensions as it is in dimension 3. To accomplish this, they extend the definition of Legendrian contact homology to $\mathbb{R}^{2n+1}$. In [4], it was shown that knotted Legendrian tori could be constructed from Lagrangian hypercube diagrams, and it was shown how to compute several invariants from such a diagram. In [16], Lambert-Cole showed how to generalize that construction to produce a product operation on Legendrian submanifolds.

1A. ***Lifts of Lagrangian immersions in $\mathbb{C}P^{n-1}$ to S^{2n-1}.*** With the appropriate setup, it is possible to construct models of Legendrian surfaces in S^5 so that the resulting cone in $\mathbb{C}^3$ is Lagrangian, and in some cases, special Lagrangian. The lifting theorem describes precisely the conditions under which an immersion into $\mathbb{C}P^{n-1}$ lifts to an embedded Legendrian submanifold of S^{2n-1} that gives rise to a Lagrangian cone.

Lifting theorem. *Let Σ be a closed, connected, smooth $(n-1)$-manifold, and $f : \Sigma \to \mathbb{C}P^{n-1}$ be a Lagrangian immersion with respect to the integral Fubini–Study symplectic form $\frac{1}{\pi}\omega_{FS}$. Let $\pi : S^{2n-1} \to \mathbb{C}P^{n-1}$ be the principle Hopf S^1-bundle with connection 1-form $\frac{i}{\pi}\alpha$ where $\alpha = i_0^*\big(\frac{1}{2}\sum_{i=1}^n x_i dy_i - y_i dx_i\big)$ for the identity map $i_0 : S^{2n-1} \to \mathbb{C}^n$. For each chart $\Psi_j : B_j \times S^1 \to S^{2n-1}$ (see Section 4), there exists a 1-form τ_j such that $\Psi_j^*(\alpha) = \frac{1}{2}(dt - \tau_j)$ where $\tau_j = -\sum_{i=1,\, i\neq j}^n (x_i dy_i - y_i dx_i)$.*
If

(1) $\Gamma \int_\gamma \tau = 0 \bmod 2\pi$ *for all* $[\gamma] \in H_1(\Sigma; \mathbb{Z})$, *and*

(2) *for all distinct points $x_1, \ldots, x_k \in \Sigma$ such that $f(x_1) = f(x_j)$ for all $j \leq k$, and a choice of path γ_j from x_1 to x_j in Σ for $2 \leq j \leq k$, the set*

$$\left\{ \left(\Gamma \int_{f(\gamma_j)} \tau \right) \bmod 2\pi \;\middle|\; 2 \leq j \leq k \right\}$$

has $k-1$ distinct values, none of which are equal to 0,

then $f : \Sigma \to \mathbb{C}P^{n-1}$ lifts to an embedding $\tilde{f} : \Sigma \to S^{2n-1}$ such that the image (the lift) $\widetilde{\Sigma}$ is a Legendrian submanifold of (S^{2n-1}, α). In turn, the cone $c\widetilde{\Sigma}$ is Lagrangian in $\mathbb{C}^n$ with respect to the standard symplectic structure $\omega_0 = \sum_{i=1}^n dx_i \wedge dy_i$.

Remark 1.1. The 1-form τ_j may be thought of as a multiple of a contact form on S^{2n-1} as observed in Section 4.

Remark 1.2. The integral $\Gamma \int_\gamma$ refers to a *lifting integral* (see Definition 4.8).

Remark 1.3. The second condition of the lifting theorem is stated for multiple points in general, but in most examples, we will only be working with double points or S^1-families of double points.

1B. ***Legendrian contact homology and Lagrangian cones.*** While the lifting theorem is quite general, it is often possible (and simpler) to work within a single chart of $\mathbb{C}P^{n-1}$. To construct a local model for special Lagrangian cones, we work in the symplectic manifold $(\mathbb{C}^n, \omega, \Omega)$ where $\mathbb{C}^n$ has complex coordinates $(z_1, \ldots, z_n)$, $\omega_0 = \frac{i}{2}(dz_1 \wedge d\overline{z_1} + \cdots + dz_n \wedge d\overline{z_n})$ is the standard Kähler form, and $\Omega = dz_1 \wedge \cdots \wedge dz_n$ is the holomorphic volume form (see [14]).

Definition 1.4. A cone $C \subset \mathbb{C}^n$ is special Lagrangian if it is Lagrangian and $\operatorname{Im} \Omega|_C \equiv 0$ or, equivalently, if C is calibrated (in the sense of [13]) with respect to $\operatorname{Re} \Omega$.

As a first step, we will focus first on the construction of Lagrangian cones. Observe that the kernel of the 1-form

$$\alpha = \tfrac{1}{2}(x_1 dy_1 - y_1 dx_1 + \cdots + x_n dy_n - y_n dx_n),$$

where $z_j = x_j + iy_j$, restricted to the unit sphere, generates the standard contact structure for S^{2n-1} and that $\alpha = \iota_R \omega$, where $R = 2\left(\sum_{i=1}^n x_i \frac{\partial}{\partial x_i} + y_i \frac{\partial}{\partial y_i}\right)$. This means that, given a Legendrian submanifold $\Sigma \subset S^{2n-1}$, the associated cone $c\Sigma$, obtained by scaling Σ by positive real numbers, is automatically Lagrangian. Moreover, any Lagrangian cone with vertex at the origin, must intersect S^{2n-1} in a Legendrian surface. Hence, with respect to the standard contact structure on S^{2n-1} and the standard symplectic form on $\mathbb{C}^n$, a given submanifold of $S^{2n-1} \subset \mathbb{C}^n$ is Legendrian if and only the associated cone in $\mathbb{C}^n$ is Lagrangian.

In knot theory, the trivial knot and the trefoil are the two simplest types of knots. Analogously, we use the lifting theorem in this paper to study the two simplest Lagrangian cones: the trivial cone and the Harvey–Lawson special Lagrangian cones. We begin by recalling the construction of the Harvey–Lawson special Lagrangian cone.

Example 1.5. Example III.3.A in [13] introduced one of the first nontrivial families of examples of special Lagrangian cones, collectively known as *the Harvey–Lawson*

cone. In particular, they proved that the cone on the $(n-1)$-tori defined by the following two sets is a special Lagrangian cone:

$$T^+ = \{(e^{i\theta_1}, \dots, e^{i\theta_n}) \in \mathbb{C}^n \mid \theta_1 + \cdots + \theta_n = 0\},$$
$$T^- = \{(e^{i\theta_1}, \dots, e^{i\theta_n}) \in \mathbb{C}^n \mid \theta_1 + \cdots + \theta_n = \pi\}.$$

Observe that we may rewrite T^+ as

$$T^+ = \{(e^{i\theta_1}, \dots, e^{i\theta_{n-1}}, e^{-i(\theta_1+\cdots+\theta_{n-1})}) \mid \theta_1, \dots, \theta_{n-1} \in S^1\}, \tag{1-1}$$

and we will call the cone on T^+ the *Harvey–Lawson cone.*

In [30], Sabloff used combinatorial methods to define a version of Legendrian contact homology for Legendrian knots in circle bundles over Riemann surfaces. We expect that similar methods give rise to a version of Legendrian contact homology in the present context as well. Sabloff's Legendrian contact homology is filtered by the "winding number" of the Reeb chord around the fiber. As such, the short Reeb chords in the 0-filtration level (i.e., those that do not wrap around the fiber) are crucial to any calculation of the homology. In this context, as an application of the lifting theorem we calculate the expected generators of the 0-filtration level of the Legendrian contact homology of the torus given by the intersection of the Harvey–Lawson special Lagrangian cone with S^5 using the standard contact structure α.

Theorem 3.16. *Let $T^2 \subset S^5$ be the torus constructed in Example 3.1, which is Legendrian isotopic to $T^+ \subset S^5$. Then the 0-filtration level of the Legendrian contact homology of T^2 is generated by four pairs of short Reeb chords, two each in gradings 4, 6, 7, and 9. These Reeb chords correspond to the double points of T^2 via the projection of T^2 under $\pi : S^5 \to \mathbb{C}P^2$ (as described in Example 3.1).*

Many of the technical calculations in this paper are devoted to proving this theorem (and Theorem 5.3). The Harvey–Lawson special Lagrangian cone has an associated Legendrian torus in S^5 that is a 3-fold cover of a (standard) Lagrangian torus in $\mathbb{C}P^2$. The isotopies that are used to place this Legendrian torus in general position are delicate and have to be done in steps: first we find projections with double point circles, and then we perturb the resulting surface to obtain one whose projection to $\mathbb{C}P^{n-1}$ has isolated transverse double points. It is only in this carefully orchestrated setup that we can count the double points, and hence the filtration level 0 generators of contact homology. We use a similar approach in Sections 3C and 3D to construct examples of Lagrangian cones arising from products of Legendrian knots.

Example 1.6. The trivial cone is simply a Lagrangian copy of $\mathbb{R}^n \subset \mathbb{C}^n$. In particular, the following is well known and easy to check:

Theorem 1.7. *If $f : \mathbb{R}^n \to \mathbb{C}^n$ is given by $(x_1, \dots, x_n) \mapsto (x_1\eta_1, \dots, x_n\eta_n)$, where $\eta = (\eta_1, \dots, \eta_n)$ is a complex vector with $\eta_j \neq 0$ for all j, then the image of f is Lagrangian with respect to the standard symplectic form ω.*

For some choices of η the trivial cone is special Lagrangian. For example, when $n = 3$ a direct calculation shows that for $\eta = (a_1 + ib_1, a_2 + ib_2, a_3 + ib_3)$, if

$$a_2a_3b_1 + a_1a_3b_2 + a_1a_2b_3 - b_1b_2b_3 = 0$$

then the map $f : \mathbb{R}^3 \to \mathbb{C}^3$ given by

$$(x_1, \dots, x_n) \mapsto (x_1(a_1 + ib_1), x_2(a_2 + ib_2), x_3(a_3 + ib_3))$$

is a special Lagrangian cone.

While the cone is just a copy of $\mathbb{R}^3 \subset \mathbb{C}^3$, its intersection with $S^5 \subset \mathbb{C}^3$ is a copy of S^2 that double covers a copy of $\mathbb{R}P^2$ under the projection $\pi : S^5 \to \mathbb{C}P^2$. For computations of Legendrian contact homology, it is desirable to perturb the cone so that, in the projection, we see only isolated transverse double points. Unlike with the Harvey–Lawson cone, whose link embeds in a single chart (see Section 2), the lift of $\mathbb{R}P^2$ used to study the trivial cone requires the full strength of the lifting theorem.

As with the Harvey–Lawson cone, we use the lifting theorem to obtain a similar theorem about the expected generators of the trivial cone's Legendrian contact homology.

Theorem 5.3. *Let $S \subset S^5$ be the Legendrian 2-sphere obtained from intersecting the trivial cone with S^5 and then perturbing it via Legendrian isotopy to one with transverse double points (see Section 5). Then the 0-filtration level of the Legendrian contact homology of S is generated by 7 pairs of short Reeb chords. These Reeb chords correspond to the double points of the projection of S under $\pi : S^5 \to \mathbb{C}P^2$.*

1C. ***Lagrangian cones given by knot diagrams.*** In [4], pairs of grid diagrams for knots were used to construct immersed Lagrangian tori in $\mathbb{R}^4$, whose lifts to $\mathbb{R}^5$ equipped with the standard contact structure are embedded Legendrian tori. In Sections 3C and 3D, we show how to adapt this construction to produce Legendrian tori in S^5 whose associated cones in $\mathbb{C}^3$ are Lagrangian. This allows us to construct infinite families of Lagrangian cones, some of which may be isotopic to special Lagrangian cones. Future research will explore the question of under what conditions this happens.

1D. ***Outline.*** The remainder of the paper is organized as follows. In Section 2, we discuss the background information leading to the statement of a useful simplification of the lifting theorem (cf. Theorem 2.2), and various examples we can construct using it. In Section 4, we prove the lifting theorem, and in Section 5 we give an example of a lift using it. Section 6 explores the implications of the lifting theorem

for the study of Legendrian submanifolds of S^{2n-1}. Finally, Section 7 introduces some questions regarding the study of Hamiltonian minimal submanifolds using the theorems and examples in this paper.

2. Lifting theorem in a single chart

In this section we develop a special case of the lifting theorem that we use for constructing examples of embedded Legendrian submanifolds of S^{2n-1} as lifts of Lagrangian immersions in $\mathbb{C}P^{n-1}$.

The local theory for lifting Lagrangian immersions into a symplectic manifold to some S^1-bundle over that manifold comes out of the theory of fiber bundles. Given a $2n$-dimensional symplectic manifold (X^{2n}, ω) with an integral symplectic form, let $\pi : L \to X^n$ be the complex line bundle such that $c_1(L) = [\omega]$. By the theory of line bundles (see [12]), we know that there is a 1-form η on the unit circle bundle $P = U(L)$ such that $d\eta = \pi^*(\omega)$. In this case, $i\eta \in \Omega^1(P; i\mathbb{R})$ is called the connection 1-form. If $f : \Sigma^n \to X^{2n}$ is a Lagrangian immersion of a connected n-dimensional manifold Σ, then $[f(\Sigma^n)] \cap [\omega] = 0$ and the pull-back of the S^1-bundle P over Σ is trivial. Given

$$\begin{array}{ccc} f^*(P) & \xrightarrow{F} & P \\ \downarrow & & \downarrow \pi \\ \Sigma & \xrightarrow{f} & X^{2n} \end{array}$$

then $f^*(P) \cong \Sigma \times S^1$. In turn, there exists a section $\sigma : \Sigma \to f^*(P)$ which gives an immersed submanifold $F(\sigma(\Sigma))$ of P (see [34]).

In this setup, η is a contact form for P. In general, $F(\sigma(\Sigma))$ will not be Legendrian with respect to η. However, we can always use η to lift a neighborhood U of $x_0 \in \Sigma$ to a Legendrian submanifold of P as follows: using the diffeomorphism $f^*(P) \cong \Sigma \times S^1$ along with the section $\sigma(x) = (x, 1)$, we can define a trivialization of $P|_U$ by (x, e^{it}) for $x \in U$ and $t \in \mathbb{R}$. For $x \in U$, let γ be a path in U from $\gamma(0) = x_0$ to $\gamma(1) = x_1$. This path gives rise to a path Γ in $P|_U$ using the holonomy of the connection 1-form $F^*(\eta)$. That is, Γ is the unique path such that $\Gamma(0) = (x_0, 1)$, $\pi(\Gamma(s)) = \gamma(s)$, and $F^*(\eta)(\Gamma'(s)) = 0$ for all $s \in (0, 1)$. Define the lift $\tilde{f} : U \to P$ by $\tilde{f}(x) = F(\Gamma(1))$.

This map is independent of the path chosen in the contractible neighborhood U because f is a Lagrangian immersion (the restricted holonomy group at x_0 is trivial).

We can write this holonomy map down explicitly in terms of $\Sigma \times S^1$ and the section σ given by coordinates (x, e^{it}) where $x \in \Sigma$ and $t \in \mathbb{R}$. Suppose

$$F^*(\eta) = k(dt - \tau),$$

where $k \in \mathbb{R}$ is a constant, and $\tau \in \Omega^1(\Sigma)$. The solution Γ is equivalent to a

path $(\gamma(x), e^{it(x)}) \in \Sigma \times S^1$ where

$$t(x) = \int_\gamma \tau$$

is obtained by integrating $dt - \tau$ along γ, setting the result equal to 0, and choosing $t(0) = 0$.

This solution defines a local Legendrian lift, $\tilde{f}$ of U into P. We get a global lift if

$$\int_\gamma \tau \in 2\pi\mathbb{Z} \quad \text{for all } [\gamma] \in H_1(\Sigma).$$

In this case, $f : \Sigma \to X$ lifts to a Legendrian immersion $\tilde{f} : \Sigma \to P$ (i.e., the local lift extends to all of Σ).

If integrating τ along any path joining a pair of double points results in a nonzero answer (mod 2π), then the lift $\tilde{f}$ is an embedding. We summarize the discussion above as follows:

Theorem 2.1. *Let Σ^n be a connected n-manifold, X^{2n} be a $2n$-dimensional symplectic manifold with integral symplectic form ω, and $f : \Sigma \to X$ be a Lagrangian immersion. Let $\pi : P \to X$ be the principle S^1-bundle with connection 1-form $i\eta$ determined by $d\eta = \pi^*(\omega)$. Suppose the section $\sigma : \Sigma \to f^*(P)$ defines coordinates (x, e^{it}) of the trivial bundle $F : f^*(P) \to P$ such that $F^*(\eta) = k(dt - \tau)$ where $k \in \mathbb{R}$ is a constant and $\tau \in \Omega^1(\Sigma)$. If*

(1) $\int_\gamma \tau \in 2\pi\mathbb{Z}$ *for all* $[\gamma] \in H_1(\Sigma; \mathbb{Z})$, *and*

(2) *for all points $x_0, x_1 \in \Sigma$ such that $f(x_0) = f(x_1)$ and any path γ from x_0 to x_1 in Σ, $\int_\gamma \tau \neq 0 \bmod 2\pi$,*

then $f : \Sigma \to X$ lifts to $\tilde{f} : \Sigma \to P$ and the image (the lift) $\tilde{\Sigma}$ is a Legendrian submanifold of P.

Theorem 2.1 is general in that it describes exactly when immersions can be lifted, but it is far from helpful in describing how to construct such lifts by hand (or with the help of a computer). For example, given a symplectic manifold X, like $\mathbb{C}P^n$ (or T^n, $E(n)$, $\mathrm{Sym}^n(\Sigma_g)$, etc), what chart system should we use to make the calculation easiest? (Note the standard chart system $U_i = \{[z_1 : \cdots : 1 : \cdots : z_n] | z_i \in \mathbb{C}\} \subset \mathbb{C}P^{n-1}$ is not convenient for constructing lifts.)

Can a chart system of X be chosen in such a way that the symplectic form ω is standard in each chart? Can a chart system be chosen so that the principal S^1-bundle trivializes over each chart in such a way that η has a nice (simple) form in each trivialization, and there is an obvious choice of sections so that τ also has a nice representation? None of these questions are answered by Theorem 2.1 (because they are specific to X), but all of them are important to being able to generate

explicit examples of lifts that satisfy the restrictive requirements needed to be able to compute invariants like the Legendrian contact homology of the lifts.

For these reasons, the following theorem is useful to us in computing the invariants of Lagrangian cones in $\mathbb{C}^n$ in this paper.

Theorem 2.2. *Let $B^{n-1} \subset \mathbb{C}^{n-1}$ be a ball, Σ be a closed, connected, smooth $(n-1)$-manifold, and $f : \Sigma \to B^{n-1}$ be a Lagrangian immersion with respect to the standard symplectic form ω_0 of $\mathbb{C}^{n-1}$. Let $\tau = -\sum_{i=1}^{n-1}(x_i dy_i - y_i dx_i)$ be a 1-form on B^{n-1}. If*

(1) *$\int_{f(\gamma)} \tau \in 2\pi\mathbb{Z}$ for all $\gamma \in H_1(\Sigma;\mathbb{Z})$, and*

(2) *for all distinct points $x_1, \dots, x_k \in \Sigma$ such that $f(x_1) = f(x_j)$ for all $j \le k$, and a choice of path γ_j from x_1 to x_j in Σ for $2 \le j \le k$, the set $\left\{\left(\int_{f(\gamma_j)} \tau\right) \bmod 2\pi \mid 2 \le j \le k\right\}$ has $k-1$ distinct values, none of which are equal to 0,*

then Σ lifts to an embedded Legendrian submanifold $\tilde{\Sigma} \subset S^{2n-1}$ whose associated cone $c\tilde{\Sigma}$ is Lagrangian in $\mathbb{C}^n$.

The lift, $\tilde{f} : \Sigma \to S^{2n-1} \subset \mathbb{C}^n$, is given by

$$\tilde{f}(x) = e^{it(x)}(f_1(x), \dots, f_{n-1}(x), \sqrt{1 - |f(x)|^2})$$

where

$$t(x) = \int_{f(\gamma)} \tau$$

for some path γ from an initial point $x_0 \in \Sigma$ to x.

Careful comparison of the calculations in Theorem 2.2 with those of Theorem 2.1 shows that Theorem 2.2 is the realization of Theorem 2.1 in the case where Σ^{n-1} is an immersion into an open unit ball, thought of as a single chart of $\mathbb{C}P^{n-1}$ (and where we do the calculations in the chart, instead of in Σ). For a proof of Theorem 2.2, see Section 4, where we prove the lifting theorem, which is a more general version of this theorem.

3. Examples of lifts using Theorem 2.2

3A. *Legendrian contact homology generators for the Harvey–Lawson cone.*

Example 3.1. Theorem 2.2 allows us to construct a family of isotopies of the famous special Lagrangian cone given by Harvey and Lawson (see Example 1.5). Choose ϵ so that $0 \le \epsilon < \sqrt{2/n}$ and define $\delta = \sqrt{1/n - \epsilon^2/2}$. Parametrize the torus T^{n-1} in the usual way with coordinates $(\theta_1, \dots, \theta_{n-1}) \in \mathbb{R}^{n-1}$. Let $r_\epsilon(\theta_1, \dots, \theta_{n-1}) = \delta + \epsilon \sin(\theta_1 + \cdots + \theta_{n-1})$, and define $f_\epsilon : T^{n-1} \to B^{n-1}$ by

$$f_\epsilon(\theta_1, \dots, \theta_{n-1}) = \left(r_\epsilon(\theta_1, \dots, \theta_{n-1})e^{i(2\theta_1+\theta_2+\cdots+\theta_{n-1})}, \dots, r_\epsilon(\theta_1, \dots, \theta_{n-1})e^{i(\theta_1+\cdots+\theta_{n-2}+2\theta_{n-1})}\right).$$

Observe that the first condition of Theorem 2.2 is satisfied. Thus, defining $t(x)$ as in Theorem 2.2, we obtain a family of Legendrian tori in $S^{2n-1} \subset \mathbb{C}^n$, each of whose associated cones are Lagrangian, given by the maps

$$\tilde{f}_\epsilon(\theta_1, \dots, \theta_{n-1}) = e^{it_\epsilon(\theta_1,\dots,\theta_{n-1})} \times \Big(r_\epsilon(\theta_1, \dots, \theta_{n-1}) e^{i(2\theta_1+\theta_2+\cdots+\theta_{n-1})}, \dots, r_\epsilon(\theta_1, \dots, \theta_{n-1}) e^{i(\theta_1+\cdots+\theta_{n-2}+2\theta_{n-1})}, \sqrt{1-(n-1)r_\epsilon^2} \Big),$$

where

$$t_\epsilon(\theta_1, \dots, \theta_{n-1}) = \int_{f_\epsilon(\gamma)} \tau,$$

as in Theorem 2.2.

Remark 3.2. The cone on the image of the lift $\tilde{f}_\epsilon$ is Lagrangian for all $\epsilon \geq 0$, but is also special Lagrangian when $\epsilon = 0$. In fact, when $\epsilon = 0$, the associated cone is the Harvey–Lawson cone (see Example 1.5).

Theorem 3.3. *The parameter t_ϵ is given by*

$$t_\epsilon(\theta_1, \dots, \theta_{n-1}) = -(\theta_1+\cdots+\theta_{n-1}) - 2n\delta\epsilon(1-\cos(\theta_1+\cdots+\theta_{n-1})) + \frac{n}{4}\epsilon^2 \sin(2(\theta_1+\cdots+\theta_{n-1})).$$

Proof. For simplicity, we work in polar coordinates and integrate the pull-back $f_\epsilon{}^*(\tau) = -n\sum_{i=1}^{n-1} r_i^2 d\theta_i$ over a path in the torus T^{n-1} for the computation below. Taking γ_i to be a path from $(\theta_1, \dots, \theta_{i-1}, 0, \dots, 0)$ to $(\theta_1, \dots, \theta_{i-1}, \theta_i, 0, \dots, 0)$, and γ to be the concatenation of these paths from $i = 1, \dots, n$, then we may solve for t_ϵ as follows:

$$\begin{aligned} t_\epsilon(\theta_1, \dots, \theta_{n-1}) &= -n \sum_{i=1}^{n-1} \int_0^{\theta_i} r_\epsilon(\theta_1, \dots, \theta_{i-1}, \alpha_i, 0, \dots, 0)^2 d\alpha_i \\ &= -n \sum_{i=1}^{n-1} \Big[\Big(\tfrac{1}{2}(2\delta^2+\epsilon^2)\alpha_i - 2\delta\epsilon \cos(\theta_1+\cdots+\theta_{i-1}+\alpha_i) \\ &\qquad - \tfrac{1}{4}\epsilon^2 \sin(2(\theta_1+\cdots+\theta_{i-1}+\alpha_i)) \Big) \Big|_0^{\theta_i} \Big] \end{aligned}$$

Observe that the sum above telescopes, and hence, we may write

$$\begin{aligned} t_\epsilon(\theta_1, \dots, \theta_{n-1}) &= -n\Big(\tfrac{1}{2}(2\delta^2+\epsilon^2)(\theta_1+\cdots+\theta_{n-1}) - 2\delta\epsilon(1-\cos(\theta_1+\cdots+\theta_{n-1})) \\ &\qquad - \tfrac{1}{4}\epsilon^2 \sin(2(\theta_1+\cdots+\theta_{n-1})) \Big) \\ &= -(\theta_1+\cdots+\theta_{n-1}) - 2n\delta\epsilon(1-\cos(\theta_1+\cdots+\theta_{n-1})) \\ &\qquad + \frac{n}{4}\epsilon^2 \sin(2(\theta_1+\cdots+\theta_{n-1})). \end{aligned}$$

□

In light of Theorem 3.3, we get the following corollary.

Corollary 3.4. *As* $\epsilon \to 0$, $\delta \to 1/\sqrt{n}$, $t_\epsilon(\theta_1, \dots, \theta_{n-1}) \to t_0(\theta_1, \dots, \theta_{n-1}) = -\theta_1 - \cdots - \theta_{n-1}$, *and*

$$\tilde{f}_\epsilon(\theta_1, \dots, \theta_{n-1}) \to \tilde{f}_0(\theta_1, \dots, \theta_{n-1}) = \frac{1}{\sqrt{n}}(e^{i\theta_1}, \dots, e^{i\theta_{n-1}}, e^{-i(\theta_1+\cdots+\theta_{n-1})}).$$

In order to verify that the second condition of Theorem 2.2 is satisfied, and consequently that the lift is embedded, we will be interested in locating the double points of f_ϵ.

Because we are mainly interested in cones of $\mathbb{C}^3$ via the SYZ conjecture, we assume $n = 3$ in the following calculation. Lemma 3.5 specifies precisely when the arguments of the exponential maps in the definition of f_ϵ all agree, a necessary condition for a double point.

Lemma 3.5. *For* $n = 3$, *if* $f_\epsilon(\theta_1, \theta_2) = f_\epsilon(\gamma_1, \gamma_2)$ *then* $\theta_1 = \gamma_1$ *and* $\theta_2 = \gamma_2$, *or* $\theta_1 - \gamma_1 = \theta_2 - \gamma_2 = \frac{2\pi}{3} \pmod{2\pi}$ *or* $\theta_1 - \gamma_1 = \theta_2 - \gamma_2 = \frac{4\pi}{3} \pmod{2\pi}$.

Proof. If $f_\epsilon(\theta_1, \theta_2) = f_\epsilon(\gamma_1, \gamma_2)$ then since the arguments of the exponential maps differ by a multiple of 2π, (θ_1, θ_2) and (γ_1, γ_2) must satisfy the equations

$$2\theta_1 + \theta_2 = 2\gamma_1 + \gamma_2 + n2\pi, \tag{3-1}$$

$$\theta_1 + 2\theta_2 = \gamma_1 + 2\gamma_2 + m2\pi, \tag{3-2}$$

for some $m, n \in \mathbb{Z}$.

Solving (3-1) and (3-2), we obtain the following:

$$\theta_1 - \gamma_1 = \frac{2n - m}{3}2\pi, \tag{3-3}$$

$$\theta_2 - \gamma_2 = \frac{2m - n}{2}\pi. \tag{3-4}$$

Since the torus T^2 is parametrized by $(\theta_1, \theta_2) \in [0, 2\pi) \times [0, 2\pi)$, it must be that $\theta_i - \gamma_i < 2\pi$ for $i = 1, 2$, and hence $\left|\frac{2m-n}{3}\right| < 1$ and $\left|\frac{2n-m}{3}\right| < 1$.

Since $n, m \in \mathbb{Z}$, we find that the possibilities for (n, m) are $\pm(1, 0)$, $\pm(0, 1)$, $\pm(1, 1)$ and $(0, 0)$. Evaluating (3-3) and (3-4), we find that either $\theta_1 = \gamma_1$ and $\theta_2 = \gamma_2$, or $\theta_1 - \gamma_1 = \theta_2 - \gamma_2 = \frac{2\pi}{3} \pmod{2\pi}$ or $\theta_1 - \gamma_1 = \theta_2 - \gamma_2 = \frac{4\pi}{3} \pmod{2\pi}$. □

In the proof above we also showed, after taking limits, that:

Scholium 3.6. *The image of* $\tilde{f}_0$ *is a* 3*-fold cover of the image of* f_0 *via the projection given by the Hopf map.*

Lemma 3.5 specifies when the arguments of the exponential maps will agree, but for a double point, the radii, determined by r_ϵ must also agree. In the following lemma, we calculate where this occurs.

Lemma 3.7. *If* $f_\epsilon(\theta_1, \theta_2) = f_\epsilon(\gamma_1, \gamma_2)$ *and either* $\theta_1 - \gamma_1 = \theta_2 - \gamma_2 = \frac{2\pi}{3} \pmod{2\pi}$ *or* $\theta_1 - \gamma_1 = \theta_2 - \gamma_2 = \frac{4\pi}{3} \pmod{2\pi}$, *then one of the following must be true*:

- $\theta_1+\theta_2=\gamma_1+\gamma_2$.
- $\theta_1+\theta_2=\frac{7\pi}{6}$ *and* $\gamma_1+\gamma_2=\frac{11\pi}{6}$.
- $\theta_1+\theta_2=\frac{5\pi}{6}$ *and* $\gamma_1+\gamma_2=\frac{\pi}{6}$.

Proof. Since $f_\epsilon(\theta_1,\theta_2)=f_\epsilon(\gamma_1,\gamma_2)$, not only must the arguments of the exponential maps differ by a multiple of 2π, but the radii in each complex factor must match, that is $r_\epsilon(\theta_1,\theta_2)=r_\epsilon(\gamma_1,\gamma_2)$. Hence one of the following equations must hold:

$$\theta_1+\theta_2=\gamma_1+\gamma_2, \tag{3-5}$$

$$\theta_1+\theta_2+\gamma_1+\gamma_2=\pi+2\pi k. \tag{3-6}$$

There are several cases. If $\theta_1+\theta_2=\gamma_1+\gamma_2$, then using (3-3) and (3-4), one can show that $n=-m$ which can only happen if $n=m=0$. Furthermore, if $\theta_1+\theta_2+\gamma_1+\gamma_2=\pi+k2\pi$, combining this with (3-3) and (3-4), we may solve the system to obtain that $\theta_1+\theta_2=\frac{7\pi}{6}$ and $\gamma_1+\gamma_2=\frac{11\pi}{6}$ or $\theta_1+\theta_2=\frac{5\pi}{6}$ and $\gamma_1+\gamma_2=\frac{\pi}{6}$. □

Remark 3.8. Lemma 3.5 rules out the possibility of multiple points of f_ϵ of multiplicity greater than 3, and Lemma 3.7 shows that for $\epsilon>0$ there are no triple points. Hence, immersion f_ϵ has only double points when $\epsilon>0$.

The families of double points identified in Lemma 3.7 form copies of S^1, and will show up not only in this example, but in others as well. Hence the following definition will be useful in some of the discussion that follows.

Definition 3.9. Let $f:\Sigma\to M$ be an immersion of a surface. Suppose C_1 and C_2 are disjoint copies of S^1 in Σ such that $f(C_1)=f(C_2)$ and $f|_{C_1\cup C_2}$ is a 2-to-1 map. Suppose further that A_1 and A_2 are disjoint annular neighborhoods of C_1 and C_2 and that $f(A_1)\cap f(A_2)=f(C_1)=f(C_2)$. If, for any pair consisting of $x_1\in C_1$ and $x_2\in C_2$ such that $f(x_1)=f(x_2)$, we have that $df_{x_1}(TA_1)\neq df_{x_2}(TA_2)$, then we call the image of C_1 and C_2 a *double point circle*.

Theorem 3.10. *The double points of f_ϵ, of the form $f_\epsilon(\theta_1,\theta_2)=f_\epsilon(\gamma_1,\gamma_2)$, consist of two double point circles such that $\theta_1-\gamma_1=\theta_2-\gamma_2=\frac{2\pi}{3}\pmod{2\pi}$ or $\theta_1-\gamma_1=\theta_2-\gamma_2=\frac{4\pi}{3}\pmod{2\pi}$ and one of the following holds*:

(1) $\theta_1+\theta_2=\frac{7\pi}{6}$ *and* $\gamma_1+\gamma_2=\frac{11\pi}{6}$.

(2) $\theta_1+\theta_2=\frac{5\pi}{6}$ *and* $\gamma_1+\gamma_2=\frac{\pi}{6}$.

Proof. Lemmas 3.5 and 3.7 demonstrate that systems of this type yield double points. All that remains is the observation that if (θ_1,θ_2) and (γ_1,γ_2) satisfy $\theta_1-\gamma_1=\theta_2-\gamma_2=\frac{2\pi}{3}\pmod{2\pi}$ or $\theta_1-\gamma_1=\theta_2-\gamma_2=\frac{4\pi}{3}\pmod{2\pi}$ but do not satisfy either (1) or (2), then $\sin(\theta_1+\theta_2)\neq\sin(\gamma_1+\gamma_2)$. For such cases, $r_\epsilon(\theta_1,\theta_2)\neq r_\epsilon(\gamma_1,\gamma_2)$ and hence $f_\epsilon(\theta_1,\theta_2)\neq f_\epsilon(\gamma_1,\gamma_2)$. □

Theorem 3.11. *The lift $\tilde{f}_\epsilon$ is an embedding.*

Proof. We already know the lift is well defined. All that remains is to check that the second condition of Theorem 2.2 is satisfied, which means that the double points of the projection are separated in the lift. This amounts to computing $\int_{f(\gamma)} \tau$ for some path γ joining a pair of double points of a double point circle. Using Theorem 3.10, suppose we have a double point such that $f_\epsilon(\theta_1, \frac{5\pi}{6} - \theta_1) = f_\epsilon\big(\theta_1 + \frac{2\pi}{3}, \frac{13\pi}{6} - (\theta_1 + \frac{2\pi}{3})\big)$. Then the integral in question is given by:

$$t_\epsilon\left(\theta_1 + \frac{2\pi}{3}, \frac{13\pi}{6} - \left(\theta_1 + \frac{2\pi}{3}\right)\right) - t_\epsilon\left(\theta_1, \frac{5\pi}{6} - \theta_1\right).$$

Using the expression for t_ϵ given in Theorem 3.3, and simplifying, we obtain

$$t_\epsilon\left(\theta_1 + \frac{2\pi}{3}, \frac{13\pi}{6} - \left(\theta_1 + \frac{2\pi}{3}\right)\right) - t_\epsilon\left(\theta_1, \frac{5\pi}{6} - \theta_1\right)$$
$$= -\frac{8\pi}{6} - 4n\delta\epsilon\cos\left(\frac{5\pi}{6}\right) + \frac{n\epsilon^2}{2}\sin\left(\frac{\pi}{3}\right).$$

Noting that $n = 3$, $0 \le \epsilon < \sqrt{2/3}$, and $\delta = \sqrt{1/3 - \epsilon^2/2}$, we have that

$$-\frac{4\pi}{3} \le t_\epsilon\left(\theta_1 + \frac{2\pi}{3}, \frac{13\pi}{6} - \left(\theta_1 + \frac{2\pi}{3}\right)\right) - t_\epsilon\left(\theta_1, \frac{5\pi}{6} - \theta_1\right) < -\frac{4\pi}{3} + \frac{\sqrt{3}}{2}.$$

The other double points are handled in a similar manner. □

Let L_ϵ be the image of f_ϵ and let $\tilde{L}_\epsilon$ be the Legendrian torus given by the lift $\tilde{f}_\epsilon$. We wish to identify the generators of the 0-filtration level of the Legendrian contact homology of $\tilde{L}_\epsilon$, which are determined by the double points of the Lagrangian projection. Recall that in this case, the double points are actually double point circles, hence we need to perturb the map so that it is chord-generic. We will demonstrate the perturbation for $n = 3$, but the general solution is similar.

Lemma 3.12. *Let $\tilde{f}_\epsilon : T^2 \to S^5$ be the Legendrian torus given by the map*

$$\tilde{f}_\epsilon(\theta_1, \theta_2) = e^{it_\epsilon(\theta_1,\theta_2)}\Big(r_\epsilon(\theta_1,\theta_2)e^{i(2\theta_1+\theta_2)}, r_\epsilon(\theta_1,\theta_2)e^{i(\theta_1+2\theta_2)}, \sqrt{1-2r_\epsilon(\theta_1,\theta_2)^2}\Big).$$

Choose a perturbation in the direction of the Reeb fiber, $s_\epsilon : T^2 \to S^1$, two perturbations in the radial directions, $s_{i,\epsilon} : T^2 \to \mathbb{R}$, for $i = 1, 2$, and define

$$\tilde{g}_\epsilon(\theta_1, \theta_2)$$
$$= e^{i(t_\epsilon(\theta_1,\theta_2)+s_\epsilon(\theta_1,\theta_2))}\Big(r_{1,\epsilon}(\theta_1,\theta_2)e^{i(2\theta_1+\theta_2)}, r_{2,\epsilon}(\theta_1,\theta_2)e^{i(\theta_1+2\theta_2)}, \sqrt{1-r_{1,\epsilon}^2-r_{2,\epsilon}^2}\Big),$$

where $r_{i,\epsilon}(\theta_1, \theta_2) = r_\epsilon(\theta_1, \theta_2) + s_{i,\epsilon}(\theta_1, \theta_2)$ for $i = 1, 2$. If

(1) $\frac{\partial s_\epsilon}{\partial \theta_1} + 2r_\epsilon(\theta_1,\theta_2)(2s_{1,\epsilon}(\theta_1,\theta_2) + s_{2,\epsilon}(\theta_1,\theta_2)) + 2s_{1,\epsilon}(\theta_1,\theta_2)^2 + s_{2,\epsilon}(\theta_1,\theta_2)^2 = 0$ *and*

(2) $\frac{\partial s_\epsilon}{\partial \theta_2} + 2r_\epsilon(\theta_1,\theta_2)(s_{1,\epsilon}1(\theta_1,\theta_2) + 2s_{2,\epsilon}(\theta_1,\theta_2)) + s_{1,\epsilon}(\theta_1,\theta_2)^2 + 2s_{2,\epsilon}(\theta_1,\theta_2)^2 = 0$

then the perturbation $\tilde{g}_\epsilon$ is a Legendrian torus having only transverse double points that is Legendrian isotopic to $\tilde{f}_\epsilon$.

Moreover, for a given choice of s_ϵ the system is solved by

$$s_{1,\epsilon}(\theta_1,\theta_2) = -r_\epsilon(\theta_1,\theta_2) + \sigma\sqrt{r_\epsilon(\theta_1,\theta_2)^2 + \frac{1}{3}\left(\frac{\partial s_\epsilon}{\partial\theta_2} - 2\frac{\partial s_\epsilon}{\partial\theta_1}\right)}$$

and

$$s_{2,\epsilon}(\theta_1,\theta_2) = -r_\epsilon(\theta_1,\theta_2) + \sigma\sqrt{r_\epsilon(\theta_1,\theta_2)^2 + \frac{1}{3}\left(\frac{\partial s_\epsilon}{\partial\theta_1} - 2\frac{\partial s_\epsilon}{\partial\theta_2}\right)},$$

where σ is ± 1.

Proof. The calculation is easiest if we work in polar coordinates and identify a neighborhood of the $\tilde{f}_\epsilon$ with $B_2 \times S^1$ (cf. the lifting theorem). Note that we may write

$$\tilde{f}_\epsilon(\theta_1,\theta_2) = \big(r_\epsilon(\theta_1,\theta_2), 2\theta_1+\theta_2, r_\epsilon(\theta_1,\theta_2), \theta_1+2\theta_2, t_\epsilon(\theta_1,\theta_2)\big),$$

and we work with the perturbation in polar coordinates as well:

$$\tilde{g}_\epsilon(\theta_1,\theta_2) = \big(r_{1,\epsilon}(\theta_1,\theta_2), 2\theta_1+\theta_2, r_{2,\epsilon}(\theta_1,\theta_2), \theta_1+2\theta_2, t_\epsilon(\theta_1,\theta_2)+s_\epsilon(\theta_1,\theta_2)\big),$$

In these coordinates, we may identify the contact form α on S^5 with $\frac{1}{2}(dt-\tau)$ (for details of this calculation see the lifting theorem). Pulling back α to T^2 via $\tilde{f}_\epsilon$ we obtain the form

$$\tilde{f}_\epsilon^*(\alpha) = \left(\frac{\partial t_\epsilon}{\partial\theta_1} + 3r_\epsilon(\theta_1,\theta_2)^2\right)d\theta_1 + \left(\frac{\partial t_\epsilon}{\partial\theta_2} + 3r_\epsilon(\theta_1,\theta_2)^2\right)d\theta_2.$$

Since $\tilde{f}_\epsilon$ is Legendrian, this is 0, and hence

$$\left(\frac{\partial t_\epsilon}{\partial\theta_1} + 3r_\epsilon(\theta_1,\theta_2)^2\right) = \left(\frac{\partial t_\epsilon}{\partial\theta_2} + 3r_\epsilon(\theta_1,\theta_2)^2\right) = 0.$$

Pulling back α using the perturbation $\tilde{g}_\epsilon$ we obtain

$$\begin{aligned}\tilde{g}_\epsilon^*(\alpha) = &\left[\left(\frac{\partial t_\epsilon}{\partial\theta_1}+3r_\epsilon(\theta_1,\theta_2)^2\right)+\frac{\partial s_\epsilon}{\partial\theta_1}+2r_\epsilon(\theta_1,\theta_2)(2s_{1,\epsilon}+s_{2,\epsilon})+2s_{1,\epsilon}^2+s_{2,\epsilon}^2\right]d\theta_1\\ &+\left[\left(\frac{\partial t_\epsilon}{\partial\theta_2}+3r_\epsilon(\theta_1,\theta_2)^2\right)\frac{\partial s_\epsilon}{\partial\theta_2}+2r_\epsilon(\theta_1,\theta_2)(s_{1,\epsilon}+2s_{2,\epsilon})+s_{1,\epsilon}^2+2s_{2,\epsilon}^2\right]d\theta_2.\end{aligned}$$

Noting that $\left(\frac{\partial t_\epsilon}{\partial\theta_1} + 3r_\epsilon(\theta_1,\theta_2)^2\right) = \left(\frac{\partial t_\epsilon}{\partial\theta_2} + 3r_\epsilon(\theta_1,\theta_2)^2\right) = 0$, we have justified (1) and (2). The last part is routine, and obtained by solving this system of equations, (1) and (2), for $s_{1,\epsilon}$ and $s_{2,\epsilon}$. □

Theorem 3.13. *The map $g_\epsilon : T^2 \to B^2$,*

$$g_\epsilon(\theta_1,\theta_2) = (r_{1,\epsilon}(\theta_1,\theta_2)e^{i(2\theta_1+\theta_2)}, r_{2,\epsilon}(\theta_1,\theta_2)e^{i(\theta_1+2\theta_2)}),$$

where

$$r_{1,\epsilon}(\theta_1,\theta_2)=\sqrt{r_\epsilon(\theta_1,\theta_2)^2-\tfrac{2}{3}\epsilon\cos(\theta_1)}\quad\text{and}\quad r_{2,\epsilon}(\theta_1,\theta_2)=\sqrt{r_\epsilon(\theta_1,\theta_2)^2+\tfrac{1}{3}\epsilon\cos(\theta_1)}$$

is a perturbation of f_ϵ having exactly two transverse double points. Moreover, the lift $\tilde{g}_\epsilon$,

$$\begin{aligned}&\tilde{g}_\epsilon(\theta_1,\theta_2)\\&=e^{i(t_\epsilon(\theta_1,\theta_2)+s_\epsilon(\theta_1,\theta_2))}\Big(r_{1,\epsilon}(\theta_1,\theta_2)e^{i(2\theta_1+\theta_2)},\, r_{2,\epsilon}(\theta_1,\theta_2)e^{i(\theta_1+2\theta_2)},\,\sqrt{1-r_{1,\epsilon}^2-r_{2,\epsilon}^2}\Big),\end{aligned}$$

is Legendrian isotopic to $\tilde{f}_\epsilon$.

Proof. Choose $s_\epsilon(\theta_1,\theta_2)=\epsilon\sin(\theta_1)$. Direct calculation shows that the conditions of Lemma 3.12 are satisfied. Moreover, the two maps $s_{1,\epsilon}$ and $s_{2,\epsilon}$ from Lemma 3.12 satisfy the following:

(1) $r_{1,\epsilon}(\theta_1,\theta_2)=r_\epsilon(\theta_1,\theta_2)+s_{1,\epsilon}(\theta_1,\theta_2)=\sqrt{r_\epsilon(\theta_1,\theta_2)^2-\tfrac{2}{3}\epsilon\cos(\theta_1)}$.

(2) $r_{2,\epsilon}(\theta_1,\theta_2)=r_\epsilon(\theta_1,\theta_2)+s_{2,\epsilon}(\theta_1,\theta_2)=\sqrt{r_\epsilon(\theta_1,\theta_2)^2+\tfrac{1}{3}\epsilon\cos(\theta_1)}$.

The remainder follows from Lemma 3.12. □

The following corollary is obvious:

Corollary 3.14. *Taking the limit as $\epsilon\to 0$, we have the following*:

(1) $t_\epsilon(\theta_1,\theta_2)\to t_0(\theta_1,\theta_2)=-\theta_1-\theta_2$.

(2) $\tilde{g}_\epsilon(\theta_1,\theta_2)\to\tilde{g}_0(\theta_1,\theta_2)=\tilde{f}_0(\theta_1,\theta_2)=\frac{1}{\sqrt{2}}(e^{i\theta_1},e^{i\theta_2},e^{-i(\theta_1+\theta_2)})$.

Corollary 3.14 shows that $\tilde{g}_0$ is the Harvey–Lawson cone (just as $\tilde{f}_0$ is). What makes $\tilde{g}_\epsilon$ useful is that although it is isotopic to the Harvey–Lawson cone, it has isolated double points. In fact, it has only four transverse double points as observed in the following corollary.

Corollary 3.15. *The double points of g_ϵ can be found directly, and we obtain 2 for each double point circle, for a total of four transverse double points*:

(1) $g_\epsilon\big(\frac{2\pi}{3},\frac{\pi}{6}\big)=g_\epsilon\big(\frac{4\pi}{3},\frac{5\pi}{6}\big)$,

(2) $g_\epsilon\big(\frac{5\pi}{3},\frac{7\pi}{6}\big)=g_\epsilon\big(\frac{\pi}{3},\frac{11\pi}{6}\big)$,

(3) $g_\epsilon\big(\frac{2\pi}{3},\frac{7\pi}{6}\big)=g_\epsilon\big(\frac{4\pi}{3},\frac{11\pi}{6}\big)$, *and*

(4) $g_\epsilon\big(\frac{5\pi}{3},\frac{\pi}{6}\big)=g_\epsilon\big(\frac{\pi}{3},\frac{5\pi}{6}\big)$.

Proof. Writing g_ϵ in polar coordinates, as in Lemma 3.5, we see that any double points must be of the form $g_\epsilon(\theta_1,\theta_2)=g_\epsilon(\theta_1+j\frac{2\pi}{3},\theta_2+j\frac{2\pi}{3})$ where j is either 1 or 2, in order that the arguments of the exponential maps both differ by a multiple

of 2π. Thus we get double points when we have the following two equations satisfied:

$$r_{1,\epsilon}(\theta_1, \theta_2) = r_{1,\epsilon}\left(\theta_1 + j\frac{2\pi}{3}, \theta_2 + j\frac{2\pi}{3}\right).$$
$$r_{2,\epsilon}(\theta_1, \theta_2) = r_{2,\epsilon}\left(\theta_1 + j\frac{2\pi}{3}, \theta_2 + j\frac{2\pi}{3}\right).$$

Solving this system of equations, we obtain the result. □

In summary, we have constructed a family of cones, each of which is isotopic to the Harvey–Lawson cone, but with the additional property that the projection to $\mathbb{C}P^2$ has only four transverse double points, unlike the actual Harvey–Lawson cone which is a 3-fold cover of its projection to $\mathbb{C}P^2$, as observed in Scholium 3.6. Although the isotopy taking the Harvey–Lawson cone to one of our perturbations does not preserve the special Lagrangian conditions, it does preserve the Legendrian link, and hence, can be used to calculate a suitably defined Legendrian contact homology [30]. Moreover, our perturbations have only transverse double points. Thus we obtain:

Theorem 3.16. *Let $T^2 \subset S^5$ be the torus constructed above, which is Legendrian isotopic to $T^+ \subset S^5$. Then the 0-filtration level of the Legendrian contact homology of T^2 is generated by four pairs of short Reeb chords, two each in gradings* 4, 6, 7, *and* 9. *These Reeb chords correspond to the double points of T^2 via the projection of T^2 under $\pi : S^5 \to \mathbb{C}P^2$.*

Remark 3.17. The lifting theorem made it possible to compute the gradings of Theorem 3.16 explicitly in Mathematica. By working in a single chart, we integrate to define the lift, and compute a unitary Lagrangian frame to obtain the Maslov index. The calculations, though long, are straightforward and therefore omitted.

Remark 3.18. While the Legendrian contact homology of the Harvey–Lawson cone is beginning to emerge in the previous theorem, it does not take into account the Reeb chords that wrap around the fiber. However, considering the gradings of the short chords, it does appear that there is nontrivial homology in gradings 4 and 9.

3B. ***Lagrangian hypercube diagrams.*** Next, we show how to generalize the calculations above to get knotted Legendrian tori in S^5 (knotted in the sense that they are the product of two Legendrian knots in $\mathbb{R}^3$; see [4]). The cones on these knotted tori are Lagrangian cones in $\mathbb{C}^3$. Therefore we begin the study of diagrammatic Lagrangian cones in $\mathbb{C}^3$.

In [4], Lagrangian hypercube diagrams were used to produce examples of Legendrian tori in the standard contact space, $(\mathbb{R}^5, \xi_{std})$, using $wxyzt$-coordinates on $\mathbb{R}^5$

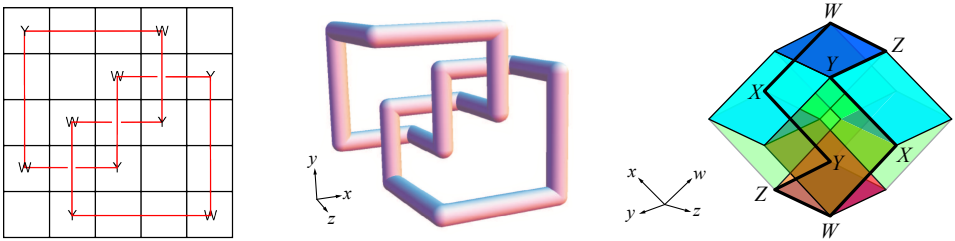


Figure 1. Grid and cube diagrams for the trefoil, and a hypercube diagram for a torus.

and letting $\xi_{std} = dt - ydw - xdz$. But they can also be adapted to produce Legendrian tori in S^5 whose cones in $\mathbb{C}^3$ are Lagrangian. Before doing so, we briefly recall some of the relevant material from [4] and refer the reader to that paper for more details.

Lagrangian hypercube diagrams are closely related to grid, cube, and hypercube diagrams. To construct a grid, cube, or hypercube diagram, one places markings in a 2-, 3-, or 4-dimensional Cartesian grid, while ensuring that certain marking conditions and crossing conditions hold (see Section 2 and 3 in [2], and Section 2 in [3]). In each case, the markings determine a link (see Figure 1). For a hypercube diagram, there is an algorithm for constructing a Lagrangian torus associated to the hypercube diagram, such as the one shown in the last picture in Figure 1 (see Theorem 5.1 in [2]).

In order to define a Lagrangian hypercube diagram, we first need to define a Lagrangian grid diagram:

Definition 3.19. A *Lagrangian grid diagram* given by $\gamma : S^1 \to \mathbb{R}^2$ where $\gamma(\theta) = (x(\theta), y(\theta))$ is an immersed grid diagram G satisfying conditions (3-7) and (3-8):

$$\int_0^{2\pi} y(\theta)x'(\theta)d\theta = 0, \tag{3-7}$$

$$\int_{\theta_0}^{\theta_1} y(\theta)x'(\theta)d\theta \neq 0 \text{ whenever } \gamma(\theta_0) = \gamma(\theta_1) \text{ and } 0 < \theta_1 - \theta_0 < 2\pi. \tag{3-8}$$

While any Lagrangian projection of a Legendrian knot satisfies (3-7) and (3-8), it is usually difficult to determine from a given diagram in the plane whether or not the diagram will lift to a Legendrian knot. The advantage with a Lagrangian grid diagram is that one merely needs to add up the signed areas of a finite number of rectangles to determine whether the diagram lifts to a Legendrian knot (see Corollary 3.10, Scholium 3.12 and Corollary 3.13 in [4]).

A Lagrangian hypercube diagram takes two Lagrangian grid diagrams and uses them to construct a *product* of two Legendrian knots (see [4] and [16]). To construct a grid diagram, one places markings in a 2-dimensional grid, subject to a set of marking conditions, and creates a knot diagram by drawing segments, joining

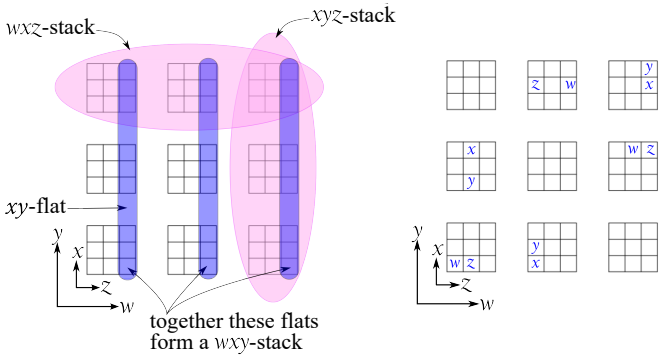


Figure 2. A schematic for displaying a Lagrangian hypercube diagram. The outer w and y coordinates indicate the "level" of each zx-flat. The inner z and x coordinates start at $(0, 0)$ for each of the nine zx-flats. With these conventions understood, one can display xy-flats, xyz-stacks, wxz-stacks, wxy-stacks, etc. The second picture is a schematic of a Lagrangian hypercube diagram.

the markings to create immersed loops. The process of creating a Lagrangian hypercube diagram is similar: there is a set of marking conditions that determine how to place markings in a 4-dimensional Cartesian grid, and the markings are joined by segments, following an algorithm to create a simple loop. Before stating the conditions, we give a few preliminaries.

A *flat* is any right rectangular 4-dimensional polytope with integer valued vertices in C such that there are two orthogonal edges at a vertex of length n and the remaining two orthogonal edges are of length 1. (Each flat is congruent to the product of a unit square and an $n \times n$ square.) Moreover, the flat will be named by the two edges of length n. Although a flat is a 4-dimensional object, the name references the fact that a flat is a 2-dimensional array of unit hypercubes. For example, an xy-flat is a flat that has a face that is an $n \times n$ square that is parallel to the xy-plane. In a hypercube of size $n = 3$, one example of a xy-flat would be the subset $[0, 1] \times [0, 3] \times [0, 3] \times [2, 3]$ (shown in Figure 2).

A *stack* is a set of n flats that form a right rectangular 4-dimensional polytope with integer vertices in C in which there are three orthogonal edges of length n at a vertex, and the remaining edge has length 1. (Each stack is the product of a cube with edges of length n and a unit interval.) A stack is named by the three edges of length n. An example of a wxz-stack in a hypercube of size 3 is the subset $[0, 3] \times [0, 3] \times [2, 3] \times [0, 3]$ (shown at the top of Figure 2). Further examples of flats and stacks may be found in Figure 2.

A marking is a labeled point in $\mathbb{R}^4$ with half-integer coordinates in C. Unit hypercubes of the 4-dimensional Cartesian grid will either be blank, or marked with a W, X, Y, or Z such that the following *marking conditions* hold:

(1) Each stack has exactly one W, one X, one Y, and one Z marking.

(2) Each stack has exactly two flats containing exactly three markings in each.

(3) For each flat containing exactly three markings, the markings in that flat form a right angle such that each ray is parallel to a coordinate axis.

(4) For each flat containing exactly three markings, the marking that is the vertex of the right angle is W if and only if the flat is a zw-flat, X if and only if the flat is a wx-flat, Y if and only if the flat is a xy-flat, and Z if and only if the flat is a yz-flat.

Condition (4) rules out the possibility of either wy-flats or a zx-flats with three markings (see Figure 2). As with oriented grid diagrams and cube diagrams, we obtain an oriented link from the markings by connecting each W marking to an X marking by a segment parallel to the w-axis, each X marking to a Y marking by a segment parallel to the x-axis, and so on.

Let $\pi_{xz}, \pi_{wy} : \mathbb{R}^4 \to \mathbb{R}^2$ be the natural projections, projecting out the x, z and w, y directions respectively. The projection $\pi_{xz}(C)$ produces an $n \times n$ square in the wy-plane. If we project the W and Y markings of the hypercube to this square as well, the markings satisfy the conditions for an immersed grid diagram, which we denote $G_{wy} := (\pi_{xz}(C), \pi_{xz}(\mathcal{W}), \pi_{xz}(\mathcal{Y}))$, where $\mathcal{W}$ and $\mathcal{Y}$ are the sets of W and Y markings, respectively. Similarly, we define $G_{zx} := (\pi_{wy}(C), \pi_{wy}(\mathcal{Z}), \pi_{wy}(\mathcal{X}))$, where $\mathcal{Z}$ and $\mathcal{X}$ are the sets of Z and X markings respectively.

In a grid diagram, one typically requires a crossing condition, namely that the vertical segment crosses over the horizontal segment. For a Lagrangian hypercube diagram, the crossing conditions are determined as follows. We require that the two immersed grid diagrams, G_{zx} and G_{wy}, are Lagrangian grid diagrams (that is, they satisfy conditions (3-7) and (3-8)). By Proposition 3.4 of [4], a Lagrangian grid diagram lifts to a smoothly embedded Legendrian knot. Hence the crossing conditions of the grid are determined by this lift. We require one additional *product lift condition* that the pair G_{zx} and G_{wy} must satisfy. In the definition below, $\Delta t(c)$ is the length of the Reeb chord associated to the crossing c).

Definition 3.20. For two Lagrangian grid diagrams, G_{wy} and G_{zx}, let $\mathcal{C} = \{c_i\}$ be the crossings in G_{zx} and $\mathcal{C}' = \{c_i'\}$ be the crossings in G_{wy}. The pair of grid diagrams is said to satisfy the *product lift condition* if $|\Delta t(c_i)| \neq |\Delta t(c_i')|$ for all i, j.

We are now ready to define a Lagrangian hypercube diagram (see [4]):

Definition 3.21. A *Lagrangian hypercube diagram*, which we denote by $H\Gamma = (C, \{\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}, G_{zx}, G_{wy})$, is a set of markings $\{\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ in C that satisfy the marking conditions, where G_{wy} and G_{zx} are Lagrangian grid diagrams, and G_{wy} and G_{zx} satisfy the product lift condition.

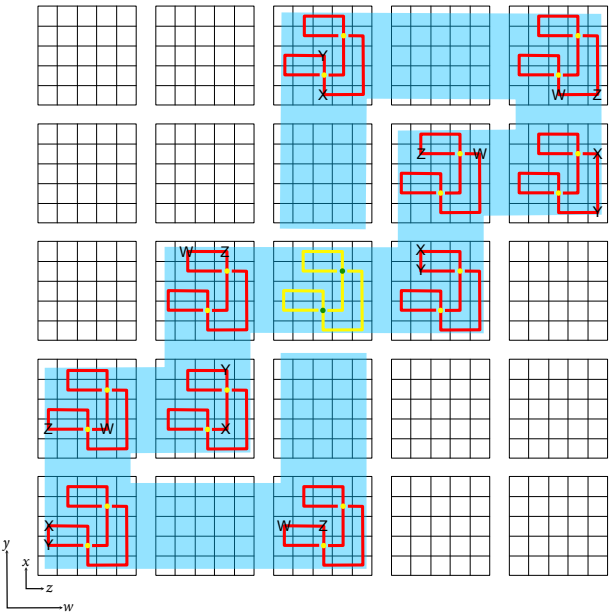


Figure 3. Lagrangian hypercube diagram with unknotted G_{zx} and G_{wy} and rotation class $(1, 0)$.

The immersed torus specified by the Lagrangian hypercube diagram is the product of G_{zx} and G_{wy}, determined as follows: place a copy of the immersed grid G_{zx} at each zx-flat on the schematic that contains a pair of markings (shown in red on Figure 3). Doing so produces a schematic with two copies of G_{zx} with the same y-coordinates and two with the same w-coordinates. For each pair of copies sharing the same w-coordinates, we may translate one parallel to the w-axis toward the other. Doing so traces out an immersed tube connecting these two copies of G_{zx}. Similarly, we may translate parallel to the y-axis to produce an immersed tube connecting two copies of G_{zx} with the same y-coordinates. Since we are connecting copies of G_{zx} in flats corresponding to the markings of G_{wy}, the tube will close to produce an immersed torus.

3C. ***Lagrangian cones in $\mathbb{C}^3$ constructed from Lagrangian hypercube diagrams.*** First, we show how to convert a grid diagram to a *radial grid diagram*. A set of concentric circles $\{C_k\}_{k=1}^{n}$ of radius $\sqrt{k/(3n)}$ will serve to represent the rows of our grid, and a set of radial lines, determined by the list of angles, $\left\{k\frac{2\pi}{n}\right\}_{k=0}^{n-1}$, to serve as columns. The counterclockwise direction is chosen to correspond to the positive x-direction in the original grid, and the outward pointing radial direction is chosen to correspond to the positive y-direction. Moreover, the radii of the concentric circles are chosen so that each annular band has area $\frac{\pi}{3n}$ and consequently, each cell, as shown in Figure 4, has equal area (in particular, each cell has area $\frac{1}{n}\cdot\frac{\pi}{3n}$).

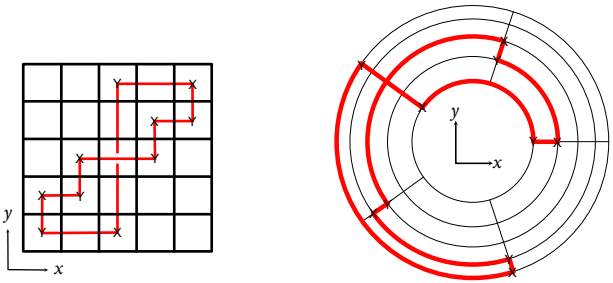


Figure 4. Converting a 5×5 Lagrangian grid diagram to a radial Lagrangian grid diagram.

For a given marking in row i and column j, we place it in the radial grid at the intersection of the circle C_i with the radial line segment determined by the angle $j\frac{2\pi}{n}$ to obtain a radial grid diagram. Join the markings in the radial grid diagram to match the original grid diagram (see Figure 4).

Remark 3.22. Notice that while the markings of the oriented grid diagram are placed in the cells of the grid, the markings of the radial grid diagram are placed at the intersections of the grid lines. This is just a shift of the markings by $\left(-\frac{1}{2}, -\frac{1}{2}\right)$.

Suppose that $\hat{G}_{x_1y_1}$ and $\hat{G}_{x_2y_2}$ are radial grid diagrams constructed (as above) from Lagrangian grid diagrams $G_{x_1y_1}$ and $G_{x_2y_2}$. We can define an immersion $f : T^2 \to B^2$ by letting $\gamma_1 : \theta_1 \mapsto (x_1(\theta_1), y_1(\theta_1))$ and $\gamma_2 : \theta_2 \mapsto (x_2(\theta_2), y_2(\theta_2))$ be the two loops corresponding to the radial grid diagrams $\hat{G}_{x_1y_1}$ and $\hat{G}_{x_2y_2}$.

We wish to lift f to a Legendrian torus in S^5 using Theorem 2.2, but to do so, it must first be smoothed. This may be remedied by following a smoothing procedure as described in Theorem 3.9, Corollary 3.10, Scholium 3.12, and Corollary 3.13 of [4], and noting that the integral used to define the lift in Theorem 2.2 results in a net area calculation here, just as it was in [4]. To see this, observe that for a path that follows a radial segment in one of the grids, the change in t is 0. For a path that follows a circular arc in one of the grids, the contribution to the change in t is given by ar^2 where a is the subtended angle of the arc (positive if the segment is oriented counterclockwise and negative otherwise), and r is the radius of the arc. That is to say, the magnitude of the change in t along such an arc is twice the area of the sector it bounds (and positive if the arc run counterclockwise, and negative otherwise). Since the radial grid is constructed so that every cell has equal area, the proofs of Theorem 3.9, Corollary 3.10, Scholium 3.12, and Corollary 3.13 in [4] may be easily adapted to this setting. Combining this with Theorem 2.2 we obtain the following:

Theorem 3.23. *Let $\hat{G}_{x_1y_1}$ and $\hat{G}_{x_2y_2}$ be radial grid diagrams constructed from Lagrangian grid diagrams $G_{x_1y_1}$ and $G_{x_2y_2}$, and let $\gamma_1 : \theta_1 \mapsto (x_1(\theta_1), y_1(\theta_1))$*

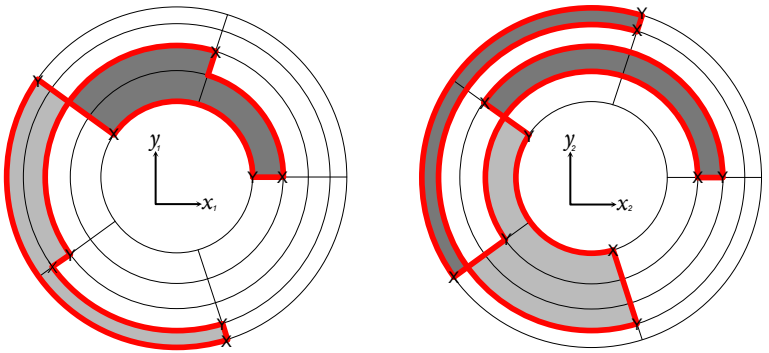


Figure 5. A pair of loops that give rise to a Lagrangian cone.

and $\gamma_2 : \theta_2 \mapsto (x_2(\theta_2), y_2(\theta_2))$ *be the immersed loops defined by these radial grid diagrams. Then the immersed torus* $f : T^2 \to B^2$,

$$f(\theta_1, \theta_2) = (x_1(\theta_1), y_1(\theta_1), x_2(\theta_2), y_2(\theta_2), \sqrt{1 - x_1^2 - y_1^2 - x_2^2 - y_2^2}, 0),$$

lifts to an immersed Legendrian torus $\tilde{f} : T^2 \to S^5 \subset \mathbb{C}^3$,

$$\tilde{f}(\theta_1, \theta_2) = e^{it(\theta_1, \theta_2)}(x_1(\theta_1), y_1(\theta_1), x_2(\theta_2), y_2(\theta_2), \sqrt{1 - x_1^2 - y_1^2 - x_2^2 - y_2^2}, 0),$$

whose cone in $\mathbb{C}^3$ *is Lagrangian.*

Consider the example shown in Figure 5. The dark shaded region of the first diagram has area $3 \cdot \frac{\pi}{75}$, as does the light shaded region. However, if we orient the two regions, using the orientation of the knot along the boundary of each, we see that the two regions have opposite orientation. The result of this is that when computing the change in t, the contributions of each region will have opposite sign. Since each contribution is equal in magnitude, the total change in t when traversing the entire knot is 0. Moreover, observe that the difference in the t coordinates at the crossing is $3 \cdot \frac{2\pi}{75}$. Similarly, one can see that the total change in t for the second grid diagram is 0, and that the difference in the t coordinates at each crossing is $2 \cdot \frac{2\pi}{75}$.

Remark 3.24. In general, beginning with two Lagrangian grid diagrams, converting to radial grid diagrams, and lifting, one produces an immersed torus, and hence an immersed Lagrangian cone. To get an embedded torus, and hence an embedded Lagrangian cone, one must check to see that the product lift condition is satisfied by the pair of Lagrangian grid diagrams (see Section 4 of [4]). This amounts to checking that condition (2) of Theorem 2.2 is satisfied. The pair of radial grid diagrams shown in Figure 5 satisfies the product lift condition, as one may check.

Remark 3.25. In Proposition 3.4 of [4] it was shown that the immersion determined by a Lagrangian grid diagram could be smoothed in such a way as to ensure that

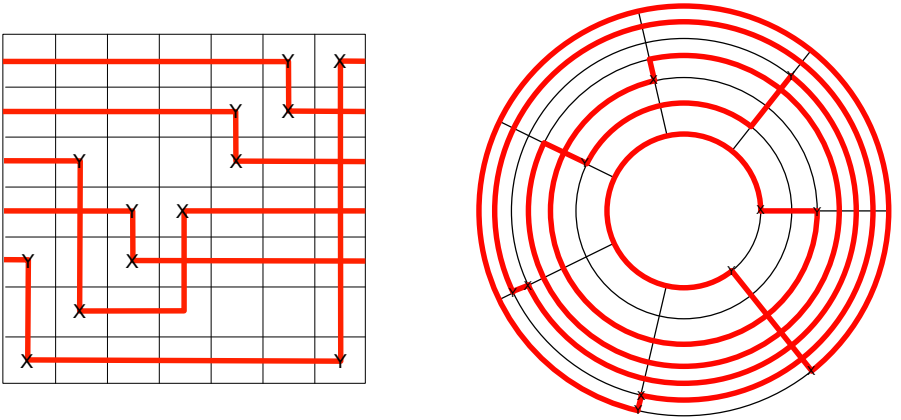


Figure 6. A 7×7 radial Lagrangian grid, with the associated grid diagram from which it is constructed.

the lift of the smoothed immersion is C^0-close to the lift of the original immersion, and that any two smoothings, sufficiently close to the original immersion, would have Legendrian isotopic lifts. The proof of that proposition depended only on the fact that the lift was determined by a net-area calculation. Since the same is true in this setting, the proof may be adapted to this situation, to produce a smoothly embedded Lagrangian cone.

The family of examples produced here is specific to the case $n = 3$, but only because the Lagrangian hypercube diagrams are constructed, at this time, only in dimension 4. Yet, it is clear that Lagrangian hypercube diagrams may be generalized to produce Lagrangian immersions $f : T^{n-1} \to B^{n-1}$.

3D. ***Examples constructed from radial hypercube diagrams.*** In the previous example, beginning with a pair of Lagrangian grid diagrams meant that for any loop on the immersed torus in B^2, in the lift, the net change in t is 0. However, this is more restrictive than necessary, since we still obtain a well-defined lift provided that the net change in t along any loop downstairs is an integer multiple of 2π. In fact, we may relax the conditions of the previous example a bit more, as follows.

Let $G_{x_1y_1}$ and $G_{x_2y_2}$ be two grid diagrams, and construct radial grid diagrams $\hat{G}_{x_1y_1}$ and $\hat{G}_{x_2y_2}$ by placing markings as in the previous example. However, to obtain an immersed loop from the diagram, we follow a slightly different procedure. Along each radial column, join the markings as in the original grid diagram. In each circular row, there are two arcs oriented from X to Y. Choose one of the two oriented arcs in each row. Figure 6 shows one example of a grid diagram, with a particular choice of connections made in each row. Thus to a given grid diagram of size n, there are 2^n distinct, immersed loops that correspond to it by following this procedure.

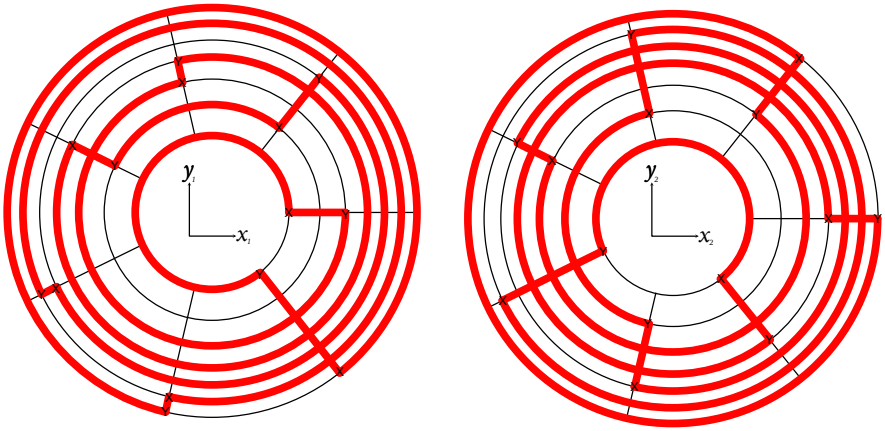


Figure 7. A pair of 7×7 radial grid diagrams that give rise to a Lagrangian cone.

Theorem 3.26. *Let $\hat{G}_{x_1y_1}$ and $\hat{G}_{x_2y_2}$ be radial grid diagrams and let*

$$\gamma_1 : \theta_1 \mapsto (x_1(\theta_1), y_1(\theta_1)) \quad \text{and} \quad \gamma_2 : \theta_2 \mapsto (x_2(\theta_2), y_2(\theta_2))$$

be the immersed loops defined by these radial grid diagrams, together with a choice of oriented circular arcs.

Suppose that $\sum_{i=1}^{n} a_i r_i^2 = 2\pi k_1$, where a_i is the angle subtended by the chosen arc in row i of $\hat{G}_{x_1,y_1}$, r_i is the radius of the corresponding circle, and $k_1 \in \mathbb{Z}$. Similarly assume that $\sum_{i=1}^{n} b_i r_i^2 = 2\pi k_2$, where b_i is the angle subtended by the chosen arc in row i of $\hat{G}_{x_2,y_2}$, r_i is the radius of the corresponding circle, and $k_2 \in \mathbb{Z}$. Then the immersed torus $f : T^2 \to B^2$,

$$f(\theta_1, \theta_2) = (x_1(\theta_1), y_1(\theta_1), x_2(\theta_2), y_2(\theta_2), \sqrt{1 - x_1^2 - y_1^2 - x_2^2 - y_2^2}, 0),$$

lifts to an immersed Legendrian torus $\tilde{f} : T^2 \to S^5 \subset \mathbb{C}^3$,

$$\tilde{f}(\theta_1, \theta_2) = e^{it(\theta_1,\theta_2)}(x_1(\theta_1), y_1(\theta_1), x_2(\theta_2), y_2(\theta_2), \sqrt{1 - x_1^2 - y_1^2 - x_2^2 - y_2^2}, 0),$$

where t is defined as in Theorem 2.2, and whose cone in $\mathbb{C}^3$ is Lagrangian.

Proof. The proof follows from Theorem 2.2 together with the observations of Theorem 3.23 that the change in t may be interpreted as a net-area calculation. The condition that $\sum_{i=1}^{n} a_i r_i^2 = 2\pi k_1$ and $\sum_{i=1}^{n} b_i r_i^2 = 2\pi k_2$ guarantees that the net-area of the loops determined by $\hat{G}_{x_1y_1}$ and $\hat{G}_{x_2y_2}$, is a multiple of 2π and hence, each loop lifts to a loop that wraps around the fiber k_1 or k_2 times. □

The two radial grid diagrams shown in Figure 7 determine an immersion that lifts to a torus whose cone is Lagrangian. A net-area calculation shows that the cone is embedded, since the two diagrams satisfy the product lift condition (see

Section 4 of [4]). Moreover, the lift has the property that each diagram lifts to a loop that wraps once around the fiber.

Remark 3.27. The pair of grid diagrams chosen at the beginning determine a structure, similar to a hypercube diagram, which we will refer to as a *radial Lagrangian hypercube diagram*.

Remark 3.28. Remark 3.25 applies in this situation as well, allowing us to produce smooth Lagrangian cones using radial Lagrangian hypercube diagrams.

In light of Example 3.1, it is natural to ask which Lagrangian hypercube diagram gives rise to the Harvey–Lawson cone. Note that the immersion given in Example 3.1 does not readily admit the structure of a Lagrangian hypercube diagram. It has only two double point circles, neither of which intersect, while any Lagrangian hypercube diagram must contain double point circles that intersect (since each Lagrangian grid diagram used to define a Lagrangian hypercube diagram must contain crossings, each of which produces a double point circle in the product). Nevertheless, it seems likely that there is a Lagrangian hypercube representation of the Harvey–Lawson cone, hence:

Conjecture 3.29. *There exists a radial Lagrangian hypercube diagram, whose associated Lagrangian cone in $\mathbb{C}^3$ is isotopic to the Harvey–Lawson cone.*

While we do not address the construction of the perturbation of a Lagrangian hypercube diagram needed to ensure that the corresponding torus in $\mathbb{C}P^2$ has only isolated transverse double points, techniques similar to those of Section 3 paired with the techniques described by Peter Lambert-Cole in [16; 17] can be used to do exactly that.

Lastly, while a radial Lagrangian hypercube diagram will not lift to a special Lagrangian cone, it may lift to a Lagrangian cone which is isotopic to a special Lagrangian cone. This leads us to pose the following question:

Question 1. What conditions on a radial Lagrangian hypercube diagram ensure that the Lagrangian cone to which it lifts is isotopic to a special Lagrangian cone? Are there any obstructions?

4. The lifting theorem

While Theorem 2.2 applies only to immersions into a unit ball, $B^{n-1} \subset \mathbb{C}^{n-1}$, thought of as a single chart of $\mathbb{C}P^{n-1}$, it can be generalized to any immersion $f : \Sigma^{n-1} \to \mathbb{C}P^{n-1}$ so that the lifting process works in much the same way as it does in Theorem 2.2. This is the content of the lifting theorem below. We build up to the lifting theorem through a series of computationally useful lemmas and definitions.

Recall that the symplectic form associated with the Fubini–Study metric is, in coordinates $z = (z_1, \ldots, z_n)$ of $\pi_{\mathbb{C}^*} : \mathbb{C}^n \setminus \{0\} \to \mathbb{C}P^{n-1}$, given by

$$\pi_{\mathbb{C}^*}^*(\omega_{FS}) = \frac{i}{2} \cdot \frac{1}{|z|^4} \sum_{k=1}^{n} \sum_{j \neq k} (\overline{z_j} z_j dz_k \wedge d\overline{z_k} - \overline{z_j} z_k dz_j \wedge d\overline{z_k}). \tag{4-1}$$

The form ω_{FS} is the form induced upon $\mathbb{C}P^{n-1}$ after quotienting by the invariant $\mathbb{C}^*$ action. One can check that

$$\int_{\mathbb{C}P^1} \omega_{FS} = \pi$$

and therefore $\frac{1}{\pi}\omega_{FS}$ is an integral symplectic form on $\mathbb{C}P^{n-1}$. Furthermore, for $i : S^{2n-1} \to \mathbb{C}^n$, it is well known that ω_{FS} is the unique form such that $i^*(\omega_0) = \pi^*(\omega_{FS})$ where $\pi : S^{2n-1} \to \mathbb{C}P^{n-1}$ is the Hopf fibration and ω_0 is the standard symplectic form on $\mathbb{C}^n$, i.e., for $z_i = x_i + iy_i$,

$$\omega_0 = \frac{i}{2} \sum_{i=1}^{n} dz_i \wedge d\overline{z_i} = \sum_{i=1}^{n} dx_i \wedge dy_i.$$

As mentioned above, the usual homogeneous, holomorphic coordinate system on $\mathbb{C}P^{n-1}$ is not suitable for our purposes. Instead, we use the hemispherical coordinate system:

Definition 4.1. Let $B_i \subset \mathbb{C}^{n-1}$ be the open unit ball and define coordinate charts $\psi_i : B_i \to \mathbb{C}P^{n-1}$, $j = 1, \ldots, n$, given by

$$\psi_i(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n) = [z_1 : \cdots : z_{i-1} : \sqrt{1 - |z|^2} : z_{i+1} : \cdots : z_n].$$

The charts, (B_i, ψ_i) are called *hemispherical charts*.

Note that we are numbering the z_i's in terms of $\mathbb{C}^n$ instead of $\mathbb{C}^{n-1}$. For example, for $n = 3$, $z \in B_2 \subset \mathbb{C}^2$ is defined by $z = (z_1, z_3)$ and is mapped to $\mathbb{C}P^3 = \mathbb{C}^3 \setminus \{0\}/\mathbb{C}^*$ as $\psi_2(z_1, z_3) = \left[z_1 : \sqrt{1 - |z|^2} : z_3\right]$ where $|z|^2 = |z_1|^2 + |z_3|^2$. We will often use the hat symbol to denote removing a term. Hence $z = (z_1, z_3)$ could also be written as $z = (z_1, \hat{z}_2, z_3)$ to simplify notation.

Also, we use U_i to refer to the image of B_i in $\mathbb{C}P^{n-1}$, i.e., $U_i = \psi_i(B_i)$. The name of the system obviously follows from the fact that the image of each chart is the image of a hemisphere in $S^{2n-1} \subset \mathbb{C}^n$ via the Hopf fibration $\pi : S^{2n-1} \to \mathbb{C}P^{n-1}$.

The hemispherical charts ψ_i are not holomorphic with respect to the natural complex structure on $\mathbb{C}P^{n-1}$. However, they do have one very nice property: the ψ_i's are Darboux charts on $\mathbb{C}P^{n-1}$.

Lemma 4.2. *If ω_0 is the standard symplectic form on $B \subset \mathbb{C}^{n-1}$ then*

$$\omega_0 = \psi_i^*(\omega_{FS}).$$

Proof. For $n = 2$, observe that in homogeneous coordinates, (4-1) translates into

$$\tilde{\omega}_{FS} = \frac{i}{2|z|^4}(\overline{z_2}z_2 dz_1 \wedge d\overline{z_1} - \overline{z_2}z_1 dz_2 \wedge d\overline{z_1} + \overline{z_1}z_1 dz_2 \wedge d\overline{z_2} - \overline{z_1}z_2 dz_1 \wedge d\overline{z_2}).$$

Observe that in B_1, $z_1 = \sqrt{1 - |z_2|^2}$. Using this observation, and changing to real coordinates, observe that in hemispherical coordinates, $\tilde{\omega}_{FS} = dx_2 \wedge dy_2$, which is ω_0 in the chart B_1. The general calculation is similar. □

Before moving on, we can characterize the sets U_i and point out that the ψ_i's are a chart system (all points of $\mathbb{C}P^{n-1}$ are in at least one chart). Let $[z_1 : \cdots : z_n] \in \mathbb{C}P^{n-1}$. At least one coordinate is nonzero, say $z_i \neq 0$. In the preimage of the quotient map for $\mathbb{C}P^{n-1} = (\mathbb{C}^n \setminus 0)/\mathbb{C}^*$, the point $(z_1, \ldots, z_n)$ is equivalent to

$$\frac{\overline{z_i}}{|z_i||z|}(z_1, \ldots, z_n)$$

where $|z| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}$. Therefore $[z_1 : \cdots : z_n] \in U_i$ and

$$U_i = \{[z_1 : \cdots : z_n] \mid z_i \neq 0\}.$$

Thus, the hemispherical chart system allows us to work with $f(\Sigma)|_{U_i} \subset B_i$ using the standard symplectic form ω_0.

Hemispherical charts also trivialize the Hopf fibration over $\mathbb{C}P^{n-1}$. In the diagram

$$\begin{array}{ccccc} B_i \times S^1 & \xrightarrow{\Psi_i} & S^{2n-1} & \hookrightarrow & \mathbb{C}^n \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi_{S^1}} & & \\ B_i & \xrightarrow{\psi_i} & \mathbb{C}P^{n-1} & & \end{array}$$

$B_i \times S^1$ is a trivialization of the S^1-bundle, $\pi : S^{2n-1} \to \mathbb{C}P^{n-1}$, given by

$$\Psi_i(z, e^{it}) = e^{it}(z_1, \ldots, z_{i-1}, \sqrt{1 - |z|^2}, z_{i+1}, \ldots, z_n) \in S^{2n-1} \subset \mathbb{C}^n.$$

The diagram commutes and Ψ_i gives a trivialization of the Hopf fibration over $U_i \subset \mathbb{C}P^{n-1}$.

As mentioned before, there is a natural contact form α on the unit sphere S^{2n-1} in $\mathbb{C}^n$. Given $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$ where $z_i = x_i + iy_i$ and $\omega_0 = \frac{i}{2}\sum_{i=1}^n dz_i \wedge d\overline{z_i} = \sum_{i=1}^n dx_i \wedge dy_i$, the form on $\mathbb{C}^n$,

$$\alpha_0 = \frac{1}{2}\left(\sum_{i=1}^n x_i dy_i - y_i dx_i\right),$$

is a contact form when restricted to S^{2n-1}. Set $\alpha = \alpha_0|_{S^{2n-1}}$. Equipped with this contact form, (S^{2n-1}, α) is a contact manifold.

We collect a few facts about α, partly to set notation for the reader, and partly to justify choices and conventions used throughout this paper.

Lemma 4.3. *For $z = (z_1, \dots, z_n) \in \mathbb{C}^n \setminus 0$ where $z_i = x_i + iy_i$, let*

$$N_z = x_1 \frac{\partial}{\partial x_1} + y_1 \frac{\partial}{\partial y_1} + \cdots + x_n \frac{\partial}{\partial x_n} + y_n \frac{\partial}{\partial y_n}$$

be the outward pointing normal vector field for any sphere of radius $r > 0$, centered at the origin in $\mathbb{R}^{2n}$, and

$$T_z = x_1 \frac{\partial}{\partial y_1} - y_1 \frac{\partial}{\partial x_1} + \cdots + x_n \frac{\partial}{\partial y_n} - y_n \frac{\partial}{\partial x_n}$$

be the vector field that generates the Hopf fibration $\pi : S^{2n-1} \to \mathbb{C}P^{n-1}$. Then the following are facts about α_0 and the contact form α:

(1) *The form α_0 is equal to $\iota_{\frac{1}{2}N_z}\omega_0$ when $|z| = 1$.*

(2) *The form α_0 also satisfies $\alpha_0(kT_z) = \frac{k}{2}|z|^2$ for k a constant, and $\iota_{T_z} d\alpha_0 = \iota_{T_z}\omega_0 = -\sum_{i=1}^{n}(x_i dx_i + y_i dy_i)$. For any vector $v \in T_z S_r^{2n-1}$ for a sphere of radius $r = |z|$,*

$$\iota_{T_z} d\alpha_0(v) = -\langle N_z, v\rangle = 0,$$

where $\langle\ ,\ \rangle$ is the usual inner product on $\mathbb{R}^{2n}$. Therefore the vector field R, defined by $R = 2T_z$ when restricted to $|z| = 1$, is the Reeb vector field of α, i.e., $\alpha(R) = 1$ and $d\alpha(R, \cdot) = 0$.

(3) *Since $i^*(\omega_0) = \pi^*(\omega_{FS})$ and $d\alpha_0 = \omega_0$, $\frac{i}{\pi}\alpha$ is the connection one-form of the integral cohomology class $\left[\frac{1}{\pi}\omega_{FS}\right]$.*

We use α for η in Theorem 2.1 to find $\Psi_i^*(\alpha)$ in the trivialization $B_i \times S^1$ with coordinates (z, e^{it}).

Lemma 4.4. *Let $B_j \subset \mathbb{C}^{n-1}$ be the unit ball with coordinates*

$$z = (z_1, \dots, z_{j-1}, z_{j+1}, \dots, z_n).$$

For a chart $\psi_j : B_j \to \mathbb{C}P^{n-1}$ and trivialization $\Psi_j : B_j \times S^1 \to S^{2n-1}$ given by $\Psi_j(z, e^{it}) = e^{it}(z_1, \dots, z_{j-1}, \sqrt{1 - |z|^2}, z_{j-1}, \dots, z_n)$,

$$\Psi_j^*(\alpha) = \tfrac{1}{2}(dt + 2\alpha_0),$$

where α_0 is the form defined above on $B_j \subset \mathbb{C}^{n-1}$.

In polar coordinates,

$$\Psi_j^*(\alpha) = \tfrac{1}{2}\big(dt + r_1^2 d\theta_1 + \cdots + \widehat{r_j^2 d\theta_j} + \cdots + r_n^2 d\theta_j\big).$$

Note that the α_0 defined on B_j has no z_j term of the form $(x_j dy_j - y_j dx_j)$ since $z \in B_j$ has coordinates $z = (z_1, \dots, \hat{z_j}, \dots, z_n)$. The proof of the lemma is a calculation, and left to the reader.

Thus we can take τ in Theorem 2.1 to be the 1-form $-2\alpha_0 \in \Omega^1(B_j)$. In each chart $B_j \times S^1$, label $\tau_j = -2\alpha_0$; note that the transition map $\Psi_{kj} : B_j \times S^1 \to B_k \times S^1$ takes

$$\Psi_{kj}^*\big(\tfrac{1}{2}(dt - \tau_k)\big) = \tfrac{1}{2}(dt - \tau_j). \tag{4-2}$$

This result follows from the next lemma.

Lemma 4.5. *Let B_j be the unit ball in $\mathbb{C}^{n-1}$ with coordinates*

$$z = (z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_n),$$

and let the function $\Psi_j : B_j \times S^1 \to S^{2n-1} \subset \mathbb{C}^n$ be given by $\Psi_j(z, e^{it}) = e^{it}(z_1, \ldots, z_{j-1}, \sqrt{1-|z|^2}, z_{j+1}, \ldots, z_n)$ where $|z| = |z_1|^2 + \cdots + \widehat{|z_j|^2} + \cdots + |z_n|^2$. For $k \neq j$, the map

$$\Psi_{kj} : B_j \setminus \{z_k = 0\} \times S^1 \to B_k \setminus \{z_j = 0\} \times S^1$$

defined by $\Psi_{kj} = \Psi_k^{-1} \circ \Psi_j$ is given by the map

$$\begin{aligned}\Psi_{kj}(z, e^{it}) &= \Bigg(z_1 \frac{\overline{z_k}}{|z_k|}, \ldots, z_{k-1}\frac{\overline{z_k}}{|z_k|}, |z_k|, z_{k+1}\frac{\overline{z_k}}{|z_k|},\\ &\qquad \ldots, z_{j-1}\frac{\overline{z_k}}{|z_k|}, \frac{\overline{z_k}}{|z_k|}\sqrt{1-|z|^2}, z_{j+1}\frac{\overline{z_k}}{|z_k|}, \ldots, z_n\frac{\overline{z_k}}{|z_k|}, e^{it}\frac{z_k}{|z_k|}\Bigg)\end{aligned}$$

In polar coordinates,

$$\begin{aligned}\Psi_{kj}(r_1, \theta_1, \ldots, \hat{r_j}, \hat{\theta_j}, \ldots, r_n, \theta_n, t) = \\ \Bigg(r_1, \theta_1 - \theta_k, r_2, \theta_2 - \theta_k, \ldots, r_{j-1}, \theta_{j-1} - \theta_k, \sqrt{1 - \sum_{i=1, i\neq j}^{n} r_i^2},\\ -\theta_k, r_{j+1}, \theta_{j+1} - \theta_k, \ldots, r_n, \theta_n - \theta_k, t + \theta_k\Bigg)\end{aligned}$$

Proof. We show the calculation for B_2, $B_3 \subset \mathbb{C}^3$. The general case is similar. The maps

$$\Psi_2 : B_2 \times S^1 \to S^7 \subset \mathbb{C}^4,$$

$$\Psi_2(z_1, z_3, z_4, e^{it}) = e^{it}(z_1, \sqrt{1-|z|^2}, z_3, z_4),$$

and

$$\Psi_3 : B_3 \times S^1 \to S^7 \subset \mathbb{C}^4,$$

$$\Psi_3(w_1, w_2, w_4, e^{it}) = e^{it}(w_1, w_2, \sqrt{1-|w|^2}, w_4)$$

give rise to $\Psi_{32} : B_2 \setminus \{z_3 = 0\} \times S^1 \to B_3 \setminus \{w_2 = 0\} \times S^1$ via $\Psi_3^{-1} \circ \Psi_2$. By

multiplying by 1 appropriately,

$$\begin{aligned}\Psi_2(z_1, z_3, z_4, e^{it}) &= e^{it}\big(z_1, \sqrt{1-|z_1|^2-|z_3|^2-|z_4|^2}, z_3, z_4\big)\\ &= e^{it}\left(\frac{z_3}{|z_3|}\frac{\overline{z_3}}{|z_3|}\right)\big(z_1, \sqrt{1-|z_1|^2-|z_3|^2-|z_4|^2}, z_3, z_4\big)\\ &= e^{it}\left(\frac{z_3}{|z_3|}\right)\left(z_1\frac{\overline{z_3}}{|z_3|}, \frac{\overline{z_3}}{|z_3|}\sqrt{1-|z_1|^2-|z_3|^2-|z_4|^2}, z_3\frac{\overline{z_3}}{|z_3|}, z_4\frac{\overline{z_3}}{|z_3|}\right)\\ &= \Psi_3(w_1, w_2, w_4, e^{it'})\end{aligned}$$

where $w_1 = z_1\,\overline{z_3}/|z_3|$, $w_2 = \overline{z_3}/|z_3|\,\sqrt{1-|z|^2}$, $w_4 = z_4\,\overline{z_3}/|z_3|$, and $e^{it'} = e^{it}z_3/|z_3|$. One can check that $(w_1, w_2, w_4, e^{it}) \in B_3 \setminus \{w_2 = 0\}$ and $e^{it}z_3/|z_3| \in S^1$ and $\sqrt{1-|w_1|^2-|w_2|^2-|w_4|^2} = |z_3|$ as desired. □

Remark 4.6. The formula for Ψ_{kj} also gives the formula for $\psi_{kj} : B_j \setminus \{z_k = 0\} \to B_k \setminus \{z_j = 0\}$ for $\psi_{kj} = \psi_k^{-1} \circ \psi_j$ by looking at the z coordinates of (z, e^{it}).

In summary, given a Lagrangian immersion $f : \Sigma \to \mathbb{C}P^{n-1}$ and $V_j = f(\Sigma) \cap B_j$, we can work with $V_j \subset B_j$ using

- the standard symplectic form ω_0 on $B_j \subset \mathbb{C}^{n-1}$,
- the standard 1-form $\tau_j = -2\alpha_0$ on $B_j \subset \mathbb{C}^{n-1}$,

and patch the V_j's together using the transition maps $\psi_{kj} : B_j \to B_k$ given by $\psi_{kj} = \psi_k^{-1} \circ \psi_j$.

In practice, this allows us to do integration and other calculations in the B_j's using standard forms in each instead of working with homogeneous coordinates and ω_{FS} in $\mathbb{C}P^{n-1}$.

This chart system also gives us new ways to build examples of Lagrangian immersions by first working with piecewise linear submanifolds in each ball B_j, pasting the pieces together, and then smoothing the result (as is done with Lagrangian hypercubes in [2] and Section 3 of [4]).

4A. ***The lifting theorem.*** The lifting theorem puts the separate pieces in the previous sections together into one result. First, we need an explicit way to calculate integrals along paths in $f(\Sigma)$.

Let $f : \Sigma \to \mathbb{C}P^{n-1}$ be a Lagrangian immersion and let $\gamma : I \to \Sigma$ be a path. In order to define the lift, we need to define a map $t : I \to \mathbb{R}/2\pi\mathbb{Z}$, which we do in pieces. Split the interval I into subintervals

$$I = \bigcup_{k=0}^{m-1} [s_k, s_{k+1}]$$

where $0 = s_0 < s_1 < \cdots < s_{m-1} < s_m = 1$ such that $f(\gamma([s_k, s_{k+1}])) \subset B_j$ for

some $j \in \{1, \dots, n\}$ (after identifying B_j with U_j using ψ_j). Index the B_j's by j_k so that $f(\gamma([s_k, s_{k+1}])) \subset B_{j_k}$ where j_k is the index of the chart in which $\gamma([s_k, s_{k+1}])$ is contained. Let $x_k = \gamma(s_k)$ so that $x_0 = \gamma(s_0)$ and $x_m = \gamma(s_m)$. Also, for convenience, use the notation $(z)_k$ to stand for the z_k coordinate of $z \in B_j$. (If $z \in B_3 \subset \mathbb{C}^3$ such that $z = (z_1, z_2, z_4)$ then $(z)_4 = z_4$.)

Since $f(\gamma([s_0, s_1])) \subset B_{j_0}$, we can integrate $\tau_{j_0} = -2\alpha_0$ (see (4-2)) along the path $f(\gamma([s_0, s_1]))$. Define $t_0 : [s_0, s_1] \to \mathbb{R}/2\pi\mathbb{Z}$ by

$$t_0(s) = \left(\int_0^s \tau_{j_0}((f \circ \gamma)'(u))\, du \right) \bmod 2\pi,$$

where $t_0(0) = 0$.

For $s \in [s_0, s_1]$ and $t(0) = a$, we can write

$$t(s) = t_0(s) + a.$$

The point $(f(\gamma(s_1)), e^{it(s_1)}) \in B_{j_0} \times S^1$ also lives as a point $\Psi_{j_1 j_0}(f(\gamma(s_1)), e^{it(s_1)}) \in B_{j_1} \times S^1$. Define $\Psi_{j_1 j_0}(t(s_1)) \in \mathbb{R}/2\pi\mathbb{R}$ to be the argument of the S^1 component of this map in $B_{j_1} \times S^1$. We can also define the point $\psi_{j_1 j_0}(f(\gamma(s_1)) \in B_{j_1}$ as the B_{j_1} component of $B_{j_1} \times S^1$ (see Remark 4.6).

Lemma 4.7. *When $t(s_k)$ is defined for $(f(\gamma(s_k)), e^{it(s_k)}) \in B_{j_{k-1}}$, then $\Psi_{j_k j_{k-1}}(t(s_k)) = t(s_k) + arg(\psi_{j_k j_{k-1}}(f(\gamma(s_k)))_{j_k})$.*

Proof. See Lemma 4.5. □

We can now continue the integration in B_{j_1}: Define $t_1 : [s_1, s_2] \to \mathbb{R}/2\pi\mathbb{Z}$ by $t_1(s_1) = 0$ and

$$t_1(s) = \left(\int_{s_1}^s \tau_{j_1}((f \circ \gamma)'(u))\, du \right) \bmod 2\pi.$$

Hence we can write $t(s)$ for $s \in [s_1, s_2]$ as

$$t(s) = t_1(s) + \Psi_{j_1 j_0}(t_0(s_1) + a).$$

Induct on k to integrate the τ_{j_k}'s over the entire path:

Definition 4.8. Let $[0, 1] \xrightarrow{\gamma} \Sigma \xrightarrow{f} \mathbb{C}P^{n-1}$ and suppose there exists an increasing sequence $0 = s_0 < s_1 < \cdots < s_{m-1} < s_m = 1$ such that $f(\gamma([s_k, s_{k+1}])) \subset B_{j_k}$ for $j_k \in \{1, \dots, n\}$ and $f(\gamma(s_k)) \neq 0$ and $f(\gamma(s_{k+1})) \neq 0$ for all $0 \leq k \leq m$. Assume $t(0) = a$ and define the *lifting integral* to be

$$\Gamma \int_\gamma \tau := \Big[t_{m-1}(s_m) + \Psi_{j_{m-1} j_{m-2}}\big(\cdots\big(t_3(s_4) + \Psi_{j_3 j_2}\big(t_2(s_3) + \Psi_{j_2 j_1}\big(t_1(s_2) + \Psi_{j_1 j_0}(t_0(s_1) + a)\big)\big)\big)\big)\Big] \bmod 2\pi.$$

Remark 4.9. See Example 5.1 for an example of a calculation of the lifting integral for the trivial cone.

In practice we usually need only $m = 1$ or $m = 2$ for most integrals. Also, since

$$\tau_{j_k} = -\sum_{i=1,\, i\neq j_k}^{n} (x_i dy_i - y_i dx_i) \quad \text{and} \quad \omega_{FS}|_{B_{j_k}} = \sum_{i=1,\, i\neq j}^{n} dx_i \wedge dy_i,$$

the calculations may be done in each chart. In summary, we obtain the lifting theorem, which says that if

(1) $\Gamma \int_\gamma \tau = 0 \bmod 2\pi$ for all $[\gamma] \in H_1(\Sigma; \mathbb{Z})$, and

(2) for all distinct points $x_1, \ldots, x_k \in \Sigma$ such that $f(x_1) = f(x_j)$ for all $j \leq k$, and a choice of path γ_j from x_1 to x_j in Σ for $2 \leq j \leq k$, the set

$$\left\{ \left(\Gamma \int_{f(\gamma_j)} \tau \right) \bmod 2\pi \mid 2 \leq j \leq k \right\}$$

has $k-1$ distinct values, none of which are equal to 0,

then $f : \Sigma \to \mathbb{C}P^{n-1}$ lifts to an embedding $\tilde{f} : \Sigma \to S^{2n-1}$ such that the image (the lift) $\tilde{\Sigma}$ is a Legendrian submanifold of (S^{2n-1}, α). Furthermore, the cone $c\tilde{\Sigma}$ is Lagrangian in $\mathbb{C}^n$ with respect to the standard symplectic structure ω_0.

5. Legendrian contact homology generators of the trivial Lagrangian cone

Example 5.1. We already saw in Example 1.6 how to obtain a trivial (special) Lagrangian cone, but, we can also construct this example using the lifting theorem, as a lift of a map $f : S^{n-1} \to \mathbb{C}P^{n-1}$.

Recall that the trivial cone is given by the map $\tilde{f} : \mathbb{R}^n \to \mathbb{C}^n$ where $(x_1, \ldots, x_n) \mapsto (x_1\eta_1, \ldots, x_n\eta_n)$, and $\eta = (\eta_1, \ldots, \eta_n)$ is a complex vector with $\eta_j \neq 0$ for all j. Clearly the trivial cone is a lift of the Lagrangian immersion $f : S^{n-1} \to \mathbb{C}P^{n-1}$ given by $f(x_1, \ldots, x_n) = [x_1\eta_1 : \cdots : x_n\eta_n]$.

Observe that the set $\{(x_1, \ldots, x_n) \subset \mathbb{R}^n \mid \sum_{k=1}^n |x_k\eta_{k,j}|^2 = 1\}$ is an $(n-1)$-dimensional sphere, S^{n-1}, for any choice of complex vector $(\eta_{1,j}, \ldots, \eta_{n,j})$ (the reason for the j-subscript will be apparent shortly). Moreover, we may cover S^{n-1} by charts of the form $\phi_j^{\pm} : V_j^{\pm} \to S^{n-1}$ where $V_j^{\pm} = \left\{(x_1, \ldots, x_{j-1}, \hat{x_j}, x_{j+1}, \ldots, x_n) \in \mathbb{R}^{n-1} \mid \sum_{k=1,\, k\neq j}^n |x_j\eta_{k,j}|^2 < 1\right\}$, and the sign indicates which hemisphere is being covered. Within each chart, after identifying $V_j^{\pm}$ with $\phi_j^{\pm}(V_j^{\pm})$, we may write $f(x)$ as $f_j^{\pm}(x)$ where $f_j^{\pm} : V_j^{\pm} \to \mathbb{C}P^{n-1}$ is given by

$$f_j^{\pm}(x_1, \ldots, x_{j-1}, \hat{x_j}, x_{j+1}, \ldots, x_n) =$$

$$\left[x_1\eta_{1,j} : \cdots : x_{j-1}\eta_{j-1,j} : \pm \sqrt{1 - \sum_{k=1,\, k\neq j}^{n} |x_k\eta_{k,j}|^2} : x_{j+1}\eta_{j+1,j} : \cdots : x_n\eta_{n,j} \right],$$

where $\eta_{k,j} = \eta_k \overline{x_j\eta_j}/|x_j\eta_j|$.

Since $H_1(S^2, \mathbb{Z})$ is trivial, the first condition of the lifting theorem is automatically satisfied. Moreover, $f_i^{\pm}$ is clearly an embedding on $V_i^{\pm}$, so within each chart the second condition is satisfied. However, observe that after patching these maps together, the antipodal points of S^{n-1} are the only ones identified by f (in fact, the image of f is a copy of $\mathbb{R}P^{n-1}$). To see that the antipodal points are separated in the lift, consider what happens when $n = 3$ and $\eta = (1, 1, 1)$. In that case, we can lift along a path γ from the origin of V_3^+ (the "north pole") to the origin of V_3^- (the "south pole"), and running diametrically through the origin of V_1^+. Notice that integrating τ along γ contributes 0 to the lift within each chart. If we transition from V_3^+ to V_1^+ at the point $(1/\sqrt{2}, 0, 1/\sqrt{2})$ and from V_1^+ to V_3^- at the point $(1/\sqrt{2}, 0, -1/\sqrt{2})$ then we pick up a factor of -1, or $e^{i\pi}$, on the S^1-factor from the transition map Ψ_{31} (the second transition map). Hence,

$$\Gamma \int_\gamma \tau = \pi.$$

The general calculation is similar. Hence, the lifting theorem guarantees the existence of an embedded lift, $\tilde{f} : S^{n-1} \to S^{2n-1} \subset \mathbb{C}^n$ such that the cone is Lagrangian in $\mathbb{C}^n$. Moreover, our discussion above clearly identifies this as a Lagrangian $\mathbb{R}^n \subset \mathbb{C}^n$, which is the trivial cone.

The trivial cone intersects S^{2n-1} in a Legendrian $(n-1)$-sphere that projects down to a copy of $\mathbb{R}P^{n-1}$ via a 2-to-1 map (the quotient by the antipodal map). This is inconvenient when one wishes to compute Legendrian contact homology, because one needs isolated transverse double points. However, we can perturb f through a family of functions f_ϵ so that for some ϵ the image of the lift, $\tilde{f}_\epsilon$, is a copy of S^{n-1} having only transverse double points when projected down to $\mathbb{C}P^{n-1}$.

For simplicity, we write down the perturbation in the case where $n = 3$ and $\eta = (1, 1, 1)$. Choose $\epsilon \geq 0$ and perturb each hemisphere of S^2 as follows:

$$f_{1,\epsilon}^{\pm}(x_2, x_3) = \left[\pm e^{\pm i\epsilon\sqrt{1-x_2^2-x_3^2}} \sqrt{1 - x_2^2 - x_3^2} : e^{i\epsilon x_2} x_2 : e^{i\epsilon x_3} x_3 \right],$$

$$f_{2,\epsilon}^{\pm}(x_1, x_3) = \left[e^{i\epsilon x_1} x_1 : \pm e^{\pm i\epsilon\sqrt{1-x_1^2-x_3^2}} \sqrt{1 - x_1^2 - x_3^2} : e^{i\epsilon x_3} x_3 \right],$$

$$f_{3,\epsilon}^{\pm}(x_1, x_2) = \left[e^{i\epsilon x_1} x_1 : e^{i\epsilon x_2} x_2 : \pm e^{\pm i\epsilon\sqrt{1-x_1^2-x_2^2}} \sqrt{1 - x_1^2 - x_2^2} \right].$$

Observe that the perturbations in each chart are consistent with the transition maps. To determine the (transverse) intersections, and hence the Reeb chords, we begin with the observation that all double points are antipodal points. We leave the proof as an exercise for the reader.

Theorem 5.2. *Let $f_\epsilon : S^2 \to \mathbb{C}P^2$ be the map determined by patching together $f_{i,\epsilon}^{\pm} : V_i^{\pm} \to \mathbb{C}P^2$ for $i = 1, 2, 3$. Let $(x_1, x_2, x_3), (y_1, y_2, y_3) \in S^2 \subset \mathbb{R}^3$ be two points such that $f_\epsilon(x_1, x_2, x_3) = f_\epsilon(y_1, y_2, y_3)$. Then $(x_1, x_2, x_3) = -(y_1, y_2, y_3))$.*

To determine the double points when $\epsilon > 0$, assume $\pm(x_1, x_2, x_3)$ map to a double point, $f_\epsilon(x_1, x_2, x_3)$. If $x_i \neq 0$ for all i, then without loss of generality we may assume $x_1 > 0$. Using the charts V_1^+ and V_1^-, we see that $f_{1,\epsilon}^+(x_2, x_3) = f_{1,\epsilon}^-(-x_2, -x_3)$, and hence

$$\left[e^{i\epsilon\sqrt{1-x_2^2-x_3^2}}\sqrt{1-x_2^2-x_3^2} : e^{i\epsilon x_2}x_2 : e^{i\epsilon x_3}x_3\right] = \left[-e^{-i\epsilon\sqrt{1-x_2^2-x_3^2}}\sqrt{1-x_2^2-x_3^2} : -e^{-i\epsilon x_2}x_2 : -e^{-i\epsilon x_3}x_3\right]. \quad (5\text{-}1)$$

Cross-multiplying in the first two homogeneous coordinates, we see that

$$e^{i\epsilon(x_2-\sqrt{1-x_2^2-x_3^2})}\sqrt{1-x_2^2-x_3^2} = e^{i\epsilon(-x_2+\sqrt{1-x_2^2-x_3^2})}\sqrt{1-x_2^2-x_3^2}.$$

If $1-x_2^2-x_3^2 \neq 0$, then for small ϵ we may equate the arguments of the exponentials, to obtain $x_2 > 0$ and $2x_2^2 + x_3^2 = 1$. Similarly, cross-multiplying in the first and third homogeneous coordinates, and applying the same reasoning, we obtain $x_3 > 0$ and $x_2^2 + 2x_3^2 = 1$. Solving this system, and recalling that $x_1 = \sqrt{1-x_2^2-x_3^2}$, we obtain that $x_1 = x_2 = x_3 = \pm 1/\sqrt{3}$.

If $1-x_2^2-x_3^2 = 0$ and $x_2 = 0$ then we get a double point at $[0:0:1]$. Similarly, if $1-x_2^2-x_3^2 = 0$ and $x_3 = 0$ then we get a double point at $[0:1:0]$. Finally, assume $x_1 = 0$ and neither x_2 nor x_3 is zero. In this case, working in the charts V_3^+ and V_3^- we obtain

$$\left[0 : e^{i\epsilon x_2}x_2 : e^{i\epsilon\sqrt{1-x_2^2}}\sqrt{1-x_2^2}\right] = \left[0 : -e^{-i\epsilon x_2}x_2 : -e^{-i\epsilon\sqrt{1-x_2^2}}\sqrt{1-x_2^2}\right].$$

For small ϵ, we may use techniques similar to the previous case to obtain that $x_2 = x_3 = \pm\frac{1}{\sqrt{2}}$. A similar discussion applies if $1-x_1^2-x_2^2 = 0$ or $1-x_1^2-x_3^2 = 0$.

From the discussion above, we obtain the following theorem.

Theorem 5.3. *Let $S \subset S^5$ be the Legendrian 2-sphere obtained from intersecting the trivial cone with S^5 and then perturbing it via Legendrian isotopy to the image of $f_{i,\epsilon}^{\pm}$ for $i \in \{1,2,3\}$, for some $\epsilon > 0$. The projection $\pi : S^5 \to \mathbb{C}P^2$ has 7 transverse double points: $\pm(1,0,0)$, $\pm(0,1,0)$, $\pm(0,0,1)$, $\pm\big(1/\sqrt{2}, 1/\sqrt{2}, 0\big)$, $\pm\big(1/\sqrt{2}, 0, 1/\sqrt{2}\big)$, $\pm\big(0, 1/\sqrt{2}, 1/\sqrt{2}\big)$, and $\pm\big(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}\big)$. Then the 0-filtration level of the Legendrian contact homology of S is generated by 7 pairs of short Reeb chords.*

In summary, we have constructed a family of Lagrangian cones, all isotopic to the trivial cone. However, for small $\epsilon > 0$ our cones have the additional property that the projection to $\mathbb{C}P^2$ has 7 transverse double points, while the trivial cone (obtained by taking $\epsilon = 0$) is a 2-to-1 cover of its projection to $\mathbb{C}P^2$.

6. Legendrian submanifolds of S^{2n-1} as lifts of Lagrangian submanifolds in $\mathbb{C}P^{n-1}$

The motivation of this paper is the study of Lagrangian cones given by lifting an immersion into $\mathbb{C}P^{n-1}$ to an embedded Legendrian submanifold of S^{2n-1}. However, Theorem 2.2 and the lifting theorem provide a way to study Legendrian submanifolds of S^{2n-1} on their own.

A lot of work has been done to study Legendrian knots in dimension 3, especially in the standard contact $\mathbb{R}^3$ (see [11; 20; 21; 22; 23; 28]), and Joshua Sabloff studied the Legendrian contact homology of knots in 3-dimensional circle bundles in [30].

Less is known about Legendrian submanifolds in higher dimensions, and much of it only in the standard contact $\mathbb{R}^{2n+1}$ (see [4; 8; 9; 10]). In [29], Legendrian submanifolds of circle bundles over orbifolds are considered, and in [1], the circle bundle $\mathbb{R}^4 \times S^1$ is considered in depth, and related to the case where $\mathbb{R}^4 \times S^1$ is identified with the Hopf bundle over a single chart of $\mathbb{C}P^2$ (the special case of Theorem 2.2 in this paper).

Theorem 2.2 allows one to study Legendrian submanifolds of S^{2n-1} just as one might study Legendrian submanifolds of $\mathbb{R}^{2n} \times S^1$ or even the standard contact $\mathbb{R}^{2n+1}$. As seen in Example 3.1, and Section 3C, the lifts function in much the same way as one might lift an exact Lagrangian to a Legendrian knot in the standard contact $\mathbb{R}^{2n+1}$, or the 1-jet space of a manifold.

Although Theorem 2.2 makes calculations simple, it fails to capture one of the most basic examples: the Legendrian sphere corresponding to the intersection of the trivial cone with S^{2n-1} (as observed in Example 5.1). The lifting theorem moves the story forward, allowing one to consider immersions into $\mathbb{C}P^{n-1}$ that do not lie in a single chart. It shows that the calculations are not much more difficult than they are in the case of Theorem 2.2, because in each chart the calculations use standard forms, and one need only to track how the lifting parameter t transitions from one chart to the next. This leads us to ask the following question:

Question 2. Sabloff showed in [30] how to compute the DGA of Legendrian knots in certain contact circle bundles over surfaces. In the context of Theorem 2.2 or the lifting theorem, is there a similar combinatorial algorithm for computing the Legendrian contact homology in higher dimensional circle bundles?

If such an algorithm can be found, one would expect the structure of a radial Lagrangian hypercube diagram to provide a setting in which such calculations would be simple, and could be automated on a computer.

7. Minimal and Hamiltonian Submanifolds

Special Lagrangian submanifolds, introduced by Harvey and Lawson in [13] have been studied extensively due to their connection with mirror symmetry. Special

Lagrangian cones in $\mathbb{C}^n$ can be studied via the equations that define in them in $\mathbb{C}^n$, as minimal Legendrians in S^{2n-1} (the link), or from the perspective of the corresponding minimal Lagrangian submanifold of $\mathbb{C}P^{n-1}$ (see [14] and [15]). While many examples have been studied, the difficulty in working with the special Lagrangian conditions has led to some weaker conditions being studied in the hope of better understanding special Lagrangians. In [26], the notion of Hamiltonian minimal (H-minimal) Lagrangian submanifolds was introduced. A Lagrangian submanifold in a Kähler manifold is said to be H-minimal if the volume is stationary under compactly supported smooth Hamiltonian deformations (see [15]).

H-minimal Lagrangian cones in $\mathbb{C}^2$ were studied and classified by Schoen and Wolfson in [31]. In particular they showed that only cones of Maslov index ± 1 are area minimizing. Moreover, they showed that if an immersed Lagrangian submanifold of a Kähler–Einstein manifold is stationary for volume, it is automatically minimal, and special Lagrangian in the Calabi–Yau case (see Lemma 8.2 of [31]).

It is already known that the trivial cone is H-minimal (see [18], [24], and [25]). The Harvey–Lawson cone is also known to be strictly Hamiltonian stable, that is, the second variation of the volume is nonnegative under every Hamiltonian deformation, (see [6] and [18]), and it is known that any Hamiltonian stable, minimal Lagrangian torus in $\mathbb{C}P^2$ is congruent to the Clifford torus (see [26], [27] and [33]).

Question 3. What are the conditions on a Lagrangian immersion into $\mathbb{C}P^{n-1}$ that guarantee it lifts to an H-minimal Lagrangian cone?

Question 4. What are the conditions on Legendrian hypercube diagrams that generate H-minimal Lagrangian cones?

References

[1] J. Asplund, *Contact homology of Legendrian knots in five-dimensional circle bundles*, Masters thesis, Uppsala University, 2016, available at https://tinyurl.com/asplund-masters.

[2] S. Baldridge, "Embedded and Lagrangian knotted tori in $\mathbb{R}^4$ and hypercube homology", 2010. arXiv 1010.3742

[3] S. Baldridge and A. M. Lowrance, "Cube diagrams and 3-dimensional Reidemeister-like moves for knots", *J. Knot Theory Ramifications* **21**:5 (2012), 1250033, 39. MR Zbl

[4] S. Baldridge and B. McCarty, "On the rotation class of knotted Legendrian tori in $\mathbb{R}^5$", *Topology Appl.* **209** (2016), 91–114. MR Zbl

[5] H. Brunn, "Über verknotete Kurven", pp. 256–259 in *Verhandlungen des ersten internationalen Mathematiker-Kongresses* (Zurich, 1897), 1897. Zbl

[6] S. Chang, "On Hamiltonian stable minimal Lagrangian surfaces in $\mathbf{C}\mathrm{P}^2$", *J. Geom. Anal.* **10**:2 (2000), 243–255. MR Zbl

[7] P. R. Cromwell, "Embedding knots and links in an open book, I: Basic properties", *Topology Appl.* **64**:1 (1995), 37–58. MR Zbl

[8] T. Ekholm, J. Etnyre, and M. Sullivan, "The contact homology of Legendrian submanifolds in $\mathbb{R}^{2n+1}$", *J. Differential Geom.* **71**:2 (2005), 177–305. MR Zbl

[9] T. Ekholm, J. Etnyre, and M. Sullivan, "Non-isotopic Legendrian submanifolds in $\mathbb{R}^{2n+1}$", *J. Differential Geom.* **71**:1 (2005), 85–128. MR Zbl

[10] T. Ekholm, J. Etnyre, and M. Sullivan, "Legendrian contact homology in $P \times \mathbb{R}$", *Trans. Amer. Math. Soc.* **359**:7 (2007), 3301–3335. MR Zbl

[11] T. Ekholm, L. Ng, and V. Shende, "A complete knot invariant from contact homology", 2016. Zbl arXiv 1606.07050

[12] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Wiley-Interscience, New York, 1978. MR Zbl

[13] R. Harvey and H. B. Lawson, Jr., "Calibrated geometries", *Acta Math.* **148** (1982), 47–157. MR Zbl

[14] M. Haskins, "Special Lagrangian cones", *Amer. J. Math.* **126**:4 (2004), 845–871. MR Zbl

[15] H. Iriyeh, "Hamiltonian minimal Lagrangian cones in $\mathbf{C}^m$", *Tokyo J. Math.* **28**:1 (2005), 91–107. MR Zbl

[16] P. Lambert-Cole, "Legendrian Products", 2013. arXiv 1301.3700

[17] P. Lambert-Cole, *Invariants of Legendrian products*, Ph.D. thesis, Louisiana State University, 2014, available at https://digitalcommons.lsu.edu/gradschool_dissertations/2909/.

[18] H. Ma and Y. Ohnita, "Differential geometry of Lagrangian submanifolds and Hamiltonian variational problems", pp. 115–134 in *Harmonic maps and differential geometry*, Contemp. Math. **542**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[19] L. L. Ng, "Computable Legendrian invariants", *Topology* **42**:1 (2003), 55–82. MR Zbl

[20] L. Ng, "Knot and braid invariants from contact homology: I", *Geom. Topol.* **9** (2005), 247–297. MR Zbl

[21] L. Ng, "Knot and braid invariants from contact homology: II", *Geom. Topol.* **9** (2005), 1603–1637. MR Zbl

[22] L. Ng, "Framed knot contact homology", *Duke Math. J.* **141**:2 (2008), 365–406. MR Zbl

[23] L. Ng and D. Thurston, "Grid diagrams, braids, and contact geometry", pp. 120–136 in *Proceedings of Gökova Geometry–Topology Conference* 2008, Gökova Geometry/Topology Conference (GGT), Gökova, 2009. MR Zbl

[24] Y.-G. Oh, "Second variation and stabilities of minimal Lagrangian submanifolds in Kähler manifolds", *Invent. Math.* **101**:2 (1990), 501–519. MR Zbl

[25] Y.-G. Oh, "Tight Lagrangian submanifolds in $\mathbf{CP}^n$", *Math. Z.* **207**:3 (1991), 409–416. MR Zbl

[26] Y.-G. Oh, "Volume minimization of Lagrangian submanifolds under Hamiltonian deformations", *Math. Z.* **212**:2 (1993), 175–192. MR Zbl

[27] H. Ono, "Hamiltonian stability of Lagrangian tori in toric Kähler manifolds", *Ann. Global Anal. Geom.* **31**:4 (2007), 329–343. MR Zbl

[28] P. Ozsvath, Z. Szabo, and D. Thurston, "Legendrian knots, transverse knots and combinatorial Floer homology", 2008. Zbl arXiv math/0611841v2

[29] J. Pati, "Contact homology of S^1-bundles over some symplectically reduced orbifolds", 2009. arXiv 0910.5934

[30] J. M. Sabloff, "Invariants of Legendrian knots in circle bundles", *Commun. Contemp. Math.* **5**:4 (2003), 569–627. MR Zbl

[31] R. Schoen and J. Wolfson, "Minimizing area among Lagrangian surfaces: the mapping problem", *J. Differential Geom.* **58**:1 (2001), 1–86. MR Zbl

[32] A. Strominger, S.-T. Yau, and E. Zaslow, "Mirror symmetry is T-duality", *Nuclear Phys. B* **479**:1-2 (1996), 243–259. MR Zbl

[33] F. Urbano, "Index of Lagrangian submanifolds of $\mathbf{C}\mathrm{P}^n$ and the Laplacian of 1-forms", *Geom. Dedicata* **48**:3 (1993), 309–318. MR Zbl

[34] J. Wolfson, "Two applications of prequantization in Lagrangian topology", *Pacific J. Math.* **215**:2 (2004), 393–398. MR Zbl

Received 31 Jan 2021. Revised 11 Nov 2021.

SCOTT BALDRIDGE: sbaldrid@math.lsu.edu
Department of Mathematics, Louisiana State University, Baton Rouge, LA, United States

BEN MCCARTY: ben.mccarty@memphis.edu
Department of Mathematical Sciences, University of Memphis, Memphis, TN, United States

DAVID VELA-VICK: shea@math.lsu.edu
Department of Mathematics, Louisiana State University, Baton Rouge, LA, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.81

L-space knots are fibered and strongly quasipositive

John A. Baldwin and Steven Sivek

We give a new, conceptually simpler proof of the fact that knots in S^3 with positive L-space surgeries are fibered and strongly quasipositive. Our motivation for doing so is that this new proof uses comparatively little Heegaard Floer-specific machinery and can thus be translated to other forms of Floer homology. We carried this out for instanton Floer homology in our article "Instantons and L-space surgeries" and used it to generalize Kronheimer and Mrowka's results on SU(2) representations of fundamental groups of Dehn surgeries.

The hat version $\widehat{HF}(Y)$ of Heegaard Floer homology, which we will take with coefficients in $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$ throughout, carries an absolute $\mathbb{Z}/2\mathbb{Z}$ grading such that

$$\chi(\widehat{HF}(Y,\mathfrak{s})) = \begin{cases} 1 & \text{if } b_1(Y) = 0, \\ 0 & \text{if } b_1(Y) \geq 1, \end{cases} \tag{1}$$

for all $\mathfrak{s} \in \mathrm{Spin}^c(Y)$ [13, Proposition 5.1]. Thus for any rational homology 3-sphere Y, we have

$$\dim \widehat{HF}(Y) \geq \chi(\widehat{HF}(Y)) = |H_1(Y;\mathbb{Z})|.$$

A rational homology 3-sphere Y is an *L-space* if

$$\dim \widehat{HF}(Y) = |H_1(Y;\mathbb{Z})|.$$

Theorem 1 [8; 10; 15; 17]. *If $S^3_r(K)$ is an L-space for some rational slope $r > 0$, then K is fibered and strongly quasipositive, and $r \geq 2g(K) - 1$.*

All proofs of Theorem 1 in the literature use at least some of the following tools: the doubly-filtered Heegaard Floer complex associated to a knot, the large integer surgery formula, the $(\infty, 0, n)$-surgery exact triangle for $n > 1$, and the Spin^c decomposition of $\widehat{HF}(Y)$ for Y a rational homology sphere. This presents a major difficulty if one wishes to port this theorem to the instanton Floer setting, where none of this machinery is available.

MSC2020: primary 57R58; secondary 57K18, 57K31.
Keywords: L-spaces, Dehn surgery, Heegaard Floer homology.

Remark 2. A primary motivation for proving an analogue of Theorem 1 in the instanton Floer setting in particular is that such an analogue can be used to prove new results about the SU(2) representation varieties of fundamental groups of 3-manifolds obtained by Dehn surgeries on knots in the 3-sphere, about which relatively little is known; see [4].

Remark 3. Some of the structure mentioned above is known to exist in monopole Floer homology, though not enough of it to translate previous proofs of Theorem 1 to that setting. The new proof of Theorem 1 presented in this article (see below) *can* be adapted directly to monopole Floer homology, with the caveat in Remark 4, to give a proof of the monopole Floer analogue of Theorem 1 which does not rely on an isomorphism between monopole Floer homology and Heegaard Floer homology.

Our goal here is to give a proof of Theorem 1 using instead: the $(\infty, 0, 1)$-surgery exact triangle, the blow-up formula for cobordism maps, the adjunction inequality for cobordism maps, the Spin^c decomposition of the maps associated to 2-handle cobordisms, and Ozsváth and Szabó's description of the contact invariant $c^+(\xi)$ as the image of a certain class under the 2-handle cobordism map

$$HF^+(-S^3_0(K)) \to HF^+(-S^3),$$

where K is a fibered knot supporting the contact structure ξ on S^3. The first four of these tools will be used to show that an L-space knot is fibered, while the last will be used to prove that an L-space knot supports the tight contact structure on S^3 and is therefore strongly quasipositive, by Hedden [8]. Strong quasipositivity will then be used to prove the $2g(K) - 1$ bound on L-space surgery slopes.

Remark 4. Ozsváth and Szabó do not prove that $c^+(\xi)$ is well-defined (and hence that it certifies that ξ is tight) directly from its description in terms of the cobordism map associated to 0-surgery on the supporting fibered knot (and it is unclear how to do so—this is an interesting problem!). They instead use the knot filtration for this, which poses a challenge for translating the strong quasipositivity argument presented here to framed instanton homology. We discovered [4] a workaround in that setting, however, by a significantly more complicated argument which involves cabling and our framed instanton contact invariant [2]. We then used that instanton contact class to prove the $r \geq 2g(K)-1$ bound, in a manner very similar to the proof of Proposition 15 here (also using results from [3] and [9]). The same difficulties and solutions apply in monopole Floer homology, using our contact invariant from [1].[1]

[1] It is reasonable to expect that Kronheimer and Mrowka's monopole Floer contact class can be characterized in terms of the 0-surgery cobordism map as above, based on Echeverria's work [5], which would allow one to circumvent the more complicated strong quasipositivity argument we have in mind.

Figure 1. An outline of the proof of Proposition 5.

In our proof of Theorem 1, we will assume that K has genus at least 2 everywhere until Proposition 15, so that we can apply Theorem 6 below to detect whether $S^3_0(K)$ is fibered. The following proposition uses a cabling trick (see Figure 1) to show that we can still conclude Theorem 1 in full generality, and moreover that it suffices to prove the theorem for integral slopes.

Proposition 5. *If Theorem 1 holds for all knots of genus at least* 2 *and integral slopes* $r \in \mathbb{Z}$, *then it is also true for knots of genus* 1 *and* $r \in \mathbb{Q}$.

Proof. The claim in Theorem 1 about the set of positive integral L-space slopes is proved in Proposition 15 without any restrictions on $g(K)$, so we will only address the other claims of Theorem 1 here.

We suppose first that K is an arbitrary nontrivial knot and that some surgery on K of nonintegral slope $r > 0$ is an L-space. We write $r = \frac{p}{q}$ for some positive integers p and $q \geq 2$. By applying [6, Corollary 7.3], we see that

$$S^3_{pq}(K_{p,q}) \cong S^3_{p/q}(K) \# S^3_{q/p}(U),$$

where $K_{p,q}$ is the cable represented by the peripheral element $\mu^p \lambda^q$ in $\pi_1(\partial N(K))$. The two summands on the right are both L-spaces; hence the Künneth formula for $\widehat{HF}$ says that pq-surgery on $K_{p,q}$ is also an L-space. We observe that

$$g(K_{p,q}) = \frac{(p-1)(q-1)}{2} + q \cdot g(K),$$

which implies that $g(K_{p,q}) \geq q \geq 2$ and which is also equivalent to

$$2g(K_{p,q}) - 1 = pq + q\left(2g(K) - 1 - \frac{p}{q}\right). \tag{2}$$

We can now apply the assumed case of Theorem 1 to pq-surgery on $K_{p,q}$ to conclude that $pq \geq 2g(K_{p,q}) - 1$, hence $r = \frac{p}{q} \geq 2g(K) - 1$ by (2); and that $K_{p,q}$

is fibered and strongly quasipositive. Since $K_{p,q}$ is fibered, K must be as well. The strong quasipositivity of K then follows from two facts:

(1) a fibered knot is strongly quasipositive if and only if its corresponding open book decomposition supports the tight contact structure on S^3 [8], and

(2) the knots K and $K_{p,q}$ support the same contact structure [7].

This concludes the proof in all cases except when K has genus 1 and r is a positive integer, say $r = n$. In this case it is automatic that $r \geq 2g(K) - 1 = 1$. Moreover, repeated application of [15, Proposition 2.1], which follows easily from the surgery exact triangle for $\widehat{HF}$, says that $S^3_{(2n+1)/2}(K)$ is an L-space. (In fact, it says that $S^3_s(K)$ is an L-space for all rational $s \geq n$.) Since $\frac{1}{2}(2n+1) \notin \mathbb{Z}$, the fiberedness and strong quasipositivity of K follow exactly as above. □

We will suppose henceforth that $K \subset S^3$ is a knot of genus $g \geq 2$.

Let Σ_0 denote the genus g surface in $S^3_0(K)$ obtained by capping off a minimal genus Seifert surface for K. Let $\mathfrak{s}_i$ be the unique Spin^c structure on $S^3_0(K)$ satisfying

$$\langle c_1(\mathfrak{s}_i), [\Sigma_0] \rangle = 2i.$$

The adjunction inequality [13, Theorem 7.1] implies that

$$HF^+(S^3_0(K), \mathfrak{s}_i) = \widehat{HF}(S^3_0(K), \mathfrak{s}_i) = 0$$

for $|i| > g - 1$. Moreover, by [11], we have

$$HF^+(S^3_0(K), \mathfrak{s}_{g-1}) \neq 0$$

and Ni proved [10] (see also [12, Corollary 4.5]) the following.

Theorem 6. *K is fibered if and only if $HF^+(S^3_0(K), \mathfrak{s}_{g-1}) \cong \mathbb{F}$.*

Recall that there is an exact triangle

$$\begin{array}{ccc} HF^+(Y,\mathfrak{s}) & \xrightarrow{\cdot U} & HF^+(Y,\mathfrak{s}) \\ & {\scriptstyle i}\nwarrow \quad \swarrow{\scriptstyle j} & \\ & \widehat{HF}(Y,\mathfrak{s}) & \end{array} \tag{3}$$

where i and multiplication by U preserve the $\mathbb{Z}/2\mathbb{Z}$ grading and j shifts it by 1. Moreover, we claim the following.

Proposition 7. *U acts trivially on $HF^+(S^3_0(K), \mathfrak{s}_{g-1})$.*

Proof. Let

$$Z : (\Sigma_0 \times S^1) \sqcup S^3_0(K) \to S^3_0(K)$$

be the cobordism obtained from $S^3_0(K) \times I$ by removing a neighborhood of Σ_0 from the interior. The Spin^c structure $\mathfrak{s}_{g-1}$ on $S^3_0(K)$ extends to a product Spin^c structure on $S^3_0(K) \times I$, and we let $\mathfrak{t}$ denote the restriction of the latter to Z. Then the induced map

$$F_{Z,\mathfrak{t}} : HF^+(\Sigma_0 \times S^1, \mathfrak{t}|_{\Sigma_0 \times S^1}) \otimes HF^+(S^3_0(K), \mathfrak{s}_{g-1}) \to HF^+(S^3_0(K), \mathfrak{s}_{g-1})$$

is surjective. Since $\Sigma_0 \times \{\mathrm{pt}\} \subset \Sigma_0 \times S^1$ is homologous in Z to $\Sigma_0 \subset S^3_0(K) \times \{0\}$, we must have

$$\langle c_1(\mathfrak{t}), [\Sigma_0 \times \{\mathrm{pt}\}] \rangle = \langle c_1(\mathfrak{t}|_{S^3_0(K) \times \{0\}}), [\Sigma_0] \rangle = 2g - 2,$$

and evidently $HF^+(\Sigma_0 \times S^1, \mathfrak{t}|_{\Sigma_0 \times S^1})$ is nonzero. Thus $\mathfrak{t}|_{\Sigma_0 \times S^1}$ must be the unique Spin^c structure, which we also denote by $\mathfrak{s}_{g-1}$, satisfying

$$\langle c_1(\mathfrak{s}_{g-1}), [\Sigma_0 \times \{\mathrm{pt}\}] \rangle = 2g - 2 \quad \text{and} \quad \langle c_1(\mathfrak{s}_{g-1}), [\gamma \times S^1] \rangle = 0$$

for all closed curves $\gamma \subset \Sigma_0$, and we have

$$HF^+(\Sigma_0 \times S^1, \mathfrak{s}_{g-1}) \cong \mathbb{F}. \tag{4}$$

For details, see [12, Theorem 9.3] and the discussion preceding it.

It follows from (4) and the surjectivity of $F_{Z,\mathfrak{t}}$ that the cobordism map

$$F_{Z,\mathfrak{t}} : HF^+(\Sigma_0 \times S^1, \mathfrak{s}_{g-1}) \otimes HF^+(S^3_0(K), \mathfrak{s}_{g-1}) \to HF^+(S^3_0(K), \mathfrak{s}_{g-1})$$

is in fact an isomorphism. Moreover, it satisfies

$$F_{Z,\mathfrak{t}}(a \otimes Ub) = F_{Z,\mathfrak{t}}(Ua \otimes b).$$

The U-action on (4) is clearly trivial, which implies the same for $HF^+(S^3_0(K), \mathfrak{s}_{g-1})$ by the relation above. □

Proposition 7 together with the exact triangle in (3) implies that

$$\widehat{HF}(S^3_0(K), \mathfrak{s}_{g-1}) \cong HF^+(S^3_0(K), \mathfrak{s}_{g-1}) \oplus HF^+(S^3_0(K), \mathfrak{s}_{g-1})[1].$$

In particular, we have the following.

Corollary 8. *If K is fibered then*

$$\widehat{HF}(S^3_0(K), \mathfrak{s}_{g-1}) \cong \mathbb{F}_0 \oplus \mathbb{F}_1,$$

where the subscripts on the right denote the $\mathbb{Z}/2\mathbb{Z}$ grading. If K is not fibered then

$$\dim \widehat{HF}(S^3_0(K), \mathfrak{s}_{g-1}) \geq 4.$$

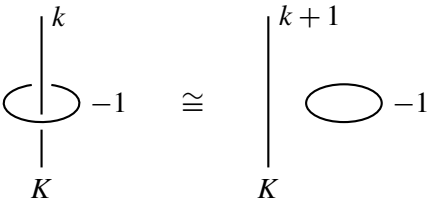


Figure 2. A handleslide showing that $W_{k+1} \circ X_k \cong X_{k+1}\#\overline{\mathbb{CP}}^2$.

We now consider the natural 2-handle cobordisms

$$S^3 \xrightarrow{X_k} S^3_k(K) \xrightarrow{W_{k+1}} S^3_{k+1}(K)$$

for each integer $k \geq 0$, where the 2-handle W_{k+1} is attached along a -1-framed meridian of K. We observe in Figure 2 that

$$W_{k+1} \circ X_k = X_k \cup_{S^3_k(K)} W_{k+1} \cong X_{k+1}\#\overline{\mathbb{CP}}^2,$$

and hence if we write

$$V_k = W_k \circ W_{k-1} \circ \cdots \circ W_1 : S^3_0(K) \to S^3_k(K)$$

then the composition

$$Z_k = V_k \circ X_0 : S^3 \to S^3_k(K)$$

is a k-fold blow-up of X_k, i.e.,

$$\begin{aligned} X_k\#k\overline{\mathbb{CP}}^2 &\cong W_k \circ (X_{k-1}\#(k-1)\overline{\mathbb{CP}}^2) \\ &\cong W_k \circ W_{k-1} \circ (X_{k-2}\#(k-2)\overline{\mathbb{CP}}^2) \\ &\cong \cdots \\ &\cong (W_k \circ \cdots \circ W_1) \circ X_0 = Z_k. \end{aligned}$$

The maps induced by X_k and W_{k+1} fit into an $(\infty, 0, 1)$-surgery exact triangle,

$$\begin{array}{ccc} \widehat{HF}(S^3) & \xrightarrow{F_{X_k}} & \widehat{HF}(S^3_k(K)) \\ & \nwarrow \quad \swarrow^{F_{W_{k+1}}} & \\ & \widehat{HF}(S^3_{k+1}(K)) & \end{array} \tag{5}$$

A Spin^c structure on X_0 is determined by its restriction to $S^3_0(K)$, or, equivalently, by the evaluation of its first Chern class on $[\Sigma_0]$. Let $\mathfrak{t}_i$ denote the unique Spin^c structure on X_0 with

$$\langle c_1(\mathfrak{t}_i), [\Sigma_0] \rangle = 2i.$$

Define

$$y_i := F_{X_0,\mathfrak{t}_i}(\mathbf{1}) \in \widehat{HF}(S^3_0(K), \mathfrak{s}_i),$$

where $\mathbf{1}$ denotes the generator of $\widehat{HF}(S^3) \cong \mathbb{F}$.

Let Σ_k denote the capped off Seifert surface in X_k, with

$$\Sigma_k \cdot \Sigma_k = k.$$

A Spin^c structure on X_k is determined by the evaluation of its first Chern class on $[\Sigma_k]$. Such Chern classes are characteristic elements, so this evaluation agrees with $k \pmod 2$. Let $\mathfrak{t}_{k,i}$ denote the unique Spin^c structure on X_k satisfying

$$\langle c_1(\mathfrak{t}_{k,i}), [\Sigma_k]\rangle + k = 2i.$$

The adjunction inequality [16, Proof of Theorem 1.5] implies that the map

$$F_{X_k,\mathfrak{t}_{k,i}} : \widehat{HF}(S^3) \to \widehat{HF}(S^3_k(K))$$

is nontrivial only if

$$|\langle c_1(\mathfrak{t}_{k,i}), [\Sigma_k]\rangle| + k \le 2g - 2,$$

or equivalently $1 - g + k \le i \le g - 1$.

Lemma 9. *Let $x_{k,i} = F_{X_k,\mathfrak{t}_{k,i}}(\mathbf{1})$ for all $k \ge 1$ and all i. Then*

$$F_{V_k}(y_i) = x_{k,i} + \binom{k}{1}x_{k,i+1} + \binom{k}{2}x_{k,i+2} + \cdots + \binom{k}{g-i-1}x_{k,g-1}$$

as elements of $\widehat{HF}(S^3_k(K))$.

Proof. Let $E_1, \ldots, E_k \subset Z_k$ denote the exceptional spheres in $Z_k \cong X_k \# k\overline{\mathbb{CP}}^2$, and $e_1, \ldots, e_k$ their Poincaré duals in $H^2(Z_k)$. Note that in Z_k, the surface Σ_0 is given by

$$\Sigma_0 = \Sigma_k - E_1 - \cdots - E_k.$$

In particular,

$$\langle c_1(\mathfrak{t}_{k,i} + a_1e_1 + \cdots + a_ke_k), [\Sigma_0]\rangle = 2i - k + a_1 + \cdots + a_k \tag{6}$$

in Z_k. We will evaluate F_{Z_k} by applying the blow-up formula for cobordism maps [16, Theorem 3.7], which says that for a Spin^c cobordism

$$(W, \mathfrak{t}) : (Y_1, \mathfrak{s}_1) \to (Y_2, \mathfrak{s}_2)$$

with blow-up $\widehat{W} = W\#\overline{\mathbb{CP}}^2$ and exceptional sphere E,

$$F_{\widehat{W},\mathfrak{t}\pm(2\ell+1)PD(E)} = \begin{cases} F_{W,\mathfrak{t}} & \text{if } \ell = 0, \\ 0 & \text{if } \ell \ne 0, \end{cases}$$

as maps on $\widehat{HF}$ for any $\ell \ge 0$.

Let F_i denote the component of $F_{Z_k} = F_{V_k} \circ F_{X_0}$ that factors through $\widehat{HF}(S^3_0(K), \mathfrak{s}_i)$. On the one hand, we have

$$F_i = F_{V_k} \circ F_{X_0, \mathfrak{t}_i} : \widehat{HF}(S^3) \to \widehat{HF}(S^3_k(K)).$$

On the other hand, if we let $e = e_1 + \cdots + e_k$, then for each i we have

$$F_i = F_{Z_k, \mathfrak{t}_{k,i}+e} + \sum_{j_1} F_{Z_k, \mathfrak{t}_{k,i+1}+e-2e_{j_1}} + \sum_{j_1<j_2} F_{Z_k, \mathfrak{t}_{k,i+2}+e-2e_{j_1}-2e_{j_2}} + \cdots + \sum_{j_1<\cdots<j_{g-i-1}} F_{Z_k, \mathfrak{t}_{k,g-1}+e-2e_{j_1}+\cdots-2e_{j_{g-i-1}}},$$

by the formula (6). From the blow-up formula, we have

$$F_{Z_k, \mathfrak{t}_{k,j} \pm e_1 \pm \cdots \pm e_k} = F_{X_k, \mathfrak{t}_{k,j}},$$

so the expression for F_i above becomes

$$F_i = F_{X_k, \mathfrak{t}_{k,i}} + \binom{k}{1} F_{X_k, \mathfrak{t}_{k,i+1}} + \binom{k}{2} F_{X_k, \mathfrak{t}_{k,i+2}} + \cdots + \binom{k}{g-i-1} F_{X_k, \mathfrak{t}_{k,g-1}}.$$

We conclude by evaluating both sides on the element $\mathbf{1} \in \widehat{HF}(S^3)$. □

Proposition 10. *For all integers $k \geq 1$, we have*

$$\ker\big(F_{V_k} : \widehat{HF}(S^3_0(K)) \to \widehat{HF}(S^3_k(K))\big) \subset \mathrm{Span}_{\mathbb{F}}(y_{1-g}, \ldots, y_{g-1}).$$

This inclusion is an equality for all $k \geq 2g-1$.

Proof. When $k = 1$, the exact triangle (5) says that

$$\ker(F_{V_1}) = \ker(F_{W_1}) = \mathrm{Im}(F_{X_0}) = \mathrm{Span}_{\mathbb{F}}\Big(\sum_{i=1-g}^{g-1} y_i\Big).$$

We prove the inclusion in general by induction on k.

Suppose that $k \geq 1$, and fix an element $z \in \ker(F_{V_{k+1}})$. Then

$$F_{W_{k+1}}\big(F_{V_k}(z)\big) = 0$$

by definition, so the exact triangle (5) tells us that $F_{V_k}(z) \in \mathrm{Im}(F_{X_k})$, or equivalently

$$F_{V_k}(z) = c \cdot F_{X_k}(\mathbf{1}) \tag{7}$$

for some $c \in \mathbb{F}$. Lemma 9 says that each element

$$x_{k,i} = F_{X_k, \mathfrak{t}_{k,i}}(\mathbf{1}) \in \widehat{HF}(S^3_k(K))$$

is a linear combination of the various $F_{V_k}(y_i)$, since the matrix of the coefficients of the system of linear equations relating $(F_{V_k}(y_i))_i$ to $(x_{k,i})_i$ is triangular and clearly invertible. In particular, summing over all i reveals that

$$F_{X_k}(\mathbf{1}) \in \mathrm{Span}_{\mathbb{F}}\big(F_{V_k}(y_{1-g}), \ldots, F_{V_k}(y_{g-1})\big). \tag{8}$$

Combining (7) and (8), there are coefficients $a_j \in \mathbb{F}$ such that

$$F_{V_k}(z) = c \cdot \sum_{j=1-g}^{g-1} a_j F_{V_k}(y_j),$$

or equivalently

$$z - \sum_{j=1-g}^{g-1} ca_j \cdot y_j \in \ker F_{V_k}. \tag{9}$$

By induction the left side of (9) lies in $\mathrm{Span}_{\mathbb{F}}(y_{1-g}, \dots, y_{g-1})$; hence the same is true of z. Since z was an arbitrary element of $\ker F_{V_{k+1}}$, this completes the inductive step.

To see that equality holds when $k \geq 2g-1$, we observe from Lemma 9 that $F_{V_k}(y_i)$ is a linear combination of various elements $x_{k,j} = F_{X_k,\mathfrak{t}_{k,j}}(\mathbf{1})$. Using the adjunction inequality, we have already noted that $x_{k,j} = 0$ unless

$$1 - g + k \leq j \leq g - 1,$$

so for $k \geq 2g-1$ the elements $x_{k,j}$ and hence the $F_{V_k}(y_i)$ are all zero. □

Proposition 11. *Suppose that $S^3_n(K)$ is an L-space for some positive integer n. Then*

$$\widehat{HF}(S^3_0(K), \mathfrak{s}_j) = \begin{cases} \mathbb{F}_0 \oplus \mathbb{F}_1 & \text{if } y_j \neq 0, \\ 0 & \text{if } y_j = 0, \end{cases}$$

for all j, where the subscripts on each copy of $\mathbb{F}$ denote the $\mathbb{Z}/2\mathbb{Z}$ grading.

Proof. We observe from (5) that

$$\dim_{\mathbb{F}} \widehat{HF}(S^3_{k+1}(K)) = \dim_{\mathbb{F}} \widehat{HF}(S^3_k(K)) + \begin{cases} 1 & \text{if } F_{X_k} = 0, \\ -1 & \text{if } F_{X_k} \neq 0, \end{cases} \tag{10}$$

for all $k \geq 0$. If m denotes the number of $k \in \{0, 1, \dots, n-1\}$ such that $F_{X_k} \neq 0$, then

$$n = \dim_{\mathbb{F}} \widehat{HF}(S^3_n(K)) = \dim_{\mathbb{F}} \widehat{HF}(S^3_0(K)) + (n-m) - m,$$

which simplifies to

$$\dim_{\mathbb{F}} \widehat{HF}(S^3_0(K)) = 2m. \tag{11}$$

Our goal is thus to compute m.

Supposing that $F_{X_k} \neq 0$ for some $k \geq 0$, then $F_{X_k}(\mathbf{1})$ is a nonzero element which spans $\ker(W_{k+1})$, and from (8) it has the form

$$F_{X_k}(\mathbf{1}) = F_{V_k}\left(\sum_{j=1-g}^{g-1} a_j y_j \right)$$

for some coefficients $a_j \in \mathbb{F}$. The sum $\sum a_j y_j$ is thus not in $\ker(F_{V_k})$, but it is in

$$\ker(F_{V_{k+1}}) = \ker(F_{W_{k+1}} \circ F_{V_k}),$$

so we have $\dim\ker(F_{V_{k+1}}) > \dim\ker(F_{V_k})$. This implies that

$$\dim\ker(F_{V_n}) \geq m.$$

Proposition 10 then implies that

$$m \leq \dim \mathrm{Span}_{\mathbb{F}}(y_j). \tag{12}$$

But the nonzero y_j are all linearly independent, since they belong to different summands $\widehat{HF}(S^3_0(K), \mathfrak{s}_j)$ of $\widehat{HF}(S^3_0(K))$, so by combining (11) and (12) we conclude that

$$\dim_{\mathbb{F}} \widehat{HF}(S^3_0(K)) \leq 2 \cdot \#\{j \mid y_j \neq 0\}. \tag{13}$$

If $y_j \neq 0$ then $\widehat{HF}(S^3_0(K), \mathfrak{s}_j)$ is nonzero, and its Euler characteristic is zero by (1), so

$$\mathbb{F}_0 \oplus \mathbb{F}_1 \subset \widehat{HF}(S^3_0(K), \mathfrak{s}_j) \quad \text{if } y_j \neq 0.$$

Thus the inequality in (13) must be an equality, and each nonzero $\widehat{HF}(S^3_0(K), \mathfrak{s}_j)$ must have the form $\mathbb{F}_0 \oplus \mathbb{F}_1$, completing the proof. □

Proposition 12. *If $S^3_n(K)$ is an L-space for some integer $n > 0$ then K is fibered.*

Proof. Corollary 8 and Proposition 11 tell us that

$$2 \leq \dim \widehat{HF}(S^3_0(K), \mathfrak{s}_{g-1}) \leq 2, \tag{14}$$

and that equality on the left holds if and only if K is fibered, so K must be fibered. □

Proposition 13. *If $S^3_n(K)$ is an L-space for some integer $n > 0$ then K is strongly quasipositive.*

Proof. We already have seen in (14) that $\widehat{HF}(S^3_0(K), \mathfrak{s}_{g-1})$ is nonzero; hence $y_{g-1} \neq 0$ by Proposition 11. Equivalently, the map

$$\widehat{HF}(S^3) \to \widehat{HF}(S^3_0(K), \mathfrak{s}_{g-1}) \tag{15}$$

induced by X_0 is nonzero. Now, we can also view X_0 as a cobordism

$$X_0 : -S^3_0(K) \to -S^3,$$

in which case the induced map

$$\widehat{HF}(-S^3_0(K), \mathfrak{s}_{1-g}) \to \widehat{HF}(-S^3)$$

is dual to that in (15). In particular, this map is also nonzero. The commutativity of

$$\begin{array}{ccc} \widehat{HF}(-S^3_0(K), \mathfrak{s}_{1-g}) & \xrightarrow{i} & HF^+(-S^3_0(K), \mathfrak{s}_{1-g}) \\ \downarrow & & \downarrow \\ \widehat{HF}(-S^3) & \xrightarrow{i} & HF^+(-S^3) \end{array}$$

where the vertical maps are those induced by X_0, together with the facts that $\widehat{HF}(-S^3) \cong \mathbb{F}$ and the bottom horizontal map is nonzero, implies that the rightmost vertical map

$$HF^+(-S^3_0(K), \mathfrak{s}_{1-g}) \to HF^+(-S^3) \tag{16}$$

is nonzero as well. But Proposition 12 says that K is fibered; hence

$$HF^+(-S^3_0(K), \mathfrak{s}_{1-g}) \cong \mathbb{F},$$

and the image of its generator under the map in (16) is the contact invariant $c^+(\xi_K)$ [14], where ξ_K is the contact structure corresponding to K. Thus, $c^+(\xi_K)$ is nonzero, which implies that ξ_K is the tight contact structure on S^3. It follows that K is strongly quasipositive, by work of Hedden [8, Proposition 2.1]. □

We will now use the fact that L-space knots are strongly quasipositive to determine the range of L-space slopes for any such knot. We begin with the following general lemma.

Lemma 14. *Let Y be a rational homology sphere with $|H_1(Y;\mathbb{Z})| = n$. Suppose that*

$$\ker\big(U : HF^+(Y) \to HF^+(Y)\big)$$

has dimension $n+k$. Then $\dim \widehat{HF}(Y) = n + 2k$.

Proof. The exact triangle (3) involving the U-action on $HF^+(Y)$ produces a short exact sequence

$$0 \to \operatorname{coker}(U) \to \widehat{HF}(Y) \to \ker(U) \to 0.$$

Thus it will suffice to show that $\dim \operatorname{coker}(U) = k$.

Since each Spin^c structure on Y is torsion, we have a short exact sequence

$$0 \to (\mathcal{T}^+)^{\oplus n} \to HF^+(Y) \xrightarrow{\pi} HF_{\mathrm{red}}(Y) \to 0,$$

of $\mathbb{F}[U]$-modules, where $\mathcal{T}^+ \cong \mathbb{F}[U, U^{-1}]/U\,F[U]$. The quotient $HF_{\mathrm{red}}(Y)$ is defined as $HF^+(Y)/\mathrm{Im}(U^d)$ for $d \gg 0$; it is finitely generated over $\mathbb{F}[U]$ and over $\mathbb{F}$, and every element is U-torsion, so it has a decomposition

$$HF_{\mathrm{red}}(Y) \cong \bigoplus_{i=1}^{r} \mathbb{F}[U]/\langle U^{n_i}\rangle,$$

with each $n_i \geq 1$. Moreover this sequence can be shown to split, so that

$$HF^+(Y) \cong (\mathcal{T}^+)^{\oplus n} \oplus \bigoplus_{i=1}^{r} \mathbb{F}[U]/\langle U^{n_i}\rangle.$$

But then it is clear that $\ker(U) \cong \mathbb{F}^{n+r}$, so that $r = k$, and then that $\operatorname{coker}(U) \cong \mathbb{F}^r = \mathbb{F}^k$, and the lemma follows immediately. □

The following proposition completes our proof of Theorem 1. The proof below is partly inspired by the work of Lidman, Pinzón-Caicedo, and Scaduto [9].

Proposition 15. *If K has genus $g \geq 1$ and $S^3_n(K)$ is an L-space for some positive integer n, then $S^3_n(K)$ is an L-space for an arbitrary integer n if and only if $n \geq 2g-1$.*

Proof. Since K is strongly quasipositive, its maximal self-linking number is $\overline{sl}(K) = 2g-1$. We take a Legendrian representative Λ of K in the standard contact S^3 with classical invariants

$$(tb(\Lambda), r(\Lambda)) = (\tau_0, r_0), \quad \tau_0 - r_0 = 2g-1,$$

and for $n \geq 1-\tau_0$, we can positively stabilize this k times and negatively stabilize it $\tau_0 + n - 1 - k$ times to get a Legendrian representative with

$$(tb, r) = (1-n, 2-2g-n+2k), \quad 0 \leq k \leq \tau_0 + n - 1.$$

For odd $n \gg 0$, these values of r include every positive odd number between 1 and $n+2g-2$.

Fixing such a large value of n, we perform Legendrian surgery on these knots Λ_i with

$$(tb(\Lambda_i), r(\Lambda_i)) = (1-n, 2i-1), \quad 1 \leq i \leq \frac{n+2g-1}{2},$$

to get contact structures

$$\xi_1, \ldots, \xi_{(n+2g-1)/2}$$

on $S^3_{-n}(K)$. If $X_{-n}(K)$ is the trace of this $-n$-surgery, and $\widehat{\Sigma} \subset X_{-n}(K)$ the union of a Seifert surface for K with the core of the 2-handle, then each ξ_i admits a Stein filling $(X_{-n}(K), J_i)$ with

$$\langle c_1(J_i), [\widehat{\Sigma}] \rangle = r(\Lambda_i) = 2i-1.$$

We can also take contact structures

$$\bar{\xi}_i = T(S^3_{-n}(K)) \cap \bar{J}_i T(S^3_{-n}(K)), \quad 1 \leq i \leq \frac{n+2g-1}{2},$$

which are filled by $X_{-n}(K)$ with the conjugate Stein structure $\bar{J}_i$ for each i. These satisfy $\langle c_1(\bar{J}_i), [\widehat{\Sigma}] \rangle = -(2i-1)$, so we have exhibited $n+2g-1$ Stein structures

$$J_1, J_2, \ldots, J_{(n+2g-1)/2}, \bar{J}_1, \bar{J}_2, \bar{J}_{(n+2g-1)/2}$$

on $X_{-n}(K)$ which are all distinguished by their first Chern classes.

A theorem of Plamenevskaya [18, Theorem 4] now tells us that the corresponding contact invariants

$$c^+(\xi_1), \dots, c^+(\xi_{(n+2g-1)/2}), c^+(\bar{\xi}_1), \dots, c^+(\bar{\xi}_{(n+2g-1)/2}) \in HF^+(-S^3_{-n}(K))$$

are linearly independent. These elements lie in $\ker(U)$, as can be seen, for example, from the fact that they are by defined by maps of the form (16) whose domains have trivial U action. Thus

$$\dim \ker(U) \geq n + 2g - 1,$$

and it follows from Lemma 14 that

$$\dim \widehat{HF}(S^3_{-n}(K)) = \dim \widehat{HF}(-S^3_{-n}(K)) \geq n + 4g - 2.$$

This same argument applies for any larger odd value of n as well, and the conclusion also holds for even values of n after making only cosmetic changes to the argument, so that $S^3_{-m}(K)$ cannot be an L-space for any $m \geq n$.

We now repeatedly apply the surgery exact triangle (5) to see that

$$\dim \widehat{HF}(S^3_{-m}(K)) \geq m + 4g - 2, \quad 0 \leq m \leq n,$$

and then that

$$\dim \widehat{HF}(S^3_m(K)) \geq 4g - 2 - m \geq m + 2, \quad 0 \leq m \leq 2g - 2.$$

Thus $S^3_m(K)$ cannot be an L-space for any integer $m < 2g(K) - 1$. On the other hand, equation (10) says that

$$\dim \widehat{HF}(S^3_{2g-1+n}(K)) = \dim \widehat{HF}(S^3_{2g-1}(K)) + n$$

for all $n \geq 0$, since the maps $F_{X_{2g-1}}, \dots, F_{X_{2g-2+n}}$ are all zero by the adjunction inequality. Thus $S^3_{2g-1+n}(K)$ is an L-space if and only if $S^3_{2g-1}(K)$ is, and this completes the proof. □

Acknowledgments

We thank Jen Hom and Tye Lidman for helpful conversations, and the referee for useful feedback. Baldwin was supported by NSF CAREER grant DMS-1454865.

References

[1] J. A. Baldwin and S. Sivek, "A contact invariant in sutured monopole homology", *Forum Math. Sigma* **4** (2016), art. id. e12. MR Zbl

[2] J. A. Baldwin and S. Sivek, "Instanton Floer homology and contact structures", *Selecta Math. (N.S.)* **22**:2 (2016), 939–978. MR Zbl

[3] J. A. Baldwin and S. Sivek, "Stein fillings and SU(2) representations", *Geom. Topol.* **22**:7 (2018), 4307–4380. MR Zbl

[4] J. A. Baldwin and S. Sivek, "Instantons and L-space surgeries", 2019. arXiv 1910.13374

[5] M. Echeverria, "Naturality of the contact invariant in monopole Floer homology under strong symplectic cobordisms", *Algebr. Geom. Topol.* **20**:4 (2020), 1795–1875. MR Zbl

[6] C. M. Gordon, "Dehn surgery and satellite knots", *Trans. Amer. Math. Soc.* **275**:2 (1983), 687–708. MR Zbl

[7] M. Hedden, "Some remarks on cabling, contact structures, and complex curves", pp. 49–59 in *Proceedings of Gökova Geometry-Topology Conference* 2007, Gökova Geometry/Topology Conference (GGT), Gökova, 2008. MR Zbl

[8] M. Hedden, "Notions of positivity and the Ozsváth–Szabó concordance invariant", *J. Knot Theory Ramifications* **19**:5 (2010), 617–629. MR Zbl

[9] T. Lidman, J. Pinzón-Caicedo, and C. Scaduto, "Framed instanton homology of surgeries on L-space knots", 2020. arXiv 2003.03329

[10] Y. Ni, "Knot Floer homology detects fibred knots", *Invent. Math.* **170**:3 (2007), 577–608. MR Zbl

[11] P. Ozsváth and Z. Szabó, "Holomorphic disks and genus bounds", *Geom. Topol.* **8** (2004), 311–334. MR

[12] P. Ozsváth and Z. Szabó, "Holomorphic disks and knot invariants", *Adv. Math.* **186**:1 (2004), 58–116. MR

[13] P. Ozsváth and Z. Szabó, "Holomorphic disks and three-manifold invariants: properties and applications", *Ann. of Math.* (2) **159**:3 (2004), 1159–1245. MR

[14] P. Ozsváth and Z. Szabó, "Heegaard Floer homology and contact structures", *Duke Math. J.* **129**:1 (2005), 39–61. MR

[15] P. Ozsváth and Z. Szabó, "On knot Floer homology and lens space surgeries", *Topology* **44**:6 (2005), 1281–1300. MR

[16] P. Ozsváth and Z. Szabó, "Holomorphic triangles and invariants for smooth four-manifolds", *Adv. Math.* **202**:2 (2006), 326–400. MR

[17] P. S. Ozsváth and Z. Szabó, "Knot Floer homology and rational surgeries", *Algebr. Geom. Topol.* **11**:1 (2011), 1–68. MR

[18] O. Plamenevskaya, "Contact structures with distinct Heegaard Floer invariants", *Math. Res. Lett.* **11**:4 (2004), 547–561. MR Zbl

Received 22 Oct 2020. Revised 17 Dec 2020.

JOHN A. BALDWIN: john.baldwin@bc.edu
Department of Mathematics, Boston College, Chestnut Hill, MA, United States

STEVEN SIVEK: s.sivek@imperial.ac.uk
Department of Mathematics, Imperial College London, London, United Kingdom

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.95

Tangles, relative character varieties, and holonomy perturbed traceless flat moduli spaces

Guillem Cazassus, Chris Herald and Paul Kirk

We prove that the restriction map from the subspace of regular points of the holonomy perturbed SU(2) traceless flat moduli space of a tangle in a 3-manifold to the traceless flat moduli space of its boundary marked surface is a Lagrangian immersion. A key ingredient in our proof is the use of composition in the Weinstein category, combined with the fact that SU(2) holonomy perturbations in a cylinder induce Hamiltonian isotopies. In addition, we show that $(S^2, 4)$, the 2-sphere with four marked points, is its own traceless flat SU(2) moduli space.

1. Introduction

We gather together some of the key symplectic properties of character varieties and traceless character varieties, as well as variants which correspond to perturbed flat moduli spaces that arise in the gauge theoretic study of 3-manifolds. Some of these results are well known to the experts, but the proofs in the literature are framed in contexts that include gauge theory, Hodge theory, and symplectic reduction. In the present exposition, we provide a general proof of the fact that, roughly, the character variety of a 3-manifold provides an immersed Lagrangian in the character variety of its boundary surface, for any compact Lie group G, whether the 3-manifold and its boundaries include tangles, or whether there are trace or other conjugacy restrictions on some meridional generators, and furthermore we extend the result to the holonomy perturbed situation. Moreover, we clarify why different natural definitions of symplectic structures on the pillowcase, arising as the character variety of the torus or the 2-sphere with four marked points, are equivalent. The results are proved using only the Poincaré-Lefschetz duality theorem, basic algebraic topology, and the notion of composition in the Weinstein category.

Herald was supported by a Simons Collaboration Grant for Mathematicians, Cazassus was funded by EPSRC grant reference EP/T012749/1.

MSC2020: primary 57K18, 57K31, 57R58; secondary 81T13.

Keywords: holonomy perturbation, Lagrangian immersion, Floer homology, flat moduli space, traceless character variety, quilted Floer homology.

Let $(X, \mathcal{L})$ be a tangle in a compact, oriented 3-manifold X; that is, assume that $\mathcal{L}$ is a properly embedded, compact 1-manifold. For our initial discussion, we consider $G = \mathrm{SU}(2)$. Let π denote some holonomy perturbation data supported in a finite disjoint union of solid tori in the interior of $X \setminus \mathcal{L}$.

This data determines two moduli spaces,

$$\mathcal{M}_\pi(X, \mathcal{L}),$$

the moduli space of π *holonomy-perturbed flat* $\mathrm{SU}(2)$ *connections on* $X \setminus \mathcal{L}$ *with traceless holonomy on small meridians of* $\mathcal{L}$, and the (well studied) moduli space

$$\mathcal{M}(\partial X, \partial \mathcal{L})$$

of *flat* $\mathrm{SU}(2)$ *connections on the punctured surface* $\partial X \setminus \partial \mathcal{L}$ *with traceless holonomy around the marked points* $\partial \mathcal{L}$.

Holonomy perturbed flat connections on a 3-manifold are flat near the boundary, and restriction to the boundary defines a map

$$r : \mathcal{M}_\pi(X, \mathcal{L}) \to \mathcal{M}(\partial X, \partial \mathcal{L}). \tag{1-1}$$

The moduli space $\mathcal{M}(\partial X, \partial \mathcal{L})$ is the cartesian product of the moduli spaces of its path components. The flat $\mathrm{SU}(2)$ moduli space of an oriented connected surface of genus g with k marked points is known to be a singular variety with smooth top stratum carrying a symplectic form called the *Atiyah–Bott–Goldman form* [3; 12]. Thus the smooth top stratum of the cartesian product $\mathcal{M}(\partial X, \partial \mathcal{L})$, denoted $\mathcal{M}(\partial X, \partial \mathcal{L})^*$, is endowed with the product symplectic form.

The main result of this article is the following theorem (Theorem 6.9 below), concerning the regular points (see Section 4C) of the perturbed moduli space.

Theorem A. *Suppose* $A \in \mathcal{M}_\pi(X, \mathcal{L})$ *is a regular point. Then* A *has a neighborhood* U *so that the restriction* $r|_U : U \subset \mathcal{M}_\pi(X, \mathcal{L}) \to \mathcal{M}(\partial X, \partial \mathcal{L})^*$ *is a Lagrangian immersion.*

This is not a surprising result; indeed, many special cases are known, for example, when $\mathcal{L}$ is empty this result is proven in [16]. Our primary aim is to provide details of the assertion that well-known arguments in the flat case extend to the holonomy-perturbed flat case when $\mathcal{L}$ is nonempty, in support of one claim of the main result of [7]. In that article it is shown that a certain process introduced by Kronheimer and Mrowka to ensure admissibility of bundles in instanton homology [21] manifests itself on the symplectic side of the Atiyah–Floer conjecture [1] (i.e., Lagrangian Floer theory of character varieties or flat moduli spaces) as a certain Lagrangian immersion of a smooth closed genus 3 surface into the smooth stratum $\mathbb{P}^* \times \mathbb{P}^*$ of the product of two pillowcases (cf. (1-3)). What is proved in [7] is that this genus 3 surface satisfies the hypotheses of Theorem A at every point.

Since it causes no extra work, we take the opportunity to provide an elementary algebraic topology proof of Theorem A *in the flat case*, for any compact Lie group G. The statement can be found in Corollary 4.5. We emphasize that the flat case of Theorem A, when $\mathcal{L}$ is nonempty, is known (e.g., to those who attended the appropriate Oxford seminars in the 1980s) and indeed discussed on pages 15–16 of Atiyah's monograph [2].

Having stated Theorem A and described its relation to the gauge theoretic literature using the language of flat and perturbed flat moduli spaces, we note that these spaces can also be identified with certain character varieties, the definitions of which do not require any of the analytical machinery of gauge theory; the proofs in this article are most simply explained without it, so we shall henceforth revert to the character variety terminology in our exposition. In the simplest situation of a connected 2- or 3-manifold M with base point x_0, for a fixed compact Lie group G, each flat G connection determines a holonomy representation from $\pi_1(M, x_0)$ to G, and this correspondence induces a bijection between the moduli space of flat connections and the set of G representations of π_1 modulo conjugation, known as the *character variety*. We describe various extensions of the notion of the character variety corresponding to traceless and perturbed flat moduli spaces below.

We also take this opportunity to provide an elementary exposition of holonomy perturbations from the perspective of fundamental groups and character varieties, in the language of composition of Lagrangian immersions. Algebraic topology arguments simplify the task of explaining how to understand the extension from the flat to the holonomy perturbed flat situation algebraically. Other arguments, which instead appeal to Hodge and elliptic theory of perturbed flat bundles over 3-manifolds with boundary, can be made when $\mathcal{L}$ is empty, and can be found in detail in [16; 27]. But when $\mathcal{L}$ is nonempty, proper treatment of the traceless condition requires more subtle analytic tools.

Our focus on holonomy perturbations is motivated by the fact that they are compatible with the analytical framework of the instanton gauge theory side of the Atiyah–Floer conjecture [10; 27]. Holonomy perturbations modify the flatness condition (i.e., the nonlinear PDE *Curvature* $= 0$) in a specific way on a collection of solid tori, as described in Lemma 8.1 of Taubes [27] (see also Lemma 16 of [16]). We translate this result into the language of character varieties in Section 6B.

In addition to proving Theorem A, we prove Theorem B, whose statement roughly says *the four punctured* 2*-sphere is its own traceless* SU(2) *moduli space*. This is a variant, for $(S^2, 4)$, of the ubiquitous mathematical statement that a torus is isomorphic to its Jacobian. We now set some notation in preparation for the formal statement. First, denote $U(1) \times U(1)$ by $\mathbb{T}$. The group $\mathrm{SL}(2, \mathbb{Z})$ acts on $\mathbb{T}$ via

$$\begin{pmatrix} p & r \\ q & s \end{pmatrix} \cdot (e^{x\boldsymbol{i}}, e^{y\boldsymbol{i}}) = (e^{(px+ry)\boldsymbol{i}}, e^{(qx+sy)\boldsymbol{i}}). \tag{1-2}$$

The 2-form $dx \wedge dy$ defines the standard symplectic structure on $\mathbb{T}$ and is invariant under the $\mathrm{SL}(2,\mathbb{Z})$ action. Note that the action of the central element $-1 \in \mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{T}$ defines the *elliptic involution*, denoted by $\iota(e^{ix}, e^{iy}) = (e^{-ix}, e^{-iy})$; this involution has four fixed points $(\pm 1, \pm 1)$. Denote by $\mathbb{T}^* \subset \mathbb{T}$ the complement of the four fixed points, on which ι acts freely. Then set

$$\mathbb{P} = \mathbb{T}/\iota \quad \text{and} \quad \mathbb{P}^* = \mathbb{T}^*/\iota. \tag{1-3}$$

The quotient $\mathrm{PSL}(2,\mathbb{Z}) = \mathrm{SL}(2,\mathbb{Z})/\{\pm 1\}$ acts on the smooth locus $\mathbb{P}^*$, a 4-punctured 2-sphere, preserving the symplectic form $dx \wedge dy$.

Next, we consider the relative character variety (defined below) of the 2-sphere with four marked points $\chi_{\mathrm{SU}(2),J}(S^2, 4)$, where the J subscript indicates the four meridians are sent to the conjugacy class J of traceless elements. This character variety is equipped with its relative Atiyah–Bott–Goldman 2-form $\omega_{(S^2,4)}$ (defined below).

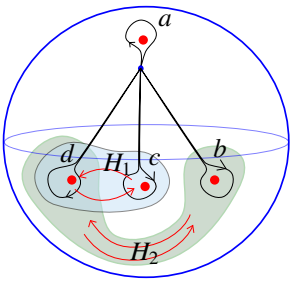


Figure 1. Half Dehn twists along the boundaries of H_1 and H_2.

From the presentation $\pi_1(S^2\backslash 4) = \langle a, b, c, d \,|\, abcd = 1\rangle$, we see that this group is freely generated by a, b, c. Hence a representation is uniquely defined by where it sends the generators a, b, c. Define the function (identifying $\mathrm{SU}(2)$ with the group of unit quaternions; see Section 2A2):

$$\hat{\rho} : \mathbb{P} \to \chi_{\mathrm{SU}(2),J}(S^2, 4), \quad [e^{ix}, e^{iy}] \mapsto \left[a \mapsto \boldsymbol{j}, b \mapsto e^{ix}\boldsymbol{j}, c \mapsto e^{iy}\boldsymbol{j}\right].$$

Theorem B. *Half Dehn twists in the two twice-punctured disks indicated in Figure 1 generate a* $\mathrm{PSL}(2,\mathbb{Z})$ *action on* $\chi(S^2, 4)$ *which preserves the Atiyah–Bott–Goldman symplectic form* $\widehat{\omega}_{(S^2,4)}$. *For some nonzero constant* c, *the map*

$$\hat{\rho} : (\mathbb{P}^*, c\, dx \wedge dy) \to (\chi(S^2, 4)^*, \widehat{\omega}_{(S^2,4)})$$

is a $\mathrm{PSL}(2,\mathbb{Z})$ *equivariant symplectomorphism.*[1]

The proof, as well as an exposition of the simpler case of $\mathrm{SU}(2)$ character variety of the torus, is contained in Section 5.

The outline of the proof of the flat case of Theorem A is the following. Holonomy identifies the flat moduli spaces with character varieties. When $\mathcal{L}$ is empty, Theorem A follows from Weil's identification of the Zariski tangent spaces of character varieties with cohomology and Poincaré-Lefschetz duality. Symplectic reduction is used to extend to the case when $\mathcal{L}$ is nonempty. We use only Poincaré

[1] The constant c equals $\frac{1}{2}$.

duality with local coefficients, which is briefly reviewed, to highlight the fact that the proof that the image of the differential is *maximal* isotropic (the subtlest part of any proof) does not depend on the deeper result that the nondegenerate 2-form is closed, i.e., symplectic.

To put Theorem A in context, notice that the Lagrangian immersion r_U depends on the perturbation data π. Building on ideas from [1; 10; 32] and many others, Wehrheim and Woodward developed quilted Lagrangian–Floer homology and Floer field theory [29; 30]; a framework that aims to produce a $2+1$ or, more generally, $(2,0)+1$ TQFT which factors as the composition of the (perturbed) flat moduli space functor, followed by passing to the Lagrangian Floer theory of the flat moduli spaces of surfaces equipped with the Atiyah–Bott symplectic form. This can be considered as an approach to realizing a *bordered* Lagrangian Floer theory of character varieties of surfaces and 3-manifolds, as was done for Heegaard–Floer theory in [23].

Even in the lucky case that one finds perturbation data π for which $\mathcal{M}_\pi(X,\mathcal{L})$ is smooth of the correct dimension and Lagrangian immerses into the smooth stratum $\mathcal{M}(\partial(X,\mathcal{L}))^*$, an understanding of how this immersion depends on π is necessary in order to extract topological information. If each perturbation curve is parallel to an embedded curve in the boundary surface, for example, then varying the perturbation parameters changes r by a Hamiltonian isotopy, and in particular has no effect on the topology of $\mathcal{M}_\pi(X,\mathcal{L})$. This is discussed in Section 6B; see also [17].

However, for general choices of perturbation data, perturbations typically change the topology of $\mathcal{M}_\pi(X,\mathcal{L})$; indeed, the primary purpose of using perturbation curves is precisely to smooth $\mathcal{M}_\pi(X,\mathcal{L})$. As we discuss in Section 6C, Floer field theory posits that, nevertheless, the resulting immersions should be independent of π in a Floer-theoretic sense (i.e., isomorphic in some Fukaya category).

2. Review of Poincaré duality and symplectic forms

2A. *Preliminaries and notation.*

2A1. *Symplectic linear algebra.* Let A denote a finite-dimensional $\mathbb{R}$-vector space. A skew symmetric bilinear form $\omega : A \times A \to \mathbb{R}$ is called a *symplectic form* if it is nondegenerate, that is, the map $A \to \mathrm{Hom}(A,\mathbb{R})$ given by $a \mapsto \omega(a,-)$ is an isomorphism. If A admits a symplectic form, then its dimension is even; denote the dimension by $2n$.

A *coisotropic subspace* $C \subset A$ is a subspace satisfying

$$\mathrm{Annihilator}(C) := \{a \in A \mid \omega(c,a) = 0 \text{ for all } c \in C\} \subset C,$$

and an *isotropic subspace* $C \subset A$ is a subspace satisfying

$$C \subset \mathrm{Annihilator}(C) := \{a \in A \mid \omega(c,a) = 0 \text{ for all } c \in C\}.$$

A *Lagrangian subspace* is a coisotropic and isotropic subspace, that is, a subspace $L \subset A$ satisfying $\text{Annihilator}(L) = L$. Equivalently, a Lagrangian subspace $L \subset A$ is a coisotropic (or isotropic) subspace satisfying $\dim L = \frac{1}{2}\dim A$.

Symplectic reduction refers to the following process. Given any coisotropic subspace $C \subset A$, the restriction

$$\omega|_C : C \times C \to \mathbb{R}$$

may be degenerate, but $\omega|_C$ descends to a symplectic form $\widehat{\omega}$ on

$$\widehat{C} := C/\text{Annihilator}(C).$$

Furthermore if $L \subset A$ is any Lagrangian subspace of (A, ω), then

$$\hat{L} := (L \cap C)/\big(L \cap \text{Annihilator}(C)\big) \subset \widehat{C}$$

is a Lagrangian subspace of $(\widehat{C}, \widehat{\omega})$. The subquotient $(\widehat{C}, \widehat{\omega})$ is called the *symplectic reduction of* (A, ω) *with respect to* C and the Lagrangian subspace $\hat{L} \subset \widehat{C}$ the *symplectic reduction of* L *with respect to* C.

2A2. *The Lie group G.* Let G be connected compact Lie group. Its Lie algebra $\mathfrak{g}$ admits a positive definite symmetric Ad-invariant bilinear form, so we fix one and denote it by $\{\ ,\ \} : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$. Fix a conjugacy class $J \subset G$. It is well known that G embeds in some $\mathbb{R}^N$ as an algebraic variety. Thus, if F_g denotes the free group on g generators, $\text{Hom}(F_g, G) = G^g$ is an affine real-set variety, with tangent space $\mathfrak{g}^g$ at the trivial homomorphism.

Our main focus, and the context for Theorem A, concerns the case when $G = \text{SU}(2)$, $\{v, w\} = -\frac{1}{2}\text{Tr}(vw)$, and $J \subset \text{SU}(2)$ is the conjugacy class of traceless matrices. To keep notation compact, we identify SU(2) with the group of unit quaternions

$$\{a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} \mid a^2 + b^2 + c^2 + d^2 = 1\}$$

and its Lie algebra su(2) with the purely imaginary quaternions $\{b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}\}$. With this identification, $\text{Re} : \text{SU}(2) \to [-1, 1]$ corresponds to $\frac{1}{2}\text{Tr}$.

2A3. *Notation used for 2- and 3-manifolds.* Throughout this article, the notation Y, S, C, V, and F is fixed as follows.

First, S denotes a possibly disconnected compact oriented surface without boundary. Denote by S_+ the path connected based space obtained by first adding a disjoint base point to S, then attaching a 1-cell from this new base point to each path component of S.

Next, C denotes a finite disjoint union of m circles in the surface S. Number its components C_i, $i = 1, \dots m$. We assume that either each C_i is oriented, or the chosen conjugacy class $J \subset G$ is invariant under inversion in G, so that the condition that a homomorphism from $\pi_1(S_+) \to G$ takes each loop C_i into J makes sense.

The pair (S, C) determines a decomposition of S into two surfaces as follows.

Denote by V a tubular neighborhood of C. Then $V \subset S$ is a disjoint union of oriented annuli, one around each C_i.

Denote by F the complementary surface, determining the decomposition:

$$S = F \cup_{\partial F = \partial V} V. \tag{2-1}$$

Orient V and F as subsurfaces of S.

Finally, Y denotes a compact, connected, and oriented 3-manifold with boundary $\partial Y = S$; that is, we assume that an orientation preserving identification of ∂Y with S is given. Fix a base point in the interior of Y, and extend the inclusion $S = \partial Y \subset Y$ to a based embedding of S_+ into Y.

2A4. *Tangles.* A tangle $(X, \mathcal{L})$ consists of a connected compact oriented 3-manifold X, and $\mathcal{L} \subset X$ a properly embedded compact 1-submanifold with boundary. Thus $\mathcal{L}$ consists of a disjoint union of n intervals and k interior circles.

A tangle $(X, \mathcal{L})$ gives rise to a triple (Y, S, C) as above by taking

$$Y = X \setminus N(\mathcal{L}),$$

where $N(\mathcal{L})$ denotes a tubular neighborhood of $\mathcal{L}$, and letting $S = \partial Y$. Then let $C \subset S$ denote a union of $m = n + k$ meridians of $\mathcal{L}$, one for each component, viewed as a curve in S. As before, Y, S, C determine V and F. Orientations of the tangle components are equivalent to orientations of the components of C.

The process $(X, \mathcal{L}) \Rightarrow (Y, S, C)$ is nearly reversible, by attaching 2-handles to Y along C and setting $\mathcal{L}$ to be union of the co-cores of the 2-handles. The resulting tangle is obtained from $(X, \mathcal{L})$ by removing k disjoint small balls from the interior of X, each meeting a different closed component of $\mathcal{L}$ in a trivial arc. This results in a tangle with no closed components, but which has the same character variety as the starting tangle. We use the notation (Y, S, C) in our arguments as it leads to simpler expressions, but state consequences in terms of the tangle $(X, \mathcal{L})$, as they are clearly seen as morphisms in a (more familiar) $(2, 0)+1$ cobordism category.

2B. ***Poincaré duality and intersection forms.*** We begin by recalling the statement of Poincaré-Lefschetz duality with local coefficients for an oriented compact connected n-manifold M with boundary ∂M equipped with a finite cell decomposition (see [8] for a careful exposition and proofs, and [24] for an elementary derivation using dual regular cell decompositions). Denote by $\xi \in C_n(M, \partial M; \mathbb{Z})$ a cellular chain representing the fundamental class.

Fix a base point in M and some homomorphism $\pi_1 M \to \Gamma$ to some group Γ, and denote by $M_\Gamma \to M$ the corresponding Γ cover, equipped with the lifted cell structure and its cellular left Γ covering action. The cellular $\mathbb{Z}$ chain complexes of M_Γ and $(M_\Gamma, \partial M_\Gamma)$ are denoted by $C_*^\Gamma(M)$ and $C_*^\Gamma(M, \partial M)$. These are free finitely

generated left $\mathbb{Z}\Gamma$ modules: a choice of $\mathbb{Z}\Gamma$ basis is given by arbitrarily choosing lifts of cells of M (respectively of cells which intersect the interior of M). A choice of cellular approximation of the diagonal determines a chain level cap product:

$$\cap\xi : \mathrm{Hom}_{\mathbb{Z}\Gamma}(C_*^\Gamma(M, \partial M), \mathbb{Z}\Gamma) \to C_*^\Gamma(M). \tag{2-2}$$

The abelian group $\mathrm{Hom}_{\mathbb{Z}\Gamma}(C_*^\Gamma(M, \partial M, \mathbb{Z}\Gamma)$ admits a left $\mathbb{Z}\Gamma$ structure by $(\gamma\cdot x)(\sigma) := x(\sigma)\cdot\gamma^{-1}$. The Poincaré duality theorem asserts that the map of (2-2) is a chain homotopy equivalence of free left $\mathbb{Z}\Gamma$ chain complexes.

Given any left $\mathbb{Z}\Gamma$ module W, we have that

$$\mathrm{Hom}_{\mathbb{Z}\Gamma}(C_*^\Gamma(M, \partial M), W) = \mathrm{Hom}_{\mathbb{Z}\Gamma}(C_*^\Gamma(M, \partial M), \mathbb{Z}\Gamma) \otimes_{\mathbb{Z}\Gamma} W$$

and hence capping with ξ induces a chain homotopy equivalence

$$\cap\xi : \mathrm{Hom}_{\mathbb{Z}\Gamma}(C_*^\Gamma(M, \partial M), W) \to C_*^\Gamma(M) \otimes_{\mathbb{Z}\Gamma} W.$$

Now suppose that W is a finite-dimensional $\mathbb{R}$ vector space equipped with a positive definite inner product $\{\ ,\ \} : W \times W \to \mathbb{R}$ and $\Gamma \to O(W)$ a representation, determining the left $\mathbb{Z}\Gamma$ module structure on W. There is an algebraic chain isomorphism of $\mathbb{R}$ chain complexes:

$$C_*^\Gamma(M) \otimes_{\mathbb{Z}\Gamma} W \cong \mathrm{Hom}_{\mathbb{R}}\big(\mathrm{Hom}_{\mathbb{Z}\Gamma}(C_*^\Gamma(M), W), \mathbb{R}\big),\ c \otimes w \mapsto \big(h \mapsto \{h(c), w\}\big).$$

Composing with the chain homotopy equivalence $\cap\xi$, and using the universal coefficient theorem for $\mathbb{R}$, we obtain an isomorphism

$$H^*(M, \partial M; W) \cong H^*(\mathrm{Hom}_{\mathbb{R}}(\mathrm{Hom}_{\mathbb{Z}\Gamma}(C_*^\Gamma(M), W), \mathbb{R})) = \mathrm{Hom}_{\mathbb{R}}(H^*(M; W), \mathbb{R})$$

whose adjoint is the (by construction nondegenerate) *cohomology intersection pairing over W*

$$H^*(M, \partial M; W) \times H^*(M; W) \to \mathbb{R}. \tag{2-3}$$

The cohomology intersection pairing can also be expressed in terms of cup products:

$$(x, y) \mapsto \{x \cup y\} \cap \xi, \tag{2-4}$$

where $\{\ \} : H^n(M, \partial M; W \otimes_{\mathbb{Z}\Gamma} W) \to H^n(M, \partial M; \mathbb{R})$ is induced by the coefficient homomorphism $W \otimes_{\mathbb{Z}\Gamma} W \to \mathbb{R}$ determined by the bilinear form $\{\ ,\ \}$.

When the boundary of M is empty, this pairing, which we denote by

$$\omega_M : H^*(M; W) \times H^*(M; W) \to \mathbb{R}, \tag{2-5}$$

is therefore a nondegenerate inner product on the $\mathbb{R}$ vector space $H^*(M; W)$.

If ∂M is nonempty, precomposing the injective adjoint

$$H^*(M; W) \to \mathrm{Hom}_{\mathbb{R}}(H^*(M, \partial M; W); \mathbb{R})$$

with the restriction map $H^*(M, \partial M; W) \to H^*(M; W)$ yields a map

$$H^*(M, \partial M; W) \to \mathrm{Hom}_{\mathbb{R}}(H^*(M, \partial M; W); \mathbb{R})$$

with kernel equal to the kernel of $H^*(M, \partial M; W) \to H^*(M; W)$. An equivalent statement is that the pairing

$$\omega_{(M,\partial M)} : H^*(M, \partial M; W) \times H^*(M, \partial M; W) \to \mathbb{R} \tag{2-6}$$

has radical equal to $\ker H^*(M, \partial M; W) \to H^*(M; W)$.

Taking gradings into account, when $\dim M = 2n$ restriction defines a pairing

$$\omega_{(M,\partial M)} : H^n(M, \partial M; W) \times H^n(M, \partial M; W) \to \mathbb{R} \tag{2-7}$$

with radical equal to $\ker H^n(M, \partial M; W) \to H^n(M; W)$.

When $\dim M = 4\ell + 2$, for example when M is a surface, the pairings ω_M and $\omega_{(M,\partial M)}$ are skew-symmetric:

$$\omega_M(x, y) = -\omega_M(y, x) \quad \text{and} \quad \omega_{(M,\partial M)}(x, y) = -\omega_{(M,\partial M)}(y, x).$$

3. Two- and three-dimensional manifolds and symplectic linear algebra

3A. ***A symplectic form on the first cohomology of surface.*** Recall that S_+ denotes the path connected CW complex obtained by adding a disjoint base point to the oriented surface S and a 1-cell connecting each path component of S to this base point. Let $\Gamma = \pi_1(S_+)$. Its universal cover $\widetilde{S}_+ \to S_+$ is a regular Γ cover, and hence so is its restriction over S,

$$\widetilde{S} \to S.$$

A representation $\rho : \Gamma \to G$ is fixed.

Then ρ determines, via the adjoint action of G, the representation $\mathrm{Ad}\,\rho : \Gamma \to O(\mathfrak{g})$, and hence cohomology groups $H^*(S; \mathfrak{g})$, $H^*(S, \partial S; \mathfrak{g})$, and $H^*(\partial S; \mathfrak{g})$, with, for example,

$$H^*(S; \mathfrak{g}) := H^*(\mathrm{Hom}_{\mathbb{Z}\Gamma}(C_*^{\Gamma}(S), \mathfrak{g})).$$

If we wish to emphasize ρ, we write $H^*(S; \mathfrak{g}_{\mathrm{Ad}\,\rho})$.

Equation (2-5) shows that, since the boundary of S is empty,

$$\omega_S : H^1(S; \mathfrak{g}) \times H^1(S; \mathfrak{g}) \to \mathbb{R} \tag{3-1}$$

is a nondegenerate skew-symmetric form.

On the other hand, if C is nonempty, so that the boundary of F is nonempty, equation (2-7) shows that

$$\omega_{(F,\partial F)} : H^1(F, \partial F; \mathfrak{g}) \times H^1(F, \partial F; \mathfrak{g}) \to \mathbb{R}$$

is in general a degenerate skew symmetric form with radical equal to

$$\ker H^1(F, \partial F; \mathfrak{g}) \to H^1(F; \mathfrak{g}).$$

A degenerate skew symmetric form induces a nondegenerate form on the quotient by its radical. The exact sequence of the pair $(F, \partial F)$ shows that $\omega_{(F,\partial F)}$ descends to a *nondegenerate* skew symmetric form

$$\widehat{\omega}_F : \widehat{H}^1(F; \mathfrak{g}) \times \widehat{H}^1(F; \mathfrak{g}) \to \mathbb{R}, \tag{3-2}$$

where

$$\widehat{H}^1(F; \mathfrak{g}) = \text{Image } H^1(F, \partial F; \mathfrak{g}) \to H^1(F; \mathfrak{g}) = \ker H^1(F; \mathfrak{g}) \to H^1(\partial F; \mathfrak{g}).$$

Summarizing:

Proposition 3.1. *Let S be a compact oriented surface without boundary and $\rho : \pi_1 S_+ \to G$ a representation. Then $(H^1(S; \mathfrak{g}_{\mathrm{Ad}\,\rho}), \omega_S)$ is a symplectic vector space. If $F \subset S$ is the complement of a nonempty disjoint union of annuli V, then $(\widehat{H}^1(F; \mathfrak{g}_{\mathrm{Ad}\,\rho}), \widehat{\omega}_F)$ is a symplectic vector space.*

In the following diagram of inclusions, the two vertical and two horizontal rows are exact, and the isomorphisms are excisions ($\mathfrak{g}_{\mathrm{Ad}\,\rho}$-coefficients omitted):

$$\begin{array}{ccccccccc}
 & & & & H^1(S, F) & \xrightarrow{\cong} & H^1(V, \partial V) & & \\
 & & & & \downarrow{\scriptstyle q_1} & & \downarrow{\scriptstyle q_2} & & \\
\longrightarrow & H^1(S, V) & \xrightarrow{\alpha} & H^1(S) & \xrightarrow{\beta} & H^1(V) & \xrightarrow{\gamma} & H^2(S, V) & \\
 & \cong\downarrow{\scriptstyle p_0} & & \downarrow{\scriptstyle p_1} & & \downarrow{\scriptstyle p_2} & & \downarrow\cong & \\
\longrightarrow & H^1(F, \partial F) & \xrightarrow[a]{} & H^1(F) & \xrightarrow[b]{} & H^1(\partial F) & \xrightarrow[c]{} & H^2(F, \partial F) &
\end{array} \tag{3-3}$$

Proposition 3.2. *The kernel of β is a coisotropic subspace of $H^1(S; \mathfrak{g})$ with annihilator* $\ker p_1$, *and hence p_1 induces a symplectomorphism*

$$\ker \beta / \ker p_1 \xrightarrow{\cong} \widehat{H}^1(F; \mathfrak{g}).$$

Proof. The surface V is a disjoint union of annuli. The composition $\{0\} \times S^1 \subset \partial(I \times S^1) \subset I \times S^1$ a homotopy equivalence, and hence the restriction $H^1(I \times S^1) \to H^1(\partial(I \times S^1))$ is injective with any coefficients. Hence p_2 is injective, and $q_2 = 0$.

If $s \in H^1(S)$ satisfies $\omega_S(s, \alpha(y)) = 0$ for all $y \in H^1(S, V)$, we then have that $\omega_{F,\partial F}(p_1(s), a \circ p_0(y)) = 0$ for all $y \in H^1(S, V)$. Hence $\omega_{F,\partial F}(p_1(s), a(z)) = 0$ for all $z \in H^1(F, \partial F)$. Since the pairing $H^1(F, \partial F) \times H^1(F) \to \mathbb{R}$ is nondegenerate (equation (2-3)), this implies that $p_1(s) = 0$. In other words, the annihilator of $\operatorname{im} \alpha = \ker \beta$ is contained in $\ker p_1$.

Since $q_2 = 0$, $\ker p_1 \subset \ker \beta$. Therefore $\ker \beta$ contains its annihilator and hence is coisotropic.

It remains to show that $\ker p_1 = \text{image}\, q_1$ is contained in the annihilator of $\ker \beta$. Given $x \in H^1(S, V)$ and $y \in H^1(S, F)$, we have that $\omega_S(\alpha(x), q_1(y)) = 0$ since the cup product

$$H^1(S, V) \times H^1(S, F) \to H^2(S)$$

factors through $H^2(S, F \cup V) = 0$ (see (2-4)). □

Corollary 3.3. *If $L \subset H^1(S; \mathfrak{g}_{\mathrm{Ad}\,\rho})$ is any Lagrangian subspace, then the image*

$$p_1(L \cap \ker \beta) \subset \widehat{H}^1(F; \mathfrak{g}_{\mathrm{Ad}\,\rho})$$

is a Lagrangian subspace.

3B. ***Restriction from a 3-manifold with boundary.*** Recall that Y is a compact, connected, oriented 3-manifold with boundary $S = \partial Y$, extended to an embedding $S_+ \subset Y$.

Assume that the representation $\rho : \pi_1(S_+) \to G$ is a restriction of a representation (of the same name) $\rho : \pi_1(Y) \to G$.

Lemma 3.4. *The image of the restriction map,*

$$L_Y := \text{Image}\, r : H^1(Y; \mathfrak{g}_{\mathrm{Ad}\,\rho}) \to H^1(S; \mathfrak{g}_{\mathrm{Ad}\,\rho}),$$

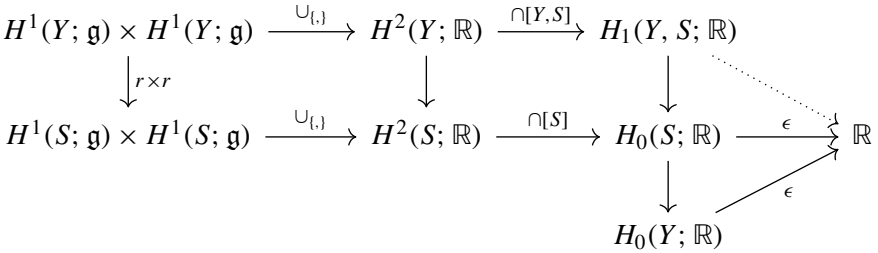
is a Lagrangian subspace of $(H^1(S; \mathfrak{g}_{\mathrm{Ad}\,\rho}), \omega_S)$.

Proof. In the following diagram, the middle row is part of the exact sequence of the pair (Y, S). The vertical arrows all isomorphisms, with the downward pointing isomorphisms Poincaré-Lefschetz duality. The diagram commutes up to sign [26]. We suppress the $\mathfrak{g}_{\mathrm{Ad}\,\rho}$ coefficients.

$$\begin{array}{ccccc}
\mathrm{Hom}(H_1(Y); \mathbb{R}) & \longrightarrow & \mathrm{Hom}(H_1(S); \mathbb{R}) & \xrightarrow{\delta^*} & \mathrm{Hom}(H_2(Y, S); \mathbb{R}) \\
\uparrow & & \uparrow & & \uparrow \\
H^1(Y) & \xrightarrow{r} & H^1(S) & \xrightarrow{\nu} & H^2(Y, S) \\
\downarrow & & \downarrow & & \downarrow \\
H_2(Y, S) & \xrightarrow{\delta} & H_1(S) & \longrightarrow & H_1(Y)
\end{array}$$

The image of r equals the kernel of ν, which is isomorphic to the kernel of δ^*. The kernel of δ^* is isomorphic to the cokernel of its dual δ, which in turn is isomorphic to the cokernel of r. Hence the image and cokernel of r are isomorphic, and so $\dim(\text{image}(r)) = \frac{1}{2} \dim H^1(S)$.

Commutativity of the following diagram is a consequence of naturality of cup product and Poincaré duality:



Exactness of the vertical sequence shows that the dotted arrow is zero, which implies that the image of r is isotropic, and therefore Lagrangian. □

Recall that the boundary $S = \partial Y$ is equipped with a embedded collection $C \subset S$ of circles, with tubular neighborhood V and complementary subsurface F, producing the decomposition $S = F \cup V$ as in (2-1). Consider the ladder of exact sequences, with all maps induced by inclusions. The $\mathfrak{g}_{\mathrm{Ad}\,\rho}$ coefficients are suppressed. The bottom two rows coincide with those of (3-3).

$$\begin{array}{ccccccccc}
\cdots & \longrightarrow & H^1(Y,V) & \xrightarrow{A} & H^1(Y) & \xrightarrow{B} & H^1(V) & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle r_0} & & \downarrow{\scriptstyle r} & & \| & & \\
\cdots & \longrightarrow & H^1(S,V) & \xrightarrow{\alpha} & H^1(S) & \xrightarrow{\beta} & H^1(V) & \longrightarrow & \cdots \\
 & & \cong\downarrow{\scriptstyle p_0} & & \downarrow{\scriptstyle p_1} & & \downarrow{\scriptstyle p_2} & & \\
\cdots & \longrightarrow & H^1(F,\partial F) & \xrightarrow{a} & H^1(F) & \xrightarrow{b} & H^1(\partial F) & \longrightarrow & \cdots
\end{array} \tag{3-4}$$

A diagram chase shows that $\mathrm{Image}\, r \cap \ker\beta = \mathrm{Image}\, \alpha \circ r_0$. Hence

$$p_1(\mathrm{Image}\, r \cap \ker\beta) = \mathrm{Image}\, a \circ p_0 \circ r_0 : H^1(Y,V) \to H^1(F) \subset \ker b = \widehat{H}^1(F)$$

Lemma 3.4 and Corollary 3.3 imply the following.

Corollary 3.5. *Suppose that $\partial Y = S = V \cup F$ with V a disjoint union of annuli. Then*

$$L_{Y,V} := \mathrm{Image}\, H^1(Y,V;\mathfrak{g}_{\mathrm{Ad}\,\rho}) \to H^1(F;\mathfrak{g}_{\mathrm{Ad}\,\rho}) \subset \widehat{H}^1(F;\mathfrak{g}_{\mathrm{Ad}\,\rho})$$

is a Lagrangian subspace. Moreover, $L_{Y,V}$ is the symplectic reduction of L_Y with respect to $\ker\beta : H^1(S;\mathfrak{g}_{\mathrm{Ad}\,\rho}) \to H^1(V;\mathfrak{g}_{\mathrm{Ad}\,\rho})$.

4. Character varieties, relative character varieties, and their tangent spaces

4A. *Character varieties.*

Definition 4.1. Given a finitely presented group Γ, its *G character variety* $\chi_G(\Gamma)$ is the real semialgebraic set defined to be the orbit space of the G-conjugation

action on the affine $\mathbb{R}$-algebraic set $\mathrm{Hom}(\Gamma, G)$. A set of g generators of Γ embeds $\mathrm{Hom}(\Gamma, G)$ in G^g equivariantly, and, since G is a compact Lie group, $\mathrm{Hom}(\Gamma, G)$ is an affine $\mathbb{R}$-algebraic set, with orbit space $\chi_G(\Gamma)$ a semialgebraic set. We call $\chi_G(\Gamma)$ the *G-character variety of* Γ. If Z is a path connected space, write $\chi_G(Z)$ instead of $\chi_G(\pi_1(Z))$. The character variety of a nonpath connected space, *by definition*, is the cartesian product of the character varieties of its path components.

As observed by Weil [31], the *formal tangent space at* the conjugacy class of *any* representation $\rho : \Gamma \to G$ to the character variety $\chi_G(\Gamma)$ is naturally identified with first cohomology:

$$T_\rho \chi_G(\Gamma) = H^1(\Gamma; \mathfrak{g}_{\mathrm{Ad}\,\rho}). \tag{4-1}$$

We take (4-1) as the definition of the formal tangent space at ρ for any $\rho \in \chi_G(\Gamma)$. Recall that $H^1(\Gamma; \mathfrak{g}_{\mathrm{Ad}\,\rho})$ and $H^1(X; \mathfrak{g}_{\mathrm{Ad}\,\rho})$ are canonically isomorphic for any space X with fundamental group Γ.

Weil's argument is based on the calculation that if a path of representations is expressed in the form $\rho_s = \exp(\alpha_s)\rho_0$ for some path $\alpha_s : \Gamma \to \mathfrak{g}$, then $\frac{d}{ds}|_{s=0}\, \alpha_s : \Gamma \to \mathfrak{g}$ is a 1-cocycle [6].

4B. ***Relative character varieties.*** As above, assume that Y is a compact connected 3-manifold with boundary $S = \partial Y$, $C \subset S$ is a union of m circles C_i, V is the tubular neighborhood of C in S, with complementary surface F.

Assume further that either J is invariant via the inversion map of G (as is the case for the conjugacy class of traceless matrices in $\mathrm{SU}(2)$) or else assume that every circle C_i is equipped with an orientation.

Then define the *relative character variety*

$$\chi_{G,J}(Y, C) \subset \chi_G(Y) \tag{4-2}$$

to be the subvariety consisting of conjugacy classes of representations $\pi_1(Y) \to G$ which send (any based representative of the homotopy class of) each circle in C into J. Define the *formal tangent space* of $\chi_{G,J}(Y, C)$ at ρ to be

$$T_\rho \chi_{G,J}(Y, C) = \ker H^1(Y; \mathfrak{g}_{\mathrm{Ad}\,\rho}) \to H^1(C; \mathfrak{g}_{\mathrm{Ad}\,\rho}). \tag{4-3}$$

Given an oriented surface F, define

$$\chi_{G,J}(F, \partial F) = \chi_{G,J}\big(F \times [0, 1], \partial F \times \{\tfrac{1}{2}\}\big).$$

Its formal tangent space is

$$T_\rho \chi_{G,J}(F, \partial F) = \ker H^1(F; \mathfrak{g}_{\mathrm{Ad}\,\rho}) \to H^1(\partial F; \mathfrak{g}_{\mathrm{Ad}\,\rho}) = \widehat{H}^1(F; \mathfrak{g}_{\mathrm{Ad}\,\rho}). \tag{4-4}$$

4C. ***Regular points.*** The term "formal tangent space" may be replaced by its usual elementary differential topology notion in neighborhoods of *regular points* of $\chi_{G,J}(Y, C)$ and $\chi_{G,J}(F, \partial F)$, as we next explain. In brief, as elsewhere in gauge theory, a regular point is one which has a neighborhood diffeomorphic to Euclidean space of the correct index-theoretic dimension. We provide a stripped-down explanation, suitable for our purposes, of what this means, for the benefit of the reader.

First, given a connected compact surface F (with possibly empty boundary) we call $\rho \in \chi_{G,J}(F, \partial F)$ a regular point provided ρ has a neighborhood $U \subset \chi_{G,J}(F, \partial F)$ so that $\dim \widehat{H}^1(F; \mathfrak{g}_{\mathrm{Ad}\,\rho'})$ is independent of $\rho' \in U$.

Next, given a disjoint union of connected compact surfaces $F = \bigsqcup_i F_i$, call

$$\rho \in \chi_{G,J}(F, \partial F) = \prod_i \chi_{G,J}(F_i, \partial F_i)$$

a regular point provided each of its components is a regular point.

Finally, for a pair (Y, C) (with C possibly empty) $\rho \in \chi_{G,J}(Y, C)$ is called a regular point provided ρ admits a neighborhood $U \subset \chi_{G,J}(Y, C)$ so that for all $\rho' \in U$:

- The restriction map $\chi_{G,J}(Y, C) \to \chi_{G,J}(F, \partial F)$ takes ρ' to a regular point, and
- $\dim T_{\rho'} \chi_{G,J}(Y, C) = \frac{1}{2} \dim T_\rho \chi_{G,J}(F, \partial F) = \frac{1}{2} \sum_i \dim T_\rho \chi_{G,J}(F_i, \partial F_i)$.

Hence, if $\rho \in \chi_{G,J}(Y, C)$ is a regular point, the map $\chi_{G,J}(Y, C) \to \chi_{G,J}(F, \partial F)$, which takes a representation of a 3-manifold group to its restriction to the boundary surface, is, near ρ, a smooth map of a smooth n-disk into $\mathbb{R}^{2n}$ for some n.

Notice that $\chi_{G,J}(Y, C)$ is the preimage of the point $(J, \dots, J)$ under the restriction map

$$\chi_G(Y) \to \chi_G(C) = \prod_{i=1}^{m} \chi_G(C_i) = (G/_{\text{conjugation}})^m,$$

and that $\chi_{G,J}(F, \partial F)$ is the preimage of the point $(J, \dots, J)$ under the restriction map

$$\chi_G(F) \to \chi_G(\partial F).$$

Since $C \subset V$ is a deformation retract, the exact sequence of the pair shows that

$$T_\rho \chi_{G,J}(Y, C) \cong \text{Image } H^1(Y, V; \mathfrak{g}_{\mathrm{Ad}\,\rho}) \to H^1(Y; \mathfrak{g}_{\mathrm{Ad}\,\rho}).$$

The image of the differential of the restriction map $\chi_{G,J}(Y, C) \to \chi_{G,J}(F, \partial F)$ at $\rho \in \chi_{G,J}(Y, C)$ is therefore identified with $L_{Y,V}$, the image of the composition

$$H^1(Y, V; \mathfrak{g}_{\mathrm{Ad}\,\rho}) \to H^1(F, \partial F; \mathfrak{g}_{\mathrm{Ad}\,\rho}) \to \widehat{H}^1(F; \mathfrak{g}_{\mathrm{Ad}\,\rho}),$$

which by Corollary 3.5 is a Lagrangian subspace of $(H^1(F; \mathfrak{g}_{\mathrm{Ad}\,\rho}), \widehat{\omega}_F)$. In sum:

Corollary 4.2. *If $\rho \in \chi_{G,J}(Y, C)$ is a regular point, then there exists a neighborhood of ρ so that the differential of the restriction $\chi_{G,J}(Y, C) \to \chi_{G,J}(F, \partial F)$ at any point ρ' in this neighborhood has image a Lagrangian subspace of $(\widehat{H}^1(F; \mathfrak{g}_{\mathrm{Ad}\,\rho'}), \widehat{\omega}_F)$.*

It is known that for surfaces, with the exception of a few low genus cases, the regular points coincide with the irreducible representations.

4D. ***Symplectic structure.*** The proof of Corollary 4.2 does not rely of the following fundamental result of Atiyah and Bott [3] and its extensions due to Goldman [12], Karshon [19], Biswas and Guruprasad [4], King and Sengupta [20], Guruprasad, Huebschmann, Jeffrey and Weinstein [14].

Theorem 4.3. *On the top stratum of regular points of $\chi_G(S)$ and $\chi_{G,J}(F, \partial F)$, the 2-forms ω_S and $\widehat{\omega}_F$ are closed, that is, are symplectic forms.*

In light of this result, Corollary 4.2 can be restated as follows.

Theorem 4.4. *Suppose that $\rho \in \chi_{G,J}(Y, C)$ is a regular point. Then there exists a neighborhood U of ρ in $\chi_{G,J}(Y, C)$ so that the restriction of r to U,*

$$r|_U : U \to \chi_{G,J}(F, \partial F),$$

is a Lagrangian embedding.

In particular, if $\chi_{G,J}(Y, C)$ contains only regular points, then the restriction map is a Lagrangian immersion.

Given a tangle $(X, \mathcal{L})$, we write $\chi_{G,J}(X, \mathcal{L})$ rather than $\chi_{G,J}(Y, C)$, where Y, S, C, F and V are determined by $(X, \mathcal{L})$ as in Section 2A4. Also write $\chi_{G,J}(\partial(X, \mathcal{L}))$ rather than $\chi_{G,J}(F, \partial F)$. Theorem 4.4 can be restated in the new notation as follows.

Corollary 4.5. *Suppose that $\rho \in \chi_{G,J}(X, \mathcal{L})$ is a regular point. Then there exists a neighborhood U of ρ in $\chi_{G,J}(X, \mathcal{L})$ so that the restriction of r to U,*

$$r|_U : U \to \chi_{G,J}(\partial(X, \mathcal{L})),$$

is a Lagrangian embedding.

In particular, if $\chi_{G,J}(X, \mathcal{L})$ contains only regular points, then the restriction map is a Lagrangian immersion.

In what follows, we simply write $\chi(A)$ for $\chi_G(A)$ and $\chi(A, B)$ for $\chi_{G,J}(A, B)$. In addition, we denote by $\chi(S)^* \subset \chi(S)$ and $\chi(F, \partial F)^* \subset \chi(F, \partial F)$ the *smooth top strata*, as real algebraic varieties, equipped with the symplectic forms ω_S, $\widehat{\omega}_F$.

5. Three pillowcases

In this section, take $G = \mathrm{SU}(2)$ viewed as the unit quaternions, with Lie algebra $\mathrm{su}(2)$ the purely imaginary quaternions. Let $J \subset \mathrm{SU}(2)$ denote the conjugacy class of unit quaternions with zero real part, so $J = \mathrm{su}(2) \cap \mathrm{SU}(2)$, a 2-sphere.

5A. ***The quotient of the torus by the elliptic involution.*** Let $\mathbb{T}$ denote $U(1)\times U(1)$ with its symplectic form $dx\wedge dy$ and symplectic $\mathrm{SL}(2,\mathbb{Z})$ action given in (1-2). Multiplication by $-1\in \mathrm{SL}(2,\mathbb{Z})$ is central and hence the quotient

$$\mathbb{P}=\mathbb{T}/\{\pm 1\}$$

inherits a $\mathrm{PSL}(2,\mathbb{Z})=\mathrm{SL}(2,\mathbb{Z})/\{\pm 1\}$ action. The quotient map $\mathbb{T}\to\mathbb{P}$ is the 2-fold branched cover of the 2-sphere with branch points the four points $\{(\pm 1,\pm 1)\}$. The complement, $\mathbb{P}^*$ of the four branch points is a smooth surface, with symplectic form $dx\wedge dy$ and symplectic $\mathrm{PSL}(2,\mathbb{Z})$ action, and the restriction

$$\mathbb{T}^*\to\mathbb{P}^*$$

is a smooth symplectic 2-fold covering map.

5B. ***The* SU(2) *character variety of the genus 1 surface.*** Consider the genus one closed oriented surface T, equipped with generators $\mu,\lambda\in\pi_1 T$ represented by a pair of oriented loops intersecting geometrically and algebraically once. The Dehn twists D_μ, D_λ about μ and λ induce the automorphisms

$$D_\mu:\big(\mu\mapsto\mu,\lambda\mapsto\mu\lambda\big)\quad\text{and}\quad D_\lambda:\big(\mu\mapsto\mu\lambda^{-1},\lambda\mapsto\lambda\big).$$

of $\pi_1(T,t_0)\cong\mathbb{Z}\mu\oplus\mathbb{Z}\lambda$ (where t_0 lies outside the support of these 2 Dehn twists). These automorphisms induce, by precomposition, homeomorphisms

$$D_\mu^*,D_\lambda^*:\chi(T)\to\chi(T).$$

Let $\mathcal{H}(T)=\mathrm{Hom}(\pi_1(T),\mathrm{SU}(2))$. Denote by

$$p:\mathcal{H}(T)\to\chi(T)$$

the (surjective) orbit map of the conjugation action.

Define

$$\rho:\mathbb{T}\to\mathcal{H}(T),\ \ \rho(e^{x\boldsymbol{i}},e^{y\boldsymbol{i}})=\big(\mu\mapsto e^{x\boldsymbol{i}},\ \lambda\mapsto e^{y\boldsymbol{i}}\big).\tag{5-1}$$

Proposition 5.1. *The Dehn twists D_μ and D_λ define a symplectic* $\mathrm{PSL}(2,\mathbb{Z})$ *action on $(\chi(T),\omega_T)$. The map ρ of (5-1) descends to a* $\mathrm{PSL}(2,\mathbb{Z})$ *equivariant homeomorphism*

$$\rho:\mathbb{P}\to\chi(T)$$

which restricts to a $\mathrm{PSL}(2,\mathbb{Z})$ *equivariant symplectomorphism*

$$(\mathbb{P}^*,c\,dx\wedge dy)\to(\chi(T)^*,\omega_T)$$

on the top strata of the SU(2) *character varieties.*

Proof. It is elementary to show that $\rho : \mathbb{P} \to \chi(T)$ is a well defined homeomorphism, as well as an analytic diffeomorphism of the smooth strata $\mathbb{P}^* \to \chi(T)^*$.

We show that $\rho : (\mathbb{P}^*, dx \wedge dy) \to (\chi(T)^*, \omega_T)$ is a symplectomorphism. Since this is a local statement, we work in $\mathbb{R}^2$ for simplicity. Define $m = \rho \circ e : \mathbb{R}^2 \to \chi(T)$, where $e(x, y) = (e^{x\boldsymbol{i}}, e^{y\boldsymbol{i}})$.

The differential of m at $(x, y) \in \mathbb{R}^2$, $dm : T_{(x,y)}\mathbb{R}^2 \to T_{m(x,y)}(\mathrm{SU}(2) \times \mathrm{SU}(2))$ is given by

$$dm\big(\tfrac{\partial}{\partial x}\big) = \big(\mu \mapsto e^{x\boldsymbol{i}}\boldsymbol{i}, \lambda \mapsto 0\big),\ dm\big(\tfrac{\partial}{\partial y}\big) = \big(\mu \mapsto 0, \lambda \mapsto e^{y\boldsymbol{i}}\boldsymbol{i}\big).$$

Following Weil [31], left translation in $\mathrm{SU}(2) \times \mathrm{SU}(2)$ identifies these with the su(2)-valued 1-cochains

$$z_x = (\mu \mapsto \boldsymbol{i}, \lambda \mapsto 0),\ z_y = (\mu \mapsto 0, \lambda \mapsto \boldsymbol{i}) \in C^1(T; \mathrm{su}(2)_{\mathrm{Ad}\, m(x,y)}).$$

The subspaces $L = \boldsymbol{i}\mathbb{R}$ and $V = \boldsymbol{j}\mathbb{R} + \boldsymbol{k}\mathbb{R}$ are invariant and complementary with respect to $\mathrm{Ad}\, m(x, y)$, and therefore

$$H^1(T; \mathrm{su}(2)_{\mathrm{Ad}\, m(x,y)}) = H^1(T; L_{\mathrm{Ad}\, m(x,y)}) \oplus H^1(T; V_{\mathrm{Ad}\, m(x,y)}).$$

Note that the action $\mathrm{Ad}\, m(x, y)$ on L is trivial, since $L = \boldsymbol{i}\mathbb{R}$ and $m(x, y)$ has values in the abelian subgroup $\{e^{\boldsymbol{i}u}\}$. Hence

$$H^1(T; L_{\mathrm{Ad}\, m(x,y)}) \cong H^1(T; \mathbb{R}) \cong \mathbb{R}^2.$$

The branched cover $m : \mathbb{R}^2 \to \chi(T)$ is a local diffeomorphism near any

$$(x, y) \in (\mathbb{R}^2)^* = \mathbb{R}^2 \setminus (\pi\mathbb{Z})^2,$$

and hence it follows that for such (x, y),

$$H^1(T; \mathbb{R}) \otimes \mathbb{R}\boldsymbol{i} = H^1(T; L_{\mathrm{Ad}\, m(x,y)}) = \mathrm{Span}\{z_x, z_y\} \quad \text{and} \quad H^1(T; V_{\mathrm{Ad}\, m(x,y)}) = 0$$

(these calculations can also be easily checked directly), so that

$$T_{m(x,y)}(\chi(T)) = H^1(T; \mathbb{R}) \otimes \mathbb{R}\boldsymbol{i}.$$

The cup product

$$H^1(T; L_{\mathrm{Ad}\, m(x,y)}) \times H^1(T; L_{\mathrm{Ad}\, m(x,y)}) \to H^2(T; \mathbb{R})$$

is thereby identified with

$$H^1(T; \mathbb{R}) \otimes \mathbb{R}\boldsymbol{i} \times H^1(T; \mathbb{R}) \otimes \mathbb{R}\boldsymbol{i} \to H^2(T; \mathbb{R}),$$
$$(z_1 \otimes \boldsymbol{i}, z_2 \otimes \boldsymbol{i}) = (z_1 \cup z_2) \cdot (-\mathrm{Re}(\boldsymbol{ii})) = z_1 \cup z_2.$$

In particular, $z_x = \mu^* \otimes \boldsymbol{i}$ and $z_y = \lambda^* \otimes \boldsymbol{i}$, where

$$\mu^*, \lambda^* \in H^1(T; \mathbb{Z}) = \mathrm{Hom}(H_1(T; \mathbb{Z}), \mathbb{Z})$$

is the basis dual to μ, λ. This basis is symplectic with respect to the (usual, untwisted) intersection form. Hence

$$\begin{aligned}\big((p\circ m)^*(\omega_T)\big)|_{(x,y)}\big(\tfrac{\partial}{\partial x}, \tfrac{\partial}{\partial y}\big) &= \omega_T(z_x, z_y) = (\mu^* \cup \lambda^*) \cap [T] = 1\\ &= dx \wedge dy\big(\tfrac{\partial}{\partial x}, \tfrac{\partial}{\partial y}\big) = 1.\end{aligned}$$

Naturality of cup products shows that D_μ^*, D_λ^* preserve the symplectic form ω_T. Next,

$$\begin{aligned}D_\mu^*(m(x, y)) &= m(x, y)\circ D_\mu\\ &= \big(\mu \mapsto m(x, y)(\mu) = e^{x\boldsymbol{i}}, \lambda \mapsto m(x, y)(\mu\lambda) = e^{(x+y)\boldsymbol{i}}\big)\end{aligned}$$

and similarly

$$D_\lambda^*(m(x, y)) = m(x, y)\circ D_\lambda = \big(\mu \mapsto e^{(x-y)\boldsymbol{i}}, \lambda \mapsto e^{y\boldsymbol{i}}\big).$$

It follows that the subgroup of $\mathrm{Homeo}(\chi(T))$ generated by D_μ^* and D_λ^* pulls back, via the homeomorphism $\rho : \mathbb{P} \to \chi(T)$, to the subgroup of $\mathrm{PSL}(2, \mathbb{Z})$ generated by

$$\rho^*(D_\mu^*) = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \quad \rho^*(D_\lambda^*) = \begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix}.$$

These two matrices generate $\mathrm{PSL}(2, \mathbb{Z})$ (see [25]), finishing the proof. □

5B1. *The solid torus and the restriction to its boundary.* Let X denote the solid torus with boundary T. Equip T with based loops μ, λ generating $\pi_1(T)$, so that μ is trivial in $\pi_1(X)$ and λ generates $\pi_1(X)$. Then

$$\chi(X) = \chi(\mathbb{Z}\lambda) = \mathrm{SU}(2)/_{\text{conjugation}}.$$

An explicit slice of the conjugation action $\mathrm{Hom}(\mathbb{Z}\lambda, \mathrm{SU}(2)) \to \chi(\mathbb{Z}\lambda)$ is given by the map

$$[0, \pi] \to \mathrm{Hom}(\mathbb{Z}\lambda, \mathrm{SU}(2)), \quad s \mapsto (\lambda \mapsto e^{\boldsymbol{i}s}) \tag{5-2}$$

with composition $[0, \pi] \to \chi(X) \xrightarrow{\mathrm{Re}_\lambda} [-1, 1]$ equal to the analytic isomorphism $\cos(s)$. Simple cohomology calculations show $\dim H^1(\mathbb{Z}; \mathrm{su}(2))$ equals 1 when $0 < s < \pi$ and equals 3 when $s = 0$ or π. This shows that the interior of the interval forms the smooth top stratum of $\chi(\mathbb{Z})$, and the endpoints are singular.

Since $\mu = 1 \in \pi_1(X)$, the restriction-to-the boundary map

$$\chi(X) \to \chi(T) \tag{5-3}$$

is easily computed, in $\mathbb{P}$, to be the smooth (necessarily Lagrangian) embedded arc given by

$$[0, \pi] \ni s \mapsto [e^{s\boldsymbol{i}}, 1] \in \mathbb{P} \tag{5-4}$$

with endpoints at $[-1, 1]$ and $[1, 1]$.

5C. ***The traceless* SU(2) *character variety of the 4-punctured 2-sphere.*** Let $4D^2 \subset S^2$ be four disjoint open disks, and set

$$F = S^2 \setminus 4D^2, \quad \partial F = 4S^1.$$

Then $\pi_1(F)$ has presentation

$$\pi_1(F) = \langle a, b, c, d \mid abcd = 1 \rangle$$

and is free on a, b, c. Set

$$\mathcal{H}(S^2, 4) = \big\{\rho \in \mathrm{Hom}(\pi_1(F), \mathrm{SU}(2)) \mid \rho(a), \rho(b), \rho(c), \rho(d) \in J\big\}.$$

Let $p : \mathcal{H}(S^2, 4) \to \chi(S^2, 4)$ denote the orbit map of the conjugation action.

Define

$$\hat{\rho} : \mathbb{T} \to \mathcal{H}(S^2, 4), \quad (e^{x\boldsymbol{i}}, e^{y\boldsymbol{i}}) \mapsto \big(a \mapsto \boldsymbol{j},\ b \mapsto e^{x\boldsymbol{i}}\boldsymbol{j},\ c \mapsto e^{y\boldsymbol{i}}\boldsymbol{j}\big). \tag{5-5}$$

A half Dehn twist of (D^2, p, q), where p, q are a fixed pair of interior points, is a homeomorphism of the disk which fixes the boundary and permutes p and q, and which generates the infinite cyclic mapping class group rel boundary of a disk with two marked interior points. The generator which veers to the right is called a positive half Dehn twist.

Proof of Theorem B. That $\hat{\rho} : \mathbb{P} \to \chi(S^2, 4)$ is a well defined homeomorphism, as well as an analytic diffeomorphism of the smooth strata $\mathbb{P}^* \to \chi(S^2, 4)^*$, is simple; its proof can be found in [22; 18].

We show that $\hat{\rho} : (\mathbb{P}^*, c\,dx \wedge dy) \to (\chi(S^2, 4)^*, \widehat{\omega}_{(S^2,4)})$ is a symplectomorphism, for some constant c. Since this is a local statement, we work in $\mathbb{R}^2$ for simplicity. Define $\widehat{m} = \hat{\rho} \circ e : \mathbb{R}^2 \to \chi(S^2, 4)$, where $e(x, y) = (e^{x\boldsymbol{i}}, e^{y\boldsymbol{i}})$.

The differential of $\widehat{m}$ at $(x, y) \in \mathbb{R}^2$ is given by

$$d\widehat{m}\big(\tfrac{\partial}{\partial x}\big) = \big(a \mapsto 0, b \mapsto e^{x\boldsymbol{i}}\boldsymbol{ij}, c \mapsto 0\big),$$
$$d\widehat{m}\big(\tfrac{\partial}{\partial y}\big) = \big(a \mapsto 0, b \mapsto 0, c \mapsto e^{y\boldsymbol{i}}\boldsymbol{ij}\big).$$

Using left translation $L_{g^{-1}} : \mathrm{SU}(2) \to \mathrm{SU}(2)$ to identify $T_g\,\mathrm{SU}(2)$ with its Lie algebra $\mathrm{su}(2) = T_1\,\mathrm{SU}(2)$, transforms $d\widehat{m}\big(\frac{\partial}{\partial x}\big)$ and $d\widehat{m}\big(\frac{\partial}{\partial y}\big)$ to the su(2)-valued 1-cochains

$$z_x = \big(a \mapsto 0, b \mapsto -\boldsymbol{i}, c \mapsto 0\big), \quad z_y = \big(a \mapsto 0, b \mapsto 0, c \mapsto -\boldsymbol{i}\big).$$

Since $\widehat{m}$ is a local diffeomorphism away from $(\pi\mathbb{Z})^2$ [15], and $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ span $T_{(x,y)}\mathbb{R}^2$, the cohomology classes $[z_x]$, $[z_y]$ span

$$\widehat{H}^1(S^2 - 4D^2; \mathrm{su}(2)_{\mathrm{Ad}\,\widehat{m}(x,y)}) = T_{[\widehat{m}(x,y)]}(\chi(S^2, 4)).$$

For any $(x, y) \in \mathbb{R}^2 \setminus (\pi\mathbb{Z})^2$, the adjoint action

$$\operatorname{Ad}\widehat{m}(x, y) : \pi_1(S^2 - 4D^2) \to \mathrm{GL}(\mathrm{su}(2))$$

reduces as the direct sum

$$\operatorname{Ad}\widehat{m}(x, y) = \operatorname{Ad}\widehat{m}(x, y)_1 \oplus \operatorname{Ad}\widehat{m}(x, y)_2 : \pi_1(S^2 - 4D^2) \to \mathrm{GL}(\mathbb{R}\boldsymbol{i}) \times \mathrm{GL}(\mathbb{C}\boldsymbol{j})$$

and hence

$$\widehat{H}^1(S^2 - 4D^2; \mathrm{su}(2)_{\operatorname{Ad}\widehat{m}(x,y)}) = \widehat{H}^1(S^2 - 4D^2; \mathbb{R}\boldsymbol{i}_{\operatorname{Ad}\widehat{m}(x,y)_1}) \oplus \widehat{H}^1(S^2 - 4D^2; \mathbb{C}\boldsymbol{j}_{\operatorname{Ad}\widehat{m}(x,y)_2}).$$

Since $\mathbb{R}\boldsymbol{i}$ and $\mathbb{C}\boldsymbol{j}$ are orthogonal, the symplectic form $\widehat{\omega}_{S^2,4}$ splits orthogonally

$$\widehat{\omega}_{S^2,4} = \widehat{\omega}^1_{S^2,4} \oplus \widehat{\omega}^2_{S^2,4}.$$

The cocyles z_x and z_y lie in the first summand, and hence they span the first summand and the second summand is zero (these two facts can also be easily calculated directly). Hence $\widehat{\omega}_{S^2,4} = \widehat{\omega}^1_{S^2,4}$.

The representation on the first summand is independent of (x, y): indeed a, b, and c (and hence also d) act by -1 for all x, y. The cocycles z_x, z_y are independent of x, y, and hence

$$\widehat{m}^*(\widehat{\omega}_{S^2,4})|_{(x,y)}\left(\tfrac{\partial}{\partial x}, \tfrac{\partial}{\partial y}\right) = \widehat{\omega}^1_{S^2,4}(z_x, z_y) = c\,dx \wedge dy\left(\tfrac{\partial}{\partial x}, \tfrac{\partial}{\partial y}\right)$$

for a nonzero constant c (since z_x, z_y span).

The half-Dehn twists along the disks H_1 and H_2 illustrated in Figure 1 induce automorphisms of $\pi_1(S^2 \setminus 4, s_0)$, for a base point chosen outside the supports of H_1 and H_2, as indicated in the figure. These automorphisms are given by

$$H_1 : a \mapsto a,\ b \mapsto b,\ c \mapsto d = \bar{c}\bar{b}\bar{a}$$
$$H_2 : a \mapsto a,\ b \mapsto cd\bar{c} = \bar{b}\bar{a}\bar{c},\ c \mapsto c$$

and induce homeomorphisms $H_1^*, H_2^* : \chi(S^2, 4) \to \chi(S^2, 4)$. Naturality of cup products shows that H_1^*, H_2^* preserve the symplectic form $\widehat{\omega}_{(S^2,4)}$.

Next,

$$\begin{aligned} &H_1^*(m(x, y)) \\ &\quad = \left(a \mapsto m(x, y)(a) = \boldsymbol{j}, b \mapsto m(x, y)(b) = e^{x\boldsymbol{i}}, c \mapsto m(x, y)(\bar{c}\bar{b}\bar{a}) = e^{(y-x)\boldsymbol{i}}\boldsymbol{j}\right) \\ &\quad = m(x, y - x) \end{aligned}$$

and similarly

$$H_2^*(m(x, y)) = \left(a \mapsto \boldsymbol{j}, b \mapsto m(x, y)(\bar{b}\bar{a}\bar{c}) = e^{(x+y)\boldsymbol{i}}\boldsymbol{j}, c \mapsto e^{y\boldsymbol{i}}\right) = m(x + y, y).$$

It follows that the subgroup of $\mathrm{Homeo}(\chi(S^2,4))$ generated by H_1^* and H_2^* pulls back, via the homeomorphism $\hat{\rho}:\mathbb{P}\to\chi(T)$, to the subgroup of $\mathrm{PSL}(2,\mathbb{Z})$ generated by

$$\hat{\rho}^*(H_1^*)=\begin{pmatrix}1&-1\\0&1\end{pmatrix},\quad \hat{\rho}^*(H_2^*)=\begin{pmatrix}1&0\\1&1\end{pmatrix}.$$

These two matrices generate $\mathrm{PSL}(2,\mathbb{Z})$ (see [25]), finishing the proof. □

Mapping classes of $(S^2,4)$ permute the four punctures. The subgroup of the mapping class group of $(S^2,4)$ which fixes the point labeled a in Figure 1 can be shown to act on $\chi(S^2,4)$; this subgroup is isomorphic to $\mathrm{PSL}(2,\mathbb{Z})$ (see, e.g., [9]), generated by these half twists.

6. Perturbations

The holonomy perturbation process is easily understood, as well as motivated, in the language of Weinstein composition of Lagrangian immersions.

6A. ***Composition.*** Given any two (set) maps

$$\alpha: A\to M \quad\text{and}\quad \beta=\beta_M\times\beta_N: B\to M\times N,$$

define the *composition* $(A\times_M B,\beta_N^\alpha)$ by

$$A\times_M B:=\{(a,b)\in A\times B\mid \alpha(a)=\beta_M(b)\}=(\alpha\times\beta_M)^{-1}(\Delta_M) \tag{6-1}$$

and

$$\beta_N^\alpha: A\times_M B\to N,\quad \beta_N^\alpha(a,b):=\beta_N(b). \tag{6-2}$$

6A1. *Composition in character varieties.* Recall that if $\rho:\Gamma\to G$ is a homomorphism,

$$\mathrm{Stab}(\rho)=\{g\in G\mid g\rho(\gamma)g^{-1}=\rho(\gamma)\text{ for all }\gamma\in\Gamma\}.$$

The proof of the following lemma can be found in [15, Lemma 4.2].

Lemma 6.1. *Suppose that Γ_0,Γ_1, H are groups, $h_0:H\to\Gamma_0$, $h_1:H\to\Gamma_1$ homomorphisms. Set $\Gamma=\Gamma_0*_H\Gamma_1$, the pushout along h_0,h_1. There is a surjection*

$$\chi(\Gamma)\to\chi(\Gamma_0)\times_{\chi(H)}\chi(\Gamma_1)$$

with fiber over $([\rho_0],[\rho_1])$ the double coset

$$\mathrm{Stab}(\rho_0)\backslash\mathrm{Stab}(\rho_0|_H)/\mathrm{Stab}(\rho_1).$$

The fibers $\mathrm{Stab}(\rho_0)\backslash\mathrm{Stab}(\rho_0|_H)/\mathrm{Stab}(\rho_1)$ are called *gluing parameters*.

Proposition 6.2. *If $Z=Z_0\cup_\Sigma Z_1$ is a decomposition of a compact 3-manifold along a closed separating surface Σ, with $\pi_1(\Sigma)\to\pi_1(Y_0)$ surjective, then $\chi(Z)\to\chi(Z_0)\times_{\chi(\Sigma)}\chi(Z_1)$ is a homeomorphism, in fact, an algebraic isomorphism.*

Proof. Choose $([\rho_0], [\rho_1]) \in \chi(Z_0) \times_{\chi(\Sigma)} \chi(Z_1)$. Since $\pi_1(\Sigma) \to \pi_1(Y_0)$ is surjective, ρ_0 and $\rho_0|_{\pi_1(\Sigma)}$ have the same image, and hence equal stabilizers. Thus $\mathrm{Stab}(\rho_0)\backslash\mathrm{Stab}(\rho_0|_\Sigma)/\mathrm{Stab}(\rho_1)$ is a single point, and the proof follows from Lemma 6.1. □

Corollary 6.3. *If $Z = Z_0 \cup_\Sigma Z_1$ with Z_0 a handlebody, then*

$$\chi(Z) = \chi(Z_0) \times_{\chi(\Sigma)} \chi(Z_1).$$

6A2. *Composition in the Weinstein category.* The Weinstein category [32] is, roughly speaking, a category with objects symplectic manifolds and morphisms Lagrangian immersions. Composition is not always defined, however. The following criterion ensures that a composition of Lagrangian immersions is defined.

Lemma 6.4 [13, §4.1; 5, Lemma 2.0.5]. *Suppose M, N are symplectic manifolds,*

$$\alpha : A \to M \quad \text{and} \quad \beta = \beta_M \times \beta_N : B \to M^- \times N$$

are Lagrangian immersions (with M^- obtained from M by reversing the sign of the symplectic form). If $\alpha \times \beta_M$ is transverse to the diagonal $\Delta_M \in M \times M$, then $A \times_M B$ is a smooth manifold and $\beta_N^\alpha : A \times_M B \to N$ is a Lagrangian immersion.

When the transversality assumption in Lemma 6.4 holds, one says *the composition*

$$(A \times_M B, \beta_N^\alpha)$$

of (A, α) and (B, β) is defined and immersed.

In cases where the transversality assumption in Lemma 6.4 does not hold, a differential topological approach to remedying the situation would be to deform either or both of the immersions α, β. In order to retain the symplectic properties, one would typically deform them by Hamiltonian flows. In the context in this article, where character varieties correspond to flat moduli spaces in the gauge theoretic framework, we seek deformations in Lemma 6.4 that correspond to *holonomy perturbations* in the gauge theory context; we describe these in the next section.

6B. SU(2) *and holonomy perturbations.* We return to the pair (Y, C), determining $S = V \cup F$ as above, so $S = \partial Y$ and $V = nbd(C)$.

Suppose that $e : D^2 \times S^1 \hookrightarrow \mathrm{Int}(Y)$ is an embedding of a solid torus. Denote its image by Y_0, the closure of the complement by Y_1, and the separating torus by T, so that $Y = Y_0 \cup_T Y_1$. Corollary 6.3 shows that

$$\chi(Y, C) = \chi(Y_0) \times_{\chi(T)} \chi(Y_1, C),$$

so that $\chi(Y, C) \to \chi(F, \partial F)$ is exhibited as the composition of

$$\alpha : \chi(Y_0) \to \chi(T) \quad \text{and} \quad \beta : \chi(Y_1, C) \to \chi(T) \times \chi(F, \partial F).$$

For a deformation $\alpha_\pi : \chi(Y_0) \to \chi(T)$ of α, or more generally a family of functions α_π, $\pi \in \mathcal{U}$ with $\mathcal{U}$ a manifold (for example, a small open interval), we can view the deformed composition

$$\chi_\pi(Y, C) := \chi(Y_0) \circ_\pi \chi(Y_1, C), \quad \beta^{\alpha_\pi} : \chi(Y_0) \circ_\pi \chi(Y_1, C) \to \chi(F, \partial F) \quad (6\text{-}3)$$

as a *perturbed character variety with perturbation data* π.

Restrict to the case when $G = \mathrm{SU}(2)$ and J is the conjugacy class of imaginary unit quaternions. Then $\chi(Y_0)$ is simply an arc $[e^{is}]$, $0 \le s \le \pi$, and the immersion to $\chi(\partial Y_0)$ is $\alpha : [e^{is}] \mapsto [e^{is}, 1]$ in the pillowcase, described in (5-3) and (5-4). This map descends from the smooth, $\mathbb{Z}_2$ equivariant map $e^{is} \to (e^{is}, 1)$ from S^1 to T.

Let $f : \mathbb{R} \to \mathbb{R}$ be a smooth, odd, 2π periodic function. Then f determines a $\mathbb{Z}_2$ equivariant Hamiltonian deformation $e^{is} \mapsto (e^{is}, e^{if(s)})$, inducing the deformation of α given by

$$\alpha_f : [0, \pi] \to \chi(S^1 \times S^1), \quad \alpha_f(s) = [e^{is}, e^{if(s)}]. \quad (6\text{-}4)$$

The definition extends easily to the setting of a finite disjoint collection of embeddings of solid tori $e = \{e_i\}_{i=1}^k$ of disjointly embedded solid tori in Y and a corresponding collection of smooth, odd, periodic functions $f_1, \dots, f_k$ as above. Denote by π this set of perturbation data $\{(e_i, f_i)\}_{i=1}^k$. Letting Y_0 denote the union of the solid tori, (6-3) defines a way to deform the character variety.

Notice that it suffices to think of e as framed link in Y, since isotopic embeddings yield equal perturbed character varieties.

Definition 6.5. Let $\mathcal{V}$ be the vector space of smooth, odd, 2π periodic functions $\mathbb{R} \to \mathbb{R}$. Fix (Y, C) as above, where C may or may not be nonempty). Given perturbation data

$$\pi = (e, f) = \left(\bigsqcup_{i=1}^k e_i : D^2 \times S^1 \subset Y,\ f = (f_i) \in \mathcal{V}^k \right),$$

define $\chi_\pi(Y, C)$ to be the resulting perturbed (traceless) character variety.

In light of (6-4), $\chi_\pi(Y, C)$ has the following explicit description.

Proposition 6.6. *Let (Y, C) be a compact oriented 3-manifold with a collection of curves C in its boundary. Let $\pi = (e, f)$ be a choice of perturbation data, and define $Y_0 \subset Y$ to be the disjoint union of solid tori $\bigsqcup_i e_i(D^2 \times S^1)$. Finally define $\lambda_i, \mu_i \in \pi_1(Y \setminus Y_0)$ to be the loops $e_i(S^1 \times \{1\})$ and $e_i(\{1\} \times \partial D^2)$, connected to the base point in some way. Then*

$$\chi_\pi(Y, C) = \{\rho \in \chi(Y \setminus Y_0, C) \mid \text{ if } \rho(\lambda_i) = e^{is}, \text{ then } \rho(\mu_i) = e^{if(s)}\}.$$

Theorem 6.7 [16; 27]. *Given a set of perturbation data $\pi = (e, f)$ as in Definition 6.5, there is a holonomy perturbation $h_{(e,f)}$ of the flatness equation on $\mathrm{SU}(2)$ connections for which the perturbed flat moduli space is identified with the perturbed character variety as described above.*

Remark 6.8. More flexible holonomy perturbations can be defined using a solid handlebody, rather than disjoint solid tori (see, for example, [10]). In light of Corollary 6.3, there is a similar composition interpretation of the perturbed character variety, with Y_0 the handlebody. But in this setting, an explicit description of the perturbed character variety and the counterpart to the restriction map in (6-4), are not known to the authors.

We can now prove the following theorem, which is equivalent to Theorem A.

Theorem 6.9. *Suppose $A \in \chi_\pi(Y, C)$ is a regular point. Then A admits a neighborhood U so that $r|_U : U \to \chi(F, \partial F)^*$ is a Lagrangian embedding.*

Proof. Apply Lemma 6.4. □

6C. ***Dependence on perturbations.*** In [16], the second author showed that (when $\mathcal{L}$ is empty) different holonomy perturbations in general yield Legendrian cobordant immersions. Unfortunately this usually does not guarantee that they have isomorphic Floer homology. We now outline how one can address this point using Wehrheim and Woodward's quilt theory [28] (a rigorous formulation of the Weinstein category [33]), as well as its extension to the immersed case as developed by Bottman and Wehrheim [5].

One can find a discussion of holonomy perturbations in cylinders $(S, p) \times [0, 1]$, with p a finite set of points in a surface S, in [17]. In particular, Theorem 6.3 of that article states the following. Take perturbation data π with framed perturbation curve obtained by pushing a simple closed curve in $S \setminus p$ into the interior of $S \times I$. Then define the 1-parameter family of holonomy perturbations $s\pi$, $s \in [0, \epsilon)$.

The restriction

$$\chi_{s\pi}((S, p) \times I) \to \chi(S_0, p) \times \chi(S_1, p)$$

can be identified with the family of graphs of a Hamiltonian isotopy of $\chi(S)$ known as the *Goldman twist flow* associated to the simple closed curve [12]. This implies that the holonomy perturbation process can be viewed as a combination of a decomposition induced by cutting a 3-manifold along a separating torus T, followed by a perturbation as in (6-3), with $\alpha_{s\pi} : \chi(S^1 \times D^2) \to \chi(T)$ the composition of the unperturbed inclusion α_0 followed by a small time flow of the Hamiltonian Goldman twist flow associated to a curve in this torus.

We formalize this in the following way. Call two Lagrangian immersions

$$\iota_0 : L_0 \looparrowright M, \quad \iota_1 : L_1 \looparrowright M$$

secretly Hamiltonian isotopic if they can be expressed as compositions with some $\beta : \Lambda \looparrowright X \times M$:

$$i_0 : L_0 \looparrowright M = \beta_M^{j_0} : L_0' \times_X \Lambda \looparrowright M,$$
$$i_1 : L_1 \looparrowright M = \beta_M^{j_1} : L_1' \times_X \Lambda \looparrowright M,$$

in such a way that $j_0, j_1 : L_0', L_1' \looparrowright X$ are Hamiltonian isotopic in X. It follows from the discussion in Section 6B that different choices of holonomy perturbations induce secretly Hamiltonian isotopic immersed Lagrangians.

Being secretly Hamiltonian isotopic does not necessarily imply being Hamiltonian isotopic. Indeed, L_0 and L_1 need not even be diffeomorphic. Moreover, in the absence of extra hypotheses (embedded composition, monotonicity, exactness, etc.), the conclusion of the Wehrheim–Woodward composition theorem [28] need not hold. In particular, given a third Lagrangian L_2, $HF(L_0, L_2)$ and $HF(L_1, L_2)$ need not be isomorphic, even if they are both well-defined.

However, provided that

- all Lagrangian immersions and symplectic manifolds satisfy suitable assumptions so to be able to define Lagrangian (quilted) Floer homology,
- all Lagrangian immersions come equipped with suitable bounding cochains, in a way consistent with composition,

then it would follow from the Bottman–Wehrheim conjecture [5, §4.4] (see also a similar statement in [11]) that the secretly Hamiltonian isotopic Lagrangian immersions L_0 and L_1, when paired with any test Lagrangian $L_2 \subset M$, produce isomorphic Floer homology groups $HF(L_0, L_2) \simeq HF(L_1, L_2)$ (since these would respectively correspond to the quilted Floer homology groups $HF(L_0', \Lambda, L_2)$ and $HF(L_1', \Lambda, L_2)$, which are isomorphic).

In that sense secretly Hamiltonian isotopic Lagrangian immersions L_0 and L_1 can be thought as being equivalent. In particular, the problem of dependence of Floer theory on the choice of holonomy perturbation is seen as a special case of the general problem of dependence of quilted Floer homology on composition.

Acknowledgements

The authors thank C. Judge, E. Toffoli, and L. Jeffrey for helpful conversations. Special thanks to A. Kotelskiy, who read early versions of this article and made important suggestions.

References

[1] M. Atiyah, "New invariants of 3- and 4-dimensional manifolds", pp. 285–299 in *The mathematical heritage of Hermann Weyl* (Durham, NC, 1987), Proc. Sympos. Pure Math. **48**, Amer. Math. Soc., Providence, RI, 1988. MR Zbl

[2] M. Atiyah, *The geometry and physics of knots*, Cambridge University Press, 1990. MR Zbl

[3] M. F. Atiyah and R. Bott, "The Yang–Mills equations over Riemann surfaces", *Philos. Trans. Roy. Soc. London Ser. A* **308**:1505 (1983), 523–615. MR Zbl

[4] I. Biswas and K. Guruprasad, "Principal bundles on open surfaces and invariant functions on Lie groups", *Internat. J. Math.* **4**:4 (1993), 535–544. MR Zbl

[5] N. Bottman and K. Wehrheim, "Gromov compactness for squiggly strip shrinking in pseudo-holomorphic quilts", *Selecta Math. (N.S.)* **24**:4 (2018), 3381–3443. MR Zbl

[6] K. S. Brown, *Cohomology of groups*, Graduate Texts in Mathematics **87**, Springer, 1994. MR

[7] G. Cazassus, C. M. Herald, P. Kirk, and A. Kotelskiy, "The correspondence induced on the pillowcase by the earring tangle", 2020. arXiv 2010.04320

[8] D. Crowley, W. Lück, and T. Macko, "Surgery theory: Foundations", 2020, available at http://thales.doa.fmph.uniba.sk/macko/surgery-book.html.

[9] B. Farb and D. Margalit, *A primer on mapping class groups*, Princeton Mathematical Series **49**, Princeton University Press, 2012. MR Zbl

[10] A. Floer, "An instanton-invariant for 3-manifolds", *Comm. Math. Phys.* **118**:2 (1988), 215–240. MR Zbl

[11] K. Fukaya, "Unobstructed immersed Lagrangian correspondence and filtered A infinity functor", 2017. arXiv 1706.02131

[12] W. M. Goldman, "The symplectic nature of fundamental groups of surfaces", *Adv. in Math.* **54**:2 (1984), 200–225. MR Zbl

[13] V. Guillemin and S. Sternberg, "Moments and reductions", pp. 52–65 in *Differential geometric methods in mathematical physics* (Clausthal, 1980), Lecture Notes in Math. **905**, Springer, 1982. MR Zbl

[14] K. Guruprasad, J. Huebschmann, L. Jeffrey, and A. Weinstein, "Group systems, groupoids, and moduli spaces of parabolic bundles", *Duke Math. J.* **89**:2 (1997), 377–412. MR Zbl

[15] M. Hedden, C. M. Herald, and P. Kirk, "The pillowcase and perturbations of traceless representations of knot groups", *Geom. Topol.* **18**:1 (2014), 211–287. MR Zbl

[16] C. M. Herald, "Legendrian cobordism and Chern–Simons theory on 3-manifolds with boundary", *Comm. Anal. Geom.* **2**:3 (1994), 337–413. MR Zbl

[17] C. M. Herald and P. Kirk, "Holonomy perturbations and regularity for traceless SU(2) character varieties of tangles", *Quantum Topol.* **9**:2 (2018), 349–418. MR Zbl

[18] M. Heusener and J. Kroll, "Deforming abelian SU(2)-representations of knot groups", *Comment. Math. Helv.* **73**:3 (1998), 480–498. MR Zbl

[19] Y. Karshon, "An algebraic proof for the symplectic structure of moduli space", *Proc. Amer. Math. Soc.* **116**:3 (1992), 591–605. MR Zbl

[20] C. King and A. Sengupta, "A symplectic structure for connections on surfaces with boundary", *Comm. Math. Phys.* **175**:3 (1996), 657–671. MR Zbl

[21] P. B. Kronheimer and T. S. Mrowka, "Khovanov homology is an unknot-detector", *Publ. Math. Inst. Hautes Études Sci.* 113 (2011), 97–208. MR Zbl

[22] X.-S. Lin, "A knot invariant via representation spaces", *J. Differential Geom.* **35**:2 (1992), 337–357. MR Zbl

[23] R. Lipshitz, P. S. Ozsvath, and D. P. Thurston, *Bordered Heegaard Floer homology*, Mem. Amer. Math. Soc. **1216**, 2018. MR Zbl

[24] J. Milnor, "A duality theorem for Reidemeister torsion", *Ann. of Math.* (2) **76** (1962), 137–147. MR Zbl

[25] J.-P. Serre, *Trees*, Springer, 1980. MR Zbl

[26] E. H. Spanier, *Algebraic topology*, Springer, 1995. MR Zbl

[27] C. H. Taubes, "Casson's invariant and gauge theory", *J. Differential Geom.* **31**:2 (1990), 547–599. MR Zbl

[28] K. Wehrheim and C. Woodward, "Quilted Floer cohomology", *Geom. Topol.* **14**:2 (2010), 833–902. MR Zbl

[29] K. Wehrheim and C. Woodward, "Floer field theory for tangles", 2015. arXiv 1503.07615

[30] K. Wehrheim and C. Woodward, "Floer field theory for coprime rank and degree", 2016. arXiv 1601.04924

[31] A. Weil, "Remarks on the cohomology of groups", *Ann. of Math.* (2) **80** (1964), 149–157. MR Zbl

[32] A. Weinstein, "The symplectic "category"", pp. 45–51 in *Differential geometric methods in mathematical physics* (Clausthal, 1980), Lecture Notes in Math. **905**, Springer, 1982. MR Zbl

[33] A. Weinstein, "Symplectic categories", *Port. Math.* **67**:2 (2010), 261–278. MR Zbl

Received 31 Jan 2021. Revised 23 Apr 2021.

GUILLEM CAZASSUS: g.cazassus@gmail.com
Mathematical Institute, University of Oxford, Oxford, United Kingdom

CHRIS HERALD: herald@unr.edu
Department of Mathematics and Statistics, University of Nevada, Reno, NV, United States

PAUL KIRK: pkirk@indiana.edu
Department of Mathematics, Indiana University, Bloomington, IN, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.123

On naturality of the Ozsváth–Szabó contact invariant

Matthew Hedden and Lev Tovstopyat-Nelip

We discuss functoriality properties of the Ozsváth–Szabó contact invariant, and expose a number of results which seemed destined for folklore. We clarify the (in)dependence of the invariant on the basepoint, prove that it is functorial with respect to contactomorphisms, and show that it is strongly functorial under Stein cobordisms.

1. Introduction

Heegaard Floer homology provides a seemingly ever-growing number of invariants for low-dimensional topology. Its influence has perhaps most firmly been felt within the realm of 3-dimensional contact geometry, upon which the Ozsváth–Szabó contact invariant [30] and its refinements have had a profound impact. In its most basic form, the contact invariant of a closed contact 3-manifold (Y, ξ) is an element residing in the Heegaard Floer homology group $\widehat{HF}(-Y)$ of the underlying manifold, equipped with the opposite orientation to the one it receives from the contact structure (a group which, perhaps more naturally, can be identified with the Floer cohomology of Y). A decade after its initial development by Ozsváth and Szabó [27; 28], Juhász, Thurston, and Zemke discovered a subtle dependence of Heegaard Floer homology on a choice of basepoint underlying its definition [19]. Indeed, they showed that Floer homology cannot associate a well-defined group to a 3-manifold alone, but only to a 3-manifold equipped with a basepoint. This raises the questions of whether the contact element is well defined, how it depends on the basepoint, and how it behaves under diffeomorphisms, questions raised but not pursued in [19] and [22, pg. 1360].

The purpose of this article is to examine these questions, and further explore functoriality properties of the contact invariant. As a first step, we show that the

Hedden gratefully acknowledges support from NSF grants DMS-1709016 and DMS-2104664.
MSC2020: primary 57K18, 57K31, 57K33, 57K43; secondary 53D05, 53D10.
Keywords: Heegaard Floer homology, contact structure, contact invariant, functoriality, contactomorphism, naturality, Weinstein, Stein.

contact element is a well-defined element (as opposed to an orbit under the group of graded automorphisms) in the Floer homology of a pointed contact 3-manifold; see Theorem 2.3. To prove this, we first establish an appropriate definition and notion of equivalence for pointed contact 3-manifolds and incorporate these ideas into the Giroux correspondence. We then revisit Ozsváth and Szabó's proof of invariance within the naturality framework of [19], ensuring that Heegaard surfaces, links, open books, basepoints, etc. can be arranged to be explicitly embedded in a fixed 3-manifold.

Having checked the aforementioned details, we turn to a refined understanding of invariance of the contact class, showing that it is functorial with respect to pointed contactomorphisms.

Theorem 1.1. *Suppose f is a pointed contactomorphism between pointed contact 3-manifolds (Y, ξ, w) and (Y', ξ', w'). Then the induced map on Floer homology*

$$f_* : \widehat{HF}(-Y, w) \to \widehat{HF}(-Y', w')$$

carries $c(\xi, w)$ to $c(\xi', w')$.

The functoriality above is an immediate consequence of the functoriality of Floer homology under pointed diffeomorphisms from [19], provided one parses the Giroux correspondence in a categorical framework, which we clarify with Proposition 2.6.

We then show that, while the group in which the contact element lives depends on the basepoint, the contact element itself does not. This can be explained as follows: the dependence of the Floer homology of Y on the basepoint is determined by a functor

$$HF(Y, -) : \Pi_1(Y) \to \mathrm{iGrp}$$

from the fundamental groupoid of Y to the isomorphism subcategory of groups. Concretely, this just means that there is a well-defined isomorphism between Floer groups $\widehat{HF}(Y, w)$ and $\widehat{HF}(Y, w')$ associated to a homotopy class of a path between w and w', which is compatible with concatenation (see the next section for more details). If one restricts to the subgroups of $\widehat{HF}(Y, w)$ spanned by contact classes, which we denote $cHF(Y, w)$, this functor yields a transitive system indexed by points in Y; that is, the isomorphisms $cHF(Y, w) \to cHF(Y, w')$ are independent of paths. We can therefore consider the direct limit of the transitive system, which we call the *contact subgroup* of Floer homology, and denote $cHF(Y)$.

Theorem 1.2. *The contact subgroup $cHF(Y)$ is a well-defined invariant of an (unpointed) 3-manifold, functorial with respect to diffeomorphisms. There is an element $c(\xi) \in cHF(Y)$, associated to a contact structure ξ on Y, and the map $f_* : cHF(Y) \to cHF(Y')$ induced by a contactomorphism $f : (Y, \xi) \to (Y', \xi')$ sends $c(\xi)$ to $c(\xi')$.*

One can view this result in two ways. On the one hand, it follows from the functoriality established in Theorem 1.1, together with the fact that we can realize the change-of-basepoint diffeomorphism associated to a homotopy class of path by a contactomorphism (see Proposition 2.9). On the other, it can be viewed as a consequence of Zemke's calculation of the representation of the fundamental group on Floer homology in terms of the $H_1(Y)/\operatorname{Tor}$ action and the basepoint action Φ_w, together with the fact that contact classes are in the kernel of the $H_1(Y)/\operatorname{Tor}$ action. Adopting the latter perspective, we see that the contact subgroup is a subgroup of a larger basepoint independent subgroup, arising as the kernel of the $H_1(Y)/\operatorname{Tor}$ action. Note that these considerations dash any naive hope that Heegaard Floer homology is generated by contact classes, much less by elements associated to taut foliations, and indicate that such a conjecture might more reasonably be made in the context of a twisted coefficient system in which the $H_1(Y)/\operatorname{Tor}$ action vanishes, e.g., totally twisted coefficients, or for the subgroup arising as the intersection of the reduced Floer homology $HF^-_{red}(Y, w)$ and the kernel of the $H_1(Y)/\operatorname{Tor}$ action. For a rational homology sphere the latter action vanishes, and the question lands back within a similar realm to the L-space conjecture.

Having clarified the definition and invariance properties of the contact element, we then show that it is functorial under Stein cobordisms in a precise way.

Theorem 1.3. *Suppose* (W, J, ϕ) *is a Stein cobordism from a contact* 3*-manifold* (Y_1, ξ_1) *to a contact* 3*-manifold* (Y_2, ξ_2)*. Then*

$$F_{W^\dagger,\mathfrak{k}}(c(Y_2, \xi_2)) = c(Y_1, \xi_1),$$

where $W^\dagger$ *indicates the* 4*-manifold* W*, viewed as a cobordism from* $-Y_2$ *to* $-Y_1$*, and* $\mathfrak{k}$ *is the canonical* Spin^c *structure associated to* J*. Moreover,*

$$F_{W^\dagger,\mathfrak{s}}(c(Y_2, \xi_2)) = 0$$

for $\mathfrak{s} \neq \mathfrak{k}$*.*

In a weaker form, such a result follows fairly easily from the existing literature, and was widely known to experts. See Section 3 for a discussion. In the present level of specificity, the proof is slightly more involved than one might initially expect, owing largely to the nature of the composition law for cobordism maps in Heegaard Floer theory. We remark that the incoming and outgoing boundaries of W are not assumed to be connected, and that Theorem 1.3 immediately yields a generalization of Plamenevskaya's independence result for contact invariants from [33]; see Corollary 3.9.

It would be interesting to know how much naturality of the contact element persists as one weakens assumptions on the cobordism. It is known, for instance, that the contact element in monopole Floer homology is natural under strong symplectic cobordisms [2, Theorem 1]; see also [24]. One would thus expect an affirmative answer to the following:

Question 1.4. Is the contact element natural under strong symplectic cobordisms?

Finally, we note that we work with Heegaard Floer homology with $\mathbb{Z}/2\mathbb{Z}$ coefficients throughout. The naturality results of [19] have been extended to projective $\mathbb{Z}$ coefficients (i.e., $\mathbb{Z}/\pm 1$) [7], but at the moment these extensions have not been established for the graph cobordism maps. Assuming they will be, the results at hand should immediately extend to the refined setting.

2. Functoriality of the contact class under diffeomorphisms

In this section we clarify the dependence of the contact class on the basepoint used in the definition of Heegaard Floer homology, and highlight how the results of Juhász, Thurston and Zemke [19] and Zemke [34] couple with Ozsváth and Szabó's argument from [30] to imply that the contact class is a well-defined invariant of a *pointed* contact 3-manifold, up to pointed isotopy (Theorem 2.3). This invariant is shown to be functorial under pointed contactomorphisms (Theorem 2.7). We then show that while, according to [19], the Heegaard Floer homology group in which the contact element lives depends in an essential way upon the basepoint, the contact invariant is essentially independent from the basepoint (Theorem 1.2). Unless otherwise specified, all 3-manifolds are assumed to be closed and oriented, and contact structures assumed to be cooriented.

Recall from [19] that the Heegaard Floer homology group of a pointed 3-manifold (Y, w) is defined as the direct limit of a transitive system of groups and isomorphisms defined by pointed Heegaard diagrams $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, w)$ embedded in (Y, w) and pointed Heegaard moves passing between them (together with auxiliary choices of almost complex structures). See [19, Theorem 1.5] and the surrounding discussion. The contact invariant should therefore be interpreted as an element in the aforementioned direct limit. As such, it would appear to depend on the basepoint, and we therefore make the following definition

Definition 2.1. A *pointed contact* 3*-manifold* is a 3-manifold Y equipped with a contact structure ξ and a distinguished basepoint w.

By Gray's theorem, an isotopy between contact structures is induced by an isotopy of the underlying (compact) 3-manifold, and two contact structures on (Y, w) will be considered equivalent if we can find such an isotopy fixing the basepoint w.

Remark 2.2. One could consider a more restrictive definition of equivalence where the isotopy fixes the contact plane at w. This differs from the present notion only by the choice of oriented plane at w, a choice parametrized by a 2-sphere, and would have no effect on our results. See [19, Lemma 2.45] for more details.

In [30], Ozsváth and Szabó defined an invariant of contact structures utilizing the Giroux correspondence between isotopy classes of contact structures on Y and

equivalence classes of fibered links in Y under Hopf plumbing. Given a fibered knot representing ξ, its knot Floer homology has a distinguished filtered subcomplex in bottommost Alexander grading, whose homology is rank one [30, Theorem 1.1]. Inclusion of this subcomplex in $\widehat{CF}(-Y)$ defines an element $c(\xi) \in \widehat{HF}(-Y)$ [30, Definition 1.2], which they showed does not depend on the particular choice of fibered knot representing ξ [30, Theorem 1.3]. Absent from the literature at that time, however, was an understanding of the dependence of the Floer homology group in which $c(\xi)$ resides on the choice of basepoint. We restate their theorem so that this dependence is explicit, and outline the elements of their proof of invariance which should be refined accordingly.

Theorem 2.3 [30, Theorem 1.3]. *Suppose two contact structures ξ, η on the pointed 3-manifold (Y, w) are equivalent. Then $c(\xi, w) = c(\eta, w) \in \widehat{HF}(-Y, w)$.*

Proof. We begin by observing that the Giroux correspondence [6; 9] has the pointed analogue

$$\frac{\{\text{pointed open books } (L, \pi_L) \text{ in } (Y^3, w)\}}{\{\text{(pointed) isotopy and positive Hopf plumbing}\}} = \frac{\{\text{contact structures on } (Y^3, w)\}}{\{\text{isotopy fixing } w\}}$$

where we emphasize that on the left-hand side we are considering *concrete* open book decompositions of Y, by which we mean an embedded link $L \subset Y$ together with a fibration on its exterior $\pi_L : Y \setminus L \to S^1$ for which the boundary of the closure of each fiber is L. In these terms a *pointed open book* for the pointed contact manifold is an open book supporting ξ for which the basepoint is contained in L. Two open books are considered equivalent if they differ by a sequence consisting of ambient isotopies of links and ambient (de)plumbings with positive Hopf bands, where isotopies and (de)plumbings are required to fix the basepoint. The pointed statement follows easily from the unpointed statement. We remark that the inclusion of ambient isotopies of the open book is essential, though typically omitted or implicit in the literature.

Ozsváth and Szabó's proof relies on two lemmas. If we denote the element associated to a fibered knot $K \subset Y$ by $c(K) \in \widehat{HF}(-Y)$, then [30, Lemma 4.1] states that this element is unchanged under connected summing with the right-handed trefoil, T; that is, $c(K\#T) = c(K)$ for any fibered knot $K \subset Y$. Then [30, Lemma 4.4] shows that the element associated to a fibered knot obtained by plumbing $2h$ right-handed Hopf bands to $K \subset Y$ is independent of the choice of plumbings. Using plumbings which realize iterated connected sums with T, the result follows.

The proof of the latter lemma goes by realizing the element associated to any genus h stabilization as the image of a fixed class in $\widehat{HF}(-Y)$ under a map induced by a cobordism W which is diffeomorphic to $Y \times [0, 1]$. To do this, one observes that a genus h stabilization can be obtained by attaching canceling 4-dimensional 1- and

2-handles to $Y \times I$, where the former add handles to the page of the open book and the latter enact Dehn twists to the monodromy. One then uses [30, Theorem 4.2], which states that the element is carried naturally under the 2-handle cobordisms that add left-handed Dehn twists to the monodromy.

According to [19], one should refine these arguments so that they use embedded Heegaard diagrams in Y, with the basepoint lying on the embedded Heegaard surface. When appealing to the functoriality with respect to cobordisms used in the proof of Ozsváth and Szabó's second lemma, one must also be careful about the embedded path in $W \cong Y \times [0, 1]$ from the basepoint to itself.

To address the first issue, consider a concrete pointed open book (K, π_K) in Y supporting (a contact structure isotopic to) ξ, with connected binding. From this, one can construct Ozsváth and Szabó's Heegaard diagram adapted to K from [30, Section 3]. This construction can be done so that the diagram is embedded in Y, built from the union of the closure of two fibers by a stabilization, and so that it contains the basepoint. The proof of naturality for knot Floer homology adapts to produce a functorial invariant from the category of pointed knots to a category whose objects are transitive systems of $\mathbb{Z}$-filtered complexes (where the maps in such systems are certain canonical filtered homotopy classes of filtered homotopy equivalences). See [14, Proposition 2.3] for the adaptation of the proof of naturality in [19] to the context of transitive systems of complexes, and [13, Proposition 2.8] for a discussion on how to apply this to knots. Naturality implies that the generator of the homology of the bottommost filtered subcomplex defined by the embedded Heegaard diagram for the knot (K, π_K) produces a well-defined invariant $c(K, w) \in \widehat{HF}(-Y, w)$, by consideration of the inclusion-induced map.

To argue that the element $c(K, w)$ is invariant under connected summing with a trefoil, we observe that one can form the connected sum of the open book with the trefoil knot *ambiently*, by embedding the trefoil and its fiber surface in a small ball near the basepoint, but in the complement of the Heegaard surface for K. One can then form an embedded Heegaard diagram adapted to $K\#T$, which is a connected sum of embedded diagrams. The Künneth theorem for the knot Floer homology of a connected sum [26, Theorem 7.1], together with the fact that the new diagram is obtained from the initial diagram by a sequence of pointed embedded Heegaard moves, shows that $c(K\#T, w) = c(K, w)$.

The second lemma from Ozsváth and Szabó's proof also goes through in the context of pointed 3-manifolds and concrete open books. Indeed, if we are given a concrete pointed open book (K', π') for (Y, w) which is obtained from (K, π) by ambiently plumbing $2h$ positive Hopf bands, we cancel the additional right-handed Dehn twists in the monodromy by attaching $2h$ 4-dimensional 2-handles along curves in the page. This results in a cobordism whose outgoing boundary is diffeomorphic to $Y\#^{2h} S^1 \times S^2$. Further attaching $2h$ 4-dimensional 3-handles

to cancel the 2-handles results in a composite cobordism which is diffeomorphic, rel boundary, to $Y \times [0, 1]$. Moreover, since the attaching regions for the 2- and 3-handles lie in the complement of the basepoint, the path traced by the basepoint in the cobordism is sent, under a diffeomorphism to $Y \times [0, 1]$, to the trivial path $w \times [0, 1]$. We can reverse the orientation of Y before performing the aforementioned handle attachments, and the resulting map on the Floer homology of $\widehat{HF}(-Y, w)$ is the identity. The result follows as in [30] by appealing to the naturality of the contact invariant under addition of Dehn twists. □

Remark 2.4. One could alternatively approach the pointed invariance of the contact element using its interpretation by Honda, Kazez and Matić [17]. Such an approach seems necessary to establish naturality of the contact invariants in sutured Floer homology defined using partial open books for 3-manifolds with convex boundary [16].

To understand the functoriality of the contact class, we observe that pointed contact 3-manifolds form the objects of a category whose morphisms are pointed isotopy classes of contactomorphisms. With respect to this structure, the (pointed) Giroux correspondence is functorial. To understand this, we make the following definition:

Definition 2.5. A (*pointed*) *diffeomorphism between concrete* (*pointed*) *open books* is an orientation-preserving diffeomorphism of pairs $f : (Y, L) \to (Y', L')$ which intertwines the fibrations on the link complements, i.e., $\pi_L = \pi_{L'} \circ f_{Y \setminus L}$ (and which maps the basepoint on L to the basepoint on L').

If the open book (L, π_L) supports a contact structure ξ_L on Y, then a diffeomorphic open book (L', π'_L) supports a contact structure $\xi_{L'}$ on Y' satisfying $f_*(\xi_L) = \xi_{L'}$. Since the contact structure induced by an open book is only well defined up to isotopy, we will regard diffeomorphisms of open books up to isotopy without loss of information. In this way, a diffeomorphism of open books defines an isotopy class of contactomorphisms.

Conversely, given an isotopy class of contactomorphisms $f : (Y, \xi) \to (Y', \xi')$, we can push-forward an open book (L, π_L) supporting ξ under f, yielding an open book $(f(L), \pi_L \circ f^{-1})$ supporting (Y', ξ'). The evident diffeomorphism of open books induces the given contactomorphism, up to isotopy. In this way, the Giroux correspondence can be lifted to an isomorphism of categories:

Proposition 2.6 (functorial Giroux correspondence). *There is an isomorphism of categories between the category of* (*pointed*) *contact* 3*-manifolds and* (*pointed*) *isotopy classes of contactomorphisms and the* ***concrete open book category***, *whose objects are* 3*-manifolds equipped with concrete* (*pointed*) *open books up to ambient* (*pointed*) *Hopf plumbing and whose morphisms are* (*pointed*) *isotopy classes of* (*pointed*) *diffeomorphisms between open books.*

Proof. The (pointed) Giroux correspondence yields a bijection between objects which, in one direction, sends an (equivalence class of (pointed)) open book to the ((pointed) isotopy class of a) contact structure supporting it. The discussion above shows that there are corresponding bijections between morphism sets. □

Using this, we can show that the contact class is functorial with respect to pointed contactomorphisms; see Theorem 1.1.

Theorem 2.7 (functoriality under contactomorphisms). *Suppose f is a pointed contactomorphism between pointed contact manifolds (Y, ξ, w) and (Y', ξ', w'). Then the map on Floer homology $f_* : \widehat{HF}(-Y, w) \to \widehat{HF}(-Y', w')$ carries $c(\xi, w)$ to $c(\xi', w')$.*

Proof. This follows easily from the definition of the map on Floer homology associated to a pointed diffeomorphism, as described in [19, Section 2.5, Definition 2.42], together with the functorial Giroux correspondence. More precisely, according to Section 2.5 of [19], the map between Floer homology groups associated to a pointed isotopy class of diffeomorphism is defined by the map on transitive systems induced by pushing forward embedded pointed Heegaard diagrams in (Y, w) (and moves between them) to (Y', w'). A pointed Heegaard diagram adapted to a concrete open book supporting (Y, ξ) is mapped, via f, to a pointed Heegaard diagram adapted to a diffeomorphic concrete open book supporting (Y', ξ'). Taking homology of these complexes gives rise to representatives for the direct limit of the transitive systems that define $\widehat{HF}(-Y, w)$ and $\widehat{HF}(-Y', w')$, respectively. Under the induced map, the cycle representing the contact element for $c(\xi, w)$ is taken to that representing $c(\xi', w')$. The result follows. □

Since the hat Floer homology groups depend on the basepoint, the above refinements are necessary in order to understand the invariance of the contact class. Having addressed this, however, we will now show that the contact class is essentially *independent* of the basepoint, relying on it only insomuch as it is required to define the group in which the class resides.

To explain this, recall that the map $\mathrm{Diff}(Y) \xrightarrow{\mathrm{ev}_w} Y$ which evaluates a diffeomorphism at a basepoint is a Serre fibration, and the fiber over w is the pointed diffeomorphism group $\mathrm{Diff}(Y, w)$. The associated long exact sequence on homotopy terminates in

$$\pi_1(Y, w) \to \pi_0(\mathrm{Diff}(Y, w)) \to \pi_0(\mathrm{Diff}(Y)) \to 1.$$

Concretely, this implies that if a pointed diffeomorphism is (unpointed) isotopic to the identity, then it is isotopic to a "point-pushing map" about a loop representing an element in $\pi_1(Y, w)$. If one considers instead the fiber of ev_w over a different basepoint, w', we see that any diffeomorphism of Y sending w to w' which is (unpointed) isotopic to the identity is isotopic, through diffeomorphisms sending

w to w', to a point-pushing map defined by a choice of arc γ from w to w'. Moreover, any two such diffeomorphisms differ, up to pointed isotopy, by a point-pushing map along an element in $\pi_1(Y, w)$. In light of this, the dependence of Floer homology on the basepoint is captured by a representation $\pi_1(Y, w) \to \mathrm{Aut}(\widehat{HF}(Y, w))$ defined by isomorphisms associated to isotopy classes of point-pushing diffeomorphisms. While this representation can be nontrivial, the following proposition implies that it acts trivially on the subspace spanned by contact elements.

Proposition 2.8. *Suppose that contact structures ξ and η on the pointed 3-manifold (Y, w) are isotopic, induced by an isotopy of Y which does not necessarily fix the basepoint. Then $c(\xi, w) = c(\eta, w) \in \widehat{HF}(-Y, w)$.*

We note that the functoriality of the contact invariant only implies $f_*(c(\xi, w)) = c(\eta, w)$, where f is the endpoint of the isotopy.

Proof. Let $\phi_t : Y \times [0, 1] \to Y$ denote the isotopy carrying ξ to η, where $\phi_0 = \mathrm{Id}_Y$ and $\phi_1 = f$ is a diffeomorphism fixing w, but where ϕ_t may not fix the basepoint for $0 < t < 1$. The discussion preceding the proposition indicates that f is isotopic to a point-pushing map along a curve γ representing an element $[\gamma] \in \pi_1(Y, w)$. More precisely, a loop γ based at w can be regarded as an isotopy of embeddings of a point into Y which, by the isotopy extension theorem, can be extended to an isotopy of Y which is the identity outside a neighborhood of the image of γ. The endpoint of this latter isotopy is a pointed diffeomorphism $f_\gamma : (Y, w) \to (Y, w)$ whose pointed isotopy class depends only on the homotopy class $[\gamma] \in \pi_1(Y, w)$, by another application of the isotopy extension theorem (or, rather its interpretation in terms of the homotopy lifting property of the map $\mathrm{Diff}(M) \to \mathrm{Diff}(N, M)$ which evaluates a diffeomorphism at a submanifold; see [20; 32]). According to the main theorem of [19], there is an induced automorphism $(f_\gamma)_*$ of the Floer homology group $\widehat{HF}(Y, w)$, and the functoriality of the contact class under pointed contactomorphisms implies

$$(f_\gamma)_*(c(\xi, w)) = c(f_{\gamma *}(\xi), w) = c(\eta, w).$$

The automorphism $(f_\gamma)_*$ will, in general, be nontrivial; indeed, Zemke shows that it can be computed via the formula [34, Theorem D]

$$(f_\gamma)_* = \mathrm{Id} + (\Phi_w)_* \circ (A_\gamma)_*,$$

where A_γ is the chain level map defining the $H_1(Y)/\mathrm{Tor}$ action on $\widehat{HF}(Y, w)$, and Φ_w is the basepoint action which, in the case of $\widehat{CF}(Y, w)$, counts J-holomorphic disks which pass through the hypersurface specified by the basepoint exactly once. The proposition will follow if we can show that contact classes are in the kernel of the $H_1(Y)/\mathrm{Tor}$ action. But this is an easy consequence of their definition. Letting $c \in H_*(\mathcal{F}(-Y, K, w, \mathrm{bot})) \cong \mathbb{F}$ denote the generator of the homology of the

bottommost nontrivial filtered subcomplex in the filtration of $\widehat{CF}(-Y, w)$ induced by the binding of a pointed open book supporting ξ, the contact class is defined as

$$c(\xi, w) := \iota_*(c),$$

where $\iota : \mathcal{F}(-Y, K, w, \text{bot}) \hookrightarrow CF(-Y, w)$ is the inclusion map. The chain map A_γ on $\widehat{CF}(-Y, w)$ respects the filtration induced by K, defined as it is by counting J-holomorphic disks which avoid w (see [13, Proof of Proposition 5.8]). It follows that A_γ maps $\mathcal{F}(-Y, K, w, \text{bot})$ to $\mathcal{F}(-Y, K, w, \text{bot})$, but since it shifts the relative $\mathbb{Z}/2\mathbb{Z}$ homological grading, and the homology of the latter subcomplex is one-dimensional, the map on homology must be trivial. Therefore the automorphism on Floer homology induced by a point-pushing map acts as the identity on any contact elements, and we have $c(\xi, w) = (f_\gamma)_*(c(\xi, w)) = c(\eta, w)$ as claimed. □

The above proposition can be used to show that change-of-basepoint maps on Floer homology induced by pushing points along arcs act on contact elements in a canonical way, i.e., $(f_\gamma)_*(c(\xi, w)) \in \widehat{HF}(Y, w')$ is independent of the choice of arc used to construct a diffeomorphism $f_\gamma : (Y, w) \to (Y, w')$. This indicates an independence of the contact class from the choice of basepoint. We can make this independence more precise. To do this, we show that the point-pushing maps along arcs can be refined to pointed contactomorphisms.

Proposition 2.9 (cf. [11]). *Given $w, w' \in Y$, there exists a contactomorphism $\phi : (Y, \xi) \to (Y, \xi)$, which is isotopic to the identity and maps w to w'.*

Proof. Let $\gamma : [0, 1] \to Y$ denote a smooth embedded path from w to w'. After a $\mathcal{C}^\infty$-small isotopy we may assume the path γ is transverse to ξ. Let $\nu(\gamma([0, 1]))$ denote a neighborhood of the transverse arc. A standard neighborhood theorem gives a contact embedding

$$\phi : (\nu(\gamma([0, 1])), \xi) \to (\mathbb{R}^3, \ker(\alpha)), \quad \text{where } \alpha = dz + r^2 d\theta,$$

which takes the image of the arc to the segment $\{(0, 0)\} \times [0, 1]$ along the z-axis; in particular $\phi(w) = (0, 0, 0)$ and $\phi(w') = (0, 0, 1)$.

Let β denote a contact 1-form for ξ which is an extension of $\phi^*\alpha$. The time one flow of the Reeb vector field R_β is then the desired contactomorphism taking w to w'. □

Corollary 2.10. *The contact class is independent of the basepoint in the following sense: given two basepoints $w, w' \subset Y$, a path γ between them induces an isomorphism $\gamma_* : \widehat{HF}(-Y, w) \to \widehat{HF}(-Y, w')$. For any choice of γ, the contact class satisfies $\gamma_*(c(\xi, w)) = c(\xi, w')$.*

Proof. Suppose γ is a path connecting w to w'. As in the proof of Proposition 2.8, the based homotopy class of γ gives rise to a well-defined pointed isotopy class

of pointed diffeomorphism $f_\gamma : (Y, w) \to (Y, w')$, whose associated isomorphism between Floer homology groups we denote γ_*. The contactomorphism constructed using γ in the proof of the preceding proposition is a representative of this pointed isotopy class. Functoriality of the invariant under pointed contactomorphisms, Theorem 2.7, then implies that $\gamma_*(c(\xi, w)) = c(\xi, w')$. □

Proof of Theorem 1.2. In light of the corollary, if we let the subgroup of $\widehat{HF}(-Y, w)$ spanned by contact elements be denoted $cHF(-Y, w)$, we obtain a transitive system (in the sense of [3, Definition 6.1]) of groups indexed by points in Y, for which the isomorphism $f_{w,w'} : cHF(-Y, w) \to cHF(-Y, w')$ is the map on Floer homology associated to the point-pushing map along *any* arc from w to w'. We call the direct limit of this transitive system the *contact subgroup* associated to Y, and denote it $cHF(-Y)$:

$$cHF(-Y) := \varinjlim cHF(-Y, w).$$

The corollary shows that it is well defined, independent of any choice of basepoint, and that a contact structure ξ on Y receives an associated element $c(\xi) \in cHF(-Y)$ defined as the image of $c(\xi, w) \in cHF(-Y, w)$ under the canonical inclusion-induced isomorphism $cHF(-Y, w) \to cHF(-Y)$. The contact subgroup is functorial with respect to (unpointed) diffeomorphisms of Y by the main theorem of [19], and an (unpointed) contactomorphism $f : (Y, \xi) \to (Y', \xi')$ sends $c(\xi)$ to $c(\xi')$ by Theorem 2.7. □

Remark 2.11. There is a contact invariant $c^+(\xi) \in HF^+(-Y)$ defined as the image of $c(\xi)$ under the map on homology induced by the inclusion of complexes $\iota : \widehat{CF} \to CF^+$, [25, Section 4]. Corresponding results for $c^+(\xi)$ follow from Theorems 1.1 and 1.2, together with the naturality of ι_* implied by [19, Theorem 1.5].

3. Functoriality of the contact class under Stein cobordisms

In this section we provide a proof of the well-known folk theorem that the Ozsváth–Szabó contact invariant is natural with respect to Stein cobordisms. Nontriviality of the contact invariant of a Stein fillable contact structure was proved in [30, Theorem 1.5]. The proof relied on a naturality result [30, Theorem 4.2] for the invariants of contact structures represented by open book decompositions which differ by a single Dehn twist. This latter result implicitly showed that the contact invariant is natural with respect to a Stein cobordism associated to a Weinstein 2-handle attachment along a Legendrian knot, a fact made more clear in [21, Theorem 2.3] (though stated there in terms of contact $+1$ surgery). These naturality results for Weinstein 2-handles consider the sum of maps associated to all the Spin^c structures on the cobordism. Together with a calculation for 1-handles, they immediately yield a weak naturality of the contact invariant under Stein cobordisms, where one

sums over all Spin^c structures. This is spelled out in [18, Theorem 11.24], under an additional topological restriction on the 1-handles.

The Spin^c refinement of naturality for the contact invariant of a Stein *filling*, viewed as a Stein cobordism to the standard structure on the 3-sphere, was established in [33, Theorem 4]. A Spin^c refinement of naturality for the contact invariant under general Weinstein 2-handle attachments along a Legendrian link was stated in [8, Lemma 2.11]. The proof relied crucially on a naturality result for the cobordism map associated to a Lefschetz fibration over an annulus. The latter was attributed to Ozsváth and Szabó, who only proved the result for Lefschetz fibrations over a disk. We spell out the proof of the required naturality in Lemma 3.6 below, and use it to establish naturality of the contact class under Weinstein 2-handle cobordism following the strategy in [8]. Given the body of literature on topological aspects of Stein surfaces and domains, exposed beautifully in [1; 4; 10], the only remaining piece necessary for the Spin^c refinement of naturality under a general Stein cobordism (Theorem 1.3) is a discussion of 1-handles, particularly those with feet in different path components.

Recall, then, that a Stein cobordism from a contact 3-manifold (Y_1, ξ_1) to (Y_2, ξ_2) is a smooth 4-manifold W with $\partial W = -Y_1 \cup Y_2$, oriented by a complex structure J for which the oriented complex lines of tangency on ∂W agree with ξ_1 and ξ_2, respectively, and which admits a *J-convex Morse function* ϕ, defined by the requirement that $-dd^{\mathbb{C}}\phi = \omega_\phi$ is symplectic. Such a manifold comes equipped with a Liouville vector field, X_ϕ, defined as the gradient of ϕ with respect to the metric induced by ω_ϕ. See [1] for an introduction.

Theorem 3.1. *Suppose (W, J, ϕ) is a Stein cobordism from a contact 3-manifold (Y_1, ξ_1) to a contact 3-manifold (Y_2, ξ_2). Then*

$$F_{W^\dagger, \mathfrak{k}}(c(Y_2, \xi_2)) = c(Y_1, \xi_1),$$

where $W^\dagger$ denotes the 4-manifold W, viewed as a cobordism from $-Y_2$ to $-Y_1$, and $\mathfrak{k}$ is the canonical Spin^c *structure associated to J. Moreover,*

$$F_{W^\dagger, \mathfrak{s}}(c(Y_2, \xi_2)) = 0$$

for $\mathfrak{s} \neq \mathfrak{k}$.

Remark 3.2. The result is equally valid for Weinstein cobordisms, which [1] shows are equivalent to Stein cobordisms for the present purposes.

Remark 3.3. Strictly speaking, the Stein cobordism should be equipped with a properly embedded graph, in the sense of [34]. In this context, the graph is obtained from the basepoints present on the incoming end of the cobordism by their image under the flow of the Liouville vector field. We pick basepoints on the incoming ends which flow to the outgoing ends, with some extra care taken in the case that

components of the boundary merge via Stein 1-handles so that all components of the boundary have a single basepoint (see Lemma 3.5 below). In light of the naturality results from the previous section, and the resulting independence of the contact class of the choice of basepoint, we can safely omit basepoints from most of the discussion and obtain a naturality result for the contact invariant which is basepoint independent.

Proof. By [4, Theorem 1.3.3], the cobordism can be decomposed as a composition of elementary cobordisms corresponding to Stein 0-, 1-, and 2-handle attachments, with the latter two attached along framed points and Legendrian curves, respectively. In this dimension, the subtleties involved with 2-handle framings were clarified by [10]. Though we could avoid it with a more cumbersome inductive argument, we can and will assume that the attachments are ordered by their indices, arising from a self-indexing plurisubharmonic Morse function. For a smooth manifold, this follows from the standard rearrangement theorem for Morse functions [23, Theorem 4.8]. The proof of that theorem, however, modifies the gradient-like vector field for the Morse function so that the stable manifold of an index λ critical point is disjoint from the unstable manifold of an index $\lambda' \geq \lambda$ critical point (achieving the Morse–Smale condition for the manifolds associated to these critical points). In the Stein setting, the gradient vector field and metric are coupled, and one cannot vary one without changing the other. Rearranging critical levels is therefore more subtle. These subtleties are nicely exposed, and dispatched with, in Chapter 10 of [1]. Of particular relevance are Proposition 10.10 and 10.1. Proposition 10.10 allows one to vary the critical values of the J-convex Morse function specifying the handle decomposition, provided the stable and unstable manifolds of the points of interest are disjoint, the Stein analogue of [23, Theorem 4.1]. Proposition 10.1 allows one to vary an isotropic submanifold of a given contact type hypersurface by an isotropic isotopy compatible with a family of J-convex Morse functions, the Stein analogue of [23, Lemma 4.7]. Applying the latter to the attaching spheres of the Stein handles allows us to assume, as in the classical case, that the stable manifold of an index λ critical point is disjoint from the unstable manifold of an index $\lambda' \geq \lambda$ critical point. Thus we can proceed by induction to order the handles and further ensure that all critical points of a given index have the same critical value. The existence of such an ordering for a 2-dimensional Stein domain (a Stein cobordism with $Y_1 = \varnothing$) is implicit in the statement of [10, Theorem 1.3], a result which itself is attributed as implicit in Eliashberg [4].

We assume then, that the cobordism is decomposed as a sequence of elementary 0- and 1-handle cobordisms, followed by a cobordism associated to a collection of Weinstein 2-handle attachments along Legendrian curves, equipped with a J-convex Morse function with a unique critical value. We will show that the contact invariant is mapped in the specified way under a single 0- or 1-handle attachment,

and similarly for a simultaneous collection of Stein 2-handle attachments. The result will then follow from the composition law for cobordism-induced maps on Heegaard Floer homology:

$$F_{W_1^\dagger,\mathfrak{t}_1} \circ F_{W_2^\dagger,\mathfrak{t}_2} = \sum_{\{\mathfrak{s}\in \mathrm{Spin}^c(W) \mid \mathfrak{s}|_{W_i}=\mathfrak{t}_i\}} F_{W^\dagger,\mathfrak{s}}.$$

Examining the law, one observes that naturality of the contact invariant for Stein cobordisms W_1 and W_2 sharing a common intermediate boundary does not imply either of the conclusions in the statement of the theorem for their union $W = W_1 \cup W_2$, if Spin^c structures on the latter are not uniquely determined by their restrictions to W_1 and W_2. The following standard lemma makes this precise:

Lemma 3.4. *Suppose $W = W_1 \cup_Y W_2$ is a 4-manifold glued along a 3-manifold Y arising as a connected component of ∂W_i (with boundary orientation of Y different for $i = 1, 2$). Then the set of* Spin^c *structures on W restricting to $\mathfrak{t}_i \in \mathrm{Spin}^c(W_i)$, provided it is nonempty, is in affine correspondence with $\delta H^1(Y)$, where δ is the connecting homomorphism in the Mayer–Vietoris sequence.*

In particular, if either W_i is a cobordism associated to a 0- or 1-handle attachment, then a Spin^c structure on W is uniquely determined by its restrictions to the pieces. This is because 0- and 1-handle cobordisms have the property that the restriction $H^1(W_i) \to H^1(\partial W_i)$ is surjective, which implies $\delta H^1(Y)$ is trivial, by exactness. Therefore, we can treat 0- and 1-handles individually. It is certainly possible, however, that a Spin^c structure on a 4-manifold composed of two or more 2-handle cobordisms will not be determined by its restrictions to the pieces. It is therefore not sufficient to prove the naturality of $c(\xi)$ with respect to a single Stein 2-handle. For this reason, we group the index 2 critical points giving rise to the 2-handles together into a single critical level, and prove naturality for such a 2-handle cobordism.

We turn to our treatment of the handles in each dimension. The fact that the contact invariant is natural under Stein 0-handle attachment follows immediately from the definition of the associated map on Floer homology, which is simply the map induced by the canonical isomorphism between Heegaard Floer chain complexes under taking disjoint union with a Heegaard diagram whose surface is a pointed 2-sphere with no curves [34, Section 11.1]. This definition, together with the fact that the contact class of the Stein fillable contact structure on the 3-sphere is nontrivial, yield, upon taking duals, the stated naturality.

The following lemma establishes naturality under Stein 1-handle cobordisms.

Lemma 3.5. *Suppose (W, J) is the cobordism associated to a Stein 1-handle attachment. Then Theorem 3.1 is true for $F_W^\dagger$.*

Proof. Unlike the case of a Stein domain, a Stein cobordism can have disconnected boundary. Thus, there are two possibilities (1) the feet of the 1-handle lay in different

components of Y_1, the incoming boundary of W, or (2) the feet of the 1-handle lay in the same component of Y_1. In the former, we may assume without loss of generality that there only two components of Y_1, since Floer homology is manifestly multiplicative under disjoint unions (i.e., groups and homomorphisms associated to disjoint unions of 3-manifolds and cobordisms, respectively, are tensor products), and product cobordisms map contact invariants naturally according to the previous section.[1] Similarly, for the second possibility, we may assume Y_1 is connected.

Naturality for 1-handles that connect two components is a consequence of a calculation for the graph cobordism map of the 1-handle cobordism, endowed with a trivalent (strong ribbon) graph that merges the two basepoints in the incoming components to a single basepoint in their outgoing connected sum. This calculation is the content of [15, Proposition 5.2], which indicates that such a graph cobordism induces a chain homotopy equivalence, and the complex associated to a connected sum is therefore homotopy equivalent to the tensor product of the complexes associated to the factors. This calculation reproved, in a functorial way, Ozsváth and Szabó's earlier connected sum formula [27, Theorem 6.2], under which the contact invariant behaves multiplicatively for contact connected sums [12, product formula]. The claimed naturality for Stein 1-handles is then immediate, provided that Ozsváth and Szabó's chain homotopy equivalence, used by the product formula for the contact invariant, agrees, up to homotopy, with the map Zemke associates to the 1-handle cobordism. But this is precisely the content of [35, Proposition 8.1]. Here, we should point out that the trivalent graph arises naturally from the Stein structure, as the stable manifold of the index one critical point of ϕ with respect to X_ϕ, union a flowline of X_ϕ from the critical point to the outgoing boundary.

The case of 1-handles with feet on the same component of the incoming boundary is simpler and, in this case, follows from Ozsváth and Szabó's definition of the 1-handle map [31, Section 4.3], together again with the fact that the contact invariant is multiplicative under contact connected sums. In this case, the outgoing manifold is contactomorphic to the connected sum $(Y\#(S^1 \times S^2), \xi\#\xi_{std})$, so it suffices to show that the image of the dual of $c(\xi)$ under Ozsváth and Szabó's map induced by the 1-handle agrees with the dual $c(\xi\#\xi_{std})$. But the 1-handle map sends $c(\xi)^*$ to $c(\xi)^* \otimes \Theta_+$. Thus the problem is reduced to a single calculation, verifying that the dual of the contact class of the standard contact structure on $S^1 \times S^2$ satisfies $\Theta_+ = (c(S^1 \times S^2, \xi_{std}))^* \in \widehat{HF}(S^1 \times S^2)$. This calculation can be done in numerous ways; see, e.g., [13, Proof of Proposition 5.19], for an explicit treatment. □

[1] Here, a product cobordism means a 4-manifold diffeomorphic to $Y \times I$, through a diffeomorphism induced by the flow of the Liouville vector field. Since the "holonomy" diffeomorphism from the outgoing boundary to the incoming boundary [1, Definition 9.40] is a contactomorphism, the naturality results of the previous section, together with [34, Theorem B.2], indicate the contact invariants are mapped in the specified way.

Next we turn to 2-handles. While the naturality statement needed here for Stein 2-handle cobordisms was stated by Ghiggini in [8, Lemma 2.11], the proof relied on a naturality result for the cobordism map associated to a Lefschetz fibration over an annulus, attributed to [29, Theorem 5.3]. The latter theorem applies only to Lefschetz fibrations over the disk. The desired result can be derived from Ozsváth and Szabó's by a capping argument, together with the composition law for cobordism induced maps on Floer homology. We spell this out explicitly.

Lemma 3.6 (cf. [29, Theorem 5.3]). *Let $\pi : W \to [0,1] \times S^1$ be a relatively minimal Lefschetz fibration over the annulus, viewed as a cobordism from Y_1 to Y_2, whose fiber F has genus $g > 1$. Then there is a unique* Spin^c *structure $\mathfrak{s}$ over W for which*

$$\langle c_1(\mathfrak{s}), [F] \rangle = 2 - 2g$$

and the induced map

$$F^+_{W,\mathfrak{s}} : HF^+(Y_1, \mathfrak{s}|_{Y_1}) \to HF^+(Y_2, \mathfrak{s}|_{Y_2})$$

is nontrivial. This is the canonical Spin^c *structure $\mathfrak{k}$, and its associated map is an isomorphism.*

Proof. Suppose that $\mathfrak{s} \in \text{Spin}^c(W)$ is as in the statement of the lemma, and induces a nontrivial map. We will show that $\mathfrak{s}$ is the canonical Spin^c structure $\mathfrak{k}$ on W, and the map is an isomorphism. By [29, Theorem 2.2] the fibration on Y_2 extends to a Lefschetz fibration on a 4-manifold W' over the disk $\pi' : W' \to D^2$, whose fiber is identified with F. Let $V = W \cup_{Y_2} W'$. Then V admits a Lefschetz fibration $\pi \cup \pi'$ over the disk.

The composition law for cobordism maps states that

$$F^+_{W'-B^4,\mathfrak{k}'} \circ F^+_{W,\mathfrak{s}} = \sum_{\{\mathfrak{t} \in \text{Spin}^c(W) \;|\; \mathfrak{t}|_W = \mathfrak{s},\ \mathfrak{t}|_{W'-B^4} = \mathfrak{k}'\}} F^+_{V-B^4,\mathfrak{t}}, \tag{1}$$

where $\mathfrak{k}'$ is the canonical Spin^c structure on W'. By [29, Theorem 5.3], the map

$$F^+_{W'-B^4,\mathfrak{k}'} : HF^+(Y_2, \mathfrak{k}'|_{Y_2}) \to HF^+(S^3)$$

is an isomorphism. Note that, according to [29, Theorem 5.2], there is a unique Spin^c structure on Y_2 whose Chern class evaluates on the class $[F]$ of the fiber to $2-2g$, and for which the Floer homology is nontrivial. It follows that if the map $F^+_{W,\mathfrak{s}}$ is nontrivial, as we've assumed, then the composite $F^+_{W'-B^4,\mathfrak{k}'} \circ F^+_{W,\mathfrak{s}}$ is also nontrivial, since the restrictions of $\mathfrak{s}$ and $\mathfrak{k}'$ to Y_2 must agree.

Now, since the Chern classes of the Spin^c structures $\mathfrak{t}'$ and $\mathfrak{s}$ evaluate to $2-2g$ on the class of the fiber $[F]$, the same is true for the Spin^c structures considered on V in the sum on the right-hand side of (1). Applying [29, Theorem 5.3] to the Lefschetz fibration on V implies that there is a unique nontrivial contribution to the sum, coming from the canonical Spin^c structure $\mathfrak{k}_V$ on V, and $F^+_{V-B^4,\mathfrak{k}_V}$ is an

isomorphism. Since $\mathfrak{k}_V$ restricts to the canonical Spin^c structure $\mathfrak{k}$ on W, it follows that $\mathfrak{s} = \mathfrak{k}$, and the corresponding map is an isomorphism. □

With this in hand, we modify the argument of [8, Lemma 2.11] to establish naturality with respect to a collection of Stein 2-handles. This boils down to another application of the composition law, together with a theorem of Eliashberg:

Lemma 3.7. *Suppose (W, J) is the cobordism associated to a collection of Stein 2-handle attachments. Then Theorem 3.1 is true for $F_{W^\dagger}$.*

Proof. As detailed above, we assume that all critical points of the plurisubharmonic Morse function on W have the same critical value, or equivalently that (W, J) is constructed by attaching a Stein 2-handle along each component of a Legendrian link L in (Y_1, ξ_1).

We may choose an open book decomposition adapted to ξ_1 such that the Legendrian link L sits naturally in a page. After positively stabilizing the open book we may assume that the pages have connected boundary and are of genus greater than one.

For $i \in \{1, 2\}$ let V_i denote the trace cobordism from Y_i to Y_i', the 3-manifold obtained by performing zero surgery along the binding of the open book. Surgery along L gives rise to the cobordism W from Y_1 to Y_2, and a cobordism W_0 from Y_1' to Y_2'. Note that both Y_1' and Y_2' are fibered 3-manifolds with fiber F obtained by capping off the boundary component of a page of the open book. W_0 admits a Lefschetz fibration over the annulus with fiber F.

Let $X = W \cup V_2 \cong V_1 \cup W_0$ denote the cobordism from Y_1 to Y_2'. Using [5, Theorem 1.1] we may extend the symplectic structure induced by the Lefschetz fibration on W_0 over the 2-handle cobordism V_1, giving a symplectic structure ω on X. The restriction of ω to W agrees with the symplectic structure on W induced by the Legendrian surgery along L; in particular, the canonical Spin^c structure $\mathfrak{k}_X \in \text{Spin}^c(X)$ of ω restricts to the canonical Spin^c-structure $\mathfrak{k}$ of (W, J).

Since V_1 can be obtained from surgery along a homologically nontrivial curve in Y_1', restriction induces an isomorphism $H^2(X, \mathbb{Z}) \to H^2(W_0, \mathbb{Z})$, so every Spin^c-structure on W_0 admits a unique extension over X. In particular, the extension of the canonical Spin^c-structure $\mathfrak{k}_0 \in \text{Spin}^c(W_0)$ is $\mathfrak{k}_X$. For $i \in \{1, 2\}$, let $\mathfrak{t}_i = \mathfrak{k}_0|_{Y_i'}$.

Ozsváth and Szabó [30] characterize the contact invariant $c^+(\xi_i) \in HF^+(-Y_i, \mathfrak{s}_{\xi_i})$ as the image of a class $c^+(\pi_i) \in HF^+(-Y_i', \mathfrak{t}_i)$ associated to the fibration under the map $F^+_{V_i^\dagger, \mathfrak{p}_i}$ where $\mathfrak{p}_i \in \text{Spin}^c(V_i)$ is the unique extension of $\mathfrak{t}_i$. Let $\mathfrak{s} \in \text{Spin}^c(W)$, then

$$F^+_{W^\dagger,\mathfrak{s}}(c^+(\xi_2)) = F^+_{W^\dagger,\mathfrak{s}} \circ F^+_{V_2^\dagger}(c^+(\pi_2)) = \sum_{\{\mathfrak{s}_X \in \text{Spin}^c(X) \;|\; \mathfrak{s}_X|_W = \mathfrak{s},\ \mathfrak{s}_X|_{V_2} = \mathfrak{p}_2\}} F^+_{X^\dagger,\mathfrak{s}_X}(c^+(\pi_2)),$$

where the last equality is given by the composition law for cobordism maps. Because

every Spin^c-structure on W_0 admits a unique extension over X, another application of the composition law shows that the above sum is equal to

$$\sum_{\{\mathfrak{s}_X \in \mathrm{Spin}^c(X) \;|\; \mathfrak{s}_X|_W = \mathfrak{s},\ \mathfrak{s}_X|_{V_2} = \mathfrak{p}_2\}} F^+_{V_1^\dagger, \mathfrak{s}_X|_{V_1}} \circ F^+_{W_0^\dagger, \mathfrak{s}_X|_{W_0}}(c^+(\pi_2)).$$

Note that $\langle c_1(\mathfrak{s}_X|_{W_0}), [F]\rangle = \langle c_1(\mathfrak{t}_2), [F]\rangle = 2 - 2g$. Lemma 3.6 implies that there is at most one nonzero contribution to this sum, coming from a term where $\mathfrak{s}_X|_{W_0}$ is the canonical Spin^c-structure $\mathfrak{k}_0$, in which case $\mathfrak{s}_X = \mathfrak{k}_X$ and $\mathfrak{s} = \mathfrak{k}$ by the preceding discussion. We have

$$F^+_{W^\dagger, \mathfrak{s}}(c^+(\xi_2)) = \begin{cases} F^+_{V_1^\dagger} \circ F^+_{W_0^\dagger, \mathfrak{k}_0}(c^+(\pi_2)) & \text{for } \mathfrak{s} = \mathfrak{k}, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, Lemma 3.6 also tells us that $F^+_{W_0^\dagger, \mathfrak{k}_0}$ is an isomorphism mapping $c^+(\pi_2)$ to $c^+(\pi_1)$, thus

$$F^+_{W^\dagger, \mathfrak{s}}(c^+(\xi_2)) = \begin{cases} F^+_{V_1^\dagger}(c^+(\pi_1)) = c^+(\xi_1) & \text{for } \mathfrak{s} = \mathfrak{k}, \\ 0 & \text{otherwise.} \end{cases}$$

This proves that the contact invariant in HF^+ satisfies the naturality claimed by Theorem 3.1 under the map induced by a Stein 2-handle cobordism. To establish the result for the contact invariant in $\widehat{HF}$, recall that $c^+(\xi)$ is defined as the image of $c(\xi)$ under the inclusion-induced map $\iota_* : \widehat{HF} \to HF^+$, and that both invariants can be characterized as the image of a particular class under the 2-handle cobordism which caps the fiber of the open book. In the case of the plus invariant, this is the distinguished class $c^+(\pi)$ associated to the fibration considered above, whereas for the hat invariant we consider the element $\widehat{c}(\pi)$ mapping to $c^+(\pi)$ under ι_*. The claimed naturality result for 2-handles now follows from naturality of ι_* with respect to the maps on Floer homology associated to cobordisms [31, Theorem 3.1, Remark 3.2]; cf. [34, Theorem A]. □

Having established the claimed naturality result for a Stein 0- or 1-handle and for a collection of Stein 2-handles, the theorem follows from the composition law for cobordism maps. □

Remark 3.8. Echoing Remark 2.4, one could alternatively approach Theorem 3.1 using the Honda–Kazez–Matić interpretation of the contact invariant. Using their Heegaard diagrams, the proof hinges on (a) showing that there exists a unique pseudoholomorphic triangle contributing to $F^+_{W^\dagger}(c(\xi_2))$ whose domain is a union of small triangles having corners at the components of a generator representing $c(\xi_1)$ and (b) identifying the Spin^c structure associated to this pseudoholomorphic triangle with the canonical one.

We conclude with the following immediate corollary of Theorem 3.1, which generalizes the main result of [33]:

Corollary 3.9 (cf. [33, Theorem 2]). *Let W be a smooth 4-manifold with boundary, equipped with two Stein structures J_1, J_2 with associated* Spin^c *structures $\mathfrak{s}_1$, $\mathfrak{s}_2$, and let ξ_1, ξ_2 be the induced contact structures on Y, the outgoing boundary of W. Suppose that the contact structure induced on the incoming boundary of W by J_1 has nonvanishing contact invariant. If the* Spin^c *structures $\mathfrak{s}_1$ and $\mathfrak{s}_2$ are not isomorphic, then the contact invariants $c(\xi_1)$, $c(\xi_2)$ are distinct elements of $\widehat{HF}(-Y)$.*

Acknowledgements

This article stemmed from the first author's work with Katherine Raoux, discussed at the BIRS workshop "Interactions of gauge theory with contact and symplectic topology in dimensions 3 and 4", and a particular application of that work which required naturality of the contact class under Stein cobordisms. It is our pleasure to thank Katherine, BIRS, and the workshop organizers for inspiring us to write this article, and Jamie Conway, Patrick Massot, Ian Zemke, and the referee for a number of helpful comments.

References

[1] K. Cieliebak and Y. Eliashberg, *From Stein to Weinstein and back*, American Mathematical Society Colloquium Publications **59**, Amer. Math. Soc., Providence, RI, 2012. Symplectic geometry of affine complex manifolds. MR Zbl

[2] M. Echeverria, "Naturality of the contact invariant in monopole Floer homology under strong symplectic cobordisms", *Algebr. Geom. Topol.* **20**:4 (2020), 1795–1875. MR Zbl

[3] S. Eilenberg and N. Steenrod, *Foundations of algebraic topology*, Princeton University Press, 1952. MR Zbl

[4] Y. Eliashberg, "Topological characterization of Stein manifolds of dimension > 2", *Internat. J. Math.* **1**:1 (1990), 29–46. MR Zbl

[5] Y. Eliashberg, "A few remarks about symplectic filling", *Geom. Topol.* **8** (2004), 277–293. MR Zbl

[6] J. B. Etnyre, "Lectures on open book decompositions and contact structures", pp. 103–141 in *Floer homology, gauge theory, and low-dimensional topology*, Clay Math. Proc. **5**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[7] M. Gartner, "Projective Naturality in Heegaard Floer Homology", 2019. arXiv 1908.06237

[8] P. Ghiggini, "Ozsváth–Szabó invariants and fillability of contact structures", *Math. Z.* **253**:1 (2006), 159–175. MR Zbl

[9] E. Giroux, "Géométrie de contact: de la dimension trois vers les dimensions supérieures", pp. 405–414 in *Proceedings of the International Congress of Mathematicians, Vol. II* (Beijing, 2002), Higher Ed. Press, Beijing, 2002. MR Zbl

[10] R. E. Gompf, "Handlebody construction of Stein surfaces", *Ann. of Math.* (2) **148**:2 (1998), 619–693. MR Zbl

[11] Y. Hatakeyama, "Some notes on the group of automorphisms of contact and symplectic structures", *Tohoku Math. J.* (2) **18** (1966), 338–347. MR Zbl

[12] M. Hedden, "An Ozsváth–Szabó Floer homology invariant of knots in a contact manifold", *Adv. Math.* **219**:1 (2008), 89–117. MR Zbl

[13] M. Hedden and K. Raoux, "Knot Floer homology and relative adjunction inequalities", 2020. arXiv 2009.05462

[14] K. Hendricks and C. Manolescu, "Involutive Heegaard Floer homology", *Duke Math. J.* **166**:7 (2017), 1211–1299. MR Zbl

[15] K. Hendricks, C. Manolescu, and I. Zemke, "A connected sum formula for involutive Heegaard Floer homology", *Selecta Math.* (*N.S.*) **24**:2 (2018), 1183–1245. MR Zbl

[16] K. Honda, W. H. Kazez, and G. Matić, "The contact invariant in sutured Floer homology", *Invent. Math.* **176**:3 (2009), 637–676. MR Zbl

[17] K. Honda, W. H. Kazez, and G. Matić, "On the contact class in Heegaard Floer homology", *J. Differential Geom.* **83**:2 (2009), 289–311. MR Zbl

[18] A. Juhász, "Cobordisms of sutured manifolds and the functoriality of link Floer homology", *Adv. Math.* **299** (2016), 940–1038. MR

[19] A. Juhász, D. Thurston, and I. Zemke, "Naturality and mapping class groups in Heegard Floer homology", pp. v+174 , 2021. MR Zbl

[20] E. L. Lima, "On the local triviality of the restriction map for embeddings", *Comment. Math. Helv.* **38** (1964), 163–164. MR Zbl

[21] P. Lisca and A. I. Stipsicz, "Ozsváth–Szabó invariants and tight contact three-manifolds, I", *Geom. Topol.* **8** (2004), 925–945. MR Zbl

[22] P. Massot, "Infinitely many universally tight torsion free contact structures with vanishing Ozsváth–Szabó contact invariants", *Math. Ann.* **353**:4 (2012), 1351–1376. MR Zbl

[23] J. Milnor, *Lectures on the h-cobordism theorem*, Princeton University Press, 1965. MR Zbl

[24] T. Mrowka and Y. Rollin, "Legendrian knots and monopoles", *Algebr. Geom. Topol.* **6** (2006), 1–69. MR Zbl

[25] P. Ozsváth and Z. Szabó, "Holomorphic disks and genus bounds", *Geom. Topol.* **8** (2004), 311–334. MR

[26] P. Ozsváth and Z. Szabó, "Holomorphic disks and knot invariants", *Adv. Math.* **186**:1 (2004), 58–116. MR

[27] P. Ozsváth and Z. Szabó, "Holomorphic disks and three-manifold invariants: properties and applications", *Ann. of Math.* (2) **159**:3 (2004), 1159–1245. MR

[28] P. Ozsváth and Z. Szabó, "Holomorphic disks and topological invariants for closed three-manifolds", *Ann. of Math.* (2) **159**:3 (2004), 1027–1158. MR

[29] P. Ozsváth and Z. Szabó, "Holomorphic triangle invariants and the topology of symplectic four-manifolds", *Duke Math. J.* **121**:1 (2004), 1–34. MR

[30] P. Ozsváth and Z. Szabó, "Heegaard Floer homology and contact structures", *Duke Math. J.* **129**:1 (2005), 39–61. MR

[31] P. Ozsváth and Z. Szabó, "Holomorphic triangles and invariants for smooth four-manifolds", *Adv. Math.* **202**:2 (2006), 326–400. MR

[32] R. S. Palais, "Local triviality of the restriction map for embeddings", *Comment. Math. Helv.* **34** (1960), 305–312. MR Zbl

[33] O. Plamenevskaya, "Contact structures with distinct Heegaard Floer invariants", *Math. Res. Lett.* **11**:4 (2004), 547–561. MR Zbl

[34] I. Zemke, "Graph cobordisms and Heegaard Floer homology", 2015. arXiv 1512.01184

[35] I. Zemke, "Duality and mapping tori in Heegaard Floer homology", 2018. Zbl arXiv 1801.09270

Received 1 Mar 2021. Revised 28 Oct 2021.

MATTHEW HEDDEN: mhedden@math.msu.edu
Department of Mathematics, Michigan State University, East Lansing, MI, United States

LEV TOVSTOPYAT-NELIP: tovstopy@msu.edu
Department of Mathematics, Michigan State University, East Lansing, MI, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.145

Dehn surgery and nonseparating two-spheres

Jennifer Hom and Tye Lidman

When can surgery on a nullhomologous knot K in a rational homology sphere produce a nonseparating sphere? We use Heegaard Floer homology to give sufficient conditions for K to be unknotted. We also discuss some applications to homology cobordism, concordance, and Mazur manifolds.

1. Introduction

One of the most fundamental constructions in three-manifold topology is Dehn surgery. By the theorems of Lickorish and Wallace, every closed, connected, oriented three-manifold is obtained by surgery on a link in S^3. Additionally, 4-dimensional 2-handle attachments induce a cobordism from a three-manifold to the result of surgery. It is therefore a fundamental question to understand the behavior of three-manifolds under Dehn surgery. In this note, we focus on surgery on knots. Two main questions are *geography* (which three-manifolds are obtained by surgery on a knot) and *botany* (which knots surger to a fixed three-manifold).

For example, Gabai's "Property R theorem" [6] shows that only 0-surgery on the unknot in S^3 can produce $S^2 \times S^1$. The proof passes through taut foliations, and as a result, shows that 0-surgery on a nontrivial knot is not $S^2 \times S^1$ *and* is prime (i.e., the 0-surgery is irreducible), giving strong geography constraints. Note that this implies that a four-manifold built with one 0-handle, one 1-handle, one 2-handle, and boundary S^3 is necessarily diffeomorphic to B^4. Similarly, Gordon and Luecke's celebrated "knot complement theorem" [10] answers the botany problem for surgeries from S^3 to S^3: only the unknot admits nontrivial S^3 surgeries. This shows that a closed four-manifold with one 0-handle, one 2-handle, and one 4-handle is necessarily diffeomorphic to $\mathbb{C}P^2$.

Hom was partially supported by NSF grant DMS-1552285. Lidman was partially supported by NSF grant DMS-1709702 and a Sloan Fellowship.
MSC2020: 57K10, 57K18, 57K30.
Keywords: Dehn surgery, Heegaard Floer homology.

In this article, we study a more general question: when can surgery on a knot in a three-manifold (other than S^3) produce an $S^2 \times S^1$ summand? In previous work of Daemi, Lidman, Vela-Vick, and Wong [3], some constraints were given on the geography problem. Here, we answer both the botany and geography problems in several different settings. While many of the arguments below are standard, we believe it is beneficial to the community for these results to be written down.

We begin with a generalization of Property R to arbitrary rational homology spheres.

Theorem 1.1. *Let Y be a rational homology sphere and K a nullhomologous knot in Y. Suppose $Y_0(K) = N\#S^2 \times S^1$. If $\dim \widehat{HF}(N) = \dim \widehat{HF}(Y)$, then $N = Y$ and K is unknotted. Otherwise, $\dim \widehat{HF}(N) < \dim \widehat{HF}(Y)$.*

Theorem 1.1 has a number of immediate applications.

Corollary 1.2. *Let K be a nullhomotopic knot in a prime rational homology sphere Y. If $Y_0(K)$ contains a nonseparating two-sphere, then K is unknotted.*

Proof. It is shown in [3, Theorem 1.8] that under these hypotheses, $Y_0(K) = Y\#S^2 \times S^1$. By Theorem 1.1, K is unknotted. □

Corollary 1.3. *Let Y be a rational homology sphere and let $W : Y \to Y$ be a rational homology cobordism with a handlebody decomposition with a total of two handles. Then, W is diffeomorphic to a product.*

Proof. Since W is a rational homology cobordism, after possibly flipping W upside down, W consists of a single 2-handle and a single 3-handle. Therefore, Y has a surgery to $Y\#S^2 \times S^1$. The result now follows from Theorem 1.1. □

Remark 1.4. It seems reasonable to conjecture that a rational homology cobordism from a 3-manifold to itself without 3-handles is homeomorphic to a product. It seems more ambitious, but still feasible, to believe that such a cobordism is diffeomorphic to a product.

Corollary 1.5. *Suppose that W is an integral homology cobordism from a rational homology sphere Y to a three-manifold Z consisting of a single 1-handle and a single 2-handle. If $\dim HF_{\mathrm{red}}(Z) = 1$, then W is diffeomorphic to a product.*

Proof. By [3, Theorem 1.19], $\dim HF_{\mathrm{red}}(Y) = 0$ or 1. If $\dim HF_{\mathrm{red}}(Y) = 1$, then $\dim \widehat{HF}(Y) = \dim \widehat{HF}(Z)$, since $\dim HF_{\mathrm{red}} = 1$ implies $\dim \widehat{HF} = |H_1| + 2$ and $|H_1(Y)| = |H_1(Z)|$. The result follows from Theorem 1.1 by applying the arguments in Corollary 1.3. (The fact that W is an integral homology cobordism implies that the relevant surgery is along a nullhomologous knot.) Next, suppose $\dim HF_{\mathrm{red}}(Y) = 0$. By the Spin^c-conjugation invariance of Heegaard Floer homology, we see that $\dim HF_{\mathrm{red}}(Z, \mathfrak{s}) = 1$ in a self-conjugate Spin^c-structure $\mathfrak{s}$. As shown by F. Lin in [15], this implies that his correction terms α, β, γ are not all equal for $\mathfrak{s}$. However,

for an L-space, they are all equal. This is a contradiction, since α, β, γ are preserved under integral homology cobordisms for each self-conjugate Spin^c structure. □

Note that the Brieskorn spheres $\Sigma(2,3,7)$ and $\Sigma(2,3,11)$ satisfy $\dim HF_{\mathrm{red}} = 1$.

Corollaries 1.3 and 1.5 can be seen as "manifold versions" of the following special case of a theorem of Gabai [7, Theorem 1]: a self-ribbon concordance with one minimum and one saddle is trivial. (This was explained to us by Maggie Miller.) In fact, one can recover a slight variant of this result using Theorem 1.1.

Corollary 1.6. *Let K be a nullhomologous knot in a rational homology sphere Y. Perform a band-sum with an unknot and denote the resulting knot by K'. Suppose K' is detected by its complement, which we additionally assume is irreducible and boundary irreducible. If $CFK^\infty(K) \cong CFK^\infty(K')$, then K' is isotopic to K and the exterior of the resulting concordance is smoothly the trivial cobordism.*

Note that if $Y = S^3$ and K is nontrivial, then the hypotheses apply for any K' by [8; 10]. For notation, we will write $E(X)$ to denote the exterior of the submanifold X. (The ambient manifold will be clear from context.)

Proof. Let $C : (Y, K) \to (Y, K')$ be the ribbon concordance in $Y \times I$ given by a single birth and saddle specified by the band-sum. Since K' is determined by its complement, it suffices to show that $E(C)$ is smoothly $E(K) \times I$.

Note that $E(C)$ is an integer homology cobordism from $E(K)$ to $E(K')$ which consists of a single 1-handle and 2-handle addition. Reversing orientation and flipping upside-down, we see that there exists a knot J knot in $E(K')$ with an $E(K)\#S^2 \times S^1$ surgery. Since K and K' are nullhomologous, we see that J is necessarily nullhomologous in $E(K')$. Note that if we can show that J is trivial, then $E(C) = E(K) \times I$ and we are done.

Write $J^{(n)}$ for the induced knot in $Y_n(K')$. Then, 0-surgery on $J^{(n)}$ results in $Y_n(K)\#S^2 \times S^1$. Since $CFK^\infty(K) \cong CFK^\infty(K')$, the large surgery formula of Ozsváth and Szabó [19, Theorem 4.4] implies $\dim \widehat{HF}(Y_n(K)) = \dim \widehat{HF}(Y_n(K'))$ for large n. Therefore, by Theorem 1.1, $J^{(n)}$ is unknotted in $Y_n(K')$ for large n. Since $Y_n(K')$ is not S^3 for large n, it follows that $E(J^{(n)}) = D^2 \times S^1 \# Y_n(K')$ is a reducible manifold for all large n.

In other words, $E(K' \cup J)$ has infinitely many reducible fillings. However, an irreducible, boundary-irreducible three-manifold with only toral boundary components has at most finitely many reducing fillings along a given boundary component (see, for example, [9]). Therefore, $E(K' \cup J)$ is either boundary reducible or reducible. Since K' is nontrivial, if $E(K' \cup J)$ is boundary reducible, then the toral boundary component coming from J must be the one that compresses, and we see that J must be unknotted in the exterior of K' completing the proof. On the other hand, if $E(K' \cup J)$ is reducible, then J must be contained in an embedded three-ball. In

this case, $E(K')_0(J) = E(K')\#S^3_0(J)$ and hence J is unknotted in the embedded three-ball. Again, J is trivial in $E(K')$ and we are done. □

Recently, Conway and Tosun [2] showed that the boundary of a nontrivial Mazur manifold is not an L-space. Ni has pointed out that an alternate proof follows from [17]. We now show how Theorem 1.1 gives another alternate proof of this fact. (Lidman and Pinzón-Caicedo have also proved the analogous result in instanton Floer homology.)

Corollary 1.7 [2, Theorem 1]. *Let $Y \neq S^3$ be a homology sphere bounding a Mazur manifold. Then Y is not an L-space.*

Proof. Suppose that Y is an L-space homology sphere which bounds a Mazur manifold. Then, there exists a knot K in Y such that $Y_0(K) = S^2 \times S^1$. Since $\dim \widehat{HF}(Y) = \dim \widehat{HF}(S^3)$, Theorem 1.1 implies that K is unknotted. Therefore, $Y_0(K) = Y\#S^2 \times S^1$, and we see that $Y = S^3$. □

We also present a symplectic analogue of Corollary 1.3. This was explained to the authors by Steven Sivek.

Corollary 1.8. *Let Y be a rational homology sphere. Let W be a Stein cobordism from (Y, ξ) to (Y, ξ') comprised of attaching single Weinstein 1- and 2-handles. If ξ' is tight, then W is deformation equivalent to the (compact) symplectization of (Y, ξ) and hence ξ and ξ' are contactomorphic contact structures.*

Proof. Consider the $(tb-1)$-framed 2-handle attachment to a Legendrian $\mathcal{K}$ in $(Y\#S^2 \times S^1, \xi\#\xi_{\text{std}})$ which results in (Y, ξ'). By reversing this picture, we see there is a Legendrian knot $\mathcal{K}'$ in (Y, ξ') with a contact $+1$-surgery to $(Y\#S^2 \times S^1, \xi\#\xi_{\text{std}})$ by [4, Proposition 8]. Note that $\mathcal{K}'$ must be nullhomologous and the framing of the surgery must be the Seifert framing in order to add a $\mathbb{Z}$-summand to H_1. Now, by Theorem 1.1, $\mathcal{K}'$ is unknotted topologically. Since $+1$-contact surgery means that the topological framing is one more than tb, we see that $tb = -1$. Because ξ' is tight, this implies $r = 0$ by [5, Theorem 1.6], and all such Legendrian unknots are Legendrian isotopic by [5, Theorem 1.5].

This implies that all Stein cobordisms from (Y, ξ'') to (Y, ξ') built out of single Weinstein 1- and 2-handles are equivalent, regardless of ξ''. However, we can produce such a cobordism by using a cancelling Weinstein 1- and 2-handle pair, i.e., the trivial cobordism. □

Finally, we give a new obstruction to a homology sphere admitting an $S^2 \times S^1$ surgery (and hence bounding a Mazur manifold).

Proposition 1.9. *Let K be a knot in a homology sphere Y with $HF_{\text{red},i}(Y) = \mathbb{F}$ for some i. Then $Y_0(K) \neq S^2 \times S^1$.*

Remark 1.10. It is easy to see that if a nullhomologous knot in a rational homology sphere admits a 0-surgery with an $S^2 \times S^1$ summand, then its Alexander polynomial is trivial (i.e., constant). We leave it as a fun exercise for the reader to deduce this fact using Heegaard Floer homology after reading the arguments in this paper.

Organization. The key idea in the proof of Theorem 1.1 comes from the special property of the twisted Heegaard Floer homology of three-manifolds with nonseparating S^2's. (This has been used in [16] and [17]; see also [11], [12], and [1].) In the next section, we review the mapping cone formula in Heegaard Floer homology, with extra attention to twisted coefficients, and prove Theorem 1.1. Lastly, we prove Proposition 1.9.

2. The mapping cone

We assume that the reader is familiar with the knot Floer chain complex of a knot CFK^∞, and the mapping cone formula for the Heegaard Floer homology of 0-surgery along a nullhomologous knot K in a rational homology sphere Y [22, Section 4.8]. We briefly recall the formula here, primarily to establish notation. Let $\mathfrak{t}$ denote a Spin^c structure on Y. As a vector space, we have that $C = CFK^\infty(Y, K, \mathfrak{t})$ decomposes as a direct sum $C = \bigoplus_{i,j\in\mathbb{Z}} C(i, j)$. For any set $X \subset \mathbb{Z}^2$ which is convex with respect to the product partial order on $\mathbb{Z}^2$ (i.e., if $a < b < c$ and $a, c \in X$, then $b \in X$), let CX denote $\bigoplus_{(i,j)\in X} C(i, j)$ which is naturally a subquotient complex of C.

Let B_s^+ (respectively, $\widehat{B}_s$) denote $C\{i \geq 0\}$ (respectively, $C\{i = 0\}$), and A_s^+ (respectively, $\widehat{A}_s$) denote $C\{\max(i, j - s) \geq 0\}$ (respectively $C\{\max(i, j - s) = 0\}$). Recall the maps $v_s^+, h_s^+ : A_s^+ \to B^+$ and $\widehat{v}_s, \widehat{h}_s : \widehat{A}_s \to \widehat{B}$. The main fact that we will need is that $\widehat{v}_s$ factors through $\widehat{v}_{s'}$ for $s' \geq s$.

Let $\widehat{F} \subset Y_0(K)$ denote the surface obtained by capping off an oriented Seifert surface F for K. As usual, we let $\mathfrak{t}_s$ denote the Spin^c structure on $Y_0(K)$ which satisfies $\langle c_1(\mathfrak{t}_s), [\widehat{F}]\rangle = 2s$ and such that $\mathfrak{t}_s$ extends $\mathfrak{t}$ over the 0-framed 2-handle cobordism from Y to $Y_0(K)$. In what follows, let $\circ$ denote either $+$ or $\widehat{\ }$.

Theorem 2.1 ([20, Theorem 9.19]; see also [22, Section 4.8]). *Let Y be a rational homology sphere and $K \subset Y$ a nullhomologous knot. With notation as above,*

$$HF^\circ(Y_0(K), \mathfrak{t}_s) \cong H_*(\mathrm{Cone}(v_s^\circ + h_s^\circ)).$$

There is a version of Theorem 2.1 with twisted coefficients, as in [20, Section 8]; see also [13, Section 2] and [14, Section 2]. Let T be a generator of $H^1(Y_0(K); \mathbb{Z})$. Consider the map

$$v_s^\circ + Th_s^\circ : A_s^\circ \otimes_{\mathbb{F}} \mathbb{F}[T, T^{-1}] \to B_s^\circ \otimes_{\mathbb{F}} \mathbb{F}[T, T^{-1}].$$

We have the following mapping cone formula with twisted coefficients. We write $HF^\circ(Y_0(K), \mathfrak{t}_s; \mathbb{F}[T, T^{-1}])$ to denote the Heegaard Floer homology with totally

twisted coefficients. We will also write $HF^\circ(Y_0(K), \mathfrak{t}_s; \mathbb{F}[\![T, T^{-1}]\!])$ to be the homology of the chain complex obtained by tensoring the twisted Heegaard Floer chain complex $CF^\circ(Y_0(K), \mathfrak{t}_s; \mathbb{F}[T, T^{-1}])$ with $\mathbb{F}[\![T, T^{-1}]\!]$ over $\mathbb{F}[T, T^{-1}]$.

Theorem 2.2 ([20, Theorem 9.23]; see also [14, Theorem 2.3]). *Let Y be a rational homology sphere and $K \subset Y$ a nullhomologous knot. With notation as above,*

$$HF^\circ(Y_0(K), \mathfrak{t}_s; \mathbb{F}[T, T^{-1}]) \cong H_*(\mathrm{Cone}(v_s^\circ + Th_s^\circ)).$$

We will be interested in the following consequence of the preceding theorem.

Corollary 2.3. *Let Y be a rational homology sphere and $K \subset Y$ nullhomologous. Then $HF^\circ(Y_0(K), \mathfrak{t}_s; \mathbb{F}[\![T, T^{-1}]\!])$ is isomorphic to the homology of the cone of*

$$v_s^\circ + Th_s^\circ : A_s^\circ \otimes_{\mathbb{F}} \mathbb{F}[\![T, T^{-1}]\!] \to B_s^\circ \otimes_{\mathbb{F}} \mathbb{F}[\![T, T^{-1}]\!].$$

Proof. The result follows from Theorem 2.2 and the fact that $\mathbb{F}[\![T, T^{-1}]\!]$ is flat over $\mathbb{F}[T, T^{-1}]$. □

We recall one key property of the Heegaard Floer homology of three-manifolds with nonseparating two-spheres. If M is a three-manifold which contains a nonseparating two-sphere S, then $HF^\circ(M; \mathbb{F}[\![T, T^{-1}]\!]) = 0$, where T denotes a generator of H^1 of the $S^2 \times S^1$ summand [16, Lemma 2.1]. Further, if $\mathfrak{s}$ is a Spin^c structure on M such that $\langle c_1(\mathfrak{s}), [S]\rangle = 0$, then $HF^\circ(M, \mathfrak{s}) \neq 0$ [20, Theorem 1.4]. With this, we analyze the mapping cone formula for knots which surger to three-manifolds with nonseparating two-spheres.

Proposition 2.4. *Let Y be a rational homology sphere and $K \subset Y$ a nullhomologous knot. Suppose that $Y_0(K) = N\#S^2 \times S^1$. Let $\circ = +$ or $\widehat{\ }$. Then $v_{s,*}^\circ + h_{s,*}^\circ : H_*(A_s^\circ) \to HF^\circ(Y, \mathfrak{t})$ is an isomorphism for all $s \neq 0$. Further, $v_{s,*}^\circ + Th_{s,*}^\circ : H_*(A_s^\circ) \otimes_{\mathbb{F}} \mathbb{F}[\![T, T^{-1}]\!] \to HF^\circ(Y, \mathfrak{t}) \otimes_{\mathbb{F}} \mathbb{F}[\![T, T^{-1}]\!]$ is an isomorphism for all s. In particular, $\dim H_*(\widehat{A}_s) = \dim \widehat{HF}(Y, \mathfrak{t})$ for all s.*

Proof. The first claim follows from Theorem 2.1 and that $Y_0(K)$ contains a nonseparating two-sphere.

Now, for the second claim, fix $\mathfrak{t}$ in Spin$^c(Y)$. Since $Y_0(K) = N\#S^2 \times S^1$, we have that $HF^+(Y_0(K), \mathfrak{t}_s; \mathbb{F}[\![T, T^{-1}]\!]) = 0$. By Corollary 2.3, we have that

$$HF^+(Y_0(K), \mathfrak{t}_s; \mathbb{F}[\![T, T^{-1}]\!]) \cong H_*(\mathrm{Cone}(v_s^+ + Th_s^+) \otimes_{\mathbb{F}[T,T^{-1}]} \mathbb{F}[\![T, T^{-1}]\!]).$$

Hence,

$$(v_s^+ + Th_s^+)_* : H_*(A_s^+ \otimes_{\mathbb{F}} \mathbb{F}[\![T, T^{-1}]\!]) \to H_*(B_s^+ \otimes_{\mathbb{F}} \mathbb{F}[\![T, T^{-1}]\!])$$

is an isomorphism of $\mathbb{F}[\![T, T^{-1}]\!]$-modules. The analogous result for the hat flavor follows immediately. □

Proof of Theorem 1.1. As before, fix $\mathfrak{t}$ in $\mathrm{Spin}^c(Y)$. Let $\mathfrak{t}'$ denote the Spin^c structure on N which is cobordant to $\mathfrak{t}$ under the homology cobordism from Y to N obtained by attaching a 3-handle to the trace of 0-surgery on K. Suppose that $\dim_{\mathbb{F}} \widehat{HF}(Y,\mathfrak{t}) \leq \dim_{\mathbb{F}} \widehat{HF}(N,\mathfrak{t}')$. We will show that equality holds and that K is the unknot. We have

$$\begin{aligned}
2\dim_{\mathbb{F}} \widehat{HF}(N,\mathfrak{t}') &= \dim_{\mathbb{F}}(\widehat{HF}(N\#S^2\times S^1, \mathfrak{t}'\#\mathfrak{s}_0)\\
&= \dim_{\mathbb{F}}(H_*(\mathrm{Cone}(\widehat{v}_0+\widehat{h}_0))\\
&= \dim_{\mathbb{F}} H_*(\widehat{A}_0) + \dim_{\mathbb{F}} \widehat{HF}(Y,\mathfrak{t}) - 2\,\mathrm{rk}(\widehat{v}_{0,*}+\widehat{h}_{0,*})\\
&= 2\dim_{\mathbb{F}} \widehat{HF}(Y,\mathfrak{t}) - 2\,\mathrm{rk}(\widehat{v}_{0,*}+\widehat{h}_{0,*})\\
&\leq 2\dim_{\mathbb{F}} \widehat{HF}(N,\mathfrak{t}') - 2\,\mathrm{rk}(\widehat{v}_{0,*}+\widehat{h}_{0,*}),
\end{aligned}$$

where the first equality follows from the Künneth formula, the second follows from Theorem 2.2, the third follows from rank-nullity (and the fact that we are working over a field), the fourth follows from Proposition 2.4, and the final inequality follows by hypothesis. Hence, we see that $\widehat{v}_{0,*} = \widehat{h}_{0,*}$. Therefore,

$$(1+T)\widehat{v}_{0,*} : H_*(\widehat{A}_0 \otimes_{\mathbb{F}} \mathbb{F}[\![T,T^{-1}]\!]) \to H_*(\widehat{B}_0 \otimes_{\mathbb{F}} \mathbb{F}[\![T,T^{-1}]\!])$$

is an isomorphism. This implies that $\widehat{v}_{0,*}$ is an isomorphism.

We now consider the case $s>0$. As mentioned above, $\widehat{v}_{0,*}$ factors through $\widehat{v}_{s,*}$. In particular, since $\widehat{v}_{0,*}$ is an isomorphism, we have that $\widehat{v}_{s,*}$ is surjective. By Proposition 2.4, $\dim H_*(\widehat{A}_s) = \dim \widehat{HF}(Y,\mathfrak{t})$, and therefore $\widehat{v}_{s,*}$ is an isomorphism. Since $\widehat{v}_{s,*}$ is an isomorphism if and only if $v^+_{s,*}$ is an isomorphism, it follows from [18, Theorem 1.2] (which holds for nullhomologous knots in arbitrary rational homology spheres) and [23, Proof of Lemma 8.1] that

$$g(K) = \min\{s \mid \widehat{v}_{i,*} \text{ is an isomorphism for all } i \geq s,\ \ \mathfrak{t}\in \mathrm{Spin}^c(Y)\} \leq 0,$$

which gives the desired result. □

Proof of Proposition 1.9. This is very similar to the proof of Theorem 1.1. After a possible orientation reversal, we may assume that $HF_{\mathrm{red},i}(Y) = \mathbb{F}$ and i is odd. By Proposition 2.4, $H_i(A_0^+) = \mathbb{F}$, and

$$v^+_{0,*} + Th^+_{0,*} : H_i(A_0^+) \otimes_{\mathbb{F}} \mathbb{F}[\![T,T^{-1}]\!] \to H_i(B_0^+) \otimes_{\mathbb{F}} \mathbb{F}[\![T,T^{-1}]\!]$$

is an isomorphism. (Here, we are using the fact that v_0^+ and h_0^+ are homogeneous of the same grading shift. This is not true for $s\neq 0$.) Restricted to this grading, this latter map can be written as $v_0^+ + Th_0^+ : \mathbb{F}[\![T,T^{-1}]\!] \to \mathbb{F}[\![T,T^{-1}]\!]$. It follows that either v_0^+ or h_0^+ must be nonzero as a map from $H_i(A_0^+)=\mathbb{F}$ to $H_i(B_0^+)=\mathbb{F}$. By conjugation invariance [21, Theorem 3.6], we have that v_0^+ is nonzero if and only if h_0^+ is nonzero, and so they must be equal. Therefore, $v^+_{0,*} = h^+_{0,*}$ as maps

from $H_i(A_0^+)$ to $H_i(B_0^+)$, and we see that the kernel of $v_{0,*}^+ + h_{0,*}^+$ contains an $\mathbb{F}$ in grading i, which is odd.

Consider the homology of the cone of $v_{0,*}^+ + h_{0,*}^+ : H_*(A_0^+) \to H_*(B_0^+)$. This has two towers: one from the kernel of $v_{0,*}^+ + h_{0,*}^+$ and one from the cokernel. We also know there is an additional generator in the kernel of $v_{0,*}^+ + h_{0,*}^+$ in degree i; this is in opposite parity of the tower found in this kernel. Consider the long exact sequence associated to a mapping cone

$$\cdots HF^+(S^2 \times S^1) \to H_*(A_0^+) \to H_*(B_0^+) \to \cdots .$$

A nontrivial element of the kernel of $v_{0,*}^+ + h_{0,*}^+$ in degree i would have to be in the image of U^n for all n, but that is ruled out by the parity of the grading. Hence, we have a contradiction. □

Acknowledgements

The authors thank Maggie Miller for helpful conversations and Steven Sivek for describing the proof of Corollary 1.8. They also thank Matt Hedden and Yi Ni for helpful comments on an earlier draft of this paper. Finally, they thank the referee for their helpful feedback.

References

[1] A. Alishahi and R. Lipshitz, "Bordered Floer homology and incompressible surfaces", *Ann. Inst. Fourier* (*Grenoble*) **69**:4 (2019), 1525–1573. MR Zbl

[2] J. Conway and B. Tosun, "Mazur-type manifolds with L-space boundary", *Math. Res. Lett.* **27**:1 (2020), 35–42. MR

[3] A. Daemi, T. Lidman, D. S. Vela-Vick, and C. M. M. Wong, "Ribbon homology cobordisms", preprint, 2019. arXiv 1904.09721

[4] F. Ding and H. Geiges, "Symplectic fillability of tight contact structures on torus bundles", *Algebr. Geom. Topol.* **1** (2001), 153–172. MR

[5] Y. Eliashberg and M. Fraser, "Topologically trivial Legendrian knots", *J. Symplectic Geom.* **7**:2 (2009), 77–127. MR Zbl

[6] D. Gabai, "Foliations and the topology of 3-manifolds, III", *J. Differential Geom.* **26**:3 (1987), 479–536. MR Zbl

[7] D. Gabai, "Genus is superadditive under band connected sum", *Topology* **26**:2 (1987), 209–210. MR Zbl

[8] C. M. Gordon, "Ribbon concordance of knots in the 3-sphere", *Math. Ann.* **257**:2 (1981), 157–170. MR Zbl

[9] C. M. Gordon, "Dehn filling: a survey", pp. 129–144 in *Knot theory* ((Warsaw, 1995)), Banach Center Publ. **42**, Polish Acad. Sci. Inst. Math., Warsaw, 1998. MR Zbl

[10] C. M. Gordon and J. Luecke, "Knots are determined by their complements", *J. Amer. Math. Soc.* **2**:2 (1989), 371–415. MR Zbl

[11] M. Hedden and Y. Ni, "Manifolds with small Heegaard Floer ranks", *Geom. Topol.* **14**:3 (2010), 1479–1501. MR Zbl

[12] M. Hedden and Y. Ni, "Khovanov module and the detection of unlinks", *Geom. Topol.* **17**:5 (2013), 3027–3076. MR Zbl

[13] S. Jabuka and T. E. Mark, "Product formulae for Ozsváth–Szabó 4-manifold invariants", *Geom. Topol.* **12**:3 (2008), 1557–1651. MR Zbl

[14] A. S. Levine and D. Ruberman, "Heegaard Floer invariants in codimension one", *Trans. Amer. Math. Soc.* **371**:5 (2019), 3049–3081. MR Zbl

[15] F. Lin, "Indefinite Stein fillings and PIN(2)-monopole Floer homology", *Selecta Math.* (*N.S.*) **26**:2 (2020), art. id. 18. MR Zbl

[16] Y. Ni, "Heegaard Floer homology and fibred 3-manifolds", *Amer. J. Math.* **131**:4 (2009), 1047–1063. MR Zbl

[17] Y. Ni, "Nonseparating spheres and twisted Heegaard Floer homology", *Algebr. Geom. Topol.* **13**:2 (2013), 1143–1159. MR Zbl

[18] P. Ozsváth and Z. Szabó, "Holomorphic disks and genus bounds", *Geom. Topol.* **8** (2004), 311–334. MR

[19] P. Ozsváth and Z. Szabó, "Holomorphic disks and knot invariants", *Adv. Math.* **186**:1 (2004), 58–116. MR

[20] P. Ozsváth and Z. Szabó, "Holomorphic disks and three-manifold invariants: properties and applications", *Ann. of Math.* (2) **159**:3 (2004), 1159–1245. MR

[21] P. Ozsváth and Z. Szabó, "Holomorphic triangles and invariants for smooth four-manifolds", *Adv. Math.* **202**:2 (2006), 326–400. MR

[22] P. S. Ozsváth and Z. Szabó, "Knot Floer homology and integer surgeries", *Algebr. Geom. Topol.* **8**:1 (2008), 101–153. MR

[23] P. S. Ozsváth and Z. Szabó, "Knot Floer homology and rational surgeries", *Algebr. Geom. Topol.* **11**:1 (2011), 1–68. MR

Received 4 Nov 2020. Revised 23 Mar 2021.

JENNIFER HOM: hom@math.gatech.edu
School of Mathematics, Georgia Institute of Technology, Atlanta, GA, United States

TYE LIDMAN: tlid@math.ncsu.edu
Department of Mathematics, North Carolina State University, Raleigh, NC, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.155

Broken Lefschetz fibrations, branched coverings, and braided surfaces

Mark C. Hughes

We discuss an important class of fibrations on smooth 4-manifolds, called broken Lefschetz fibrations. We outline their connection to symplectic and near-symplectic structures, describe their topology, and discuss several approaches to their construction. We focus on new techniques involving branched coverings and braided surfaces with folds, and provide explicit examples of fibrations constructed using these approaches.

1. Fibrations on 4-manifolds

Fibrations on smooth manifolds have played an important role in the development of low-dimensional topology. These fibrations show up naturally from the viewpoint of algebraic geometry, but have broad generalizations that extend outside of their algebrogeometric origins. They provide very useful topological frameworks to study geometric objects, like contact, symplectic, and Stein manifolds. Furthermore, they can be used to describe 3- and 4-dimensional manifolds in terms of diffeomorphism groups of surfaces, a viewpoint which can be especially fruitful.

In this paper we discuss Lefschetz fibrations and broken Lefschetz fibrations, and survey the main results on these structures. After defining them and describing their connection to symplectic and near-symplectic structures, we will outline several important constructions and provide examples. These examples focus on explicit constructions using branched coverings and braided surfaces. Although these techniques can often be used to construct explicit broken Lefschetz fibrations on 4-manifolds directly from a given handle decomposition, they rely on the construction of a certain branched covering with orientable branch locus and prescribed boundary, which cannot always be achieved.

MSC2020: primary 57K40, 57K43; secondary 57R35.
Keywords: Broken Lefschetz fibrations, 4-manifolds, braided surfaces.

This first section contains definitions of the various fibration structures that we will be concerned with on 3- and 4-manifolds, as well as descriptions of their topology.

1A. ***Singular fibrations on 4-manifolds.*** Let X be a smooth, compact, connected, oriented 4-manifold, Σ be a compact surface, and let $f : X \to \Sigma$ be a smooth map. A critical point p of f is called a *positive Lefschetz critical point* if there are orientation-preserving local complex coordinates about p on which $f : \mathbb{C}^2 \to \mathbb{C}$ is modeled as $f(u, v) = u^2 + v^2$. If the coordinates around the critical point are instead orientation-reversing, then it is called a *negative Lefschetz critical point*. We will often omit the adjective *positive*, and refer to a positive Lefschetz critical point simply as a *Lefschetz critical point*.

An embedded circle $C \subset X$ of critical points of f is called an *indefinite fold singularity* if f is modeled near points of C by the map

$$(\theta, x, y, z) \mapsto (\theta, x^2 + y^2 - z^2)$$

from $\mathbb{R} \times \mathbb{R}^3 \to \mathbb{R} \times \mathbb{R}$, where C is given locally by $x = y = z = 0$. Indefinite fold singularities are sometimes referred to as *round* 1*-handle singularities* or *broken singularities* in the literature.

A surjective map $f : X \to \Sigma$ is called a *Lefschetz fibration* if all critical points of f are in the interior of X and are positive Lefschetz critical points. It is called an *achiral Lefschetz fibration* if we also allow negative Lefschetz critical points. Finally, we add the adjective *broken* to either of these names to indicate that we also allow indefinite fold singularities in the set of critical points of f. When discussing these maps we will sometimes use the generic term *fibration* to describe a map which can be any of the types defined above.

1B. ***Boundary behavior of fibrations.*** Let M be a 3-dimensional closed smooth oriented manifold. An *open book decomposition* on M is a smooth map $\lambda : M \to D^2$ such that $\lambda^{-1}(\partial D^2)$ is a compact 3-dimensional submanifold on which λ restricts as a surface bundle over $S^1 = \partial D^2$. Furthermore, we require that the closure of $\lambda^{-1}(\text{int } D^2)$ be the disjoint union of solid tori, on which λ is the projection $D^2 \times S^1 \to D^2$. We say that $\lambda^{-1}(0)$ is the *binding* of the open book on M, and for any $p \in S^1$ the compact surface $\Sigma_p = \lambda^{-1}(\{\alpha p \mid 0 \leq \alpha \leq 1\})$ is the *page* over p. The surface bundle structure on $\lambda^{-1}(\partial D^2)$ induces a monodromy map on the pages of λ.

By a celebrated theorem of Giroux [19], open book decompositions on a closed 3-manifold M (up to a stabilization operation) are in one-to-one correspondence with contact structures on M (up to isotopy). Thus open book decompositions provide a useful topological setting in which to study contact structures on a given closed 3-manifold.

Now suppose that X is a smooth 4-manifold and Σ is a compact surface, that $\partial X \neq \varnothing$ is connected, and that $f : X \to \Sigma$ is a fibration. Then we say that f is *convex*, if

- $\Sigma = D^2$,
- $f(\partial X) = D^2$, and
- $f|_{\partial X} : \partial X \to D^2$ is an open book decomposition on ∂X.

We say that f is *concave* if there is a disk $D \subset \operatorname{int} \Sigma$ such that

- $f(\partial X) = D$, and
- $f|_{\partial X} : \partial X \to D$ is an open book decomposition on ∂X.

Finally, f is said to be *flat* if

- $f(\partial X) = \partial \Sigma$, and
- $f|_{\partial X} : \partial X \to \partial \Sigma$ is a nonsingular fiber bundle.

The fibers of a flat fibration are all closed surfaces, and the boundary ∂X consists of the fibers above $\partial \Sigma$. The fibers of a convex fibration all have boundary, and ∂X is comprised of the fibers above $\partial \Sigma = \partial D^2$, along with the boundaries of the fibers above int D^2. In contrast, concave fibrations will have both closed fibers and fibers with boundary. Indeed, the fibers above int $D \subset \Sigma$ will have boundary, while all other fibers will be closed.

Suppose now that $f_1 : X_1 \to \Sigma$ is a concave fibration, $f_2 : X_2 \to D^2$ is a convex fibration, and that there is an orientation-reversing diffeomorphism $\phi : \partial X_1 \to \partial X_2$ which respects the open book decompositions. Then f_1 and f_2 can be glued together, to give a fibration $f : X_1 \cup_\phi X_2 \to \Sigma$. This gives a very useful method for constructing fibrations on closed 4-manifolds. Indeed, one effective strategy is to divide the closed manifold X into simpler pieces X_1 and X_2, on which convex and concave fibrations can be constructed. In general these maps will induce different open book decompositions along their common boundary. If, however, these fibrations can be modified so that they agree along $\partial X_1 = \partial X_2$, then they can be glued to give a fibration on all of X. See [1; 17; 18] for approaches to matching these boundary fibrations which make use of Giroux's theorem and Eliashberg's classification of overtwisted contact structures.

1C. ***Monodromy around Lefschetz critical points.*** The regular fibers of a flat or convex (achiral) Lefschetz fibration $f : X \to \Sigma$ will all be surfaces of the same diffeomorphism type, which we call the *genus of* f. Lefschetz fibrations of genus $g \geq 2$ can be determined entirely by their *monodromy representations*. Let $\Sigma^* \subset \Sigma$ denote the set of regular values of f, and let $p \in \Sigma \setminus \Sigma^*$ be a critical value. If $\gamma \subset \Sigma^*$ is an oriented loop based at $q \in \Sigma^*$ which travels counterclockwise

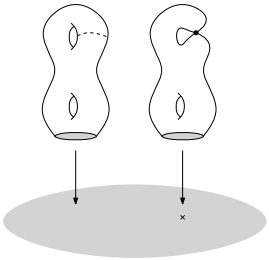


Figure 1. Vanishing cycle of Lefschetz critical point.

around p and no other critical values, then a trivialization of the bundle $f^{-1}(\gamma)$ over γ induces a diffeomorphism of the fiber F_q above q. This diffeomorphism will be a positive (negative) Dehn twist if p corresponds to a positive Lefschetz critical point (respectively, negative Lefschetz critical point). The cycle along which this Dehn twist takes place is called the *vanishing cycle* associated to the critical point. As we approach the critical fiber F_p, the corresponding vanishing cycles in nearby regular fibers shrink down to a single transverse intersection in F_p (see Figure 1 where the vanishing cycle is denoted with a dashed line).

The monodromy of a regular fiber provides a useful way to describe Lefschetz fibrations. More precisely, suppose that $f : X \to \Sigma$ is a Lefschetz fibration that has m critical points and that f is injective on the set of critical points. Suppose further that Σ is either S^2 or D^2, and hence Σ^* is an m-times punctured sphere or disk. Fix a basepoint $q \in \Sigma^*$, and a collection of oriented simple closed curves $\gamma_1, \ldots, \gamma_m$ based at q, which are disjoint away from q, where γ_j travels counterclockwise around the j-th puncture of Σ^* and no other punctures. Note that the loops $\gamma_1, \ldots, \gamma_m$ then generate $\pi_1(\Sigma^*; q)$, and that we can order them so that the product $\gamma_1 \cdots \gamma_m$ is null-homotopic when $\Sigma = S^2$, and homotopic to $\partial\Sigma$ when $\Sigma = D^2$. Finally, let $F_q = f^{-1}(q)$ be the fiber above $q \in \Sigma^*$, and let $\mathcal{M}(F_q)$ denote the mapping class group of F_q (i.e., the group of orientation-preserving diffeomorphisms of F_q fixing the boundary pointwise, mod isotopy rel boundary).

Then to each loop γ_j we can associate an element $\varphi_j \in \mathcal{M}(F_q)$, which is represented by a positive Dehn twist along the corresponding vanishing cycle. In the case when $\Sigma = S^2$ these elements must additionally satisfy $\varphi_1\varphi_2 \cdots \varphi_m = 1 \in \mathcal{M}(F_q)$, since the product of the loops $\gamma_1, \ldots, \gamma_m$ is trivial in $\pi_1(\Sigma^*; q)$. The fibration f then determines a homomorphism $\Lambda : \pi_1(\Sigma^*; q) \to \mathcal{M}(F_q)$, called the *monodromy representation* of f. Note that Λ is only determined by f up to conjugation by a fixed element in $\mathcal{M}(F_q)$ along with changes in the set of the generating loops γ_j. Conversely, given a set of generating loops $\gamma_1, \ldots, \gamma_m$ as above, and a collection of positive Dehn twists $\tau_1, \ldots, \tau_m \in \mathcal{M}(F_q)$, we can construct a Lefschetz fibration $f : X \to \Sigma$ whose monodromy representation satisfies $\Lambda(\gamma_j) = \tau_j$ for each j (in the case when $\Sigma = S^2$ we must additionally require that $\tau_1 \cdots \tau_m = 1 \in \mathcal{M}(F_q)$).

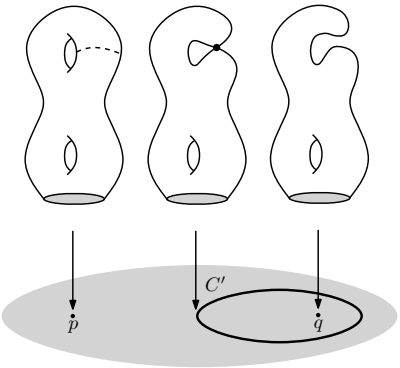


Figure 2. Passing an indefinite fold singularity.

Thus problems involving Lefschetz fibrations can be reformulated and successfully studied in terms of factorizations of mapping class group elements.

This monodromy description of a Lefschetz fibration on X can be adapted further to encode information about embedded surfaces in X. By selecting points $s_1, \ldots, s_k \in F_q$, we can instead consider the group $\mathcal{M}(F_q; \{s_1, \ldots, s_k\})$ of isotopy classes of orientation-preserving diffeomorphisms of F_q which fix the boundary pointwise and preserve $\{s_1, \ldots, s_k\}$ setwise. If we think of our monodromy representation as taking values in $\mathcal{M}(F_q; \{s_1, \ldots, s_k\})$, the trace of the marked points under the monodromy can be completed to an embedded surface in X, called a *multisection* of the fibration. See Baykur and Hayano's work in [7] or [8] for more details.

1D. ***The topology of broken Lefschetz fibrations.*** Suppose now that $f : X \to \Sigma$ is a broken fibration, with indefinite fold singularity along an embedded circle C. Suppose that $C' \subset \Sigma$ is the image of C under f, and that C' is embedded. Let p and q be nearby regular points sitting on opposite sides of C'. Suppose for concreteness that $p = (\theta, -1)$ and $q = (\theta, 1)$ for some $\theta \in S^1$ in the coordinate charts described above. Then the fiber F_q above q can be obtained from F_p by 0-surgery along a pair of points in F_p. Equivalently, F_p can be obtained from F_q by 1-surgery along a simple closed curve (see Figure 2). Indeed, we can think of the coordinate charts describing the indefinite fold singularity as defining an S^1-family of local Morse functions, each with a single index 1 critical point. In particular, for a broken fibration with connected fiber, the genus of the fiber changes by ± 1 each time we cross the image of an indefinite fold singularity in Σ.

Now suppose that $f : X \to D^2$ is a Lefschetz fibration, possibly achiral, possibly broken. Let K be a framed knot in $f^{-1}(\partial D^2) \subset \partial X$, which can be isotoped so that it lies entirely on the interior of a single fiber. Then we can attach a 2-handle along K to yield a new manifold with boundary which we denote X'. If we choose the framing along K so that it is one less than the induced fiber framing, then f will

extend to a fibration on X' with a new Lefschetz critical point in the newly added 2-handle. If we instead choose K to have framing one greater than the induced fiber framing, f will instead extend to a fibration on X' with an additional negative Lefschetz critical point (see, e.g., [17]).

Suppose again that $f : X \to D^2$ is a fibration as above, but that we have now chosen two disjoint knots K_1 and K_2 in ∂X, each of which gives a section of f restricted to $f^{-1}(\partial D^2) \subset \partial X$. Then we obtain a new manifold X'' by attaching $S^1 \times D^1 \times D^2$ to ∂X along K_1 and K_2, by identifying $S^1 \times \{-1\} \times D^2$ and $S^1 \times \{1\} \times D^2$ with tubular neighborhoods of K_1 and K_2, respectively. In this case the fibration f will extend to X'', with a single indefinite fold singularity along $S^1 \times \{0\} \times \{0\}$. Indeed, the knots K_1 and K_2 intersect each of the boundary fibers in a pair of points, which specify the locations of the 0-surgeries that take place as we pass the indefinite fold image. We will sometimes refer to this procedure as attaching a *round* 1*-handle* to X, as $S^1 \times D^1 \times D^2$ can be thought of as an S^1-family of 3-dimensional 1-handles $D^1 \times D^2$, which are attached to X fiberwise along the boundary. Alternatively, we can split $S^1 \times D^1 \times D^2$ into a 4-dimensional 1-handle and 2-handle pair, where the 2-handle runs over the 1-handle twice geometrically, but zero times algebraically.

The effect on X of adding a round 1-handle is the same as gluing a fibered cobordism to ∂X, where each fiber over S^1 is the standard Morse theoretic cobordism obtained by adding a 3-dimensional 1-handle to a thickened surface. Broken Lefschetz fibrations and round 1-handle attachments are studied in detail by Baykur in [4], where he also defines generalized n-dimensional round j-handles, for any index j in any dimension n. In what follows we will sometimes find it convenient to refer to 4-dimensional round 2-handles, which are the product of a 3-dimensional 2-handle with S^1 (these are, of course, just upside-down round 1-handles, and will not warrant any further discussion here).

As in the case of Lefschetz critical points, we also obtain monodromy descriptions of the indefinite fold singularities. The monodromy of the fibration outside a new indefinite fold singularity will depend on the framings of the tubular neighborhoods of K_1 and K_2, or alternatively, on the framing k of the 2-handle in the 4-dimensional handle pair description. Indeed, suppose that F is a fiber of the fibration f before attaching the round 1-handle, and that the monodromy around the boundary ∂D^2 is given by a map $\varphi : F \to F$. Then adding the new round 1-handle changes the fibers along the boundary by replacing two disks D_1 and D_2 in F with $S^1 \times [0, 1]$. The new monodromy will be given by the restriction of φ to $F \backslash (D_1 \cup D_2)$, with $-k$ Dehn twists along the cycle $S^1 \times \left\{\frac{1}{2}\right\}$ (i.e., $|k|$ positive Dehn twists if k is negative, and $|k|$ negative Dehn twists if k is positive).

Combining the above monodromy descriptions of indefinite fold singularities with those of Lefschetz critical points gives monodromy representations of broken

Lefschetz fibrations. Returning to the notation from Section 1C, suppose that $f : X \to \Sigma$ has a single indefinite fold singularity along a loop C, that $f(C)$ is embedded in Σ, and that the basepoint q and the images of each Lefschetz critical point are on the side of $f(C)$ with higher genus fibers (or lower Euler characteristic, in the case of disconnected fibers). Suppose that in addition to the loops $\gamma_1, \ldots, \gamma_m$ in Σ^*, we have also selected an embedded arc γ in Σ from the basepoint q to the image of the round 1-handle singularity, which is disjoint from the loops γ_i away from q. Then given $\gamma, \gamma_1, \ldots, \gamma_m$, the manifold X can be reconstructed from the mapping class group elements $\varphi_1, \ldots, \varphi_m \in \mathcal{M}(F_q)$, together with a loop in F_q specifying the location of the 1-surgery that corresponds to crossing the image of the indefinite fold singularity from the high genus side to the low genus side along γ. Such monodromy descriptions of *simplified broken Lefschetz fibrations* are studied in detail by Baykur and Hayano [6].

1E. ***Symplectic and near-symplectic structures.*** Lefschetz fibrations are of great interest in 4-manifold topology, in large part due to theorems of Donaldson [14] and Gompf [20] relating them to symplectic 4-manifolds. A *symplectic form* on a smooth oriented 4-manifold X is a closed, nondegenerate 2-form ω, whose wedge product square $\omega \wedge \omega$ is a volume form inducing the given orientation on X. A *symplectic manifold* is a manifold equipped with a symplectic form.

Donaldson proved that any symplectic 4-manifold admits a *Lefschetz pencil*. That is, there is a finite set of points $B \subset X$ and a smooth map $F : X \backslash B \to \mathbb{CP}^1$ which is a Lefschetz fibration, and around each point of B the map F is locally modeled by the projectivization map $\mathbb{C}^2 \backslash \{0\} \to \mathbb{CP}^1$. Blowing up at the points in B gives an honest Lefschetz fibration; thus Donaldson's result can be restated by saying that any symplectic 4-manifold admits a Lefschetz fibration after blow-ups. Gompf proved the converse to this, by showing that any manifold which admits a Lefschetz pencil also admits a symplectic structure.

A similar relationship exists between broken Lefschetz fibrations and *near-symplectic structures*. Let ω be a smooth closed 2-form with $\omega^2 \geq 0$, and set $Z = \{\omega = 0\}$. Then ω is called a *near-symplectic structure* on X if $\omega^2 > 0$ on the complement of Z, and for each point in Z there is a neighborhood U such that the map $U \to \Lambda^2(T^*U)$ induced by ω has rank 3. This implies that the zero locus Z is a family of embedded circles. Manifolds admitting near-symplectic structures are quite common. Indeed, any closed oriented smooth 4-manifold with $b_2^+(X) > 0$ admits a near-symplectic structure (see [22]).

Analogous to the relationship between Lefschetz pencils and symplectic structures, Auroux, Donaldson, and Katzarkov [2] proved the following: a smooth 4-manifold X admits a near-symplectic structure with zero locus Z if and only if it admits a broken Lefschetz pencil f with indefinite fold singularities along Z, and there is a class $\omega \in H^2(X)$ that evaluates positively on every component of every

fiber of f. Here, a broken Lefschetz pencil on X is a finite set of points $B \subset X$, and a smooth map $F : X \setminus B \to \mathbb{CP}^1$ which is a broken Lefschetz fibration, and around each point of B the map F is locally modeled by the projectivization map as above. These structures can be chosen to be compatible, in the sense that if we specify either a near-symplectic structure or broken Lefschetz pencil, then the other object may be chosen so that the regular fibers of the pencil are symplectic away from the singular locus.

Broken Lefschetz fibrations and near-symplectic structures gained a great deal of attention following [2], due in part to constructions of new Floer theoretic invariants and a conjectured relationship to gauge theory and the Seiberg–Witten invariants of 4-manifolds. In [34; 35], Perutz defines and studies the *Lagrangian matching invariant*, which counts pseudoholomorphic multisections of a broken Lefschetz fibration, subject to certain Lagrangian boundary conditions (alternatively, these can be thought of as pseudoholomorphic sections of an associated family of symmetric products of the nonsingular fibers). The Lagrangian matching invariant is a near-symplectic generalization of the Donaldson–Smith invariants defined on symplectic Lefschetz fibrations, which were shown by Usher to be equivalent to the Seiberg–Witten invariants of the underlying 4-manifold for fibrations of high degree [42]. Similarly, the Lagrangian matching invariant can also be compared to the Seiberg–Witten invariants, and Perutz conjectures these invariants are in fact equivalent. The relationship between broken Lefschetz fibrations and Seiberg–Witten invariants is studied further by Baykur in [4], where he discusses vanishing results for Seiberg–Witten invariants under a near-symplectic fiber sum operation and presents numerous examples.

Broken Lefschetz fibrations have also been used to define invariants outside of the Floer and gauge theoretic worlds. In [5], Baykur defines the *broken genera* of an oriented 4-manifold X, which are diffeomorphism invariants constructed using a family of simplified broken Lefschetz fibrations on X. These are defined in terms of the minimal genus of a regular fiber among all simplified broken Lefschetz fibration on X (or an associated blow-up of X), whose fiber realizes a certain homology class in $H_2(X; \mathbb{Z})$. In addition to defining these invariants, Baykur shows that these invariants are able to distinguish infinitely many exotic smooth structures among manifolds of the same homeomorphism type.

1F. ***Existence of fibrations on closed 4-manifolds.*** Besides establishing a relationship between near-symplectic structures and broken Lefschetz fibrations, Auroux, Donaldson, and Katzarkov also constructed a fibration on S^4 with a single indefinite fold singularity, and no other critical points. As S^4 is clearly not near-symplectic, this raised the question of determining which smooth oriented 4-manifolds admit broken Lefschetz fibrations.

This question, and related ones, were answered in stages by several authors. In [17], Etnyre and Fuller proved that after surgery along an embedded circle

every smooth closed 4-manifold admits an achiral Lefschetz fibration. Gay and Kirby proved in [18] that every smooth closed 4-manifold admits a broken achiral Lefschetz fibration. Building on the work of Saeki in [41], Baykur used singularity theory to prove that all closed orientable smooth 4-manifolds admit broken Lefschetz fibrations in [3]. Moreover, Akbulut and Karakurt [1], Baykur [30, Appendix B], and Lekili [30] demonstrated that the negative Lefschetz singularities in Gay and Kirby's construction can be eliminated, and hence provided an alternate proof that every closed oriented smooth 4-manifold admits a broken Lefschetz fibration.

The constructions in [1; 17; 18] each involve cutting X up into pieces and constructing the desired fibrations on the pieces separately as described in Section 1B, before regluing. The main differences lie in the modifications that are made to the fibrations to match the boundary open book decompositions. In either approach however, the core argument is the same, relying on machinery from contact topology to ensure that the open book decompositions match along the boundaries before the pieces are reglued. More precisely, the fibrations are first modified to ensure that both boundary open book decompositions support overtwisted contact structures, and then to arrange that both of these contact structures are homotopic. By Eliashberg's classification of overtwisted contact structures [16], the two contact structures must then be isotopic, and hence by Giroux's theorem [19], the boundary open book decompositions will agree after some number of positive stabilizations (which can be realized by further modifications to the fibrations). This process is, of course, nonconstructive due to its reliance on these deep classification results.

Baykur and Lekili's constructions instead focused on studying deformations of generic maps near their singularities. More precisely, they both show that a generic indefinite surjective map $X \to S^2$ can be modified near its critical points to obtain a broken Lefschetz fibration $f : X \to S^2$. These early singularity theory constructions did not (in general) produce broken Lefschetz fibrations with embedded images of their indefinite fold singularities, however. More recent work of Baykur and Saeki [9; 10] improves upon these techniques, by presenting explicit algorithms to convert an arbitrary broken Lefschetz fibration into one with connected fibers and a single indefinite fold singularity with embedded image.

In the case when $b_2^+(X) > 0$, or equivalently when X is near-symplectic, the near-symplectic structure can be used to construct broken Lefschetz fibrations and pencils with additional desired properties. For example, it can be shown that any near-symplectic structure is cohomologous to a near-symplectic form which has connected zero locus, and this can be used to show that in this case X admits a broken Lefschetz pencil with connected fibers and at most one indefinite fold singularity, and that the indefinite fold image is embedded.

Aside from the existence results mentioned above, the uniqueness question for broken Lefschetz fibrations has also been studied. In [43], Williams establishes a set

of modifications to broken Lefschetz fibrations which preserve the homotopy class of the fibration map, and proves that they are sufficient to relate any two broken Lefschetz fibrations in the same homotopy class. He also obtains a set of moves relating all broken Lefschetz fibrations on a given 4-manifold (even nonhomotopic ones) by adding in an additional projection move.

Finally, it is worth noting that Lefschetz fibrations have also been extended to certain nonorientable manifolds. In [33], Miller and Ozbagci show that any nonorientable handlebody without 3- and 4-handles admits a Lefschetz fibration over the disk. The fibers of these fibrations are nonorientable surfaces with nonempty boundary.

2. Braided surfaces in $D^2 \times D^2$

In addition to the constructions described in Section 1F, (broken) Lefschetz fibrations can also be obtained by way of branched coverings and braided surfaces. More precisely, fibrations on X can be obtained by constructing and modifying certain coverings $h : X \to D^2 \times D^2$, which are branched along properly embedded surfaces in $D^2 \times D^2$. To obtain a fibration, we will require that these branch loci are braided surfaces with folds in $D^2 \times D^2$. This approach can be carried out directly on a given handle decomposition of the 4-manifold, and yield explicit broken Lefschetz fibrations. In this section we define braided surfaces and a generalization, before outlining this technique and providing explicit examples in later sections.

2A. ***Braided ribbon surfaces.*** Rudolph defined a *braided surface* [37] to be a smooth properly embedded oriented surface $S \subset D^2 \times D^2$ on which the projection to the second factor $\mathrm{pr}_2 : D^2 \times D^2 \to D^2$ restricts as a simple branched covering. Examples of these braided surfaces can be obtained by taking intersections of nonsingular complex plane curves with 4-balls in $\mathbb{C}^2$, and they can be used to study the links that arise as their boundaries in $S^3 = \partial D^4$ (see, e.g., [38; 39; 40]). See Figure 3. The boundary of a braided surface will be a closed braid in the solid torus $D^2 \times S^1 \subset \partial(D^2 \times D^2)$.

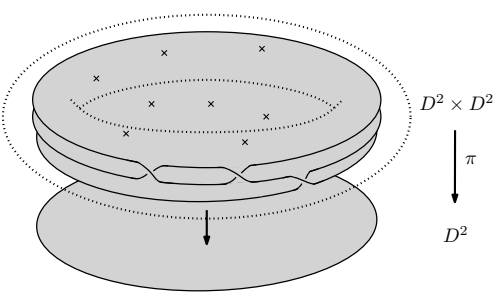


Figure 3. Braided ribbon surface.

Let S be a braided surface. In a neighborhood of any branch point p of the covering $\mathrm{pr}_2|_S$, there are local complex coordinates u and v on D^2 such that S is given by the equation $u^2 = v$ in the coordinates (u, v) on $D^2 \times D^2$. We say that p is a *positive* branch point if these coordinates can be taken to be orientation-preserving, and a *negative* branch point otherwise.

One feature of Rudolph's braided surfaces is that they are all necessarily *ribbon*. A properly embedded surface S in $D^4 = \{(z, w) : |z|^2 + |w|^2 \leq 1\}$ is said to be *ribbon embedded* if the function $|z|^2 + |w|^2$ restricts to S as a Morse function with no local maximal points on $\operatorname{int} S$. A properly embedded surface in D^4 is said to be *ribbon* if it is isotopic to a surface which is ribbon embedded. By fixing an identification of $D^2 \times D^2$ with D^4, we can similarly consider ribbon surfaces in $D^2 \times D^2$ (the definition of ribbon embeddings in $D^2 \times D^2$ will depend on our choice of identification, though the resulting class of ribbon surfaces will not).

Rudolph proved that any orientable ribbon surface in $D^2 \times D^2$ is isotopic to a braided surface, though in general this isotopy cannot be chosen to fix ∂S even if ∂S is already a closed braid in $D^2 \times S^1 \subset \partial(D^2 \times D^2)$. Rudolph's braiding algorithm involves manipulating a ribbon *immersed* surface in $\mathbb{R}^3$, and hence can't be applied to nonribbon surfaces in $D^2 \times D^2$.

Viro defined a similar notion which he called a 2-*braid*, by additionally requiring that $\partial S \subset D^2 \times S^1$ be a trivial closed braid (i.e., $\partial S = P \times S^1$ for some finite subset $P \subset D^2$). Viro's 2-braids come equipped with a closure operation yielding closed surfaces in S^4, and in a September 1990 lecture at Osaka City University, Viro proved a 4-dimensional Alexander theorem by showing that every closed oriented surface in S^4 is isotopic to the closure of a 2-braid. These 2-braids were also studied extensively by Kamada [24; 25; 26; 27; 28], who proved a 4-dimensional Markov theorem relating any two 2-braids with isotopic closures.

Braided surfaces admit monodromy representations, similar to the multisections described in Section 1C. Let $S \subset D^2 \times D^2$ be a braided surface, and let D^* be the regular values of the restriction $\mathrm{pr}_2|_S$ (i.e., the complement of the images of the branched points of S under $\mathrm{pr}_2|_S$). Then D^* will be a punctured disk, and after fixing a basepoint $q \in \partial D^*$ we can choose a collection of oriented simple closed curves $\gamma_1, \ldots, \gamma_m$ based at q, which are disjoint away from q, and such that γ_j travels counterclockwise around the j-th puncture of D^* and no other punctures. We can order the loops $\gamma_1, \ldots, \gamma_m$ so that the product $\gamma_1 \cdots \gamma_m$ is homotopic to ∂D^2.

If the restriction $\mathrm{pr}_2|_S$ is an n-sheeted branched covering of D^2, then for each j the set $(\mathrm{pr}_2|_S)^{-1}(\gamma_j)$ will be a closed n-stranded braid in the solid torus $(\mathrm{pr}_2)^{-1}(\gamma_j)$. Each of the $(\mathrm{pr}_2|_S)^{-1}(\gamma_j)$ will be the closure of a braid of the form $\alpha_j = \beta_j^{-1}\sigma_{i_j}^{\pm 1}\beta_j$, where σ_{i_j} is one of the standard Artin generators of the n-strand braid group B_n, and $\beta_j \in B_n$. Equivalently, if D_q denotes the fiber of pr_2 above the point q, and $(\mathrm{pr}_2|_S)^{-1}(q) = \{s_1, \ldots, s_n\} \subset D_q$, then each loop γ_j will induce a monodromy

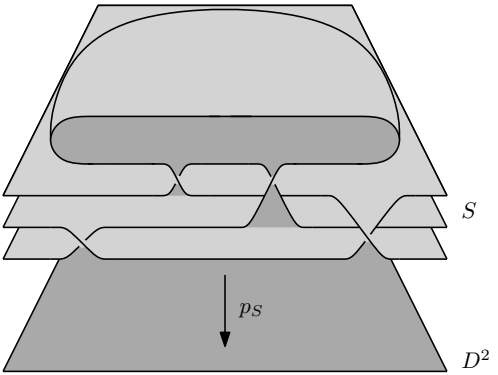


Figure 4. Cross section of a braided surface with folds.

map $\varphi_j \in \mathcal{M}(D_q; \{s_1, \ldots, s_n\})$, which swaps precisely two points in $\{s_1, \ldots, s_n\}$ along some arc in D_q, and leaves the others points fixed. The family of monodromy maps $\varphi_1, \ldots, \varphi_m$ (resp. family of braids $\alpha_1, \ldots, \alpha_m$) will be determined by S up to conjugation by a fixed element of $\mathcal{M}(D_q; \{s_1, \ldots, s_n\})$ (resp. conjugation by a fixed braid in B_n), as well as changes in the choice of loops $\gamma_1, \ldots, \gamma_m$. Conversely, a family of such maps in $\mathcal{M}(D_q; \{s_1, \ldots, s_n\})$ or braids in B_n define a braided surface in $D^2 \times D^2$ up to isotopy through braided surfaces.

2B. ***Braided surfaces with folds.*** The surfaces in $D^2 \times D^2$ we use to construct broken Lefschetz fibrations will not in general be ribbon, and hence cannot be braided via Rudolph's algorithm. We thus consider a less restrictive notion of braiding, which we define now.

Let $\phi : F \to \Sigma$ be a smooth map of oriented surfaces. Then a *fold of F with respect to ϕ* is an embedded circle $C \subset F$, so that

(1) ϕ restricts to an embedding on C,

(2) F and Σ both admit coordinate charts of the form $S^1 \times [-1, 1]$ around $C = S^1 \times \{0\}$ and $\phi(C) = S^1 \times \{0\}$, on which ϕ is given by $(\theta, t) \mapsto (\theta, t^2)$,

Now let $S \subset D^2 \times D^2$, and let pr_S denote the restriction of pr_2 to S. We say that S is a *braided surface with folds* if the critical points of pr_S all correspond either to isolated simple branch points or folds of S with respect to pr_S. Moreover, we will often assume that the critical values in D^2 form a set of embedded concentric circles (corresponding to folds), with isolated critical values lying inside the innermost circle. See Figure 4 for a cross sectional diagram of a braided surface with a single fold. We prove the following.

Theorem 2.1. *Let S be a smooth oriented surface properly embedded in $D^2 \times D^2$. Then S is isotopic to a braided surface with folds and only positive branch points. If ∂S is already a closed braid, then the isotopy can be chosen rel ∂S.*

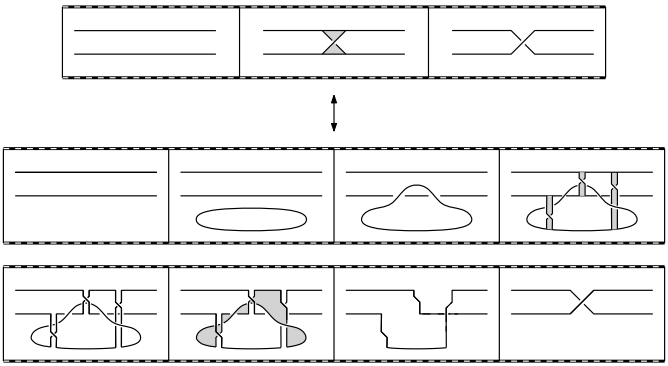


Figure 5. Replacing a negative branch point with three positive branch points and a fold.

Proof. In [23], the author proves that every such surface $S \subset D^2 \times D^2$ is isotopic to a braided surface with caps. Here, a cap is an embedded disk in S on which the projection pr_2 restricts as an embedding, and whose boundary is a fold circle as defined above. Moreover, the isotopy arranging S as a braided surface with caps can be taken rel ∂S is ∂S if already a closed braid. Alternatively, one could start with a bridge trisection of the surface S in the standard genus zero trisection of $D^4 = D^2 \times D^2$, which Meier [32] proved is always possible.

In order to ensure only positive branch points, we replace any negative branch points as shown in Figure 5. More precisely, if $p \in S$ is a negative branch point, then we can choose some parametrized neighborhood V around p so that $S \cap V$ is described locally by the motion picture diagram shown at the top of Figure 5. We can then remove $S \cap V$ from S, and replace it with the surface whose local motion picture description is shown at the bottom of Figure 5. This removes the original negative branch point p, and replaces it with three positive branch points, and a single fold circle. To see that these two surfaces are isotopic rel ∂S, we construct the isotopy shown in Figure 6 using band slides. □

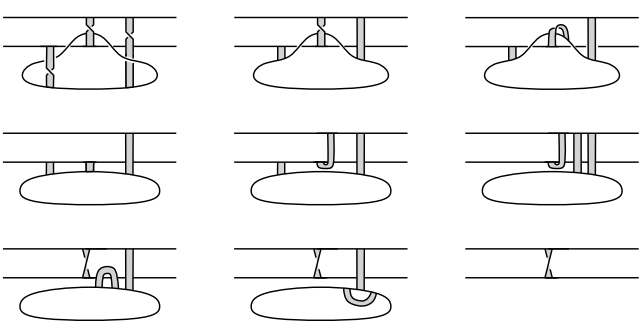


Figure 6. An isotopy which takes the replacement surface to the original surface with a single negative branch point.

3. Broken Lefschetz fibrations from branched coverings

In this section we describe how (broken) Lefschetz fibrations can be constructed via branched coverings and braided surfaces. We then provide examples of this construction in Section 4.

3A. ***Lefschetz fibrations from branched coverings.*** Suppose X is an oriented 4-dimensional 2-handlebody with a fixed handle decomposition that has no 3- or 4-handles. Then we can construct a simple branched covering $H : X \to D^2 \times D^2$ branched along an orientable ribbon surface S. Further, we can assume S is a braided surface by Rudolph's algorithm. Then the composition $X \xrightarrow{H} D^2 \times D^2 \xrightarrow{\mathrm{pr}_2} D^2$ is an achiral Lefschetz fibration, with a positive Lefschetz critical point (resp. negative Lefschetz critical point) for each positive (resp. negative) branch point of $S \to D^2$. Thus if S has only positive branch points, we obtain a true Lefschetz fibration. In fact, Loi and Piergallini [31] show that any sufficiently nice Lefschetz fibration over D^2 necessarily factors in this way.

Using these constructions, Loi and Piergallini also prove that for an oriented connected compact 4-manifold X with boundary, the existence of a Stein structure is equivalent to the existence of a Lefschetz fibration over D^2 with all vanishing cycles nonseparating in the fiber. By considering the associated simple branched covering restricted to ∂X, it follows that a 3-manifold is Stein fillable if and only if it admits a positive open book decomposition.

Now suppose we start instead with a handlebody description of a 4-manifold X which has 3- and 4-handles. As noted above we can construct a branched covering of the 0-,1-, and 2-handles over $D^2 \times D^2$, branched along a ribbon surface. Once we try to extend this covering to the 3- and 4-handles however, the branch locus is no longer ribbon, and may additionally have cusp and node singularities.

3B. ***Broken Lefschetz fibrations from branched coverings.*** Our method for creating broken Lefschetz fibrations on handlebodies with 3- and 4-handles is based on Proposition 3.1, which takes as input a simple branched covering $h : X \to D^2 \times D^2$ with orientable branch locus, and yields a broken Lefschetz fibration $g : X \to D^2$. This approach can then be combined with techniques of Gay and Kirby to produce broken Lefschetz fibrations over S^2 on many closed 4-manifolds.

Proposition 3.1 is a generalization of Proposition 1.2 of [31] to branched coverings with nonribbon branch loci.

Proposition 3.1. *Suppose that X is a smooth 4-manifold with boundary, and that $h : X \to D^2 \times D^2$ is a simple branched covering with branch locus $B_h \subset D^2 \times D^2$ an embedded orientable surface. Then there is an isotopy $\phi_t : D^2 \times D^2 \to D^2 \times D^2$, $\phi_0 = \mathrm{id}_{D^2 \times D^2}$, such that $\mathrm{pr}_2 \circ \phi_1 \circ h : X \to D^2$ is a broken Lefschetz fibration.*

Proof. By Theorem 2.1, B_h is isotopic in $D^2 \times D^2$ to a braided surface with folds and only positive branch points. Let ϕ_t be an isotopy of $D^2 \times D^2$ which takes B_h to such a surface. Let $H = \phi_1 \circ h$ denote the isotoped branched covering, and let B_H denote its branch locus. Away from the preimages of the critical points of $\mathrm{pr}_2|_{B_H}$, the composition $g = \mathrm{pr}_2 \circ H$ is a regular map. By [31] the map g has a Lefschetz critical point for every positive branch point of $\mathrm{pr}_2|_{B_H}$.

To see that the fold lines of B_H give indefinite fold singularities, note that along these fold lines B_H is locally embedded as $\mathbb{R}^2 \to \mathbb{R}^2 \times \mathbb{R}^2$, by $(s, r) \mapsto (0, r, s, r^2)$. Furthermore, near nonsingular points of B_H, H can be written in complex coordinates as $(u, v) \mapsto (u^2, v)$, where B_H is given locally by $u = 0$. Combining these two local models yields a map of the required local form. Furthermore, the folds of B_H can be pushed out so that they lie above a neighborhood of the boundary of D^2, so that their images form a collection of concentric circles in D^2 which enclose the Lefschetz critical values. □

Remark 3.2. Note that Proposition 3.1 holds more generally than stated above. Indeed, by [3; 30] any generic map $X \to D^2$ can be perturbed to become a broken Lefschetz fibration, and by [9; 10] any map $X \to S^2$ can be converted to a broken Lefschetz fibration whose fibers are all connected, and whose indefinite fold singularities are connected with embedded image. The proof of Proposition 3.1 is what will be most useful to us, since the branched covering $h : X \to D^2 \times D^2$ and the isotopy $\phi_t : D^2 \times D^2 \to D^2 \times D^2$ can often be constructed by hand from a given Kirby diagram of X (see Section 4).

3C. ***Broken Lefschetz fibrations on closed 4-manifolds.*** We will now show how Proposition 3.1 can be used in many cases to construct a broken Lefschetz fibration $f : X \to S^2$ on a closed orientable 4-manifold X from a given handle decomposition. Let $F \subset X$ be a closed surface with $F \cdot F = 0$, and consider a tubular neighborhood νF of F. For simplicity, we describe first the construction in the case that $F \cong S^2$, and hence $\nu F \cong S^2 \times D^2$. Such a neighborhood can sometimes be identified in the handle diagram of X as a 2-handle attached along a 0-framed unknot together with the 0-handle of X. If no such $S^2 \times D^2$ can be identified, it can be added to the diagram by adjoining a canceling 2- and 3-handle pair, where the 2-handle is attached along a 0-framed unknot. We will think of the union of this 2-handle with the 0-handle to which it is attached as forming νF.

3D. ***Building the concave piece.*** We describe how to construct a concave broken fibration $f : \nu F \to S^2$ with a single indefinite fold singularity and no Lefschetz critical points. This construction is originally due to Auroux, Donaldson, and Katzarkov [2], as part of their construction of a broken Lefschetz fibration on S^4, though our description follows that in [18].

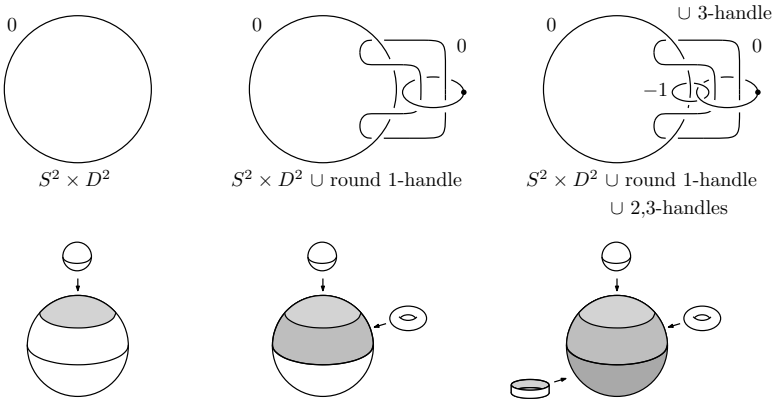


Figure 7. Concave broken fibration on $S^2 \times D^2$.

We begin by identifying the target of the projection $\mathrm{pr}_2 : S^2 \times D^2 \to D^2$ with the northern polar cap in S^2. This defines a fibration of $S^2 \times D^2$ with fiber S^2 over this region (see the bottom-left diagram in Figure 7). Expressing $S^2 \times D^2$ with the usual handlebody diagram (top-left, Figure 7), we can add a 1-handle and 0-framed 2-handle to this diagram, as in the top middle diagram. Taken together, these two handles can be interpreted as a round 1-handle, which is attached to $S^2 \times D^2$ along two sections of the existing fibration restricted to the boundary. We can thus extend this fibration over the round 1-handle, giving a fibration over the northern hemisphere with an indefinite fold singularity over the arctic circle. Fibers between the equator and the arctic circle will be obtained from the polar fibers by 0-surgery, and hence will be tori. Note that the fibration we have constructed so far is flat along its boundary.

Finally, we add an additional 2-handle H_2, and a 3-handle H_3 to our diagram (top-right, Figure 7). The attaching circle of H_2 is a section of the flat fibration restricted to the boundary, and hence the fibration can be extended over H_2, by projecting it to the southern hemisphere (with fiber D^2). In other words, thinking of H_2 as $D^2 \times D^2$ attached along $\partial D^2 \times D^2$, we think of H_2 as sitting as a D^2-bundle over the southern hemisphere, with projection map $D^2 \times D^2 \to D^2$ being given by projection onto the first factor. Note that we choose the attaching circle of H_2 so that it runs over the existing 1-handle from the round 1-handle once, and has framing -1. While these choices are not necessary to ensure the fibration extends, they are made to allow for the handle cancellations described below.

After extending over H_2 the resulting fibration is concave. The page of the boundary open book decomposition is a torus with a single hole (which resulted from attaching the 2-handle H_2), while its binding will be the belt-sphere of H_2.

The attaching sphere of the new 3-handle H_3 is arranged so that it intersects the binding at its north and south poles, and so that it intersects each page in a properly

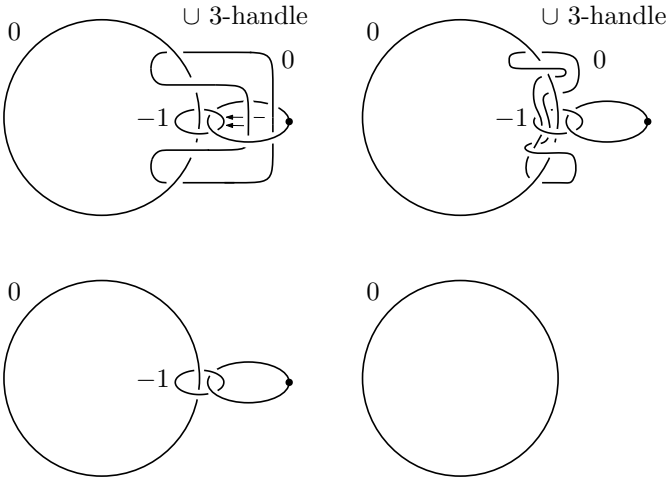


Figure 8. Sliding handles in the concave fibration to identify canceling pairs. After sliding one of the 0-framed 2-handles over the (-1)-framed 2-handle H_2 twice, the slid handle becomes a 0-framed unknot that can be canceled with the 3-handle. The handle H_2 can then be canceled with the 1-handle, leaving behind $S^2 \times D^2$.

embedded arc. The fibration can then be extended across H_3, resulting in no new critical points. This extension changes the D^2 fibers over the southern hemisphere by adding a 2-dimensional 1-handle, yielding annular fibers. On the other hand, the pages of the boundary open book change by the *removal* of a neighborhood of a properly embedded arc from the puncture torus pages (the intersection of the original page with the attaching sphere of H_3), yielding annular pages.

This gives a concave broken fibration as depicted in the bottom-right diagram of Figure 7, with a single indefinite fold singularity, and no positive or negative Lefschetz critical points. Moreover, after sliding the 0-framed 2-handle from the round 1-handle over H_2 twice as shown in Figure 8, we find that the added 1-,2-, and 3-handles all form canceling pairs. Hence the total space of our fibration is diffeomorphic to $S^2 \times D^2 \cong \nu F$. Notice that the induced open book decomposition on $\partial(\nu F)$ will have disconnected binding, which may cause problems when we try to construct a matching convex fibration on $X \backslash \nu F$. We thus instead think of the lone canceling 3-handle as being attached as a 1-handle to $X \backslash \nu F$, and construct a concave fibration f_1 on $X_1 = \nu F \backslash \{\text{3-handle}\}$, whose boundary open book decomposition has punctured torus page and connected binding (see Figure 9).

If instead F has genus $g \geq 1$ we can proceed much as before, either identifying a neighborhood νF in the handle diagram of X, or by adding a standard diagram of $F \times D^2$ with additional 2- and 3-handles to cancel the 1- and 2-handles of νF. More

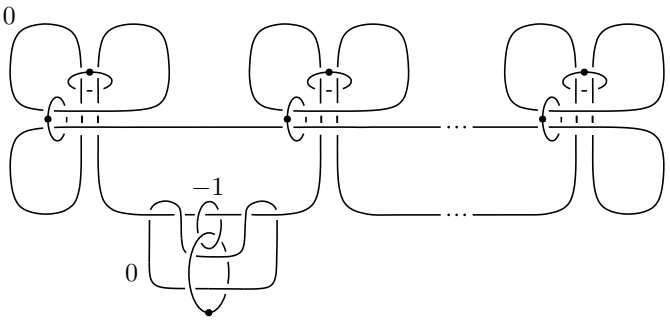


Figure 9. Neighborhood of $F \subset X$ with an extra 2-handle and round 1-handle added.

precisely, ignoring the (-1)-framed 2-handle in Figure 9, and the 1- and 2-handles coming from the round 1-handle, the remaining handles (together with the 0-handle) give a diagram for $F \times D^2$. This diagram can be placed in any handlebody diagram of a 4-manifold X without changing the diffeomorphism type of X, provided we also add a 0-framed 2-handle attached along a meridian for each 1-handle circle, along with a single extra 3-handle. To this diagram we could then add a round 1-handle and pair of 2- and 3-handles (see Figure 9) and continue as above.

3E. ***Building the convex piece.*** Let $Y = X \backslash X_1$. We now discuss how to use a handle structure on Y to build a convex fibration $g : Y \to D^2$, so that it extends the open book decomposition $\lambda : \partial Y \to D^2$ induced by the concave fibration $f_1 : \nu F \to S^2$. We attempt to do this in three steps:

(1) Express the open book decomposition λ as $\lambda = \mathrm{pr}_2 \circ h$, where $h : \partial Y \to \partial(D^2 \times D^2)$ is a simple covering branched along a closed braid in $\partial(D^2 \times D^2)$, and $\mathrm{pr}_2 : D^2 \times D^2 \to D^2$ is the projection.

(2) Extend the branched covering h to a covering $H : Y \to D^2 \times D^2$ branched along an orientable surface.

(3) Use Proposition 3.1 to obtain the desired broken Lefschetz fibration.

Part (1) is always possible. Indeed, let P be the page of λ, with monodromy $\tau : P \to P$. Then by choosing a suitable (degree ≥ 3 and simple) branched covering $\alpha : P \to D^2$, the map τ is the lift of a map $\widehat{\tau} : D^2 \to D^2$ which fixes the branch locus of α setwise [21; 29]. If K is the binding of λ, then this allows us to write $\partial Y \backslash \nu K$ as a branched covering of the solid torus $D^2 \times \partial D^2$ branched over a closed braid. A matching (unbranched) covering $\nu K \to \partial D^2 \times D^2$ can be glued to this covering to give the desired map $h : \partial Y \to \partial(D^2 \times D^2)$.

Problems may arise when we try to carry out part (2) of the above process, however. The covering h can always be extended to a branched covering $H : Y \to D^2 \times D^2$, though the branch locus may not be an orientable embedded surface.

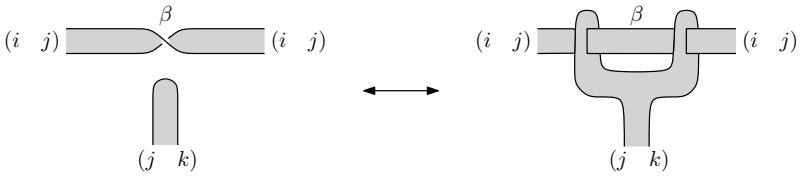


Figure 10. Fixing a nonorientable band in the branch locus of $\widehat{h}$.

To see how this covering is constructed, fix some choice of relative handle decomposition for the pair $(Y, \partial Y)$. The covering $h : \partial Y \to \partial(D^2 \times D^2)$ can be extended to a covering

$$\widehat{h} : \partial Y \times [0, 1] \to \partial(D^2 \times D^2) \times [0, 1]$$

in the usual way. Here, $\partial Y \times [0, 1]$ and $\partial(D^2 \times D^2) \times [0, 1]$ are thought of as collar neighborhoods of ∂Y and $\partial(D^2 \times D^2)$ respectively. Identify ∂Y with $\partial Y \times \{0\}$, and let $\partial_+ Y = \partial Y \times \{1\}$. We now attempt to extend this covering over the handles of Y to construct the desired covering H.

Let $\sigma_1 = D^1 \times D^3$ be a 1-handle, and let $\tau_1 : \sigma_1 \to \sigma_1$ be the involution defined by

$$\tau_1 : (t, x, y, z) \longmapsto (-t, -x, y, z).$$

If σ_1 is a 1-handle in our handle decomposition of $(Y, \partial Y)$, then we can isotope its attaching map $\alpha_1 : S^0 \times D^3 \to \partial_+ Y$ so that it is symmetric with respect to $\widehat{h}$, i.e., so that $\widehat{h} \circ \alpha_1 \circ \tau = \widehat{h} \circ \alpha_1$. Once this is done, by Lemma 6.1 of [11] we can extend $\widehat{h}$ over the 1-handle σ_1, using the quotient induced by τ_1. The result is a branched covering of $(\partial Y \times [0, 1]) \cup \sigma_1$ over $\partial(D^2 \times D^2) \times [0, 1]$, where the new branch locus is obtained by adding a disjoint disk to the branch locus of $\widehat{h}$ in $\partial(D^2 \times D^2) \times [0, 1]$.

Similarly, if $\sigma_2 = D^2 \times D^2$ is instead a 2-handle attached to $\partial_+ Y$ by some attaching map $\alpha_2 : S^1 \times D^2 \to \partial_+ Y$, by [15] we can isotope α_2 so that it becomes symmetric with respect to the involution $\tau_2 : D^2 \times D^2 \to D^2 \times D^2$, defined by

$$\tau_2 : (t, s, x, y) \longmapsto (-t, s, -x, y).$$

Here, the attaching circle of σ_2 will intersect the branching set of $\widehat{h}$ in two points, say p_1 and p_2. Then we can extend the covering $\widehat{h}$ to a branched covering of $(Y \times [0, 1]) \cup \sigma_2$ over $\partial(D^2 \times D^2)$, where the new branch locus is obtained by attaching a single band to the branch locus of $\widehat{h}$ at the points $\widehat{h}(p_1)$ and $\widehat{h}(p_2)$. This band will have n half-twists in it, where n is the framing of σ_2.

When extending $\widehat{h}$ over a 2-handle σ_2, it is possible that the corresponding band β may be attached to the branch locus $B \subset \partial(D^2 \times D^2) \times [0, 1]$ in a nonorientable way. By [12] this can be remedied, by adding (or removing) a half-twist in β as in Figure 10. In this local picture we have pushed B entirely into the 3-dimensional space $\partial(D^2 \times D^2) \times \{1\}$, where it can be depicted as an immersed surface with

only ribbon double points. The labels on the components denote the associated monodromy action on the sheets of $\widehat{h}$.

Let Y_2 denote the union of $\partial Y \times [0, 1]$ with the 1- and 2-handles. We can thus extend the branched covering $h : \partial Y \to \partial(D^2 \times D^2)$ to a covering

$$\tilde{h} : Y_2 \to \partial(D^2 \times D^2) \times [0, 1],$$

where the associated branch locus $\widetilde{B} \subset \partial(D^2 \times D^2) \times [0, 1]$ is an embedded orientable surface. If the intersection

$$\widetilde{B}_1 = \widetilde{B} \cap \partial(D^2 \times D^2) \times \{1\}$$

is an unlink, then $\tilde{h}$ can be extended across the 3- and 4-handles to give a branched covering $H : Y \to D^2 \times D^2$ with orientable embedded branch locus. This can be seen by noting that the union of the 3- and 4-handles is a thickened bouquet of circles, which can be expressed as a branched covering of D^4 with branch locus a collection of properly embedded disjoint disks. If $\widetilde{B}_1$ is an unlink, this covering can be glued to $\tilde{h} : Y_2 \to \partial(D^2 \times D^2) \times [0, 1]$ to obtain the desired covering H.

In general however, $\widetilde{B}_1$ will not be an unlink. By [36] we can modify the covering by adding cusp and node singularities on the interior of $\widetilde{B}$ so that $\widetilde{B}_1$ becomes an unlink, though doing so may fail to preserve the required orientability of the branch locus B. When this can be avoided, we can proceed with the rest of the construction to obtain a broken Lefschetz fibration of X over S^2.

3F. ***Broken Lefschetz fibrations on doubles of 4-manifolds.*** We now discuss a situation in which the above construction will always be possible. Let U be a handlebody with single 0-handle and no 4-handles. The *double* of U is the manifold $X = U \cup_{\mathrm{Id}_{\partial U}} \overline{U}$, where $\overline{U}$ denotes the handlebody U with reversed orientation. The handle structure on U induces a handle structure on X in a natural way, by turning the j-handles of $\overline{U}$ upside-down and attaching them as $(4-j)$-handles to U.

Theorem 3.3. *Let X be a smooth, closed, orientable 4-manifold, with handle structure coming from the double of a handlebody U. Then the procedure described in Section 3E will produce a broken Lefschetz fibration $f : X \to S^2$.*

Proof. If $F = S^2$ is a trivially embedded sphere in the 0-handle of X, we can construct a concave fibration of νF over S^2 as in Section 3D. Let $Y = X \backslash \nu F$, and let $\lambda : \partial Y \to D^2$ be the induced open book decomposition. By [18] the monodromy of λ is trivial, and hence it factors through a simple branched covering $h : \partial Y \to \partial(D^2 \times D^2)$ of degree ≥ 3, whose branch locus is a trivial closed braid in $D^2 \times \partial D^2$.

We now proceed to extend the covering h to a covering

$$\tilde{h} : Y_2 \to \partial(D^2 \times D^2) \times [0, 1]$$

with branch locus $\widetilde{B}$. Again we let $\widetilde{B}_1$ be the intersection of $\widetilde{B}$ with $\partial(D^2 \times D^2) \times \{1\}$.

Each 1-handle we extend over contributes an unknot component to $\widetilde{B}_1$ which is unlinked from the other components.

Before extending h across the 2-handles, note that in the induced handle structure on Y, the 2-handles occur pairs. Every 2-handle σ from U is paired with a 2-handle σ' from $\overline{U}$, where σ' is attached along a 0-framed meridian of the attaching circle of σ (see [20]). We can also think of σ' as being attached along the belt sphere of σ.

Extending h over a 2-handle from U changes $\widetilde{B}_1$ by oriented surgery along a band β. On the other hand, since the belt sphere of σ is symmetric with respect to the involution $\tau_2 : \sigma \to \sigma$, extending h across σ' will change $\widetilde{B}_1$ by oriented surgery along a band β' which cancels β. Hence the net effect of extending h across σ and σ' does not change $\widetilde{B}_1$, which thus remains an unlink. □

Any orientable S^2-bundle over a (possibly nonorientable) surface Σ is the double of a D^2-bundle over Σ. See [18] for an alternate construction of broken Lefschetz fibrations on doubles of 2-handlebodies.

3G. ***Connected sums.*** The procedure outlined in Section 3E respects connected sums in the following sense:

Proposition 3.4. *Suppose that X_1 and X_2 are two handlebodies for which the procedure in Section 3E yields broken Lefschetz fibrations $f_1 : X_1 \to S^2$ and $f_2 : X_2 \to S^2$. Then the same procedure can be used to obtain a broken Lefschetz fibration $f : X_1 \# X_2 \to S^2$ which restricts to a concave fibration on $X_1 \backslash D^4 \subset X_1 \# X_2$ and to a convex fibration on $X_2 \backslash D^4 \subset X_1 \# X_2$. Moreover, the ball $D^4 \subset X_1$ can be chosen so that $f|_{X_1 \backslash D^4} = f_1|_{X_1 \backslash D^4}$.*

Proof. The handle structures on X_1 and X_2 yield a handle decomposition of $X_1 \# X_2$ by starting with the 0-handle of X_1 and attaching all 1-,2- and 3-handles of X_1, followed by the 1-, 2-, 3- and 4-handles of X_2.

Cut out a neighborhood of an S^2 from X_1, and construct the concave fibration on νS^2 and the branched covering h as above. The map h can be extended across the 1,2 and 3-handles of X_1 to give a covering

$$h' : X_1 \backslash (\nu S^2 \cup \text{4-handle}) \to \partial(D^2 \times D^2) \times [0, 1].$$

We identify $\partial(D^2 \times D^2) \times [0, 1]$ with $(D^2 \times D^2) \backslash (D' \times D')$, where $D' \subset D^2$ is a small disk containing the origin. Then by Theorem 2.1 the branch locus B' of h' can be braided rel $\partial B'$ so that it is a braided surface with folds in $(D^2 \times D^2) \backslash (D' \times D')$ and only positive branch points. Gluing the map $\mathrm{pr}_2 \circ h'$ to the concave fibration on νS^2 gives a concave fibration $X_1 \backslash$4-handle $\to S^2$. This fibration can either be continued across the 4-handle of X_1 to obtain the fibration $f_1 : X_1 \to S^2$, or across the 1-, 2-, 3- and 4-handles of X_2 to give a fibration $f : X_1 \# X_2 \to S^2$. □

Broken Lefschetz fibrations on connected sums were also built by Baykur in [4] (based on an observation by Perutz [35]), where he also defines a generalization of the symplectic fiber sum operation on near-symplectic broken Lefschetz fibrations.

4. Examples

In this section we compute a few simple examples, to illustrate how the above procedure is carried out.

4A. ***Broken Lefschetz fibration on S^4.*** Consider the diagram of S^4 in Figure 11. As in Figure 7, the union of all 0-, 1-, and 2-handles in this decomposition gives a neighborhood of an unknotted $S^2 \subset S^4$, together with an additional round 1-handle and (ordinary) 2-handle attached. Call the union of these handles X_1, and set $X_2 = S^4 \backslash X_1$. The open book decomposition on $\partial X_1 = \partial X_2$ induced by the concave fibration $f_1 : X_1 \to S^2$ from the above proof will have a punctured torus page with trivial monodromy (see [18]). Hence it can be represented by a 3-fold simple branched covering $h : \partial X_2 \to \partial(D^2 \times D^2)$, and whose branch locus in $\partial(D^2 \times D^2)$ is the closure of the trivial 4-strand braid in $D^2 \times \partial D^2$ (h can be described on each page by the branched covering in Figure 12).

The branched covering h extends to a covering $H : X_2 \to D^4$, which is built by turning the handle decomposition from Figure 11 upside-down, and viewing X_2 as a 0-handle with two 1-handles attached. The 0-handle can be expressed as a 3-fold covering of D^4 branched over two properly embedded unknotted disks. For each 1-handle we extend this covering over, a properly embedded unknotted disk is added to the branch locus. Hence the branch locus B_H of H in $D^4 \cong D^2 \times D^2$ is isotopic to the braided surface $\{p_1, \ldots, p_4\} \times D^2$, for some collection of disjoint points $\{p_1, \ldots, p_4\} \subset D^2$. The only critical points in the resulting broken Lefschetz fibration $f : S^4 \to S^2$ will thus lie along the indefinite fold singularity in X_1, and we recover Auroux, Donaldson, and Katzarkov's example in [2].

4B. ***S^2-bundles over orientable surfaces.*** Let X be an S^2-bundle over an closed orientable surface of genus g. For simplicity, we consider first the case when $g = 1$.

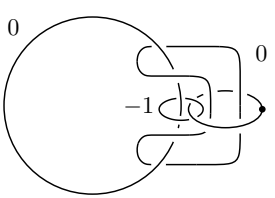


Figure 11. Handlebody structure of a neighborhood of S^2 in S^4.

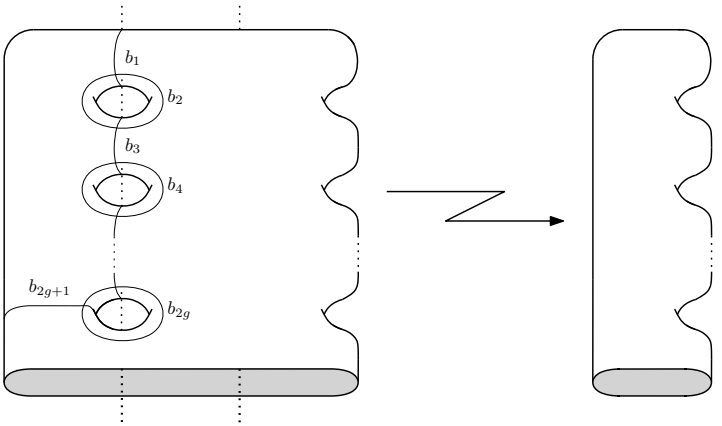


Figure 12. 3-fold branched cover $\Sigma \to D^2$.

Consider the diagram of X in Figure 13, where each 1-handle attaching sphere is paired with the sphere directly across from it. Notice that the diffeomorphism type of X depends only on the parity of n. Assume first that $n=0$. In this case $X \cong S^2 \times T^2$. While there is an obvious fibration $S^2 \times T^2 \to S^2$, the construction below has the advantage that it can be iterated to construct broken Lefschetz fibrations on connected-sums of S^2-bundles, and generalizes to the twisted bundle $S^2 \tilde{\times} T^2$. Note that broken Lefschetz fibrations on S^2 bundles over T^2 can also be obtained by converting trisection examples given in [10] or [13] to broken Lefschetz fibrations.

We begin by adding a copy of the diagram in Figure 11 (minus the 4-handle) to the diagram of X, which does not change the diffeomorphism type of X. Again, let X_1 denote the union of the 0-handle with the newly added 1-handle and 2-handles, and let $X_2 = X \backslash X_1$. As above, X_1 admits a concave fibration over S^2 and induces an open book decomposition on ∂X_2 with punctured torus page and trivial monodromy. The associated 3-fold branched covering

$$h : \partial X_2 \to \partial(D^2 \times D^2)$$

has branch locus a trivial 4-strand closed braid in $D^2 \times \partial D^2$. We need to extend h

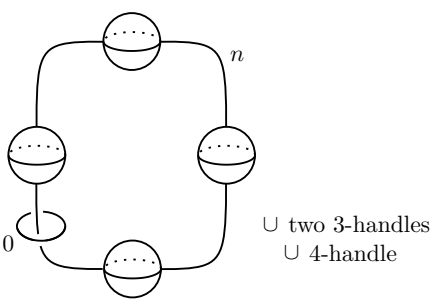


Figure 13. S^2-bundle over torus.

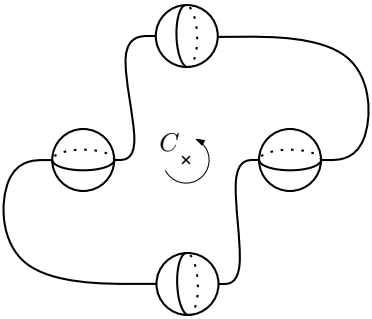


Figure 14. Symmetrizing the handles.

over the handles in Figure 13, as well as the additional 3-handles we introduced when adding the diagram in Figure 11.

In ∂X_2 there are four circles of branch points, corresponding to the four components of the branch locus in $\partial(D^2 \times D^2)$. We can isotope the handle attaching maps so that one of these four circles C skewers the diagram in Figure 13, so that locally the covering looks like rotation of π about the center of the diagram. We first focus on extending the covering over the 1-handles σ_1 and σ_1', and over the 2-handle σ_2 coming from the handle structure on T^2.

Isotope the attaching maps of these handles so that they are symmetric with respect to rotation by π around C, as in Figure 14. We can thus extend the covering $\tilde{h}$ over σ_1, σ_1', and σ_2. Extending over the 1-handles adds a pair of disks to the branch locus, while extending over the 2-handle adds a band. Notice that when the attaching circle of σ_2 runs along the horizontal 1-handle σ_1, it will intersect the branch set in precisely two points. The branch $\widetilde{B}$ locus of

$$\tilde{h} : (\partial X_2 \times [0,1]) \cup \sigma_1 \cup \sigma_1' \cup \sigma_2 \to \partial(D^2 \times D^2) \times [0,1]$$

will be as in Figure 15.

More precisely, let $\widetilde{B}_t = \widetilde{B} \cap (\partial(D^2 \times D^2) \times \{t\})$ for $t \in [0,1]$. The leftmost frame represents $\widetilde{B}_0$, the branch locus of h, where we have suppressed all of the components except for $h(C)$. As t increases, we see two unknotted components appear, corresponding to the 1-handles σ_1 and σ_1', followed by a band surgery corresponding to the 2-handle σ_2. Extending $\tilde{h}$ across the remaining 2-handle in Figure 13 results in an additional band surgery which cancels the first. Note that all of the components in Figure 15 will have the same monodromy as $h(C)$.

The branch locus $\widetilde{B}_1 = \widetilde{B} \cap (\partial(D^2 \times D^2) \times \{1\})$ is thus a six component unlink (three components from Figure 15 and three additional components from $h : \partial X_2 \to \partial(D^2 \times D^2)$ which were suppressed from the diagrams). It only remains to extend this covering over the four 3-handles and unique 4-handle of X_2. It is not hard to see that the union of these higher index handles admits a 3-fold simple

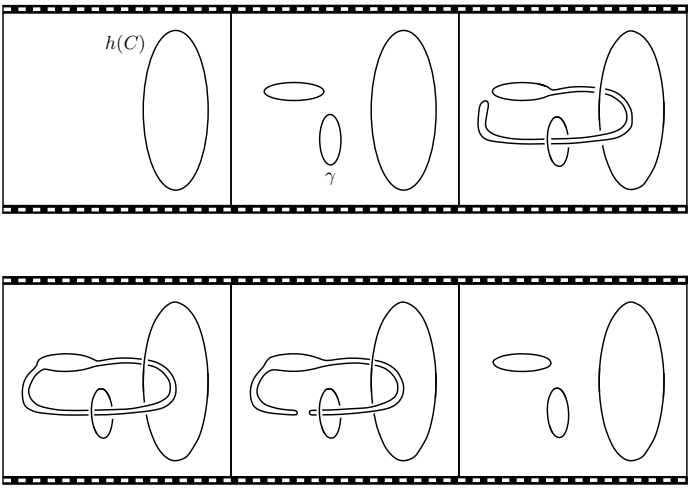


Figure 15. Branch locus $\widetilde{B}$ after extending over 1-handles and σ_2.

branched covering over D^4, with branch locus consisting of six disjoint properly embedded disks in D^4. This covering can thus be glued to $\tilde{h}$ to give a covering $H : X_2 \to D^2 \times D^2$, where these six disks cap off the six component unlink $\widetilde{B}_1$. Let B_H denote the branch locus of H, which consists of $\widetilde{B}$ capped off with these six disks.

Finally, in order to apply Proposition 3.1, we must arrange B_H as a braided surface with folds. By [23] this is equivalent to arranging $B_H \subset D^2 \times D^2$ so that it sits in a collar neighborhood $\partial(D^2 \times D^2) \times [0, 1]$ such that

(1) the restriction to B_H of the projection $\rho : \partial(D^2 \times D^2) \times [0, 1] \to [0, 1]$ is a Morse function, and

(2) $(\rho|_{B_H})^{-1}(t)$ a closed braid in $\partial(D^2 \times D^2) \times \{t\}$ for all regular values t.

Figure 16 shows how this can be done. Again we start with the component $h(C)$ (hiding the three other components), and introduce two new unknots corresponding to extending the branched covering over the 1-handles. The key difference now is that at every regular level the branch locus must be a closed braid. Hence, the band corresponding to σ_2 now shows up first as a maximal point, which is then completed by adding two half-twisted bands via saddle points in the seventh frame. The second band surgery takes place in the ninth frame. Finally the branch locus is simplified to the trivial 3-strand braid, which is capped of by three minimal points (the other unseen three unknot components are similarly capped off).

The resulting broken achiral Lefschetz fibration $\mathrm{pr}_2 \circ H : X_2 \to D^2$ has an indefinite fold singularity for each maximal point of B_H (which shows up along the boundary of the maximal disk), and a positive or negative Lefschetz critical point for each saddle point. Hence $\mathrm{pr}_2 \circ H$ has three indefinite fold singularities, two

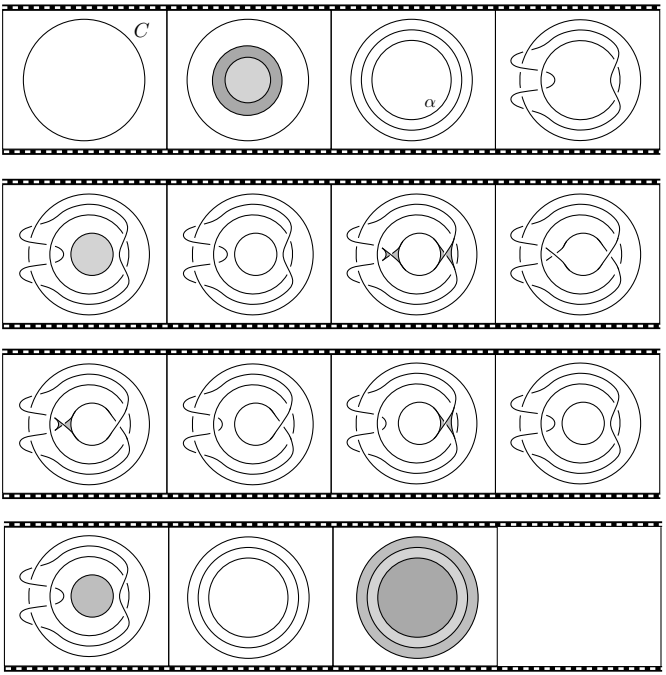


Figure 16. Branch locus B_H as a braided surface with folds.

positive Lefschetz critical points, and two negative Lefschetz critical points. The negative Lefschetz critical points can be replaced by the isotopy in Theorem 2.1, and the monodromy information of the fibration can be read off of Figure 16.

Now suppose that X is the S^2-bundle over T^2 given by Figure 13 with $n = 1$, i.e., $X \cong S^2 \tilde{\times} T^2$. Then the branch locus $\widetilde{B}$ will be as in Figure 15, except that the band corresponding to σ_2 will have a single half-twist, and hence $\widetilde{B}$ will be nonorientable. This can be remedied by involving another component of the branch locus $h : \partial X_2 \to \partial(D^2 \times D^2)$, and performing a move as in Figure 10 (see Figure 17, where the monodromy information must be chosen to agree with the labels in Figure 10).

When braided, this move introduces a new local maximal point, and two new saddle points (one of each sign). Hence the resulting broken achiral Lefschetz fibration has an additional indefinite fold singularity, positive Lefschetz critical point, and negative Lefschetz critical point when compared to the fibration constructed on $S^2 \times T^2$.

If X is a S^2-bundle over a higher surface of genus $g > 1$, we can start instead with the diagram in Figure 18. The associated branch locus will be as in Figure 16, except that the innermost strand α will be replaced by $2g - 1$ parallel strands, and hence the fibration $\mathrm{pr}_2 \circ H : X_2 \to D^2$ will now have $2g + 1$ indefinite fold singularities. The number of saddle points will be the same as in Figure 16, and

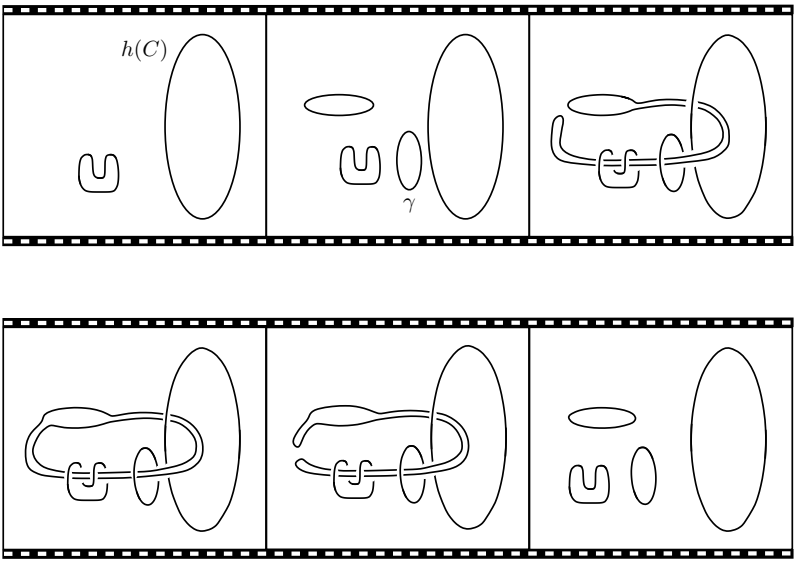


Figure 17. Branch locus $\widetilde{B}$ for $S^2 \tilde{\times} T^2$.

hence the resulting fibration will have two positive Lefschetz critical points and two negative Lefschetz critical points (where the negative critical points can be replaced as described above).

4C. ***S^2-bundles over $\mathbb{RP}^2$.*** We now consider S^2-bundles over $\mathbb{RP}^2$, which can be described by the diagram in Figure 19. Proceeding as above, we can arrange the component C of the branch set so that it sits vertically in the diagram between the two strands of the attaching circle of the n-framed 2-handle σ_2, and so that the attaching maps of σ_2 and the 1-handle σ_1 are symmetric with respect to rotation about C. For $n = 0$ and $n = 1$ the branch locus $\widetilde{B}$ will be as in Figures 15 and 17, respectively,

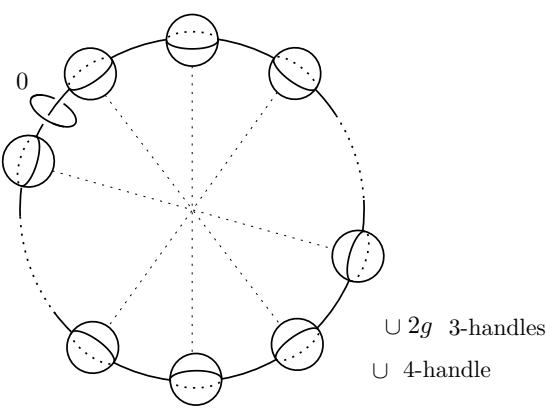


Figure 18. S^2-bundle over genus g surface.

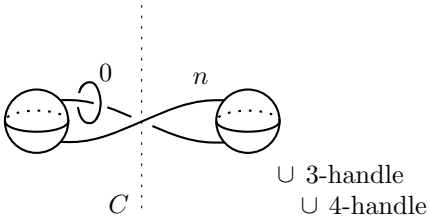


Figure 19. S^2-bundle over $\mathbb{RP}^2$.

except that the second unknot components (labelled by γ and corresponding to the extra 1-handle) will not be present. After filling in the higher index handles and braiding the resulting branch locus B_H, the result will be the same as in Figure 16, except that in the second still only the outermost new component will appear.

4D. ***Connected sums.*** The above constructions can be repeated to give broken Lefschetz fibrations on connected sums. For example, instead of capping off the unknot components in the third to last still of Figure 16, the movie (or another similar braided movie) could be repeated.

Acknowledgements

The author would like to thank Oleg Viro for helpful conversations while developing the constructions described above, as well as the anonymous referee for many valuable comments and suggestions that helped greatly improve this paper.

References

[1] S. Akbulut and c. Karakurt, "Every 4-manifold is BLF", *J. Gökova Geom. Topol. GGT* **2** (2008), 83–106. MR Zbl

[2] D. Auroux, S. K. Donaldson, and L. Katzarkov, "Singular Lefschetz pencils", *Geom. Topol.* **9** (2005), 1043–1114. MR Zbl

[3] R. I. Baykur, "Existence of broken Lefschetz fibrations", *Int. Math. Res. Not.* **2008** (2008), art. id. rnn101. MR Zbl

[4] R. I. Baykur, "Topology of broken Lefschetz fibrations and near-symplectic four-manifolds", *Pacific J. Math.* **240**:2 (2009), 201–230. MR Zbl

[5] R. I. Baykur, "Broken Lefschetz fibrations and smooth structures on 4-manifolds", pp. 9–34 in *Proceedings of the Freedman Fest*, Geom. Topol. Monogr. **18**, Geom. Topol. Publ., Coventry, 2012. MR Zbl

[6] R. I. Baykur and K. Hayano, "Broken Lefschetz fibrations and mapping class groups", pp. 269–290 in *Interactions between low-dimensional topology and mapping class groups*, Geom. Topol. Monogr. **19**, Geom. Topol. Publ., Coventry, 2015. MR Zbl

[7] R. I. Baykur and K. Hayano, "Hurwitz equivalence for Lefschetz fibrations and their multisections", pp. 1–24 in *Real and complex singularities*, Contemp. Math. **675**, Amer. Math. Soc., Providence, RI, 2016. MR Zbl

[8] R. I. Baykur and K. Hayano, "Multisections of Lefschetz fibrations and topology of symplectic 4-manifolds", *Geom. Topol.* **20**:4 (2016), 2335–2395. MR Zbl

[9] R. I. Baykur and O. Saeki, "Simplifying indefinite fibrations on 4-manifolds", 2017. arXiv

[10] R. I. Baykur and O. Saeki, "Simplified broken Lefschetz fibrations and trisections of 4-manifolds", *Proc. Natl. Acad. Sci. USA* **115**:43 (2018), 10894–10900. MR Zbl

[11] I. Berstein and A. L. Edmonds, "On the construction of branched coverings of low-dimensional manifolds", *Trans. Amer. Math. Soc.* **247** (1979), 87–124. MR Zbl

[12] I. Bobtcheva and R. Piergallini, "Covering moves and Kirby calculus", 2007. arXiv

[13] N. A. Castro and B. Ozbagci, "Trisections of 4-manifolds via Lefschetz fibrations", *Math. Res. Lett.* **26**:2 (2019), 383–420. MR Zbl

[14] S. K. Donaldson, "Lefschetz fibrations in symplectic geometry", pp. 309–314 in *Proceedings of the International Congress of Mathematicians*, *Vol. II* (Berlin, 1998), 1998. MR Zbl

[15] A. L. Edmonds, "Extending a branched covering over a handle", *Pacific J. Math.* **79**:2 (1978), 363–369. MR Zbl

[16] Y. Eliashberg, "Classification of overtwisted contact structures on 3-manifolds", *Invent. Math.* **98**:3 (1989), 623–637. MR Zbl

[17] J. B. Etnyre and T. Fuller, "Realizing 4-manifolds as achiral Lefschetz fibrations", *Int. Math. Res. Not.* **2006** (2006), art. id. 70272. MR Zbl

[18] D. T. Gay and R. Kirby, "Constructing Lefschetz-type fibrations on four-manifolds", *Geom. Topol.* **11** (2007), 2075–2115. MR Zbl

[19] E. Giroux, "Géométrie de contact: de la dimension trois vers les dimensions supérieures", pp. 405–414 in *Proceedings of the International Congress of Mathematicians*, *Vol. II* (Beijing, 2002), Higher Ed. Press, Beijing, 2002. MR Zbl

[20] R. E. Gompf and A. I. Stipsicz, *4-manifolds and Kirby calculus*, Graduate Studies in Mathematics **20**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[21] H. M. Hilden, "Three-fold branched coverings of S^3", *Amer. J. Math.* **98**:4 (1976), 989–997. MR Zbl

[22] K. Honda, "Transversality theorems for harmonic forms", *Rocky Mountain J. Math.* **34**:2 (2004), 629–664. MR Zbl

[23] M. C. Hughes, "Braiding link cobordisms and non-ribbon surfaces", *Algebr. Geom. Topol.* **15**:6 (2015), 3707–3729. MR Zbl

[24] S. Kamada, "2-dimensional braids and chart descriptions", pp. 277–287 in *Topics in knot theory* (Erzurum, 1992), NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci. **399**, Kluwer Acad. Publ., Dordrecht, 1993. MR Zbl

[25] S. Kamada, "Alexander's and Markov's theorems in dimension four", *Bull. Amer. Math. Soc. (N.S.)* **31**:1 (1994), 64–67. MR Zbl

[26] S. Kamada, "On braid monodromies of non-simple braided surfaces", *Math. Proc. Cambridge Philos. Soc.* **120**:2 (1996), 237–245. MR Zbl

[27] S. Kamada, "Arrangement of Markov moves for 2-dimensional braids", pp. 197–213 in *Low-dimensional topology* (Funchal, 1998), Contemp. Math. **233**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[28] S. Kamada, *Braid and knot theory in dimension four*, Mathematical Surveys and Monographs **95**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[29] C. Labruère and L. Paris, "Presentations for the punctured mapping class groups in terms of Artin groups", *Algebr. Geom. Topol.* **1** (2001), 73–114. MR Zbl

[30] Y. Lekili, "Wrinkled fibrations on near-symplectic manifolds", *Geom. Topol.* **13**:1 (2009), 277–318. MR Zbl

[31] A. Loi and R. Piergallini, "Compact Stein surfaces with boundary as branched covers of B^4", *Invent. Math.* **143**:2 (2001), 325–348. MR Zbl

[32] J. Meier, "Filling braided links with trisected surfaces", 2020. arXiv

[33] M. Miller and B. Ozbagci, "Lefschetz fibrations on nonorientable 4-manifolds", *Pacific J. Math.* **312**:1 (2021), 177–202. MR Zbl

[34] T. Perutz, "Lagrangian matching invariants for fibred four-manifolds, I", *Geom. Topol.* **11** (2007), 759–828. MR Zbl

[35] T. Perutz, "Lagrangian matching invariants for fibred four-manifolds, II", *Geom. Topol.* **12**:3 (2008), 1461–1542. MR Zbl

[36] R. Piergallini, "Four-manifolds as 4-fold branched covers of S^4", *Topology* **34**:3 (1995), 497–508. MR Zbl

[37] L. Rudolph, "Braided surfaces and Seifert ribbons for closed braids", *Comment. Math. Helv.* **58**:1 (1983), 1–37. MR Zbl

[38] L. Rudolph, "Special positions for surfaces bounded by closed braids", *Rev. Mat. Iberoamericana* **1**:3 (1985), 93–133. MR Zbl

[39] L. Rudolph, "Quasipositivity as an obstruction to sliceness", *Bull. Amer. Math. Soc.* (*N.S.*) **29**:1 (1993), 51–59. MR Zbl

[40] L. Rudolph, "Knot theory of complex plane curves", pp. 349–427 in *Handbook of knot theory*, Elsevier B. V., Amsterdam, 2005. MR Zbl

[41] O. Saeki, "Elimination of definite fold", *Kyushu J. Math.* **60**:2 (2006), 363–382. MR Zbl

[42] M. Usher, "The Gromov invariant and the Donaldson–Smith standard surface count", *Geom. Topol.* **8** (2004), 565–610. MR Zbl

[43] J. Williams, "The h-principle for broken Lefschetz fibrations", *Geom. Topol.* **14**:2 (2010), 1015–1061. MR Zbl

Received 10 Feb 2021. Revised 26 Sep 2021.

MARK C. HUGHES: hughes@mathematics.byu.edu
Department of Mathematics, Brigham Young University, Provo, UT, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.185

Small exotic 4-manifolds and symplectic Calabi–Yau surfaces via genus-3 pencils

R. İnanç Baykur

We introduce a strategy to produce exotic rational and elliptic ruled surfaces, and possibly new symplectic Calabi–Yau surfaces, via constructions of symplectic Lefschetz pencils using a novel technique we call breeding. We deploy our strategy to breed explicit symplectic genus-3 pencils, whose total spaces are homeomorphic but not diffeomorphic to the rational surfaces $\mathbb{CP}^2\#p\overline{\mathbb{CP}}^2$ for $p = 6, 7, 8, 9$. Similarly, we breed explicit genus-3 pencils, whose total spaces are symplectic Calabi–Yau surfaces that have $b_1 > 0$ and realize all the integral homology classes of torus bundles over tori.

1. Introduction

Since the advent of Gauge theory, many construction techniques, such as knot surgery, rational blowdowns, generalized fiber sums and Luttinger surgery, have been introduced and successfully employed to produce exotic smooth structures on 4-manifolds, primarily through constructions of symplectic 4-manifolds homeomorphic but not diffeomorphic to smooth connected sums of standard 4-manifolds, where those with small topology (i.e., small second homology) have proven to be the most challenging.

In this article, we deploy a strategy to produce small symplectic 4-manifolds as total spaces of Lefschetz pencils,[1] which correspond to *small positive factorizations* (i.e., a small number of Dehn twists) we construct using a new technique we will discuss below. Recall that by the celebrated work of Donaldson [20] any compact symplectic 4-manifold admits a Lefschetz pencil, and in turn, corresponds to a positive factorization in the mapping class group of an orientable surface [44; 52; 54].

MSC2020: primary 57R55; secondary 57K20, 57K43.
Keywords: Lefschetz pencil, exotic 4-manifold, symplectic Calabi–Yau.

[1] Conventions: We assume that Lefschetz *pencils*, unlike Lefschetz fibrations, always have base points, whereas both have critical points and no exceptional spheres contained in the fibers.

For the small 4-manifolds we consider, the additional information presented by the pencil structure will be crucial to detect the exotic smooth structures, as illustrated by our first theorem:

Theorem A. *Let X be a symplectic 4-manifold homeomorphic to a rational or ruled surface Z with $c_1^2(Z) \geq 0$. Then X is an exotic Z if and only if it admits a genus-g Lefschetz pencil with number of base points $b \leq 2g-2-\chi_h(Z)$.*

Here $c_1^2 = 2\mathrm{e}+3\sigma$ is the first Chern number and $\chi_h = \frac{1}{4}(\mathrm{e}+\sigma)$ is the holomorphic Euler characteristic, where e and σ are the Euler characteristic and the signature of the 4-manifold. Rational and ruled surfaces satisfying the $c_1^2 \geq 0$ condition are the rational surfaces $\mathbb{CP}^2\# p\,\overline{\mathbb{CP}}^2$, for $p \leq 9$, and $S^2 \times S^2$ (which have $\chi_h = 1$), and the minimal elliptic ruled surfaces $T^2 \times S^2$ and $T^2 \tilde{\times} S^2$ (which have $\chi_h = 0$). The existence of the Lefschetz pencils in the statement of the theorem is granted by Donaldson, whereas our proof of the essential constraints on the topology of the pencils uses Seiberg–Witten theory, and builds on the works of Taubes [63; 64], McDuff [53] and Li and Liu [51]. We note that while there are numerous constructions of minimal symplectic 4-manifolds homeomorphic but not diffeomorphic to the rational surfaces $\mathbb{CP}^2\# p\overline{\mathbb{CP}}^2$, for $p \geq 2$, there are no known examples of exotic irrational ruled surfaces to date.

The homeomorphism types of the rational and elliptic ruled surfaces are easily determined by their fundamental group and intersection form by Freedman [30], and Hambleton and Kreck [40], respectively. Thus, powered by Theorem A, one can produce exotic copies of these small 4-manifolds by constructing Lefschetz pencils with the right algebraic invariants and small number of base points relative to the fiber genus. As a successful implementation of this approach, we show that:

Theorem B. *There are symplectic genus-3 Lefschetz pencils $\{(X_{i,\phi}, f_{i,\phi})\}$ whose total spaces have $\chi_h(X_{i,\phi}) = 1$ and $c_1^2(X_{i,\phi}) = 3-i$, and they include exotic rational surfaces $\mathbb{CP}^2\#(6+i)\overline{\mathbb{CP}}^2$ as well as infinitely many symplectic 4-manifolds which are not homotopy equivalent to any complex surface, for each $i = 0, 1, 2, 3$.*

The index ϕ for the family of pencils $\{(X_{i,\phi}, f_{i,\phi})\}$ takes values in a certain infinite subgroup of the mapping class group $\mathrm{Mod}(\Sigma_3^1)$ for each $i = 0, 1, 2, 3$.

Each symplectic 4-manifold $X_{i,\phi}$ in the theorem is "almost minimal", that is, it is either minimal or at most one blow-up of a minimal symplectic 4-manifold; see Remark 8. Notably, our family of genus-3 pencils with $c_1^2 = 3$ are all hyperelliptic, and therefore, by the work of Siebert-Tian [58], each $X_{0,\phi}$, including our exotic $\mathbb{CP}^2\# 6\,\overline{\mathbb{CP}}^2$, admits a symplectic involution and is a blow-down of a symplectic double branched covering of a rational surface; see Remark 10. We should also note that $g = 3$ is the smallest fiber genus for any Lefschetz pencil on an exotic rational surface, and we moreover suspect that our examples in Theorem B are also optimal in regard to the smallest exotic rational surfaces one can obtain via genus-3 pencils;

see Remark 9. While in this article we only study pencils of genus $g = 3$, one can obtain much sharper results even with $g = 4$ or 5 pencils, as demonstrated in our forthcoming work [11].

We describe our Lefschetz pencils in Theorem B in terms of their monodromy factorizations, which amount to positive Dehn twist factorizations of the boundary multitwist in the mapping class group of an orientable surface. We build these pencils out of lower genera pencils, using a novel technique we call *breeding*, which consists of carefully embedding the positive factorizations for lower genera pencils into the mapping class group of a higher genus surface in such a way that one can cancel all the negative Dehn twists (along nonboundary parallel curves) against positive ones at the end. It is worth noting that, although we use the breeding technique to derive new symplectic 4-manifolds from smaller ones, it is not an inherently symplectic operation. In the intermediary steps we get achiral Lefschetz pencils and fibrations which do contain negative nodes, but then we match them with positive nodes and remove all these pairs, which corresponds to surgering out self-intersection zero spheres contained in the fibers.

In an unpublished note with Korkmaz, we used a simpler version of the breeding technique to produce hyperelliptic genus-g Lefschetz fibrations with $5g - 3$ critical points, which yield the smallest hyperelliptic Lefschetz fibrations when $g = 3$. Since the appearance of the first version of this paper on the arxiv, the breeding technique has been used to produce several new Lefschetz pencils and fibrations (e.g., [38; 4; 10; 11]) and especially played a vital role in the recent resolution of Stipsicz's conjecture on the signature of Lefschetz fibrations in [10].

In the last portion of our paper, we turn to symplectic Calabi–Yau surfaces. Recall that a symplectic 4-manifold is called a *symplectic Calabi–Yau surface* if it has trivial canonical class, in obvious analogy with complex Calabi–Yau surfaces. The works of Li and Bauer established that any symplectic Calabi–Yau surface with $b_1 > 0$ has the rational homology type of a torus bundle over a torus [7; 48; 49], and it remains an open question whether torus bundles over tori exhaust all the diffeomorphism types of symplectic Calabi–Yau surfaces with $b_1 > 0$ [22; 49]. As stated by Li [50], a posteriori reasoning for an affirmative answer to this question often seems to stem from the lack of any new constructions of symplectic Calabi–Yau surfaces. The surgical operations like knot surgery, simplest rational blow-downs, generalized fiber sums or Luttinger surgery, do not produce any new symplectic Calabi–Yau surfaces [23; 43; 50; 66].

Akin to our strategy for producing exotic rational and elliptic ruled surfaces, in [12; 16] we implemented a strategy to construct (possibly new) symplectic Calabi–Yau surfaces via positive factorizations for pencils. The breeding technique, which is particularly effective for getting small positive factorizations, allows us to produce small symplectic Calabi–Yau surfaces as well:

Theorem C. *There are symplectic genus-3 Lefschetz pencils $\{(X_\phi, f_\phi)\}$ whose total spaces are symplectic Calabi–Yau surfaces that realize all integral homology types of torus bundles over tori, and they include a symplectic Calabi–Yau surface homeomorphic to the 4-torus and fake symplectic $T^2 \times S^2$s.*

The index ϕ for the family of pencils $\{(X_\phi, f_\phi)\}$ takes values in a certain infinite subgroup of the mapping class group $\mathrm{Mod}(\Sigma_3^4)$. A fake $T^2 \times S^2$ is a 4-manifold which has the same homology type as $T^2 \times S^2$ but is not diffeomorphic to it.

We describe the Lefschetz pencils in Theorem C in terms of their monodromy factorizations given in (47), which feeds into Donaldson's proposal of analyzing monodromies of pencils on symplectic Calabi–Yau surfaces [21, Problem 5]. These are the first explicit monodromy factorizations of pencils on symplectic Calabi–Yau surfaces with $b_1 > 0$ in the literature, whereas many examples on symplectic Calabi–Yau surfaces with $b_1 = 0$ were obtained in [12; 16]. Following the arxiv posting of an earlier version of this paper, similar examples were obtained by Hamada and Hayano in [38], also by employing the breeding technique.

Since symplectic Calabi–Yau surfaces with $b_1 > 0$ have the same Seiberg–Witten invariants as torus bundles over tori, detecting any new symplectic Calabi–Yau surfaces among $\{X_\phi\}$ hangs on essentially the possibility of detecting a $\pi_1(X_\phi)$ that is not a torus bundle group; see Remark 14. At the time of writing, we have not been able to determine whether all $\pi_1(X_\phi)$ we get are torus bundle groups. Likewise, we have not been able to spot any fake symplectic $T^2 \times S^2$ among $\{X_\phi\}$ with $\pi_1(X_\phi) = \mathbb{Z}^2$, which would make it homeomorphic to $T^2 \times S^2$, and thus an exotic elliptic ruled surface. (There are torus bundle over tori which have the same homology type as $T^2 \times S^2$.) On the other hand, Hamada and Hayano were able to show in [38] that our symplectic Calabi–Yau surface homeomorphic to the 4-torus is in fact diffeomorphic to it, by comparing our example with a holomorphic pencil on the standard 4-torus described by Smith; see Remark 15. While we do not know if any other X_ϕ is standard, it is worth noting that if our family of symplectic Calabi–Yau surfaces $\{X_\phi\}$ were to fully overlap with torus bundles over tori, then an additional feature of our construction would imply that any of these bundles can be equipped with a symplectic structure so that it is obtained via Luttinger surgeries from the standard 4-torus [43, Conjecture 4.9]; see Remark 16.

Outline of the paper: We review the basic definitions and preliminary results on Lefschetz pencils and fibrations, mapping class groups and positive factorizations, and symplectic 4-manifolds and Calabi–Yau surfaces in Section 2. In Section 3, we provide a characterization of small symplectic exotic rational surfaces (Theorem 3) and that of exotic minimal ruled surfaces (Theorem 6), which together give Theorem A. We breed our genus-3 pencils on exotic rational surfaces in Section 4, and on symplectic Calabi–Yau surfaces with $b_1 > 0$ in Section 5, which yield Theorem B and Theorem C, respectively.

2. Preliminaries

Here we quickly review the definitions and the basic properties of Lefschetz pencils and fibrations, Dehn twist factorizations in mapping class groups of surfaces, and symplectic 4-manifolds. The reader can turn to [13; 36; 47] for more details.

2.1. ***Lefschetz pencils and fibrations.*** A *Lefschetz pencil* on a closed, smooth, oriented 4-manifold X is a smooth surjective map $f : X \setminus \{b_j\} \to S^2$, defined on the complement of a nonempty finite collection of points $\{b_j\}$, such that around every *base point* b_j and *critical point* p_i there are local complex coordinates (compatible with the orientations on X and S^2) with respect to which the map f takes the forms $(z_1, z_2) \mapsto z_1/z_2$ and $(z_1, z_2) \mapsto z_1 z_2$, respectively. A *Lefschetz fibration* is defined similarly for $\{b_j\} = \varnothing$. Blowing-up at each base point b_j of a pencil (X, f), one obtains a Lefschetz fibration $(\widetilde{X}, \widetilde{f})$ with disjoint (-1)-sphere sections S_j corresponding to each b_j, and conversely, blowing down disjoint (-1)-sphere sections of a Lefschetz fibration, one obtains a pencil.

We say (X, f) is a *genus-g Lefschetz pencil* or *fibration* for g the genus of a regular fiber F of f. The fiber containing the critical point p_i has a nodal singularity at p_i, which locally arises from shrinking a simple loop c_i on F, called a *vanishing cycle*. A singular fiber of (X, f) is called *reducible* if c_i is separating. When c_i is null-homotopic on F, one of the fiber components becomes an *exceptional sphere*, an embedded 2-sphere of self-intersection -1, which one can blow down without altering the rest of the fibration.

In this paper we use the term Lefschetz fibration only when the set of critical points $\{p_i\}$ is nonempty, i.e., when the Lefschetz fibration is nontrivial. We moreover assume that the fibration is *relatively minimal*, i.e., there are no exceptional spheres contained in the fibers, and also that the critical points p_i lie in distinct singular fibers, which can be always achieved after a small perturbation.

Allowing the local model $(z_1, z_2) \mapsto z_1 \bar{z}_2$ around the critical points p_i, which give rise to negative nodes, all of the above notions extend to so-called *achiral* Lefschetz pencils and fibrations.

2.2. ***Positive factorizations.*** Let Σ_g^b denote a compact, connected, oriented surface of genus g with b boundary components, and simply write Σ_g when there is no boundary. We denote by $\mathrm{Mod}(\Sigma_g^b)$ its *mapping class group*; the group composed of orientation-preserving self-homeomorphisms of Σ_g^m which restrict to the identity along $\partial\Sigma_g^b$, modulo isotopies that also restrict to the identity along $\partial\Sigma_g^b$. Let $\mathrm{Mod}(\Sigma_g^b, S)$ denote the stabilizer subgroup of $\mathrm{Mod}(\Sigma_g^b)$ which consists of elements fixing the subset $S \subset \Sigma_g^b$ pointwise. Denote by $t_c \in \mathrm{Mod}(\Sigma_g^b)$ the positive (right-handed) Dehn twist along the simple closed curve $c \subset \Sigma_g^m$. Its inverse t_c^{-1} is the negative (left-handed) Dehn twist along c.

Let $\{c_i\}$ be a nonempty collection of simple closed curves on Σ_g^b, which do not become null-homotopic when $\partial\Sigma_g^b$ is capped off by disks, and let $\{\delta_j\}$ be a collection of curves parallel to distinct boundary components of Σ_g^b. If the relation

$$t_{c_l}\cdots t_{c_2}t_{c_1} = t_{\delta_1}\cdots t_{\delta_b} \tag{1}$$

holds in $\mathrm{Mod}(\Sigma_g^b)$, we call the word W on the left-hand side a *positive factorization* of the *boundary multitwist* $\Delta = t_{\delta_1}\cdots t_{\delta_b}$ in Γ_g^b. (We will also use ∂_i instead of δ_i when there are several surfaces with boundaries involved in our discussion.) Capping off all the boundary components of Σ_g^b with disks induces a homomorphism $\mathrm{Mod}(\Sigma_g^b)\to\mathrm{Mod}(\Sigma_g)$, under which W maps to a similar positive factorization of the identity element $1\in\mathrm{Mod}(\Sigma_g)$.

The positive factorization in (1) gives rise to a genus-g Lefschetz fibration $(\widetilde{X},\widetilde{f})$ with b disjoint (-1)-sections S_j, and therefore a genus-g Lefschetz pencil (X,f) with b base points. Identifying the regular fiber F with Σ_g, we can view the vanishing cycles of the fibration as the Dehn twist curves $\{c_i\}$. Every Lefschetz pencil and fibration can be described by such a positive factorization, which is called its *monodromy factorization* [36; 44; 52].

Let W be a positive factorization of the form $W = PP'$ in $\mathrm{Mod}(\Sigma_g^b)$, where P and P' are some products of positive Dehn twists along curves which do not become null-homotopic when $\partial\Sigma_g^b$ is capped off. If $P=\Pi_i t_{c_i}$, as a mapping class, commutes with some element $\phi\in\mathrm{Mod}(\Sigma_g^b)$, we can then produce a new positive factorization $W_\phi = P^\phi P'$, where P^ϕ denotes the conjugate factorization $\phi P\phi^{-1} = \Pi_i(\phi t_{c_i}\phi^{-1}) = \Pi_i t_{\phi(c_i)}$. In this case, we say W_ϕ is obtained from W by a partial conjugation ϕ along P.

Allowing negative Dehn twists, which correspond to negative nodes, we can more generally work with factorizations for achiral Lefschetz fibrations and pencils. All of the above definitions and results extend to this more general setting.

2.3. ***Symplectic* 4*-manifolds and Kodaira dimension.*** It was shown by Donaldson that every symplectic 4-manifold (X,ω) admits a symplectic Lefschetz pencil whose fibers are symplectic with respect to ω [20]. Conversely, generalizing a construction of Thurston, Gompf showed that the total space of a Lefschetz pencil and fibration always admits a symplectic form ω with respect to which all regular fibers and any preselected collection of disjoint sections are symplectic [36]. Whenever we take a symplectic form ω on a Lefschetz pencil or fibration (X,f), we will assume it is of Thurston–Gompf type, with respect to which any explicitly mentioned sections will be assumed to be symplectic as well.

The Kodaira dimension for projective surfaces can be extended to symplectic 4-manifolds as follows: Let $K_{X_{\min}}$ be the canonical class of a minimal model $(X_{\min},\omega_{\min})$ of (X,ω). The *symplectic Kodaira dimension* of (X,ω), denoted

by $\kappa = \kappa(X, \omega)$ is then defined as

$$\kappa(X,\omega) = \begin{cases} -\infty & \text{if } K_{X_{\min}} \cdot [\omega_{\min}] < 0 \text{ or } K^2_{X_{\min}} < 0, \\ 0 & \text{if } K_{X_{\min}} \cdot [\omega_{\min}] = K^2_{X_{\min}} = 0, \\ 1 & \text{if } K_{X_{\min}} \cdot [\omega_{\min}] > 0 \text{ and } K^2_{X_{\min}} = 0, \\ 2 & \text{if } K_{X_{\min}} \cdot [\omega_{\min}] > 0 \text{ and } K^2_{X_{\min}} > 0. \end{cases}$$

Remarkably, not only is κ independent of the minimal model $(X_{\min}, \omega_{\min})$ but also it is independent of the particular symplectic form ω on X; so it is a smooth invariant of the 4-manifold X [49]. Symplectic 4-manifolds with $\kappa = -\infty$ are classified up to symplectomorphisms, which are precisely the rational and ruled surfaces [47].

Symplectic 4-manifolds with $\kappa = 0$, which are the analogues of the Calabi–Yau surfaces, are those with torsion canonical class [49]. It was shown by Tian-Jun Li, and independently by Stefan Bauer [49; 7], that the rational homology type of any minimal symplectic 4-manifold with $\kappa = 0$ is that of a torus bundle over a torus, the K3 surface or the Enrique surfaces. In the first two cases we have symplectic Calabi–Yau surfaces, which have trivial canonical class, whereas in the last case the canonical class is torsion.

We have the following topological characterization of Lefschetz pencils on minimal symplectic 4-manifolds with $\kappa = 0$, which can be easily derived from the more general characterization for Lefschetz fibrations on symplectic 4-manifolds with $\kappa = 0$ given in [12, Theorem 4.1], [57, Theorem 5.12]:

Proposition 1. *Let (X, f) be a genus-g Lefschetz pencil with b base points, where X is neither rational nor ruled. Then there is a symplectic form ω on X so that (X, ω) is a symplectic Calabi–Yau or a rational homology Enriques surface if and only if $b = 2g - 2$.*

3. Topology of pencils on rational and elliptic ruled surfaces

In this section we will prove two theorems that might be of independent interest; one on the topology of Lefschetz pencils and fibrations on (small) rational surfaces, and one on (small) irrational ruled surfaces. These results enable one to tackle producing exotic smooth structures on the rational surfaces $\mathbb{CP}^2 \# p\,\overline{\mathbb{CP}}^2$, $S^2 \times S^2$, and the minimal elliptic ruled surfaces $T^2 \times S^2$ and $T^2 \tilde{\times} S^2$, via constructions of new positive factorizations, as we will try to demonstrate in the later sections.

3.1. ***Lefschetz pencils and fibrations on rational surfaces.*** We first prove the following lemma, which shows that pencils on rational surfaces always have a lot of base points with respect to the fiber genera:

Lemma 2. *The rational surfaces $\mathbb{CP}^2 \# p\,\overline{\mathbb{CP}}^2$, for $p \le 9$, or $S^2 \times S^2$, do not admit any genus-g pencil with $b < 2g - 2$ base points or any Lefschetz fibration of $g \ge 2$.*

Proof. We claim that the statement of the lemma holds even for nonrelatively minimal pencils and fibrations. With this in mind, it suffices to prove our claim for $X = \mathbb{CP}^2 \# 9\,\overline{\mathbb{CP}}^2$, because we can blow-up on the fibers of a given genus-g Lefschetz pencil or fibration on $\mathbb{CP}^2 \# p\,\overline{\mathbb{CP}}^2$ with $p < 9$ or $S^2 \times S^2$ to get one on X.

Now suppose for contradiction that $X = \mathbb{CP}^2 \# 9\,\overline{\mathbb{CP}}^2$ admits a genus-g pencil with $b < 2g-2$ base points or a Lefschetz fibration of genus $g \geq 2$. Note that in either case, $g \geq 2$. For our arguments to follow, it will be convenient to allow b to be a nonnegative integer so that $b = 0$ marks the fibration case.

Let $F = aH - \sum_{i=1}^{9} c_i E_i$ be the fiber class, where $H_2(X)$ is generated by the hyperplane class H and the exceptional classes $E_1, \ldots, E_9$, which satisfy $H^2 = 1$, $E_i \cdot E_j = -\delta_{ij}$, and $H \cdot E_i = 0$. Since $F^2 = b$, we have

$$a^2 = b + \sum_{i=1}^{9} c_i^2.$$

We can equip X with a Thurston–Gompf symplectic form ω which makes the fibers symplectic. Moreover, we can choose an ω-compatible almost complex structure J, even a generic one in the sense of Taubes, with respect to which the pencil/fibration is J-holomorphic for a suitable choice of almost complex structure on the base 2-sphere; see, e.g., [65]. It was shown by Li and Liu [51] that for a generic ω-compatible J, the class H in the rational surface X has an embedded J-holomorphic representative. Hence, F and H both have J-holomorphic representatives, which implies that $F \cdot H = a \geq 0$.

Since there is a unique symplectic structure on X up to deformation and symplectomorphisms [51], we can apply the adjunction formula to get

$$2g-2 = F^2 + K \cdot F = b + \left(-3H + \sum_{i=1}^{9} E_i\right) \cdot \left(aH - \sum_{i=1}^{9} c_i E_i\right) = b - 3a + \sum_{i=1}^{9} c_i.$$

Since $a, b \geq 0$, and $g \geq 2$, from the above equalities we have

$$\begin{aligned} 3a = \sqrt{9a^2} &= \sqrt{9\big(b + \textstyle\sum_{i=1}^{9} c_i^2\big)} \geq \sqrt{9\big(\textstyle\sum_{i=1}^{9} c_i^2\big)} \\ &= \sqrt{\big(\textstyle\sum_{i=1}^{9} 1\big)\big(\textstyle\sum_{i=1}^{9} c_i^2\big)} \geq \sqrt{\big|\textstyle\sum_{i=1}^{9} c_i\big|^2}, \end{aligned}$$

where the last inequality is by Cauchy–Schwarz. In turn, we get:

$$3a \geq \sqrt{\big|\textstyle\sum_{i=1}^{9} c_i\big|^2} = \big|\textstyle\sum_{i=1}^{9} c_i\big| = |2g-2-b+3a| = 2g-2-b+3a,$$

which implies that $b \geq 2g-2$. The contradiction shows that there is no such fiber class F. In turn, there is no such Lefschetz pencil or fibration. □

The statement as stated is obviously not true for $p > 9$; for example there is a genus-2 Lefschetz fibration on $\mathbb{CP}^2 \# 13\,\overline{\mathbb{CP}}^2$ (which in fact is the blow-up of

a genus-2 pencil on $S^2 \times S^2$). Otherwise, one can generalize the above result to rational surfaces $\mathbb{CP}^2 \# \, p\, \overline{\mathbb{CP}}^2$ with $p > 9$, under particular assumptions for b and g with respect to the number of blow-ups p.

Any symplectic, exotic rational surface would admit a pencil of genus $g \geq 2$ by Donaldson's result and by the fact that only rational surfaces admit genus-0 or genus-1 pencils. On the other hand, the regular fiber of a pencil of genus $g \geq 2$ with $b \geq 2g-2$ base points would violate the Seiberg–Witten adjunction inequality (which holds for any symplectic 4-manifold that is not a rational or a ruled surface). We can thus conclude that:

Theorem 3. *A symplectic 4-manifold X in the homeomorphism class of $\mathbb{CP}^2 \# \, p\, \overline{\mathbb{CP}}^2$ with $p \leq 9$ or $S^2 \times S^2$ is an exotic rational surface if and only if it admits a genus-g pencil with $b < 2g-2$ base points or a Lefschetz fibration of genus $g \geq 2$.*

As we mentioned earlier, there are numerous examples of symplectic 4-manifolds homeomorphic but not diffeomorphic to $\mathbb{CP}^2 \# \, p\, \overline{\mathbb{CP}}^2$, for $2 \leq p \leq 9$, and they should all admit genus-g pencils with $b < 2g-2$ base points by the above theorem. However, in the literature there appear to be no examples of Lefschetz pencils (with base points, no multiple fibers) on these 4-manifolds, even on the complex algebraic ones. We will provide some novel symplectic examples admitting genus-3 pencils in the next section.

As for fibrations, for K any fibered knot of genus $g \geq 1$, knot surgered elliptic surfaces $E(1)_K$ of Fintushel and Stern yield exotic $E(1) = \mathbb{CP}^2 \# \, 9\, \overline{\mathbb{CP}}^2$, which admit symplectic genus-$2g$ Lefschetz fibrations [28]. Moreover, there are genus-2 symplectic Lefschetz fibrations in the homeomorphism classes of $\mathbb{CP}^2 \# \, p\, \overline{\mathbb{CP}}^2$ for $p = 7, 8, 9$ [13] and even holomorphic ones for $p = 8, 9$ [56].

Remark 4. When X is an exotic $\mathbb{CP}^2 \# \, p\, \overline{\mathbb{CP}}^2$, with $p \leq 8$, we can strengthen the statement of Theorem 3 a bit. If the pencil of genus $g \geq 2$ on X had $b = 2g-3$ base points, blowing up all of them, we would get a Lefschetz fibration with b disjoint (-1)-sphere sections. It then follows from [57, Theorem 5-12] that $K^2_{X_{\min}} = 0$, which cannot be the case here since $K^2_{X_{\min}} \geq K^2_X = 9 - p > 0$. Hence, any pencil on such an exotic rational surface X can have at most $2g-4$ base points.

3.2. ***Lefschetz pencils and fibrations on minimal elliptic ruled surfaces.*** We now show that pencils on minimal elliptic ruled surfaces also have a lot of base points with respect to the fiber genera:

Lemma 5. *The minimal elliptic ruled surfaces $T^2 \times S^2$ or $T^2 \tilde{\times} S^2$ do not admit any genus-g Lefschetz pencil with $b \leq 2g-2$ base points or any Lefschetz fibration.*

Proof. These minimal elliptic surfaces do not admit Lefschetz fibrations of genus $g < 2$ for fairly elementary reasons (which do not require classification results): any genus-0 Lefschetz fibration has a simply connected total space, and the Euler

characteristic of any genus-1 Lefschetz fibration is equal to the number of critical points, and therefore it is positive. Clearly, neither one of these two implications work for $T^2 \times S^2$ or $T^2 \tilde{\times} S^2$.

Now suppose for a contradiction that $X = T^2 \times S^2$ or $T^2 \tilde{\times} S^2$ admits a genus-g pencil with $b \leq 2g-2$ base points or a Lefschetz fibration. Once again, it will be convenient here to let b be a nonnegative integer so that $b = 0$ marks the fibration case. By our observation in the previous paragraph, we can assume that $g \geq 2$.

We equip X with a Thurston–Gompf symplectic form ω which makes the fibers, and in particular a regular fiber F, of the pencil/fibration symplectic. We can choose an ω-compatible almost complex structure J with respect to which the pencil/fibration is J-holomorphic, so in particular F is a J–holomorphic curve. Because there is a unique symplectic structure on a minimal ruled surface up to deformations and symplectomorphisms [51], we will be able to once again apply the adjunction formula using a standard canonical class in each case. Furthermore, it will be important for our arguments that it was also shown in [51] that for any ω-compatible almost complex structure J, the sphere fiber of the ruling on the elliptic surface has a J–holomorphic representative. Therefore the algebraic intersection of F with the sphere fiber is nonnegative. Akin to our proof of Lemma 2, we will show that neither one of the minimal elliptic ruled surfaces contains an embedded genus-g symplectic surface with self-intersection $\leq 2g-2$, whereas F is such.

We will run our arguments for the spin and nonspin cases separately:

$\underline{X = T^2 \times S^2}$: Here $H_2(X) \cong \mathbb{Z}^2$ is generated by $S = \{pt\} \times S^2$ and $T = T^2 \times \{pt\}$, where $S \cdot S = 0$, $T \cdot T = 0$, and $S \cdot T = 1$. By a slight abuse of notation, we denote the homology class of the fiber also by F, so $F = xS + yT$ for some $x, y \in \mathbb{Z}$.

As remarked above, the algebraic intersection of F with S is nonnegative, which means that $F \cdot S = y \geq 0$. Since $F^2 = b$, we have

$$b = 2xy,$$

where $b \geq 0$ and $y \geq 0$ imply that $x \geq 0$.

On the other hand, by the adjunction formula we get

$$2g-2 = F^2 + K_X \cdot F = b + (-2T) \cdot (xS + yT) = b - 2x,$$

which implies that $2x = b - (2g-2) \leq 0$ by our assumption on b. It follows that $x = 0$, and in turn, $b = 0$ by the first equality, and $g = 1$ by the second, which is a contradiction.

$\underline{X = T^2 \tilde{\times} S^2}$: Now $H_2(X) \cong \mathbb{Z}^2$ is generated by the fiber S and section T of the degree–1 ruling on X, where $S \cdot S = 0$, $T \cdot T = 1$, and $S \cdot T = 1$. Let the fiber class F be given by $F = xS + yT$, for some $x, y \in \mathbb{Z}$.

Since the algebraic intersection of F with S is nonnegative, we have $F\cdot S = y \geq 0$. Since $F^2 = b$, we now have

$$b = 2xy + y^2 = (2x+y)y,$$

where $b \geq 0$ and $y \geq 0$ imply that $2x + y \geq 0$.

We apply the adjunction formula to get

$$2g-2 = F^2 + K_X \cdot F = b + (S-2T)\cdot(xS+yT) = b - 2x - y,$$

which means that $2x + y = b - (2g-2) \leq 0$. As we also had $2x + y \geq 0$, it follows that $2x + y = 0$, and then $b = 0$ by the first equality, and $g = 1$ by the second, which once again contradicts our assumption on the fiber genus — that g is greater than or equal to 2. □

The above result, at least as stated, does not generalize to pencils on other ruled surfaces. First of all, there exist genus-g Lefschetz pencils with $b = 2g-2$ base points on nonminimal elliptic ruled surfaces, even after a single blow-up; an example with $g = b = 2$ can be found in the next section. Secondly, there are pencils on the minimal ruled surfaces $\Sigma_h \times S^2$ and $\Sigma_h \tilde{\times} S^2$ with fiber genus $g = 2h$ and $b = 4$ base points [37], so the statement fails for any $h \geq 2$ in both spin and nonspin cases.

Using Donaldson's result on the existence of Lefschetz pencils on symplectic 4-manifolds, we moreover conclude that:

Theorem 6. *A symplectic 4-manifold X in the homeomorphism class of $T^2 \times S^2$ or $T^2 \tilde{\times} S^2$ is an exotic elliptic ruled surface if and only if it admits a genus-g pencil with $b \leq 2g-2$ base points or a Lefschetz fibration.*

It is worth noting that to this date there are no known examples of exotic elliptic ruled surfaces, despite their topological types being amenable to Freedman type arguments [41]. While we plan to explore this direction elsewhere, in Section 5, through positive factorizations for pencils, we will provide examples of fake symplectic elliptic ruled surfaces, which have the same cohomology as $T^2 \times S^2$, but are not diffeomorphic to it.

4. Exotic rational surfaces via symplectic genus-3 pencils

Here we construct positive factorizations for symplectic genus-3 Lefschetz pencils, whose total spaces are homeomorphic but not diffeomorphic to rational surfaces. These will be bred from genus-2 pencils on elliptic ruled surfaces. For a better exposition, we first present our examples with $\chi_h = 1$ and $c_1^2 = 0, 1, 2$, and we discuss our examples with $\chi_h = 1$ and $c_1^2 = 3$, whose constructions are a bit more involved, afterwards.

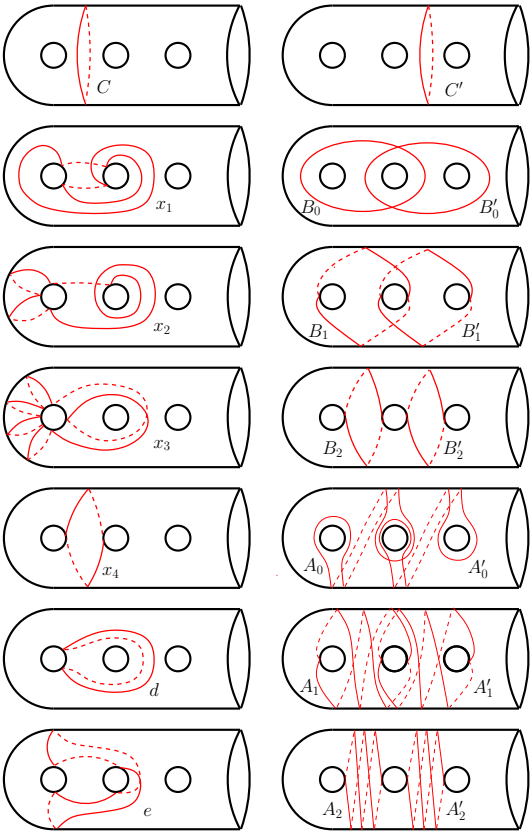


Figure 1. The curves C, x_1, x_2, x_3, x_4, d, e of the first embedding are given on the left. The curves B_0, B_1, B_2, A_0, A_1, A_2 of the second embedding (except for C, which is already given on the left) and C', B_0', B_1', B_2', A_0', A_1', A_2' of the third embedding are on the right.

4.1. ***Breeding pencils with*** $\chi_h = 1$ ***and*** $c_1^2 = 0, 1, 2$**.** In [13], Korkmaz and the author obtained the following relation in $\mathrm{Mod}(\Sigma_2^1)$:

$$t_e t_{x_1} t_{x_2} t_{x_3} t_d t_C t_{x_4} = t_\delta,$$

which is a positive factorization for a genus-2 pencil with one base point on $T^2 \times S^2 \# 2\,\overline{\mathbb{CP}}^2$. See Figure 2 for the curves x_i, C, d, e (where the boundary component δ is obtained by carving out a disk neighborhood of the marked point on right end of the surface). Consider the embedding of Σ_2^1 into Σ_3^1 given by mapping the boundary $\delta = \partial \Sigma_2^1$ to the curve C', and the remaining Dehn twist curves x_i, C, d, e to the ones shown in Figure 1, denoted by the same letters. After a single Hurwitz move (see, e.g., [36]), and collecting all the Dehn twists on the

same side, we get the following relation in $\mathrm{Mod}(\Sigma_3^1)$:

$$t_e t_{x_1} t_{x_2} t_{x_3} t_d t_{B_2} t_C t_{C'}^{-1} = 1, \tag{2}$$

where $B_2 = t_C(x_4)$. Rewrite this relation as $P_1 t_C t_{C'}^{-1} = 1$, for $\underline{P_1 = t_e t_{x_1} t_{x_2} t_{x_3} t_d t_{B_2}}$. Note that P_1, t_C and $t_{C'}$ all commute with each other.

Next, we take the following lift of the positive factorization for Matsumoto's genus-2 Lefschetz fibration to $\mathrm{Mod}(\Sigma_2^2)$ obtained by Hamada in [37]:

$$(t_{B_0} t_{B_1} t_{B_2} t_C)^2 = t_{\delta_1} t_{\delta_2},$$

where δ_i are the boundary parallel curves, and the curves B_i and C are as shown on the left-hand side of Figure 6. This is a positive factorization for a genus-2 pencil with two base points on $T^2 \times S^2 \# 2\,\overline{\mathbb{CP}}^2$. After Hurwitz moves, and collecting all the Dehn twists on the same side, we get the following relation in $\mathrm{Mod}(\Sigma_2^2)$:

$$t_{B_0} t_{B_1} t_{B_2} t_{A_0} t_{A_1} t_{A_2} t_C^2 t_{\delta_1}^{-1} t_{\delta_2}^{-1} = 1,$$

where each $A_j = t_C(B_j)$, for $j = 0, 1, 2$, are as shown in Figure 6. We will describe two different embeddings of this relation into $\mathrm{Mod}(\Sigma_3^1)$.

Cap off the boundary component δ_1 of Σ_2^2, and then embed the resulting copy of Σ_2^1 into Σ_3^1 via the embedding we used to derive the relation (2) above, so the boundary δ_2 is mapped to C', and all the other Dehn twist curves are as shown in Figure 1, once again denoted by the same letters. So we have the following relation in $\mathrm{Mod}(\Sigma_3^1)$:

$$t_{B_0} t_{B_1} t_{B_2} t_{A_0} t_{A_1} t_{A_2} t_C^2 t_{C'}^{-1} = 1, \tag{3}$$

which we rewrite as $P_2 t_C^2 t_{C'}^{-1} = 1$, for $\underline{P_2 = t_{B_0} t_{B_1} t_{B_2} t_{A_0} t_{A_1} t_{A_2}}$. Here P_2, t_C and $t_{C'}$ all commute with each other.

Lastly, consider an embedding of Σ_2^2 into Σ_3^1 so that δ_1 is mapped to c, δ_2 is mapped to $\partial = \partial \Sigma_3^1$, and the remaining curves are as shown in Figure 1, where we use a prime symbol when denoting the curves by the same letters. This gives a third relation in $\mathrm{Mod}(\Sigma_3^1)$:

$$t_{B_0'} t_{B_1'} t_{B_2'} t_{A_0'} t_{A_1'} t_{A_2'} t_{C'}^2 t_C^{-1} = t_\partial, \tag{4}$$

which we rewrite as $P_2' t_{C'}^2 t_C^{-1} = t_\delta$, for $\underline{P_2' = t_{B_0'} t_{B_1'} t_{B_2'} t_{A_0'} t_{A_1'} t_{A_2'}}$. Similarly, P_2', t_C and $t_{C'}$ all commute with each other.

With these three embeddings in hand, we can now describe our positive factorizations. Let ϕ be any mapping class in $\mathrm{Mod}(\Sigma_3^1, S)$, the subgroup of $\mathrm{Mod}(\Sigma_3^1)$ which consists of elements fixing the set $S := \{C, C'\}$ pointwise. Then we have

$$(P_1)^\phi P_1 P_2' t_C = (P_1 t_C t_{C'}^{-1})^\phi P_1 t_C t_{C'}^{-1} P_2' t_{C'}^2 t_C^{-1} = 1 \cdot t_\partial \cdot 1 = t_\partial,$$

where the first equality follows from the commutativity relations noted above

and the fact that ϕ commutes with $t_C t_{C'}^{-1}$. The second equality follows from the relations (2)–(4). Therefore

$$\underline{W_{1,\phi} = (P_1)^{\phi} P_1 P_2' t_C}$$

is a positive factorization of the boundary twist t_∂ in $\mathrm{Mod}(\Sigma_3^1)$. By identical arguments, we see

$$\underline{W_{2,\phi} = (P_1)^{\phi} P_2 P_2' t_C^2} \quad \text{and} \quad \underline{W_{3,\phi} = (P_2)^{\phi} P_2 P_2' t_C^3}$$

are also positive factorizations of t_∂ in $\mathrm{Mod}(\Sigma_3^1)$.

Each $W_{i,\phi}$ prescribes a symplectic genus-3 Lefschetz pencil $(X_{i,\phi}, f_{i,\phi})$ with one base point, equipped with a Thurston–Gompf symplectic form. We claim that $\chi_h(X_{i,\phi}) = 1$ and $c_1^2(X_{i,\phi}) = 3 - i$ for each $i = 1, 2, 3$.

The Euler characteristic of $X_{i,\phi}$ is given by

$$\mathrm{e}(X_{i,\phi}) = 4 - 4g + \ell - b = 4 - 4 \cdot 3 + (18 + i) - 1 = 9 + i,$$

where g and b are the genus and the number of base points of the pencil, and ℓ is the number of critical points, which is the same as the number of Dehn twists in the positive factorization $W_{i,\phi}$.

Since we have explicit positive factorizations for the pencils $(X_{i,\phi}, f_{i,\phi})$, the signature of each $X_{i,\phi}$ can be easily calculated using the work of Endo and Nagami [24], which states that the signature of the pencil is equal to the algebraic sum of the signatures of the mapping class group relations used to derive this positive factorization from the trivial word in $\mathrm{Mod}(\Sigma_3^1)$. Since the signature of any embedding of a relation into a higher genus surface is the same, and since Hurwitz moves, conjugations and cancellations of positive–negative Dehn twist pairs do not change the signature, we just need to understand the signatures of the genus-2 relations we used as our building blocks. The signature of the relation (2) is the same as the signature of the genus-2 pencil with one base point on $T^2 \times S^2 \# 2\,\overline{\mathbb{CP}}^2$, which is -2. The signature of the relation (3) is that of the genus-2 pencil with one base point on $T^2 \times S^2 \# 3\,\overline{\mathbb{CP}}^2$ (recall that we capped off one of the boundaries first), which is -3. Finally, the signature of the relation (4) is that of the genus-2 pencil with two base points on $T^2 \times S^2 \# 2\,\overline{\mathbb{CP}}^2$, which is -2. We conclude that $\sigma(X_{i,\phi}) = -5 - i$.

Hence, $\chi_h(X_{i,\phi}) = \frac{1}{4}(\mathrm{e}(X_{i,\phi}) + \sigma(X_{i,\phi})) = \frac{1}{4}(9 + i - 5 - i)) = 1$ for each $i = 1, 2, 3$, whereas $c_1^2(X_{i,\phi}) = 2\mathrm{e}(X_{i,\phi}) + 3\sigma(X_{i,\phi}) = 2(9 + i) + 3(-5 - i) = 3 - i$, as claimed. Note that the only rational or ruled surfaces which have the same invariants are the rational surfaces $\mathbb{CP}^2 \# (6 + i)\overline{\mathbb{CP}}^2$, which by Lemma 2, cannot admit such pencils.

4.2. ***Breeding pencils with $\chi_h = 1$ and $c_1^2 = 3$.*** In our next construction we strive to get *hyperelliptic* pencils. While getting hyperelliptic positive factorizations at every

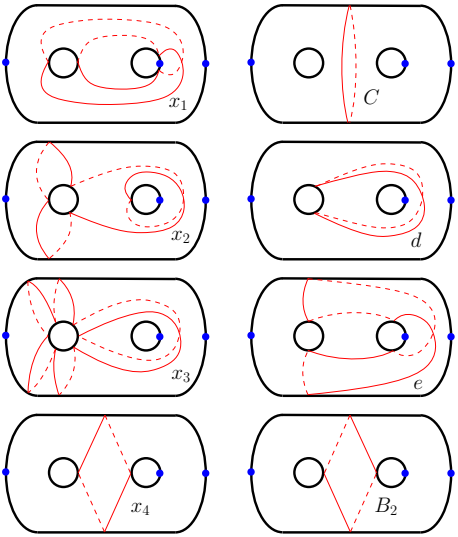


Figure 2. The curves C, x_1, x_2, x_3, x_4, d, e in the lift of Baykur–Korkmaz genus-2 positive factorization to $\mathrm{Mod}(\Sigma_2^3)$, along with the curve B_2 one gets after a Hurwitz move

step will constrain some of the freedom we have in our breeding constructions, we will also leverage this additional property whenever we can.

Our construction will be comparable to that of the positive factorization $W_{3,\phi}$ in the previous section, where we employed three embeddings of (various lifts of) the positive factorization for Matsumoto's genus-2 Lefschetz fibration. Here we will get our examples using three different embeddings of (various lifts of) the positive factorization for the genus-2 Lefschetz fibration of Korkmaz and the author in [13]. This positive factorization, after a single Hurwitz move as before, has the following lift in $\mathrm{Mod}(\Sigma_2^3)$:

$$t_e t_{x_1} t_{x_2} t_{x_3} t_d t_{B_2} t_C = t_{\delta_1} t_{\delta_2} t_{\delta_3}, \tag{5}$$

where the curves x_i, B_2, C, d, e are as shown in Figure 2. We will simply use the same labels for the Dehn twist curves x_i, B_2, C, d, e for any other relation we derive from (5) by capping off some of the boundary components δ_1, δ_2, δ_3.

A comprehensive proof of the relation (5) is given in [11], where it is also shown that this is a positive factorization for a genus-2 pencil with two marked points on the elliptic ruled surface $T^2 \tilde{\times} S^2$. It can also be verified in a straightforward fashion using the Alexander method [25]. Below we sketch yet another argument based on the hyperelliptic symmetry of the monodromy curves. This line of arguments can be proved to be useful for similar calculations in general.

Let (X, f) be the hyperelliptic genus-2 Lefschetz fibration corresponding to the positive factorization $t_e t_{x_1} t_{x_2} t_{x_3} t_d t_{B_2} t_C = 1$ in $\mathrm{Mod}(\Sigma_2)$, where $X \cong T^2 \times S^2 \# 3\,\overline{\mathbb{CP}}^2$

is equipped with a Thurston–Gompf symplectic form [13]. As shown in [33; 58], there is a symplectic involution on X extending the hyperelliptic involution on the fibers, and f is the relative minimalization of a Lefschetz fibration obtained via the induced symplectic double branched cover $X\#\,3\,\overline{\mathbb{CP}}^2 \to S^2\times S^2\#\,6\,\overline{\mathbb{CP}}^2$ (where the blow-ups are for the reducible fibers). The branch set consists of a multisection B of the latter fibration, which intersects every fiber at the fixed points of the hyperelliptic involution, and B_E that consists of exceptional spheres contained in the reducible fibers. Now, observe that when we isotope the monodromy curves of (X, f) so that they are symmetric under the obvious hyperelliptic involution obtained by rotating the surface Σ_2 in Figure 2 by a π–degree rotation along the x–axis (taking the z–axis to be perpendicular to the page), they miss the three marked points (drawn in blue in the figure) of the fixed points of the hyperelliptic involution, whereas the four nonseparating curves go through the other three points. We can deduce the topology of the branch set from this very data, and in particular conclude that the multisection B consists of three disjoint (-1)-sphere sections E_1, E_2, E_3 (one for each marked point) and a 3-section which is a square zero symplectic 2-sphere (going through the other three fixed points). Circling back to our original discussion, the (-1)-sections E_1, E_2, E_3 yield the lift (5).

We are now ready to describe our three embeddings.

Note that we have now drawn the surface Σ_3^1 so that its boundary curve ∂ is as shown in Figure 3. With this in mind, our first embedding is essentially the same as the one yielded by the relation (2) in $\mathrm{Mod}(\Sigma_3^1)$: Cap off the boundary components δ_1 and δ_2 of Σ_2^3 and then embed it into Σ_3^1 so that δ_3 maps to C' and the rest of the Dehn twist curves are as shown in Figure 1, except the boundary $\partial\Sigma_3$, which is outside of their support, is shifted. Using the same notation as before, we get the relation $P_1 t_C t_{C'}^{-1} = 1$ in $\mathrm{Mod}(\Sigma_3^1)$, where $\underline{P_1 = t_e t_{x_1} t_{x_2} t_{x_3} t_d t_{B_2}}$.

For our second embedding, cap off the boundary component δ_3 of Σ_2^3, and then embed the resulting copy of Σ_2^2 into Σ_3^1 so that

$$\delta_1 \mapsto C, \quad \delta_2 \mapsto \partial, \quad C \mapsto C',$$

and the remaining curves are as shown in Figure 3, where we once again use a prime symbol when denoting the curves by the same letters. For our arguments to follow, here it is more convenient to take the genus-2 relation as $t_{x_4} t_e t_{x_1} t_{x_2} t_{x_3} t_d t_C = t_{\delta_1} t_{\delta_2}$, where we moved t_{B_2} back over t_C by a Hurwitz move and then applied a cyclic permutation. So we get the following relation in $\mathrm{Mod}(\Sigma_3^1)$:

$$t_{x_4'} t_{e'} t_{x_1'} t_{x_2'} t_{x_3'} t_{d'} t_{C'} t_C^{-1} = t_\partial, \tag{6}$$

which we rewrite as $P_1' t_{d'} t_{C'} t_C^{-1} = t_\delta$, for $\underline{P_1' = t_{x_4'} t_{e'} t_{x_1'} t_{x_2'} t_{x_3'}}$. Notably, $P_1' t_{d'}$, t_C, $t_{C'}$ and t_∂ all commute with each other.

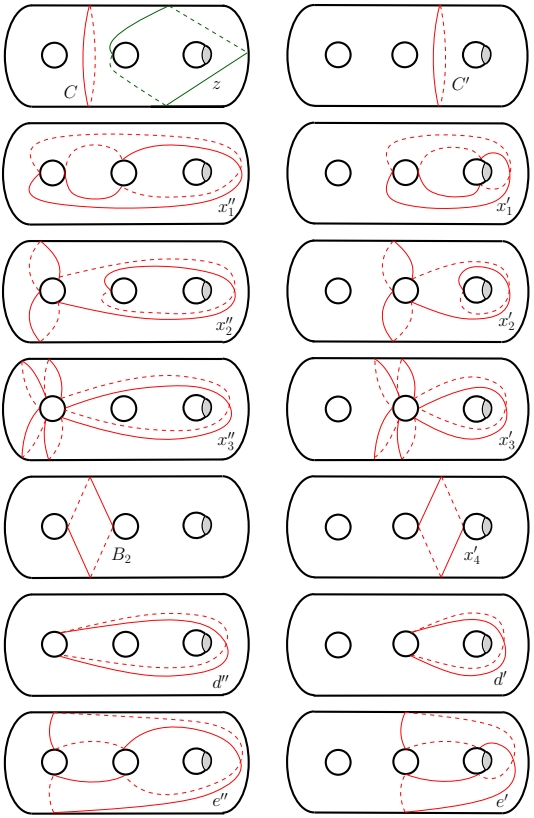


Figure 3. The surface Σ_3^1 with shifted boundary. The curves C', x_1', x_2', x_3', x_4', d', e' of the second embedding are on the right, and the curves C, x_1'', x_2'', x_3'', B_2, d'', e'' of the third embedding are given on the left. The Dehn twist curve z in the conjugation ϕ is on the top left (in green).

For our last embedding, cap off the boundary components δ_1 and δ_2 of Σ_2^3, and then embed the resulting copy of Σ_2^1 into Σ_3^1 so that δ_3 maps to the curve d' above, where the curves e'', x_1'', x_2'', x_3'', d'', B_2, C, d' are as shown in Figure 3. (Note that we get the same B_2, C curves.) So we obtain another relation in $\mathrm{Mod}(\Sigma_3^1)$:

$$t_{e''}t_{x_1''}t_{x_2''}t_{x_3''}t_{d''}t_{B_2}t_C t_{d'}^{-1} = 1, \tag{7}$$

which we rewrite as $P_1'' t_{d'}^{-1} = 1$, for $P_1'' = t_{e''}t_{x_1''}t_{x_2''}t_{x_3''}t_{d''}t_{B_2}t_C$. Here P_1'' and $t_{d'}$ commute.

We can now describe our positive factorizations using the three embeddings above. Let ϕ be any mapping class in the stabilizer group $\mathrm{Mod}(\Sigma_3^1, d')$. Then

$$P_1 P_1' (P_1'')^{\phi} = (P_1 t_C t_{C'}^{-1})(P_1' t_{d'} t_{C'} t_C^{-1})(P_1'' t_{d'}^{-1})^{\phi} = 1 \cdot t_\partial \cdot 1 = t_\partial.$$

Here the first equality follows from the commutativity relations we noted above, along with our choice of ϕ as follows: In the middle, the multitwist $t_{C'}t_C^{-1}$ commutes with P_1', so we can bring it to its left and cancel it against $t_C t_{C'}^{-1}$. Since ϕ stabilizes the curve d', we have $t_{d'} = (t_{d'})^{\phi}$, so we can now take the $t_{d'}$ factor into the conjugated expression, and then, because it commutes with P_1'', we can move it to its right and cancel against $t_{d'}^{-1}$ within the parentheses. The second equality is the product of the equalities (2), (6) and (7).

So, we have obtained a positive factorization $W_{0,\phi} = P_1 P_1'(P_1'')^{\phi}$ of the boundary twist t_δ in $\mathrm{Mod}(\Sigma_3^1)$. Each $W_{0,\phi}$ prescribes a symplectic genus-3 Lefschetz pencil $(X_{0,\phi}, f_{0,\phi})$ with one base point, equipped with a Thurston–Gompf symplectic form.

As before, we can calculate the Euler characteristic of $X_{0,\phi}$ as

$$\mathrm{e}(X_{0,\phi}) = 4 - 4g + \ell - b = 4 - 4\cdot 3 + 18 - 1 = 9,$$

and the signature of $X_{0,\phi}$ as the signature of the relation $W_{0,\phi}$ after [24]. The latter is equal to the sum of the signatures of the three relations (2), (6) and (7), which correspond to pencils on $T^2 \tilde{\times} S^2 \# 2\,\overline{\mathbb{CP}}^2$, $T^2 \tilde{\times} S^2 \# \overline{\mathbb{CP}}^2$ and $T^2 \tilde{\times} S^2 \# 2\overline{\mathbb{CP}}^2$, respectively. (In the case of the second embedding, since its second boundary twist t_{δ_2} was mapped to t_∂, it now corresponds to the base point of the genus-3 pencil.) So we get $\sigma(X_{0,\phi}) = -2 - 1 - 2 = -5$. Therefore, $\chi_h(X_{0,\phi}) = 1$ and $c_1^2(X_{0,\phi}) = 3$, as claimed. The only rational or ruled surface which has the same invariants is $\mathbb{CP}^2 \# 6\,\overline{\mathbb{CP}}^2$, which by Lemma 2, cannot admit such pencils.

Lastly, observe that the Dehn twist curves in all three factors P_1, P_1' and P_1'' involved in $W_{0,\phi}$ commute with the obvious involution on Σ_3^1 given by a π–rotation of the surface along the x–axis in Figure 3 (taking the z–axis perpendicular to the page). If we let $\mathrm{HMod}(\Sigma_3^1)$ denote the symmetric mapping class group with respect to this involution [25], then for any ϕ in the subgroup $\mathrm{Mod}(\Sigma_3^1, d') \cap \mathrm{HMod}(\Sigma_3^1)$, we get a positive factorization $W_{0,\phi}$ prescribing a hyperelliptic pencil $(X_{0,\phi}, f_{0,\phi})$.

4.3. ***Homeomorphism and homology types of $c_1^2 = 0, 1, 2$ examples.*** Recall that we have $\mathrm{e}(X_{i,\phi}) = 9 + i$ and $\sigma(X_{i,\phi}) = -5 - i$, for $i = 0, 1, 2, 3$. None of our examples have even intersection forms, which can be easily seen by the existence of reducible fibers in $X_{i,\phi}$, which have self-intersection -1. To be able to pin down the homeomorphism and integral homology types of these 4-manifolds, it remains to determine their fundamental groups and the first integral homology groups, which we will do for particular choices of ϕ.

Below, we will first carry out these calculations for our examples $(X_{i,\phi}, f_{i,\phi})$ with $i = 1, 2, 3$, and then do the same for the $i = 0$ case in the next subsection. Here we aspire to keep our calculations simple but also generate as many fundamental groups as possible. Finding the right balance will come at a cost of getting a somewhat asymmetric picture; the fundamental groups of $X_{i,\phi}$ will realize any quotient of $\mathbb{Z}^2$ when $i = 1, 3$, and any quotient of $\mathbb{Z}$ when $i = 0, 2$.

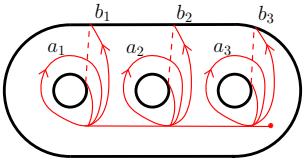


Figure 4. Generators a_j, b_j of $\pi_1(\Sigma_3)$

Let $\phi = t_{b_1}^{-m_1} t_{a_2}^{m_2}$, where b_1, a_2 are as in Figure 4. Since b_1 and a_2 are disjoint from C and C', we have $\phi \in \mathrm{Mod}(\Sigma_3^1, S)$ for $S = \{C, C'\}$, as required. Denote the positive factorizations in this case by $W_{i,m} := W_{i,\phi}$, where for $i = 1, 3$, we take $\phi = t_{b_1}^{-m_1} t_{a_2}^{m_2}$ and $m = (m_1, m_2) \in \mathbb{N}^2$, whereas for $i = 2$, we take $\phi = t_{b_1}^{-5} t_{a_2}^{m}$ and $m \in \mathbb{N}$. Now, set $(X_{i,m}, f_{i,m}) := (X_{i,\phi}, f_{i,\phi})$, and further set $(X_i, f_i) := (X_{i,m}, f_{i.m})$ in the specific cases of $m = (1, 1)$ when $i = 1, 3$, and $m = 1$ when $i = 2$.

We claim that $\pi_1(X_{i,m})$ is $(\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z})$, for $i = 1, 3$, and $\mathbb{Z}/m\mathbb{Z}$, for $i = 2$. In particular each X_i is simply connected.

Let $(\widetilde{X}_{i,m}, \widetilde{f}_{i,m})$ be the Lefschetz fibration we obtain by blowing-up the base points of the pencil $(X_{i,m}, f_{i,m})$. Let $\{a_j, b_j\}$ be the standard generators of $\pi_1(\Sigma_g)$ as shown in Figure 4. Using the standard handlebody decomposition for a Lefschetz fibration with a section, we obtain a finite presentation for $\pi_1(X_{i,m}) = \pi_1(\widetilde{X}_{i,m})$ of the form

$$\langle a_1, b_1, a_2, b_2, a_3, b_3 \mid [a_1, b_1][a_2, b_2][a_3, b_3],\, R_{i,m,1}, \ldots, R_{i,m,18+i} \rangle, \tag{8}$$

where $\{R_{i,m,k}\}_{k=1}^{18+i}$ are relators obtained by expressing the Dehn twist curves in the positive factorization $W_{i,m}$ in the basis $\{a_j, b_j\}_{j=1}^{3}$. We denote the inverse of any fundamental group element x by $\bar{x}$.

We will first show that a subset of these relators, which come from Dehn twist curves that are all present in each factorization $W_{i,m}$, for $i = 1, 2, 3$, already yield an abelian quotient. Since any further quotient will also be abelian, at that point it will suffice to consider only the abelianizations of all the relators $\{R_{i,m,k}\}_{k=1}^{18+i}$.

Each positive factorization $W_{i,m}$ contains the factor $P_2' t_C$. So the following relations hold for the finite presentations we have for each $\pi_1(X_{i,m})$:

$$[a_1, b_1][a_2, b_2][a_3, b_3] = 1, \tag{9}$$

$$[a_1, b_1] = 1, \tag{10}$$

$$a_2 a_3 = 1, \tag{11}$$

$$a_2 \bar{b}_2 a_3 \bar{b}_3 = 1, \tag{12}$$

$$b_3 b_2 = 1, \tag{13}$$

where the relators (10)–(13) come from the vanishing cycles C, B_0', B_1', B_2', respectively. We have $\underline{a_3 = \bar{a}_2}$ from (11) and $\underline{b_3 = \bar{b}_2}$ from (13). Together with (12), these imply $\underline{[a_2, b_2] = 1}$ and $[a_3, b_3] = 1$. We conclude $[a_j, b_j] = 1$ for every $j = 1, 2, 3$.

From the factor P_1, we get (among many others) the relators

$$a_1(\bar{b}_1 a_2 b_2)^2 = 1, \tag{14}$$

$$a_1 \bar{b}_1^3 a_2 b_2 a_2 = 1, \tag{15}$$

$$a_1 \bar{b}_1^5 a_2 [b_2, a_2] b_1 a_2 = 1, \tag{16}$$

$$b_2 b_1 [b_3, a_3] = 1, \tag{17}$$

induced by the vanishing cycles x_1, x_2, x_3 and B_2, respectively. Adding these to the previous relators from the factor P_2', we immediately see that the commutativity of a_3 and b_3 and (17) imply $\underline{b_1 = \bar{b}_2}$. So (14) implies that $a_1 = (\bar{b}_2 \bar{a}_2 \bar{b}_2)^2$, and since a_2 and b_2 commute, we get $\underline{a_1 = \bar{a}_2^2 \bar{b}_2^4}$

On the other hand, if we have the factor P_2 instead, we get (again, among many others) the relators

$$a_1 a_2 = 1, \tag{18}$$

$$b_2 \bar{a}_2 b_1 \bar{a}_1 [b_3, a_3] = 1, \tag{19}$$

$$b_2 b_1 [b_3, a_3] = 1,$$

induced by the vanishing cycles B_0, B_1 and B_2, respectively. We get $\underline{a_1 = \bar{a}_2}$, and together with the relators from P_2' we once again get $\underline{b_1 = \bar{b}_2}$, since $[a_3, b_3] = 1$.

Now, since the positive factorization $W_{1,m}$ contains the factor $P_1 P_2' t_C$ and the positive factorizations $W_{2,m}$ and $W_{3,m}$ both contain the factor $P_2 P_2' t_C$, the above discussion shows that every $\pi_1(X_{i,m})$ is a quotient of an abelian group generated by a_2 and b_2. It therefore remains to look at the abelianizations of the relators coming from the remaining Dehn twist curves, i.e., we can simply look at the homology classes of the vanishing cycles.

Without the conjugated factor, we have the abelianized relations

$$a_3 = -a_2 \text{ and } b_3 = -b_2 = b_1 \quad \text{for all } W_{i,m} \tag{20}$$

and depending on whether W_i contains the factor P_1 or P_2, either

$$a_1 + 2a_2 + 4b_2 = 0 \quad \text{for } W_{1,m}, \text{ or} \tag{21}$$

$$a_1 + a_2 = 0 \quad \text{for } W_{2,m} \text{ and } W_{3,m}, \tag{22}$$

where we used (20) to simplify the relators. These relators amount to all the other generators being obtained from a_2 and b_2.

In fact, there are no other relations coming from the nonconjugated factors P_1, P_2 or P_2': This is easy to see by abelianizing the relators (9)–(19), which include all the relators induced by the curves x_1, x_2, x_3, B_0, B_1, B_2, B_0', B_1', B_2'. Missing are the relators induced by the separating curves d, e from P_1, the curves A_0, A_1, A_2 from P_1, and the curves A_0', A_1', A_2' from P_2'. The first two are trivial in homology,

so they have no contribution to the list of relators we already have. On the other hand, for each $j = 0, 1, 2$, A_j is homologous to B_j, because $[A_j] = [t_C(B_j)] = [B_j] + (C \cdot B_j)[C]$, where C is a separating cycle. Similarly each A'_j is homologous to B'_j. Therefore the abelianized relations they induce are identical to those we already had from B_j, B'_j.[2]

It remains to look at the abelianizations of the relators coming from the conjugated factors P_1^ϕ or P_2^ϕ. When $i = 1, 3$, for $\phi = t_{b_1}^{-m_1} t_{a_2}^{m_2}$, we easily check using the Picard–Lefschetz formula that we get the additional relators

$$b_1 + m_2 a_2 + b_2 = 0 \quad \text{for } W_{1,m} \text{ and } W_{3,m}, \tag{23}$$

$$a_1 + m_1 b_1 + (2 + 4m_2)a_2 + 4b_2 = 0 \quad \text{for } W_{1,m}, \tag{24}$$

$$a_1 + m_1 b_1 + a_2 = 0 \quad \text{for } W_{3,m}. \tag{25}$$

The relations (20) and (23) imply that $\underline{m_2 a_2 = 0}$. The remaining relators involved in $W_{1,m}$ or $W_{3,m}$ then easily give $\underline{m_1 b_2 = 0}$. Hence, for $i = 1, 3$, we have

$$\pi_1(X_{i,m}) = (\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z}),$$

as claimed.

On the other hand, when $i = 2$, for $\phi = t_{b_1}^{-5} t_{a_2}^{m}$, we get the following additional relators in $W_{2,m}$:

$$b_1 + m a_2 + b_2 = 0, \tag{26}$$

$$a_1 + 5b_1 + (2 + 4m)a_2 + 4b_2 = 0. \tag{27}$$

This time, the relations (20) and (26) imply that $\underline{m a_2 = 0}$, but then if we use this identity and substitute $a_1 = -a_2$ and $b_1 = -b_2$ into the relator (27), we get $b_2 = a_2$. Therefore, the m–torsion element a_2 generates the whole group. So we have

$$\pi_1(X_{2,m}) = \mathbb{Z}/m\mathbb{Z}.$$

In particular, when $i = 1, 3$, we get a trivial group for $(m_1, m_2) = (\pm 1, \pm 1)$, and when $i = 2$, we get a trivial group for $m = 1$. So X_i is simply connected for each $i = 1, 2, 3$. By Freedman's celebrated work [30], each X_i is homeomorphic to $\mathbb{CP}^2 \# (6+i)\overline{\mathbb{CP}}^2$, for $i = 1, 2, 3$.[3] However, they are not diffeomorphic by Theorem 3.

[2] For the proof of the simply connected case, one could skip this whole paragraph, since we would only need to find enough relations to kill the fundamental group.

[3] Homeomorphism types of other $X_{i,m}$ can also be determined using extensions of Freedman's work by Hambleton, Kreck and Teichner for respective fundamental groups; for example by [40] we can see that when $i = 1, 3$, for $m = (p, 1)$, and when $i = 2$, for $m = p$, each $X_{i,m}$ is homeomorphic to $\mathbb{CP}^2 \# (6+i)\overline{\mathbb{CP}}^2 \# L_p$, where L_p is the spun of the Lens space $L(p, 1)$.

4.4. ***Homeomorphism and homology types of $c_1^2 = 3$ examples.*** Now we take $\phi = t_{b_1}^{-m-10} t_z$, where b_1 and z are as in Figures 4 and 3. Since

$$\phi \in \mathrm{Mod}(\Sigma_3^1, d') \cap \mathrm{HMod}(\Sigma_3^1),$$

for each such ϕ, the positive factorization $W_{0,\phi}$ prescribes a hyperelliptic genus-3 pencil $(X_{0,\phi}, f_{0,\phi})$. To sync up our notation with the $c_1^2 = 0, 1, 2$ examples, set $W_{0,m} := W_{0,\phi}$ and $(X_{0,m}, f_{0,m}) := (X_{0\phi}, f_{0,\phi})$, while noting that the parameter m takes values in $\mathbb{N}$ (rather than in $\mathbb{N}^2$). Finally, let $(X_0, f_0) := (X_{0,1}, f_{0,1})$. We claim that $\pi_1(X_{0,m}) = \mathbb{Z}/m\mathbb{Z}$, and in particular X_0 is simply connected.

As before, we calculate the fundamental group using the presentation of the form (8) induced by the pencil structure. We will first write down only some of the relators we get from the Dehn twist curves in the positive factorization $W_{0,m}$ and observe that any $\pi_1(X_{0,m})$ will be a quotient of an abelian group. It will then suffice to look at the abelianized relators induced by the remaining Dehn twist curves, and run the calculation at the level of homology.

The following relations hold in $\pi_1(X_{0,m})$:

$$[a_1, b_1][a_2, b_2][a_3, b_3] = 1, \tag{28}$$

$$a_1(\bar{b}_1 a_2 b_2)^2 = 1, \tag{29}$$

$$a_1 \bar{b}_1^3 a_2 b_2 a_2 = 1, \tag{30}$$

$$[b_1, a_2 b_2 a_1] = 1, \tag{31}$$

$$[a_2, \bar{b}_1 a_2 b_2] = 1, \tag{32}$$

$$a_2(\bar{b}_2 a_3 b_3)^2 = 1, \tag{33}$$

$$a_3 \bar{b}_3 \bar{a}_3 \bar{b}_2 [a_3, b_3] = 1, \tag{34}$$

$$[a_1, b_1] = 1, \tag{35}$$

where the first one is the surface relation, and (29)–(32) are induced by x_1, x_2, e, d coming from the P_1 factor, (33)–(34) by x_1', x_4' from P_1' and (35) by C from $(P_1'')^\phi$.

We can rewrite (30) as $a_1 \bar{b}_1^2 a_2 \bar{b}_1 a_2 b_2 = 1$ using the commutativity relation (32). Setting this relator equal to the relator (29), we get: $a_1 \bar{b}_1^2 a_2 \bar{b}_1 a_2 b_2 = a_1 \bar{b}_1 a_2 b_2 \bar{b}_1 a_2 b_2$, which, through cancellations, give $\bar{b}_1 a_2 = a_2 b_2$. Using this last identity, we can rewrite (31) as $[b_1, \bar{b}_1 a_2 a_1] = 1$. This implies that $[b_1, a_2 a_1] = 1$. However, by (35), b_1 commutes with a_1, so we can further conclude that $[b_1, a_2] = 1$. Since a_2 commutes with b_1, we derive from (32) that $[a_2, b_2] = 1$. In turn, using (28) and (35) we conclude that $[a_3, b_3] = 1$ as well.

We are now ready to show that a_3 and b_1 generate the whole group. Since we saw that $\bar{b}_1 a_2 = a_2 b_2$, the commutativity of a_2 and b_2 implies that $\underline{b_2 = \bar{b}_1}$. Since a_3 and b_3 commutes, (34) gives $b_3 = \bar{b}_2$, which in turn means $\underline{b_3 = b_1}$. Note that this last identity and the commutativity of a_3 and b_3 now show that $\underline{[a_3, b_1] = 1}$. Now by (33),

we have $a_2 = (\bar{b}_3\bar{a}_3 b_2)^2$, which implies that $a_2 = (\bar{b}_1\bar{a}_3\bar{b}_1)^2$. After commuting the factors, we can rewrite the last identity as $\underline{a_2 = \bar{a}_3^2\bar{b}_1^4}$. Similarly, by (29), we have $a_1 = (\bar{b}_2\bar{a}_2 b_1)^2$, which, after substitutions becomes $a_1 = (b_1(b_1^4 a_3^2)b_1)^2$, so $\underline{a_1 = b_1^{12}a_3^4}$.

Underlined equalities we obtained above show that a_3 and b_1 generate the whole group and commute with each other. Hence, $\pi_1(X_{0,m})$ is the quotient of an abelian group with two generators. To finish our calculation of $\pi_1(X_{0,m})$, it now suffices to write out the abelianizations of the relators induced by all the Dehn twist curves in the positive factorization $W_{0,m}$. Clearly the separating Dehn twists do not contribute any nontrivial abelianized relators, whereas each quadruple of nonseparating Dehn twists coming from the factors P_1, P_1' and $(P_2'')^\phi$, respectively, can be seen to give only two linearly independent abelianized relators. For instance, the curves x_1, x_2, x_3, B_2 in the P_1 factor yield the relators

$$a_1 - 2b_1 + 2a_2 + 2b_2 = 0, \tag{36}$$
$$a_1 - 3b_1 + 2a_2 + b_2 = 0, \tag{37}$$
$$a_1 - 4b_1 + 2a_2 = 0, \tag{38}$$
$$b_1 + b_2 = 0, \tag{39}$$

where (38) and (39) generate them all. Similarly, the abelianized relators we get from x_1', x_2', x_3', x_4' in the P_1' factor are generated by

$$a_2 - 4b_2 + 2a_3 = 0, \tag{40}$$
$$b_2 + b_3 = 0, \tag{41}$$

and those we get from x_1'', x_2'', x_3'', B_2 in the nonconjugated P_1'' are generated by

$$a_1 - 4b_1 + 2a_2 + 2a_3 = 0,$$
$$b_1 + b_2 = 0.$$

By the Picard-Lefschetz formula, conjugating P_1'' with $\phi = t_{b_1}^{-m-10}t_z$ yields the following additional relators:

$$a_1 + (m+6)b_1 - 2b_2 = 0, \tag{42}$$
$$b_1 + a_2 + 2b_2 + a_3 = 0. \tag{43}$$

Note that, taking an auxiliary orientation on the twisting curve z, here we have $[z] = a_2 + b_2 + a_3$ in homology.

We can replace the two relations (39) and (41) with $\underline{b_2 = -b_1}$ and $\underline{b_3 = b_1}$. Then (40) can be changed to $\underline{a_2 = -4b_1 - 2a_3}$, and in turn (38) can be changed to $\underline{a_1 = 12b_1 + 4a_3}$. As we express all the other generators in terms of a_3 and b_1, (43) becomes $\underline{a_3 = -5b_1}$. Finally, expressing all the generators in terms of b_1, the

remaining relation (42) now reads as $\underline{mb_1 = 0}$. We conclude that

$$\pi_1(X_{0,m}) = \mathbb{Z}/m\mathbb{Z}$$

as claimed. When $m = \pm 1$, we get a trivial group, so in particular, $X_0 = X_{0,1}$ is simply connected. Since we have $e(X_0) = 9$ and $\sigma(X_0) = -5$, by Freedman [30], X_0 is homeomorphic to $\mathbb{CP}^2 \# 6\,\overline{\mathbb{CP}}^2$, but not diffeomorphic to it by Theorem 3.

4.5. ***The theorem and ancillary remarks.*** Combining the results of the previous four subsections, we have:

Theorem 7. *$\{(X_{i,\phi},\, f_{i,\phi})\}$ are symplectic genus-3 Lefschetz pencils whose total spaces have $\chi_h(X_{i,\phi}) = 1$ and $c_1^2(X_{i,\phi}) = 3 - i$, and they include exotic rational surfaces $\mathbb{CP}^2 \# (6+i)\,\overline{\mathbb{CP}}^2$ as well as infinitely many symplectic 4-manifolds which are not homotopy equivalent to any complex surface, for each $i = 0, 1, 2, 3$.*

The additional claim regarding the examples which are not homotopy equivalent to any complex surface follows from standard arguments: The family of symplectic 4-manifolds $\{X_{i,\phi}\}$ contains $\{X_{i,m}\}$ we studied in detail, and the fundamental groups of the latter family realize any $(\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z})$ for $i = 1, 3$, and any $\mathbb{Z}/m\mathbb{Z}$, for $i = 0, 2$. For $i = 1, 3$, we get infinitely many examples with $b_1(X_{i,m}) = 1$ (and $b^+(X_{i,m}) > 0$) by setting $m_1 = 0$ and varying m_2. These cannot be homotopy equivalent to any complex surface; see, e.g., [8, Lemma 2]. For $i = 0, 2$, we have $\kappa(X_{i,m}) = 2$. However, there are only finitely many deformation classes of compact complex surfaces of general type with the same χ_h and c_1^2 invariants [34], so all but finitely many of these $X_{i,m}$ cannot have the homotopy type of a complex surface.

Remark 8. We claim that the $X_{i,\phi}$ are either minimal or at most one blow-up of a minimal symplectic 4-manifold. This follows from the following more general observation (cf. Remark 4): For any pencil $(X,\, f)$, where X is not rational or ruled, the collection of all exceptional classes in the corresponding Lefschetz fibration $(\widetilde{X},\, \widetilde{f})$ can be represented by disjoint multisections S_j, each one of which intersects the regular fiber F positively. By [57, Theorem 5-12], $\kappa(X) = 2$ and $g \geq 3$ implies that $\left(\sum S_j\right) \cdot F \leq 2g-4$, which in turn means X can have at most $2g-4$ exceptional classes. Note that if $X_{3,\phi}$ is not minimal, then $\kappa(X_{3,\phi}) = 2$, like the other $X_{i,\phi}$, for $i = 0, 1, 2$. Now for every $i = 0, 1, 2, 3$, since the genus-3 pencil $(X_{i,\phi},\, f_{i,\phi})$ already has one base point (yielding an exceptional class in the corresponding fibration), there can be at most one more exceptional class, proving our claim.

Remark 9. We suspect that the smallest exotic rational surface one can get via genus-3 pencils has $c_1^2 = 3$ or 4 and our example $(X_0,\, f_0)$ might very well be optimal. Our subsequent work in [11] shows that one can already get sharper results with pencils of genus $g = 4$ or 5, which is in part due to having room for more base points, since the number of base points b is less than or equal to $2g - 4$ by the previous

remarks. It is also worth noting that no exotic rational surface admits a pencil of genus $g \leq 2$. The total space of any pencil of genus $g \leq 1$ is a rational surface, and that of any genus-2 pencil has $\kappa \geq 1$ by Lemma 2 and Proposition 1. Moreover, by [57, Theorem 5-5(iii)], a genus-2 pencil with $\kappa = 1$ should have only one reducible fiber, which is not possible when the total space has Euler characteristic smaller than 14 by [13, Lemmas 4 and 5]. Hence, the smallest fiber genera for pencils on minimal exotic rational surfaces is $g = 3$. In contrast, there exist genus-2 Lefschetz fibrations on minimal exotic rational surfaces with $c_1^2 = 0, 1, 2$, and in fact for no other c_1^2 [13].

Remark 10. By the work of Siebert and Tian [58], the hyperellipticity of our genus-3 pencil (X_0, f_0) implies that the exotic rational surface X_0 with $c_1^2 = 3$ is the blow-down of a symplectic double cover of a rational ruled surface. The exotic rational surfaces we built in [13] with $c_1^2 = 0, 1, 2$ via hyperelliptic genus-2 Lefschetz fibrations have the same property. In particular, all these exotic rational surfaces admit symplectic involutions.

Remark 11. There are many prior constructions of Kähler surfaces and symplectic 4-manifolds in the homeomorphism classes of the rational surfaces in Theorem 7. The first examples with $c_1^2 = 0$ and 1 were the Dolgachev surfaces and the Barlow surface, as shown by Donaldson [19] and Kotschick [46], respectively, in the late 1980s. The first examples with $c_1^2 = 2$ and 3 were obtained around 2005 via generalized rational blowdowns by J. Park [55] and Stipsicz and Szabó [61], respectively. Infinitely many distinct smooth structures in these homeomorphism classes were constructed using logarithmic transforms, knot surgeries and Luttinger surgeries; see, e.g., [1; 2; 3; 27; 29; 32; 62] (all of which are indeed instances of surgeries along tori [14].) However, it remains an open question whether there are two distinct minimal symplectic 4-manifolds homeomorphic but not diffeomorphic to the same rational surface with $c_1^2 < 9$; see [60, Problem 11]. As observed by Stipsicz and Szabó, Seiberg–Witten invariants cannot distinguish these symplectic 4-manifolds [61, Corollary 4.4]. It is thus desirable to have examples with more structure like ours, in hope of addressing this intriguing question.

Remark 12. It follows from the works of Donaldson [20] and Gompf [35] that every finitely presented group is the fundamental group of a symplectic Lefschetz pencil; also see [5; 39; 45] for direct constructions. One can thus define an invariant m_g of finitely presented groups, where for any such G, $m_g(G)$ is the smallest g among all the genus-g pencils with $\pi_1 = G$. Well-known examples of pencils of genus $g = 0, 1, 2$ show that for the groups $G = 1, \mathbb{Z}_2$ and $\mathbb{Z}^2$, we have $m_g = 0, 2, 2$, realized by pencils on $\mathbb{CP}^2$, the Enriques surface, and $T^2 \times S^2$, respectively. We conjecture that $m_g(G) = 3$ for all the other $G \cong (\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z})$, which are realized by our genus-3 pencils $(X_{i,m}, f_{i,m})$, when $i = 0, 1, 2$.

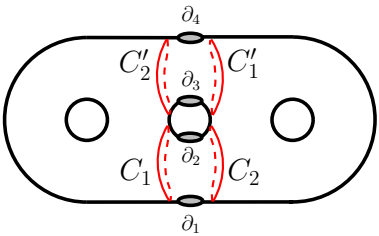


Figure 5. The curves involved in our embeddings of $\partial\Sigma_4^2$ into $\partial\Sigma_3^4$.

5. Symplectic Calabi–Yau surfaces with $b_1 > 0$ via genus-3 pencils

In this section we will give a new construction of an infinite family of symplectic Calabi–Yau surfaces with $b_1 > 0$ in all possible rational homology classes allowed by the rational homology classification of symplectic Calabi–Yaus [7; 49]. These examples will come from our construction of new positive factorizations of boundary multitwists in $\mathrm{Mod}(\Sigma_3^4)$ corresponding to symplectic genus-3 Lefschetz pencils.

5.1. ***Breeding symplectic Calabi–Yau pencils.*** The positive factorization for Matsumoto's genus-2 Lefschetz fibration has the following further lift to $\mathrm{Mod}(\Sigma_2^4)$, which was obtained by Hamada in [37]:

$$t_{B_{0,1}} t_{B_{1,1}} t_{B_{2,1}} t_{C_1} t_{B_{0,2}} t_{B_{1,2}} t_{B_{2,2}} t_{C_2} = t_{\delta_1} t_{\delta_2} t_{\delta_3} t_{\delta_4}, \tag{44}$$

where δ_i are boundary parallel curves, and $B_{j,i}$ and C_i are as shown on the right-hand side of Figure 6. This relation will be the main building block in our construction.

After Hurwitz moves, we can rewrite the relation (44) as

$$t_{B_{0,1}} t_{B_{1,1}} t_{B_{2,1}} t_{A_{0,2}} t_{A_{1,2}} t_{A_{2,2}} t_{C_1} t_{C_2} t_{\delta_1}^{-1} t_{\delta_2}^{-1} = t_{\delta_3} t_{\delta_4},$$

where each $A_{j,2} = t_{C_1}(B_{j,2})$ for $j = 0, 1, 2$. Note that if we cap off the two boundary components δ_3 and δ_4, the curves $B_{j,i}$ descend to the curves B_j and C_i to C given on the left-hand side of Figure 6, for each $j = 0, 1, 2$ and $i = 1, 2$.

Consider the embedding of Σ_2^4 into Σ_3^4 obtained by attaching a Σ_0^4 along two of the boundary components of Σ_2^4. We choose an embedding of Σ_2^4 such that we map

$$\delta_1 \mapsto C_2', \quad \delta_2 \mapsto C_1', \quad \delta_3 \mapsto \partial_2, \quad \delta_4 \mapsto \partial_1,$$

where C_i, C_i' are as shown in Figure 5. We then map the interior of Σ_2^4 so that the curves $B_{j,1}$ and $A_{j,2}$ all map to the curves B_j and A_j in Figure 7 when the boundary components $\partial_1, \ldots, \partial_4$ are capped off.[4] Thus, the following relation holds

[4]At the end of our construction, the four boundary components of Σ_3^4 will correspond to disk neighborhoods of the four base points of our genus-3 pencils, so knowing the isotopy classes of these Dehn twist curves after we cap off all ∂_i will be enough for our π_1 and H_1 calculations.

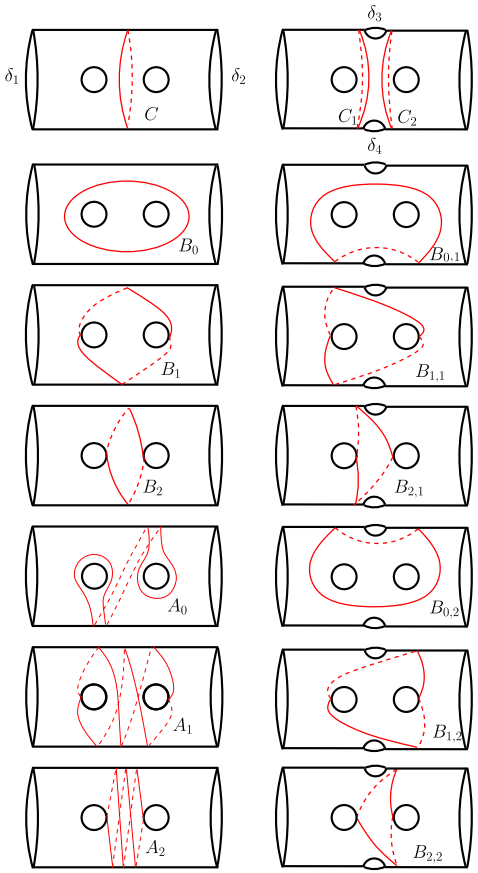


Figure 6. The curves B_j, C, $B_{j,i}$, C_i in Hamada's lifts. On the left are the curves of the positive factorization in $\mathrm{Mod}(\Sigma_2^2)$, along with the curves A_j we got after the Hurwitz moves. On the right are the curves of the further lift in $\mathrm{Mod}(\Sigma_2^4)$.

in $\mathrm{Mod}(\Sigma_3^4)$:

$$t_{B_{0,1}} t_{B_{1,1}} t_{B_{2,1}} t_{A_{0,2}} t_{A_{1,2}} t_{A_{2,2}} t_{C_1} t_{C_2} t_{C_2'}^{-1} t_{C_1'}^{-1} = t_{\partial_1} t_{\partial_2}, \tag{45}$$

which we can rewrite as

$$P t_{C_1} t_{C_2} t_{C_2'}^{-1} t_{C_1'}^{-1} = t_{\delta_1} t_{\delta_2}, \quad \text{for } \underline{P = t_{B_{0,1}} t_{B_{1,1}} t_{B_{2,1}} t_{A_{0,2}} t_{A_{1,2}} t_{A_{2,2}}}.$$

A similar embedding of Σ_2^4 into Σ_3^4 can be given by mapping

$$\delta_1 \mapsto C_2, \quad \delta_2 \mapsto C_1, \quad \delta_3 \mapsto \partial_3, \quad \delta_4 \mapsto \partial_4,$$

where the interior is mapped in a similar fashion to before, so we get the curves B_j' and A_j' in Figure 7 when the boundary components $\partial_1, \ldots, \partial_4$ are capped off. Note

that this second embedding can be obtained from the first one by a rotation of the surface Σ_3^4 in Figure 5. So we get another relation in $\mathrm{Mod}(\Sigma_3^4)$:

$$t_{B'_{0,1}} t_{B'_{1,1}} t_{B'_{2,1}} t_{A'_{0,2}} t_{A'_{1,2}} t_{A'_{2,2}} t_{C'_1} t_{C'_2} t_{C_2}^{-1} t_{C_1}^{-1} = t_{\partial_3} t_{\partial_4}, \tag{46}$$

which we can rewrite as

$$P' t_{C'_1} t_{C'_2} t_{C_2}^{-1} t_{C_1}^{-1} = t_{\delta_3} t_{\delta_4}, \quad \text{for } \underline{P' = t_{B'_{0,1}} t_{B'_{1,1}} t_{B'_{2,1}} t_{A'_{0,2}} t_{A'_{1,2}} t_{A'_{2,2}}}.$$

Let ϕ be any mapping class in $\mathrm{Mod}(\Sigma_3^4)$ which fixes the set $S := \{C_1, C_2, C'_1, C'_2\}$ pointwise, i.e., $\phi \in \mathrm{Mod}(\Sigma_3^4, S)$. Then the product of P^ϕ and P' yield

$$P^\phi P' = P^\phi t_{C_1} t_{C_2} t_{C'_2}^{-1} t_{C'_1}^{-1}\ P' t_{C'_1} t_{C'_2} t_{C_2}^{-1} t_{C_1}^{-1} = (P t_{C_1} t_{C_2} t_{C'_2}^{-1} t_{C'_1}^{-1})^\phi\ P' t_{C'_1} t_{C'_2} t_{C_2}^{-1} t_{C_1}^{-1} = \Delta,$$

where $\Delta = t_{\partial_1} t_{\partial_2} t_{\partial_3} t_{\partial_4}$ is the boundary multitwist. Here, in the first equality we used the commutativity of disjoint Dehn twists t_{C_1}, t_{C_2}, $t_{C'_1}$, $t_{C'_2}$ and that they all commute with P and P'. The second equality holds since ϕ commutes with the Dehn twists along C_1, C_2, C'_1 and C'_2.

Therefore $\underline{W_\phi = P^\phi P'}$ is a positive factorization of the boundary multitwist $\Delta = t_{\partial_1} t_{\partial_2} t_{\partial_3} t_{\partial_4}$ in $\mathrm{Mod}(\Sigma_3^4)$ for any ϕ as above. Under the boundary capping homomorphism $\mathrm{Mod}(\Sigma_3^4) \to \mathrm{Mod}(\Sigma_3)$ this maps to a positive factorization

$$(t_{B_0} t_{B_1} t_{B_2} t_{A_0} t_{A_1} t_{A_2})^\psi t_{B'_0} t_{B'_1} t_{B'_2} t_{A'_0} t_{A'_1} t_{A'_2} = 1, \tag{47}$$

where ψ is the image of the mapping class ϕ under this homomorphism.

Let (X_ϕ, f_ϕ) denote the symplectic genus-3 Lefschetz pencil corresponding to the positive factorization W_ϕ. We claim that each X_ϕ is a symplectic Calabi–Yau surface.

The Euler characteristic of X_ϕ is easily calculated as

$$\mathrm{e}(X_\phi) = 4 - 4g + \ell - b = 4 - 4 \cdot 3 + 12 - 4 = 0,$$

where g and b are the genus and the number of base points of the pencil, and ℓ is the number of critical points, which is equal to the number of Dehn twists in the positive factorization W_ϕ.

As we have an explicit positive factorization (47) for the pencil (X_ϕ, f_ϕ), the signature of X_ϕ can be once again easily calculated using the work of Endo and Nagami in [24]. The signature of the relation (44) we used as our main building block, which corresponds to a pencil on a minimal ruled surface, is zero, and so is the signature of any embedding of this relation into a higher genus surface. Since Hurwitz moves, conjugations and cancellations of positive–negative Dehn twist pairs have no effect on the signature, the signature of the final relation (47) is also zero. Therefore $\sigma(X_\phi) = 0$.

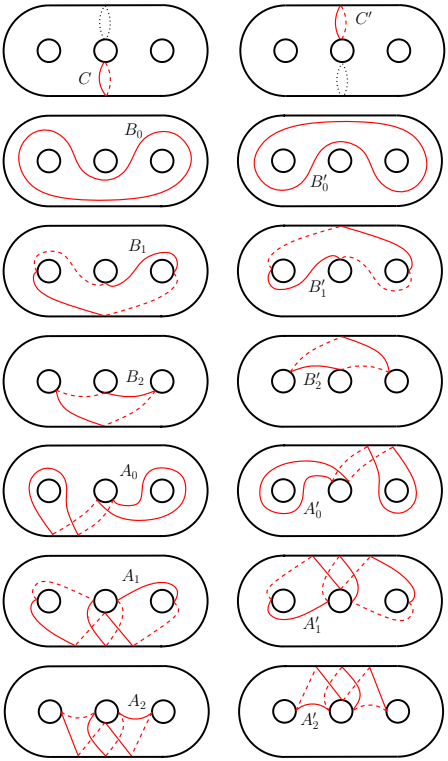


Figure 7. The curves B_j, A_j, B'_j, A'_j of the genus-3 pencil (X, f). On the left are the curves coming from the factorization P and on the right are those coming from P', which correspond to the two different embeddings of the factorization in $\mathrm{Mod}(\Sigma_2^2)$ into $\mathrm{Mod}(\Sigma_3)$. (Dotted lines are the identified images of δ_1 and δ_2 under these two embeddings.)

The only rational or ruled surfaces that have the same Euler characteristic and signature as X_ϕ are $T^2 \times S^2$ and $T^2 \tilde{\times} S^2$. However, by Lemma 5 they do not admit pencils with $b = 2g-2$ base points. Hence, we can apply Proposition 1 to conclude that $\kappa(X_\phi) = 0$. Since X_ϕ clearly does not have the same rational homology as the K3 surface or the Enriques surface, we can already tell that it is a symplectic Calabi–Yau surface with $b_1 > 0$.

5.2. ***Homeomorphism and homology types.*** We will first calculate the fundamental group of X_ϕ in the extremal case: when ϕ is the identity and $b_1(X_\phi) = 4$. We will show that the 4-manifold we simply denote by X in this case has $\pi_1(X) = H_1(X) \cong \mathbb{Z}^4$, and we will in fact conclude that X is homeomorphic to the 4-torus. After this detailed calculation, we will calculate $H_1(X_\phi)$ for a certain family of $\phi \in \mathrm{Mod}(\Sigma_3^4, S)$, where $S = \{C_1, C_2, C'_1, C'_2\}$, to cover all rational homology types of symplectic

Calabi–Yau surfaces with $b_1 > 0$. For any choice of ϕ, one can easily derive a presentation for $\pi_1(X_\phi)$ from that of $\pi_1(X)$, which we will leave to the reader.

Let $(\widetilde{X}, \widetilde{f})$ be the Lefschetz fibration we obtain by blowing-up the base points of the pencil (X, f). Let $\{a_j, b_j\}$ be the standard generators of $\pi_1(\Sigma_g)$ as shown in Figure 4. Once again, we have a finite presentation for $\pi_1(\widetilde{X})$ of the form

$$\langle a_1, b_1, a_2, b_2, a_3, b_3 \mid [a_1, b_1][a_2, b_2][a_3, b_3], R_1, \dots, R_{12}\rangle,$$

where each $\{R_k\}_{k=1}^{12}$ is a relation obtained by expressing the Dehn twist curves in the positive factorization (47) in the basis $\{a_j, b_j\}_{j=1}^{3}$.

So we have the following relations induced by B_0, B_1, B_2, A_0, A_1, A_2, B_0', B_1', B_2', A_0', A_1', A_2' (see Figure 7), in the same order:

$$a_1 a_3 = 1, \tag{48}$$

$$a_1 \bar{b}_1 a_2 b_2 \bar{a}_2 a_3 \bar{b}_3 = 1, \tag{49}$$

$$\bar{b}_1 a_2 b_2 \bar{a}_2 \bar{b}_3 = 1, \tag{50}$$

$$a_1 [b_3, a_3] b_2 a_3 \bar{b}_2 [a_3, b_3] = 1, \tag{51}$$

$$a_3 \bar{b}_3 \bar{b}_2 [a_3, b_3] a_1^2 \bar{b}_1 \bar{a}_1 [b_3, a_3] b_2 [b_3, a_3] b_2 = 1, \tag{52}$$

$$a_1 \bar{b}_1 \bar{a}_1 [b_3, a_3] b_2 [b_3, a_3] b_2 \bar{b}_3 \bar{b}_2 [a_3, b_3] = 1, \tag{53}$$

$$\bar{a}_2 a_1 a_2 a_3 = 1, \tag{54}$$

$$a_1 \bar{b}_1 a_2 a_3^2 \bar{b}_3 \bar{a}_3 b_2 \bar{a}_2 = 1, \tag{55}$$

$$b_1 a_2 \bar{b}_2 a_3 b_3 \bar{a}_3 \bar{a}_2 = 1, \tag{56}$$

$$a_1 a_2 \bar{b}_2 a_3 b_2 \bar{a}_2 = 1, \tag{57}$$

$$a_1 \bar{b}_1 a_2 \bar{b}_2 a_3^2 \bar{b}_3 \bar{a}_3 b_2^2 \bar{a}_2 = 1, \tag{58}$$

$$\bar{b}_1 a_2 \bar{b}_2 a_3 \bar{b}_3 \bar{a}_3 b_2^2 \bar{a}_2 = 1. \tag{59}$$

First observe that, when abelianized, the relations coming from each triple $\{B_0, B_1, B_2\}$, $\{A_0, A_1, A_2\}$, $\{B_0', B_1', B_2'\}$, $\{A_0', A_1', A_2'\}$ yield the same three relations

$$\begin{aligned} a_1 + a_3 &= 0, \\ a_1 - b_1 + b_2 + a_3 - b_3 &= 0, \\ b_1 - b_2 + b_3 &= 0, \end{aligned}$$

where we identified the abelianized images of the π_1 generators with the same letters. Any two of these relations imply the third. Since $a_1 = -a_3$ and $b_1 = b_2 - b_3$, we can eliminate a_1 and b_1 (and these relations) from the presentation, and we get a free abelian group of rank 4, generated by a_2, b_2, a_3 and b_3.

Now, going back to the presentation we had for $\pi_1(\widetilde{X})$, we see that it is also generated by a_2, b_2, a_3, b_3, for $a_1 = \bar{a}_3$ by (48) and $b_1 = a_2 b_2 \bar{a}_2 \bar{b}_3$ by (50). Therefore, to

conclude that $\pi_1(\widetilde{X}) = \mathbb{Z}^4$, it suffices to show that a_2, b_2, a_3 and b_3 all commute with each other, which is what do next: Replacing a_1 with $\bar{a}_3$ in (54) gives $\underline{[a_2, a_3] = 1}$. From (50) we have $\bar{b}_1 a_2 b_2 \bar{a}_2 = b_3$. Substituting this in (49), and replacing a_1 with $\bar{a}_3$, we get $\underline{[a_3, b_3] = 1}$. With $a_1 = \bar{a}_3$ and $[a_3, b_3] = 1$, the relation (51) simplifies to $\underline{[b_2, a_3] = 1}$. So a_3 commutes with a_2, b_2 and b_3, and therefore, with everything. Since $a_1 = \bar{a}_3$, the surface relation $[a_1, b_1][a_2, b_2][a_3, b_3] = 1$ becomes $\underline{[a_2, b_2] = 1}$. Recall that $b_1 = a_2 b_2 \bar{a}_2 \bar{b}_3$, which now becomes $b_1 = b_2 \bar{b}_3$. Substituting $a_1 = \bar{a}_3$ and $b_1 = b_2 \bar{b}_3$ into (53), and then simplifying it using all the commutativity relations we have so far, we get $\underline{[b_2, b_3] = 1}$. Finally, commuting and canceling the a_1 and a_3 terms in the relation (55) we get $\bar{b}_1 a_2 \bar{b}_3 b_2 \bar{a}_2 = 1$, which, we can rewrite as $b_3 \bar{b}_2 a_2 \bar{b}_3 b_2 \bar{a}_2 = 1$ by substituting $b_1 = b_2 \bar{b}_3$. Since b_2 commutes with both a_2 and b_3, we can simplify the last relation to get $\underline{[a_2, b_3] = 1}$.

Hence $\pi_1(X) = \pi_1(\widetilde{X}) = \mathbb{Z}^4$, generated by a_2, b_2, a_3 and b_3.

Since $\pi_1(X) = \mathbb{Z}^4$ is a virtually poly-$\mathbb{Z}$ group, the Borel conjecture holds in this case by the work of Farrell and Jones [26]. As observed by Friedl and Vidussi, this implies that a symplectic Calabi–Yau surface with $\pi_1 = \mathbb{Z}^4$ is unique up to homeomorphism [31]. So X is homeomorphic to the 4-torus.

Lastly, we will show that for suitable choices of ϕ, we can get X_ϕ realizing all possible rational homology types of symplectic Calabi–Yau surfaces with $b_1 > 0$, which are precisely the rational homology types of torus bundles over tori [48]. In fact, we will get an infinite family realizing all integral homology types of torus bundles over tori. Because the Euler characteristic and the signature are fixed (both zero), the first homology groups determine all the others. Therefore, it will suffice to show that we can get X_ϕ with $H_1(X_\phi) = \mathbb{Z}^2 \oplus (\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z})$ for any given $m_1, m_2 \in \mathbb{N}$.

Let us take $\phi = t_{b_1}^{-m_1} t_{a_3}^{m_2}$, where b_1 and a_3 are as in Figure 4. Note that b_1 and a_3 are disjoint from C_1, C_2, C_1', C_2', so ϕ fixes this set of curves pointwise. For $m = (m_1, m_2)$ any pair of nonnegative integers, let us denote the genus-3 pencil we obtain this way by (X_m, f_m) and its positive factorization by $\underline{W_m = P^\phi P'}$, where $\phi = t_{b_1}^{-m_1} t_{a_3}^{m_2}$. Note that $X_{(0,0)} = X$.

Recall that, every triple of vanishing cycles

$$\{B_0, B_1, B_2\}, \quad \{A_0, A_1, A_2\}, \quad \{B_0', B_1', B_2'\}, \quad \{A_0', A_1', A_2'\}$$

yield the two linearly independent relations

$$a_1 + a_3 = 0, \tag{60}$$

$$b_1 - b_2 + b_3 = 0. \tag{61}$$

There are two conclusions to draw: first, the Dehn twist curves coming from the nonconjugated factor P' induce exactly these relations in $H_1(X_m)$. Second,

the vanishing cycles coming from the conjugated factor P^{ϕ} induce the following relations, which we can easily derive using the Picard–Lefschetz formula:

$$a_1 + m_1 b_1 + a_3 = 0, \tag{62}$$

$$b_1 - b_2 + m_2 a_3 + b_3 = 0. \tag{63}$$

We can now easily see that (60) and (62) together imply $m_1 b_1 = 0$, whereas (61) and (63) imply $m_2 a_3 = 0$. From the relations $a_3 = -a_1$ and $b_3 = b_2 - b_1$, we then conclude that $H_1(X_m)$ is generated by a_1, b_1, a_2, b_2 with only two relations: $m_1 b_1 = 0$ and $m_2 a_1 = 0$. Hence, $H_1(X_m) = \mathbb{Z}^2 \oplus (\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z})$, as claimed.

Note that when $m_1 = m_2 = \pm 1$, we get symplectic Calabi–Yau surfaces with the same integral homology type as $T^2 \times S^2$, but obviously not diffeomorphic to it, as they have different Kodaira dimensions.

5.3. ***The theorem and final remarks.*** For $\{(X_\phi, f_\phi)\}$ symplectic genus-3 pencils prescribed by the positive factorizations W_ϕ, for $\phi \in \mathrm{Mod}(\Sigma_3^4, S)$, we have now proved that:

Theorem 13. *The* $\{(X_\phi, f_\phi)\}$ *are symplectic genus-*3 *Lefschetz pencils whose total spaces are symplectic Calabi–Yau surfaces that realize all integral homology types of torus bundles over tori, and they include a symplectic Calabi–Yau surface homeomorphic to the* 4*-torus and fake symplectic* $T^2 \times S^2$*s.*

We finish with a few observations and comparisons regarding our examples.

Remark 14. The most curious question about our examples is whether every X_ϕ is a torus bundle over a torus, as they are commonly conjectured to exhaust all the diffeomorphism types of symplectic Calabi–Yau surfaces with $b_1 > 0$. After the first version of our paper was publicized, Hamada and Hayano succeeded in proving that our symplectic Calabi–Yau surface that is homeomorphic to the 4-torus [38], is in fact diffeomorphic to it, by comparing the pencil we described on it with a pencil described by Ivan Smith on the standard 4-torus [59] (more on this below). This is so far the only example we know to be standard within this infinite family of examples.[5] If for any conjugation $\phi \neq 1$, it turns out that $\pi_1(X_\phi)$ is not a 4-dimensional solvmanifold group [31; 42], this would imply that X_ϕ is *not* a torus bundle over a torus, and is a new symplectic Calabi–Yau surface. As our arguments in the proof of Theorem 13 show, more generally, if any partial conjugation along any Hurwitz equivalent factorization to the positive factorization W_ϕ results in a pencil with a fundamental group which is not a solvmanifold group, we can arrive

[5]It might be possible to use the recent works of W. Chen in [18; 17] to conclude that some other X_ϕ are also standard by finding finite symplectic symmetries on them. In the special case of trivial ϕ, one can in fact see that the monodromy of the pencil (X, f) with $b_1(X) = 4$ has a $\mathbb{Z}_2$–symmetry under cyclic permutation, which gives rise to a symplectic involution on X.

at a similar conclusion. So far, a handful of examples we examined seem to have the same group theoretic properties as their infrasolvmanifold counter-parts; e.g., they are poly-$\mathbb{Z}$ of Hirsch length 4. In particular, we don't know at this point if any of the fake symplectic $T^2 \times S^2$ we get has $\pi_1 = \mathbb{Z}^2$ so that it would be exotic.

Remark 15. In [59], Ivan Smith constructed genus-3 pencils on torus bundles over tori admitting sections (not all do), by generalizing the algebraic geometric construction of holomorphic genus-3 pencils on abelian surfaces. It is natural to ask whether our examples overlap with Smith's. As Hamada and Hayano showed in [38], this is the case for our 4-torus example, but we don't know much about it beyond that. There are however reasons to think that our family of genus-3 pencils (X_ϕ, f_ϕ) is at least larger than Smith's examples. For comparison, note that any torus bundle over a torus with a section would admit a second disjoint section as well; for any section of a surface bundle over a torus has self-intersection zero [15] and can be pushed off itself. So the family of genus-3 pencils of Smith are determined by a pair $\phi_1, \phi_2 \in \mathrm{Mod}(\Sigma_1^2)$ subject to the relation $[\phi_1, \phi_2] = 1$. On the other hand, our family of genus-3 pencils are parametrized by $\phi \in \mathrm{Mod}(\Sigma_3^4, S)\}$, where $S := \{C_1, C_2, C_1', C_2'\}$, which has a proper subgroup that consists of mapping classes which fix each one of the curves C_1, C_2, C_1', C_2'. Any ϕ in the latter stabilizer group has disjoint support in two copies of Σ_1^2 embedded in Σ_3^4 (the left and the right sides of the surface in Figure 5). So a subset of our family of examples are also parametrized by $\phi_1, \phi_2 \in \mathrm{Mod}(\Sigma_1^2)$, but with no relation to each other whatsoever.

Remark 16. The subfamily of pencils $\{(X_m, f_m) \mid m = (m_1, m_2) \in \mathbb{N}^2\}$ we studied in the proof of Theorem 13 have the following property: they can all be obtained from the 4-torus pencil (X, f) through *fibered Luttinger surgeries* [6; 9]. To see this, first observe that for $\phi = t_{b_1}^{-m_1} t_{a_3}^{m_2}$, we have the positive factorizations

$$W_m = (t_{b_1}^{-m_1} t_{a_3}^{m_2} P t_{a_3}^{-m_2} t_{b_1}^{m_1}) P' = (t_{b_1}^{-1} \cdots t_{b_1}^{-1} t_{a_3} \cdots t_{a_3} P t_{a_3}^{-1} \cdots t_{a_3}^{-1} t_{b_1} \cdots) t_{b_1} P'$$

which are obtained by a sequence of partial conjugations by t_{b_1} and t_{a_3}. Since b_1 and a_3 are disjoint from C_1, C_1', C_2, C_2', they are stabilized by P, which, as a mapping class, is equal to $t_{C_1}^{-1} t_{C_2}^{-1} t_{C_1'} t_{C_2'}$. So each conjugation by a factor of $t_{b_1}^{\pm 1}$ or $t_{a_3}^{\pm 1}$ amounts to performing a Luttinger surgery along a Lagrangian torus swept off by b_1 or a_3 on the regular fibers, over a loop on the base [6; 9]. One can easily see how this observation generalizes to more general conjugations (but perhaps requiring Luttinger surgeries along Lagrangian Klein bottles). With this in mind, we see that if $\{(X_\phi, f_\phi)\}$ contains all the torus bundles over tori, then one would immediately get a proof of an improved version of a conjecture by Ho and Li: that every torus bundle over a torus admits a symplectic structure so that it is obtained via Luttinger surgeries along tori from the 4-torus equipped with the standard product symplectic structure [43, Conjecture 4.9], or, we add, Klein bottles.

Acknowledgements

We would like to thank Weimin Chen, Terry Fuller, Bob Gompf, Jonathan Hillman, Mustafa Korkmaz, Tian-Jun Li and Stefano Vidussi for their interest and helpful conversations on our work presented in this paper. We thank Noriyuki Hamada for his careful reading and feedback on an earlier draft of the paper. The author was partially supported by the National Science Foundation Grants DMS-1510395 and DMS-200532, and the Simons Foundation Grant 634309.

References

[1] A. Akhmedov and B. D. Park, "Exotic smooth structures on small 4-manifolds", *Invent. Math.* **173**:1 (2008), 209–223. MR Zbl

[2] A. Akhmedov and B. D. Park, "Exotic smooth structures on small 4-manifolds with odd signatures", *Invent. Math.* **181**:3 (2010), 577–603. MR Zbl

[3] A. Akhmedov, R. I. Baykur, and B. D. Park, "Constructing infinitely many smooth structures on small 4-manifolds", *J. Topol.* **1**:2 (2008), 409–428. MR Zbl

[4] T. Altunöz, "Genus-3 Lefschetz fibrations and exotic 4-manifolds with $b_2^+ = 3$", *Mediterr. J. Math.* **18**:3 (2021), Paper No. 102, 31. MR

[5] J. Amorós, F. Bogomolov, L. Katzarkov, and T. Pantev, "Symplectic Lefschetz fibrations with arbitrary fundamental groups", *J. Differential Geom.* **54**:3 (2000), 489–545. MR

[6] D. Auroux, "Mapping class group factorizations and symplectic 4-manifolds: some open problems", pp. 123–132 in *Problems on mapping class groups and related topics*, Proc. Sympos. Pure Math. **74**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[7] S. Bauer, "Almost complex 4-manifolds with vanishing first Chern class", *J. Differential Geom.* **79**:1 (2008), 25–32. MR Zbl

[8] R. I. Baykur, "Non-holomorphic surface bundles and Lefschetz fibrations", *Math. Res. Lett.* **19**:3 (2012), 567–574. MR Zbl

[9] R. I. Baykur, "Inequivalent Lefschetz fibrations and surgery equivalence of symplectic 4-manifolds", *J. Symplectic Geom.* **14**:3 (2016), 671–686. MR Zbl

[10] R. I. Baykur and N. Hamada, "Lefschetz fibrations with arbitrary signature", 2020. arXiv 2010.11916

[11] R. I. Baykur and N. Hamada, "A small exotic symplectic rational surface", in preparation.

[12] R. I. Baykur and K. Hayano, "Multisections of Lefschetz fibrations and topology of symplectic 4-manifolds", *Geom. Topol.* **20**:4 (2016), 2335–2395. MR Zbl

[13] R. I. Baykur and M. Korkmaz, "Small Lefschetz fibrations and exotic 4-manifolds", *Math. Ann.* **367**:3-4 (2017), 1333–1361. MR Zbl

[14] R. I. Baykur and N. Sunukjian, "Round handles, logarithmic transforms and smooth 4-manifolds", *J. Topol.* **6**:1 (2013), 49–63. MR Zbl

[15] R. I. Baykur, M. Korkmaz, and N. Monden, "Sections of surface bundles and Lefschetz fibrations", *Trans. Amer. Math. Soc.* **365**:11 (2013), 5999–6016. MR Zbl

[16] R. I. Baykur, K. Hayano, and N. Monden, "Unchaining surgery and topology of symplectic 4-manifolds", 2019. arXiv 1903.02906

[17] W. Chen, "Finite group actions on symplectic Calabi–Yau 4-manifolds with $b_1 > 0$", 2020. Zbl arXiv 2002.12849

[18] W. Chen, "On a class of symplectic 4-orbifolds with vanishing canonical class", *J. Gökova Geom. Topol. GGT* **14** (2020), 55–90. MR Zbl

[19] S. K. Donaldson, "Irrationality and the h-cobordism conjecture", *J. Differential Geom.* **26**:1 (1987), 141–168. MR Zbl

[20] S. K. Donaldson, "Lefschetz pencils on symplectic manifolds", *J. Differential Geom.* **53**:2 (1999), 205–236. MR Zbl

[21] S. K. Donaldson, "Lefschetz pencils and mapping class groups", pp. 151–163 in *Problems on mapping class groups and related topics*, Proc. Sympos. Pure Math. **74**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[22] S. K. Donaldson, "Some problems in differential geometry and topology", *Nonlinearity* **21**:9 (2008), T157–T164. MR Zbl

[23] J. G. Dorfmeister, "Kodaira dimension of fiber sums along spheres", *Geom. Dedicata* **177** (2015), 1–25. MR Zbl

[24] H. Endo and S. Nagami, "Signature of relations in mapping class groups and non-holomorphic Lefschetz fibrations", *Trans. Amer. Math. Soc.* **357**:8 (2005), 3179–3199. MR Zbl

[25] B. Farb and D. Margalit, *A primer on mapping class groups*, Princeton Mathematical Series **49**, Princeton University Press, 2012. MR Zbl

[26] F. T. Farrell and L. E. Jones, *Classical aspherical manifolds*, CBMS Regional Conference Series in Mathematics **75**, Amer. Math. Soc., Providence, RI, 1990. MR Zbl

[27] R. Fintushel and R. J. Stern, "Knots, links, and 4-manifolds", *Invent. Math.* **134**:2 (1998), 363–400. MR Zbl

[28] R. Fintushel and R. J. Stern, "Families of simply connected 4-manifolds with the same Seiberg–Witten invariants", *Topology* **43**:6 (2004), 1449–1467. MR Zbl

[29] R. Fintushel and R. J. Stern, "Double node neighborhoods and families of simply connected 4-manifolds with $b^+ = 1$", *J. Amer. Math. Soc.* **19**:1 (2006), 171–180. MR Zbl

[30] M. H. Freedman, "The topology of four-dimensional manifolds", *J. Differential Geometry* **17**:3 (1982), 357–453. MR Zbl

[31] S. Friedl and S. Vidussi, "On the topology of symplectic Calabi–Yau 4-manifolds", *J. Topol.* **6**:4 (2013), 945–954. MR Zbl

[32] R. Friedman, "Vector bundles and SO(3)-invariants for elliptic surfaces", *J. Amer. Math. Soc.* **8**:1 (1995), 29–139. MR Zbl

[33] T. Fuller, "Hyperelliptic Lefschetz fibrations and branched covering spaces", *Pacific J. Math.* **196**:2 (2000), 369–393. MR Zbl

[34] D. Gieseker, "Global moduli for surfaces of general type", *Invent. Math.* **43**:3 (1977), 233–282. MR Zbl

[35] R. E. Gompf, "A new construction of symplectic manifolds", *Ann. of Math.* (2) **142**:3 (1995), 527–595. MR Zbl

[36] R. E. Gompf and A. I. Stipsicz, *4-manifolds and Kirby calculus*, Graduate Studies in Mathematics **20**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[37] N. Hamada, "Sections of the Matsumoto–Cadavid–Korkmaz Lefschetz fibration", 2016. arXiv 1610.08458

[38] N. Hamada and K. Hayano, "Topology of holomorphic Lefschetz pencils on the four-torus", *Algebr. Geom. Topol.* **18**:3 (2018), 1515–1572. MR Zbl

[39] N. Hamada, R. Kobayashi, and N. Monden, "Nonholomorphic Lefschetz fibrations with (-1)-sections", *Pacific J. Math.* **298**:2 (2019), 375–398. MR Zbl

[40] I. Hambleton and M. Kreck, "On the classification of topological 4-manifolds with finite fundamental group", *Math. Ann.* **280**:1 (1988), 85–104. MR Zbl

[41] I. Hambleton, M. Kreck, and P. Teichner, "Topological 4-manifolds with geometrically two-dimensional fundamental groups", *J. Topol. Anal.* **1**:2 (2009), 123–151. MR Zbl

[42] J. A. Hillman, *Four-manifolds, geometries and knots*, Geometry & Topology Monographs **5**, Geometry & Topology Publications, Coventry, 2002. MR Zbl

[43] C.-I. Ho and T.-J. Li, "Luttinger surgery and Kodaira dimension", *Asian J. Math.* **16**:2 (2012), 299–318. MR Zbl

[44] A. Kas, "On the handlebody decomposition associated to a Lefschetz fibration", *Pacific J. Math.* **89**:1 (1980), 89–104. MR Zbl

[45] M. Korkmaz, "Lefschetz fibrations and an invariant of finitely presented groups", *Int. Math. Res. Not.* **2009**:9 (2009), 1547–1572. MR Zbl

[46] D. Kotschick, "On manifolds homeomorphic to $\mathbb{C}\mathrm{P}^2\#8\overline{\mathbb{C}\mathrm{P}}^2$", *Invent. Math.* **95**:3 (1989), 591–600. MR Zbl

[47] T.-J. Li, "The Kodaira dimension of symplectic 4-manifolds", pp. 249–261 in *Floer homology, gauge theory, and low-dimensional topology*, Clay Math. Proc. **5**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[48] T.-J. Li, "Quaternionic bundles and Betti numbers of symplectic 4-manifolds with Kodaira dimension zero", *Int. Math. Res. Not.* **2006** (2006), art. id. 37385. MR Zbl

[49] T.-J. Li, "Symplectic 4-manifolds with Kodaira dimension zero", *J. Differential Geom.* **74**:2 (2006), 321–352. MR Zbl

[50] T.-J. Li, "Symplectic Calabi–Yau surfaces", pp. 231–356 in *Handbook of geometric analysis, No.* 3, Adv. Lect. Math. (ALM) **14**, International Press, Somerville, MA, 2010. MR Zbl

[51] T. J. Li and A. Liu, "Symplectic structure on ruled surfaces and a generalized adjunction formula", *Math. Res. Lett.* **2**:4 (1995), 453–471. MR Zbl

[52] Y. Matsumoto, "Lefschetz fibrations of genus two—a topological approach", pp. 123–148 in *Topology and Teichmüller spaces* (Katinkulta, 1995), World Sci. Publ., River Edge, NJ, 1996. MR Zbl

[53] D. McDuff, "Erratum to: "The structure of rational and ruled symplectic 4-manifolds" [J. Amer. Math. Soc. **3** (1990), no. 3, 679–712; MR1049697 (91k:58042)]", *J. Amer. Math. Soc.* **5**:4 (1992), 987–988. MR

[54] B. Moishezon, *Complex surfaces and connected sums of complex projective planes*, Lecture Notes in Mathematics **603**, Springer, 1977. MR Zbl

[55] J. Park, "Simply connected symplectic 4-manifolds with $b_2^+ = 1$ and $c_1^2 = 2$", *Invent. Math.* **159**:3 (2005), 657–667. MR Zbl

[56] J. Rana, J. Tevelev, and G. Urzúa, "The Craighero–Gattazzo surface is simply connected", *Compos. Math.* **153**:3 (2017), 557–585. MR Zbl

[57] Y. Sato, "Canonical classes and the geography of nonminimal Lefschetz fibrations over S^2", *Pacific J. Math.* **262**:1 (2013), 191–226. MR Zbl

[58] B. Siebert and G. Tian, "On the holomorphicity of genus two Lefschetz fibrations", *Ann. of Math.* (2) **161**:2 (2005), 959–1020. MR Zbl

[59] I. Smith, "Torus fibrations on symplectic four-manifolds", *Turkish J. Math.* **25**:1 (2001), 69–95. MR Zbl

[60] R. J. Stern, "Will we ever classify simply-connected smooth 4-manifolds?", pp. 225–239 in *Floer homology, gauge theory, and low-dimensional topology*, Clay Math. Proc. **5**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[61] A. I. Stipsicz and Z. Szabó, "An exotic smooth structure on $\mathbb{CP}^2\#6\overline{\mathbb{CP}^2}$", *Geom. Topol.* **9** (2005), 813–832. MR

[62] Z. Szabó, "Exotic 4-manifolds with $b_2^+ = 1$", *Math. Res. Lett.* **3**:6 (1996), 731–741. MR

[63] C. H. Taubes, "The Seiberg–Witten and Gromov invariants", *Math. Res. Lett.* **2**:2 (1995), 221–238. MR Zbl

[64] C. H. Taubes, "SW $\Rightarrow$ Gr: from the Seiberg–Witten equations to pseudo-holomorphic curves", *J. Amer. Math. Soc.* **9**:3 (1996), 845–918. MR Zbl

[65] M. Usher, "The Gromov invariant and the Donaldson–Smith standard surface count", *Geom. Topol.* **8** (2004), 565–610. MR Zbl

[66] M. Usher, "Kodaira dimension and symplectic sums", *Comment. Math. Helv.* **84**:1 (2009), 57–85. MR Zbl

Received 31 Jan 2021. Revised 9 Aug 2021.

R. İNANÇ BAYKUR: baykur@math.umass.edu
Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA, United States

THE OPEN BOOK SERIES **5** (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.223

Khovanov homology and strong inversions

Artem Kotelskiy, Liam Watson and Claudius Zibrowius

There is a one-to-one correspondence between strong inversions on knots in the three-sphere and a special class of four-ended tangles. We compute the reduced Khovanov homology of such tangles for all strong inversions on knots with up to 9 crossings, and discuss these computations in the context of earlier work by the second author (*Adv. Math.* **313** (2017), 915–946). In particular, we provide a counterexample to Conjecture 29 therein, as well as a refinement of and additional evidence for Conjecture 28.

The Brieskorn spheres $\Sigma(2, q, 2nq \mp 1)$ may be obtained by Dehn surgery on a torus knot in the three-sphere, namely, these are the integer homology spheres $S^3_{\pm 1/n}(T_{2,q})$, where $T_{2,q}$ is the positive $(2, q)$ torus knot. These homology spheres admit Seifert fibrations, with base orbifold $S^2(2, q, 2nq \mp 1)$. Denoting by $\mathbf{\Sigma}(A, b)$ the two-fold branched cover of A with branch set b, each of these manifolds admits two descriptions as a two-fold branched cover:

$$S^3_{\pm 1/n}(T_{2,q}) \cong \mathbf{\Sigma}(S^3, T^*_{q,2qn\mp 1}) \cong \mathbf{\Sigma}(S^3, \tau(\pm 1/n)).$$

This construction might be best termed as classical; for our purposes it is helpful to review the notation introduced in [9]. The first of these two-fold branched covers results from an involution on the Seifert fibred space that preserves an orientation on the fibres — we refer to this as the Seifert involution. The second of these arises from the Montesinos involution, which reverses an orientation on the fibres; the branch set in question arises from the Montesinos trick, that is, by first constructing a tangle (B^3, τ) over which the exterior of $T_{2,q}$ is realized as a two-fold branched cover. We review the construction below as it is central to our enumeration of

Kotelskiy is supported by an AMS-Simons travel grant. Watson is supported by an NSERC discovery/accelerator grant. Zibrowius is supported by the Emmy Noether Programme of the DFG, Project number 412851057.
MSC2020: 57K10, 57K18.
Keywords: Khovanov homology, strong inversion, tangle, immersed curves.

tangles. In particular, the branch set $\tau(\pm 1/n)$ is an explicit Montesinos link, which despite the notation depends on q. Generically, these two branch sets are distinct knots. For example, when $q = 5$ and $n = 1$ we have

$$S^3_{+1}(T_{2,5}) \cong \mathbf{\Sigma}(S^3, T^*_{5,9}) \cong \mathbf{\Sigma}(S^3, \tau(+1))$$

and it can be calculated that

$$\dim \widetilde{\mathrm{Kh}}(T^*_{5,9}) = 57 > 15 = \dim \widetilde{\mathrm{Kh}}(\tau(+1)),$$

where the reduced Khovanov homology is taken over the two-element field $\mathbb{F}$. (The Montesinos branch set when $q = 5$ is shown on the right, and the construction of the tangle in this case is reviewed in Figure 1.) This example serves to answer a question due to Ozsváth in the negative: The total dimension of the mod 2 reduced Khovanov homology is not an invariant of two-fold branched covers [9].

This paper focusses on tangles admitting knot exteriors as two-fold branched covers, and the immersed curves that arise as the reduced Khovanov invariants of these tangles [5]. Coefficients are restricted to $\mathbb{F}$ throughout.

1. Strong inversions

A knot K in S^3 is invertible if it admits an isotopy exchanging a choice of orientation on K with the reverse of this choice. A strong inversion is an inversion realized by an involution of the three-sphere. Note that invertible knots which are not strongly invertible exist [4], but that when restricting to hyperbolic knots the two symmetries are equivalent. Following Sakuma [8], given a knot K and a strong inversion h, we will call the pair (K, h) a strongly invertible knot. Strongly invertible knots (K, h) and (K', h') are equivalent if there exists an orientation preserving homeomorphism f on S^3 for which $f(K) = K'$ (so that K and K' are equivalent knots) and $h = f^{-1} \circ h' \circ f$. Alternatively, a strong inversion on a given knot may be viewed as the conjugacy class of an order-two element of the mapping class group of the exterior of the knot; this mapping class group is known as the symmetry group of the knot K. For hyperbolic knots, this group is a subgroup of a dihedral group; see [8, Proposition 3.1].

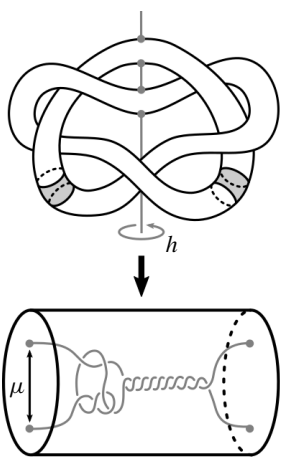


Figure 1. The cinqfoil exterior as a 2-fold branched cover.

A strong inversion gives rise to an involution on a knot's exterior with one-dimensional fixed point set meeting the boundary in exactly 4 points. Note that, according to the Smith conjecture, given an involution on the three-sphere with one-dimensional fixed point set, the fixed point set is unknotted; as a result the fixed point set of the strong inversion on the exterior is a pair of unknotted arcs. Taking the quotient of the order two action on the knot exterior gives rise to a three ball $B^3 = (S^3 \setminus \nu(K))/h$, and the image of the fixed point set of the strong inversion gives rise to a pair of properly embedded arcs $\tau = \mathrm{Im}(\mathrm{Fix}(h))$. Therefore, given a strongly invertible knot, the knot exterior may be viewed as a two-fold branched cover over a four-ended tangle τ in a three ball: $h \curvearrowright (S^3, K)$ so that $S^3 \setminus \nu(K) \cong \mathbf{\Sigma}(B^3, \tau)$. We refer to the tangle $T = (B^3, \tau)$ as the associated quotient tangle to a given knot with strong inversion; see Figure 1 for an explicit example.

In this context, the natural notion of equivalence on tangles is homeomorphism of the pair (B^3, τ) where the boundary is fixed only set-wise. Note that this is, of course, more flexible than the requirement that the boundary be fixed point-wise as is perhaps more common when considering tangle diagrams. In order to be clear about the distinction, we will refer to this latter as a framed tangle. We remark that [9, Definition 3] gives a third notion of equivalence by introducing sutured tangles. This object will not play an explicit role here, however, consulting Figure 1 the reader may find that sutures are a helpful tool for tracking the image of the knot meridian μ in the quotient.

From this point of view, there is a distinguished *trivial* tangle obtained as the quotient of a solid torus, that is, the associated quotient tangle for the trivial knot. The equivalence class of this tangle is the rational tangle. We will have need for choices of representatives $Q_{p/q}$ as described in Figure 2: For any $p/q \in \mathbb{Q}\mathrm{P}^1$, there is a framed tangle diagram $Q_{p/q}$. This choice will allow us to make use of the Montesinos trick:

$$S^3_{p/q}(K) \cong \mathbf{\Sigma}(S^3, \tau(p/q)),$$

where p/q-surgery along the knot K corresponds to the knot $\tau(p/q)$ obtained by gluing the $-p/q$-rational tangle $Q_{-p/q}$ to the tangle τ as in Figure 2 (bottom).

In particular, we are fixing a preferred representative for our associated quotient tangle once and for all: Given (K, h), this is the tangle $T = (B^3, \tau)$ such that the rational closure $\tau(0)$ corresponds to surgery along the Seifert longitude and $\tau(\infty)$ corresponds to $S^3_\infty(K) \cong S^3$. The former implies that the determinant of $\tau(0)$ vanishes, while the latter implies that $\tau(\infty)$ is the unknot, as observed above. (These are also the conditions we check in practice to determine the correct framing.) In fact, there is a bijection between equivalence classes of nontrivial strongly invertible knots (K, h) and tangles (B^3, τ) for which $\tau(\infty)$ is unknotted (this follows from [3, Theorem 2]; compare [10, Proposition 9]). As such, associated quotient tangles provide an invariant of strong inversions.

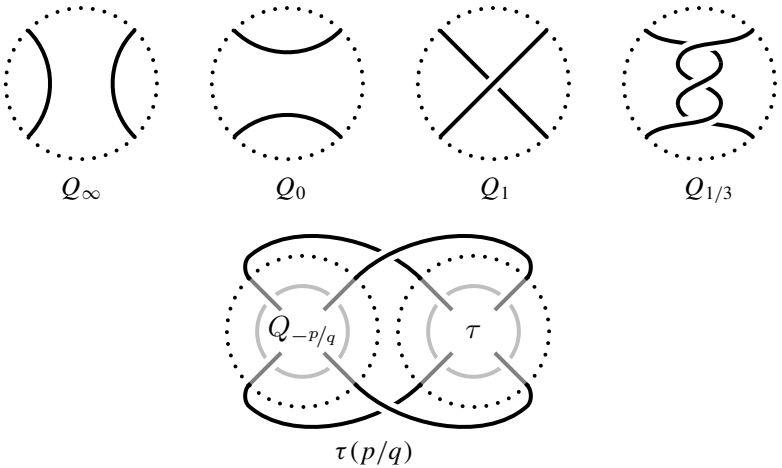


Figure 2. Examples of framed rational tangles (top) and the rational closure of a four-ended tangle τ (bottom).

Strong inversions give a means of enumerating interesting tangles, and we use this strategy below. Observe that, given a hyperbolic knot admitting a pair of distinct strong inversions h_1 and h_2, by Thurston's hyperbolic Dehn filling theorem, $S^3_{p/q}(K)$ is hyperbolic for all but finitely many slopes p/q. Moreover

$$\mathbf{\Sigma}(S^3, \tau_1(p/q)) \cong S^3_{p/q}(K) \cong \mathbf{\Sigma}(S^3, \tau_2(p/q))$$

and, generically, the branch sets $\tau_1(p/q)$ and $\tau_2(p/q)$ are distinct knots. It is a striking fact that for all K with fewer than 9 crossings

$$\dim \widetilde{\mathrm{Kh}}(\tau_1(p/q)) = \dim \widetilde{\mathrm{Kh}}(\tau_2(p/q)),$$

that is, finding a negative answer to Ozsváth's question in the hyperbolic setting is surprisingly difficult. This observation follows from Theorem 3.2 below.

A range of examples of strong inversions and associated quotient tangles are considered in [10]. We will expand this list in a systematic way below, and highlight in particular the knot 9_{46} in the Rolfsen table; see Figure 3. It is worth noting that these tangles distinguish the strong inversions in question: Ignoring one strand, the second tangle contains 8_{19} as a subknot, while the only subknots in the first tangle are trefoils.

In the context of the discussion above, it seems natural to try and articulate a relative version of Ozsváth's question. In particular, are there Khovanov-type tangle invariants that are invariants of the knot K independent on the chosen strong inversion h? We will see that 9_{46} dashes any hope of this (see Counterexample 3.8) and, furthermore, provides hyperbolic counterexamples to Ozsváth's original question (see Remark 3.3).

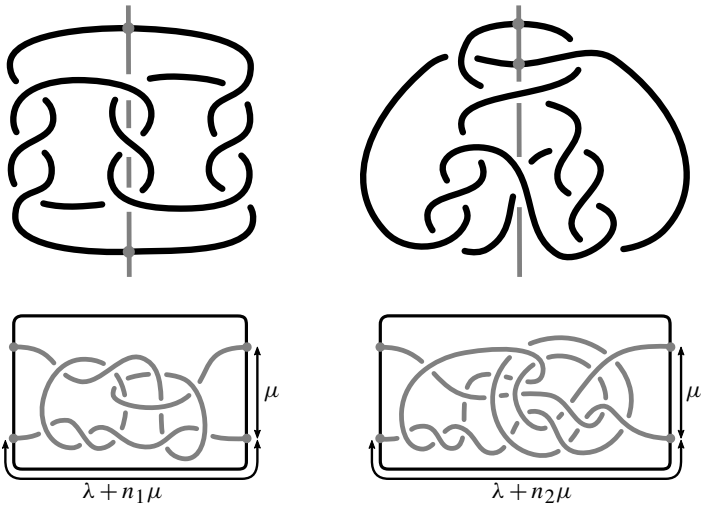


Figure 3. Two strong inversions on the knot 9_{46} (top row) together with the associate quotient tangle for each (bottom row). In order to present minimal crossing diagrams for the associated quotient tangles, this is the one place where we present something other than the preferred tangle representative: The first symmetry gives rise to a tangle T_1 shown with framing $n_1 = 2$, while the second symmetry gives rise to a tangle T_2 with framing $n_2 = -6$.

Interestingly, 9_{46} also makes a star appearance in the recent work of Boyle and Issa [2]. In particular, this knot is slice, but it has nonzero equivariant four-genus; see [2, Figure 14] and the surrounding discussion.

2. Review of the tangle invariants $\widetilde{\mathrm{Kh}}$ and $\widetilde{\mathrm{BN}}$

Let T be a four-ended tangle in the three-ball B^3. In earlier work we interpreted the Bar–Natan tangle invariant $[\![T]\!]_{/l}$ in terms of *multicurves*, that is, isotopy classes of collections of immersed curves in the four-punctured sphere $\partial B^3 \smallsetminus \partial T$ [1; 5]. After choosing a distinguished tangle end $*$ of T, the construction outputs two multicurve-valued invariants $\widetilde{\mathrm{Kh}}(T)$ and $\widetilde{\mathrm{BN}}(T)$:

$$T \overset{[1]}{\longmapsto} [\![T]\!]_{/l} \overset{[5]}{\longmapsto} \widetilde{\mathrm{Kh}}(T),\ \widetilde{\mathrm{BN}}(T) \looparrowright S^2_{4,*} = \partial B^3 \smallsetminus \partial T.$$

These multicurves are equipped with a bigrading, which we expand on below. Also, strictly speaking, components of $\widetilde{\mathrm{BN}}(T)$ come equipped with local systems. However, in this paper, we will suppress this subtlety, since over $\mathbb{F}$ these local systems are trivial in all known examples.

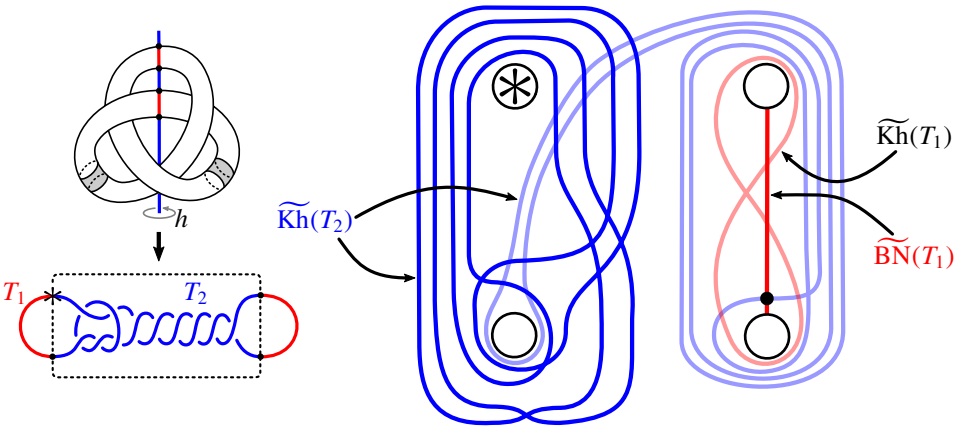


Figure 4. Illustrated on the left is the trefoil with strong inversion, highlighting the decomposition of the unknot into tangles T_1 and T_2, where T_2 is the preferred representative of the associated quotient tangle to the strong inversion. The corresponding intersection picture is shown on the right, where all of the relevant information has been projected to the front face of the sphere: The Lagrangian Floer homology $\mathrm{HF}(\widetilde{\mathrm{BN}}(T_1), \widetilde{\mathrm{Kh}}(T_2)) = \mathbb{F}$ recovers the one-dimensional reduced Khovanov homology of the unknot.

The curve invariants enjoy the following gluing property [5, Theorem 1.9]: Given a decomposition of a knot $K = T_1 \cup T_2$ into four-ended tangles, the reduced Khovanov homology of the knot is recovered via Lagrangian Floer homology:

$$\widetilde{\mathrm{Kh}}(K) \cong \mathrm{HF}(\mathrm{m}(\widetilde{\mathrm{BN}}(T_1)), \widetilde{\mathrm{Kh}}(T_2)) \tag{1}$$

where m is the map identifying the two four-punctured spheres. An example illustrating this gluing formula is given in Figure 4.

Let us recall some basic facts about $\widetilde{\mathrm{Kh}}(T)$ and $\widetilde{\mathrm{BN}}(T)$ from [5, Section 6]: Both invariants may consist of multiple components. A component is either an immersion of a circle or an immersion of an interval. In the first case we call a component compact; in the second, noncompact. The invariant $\widetilde{\mathrm{BN}}(T)$ contains exactly one noncompact component (unless the tangle T contains some closed component, which is not the case in the present context), whereas $\widetilde{\mathrm{Kh}}(T)$ consists only of compact components. These curves become easier to manage when considered in a certain covering space of $S^2_{4,*}$, namely the planar cover that factors through the toroidal two-fold cover:

$$(\mathbb{R}^2 \smallsetminus \mathbb{Z}^2) \to (T^2 \smallsetminus 4\mathrm{pt}) \to S^2_{4,*}$$

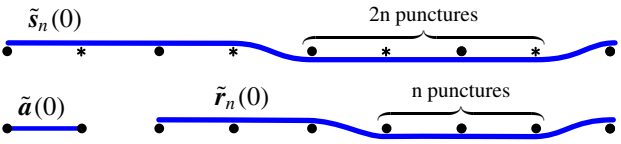


Figure 5. The curves $\tilde{\boldsymbol{a}}(0)$, $\{\tilde{\boldsymbol{s}}_n(0)\}_{n\geqslant 1}$, and $\{\tilde{\boldsymbol{r}}_n(0)\}_{n\geqslant 1}$.

Definition 2.1. Given an immersed curve $c \looparrowright S^2_{4,*}$, denote by $\tilde{c}$ a lift of c to the cover $\mathbb{R}^2 \smallsetminus \mathbb{Z}^2$. For $n \in \mathbb{N}$, let $\boldsymbol{a}(0)$, $\boldsymbol{r}_n(0)$, and $\boldsymbol{s}_{2n}(0)$ be the immersed curves in $S^2_{4,*}$ that respectively admit lifts to the curves $\tilde{\boldsymbol{a}}(0)$, $\tilde{\boldsymbol{r}}_n(0)$, and $\tilde{\boldsymbol{s}}_{2n}(0)$ in Figure 5. For every $p/q \in \mathbb{Q}\mathrm{P}^1$, we respectively define the curves $\boldsymbol{a}(p/q)$, $\boldsymbol{r}_n(p/q)$, and $\boldsymbol{s}_{2n}(p/q)$ as the images of $\boldsymbol{a}(0)$, $\boldsymbol{r}_n(0)$, and $\boldsymbol{s}_{2n}(0)$ under the action of

$$\begin{bmatrix} q & r \\ p & s \end{bmatrix}$$

considered as an element of the mapping class group fixing the special puncture $\mathrm{Mod}(S^2_{4,*}) \cong PSL(2,\mathbb{Z})$, where $qs - pr = 1$. (This transformation maps straight lines of slope 0 to straight lines of slope p/q.) We call $\boldsymbol{a}(p/q)$ a *rational arc of slope* p/q; we call $\boldsymbol{r}_n(p/q)$ a curve of *rational type*, *slope* p/q, *and length* n; and $\boldsymbol{s}_{2n}(p/q)$ a curve of *special type*, *slope* p/q, *and length* $2n$.

Example 2.2. The curve $\widetilde{\mathrm{Kh}}(T_2)$ from Figure 4 consists of the special component $\boldsymbol{s}_4(\infty)$ and the rational component $\boldsymbol{r}_1(4)$. Furthermore, $\widetilde{\mathrm{BN}}(T_1) = \{\boldsymbol{a}(\infty)\}$ is the red vertical arc, while $\widetilde{\mathrm{Kh}}(T_1) = \{\boldsymbol{r}_1(\infty)\}$ consists of the figure-eight curve lying in a small neighbourhood of this arc. More generally, justifying the terminology, naturality of the invariants under the mapping class group action [5, Theorem 1.13] implies that $\widetilde{\mathrm{BN}}(Q_{p/q}) = \{\boldsymbol{a}(p/q)\}$ and $\widetilde{\mathrm{Kh}}(Q_{p/q}) = \{\boldsymbol{r}_1(p/q)\}$ — so, rational tangles have rational invariants.

The invariants $\widetilde{\mathrm{BN}}(T)$ and $\widetilde{\mathrm{Kh}}(T)$ are topological interpretations of algebraic invariants $\mathit{Д}(T)$ and $\mathit{Д}_1(T)$, respectively, which are type D structures over the algebra $\mathcal{B}$. For example, the curve $\boldsymbol{s}_4(\infty)$, which is the blue curve on the left of Figure 7, corresponds to the type D structure in Figure 6; the arc $\boldsymbol{a}(0)$ corresponds to the one-generator complex $[{}^0\bullet_0]$. To translate between these two viewpoints, we need to fix a parametrization of the four-punctured sphere $S^2_{4,*}$, which is indicated by the two grey dashed arcs in Figure 7. Briefly, intersection points of these arcs with a given curve give rise to the generators of the corresponding type D structure and paths between intersection points determine the differential; for more details, see [5, Example 1.6]. The generators of the type D structures carry a bigrading, that is a quantum grading q and a homological grading h. We indicate these gradings on generators $\bullet$ and the corresponding intersection points by super- and subscripts like so: ${}^q\bullet_h$. Also the algebra $\mathcal{B}$ is equipped with quantum and homological

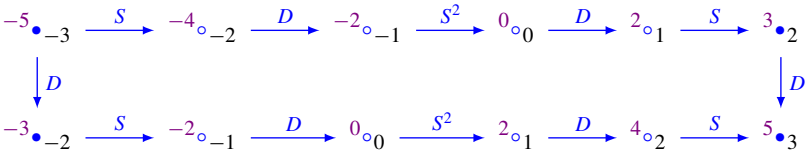


Figure 6. The symmetrically bigraded type D structure corresponding to the special curve $s_{4n}(\infty)$ for $n = 1$. The type D structures for $n > 1$ look similar. The total number of generators in idempotent $\circ$ is equal to $8n$. The absolute bigrading is fixed by requiring that the minimal and maximal homological and quantum gradings are $\pm(2n+1)$ and $\pm(4n+1)$, respectively.

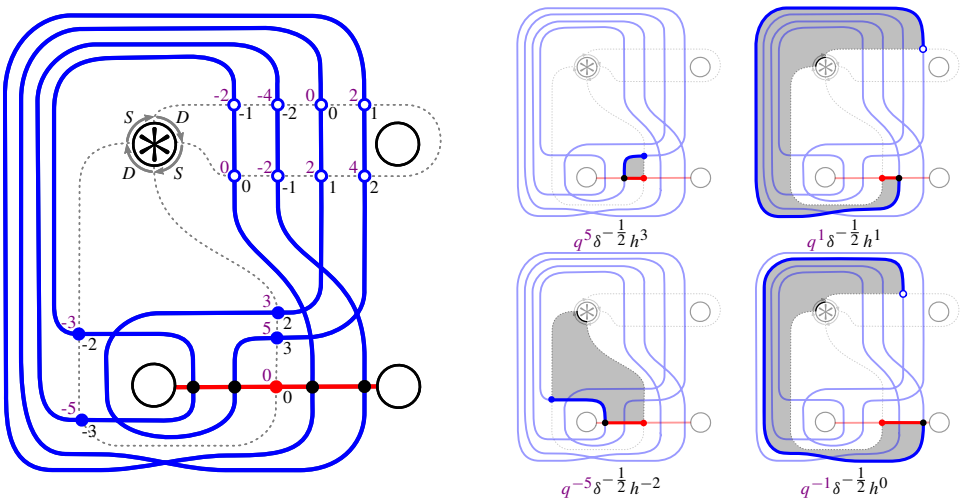


Figure 7. Computing the gradings associated with the four-dimensional vector space $\mathrm{HF}(\boldsymbol{a}(0), s_4(\infty))$, according to Lemma 2.5.

gradings, which are determined by $q(D) = 2q(S) = -2$, $h(D) = h(S) = 0$. The differentials of type D structures reduce homological grading by 1 and preserve the quantum grading, namely if there is an arrow $x \xrightarrow{a} y$ in the differential, then $q(y) + q(a) - q(x) = 0$, $h(y) + h(a) - h(x) = 1$. Thus, the bigrading of any given component of a multicurve is determined by the bigrading of a single generator which lies on this component.

The bigradings on $\widetilde{\mathrm{BN}}(T)$ and $\widetilde{\mathrm{Kh}}(T)$ depend on an orientation of the tangle T. This dependence is discussed in [5, Proposition 4.8]. If T is an unoriented tangle, $\widetilde{\mathrm{BN}}(T)$ and $\widetilde{\mathrm{Kh}}(T)$ only carry *relative* bigradings, i.e., bigradings that are well-defined up to an overall shift. For tangles whose components have linking number 0, such as the trivial tangle  , all orientations induce the same bigrading.

Theorem 2.3 [6, Theorem 2.15]. *For any pointed Conway tangle T, every component of $\widetilde{\mathrm{Kh}}(T)$ is equal to $\boldsymbol{r}_n(p/q)$ or $\boldsymbol{s}_{2n}(p/q)$ for some $n \in \mathbb{N}$ and $p/q \in \mathbb{QP}^1$, up to some bigrading shift. In other words, components of $\widetilde{\mathrm{Kh}}(T)$ are completely classified by their type, slope, length, and bigrading.*

In the context of Heegaard Floer homology, analogous properties are known for the tangle invariant $\mathrm{HFT}(T)$ [11]. Note, however, that components of the invariant $\widetilde{\mathrm{BN}}(T)$ are known to be much more complicated even when restricting to only compact components.

We conclude this section with a computation of the Lagrangian Floer intersection pairing between a simple arc and the special curves $\boldsymbol{s}_{4n}(\infty)$, which will play a central role later.

Definition 2.4. Define $V_\varkappa$ as the four-dimensional bigraded vector space supported in δ-grading 0 and quantum gradings -5, -1, $+1$, and $+5$.

Lemma 2.5. *We have*

$$\mathrm{HF}(\boldsymbol{a}(0), \boldsymbol{s}_4(\infty)) \cong q^0\delta^{-\frac{1}{2}}h^{\frac{1}{2}}V_\varkappa.$$

More generally, for any positive integer n,

$$\mathrm{HF}(\boldsymbol{a}(0), \boldsymbol{s}_{4n}(\infty)) \cong \bigoplus_{i=1}^{n} q^{4(2i-n-1)}\delta^{-\frac{1}{2}}h^{\frac{1}{2}+2(2i-n-1)}V_\varkappa.$$

In both cases, the isomorphism holds as absolutely bigraded vector spaces if $\boldsymbol{s}_{4n}(\infty)$ is symmetrically bigraded (in the sense of Figure 6) *and $\boldsymbol{a}(0)$ agrees with $\widetilde{\mathrm{BN}}(\infty)$ as a **bigraded** curve.*

Proof. The total dimension of $\mathrm{HF}(\boldsymbol{a}(0), \boldsymbol{s}_{4n}(\infty))$ is $4n$, since the minimal number of intersection points between the two curves is $4n$. For the computation of the bigrading, we will focus on the case $n = 1$; the case $n > 1$ is similar. Recall how this works [5, Section 7.2]: The bigrading of an intersection point $\bullet$ generating $\mathrm{HF}(\boldsymbol{a}(0), \boldsymbol{s}_4(\infty))$ is computed by considering a path on $\boldsymbol{a}(0) \cup \boldsymbol{s}_4(\infty)$ which starts at $x = {}^{0}\!\bullet_{0}$, turns right at the intersection point $\bullet$ and ends at the nearest intersection point $y = {}^{q}\!\circ_{h}$ or $y = {}^{q}\!\bullet_{h}$ of $\boldsymbol{s}_4(\infty)$ with the parametrization. These paths are illustrated on the right of Figure 7. Each of these paths in $S^2_{4,*}$ is homotopic relative to the parametrizing arcs to some path γ on the boundary of the special puncture; these homotopies are indicated on the right of Figure 7 by the shaded disks. We set $q(\gamma) = 0, -1, -2$, depending on whether the path γ is constant, equal to S, or equal to D, respectively. (More generally, these paths correspond to some algebra elements in $\mathcal{B}$, whose quantum gradings define $q(\gamma)$.) Then

$$h(\bullet) = h(y) - h(x) \quad \text{and} \quad q(\bullet) = q(y) - q(x) + q(\gamma).$$

Finally, we compute the δ-grading from the identity $\delta + h = \frac{1}{2}q$. □

3. The Khovanov homology of strong inversions

A strongly invertible knot (K, h) gives rise to an associated four-ended quotient tangle $T = (B^3, \tau)$. (When we have need for it, we will use the notation $T_{K,h}$ to remember the dependence on the strongly invertible knot.) We use the same preferred framing as in Section 1; in particular, $\tau(\infty)$ is the unknot. This imposes strong restrictions on the curve invariant $\widetilde{\mathrm{Kh}}(T)$, namely, it implies that

$$\widetilde{\mathrm{Kh}}(\tau(\infty)) = \mathrm{HF}\big(\widetilde{\mathrm{BN}}(\,)\,, \widetilde{\mathrm{Kh}}(T)\big) = \mathrm{HF}(\boldsymbol{a}(\infty), \widetilde{\mathrm{Kh}}(T)) = \mathbb{F}.$$

In other words, $\widetilde{\mathrm{Kh}}(T)$ is homotopic to a multicurve that intersects the arc $\boldsymbol{a}(\infty)$ only once, as is the case for the example $\widetilde{\mathrm{Kh}}(T_2)$ from Figure 4. In the light of the classification given in Theorem 2.3, this implies:

Theorem 3.1. *Let h be a strong inversion on a knot $K \subset S^3$. Then*

$$\widetilde{\mathrm{Kh}}(T_{K,h}) = \boldsymbol{r}_1(k) \cup \boldsymbol{s}_{2n_1}(\infty) \cup \cdots \cup \boldsymbol{s}_{2n_N}(\infty)$$

for some integers $k \in \mathbb{Z}$ and $N, n_1, \ldots, n_N \in \mathbb{Z}_{>0}$. □

In fact, with the help of a computer [12], we show the following:

Theorem 3.2. *Let h be a strong inversion on a knot $K \subset S^3$ with at most 9 crossings. With notation as in Theorem 3.1, k is divisible by 4 and $n_i = 2$ for $i = 1, \ldots, N$.* □

Table 1 shows the ungraded invariants for all pairs (K, h) from Theorem 3.2; the bigraded invariants are listed in Section 4. To determine the absolute bigrading, we use the braid-like orientation on $T_{K,h}$. Note that, in order to match Sakuma's table [8], the first row in Table 1 describes the left-hand trefoil and its rational component has slope -4, while Figure 4 depicts the right-hand trefoil and its rational component has slope 4. Among tabulated knots through 9 crossings, there are 57 knots that admit two distinct strong inversions. For all but one of these knots with two distinct strong inversions, the ungraded invariants $\widetilde{\mathrm{Kh}}(T)$ agree; these pairs of strong inversions are indicated in Table 1 by the superscript ♠.

Remark 3.3. The knot 9_{46} is the only knot in this collection whose strong inversions 9^1_{46} and 9^2_{46} can be distinguished by their ungraded invariants $\widetilde{\mathrm{Kh}}(T)$, as shown in the two highlighted rows in Table 1. Here, we follow the same numbering convention for strong inversions as in [8]. Note that both the slope k and the number n of special components $\boldsymbol{s}_4(\infty)$ distinguish 9^1_{46} and 9^2_{46}. With this example in hand, one can check that $\dim \widetilde{\mathrm{Kh}}(\tau_1(n)) \neq \dim \widetilde{\mathrm{Kh}}(\tau_2(n))$ by applying the pairing formula (1), despite the fact that $\boldsymbol{\Sigma}(S^3, \tau_1(n)) \cong \boldsymbol{\Sigma}(S^3, \tau_2(n))$ by construction. Calculations for $\dim \widetilde{\mathrm{Kh}}(\tau_i(p/q))$ with $i = 1, 2$ can be obtained as well with a little more patience.

3_1	−4	1	$8_4^\spadesuit$	−4	9	9_1	−16	4	$9_{18}^\spadesuit$	−8	20	$9_{37}^\spadesuit$	0	22
$4_1^\spadesuit$	0	2	$8_5^\spadesuit$	8	10	$9_2^\spadesuit$	−4	7	$9_{19}^\spadesuit$	0	20	9_{38}	−8	28
5_1	−8	2	$8_6^\spadesuit$	−4	11	$9_3^\spadesuit$	12	9	$9_{20}^\spadesuit$	−8	20	9_{39}	4	27
$5_2^\spadesuit$	−4	3	$8_7^\spadesuit$	4	11	$9_4^\spadesuit$	−8	10	$9_{21}^\spadesuit$	4	21	$9_{40}^\spadesuit$	−4	37
$6_1^\spadesuit$	0	4	$8_8^\spadesuit$	0	12	$9_5^\spadesuit$	4	11	9_{22}	4	21	9_{41}	0	24
$6_2^\spadesuit$	−4	5	$8_9^\spadesuit$	0	12	$9_6^\spadesuit$	−12	13	$9_{23}^\spadesuit$	−8	22	9_{42}	0	4
$6_3^\spadesuit$	0	6	8_{10}	4	13	$9_7^\spadesuit$	−8	14	9_{24}	0	22	9_{43}	8	6
7_1	−12	3	$8_{11}^\spadesuit$	−4	13	$9_8^\spadesuit$	−4	15	9_{25}	−4	23	9_{44}	0	8
$7_2^\spadesuit$	−4	5	$8_{12}^\spadesuit$	0	14	$9_9^\spadesuit$	−12	15	$9_{26}^\spadesuit$	4	23	9_{45}	−4	11
$7_3^\spadesuit$	8	6	$8_{13}^\spadesuit$	0	14	$9_{10}^\spadesuit$	8	16	$9_{27}^\spadesuit$	0	24	9_{46}^1	0	4
$7_4^\spadesuit$	4	7	$8_{14}^\spadesuit$	4	15	$9_{11}^\spadesuit$	8	16	$9_{28}^\spadesuit$	−4	25	9_{46}^2	−4	7
$7_5^\spadesuit$	−8	8	$8_{15}^\spadesuit$	−8	16	$9_{12}^\spadesuit$	−4	17	9_{29}	−4	25	9_{47}	4	13
$7_6^\spadesuit$	−4	9	8_{16}	−4	17	$9_{13}^\spadesuit$	8	18	9_{30}	0	26	$9_{48}^\spadesuit$	−4	13
$7_7^\spadesuit$	0	10	$8_{18}^\spadesuit$	0	22	$9_{14}^\spadesuit$	0	18	$9_{31}^\spadesuit$	4	27	9_{49}	8	12
$8_1^\spadesuit$	0	6	8_{19}	8	2	$9_{15}^\spadesuit$	4	19	9_{34}	0	34			
$8_2^\spadesuit$	−8	8	8_{20}	0	4	$9_{16}^\spadesuit$	12	19	$9_{35}^\spadesuit$	−4	13			
$8_3^\spadesuit$	0	8	$8_{21}^\spadesuit$	−4	7	$9_{17}^\spadesuit$	−4	19	9_{36}	8	18			

Table 1. The ungraded reduced Khovanov homology of the tangles $T_{K,h}$ associated with strong inversions h on all prime knots K with up to 9 crossings. There are three columns in this table. The first specifies the knots K. Those marked by a superscript $\spadesuit$ admit two distinct strong inversions, and the associated ungraded invariants agree (and are hence listed together). The knot 9_{46} admits two strong inversions, but their invariants are distinct; see the two highlighted rows. All remaining knots admit only a single strong inversion. The second column specifies the slope k of the single rational component $\boldsymbol{r}_1(k)$ in $\widetilde{\mathrm{Kh}}(T_{K,h})$, and the third column gives the number N of special components, all of which are equal to $\boldsymbol{s}_4(\infty)$.

As stated in Section 2, $\widetilde{\mathrm{Kh}}(T)$ for a general four-ended tangle only consists of rational and special curves; it would be interesting to give a geometric interpretation for the slopes of these curves. For quotient tangles $T_{K,h}$ of strong inversions (K, h), the existence of the unknot closure implies that the slope of any special curve is fixed and the slope of the rational component is an integer. As we can see from the second column in Table 1, this integer may be nonzero; in other words, the slope of the rational component of $\mathrm{Kh}(T_{K,h})$ need not agree with the slope given by the rational longitude of the knot K. Our computations raise the following:

Questions 3.4. *Is there a geometric/topological meaning of the slope of the rational component of* $\widetilde{\mathrm{Kh}}(T_{K,h})$? *Is the slope always divisible by* 4?

We now focus on the special components:

Definition 3.5. Given a knot with a strong inversion (K, h), let $\boldsymbol{s}(K, h)$ be the set of special components of $\widetilde{\mathrm{Kh}}(T_{K,h})$. Let $\boldsymbol{a}(0) = \widetilde{\mathrm{BN}}\left(\text{  }\right)$ as in Figure 7. Then define

$$\varkappa(K, h) := \mathrm{HF}(\boldsymbol{a}(0), \boldsymbol{s}(K, h))$$

In [10], the second author defined an invariant of the same name. The following lemma justifies our notation:

Lemma 3.6. *The invariant* $\varkappa(K, h)$ *agrees with the invariant from* [10] *as a relatively bigraded vector space.*

Proof. The original invariant $\varkappa(K, h)$ was defined as a certain finite-dimensional quotient of the inverse limit $\varprojlim \widetilde{\mathrm{Kh}}(\tau(n))$ associated with the maps

$$\widetilde{\mathrm{Kh}}(\tau(n+1)) \to \widetilde{\mathrm{Kh}}(\tau(n))$$

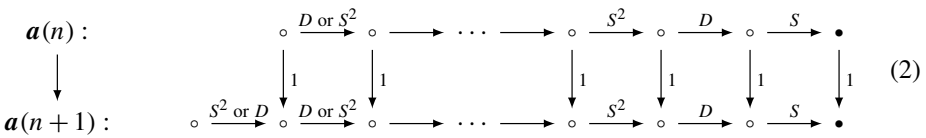
induced by resolving a crossing [10, Proposition 11]. Since this inverse limit is determined by the maps for positive n we can assume in the following that $n > 0$. By the gluing property of the multicurve tangle invariants,

$$\widetilde{\mathrm{Kh}}(\tau(n)) \cong \mathrm{HF}(\boldsymbol{a}(n), \widetilde{\mathrm{Kh}}(\tau))$$

The right-hand side can also be interpreted as the homology of the morphism space between the type D structures associated with $\boldsymbol{a}(n)$ and $\widetilde{\mathrm{Kh}}(\tau)$ [5, Theorem 1.5]. Gluing is functorial; in fact, the map

$$\mathrm{HF}(\boldsymbol{a}(n+1), \widetilde{\mathrm{Kh}}(\tau)) \xrightarrow{\cong} \widetilde{\mathrm{Kh}}(\tau(n+1)) \to \widetilde{\mathrm{Kh}}(\tau(n)) \xrightarrow{\cong} \mathrm{HF}(\boldsymbol{a}(n), \widetilde{\mathrm{Kh}}(\tau))$$

is induced by precomposition with the following morphism from the type D structure for $\boldsymbol{a}(n)$ to the type D structure for $\boldsymbol{a}(n+1)$:



(2)

This can be seen explicitly by induction or by observing that the Lagrangian Floer homology $\mathrm{HF}(\boldsymbol{a}(n), \boldsymbol{a}(n+1))$ computes the Bar–Natan homology of the unknot, so there is only one nonzero equivalence class of morphisms in the correct quantum grading.

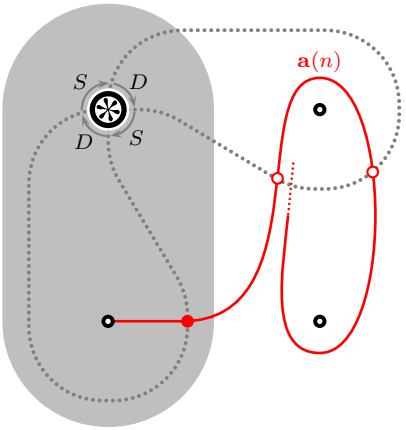


Figure 8. A schematic for the Lagrangian Floer pairing $\mathrm{HF}(\boldsymbol{a}(n), \boldsymbol{s}_{2i}(\infty))$ in the proof of Lemma 3.6: The curve $\boldsymbol{s}_{2i}(\infty)$ can be drawn such that it lies in the shaded region, and so do all the intersection points with $\boldsymbol{a}(n)$. Precomposition with the map (2) induces the identity on the corresponding morphisms.

Now apply the (contravariant) functor $\mathrm{HF}(-, \widetilde{\mathrm{Kh}}(\tau))$ to this map. We can do this for each component of $\widetilde{\mathrm{Kh}}(\tau)$ separately, using Theorem 3.1. First, for any integer k, $\varprojlim \mathrm{HF}(\boldsymbol{a}(n), \boldsymbol{r}_1(k))$ is isomorphic to $\varprojlim \mathrm{HF}(\boldsymbol{a}(n-k), \boldsymbol{r}_1(0))$. The latter computes the invariant of the unknot (with its unique strong inversion), which vanishes by [10, Theorem 1]. Moreover, a simple computation shows that the map

$$\mathrm{HF}(\boldsymbol{a}(n+1), \boldsymbol{s}_{2i}(\infty)) \to \mathrm{HF}(\boldsymbol{a}(n), \boldsymbol{s}_{2i}(\infty))$$

induced by (2) is an isomorphism for any integer $i > 0$; this is illustrated in Figure 8. Thus

$$\varprojlim \mathrm{HF}(\boldsymbol{a}(n), \boldsymbol{s}_{2i}(\infty)) \cong \mathrm{HF}(\boldsymbol{a}(0), \boldsymbol{s}_{2i}(\infty)).$$

Moreover, the subspace that we need to quotient by to obtain the invariant from [10] vanishes. This proves the claim. □

Remark 3.7. Since the maps for the inverse limit $\varprojlim \widetilde{\mathrm{Kh}}(\tau(n))$ preserve the homological grading, the original invariant $\varkappa(K, h)$ carries an absolute (co)homological grading. The identification with our invariant respects this grading, so the two definitions actually give rise to the same absolute homological grading. However, the quantum and δ-gradings on the original invariant are only well-defined as relative gradings. An ad hoc construction to lift these to absolute gradings is discussed in [10, Section 7]. We have not attempted to relate these lifts to the absolute bigrading on our invariant; in general, they do not agree.

Counterexample 3.8. Returning to the pair of strong inversions on the knot 9_{46}, we calculate that

$$\dim \varkappa(K, h_1) = 16 < 28 = \dim \varkappa(K, h_2)$$

contrary to [10, Conjecture 29]. This conjecture was originally posed in the hope that a relative variant of Ozsváth's question might help to explain the examples described in [9], which depend heavily on the Seifert structure. Comparing with Remark 3.3 and the proof of Lemma 3.6, it is worth noting a simple relationship between $\varkappa(K, h)$ and $\widetilde{\mathrm{Kh}}(\tau(n))$ when restricting to the preferred framing: According to Definition 3.5,

$$\dim \widetilde{\mathrm{Kh}}(\tau(n)) = \dim \varkappa(K, h) + \dim \mathrm{HF}(\boldsymbol{a}(n), \boldsymbol{r}_1(k)),$$

where k is as in Theorem 3.2. Note that the second summand is the dimension of the reduced Khovanov homology of the $(2, n-k)$-torus link.

On the other hand, certain structural properties appear to persist:

Conjecture 3.9 [10, Conjecture 28]. *For any strong inversion (K, h), the vector space $\varkappa(K, h)$ is a direct sum of copies of $V_\varkappa$. In particular,* $\dim \varkappa(K, h)$ *is divisible by* 4.

By Lemma 2.5, the vector space $V_\varkappa$ is precisely the Lagrangian Floer homology of the arc $\boldsymbol{a}(0)$ and the special curve $\boldsymbol{s}_4(\infty)$. So a strong inversion (K, h) satisfies Conjecture 3.9 if all special components of $\widetilde{\mathrm{Kh}}(T_{K,h})$ are equal to $\boldsymbol{s}_4(\infty)$ up to a grading shift. In particular, the calculations summarized in Table 1 verify Conjecture 3.9 for all strong inversions on tabulated knots through 9 crossings.

We remark that Conjecture 3.9 implies that the slope of the rational component of $\widetilde{\mathrm{Kh}}(T_{K,h})$ should be divisible by 4 (Questions 3.4, second part). This can be seen as follows. First note that the group $H_1(\boldsymbol{\Sigma}(S^3, \tau(0))) \cong H_1(S^3_0(K))$ has positive rank hence $\det(\tau(0)) = 0$. The determinant of a link can also be computed by evaluating the Jones polynomial at -1, which agrees with the Euler characteristic of Khovanov homology with respect to the δ-grading:

$$\begin{aligned} 0 = \det(\tau(0)) &= |V(-1)| \\ &= |\chi_\delta \widetilde{\mathrm{Kh}}(\tau(0))| = |\chi_\delta \mathrm{HF}(\boldsymbol{a}(0), \widetilde{\mathrm{Kh}}(T))|. \end{aligned}$$

This means that the δ-graded intersections of $\boldsymbol{a}(0)$ with special components cancel out intersections with the rational component. Conjecture 3.9 implies that intersections with special components come in groups of 4 concentrated in a single δ-grading, and so the number of intersections with the rational component — also concentrated in a single δ-grading — should be divisible by 4.

Incorporating Lemma 2.5 gives the following refinement of Conjecture 3.9:

Conjecture 3.10. *The length of any special component of* $\widetilde{\mathrm{Kh}}(T_{K,h})$ *for any strongly invertible knot* (K, h) *is divisible by* 4.

The special components in all examples that we have tabulated are equal to $\boldsymbol{s}_4(\infty)$, but there exist strongly invertible knots (K, h) for which $\widetilde{\mathrm{Kh}}(T_{K,h})$ contains special components $\boldsymbol{s}_{4n}(\infty)$ with $n \neq 1$.

As we can see from Table 1, the ungraded invariant $\varkappa(K, h)$ cannot tell all strong inversions apart. However, the *absolutely bigraded* invariant $\varkappa(K, h)$ can distinguish any pair of strong inversion for knots up to 9 crossings.

Question 3.11 [10, Question 19]. *Is there a hyperbolic knot* $K \subset S^3$ *with two distinct strong inversions that cannot be distinguished by the absolutely bigraded invariant* $\varkappa$?

The restriction to hyperbolic is important here, and was omitted in error in the statement of [10, Question 19]. Note that the question is not interesting for torus knots, as these admit unique strong inversions. However, it is possible to generate satellite knots with strong inversions whose associated quotient tangles differ by mutation on a subtangle. Owing to the insensitivity of Khovanov homology to mutation, such symmetries cannot be separated.

Remark 3.12. Although our focus has been entirely on four-ended tangles admitting an unknot closure, we expect Conjecture 3.10 to hold more generally: For any four-ended tangle T, any special component of $\widetilde{\mathrm{Kh}}(T)$ should be equal to $\boldsymbol{s}_{4n}(p/q)$ for some $p/q \in \mathbb{Q}\mathrm{P}^1$ and $n > 0$.

Remark 3.13. In this paper, we have worked exclusively over the field $\mathbb{F}$ of two elements; all of the above questions should be read with this coefficient system in place. We conclude with a comment about other fields of coefficients. Unlike knot Floer homology, Khovanov homology is known to behave very differently over fields of different characteristic. Many of the conjectures above fail if we work over a field different from $\mathbb{F}$. For example, consider the second strong inversion 7_4^2 on the knot 7_4 in Sakuma's table [8]. Its invariant $\widetilde{\mathrm{Kh}}(T_{7_4^2}; \mathbb{Z}/3)$ consists of a single rational component $\boldsymbol{r}_1(6)$, two special components $\boldsymbol{s}_4(\infty)$, but also five special components $\boldsymbol{s}_2(\infty)$. So the corresponding version of Conjecture 3.9 over $\mathbb{Z}/3$ does not hold. Also note that the slope of the rational component is no longer divisible by 4.

4. A table of invariants

In the following, we list the absolutely bigraded invariants $\varkappa$ for all strong inversions whose underlying knots have at most 9 crossings. We do this as follows: We subdivide the drawing plane into a grid and associate with each point on the plane a bigrading, namely the homological grading is equal to the x-coordinate and the

δ-grading is equal to the y-coordinate. Moving a point right and upwards increases both gradings. For each pair (K, h), we can write

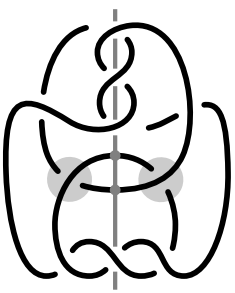
$$\varkappa(K, h) = \bigoplus_i h^{a_i} \delta^{b_i} V_\varkappa$$

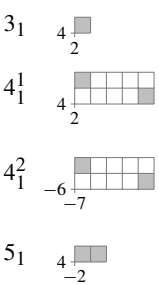
since all special components of the tangle invariants have length 2. Then, for each i, we place a grey box centred at the point (a_i, b_i). If multiple boxes line up at the same bigrading, we add a label which indicates the number of such boxes. Finally, the absolute bigrading is specified by the two numbers in the bottom left corner, which indicate the absolute coordinate of the bottom-and-left-most intersection point of the grid: The top-left number is the δ-grading, the bottom-right number is equal to the homological grading. For example,

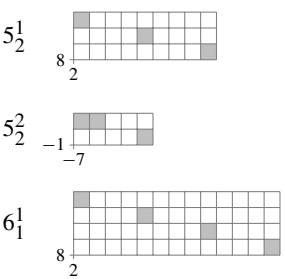
$$\varkappa(4_1^1) = h^{2.5}\delta^{5.5} V_\varkappa \oplus h^{6.5}\delta^{4.5} V_\varkappa.$$

The superscripts $\spadesuit = 1, 2$ indicate the strong inversions, using the same numbering convention as in [8]. We follow the grading conventions in [5]; when comparing these computations to [10], note that our δ-grading has the opposite sign, and our quantum grading is twice as large.

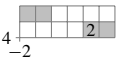
Finally, we note that there is a mistake in the diagram for 9_{43} in Sakuma's tabulation [8], where the diagram given is an alternating knot of determinant 43 and Rasmussen's s-invariant ± 2. The only such knots with fewer than 10 crossings sharing these properties are 9_{21} and 9_{22}, according to knotinfo [7]; by comparing the $\varkappa$-invariant of the quotients, we see that Sakuma's diagram shows 9_{22}. A diagram for the knot 9_{43} is shown on the right. (Its determinant is 13 and $s = \pm 4$. The only other knot with fewer than 10 crossings with these properties is 7_3, but its strong inversions have different $\varkappa$ than the ones we compute for 9_{43}.) Changing the pair of shaded crossings indicated in the image on the right gives the diagram in Sakuma's table.

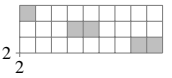


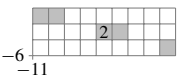


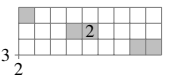


6_1^2 

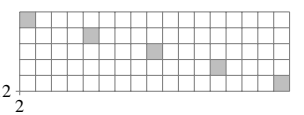
6_2^1 

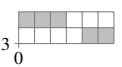
6_2^2 

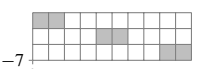
6_3^1 

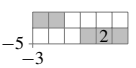
6_3^2 

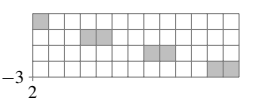
7_1 

7_2^1 

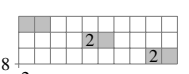
7_2^2 

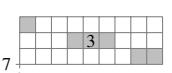
7_3^1 

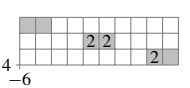
7_3^2 

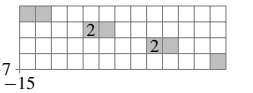
7_4^1 

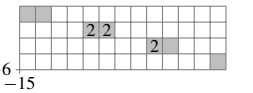
7_4^2 

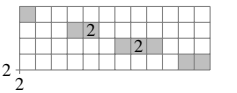
7_5^1 

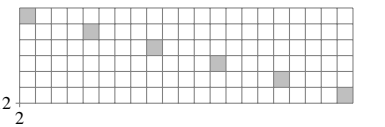
7_5^2 

7_6^1 

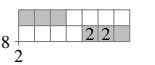
7_6^2 

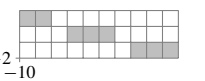
7_7^1 

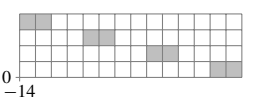
7_7^2 

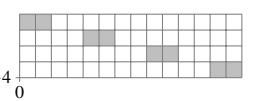
8_1^1 

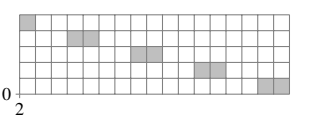
8_1^2 

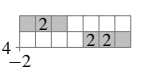
8_2^1 

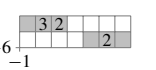
8_2^2 

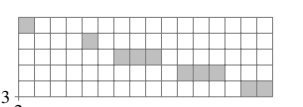
8_3^1 

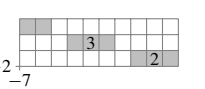
8_3^2 

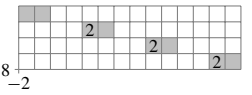
8_4^1 

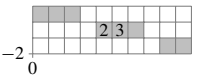
8_4^2 

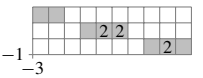
8_5^1 

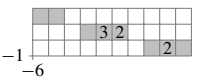
8_5^2 

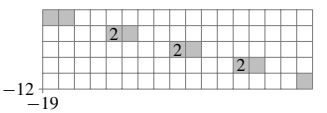
8_6^1 

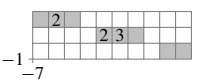
8_6^2 

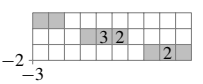
8_7^1 

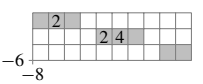
8_7^2 

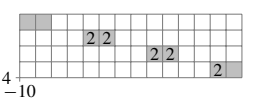
8_8^1 

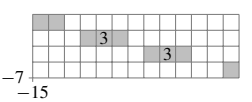
8_8^2 

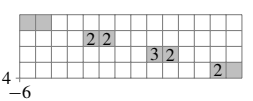
8_9^1 

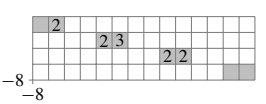
8_9^2 

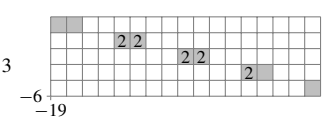
8_{10} 

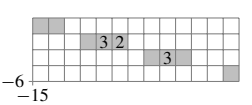
8_{11}^1 

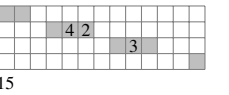
8_{11}^2 

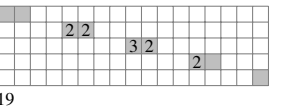
8_{12}^1 

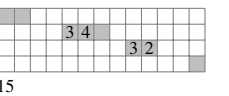
8_{12}^2 

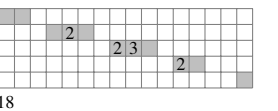
8_{13}^1 

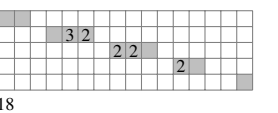
8_{13}^2 

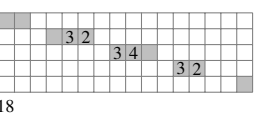
8_{14}^1 

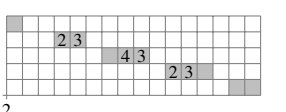
8_{14}^2 

8_{15}^1 

8_{15}^2 

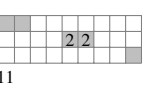
8_{16} 

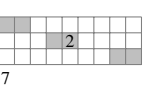
8_{18}^1 

8_{18}^2 

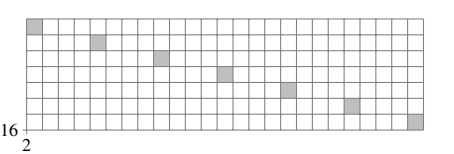
8_{19} 

8_{20} 

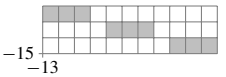
8_{21}^1 

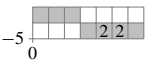
8_{21}^2 

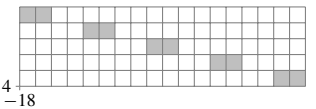
9_1 

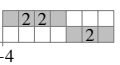
9_2^1 

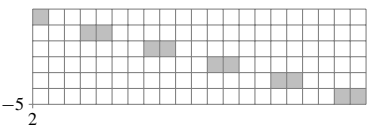
9_2^2


9_3^1


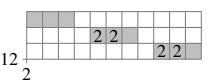
9_3^2


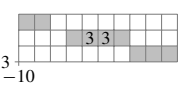
9_4^1


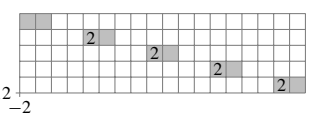
9_4^2


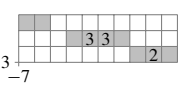
9_5^1


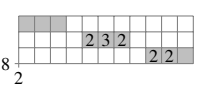
9_5^2

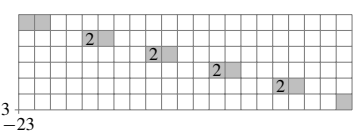

9_6^1


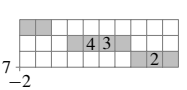
9_6^2


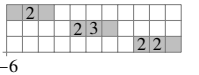
9_7^1


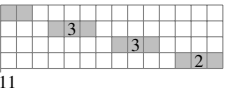
9_7^2


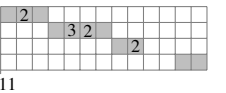
9_8^1


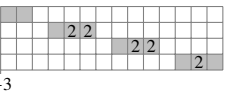
9_8^2


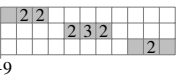
9_9^1


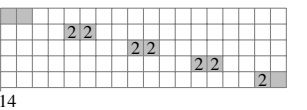
9_9^2


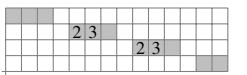
9_{10}^1


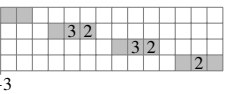
9_{10}^2


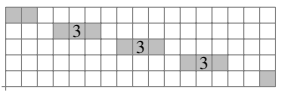
9_{11}^1


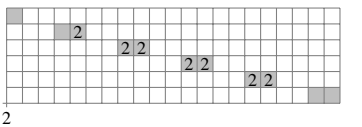
9_{11}^2


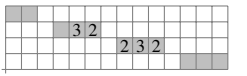
9_{12}^1


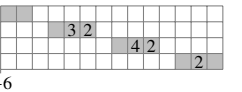
9_{12}^2


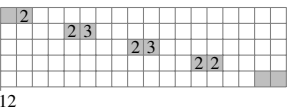
9_{13}^1


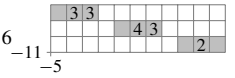
9_{13}^2


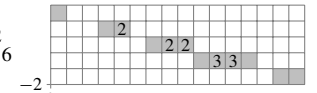
9_{14}^1


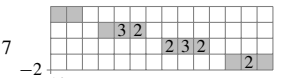
9_{14}^2


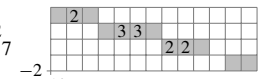
9_{15}^1


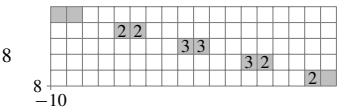
9_{15}^2


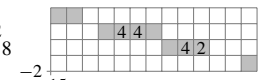
9^1_{16}


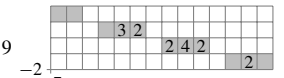
9^2_{16}


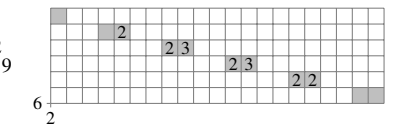
9^1_{17}


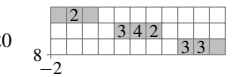
9^2_{17}


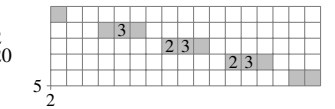
9^1_{18}


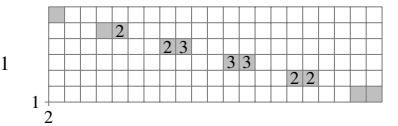
9^2_{18}


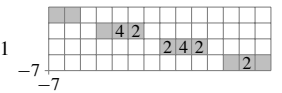
9^1_{19}


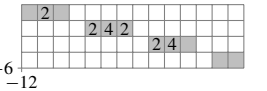
9^2_{19}


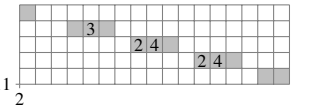
9^1_{20}


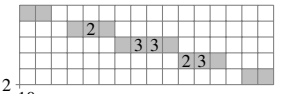
9^2_{20}


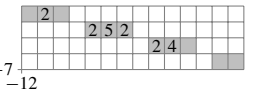
9^1_{21}


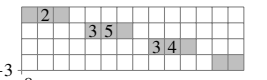
9^2_{21}


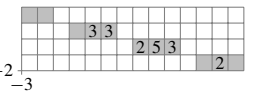
9_{22}


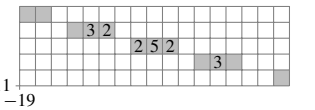
9^1_{23}


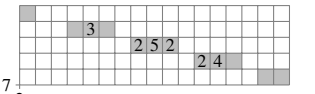
9^2_{23}


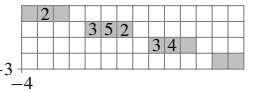
9_{24}


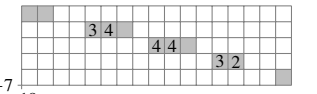
9_{25}


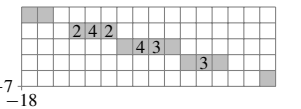
9^1_{26}


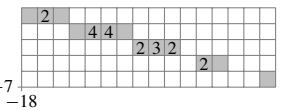
9^2_{26}


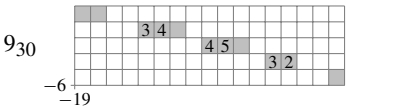
9^1_{27}


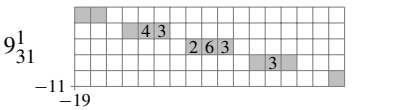
9^2_{27}


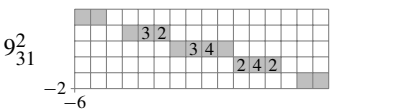
9^1_{28}


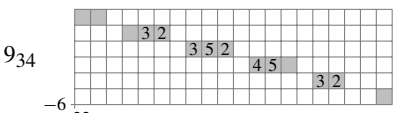
9^2_{28}


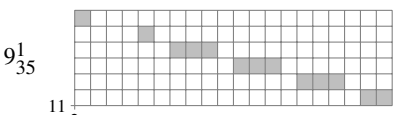
9_{29}


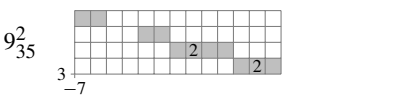

9_{30}
3 4
4 5
3 2
−6
−19

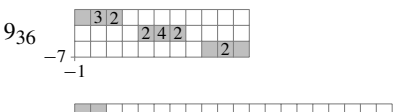

9^1_{31}
4 3
2 6 3
3
−11
−19

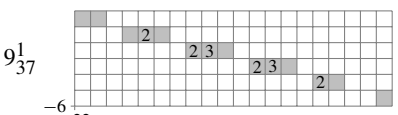

9^2_{31}
3 2
3 4
2 4 2
−2
−6

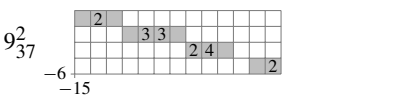

9_{34}
3 2
3 5 2
4 5
3 2
−6
−22

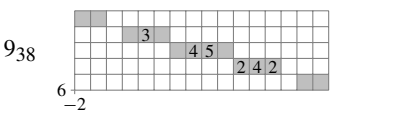

9^1_{35}
11
2

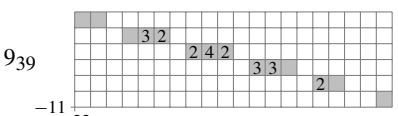

9^2_{35}
2
2
3
−7

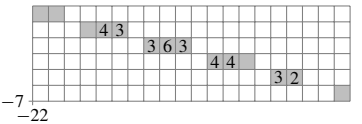

9_{36}
3 2
2 4 2
2
−7
−1

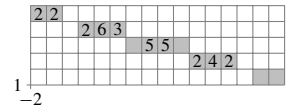

9^1_{37}
2
2 3
2 3
2
−6
−22

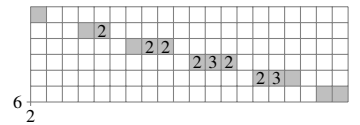

9^2_{37}
2
3 3
2 4
2
−6
−15


9_{38}
3
4 5
2 4 2
6
−2


9_{39}
3 2
2 4 2
3 3
2
−11
−22

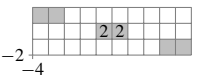

9^1_{40}
4 3
3 6 3
4 4
3 2
−7
−22

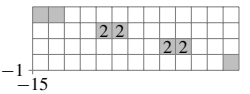

9^2_{40}
2 2
2 6 3
5 5
2 4 2
1
−2

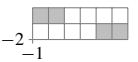

9_{41}
2
2 2
2 3 2
2 3
6
2

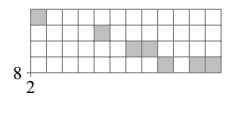

9_{42}
−2
3

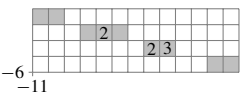

9_{43}
−5
−8

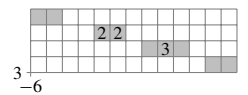

9_{44}
2 2
−2
−4

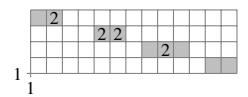

9_{45}
2 2
2 2
−1
−15

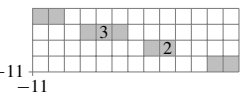

9^1_{46}
−2
−1


9^2_{46}
8
2


9_{47}
2
2 3
−6
−11


9^1_{48}
2 2
3
3
−6


9^2_{48}
2
2 2
2
1
1


9_{49}
3
2
−11
−11

Acknowledgements

The computer program [12] used in this work represents a major overhaul to an earlier piece of software, developed as part of a UBC undergraduate research project with Gurkeerat Chhina.

References

[1] D. Bar-Natan, "Khovanov's homology for tangles and cobordisms", *Geom. Topol.* **9** (2005), 1443–1499. MR Zbl

[2] K. Boyle and A. Issa, "Equivariant 4-genera of strongly invertible and periodic knots", 2021. arXiv 2101.05413

[3] C. M. Gordon and J. Luecke, "Knots are determined by their complements", *J. Amer. Math. Soc.* **2**:2 (1989), 371–415. MR Zbl

[4] R. Hartley, "Knots and involutions", *Math. Z.* **171**:2 (1980), 175–185. MR Zbl

[5] A. Kotelskiy, L. Watson, and C. Zibrowius, "Immersed curves in Khovanov homology", 2019. arXiv 1910.14584

[6] A. Kotelskiy, L. Watson, and C. Zibrowius, "Khovanov multicurves are linear", 2022. arXiv 2202.01460v1

[7] C. Livingston and A. H. Moore, "KnotInfo: table of knot invariants", 2021, available at https://knotinfo.math.indiana.edu.

[8] M. Sakuma, "On strongly invertible knots", pp. 176–196 in *Algebraic and topological theories* (Kinosaki, 1984), Kinokuniya, Tokyo, 1986. MR Zbl

[9] L. Watson, "A remark on Khovanov homology and two-fold branched covers", *Pacific J. Math.* **245**:2 (2010), 373–380. MR Zbl

[10] L. Watson, "Khovanov homology and the symmetry group of a knot", *Adv. Math.* **313** (2017), 915–946. MR Zbl

[11] C. Zibrowius, "On symmetries of peculiar modules; or, δ-graded link Floer homology is mutation invariant", 2019. To appear in *J. Eur. Math. Soc.* arXiv 1909.04267

[12] C. Zibrowius, "`kht++`, a program for computing Khovanov invariants for links and tangles", April 2021, available at https://cbz20.raspberryip.com/code/khtpp/docs/.

Received 3 Mar 2021. Revised 18 Aug 2021.

ARTEM KOTELSKIY: `artofkot@gmail.com`
Mathematics Department, Stony Brook University, Stony Brook, NY, United States

LIAM WATSON: `liam@math.ubc.ca`
Department of Mathematics, University of British Columbia, Vancouver, BC, Canada

CLAUDIUS ZIBROWIUS: `claudius.zibrowius@posteo.net`
Faculty of Mathematics, University of Regensburg, Regensburg, Germany

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.245

Lecture notes on trisections and cohomology

Peter Lambert-Cole

We describe several geometric interpretations of $H_2(X)$ when X is a trisected 4-manifold. The main insight is that, by analogy with Hodge theory and sheaf cohomology in algebraic geometry, classes in $H_2(X)$ can be usefully interpreted as "(1,1)"-classes. First, we reinterpret work of Feller, Klug, Schirmer and Zemke and of Florens and Moussard on the (co)homology of trisected 4-manifolds in terms of the Čech cohomology of presheaves on X, in both the case of singular and de Rham cohomology. We then discuss complex line bundles, almost-complex structures, spin structures and $\mathrm{Spin}^{\mathbb{C}}$-structures on trisected 4-manifolds.

1. Introduction

A motivating question in 4-manifold topology is the following:

Question 1.1. To what extent are general 4-manifolds similar to projective complex surfaces?

Donaldson showed that, like projective surfaces, every closed symplectic manifold admits a Lefschetz pencil [4]. Later, Auroux, Donaldson and Katzarkov showed that near-symplectic manifolds admit so-called broken Lefschetz pencils[1] [1]. Baykur then proved that every closed, oriented, smooth 4-manifold admits a broken Lefschetz fibration over S^2 [2]. This gives one sense in which all such 4-manifolds are similar to projective surfaces.

It is a classical fact, known as Theorem B, that over a Stein domain, coherent sheaves have no higher cohomology. That is, if Z is Stein and $\mathcal{F}$ is a coherent sheaf, then $H^i(Z;\mathcal{F}) = 0$ for $i > 0$. A consequence is that if X is a complex manifold, $\mathcal{F}$ is a coherent sheaf, and $\mathcal{Z} = \{Z_i\}$ is an open cover of X by Stein domains, then

MSC2020: 57K40.
Keywords: 4-manifolds, trisections.

[1]The term "singular Lefschetz pencil" was used in [1].

the sheaf cohomology of $\mathcal{F}$ can be computed by the Čech complex with respect to the open cover $\mathcal{Z}$:

$$H^*(X;\mathcal{F}) \cong \check{H}^*(\mathcal{Z};\mathcal{F}).$$

On a projective surface, Hodge theory implies that Dolbeault cohomology refines de Rham cohomology. Specifically, there is an isomorphism

$$H^k(X;\mathbb{C}) \cong \bigoplus_{i+j=k} H^{i,j}_{\bar{\partial}}(X;\mathbb{C}).$$

In addition, Dolbeault's theorem states that Dolbeault cohomology is isomorphic to the cohomology of the sheaf of holomorphic differential forms:

$$H^{i,j}_{\bar{\partial}}(X;\mathbb{C}) \cong H^i(X;\Omega^j).$$

Moreover, applying Serre duality to the constant sheaf $\underline{\mathbb{C}}$ shows that there is an isomorphism

$$H^{i,j}_{\bar{\partial}}(X;\mathbb{C}) \cong H^{n-i,n-j}_{\bar{\partial}}(X;\mathbb{C}),$$

where n is the complex dimension of X.

Interestingly, trisections of 4-manifolds reveal similar results for singular and de Rham cohomology. The four-dimensional handlebody $\natural_k S^1 \times B^3$ admits a Stein structure. Thus, since every closed 4-manifold admits a trisection, it can be covered by three domains that admit Stein structures. In addition, by slightly enlarging the sectors of trisection, we get an open cover $\mathcal{T} = \{U_1, U_2, U_3\}$, where

(1) U_i is diffeomorphic to $\natural_{k_i} S^1 \times B^3$,

(2) $U_i \cap U_j$ is diffeomorphic to $\natural_g S^1 \times B^3$, and

(3) $U_1 \cap U_2 \cap U_3$ is diffeomorphic to $\Sigma_g \times D^2$.

Let $\mathcal{C}^i$ denote the presheaf on X defined as

$$\mathcal{C}^i(U) := H^i(U;\mathbb{Z}).$$

It is clear that $\mathcal{C}^i$ is a presheaf. However, in general it is not a sheaf as it satisfies the gluing axiom but not the locality axiom. In particular, it is not separated. Nonetheless, we can compute the Čech cohomology $\check{H}^*(\mathcal{T},\mathcal{C}^i)$ of the presheaf $\mathcal{C}^i$ with respect to the open cover $\mathcal{T}$.

Methods to compute the homology of 4-manifolds from a trisection have been given by Feller, Klug, Schirmer and Zemke [5] and by Florens and Moussard [6]. Reinterpreting their results, we get the following theorems:

Theorem 1.2 (Hodge/Dolbeault theorem). *There is an isomorphism*

$$H^k(X;\mathbb{Z}) \cong \bigoplus_{i+j=k} \check{H}^i(\mathcal{T},\mathcal{C}^j).$$

Moreover, we have the following "Hodge diamond" for the cohomology of a trisected 4*-manifold*:

$$\begin{array}{ccccc} & & H^4(X;\mathbb{Z}) & & \\ & 0 & & H^3(X;\mathbb{Z}) & \\ 0 & & H^2(X;\mathbb{Z}) & & 0 \\ & H^1(X;\mathbb{Z}) & & 0 & \\ & & H^0(X;\mathbb{Z}) & & \end{array}$$

In particular, the Čech complex $\check{C}^*(\mathcal{T}, \mathcal{C}^1)$ — representing the middle diagonal of the Hodge diamond — is essentially given in [6, Section 2.1] but not described as such.

We can also interpret the symmetry of the Hodge diamond as Serre duality.

Theorem 1.3 (Serre duality). *There is an isomorphism*

$$\check{H}^i(\mathcal{T}, \mathcal{C}^j) \otimes \mathbb{R} \cong \check{H}^{2-i}(\mathcal{T}, \mathcal{C}^{2-j}) \otimes \mathbb{R}.$$

1A. ***Second cohomology as* (1, 1)*-classes.*** By analogy with complex geometry, we refer to any class in $\check{H}^1(\mathcal{T}, \mathcal{C}^1) \cong H^2(X;\mathbb{Z})$ as a $(1,1)$*-class*. On a projective surface, the Lefschetz theorem states that the integral (1,1)-classes are precisely those that can be represented by a divisor. The proof of Theorem 1.2 further implies that every class of is a (1,1)-class.

Theorem 1.4. *Every class in* $H^2(X;\mathbb{Z})$ *is a* (1,1)*-class with respect to the trisection* $\mathcal{T}$*. Specifically*,

$$H^2(X;\mathbb{Z}) \cong \check{H}^1(\mathcal{T}, \mathcal{C}^1).$$

Unpacking the definition of Čech cohomology, this means that every element of $H^2(X)$ is represented by a triple $(\beta_1, \beta_2, \beta_3)$ where β_λ is a 1-dimensional cohomology class on the handlebody H_λ of the trisection. We will describe several geometric interpretations of this.

(1) **De Rham cohomology**: Every class $\omega \in H^2_{DR}(X)$ can be represented by a triple $(\beta_1, \beta_2, \beta_3)$ where β_λ is a closed 1-form on H_λ.

(2) $\mathbb{C}$**-line bundles**. Recall that isomorphism classes of $\mathbb{C}$-line bundles over X are classified by $H^2(X;\mathbb{Z})$ and homotopy classes of maps from H_λ to S^1 are classified by $H^1(H_\lambda;\mathbb{Z})$. Take a $\mathbb{C}$-line bundle E with first Chern class $c_1(E)$. Then E can be trivialized over each sector Z_λ of the trisection and the triple $(\beta_1, \beta_2, \beta_3)$ corresponding to $c_1(E)$ determines the transition maps (up to homotopy).

(3) **Spin$^{\mathbb{C}}$-structures**. The set of Spin$^{\mathbb{C}}$-structures on X is an affine copy of $H^2(X;\mathbb{Z})$. Following Gompf, we show how to interpret a Spin$^{\mathbb{C}}$-structure as an almost-complex structure on the spine of the trisection. Then, the action of $H^2(X;\mathbb{Z})$ can be described in terms of "Lutz twists" along a collection of curves representing homology classes in $H_1(H_\lambda)$ that are hom-dual to $(\beta_1, \beta_2, \beta_3)$.

2. Singular cohomology

2A. ***Sheaves.*** We first review the basics of sheaves and Čech cohomology. Let X be a topological space and let R be a commutative ring.

Definition 2.1. A *presheaf of R-modules* $\mathcal{F}$ on X consists of

(1) an R-module $\mathcal{F}(U)$ for each open set U, and

(2) a restriction map $\rho_{U,V} : \mathcal{F}(U) \to \mathcal{F}(V)$ if V is contained in U.

Furthermore, the restriction maps satisfy the relations

(1) $\rho_{U,U} : \mathcal{F}(U) \to \mathcal{F}(U)$ is the identity homomorphism, and

(2) $\rho_{U,W} = \rho_{V,W} \circ \rho_{U,V}$ if $W \subset V \subset U$.

If $s \in \mathcal{F}(U)$ and $V \subset U$, then we will denote $\rho_{U,V}(s)$ by $s|_V$.

Exercise 2.2. Suppose that X is a smooth manifold. Check that the following are presheaves:

(1) $\mathcal{R}$ is the constant presheaf, where $\mathcal{R}(U) = R$ and the restriction map is the identity.

(2) $\mathcal{C}^i$ is a presheaf of $\mathbb{Z}$-modules, where $\mathcal{C}^i(U) = H^i(U;\mathbb{Z})$ and the restriction maps are given by the inclusion map.

(3) $\mathcal{DR}^i$ is a presheaf of $\mathbb{R}$-modules, where $\mathcal{DR}^i(U) = H^i_{DR}(U, \mathbb{R})$ and the restriction maps are given by the inclusion map.

(4) Ω^p is a sheaf of $\mathbb{R}$-modules, where $\Omega^p(U)$ is the set of p-forms on U and the restriction maps are given by the inclusion map.

Definition 2.3. A *sheaf of R-modules* is a presheaf of R-modules that satisfy the further conditions:

(1) (*locality*) If $\{U_i\}_{i\in I}$ is an open covering of U and if $s, t \in \mathcal{F}(U)$ satisfy $s|_{U_i} = t|_{U_i}$ for all $i \in I$, then $s = t$.

(2) (*gluing*) Let $\{U_i\}_{i\in I}$ be an open covering of U and let $\{s_i \in \mathcal{F}(U_i)\}_{i\in I}$ be a collection of local sections such that

$$s_i|_{U_i\cap U_j} = s_j|_{U_i\cap U_j}$$

for all $i, j \in I$. Then there is a section $s \in \mathcal{F}(U)$ such that $s|_{U_i} = s_i$ for all $i \in I$.

Exercise 2.4. Show that $\mathcal{R}$ and Ω^p are sheaves, but $\mathcal{C}^i$ and $\mathcal{DR}^i$ are not sheaves in general.

2B. ***Čech cohomology.*** We refer the reader to [3] for a discussion of Čech cohomology in general. To simplify the exposition, we restrict to open covers consisting of at most three open sets.

Let X be a smooth manifold, let $\mathcal{F}$ be a presheaf of R-modules, and let $\mathcal{U} = \{U_1, U_2, U_3\}$ be an open cover of X. The *Čech cochain groups* are defined to be

$$\begin{aligned}
\check{C}^0(\mathcal{U}, \mathcal{F}) &= \mathcal{F}(U_1) \oplus \mathcal{F}(U_2) \oplus \mathcal{F}(U_3),\\
\check{C}^1(\mathcal{U}, \mathcal{F}) &= \mathcal{F}(U_1 \cap U_2) \oplus \mathcal{F}(U_1 \cap U_3) \oplus \mathcal{F}(U_2 \cap U_3),\\
\check{C}^2(\mathcal{U}, \mathcal{F}) &= \mathcal{F}(U_1 \cap U_2 \cap U_3).
\end{aligned}$$

For $1 \leq i < j \leq 3$, denote the restriction maps by

$$\begin{aligned}
\rho_{i,ij} &: \mathcal{F}(U_i) \to \mathcal{F}(U_i \cap U_j),\\
\rho_{ij,123} &: \mathcal{F}(U_i \cap U_j) \to \mathcal{F}(U_1 \cap U_2 \cap U_3).
\end{aligned}$$

The *Čech coboundary map* is defined to be

$$\begin{aligned}
\delta^{-1} &= 0,\\
\delta^0 &= (\rho_{1,12} - \rho_{2,12}) \oplus (\rho_{1,13} - \rho_{3,13}) \oplus (\rho_{2,23} - \rho_{3,23}),\\
\delta^1 &= \rho_{12,123} - \rho_{13,123} + \rho_{23,123},\\
\delta^2 &= 0.
\end{aligned}$$

The *Čech cohomology* $\check{H}^*(\mathcal{U}, \mathcal{F})$ of the presheaf $\mathcal{F}$ with respect to the open cover $\mathcal{U}$ is defined to be

$$\check{H}^i(\mathcal{U}, F) = \frac{\ker(\delta^i)}{\operatorname{Im}(\delta^{i-1})}.$$

Exercise 2.5. Find open covers $\mathcal{U} = \{U_1, U_2, U_3\}$ and compute the Čech cohomology of the sheaf $\mathcal{R}$ for the following topological spaces:

(1) S^1.

(2) $S^1 \vee S^1$.

(3) S^2.

2C. ***Notational setup.*** Let $X = Z_1 \cup Z_2 \cup Z_3$ be a trisection of X, let $Y_\lambda = \partial Z_\lambda$ and let $H_\lambda = \mathbb{Z}_{\lambda-1} \cap Z_\lambda$. Let Σ be the central surface. The inclusion

$$\iota_\lambda : \Sigma \to H_\lambda$$

induces two maps

$$(\iota_\lambda)_* : H_1(\Sigma) \to H_1(H_\lambda), \qquad (\iota_\lambda)^* : H^1(H_\lambda) \to H^1(\Sigma).$$

Define subspaces

$$L_\lambda := \ker((\iota_\lambda)_*) \subset H_1(\Sigma), \qquad M_\lambda := \mathrm{Im}((\iota_\lambda)^*) \subset H^1(\Sigma).$$

We can use the intersection pairing $\langle -, - \rangle_\Sigma$ on $H_1(\Sigma)$ to define an isomorphism $\pi : H_1(\Sigma) \to H^1(\Sigma)$ by setting

$$\pi(x) = \langle -, x \rangle_\Sigma.$$

Furthermore, we have inclusion maps $\kappa_{i,j} : H_j \hookrightarrow Y_i$ and $\rho_i : Y_i \to Z_i$ for $i = 1, 2, 3$ and $j = i - 1, i$. These induce maps

$$(\kappa_{i,j})_* : H_1(H_j) \to H_1(Y_i), \qquad (\rho_i)_* : H_1(Y_i) \to H_1(Z_i),$$
$$(\kappa_{i,j})^* : H^1(Y_i) \to H^1(H_j), \qquad (\rho_i)^* : H^1(Z_i) \to H^1(Y_i).$$

2D. ***Hodge diamond.*** The results in [5; 6] compute homology. In particular, we have the following expression for $H_*(X)$.

Theorem 2.6 [6]. *The homology of X with $\mathbb{Z}$-coefficients is the homology of the complex*

$$0 \to \mathbb{Z} \xrightarrow{0} (L_1 \cap L_2) \oplus (L_2 \cap L_3) \oplus (L_3 \cap L_1) \xrightarrow{\zeta} L_1 \oplus L_2 \oplus L_3 \xrightarrow{\iota} H_1(\Sigma) \xrightarrow{0} \mathbb{Z} \to 0,$$

where $\zeta(a, b, c) = (c - a, a - b, b - c)$ and $\iota(a, b, c) = a + b + c$.

The middle terms of this complex are essentially the Čech complex.

Proposition 2.7. *There is a chain complex isomorphism*

$$\begin{array}{ccccccccc}
0 & \longrightarrow & (L_1 \cap L_2) \oplus (L_2 \cap L_3) \oplus (L_3 \cap L_\alpha) & \xrightarrow{\zeta} & L_\alpha \oplus L_2 \oplus L_3 & \xrightarrow{\iota} & H_1(\Sigma) & \longrightarrow & 0 \\
\downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle \phi_1} & & \downarrow{\scriptstyle \phi_2} & & \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle 0} \\
0 & \longrightarrow & \bigoplus_\lambda H^1(Z_\lambda) & \xrightarrow{\delta_1} & \bigoplus_\lambda H^1(H_\lambda) & \xrightarrow{\delta_2} & H^1(\Sigma) & \longrightarrow & 0
\end{array}$$

The second complex of this proposition is exactly the Čech complex of $\mathcal{C}^1$ with respect to $\mathcal{T}$, thus by applying Poincaré duality we obtain the following corollary.

Corollary 2.8. *For $i = 1, 2, 3$, there are isomorphisms*

$$H_{4-i}(X; \mathbb{Z}) \cong H^i(X; \mathbb{Z}) \cong \check{H}^{i-1}(\mathcal{T}, \mathcal{C}^1).$$

Proof of Proposition 2.7. By definition, $Z_\lambda = \natural_{k_\lambda} S^1 \times B^3$ and $Y_\lambda = \partial Z_\lambda = \#_{k_\lambda} S^1 \times S^2$. In particular,

$$H_1(Z_\lambda) \cong H_1(Y_\lambda) \cong \mathbb{Z}^{k_\lambda}.$$

We can apply the Mayer–Vietoris sequence to the Heegaard splitting $Y_\lambda = H_\lambda \cup H_{\lambda+1}$ to get the sequence

$$\to H_2(H_\lambda) \oplus H_2(H_{\lambda+1}) \to H_2(Y_\lambda) \to H_1(\Sigma) \\ \to H_1(H_\lambda) \oplus H_1(H_{\lambda+1}) \to H_1(Y_\lambda) \to H_0(\Sigma).$$

Since $H_2(H_\lambda) = H_2(H_{\lambda+1}) = 0$, we see that

$$H^1(Y_\lambda) \cong H_2(Y_\lambda) \cong \ker(H_1(\Sigma) \to H_1(H_\lambda) \oplus H_1(H_{\lambda+1})) \cong L_\lambda \cap L_{\lambda+1},$$

where the first isomorphism follows by Poincaré duality. This defines ϕ_1.

Using the long exact sequence of the pair (H_λ, Σ) we obtain

$$H_2(H_\lambda) \to H_2(H_\lambda, \Sigma) \to H_1(\Sigma) \to H_1(H_\lambda) \to .$$

Since $H_2(H_\lambda) = 0$, we see that

$$H^1(H_\lambda) \cong H_2(H_\lambda, \Sigma) \cong \ker(H_1(\Sigma) \to H_1(H_\lambda)) = L_\lambda.$$

This defines ϕ_2. □

The remaining cohomology groups are straightforward to calculate.

Proposition 2.9. *The cohomology groups of $\mathcal{C}^0$ are*

$$\check{H}^0(\mathcal{T}, \mathcal{C}^0) \cong \mathbb{Z}, \qquad \check{H}^1(\mathcal{T}, \mathcal{C}^0) \cong 0, \qquad \check{H}^2(\mathcal{T}, \mathcal{C}^0) \cong 0.$$

Proof. Each open set U_i and each double and triple intersection is connected and so

$$H^0(U_i; \mathbb{Z}) \cong H^0(U_i \cap U_j; \mathbb{Z}) \cong H^0(U_1 \cap U_2 \cap U_3) \cong \mathbb{Z}.$$

The Čech complex is therefore

$$0 \to \mathbb{Z}^3 \to \mathbb{Z}^3 \to \mathbb{Z} \to 0.$$

If $\{a, b, c\}$ is a chain in $\check{C}^0(\mathcal{T}, \mathcal{C}^0)$ then

$$\delta^0\{a, b, c\} = \{a - b, b - c, c - a\}.$$

Thus, this chain is coclosed if and only if $a = b = c$. Thus, $\check{H}^0(\mathcal{T}, \mathcal{C}^0) \cong \mathbb{Z}\langle\{a, a, a\}\rangle \cong \mathbb{Z}$. If $\{a, b, c\}$ is a chain in $\check{C}^1(\mathcal{T}, \mathcal{C}^0)$, then

$$\delta^1\{a, b, c\} = \{a + b + c\}.$$

The chain is coclosed if and only if it has the form

$$\{a, b, -a - b\} = a\{1, 0, -1\} + b\{0, 1, -1\}.$$

Both elements $\{1, 0, -1\}$ and $\{0, 1, -1\}$ are in the image of δ^0, so $\check{H}^1(\mathcal{T}, \mathcal{C}^0) \cong 0$. Finally, the differential δ^1 is surjective so $\check{H}^1(\mathcal{T}, \mathcal{C}^0) \cong 0$ as well. □

Proposition 2.10. *The cohomology groups of $\mathcal{C}^2$ are*

$$\check{H}^0(\mathcal{T}, \mathcal{C}^2) \cong 0, \qquad \check{H}^1(\mathcal{T}, \mathcal{C}^2) \cong 0, \qquad \check{H}^2(\mathcal{T}, \mathcal{C}^2) \cong \mathbb{Z}.$$

Exercise 2.11. Prove the proposition. [Hint: what is the rank of $\check{C}^i(\mathcal{T}, \mathcal{C}^2)$?]

3. De Rham

Let $\mathcal{DR}^i$ denote the presheaf on X defined as

$$\mathcal{DR}^i(U) := H^i_{DR}(U; \mathbb{R})$$

3A. *De Rham to Čech isomorphism.*

Theorem 3.1. *There are isomorphisms*

$$H^1_{DR}(X; \mathbb{R}) \cong \check{H}^0(\mathcal{T}, \mathcal{DR}^1), \qquad H^0_{DR}(X; \mathbb{R}) \cong \check{H}^0(\mathcal{T}, \mathcal{DR}^0),$$
$$H^2_{DR}(X; \mathbb{R}) \cong \check{H}^1(\mathcal{T}, \mathcal{DR}^1), \qquad H^4_{DR}(X; \mathbb{R}) \cong \check{H}^2(\mathcal{T}, \mathcal{DR}^2),$$
$$H^3_{DR}(X; \mathbb{R}) \cong \check{H}^2(\mathcal{T}, \mathcal{DR}^1).$$

We break up the proof by the degree of the cohomology group:

Degree 0: The cohomology group $H^0_{DR}(X; \mathbb{R})$ consists of constant functions. Given a constant function $C : X \to \mathbb{R}$, its restriction to U_λ is also a constant function $C : U_\lambda \to \mathbb{R}$ and therefore an element of $H^0_{DR}(U_\lambda; \mathbb{R})$. The isomorphism from de Rham to Čech is given by $C \mapsto (C, C, C)$.

Conversely, an element of $\check{H}^0(\mathcal{T}, \mathcal{DR}^0)$ is a triple (C_1, C_2, C_3) of constant functions whose restrictions to the pairwise intersections agree. In other words, $C_1 = C_2 = C_3 = C$. The inverse isomorphism is therefore $(C, C, C) \mapsto C$.

Degree 1: The map from de Rham to Čech is identical to the degree 0 case above. Given some closed 1-form β, the corresponding element in Čech cohomology is given by restricting β to each U_λ.

The inverse isomorphism is more complicated. In particular, an element of $\check{H}^0(\mathcal{T}, \mathcal{DR}^1)$ is a triple $([\beta_1], [\beta_2], [\beta_3])$ of cohomology classes, not specific closed forms. Choose representative closed 1-forms $\beta_1, \beta_2, \beta_3$. By assumption, the restrictions satisfy

$$[\beta_{\lambda-1}] = [\beta_\lambda] \in H^1_{DR}(U_{\lambda-1} \cap U_\lambda; \mathbb{R}).$$

Therefore, $\beta_\lambda - \beta_{\lambda-1} = dg_\lambda$ for some function $g_\lambda : U_{\lambda-1} \cap U_\lambda \to \mathbb{R}$.

Exercise 3.2. Show that there exist functions $f_\lambda : U_\lambda \to \mathbb{R}$ such that, on $U_{\lambda-1} \cap U_\lambda$,

$$\beta_{\lambda-1} + df_{\lambda-1} = \beta_\lambda + df_\lambda.$$

Consequently, we can represent our original Čech class by the triple

$$(\beta_1 + df_1, \beta_2 + df_2, \beta_3 + df_3)$$

and these 1-forms glue into a global 1-form β.

Degree 2: In this case, the maps in both directions are more complicated and we need to check that they are in fact isomorphisms. First, choose a class $[\omega] \in H^2_{DR}(X; \mathbb{R})$ and represent it by a closed 2-form ω. The restriction $\omega|_{U_\lambda}$ is exact since $H^2_{DR}(U_\lambda; \mathbb{R}) = 0$, thus we can choose a primitive α_λ for $\omega|_{U_\lambda}$. Over the double intersection $U_{\lambda-1} \cap U_\lambda$, the restrictions $\alpha_{\lambda-1}$ and α_λ are both primitives for ω, therefore their difference $\alpha_\lambda - \alpha_{\lambda-1}$ is closed. Consequently, the map from de Rham to Čech is given by

$$\omega \mapsto (\alpha_1 - \alpha_3, \alpha_2 - \alpha_1, \alpha_3 - \alpha_2).$$

There were three sources of indeterminancy:

(1) We could replace α_λ by $\alpha_\lambda + df_\lambda$ for some function $f_\lambda : U_\lambda \to \mathbb{R}$.

(2) We could replace ω by $\omega + d\mu$ for some global 1-form μ.

(3) We could replace the primitive α_λ with $\alpha_\lambda + \rho_\lambda$, where ρ is a closed 1-form on U_λ.

Exercise 3.3. (1) Show that modifying the primitives $\{\alpha_\lambda\}$ by exact 1-forms results in the same Čech cochain.

(2) Show that we can choose primitives for $\omega + d\mu$ that result in the same Čech cochain.

(3) Show that modifying the primitives $\{\alpha_\lambda\}$ by closed 1-forms $\{\rho_\lambda\}$ changes the Čech cochain by a Čech coboundary.

Conversely, given a class in $\check{H}^0(\mathcal{T}, \mathcal{DR}^1)$, choose a fixed cochain $([\beta_1], [\beta_2], [\beta_3])$ and fixed closed 1-forms $\{\beta_1, \beta_2, \beta_3\}$ to represent this class.

Exercise 3.4. (1) There exists a triple of 1-forms $\{\alpha_\lambda\}$ on the open sets $\{U_\lambda\}$ such that $\alpha_\lambda - \alpha_{\lambda-1} = \beta_\lambda$.

(2) The 2-forms $\{d\alpha_1, d\alpha_2, d\alpha_3\}$ glue together to give a global 2-form ω.

(3) Modifying the choices — modifying the Čech cochain by a coboundary, modifying the closed 1-forms $\{\beta_\lambda\}$ by exact 1-forms, and modifying the choices of $\{\alpha_\lambda\}$ — results in a cohomologous 2-form ω'.

Degree 3: Given a class $[\mu] \in H^3_{DR}(X; \mathbb{R})$, represent it by a closed 3-form μ. Since $H^3_{DR}(U_\lambda; \mathbb{R}) = 0$, we can choose a primitive ω_λ for μ over each U_λ. The differences $\omega_\lambda - \omega_{\lambda-1}$ are closed and represent elements of $H^3_{DR}(U_\lambda \cap U_{\lambda-1}; \mathbb{R}) = 0$. In particular, these forms are also exact and we can choose further primitive 1-forms $\{\beta_\lambda\}$.

Restricting to the triple intersection $U_1 \cap U_2 \cap U_3$ we get a 1-form $\beta = \beta_1 + \beta_2 + \beta_3$ that is closed since

$$d\beta = d\beta_1 + d\beta_2 + d\beta_3 = (\omega_1 - \omega_3) + (\omega_2 - \omega_1) + (\omega_3 - \omega_2) = 0.$$

Thus, $[\mu]$ is sent to an element $[\beta] \in H^1_{DR}(U_1 \cap U_2 \cap U_3; \mathbb{R})$ and therefore represents a Čech 2-cocycle.

Exercise 3.5. (1) Show that changing ω_λ by a closed 2-form results in the same Čech 2-cocycle

(2) Show that changing β_λ by a closed 1-form modifies the resulting Čech 2-cocycle by a Čech 2-coboundary.

The inverse map can be constructed by an argument similar to the degree 2 case; we leave it as an exercise.

Exercise 3.6. Construct the inverse map $\check{H}^2(\mathcal{T}, \mathcal{C}^1) \to H^3_{DR}(X; \mathbb{R})$ and show that it is well defined.

Degree 4: The isomorphism is constructed in a analogous method to the degree 3 case and we leave it as an exercise to the reader.

Exercise 3.7. Construct the isomorphism $H^4_{DR}(X; \mathbb{R}) \cong \check{H}^2(\mathcal{T}, \mathcal{C}^2)$.

3B. ***Intersection pairing.*** The intersection pairing on de Rham cohomology can also be expressed in terms of the Čech cohomology of the de Rham presheafs. In particular, we can describe the pairings

$$H^2_{DR}(X) \times H^2_{DR}(X) \to \mathbb{R},$$
$$H^3_{DR}(X) \times H^1_{DR}(X) \to \mathbb{R}.$$

Moreover, we can describe the pairing obtained by integrating a closed p-form over a closed p-dimensional submanifold:

$$H^2_{DR}(X) \times H_2(X; \mathbb{Z}) \to \mathbb{R},$$
$$H^3_{DR}(X) \times H_3(X; \mathbb{Z}) \to \mathbb{R},$$
$$H^4_{DR}(X) \times H_4(X; \mathbb{Z}) \to \mathbb{R}.$$

Theorem 3.8 (intersection pairing). *Let X be a trisected 4-manifold.*

(1) *Let ω_1, ω_2 be a pair of closed 2-forms. Suppose that under the de Rham–Čech isomorphism we have*

$$[\omega_1] \mapsto (\alpha_1, \alpha_2, \alpha_3), \qquad [\omega_2] \mapsto (\beta_1, \beta_2, \beta_3).$$

Then

$$\int_X \omega_1 \wedge \omega_2 = \int_\Sigma \alpha_1 \wedge \beta_2 = \int_\Sigma \alpha_2 \wedge \beta_3 = \int_\Sigma \alpha_3 \wedge \beta_1.$$

(2) *Let μ be a closed 3-form and α be a closed 1-form. Suppose that under the de Rham–Čech isomorphism, we have that $[\mu] \mapsto [\beta]$. Then*

$$\int_X \mu \wedge \alpha = \int_\Sigma \beta \wedge \alpha|_\Sigma.$$

Exercise 3.9. Prove these statements. [Hint: use Stokes' theorem combined with the arguments in the previous subsection.]

To describe the integration pairing, we first fix some notation.

(1) Let $\mathcal{K}$ be an embedded, oriented closed surface in general position with respect to the trisection. Let $\tau_\lambda^{\mathcal{K}}$ denote the tangle $\mathcal{K} \pitchfork H_\lambda$. We orient τ_λ as follows: since $\mathcal{K}$ is oriented, the intersection $F_\lambda = \mathcal{K} \cap Z_\lambda$ is oriented. The boundary ∂F_λ inherits an orientation from F_λ; the tangle $\tau_\lambda^{\mathcal{K}}$ is a subset of this boundary and inherits an orientation.

(2) Let $\mathcal{M}$ be an embedded, oriented, closed hypersurface in general position with respect to the trisection. In particular, the intersection $\mathcal{M} \pitchfork \Sigma$ is a simple closed curve $\gamma_{\mathcal{M}}$.

Theorem 3.10 (integration pairing). *Let X be a trisected 4-manifold.*

(1) *Let ω be a closed 2-form on X that maps to $(\beta_1, \beta_2, \beta_3)$ under the de Rham–Čech isomorphism and let $\mathcal{K}$ be an embedded, oriented closed surface. Then,*

$$\int_{\mathcal{K}} \omega = \sum_{\lambda=1,2,3} \int_{\tau_\lambda^{\mathcal{K}}} \beta_\lambda.$$

(2) *Let μ be a closed 3-form on X that maps to $\beta \in H^1_{DR}(\Sigma)$ under the de Rham–Čech isomorphism and let $\mathcal{M}$ be an embedded, oriented, closed hypersurface. Then,*

$$\int_{\mathcal{M}} \mu = \int_{\gamma_{\mathcal{M}}} \beta.$$

(3) *Let Ω be a closed 4-form on X that maps to $\omega \in H^2_{DR}(\Sigma)$ under the de Rham–Čech isomorphism. Then,*

$$\int_X \Omega = \int_\Sigma \omega.$$

Exercise 3.11. Prove these statements [Hint: again, use Stokes' theorem].

4. Complex line bundles

4A. ***Algebraic topology.*** First, we recall some facts from algebraic topology.

(1) The circle S^1 is a $K(\mathbb{Z}, 1)$. In particular, there is a one-to-one correspondence between classes in $H^1(X; \mathbb{Z})$ and homotopy classes of maps $f : X \to S^1$.

(2) The space $\mathbb{CP}^\infty$ is a $K(\mathbb{Z}, 2)$. In particular, there is a one-to-one correspondence between classes in $H^2(X; \mathbb{Z})$ and homotopy classes of maps $f : X \to \mathbb{CP}^\infty$. The cohomology ring of $\mathbb{CP}^\infty$ is $\mathbb{Z}[\alpha]$, where α has degree 2, and the identification between maps and cohomology classes is given by

$$f \leftrightarrow f^*(\alpha).$$

(3) The space $\mathbb{CP}^\infty$ is the classifying space for $U(1)$ (equivalently $\mathbb{C}$-line) bundles. In particular, there is a one-to-one correspondence between $\mathbb{C}$-line bundles on X, up to isomorphism, and homotopy classes of maps $f : X \to \mathbb{CP}^\infty$. There is a *tautological line bundle* $E \to \mathbb{CP}^\infty$ and the correspondence between maps and $\mathbb{C}$-bundles is given by

$$f \leftrightarrow f^*(E).$$

(4) The *first Chern class* is a complete invariant of $\mathbb{C}$-line bundles and connects (2) and (3) above. In particular, for the tautological bundle E on $\mathbb{CP}^\infty$ we have

$$c_1(E) = \alpha.$$

Moreover, since Chern classes are characteristic, they are natural with respect to pullbacks and therefore

$$c_1(f^*(E)) = f^*(c_1(E)) = f^*(\alpha).$$

4B. ***Chern classes of*** $\mathbb{C}$***-line bundles.*** Using a trisection $\mathcal{T}$ of X, we can explicitly see the equivalence

$$\{\mathbb{C}\text{-line bundles on X}\}/\{\text{equivalence}\} = \check{H}^1(\mathcal{T}, \mathcal{C}^1) \cong H^2(X; \mathbb{Z}),$$

where an element of $\check{H}^1(\mathcal{T}, \mathcal{C}^1)$ is a "(1, 1)-class".

Line bundles to (1, 1)*-classes.* Take a line bundle E on X. Since each sector Z_λ of a trisection is a 1-handlebody, we can choose a trivialization s_λ of E over Z_λ. Up to homotopy, the potential choices of trivializations are in one-to-one correspondence with elements of $H^1(Z_\lambda; \mathbb{Z}) \cong \mathbb{Z}^{k_\lambda}$. Over the double intersection H_λ, we have two trivializations $s_{\lambda-1}, s_\lambda$. Taking their quotient, we obtain a map

$$g_\lambda := \frac{s_\lambda}{s_{\lambda-1}} \to \mathbb{C}^*.$$

Composing this with the homotopy equivalence $\mathbb{C}^* \simeq S^1$, the map g_λ determines a homotopy class of maps from H_λ to S^1. In other words, the transition function g_λ determines a unique element β_λ of $H^1(H_\lambda; \mathbb{Z})$. Moreover, since

$$g_1 g_2 g_3 = \frac{s_1}{s_3}\frac{s_2}{s_1}\frac{s_3}{s_2} = 1,$$

the resulting triple $(\beta_1, \beta_2, \beta_3)$ is a Čech 1-cocycle in $\check{C}^*(\mathcal{T}, \mathcal{C}^1)$. Modifying the trivialization s_λ by some element of $H^1(Z_\lambda; \mathbb{Z})$ changes the resulting cocycle

by a Čech coboundary. In particular, we obtain a well-defined element $c_1(E) \in \check{H}^1(\mathcal{T}, \mathcal{C}^1)$.

(1, 1)-*classes to line bundles.* Given a (1,1)-class $(\beta_1, \beta_2, \beta_3) \in \check{H}^1(\mathcal{T}, \mathcal{C}^1)$, we can represent $\beta_\lambda \in H^1(H_\lambda; \mathbb{Z})$ by a map $g_\lambda : H_\lambda \to S^1$. Moreover, given the cocycle condition $\beta_1 + \beta_2 + \beta_3 = 0$ we can assume that $g_1 g_2 g_3 = 1$. In particular, the triple $\{g_1, g_2, g_3\}$ determines a triple of transition functions that allow us to construct a $\mathbb{C}$-bundle over X.

5. Almost-complex structures

An *almost-complex structure* J on X is a fiberwise homomorphism $J : TX \to TX$ such that $J^2 = -I$. This turns every fiber $T_x X$ into a complex vector space, where J is multiplication by i. Consequently, the almost-complex structure determines Chern classes $c_i(TX, J) \in H^{2i}(X; \mathbb{Z})$. The goal of this section is to describe almost-complex structures on the spine of a trisection.

5A. ***Field of complex tangencies.*** Let $Y^3 \subset X^4$ be a smooth hypersurface and let J be an almost-complex structure. The *field of J-complex tangencies* is defined to be

$$\xi := J(TY) \cap TY$$

Exercise 5.1. Show that ξ has rank 2 at every point. [Hint: ξ_x is a J-complex line in $T_x X$.] In particular, ξ is an *oriented* plane field.

Exercise 5.2. Let $\phi : X \to \mathbb{R}$ be a function such that $Y = \phi^{-1}(0)$. Show that the field of J-tangencies is the kernel of the 1-form $d^{\mathbb{C}}\phi = d\phi(J-)$, restricted to Y.

Proposition 5.3. *Let Y be a 3-manifold. Homotopy classes of almost-complex structures on $Y \times [0, 1]$ are in one-to-one correspondence with homotopy classes of (coorientable) 2-plane fields on Y.*

Proof. Let J be an almost-complex structure on $Y \times [0, 1]$ and let ξ_t denote the field of J-tangencies along $Y \times \{t\}$. It is immediately clear that $\{\xi_t\}$ is a homotopy of 2-plane fields. Furthermore, let J_s be a family of almost-complex structures and let $\xi_{s,t}$ denote the field of J_s-tangencies along $Y \times \{t\}$. Again, this clearly gives a 2-parameter homotopy of plane fields on Y.

Now let ξ be an oriented, coorientable 2-plane field and choose a fiberwise metric g on ξ. We can define an almost-complex structure $J : \xi \to \xi$ using the metric as follows. Locally, we can choose an oriented, orthonormal frame $\{e_1, e_2\}$ and define

$$J(e_1) = e_2 \quad \text{and} \quad J(e_2) = -e_1$$

and extend linearly.

Exercise 5.4. Show that, up to homotopy, this J does not depend on the metric g or the local orthonormal frame.

Next, let Λ be an oriented line field that coorients ξ. After choosing a metric h on Λ, we get a unit-length section σ of Λ and can extend J from ξ to TX by defining

$$J(\partial_t) = \sigma \quad \text{and} \quad J(\sigma) = -\partial_t.$$

Exercise 5.5. Show that, up to homotopy, this J does not depend on the homotopy class of $J|_\xi$, the homotopy class of Λ, or the metric h.

Finally, we have to check that every J on $Y \times [0, 1]$ can be constructed in this way. Choose some J and define $E = \langle \partial_t, J(\partial_t) \rangle$ and $\Lambda = TY \cap E$. Choose a nonvanishing section σ of Λ. Then,

$$J(\partial_t) = f\partial_t + g\sigma$$

for some functions f, g. By assumption, $\{\partial_t, J\partial_t\}$ is an oriented basis for E and therefore $g > 0$. Since J preserves ξ, we can define a family J_s of almost-complex structures for $s \in [0, 1]$ by defining

$$J_s|_\xi = J \quad \text{and} \quad J_s(\partial_t) = sf\partial_t + g\sigma.$$

After scaling the metric so that $|g\sigma| = 1$, we have that J_0 is an almost-complex structure of the form constructed above and J_1 is our original J. □

Exercise 5.6. Show that $\Sigma \times D^2$ admits an almost-complex structure J with $c_1(J) = 0$. [Hint: embed Σ in $\mathbb{C}^2$.]

Lemma 5.7. *The spine of a trisection admits an almost-complex structure J.*

Proof. By the previous exercise, we can choose some J on a tubular neighborhood of the central surface Σ. The remaining task is to extend it across each handlebody H_λ. The almost-complex structure J determines a hyperplane field ξ_λ in a neighborhood of $\partial H_\lambda = \Sigma$.

Exercise 5.8. Show $\langle e(\xi_\lambda), [\Sigma] \rangle = \langle c_1(J), [\Sigma] \rangle = 0$. [Hint: Choose a section σ of ξ_λ and a normal vector field ν to H_λ. Then $\det(\nu, \sigma) = 0$ precisely where $\sigma = 0$.]

Consequently, it is possible to extend ξ_λ across H_λ and by Proposition 5.3, this determines a homotopy class of J in a neighborhood of H_λ. □

5B. ***First Chern class of J.*** Given some J on the spine of a trisection, we can construct a 1-complex C_J in the spine that represents the Poincaré dual to $c_1(TX, J)$.

The central surface Σ is canonically framed. In particular, we can choose coordinates (s, t) on D^2 such that pulling back the coordinates by the projection

$$\pi : \nu(\Sigma) \cong \Sigma \times D^2 \to D^2$$

we have that

$$\Sigma = \pi^{-1}(0), \qquad H_2 = \pi^{-1}(0, t) \text{ for } t \geq 0,$$
$$H_1 = \pi^{-1}(s, 0) \text{ for } s \leq 0, H_3 = \pi^{-1}(-x, x) \text{ for } x \geq 0.$$

Consider the *conormal sequence* for the central surface Σ:

$$0 \to N^*\Sigma \to T^*X \to T^*\Sigma \to 0.$$

A *coframing* of Σ is a trivialization of its conormal bundle. Since $N^*\Sigma$ is an $\mathbb{R}^2$-bundle, a coframing is determined by a single, nowhere-vanishing section. Moreover, it is clear from the conormal sequence that such a section is given by a nowhere-vanishing 1-form whose restriction to Σ is identically 0. An almost-complex structure J determines a dual almost-complex structure $J^t : T^*X \to T^*X$. Inserting this, we get a (nonexact) sequence

$$N^*\Sigma \to T^*X \xrightarrow{J^t} T^*X \to T^*\Sigma.$$

Given a section α of $N^*\Sigma$, we can push it through this sequence to get a 1-form $\widetilde{\alpha}$ on Σ, defined to be

$$\widetilde{\alpha} = \alpha(J-)|_\Sigma.$$

Exercise 5.9. A *complex point* of Σ is a point $x \in \Sigma$ such that $J(T_x\Sigma) = T_x\Sigma$. Show that $\widetilde{\alpha}$ vanishes at precisely the complex points of Σ.

Exercise 5.10. By a C^∞-small perturbation of Σ, we can assume that Σ has finitely many complex points [Hint: what are the dimensions of the Grassmannians $\mathrm{Gr}_{\mathbb{R}}(2,4)$ and $\mathrm{Gr}_{\mathbb{C}}(1,2)$?]

Recall the normal coordinates (s,t) on $\Sigma \times D^2$. The pair ds, dt of 1-forms gives a coframing of Σ. Define

$$\beta_1 := \widetilde{ds}, \qquad \beta_2 := \widetilde{dt}, \qquad \beta_3 = -\widetilde{ds} - \widetilde{dt}.$$

Exercise 5.11. Show that $\beta_1 \wedge \beta_2 \neq 0$, except at the complex points of Σ. In particular, β_1 vanishes at $x \in \Sigma$ if and only if β_2 vanishes at x.

Exercise 5.12. Suppose that β_1, β_2, viewed as sections of $T^*\Sigma$, are transverse to the 0-section. Show that at each complex point $x \in \Sigma$, the indices of the vanishing of β_1 and β_2 at x agree.

Exercise 5.13. Show that β_λ extends to a 1-form on the handlebody H_λ of the trisection such that $\ker(\beta_\lambda)$ is the field of J-complex tangencies along H_λ.

Choose vector fields $\{v_1, v_2\}$ on Σ such that

$$\beta_1(v_1) = 0, \quad \beta_2(v_1) = \beta_1(v_2) \geq 0, \quad \beta_2(v_2) = 0$$

and set $v_3 = -v_1 - v_2 \in \ker(\beta_3)$. Since $v_\lambda \in \ker(\beta_\lambda)$, we can extend v_λ to a section of ξ_λ over H_λ.

For notational purposes, let ν_λ be a normal vector fields to H_λ such that near Σ, we have

$$u_1 = \partial_s, \quad u_2 = \partial_t, \quad u_3 = -\partial_s - \partial_t.$$

Exercise 5.14. Show that the pairs

$$\{u_1, v_1\}, \quad \{u_2, v_2\}, \quad \{u_3, v_3\}$$

determine the same section of $\det(TX, J)$ over Σ.

Proposition 5.15. *Let J be an almost-complex structure on the spine of a trisection $\mathcal{T}$ of X. Choose vector fields $\{v_\lambda \subset \xi_\lambda\}$ as above and let $\tau_\lambda = v_\lambda^{-1}(0)$. The 1-complex*

$$C_J = \tau_1 \cup \tau_2 \cup \tau_3$$

is the intersection of $PD(c_1(J))$ with the spine of the trisection $\mathcal{T}$.

Proof. The bivector $u_\lambda \wedge v_\lambda$ determines a section of the determinant line bundle over H_λ. The vector u_λ is everywhere normal to H_λ and nonvanishing, while v_λ is tangent and vanishes along τ_λ. By the previous exercise, we obtain a section of the determinant bundle on the entire spine that vanishes precisely along the 1-complex C_J. □

6. Spin$^{\mathbb{C}}$-structures

A standard interpretation of a spin structure on a manifold X is a trivialization of TX over the 1-skeleton that extends across the 2-skeleton. A similar interpretation of Spin$^{\mathbb{C}}$-structures, due to Gompf, is an almost-complex structure over the 2-skeleton that extends across the 3-skeleton.

6A. ***Handle decompositions.*** Every trisection $\mathcal{T}$ of X determines an *inside-out* handle decomposition as follows.

(1) Start with a neighborhood $\nu(\Sigma)$ of the central surface. This is diffeomorphic to $\Sigma \times D^2$ and can be built in the standard way using a 0-handle, $2g$ 1-handles, and a 2-handle. The boundary of this neighborhood is $\Sigma \times S^1$.

(2) Next, attach a neighborhood $\nu(H_\lambda)$ of each 3-dimensional piece of the trisection. The solid handlebody H_λ is built from a single 3-dimensional 0-handle and g 3-dimensional 1-handles. Upside-down, this becomes g 2-handles and a single 3-handle. Fix some distinct angular points $\theta_1, \theta_2, \theta_3 \in S^1$ in positive cyclic order. Then attaching $\nu(H_\lambda)$ is equivalent to the following. Attach g 2-handles along a cut system of curves on $\Sigma \times \{\theta_\lambda\}$ with surface framing. After this surgery, the surface $\Sigma \times \{\theta_\lambda\}$ is now an essential 2-sphere and the 4-dimensional 3-handle is attached along this 2-sphere. The resulting boundary of the 4-manifold has three components Y_1, Y_2, Y_3 with $Y_3 \cong \#_{k_i} S^1 \times S^2$.

(3) Last, attach the 4-dimensional sectors. These are 4-dimensional 1-handlebodies; upside-down they consist of k_i 3-handles and a single 4-handle. The 3-handles are attached along the essential spheres in $\#_{k_i} S^1 \times S^2$. The resulting boundary is three copies of S^3, which is where the 4-handles are attached.

The *outside-in* handle decomposition determined by $\mathcal{T}$ is the handle decomposition obtained by turning the inside-out handle decomposition upside-down.

6B. ***Spin structures.*** A standard interpretation of a spin structure on a manifold X is a trivialization of TX over the 1-skeleton that extends across the 2-skeleton. Now, consider the inside-out handle decomposition of X determined by a trisection $\mathcal{T}$. The 1-skeleton of X is contained in the 1-skeleton of $\nu(\Sigma)$. Thus, every spin structure of X restricts to a spin structure on $\nu(\Sigma)$; moreover, since spin structures are stable, every spin structure of X restricts to a spin structure on the central surface Σ.

Recall that there exist two spin structures on S^1 and exactly one extends across D^2. The spin structures on a closed, oriented surface Σ are classified by maps

$$q : H_1(\Sigma; \mathbb{Z}/2\mathbb{Z}) \to \mathbb{Z}/2,$$

where $q(\gamma) = 0$ if the spin structure, restricted to a curve representing γ, is the spin structure that extends across the disk. This map is a quadratic enhancement of the intersection form on $H_1(\Sigma)$; in particular, it satisfies the relation

$$q(x + y) = q(x) + q(y) + \langle x, y \rangle \mod 2. \tag{1}$$

Let $\boldsymbol{\alpha} = \{\alpha_i\}$ be a cut system of curves on Σ. We say that $q(\boldsymbol{\alpha}) = 0$ if $q(\alpha_i) = 0$ for every $\alpha_i \in \boldsymbol{\alpha}$. Note that by the relation in (1), if $q(\boldsymbol{\alpha}) = 0$, then for every cut system $\boldsymbol{\alpha}'$ obtained by handlesliding some curves in $\boldsymbol{\alpha}$, we also have $q(\boldsymbol{\alpha}') = 0$.

Proposition 6.1. *Let $\mathcal{T}$ be a trisection of X with trisection diagram $(\Sigma, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$. Then X admits a spin structure if and only if there exists a quadratic enhancement $q : H_1(\Sigma; \mathbb{Z}/2\mathbb{Z}) \to \mathbb{Z}/2\mathbb{Z}$ such that*

$$q(\boldsymbol{\alpha}) = q(\boldsymbol{\beta}) = q(\boldsymbol{\gamma}) = 0.$$

Moreover, the set of spin structures is in one-to-one correspondence with such quadratic enhancements.

Proof. Each q corresponds to a spin structure on Σ and therefore a trivialization of TX over its 1-skeleton. In the inside-out handle decomposition, we have $3g + 1$ 2-handles. One 2-handle corresponds to the 2-handle of Σ; by assumption the trivialization extends over this handle. The remaining 2-handles are attached along the curves of $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ with surface framing. Consequently, the trivialization of TX extends across such a handle if and only if the spin structure, restricted to the attaching circle, is the spin structure on S^1 that extends across the disk. □

6C. ***Lutz twists.*** A Lutz twist is a method for modifying a 2-plane field ξ along an embedded curve γ.

Fix a metric and orthonormal framing of TH_λ. Let ξ be a 2-plane field on H_λ. Then ξ determines a map $\psi : H_\lambda \to S^2$, by sending the unit normal vector to ξ to its

direction in $\mathbb{R}^3$ using the framing of TH_λ. Now let γ be an embedded curve in H_λ. The image $\psi(\gamma)$ is a closed loop S^2, which is contractible and therefore this path is homotopic to a constant path at the north pole. Consequently, we can homotope ξ and assume that $\psi(\gamma)$ is the constant map to the north pole. Geometrically, this means that tangent vector γ' is perpendicular to ξ at every point along γ.

Definition 6.2. A *Lutz twist* of ξ consists of the following operation. Choose a framed neighborhood of γ, with coordinates (r, θ, t). Assume that $\xi = \ker(dt)$. Now, choose smooth functions f, g such that

(1) $f : [0, 2\epsilon] \to \mathbb{R}$ is identically 0 near the endpoints and nonnegative,

(2) $g : [0, 2\epsilon] \to \mathbb{R}$ is increasing, identically -1 near 0, identically 0 near ϵ, and identically 1 near 2ϵ.

Replace ξ with

$$\widehat{\xi} = \ker(g\,dt + f\,d\theta).$$

Exercise 6.3. (1) Show that applying two Lutz twists along γ is homotopic to the identity.

(2) We have described a *left-handed* Lutz twist; i.e., the planes make a single left-handed turn along every diameter of the normal disk to γ. We could alternatively do a *right-handed* Lutz twist by choosing f to be nonpositive. Show that left-handed and right-handed Lutz twists result in homotopic plane fields.

A Lutz twist changes the relative Euler class of the plane field ξ_λ. Let τ denote a fixed trivialization of ξ_λ along Σ and define the relative Euler class $e(\xi_\lambda, \tau) \in H^2(H_\lambda, \Sigma) \cong H_1(H_\lambda)$.

Lemma 6.4. *For a Lutz twist along γ, the relative Euler classes satisfy*

$$e(\xi, \tau) - e(\widehat{\xi}, \tau) = 2[\gamma] \in H_1(H_\lambda).$$

Proof. We can extend τ to a framing that is $\{\partial_r, \partial_\theta\}$ in a tubular neighborhood of γ. This framing must vanish along γ and so $e(\xi_\lambda, \tau) = A + [\gamma]$ for some $A \in H_1(H_\lambda)$. However, after the Lutz twist, we can use the same framing, which still vanishes along γ, except with opposite sign. Thus $e(\widehat{\xi_\lambda}, \tau) = A - [\gamma]$. □

6D. *Action of $H^2(X; \mathbb{Z})$.* The set of Spin$^{\mathbb{C}}$-structures on X is an affine copy of $H^2(X; \mathbb{Z})$. This means that $H^2(X; \mathbb{Z})$ acts freely and transitively on the set of Spin$^{\mathbb{C}}$-structures. That is, given a Spin$^{\mathbb{C}}$-structure $\mathfrak{s}$ and some nonzero $A \in H^2(X; \mathbb{Z})$, there is a distinct Spin$^{\mathbb{C}}$-structure $\mathfrak{s}' = \mathfrak{s} + A$. Furthermore, the first Chern classes satisfy

$$c_1(\mathfrak{s} + A) = c_1(\mathfrak{s}) + 2A.$$

To describe the action of $H^2(X;\mathbb{Z})$ on the set of Spin$^{\mathbb{C}}$-structures, we interpret $H^2(X)$ by dualizing the complex in Proposition 2.7. This is a complex

$$0 \to H_1(\Sigma) \to \bigoplus_\lambda H_1(H_\lambda) \to \bigoplus_\lambda H_1(Z_\lambda) \to 0$$

whose middle homology group is isomorphic to $H^2(X;\mathbb{Z})$. In particular, it consists of triples $(a,b,c) \in \bigoplus_\lambda H_1(H_\lambda)$ such that

$$a-b=0 \in H_1(Z_1), \quad b-c=0 \in H_1(Z_2), \quad c-a=0 \in H_1(Z_3)$$

modulo the image of $H_1(\Sigma)$.

In order to move from almost-complex structures to Spin$^{\mathbb{C}}$-structures, we need the following facts.

Lemma 6.5. *Let X be a closed 4-manifold with a handle decomposition. Let J be an almost-complex structure on the 2-skeleton X_2 and let ξ be the field of J-tangencies along the boundary $Y_2 := \partial X_2$. In particular, ξ is the 2-plane field $TY_2 \cap J(TY_2)$. Then J extends across a 3-handle attached along a 2-sphere $S \subset Y_2$ if and only if $\langle e(\xi), [S]\rangle = 0$.*

Proof. One direction is obvious: if a 3-handle is attached along S then $[S]=0$ in $H_2(X;\mathbb{Z})$. Thus $\langle e(\xi),[S]\rangle = \langle c_1(J),[S]\rangle = 0$.

Conversely, suppose that $\langle e(\xi),[S]\rangle = 0$. There is a homotopy $\{\xi_t\}$ of 2-plane fields from $\xi=\xi_0$ to ξ_1 such that ξ_1 is the standard, *negative* tight contact structure in a neighborhood of S. There is an almost-complex structure J on $Y\times[0,1]$ whose restriction to $Y\times\{t\}$ is precisely ξ_t.

Finally, we can attach a 3-handle by turning a Stein 1-handle upside-down. A 1-handle addition is cobordism from $S^0\times B^3$ to $B^1\times S^2$; when $S^0\times B^3$ has the standard tight contact structure, the almost-complex structure can be extended across the cobordism and induces the standard tight (positive) contact structure on $B^1\times S^2$. Turning this upside-down, the induced contact structure is *negative* on the neighborhood of S^2. Therefore, the almost-complex structure extends across the 3-handle. □

Choose a thickening of the spine and let $\{\widehat{Y}_\lambda\}$ denote its boundary components. If J is an almost-complex structure on the spine, let $\{\widehat{\xi}_\lambda\}$ denote the fields of J-complex tangencies.

Corollary 6.6. *An almost-complex structure J on the spine of the trisection $\mathcal{T}$ of X is a* Spin$^{\mathbb{C}}$*-structure if and only if the plane field $\widehat{\xi}_\lambda$ satisfies $e(\widehat{\xi}_\lambda)=0$ for all $\lambda=1,2,3$.*

We can now define the action of $H^2(X;\mathbb{Z})$ on a Spin$^{\mathbb{C}}$-structure $\mathfrak{s}$.

(1) We can view $\mathfrak{s}$ as an almost-complex structure on the spine such that the Euler classes $e(\widehat{\xi}_\lambda)$ all vanish.

(2) Given $A \in H^2(X;\mathbb{Z})$, represent it by a triple (a,b,c) in $\bigoplus_\lambda H_1(H_\lambda)$. We can represent each element a, b, c, by an embedded collection of curves $\{\gamma_\lambda \subset H_\lambda\}$.

(3) Modify J by a Lutz twist on every component of γ_λ for $\lambda = 1, 2, 3$.

Exercise 6.7. Show that after the Lutz twists, we still have that $e(\widehat{\xi}_\lambda) = 0$ for each $\lambda = 1, 2, 3$.

Consequently, the resulting almost-complex structure also extends across the 3-handles and determines a Spin$^{\mathbb{C}}$-structure.

Exercise 6.8. Show that modifying a Spin$^{\mathbb{C}}$-structure $\mathfrak{s}$ by the Lutz twist along $A \sim (a,b,c)$, the first Chern classes satisfy

$$c_1(\mathfrak{s} + A) = c_1(\mathfrak{s}) + 2A.$$

[Hint: how does the Lutz twist affect the 1-complex C_J from Proposition 5.15?]

References

[1] D. Auroux, S. K. Donaldson, and L. Katzarkov, "Singular Lefschetz pencils", *Geom. Topol.* **9** (2005), 1043–1114. MR Zbl

[2] R. I. Baykur, "Existence of broken Lefschetz fibrations", *Int. Math. Res. Not.* **2008** (2008), art. id. rnn101. MR Zbl

[3] G. E. Bredon, *Sheaf theory*, vol. 170, 2nd ed., Graduate Texts in Mathematics, Springer-Verlag, New York, 1997. MR

[4] S. K. Donaldson, "Lefschetz pencils on symplectic manifolds", *J. Differential Geom.* **53**:2 (1999), 205–236. MR Zbl

[5] P. Feller, M. Klug, T. Schirmer, and D. Zemke, "Calculating the homology and intersection form of a 4-manifold from a trisection diagram", *Proc. Natl. Acad. Sci. USA* **115**:43 (2018), 10869–10874. MR Zbl

[6] V. Florens and D. Moussard, "Torsions and intersection forms of 4-manifolds from trisection diagrams", *Canad. J. Math.* **74**:2 (2022), 527–549. MR

Received 27 Jan 2021. Revised 21 Oct 2021.

PETER LAMBERT-COLE: plc@uga.edu
Department of Mathematics, University of Georgia, Boyd Graduate Studies Research Center, Athens, GA, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.265

A remark on quantum Hochschild homology

Robert Lipshitz

We observe that quantum Hochschild homology is a composition of two familiar operations, and give a short proof that it gives an invariant of annular links, in some generality. Much of this is implicit in Beliakova, Putyra and Wehrli's work.

Beliakova, Putyra and Wehrli studied various kinds of traces, in relation to annular Khovanov homology [2]. In particular, to a graded algebra and a graded bimodule over it, they associate a quantum Hochschild homology of the algebra with coefficients in the bimodule, and use this to obtain a deformation of the annular Khovanov homology of a link. A spectral refinement of the resulting invariant was recently given by Akhmechet, Krushkal and Willis [1].

Before giving our reformulation, we recall Beliakova, Putyra, and Wehrli's definition.

Definition 1 [2, Section 3.8.5]. Let A be a graded ring, M a chain complex of graded A-bimodules (so M is bigraded), and $q \in A$ an invertible central element with grading 0. The *quantum Hochschild complex* of A with coefficients in M and parameter q has chain groups $qCH_n(A; M) = M \otimes_{\mathbb{Z}} A^{\otimes_{\mathbb{Z}} n}$ and differential

$$\partial(m \otimes a_1 \otimes \cdots \otimes a_n) = ma_1 \otimes a_2 \otimes \cdots \otimes a_n + \sum_{i=1}^{n-1} (-1)^i m \otimes a_1 \otimes \cdots \otimes a_i a_{i+1} \otimes \cdots \otimes a_n + (-1)^n q^{-|a_n|} a_n m \otimes a_1 \otimes \cdots \otimes a_{n-1}.$$

The homology of this complex is the *quantum Hochschild homology* $qHH_\bullet(A; M)$ of A with coefficients in M and parameter q.

This work was supported by NSF grant DMS-1810893.
MSC2020: 16E40, 57K16.
Keywords: Hochschild homology, trace decategorification, Chen–Khovanov algebras, annular Khovanov homology.

The goal of this note is to reformulate this operation and deduce that it often leads to annular link invariants. The data of A and q specifies a ring homomorphism $f_q : A \to A$ defined on homogeneous elements a of A by

$$f_q(a) = q^{-|a|}a,$$

where $|a|$ denotes the grading of a. We can twist the left action on the A-bimodule M by f_q to obtain a new bimodule ${}_{f_q}M$ which is equal to M as a right A-module and has left action given by the composition $A \otimes_{\mathbb{Z}} {}_{f_q}M \xrightarrow{f_q \otimes \mathbb{I}} A \otimes_{\mathbb{Z}} M \xrightarrow{m} M = {}_{f_q}M$. This operation is a special case of tensor product:

$${}_{f_q}M \cong {}_{f_q}A \otimes_A M$$

(compare [2, Section 3.8.3]).

Our first observation is:

Proposition 2. *The quantum Hochschild homology of A with coefficients in M is isomorphic to the ordinary Hochschild homology of A with coefficients in ${}_{f_q}M$.*

Proof. This is immediate from the definitions. □

Call a chain complex of graded A-bimodules M *weakly central* if for any graded A-bimodule N there is a quasi-isomorphism $M \otimes_A^L N \simeq N \otimes_A^L M$.

Lemma 3. *The bimodule ${}_{f_q}A$ is weakly central.*

Proof. The isomorphism $M \otimes_A {}_{f_q}A \to {}_{f_q}A \otimes M$ sends m to $q^{-|m|}m$. □

We turn next to annular link invariants. Consider the category Tan with one object for each even integer and $\mathrm{Hom}(2m, 2n)$ given by the set of isotopy classes of $(2m, 2n)$-tangles (embedded in $D^2 \times [0, 1]$). Given a (graded) algebra A, a *very weak action* of Tan on the category of A-modules is a choice of quasi-isomorphism class of chain complex of (graded) A-bimodules $C(T)$ for each $T \in \mathrm{Hom}(2m, 2n)$ so that $C(T_2 \circ T_1)$ is quasi-isomorphic to $C(T_2) \otimes_A^L C(T_1)$. For example, if we take A to be the direct sum of the Khovanov arc algebras [4] then Khovanov defined a very weak action of Tan on ${}_A\mathsf{Mod}$, and if we define A to be the direct sum of the Chen–Khovanov algebras [3] then Chen and Khovanov defined a very weak action of Tan on ${}_A\mathsf{Mod}$. (In fact, in both cases, they did more; cf. Remark 6.)

Any $(2n, 2n)$-tangle $T \subset D^2 \times [0, 1]$ has an *annular closure* in $D^2 \times S^1$.

Proposition 4. *Fix a very weak action of Tan on ${}_A\mathsf{Mod}$ and fix a weakly central A-bimodule P. Then, for any $(2n, 2n)$-tangle T, the isomorphism class of $HH_*(A; C(T) \otimes_A^L P)$ is an invariant of the annular closure of T.*

(Compare [2, Corollary 3.23].)

Proof. This is immediate from the definitions and the trace property of Hochschild homology, i.e., that, given A-bimodules M and N,

$$HH_*(A; M \otimes_A^L N) \cong HH_*(A; N \otimes_A^L M). \qquad \square$$

The following is part of Beliakova, Putyra and Wehrli's Theorem B [2]:

Corollary 5. *Up to isomorphism, the quantum Hochschild homology of the Chen–Khovanov bimodule associated to a* $(2n, 2n)$*-tangle* T *is an invariant of the annular closure of* T*.*

Proof. This is immediate from Lemma 3, Proposition 4, and the fact that the Chen–Khovanov bimodules induce a very weak action of Tan [3]. $\square$

Remark 6. To keep this note short, we have not discussed functoriality of these annular link invariants under annular cobordisms. To do so, one replaces Tan by an appropriate 2-category of tangles and weak centrality by a notion keeping track of the isomorphisms. See Beliakova, Putyra and Wehrli [2] for further discussion.

References

[1] R. Akhmechet, V. Krushkal, and M. Willis, "Stable homotopy refinement of quantum annular homology", *Compos. Math.* **157**:4 (2021), 710–769. MR Zbl

[2] A. Beliakova, K. K. Putyra, and S. M. Wehrli, "Quantum link homology via trace functor I", *Invent. Math.* **215**:2 (2019), 383–492. MR Zbl

[3] Y. Chen and M. Khovanov, "An invariant of tangle cobordisms via subquotients of arc rings", *Fund. Math.* **225**:1 (2014), 23–44. MR Zbl

[4] M. Khovanov, "A functor-valued invariant of tangles", *Algebr. Geom. Topol.* **2** (2002), 665–741. MR Zbl

Received 28 Oct 2020. Revised 31 Jan 2021.

ROBERT LIPSHITZ: lipshitz@uoregon.edu
Department of Mathematics, University of Oregon, Eugene, OR, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.269

On uniqueness of symplectic fillings of links of some surface singularities

Olga Plamenevskaya

We consider the canonical contact structures on links of rational surface singularities with reduced fundamental cycle. These singularities can be characterized by their resolution graphs: the graph is a tree, and the weight of each vertex is no greater than its negative valency. The contact links are given by the boundaries of the corresponding plumbings. In a joint work with L. Starkston, we have previously shown that if the weight of each vertex in the graph is at most -5, the contact structure has a unique symplectic filling (up to symplectic deformation and blow-up); the proof was based on a symplectic analog of de Jong and van Straten's description of smoothings of these singularities. Here, we give a short self-contained proof of the uniqueness of fillings, via analysis of positive monodromy factorizations for planar open books supporting these contact structures.

1. Introduction

In this note, we consider links of complex surface singularities, equipped with their canonical contact structures. Let $X \subset \mathbb{C}^N$ be a singular complex surface with an isolated singularity at the origin. For small $r > 0$, the intersection $Y = X \cap S_r^{2N-1}$ with the sphere $S_r^{2N-1} = \{|z_1|^2 + |z_2|^2 + \cdots + |z_N|^2 = r\}$ 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.

Our main result, Theorem 1, states that for a certain class of singularities, the canonical contact structure on the link has a unique symplectic filling (up to blow-up and symplectic deformation). This theorem was originally proved in [23]; here, we

Partially supported by NSF grant DMS-1906260.
MSC2020: 57K33, 57K43.
Keywords: links of singularities, symplectic fillings.

will give a new proof, from a different perspective. The sufficient condition will be stated in terms of the dual resolution graph of the singularity. Recall that for a normal surface singularity $(X, 0)$, this graph is defined as follows. Consider a resolution of the singularity, i.e., a proper birational morphism $\pi : \widetilde{X} \to X$ such that $\widetilde{X}$ is smooth. 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. The (dual) resolution graph encodes the topology of the resolution: the vertices $E \in G$ correspond to the exceptional curves and are weighted by the self-intersection $E \cdot E$ of the corresponding curve, while the edges of G record intersections of different irreducible components. Up to contactomorphism, the link of the singularity with its canonical contact structure can be reconstructed from the graph G and the data of self-intersections and genera of exceptional curves, as the boundary of the plumbing of symplectic disk bundles over surfaces according to G.

In this paper, we only work with *rational* singularities; then G is always a tree, and each exceptional curve has genus 0. The following assumption plays the key role in this paper: for every exceptional curve E, we require that the self-intersection $E \cdot E$ and the valency $a(E)$ of the corresponding vertex in G satisfy the inequality

$$a(E) \leq -E \cdot E. \tag{1}$$

Plumbing graphs with this property are sometimes referred to as "graphs with no bad vertices" in low-dimensional topology; a bad vertex, by definition, has valency greater than its negative weight. (The boundary of the corresponding plumbing is a Heegaard Floer L-space [21].) If the dual resolution graph is a tree with the above property, $(X, 0)$ is a *rational singularity with reduced fundamental cycle*. In the literature, singularities of this type are also known as *minimal* singularities [13].

We will give a direct new proof of the following theorem, first established in [23]:

Theorem 1 [23]. *Suppose that $(X, 0)$ is a rational surface singularity with reduced fundamental cycle, and assume additionally that every exceptional curve in its resolution has self-intersection at most -5. Then the contact link (Y, ξ) of $(X, 0)$ has a unique minimal weak symplectic filling, which is Stein.*

In the special case where the resolution graph is star-shaped with three legs, this fact is proved in [5, Theorem 2.7, Remark 2.8], by a different method.

Symplectic and Stein fillings of links of surface singularities are of interest because of the connection to algebrogeometric questions, namely to the smoothings of the singularity. The Milnor fiber of each smoothing of $(X, 0)$ gives a Stein filling of its link (Y, ξ); another Stein filling can be provided by the minimal resolution of the singularity, after deforming the symplectic form. (Rational surface singularities

are always smoothable, with an "Artin smoothing component" whose Milnor fiber gives the same Stein filling as the resolution. In particular, for the singularities in Theorem 1, the filling can be viewed as the resolution or as the Milnor fiber for the Artin smoothing.) An important question is whether *all* Stein fillings of a given surface singularity arise in this way [16]. Although this correspondence breaks down when the singularity is sufficiently complicated [1; 2; 23], the answer is positive for certain simple classes of singularities. Namely, all Stein fillings come from Milnor fibers or the minimal resolution for (S^3, ξ_{std}) [6], for links of simple and simple elliptic singularities [19; 20], for lens spaces (links of cyclic quotient singularities) [15; 17], and in general for quotient singularities [4; 22]. Theorem 1 significantly extends this list.

Our interest in the question of Theorem 1 was motivated by a (very special case of) a conjecture of Kollár on deformations of rational surface singularities [14]. The conjecture asserts that every exceptional curve has self-intersection at most -5 in the resolution of a rational singularity, then the base space of a semiuniversal deformation of this singularity has a unique component. For the case of rational singularities with reduced fundamental cycle, the conjecture was established by de Jong and van Straten [12]; in particular, it follows that under the hypotheses of Theorem 1, the singularity has a unique smoothing component. Our Theorem 1 gives the symplectic analog of this statement.

The proof we gave in [23] comes as a side product of the theory developed in that article, where we describe symplectic fillings of the corresponding class of singularities via a symplectic analog of de Jong and van Straten's construction. Fillings are encoded by certain configurations of symplectic disks in $\mathbb{C}^2$; we were then able to apply a lemma of de Jong and van Straten to establish "combinatorial uniqueness" of the corresponding disk arrangements, and then finish the argument via topological considerations.

In this paper, we will instead give a direct proof of Theorem 1, working with open book factorizations. As a corollary, we get a symplectic proof that all smoothings of the corresponding singularity are diffeomorphic. We will assume that the reader is familiar with the basics of open book decompositions for contact 3-manifolds; see [7] for a survey. Under the hypotheses of the theorem, the canonical contact structures on the links of singularities admit planar open books. (This follows from a construction of Gay and Mark [10]; see Section 2. Planarity was also a key ingredient that allowed us to build an analog of the de Jong–van Straten theory in [23].) In the planar case, symplectic fillings can be studied via theorems of Wendl and Niederkruger [18; 25]: every minimal symplectic filling is symplectic deformation equivalent to a Lefschetz fibration over a disk with the same planar fiber P. The classification of fillings then reduces to enumerating positive factorizations of the monodromy of the given open book.

In general, finding all positive factorizations of the monodromy is a daunting task, even in the planar case. The question is much easier if one only seeks to determine the image of the Dehn twists of the factorization in the *abelianization* of the mapping class group of the page. This is equivalent to finding the *homology classes* of the curves about which the Dehn twists are performed; we also disregard the order of the twists. This easier question can be studied by counting how many times the Dehn twists enclose each hole in the planar page, and how many times they enclose each pair of holes. (The planar page is a disk with holes, and we say that a simple closed curve in a disk with holes encloses a hole if the curve separates the hole from the outer boundary component of the disk.) If P is planar, any two factorizations of the boundary-fixing monodromy $\phi : P \to P$ can be connected by a sequence of lantern relations, and it follows that the number of Dehn twists enclosing a given hole (or a given pair of holes) is *independent* of the factorization of ϕ. Thus, we can introduce the multiplicity $m(v)$ of a hole v with respect to the monodromy ϕ, and similarly the joint multiplicity $m(v_1, v_2)$ of a pair of holes v_1, v_2. Knowing these multiplicities, one can attempt to describe possible other factorizations of ϕ, by examining the combinatorics of how the Dehn twists can enclose the holes. This method was introduced in [24] to classify fillings of certain lens spaces.

Once we understand the Dehn twists in the factorization at the level of homology classes of the curves, additional information is needed to find the isotopy classes of the curves. In the case at hand, this step is possible because the given monodromy admits a positive factorization into Dehn twists about *disjoint* curves.

In [23], the combinatorial part of the proof was based on the description of fillings via a symplectic analog of the de Jong–van Straten construction. We then used the result of [12] asserting uniqueness of a combinatorial solution for a certain curve arrangement problem. For the second part of the proof, we gave a direct mapping class argument. The purpose of the note is to give a direct multiplicity-count argument for the first part; see Lemma 2. For the second part, we essentially repeat the reasoning from [23]; this argument, based on right-veering properties, is given in Lemma 3 for completeness.

It is interesting to note that our direct argument for the combinatorics of Dehn twists follows the strategy of [12, Theorem 6.23]: we translate their proof from the incidence matrices to multiplicities of holes, and provide some extra details where needed.

2. Proof of Theorem 1

To begin, we recall the construction of the open books supporting the canonical contact structures for the class of singularities that satisfy (1) [10, Theorem 1.1]. Starting with the plumbing graph G, the construction given by Gay and Mark

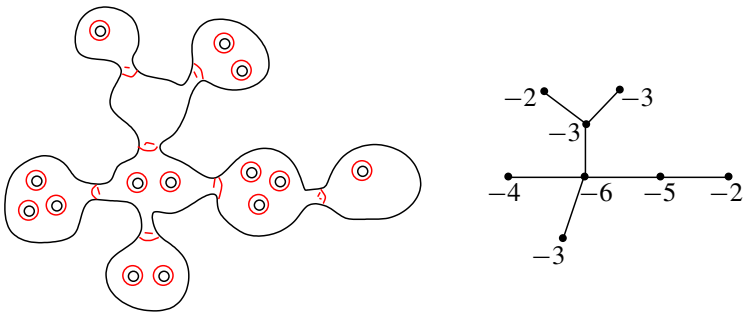


Figure 1. The Gay–Mark open book supporting the canonical contact structure on the link of the singularity with dual resolution graph shown on the right. The page of the open book has genus 0 and is constructed from the spheres with holes corresponding to the vertices of the graph. Each sphere is connected to the other spheres by necks that correspond to the edges; the total number of holes and necks for each sphere equals the negative self-intersection of the vertex. The monodromy is the product of the positive Dehn twists about the boundaries of the holes and the meridians of the necks; these curves are shown in red.

produces a planar Lefschetz fibration compatible with the symplectic resolution of a rational singularity $(X, 0)$ with reduced fundamental cycle. (The symplectic structure on the plumbing can be deformed to the corresponding Stein structure.) We describe the induced planar open book on the link (Y, ξ). To construct the page of the Gay–Mark open book, take a sphere S_E for each vertex $E \in G$ and cut $-a(E) - E \cdot E \geq 0$ disks out of this sphere. (As before, $a(E)$ is the valency of the vertex E; the number of disks is nonnegative by (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_E corresponding to the vertex E, the number of necks equals the number of edges adjacent to E, i.e., its valency $a(E)$. The resulting surface S has genus 0 because G is a tree. See Figure 1 for an example. The open book monodromy is given by the product of positive Dehn twists around each of the holes and around the meridians of the necks. We will call this product the *standard* factorization of the Gay–Mark monodromy.

To examine positive factorizations of this open book, we first we put the resolution graph in the following special form, as in [12]. We choose the vertices $E_1, E_2, \ldots E_k$ and partition the remaining vertices into subsets $R_2, \ldots, R_k$ as shown in Figure 2, so that for any vertex $F \in R_j$, the length $l(E_j, F)$ of the chain from F to E_j satisfies $l(E_j, F) \leq j - 1$. Here, the length of chain means the number of edges; for example, the statement means that every vertex in R_2 is directly connected to E_2 by a single edge. This can always be achieved via the following procedure. We choose E_1 to be the endpoint of a longest chain C in the graph; then E_1 is necessarily a leaf

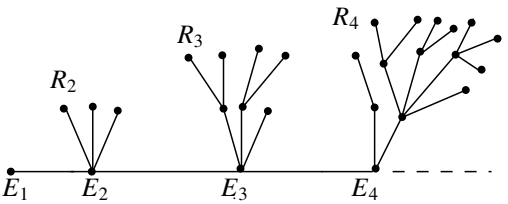


Figure 2. A graph with conveniently arranged vertices: after removing $E_1, E_2, \ldots$, the remaining vertices are partitioned into subsets $R_2, R_3, \ldots$. Every vertex in the set R_j connects to the vertex E_j by a chain with fewer than j edges.

vertex of G. Let E_2 be its adjacent vertex, and let E_3 be the vertex on the chain C that is adjacent to E_2. Removing E_2 from G, we get one connected component consisting of E_1, another that contains E_3, and possibly a number of other vertices in the remaining components. Let R_2 be the set of these remaining vertices. Each vertex $F \in R_2$ must be a leaf vertex connected to E_2 (otherwise we can build a chain longer than C by going to R_2 instead of E_1); thus the condition $l(E_2, F) \leq 1$ is satisfied. If E_3 is a leaf vertex, it can be included in R_2, and the procedure is over. If E_3 is not a leaf vertex, and every other vertex in $G \setminus (E_1 \cup R_2)$ is connected to E_3 by a path of at most 2 edges, then all remaining vertices can be included in E_3. Otherwise we consider the vertex E_4 preceding E_3 on the path C. Removing E_3, we set aside the two components of $G \setminus E_3$ that contain E_1 resp. E_4 and let E_3 be the set of all remaining vertices. Again, since no chain in the graph G can be longer than C, every vertex $F \in R_3$ must be connected to E_3 by a path of no more than 2 edges, satisfying $l(F, E_3) \leq 2$. We continue this process to define E_5, R_4, etc., stopping when we reach k such that all the remaining vertices in G can be connected to E_k by a path no longer than $(k-1)$ edges, and thus placed in R_k. See Figure 2. We will say that vertices of the graph are *conveniently arranged* if they are partitioned into subsets as above.

For the resolution graph G with conveniently arranged vertices, we build the Gay–Mark open book as in Figure 1. We will identify the planar page of this open book with a disk with holes, so that the outer boundary of the disk corresponds to the boundary of one of the holes associated to the vertex E_1. This identification and the choice of the outer boundary component of the disk will be fixed from now on, for the statement and the proof of Lemma 2. In the standard factorization of the Gay–Mark monodromy, there is a sequence of the Dehn twists $D_1, D_2, \ldots, D_k$ around a nested collection of curves $\gamma_1, \ldots, \gamma_k$, such that

(i) D_1 is the twist around the outer boundary component γ_1 of the page, and therefore D_1 encloses all the holes;

(ii) D_j is the twist around the neck γ_j between E_{j-1} and E_j for $j = 2, \ldots, k$, so that D_j encloses all the holes corresponding to $E_j, R_j, E_{j+1}, R_{j+1}, \ldots, E_k, R_k$.

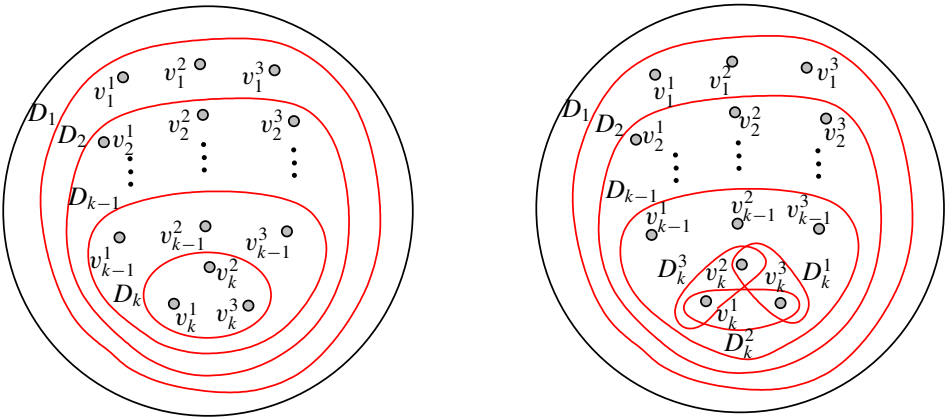


Figure 3. Property F1 (left) and Property F2 (right) for a chosen subset of holes $v_1^1, v_1^2, v_1^3, \dots v_k^1, v_k^2, v_k^3$ and the Dehn twists that enclose them in a factorization. Note that we only require that the *homology classes* of the curves are as schematically shown; isotopy classes may look different from the picture.

The curves $\{\gamma_j\}_{j=1}^k$ cut the disk page into annular domains $\{V_j\}_{j=1}^{k-1}$ such that V_j is bounded by γ_j, γ_{j+1} for $j = 1, \dots, k-1$, and a disk V_k bounded by γ_k. It follows that the joint multiplicity of any two holes from V_j is at least j. In the Gay–Mark construction, the holes from V_j are associated to vertices E_j, R_j of the graph G.

We now make a choice of a certain ordered subset of holes in the page. Because the valency of E_1 is one, and the self-intersection $E_1 \cdot E_1$ is at most -5, the corresponding annular domain V_1 contains $-E_1 \cdot E_1 - 2 \geq 3$ holes. We label three of these holes as v_1^1, v_1^2, v_1^3. Next, again because self-intersections of vertices are at most -5, we can pick three holes v_2^1, v_2^2, v_2^3 in the domain V_2. We require that v_2^1, v_2^2, v_2^3 satisfy an additional condition $m(v_2^r, v_2^s) = 2$: if R_2 is nonempty, we make sure that no two holes are in the same branch of R_2 to avoid higher joint multiplicities. For $j = 3, \dots k$, we proceed to pick v_j^1, v_j^2, v_j^3 in the domain V_j, choosing different branches of R_j if R_j is nonempty, so that $m(v_j^r, v_j^s) = j$ for any pair of indices $r, s = 1, 2, 3$. By construction, we have

$$m(v_i^r, v_j^s) = \min(i, j) \tag{2}$$

for any two chosen holes v_i^r, v_j^s.

The choice of the holes $v_1^1, v_1^2, v_1^3, \dots, v_k^1, v_k^2, v_k^3$ will be fixed. By construction, the standard factorization of the Gay–Mark open book satisfies the following:

Property F1. The factorization includes Dehn twists $D_1, \dots, D_k$ such that

- the Dehn twist D_j encloses the holes v_i^1, v_i^2, v_i^3 for all $i \geq j$.

This is illustrated in Figure 3. Note that we have only listed the Dehn twists that correspond to the edges of the chain $E_1, \dots, E_k$. The Dehn twists that correspond

to edges the sets R_j are not listed above; for each R_j, the corresponding Dehn twists enclose holes in the domain V_j. While these Dehn twists may be nested, the hypothesis that chains of edges in R_j connecting to E_j have length at most $j-1$ gives a bound on a number of twists enclosing any hole w in V_j in the standard factorization: there are at most j nested Dehn twists inside V_j (including the Dehn twist around the boundary of w), in addition to $D_1, \dots, D_j$. This means that in ϕ, the multiplicity of the hole w is at most $2j$.

For an inductive step in our proof, we will consider graphs G satisfying a weaker hypothesis: when the vertices of G are conveniently arranged, we require that the self-intersection $E_k \cdot E_k$ be less than or equal to -4, whereas $E \cdot E \leq -5$ for every other vertex $E \in G$. Note that in the case where $E_k \cdot E_k = -4$, we can still choose a labeled collection of holes $v_1^1, v_1^2, v_1^3, \dots, v_k^1, v_k^2, v_k^3$ as above. Because $E_k \cdot E_k = -4$, we can use the lantern relation to replace the product of D_k and three other Dehn twists (around holes or necks in the sphere corresponding to E_k) by three Dehn twists D_k^1, D_k^2, D_k^3. For this new factorization, we have:

Property F2. The factorization includes Dehn twists $D_1, \dots, D_{k-1}, D_k^1, D_k^2, D_k^3$ such that:

- D_j encloses the holes v_i^1, v_i^2, v_i^3 for all $i \geq j$, for each for each $j = 1, \dots, k-1$.
- D_k^1 encloses v_k^2, v_k^3 but not v_k^1.
- D_k^2 encloses v_k^1, v_k^3 but not v_k^2.
- D_k^3 encloses v_k^1, v_k^2 but not v_k^3.

Under the hypotheses of the following lemma, we will show that an *arbitrary* factorization of the monodromy of the Gay–Mark open book must have Property F1 or Property F2. This will be a step in the argument showing that any monodromy factorization must be standard if the self-intersection of each vertex of G is at most -5.

Lemma 2. *Suppose that the vertices of the graph G are conveniently arranged, with distinguished vertices $E_1, E_2, \dots, E_k$, and the corresponding sets $R_2, \dots, R_k$. Assume that the self-intersections of all vertices in the graph are at most -5, except possibly E_k, which has self-intersection at most -4. Suppose also that there is a collection of holes $\{v_j^1, v_j^2, v_j^3\}_{j=1}^k$, chosen as above. Then*:

(1) *If $E_k \cdot E_k = -4$, then **every** monodromy factorization includes*
   (a) *the Dehn twist around the outer boundary component of the page,*
   (b) *a collection of Dehn twists (containing the twist around the outer boundary) that has Property F1 or Property F2.*

(2) *If $E_k \cdot E_k \leq -5$, then every monodromy factorization is homologically equivalent to the standard one.*

Proof. We will build an inductive argument, with double induction on k and the number of vertices in the graph.

The base of induction is given by $k = 1$ and $k = 2$. The case of $k = 1$ corresponds to lens spaces and was treated in [24]. (This is an easy exercise on computing multiplicities.) The case $k = 2$ is straightforward but more tedious; we check it after explaining the induction step.

For now, assume that $k \geq 2$, and that the statement of the lemma is established for all graphs where the chain of distinguished vertices $E_1, E_2, \ldots$ has length at most k. Consider the graph G with conveniently arranged vertices with a longer chain $E_1, \ldots, E_{k+1}$, and the remaining vertices partitioned into the sets $R_2, \ldots, R_{k+1}$. Take a new graph G', obtained from G by removing all vertices of R_{k+1} and E_{k+1}, increasing by 1 the self-intersection of E_k, and keeping the same self-intersection for all other vertices. The Gay–Mark open books (P, ϕ) and (P', ϕ'), representing, respectively, the contact links of singularities with graphs G and G', are related as follows. To obtain the page P' from P, we cap off all the holes in P associated to E_{k+1} and R_{k+1}; the boundary-fixing diffeomorphism ϕ then induces ϕ'. (Note that $E_k \cdot E_k$ increases by 1 in the graph G' since the corresponding subsurface in the page has fewer boundary components now: removing E_{k+1}, R_{k+1} is the same as pinching off the neck connecting the E_k-sphere to the E_{k+1}-sphere.)

Fix an arbitrary factorization Φ of the Gay–Mark open book for G. When the holes in P are capped off to obtain P', Φ induces the factorization Φ' for the open book (P', ϕ'). Since by assumption $E_k \cdot E_k \leq -5$ in G, the self-intersection of the corresponding vertex is at most -4 in G'. The induction hypothesis applies to the graph G', and therefore, the conclusion of the lemma holds for the factorization Φ' of the monodromy ϕ'. In particular, there is a Dehn twist $T' = T'_1$ around the outer boundary component of P' in the factorization Φ', and moreover, there are Dehn twists $T'_2, T'_3, \ldots, T'_{k-1}$, and T'_k (or $T'_{k,1}, T'_{k,2}, T'_{k,3}$) that have Property F1 (or, respectively, Property F2). These Dehn twists must be induced by the corresponding Dehn twists $T = T_1, T_2, \ldots, T_{k-1}$ and T_k (or $T_{k,1}, T_{k,2}, T_{k,3}$) in the factorization Φ of $\phi : P \to P$.

To show that Φ has a Dehn twist around the outer boundary component of P, we need to check that T encloses all the holes corresponding to E_{k+1} and R_{k+1} that were removed from P to obtain P'; we already know that T encloses all the holes that P inherits from P'. For the sake of contradiction, let $v = v^s_{k+1}$ be a hole associated to E_{k+1} or R_{k+1}, and suppose that it is not enclosed by T. First assume that the factorization Φ' has property F1, so that the factorization Φ includes Dehn twists $T = T_1, T_2, \ldots, T_{k-1}, T_k$ as above. We examine the multiplicities of the selected holes. Because these multiplicities can be computed from the standard factorization of $\phi : P \to P$, by (2) we know that the joint multiplicity $m(v, v^s_j) = j$ for $i = 1, 2, 3$ and $j = 1, \ldots, k$.

If $T = T_1$ does not enclose v, the holes v and v_j^s are enclosed together by at most $j-1$ of the Dehn twists $T_1, T_2, \ldots, T_{k-1}, T_k$. Even if v is enclosed by all of $T_2, \ldots, T_{k-1}, T_k$, it follows that there must be an additional Dehn twist τ_j^s enclosing both v and v_j^s. Observe that the Dehn twists τ_j^i must be all distinct (that is, $\tau_j^s = \tau_i^r$ only if $i = j$, $r = s$): the joint multiplicity $m(v_j^s, v_i^r)$ of any two distinct holes v_j^s, v_i^r is already realized by $T_1, T_2, \ldots, T_{k-1}, T_k$, so no additional Dehn twist can enclose them both. It follows that the hole v must be enclosed by at least $3k$ distinct Dehn twists $\tau_1^1, \tau_1^2, \tau_1^3, \ldots, \tau_k^1, \tau_k^2, \tau_k^3$ in the factorization Φ, in addition to $k-1$ Dehn twists $T_2, \ldots, T_{k-1}, T_k$. It is not hard to see that if v is not enclosed by some of the twists among $T_2, \ldots, T_{k-1}, T_k$, then each missing twist will need to be replaced by several individual twists to achieve $m(v, v_j^s) = j$. It follows that v is enclosed by at least $3k + k - 1 = 4k - 1$ twists. To obtain a contradiction, we compute the multiplicity $m(v)$ in the monodromy ϕ. The hole v is associated to E_{k+1} or to some vertex E in R_{k+1}; in the standard factorization of ϕ, it is enclosed by the small twist around the hole v, by the outer boundary twist, as well as by the Dehn twists corresponding to the edges in the chain in G from E_1 to E_{k+1} and then the chain from E_{k+1} to E, if the latter chain is present. Since $E \in R_{k+1}$, and by construction the length of the chain from E_{k+1} to any vertex in R_{k+1} is at most k, we see that $m(v) \leq 2k + 2$. This is a contradiction since $2k + 2 < 4k - 1$ for $k \geq 2$.

Similar reasoning leads to the same conclusion in the case where the factorization Φ' has Property F2 instead of Property F1. As above, we see that if v is not enclosed by $T = T_1$, there must be at least $3(k-1)$ distinct Dehn twists $\tau_1^1, \tau_1^2, \tau_1^3, \ldots, \tau_{k-1}^1, \tau_{k-1}^2, \tau_{k-1}^3$ in the factorization Φ to achieve $m(v, v_j^s) = j$ for $j = 1, \ldots, k-1$, $s = 1, 2, 3$. The holes v_k^1, v_k^2, v_k^3 need a bit more attention. Indeed, when v is not enclosed by $T = T_1$, there are at most $k-2$ twists among $T_2, \ldots, T_{k-1}$ enclosing v and v_k^s for each $s = 1, 2, 3$. The joint multiplicity $m(v, v_k^s) = k$ can be achieved if v is enclosed by all three twists $T_{k,1}, T_{k,2}, T_{k,3}$, in addition to all of $T_2, \ldots, T_{k-1}$ and $\tau_1^1, \tau_1^2, \tau_1^3, \ldots, \tau_{k-1}^1, \tau_{k-1}^2, \tau_{k-1}^3$. This would give $m(v) \geq 3k + (k-1)$ as before. Another case is when one of the twists $T_{k,1}, T_{k,2}, T_{k,3}$ (say $T_{k,3}$) does not enclose v. In that case, two additional twists, distinct from all of the above, enclosing, respectively, v and v_k^1 (but not v_k^2 or v_k^3) and v and v_k^2 (but not v_k^1 or v_k^3), are needed (again for the reason of joint multiplicities). This would still yield $m(v) > 3k + (k-1)$. As in the case of Property F1, the multiplicity of v will be even higher if v is not enclosed by some of the twists among $T_2, \ldots, T_{k-1}, T_k$. As above, we get a contradiction since $m(v) \leq 2k + 2$, as computed from the standard factorization of Φ.

At this point, we have shown that the Gay–Mark open book for the resolution graph G must have an outer boundary twist T in any factorization Φ, assuming that the smaller graph G' satisfies the conclusion of the lemma. To prove the other statements of the lemma for G, we will now reduce to a different smaller graph $\widetilde{G}$.

In the page P, consider all the holes associated to the vertex $E_1 \in G$. We know that these holes have joint multiplicity 1 with any other hole in P, thus they cannot be enclosed by any twists other than T that involve several holes. Since there is one boundary Dehn twist δ_i around each of these holes in the standard factorization, so that the multiplicity of each hole is 2, these boundary twists must be present in Φ as well. It follows that the factorization Φ has the form $\Phi = T\delta_1\delta_2 \cdots \delta_m \circ \widetilde{\Phi}$, where $\widetilde{\Phi}$ is supported in $\widetilde{P} = P \setminus V_1$, the part of the page P associated to $G \setminus E_1$. Remove the vertex E_1 and its connecting edge from the graph G, and consider the resulting graph $\widetilde{G}$, keeping the same self-intersections of vertices. Clearly, $\widetilde{\Phi}$ gives a factorization of the Gay–Mark open book associated to $\widetilde{G}$. If self-intersections of all vertices of G are at most -5, the same holds for $\widetilde{G}$. The graph $\widetilde{G}$ has fewer vertices, so by the induction hypothesis, the factorization $\widetilde{\Phi}$ must be standard. It follows that the factorization Φ of the Gay–Mark open book for G is standard as well, proving part (2) of the lemma.

To make the induction step work for part (1), assume that in G, the vertex E_{k+1} has self-intersection at most -4, while all the other vertices have self-intersection at most -5. When we remove E_1 to form the graph $\widetilde{G}$, the vertices of $\widetilde{G}$ may no longer be conveniently arranged; after vertices are rearranged, we need to have the (-4) vertex at the appropriate position to apply the induction hypothesis. We must rearrange the vertices of $\widetilde{G}$ to have a chain $\widetilde{E}_1, \widetilde{E}_2, \ldots$, with the other vertices partitioned into the sets $\widetilde{R}_1, \widetilde{R}_2, \ldots$, so that the length of any chain in E_j, R_j is at most $j-1$. Consider the vertex E_2 in G. If R_2 is not empty in G, we can pick a vertex of R_2 to play the role $\widetilde{E}_1$ in $\widetilde{G}$, let $\widetilde{R}_2$ be the remaining vertices of R_2, and set $\widetilde{E}_2 = E_2$, $\widetilde{E}_3 = E_3, \ldots, \widetilde{R}_3 = R_3$, $\widetilde{R}_4 = R_4, \ldots$. In this case, the (-4) vertex E_{k+1} in G becomes the vertex $\widetilde{E}_{k+1}$ at the end of the chain $\widetilde{E}_1, \widetilde{E}_2, \ldots$ in $\widetilde{G}$, as required. If R_2 is empty in G, we check if R_3 has a chain of (maximum possible) length 2. If so, we rearrange the vertices: let $\widetilde{E}_3 = E_3$, pick $\widetilde{E}_1$ and $\widetilde{E}_2$ forming a length 2 chain in R_3 (with $\widetilde{E}_1$ being the leaf vertex). Let $\widetilde{R}_2$ consist of all vertices other than $\widetilde{E}_1$, $\widetilde{E}_3$, and let $\widetilde{R}_3$ consist of all remaining vertices of R_3, together with the old vertex E_2. For $j \geq 4$, we have $\widetilde{E}_j = E_j$, $\widetilde{R}_j = R_j$, so the (-4) vertex remains in the right place for the graph $\widetilde{G}$, which is now conveniently arranged. See Figure 4. If there are no chains of length 2 in R_3, we similarly examine R_4 to see if there are chains of length 3. If so, we flip the graph to put this length 3 chain into the position of vertices $\widetilde{E}_1, \widetilde{E}_2, \widetilde{E}_3$, make the vertices E_2, E_3 and all of R_3 be part of the new set R_4; the graph is now conveniently arranged, and the (-4) vertex does not move. If there are no length 3 chains in R_4, we look at R_5, etc. To summarize, the above procedure means that we can conveniently rearrange the vertices of $\widetilde{G}$ without moving the (-4) vertex whenever for some $j = 2, \ldots, k$, the set R_j in G has a chain of the maximum possible length $j-1$. If such a chain does not exist, we check if R_{k+1} has a chain of length k. If so, this chain will become the

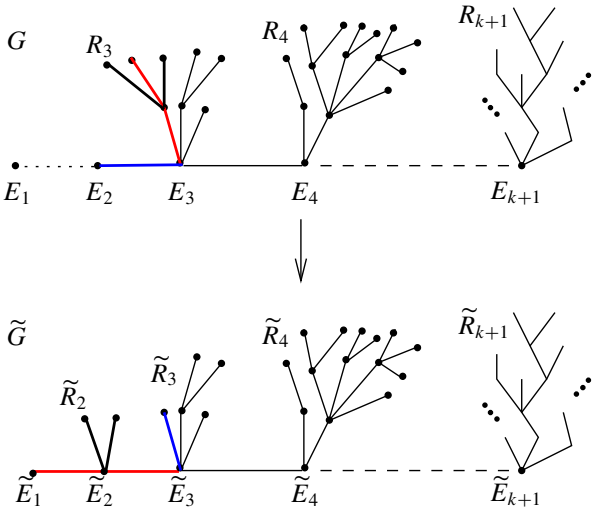


Figure 4. After removing the vertex E_1 from G, we flip a chain in the graph to make the new graph $\widetilde{G}$ conveniently arranged, while keeping in place the vertex E_k. The graph G is shown at the top, the new graph $\widetilde{G}$ at the bottom. The picture illustrates the situation where R_2 is empty, and we flip a length 2 chain in R_3, together with all the edges and vertices attached to this chain in R_3. The vertex E_2 becomes part of the new set $\widetilde{R}_3$.

new chain $\widetilde{E}_1, \widetilde{E}_2, \dots$, the old vertices $E_2, \dots, E_k$ as well as the sets $R_2, \dots, R_k$ will all be in $\widetilde{R}_{k+1}$, and the graph will be conveniently rearranged without moving the (-4) vertex $E_{k+1} = \widetilde{E}_{k+1}$. Lastly, if each chain R_j, $j = 2, \dots, k+1$ has length at most $j-2$ in G, we can set $\widetilde{E}_1 = E_2, \widetilde{E}_2 = E_3, \dots, \widetilde{E}_k = E_{k+1}$ and $\widetilde{R}_2 = R_3, \dots, \widetilde{R}_k = R_{k+1}$, so that the graph $\widetilde{G}$ will be conveniently arranged, and the vertex $\widetilde{E}_k \in \widetilde{G}$ will have self-intersection -4.

With the rearrangement in place, part (1) follows by induction: if the factorization Φ of the Gay–Mark open book for G has the form $\Phi = T\delta_1\delta_2 \cdots \delta_m \circ \widetilde{\Phi}$, where $\widetilde{\Phi}$ is the factorization of the Gay–Mark open book for $\widetilde{G}$, and part (1) of the lemma holds for $\widetilde{\Phi}$, then clearly the same is true for Φ.

We now return to the base of induction and check the case $k = 2$. In this case, the graph is star-shaped with legs of length 1, with E_2 in the center. As above, the page P is identified with the disk whose outer boundary corresponds to one of the holes associated to E_1; there are at least three holes v_1^1, v_1^2, v_1^3 in V_1. First, we claim that any factorization has a Dehn twist enclosing all of these holes. If not, we must have two distinct Dehn twists τ_1 and τ_2, τ_1 enclosing v_1^1, v_3^1 and τ_2 enclosing v_2^1, v_3^1, because $m(v_1^r, v_1^s) = 1$. Since $m(v_3^1) = 2$ and there are at least 4 holes in P having joint multiplicity 2 with v_3^1, one of these Dehn twists, say τ_1, contains an additional hole w. But then there must be two additional Dehn twists

in the factorization, enclosing, respectively, v_2^1, v_1^1 and v_2^1, w, which is impossible since $m(v_2^2) = 2$. Thus, there is a Dehn twist τ enclosing v_1^1, v_1^2, v_1^3. We would like to show that τ encloses all the holes in P. Suppose not, and let w be a hole not enclosed by τ. Then there must be distinct Dehn twists τ_1, τ_2, τ_3, enclosing, respectively, w, v_1^1, w, v_1^2, and w, v_1^3. Since $m(v_1^1) = m(v_1^2) = m(v_1^3) = 2$, there cannot be any other Dehn twists enclosing v_1^1, v_1^2, v_1^3, and since $m(v_1^r, v) = 1$ for $r = 1, 2, 3$ and any other hole v, every hole v must be either in τ or in *each* of τ_1, τ_2, τ_3 (but not simultaneously in τ and τ_i, $i = 1, 2, 3$). For the hole w, two cases are possible: (1) w belongs to E_2, in which case $m(w) = 3$, so there are no Dehn twists except τ_1, τ_2, τ_3 enclosing w; or (2) w belongs to R_2, in which case $m(w) = 4$, and there is exactly one additional Dehn twist τ'. In either case, there must exist another hole w' such that $m(w, w') = 2$; however, we can only get $m(w, w') = 3$ (if w' is in all three of τ_1, τ_2, τ_3) or $m(w, w') = 1$ (if w' is in τ and τ'). It follows that the Dehn twist τ must enclose all holes in P.

We conclude that the factorization includes Dehn twists around the curves that are homologous, and therefore isotopic, to the boundaries of all the holes associated to E_1. As above, these can be removed from consideration. The same argument works for any leaf vertex of the graph, reducing the question to the situation of only one vertex, E_2. This is the case $k = 1$ representing an open book for a lens space as in [24]; if $E_2 \cdot E_2 \leq -5$, there is a unique factorization, and if $E_2 \cdot E_2 = -4$, then the only other option for the homology classes of curves comes from the lantern relation. □

By Lemma 2, we now know that under the hypotheses of Theorem 1, the Dehn twists in every positive factorization are performed about the curves in the same *homology* classes as the Dehn twists in the standard factorization. We now show that the curves are in the same *isotopy* classes.

Lemma 3. *Let (P, ϕ) be a planar open book whose monodromy ϕ admits a factorization Φ into a product of positive Dehn twists about* ***disjoint*** *simple closed curves in P. Suppose that Φ' is another positive factorization of ϕ, such that Φ is homologically equivalent to Φ'. Then the factorizations Φ and Φ' are the same, up to the order of Dehn twists.*

Proof. After reordering, we can write $\Phi = D_1 D_2 \cdots D_l \delta_1 \delta_2 \cdots \delta_n$, where the δ_i's are Dehn twists about the boundary-parallel curves, and $D_1, D_2, \ldots, D_l$ are the Dehn twists around disjoint curves $\gamma_1, \ldots, \gamma_l$ in P that are not boundary parallel.

Then, again after reordering, we have $\Phi' = T_1 T_2 \cdots T_l \delta_1 \delta_2 \cdots \delta_n$, where the Dehn twists D_j and T_j are performed about homologous curves in P: indeed, every boundary-parallel curve γ_j is determined by its homology class, uniquely up to isotopy. We can thus remove the Dehn twists $\delta_1, \delta_2, \ldots, \delta_n$ from consideration. We will use the same notation, $\Phi = D_1 D_2 \cdots D_l$ and $\Phi' = T_1 T_2 \cdots T_l$ for the two

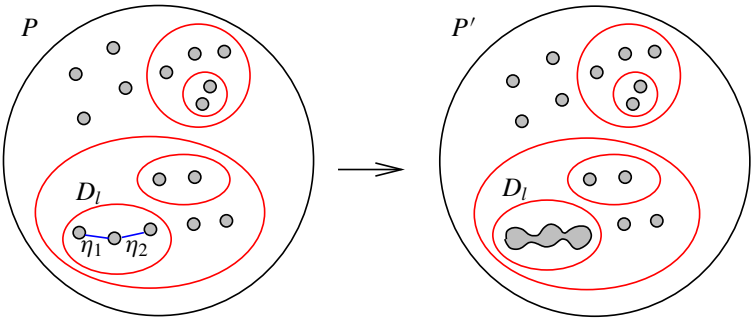


Figure 5. After cutting P along arcs $\eta_1, \eta_2, \ldots$, the Dehn twist D_l becomes boundary-parallel in the new surface P'.

factorizations of the diffeomorphism

$$\phi = D_1 D_2 \cdots D_l = T_1 T_2 \cdots T_l. \tag{3}$$

We will prove the lemma by induction on the number l of the nonboundary parallel Dehn twists. Identifying P with a disk with holes, we can assume that γ_l is an innermost curve in the collection $\gamma_1, \gamma_2, \ldots, \gamma_l$. Suppose that γ_l encloses r holes. Choose a collection of arcs $\eta_1, \eta_2, \ldots, \eta_{r-1}$ connecting these holes and disjoint from γ_l, so that after cutting along these arcs, the holes become a single hole, and the domain enclosed by γ_l becomes an annulus (which deformation retracts to γ_l). See Figure 5. By construction, the arcs $\eta_1, \eta_2, \ldots, \eta_{r-1}$ are disjoint from the support of each of the Dehn twists $D_1, D_2, \ldots, D_l$, thus the diffeomorphism $\phi = D_1 D_2 \cdots D_l$ fixes each of these arcs. As in [3, Proposition 3] and [9, Section 2], we now make the following key observation: after an isotopy removing nonessential intersections, all arcs $\eta_1, \ldots, \eta_{r-1}$ must be also disjoint from the support of each of the Dehn twists $T_1, T_2, \ldots, T_l$. To see this, we recall that each right-handed Dehn twist is a right-veering diffeomorphism of the oriented surface P [11]. 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 P. The right-veering property of a boundary-fixing map $\tau : P \to P$ means that for every simple arc α with endpoints on ∂P, 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 the support of the Dehn twist T_j essentially intersects one of the arcs, say η_1. Then the curve $T_j(\eta_1)$ is not isotopic to η_1 (see, e.g., [8, Proposition 3.2]), so $T_j(\eta_1)$ lies to the right of η_1. Since the composition of right-veering maps is right-veering, we can only get curves that lie further to the right of η_1 after composing with the other Dehn twists $T_1, \ldots, T_l$. However, the composition $\phi = T_1 T_2 \cdots T_j \cdots T_l$ fixes η_1, a contradiction.

Once we know that the support of all the Dehn twists is disjoint from all of the arcs $\eta_1, \ldots, \eta_{r-1}$, we can cut the page P along these arcs, and consider the

image of the relation (3) in the resulting cut-up surface P'. In P', we have that (the induced diffeomorphisms) T_l and D_l are Dehn twists around the curve homologous to the boundary of the same hole, and therefore, $T_l = D_l$ as Dehn twists in P'. It follows that for the Dehn twists (induced by) $D_1, \dots, D_{l-1}$ and $T_1, \dots, T_{l-1}$ in P', we have

$$D_1 D_2 \cdots D_{l-1} = T_1 T_2 \cdots T_{l-1}.$$

By the induction hypothesis, we can conclude that for each $j = 1, \dots, l-1$, the Dehn twists D_j and T_j are performed about isotopic curves in P'. It follows that each pair D_j, T_j gives the same Dehn twists in P, for each $j = 1, \dots, l$. □

Proof of Theorem 1. Under the hypotheses of Theorem 1, the contact 3-manifold (Y, ξ) is supported by an open book with planar page P. Theorems of Wendl and Niederkruger then imply that up to blow-up and deformation of the symplectic form, every weak symplectic filling has a Lefschetz fibration whose fiber is given by P; the monodromy of the fibration is the monodromy of the open book. The Lefschetz fibration is described by its vanishing cycles, or, equivalently, by a positive factorization of the monodromy. Lemmas 2 and 3 show that the positive monodromy factorization is unique. □

Acknowledgements

Originally, Theorem 1 was proved in joint work with Laura Starkston [23]. The alternative proof given here resulted from the author's attempts to understand the combinatorial constructions of [12, Theorem 6.23] to open book factorizations.

This article was written for Proceedings of the 2020 BIRS workshop on Interactions of Gauge Theory and Contact and Symplectic Topology. The author benefited greatly from the series of the BIRS workshops on these topics and would like to thank the organizers of all the past workshops of the series.

References

[1] A. Akhmedov and B. Ozbagci, "Singularity links with exotic Stein fillings", *J. Singul.* **8** (2014), 39–49. MR Zbl

[2] A. Akhmedov and B. Ozbagci, "Exotic Stein fillings with arbitrary fundamental group", *Geom. Dedicata* **195** (2018), 265–281. MR Zbl

[3] R. I. Baykur, N. Monden, and J. Van Horn-Morris, "Positive factorizations of mapping classes", *Algebr. Geom. Topol.* **17**:3 (2017), 1527–1555. MR Zbl

[4] M. Bhupal and K. Ono, "Symplectic fillings of links of quotient surface singularities", *Nagoya Math. J.* **207** (2012), 1–45. MR Zbl

[5] M. Bhupal and A. I. Stipsicz, "Smoothings of singularities and symplectic topology", pp. 57–97 in *Deformations of surface singularities*, Bolyai Soc. Math. Stud. **23**, János Bolyai Math. Soc., Budapest, 2013. MR Zbl

[6] Y. Eliashberg, "On symplectic manifolds with some contact properties", *J. Differential Geom.* **33**:1 (1991), 233–238. MR Zbl

[7] J. B. Etnyre, "Lectures on open book decompositions and contact structures", pp. 103– 141 in *Floer homology, gauge theory, and low-dimensional topology*, Clay Math. Proc. **5**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[8] B. Farb and D. Margalit, *A primer on mapping class groups*, Princeton Mathematical Series **49**, Princeton University Press, 2012. MR Zbl

[9] E. Fossati, "Contact surgery on the Hopf link: classification of fillings", 2019. arXiv 1905.13026

[10] D. Gay and T. E. Mark, "Convex plumbings and Lefschetz fibrations", *J. Symplectic Geom.* **11**:3 (2013), 363–375. MR Zbl

[11] K. Honda, W. H. Kazez, and G. Matić, "Right-veering diffeomorphisms of compact surfaces with boundary", *Invent. Math.* **169**:2 (2007), 427–449. MR Zbl

[12] T. de Jong and D. van Straten, "Deformation theory of sandwiched singularities", *Duke Math. J.* **95**:3 (1998), 451–522. MR Zbl

[13] J. Kollár, "Toward moduli of singular varieties", *Compositio Math.* **56**:3 (1985), 369–398. MR

[14] J. Kollár, "Flips, flops, minimal models, etc", pp. 113–199 in *Surveys in differential geometry* ((Cambridge, MA, 1990)), Lehigh Univ., Bethlehem, PA, 1991. MR

[15] P. Lisca, "On symplectic fillings of lens spaces", *Trans. Amer. Math. Soc.* **360**:2 (2008), 765–799. MR Zbl

[16] A. Némethi, "Five lectures on normal surface singularities", pp. 269–351 in *Low dimensional topology* ((Eger, 1996/Budapest, 1998)), Bolyai Soc. Math. Stud. **8**, János Bolyai Math. Soc., Budapest, 1999. MR Zbl

[17] A. Némethi and P. Popescu-Pampu, "On the Milnor fibres of cyclic quotient singularities", *Proc. Lond. Math. Soc.* (3) **101**:2 (2010), 554–588. MR Zbl

[18] K. Niederkrüger and C. Wendl, "Weak symplectic fillings and holomorphic curves", *Ann. Sci. Éc. Norm. Supér.* (4) **44**:5 (2011), 801–853. MR Zbl

[19] H. Ohta and K. Ono, "Symplectic fillings of the link of simple elliptic singularities", *J. Reine Angew. Math.* **565** (2003), 183–205. MR Zbl

[20] H. Ohta and K. Ono, "Simple singularities and symplectic fillings", *J. Differential Geom.* **69**:1 (2005), 1–42. MR Zbl

[21] P. Ozsváth and Z. Szabó, "On the Floer homology of plumbed three-manifolds", *Geom. Topol.* **7** (2003), 185–224. MR

[22] H. Park, J. Park, D. Shin, and G. Urzúa, "Milnor fibers and symplectic fillings of quotient surface singularities", *Adv. Math.* **329** (2018), 1156–1230. MR Zbl

[23] O. Plamenevskaya and L. Starkston, "Unexpected Stein fillings, rational surface singularities, and plane curve arrangements", 2020. arXiv 2006.06631

[24] O. Plamenevskaya and J. Van Horn-Morris, "Planar open books, monodromy factorizations and symplectic fillings", *Geom. Topol.* **14**:4 (2010), 2077–2101. MR Zbl

[25] C. Wendl, "Strongly fillable contact manifolds and J-holomorphic foliations", *Duke Math. J.* **151**:3 (2010), 337–384. MR Zbl

Received 17 Feb 2021.

OLGA PLAMENEVSKAYA: olga@math.stonybrook.edu
Stony Brook University, Department of Mathematics, Stony Brook, NY, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.285

On the spectral sets of Inoue surfaces

Daniel Ruberman and Nikolai Saveliev

We study the Inoue surfaces S_M with the Tricerri metric and the canonical spin^c structure, and the corresponding chiral Dirac operators twisted by a flat $\mathbb{C}^*$-connection. The twisting connection is determined by $z \in \mathbb{C}^*$, and the points for which the twisted Dirac operators $\mathcal{D}_z^{\pm}$ are not invertible are called spectral points. We show that there are no spectral points inside the annulus $\alpha^{-1/4} < |z| < \alpha^{1/4}$, where $\alpha > 1$ is the only real eigenvalue of the matrix M that determines S_M, and find the spectral points on its boundary. Via Taubes' theory of end-periodic operators, this implies that the corresponding Dirac operators are Fredholm on any end-periodic manifold whose end is modeled on S_M.

1. Introduction

Inoue surfaces are compact complex surfaces with zero second Betti number which are most remarkable in that they contain no holomorphic curves. These surfaces, constructed by Inoue [18], belong to the class VII_0 in Kodaira's classification [6], which is to say that they are minimal connected compact complex surfaces X with Kodaira dimension $\kappa(X) = -\infty$ and the first Betti number $b_1(X) = 1$. In fact, any class VII_0 surface with vanishing second Betti number and no holomorphic curves is biholomorphic to an Inoue surface; see Bogomolov [7; 8] and Teleman [28]. Inoue surfaces, which are not Kähler because their first Betti number is odd, have been extensively studied from the viewpoints of both algebraic and differential geometry.

In this paper, we restrict ourselves to the Inoue surfaces X of class S_M associated with certain integral matrices $M \in \mathrm{SL}(3, \mathbb{Z})$ with one real eigenvalue $\alpha > 1$ and two complex eigenvalues $\beta \neq \bar{\beta}$. These surfaces, described in detail in Section 2, are known to be diffeomorphic to the mapping torus of a self-diffeomorphism of the 3-torus induced by M. It is in this incarnation that the surfaces S_M are best known

Ruberman was partially supported by NSF Grants DMS-1811111 and DMS-1952790, and Saveliev was partially supported by NSF Grant DMS-1952762.
MSC2020: primary 32J15, 53C55; secondary 57R57, 58J50.
Keywords: Dirac operator, Seiberg–Witten, Inoue surface, Tricerri metric.

to topologists. In particular, Cappell and Shaneson [9; 10] independently used some of the matrices M to construct a fake $\mathbb{R}P^4$ and interesting fibered 2-spheres in a homotopy 4-sphere. From this point of view, the manifolds S_M are given by surgery on this homotopy 4-sphere along those knots. The question of when this homotopy 4-sphere is in fact diffeomorphic to S^4 has received considerable attention [2; 1; 3; 14].

Inoue surfaces are an intriguing class of examples to which to apply our work on the Seiberg–Witten invariants [24] and the end-periodic index theorem [25]. Spectral properties of chiral Dirac operators $\mathcal{D}^{\pm}(X)$ play an important role in determining the index of associated Dirac operators on end-periodic manifolds whose end is modeled on an infinite cyclic cover of X. In applications of those papers to date [21; 20; 22], the infinite cyclic cover was a Riemannian product of the real line and a 3-manifold. In the case of an Inoue surface, while this cover is topologically the product of the real line and a 3-torus, it is not a metric product. (This is related to the fact that the monodromy of the bundle $X \to S^1$ has infinite order.) Since the end-periodic index is metric dependent, this makes for an index problem that must be investigated analytically. We study this problem for the Tricerri metric on X, which makes it into a locally conformal Kähler manifold, and the canonical spin^c structure; see Section 2.

More specifically, we are interested in the spectral sets of the associated chiral Dirac operators $\mathcal{D}^{\pm}(X)$. Recall from [24] that $z \in \mathbb{C}^*$ is a spectral point of $\mathcal{D}^{\pm}(X)$ if and only if the operator

$$z^f \circ \mathcal{D}^{\pm}(X) \circ z^{-f} = \mathcal{D}^{\pm}(X) - \ln z \cdot df$$

has nonzero kernel, where $f : X \to S^1$ is a smooth function realizing a generator of $H^1(X; \mathbb{Z}) = \mathbb{Z}$, and df operates by Clifford multiplication. One can easily check that the spectral sets of $\mathcal{D}^+(X)$ and $\mathcal{D}^-(X)$ are obtained from each other by inversion $\tau(z) = 1/\bar{z}$ with respect to the unit circle. The following theorem, which was announced in [25, Section 6.4], is the main result of this paper.

Theorem 1.1. *The operators $\mathcal{D}^{\pm}(X)$ have no spectral points in the annulus $\alpha^{-1/4} < |z| < \alpha^{1/4}$. Furthermore, the only spectral points of $\mathcal{D}^+(X)$ on the circles $|z| = \alpha^{-1/4}$ and $|z| = \alpha^{1/4}$ are, respectively, $z = \alpha^{1/4}\beta$ and $z = \alpha^{1/4}$.*

Let Z_∞ be a spin^c end-periodic manifold whose end is modeled on the infinite cyclic cover of an Inoue surface X. According to Taubes [27, Lemma 4.3], the Dirac operators $\mathcal{D}^{\pm}(Z_\infty)$ are Fredholm in the usual Sobolev L^2 completion if and only if their spectral sets are disjoint from the unit circle $|z| = 1$.

Corollary 1.2. *The operators $\mathcal{D}^{\pm}(Z_\infty) : L^2_1(Z_\infty) \to L^2(Z_\infty)$ are Fredholm on any end-periodic* spin^c *manifold Z_∞ whose end is modeled on an Inoue surface X of type S_M.*

Remark 1.3. Inoue surfaces do not admit metrics of positive scalar metric, as was proved by Albanese [4, Theorem 4.5]. This also follows from Cecchini and Schick [11], making use of the fact that Inoue surfaces are solvmanifolds (see Wall [31; 32] and Hasegawa [16]) and hence are enlargeable in the sense of Gromov and Lawson [15]. In particular, one cannot prove that the operators $\mathcal{D}^{\pm}(Z_\infty)$ of Corollary 1.2 are Fredholm by using the (uniformly) positive scalar curvature at infinity condition as in [15].

Once we establish that the operators $\mathcal{D}^{\pm}(Z_\infty)$ are Fredholm, their index can in principle be calculated as in [25] in terms of an integral term and the periodic eta-invariant $\eta(X)$. The latter is a spectral invariant which generalizes the eta-invariant of Atiyah, Patodi, and Singer [5] and which can be viewed as a regularized count of points in the spectral set of $\mathcal{D}^{\pm}(X)$. The partial information about the spectral set we obtain in this paper is not sufficient to calculate $\eta(X)$ or the associated index of $\mathcal{D}^{\pm}(Z_\infty)$. However, even this modest attempt leads to some fascinating analysis which we felt was worth sharing.

It is worth mentioning that our original interest in end-periodic index theory grew out of our work [24] with Mrowka on Seiberg–Witten theory for 4-manifolds X with $b_2(X) = 0$ and $b_1(X) = 1$. In that paper, a Seiberg–Witten invariant $\lambda_{\mathrm{SW}}(X)$ was defined as a sum of two metric dependent terms. One is a count of solutions to the Seiberg–Witten equations, and the other is an index-theoretic correction term, whose most important part is the index of the Dirac operator $\mathcal{D}^{+}(Z_\infty)$.

Evaluating $\lambda_{\mathrm{SW}}(X)$ for an Inoue surface X presents quite a challenge. One can actually solve a modified version of the Seiberg–Witten equations for the Tricerri metric; see [26; 23]. However, the modification involves a certain twisting of the Dirac operator used in the formulation of the Seiberg–Witten equations. In order to turn this into a calculation of $\lambda_{\mathrm{SW}}(X)$, one would have to first relate this modified Seiberg–Witten equation to the one used in [25]. The second step would be to evaluate the correction term; this is essentially the same as finding the invariant $\eta(X)$. As mentioned above, we are quite far from achieving this.

In conclusion, we mention that a recent paper of Holt and Zhang [17] uses related techniques to investigate $\bar{\partial}$-harmonic forms on a different non-Kähler complex manifold, the Kodaira–Thurston surface [19; 29].

2. Inoue surfaces

The Inoue surfaces X we are interested in are all compact quotients of $\mathcal{H} \times \mathbb{C}$, where $\mathcal{H} = \{w = w_1 + i w_2 \in \mathbb{C} \mid w_2 > 0\}$ is the upper complex half-plane. To construct X, start with an integral matrix $M \in \mathrm{SL}(3, \mathbb{Z})$ with one real eigenvalue $\alpha > 1$ (which must therefore be irrational) and two complex conjugate eigenvalues $\beta \neq \bar{\beta}$. For

example, the matrices

$$A_m = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & m+1 \end{pmatrix},$$

which are equivalent to the Cappell and Shaneson [10] family, will do as long as $-2 \le m \le 3$. Let $a = (a_1, a_2, a_3)$ be a real eigenvector corresponding to α, and $b = (b_1, b_2, b_3)$ a complex eigenvector corresponding to β. Let G_M be the group of complex analytic transformations of $\mathcal{H} \times \mathbb{C}$ generated by

$$\begin{aligned} g_0(w, z) &= (\alpha w, \beta z), \\ g_i(w, z) &= (w + a_i, z + b_i), \quad i = 1, 2, 3. \end{aligned}$$

The group G_M acts on $\mathcal{H} \times \mathbb{C}$ freely and properly discontinuously so that the quotient $X = (\mathcal{H} \times \mathbb{C})/G_M$ is a compact complex surface.

Inoue [18] showed that, as a smooth manifold, X is a 3-torus bundle over a circle whose monodromy is given by the matrix M, and that $b_1(X) = 1$ and $b_2(X) = 0$. One can check, for example, that $H_*(X) = H_*(S^1 \times S^3)$ for all manifolds X obtained from the Cappell–Shaneson matrices A_m. Define a function $f : \mathcal{H} \times \mathbb{C} \to \mathbb{R}$ by the formula $f(w, z) = \ln w_2 / \ln \alpha$. One can easily see that df is a well-defined 1-form on X, whose cohomology class generates $H^1(X; \mathbb{Z}) = \mathbb{Z}$.

The complex surface X admits no global Kähler metric. We will however consider the following Hermitian metric on $\mathcal{H} \times \mathbb{C}$, called the Tricerri metric:

$$g = \frac{dw \otimes d\bar{w}}{w_2^2} + w_2 dz \otimes d\bar{z},$$

see [30; 12]. Let ω be the Kähler form associated with this metric, then $d\omega = d \ln w_2 \wedge \omega$, with the torsion form $d \ln w_2 = \ln \alpha \cdot df$. The metric g is G_M-invariant; hence it defines a metric on X which makes X into a locally conformal Kähler manifold (or l.c.K. manifold, for short).

The complex surface X admits a canonical spin^c structure with respect to which

$$\mathcal{S}^+ = \Lambda^{0,0}(X) \oplus \Lambda^{0,2}(X) \quad \text{and} \quad \mathcal{S}^- = \Lambda^{0,1}(X).$$

Let $\mathcal{D}^{\pm}(X)$ be the chiral Dirac operators associated with the Tricerri metric and the canonical spin^c structure on X. These are the operators that Theorem 1.1 is concerned with. The proof of Theorem 1.1 will take up the rest of these notes.

3. Reduction to the Dirac–Dolbeault operator

Let $\mathcal{D}^-(X)$ be the negative chiral Dirac operator associated with the Tricerri metric and the canonical spin^c structure on X. According to Gauduchon [13, page 283],

there is an isomorphism

$$\mathcal{D}^-(X) + \tfrac{1}{4}\ln\alpha \cdot df = \sqrt{2}(\bar\partial \oplus \bar\partial^*), \tag{1}$$

where

$$\bar\partial \oplus \bar\partial^* : \Omega^{0,1}(X) \to \Omega^{0,2}(X) \oplus \Omega^{0,0}(X) \tag{2}$$

is the Dirac–Dolbeault operator on the complex surface X. To prove Theorem 1.1, it will suffice to compute the spectral set of (2). The spectral set of $\mathcal{D}^-(X)$ will be obtained from it via multiplication by $\alpha^{-1/4}$, and the spectral set of $\mathcal{D}^+(X)$ by further inversion.

4. The periodic boundary value problem

To compute the spectral set of (2), we will complete the operator (2) to an operator $L_1^2 \to L^2$ and look for $z = e^\mu \in \mathbb{C}^*$ such that the kernel of the operator

$$e^{\mu f} \circ (\bar\partial \oplus \bar\partial^*) \circ e^{-\mu f} = (\bar\partial \oplus \bar\partial^*) - \mu \cdot df$$

on X is nonzero. Equivalently, after passing to the universal covering space $\mathcal{H} \times \mathbb{C} \to X$, we will look for μ such that the following periodic boundary problem on $\mathcal{H} \times \mathbb{C}$ has a nonzero solution $\omega \in \Omega^{0,1}(\mathcal{H} \times \mathbb{C})$:

$$(\bar\partial \oplus \bar\partial^*)(\omega) = 0, \quad \text{where } g_i^*\omega = \omega \text{ for } i = 1, 2, 3, \text{ and } g_0^*\omega = e^{-\mu} \cdot \omega.$$

Let us restate this periodic boundary problem by writing $\omega = a\,d\bar w + b\,d\bar z$ on $\mathcal{H} \times \mathbb{C}$. The equation $(\bar\partial \oplus \bar\partial^*)(\omega) = 0$ turns into the system

$$\begin{cases} \dfrac{\partial a}{\partial \bar z} - \dfrac{\partial b}{\partial \bar w} = 0, \\ \dfrac{\partial (w_2 a)}{\partial w} + \dfrac{1}{w_2^2} \cdot \dfrac{\partial b}{\partial z} = 0, \end{cases}$$

and, after introducing the new function $c = w_2 a$ and the new variable $t = \ln w_2$ into the system,

$$\left(\frac{\partial}{\partial t} + i B_t\right)\begin{pmatrix} b \\ c \end{pmatrix} = 0 \tag{3}$$

with

$$B_t = \begin{pmatrix} -e^t \dfrac{\partial}{\partial w_1} & 2\dfrac{\partial}{\partial \bar z} \\ 2e^{-t}\dfrac{\partial}{\partial z} & e^t \dfrac{\partial}{\partial w_1} \end{pmatrix}.$$

Taking into account the periodic boundary conditions $g_i^*\omega = \omega$ for $i = 1, 2, 3$, this can be viewed as a system on the product $\mathbb{R} \times T^3$, with the coordinates t on

the real line and (w_1, z_1, z_2) on the torus T^3. The remaining periodic boundary condition $g_0^*\omega = e^{-\mu}\cdot\omega$ can be expressed in the language of (0, 1)-forms as

$$g_0^*(a(w,z)d\bar{w} + b(w,z)d\bar{z}) = e^{-\mu}\cdot(a(w,z)d\bar{w} + b(w,z)d\bar{z}).$$

After switching to $c = w_2\cdot a$, this turns into

$$\bar{\beta}\cdot b(\alpha w, \beta z) = e^{-\mu}\cdot b(w,z) \quad \text{and} \quad c(\alpha w, \beta z) = e^{-\mu}\cdot c(w,z). \tag{4}$$

It is the periodic boundary value problem (3), (4) on the manifold $\mathbb{R}\times T^3$ that we now wish to solve.

5. Fourier analysis

We will use Fourier analysis on the 3-torus to solve the system (3). First, consider the following basis in $\mathbb{R}^3$:

$$\xi = (a_1,\ \operatorname{Re} b_1,\ \operatorname{Im} b_1),$$
$$\eta = (a_2,\ \operatorname{Re} b_2,\ \operatorname{Im} b_2),$$
$$\zeta = (a_3,\ \operatorname{Re} b_3,\ \operatorname{Im} b_3),$$

where $a = (a_1, a_2, a_3)$ and $b = (b_1, b_2, b_3)$ are, as before, the eigenvectors of M corresponding to the eigenvalues α and β. The quotient of $\mathbb{R}^3$ by the integer lattice spanned by the vectors ξ, η, ζ is our 3-torus. The matrix whose rows are the vectors ξ, η, ζ will be called Y so that

$$Y = \begin{pmatrix} \xi_1 & \xi_2 & \xi_3 \\ \eta_1 & \eta_2 & \eta_3 \\ \zeta_1 & \zeta_2 & \zeta_3 \end{pmatrix}.$$

Without loss of generality, we will assume that $\det Y = 1$. The columns of the matrix

$$Y^{-1} = \begin{pmatrix} \xi_1^* & \eta_1^* & \zeta_1^* \\ \xi_2^* & \eta_2^* & \zeta_2^* \\ \xi_3^* & \eta_3^* & \zeta_3^* \end{pmatrix}$$

form the dual basis ξ^*, η^*, ζ^* with respect to the usual dot product $(\cdot,\cdot)$ on $\mathbb{R}^3$. One can easily check that the functions $T^3 \to \mathbb{C}$ defined by

$$\theta \to \exp(2\pi i(\theta, k\xi^* + \ell\eta^* + m\zeta^*)) \quad \text{for all } (k,\ell,m)\in\mathbb{Z}^3, \tag{5}$$

where $\theta = (\theta_1, \theta_2, \theta_3) = (w_1, z_1, z_2)$, form an orthonormal basis in the L^2-space of complex-valued functions on the 3-torus.

For each $t\in\mathbb{R}$, expand the functions $b(t,\theta)$ and $c(t,\theta) : T^3 \to \mathbb{C}$ into Fourier series,

$$b(t,\theta) = \sum_{k,\ell,m} b_{k\ell m}(t)\exp(2\pi i(\theta, k\xi^* + \ell\eta^* + m\zeta^*))$$

and

$$c(t,\theta)=\sum_{k,\ell,m} c_{k\ell m}(t)\exp(2\pi i(\theta,k\xi^*+\ell\eta^*+m\zeta^*)),$$

and plug them into (3). For each individual triple of integers (k,ℓ,m), we obtain the system

$$\begin{pmatrix} b'_{k\ell m}\\ c'_{k\ell m}\end{pmatrix}=\begin{pmatrix} -e^t P_{k\ell m} & Q_{k\ell m}\\ e^{-t}\bar{Q}_{k\ell m} & e^t P_{k\ell m}\end{pmatrix}\begin{pmatrix} b_{k\ell m}\\ c_{k\ell m}\end{pmatrix}, \tag{6}$$

where the prime stands for the t-derivative,

$$P_{k\ell m}=2\pi(k\xi_1^*+\ell\eta_1^*+m\zeta_1^*)\in\mathbb{R},\quad \text{and}$$

$$Q_{k\ell m}=2\pi(k\xi_2^*+\ell\eta_2^*+m\zeta_2^*)+2\pi i(k\xi_3^*+\ell\eta_3^*+m\zeta_3^*)\in\mathbb{C}.$$

This is a linear system of ordinary differential equations with nonconstant coefficients. Note that $P_{k\ell m}$ and $Q_{k\ell m}$ are actually constants so the only dependence of the coefficients on t comes from the factors of e^t and e^{-t}. For future use, we make the following observation.

Lemma 5.1. *For no choice of* $(k,\ell,m)\neq(0,0,0)$ *can* $Q_{k\ell m}$ *be equal to zero.*

Proof. Observe that

$$Y\begin{pmatrix} P_{k\ell m}\\ \operatorname{Re} Q_{k\ell m}\\ \operatorname{Im} Q_{k\ell m}\end{pmatrix}=2\pi\begin{pmatrix} k\\ \ell\\ m\end{pmatrix}.$$

If $Q_{k\ell m}=0$, the first column of Y, which is an eigenvector of M with the eigenvalue α, is proportional to the vector with integral coordinates k, ℓ, and m. The latter vector is then also an eigenvector of $M\in\mathrm{SL}(3,\mathbb{Z})$ with the eigenvalue α, which contradicts the fact that α is irrational. □

Next, we need to take care of the boundary conditions (4). In our θ-notation, we have $\beta z=(\beta_1+i\beta_2)(z_1+iz_2)=(\beta_1+i\beta_2)(\theta_2+i\theta_3)=(\beta_1\theta_2-\beta_2\theta_3)+i(\beta_2\theta_2+\beta_1\theta_3)$ and $\alpha w=\alpha(w_1+iw_2)=\alpha\theta_1+ie^{t+\ln\alpha}$. To simplify notation, introduce the matrix

$$A=\begin{pmatrix}\alpha & 0 & 0\\ 0 & \beta_1 & -\beta_2\\ 0 & \beta_2 & \beta_1\end{pmatrix},$$

then the boundary conditions (4) become

$$\bar{\beta}\cdot b(t+\ln\alpha,A(\theta))=e^{-\mu}\cdot b(t,\theta),\quad c(t+\ln\alpha,A(\theta))=e^{-\mu}\cdot c(t,\theta).$$

In order to rewrite these in terms of the Fourier coefficients $b_{k\ell m}$ and $c_{k\ell m}$, we need the following technical result.

Lemma 5.2. *For any integers k, ℓ and m, we have* $(A(\theta), k\xi^* + \ell\eta^* + m\zeta^*) = (\theta, k'\xi^* + \ell'\eta^* + m'\zeta^*)$, *where*

$$\begin{pmatrix} k' \\ \ell' \\ m' \end{pmatrix} = M \begin{pmatrix} k \\ \ell \\ m \end{pmatrix}. \tag{7}$$

Proof. A straightforward calculation with matrices shows that $MY = YA^t$. Viewing θ as a column, we obtain

$$\begin{aligned}(A(\theta), k\xi^* + \ell\eta^* + m\zeta^*) &= \theta^t A^t Y^{-1} \begin{pmatrix} k \\ \ell \\ m \end{pmatrix} = \theta^t Y^{-1} M \begin{pmatrix} k \\ \ell \\ m \end{pmatrix} = \theta^t Y^{-1} \begin{pmatrix} k' \\ \ell' \\ m' \end{pmatrix} \\ &= (\theta, k'\xi^* + \ell'\eta^* + m'\zeta^*). \quad \square\end{aligned}$$

Now, substitute the Fourier expansions of $b(t, \theta)$ and $c(t, \theta)$ into the boundary conditions to obtain

$$\begin{aligned}\bar\beta \cdot b(t + \ln\alpha, A(\theta)) &= \bar\beta \sum_{k,\ell,m} b_{k\ell m}(t + \ln\alpha) \exp(2\pi i (A(\theta), k\xi^* + \ell\eta^* + m\zeta^*)) \\ &= \bar\beta \sum_{k,\ell,m} b_{k\ell m}(t + \ln\alpha) \exp(2\pi i (\theta, k'\xi^* + \ell'\eta^* + m'\zeta^*)) \\ &= e^{-\mu} \sum_{k',\ell',m'} b_{k'\ell'm'}(t) \exp(2\pi i (\theta, k'\xi^* + \ell'\eta^* + m'\zeta^*)),\end{aligned}$$

and similarly for c. A term-by-term comparison of the coefficients allows us to conclude that

$$\bar\beta \cdot b_{k\ell m}(t + \ln\alpha) = e^{-\mu} \cdot b_{k'\ell'm'}(t), \quad c_{k\ell m}(t + \ln\alpha) = e^{-\mu} \cdot c_{k'\ell'm'}(t), \tag{8}$$

where the triples (k, ℓ, m) and (k', ℓ', m') are related by (7). Therefore, to fit $b_{k\ell m}(t)$ and $c_{k\ell m}(t)$ together into a Fourier series solution, we need to know how M acts on the triples (k, ℓ, m).

6. Finite orbits

The infinite cyclic subgroup of $\mathrm{SL}(3, \mathbb{Z})$ generated by the matrix M acts on the lattice $\mathbb{Z}^3$. The only finite orbit of this action consists of the triple $(k, \ell, m) = (0, 0, 0)$. The solutions of (6) corresponding to this triple must be constant; we will denote them by b and c. The boundary conditions (8) then translate into $\bar\beta b = e^{-\mu} b$ and $c = e^{-\mu} c$, resulting in exactly two choices for the spectral point $z = e^{\mu}$ of the operator (2), namely, $z = 1$ and $z = 1/\bar\beta = \alpha\beta$. These correspond to the spectral points $z = \alpha^{1/4}$ and $z = \alpha^{1/4}\beta$ of the operator $\mathcal{D}^+(X)$ as claimed in Theorem 1.1.

7. Infinite orbits

For any fixed triple of integers $(k_0, \ell_0, m_0) \neq (0,0,0)$, the triples (k_n, ℓ_n, m_n), $n \in \mathbb{Z}$, in its orbit can be found from the equation

$$\begin{pmatrix} k_n \\ \ell_n \\ m_n \end{pmatrix} = M^n \begin{pmatrix} k_0 \\ \ell_0 \\ m_0 \end{pmatrix}.$$

Denote $b_n(t) = b_{k_n \ell_n m_n}(t)$ and $c_n(t) = c_{k_n \ell_n m_n}(t)$. It follows from (8) that, once we know $b_0(t)$ and $c_0(t)$, the rest of $b_n(t)$ and $c_n(t)$ can be determined uniquely from the recursive relation

$$b_{n+1}(t) = \bar{\beta} \cdot e^{\mu} \cdot b_n(t + \ln \alpha), \quad c_{n+1}(t) = e^{\mu} \cdot c_n(t + \ln \alpha).$$

Therefore, each infinite orbit gives rise to the infinite series

$$b(t,\theta) = \sum_{n \in \mathbb{Z}} \bar{\beta}^n \cdot e^{n\mu} \cdot b_0(t + n \ln \alpha) \cdot \exp(2\pi i (\theta, k_n \xi^* + \ell_n \eta^* + m_n \zeta^*)),$$
$$c(t,\theta) = \sum_{n \in \mathbb{Z}} e^{n\mu} \cdot c_0(t + n \ln \alpha) \cdot \exp(2\pi i (\theta, k_n \xi^* + \ell_n \eta^* + m_n \zeta^*)).$$

The question becomes whether these formal series solutions converge to a solution of (3). We will show that, for certain values of μ, the series cannot converge in L^2 norm unless $b_0(t) = c_0(t) = 0$; this will imply that the corresponding $z = e^{\mu}$ are not in the spectral set of the operator $\bar{\partial} \oplus \bar{\partial}^*$. To this end, denote by δ the real number

$$\delta = \operatorname{Re} \mu / \ln \alpha - 1/4$$

and introduce the notation

$$u(t) = b_0(t) \quad \text{and} \quad v(t) = e^{t/2} c_0(t).$$

Lemma 7.1. *The above Fourier series for $b(t,\theta)$ and $c(t,\theta)$ converge to L^2_1 sections on X if and only if both $u(t)$ and $v(t)$ belong to $L^2_{1,\delta-1/4}(\mathbb{R})$.*

Proof. Let $z = \bar{\beta} \cdot e^{\mu}$, then $z^{t/\ln \alpha} \cdot b(t,\theta)$ is the Fourier–Laplace transform [24] of the function $u(t) \exp(2\pi i (\theta, k_0 \xi^* + \ell_0 \eta^* + m_0 \zeta^*))$ on $\mathbb{R} \times T^3$ with respect to the covering translation $(t,\theta) \to (t + \ln \alpha, A(\theta))$. One can easily check that

$$|z^{1/\ln \alpha}| = e^{\delta - 1/4}.$$

From this point on, we follow the proof of [24, Proposition 4.2] and use the fact that the functions $\exp(2\pi i (\theta, k_n \xi^* + \ell_n \eta^* + m_n \zeta^*))$ form an orthonormal basis on

the fibers $\{t\} \times T^3$. For example, it follows by direct calculation that

$$\|z^{t/\ln\alpha} \cdot b(t,\theta)\|^2_{L^2(X)} = \sum_{n\in\mathbb{Z}} \int_0^{\ln\alpha} |z|^{2(n+t/\ln\alpha)} \cdot |u(t+n\ln\alpha)|^2\, dt$$
$$= \int_{-\infty}^{\infty} |z|^{2t/\ln\alpha} \cdot |u(t)|^2\, dt = \|u\|^2_{L^2_{\delta-1/4}(\mathbb{R})}.$$

The proof for the function $c(t,\theta)$ is similar. □

One can easily check using (6) that the functions $u(t)$ and $v(t)$ solve the system of ordinary differential equations

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} -Pe^t & Qe^{-t/2} \\ \bar{Q}e^{-t/2} & 1/2 + Pe^t \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \tag{9}$$

where $P = P_{k_0\ell_0 m_0} \in \mathbb{R}$ and $Q = Q_{k_0\ell_0 m_0} \in \mathbb{C}$. Because of Lemma 7.1, we are only interested in solutions $u(t)$ and $v(t)$ which belong to $L^2_{1,\delta-1/4}(\mathbb{R})$.

Proposition 7.2. *Suppose that $-1/4 \le \delta \le 1/4$, then all solutions $u(t)$, $v(t)$ of the system (9) which belong to $L^2_{1,\delta-1/4}(\mathbb{R})$ are identically zero.*

Proof. Decoupling (9) turns it into the following pair of Sturm–Liouville problems:

$$-u'' + (Pe^t(Pe^t - 1) + |Q|^2 e^{-t})u = 0 \quad \text{and} \tag{10}$$
$$-v'' + (Pe^t(Pe^t + 2) + |Q|^2 e^{-t} + 1/4)v = 0. \tag{11}$$

Without loss of generality, we will assume that u and v are real-valued functions. We will separate our argument into three cases, depending on whether P is positive, negative, or zero.

If $P < 0$, introduce the positive real numbers $p = -P$ and $q = |Q|$ and rewrite (10) in the form $-u'' + U(t)u = 0$ with the everywhere-positive potential $U(t) = pe^t(pe^t + 1) + q^2 e^{-t}$. For any choice of $a < b$, we then have

$$-\int_a^b u''(t)u(t)\, dt + \int_a^b U(t)u^2(t)\, dt = 0$$

and, after integration by parts,

$$\int_a^b u'(t)^2\, dt + u(a)u'(a) - u(b)u'(b) + \int_a^b U(t)u(t)^2\, dt = 0. \tag{12}$$

The first and the last terms in this formula are nonnegative for any choice of $a < b$. We will show that there exist a arbitrarily close to $-\infty$ and b arbitrarily close to $+\infty$ such that the other two terms in (12) are nonnegative as well. This will imply that $u(t) = 0$. Plugging $u(t) = 0$ back into (9) will then imply that $v(t) = 0$ because $Q \neq 0$ by Lemma 5.1.

We first show that for any a_0 there exists $a \le a_0$ such that $u(a)u'(a) \ge 0$. If $u(a_0) = 0$, we are finished. Otherwise, suppose that $u(t)u'(t) < 0$ for all $t \le a_0$. Then $(u^2(t))' = 2u(t)u'(t) < 0$ so that $u^2(t)$ is a decreasing function and hence $u^2(t) \ge u^2(a_0) > 0$ for all $t \le a_0$. This contradicts the fact that $u \in L^2_{\delta-1/4}(\mathbb{R})$ with $\delta \le 1/4$.

Next, we show that for any b_0 there exists $b \ge b_0$ such that $u(b)u'(b) \le 0$. If $u(b_0) = 0$ we are finished. Otherwise, suppose that $u(t)u'(t) > 0$ for all $t \ge b_0$. Then $(u^2(t))' = 2u(t)u'(t) > 0$ so that $u^2(t)$ is an increasing function and hence $u^2(t) \ge u^2(b_0) > 0$ for all $t \ge b_0$. Using the formula (12) with $a = b_0$ we obtain the estimate

$$u(b)u'(b) \ge \int_{b_0}^{b} U(t)u^2(t)\,dt \ge u^2(b_0)\int_{b_0}^{b} U(t)\,dt,$$

and using the fact that $U(t) \ge p^2e^{2t}$ for all t, the estimate

$$u(b)u'(b) \ge \tfrac{1}{2}p^2u^2(b_0)(e^{2b} - e^{2b_0}) \quad \text{for all } b \ge b_0.$$

Since $u(t)$ and $u'(t)$ belong to $L^2_{\delta-1/4}(\mathbb{R})$, it follows from the Hölder inequality that $u(t)u'(t) \in L^1_{2(\delta-1/4)}(\mathbb{R})$. This contradicts the above estimate for $\delta \ge -1/4$.

If $P > 0$, essentially the same argument using (11) shows that $v(t) = 0$. After plugging $v(t) = 0$ back in (9), we see that $u(t) = 0$ as well.

In the remaining case of $P = 0$, both (10) and (11) admit explicit solutions in terms of Bessel functions. To be precise, the general solution of (11) is of the form

$$C_1 \cdot I_1(2qe^{-t/2}) + C_2 \cdot K_1(2qe^{-t/2}), \tag{13}$$

where $I_1(x)$ and $K_1(x)$ are the modified Bessel functions of the first and second kind, solving the equation $x^2y'' + xy' - (x^2+1)y = 0$. One can check that the zero function is the only function among (13) that belongs to $L^2_{\delta-1/4}(\mathbb{R})$ with $-1/4 \le \delta \le 1/4$. □

Proposition 7.2 together with the discussion in Section 6 completes the proof of Theorem 1.1.

Acknowledgments

We thank Tom Mrowka, Leonid Parnovski, and Andrei Teleman for generously sharing their expertise, and the anonymous referee for useful comments.

References

[1] I. R. Aitchison and J. H. Rubinstein, "Fibered knots and involutions on homotopy spheres", pp. 1–74 in *Four-manifold theory* (Durham, N.H., 1982), Contemp. Math. **35**, Amer. Math. Soc., Providence, RI, 1984. MR Zbl

[2] S. Akbulut and R. Kirby, "An exotic involution of S^4", *Topology* **18**:1 (1979), 75–81. MR Zbl

[3] S. Akbulut and R. Kirby, "A potential smooth counterexample in dimension 4 to the Poincaré conjecture, the Schoenflies conjecture, and the Andrews–Curtis conjecture", *Topology* **24**:4 (1985), 375–390. MR Zbl

[4] M. Albanese, "The Yamabe invariants of Inoue surfaces, Kodaira surfaces, and their blowups", *Ann. Global Anal. Geom.* **59**:2 (2021), 179–195. MR Zbl

[5] M. F. Atiyah, V. K. Patodi, and I. M. Singer, "Spectral asymmetry and Riemannian geometry, I", *Math. Proc. Cambridge Philos. Soc.* **77** (1975), 43–69. MR Zbl

[6] W. P. Barth, K. Hulek, C. A. M. Peters, and A. Van de Ven, *Compact complex surfaces*, 2nd ed., Ergeb. Math. Grenzgeb. **4**, Springer, 2004. MR Zbl

[7] F. A. Bogomolov, "Classification of surfaces of class VII_0 with $b_2 = 0$", *Izv. Akad. Nauk SSSR Ser. Mat.* **40**:2 (1976), 273–288. MR

[8] F. A. Bogomolov, "Surfaces of class VII_0 and affine geometry", *Izv. Akad. Nauk SSSR Ser. Mat.* **46**:4 (1982), 710–761. MR

[9] S. E. Cappell and J. L. Shaneson, "Some new four-manifolds", *Ann. of Math.* (2) **104**:1 (1976), 61–72. MR Zbl

[10] S. E. Cappell and J. L. Shaneson, "There exist inequivalent knots with the same complement", *Ann. of Math.* (2) **103**:2 (1976), 349–353. MR Zbl

[11] S. Cecchini and T. Schick, "Enlargeable metrics on nonspin manifolds", *Proc. Amer. Math. Soc.* **149**:5 (2021), 2199–2211. MR Zbl

[12] S. Dragomir and L. Ornea, *Locally conformal Kähler geometry*, Progress in Mathematics **155**, Birkhäuser Boston, Boston, 1998. MR Zbl

[13] P. Gauduchon, "Hermitian connections and Dirac operators", *Boll. Un. Mat. Ital. B* (7) **11**:2 (1997), 257–288. MR Zbl

[14] R. E. Gompf, "Killing the Akbulut-Kirby 4-sphere, with relevance to the Andrews–Curtis and Schoenflies problems", *Topology* **30**:1 (1991), 97–115. MR Zbl

[15] M. Gromov and H. B. Lawson, Jr., "Spin and scalar curvature in the presence of a fundamental group, I", *Ann. of Math.* (2) **111**:2 (1980), 209–230. MR Zbl

[16] K. Hasegawa, "Complex and Kähler structures on compact solvmanifolds", pp. 749–767 in *Conference on Symplectic Topology*, vol. 3, 2005. MR Zbl

[17] T. Holt and W. Zhang, "Harmonic forms on the Kodaira–Thurston manifold", *Adv. Math.* **400** (2022), art. id. 108277. MR

[18] M. Inoue, "On surfaces of Class VII_0", *Invent. Math.* **24** (1974), 269–310. MR Zbl

[19] K. Kodaira, "On the structure of compact complex analytic surfaces, I", *Amer. J. Math.* **86** (1964), 751–798. MR Zbl

[20] J. Lin, D. Ruberman, and N. Saveliev, "On the Frøyshov invariant and monopole Lefschetz number", preprint, 2018. To appear in *J. Differential Geom.* arXiv 1802.07704

[21] J. Lin, D. Ruberman, and N. Saveliev, "A splitting theorem for the Seiberg–Witten invariant of a homology $S^1 \times S^3$", *Geom. Topol.* **22**:5 (2018), 2865–2942. MR Zbl

[22] J. Lin, D. Ruberman, and N. Saveliev, "On the monopole Lefschetz number of finite-order diffeomorphisms", *Geom. Topol.* **25**:7 (2021), 3591–3628. MR Zbl

[23] P. Lupascu, "The Seiberg–Witten equations on Hermitian surfaces", *Math. Nachr.* **242** (2002), 132–147. MR Zbl

[24] T. Mrowka, D. Ruberman, and N. Saveliev, "Seiberg–Witten equations, end-periodic Dirac operators, and a lift of Rohlin's invariant", *J. Differential Geom.* **88**:2 (2011), 333–377. MR Zbl

[25] T. Mrowka, D. Ruberman, and N. Saveliev, "An index theorem for end-periodic operators", *Compos. Math.* **152**:2 (2016), 399–444. MR Zbl

[26] C. Okonek and A. Teleman, "Seiberg–Witten invariants for 4-manifolds with $b_+ = 0$", pp. 347–357 in *Complex analysis and algebraic geometry*, de Gruyter, Berlin, 2000. MR Zbl

[27] C. H. Taubes, "Gauge theory on asymptotically periodic 4-manifolds", *J. Differential Geom.* **25**:3 (1987), 363–430. MR Zbl

[28] A. D. Teleman, "Projectively flat surfaces and Bogomolov's theorem on class VII_0 surfaces", *Internat. J. Math.* **5**:2 (1994), 253–264. MR Zbl

[29] W. P. Thurston, "Some simple examples of symplectic manifolds", *Proc. Amer. Math. Soc.* **55**:2 (1976), 467–468. MR Zbl

[30] F. Tricerri, "Some examples of locally conformal Kähler manifolds", *Rend. Sem. Mat. Univ. Politec. Torino* **40**:1 (1982), 81–92. MR Zbl

[31] C. T. C. Wall, "Geometries and geometric structures in real dimension 4 and complex dimension 2", pp. 268–292 in *Geometry and Topology* (College Park, Md., 1983/84), Lecture Notes in Math. **1167**, Springer, 1985. MR Zbl

[32] C. T. C. Wall, "Geometric structures on compact complex analytic surfaces", *Topology* **25**:2 (1986), 119–153. MR Zbl

Received 23 Jan 2021. Revised 5 May 2021.

DANIEL RUBERMAN: ruberman@brandeis.edu
Department of Mathematics, Brandeis University, Waltham, MA, United States

NIKOLAI SAVELIEV: saveliev@math.miami.edu
Department of Mathematics, University of Miami, Coral Gables, FL, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.299

A note on thickness of knots

András I. Stipsicz and Zoltán Szabó

We introduce a numerical invariant $\beta(K) \in \mathbb{N} \cup \{0\}$ of a knot $K \subset S^3$ which measures how nonalternating K is. We prove an inequality between $\beta(K)$ and the (knot Floer) thickness $th(K)$ of a knot K. As an application we show that all Montesinos knots have thickness at most one.

1. Introduction

A knot $K \subset S^3$ is *alternating* if it admits a diagram with the property that when traversing through the diagram, we alternate between over- and under-crossings. (An intrinsic definition of alternating knots has been recently found by Greene and Howie [5; 6].) A diagram of K partitions the plane into domains (the connected components of the complement of the projection), and the alternating property can be rephrased by saying that on the boundary of each domain each edge connects an under-crossing with an over-crossing. Indeed, this observation provides a way to measure how far a knot is from being alternating. We introduce the following definition:

Definition 1.1. Suppose that D is the diagram of a given knot $K \subset S^3$. A domain d of D is *good* if any edge on the boundary of d connects an over- and an under-crossing. The domain d is *bad* if it is not good. The number of bad domains of the diagram D is denoted by $B(D)$.

Clearly, the diagram D is alternating if and only if $B(D) = 0$. Indeed, by taking

$$\beta(K) = \min\{B(D) \mid D \text{ is a diagram for } K\},$$

we get a knot invariant, which satisfies $\beta(K) = 0$ if and only if K is an alternating knot. As it is typical for knot invariants given by minima of quantities over all diagrams, it is easy to find an upper bound on $\beta(K)$ (by determining $B(D)$ for a diagram of K), but it is harder to actually compute its value.

MSC2020: 57K10.
Keywords: knot Floer homology, thickness, alternating knots.

As it turns out, knot Floer homology provides a lower bound for $\beta(K)$ through the *thickness* of K. Recall that $\widehat{\mathrm{HFK}}(K)$, the hat-version of knot Floer homology of K, is a finite-dimensional bigraded vector space over the field $\mathbb{F}$ of two elements. By collapsing the Maslov and Alexander gradings M and A on $\widehat{\mathrm{HFK}}(K)$ to $\delta = A - M$, we get a graded vector space $\widehat{\mathrm{HFK}}^{\delta}(K)$. The thickness $th(K)$ of K is the largest possible difference of δ-gradings of two homogeneous (nonzero) elements of this vector space. It is known that for an alternating knot K the δ-graded Floer homology is in a single δ-grading (determined by the signature of the knot); hence if K is alternating, then $th(K) = 0$. (Knots satisfying $th(K) = 0$ are called *thin* knots, hence alternating knots are thin.)

With this definition in place, the main result of this paper is as follows:

Theorem 1.2. *Suppose that $K \subset S^3$ is a nonalternating knot. Then*

$$th(K) \leq \tfrac{1}{2}\beta(K) - 1. \tag{1-1}$$

While the thickness of K can be used to estimate how nonalternating K is, (1-1) can also be used to estimate $th(K)$ by finding appropriate diagrams of K. In particular, the formula can be applied to show the following:

Corollary 1.3 (Lowrance [7]). *Suppose K is a Montesinos knot. Then, $th(K) \leq 1$.*

Remark 1.4. • A quantity similar to $\beta(K)$ has been introduced by Turaev [10], now called the *Turaev genus* $g_T(K)$. An inequality similar to (1-1) for the Turaev genus and the (knot Floer) thickness $th(K)$ was shown by Lowrence in [7]. As the Turaev genus of nonalternating Montesinos knots is known to be equal to 1 [1; 2], our Corollary 1.3 also follows from [7].

• Indeed, a simple argument (due to Adam Lowrence (personal communication, 2020)) shows that

$$g_T(K) \leq \tfrac{1}{2}\beta(K) - 1,$$

since by [7, Theorem 4.1] for a diagram D of K we have $g_T(D) = th(C^{\delta}_{D,p})$ (with the notation of Section 2).

• Similar observations regarding the relation between the Turaev genus g_T and β have been communicated to us by Homayun Karimi and Seungwon Kim (2020).

The formula (1-1) can be used in a further way: by a recent result of Zibrowius [11], mutation does not change $\widehat{\mathrm{HFK}}^{\delta}(K)$, and hence leaves $th(K)$ unchanged. Consequently, besides isotopies, we can change a diagram by mutations to get better estimates for $th(K)$ through $B(D)$ for a diagram D of a mutant.

The paper is organized as follows. In Section 2, we recall basics of knot Floer homology and prove the theorem stated above. In Section 3, we give the details of the proof of Corollary 1.3, and finally in Section 4, we list some further properties and questions regarding β.

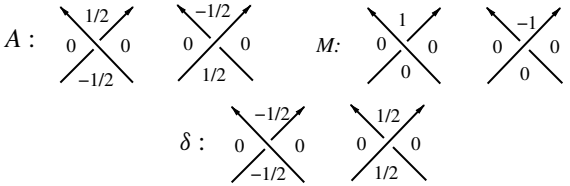


Figure 1. The local contributions for A, M and $\delta = A - M$ at a crossing. The Kauffman state distinguishes a corner at the crossing, and we take the value in that corner as a contribution of the crossing to A, M or δ of the Kauffman state at hand.

2. The knot Floer homology thickness of knots

Suppose that $V = \sum_a V_a$ is a finite-dimensional graded vector space, where $V_a \subset V$ is the subspace of homogeneous elements of grading $a \in \mathbb{R}$. The *thickness* $th(V)$ of V is by definition the largest possible difference between gradings of (nonzero) homogeneous elements:

$$th(V) = \max\{a \in \mathbb{R} \mid V_a \neq 0\} - \min\{a \in \mathbb{R} \mid V_a \neq 0\}.$$

Suppose now that the graded vector space V is endowed with a boundary operator ∂ of degree 1; then the homology $H(V, \partial)$ also admits a natural grading from the grading of V. As $H(V, \partial)$ is the quotient of a subspace of V, it is easy to see that

$$th(H(V, \partial)) \leq th(V).$$

The hat version of knot Floer homology (over the field $\mathbb{F}$ of two elements) of a knot $K \subset S^3$ is a finite-dimensional bigraded vector space $\widehat{\mathrm{HFK}}(K) = \sum_{M,A} \widehat{\mathrm{HFK}}_M(K, A)$. By collapsing the two gradings to $\delta = A - M$, we get the δ-graded invariant $\widehat{\mathrm{HFK}}^\delta(K)$. The thickness of $\widehat{\mathrm{HFK}}^\delta(K)$ is by definition the thickness $th(K)$ of K.

Knot Floer homology is defined as the homology of a chain complex, which we can associate to a diagram of the knot (and some further choices). Indeed, for a given diagram D of a knot K, fix a marking, that is, a point of D which is not a crossing. Consider the bigraded vector space $C_{D,p}$ (graded by the Alexander and the Maslov gradings A and M) associated to the marked diagram (D, p), which is generated over $\mathbb{F}$ by the Kauffman states of the marked diagram, a concept which we recall below.

Suppose that for the marked diagram (D, p) of the knot K, the set of crossings is denoted by $Cr(D)$, the set of domains by $Dom(D)$, and $Dom_p(D)$ denotes the set of those domains which do not contain p on their boundary. A *Kauffman state* κ is a bijection $\kappa : Cr(D) \to Dom_p(D)$ with the property that for a crossing $c \in Cr(D)$ the value $\kappa(c)$ is one of the (at most four) domains meeting at c. The Alexander, Maslov and δ-gradings of a Kauffman state are computed by summing the local contributions at each crossing, as given by the diagrams of Figure 1.

According to [8] there is a boundary map $\partial : C_{D,p} \to C_{D,p}$ of bidegree $(-1, 0)$ (in the bigrading (M, A)) with the property that $H(C_{D,p}, \partial)$ is isomorphic to the knot Floer homology $\widehat{\mathrm{HFK}}(K)$ of K (as a bigraded vector space). By collapsing the two gradings A and M to $\delta = A - M$, we get the graded vector spaces $(C^{\delta}_{D,p}, \partial)$ and its homology $\widehat{\mathrm{HFK}}^{\delta}(K)$. As $\widehat{\mathrm{HFK}}^{\delta}(K)$ is the quotient of a subspace of $C^{\delta}_{D,P}$, we have that

$$th(\widehat{\mathrm{HFK}}^{\delta}(K)) \leq th(C^{\delta}_{D,p}, \partial).$$

Proposition 2.1. *Suppose that D is a diagram of the knot K. If D is not an alternating diagram, then*

$$th(C^{\delta}_{D,p}) \leq \tfrac{1}{2}B(D) - 1.$$

Proof. Fix a marked point p on D, and consider the δ-graded chain complex $(C^{\delta}_{D,p}, \partial)$ generated by the Kauffman states of (D, p).

The δ-grading at a positive crossing is either 0 or $\frac{1}{2}$, and at a negative crossing it is either 0 or $-\frac{1}{2}$. So we can express the δ-grading of a Kauffman state κ as the sum

$$\tfrac{1}{4}\mathrm{wr}(D) + \sum_{c \in Cr} f(\kappa(c)),$$

where wr is the writhe of the diagram, and f is a function on the Kauffman corners, which is either $\frac{1}{4}$ or $-\frac{1}{4}$ (depending on the chosen quadrant at the crossing c).

Simple computation shows that for a good domain each corner in the domain gives the same f-value; hence for different Kauffman states the contributions from this particular domain are the same. This is no longer true for a bad domain, but the difference of two contributions is at most $\frac{1}{2}$. When determining the possible maximum of $\delta(x) - \delta(x')$ for two homogeneous elements $x, x' \in C^{\delta}_{D,p}$, the contributions from the writhe cancel, and so do the contributions from good domains, while bad domains contribute at most $\frac{1}{2}$. This shows that $th(C^{\delta}_{D,p}) \leq \frac{1}{2}B(D)$.

By assumption, D is not alternating; hence there is a bad domain, with an edge showing that it is bad. Choose the marking p on such an edge. Since this edge guarantees that the two domains having it on their boundary are both bad, while these two bad domains do not get Kauffman corners, we get that $th(C_{D,p})$ is bounded by

$$\tfrac{1}{2}(B(D) - 2) = \tfrac{1}{2}B(D) - 1,$$

concluding the proof. □

Proof of Theorem 1.2. Suppose that K is not alternating. Then any diagram D of K is nonalternating; hence we have that

$$th(K) \leq th(C^{\delta}_{D,p}) \leq \tfrac{1}{2}B(D) - 1.$$

Since $\beta(K)$ is computed from the minimum of the right-hand side of this inequality, the proof follows at once. □

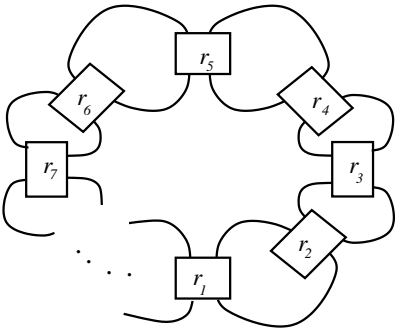


Figure 2. The Montesinos knot $M(r_1, \ldots, r_n)$. The box containing r_i denotes the algebraic tangle determined by the rational number $r_i = \beta_i/\alpha_i$ (cf. Figure 3). In order to have a knot, at most one of the α_i can be even.

3. Montesinos knots

Montesinos knots are straightforward generalizations of pretzel knots; a diagram involving rational tangles defining the Montesinos knot $M(r_1, \ldots, r_n)$ is shown by Figure 2. (A box with a rational number r_i in it symbolizes the tangle shown by Figure 3.) We allow any of the r_i to be equal to ± 1. Notice that the order of $(r_1, \ldots, r_n)$ is important; those r_i which are equal to ± 1 can be commuted with any other parameter through a simple isotopy of the diagram.

Lemma 3.1. *Consider the diagram of the Montesinos knot $M(r_1, \ldots, r_n)$ given by Figure 2. It can be isotoped to a diagram with at most four bad domains.*

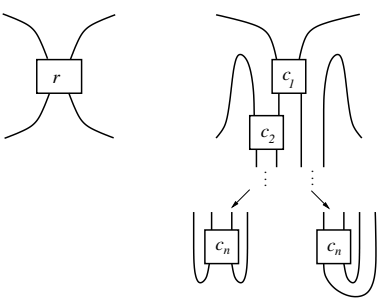


Figure 3. The rational tangle corresponding to $r \in \mathbb{Q}$. The rational number r determines the coefficients c_i through its continued fraction expansion. The boxes with $c_i \in \mathbb{Z}$ on the right denote $|c_i|$ half twists (right-handed for positive, left-handed for negative c_i). Depending on the parity of n (the number of c_i's) we have two different finishing forms. The tangle is alternating (as part of a knot or link) if the c_i alternate in sign.

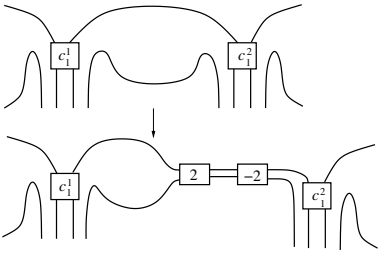


Figure 4. The introduction of cancelling twists to turn domains between tangles to be good.

Proof. Recall that a rational tangle has the form given by Figure 3. Adapting the isotopies described in [4], we can achieve that all tangles are alternating; hence the potentially bad domains are the ones between the tangles, together with the central and the unbounded domains. The number of bad domains between the tangles can be reduced by the following observation. The domain between two tangles is bad if the first coefficients c_1^1 and c_1^2 of the two rational numbers determining the tangles have opposite signs, say $c_1^1 > 0$ and $c_1^2 < 0$. Then by Reidemeister-II moves we can introduce canceling twistings, as shown by Figure 4, and then commute the first twisting (in the figure given by the box with 2 in it) between the first and second tangles of the Montesinos knot. All domains between the boxes will become good, except the ones connecting the first tangle with the newly introduced twists and the second tangle also connecting it with the newly introduced twists. After these alterations, make sure that (by the adaptation of [4]) all tangles are isotoped to be alternating. In total the new diagram then has four bad domains, concluding the proof. □

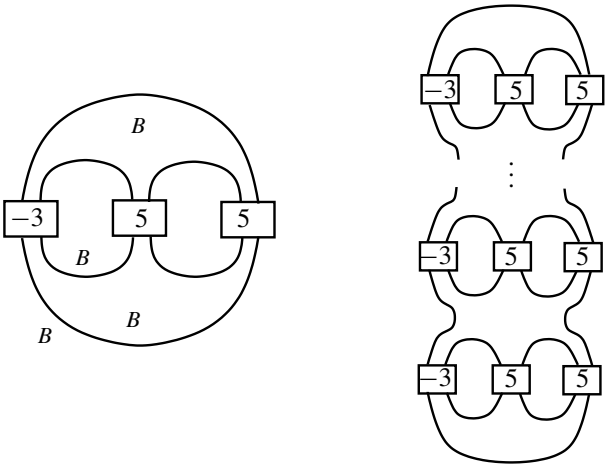
Proof of Corollary 1.3. For a Montesinos knot $M(r_1, \dots, r_n)$, an appropriate isotopy of the diagram of Figure 2 (as given by Lemma 3.1) gives a diagram with at most four bad domains. The application of Theorem 1.2 concludes the argument. □

Remark 3.2. Using the mutation invariance of $th(K)$, Lemma 3.1 can be avoided: by mutations, any Montesinos knot $M(r_1, \dots, r_n)$ can be moved to $M(q_1, \dots, q_n)$ with the same rational parameters in a different order so that q_i and q_{i+1} have the same sign with at most one exception. Isotoping the diagram so that the tangles are alternating, the mutated diagram then has at most 4 bad domains. Using the result of [11, Theorem 0.1] then the corollary follows as before.

4. Further properties

It is a standard fact that the knot Floer homology of the connected sum of two knots is the tensor product of the knot Floer homologies:

$$\widehat{\mathrm{HFK}}(K_1 \# K_2) \cong \widehat{\mathrm{HFK}}(K_1) \otimes \widehat{\mathrm{HFK}}(K_2).$$



(a) (b)

Figure 5. The knot K_n. In (a) the pretzel knot $P(-3, 5, 5)$ is shown. The B symbols signify the bad domains. (A box containing the integer n denotes $|n|$ half twists, right-handed for $n > 0$ and left-handed for $n < 0$.) In (b) we provide a diagram of K_n, where the connected sum is taken at bad domains.

From this (bigraded) isomorphism it follows that

$$th(K_1\#K_2) = th(K_1) + th(K_2).$$

The behaviour of $\beta(K)$ is less clear under connected summing. Suppose that K_1, K_2 are both nonalternating knots. By taking the connected sum of two diagrams D_1, D_2 for these knots at bad edges (i.e., arcs on the boundary of bad domains verifying that the domains are bad), we get that

$$B(D_1\#D_2) = B(D_1) + B(D_2) - 2,$$

immediately implying that

$$\beta(K_1\#K_2) \leq \beta(K_1) + \beta(K_2) - 2.$$

Motivated by the equality for the thickness th, we arrive at the following conjecture:

Conjecture 4.1. If K_1, K_2 are two nonalternating knots, then

$$\beta(K_1\#K_2) = \beta(K_1) + \beta(K_2) - 2.$$

Sharpness. It is not hard to find knot diagrams for which (1-1) is sharp. Indeed, the standard diagram of the pretzel knot $P(-3, 5, 5)$ admits four bad domains (see Figure 5(a)), while an explicit calculation of $\widehat{\mathrm{HFK}}(P(-3, 5, 5))$ shows

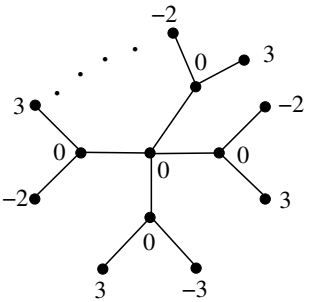


Figure 6. The planar weighted tree defining the knot C_n. The central vertex is of framing 0, and has $n+1$ neighbours, all with framing 0. The first of these vertices is connected to two leaves of framings 3 and -3, while the further vertices are connected to two leaves with framings 3 and -2.

that $th(P(-3,5,5))=1$. Consider the n-fold connected sum $K_n=\#_n P(-3,5,5)$; connect summing the diagrams at bad edges (in the above sense) we get a sequence of knots K_n and diagrams D_n for them with the properties that $th(K_n)=n$ and $B(D_n)=2n+2$; see Figure 5(b). The nonalternating knots K_n then satisfy $n=th(K_n)=\frac{1}{2}\beta(K_n)-1$.

Arborescent examples. A family of knots (and links) can be specified by combinatorial means as follows. Consider a planar tree (a graph with no circles), with an integer attached to each vertex. An embedded surface can be constructed from the tree by the following algorithm: for each vertex consider a twisted band, with the integer attached to the vertex prescribing the number of half-twists introduced. (The boundary of such a band is the $T_{2,n}$ torus knot or link, where $n\in\mathbb{Z}$ is the decoration of the vertex.) If two vertices are connected in the tree by an edge, plumb the two surfaces together. The boundary of the resulting surface is an arborescent knot (or link). To make the definition precise (i.e., to get a well-defined knot or link) further information is needed, prescribing the location of the plumbing on each band, relative to the twisting; see [3]. We will not make this distinction here for two reasons: (a) the different choices one can make for a given graph result in mutation equivalent knots, and since the thickness is mutation invariant, different choices make no effect on our calculations, and (b) in the example we will show below, the nodes (i.e., vertices of degree more than 2) have framing 0, hence the above mentioned choice makes no difference.

It is easy to see that pretzel knots (and more generally Montesinos knots) are all arborescent; these knots correspond to graphs with a single node. (Such graphs are called star-shaped.) Consider the family of knots C_n defined by the graph of Figure 6. (For diagrams of the knots C_n, see Figure 7.) Computer calculations [9]

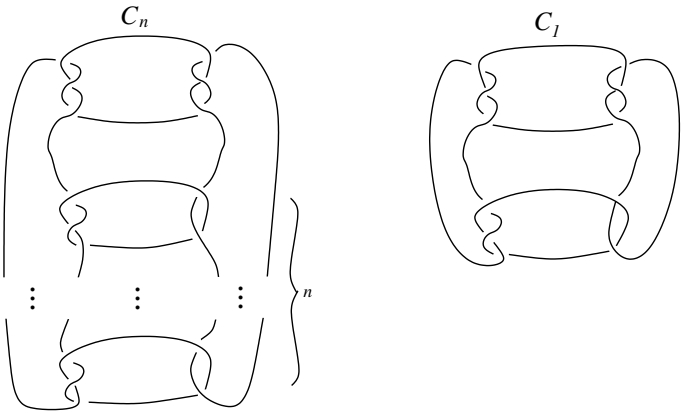


Figure 7. The knot C_n in general and C_1 in particular.

show that C_n has thickness n once $n \leq 4$. (For $n = 1$ the knot C_1 is a pretzel knot having thickness equal to 1, while for $n = 0$ the knot C_0 is the connected sum of a right-handed and a left-handed trefoil; hence it is thin.) These cases lead us to expect that $th(C_n) = n$ holds in general. Indeed, it is not hard to find a diagram for C_n with $2n + 2$ bad domains; hence $th(C_n) \leq n$ follows from our main result, and the above mentioned calculations suggest that we have equality here. More generally, it would be interesting to see if there is a simple relation between the number of nodes of a (weighted) tree and the thickness of a corresponding knot; maybe the thickness is at most the number of nodes.

Acknowledgements

A. Stipsicz was partially supported by the Élvonal (Frontier) project of the NKFIH (KKP126683). Z. Szabó was partially supported by NSF Grants DMS-1606571 and DMS-1904628. We would like to thank Jen Hom and Tye Lidman for a motivating discussion.

References

[1] A. Champanerkar and I. Kofman, "A survey on the Turaev genus of knots", *Acta Math. Vietnam.* **39**:4 (2014), 497–514. MR Zbl

[2] O. T. Dasbach, D. Futer, E. Kalfagianni, X.-S. Lin, and N. W. Stoltzfus, "The Jones polynomial and graphs on surfaces", *J. Combin. Theory Ser. B* **98**:2 (2008), 384–399. MR Zbl

[3] D. Gabai, *Genera of the arborescent links*, Mem. Amer. Math. Soc. **339**, 1986. MR Zbl

[4] J. R. Goldman and L. H. Kauffman, "Rational tangles", *Adv. in Appl. Math.* **18**:3 (1997), 300–332. MR Zbl

[5] J. E. Greene, "Alternating links and definite surfaces", *Duke Math. J.* **166**:11 (2017), 2133–2151. MR Zbl

[6] J. A. Howie, "A characterisation of alternating knot exteriors", *Geom. Topol.* **21**:4 (2017), 2353–2371. MR Zbl

[7] A. M. Lowrance, "On knot Floer width and Turaev genus", *Algebr. Geom. Topol.* **8**:2 (2008), 1141–1162. MR Zbl

[8] P. Ozsváth and Z. Szabó, "Heegaard Floer homology and alternating knots", *Geom. Topol.* **7** (2003), 225–254. MR

[9] Z. Szabó, "Knot Floer homology calculator", computer program, 2020, available at https://web.math.princeton.edu/~szabo/HFKcalc.html.

[10] V. G. Turaev, "A simple proof of the Murasugi and Kauffman theorems on alternating links", *Enseign. Math.* (2) **33**:3-4 (1987), 203–225. MR Zbl

[11] C. Zibrowius, "On symmetries of peculiar modules; or, δ-graded link Floer homology is mutation invariant", preprint, 2019. arXiv 1909.04267

Received 17 Jan 2021. Revised 22 Feb 2021.

ANDRÁS I. STIPSICZ: stipsicz.andras@renyi.hu
Rényi Institute of Mathematics, Budapest, Hungary

ZOLTÁN SZABÓ: szabo@math.princeton.edu
Department of Mathematics, Princeton University, Princeton, NJ, United States

THE OPEN BOOK SERIES 5 (2022)
Gauge theory and low-dimensional topology: progress and interaction
https://doi.org/10.2140/obs.2022.5.309

Morse foliated open books and right-veering monodromies

Vera Vértesi and Joan E. Licata

Morse foliated open books were introduced by the authors (arXiv 2002.01752v1), along with abstract and embedded versions, as a tool for studying contact manifolds with boundary. This article illustrates the advantages of the Morse perspective. We use this to extend the definition of *right-veering* to foliated open books and we show that it plays a similar role in detecting overtwistedness as in other versions of open books.

1. Introduction

Three flavors of foliated open books were introduced in [11], each a topological decomposition of a manifold with boundary which determines an equivalence class of contact structures on the ambient manifold. Embedded foliated open books provide an intuitive construction: cut a traditional open book along a generic separating surface and the result is a pair of embedded foliated open books. Abstract foliated open books were explored further in [1], where they were used to defined a contact invariant in bordered sutured Floer homology. In this article we turn attention to Morse foliated open books, illustrating the benefits of this perspective by extending the established notion of a right-veering monodromy to the open book setting.

We also admit to a fondness for Morse foliated open books. There are technical advantages to having three versions to select from, but when foliated open books existed only as chalked pictures on a board, their intrinsic data always included a circle-valued Morse function. It is therefore a pleasure to return to this approach now. For a topologist, one of the beautiful applications of Morse theory is the metamorphosis it facilitates from differential geometry to geometric topology, turning a smooth manifold into a handle structure. Open books of all flavours

MSC2020: 57K33.
Keywords: contact structure, partial open book, open book foliation, Morse function, monodromy, right-veering, tight, overtwisted.

serve a similar purpose in contact geometry, encoding a non-integrable plane field via a topological decomposition. Here, we equip the complement of a properly embedded one-manifold with an S^1-Morse function, all of whose critical points lie on the boundary. As in standard Morse theory, such a critical point corresponds to a change in the topology of the level sets of the Morse function, but the level sets are now interpreted as pages of an open book decomposition. We may thus see an unlimited number of topologically distinct page types, but the transition between any two page types is tightly controlled. Furthermore, a catalog of these changes is recorded on the boundary, where each critical point of the original Morse function is a critical point of its restriction to the boundary. As previously shown, this boundary data alone determines a family of compatible Morse functions on the original contact manifold, so that a relatively small amount of data captures a broad and flexible set of decompositions. Precise definitions and key theorems are recalled in the next section, and we briefly outline some dividends from this perspective.

The step from smooth manifold to handle decomposition requires a choice of gradient-like vector field, and such a choice is similarly useful in the case of Morse foliated open books. We choose a particular class of gradient-like vector fields characterized by their flow on the restriction to the boundary of the manifold. With this class fixed, we define the monodromy of a Morse foliated open book as the first return map of this flow relative to a fixed page. This is defined only on a subsurface of the page, just as in the case of the partial open books defined by Honda–Kazez–Matić, and in some cases our notion of monodromy can be directly identified with their notion [7]. However, the fact that foliated open books admit a quite flexible notion of pages leads to Morse foliated open books which may not immediately be interpreted as partial open books. In this case, the monodromy of a foliated open book is a strict generalization of the partial open book monodromy, and Section 3 explores the consonance of the two versions through a family of examples. We also show — perhaps unsurprisingly — that veering monodromies play a similar role in foliated open books as in their classical and partial counterparts.

The study of right-veering monodromies initiated by Goodman and developed by Honda–Kazez–Matić relates the tightness of a contact structure to a measure of positivity in its supporting open books [5; 8; 7]. The definition of right-veering extends verbatim to foliated open book monodromies, and with similar consequences: a contact manifold is tight if and only if all of its supporting Morse foliated open books have right-veering monodromy. In the case that a Morse foliated open book admits a left-veering arc, one may construct an overtwisted disc as in [10].

Acknowledgements. We are grateful to BIRS for hosting the workshop *Interactions of gauge theory with contact and symplectic topology in dimensions* 3 *and* 4. We'd also like to thank the referee for helpful feedback.

2. Foliated open books

Definition 2.1. A (*Morse*) *foliated open book* for a three-manifold with boundary $(M, \partial M)$ consists of (B, π), where B is an oriented, properly embedded 1-manifold and $\pi : M \setminus B \to S^1$ satisfies the following properties:

(1) π is an S^1-valued Morse function whose critical points all lie on ∂M.

(2) $\widetilde{\pi} := \pi|_{\partial M}$ is Morse with the same set of critical points as π.

(3) π has a unique critical point for each critical value.

One of the first applications of classical Morse theory is to relate the topology of sublevel sets to the critical points of the Morse function. Because the present function π has only boundary critical points, the critical points instead detect changes in the topology of the level sets of π. Equivalently, the critical points of $\widetilde{\pi}$ detect changes in the topology of the level sets of π, and we record this on ∂M. Specifically, let $\mathcal{F}_{\widetilde{\pi}}$ be the singular foliation whose leaves are the level sets of $\widetilde{\pi}$, oriented as the boundary of the level sets of π, and each singular point comes with a sign: elliptic points are distinguished as positive sources and negative sinks, while (four-pronged) hyperbolic points are distinguished by the index of the critical point of π. An index two critical point of π gives rise to a positive hyperbolic point of the singular foliation and corresponds to cutting a level set of π along an arc. An index one critical point of π gives rise to a negative hyperbolic point of the singular foliation and corresponds to adding a one-handle to a level set of π. See Figure 1.

Definition 2.2. If (B, π) is a foliated open book on $(M, \partial M)$, then $\mathcal{F}_{\widetilde{\pi}}$ is the associated *open book foliation* on ∂M.

By a slight abuse of the notation we call any (singular) foliation that comes from this construction an open book foliation, sometimes without a reference to the enclosing open book and manifold.

Open book foliations were first studied by Pavelescu [12] and by Ito–Kawamuro

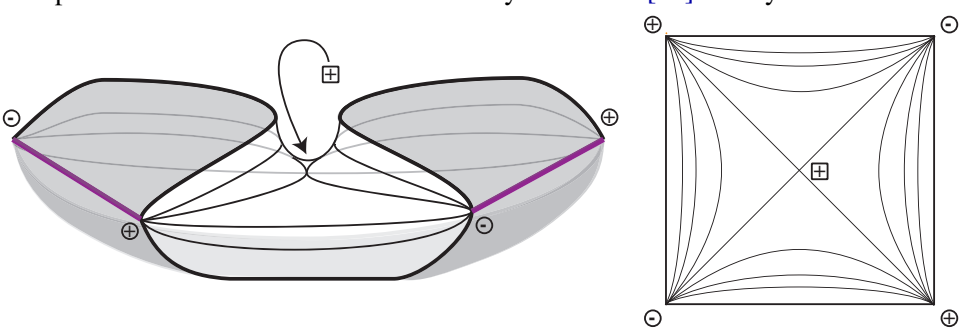


Figure 1. Left: a local cobordism between level sets near an index two boundary critical point of π. Right: the boundary foliation $\mathcal{F}_{\widetilde{\pi}}$ near the positive hyperbolic point.

[9] as a generalization of the braid foliations introduced by Birman–Menasco [4] that in turn rest on the ideas of Bennequin [3]. Our definition differs slightly from that of Ito–Kawamuro in the requirement that the foliation is by level sets of a Morse function, but this may be imposed in their setting without penalty.

When $\mathcal{F}_{\widetilde{\pi}}$ has no circle leaves, it admits a *dividing set*, a curve Γ defined up to isotopy as the boundary of a neighborhood of the positive separatrices from positive hyperbolic points. If there is a (not necessarily smooth) isotopy of the surface taking one signed singular foliation to another through a path of foliations with the same cyclic order on hyperbolic points, we say that the two foliations are *strongly topologically conjugate*. This allows us to state the relationship between foliated open books and contact structures:

Definition 2.3. [11, Definition 3.7] The Morse foliated open book (B, π) *supports* the contact structure ξ on $(M, \partial M)$ if ξ is the kernel of some one-form α on M satisfying the following properties:

(1) $\alpha(TB) > 0$.

(2) $d\alpha|_{\pi^{-1}(t)}$ is an area form for all t.

(3) $\mathcal{F}$ is strongly topologically conjugate to the characteristic foliation of ξ.

Note that condition (3) imposes that $\mathcal{F}_{\widetilde{\pi}}$ has no circle leaves.

Definition 2.4. A *foliated contact three-manifold* $(M, \xi, \mathcal{F})$ is a manifold with foliated boundary together with a contact structure ξ on M such that $\mathcal{F}$ is an open book foliation that is strongly topologically conjugate to $\mathcal{F}_{\xi}$.

Theorem 2.5. [11, Theorems 3.10, 6.9, 7.1, 7.2] *Any foliated open book supports a unique isotopy class of contact structures, and any foliated contact three-manifold $(M, \xi, \mathcal{F})$ admits a supporting foliated open book. Two foliated open books for the same foliated contact three-manifold are related by positive stabilization.*

Although we have not yet defined positive stabilization here, we note that it is an internal operation analogous to stabilization on other forms of open books.

2.1. ***Preferred gradient-like vector fields and the monodromy.*** As in the case of standard Morse theory, the benefit of a Morse function is fully realized only in the presence of a gradient-like vector field. We will designate a class of gradient-like vector fields as *preferred* based on their compatibility with the open book foliation on ∂M. Suppose that $\mathcal{F}_{\widetilde{\pi}}$ is the open book foliation on a manifold M supporting some contact structure with convex boundary. As $\mathcal{F}_{\widetilde{\pi}}$ has no circle leaves, it decomposes the surface as a union of square tiles, each of which contains a single hyperbolic singularity in the middle, four elliptic singularities at the corners, and four connected components of leaves as edges. In fact, one may recover the entire open book foliation up to topological equivalence by decomposing a surface into

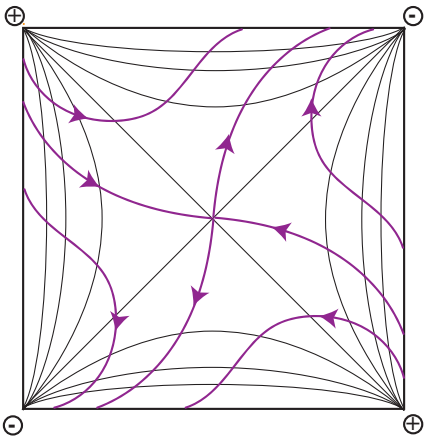


Figure 2. Flowlines of a preferred gradient-like vector field on the tile defined by a hyperbolic point. The hyperbolic point could be positive or negative.

squares labeled with signs, signed corners, and an order in which the hyperbolic singularities appear. See Figure 2 for an illustration of a single tile and Figure 10 for an example of a surface decomposed into squares.

Definition 2.6. A gradient-like vector field $\nabla\pi$ is *preferred* if it is tangent to ∂M on ∂M and if the flowlines on each tile of the foliated surface ∂M are isotopic to those shown in Figure 2.

We will assume henceforth that $\nabla\pi$ is always preferred.

Said slightly differently, $\nabla\pi$ is preferred if it is the extension to M of some $\nabla\widetilde{\pi}$ with prescribed properties on ∂M. The condition on flowlines is chosen so that topological properties of the original foliation are reflected in topological properties of the flowlines of $\nabla\widetilde{\pi}$. Specifically, we may consider the graph formed by the positive separatrices of positive hyperbolic points, just as above in the definition of Γ. The flowlines of the gradient-like vector field spiral around the elliptic points

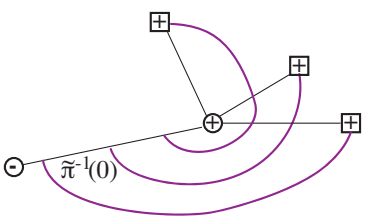


Figure 3. The squares represent positive hyperbolic points with a positive separatrix terminating at the shown positive elliptic point (circle). The separatrices of $\nabla\widetilde{\pi}$ intersect $\widetilde{\pi}^{-1}(0)$ in the same order that the separatrices of $\mathcal{F}$ intersect the elliptic point.

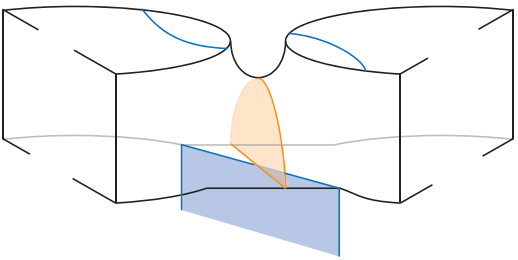


Figure 4. The lightly shaded descending critical submanifold intersects the darker-shaded ascending critical submanifold, cutting the intersection curve on all subsequent pages.

infinitely many times, but we may nevertheless define a similar graph: replace each positive elliptic point by the connected component of $\widetilde{\pi}^{-1}(0)$ which terminates at it, and let the graph be this component together with the union of positive flowlines of $\nabla\widetilde{\pi}$ from positive hyperbolic points truncated when they hit these intervals first. See Figure 3. After a deformation retract of the leaf to its positive endpoint, these two graphs are isotopic. The analogous statement holds for the analogously defined graph constructed from negative separatrices of negative elliptic points.

We now assume every foliated open book is equipped with a preferred gradient-like vector field, and we write $(B, \pi, \mathcal{F}_{\widetilde{\pi}}, \nabla\pi)$ to denote the additional data described above. Choosing a gradient-like vector field has the immediate consequence of determining critical submanifolds associated to each hyperbolic point. In fact, we will truncate each critical submanifold at its first intersection with $\pi^{-1}(0)$. If h^+ is a positive hyperbolic point of $\mathcal{F}_{\widetilde{\pi}}$ (i.e., an index-two critical point of π) with $\pi(h^+) = c_+$, its stable submanifold intersects each level set of $\pi^{-1}[0, c_+]$. Denote these intersections by γ^+, specifying the level set containing γ^+ if necessary. Similarly, the unstable critical submanifold of a negative hyperbolic point h^- (i.e., an index-one critical point of π) with $\pi(h^-) = c_-$ intersects each level set of $\pi^{-1}[c_-, 1]$ and we denote each of these intersections by γ^-. Observe, as in Figure 4, that if the unstable and stable critical submanifolds intersect, then a single critical submanifold may be represented on subsequent pages by multiple arcs.

Example 2.7. As a first example, we consider a Morse foliated open book for a solid torus. Begin with an embedding of the solid torus in S^3 as shown in Figure 5. As proven in [11], the function defined as the radial coordinate of the embedding may be perturbed near the boundary of the solid torus so that the resulting restriction defines a Morse foliated open book with the same critical points on the boundary. There are four critical points, alternating in sign, and S_1, S_2, and S_3 in Figure 5 are pages — that is, level sets — for regular values separated by critical points. Pages S_0 and S_4 are not separated by a critical level, but we include both in order to

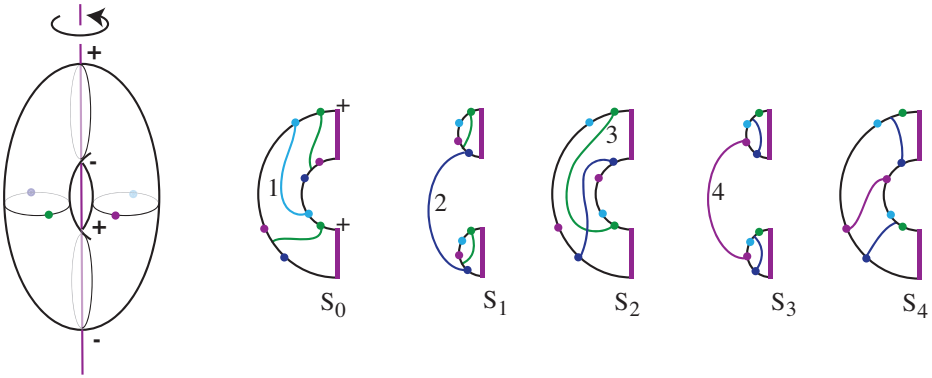


Figure 5. Level sets $S_q := \pi^{-1}(\frac{q\pi}{2} + \epsilon)$ for a Morse foliated open book on a solid torus, together with their intersections with the truncated critical submanifolds. The bold dots along the non-binding boundary of the pages indicate the relative positions of the intersection arcs $\gamma^{\pm}$, a consequence of choosing $\nabla\pi$ to be preferred.

see the intersections of both the ascending (on S_4) and descending (on S_0) critical submanifolds.

Lemma 2.8. *The truncated stable critical submanifolds are disjoint. Similarly, the truncated unstable critical submanifolds are disjoint.*

Proof. The lemma follows from the fact that $\nabla\widetilde{\pi}$ is Morse–Smale. □

We now consider the flowline through an arbitrary point p in the interior of $\pi^{-1}(0)$. If $p \in \gamma_i^+$ for some positive hyperbolic point h_i^+, then the flowline through p will terminate at h_i^+. However, for all points in the complement of the $\{\gamma_i^+\}$, there is a well-defined first-return map. Define $P := \pi^{-1}(0) \setminus (\cup_i \gamma_i^+)$.

Definition 2.9. The *monodromy* $H : P \to P$ of a foliated open book is the first return map of $\nabla\pi$.

Remark 2.10. This use of the term monodromy differs slightly from the map h called the monodromy in [11; 1]. The domain of H is a proper subset of the S_0 page, while h is a homeomorphism from the final page to the initial one. On P, $H = h \circ \iota$.

2.2. ***From foliated to partial open books.*** Foliated and partial open books are alternative ways to decompose contact manifolds with boundary, and unsurprisingly, they are closely related. To each foliated open book we associate a triple (S, P, H), where $S := \pi^{-1}(0)$, and P and H are as above. Although S and P may be viewed as abstract surfaces, rather than embedded ones, we retain the identifications $P \subset S$ and $\partial S = B \cup \widetilde{\pi}^{-1}(0)$, where $B = \partial S \cap \partial P$. Under certain circumstances described

below, the triple associated to a foliated open book in fact defines a partial open book for the same contact manifold.

Remark 2.11. In [11], P is defined by removing a neighborhood of the arcs γ_i^+ together with the non-binding boundary of S. However, the resulting subsurface is isotopic to the P defined above and the results stated for triples are all independent of this choice.

Definition 2.12. A Morse foliated open book $(B, \pi, \mathcal{F}_{\widetilde{\pi}}, \nabla\pi)$ is *sorted* if there are no flowlines contained in $\pi^{-1}(0, 1)$ between distinct critical points.

Equivalently, a foliated open book is sorted if the set of all truncated critical submanifolds is disjoint. We note that this definition highlights the role of preferred gradient-like vector fields, as the constraints on the boundary will force intersections that could otherwise be avoided.

Proposition 2.13 [11, Proposition 8.11]. *Suppose that* (S, P, H) *is the triple associated to a sorted foliated open book. If* S *may be built up from* $S \setminus P$ *by attaching one-handles along* 0*-spheres embedded in* P*, then the triple* (S, P, H) *defines a partial open book for* $(M, \partial M, \Gamma)$.

As in the case of other forms of open books, foliated open books admit an operation called *positive stabilization* which preserves the supported contact structure. Positive stabilization is equivalent to taking an appropriate connect sum with a foliated open book cut from the standard tight S^3; see Figure 6. A positive stabilization is determined up to equivalence by a choice of properly embedded arc $\gamma \subset \pi^{-1}(t)$ with $\partial\gamma \subset B$, and it changes each level set of π by attaching a 1-handle along $\partial\gamma$ which becomes part of P. The monodromy of the foliated open book changes by a positive Dehn twist along the circle formed by γ and core of the added 1-handle, restricted to P.

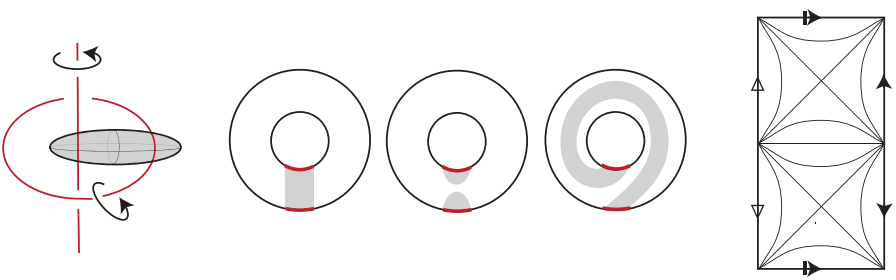


Figure 6. The complement of the shaded ball on the left is a ball in the standard tight S^3. The unshaded portions of the annuli in the center are the topologically distinct pages of the resulting foliated open book, and the boundary foliation is shown on the right. This example appears as Figure 16 in [11] where it is discussed in detail.

3. Right-veering monodromies and examples

Definition 3.1. Let γ, δ be properly embedded arcs with $\partial\gamma = \partial\delta$. We write $\delta < \gamma$ if, after isotoping the two arcs relative to their shared boundary so that they intersect minimally, δ does not lie to the left of γ near either endpoint. If $\delta < \gamma$ or δ is isotopic to γ, then we write $\delta \leq \gamma$.

Consider a surface S and a subsurface $P \subset S$. Let $H : P \to S$ be an embedding which restricts to the identity on $\partial P \cap \partial S$.

Definition 3.2. Given (S, P, H) as above, $H : P \to S$ is *right-veering* if for every γ properly embedded in P with $\partial\gamma \subset (\partial P \cap \partial S)$, $H(\gamma) \leq \gamma$.

Definition 3.3. The foliated open book $(B, \pi, \mathcal{F}_{\widetilde{\pi}}, \nabla\pi)$ is *right-veering* if the associated triple (S, P, H) is right-veering.

We note that this definition does not depend on the choice of preferred $\nabla\pi$. Any two preferred gradient-like vector fields for a fixed π agree near ∂M, and they are connected by a path of preferred gradient-like vector fields all fixed near the boundary. In particular, this implies that the flowlines on the boundary are preserved throughout the interpolation. It follows that the arcs γ_i^+ on S may change only by isotopy, as an arc slide would required the Morse–Smale condition to fail at some point.

When the triple (S, P, H) defines a partial open book, the previous two definitions exactly coincide with those of Honda–Kazez–Matić. However, we observe that the definitions here are broader in scope, applying to the triple (S, P, H) associated to an arbitrary Morse foliated open book.

Example 3.4. Consider the solid torus of Example 2.7. As seen in Figure 5, P consists of a union of four discs. It follows that the monodromy H restricted to each component is trivial, so the Morse foliated open book is necessarily right-veering. Note, however, that this triple does not define a partial open book; since some components of P meet $S \setminus P$ along a single curve, S cannot be built up from $S \setminus P$ by one-handle addition.

We next show an advantage of extending these definitions; informally, it allows one to recognize non-right-veering monodromy (and hence, overtwistedness) in simpler objects. The following example begins with a non-right-veering Morse foliated open book that does not define a partial open book. However, after performing a single stabilization, the resulting Morse foliated open book defines a right-veering partial open book.

Example 3.5. We describe $(B, \pi, \mathcal{F}_{\widetilde{\pi}}, \nabla\pi)$ via three of its regular pages; in fact, this determines several distinct foliated open books depending on the order in which the positive (respectively, negative) hyperbolic points appear in the foliation, but the

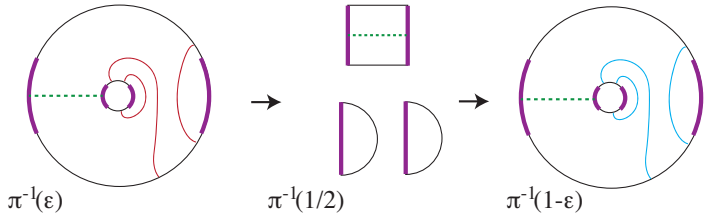


Figure 7. Selected regular pages decorated with their intersections with the truncated critical submanifolds.

example does not depend on this choice. Suppose that the map $h : \pi^{-1}(1-\epsilon) \to \pi^{-1}(\epsilon)$ is a left-handed Dehn twist relative to the page, so that the dotted arc in Figure 7 is evidently non-right-veering. Note, too, that the associated triple does not define a partial open book, as the bigon components of P cannot be built up from the rest of the page by attaching one-handles.

Now stabilize the foliated open book along $\gamma \subset \pi^{-1}(1-2\epsilon)$ of Figure 8, adding a one-handle to P on each page and changing the monodromy by a positive Dehn twist on $\pi^{-1}(1-2\epsilon)$. After the dotted arc flows through this page, it is disjoint from the core of the original annulus, so it remains unaffected by the negative Dehn twist. Since P consists of just two rectangular components, it is easy to verify that there are no arcs which veer strictly left.

The appeal of this example lies in the fact that it detects non-right-veering behavior — and hence, overtwistedness — in a simple object, a foliated open book whose associated triple satisfies weaker conditions than those required by a partial open book. However, it also highlights the difference between equivalence classes of contact manifolds with foliated boundary and those with merely convex boundary. Honda–Kazez–Matić have shown that every overtwisted contact manifold with convex boundary is supported by some non-right-veering open book; in fact, the partial open book of Figure 8 is a stabilization of a non-right-veering open book, but this stabilization changes $|B|$, and hence does not preserve the foliated boundary.

Nevertheless, a non-right-veering arc in a foliated open book does play the same role as in other forms of open books:

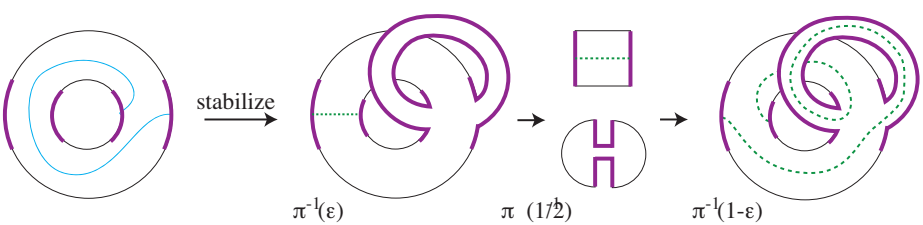


Figure 8. After stabilizing along the indicated arc, the flow changes by a positive Dehn twist.

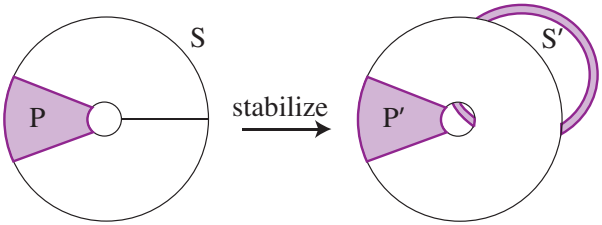


Figure 9. This partial open book stabilization preserves the equivalence class of the convex boundary, but not the foliated boundary.

Theorem 3.6. *A contact three-manifold with foliated boundary is overtwisted if and only if it is supported by a foliated open book that is not right-veering.*

Remark 3.7. The overtwistedness of a contact 3-manifold with foliated boundary does not depend on the particular foliation on the boundary, but only on the associated dividing set Γ. By Theorem 3.6, then, if two contact structures differ only near $\partial M \times I$, and there they are both foliated by convex surfaces $\partial M \times \{t\}$, then one of them is supported by a non-right-veering foliated open book if and only if the other is.

The "if" direction of Theorem 3.6 follows from the following result:

Proposition 3.8. *If a foliated open book is non-right-veering, then the supported contact structure is overtwisted.*

Proof. We prove this result via the analogous statement for partial open books. Specifically, we will show that if a foliated open book has a non-right-veering arc, then we may construct a partial open book for the same contact manifold that also has a non-right-veering arc. By the work of Honda–Kazez–Matić, this implies the supported contact structure is overtwisted. If the triple associated to the foliated open book already defines a partial open book, there is nothing to do, so we consider the following case.

Suppose that $(B, \pi, \mathcal{F}_{\widetilde{\pi}}, \nabla\pi)$ has a non-right-veering arc but the associated triple (S, P, H) does not define a partial open book. It is always possible to stabilize the foliated open book so that the triple defines a partial open book, and we show that these stabilizations may be chosen to preserve the non-right-veering arc.

A stabilization is completely determined by a choice of stabilizing arc on a page. We show that under the conditions of the proposition, a sequence of stabilizing arcs may be chosen that are both disjoint from the fixed non-right-veering arc $\gamma \subset P$ and which produce a partial open book as the associated triple.

Note first that there are two ways in which the triple associated to a Morse foliated open book may fail to define a partial open book. First, the Morse foliated open book may not be sorted, and second, if may be impossible to build S up from $S \setminus P$ by one-handle addition.

If a foliated open book is not sorted, [11] describes how to choose stabilizing

arcs to produce a sorted version. Starting from $t = 0$, increase the t value until the first time when a flowline between a pair of critical points appears. The stabilization should be performed on a page intersecting this flowline, and except for at its endpoints, the arc should be chosen to lie in a neighborhood of the non-binding boundary and the intersections between critical submanifolds and the chosen page, $\pi^{-1}(t_0)$. A non-right-veering arc on $\pi^{-1}(0)$ lies completely in P, and by definition, it will flow to an arc in $\pi^{-1}(t_0)$ which is disjoint from the critical submanifolds and the non-binding boundary of the page. The stabilizing arcs involved in rendering a foliated open book sorted may be assumed disjoint from any non-right-veering arcs.

If (S, P, H) fails to define a partial open book because S cannot be built up from $S \setminus P$ by one-handle addition, then there are components of P which intersect $S \setminus P$ along a single interval in ∂P. Fix one such component, and choose a next component with respect the boundary orientation around S. Choose a boundary parallel stabilizing arc that connects the two components; this may clearly be done in the complement of a non-right-veering arc that is properly embedded in P. This argument misses the case when P consists of single component; however, since P is cut from S by a two-sided arc, such a subsurface meets $S \setminus P$ in two components. □

The proof of the "only if" direction in Theorem 3.6 is an amalgamation of the constructions of [11, Section 8.5] and [7, Proposition 4.1.]. Although these require the use of embedded foliated open books, rather than the Morse foliated open books highlighted in this article, the statement is such a close fit to the topic that we have elected to include it anyway.

Proof of Theorem 3.6:. Section 8.5 of [11] proves the existence of a supporting foliated open book for a given contact manifold via a partial open book adapted to the given foliation near the boundary. The proof uses the foliation to construct this adapted partial open book near the boundary and then extends it into the interior of the manifold in a standard way using a contact cell decomposition. On the other hand, the proof of [7, Proposition 4.1.] considers a (non-right-veering) partial open book for a neighborhood of an overtwisted disc, connects it with a Legendrian arc to the portion of the partial open book that is constructed near the boundary, and then extends it in a standard way using a contact cell decomposition for the complement.

We now prove that an overtwisted contact three-manifold with foliated boundary $(M, \xi, \mathcal{F})$ has a non-right-veering foliated open book. Take a partial open book for a neighborhood of an overtwisted disc, connect it with a Legendrian arc to the portion of the partial open book near the boundary that is adapted to the foliation, and then extend it in a standard way using a contact cell decomposition for the complement. This partial open book now can be extended to a foliated open book that supports $(M, \xi, \mathcal{F})$ and the non-right-veering arc is preserved throughout the extension process. □

Corollary 3.9. *If $(B, \pi, \mathcal{F}_{\widetilde{\pi}}, \nabla\pi)$ is non-right-veering, then the bordered sutured contact invariant $c(\xi)$ vanishes.*

Proof. Theorem 3 in [1] shows that under a certain isomorphism between the bordered sutured Floer homology of a foliated open book and the sutured Floer homology associated to the corresponding partial open book, the bordered sutured contact invariant $c(\xi)$ maps to the Honda–Kazez–Matić invariant $EH(M, \Gamma, \xi)$. The EH invariant vanishes for overtwisted contact manifolds, and the result follows. □

Example 3.10. Here, we describe a Morse foliated open book for an overtwisted three-ball and then show how the existence of a non-right-veering arc guides the construction of a transverse overtwisted disc. This is essentially the construction described in [10], but some adaptation is required because the topological type of the page changes with the S^1 parameter.

Figure 10 shows an S^2 decomposed into ten squares, each of which represents the square tile in the boundary open book foliation defined by a single hyperbolic point. The integers indicate the order of the hyperbolic points, so that the first is a saddle resolution which transforms AJ and LE leaves into AE and LJ leaves. The first five hyperbolic points are all positive, so each of the first five changes to the topology of the page is given by cutting along a single arc where the corresponding critical submanifold intersects the page. These arcs are shown together on the first page in Figure 11, showing that P consists of five discs. There is a unique non-boundary-parallel arc, shown as a dotted curve, in the disc with corners labeled GLKH.

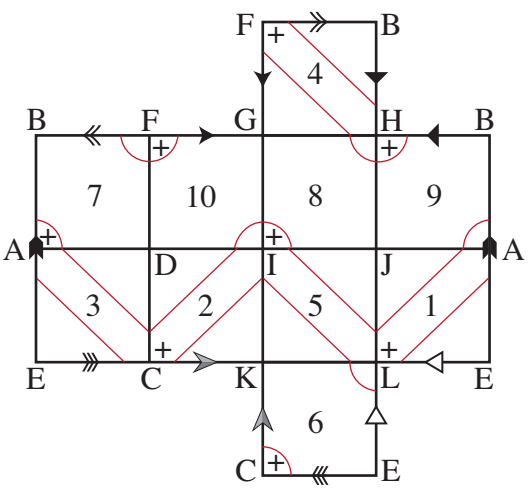


Figure 10. Hyperbolic tiles of the open book foliation on the boundary of an overtwisted three-ball. The dividing set Γ is shown in thinner (red) lines.

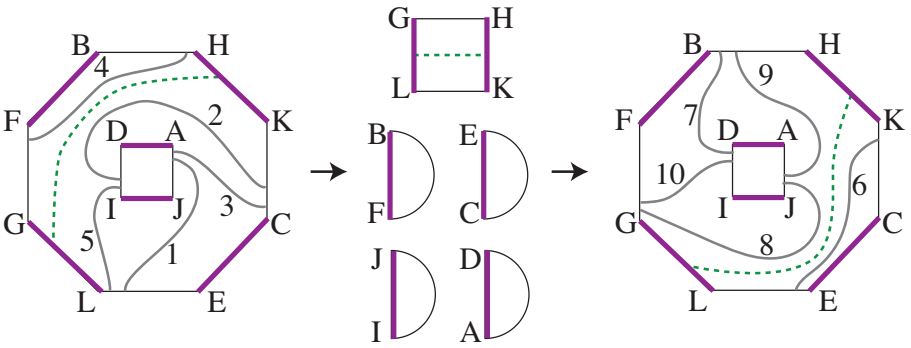


Figure 11. Selected pages of the Morse foliated open book decorated with their intersections with the truncated critical submanifolds, each labeled by the corresponding hyperbolic point. The dotted arc is non-right-veering under the foliated open book monodromy that identifies the right-hand page with the left-hand page by translation.

It is not difficult to show that the manifold defined by these pages is a ball; to see that it's an overtwisted ball, recall from Definition 2.3 that the open book foliation is topologically conjugate to the characteristic foliation of a supported contact structure. Thus, the dividing set on S^2 is the boundary of a neighborhood of the positive separatrices of the positive hyperbolic points. Shown in thinner (red) lines on Figure 10, the resulting Γ has three components.

Following the approach of [10], we construct a transverse overtwisted disc by describing how it intersects each page. As t changes, this intersection changes either by isotopy or by a saddle resolution, as shown in Figure 12.

Although the existence of a non-right-veering arc allows us to construct an overtwisted disc, certainly there are foliated open books for overtwisted contact manifolds that don't have non-right-veering arcs. There are several constructions in the literature that show how a non-right-veering open book may become a right-veering open book via a sequence of positive stabilizations but the next example illustrates that even an unstabilized foliated open book for an overtwisted contact manifold may be right-veering [6].

Example 3.11. Finally, we turn to a "minimal" neighborhood of an overtwisted disc. Beginning with the simplest open book foliation on a transverse overtwisted disc, we first construct something that is almost a foliated book by thickening the disc and taken the thickened leaves of the foliation as pages. This general approach for constructing foliated open books described in Section 4.2 of [11], and this specific case is explored in detail as Example 4.7 of [2].

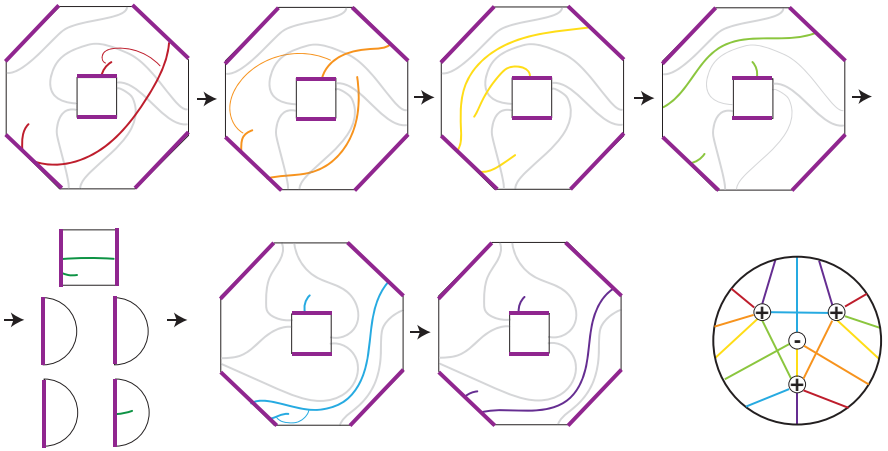


Figure 12. Movie presentation of the transverse overtwisted disc shown on the lower right. Steps between slices are given by isotopy of the bold arcs and saddle resolutions guided by the thin arcs.

In this construction, the function to S^1 is induced from the S^1 function on the foliation, but when we thicken the surface, each of the two critical pages constructed thus has a pair of critical points. We therefore locally perturb the Morse function so that $\pi(h_1) < \pi(h_2) < \pi(h_3) < \pi(h_4)$. This yields the regular pages shown in Figure 13. Each page is decorated with its intersections with the ascending critical

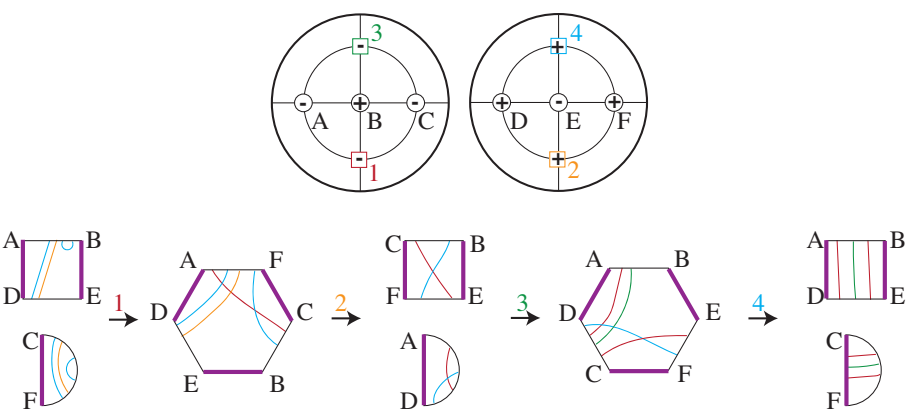


Figure 13. Top: selected leaves of the open book foliation on the front (left) and reverse (right) of a transverse overtwisted disc. Circles are elliptic points and squares are hyperbolic points. Bottom: Regular pages of a Morse foliated open book for a neighborhood of the disc, decorated with the their intersections with the truncated critical submanifolds.

submanifold of hyperbolic points with lower values and the descending critical submanifold of hyperbolic points with greater values.

Although this ball is overtwisted, the Morse foliated open book is right-veering. To see this, observe in Figure 13 that P consists of eight discs, none of which has more than one component of B on the boundary. Thus all arcs in P are boundary parallel and the monodromy is trivial.

References

[1] A. Alishahi, V. Földvári, K. Hendricks, J. E. Licata, I. Petkova, and V. Vértesi, "Bordered Floer homology and contact structures", 2021. arXiv 2011.08672

[2] A. Alishahi, J. E. Licata, I. Petkova, and V. Vértesi, "A friendly introduction to the bordered contact invariant", 2021. arXiv 2104.07616

[3] D. Bennequin, "Entrelacements et équations de Pfaff", pp. 87–161 in *Third Schnepfenried geometry conference* (Schnepfenried, 1982), vol. 1, Astérisque **107**, Soc. Math. France, Paris, 1983. MR

[4] J. S. Birman and W. W. Menasco, "Studying links via closed braids, I: A finiteness theorem", *Pacific J. Math.* **154**:1 (1992), 17–36. MR

[5] N. Goodman, "Overtwisted open books from sobering arcs", *Algebr. Geom. Topol.* **5** (2005), 1173–1195. MR

[6] K. Honda, W. H. Kazez, and G. Matić, "Right-veering diffeomorphisms of compact surfaces with boundary", *Invent. Math.* **169**:2 (2007), 427–449. MR

[7] K. Honda, W. H. Kazez, and G. Matić, "The contact invariant in sutured Floer homology", *Invent. Math.* **176**:3 (2009), 637–676. MR

[8] K. Honda, W. H. Kazez, and G. Matić, "On the contact class in Heegaard Floer homology", *J. Differential Geom.* **83**:2 (2009), 289–311. MR

[9] T. Ito and K. Kawamuro, "Open book foliation", *Geom. Topol.* **18**:3 (2014), 1581–1634. MR

[10] T. Ito and K. Kawamuro, "Visualizing overtwisted discs in open books", *Publ. Res. Inst. Math. Sci.* **50**:1 (2014), 169–180. MR

[11] J. E. Licata and V. Vertesi, "Foliated open books", 2020. arXiv 2002.01752

[12] E. Pavelescu, *Braids and open book decompositions*, Ph.D. Thesis, University of Pennsylvania, 2008, available at www.proquest.com/docview/304492097. MR

Received 19 Feb 2021. Revised 27 Oct 2021.

VERA VÉRTESI: vera.vertesi@univie.ac.at
Faculty of Mathematics, University of Vienna, 1090 Vienna, Austria

JOAN E. LICATA: joan.licata@anu.edu.au
Mathematical Sciences Institute, The Australian National University, Canberra ACT 0200, Australia

THE OPEN BOOK SERIES 5

Gauge Theory and Low-Dimensional Topology: Progress and Interaction

This volume is a proceedings of the 2020 BIRS workshop *Interactions of gauge theory with contact and symplectic topology in dimensions 3 and 4*. This was the 6th iteration of a recurring workshop held in Banff. Regrettably, the workshop was not held onsite but was instead an online (Zoom) gathering as a result of the Covid-19 pandemic. However, one benefit of the online format was that the participant list could be expanded beyond the usual strict limit of 42 individuals. It seemed to be also fitting, given the altered circumstances and larger than usual list of participants, to take the opportunity to put together a conference proceedings.

The result is this volume, which features papers showcasing research from participants at the 6th (or earlier) *Interactions* workshops. As the title suggests, the emphasis is on research in gauge theory, contact and symplectic topology, and in low-dimensional topology. The volume contains 16 refereed papers, and it is representative of the many excellent talks and fascinating results presented at the Interactions workshops over the years since its inception in 2007.

TABLE OF CONTENTS