EXISTENCE OF TWO-PARAMETER CROSSINGS, WITH APPLICATIONS

JONATHAN D. WILLIAMS

ABSTRACT. A Morse 2-function is a generic smooth map f from a manifold M of arbitrary finite dimension to a surface B. Its critical set maps to an immersed collection of cusped arcs in B. The aim of this paper is to explain exactly when it is possible to move these arcs around in B by a homotopy of f and to give a library of examples when M is a closed 4-manifold. The last two sections give applications to the theory of surface diagrams of smooth 4-manifolds.

CONTENTS

1. INTRODUCTION

The theory of singular fibrations of n-manifolds for $n \geq 3$ has many applications in low-dimensional topology. For instance Morse functions, open books and Lefschetz fibrations all lead to depictions of 3- and 4-manifolds using surfaces (which are regular fibers) decorated with circles, and such discrete information has been a rich playground. Though its definitions go at least as far back as catastrophe theory, the Morse 2-function as a fibration structure is a promising invention of [ADK], with recent development such as [AK, B1, B2, Be, BH, GK2, H, Le, W1, W2]. The image of the critical set of a Morse 2-function is an immersed collection of cusped 1-submanifolds, and the main goal of this paper is to explain when it is possible to move the critical image around by a homotopy of the fibration map (Theorem 2.4). The subsequent sections give applications of the ideas from Section 2: Section 3 presents an algorithm for converting an arbitrary indefinite Morse 2-function into one that is *simplified* in the sense of [BS], with a significantly shorter proof. Section 4 has refinements of the main result of [W2], which is a uniqueness result for surface diagrams of smooth 4-manifolds (see Definition 4.1). One of these refinements states that if two surface diagrams come from homotopic fibration maps and have the same genus, then one may be transformed into the other using genus-preserving moves. This is in contrast to the corresponding statement for pairs of Heegaard diagrams of 3-manifolds, which in general require stabilizations, and it simplifies any effort to define a diffeomorphism invariant calculated using surface diagrams.
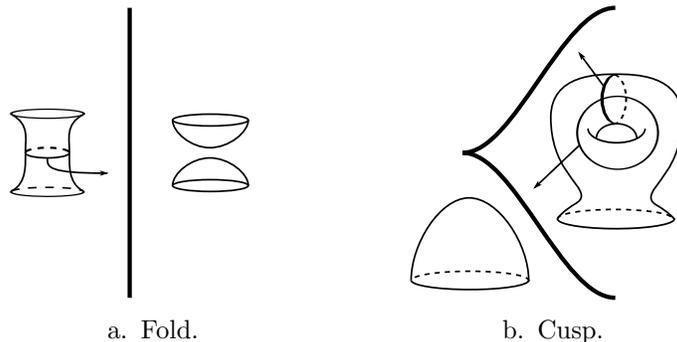
a. Fold.   b. Cusp.

FIGURE 1. Critical points of purely wrinkled fibrations.

Definition 1.1. An *indefinite Morse 2-function* is a smooth map f from an n-manifold M to a 2-manifold B such that each critical point has the local model of an index k *indefinite fold* or an index k *indefinite cusp*, given by

$$(x_1, \ldots, x_n) \mapsto \left(x_1, -\sum_{i=2}^{k+1} x_i^2 + \sum_{i=k+2}^{n} x_i^2 \right) \qquad \text{(fold)}$$

$$(x_1, \ldots, x_n) \mapsto \left(x_1, x_2^3 - 3x_1 x_2 - \sum_{i=3}^{k+2} x_i^2 + \sum_{i=k+3}^{n} x_i^2 \right) \qquad \text{(cusp)}$$

for $k \in \{1, \ldots, n-2\}$. In this paper, a *deformation* of a Morse 2-function is a homotopy of maps from a smooth, closed oriented 4-manifold M to a smooth, closed oriented surface B given by a sequence of merges, births, flips and 2-parameter crossings as described in [W2].

The local models for $k = 0$ and $k = n - 1$ (the *definite* fold and cusp) are omitted, though Theorem 2.4 could be easily adapted to handle those critical points as well. The reason for this is the definite locus can be eliminated from smooth maps when $n = 4$ by homotopy (see for example the earliest proof of this fact in [S]), and there are general existence results for indefinite one-parameter families of maps connecting any pair of homotopic indefinite Morse 2-functions when $n \geq 4$ [W2, GK2]. To help understand folds and cusps, Figure 1 shows two disks in B, schematically depicting the fibration structure near a fold and a cusp for $n = 4$. Bold arcs represent the image of the critical locus and a surface is pictured in the region of regular values that have that surface as their preimage. Tracing point preimages above a horizontal arc from left to right gives the foliation of $\mathbb{R}^3$ by hyperboloids, with a double cone above the fold point. The circle (drawn on the cylinder to the left) that shrinks to the cone point is called the *vanishing cycle* for that critical arc. Each cusp point is a common endpoint of two open arcs of fold points for which the two vanishing cycles must transversely intersect at a unique point in the fiber. See [GK2] for a more detailed background on these maps.

Section 2 gives the complete characterization of when a proposed movement in the critical image of a wrinkled fibration, similar to a sequence of Reidemeister moves for knot projections, can be achieved by a deformation. Such homotopies are sequences of *2-parameter crossings*, a term that appeared in [GK2, Section 2]. In their paper and in this paper, it replaces the less precise term *isotopy* that
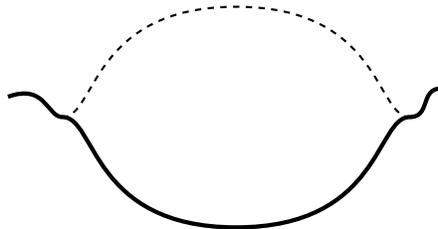
FIGURE 2. A fold arc (in bold) runs along the boundary of a disk Δ in the target space of f_0 in preparation to move across Δ by some deformation of f_0, eventually to coincide with the dotted line (other critical points omitted).

appeared in [Le] and [W1, Section 2.4.1], which in those papers was a sequence of 2-parameter crossings. To describe the general situation, suppose there is a purely wrinkled fibration $f_0 \colon M^4 \to D^2$ and an embedded disk

$$\Delta \colon \{z \in \mathbb{C} : |z| \leq 1\} \to D^2$$

smoothly embedded except at the two corners on either side, such that, for some fold arc $\varphi_0 \subset M_0$, the restriction

$$f_0|_{\varphi_0} \colon \varphi_0 \to \Delta \left(\{|z| = 1, \ \mathrm{Im}\, z \leq 0\}\right)$$

is a diffeomorphism (see Figure 2; sometimes Δ also denotes its image). The problem is to determine when there is a deformation $f_t \colon M_I \to D^2_I$ such that the image of each point $\Delta(z) \in f_0(\varphi_0)$ in this fold arc follows the path $\Delta(t\bar{z} + (1-t)z)$ while the rest of the critical image remains fixed. Theorem 2.4 gives a necessary and sufficient criterion for this to occur, and the following propositions present a few standard cases in which a proposed 2-parameter crossing always exists.

Acknowledgments. The author would like to thank Denis Auroux, David Gay and Yankı Lekili for helpful conversations during the preparation of this work.

2. EXISTENCE OF 2-PARAMETER CROSSINGS

Consider a Morse function $f \colon M^n \to [0, 4]$ with two isolated critical points p_1 and p_2, with $f(p_i) = i$, and equip M with a Riemannian metric. If the ascending manifold Z of p_1 is disjoint from the descending manifold Y of p_2 (where critical points are considered part of their ascending and descending manifolds), then there is a neighborhood νZ on which the restriction of f has exactly one critical point p_1, and a homotopy of f supported on νZ that sends $f(p_1)$ to 3. In Figure 2, there is a one-parameter family of smooth functions, finitely many of which may not be Morse, and a one-parameter family Z of ascending manifolds for the one-parameter family of Morse critical points given by φ_0. If Z is disjoint from the one-parameter family of descending manifolds coming from all other critical points in Δ, then the proposed homotopy exists over every vertical arc $\{z = const\}$, allowing the fold arc to pass over Δ. Here follows a more precise statement using the language of [BH, Section 2.3].

Definition 2.1 (Disjointness condition)**.** Let $\mathcal{H}$ denote a horizontal distribution for the Morse 2-function f_0 over a neighborhood of Δ, obtained by taking the orthogonal complement $\ker(df_0)^\perp$ with respect to some Riemannian metric on $f_0^{-1}(\Delta)$. Consider the union over all regular values p of f_0,

$$V = \bigcup_p V_0^{\mathcal{H}}(\gamma_p^{up}) \cup V_0^{\mathcal{H}}(\gamma_p^{down}),$$

where the smooth embeddings $\gamma_p^{up}, \gamma_p^{down} \colon [0,1] \to \Delta$ travel from p vertically up or down, respectively, to the two boundary arcs of Δ. Let Z be the points in V that run into φ_0 and let Y be the points in V that run into $\mathrm{crit}(f_0)\backslash \varphi_0$. The disjointness condition is that there exists an $\mathcal{H}$ such that Y and Z are bounded away from each other in the preimage of Δ.

In this definition, it is acceptable for the terminal point of γ_p^{down} to be a singular fiber over $f(\varphi_0)$ because we only consider the vanishing set V_0 of points in the regular fiber $f_0^{-1}(p)$ that run into critical points (not the ones that emerge from critical points). Similarly, there may be arcs γ_p^{up} whose endpoint is a critical value. Vanishing sets appear in the definition instead of ascending and descending manifolds because they are defined. For example, suppose γ_p^{up} is tangent to the image of a fold arc. In that case, the restriction of f_0 to the preimage of γ_p^{up} is not a Morse function, so it is not appropriate to speak of the descending manifold of the critical point of f_0 above that tangency. On the other hand, the vanishing set is defined in [BH] and has the same properties related to pushing around critical values.

Remark 2.2. A disk Δ and a horizontal distribution $\mathcal{H}$ satisfying the disjointness condition over Δ can be used to specify a deformation suggested by Δ. Let N be a neighborhood of the vanishing set of φ_0 over a neighborhood $\nu\Delta$ of Δ (note N is a ball, as discussed in [W1, Section 2.3.1]). Choose local coordinates near N so that $f_0|_N$ is given by the local model for folds. Then x_1 parameterizes φ_0, and the level sets for suitably parameterized vertical arcs foliating $\nu\Delta$ are the level sets of $-\sum_{i=2}^{k+1} x_i^2 + \sum_{i=k+2}^{n} x_i^2$. Such local coordinates exist because N consists of regular points, with the exception of φ_0, by the disjointness condition. Now, restricting to N, Figure 2 is merely the image of the local model for folds, and there is certainly a deformation of the local model $N \to \nu\Delta$, fixing the map near $\partial \overline{N}$, realizing the suggested movement in S^2. Call it f_t^N, $t \in [0,1]$. Then the deformation is given by

$$f_t(x) = \begin{cases} f_t^N(x) & x \in N \\ f_0(x) & x \notin N. \end{cases}$$

Remark 2.3. One further observation is that $Z \cap Y$ always maps to a family of vertical arcs connecting the two sides of Δ, because the union of paths traced by any point in $f_0^{-1}(p)$ above γ_p^{up} and its reverse coincides with that of γ_p^{down}, and its reverse.

Theorem 2.4. The disjointness condition of Definition 2.1 is equivalent to the existence of a deformation as suggested by Figure 2.

Proof. The disjointness condition is merely a restatement of the requirement that the relevant critical sets of the restrictions of f_0 to the vertical arcs in Δ simultaneously satisfy the same disjointness condition for switching critical points of smooth real-valued functions. $\square$
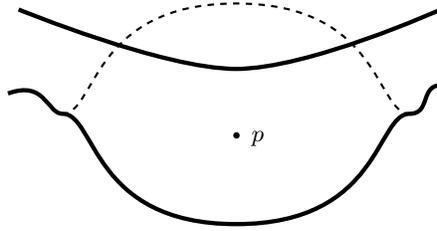
FIGURE 3. Preparing to push a bit of a fold arc past another, in what is commonly called a *finger move*, for Proposition 2.5(1).

Here follow some examples of how to use Theorem 2.4 to verify a proposed movement of critical arcs in the base when $n = 4$. Note that indefinite folds in this case have a higher-genus and lower-genus side because the regular fibers of such maps are closed, orientable surfaces. Existence results for various types of what are called *Reidemeister-2 fold crossings* in [GK2] and [W2] are collected into the following proposition.

Proposition 2.5 (Existence of Reidemeister-2 fold crossings)**.**
   (1) There exists a deformation as suggested in Figure 3 if one of the following two conditions occur.
       (a) The point p is on the lower-genus side of at least one of the pictured fold arcs.
       (b) The point p is on the higher-genus side of both fold arcs, and there exist embedded reference paths from p to each fold arc within Δ for which the measured vanishing cycles in $f_0^{-1}(p)$ are disjoint.
   (2) In Figure 4, suppose the fiber genus of each fiber component above p is at least 2, and that p lies on the higher genus side of exactly one of the fold arcs, say φ_0. Then the suggested R_2 deformation exists if and only if the vanishing cycles at $\Delta(\pm 1)$ match, as measured with reference paths from $\Delta(i)$.
   (3) There always exists an R_2 deformation as suggested by Figure 4 when p is on the lower-genus side of both fold arcs, even when there are cusps in the boundary of the lune.
   (4) In Figure 4, suppose p lies on the higher-genus side of both fold arcs, all regular fibers are connected, and the fiber above p has genus at least 2. Then there exists an R_2 deformation as suggested by the figure.

Proof. The typical approach for a statement like Assertion (1) will be to show that there is a choice of horizontal distribution such that $Z \cap Y = \emptyset$. First, note that p must either lie on the lower-genus side of at least one fold arc which we will call A (case (a)), or on the higher-genus side of both (case (b)). In either case, the deformation exists if and only if there is a vertical arc $v \subset \Delta$ connecting the two fold arcs that is disjoint from $f_0(Z \cap Y)$, because in that case one could shrink Δ so that its interior is a tubular neighborhood of v whose preimage satisfies the disjointness condition.

For case (1a), $Z \cap Y$ is 1-dimensional if p lies on the lower-genus side of exactly one of the fold arcs, and by Remark 2.3 it maps to a disjoint union of vertical
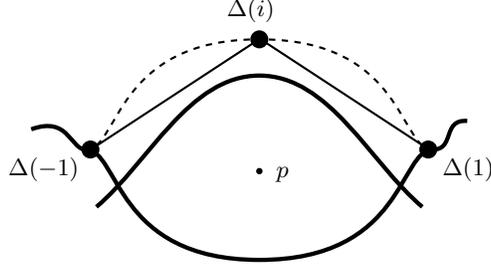
FIGURE 4. The disk Δ with lower fold arc $f_0(\varphi_0)$ in preparation
for an R_2 deformation in items (2)-(4) of Proposition 2.5.

arcs in Δ. If p lies on the lower-genus side of both of the fold arcs then $Z \cap Y$ is
0-dimensional, the intersection of two surfaces in M_I. However, Remark 2.3 implies
the dimension of $Z \cap Y$ is at least 1, for any $\mathcal{H}$. For this reason, $Z \cap Y$ must be
empty in this case.

Remark 2.6. If the previous paragraph sounds strange, here is a digression, giving
another way to explain its conclusion: Any arc in $Z \cap Y$ is simultaneously the cocore
of a three-dimensional 2-handle and the core of a 3-dimensional 1-handle, which is
not only non-generic in the space of Morse functions because it is not Morse-Smale,
but is also non-generic in the space of one-parameter families of Morse functions
(parameterized by the horizontal direction s in Figure 3): Over any horizontal line
$\ell(s)$, the pairs of points that run into the top and bottom fold arcs trace paths in the
3-manifold (fiber) $\times \ell$, and an arc in $Z \cap Y$ is an arc of intersections between these
paths. Given such an arc $\gamma \subset Z \cap Y$, one merely perturbs $\mathcal{H}$ so that the s-value
at which one pair of paths crosses γ is different than the s-value for the other, and
this can be achieved all along γ by performing essentially the same perturbation
crossed with γ. This concludes the digression.

In either case, the dimension of $Z \cap Y$ is no greater than 1, so there is at least
one vertical arc in Δ whose preimage in Z is disjoint from Y, and so this arc has a
neighborhood whose preimage in Z is disjoint from Y, and we perform the finger
move in this neighborhood.

For case (1b), the assumption is that $Z \cap Y = \emptyset$ in the fiber above p. In
the language of handlebodies, we have a three-dimensional 1-handle attachment
followed by a 2-handle attached along a circle that can up to isotopy is disjoint
from the belt circle of the 1-handle. In this case, it is also known that the order
of handle attachment can be switched by a homotopy of the corresponding Morse
function, then un-switched, which similarly gives the suggested R_2 deformation.

For the ($\Leftarrow$) part of Assertion (2), assume the fiber above p is connected, since
the result obviously holds if the vanishing cycles of the two fold arcs, as measured
from $\Delta(i)$, lie in different fiber components. Denote by F the generic fiber over $\Delta(i)$
and F' the fiber over p. Also, denote by q, q' the pair of points in F' corresponding
to the fold arc not containing φ_0, again assumed to be stationary. Because of
the matching vanishing cycles, it makes sense to talk about a single vanishing
cycle v of the fold arc containing φ_0 that lives in F, measured near either end
of φ (and also in F' for those points actually in φ). Choosing small $\epsilon > 0$, the
representatives of v as measured at $\Delta(\pm 1 \mp \epsilon)$ are isotopic in $F' \setminus \{q, q'\}$ because
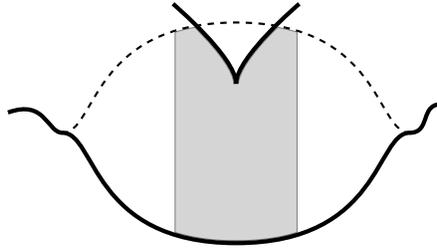
FIGURE 5. Pushing a fold across a cusp. Here, Δ could lie on the higher or lower genus side of $f_0(\varphi_0)$.

they are isotopic in F. The surface $F' \setminus \{q, q'\}$ has free fundamental group, in which elements whose representatives differ by crossing q, q' are distinct. This implies that each crossing of v over q or q' (that is, each point in $Z \cap Y$, which is generically a 1-submanifold of M) is accompanied by a canceling crossing (paired, say, by letting the endpoints of reference paths travel along horizontal lines across the central lune to get one-parameter families of points and vanishing cycles in F'). This allows the disjointness condition to be satisfied by a homotopy of $\mathcal{H}$ that cancels the crossings by identifying vertical arcs in Δ in pairs. The ($\Rightarrow$) part of Assertion (2) is clear; see Example 2.8 for a few examples.

For Assertion (3), after moving all cusps into the higher-genus sides of the two folds, the proof of the analogous case of Assertion (1a) (the last two sentences above Remark 2.6) goes through word-for-word.

Now for Assertion (4). If Z and Y have no intersections above the lune $B \subset \Delta$ bounded by the fold arcs, then it is possible to choose Δ small enough to make $Z \cap Y = \emptyset$. Assume without loss of generality that Z and Y intersect transversely, and let $p \in L$. Since all fibers are connected, the two vanishing cycles as measured at p using the reference paths $\gamma_p^{up}, \gamma_p^{down}$ are homologically distinct and disjoint for all p sufficiently close to the fold crossings. In this situation, any point of intersection between Z and Y is reflected by an intersection between the vanishing cycles in the fiber using vertical reference paths, and this point is one of a pair that forms the corners of a family of lunes, one in in each fiber, each of whose boundary is a pair of arcs, one in each of the two vanishing cycles. By a homotopy of $\mathcal{H}$, these lunes can be simultaneously contracted, eliminating the intersections of Z and Y from L. Now the proposition follows from Theorem 2.4. $\qquad\square$

Here is one more existence result for finger moves involving cusps.

Proposition 2.7. *There always exists a* cusp-fold crossing *deformation as suggested by Figure 5.*

Proof. The proof of this result is just the same as that of Proposition 2.5(1a), except that Z, instead of being a pair of embedded disks given by the fiberwise feet of a 3-dimensional 1-handle, is an immersed pair of disks whose immersion locus is an arc of double points limiting to the cusp, and the argument applies after perturbing $\mathcal{H}$ to eliminate intersections between Z and Y along this arc similarly to Remark 2.6. $\qquad\square$

Example 2.8. Here are a few examples illustrating subtleties surrounding Proposition 2.5(2). Suppose that above $\Delta^{-1}(\{\operatorname{Re} z = 0\})$ there is an arc of intersection

between Z and Y, where the vanishing cycle of φ_0 crosses once over one of the points that runs into the other fold arc, traveling along the reference path $\{\operatorname{Re} z = 0\}$. Then up to isotopy the two vanishing cycles mentioned in Proposition 2.5(2) differ by sliding over a disk bounded by the other fold arc's vanishing cycle, so that the R_2 deformation would result in the particular impossibility that validates the $(\Rightarrow)$ part of the statement: φ_0 would be free of intersection points with other folds and have vanishing cycles at either end that are not isotopic in the fiber above $\Delta(i)$.

Another more subtle problem occurs when the regular fiber over the lune is a torus or a sphere. In this case, there can be many vertical arcs in $f_0(Z \cap Y) \cap \Delta$ in which the vanishing cycle crosses one of the points with the same orientation (think of a meridian of a torus traveling around and around the torus, repeatedly crossing some marked point). Unlike the previous example, the vanishing cycles of φ_0 at $\Delta(\pm 1)$ can match, and the base diagram one might expect from such a move might not have obvious inconsistencies, but the proposed movement of fold arcs in the base is still not valid by Theorem 2.4.

In all these cases, $Z \cap Y$ is generically 1-dimensional: over every vertical arc (oriented downward) in Δ the Morse function corresponds to a pair of consecutive 3-dimensional 2-handle attachments, and every point of $Z \cap Y$ is an intersection between the cocore of the first 2-handle (1-dimensional in each arc, sweeping out a 2-dimensional subset of M_0) and the core of the second (2-dimensional in each arc, sweeping out a 3-dimensional subset of M_0). Certainly there is no way to switch the order of handle attachments while keeping the intersection between core and cocore inside the lune, and (as the first example shows) there is no general way to move a vertical arc of intersections out of Δ by a homotopy.
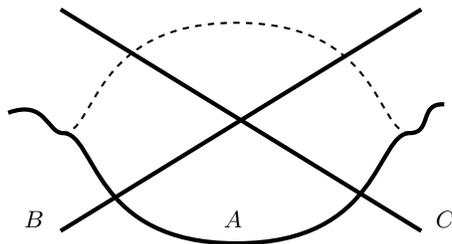

FIGURE 6. A Reidemeister-3 fold crossing.

Proposition 2.9 (Existence of Reidemeister-3 fold crossings)**.** A Reidemeister-3 fold crossing exists if there are no cusps in the triangle and the minimal fiber genus in a neighborhood of the triangle is at least 2.

Proof. Figure 6 depicts the closure Δ of a neighborhood of a triangle of fold arcs. Foliate Δ by a family of vertical paths, each of which can be considered the target of a member of a one-parameter family of Morse functions. There are four cases, according to the number n of fold arcs that have the central triangle on their lower-genus sides. In all cases, the restriction of the map to the preimage of each leaf is a Morse function with three critical points and two or three critical values. The depicted Reidemeister-3 fold crossing can be thought of as affecting the vertical order of the critical values of these Morse functions in various ways. For any Reidemeister-3 fold crossing, we let the disk Δ be the closure of a regular

neighborhood of the triangle formed by the three fold arcs, with $f_0(\varphi_0)$ coinciding with A. Since the fiber is connected with genus at least 2 over every regular value, the vanishing cycles of any one of the three folds as measured by different reference paths from the some arbitrarily chosen point can be assumed to be equal by appropriate choice of horizontal distribution, not just isotopic. Example 2.10 shows why this is a necessary assumption.

The case $n = 3$ can be satisfied as in Proposition 2.5(3): though they would be intersections between two surfaces in a 4-manifold, intersections between the ascending manifold of A and the descending manifold of B or C are not generic.

For $n = 2$, without loss of generality suppose the triangle is on the higher-genus side of B and the lower genus sides of A and C. As in the paragraph above Proposition 2.9, generically the intersection between the ascending manifold of A and the descending manifold of B maps to a collection of vertical line segments that are not clearly removable. However, it is possible to move such a line segment to the left across C, out of the triangle, to lie in the higher-genus side of C. Intersections coming from A and C are ruled out as in the case $n = 3$.

For $n = 1$, without loss of generality suppose the triangle is on the higher-genus sides of B and C and the on the lower genus side of A. As in the case $n = 2$, $Z \cap Y$ is a finite number of line segments connecting A to B (or A to C). Since B and C intersect, their vanishing cycles b_p, c_p (as measured from any reference point p in their higher-genus sides) are isotopic to disjoint circles. For this reason, for every p, every transverse intersection $b_p \cap c_p$ is one of a pair of corners of a lune in F_p, with one side lying in b_p and the other side lying in c_p. By smoothness of the fibration map, these lunes can be chosen to vary smoothly with p. For all p on the higher genus side of both B and C, simultaneously collapse the lunes by an isotopy sending the b_p side to the c_p side of each. What results in each fiber is a collection of arcs (perhaps even all of B and C if they are isotopic) on which the intersections of the fiber with B and C coincide. These arcs vary smoothly with p and it is possible to perturb b_p smoothly over all p to replace each arc with two parallel arcs, one for b_p and one for c_p. Now points in $Z \cap Y$ coming from A and B are necessarily disjoint from c_p for all p in the triangle, so they can be pushed out of the triangle across C, and those points in $Z \cap Y$ coming from A and C can be pushed out across B.

The $n = 0$ case, which is not used in this paper because it is strictly genus-decreasing, is included for completeness. Here, $Z \cap Y$ is generically a surface in M mapping to a codimension 0 subset of Δ, naturally interpreted as a collection of intersection points between vanishing cycles. Let a_p, b_p, c_p be the three vanishing cycles for A, B, C, respectively, measured from a reference fiber F_p over a point p in the triangle. All three fold arcs intersect pairwise, so that a_p, b_p, c_p are isotopic to pairwise disjoint circles in F. As in the $n = 1$ case, it is possible to arrange for b_p and c_p to be disjoint in every fiber on the higher-genus sides of B and C. In particular, this eliminates any triple intersections in $Z \cap Y$, so that points in that set coming from A and C are disjoint from B, so they can be pushed across B, and similarly those coming from A and B can be pushed across C. The three remaining bits at the angles opposite the corners of the triangle can be assumed free of intersections by choosing Δ small enough. □

Example 2.10. This example, adapted from one due to an anonymous referee for [W2], shows why the genus assumption in Proposition 2.9 may be necessary.

Suppose $n = 0$, so that the triangle is on the higher-genus side of all three fold arcs, the regular fiber over any point inside the triangle has genus 1, and that the vanishing cycles for all three fold arcs are isotopic in the central fiber to a meridian. Choose the reference fiber that lies over the point p at the center of the triangle, which we suppose is equilateral. For a short interval of radial reference paths from p to points in the interior of one of the sides of the triangle, suppose the vanishing cycle travels once around the torus. In this situation, it seems the disjointness condition between Z and Y cannot be arranged by pushing intersections out of the triangle.

Proposition 2.11 (Genus-increasing movements of an isolated fold arc)**.** There exists a sequence of two-parameter crossings as suggested by Figure 2 when the lune is on the lower-genus side of $f(\varphi_0)$ and $\iota \cap \varphi_0 = \emptyset$ and all components of all regular fibers inside the lune have genus at least 2.

Proof. The proposed movement exists by the genus-increasing 2-parameter crossing propositions, namely Proposition 2.5(1a,2,3) for R_2 deformations, Proposition 2.7 for cusp-fold crossings, and $n \neq 0$ cases of Proposition 2.9 for Reidemeister-3 fold crossings. The hypotheses of these results are satisfied automatically, except those of Proposition 2.9 (simply observe that any cusps in the triangle can be first pushed into the higher-genus side of φ_0, which by hypothesis has no cusps) and Proposition 2.5(2).

In the context of Proposition 2.5(2) and in Figure 4, the fold arc φ_0 for the present proposition (call it φ^{present}) would be the upper fold arc, while the φ_0 mentioned in Proposition 2.5(2) (call it $\varphi^{(2)}$) would be the lower fold arc. One may choose the movement of φ^{present} as beginning with a pair of finger moves that exist by Proposition 2.5(1a), so that the first pair of intersections as in Proposition 2.5(2) to form during the movement comes from the two finger moves. By appropriately choosing a path in M_0 for the tip of the second finger move to trace, one may choose the location of the two points that become identified in the fiber when φ^{present} passes by; in particular, near the crossings of Figure 4. In this way, it is possible to arrange for the two vanishing cycles of $\varphi^{(2)}$ to match, allowing an application of Proposition 2.5(2) to unify the two fingers into one thick finger. The movement continues with another finger move, chosen appropriately as before, and so on until the proposed movement is finished. $\square$

3. Simplifying a Morse 2-function to the 2-sphere

This section gives a new proof of the existence result [W1, Corollary 2]. The argument in that paper begins with the main theorem of the 41-page paper [GK1], an existence result that in turn relies on the Giroux correspondence of contact geometry [Et] and Eliashberg's classification of overtwisted contact structures on 3-manifolds [E]. Recently, there appeared another algorithmic existence result in [BS]. This argument is significantly shorter than both, relying instead entirely on Section 2.

It is a classical result that any continuous map from a smooth, closed oriented 4-manifold M to the 2-sphere becomes a smooth map whose critical locus consists of a smooth disjoint union of circles in M, stratified into a collection of definite and indefinite fold arcs which meet at cusp points. The main result of [S] states that the definite locus may be eliminated by homotopy, resulting in what was named

an *indefinite Morse 2-function* in [GK2]. This section is about indefinite Morse 2-functions from a closed, orientable 4-manifold M to S^2. The critical locus of such a map may have more than one path component and its image may have transverse double points.

Definition 3.1. An *equivalence* f_t of indefinite Morse 2-functions f_0, f_1 is a deformation which begins with f_0 and ends with f_1. A Morse 2-function is *simple* if its critical set is connected, and it is injective on its critical set.

Theorem 3.2. For any indefinite Morse 2-function $f\colon M \to S^2$, there is an algorithm that produces an equivalence between f and a simple indefinite Morse 2-function.

Proof. The algorithm is to eliminate the crossings between indefinite fold arcs (Lemma 3.3), then unify the resulting collection of circles without introducing further crossings (Lemma 3.4). Figure 7 gives the recipe for Lemma 3.3, and its explanation is the main content of the section. $\square$

Lemma 3.3 (Resolving crossings)**.** The movements suggested by Figure 7 come from an equivalence.

Proof. The following argument uses Proposition 2.11, so the first step is to make sure the assumption on genus in that proposition is satisfied, using what will be called *genus inflation* in Remark 3.5 below. Genus inflation begins by first performing birth moves, then performing finger moves (Proposition 2.5(1a)), cusp-fold crossings (Proposition 2.7), and Reidemeister-3 fold crossings (Proposition 2.9) so that the lunes in S^2 introduced by the birth moves (that is, the disks in S^2 bounded by the images of the critical circles) have sufficient overlap to increase the genus of all fiber components to at least 2. The use of Proposition 2.9 in that process has its own fiber genus requirement, which can be satisfied by the following construction: For a chosen fiber component F of the fibration map f, choose an open 2-dimensional disk $D \subset M$ of regular points passing through F and such that $f|_D$ is a diffeomorphism onto a neighborhood of Δ (such D exists because any transverse intersection between a 2-dimensional disk and crit f is empty). Then there is a neighborhood νD of D on which f is a submersion, so it is possible to perform a birth move on $f|_{\nu D}$ whose critical circle has image approximating $f(\partial D)$, increasing the fiber genus in Δ by a self-connect sum at each point $D \cap f^{-1}(x)$, $x \in \Delta$.

Figure 7a depicts a fold crossing, with arrows as in [GK2] indicating that the restriction of f to the preimage of each indicated transverse oriented arc is a Morse function with a single index 2 critical point. In a move inspired by [Le, Figure 10], perform two flips to obtain 7b, followed by a cusp merge that exists as discussed in [BH, Le] to obtain 7c. Figure 7d is an application of Proposition 2.11 according to the dotted line in Figure 7c.

In this paragraph, let f_0 denote the fibration depicted in Figure 7c and let φ_0 denote the fold arc forming part of $\partial\Delta$, with the rest of $\partial\Delta$ depicted by the dotted line in that figure. As with any movement of a fold arc, the movement depicted in Figures 7c-7d is encoded by the ascending manifold Z of φ_0 over Δ. This movement is genus-increasing, so Z intersects each fiber over $\Delta \setminus \varphi_0$ in two points. As discussed in the previous section, a change in Z is equivalent to a change in the horizontal distribution $\mathcal{H}$ used to define the deformation which moves φ_0 (see also [BH, Lemma 4.6]). With this in mind, the claim is that it is possible to choose Z

a. A crossing.   b. Two flips.   c. Cusp merge.   d. Prop. 2.11

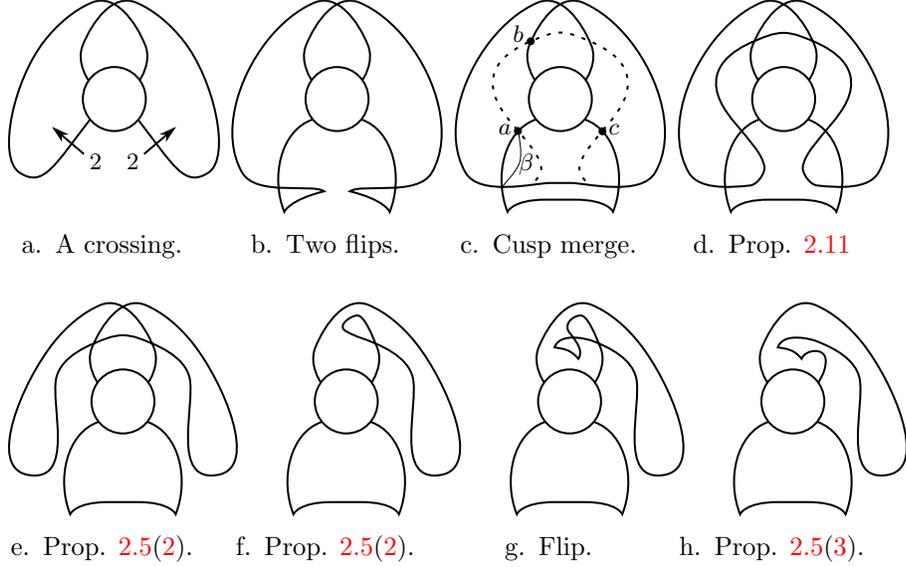e. Prop. 2.5(2).   f. Prop. 2.5(2).   g. Flip.   h. Prop. 2.5(3).

FIGURE 7. Resolving a crossing between the images of two fold arcs. These are base diagrams, where S^2 is punctured at a point near the crossing and the rest of the critical image is in the circle in each figure. The dots and the curve β in 7d correspond to parts of $\partial\Delta$ along which the suggested movement is required to satisfy an extra condition.

for the movement depicted in Figures 7c-7d so that the matching requirement of Proposition 2.5(2) is satisfied for the movements in Figures 7d-7f. To prove this claim, consider the lune containing the smooth injective curve $\beta\colon [0,1] \to S^2$ in Figure 7c. The left side is a fold arc and the right side comprises two arcs, one in φ_0 and the other in $\partial\Delta$. Foliate this lune by line segments $\ell_t(s)\colon [0,1] \to S^2$ transverse to β, where $\ell_t \cap \beta = \beta(t)$. Parameterize ℓ_t and β so that $\ell_t(0) \in \partial\Delta$ for $t \geq 1/2$. In the fiber $f_0^{-1}(\beta(1/2))$, let C_t be the vanishing cycle measured from $\ell_t(1)$, using a reference path that travels along β from $\beta(1/2)$ to $\beta(t)$ and then along ℓ_t. Let $(z_1, z_2)_t$ be the pair of points that run into either Z or the critical point mapping to $\ell_t(0)$ using the analogous reference path. This data specifies a family $\mathcal{F}_t = (C, z_1, z_2)_t$, $t \in [0,1]$. In other words, $\mathcal{F}_t$ is a circle and two points, with each one moving around the regular fiber by isotopy, perhaps intersecting each other, with the paths taken by z_1 and z_2 for $t > 1/2$ dictated entirely by the placement of Z. Now isotope Z in fibers near $f_0^{-1}(a)$, changing $\mathcal{F}_t$ for $t \in (1/2, 1)$, so that $\mathcal{F}_t$ is isotopic to a family $\mathcal{F}'_t$ satisfying $\mathcal{F}'_t = \mathcal{F}'_{1-t}$. With this achieved, the required matching condition is satisfied, and there are analogous modifications near the points x and y for the other two applications of Proposition 2.5(2) depicted in 7d-7f.

A homotopy inducing the rest of the movements (7f-7h) exists by the existence of the flipping move and Proposition 2.5(3).   □

Lemma 3.4. The critical circles from Lemma 3.3 can be unified without introducing new intersection points.

Proof. Given a fibration that is injective on its critical locus and has more than one critical circle, there is a pair of circles S_1 and S_2 of critical values bounding disjoint embedded open disks D_1, D_2, respectively, with a smooth embedded path β of regular values connecting $p_1 \in S_1$ to $p_2 \in S_2$.

If β lies on the higher-genus side of each circle, then the circles can be unified by two flips and a cusp merge as in Figures 7a-7c. Each disk D_i now has one crossing in its boundary coming from the flips. These can be eliminated by a flip and an application of Lemma 2.11 followed by 2.5(3), as in Figures 7e-7h (D_i corresponds to the lune in Figure 7g which is gone in Figure 7h).

If β lies on the lower-genus side of S_i, pick a small neighborhood $\nu p_i \subset S^2$, define $\Delta = S^2 \setminus (\nu p_i \cup D_i)$ and apply Proposition 2.11. This causes β to lie on the higher-genus side of S_i. In this way, reduce to the previous case. $\square$

Remark 3.5. It is not necessary for this paper, but it seems likely that the modifications subsequent to genus inflation up to this remark could be performed using homotopies whose support is disjoint from that of the genus inflation. In that case, the genus inflation would have been unnecessary and Example 2.10 would present a non-issue.

4. Refined uniqueness results for surface diagrams

This section uses a number of ideas and results from [W2], which in turn uses results from Section 2. It also relies on the excellent papers of Behrens and Hayano [H, BH] as a starting point for its main results; some familiarity with these three papers will be necessary to follow its arguments. Terminology from [GK2] is used as much as possible, though the content of that paper is not required for this section.

Continuing the theme of this paper, this section is about modifying certain fibration maps using existence results for homotopies of those maps. These modifications are used to better understand how much variety is available to the collection of *surface diagrams* coming from a fixed homotopy class of maps $M \to S^2$.

Definition 4.1. Suppose $f \colon M \to S^2$ is an indefinite Morse 2-function which is injective on its critical set, and its critical set is connected. Choose a reference fiber Σ on the higher genus side and let $\Gamma = (\gamma_i)_{i \in \mathbb{Z}/k\mathbb{Z}}$ be the cyclically indexed sequence of vanishing cycles up to isotopy measured from Σ. If the genus of Σ is at least 3, then the pair (Σ, Γ) is called a *surface diagram* of M.

The map f in the definition above is known as a simple wrinkled fibration in [BH] or a simplified purely wrinkled fibration in [Le, W1]; this section freely uses language from [BH, GK2, W2]. It is known that every smooth closed oriented 4-manifold has surface diagrams coming from maps in every homotopy class of maps to S^2 [W1], and that any surface diagram specifies M up to orientation-preserving diffeomorphism [W1]. The uniqueness result this section aims to refine is as follows.

Theorem 4.2 ([W2])**.** Suppose $\alpha_0, \alpha_1 \colon M \to S^2$ are two homotopic simplified purely wrinkled fibrations with surface diagrams (Σ^i, Γ^i), $i = 0, 1$. Then there is a finite sequence of surface diagrams, beginning with (Σ^0, Γ^0) and ending with (Σ^1, Γ^1), obtained by performing stabilizations, handleslides, shifts, multislides, and their inverses.

In view of the goal of establishing smooth 4-manifold invariants taking surface diagrams as input, it would be useful to pare down that list. The diagrammatic

moves seem at first to be diverse and complicated, but it turns out there are re-lations between the moves in the sense that there are many sequences of moves for which it is possible to eliminate, say, all (de)stabilizations (Theorem 4.7), and one may unify some of the above moves into a move of a single type (Proposition 4.9) or into a sequence of moves of a rather simple and familiar type called a *slide*, with moves corresponding to restrictions on the choices available to build such a sequence (Theorem 4.3). With these results in mind, establishing any smooth 4-manifold invariant calculated using surface diagrams of a fixed genus largely reduces to proving invariance under slides.

Descriptions of how the four moves affect a surface diagram were first given by Behrens and Hayano as applications of more general results in [H, BH]. It will be useful for the reader to also see [W2, Section 2] for a unified, concise catalog of the moves on diagrams and their corresponding deformations, almost entirely following the previous two papers. The first result of this section involves a *slide* between two elements $\gamma_i, \gamma_j \in \Gamma$, which resembles the handleslide move for Heegaard diagrams of 3-manifolds. To slide γ_i over γ_j, choose an embedded path from a point in γ_i to a point in γ_j, whose intersection with γ_i and γ_j is precisely its endpoints, and then replace γ_i with the connect sum $\gamma_i \# \gamma_j$ in Σ specified by that path.

Theorem 4.3. The handleslide and multislide moves affect Γ by a sequence of slides. The shift move is also a sequence of slides, followed by a reordering of the elements of Γ.

Proof. In [H, Section 2.2] there appears a *surgery homomorphism*

$$\Phi_c \colon \ \mathrm{MCG}(\Sigma)(c) \to \mathrm{MCG}(\Sigma'),$$

where $\mathrm{MCG}(\Sigma)(c)$ is the mapping class group of orientation-preserving diffeomor-phisms of the genus g oriented closed surface Σ that fix the unoriented isotopy class of the embedded circle c, and $\mathrm{MCG}(\Sigma')$ is the mapping class group of the genus $g-1$ surface Σ' (see also the introduction of [B1, Section 2.3], where the map is called ψ_γ). According to [BH, Proposition 6.2] and [H, Theorem 3.9] respectively, any multislide or handleslide, applied using the pair $\gamma_1, \gamma_n \in \Gamma$, is realized by apply-ing an element of $\mathrm{Ker}\,\Phi_{\gamma_1} \cap \mathrm{Ker}\,\Phi_{\gamma_n}$ to each element of $\{\gamma_2, \dots, \gamma_{n-1}\} \subset \Gamma$, which we will call the *nonstationary set*. In [BH], $\mathrm{Ker}\,\Phi_{\gamma_1} \cap \mathrm{Ker}\,\Phi_{\gamma_n}$ is denoted $\mathcal{K}(c, d)$, where $c = \gamma_1$ and $d = \gamma_n$.

For handleslides, γ_1 and γ_n are disjoint and their union is nonseparating. Con-sider the genus $g-2$ surface Σ'' obtained from replacing γ_1 and γ_n with pairs of disks whose centers are labeled v_1, v_2 for the disks coming from γ_1 and w_1, w_2 for the disks coming from γ_n. Hayano constructs an element of $\mathrm{Ker}\,\Phi_{\gamma_1} \cap \mathrm{Ker}\,\Phi_{\gamma_n}$ in [H, Section 3] as the lift (according to the pair of surgeries specified by the pair of 0-spheres (v_1, v_2) and (w_1, w_2)) to $\mathrm{MCG}(\Sigma)$ of the point pushing map defined by an oriented embedded arc $\eta \subset \Sigma''$ connecting v_i to w_j, for $i, j \in \{1, 2\}$. With this lift understood, it is easy to see that the lemma is proved for handleslides. However, there is another way to see it: In [H, Lemma 3.8], Hayano proves this lift can be factored into a product of mapping classes, each of which is a product of Dehn twists $t_{\tilde{\delta}(\eta)} \cdot t_{\gamma_1}^{-1} \cdot t_{\gamma_n}^{-1}$, where $\tilde{\delta}(\eta)$ is the obvious lift of the embedded circle $\partial \overline{\nu \eta}$ to Σ, where $\nu \eta$ is a regular neighborhood of η. One can then deduce the lemma by checking it for each of the two ways an embedded circle can intersect the triple $\tilde{\delta}(\eta), \gamma_1, \gamma_n$ as in Figure 8.
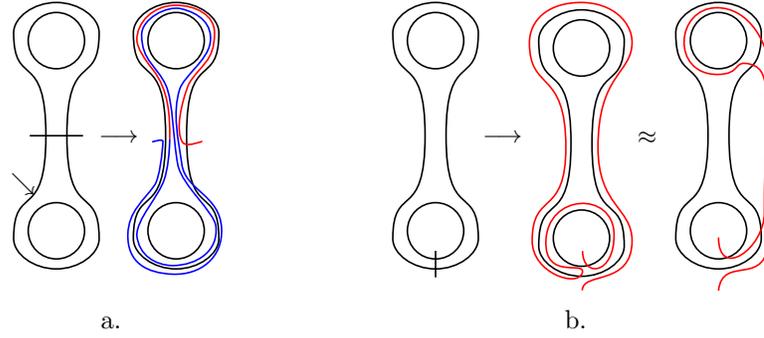
FIGURE 8. Checking that the handleslide and multislide are sequences of slides. For handleslide, the two smaller circles are interchangeably γ_1 and γ_n and the larger circle is $\tilde{\delta}(\eta)$. For multislide, a is the lower circle, $\partial\Sigma^\circ$ is the upper circle, and the larger circle is $a\#_\zeta\partial\Sigma^\circ$.
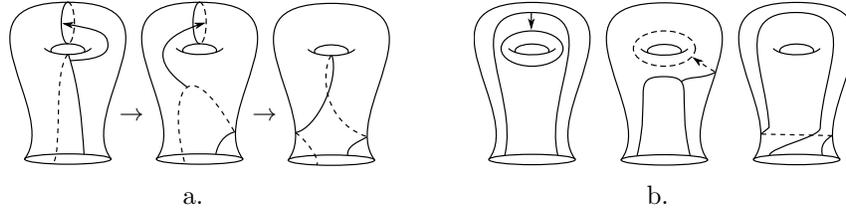

FIGURE 9. Checking that $\Delta_{c,d}$, a part of the multislide move, affects a nonstationary circle by a sequence of slides over c and d.

For multislides, γ_1 and γ_n intersect at one transverse point. By [BH, Lemma 5.4], $\mathrm{Ker}\,\Phi_{\gamma_1} \cap \mathrm{Ker}\,\Phi_{\gamma_n}$ is generated by the mapping classes $\Delta_{\gamma_1,\gamma_n} = (t_{\gamma_1}t_{\gamma_n})^3$ and $\vartheta_{a,\zeta} = t_a t_{a\#_\zeta\partial\Sigma^\circ}^{-1}$. The circles a and $a\#_\zeta\partial\Sigma^\circ$ for the Dehn twists in the factorization of $\vartheta_{a,\zeta}$ are constructed as follows. Let $\Sigma^\circ \subset \Sigma$ be the complement of a regular neighborhood of $\gamma_1 \cup \gamma_n$. Then a is an arbitrary simple closed curve in the interior of Σ° and $a\#_\zeta\partial\Sigma^\circ$ is the simple closed curve given by the connected sum of a and $\partial\Sigma^\circ$ along an embedded arc ζ whose interior is disjoint from a. It is straightforward to use Alexander's method [FM, Proposition 2.8] to show that $\Delta_{\gamma_1,\gamma_n}$ is isotopic to a half-twist about $\partial\Sigma^\circ$, and Figure 9 shows that such a mapping class has the effect of slides over γ_1 and γ_n on any nonstationary circle. As for $\vartheta_{a,\zeta}$, if a nonstationary circle intersects ζ, the effect of $\vartheta_{a,\zeta}$ is to apply a slide over $\partial\Sigma^\circ$ as specified by the subarc of ζ connecting that circle to $\partial\Sigma^\circ$ as in Figure 8a, where $\partial\Sigma^\circ$ is the upper small circle and $a\#_\zeta\partial\Sigma^\circ$ is the larger circle. Such a slide is realized by two slides over γ_1 (and also by two slides over γ_2).

The last case is the shift move. To make it easier to read, the rest of this proof uses the indexing conventions of [BH, Section 6.2], in which $\Gamma = (\gamma_1, \ldots, \gamma_l)$, the initial fold merge occurs between the fold arcs corresponding to elements $\gamma_k, \gamma_l \in \Gamma$ for $1 < k < l$ and the cusp merge involves the cusp between the fold arcs corresponding to γ_1 and γ_l. Like multislide, shift is a move one can perform for any pair γ_1, γ_k that intersects at one transverse point, and according to [BH, Lemmas
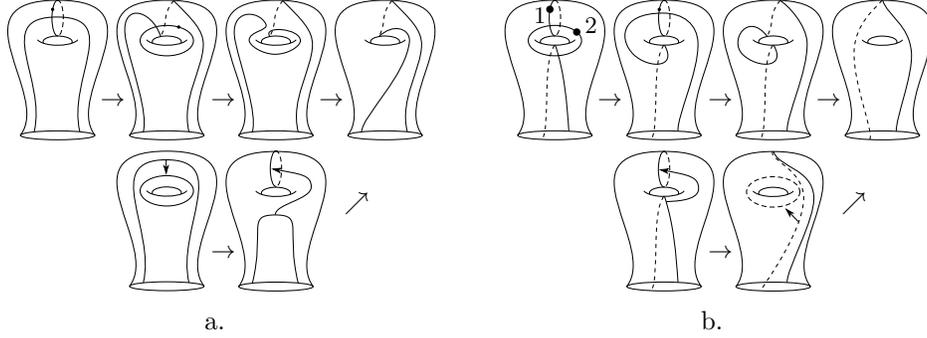
a.   b.

FIGURE 10. Realizing $(t_{\gamma_k} t_{\gamma_l} t_{\gamma_k})^{2m+1}$, part of the shift move, by slides over γ_k and γ_l: (10a) is the case in which an arc of a non-stationary circle initially intersects γ_k; (10b) is the case in which it initially intersects γ_l.
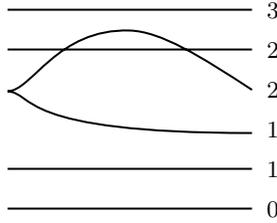

FIGURE 11. A cerf graphic corresponding to a stabilization followed by handleslide in a Heegaard diagram.

6.3 and 6.6] it sends

$$(\gamma_1, \ldots, \gamma_k, \ldots, \gamma_l) \mapsto (\gamma_1, \ldots, \gamma_k, \gamma_l, \chi_i(\gamma_{k+1}), \ldots, \chi_i(\gamma_{l-1})),$$

where i can be chosen from $\{1, 2\}$. Toward the end of the proof of [BH, Lemma 6.6], the mapping class χ_i has the formula

$$\chi_i = \varphi_i \cdot (t_{\gamma_k} t_{\gamma_l} t_{\gamma_k})^{2m+1},$$

for some $m \in \mathbb{Z}$. Figure 10 shows that $(t_{\gamma_k} t_{\gamma_l} t_{\gamma_k})^{2m+1}$ affects nonstationary circles by a sequence of slides, so it remains to prove the same for φ_i. According to [BH, Lemma 6.4] and the paragraph directly above [BH, Lemma 6.4], the map φ_i is constructed as follows. First surger Σ along γ_l, replacing that circle with two marked points v_1, v_2 to obtain a surface Σ'_{γ_l}. Since γ_1 and γ_k each intersected γ_l in Σ unique transverse points, they appear as embedded arcs γ'_1, γ'_k in Σ'_{γ_l}, and each has the endpoints v_1, v_2. Together these arcs form a circle containing v_1 and v_2, and by slight perturbations they obtain two loops, $\eta_1 \subset \Sigma'_{\gamma_l} \setminus \{v_2\}$ and $\eta_2 \subset \Sigma'_{\gamma_l} \setminus \{v_1\}$, where η_i is based at v_i and is oriented so that it leaves v_i along γ'_k and returns along γ'_1. With that understood, [BH, Lemma 6.4] can be restated to say that, to apply φ_i, first replace γ_l with a pair of marked disks to obtain Σ'_{γ_l}, then apply a point-pushing map along η_i to the nonstationary circles (or arcs, if they intersected γ_l in Σ), then perform surgery along v_1, v_2 to get Σ back. This is clearly a collection of slides over γ_l.   □

It is known that there are pairs of Heegaard splittings of closed 3-manifolds that require stabilization before they become isotopic. In other words, there are pairs of Heegaard diagrams D, D' whose diffeomorphism types can be connected by a finite sequence of handleslides and isotopies of individual α and β circles only after some positive number of stabilizations, even when D and D' have the same genus. One result below, Theorem 4.7, states this is not the case for surface diagrams. Here are a few observations that might make this less surprising. In [HTT] and every other construction known to the author, two Heegaard diagrams for the same 3-manifold coming from non-isotopic splittings of the same genus are related by switching the roles of the α and β circles (equivalently, replacing a compatible Morse function $f \colon Y \to [0,1]$ with $1 - f$). This is a modification that does not exist for surface diagrams. It also seems that the obstruction to performing R_2 deformations discussed in [W2, Example 2.8] could be exploited to prove a similar non-isotopy result, working with the Cerf graphic of a carefully chosen deformation of Morse functions: Here a stabilization or its inverse occurs at every cusp, and handleslides occur precisely as instances of the mentioned obstruction (see [La, Section 3] for discussion of Cerf graphics). This is another kind of behavior that would not apply to surface diagrams. Finally, in the proof for Theorem 4.7 below, a crucial step is to move a stabilization to occur before or after any genus-preserving move by a homotopy of the deformation; such a technique is not generally available to Cerf graphics (see for example Figure 11, a Cerf graphic corresponding to the stabilization of a genus-one Heegaard diagram, introducing circles α and β, immediately followed by a handleslide involving α or β. Such a modification is impossible there simply because of the number of attaching circles: No handleslides can be performed without first performing a stabilization).

One analogue of a Cerf graphic for a family of surface diagrams appears in Figure 12. That figure depicts the critical set of a deformation, that is, a copy of $S^1 \times [0,1]$ smoothly properly embedded into $M \times [0,1]$. Any deformation with this decorated critical surface is a stabilization deformation followed by handleslide and destabilization deformations (see also [W2, Figure 4]). The dots are swallowtails, the heavier arcs are cusp arcs, and the lighter arcs are the preimage of fold crossings in $S^2 \times [0,1]$ (Figure 14a is a depiction of the same critical surface omitting the cusps). See [W2, Sections 3.1 and 3.5.1] for an introduction to such figures and results concerning movements of their decorations. The rest of this paper freely uses ideas and terminology from those sections.
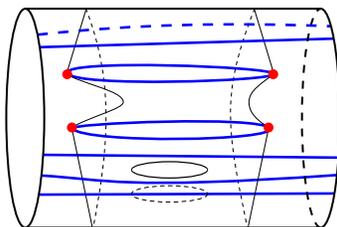

FIGURE 12. The decorated critical surface $S^1 \times [0,1] \subset M \times [0,1]$ for a stabilization deformation followed by handleslide and finally a destabilization at the same fold arc as the stabilization. Here, t increases to the right.

We begin with a few results used in the proof of Theorem 4.7. The first implies that the location of a stabilization is irrelevant to the resulting surface diagram up to diffeomorphism.

Proposition 4.4. Enumerate the fold arcs of an SPWF counter-clockwise according to $\mathbb{Z}_k$ as $\varphi_1, \varphi_2, \ldots, \varphi_k$ and suppose α is a stabilization deformation whose initial flips occur at φ_n. Then α is homotopic, relative to its endpoints, to a stabilization whose initial flips occur on φ_{n+1}.

Proof. In [W2, Figure 23] there appears a way to change the fold arc on which a flip occurs by a homotopy that fixes the endpoints of the given deformation. Use this to move the swallowtails into the desired region of fold points in crit α. The immersion arcs can be pushed in a genus-increasing manner so that the R_2 deformation that concludes the stabilization deformation occurs between the fold arc bounded by the two swallowtails and some other fold arc: This movement of the immersion locus merely corresponds to moving cusps into the higher-genus side of an intersecting fold arc at earlier values of t than in the original deformation. $\square$

Proposition 4.5. If α is a stabilization deformation followed by a destabilization deformation, then α is homotopic to a constant deformation.

Proof. Use Proposition 4.4 to arrange for the stabilization to occur at the same fold arc as the destabilization, and use [W2, Lemma 3.13] to isotope the two immersion arcs corresponding to the destabilization forward and backward in t so that their t-minimal points (the pair of points at which the second Reidemeister-II fold crossing occurs) lie in the same fold arc as the t-maximal points coming from the destabilization. In this way, the Reidemeister-II fold crossings for the stabilization and destabilization occur between the same pair of fold arcs; Figure 13 gives base diagrams for the resulting deformation.

As arranged in the previous paragraph, near the end of the stabilization deformation (at step 3), a fold arc φ cancels its two intersections with another fold arc in a genus-increasing R_2 deformation. In step 4, φ re-acquires its intersections at the beginning of the destabilization deformation. During steps 3 and 4, a reference fiber chosen above p with reference arcs crossing φ gives a surgered surface diagram that is unchanged during those steps (see [W2, Remark 1.7] or [BH, Section 6.1] about surgered surface diagrams). For this reason, the interval in t between steps 3 and 4 can be eliminated by homotopy of α, resulting in a new deformation with
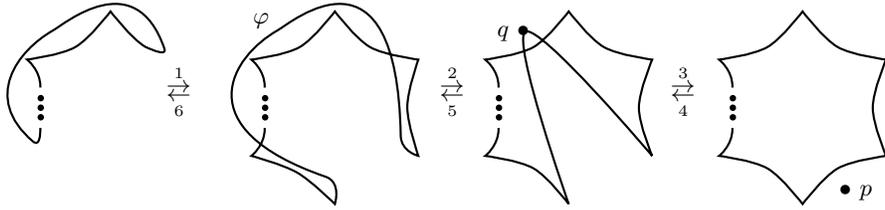

FIGURE 13. Base diagrams for a stabilization deformation followed by a destabilization deformation. Steps 1 and 6 involve flips, the rest are 2-parameter crossings between the fold arc φ and others.

steps 1-2-5-6. Going further, steps 2 and 5 can similarly be eliminated, using the same argument with one additional reference path from the same reference fiber that goes around the back of the sphere to the point q. What remains is a pair of flipping moves, followed by a pair of inverse flipping moves, which can also be eliminated by homotopy of α. To explain this last step, choose a smooth path γ in the fold locus of α connecting two of the swallowtail points such that $\gamma(t) \in \mathrm{crit}\, \alpha_t$. For example, $\alpha(\gamma)$ could coincide with a crossing introduced by one flip at the beginning of $\alpha(\gamma)$, (this crossing is then eliminated by the inverse flip at the end of $\alpha(\gamma)$). According to the local model of a swallowtail as a deformation of maps $B^4 \to B^2$, the 2-cusped loop introduced at the beginning of γ bounds a one-parameter family of disks in S^2, each of which is the image of a 4-ball neighborhood of a point of γ. These 4-balls trace out a neighborhood U of γ such that the restriction $\alpha|_U$ is simply a flip followed by its inverse. Such a deformation, called a *detour* in this paper and in [W2, Definition 2.10], is clearly a nullhomotopic loop in the space of maps $M \to S^2$. $\qquad\square$

Proposition 4.6. *If α is a destabilization deformation followed by a stabilization deformation then α is homotopic to a sequence of handleslide deformations.*

Proof. Apply Proposition 4.4 to arrange for the stabilization to occur at the same fold arc as the destabilization. Base diagrams for the result start at the right of Figure 13, then proceed 4-5-6-1-2-3. Now apply the modification from [W2, Figure 21] to cancel the four swallowtails. The result is what is called a *genus-decreasing immersion pair* in [W2, Definition 3.11]. Use [W2, Lemma 3.16(b)] to break this pair into many smaller genus-decreasing pairs, each of which is disjoint from the cusp locus. Because the supports of each pair of resulting R_2 deformations are disjoint (they occur in the preimages of distinct regions in S^2), these R_2 deformations can be performed in sequence, yielding a sequence of handleslides. $\qquad\square$

Theorem 4.7. *Suppose two surface diagrams (Σ, Γ) and (Σ, Γ') are related by one of the following sequences of moves.*

1. *Stabilize, then perform genus-preserving moves, then destabilize.*
2. *Destabilize, then perform genus-preserving moves, then stabilize.*

Then there is a sequence of genus-preserving moves relating (Σ, Γ) and (Σ, Γ').

Proof. For both cases the argument is to prove it is possible to switch the order in which the model deformations for (de)stabilizations and genus-preserving moves appear in α without changing the fibration maps α_0 and α_1, then to apply Propositions 4.5 and 4.6.

Case 1. First, the argument for handleslide. As remarked in the proof of Theorem 4.3 and [H, BH, W2], these moves (using vanishing cycles γ_1 and γ_n) affect a *nonstationary set* of vanishing cycles $\{\gamma_2, \ldots, \gamma_{n-1}\}$ by a lift to $MCG(\Sigma)$ of a pushing map in $MCG(\Sigma')$, where Σ' is the result of capping off $\Sigma \setminus (\gamma_1 \cup \gamma_n)$ with two disks (one disk for multislide and shift). The pushing maps come from a choice of properly embedded arc (call it $\tilde{c}$) in the surface $\Sigma \setminus (\gamma_1 \cup \gamma_n)$. The move affects an element of the nonstationary set by sliding over γ_1 or γ_n (or both) along paths that are parallel to $\tilde{c}$, and a nonstationary vanishing cycle is modified by the move exactly when its image in $\Sigma \setminus (\gamma_1 \cup \gamma_n)$ intersects $\gamma_1 \cup \gamma_n \cup \tilde{c}$. For this reason, copies of the image of $\tilde{c}$ will appear in any nonstationary vanishing cycle that is modified by the move. Since the surface diagram destabilizes after the genus-preserving

move, there is a fold arc φ that sweeps across all of the higher-genus region of S^2 in a genus-decreasing R_2 deformation. Because of this, for some chosen reference fiber in the higher genus side of the fibration just before destabilizing, Theorem 2.4 implies the vanishing cycle γ_i of this fold arc is disjoint from all the other vanishing cycles except $\gamma_{i\pm1}$ up to isotopy. Use Proposition 4.4 to cause φ to also be the fold arc that sweeps across the back of S^2 in the stabilization deformation. Now γ_i is disjoint from the image of $\tilde{c}$ and all other vanishing cycles except $\gamma_{i\pm1}$ before and after the handleslide. For this reason, the supports of the stabilization deformations and the handleslide deformation are disjoint and the deformations' effects on Γ commute: the handleslide can be moved into the lower-genus side of either of the stabilization deformations.

The multislide deformation similarly corresponds to a choice of loop $\tilde{c}$ based at a point in the circle $\partial(\Sigma \setminus (\gamma_1 \cup \gamma_n))$, along which nonstationary vanishing cycles are slid over γ_1 and γ_n, in addition to the half-twists about $\partial(\Sigma \setminus (\gamma_1 \cup \gamma_n))$ considered in Figure 9. As before, use Proposition 4.4 to cause both the stabilization and destabilization to occur on a fold arc such that the image of its vanishing cycle in $\Sigma \setminus (\gamma_1 \cup \gamma_n)$ is disjoint from the circles $\tilde{c}$ and $\partial(\Sigma \setminus (\gamma_1 \cup \gamma_n))$, allowing the multislide deformation to move into the lower-genus side of, say, the stabilization deformation.

As for shift, essentially the same argument applies, except that there is no choice of curve in $\Sigma \setminus (\gamma_1 \cup \gamma_n)$ – the only requirement is for φ to not be one of the fold arcs involved in either merge move. Once this is achieved, apply Proposition 4.5 to cancel the pair of stabilization deformations. This concludes Case 1.

Case 2. This is easier than the previous case, because it is a genus-increasing modification of the deformation in the sense that the genus-preserving moves are all transferred into higher-genus diagrams. There is initially a vanishing cycle γ_i, disjoint from all others except $\gamma_{i\pm1}$, whose corresponding fold sweeps across S^2 in a genus-decreasing direction for the destabilization. After the destabilization deformation (following [H, Theorem 4.1], in which γ_i is called d), γ_i survives as the endpoints of a short arc (which we will call a, but Hayano calls α) which is transverse to the vanishing cycle on which the inverse flipping moves occurred: Choosing a horizontal distribution for the fibration structure over $[0,1] \times S^2$ given by α, this fiberwise pair of points is precisely the set of points that runs into φ in the direction $-\partial/\partial t$. Generically, the support of the genus-preserving moves is disjoint from a, so the deformation for the genus-preserving move may be moved backward in t to occur before the destabilization deformation. Once this is achieved, apply Proposition 4.6 to the remaining destabilization-stabilization deformation. $\square$

The next two results of the paper eliminate handleslides in favor of stabilizations and consolidate the other two genus-preserving moves.

Proposition 4.8. Any handleslide deformation α is homotopic to a stabilization deformation followed by an inverse stabilization deformation, so that handleslides may be omitted from the list in Theorem 4.2.

Proof. Using language from [W2], the handleslide deformation is defined to be a genus-decreasing R_2 deformation that introduces two intersection points in the critical image, immediately followed by another R_2 deformation that cancels the intersections. The first step in the proposed homotopy of α is to smoothly insert a
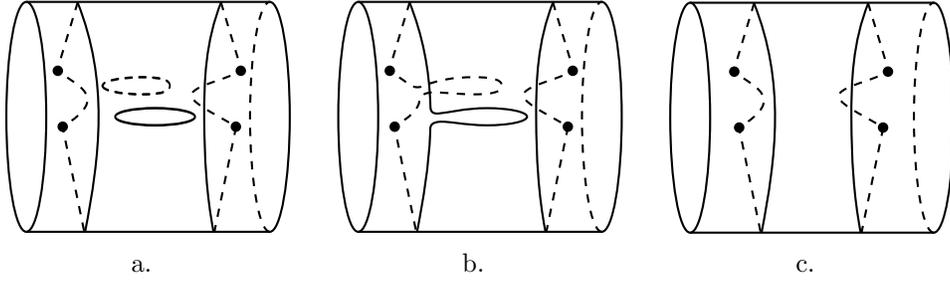
FIGURE 14. Including a handleslide deformation into a stabilize-destabilize pair.

deformation into α at a t-value t_0 immediately before the initial R_2 deformation occurs: A stabilization deformation, immediately followed by its reverse. The inserted deformation, called a *detour* in [W2, Definition 2.10], forms a loop in the space of maps $M \to S^2$ (based at α_{t_0}) which is clearly nullhomotopic, so the proposed modification can be achieved by a homotopy of α. As in Case 2 of Theorem 4.7, move the handleslide deformation into the higher-genus region between the two stabilizations, resulting in Figure 14a. The first stabilization deformation ends with an R_2 deformation canceling intersections between two fold arcs, and the handleslide deformation begins with an R_2 deformation that creates two intersections between a pair of fold arcs. It is possible to choose the first stabilization above so that these pairs of fold arcs coincide. With this understood, consider the image of these fold arcs in S^2 at the end of the stabilization and the beginning of the handleslide: It is the contraction and immediate re-formation of a lune of regular values bounded by those two fold arcs. The next step is a homotopy of α in which this lune disappears for decreasing t-increments until the two fold arcs do not separate at all (at least, until the R_2 deformation coming from the end of the handleslide deformation). This results in Figure 14b, which becomes clearly a stabilization-destabilization pair after possibly applying [W2, Proposition 3.10], yielding Figure 14c.   □

In [BH, Section 4.1.2], there appeared a *generalized shift deformation.* Proposition 4.5 of the same paper gave its effect on a surface diagram. The following proposition can also be interpreted as a converse of [W2, Remark 2.9].

Proposition 4.9. The shift and multislide deformations in Theorem 4.2 are homotopic to sequences of generalized shift deformations or length 1 and 2, respectively.

Proof. Since shift deformations are a special case of generalized shift deformations, it remains to show that any multislide deformation can be decomposed into a pair of generalized shift deformations. This is easy to arrange: At a point t_0 just after the initial fold merge, which produces two cusps c_1 and c_2, perform an arbitrary cusp merge between c_1 and a cusp which is adjacent to c_2, then immediately reverse this cusp merge, and continue on with the deformation as it proceeded originally. The first and second pair of merge moves now each form a generalized shift deformation. The inserted deformation is a detour, so the proposed modification can be achieved by a homotopy of α.   □

In [H], Hayano describes how a stabilization deformation affects a surface diagram. The description explicitly makes use of the fold arc on which the initial flips

occur, though according to the following proposition, one can change this choice of fold arc without changing the diffeomorphism class of the resulting diagram.

The last result gives a simple criterion for when a surface diagram destabilizes.

Proposition 4.10. The surface diagram (Σ, Γ) destabilizes if and only if Γ has an element γ_i which is disjoint from all other elements except $\gamma_{i\pm 1}$.

Proof. In this case, [W2, Propositions 2.5(4) and 2.7] imply the existence of a deformation corresponding to steps 4 and 5 of Figure 13. The vanishing cycles measured from a reference fiber on the higher-genus sides of a pair intersecting folds are disjoint, and the vanishing cycles on either side of a cusp must intersect transversely at a single point so the vanishing cycles for the three fold arcs in each of the two triangles after step 5 (or, indeed, any such triangle) appear as in [W1, Figure 5] and so are suitable for inverse flipping moves that complete the destabilization. □

References

[ADK] D. Auroux, S. Donaldson, and L. Katzarkov, Singular Lefschetz pencils, *Geom. Topol.* **9** (2005), 1043-1114. doi:10.2140/gt.2005.9.1043 MR2140998

[AK] S. Akbulut and Ç. Karakurt, Every 4-manifold is BLF, *J. Gökova Geom. Topol.* **2** (2008) 40-82. MR2466002

[B1] R. I. Baykur, Topology of broken Lefschetz fibrations and near-symplectic 4-manifolds, *Pac. J. Math* **240** No. 2 (2009), 201-230. doi:10.2140/pjm.2009.240.201 MR2485463

[B2] R. I. Baykur, Existence of broken Lefschetz fibrations, *Int. Math. Res. Notices* **2008** (2008). doi:10.1093/imrn/rnn101 MR2439543

[BS] R. I. Baykur, O. Saeki, Simplifying indefinite fibrations on 4-manifolds, 2017 preprint.

[Be] S. Behrens, On 4-manifolds, folds and cusps, *Pac. J. Math* **264**(2) (2013), 257-306. doi:10.2140/pjm.2013.264.257 MR3089398

[BH] S. Behrens and K. Hayano, Elimination of cusps in dimension 4 and its applications, *Proc. London Math Soc.* **113** (5), 674-724. doi:10.1112/plms/pdw042 MR3570242

[E] Y. Èliashberg, Classification of overtwisted contact structures on 3-manifolds, *Invent. Math.* **98**(3) (1989) 623-637. MR1022310

[Et] J. Etnyre, Lectures on open book decompositions and contact structures, from: "Floer homology, gauge theory, and low-dimensional topology", (D. Ellwood, P. Ozsvth, A. Stipsicz, Z. Szabó, editors), Clay Math. Proc. **5**, Amer. Math. Soc. (2006) 103-141 MR2249250

[FM] B. Farb and D. Margalit, A Primer on Mapping Class Groups, Princeton Mathematical Series **49**, Princeton University Press, Princeton, NJ 2012. MR2850125

[GK1] D. Gay and R. Kirby Constructing Lefschetz-type fibrations on 4-manifolds, *Geom. Topol.* **11** (2007), 2075-2115. DOI 10.2140/gt.2007.11.2075 MR2350472

[GK2] D. Gay and R. Kirby, Indefinite Morse 2-functions; broken fibrations and generalizations, *Geom. Topol.* **19** (2015), 2465-2534. doi:gt.2015.19.2465 MR3416108

[HTT] J. Hass, A. Thompson, W. Thurston, Stabilization of Heegaard splittings, *Geom. Topol.* **13** (2009), 2029-2050. doi:10.2140/gt.2009.13.2029 MR2507114

[H] K. Hayano, Modification rule of monodromies in an R_2–move, Alg. Geom. Topol **14** (2014), 2181–2222. doi:10.2140/agt.2014.14.2181 MR3331613

[La] F. Laudenbach, A proof of Reidemeister-Singer's theorem by Cerf's methods, Annales de la Faculté des sciences de Toulouse Mathématiques XXIII, **1** (2014) 197–221. doi:10.5802/afst.1404 MR3204738

[Le] Y. Lekili, Wrinkled fibrations on near-symplectic manifolds, *Geom. Topol.* **13** (2009), 277-318. doi:10.2140/gt.2009.13.277 MR2469519

[S] O. Saeki, Elimination of definite fold, *Kyushu J. Math.* **60** (2006), 363-382. doi:10.2206/kyushujm.60.363 MR2268242

[W1] J. Williams, The h-principle for broken Lefschetz fibrations, *Geom. Topol.* **14** no. 2 (2010), 1015-1061. doi:10.2140/gt.2010.14.1015 MR2629899

[W2] J. Williams, Uniqueness of surface diagrams of smooth 4-manifolds, 2011 preprint.

Department of Mathematical Sciences, Binghamton University

Email address: jdw.math@gmail.com