

Reconstructing 4–manifolds from Morse 2–functions

DAVID GAY
ROBION KIRBY

Given a Morse 2–function $f: X^4 \rightarrow S^2$, we give minimal conditions on the fold curves and fibers so that X^4 and f can be reconstructed from a certain combinatorial diagram attached to S^2 . Additional remarks are made in other dimensions.

57M50; 57R45, 57R65

Let X^n and Σ^2 be smooth, oriented, closed, connected manifolds. A *Morse 2–function* $f: X^n \rightarrow \Sigma^2$ is a smooth function such that each point $x_0 \in X^n$ has a coordinate chart $\mathbb{R} \times \mathbb{R}^{n-1}$ and $f(x_0)$ has a coordinate chart $\mathbb{R} \times \mathbb{R}$ for which $f(x) = f(t, y) = (t, f_t(y))$ where $f_t: \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ is a generic 1–parameter family of smooth functions with non-degenerate critical points and births and deaths (Gay and Kirby [3]). The image of the set of critical points of f_t is often called a *Cerf graph* [1]. The set of points in X for which the rank of Df is only one forms a smoothly embedded 1–manifold Z whose image in Σ is an immersed 1–manifold Z' (with cusps) of critical values of f . Z' is called the set of *fold curves* in Σ . Under appropriate conditions on the homotopy class of the Morse 2–function f , in particular when $\Sigma = S^2$, f is homotopic to a Morse 2–function for which all fibers are connected and all critical points are indefinite [3]. That is, in local coordinates, $f(t, y) = (t, -y_1^2 \pm y_2^2 \pm \cdots \pm y_{n-2}^2 + y_{n-1}^2)$. From now on we assume all Morse 2–functions are indefinite and fiber-connected.

In the case $n = 4$ which we focus on in this paper, all the fibers (preimages of points under f) are oriented surfaces, and when crossing a fold curve in Σ the genus changes by one. On the higher genus side is a circle in the fiber to which a 2–handle is attached (we call this the *attaching circle* for the fold) and on the lower genus side are a pair of points to which a 1–handle is attached; these are the ascending and descending spheres of the critical point of f_t at a fold curve. At a cusp, the pair of circles, one from each side of the cusp, must meet in exactly one point in the fiber; at a crossing of fold curves, they must be disjoint (see Figure 1).

Theorem 1 *Suppose that all the fibers of a given Morse 2–function $f: X \rightarrow \Sigma$ have genus > 1 , and all the regions in Σ bounded by fold curves are simply connected. Then the following data suffice to reconstruct X and f up to diffeomorphism:*

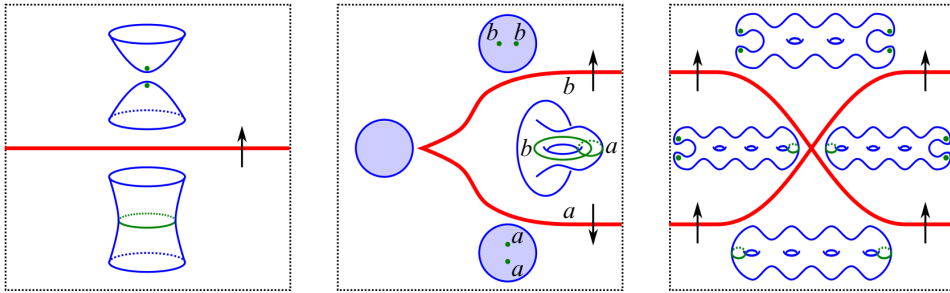


Figure 1: Local models for indefinite fiber-connected Morse 2-functions on 4-manifolds; folds are in red, fibers (or local parts of fibers) are in blue, attaching S^0 's and S^1 's are in green. Arrows transverse to folds indicate the direction of decreasing genus, ie the direction in which a fold has index 2.

- (1) The fold graph Γ , the image of the fold curves in Σ ; this is a co-oriented graph with vertices of valence 2 and 4 dividing Σ into polygonal regions; the co-orientation is in the direction of decreasing genus. When $\partial\Sigma \neq \emptyset$, the vertices should be in the interior and edges should be transverse to the boundary.
- (2) The standard fiber in each region R , a drawing of a standard genus g_R surface F_R ; this is the fiber over a point in the interior of R .
- (3) The attaching circle for each outward oriented edge e of a region R , a simple closed curve C_e drawn on F_R ; this is the attaching circle for the 2-handle attached when crossing e .
- (4) The gluing data for each outward oriented edge e of a region R , a basis of $2(g_R - 1)$ simple closed curves $A_{(1,e)}, B_{(1,e)}, \dots, A_{((g_R-1),e)}, B_{((g_R-1),e)}$ disjoint from C_e , with $A_i \cap B_i$ a single transverse point of intersection and $(A_i \cup B_i) \cap (A_j \cup B_j) = \emptyset$ when $i \neq j$; this specifies that the genus $g_R - 1$ surface obtained from surgery along C_e should be identified with the standard fiber in the adjacent region across edge e so that the basis $A_{(1,e)}, B_{(1,e)}, \dots, A_{((g_R-1),e)}, B_{((g_R-1),e)}$ maps to the standard basis for the standard genus $g_R - 1$ fiber. (Figure 2 illustrates this standard basis.)

Remark 2 Without item 4 in the list above (the gluing data) and with $\Sigma = S^2$, this is our Theorem 3 from [4]. *That theorem is in fact false.* We will present explicit counterexamples showing that, in general, one does need to record *at least some* of the gluing data. However, in the next theorem we will see that one should expect many edges across which one does *not* need to record the gluings.

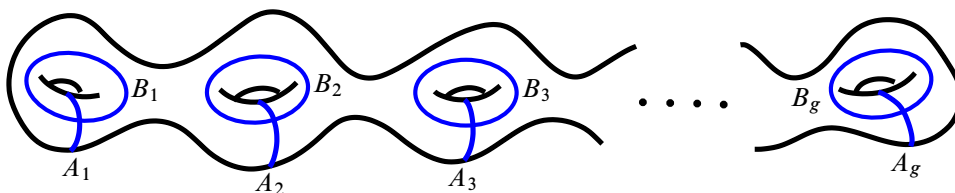


Figure 2: The standard genus g surface and the standard basis $A_1, B_1, \dots, A_g, B_g$

Remark 3 If one draws an arbitrary co-oriented 2- and 4-valent graph in a surface Σ with data as in the theorem, there are many obvious constraints that would prevent such data from describing a closed 4-manifold with a Morse 2-function f . Obvious constraints include that the genus should drop by one across each edge, that co-orientations should be consistent at 4-valent vertices, and that attaching circles at crossings should be disjoint and at cusps should intersect once. There are also much less obvious constraints on both attaching circles and gluing maps related to the fact that, after reconstructing X over a neighborhood of a dual graph, we need to be able to fill in the inverse images of disks containing the cusps and crossings.

Remark 4 It is almost immediate that the data in Theorem 1 will reconstruct (X, f) over the complement of disk neighborhoods of the cusps and crossings, and thus the real content of the theorem is that, once this reconstruction is done, the extension over the cusps and crossings is unique if it exists.

Theorem 5 *The following are situations under which we only need to record the gluing data across a subset of the edges of the fold graph Γ . See Figure 3.*

- (1) Define a chain of edges to be a sequence of edges connected by cusps, and thus all adjacent to the same two regions. For any chain of edges, the gluing data across one edge in the chain determines the gluing data across all other edges in the chain.
- (2) At a crossing involving one region of genus $g + 2$, two of genus $g + 1$ and one of genus g , consider the four edges meeting at that crossing. The gluing data across the two edges between genus $g + 2$ and genus $g + 1$ together with the gluing data across one of the edges between genus $g + 1$ and genus g determines the gluing data across the remaining edge (between genus $g + 1$ and genus g).
- (3) If R is a locally genus minimizing region, so that all the edges of R are co-oriented inwards, then we do not need to record any gluing data across any edges of R . In other words, the gluing data across all the other edges meeting R at vertices (crossings) determines the gluing data across the edges of R up to an

automorphism of the fiber F_R , and changing the gluing data across the edges of R by such an automorphism does not change (X, f) .

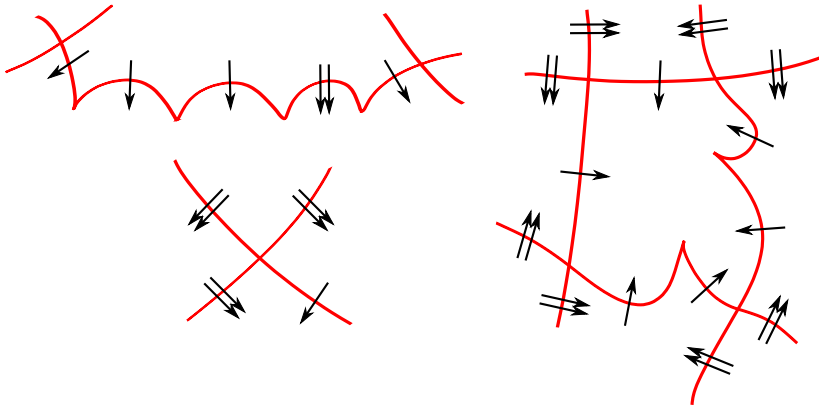


Figure 3: Cases in which we do not need to record all the gluing data; the edges across which the gluing data needs to be recorded are labelled with double transverse arrows; the single transverse arrows indicate edges for which the gluing data can be recovered from the gluing data along the double arrows.

Remark 6 Theorem 5 implies that in many cases, one only needs to record the gluing data along a rooted maximal tree T in the oriented dual graph Γ^* to Γ . The condition we need is that, in the process of adding edges to turn T into Γ^* , whenever one extends across a quadrilateral (which corresponds to a crossing in Γ), the edge being added is one of the two lower-genus edges of the quadrilateral. For many graphs such a maximal tree will not exist.

To warm up, consider a traditional Morse function $g: X^n \rightarrow \mathbb{R}$. The analogues of fold curves are critical values, and the analogues of fibers are level sets, and the analogues of circles and pairs of points are the ascending and descending spheres. This is not enough information to reconstruct X^n in general. One needs to know either

- (1) how to identify the level set just above a critical value with the level set just below the next higher critical value (this can be done with a choice of metric which gives a gradient flow), or
- (2) how to see all the attaching maps at once in the boundary (= level set) of the 0-handle, for all the handles (including even the last n -handle, as can be seen by considering exotic smooth structures on spheres).

In the special case of dimension four, X^4 can be determined by drawing the attaching maps of 1–handles as pairs of points in S^3 , the 2–handles as framed links in S^3 which may go over 1–handles, and then using the homotopy type of X^4 to determine how to add the 3–handles and 4–handle (see Laudenbach and Poénaru [5]).

When considering a circle valued Morse function, $g: X^n \rightarrow S^1$, even the information of (1) and (2) is not enough for g may be a bundle map with no critical values and hence nothing to determine the monodromy of the bundle.

Right off we see that a non-trivial circle bundle $F^2 \rightarrow M^3 \rightarrow S^1$ produces a Morse 2–function $M^3 \times S^1 \rightarrow S^1 \times S^1$ with no fold curves and therefore no information to distinguish $M^3 \times S^1$ from, say, $F \times S^1 \times S^1$.

Traditionally we define bundles $p: X^n \rightarrow \Sigma^2$ by giving an open cover $\{U_\alpha\}$ of Σ and local trivializations $h_\alpha: p^{-1}(U_\alpha) \rightarrow U_\alpha \times F^{n-2}$ together with clutching functions $\phi_{\alpha\beta}: h_\alpha p^{-1}(x) \rightarrow h_\beta(p^{-1}(x))$ for $x \in U_\alpha \cap U_\beta$, which satisfy the cocycle condition, $\phi_{\alpha\beta}\phi_{\beta\gamma} = \phi_{\alpha\gamma}$.

This definition can be extended to what might be called a broken bundle $p: S^n \rightarrow \Sigma^2$ in which Σ has the usual fold curves with their data of ascending and descending spheres specified by giving an open cover $\{U_\alpha\}$ of Σ , and local trivializations of the follow four kinds:

- (1) $h: p^{-1}(U_\alpha) \rightarrow U_\alpha \times F^{n-2}$,
- (2) $h: p^{-1}(U_\alpha) \rightarrow B$ where B is an interval cross an $(n-1)$ –dimensional bordism with one critical point,
- (3) $p^{-1}(U_\alpha)$ is the obvious preimage of a birth or death of a canceling pair of critical points,
- (4) $p^{-1}(U_\alpha)$ is the obvious preimage of a crossing of fold curves.

In addition, clutching functions must be given for the traditional case with no folds involved, or for gluing along a fold curve. Again the cocycle condition must be satisfied, and then the broken bundle, hence X^n , is well defined.

With these remarks behind us, we now give the proof of Theorem 1.

Proof Clearly the fold graph Γ and the standard fibers determine X and f over a disk in the interior of each region, the blue disks in Figure 4. Enlarge each of these disks by sending out fingers (green in Figure 4) which intersect each outward-pointing edge along an arc in the interior of the edge, but are still disjoint from the neighboring disks; the attaching circles then determine X and f over these enlarged disks because,

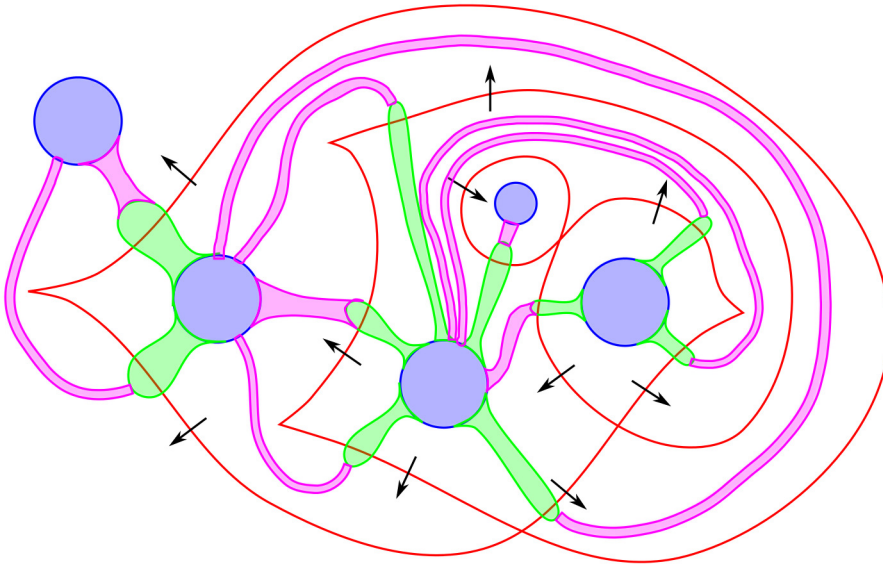


Figure 4: Using the basic data to build X , first over blue disks, then green fingers, then pink bridges

in X , adding each finger corresponds to attaching I cross a 3-dimensional 2-handle cobordism for each edge, with the 2-handle attached along the given attaching circle. The gluing data are then used to attach the tips of the fingers to the adjacent disks, along the pink bridges in Figure 4: Consider two regions R and R' meeting along an edge e co-oriented from R to R' . Thus $g_{R'} = g_R - 1$. Let $F_{R,e}$ be the result of surgering F_R along C_e . The attaching data $A_1, B_1, \dots, A_{g_R-1}, B_{g_R-1}$ form a basis for $F_{R,e}$. Meanwhile, $F_{R'}$ is drawn in a standard way which implies a standard basis, and the gluing map is then constructed to send the given basis of $F_{R,e}$ to the standard basis on $F_{R'}$.

Thus the full data determines (X, f) over the complement of disk neighborhoods of the cusps and crossings. We now need to show that the extensions over the cusps and crossings are unique. We know that the extensions exist by the assumption that this data comes from a given (X, f) .

We will first show the uniqueness of three kinds of extensions across disks, where the fibration in each case is given over half of the boundary, as in Figure 5. In the three cases in Figure 5 we have chosen local coordinates so that we have Cerf diagrams.

In Figure 5(a) and (b) we are extending an ordinary Morse function $g: M^3 \rightarrow [0, 1]$ to a stable 1-parameter family $g_t: M \rightarrow [0, 1]$ with $g_0 = g$, with a prescribed Cerf graphic for g_t , either a cusp or a crossing. In the cusp case, the uniqueness assertion is

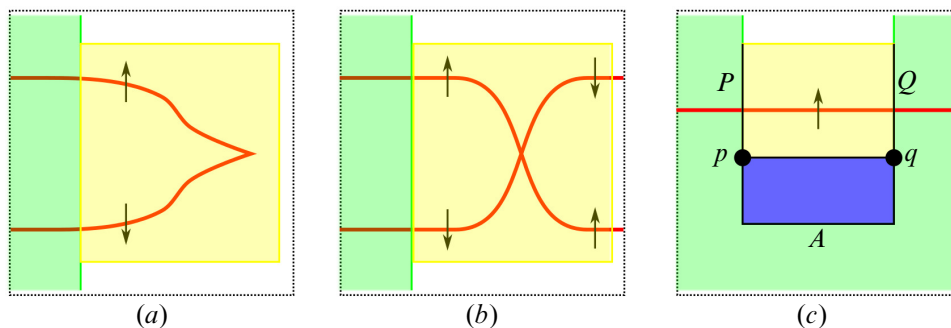


Figure 5: Three cases in which an extension across a disk is unique; in the first two cases we are extending from the green region to the yellow region; in the last case we are extending first from the green to the yellow and then to the blue.

then just the assertion that, if two critical points of index 1 and 2, with adjacent critical values, can be cancelled then there is a unique stable 1-parameter family of Morse functions achieving this cancellation. In the crossing case, the direction in which we are extending is such that we start with a Morse function with an index 1 critical point below an index 2 critical point, and we know that the descending manifold for the index 2 critical point is disjoint from the ascending manifold for the index 1 point, and then we extend to a 1-parameter family in which the index 1 point rises above the index 2 point. Up to diffeomorphism this is unique because we do not have any handle slides between the two critical points to worry about. (By comparison, if we were extending in a direction in which both critical points had the same index, there would be many different extensions depending on whether we put handle slides in before the critical point crossing.)

In Figure 5(c), we have illustrated the extension in two steps, first from green to yellow and then to blue. In the first step (yellow), we need to choose a diffeomorphism ϕ from the fiber F_p at p to the fiber F_q at q , sending the attaching circle $C_p \subset F_p$ for the 2-handle cobordism H_p sitting over the arc P to the attaching circle $C_q \subset F_q$ for the 2-handle cobordism H_q sitting over the arc Q . The results of the first step only depend upon ϕ up to isotopy through diffeomorphisms sending C_p to C_q . After the first step, the monodromy around the enclosed disk (blue) should be trivial. Thus ϕ is determined up to isotopy (not necessarily rel. C_p and C_q) by the existing identification between F_p and F_q given by going around the lower side of the blue disk, the arc A . (This uses the fact that $\pi_1(\text{Diff}_0 \Sigma_g) = 0$ when $g > 1$, Earle and Eells [2].) Now, since all fibers have genus larger than 1, maps sending C_p to C_q which are isotopic

are actually isotopic rel. C_p and C_q , and thus the isotopy class of ϕ relative to the attaching curves is determined.

Given the uniqueness of the three extensions above, the uniqueness of extension across a cusp or crossing from the full boundary of a disk to the interior of the disk follows as in Figure 6, where the last extension across a disk containing no folds is again unique because $\pi_1(\text{Diff}_0 \Sigma_g) = 0$. It is important that we apply all the extensions from Figure 5 in the correct direction, ie in the direction of decreasing genus. \square

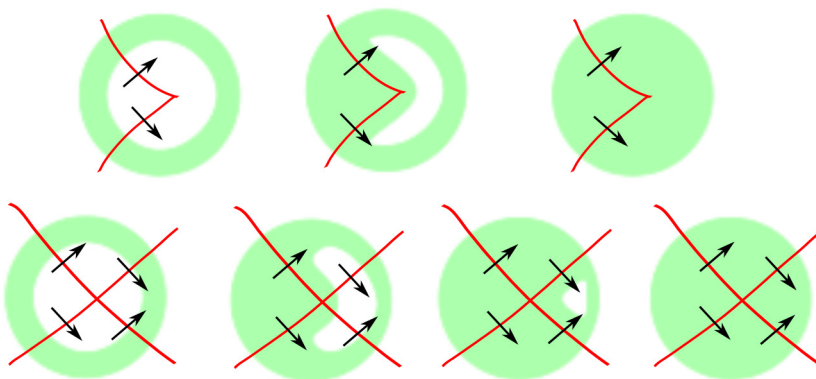


Figure 6: Full extensions across cusps and crossings

Proof of Theorem 5 The first two cases (chains of edges and crossings) both follow immediately from the uniqueness of the three extensions across disks shown in Figure 5. The case of a locally genus minimizing region is also immediate: It is clear from the first two cases that the gluing data across one edge of R determines the gluing data across all other edges. Since there are no outward co-oriented edges of R , changing this gluing data by an automorphism of F_R changes no other data in the diagram, and any two gluing data across one fixed edge of R are related by such an automorphism. \square

We next present two interesting counterexamples to our original false claim (Theorem 3 of [4]), that in the case of maps to S^2 the gluing data is unnecessary. The first is a local picture that can be embedded in a S^2 -valued Morse 2-function on some closed 4-manifold, and the second is a more global construction.

In Figure 7 we show the fold graph Γ in a disk, with the attaching circles indicated, and the gluing data specified around the boundary. To avoid excess subscripts, the attaching circles are labelled with the same label as the edges. The fiber is drawn as a

surface with boundary, the understanding being that this can be glued onto a surface of any genus and the gluing data extended by the identity map. Recall that the gluing data across an edge e is a basis $A_{(1,e)}, B_{(1,e)}, \dots, A_{((g_R-1),e)}, B_{((g_R-1),e)}$ disjoint from the attaching circle for that edge. Thus for the upper edges, going from genus 1 to genus 0, there is nothing to draw, and for the lower edges, we should draw one pair (A_e, B_e) for each edge e . The gluing data then means that, after surgering the genus 2 surface along the attaching circle for edge e , we glue the resulting genus 1 surface to the standard genus 1 surface sending (A_e, B_e) to the standard basis (see Figure 2).

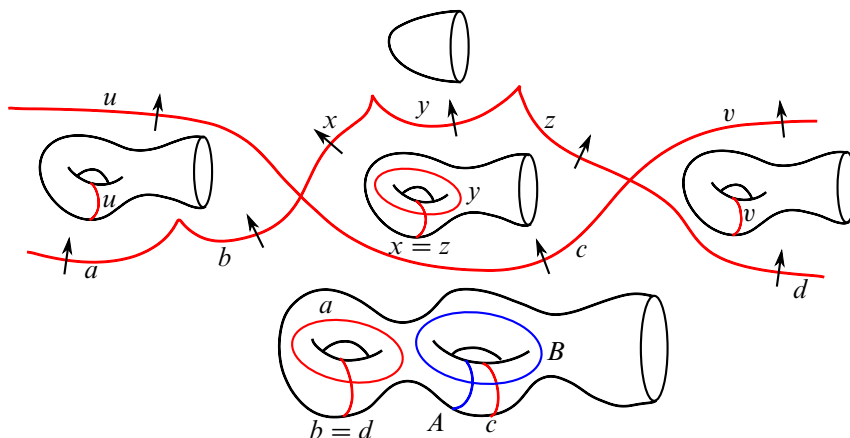


Figure 7: The local counterexample. Curves in red are attaching circles. The curves in blue labelled A and B are the gluing data for both edges a and d . In other words, $A_a = A_d = A$ and $B_a = B_d = B$. No gluing data is shown for other edges.

The ambiguity lies in the gluing data across edge c at the bottom, and in particular, the curve B_c in the basis (A_c, B_c) . The pair (A_c, B_c) should be a basis for what's left after one surgers along c , and thus should be disjoint from c . For example, (A_c, B_c) can be either of the two examples illustrated in Figure 8. To see that the results are different, pull back the attaching circle for edge y from the genus 1 surface to the genus 2 surface via the gluing map; in one case, y becomes a circle that intersects a once, transversely, while in the other case, y is disjoint from a . Thus, if one considers the S^1 -valued Morse function sitting over the green loop in Figure 8, either the critical points a and y cancel or they do not, and so the homology of the 3-manifold sitting over the green circle is different.

In Figure 9 we show a Morse 2–function over S^2 with annular regions between the folds, and we indicate the possibilities for monodromy around those annuli. We think of this monodromy as being a self-gluing of the fiber along the vertical dotted lines. Notice

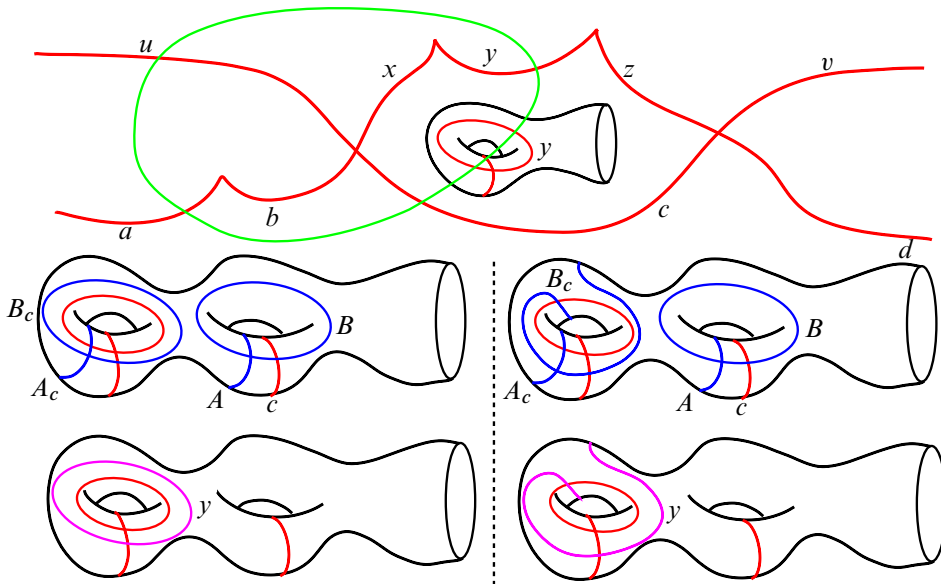


Figure 8: The local counterexample, part 2 Two choices for the gluing data (A_c, B_c) across edge c are indicated in blue, one to the left of the dotted line, the other to the right. The pullbacks by these two gluings of the attaching circle for edge y are indicated in pink on the lower surfaces. The curves A and B are the gluing data across edges a, b and d , ie $A = A_a = A_b = A_d$ and $B = B_a = B_b = B_d$. All unlabelled curves should be taken to be labelled as in Figure 7.

that we cannot choose any monodromy, because the monodromies should become trivial after surgering along the attaching circles for the edges, so that the disks can be filled in. We have indicated here a few of the many possible choices satisfying this condition. In Figure 10 we attempt to hide the monodromy by making all regions simply connected. The Reidemeister II type moves that we have done in the middle to cut the annuli into disks are legitimate because all the attaching circles involved are disjoint. Since the monodromy is now “happening” in the two small quadrilateral regions in the middle, it does not affect any of the attaching circles for the edges, and thus the attaching circles alone cannot possibly be sufficient to reconstruct the monodromy.

This last example leaves us with an interesting question: How can one formalize the notion of nontrivial monodromy when all regions and even the entire base are simply connected, without reference to a homotopic fibration with non simply connected regions?

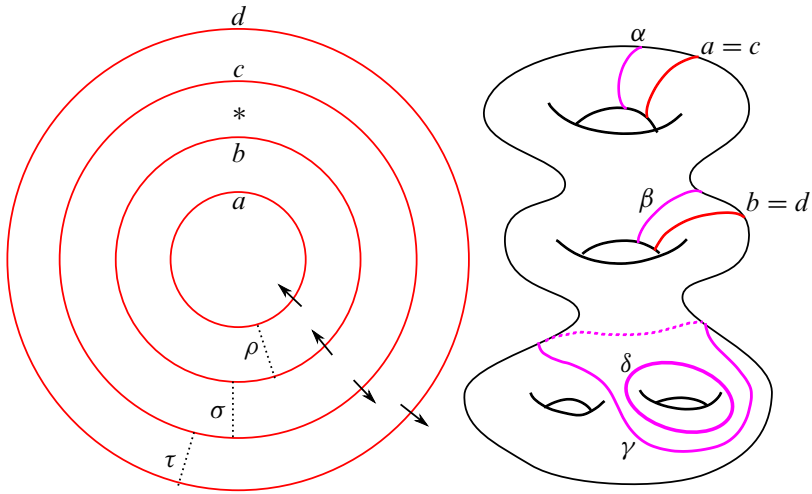


Figure 9: The global counterexample The genus 4 surface drawn at right is the fiber at the point labelled $*$, with all attaching circles drawn in red. The pink curves are Dehn twist curves representing the monodromy across the dotted lines: $\rho = \alpha^k \gamma^m \delta^{-m}$, $\tau = \beta^l \gamma^m \delta^{-m}$, $\sigma = \alpha^k \beta^l \gamma^m \delta^{-m}$, for some $k, l, m \in \mathbb{Z}$.

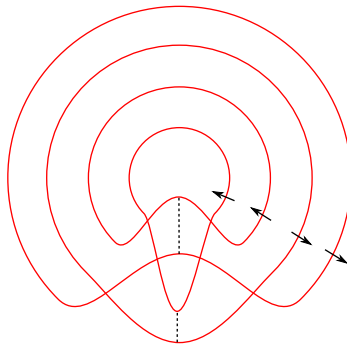


Figure 10: The global counterexample, part 2

Acknowledgements This work was partially supported by a grant from the Simons Foundation (#210381 to David Gay). The second author was partially supported by NSF grant DMS-0838703.

References

- [1] **J Cerf**, *La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie*, Inst. Hautes Études Sci. Publ. Math. (1970) 5–173 MR0292089

- [2] **C J Earle, J Eells**, *The diffeomorphism group of a compact Riemann surface*, Bull. Amer. Math. Soc. 73 (1967) 557–559 MR0212840
- [3] **D T Gay, R C Kirby**, *Indefinite Morse 2–functions; broken fibrations and generalizations* arXiv:1102.0750
- [4] **D T Gay, R C Kirby**, *Fiber-connected, indefinite Morse 2–functions on connected n –manifolds*, Proc. Natl. Acad. Sci. USA 108 (2011) 8122–8125 MR2806648
- [5] **F Laudenbach, V Poénaru**, *A note on 4–dimensional handlebodies*, Bull. Soc. Math. France 100 (1972) 337–344 MR0317343

DG: *Euclid Lab, 500 Willow St, Athens GA 30601, USA*

Department of Mathematics, University of Georgia, Athens GA 30602, USA

RK: *Department of Mathematics, University of California, Berkeley CA 94720-3840, USA*

d.gay@euclidlab.org, dgay@uga.edu, kirby@math.berkeley.edu

<http://euclidlab.org/david-gay/>

Received: 15 February 2012 Revised: 5 May 1012