

Trisecting 4-manifolds

DAVID T. GAY, ROBION KIRBY *

Euclid Lab, 500 Willow St, Athens, GA 30601

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

and

University of California, Berkeley, CA 94720

Email: d.gay@euclidlab.org and kirby@math.berkeley.edu

Abstract We show that, for any closed, oriented 4-manifold X , there are integers $0 \leq k \leq g$ such that X decomposes as a union of three copies of $\natural^k(S^1 \times B^3)$ glued together along genus g handlebodies. These genus g handlebodies are the halves of the standard $(g-k)$ -times stabilization of the standard genus k Heegaard splitting of $\sharp^k(S^1 \times S^2) = \partial(\natural^k(S^1 \times B^3))$. This is analogous to the fact that every 3-manifold splits into two handlebodies, and is a natural generalization of the fact that $\mathbb{CP}^2$ splits as three balls glued along solid tori. We prove our result using Morse 2-functions.

AMS Classification 57M50; 57R45, 57R65

Keywords Morse 2-function, 4-manifold, Heegaard splitting

Given an integer $k \geq 0$, let $Z_k = \natural^k(S^1 \times B^3)$ with $Y_k = \partial Z_k = \sharp^k(S^1 \times S^2)$. Given an integer $g \geq k$, let $Y_{k,g}^+ \cup Y_{k,g}^- = Y_k$ be the standard genus g Heegaard splitting of Y_k obtained by stabilizing the standard genus k Heegaard splitting $g - k$ times.

Theorem 1 *Given a closed, oriented 4-manifold X , there are integers $0 \leq k \leq g$ and a decomposition of X into three submanifolds $X = X_1 \cup X_2 \cup X_3$ satisfying the following properties:*

- (1) *For each $i = 1, 2, 3$, there is a diffeomorphism $\phi_i: X_i \rightarrow Z_k$.*
- (2) *For each $i = 1, 2, 3$, taking indices mod 3, $\phi_i(X_i \cap X_{i+1}) = Y_{k,g}^-$ and $\phi_i(X_i \cap X_{i-1}) = Y_{k,g}^+$.*

Note that the 4-manifold X is then determined up to diffeomorphism by the data of k, g and two self-diffeomorphisms $\beta_1, \beta_2: \natural^g(S^1 \times B^2) \rightarrow \natural^g(S^1 \times B^2)$ of the standard genus g handlebody. These determine the gluing maps from X_1 to X_2 and from X_2 to X_3 ; the gluing map from X_3 to X_2 is then given by $\beta_2^{-1} \circ$

*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.

β_1^{-1} . In fact, *any* such choice of (g, k, β_1, β_2) determines a closed 4-manifold. Thus we avoid the problems related to whether or not the construction “caps off” which arise when describing 4-manifolds via framed links or Lefschetz fibrations, for example.

For some simple examples, note that the theorem works for S^4 with $k = g = 0$ and for $\mathbb{CP}^2$ and $\overline{\mathbb{CP}^2}$ with $k = 0$ and $g = 1$. On the other hand, simple 4-manifolds do not necessarily have simple trisections. For example, $S^2 \times S^2$ has a natural decomposition into four B^4 ’s or two $(S^2 \times B^2)$ ’s, but its trisection does not seem to be so natural. The method of the construction in the proof below, applied to $S^2 \times S^2$ with the standard handle decomposition, produces a trisection with $k = 2$ and $g = 6$, and the gluing maps are not obvious.

We use Morse 2-functions in our proof, which is essentially an application of tools developed in [2]; with some effort this proof could be rewritten entirely in terms of ordinary Morse functions and handle decompositions, but the trisection is so natural from the point of Morse 2-functions that such a project might not be worth the effort. However, to give the basic idea for those most comfortable with the language of handle decompositions, our construction ends up putting the 0- and 1-handles of X into X_1 , the 3- and 4-handles into X_3 , and the 2-handles together with some “connective tissue” into X_2 .

A Morse 2-function is a smooth, stable map $G: X^n \rightarrow \Sigma^2$; in this paper we will always map to $\mathbb{R}^2$. (Stable implies generic when mapping to dimension two.) Just like Morse functions, Morse 2-functions can be characterized by local models, and we now give these local models only in the case of $n = 4$, i.e. we are considering an $\mathbb{R}^2$ -valued Morse 2-function G on a 4-manifold X :

- (1) Each regular value $q \in \mathbb{R}^2$ has a coordinate neighborhood over which G looks like $F^2 \times B^2 \rightarrow B^2$ for some closed fiber surface F .
- (2) The set of critical points of G is a smooth 1-dimensional submanifold $\text{Crit}_G \subset X$ such that $G: \text{Crit}_G \rightarrow \mathbb{R}^2$ is an immersion with isolated semicubical cusps and crossings. The non-cusp points of Crit_G are called *fold points*, and arcs of such points are called *folds*.
- (3) Each point $q \in G(\text{Crit}_G)$ which is not a cusp or crossing has a neighborhood $U = I \times I$ with coordinates (t, y) , with $G^{-1}(U)$ diffeomorphic to $I \times M^3$ for a 3-dimensional cobordism M , so that $G(t, p) = (t, g(p))$, where $g: M \rightarrow I$ is a Morse function on M with one critical point. The index of this critical point is then called the index of the fold, although this is only well-defined up to $i \mapsto 3 - i$. When the image of the fold is co-oriented, the index is well-defined by insisting that the y -coordinate on $I \times I$ increases in the direction of this co-orientation.

- (4) Each cusp point $q \in G(\text{Crit}_G)$ has a neighborhood $U = I \times I$ with coordinates (t, y) , with $G^{-1}(u) = I \times M^3$, so that $G(t, p) = (t, g_t(p))$, where g_t is a 1-parameter family of Morse functions on M with no critical points for $t = 0$ and a birth of a cancelling pair of critical points at $t = 1/2$. In our examples, these two critical points will always be of index 1 and 2.
- (5) Each crossing point $q \in G(\text{Crit}_G)$ has a neighborhood $U = I \times I$ with coordinates (t, y) , with $G^{-1}(u) = I \times M^3$, so that $G(t, p) = (t, g_t(p))$, where g_t is a 1-parameter family of Morse functions on M with two critical points for all t , such that the critical values cross at $t = 1/2$. In our examples, these two critical points will never be of index 0 or 3.

The basic example of a Morse 2-function is $(t, p) \mapsto (t, g_t(p))$ for an arbitrary generic homotopy g_t between two given Morse functions $g_0, g_1: M^3 \rightarrow [0, 1]$, and the message of the above local models is that Morse 2-functions look locally like homotopies between Morse functions, but globally we may not have a preferred “time” direction. When G is of the form $(t, p) \mapsto (t, g_t(p))$, we call $G(\text{Crit}_G)$ a *Cerf graphic* [1]. Conversely, given a Morse 2-function $G: X^4 \rightarrow \mathbb{R}^2$ and a rectangle $I \times I \subset \mathbb{R}^2$ in which $G(\text{Crit}_G)$ has no vertical tangencies, we can find coordinates in which G is of this form $(t, p) \mapsto (t, g_t(p))$, and so again we will say that $G(\text{Crit}_G)$ is a Cerf graphic in this rectangle.

Proof Throughout we will use coordinates (t, z) on $\mathbb{R}^2$, with t horizontal and z vertical. Here is an outline of the proof:

- (1) First we will show that there is a Morse 2-function $G_1: X \rightarrow \mathbb{R}^2$ such that the image of the fold locus is as in Figure 1. In this and the following figures, three dots between two curves indicate that there are some number of parallel copies of the two curves in between. Fold indices are indicated with labelled transverse arrows. Boxes with folds coming in from the left and out at the right represent arbitrary Cerf graphics, with the left-right axis being time. Note that a Cerf graphic may contain left-cusps, right-cusps and crossings, but may not contain any vertical tangencies of the image of the fold locus.
- (2) In Figure 1, the vertical tangencies of the folds are highlighted and projected to the horizontal t -axis; these become critical points of the projection $t \circ G_1: X \rightarrow \mathbb{R}$. These critical values in $\mathbb{R}$ are also indicated at the bottom of the diagram along the t -axis, with their indices.
- (3) After constructing G_1 , we will show to homotope G_1 to G_2 such that the image of the fold locus for G_2 is as in Figure 2. Here the two Cerf

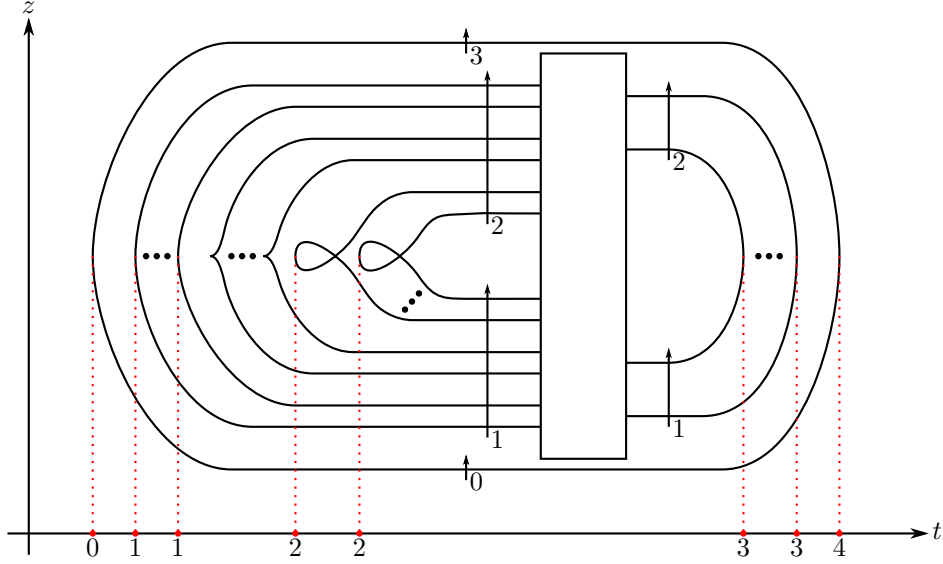


Figure 1: The image of the fold locus for G_1 .

graphics have no cusps. We have achieved two goals here: (1) Splitting the Cerf graphic into two, each involving only critical points of the same index and no cusps. (2) Replacing each kink that corresponds to an index 2 critical point of $t \circ G_1$ with a pair of cusps.

- (4) Figure 3 is simply a redrawing of Figure 2 that highlights a natural tri-section of $\mathbb{R}^2$ into three sectors $\mathbb{R}_1^2$, $\mathbb{R}_2^2$ and $\mathbb{R}_3^2$. Note that the critical locus over each sector consists of g components, where g is the genus of the central fiber. Also, each such component has at most one cusp. We no longer indicate the indices of the folds; the outermost fold is index 0 pointing inwards, and all other folds are index 1 pointing in.
- (5) The form of the folds in Figure 3 is a special case of the form shown in Figure 4, where now we are not paying attention to which folds in a given sector, with or without cusps, connect to which folds in the next sector, with or without cusps, and we allow for arbitrary Cerf graphics (without cusps) between the sectors.
- (6) Now we have G_2 such that the image of the fold locus is as in Figure 4. At this point we could take $X_i = G_2^{-1}(\mathbb{R}_i^2)$ and we would have each X_i diffeomorphic to $\mathbb{H}^{k_i} S^1 \times B^3$ for different k_i 's. There is one last step to arrange that the k_i 's are equal: In fact, k_i is equal to the number of folds in sector X_i *without cusps*. We will show how to add a fold without a cusp to any one sector while adding a fold with a cusp to each of the

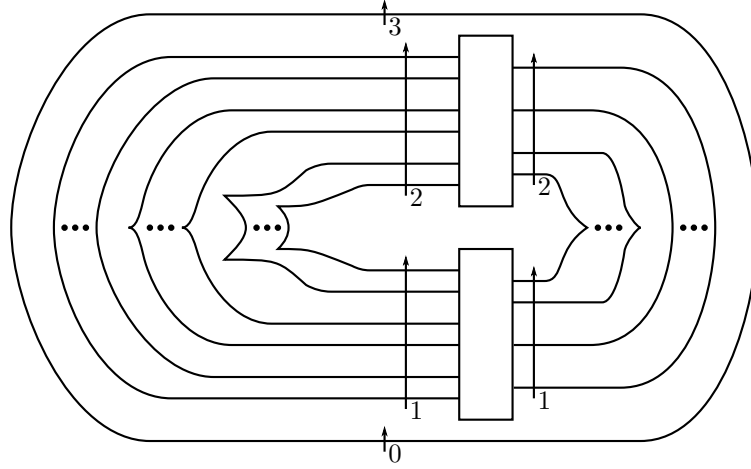


Figure 2: The image of the fold locus for G_2 .

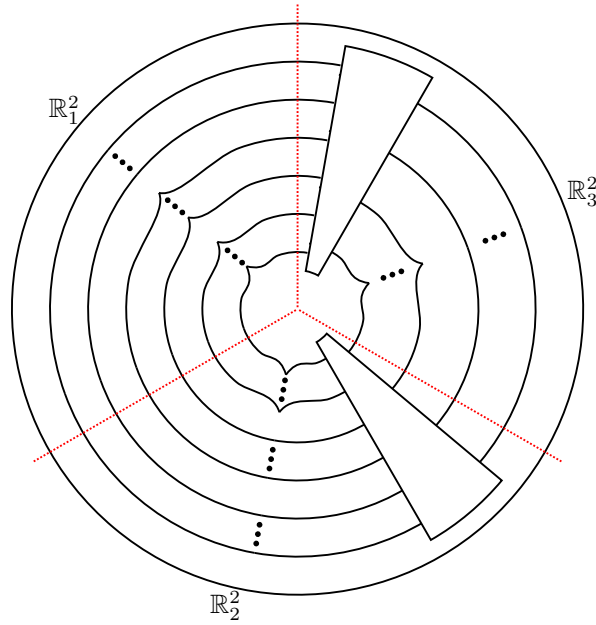


Figure 3: A more symmetric drawing of the image of the fold locus for G_2 . We no longer indicate the indices of the folds; the outermost fold is index 0 going inwards, the others are index 1 going inwards.

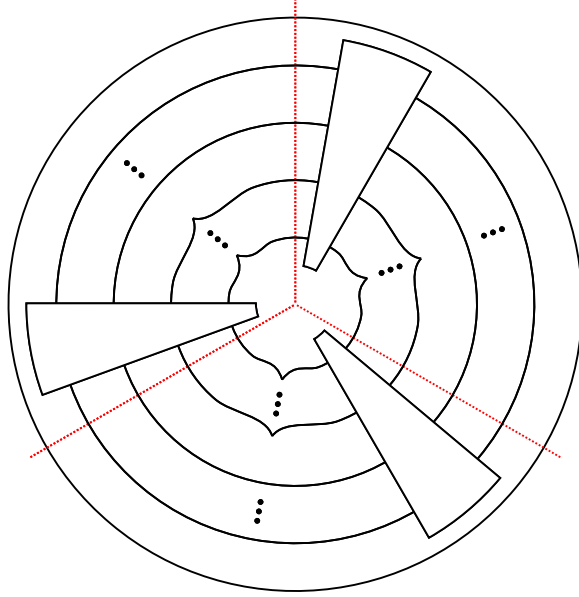


Figure 4: A slightly more general form for the image of the fold locus, which fits G_2 .

other two sectors. This allows us to construct a homotopy from G_2 to G_3 , such that G_3 has the image of its fold locus of the same form as G_2 (i.e. as in Figure 4), with the same number of folds without cusps in each sector, i.e. $k_1 = k_2 = k_3 = k$.

- (7) Finally we will justify the claim that each $X_i = G_3^{-1}(\mathbb{R}_i^2)$ is diffeomorphic to $\natural^k S^1 \times B^3$ with overlap maps as advertised.

We now fill in the details.

Begin with a handle decomposition of X with one 0-handle, i_1 1-handles, i_2 2-handles, i_3 3-handles and one 4-handle. The union of the 0- and 1-handles, X_1 is diffeomorphic to $I \times \natural^{i_1}(S^1 \times B^2)$. Map this to $I \times I$ by $(t, p) \mapsto (t, g(p))$ where $g: \natural^{i_1}(S^1 \times B^2) \rightarrow I$ is the standard Morse function with one index 0 critical point and i_1 index 1 critical points. Post-compose this map with a diffeomorphism from $I \times I$ to a half-disk and we have constructed G_1 on the union of the 0- and 1-handles so that the image of the fold locus is as in the right half of Figure 5.

Now note that $\partial X_1 = \natural^{i_1}(S^1 \times S^2)$ sits over the right edge of the half disk in Figure 5 and that the vertical Morse function on ∂X_1 , i.e. $z \circ G_1|_{\partial X_1}$ is the

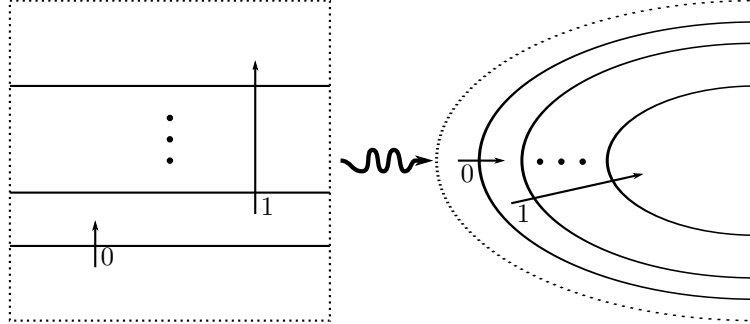


Figure 5: The first Morse 2-function, G_1 , on the 0- and 1-handles of X .

standard Morse function with i_1 index 1 critical points and i_1 index 2 critical points, inducing the standard genus i_1 splitting of ∂X_1 , with Heegaard surface Σ .

Consider the framed attaching link $L \subset \partial X_1$ for the 2-handles of X . Generically L will be disjoint in ∂X_1 from the ascending 1-manifolds of the index 2 critical points of $z \circ G_1|_{\partial X_1}$ as well as the descending 1-manifolds of the index 1 critical points. Thus L can be projected onto the Heegaard surface Σ along gradient flow lines to give an immersed curve $\overline{L}$ in Σ with at worst double points. By adding kinks if necessary, we can assume that the handle framing of L agrees with the “blackboard framing” coming from $\overline{L} \subset \Sigma$. Then by stabilizing this Heegaard splitting once for each crossing of $\overline{L}$, we can resolve these crossings and get L to lie in the Heegaard surface with framing coming from the surface. This process translates into an extension of the thus-far constructed G_1 from X_1 to $X_1 \cup ([0, 1] \times \partial X_1)$ with fold locus as in Figure 6, with one cusp for each stabilization. In other words, the sequence of stabilizations translates into a homotopy g_t from g_0 , the standard Morse function on $\sharp^{i_1}(S^1 \times S^2)$, to g_1 , the stabilized Morse function. This homotopy then becomes a Morse 2-function on the collar $[0, 1] \times \partial X_1$.

Now let Σ refer to the stabilized Heegaard surface, in which L lies. Attaching a 4-dimensional 2-handle to X_1 along a component K of L is the same as attaching I times a 3-dimensional 2-handle to X_1 along $I \times K \subset I \times \Sigma \subset \partial X_1$. In Figure 7 we show the resulting Morse 2-function at the left, where the handle sits over a vertical rectangle. Next we bend this rectangle to make the image again a half-disk. Finally, noting that the vertical Morse function at the right edge now has an index 2 critical value below an index 1 critical value, we switch these values to get the Morse 2-function at the right side of Figure 7.

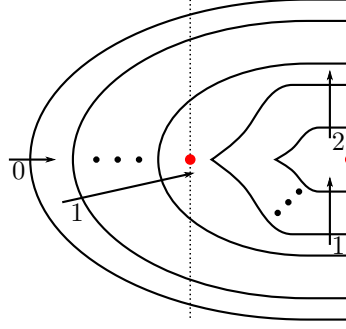


Figure 6: G_1 extended to a collar on ∂X_1 . In the two vertical slices shown, both diffeomorphic to $\sharp^n(S^1 \times S^2)$, the Heegaard surface sits over the highlighted red points. The framed attaching link L for the 2–handles of X lies in the Heegaard surface for the right–most Morse function, i.e. over the right–most red point, with framing coming from the surface.

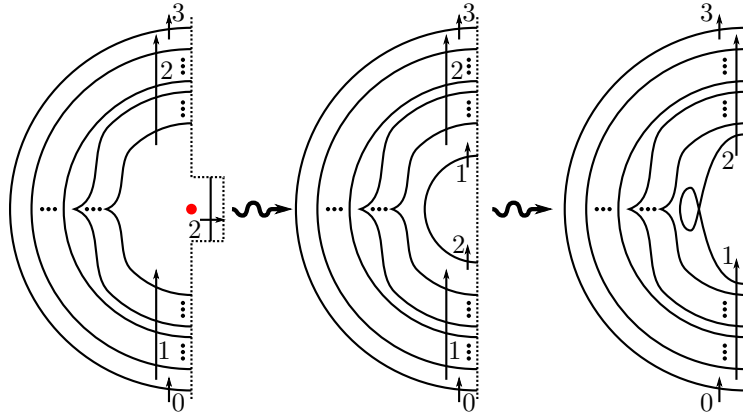


Figure 7: G_1 after attaching a 4–dimensional 2–handle.

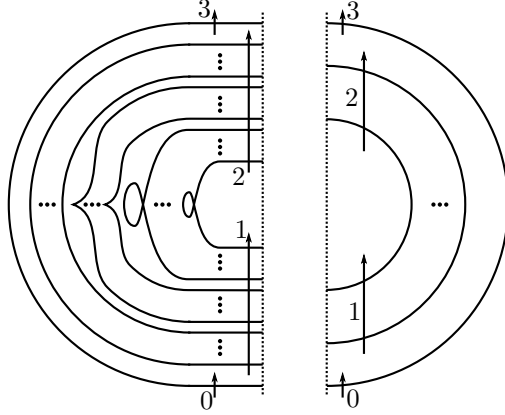


Figure 8: Two halves of G_1 : the 0-, 1- and 2-handles on the left and the 3- and 4-handles on the right. Connecting them with a Cerf graphic gives Figure 1.

Note that everything in the preceding paragraph happened in a neighborhood of K , so that the rest of L still lies in the middle Heegaard surface for the Morse function at the right edge of the final diagram in Figure 7. Thus we can attach each 4-dimensional 2-handle this way to get the Morse 2-function at the left side of Figure 8. Each 2-handle of X corresponds to a kink in the image of the folds, i.e. a smoothly immersed arc with a single transverse double point. Repeating our construction for X_1 with the union of the 3- and 4-handles, we construct the Morse 2-function at the right side of Figure 8. The two halves give vertical Morse functions on the boundary of the union of the 3- and 4-handles, which are related by some Cerf graphic. Putting this Cerf graphic in between the two parts of Figure 8 gives us G_1 as in Figure 1.

To get to Figure 2, first we take the Cerf graphic section of Figure 1 and pull the births (left-cusps) to the left of the Cerf graphic and the deaths (right-cusps) to the right, and then pull all index 1 critical points below all index 2 critical points. Then the left-cusps can be pulled further left, past the kinks which correspond to 4-dimensional 2-handles, because the 4-dimensional 2-handle attachments are independent of the 3-dimensional stabilizations corresponding to the cusps. This is shown in Figure 9.

Next we homotope the kinks into pairs of cusps as in Figure 10. The first step of Figure 10 introduces a swallowtail at the vertical tangency of the kink; this move has been discussed extensively elsewhere [3] and is a standard singularity that occurs in a homotopy between homotopies between Morse functions. The second step moves an arc of index 1 critical points in a homotopy (Cerf graphic)

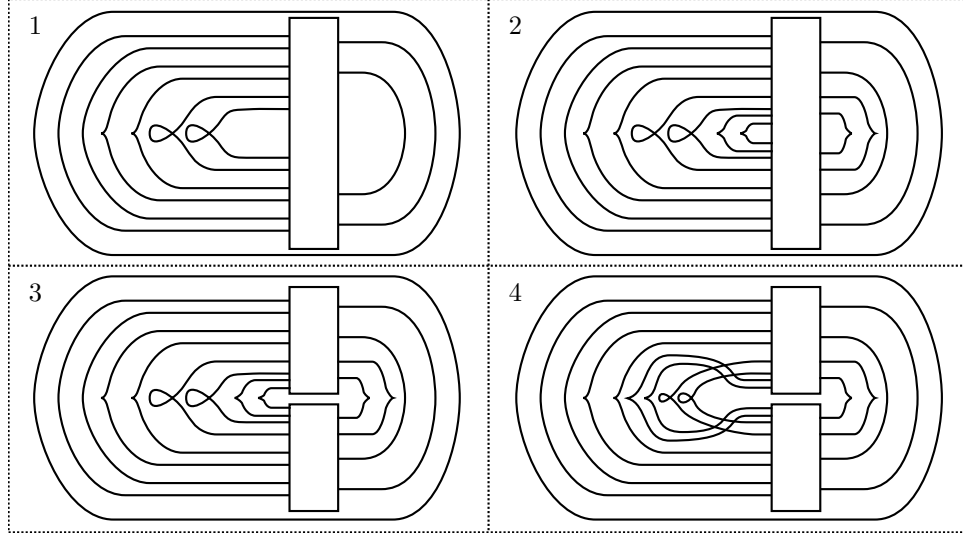


Figure 9: Pulling cusps out of the Cerf graphic. Here we suppress the “three dots” notation as well as the indices of the folds, as these are understood from earlier figures.

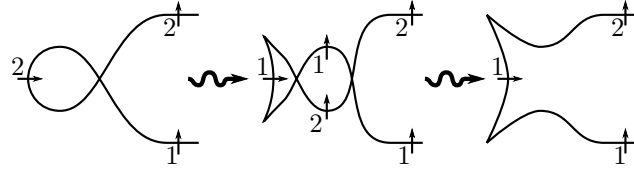


Figure 10: Turning kinks into pairs of cusps.

below an arc of index 2 critical points. This is also standard and is possible because the descending manifold for the index 1 point remains disjoint from the ascending manifold for the index 2 point throughout the homotopy. (Equivalently, in homotopies between Morse functions we never expect 1–handles to slide over 2–handles.)

Finally, Figure 11 shows how to add folds and cusps to a Morse 2–function as in Figure 4 so as to increase the number of folds without cusps in one of the three sectors. This uses another standard homotopy of Cerf graphics that introduces a birth of a cancelling pair of critical points immediately followed by the death of the same pair; depending on how we orient this Cerf graphic with respect to the trisection of $\mathbb{R}^2$, we either add the fold without cusps to $\mathbb{R}_1^2$, $\mathbb{R}_2^2$, or $\mathbb{R}_3^2$. Note that if we do this operation three times, once for each sector, we obtain a

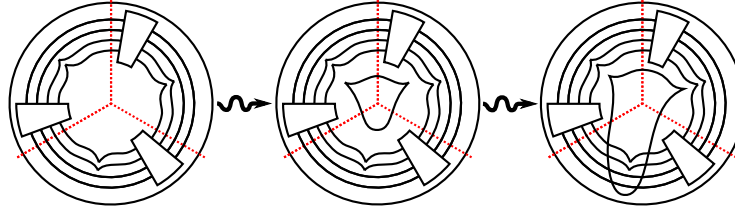


Figure 11: Stabilization; again we suppress the “three dots” notation and the fold indices.

natural stabilization operation for the splittings produced by this theorem.

Now we need to show that, having put our Morse 2-function finally into the form of Figure 4, with k folds in each sector without cusps and $g - k$ folds with cusps, then for each i , $G^{-1}(\mathbb{R}_i^2) = X_i \cong \natural^k(S^1 \times B^3)$. However, we have already seen this: Each sector, ignoring the Cerf graphic block, looks just like Figure 6, which we already know is $\natural^k(S^1 \times B^3)$ with a $(g - k)$ -times stabilized standard Heegaard splitting on the boundary. The Cerf graphic block connecting one sector to another is a product which does not interfere with the Heegaard splitting.

□

References

- [1] Cerf, Jean La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie, *Inst. Hautes Études Sci. Publ. Math.* 39:5–173, 1970.
- [2] Gay, David T. and Kirby, Robion C. Indefinite Morse 2-functions: Broken fibrations and generalizations. preprint arXiv:1102.0750 (2011).
- [3] Yankı Lekili. Wrinkled fibrations on near-symplectic manifolds. *Geom. Topol.*, 13(1):277–318, 2009. Appendix B by R. İnanç Baykur.