Morse theory for manifolds with boundary

MACIEJ BORODZIK
ANDRÁS NÉMETHI
ANDREW RANICKI

We develop Morse theory for manifolds with boundary. Beside standard and expected facts like the handle cancellation theorem and the Morse lemma for manifolds with boundary, we prove that under suitable connectedness assumptions a critical point in the interior of a Morse function can be moved to the boundary, where it splits into a pair of boundary critical points. As an application, we prove that every cobordism of connected manifolds with boundary splits as a union of left product cobordisms and right product cobordisms.

1 Introduction

For some time now, Morse theory has been a very fruitful tool in the topology of manifolds. One of its milestones was the h–cobordism theorem of Smale [17], and its Morse-theoretic exposition by Milnor [11; 12]. Recently, Morse theory has become even more popular, for two reasons: its connections with Floer homology (see eg Salamon [15], Witten [18], Nicolaescu [13] and Kronheimer and Mrowka [9]) and the stratified Morse theory developed by Goresky and MacPherson [6]. In the last 20 years Morse theory has also had an enormous impact on the singularity theory of complex algebraic and analytic varieties.

Morse theory for manifolds with boundary was studied in the seventies by Braess [5], Jankowski and Rubinsztein [8] and Hajduk [7]. Recently it was developed by Kronheimer and Mrowka [9]. Since then the theory has experienced very fast development, as witnessed by the papers of Bloom [2] and Laudenbach [10]. Our paper is another contribution.

In this paper we prove some new results in Morse theory for manifolds with boundary. Besides some standard and expected results, like the boundary handle cancellation theorem (Theorem 5-1) and the topological description of passing critical points on the boundary (using the notions of right and left half-handles introduced in Section 2)

we describe another phenomenon; see Theorem 3-1. An interior critical point can be moved to the boundary and there split into two boundary critical points. A related result was stated in Hajduk [7, Theorem 5]; we provide a rigorous proof under much weaker assumptions.

In particular, if we have a cobordism of manifolds with boundary, then under a natural topological assumption we can find a Morse function which has only boundary critical points. We use this result to prove a structure theorem for connected cobordisms of connected manifolds with connected nonempty boundary: such a cobordism splits as a union of left and right product cobordisms. This is a topological counterpart to the algebraic splitting of cobordisms obtained by the authors in [3, Main theorem 1]: an algebraic splitting of the chain complex cobordism of a geometric cobordism can be realized topologically by a geometric splitting. This algebraic splitting is used to study the algebraic properties of the Seifert matrices of isotopic nonspherical $(2n-1)-$dimensional links in S^{2n+1}. This will provide the algebraic background to our proof that the semicontinuity of mod 2 spectra of hypersurface singularities is a purely topological phenomenon (see the authors' [4], especially the paragraph before proof of Theorem 2.1.8).

The structure of the paper is as follows. After preliminaries in Section 1.1 we study in Section 2 the changes in the topology of the level sets when crossing a boundary critical point. Theorem 2-27 is the main result of this section: passing a boundary stable (unstable) critical point produces a left (right) half-handle attachment. In Section 3 we prove Theorem 3-1, which moves interior critical points to the boundary. This is the first main result of this article. Then we pass to some more standard results, namely rearrangements of critical points in Section 4. We finish the section with our most important — up to now — application, Theorem 4-18, about the splitting of a cobordism into left product and right product cobordisms. Finally, in Section 5 we discuss the possibility of canceling a pair of critical points. We include this part for completeness, the main result of this section was proved in Hajduk [7, Theorem 1].

1.1 Notes on gradient vector fields

To fix notation, let us recall what a cobordism of manifolds with boundary is.

Definition 1-1 Let Σ_0 and Σ_1 be compact oriented, n–dimensional manifolds with nonempty boundaries M_0 and M_1. We shall say that (Ω, Y) is a *cobordism* between (Σ_0, M_0) and (Σ_1, M_1) if Ω is a compact oriented $(n+1)$–dimensional manifold with boundary $\partial\Omega = Y \cup \Sigma_0 \cup \Sigma_1$, where Y is nonempty, $\Sigma_0 \cap \Sigma_1 = \varnothing$, and $Y \cap \Sigma_0 = M_0$, $Y \cap \Sigma_1 = M_1$.

Remark 1-2 Strictly speaking, Ω is a manifold with corners, so around a point $x \in M_0 \cup M_1$ it is locally modeled on $\mathbb{R}^{n-1} \times \mathbb{R}^2_{\geq 0}$. Accordingly, sometimes we write that Σ_0, Σ_1 and Y, as manifolds with boundary, have tubular neighborhoods in Ω of the form $\Sigma_0 \times [0, 1)$, $\Sigma_1 \times [0, 1)$, or $Y \times [0, 1)$, respectively. Nevertheless, in most cases it is safe (and more convenient) to assume that Ω is a manifold with boundary, ie that the corners are smoothed along M_0 and M_1. Whenever possible we make this simplification in order to avoid unnecessary technicalities.

Example 1-3 Given a manifold with boundary (Σ, M), we call $(\Sigma, M) \times [0, 1]$ a trivial cobordism, with $\Omega = \Sigma \times [0, 1]$, $Y = M \times [0, 1]$, $\Sigma_i = \Sigma \times \{i\}$, $M_i = M \times \{i\}$ for $i = 0, 1$.

We recall the notion of a Morse function (in [7] they are called m–functions). For this it is convenient to fix a Riemannian metric g on Ω.

Definition 1-4 Let $F\colon \Omega \to [0, 1]$ be a smooth function. A critical point z of F is called *Morse* if the Hessian of F at z is nondegenerate. The function $F\colon \Omega \to [0, 1]$ is called a *Morse function* on the cobordism (Ω, Y) if $F(\Sigma_0) = 0$, $F(\Sigma_1) = 1$, F has only Morse critical points, the critical points are not situated on $\Sigma_0 \cup \Sigma_1$, and ∇F is everywhere tangent to Y.

There are two ways of doing Morse theory on manifolds. One can either consider the gradient flow of ∇F associated with F and the Riemannian metric (in the Floer theory, one often uses $-\nabla F$), or the so-called gradient-like vector field.

Definition 1-5 (See Milnor [12, Definition 3.1]) Let F be a Morse function on a cobordism (Ω, Y). Let ξ be a vector field on Ω. We shall say that ξ is *gradient-like with respect to* F if the following conditions are satisfied:

(a) $\xi \cdot F > 0$ away from the set of critical points of F.

(b)   If p is a critical point of F of index k, then there exist local coordinates $x_1, \ldots, x_{n+1}$ in a neighborhood of p, such that

$$F(x_1, \ldots, x_{n+1}) = F(p) - (x_1^2 + \cdots + x_k^2) + (x_{k+1}^2 + \cdots + x_{n+1}^2)$$

and

$$\xi = (-x_1, \ldots, -x_k, x_{k+1}, \ldots, x_{n+1}) \quad \text{in } U.$$

(b$'$)   Also, if p is a boundary critical point, then the above coordinate system can be chosen so that $Y = \{x_j = 0\}$ and $U = \{x_j \geqslant 0\}$ for some $j \in \{1, \ldots, n+1\}$.

(c) ξ is everywhere tangent to Y.

The conditions (a) and (b) are the same as in the classical case. Condition (b$'$) is an analogue of condition (b) in the boundary case; cf Lemma 2-6.

Smale [16] noticed that for any gradient-like vector field ξ for a function F there exists a Riemannian metric such that $\xi = \nabla F$ in that metric. The situation is identical in the boundary case. This is stated explicitly in the following lemma, whose proof is straightforward and will be omitted.

Lemma 1-6   *Let U be a paracompact k–dimensional manifold and $F \colon U \to \mathbb{R}$ a Morse function without critical points. Assume that ξ is a gradient–like vector field on U. Then there exists a metric g on U such that $\xi = \nabla F$ in that metric. A similar statement holds if U has boundary and ξ is everywhere tangent to the boundary.*

Hence the two approaches — by gradients and gradient-like vector fields — are equivalent. We shall need both of them. In Section 3 we use gradients of functions and a specific choice of a metric, which make the arguments slightly simpler. In Section 5 we follow Milnor [12] very closely; as he uses gradient-like vector fields, we use them as well.

The next result shows that the condition from Definition 1-4 that ∇F is everywhere tangent to Y can be relaxed. We shall use this result in Proposition 4-1.

Lemma 1-7   *Let Ω be a compact Riemannian manifold of dimension $n + 1$ and let $Y \subset \partial \Omega$ be compact as well. Let g denote the metric. Suppose that there exists a function $F \colon \Omega \to \mathbb{R}$, and a relative open subset $U \subset Y$ such that ∇F is tangent to Y at each point $y \in U$. Suppose furthermore that for any $y \in Y \setminus U$ we have*

(1-8) $$T_y Y \not\subset \ker dF.$$

Then, for any open neighborhood $W \subset \Omega$ of $Y \setminus U$, there exist a metric h on Ω, agreeing with g away from W, such that $\nabla_h F$ (the gradient in the new metric) is everywhere tangent to Y.

Proof Let us fix a point $y \in Y \setminus U$ and consider a small open neighborhood V_y of y in W, in which we choose local coordinates $x_1, \ldots, x_{n+1}$ such that $Y \cap V_y = \{x_{n+1} = 0\}$ and $V_y \subset \{x_{n+1} \geq 0\}$. In these coordinates we have $dF = \sum_{i=1}^{n+1} f_i(x) dx_i$ for some smooth functions $f_1, \ldots, f_{n+1}$. By (1-8), for each $x \in V_y$, there exists $i \leq n$ such that $f_i(x) \neq 0$. Shrinking V_y if needed, we may assume that the index i is the same for each $x \in V_y$. In the following we suppose that $i = 1$, that is for every $x \in V_y$ we have $\pm f_1(x) > 0$ and the sign $\pm$ is the same for every x. Let us choose a symmetric positive-definite matrix $A_y = \{a_{ij}(x)\}_{i,j=1}^{n+1}$ so that $a_{11} = \pm f_1(x)$ and for $i > 1$, $a_{1i} = a_{i1} = f_i(x)$. A_y defines a metric h_y on V_y such that $\nabla_{h_y} F = (\pm 1, 0, \ldots, 0) \subset TY$ in that metric.

Now let us choose an open subset V of $\Omega \setminus (Y \setminus U)$ such that $V \cup \bigcup_{y \in Y \setminus U} V_y$ is a covering of Ω. Let $\{\phi_V\} \cup \{\phi_y\}_{y \in Y \setminus U}$ be a partition of unity subordinate to this covering. Define

$$h = \phi_V \cdot g + \sum_{y \in Y \setminus U} \phi_y h_y.$$

Then h is a metric, which agrees with g away from W. Moreover, as for each metric h_y, and $x \in V_y \cap Y$ we have $\nabla_{h_y} F(x) \in T_x Y$ by construction, the same holds for a convex linear combination of metrics. $\square$

2 Boundary stable and unstable critical points

Most of the results of this section appeared previously in [5; 8; 7]. We provide them for completeness of exposition and for the convenience of the reader. In what follows, we use notation and terminology of [9].

2.1 Morse function for manifolds with boundary

The first question concerns the existence of Morse functions. While the condition that the function has only critical points of Morse type is open and dense, it requires a little argument to show that there are many functions generic in the interior such that their gradient, when restricted to Y, is tangent to Y.

Lemma 2-1 *Morse functions exist. In fact, for any Morse function $f: Y \to [0,1]$ with $f(M_0) = 0$, $f(M_1) = 1$ there exists a Morse function $F: \Omega \to [0,1]$ whose restriction to Y is f.*

Proof Let $f\colon Y \to [0,1]$ be a Morse function on the boundary, such that $f(M_0) = 0$ and $f(M_1) = 1$. We want to extend f to a Morse function on Ω.

First, let us choose a small tubular neighborhood U of Y and a diffeomorphism $U \cong Y \times [0, \varepsilon)$ for some $\varepsilon > 0$. Let $\widetilde{F}\colon U \to [0,1]$ be given by the formula

$$(2\text{-}2) \qquad U \cong Y \times [0, \varepsilon) \ni (x,t) \to \widetilde{F}(x,t) = f(x) - f(x)(1 - f(x))t^2.$$

The factor $f(x)(1 - f(x))$ ensures that $\widetilde{F}$ attains values in the interval $[0,1]$ and $\widetilde{F}^{-1}(i) \subset \Sigma_i$ for $i \in \{0,1\}$. It is obvious that there exists a smooth function $F\colon \Omega \to [0,1]$, which agrees on $Y \times [0, \varepsilon/2)$ with $\widetilde{F}$, and it satisfies the Morse condition on the whole of Ω. The gradient ∇F is everywhere tangent to Y. $\qquad \square$

Remark 2-3 The above construction yields a function with the property that all its boundary critical points are boundary stable (see Definition 2-4 below). This is due to the choice of sign -1 in front of $f(x)(1 - f(x))t^2$ in (2-2). If we change the sign to $+1$, we obtain a function with all boundary critical points boundary unstable.

We fix a Morse function $F\colon \Omega \to [0,1]$ and we start to analyze its critical points. Let z be such a point. If $z \in \Omega \setminus Y$, we shall call it an *interior* critical point. If $z \in Y$, it will be called a *boundary* critical point. There are two types of boundary critical points.

Definition 2-4 Let z be a boundary critical point. We shall call it *boundary stable*, if the tangent space to the unstable manifold of z lies entirely in $T_z Y$, otherwise it is called *boundary unstable*.
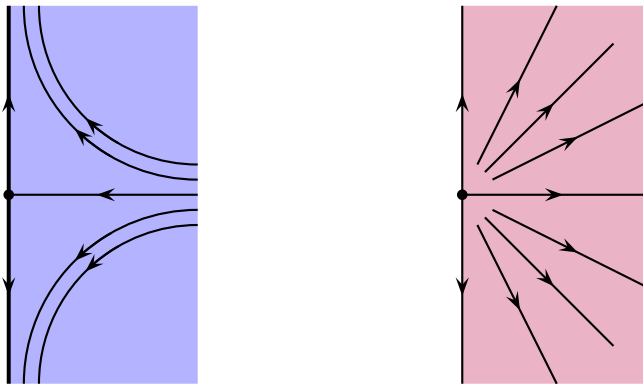

Figure 1: Boundary stable (on the left) and unstable critical points.

The index of the boundary critical point z is defined as the dimension of the stable manifold W_z^s. If z is boundary unstable, this is the same as the index of z regarded

as a critical point of the restriction f of F on Y. If z is boundary stable, we have $\operatorname{ind}_F z = \operatorname{ind}_f z + 1$. In particular, there are no boundary stable critical point with index 0, nor boundary unstable critical points of index $n + 1$.

Remark 2-5 We point out that we use the flow of ∇F and not of $-\nabla F$ as Kronheimer and Mrowka [9] do, hence our definitions and formulae are slightly different from theirs.

We finish this subsection with three standard results.

Lemma 2-6 (Boundary Morse lemma) *Assume that F has a critical point $z \in Y$ such that the Hessian $D^2 F(z)$ at z is nondegenerate, and ∇F is everywhere tangent to Y. Then there are local coordinates $(x_1, \ldots, x_{n+1})$ in an open neighborhood $U \ni z$ such that $U = \{x_1^2 + \cdots + x_{n+1}^2 \leqslant \varepsilon^2\} \cap \{x_1 \geqslant 0\}$ and $U \cap Y = \{x_1 = 0\}$ for some $\varepsilon > 0$, and F in these coordinates has the form $\pm x_1^2 \pm x_2^2 \pm \cdots \pm x_{n+1}^2 + F(z)$.*

Proof We choose a coordinate system $y_1, \ldots, y_{n+1}$ in a neighborhood $U \subset \Omega$ of z such that $z = (0, \ldots, 0)$, $Y = \{y_1 = 0\}$, $U = \{y_1 \geqslant 0\}$, and the vector field $\frac{\partial}{\partial y_1}$ is orthogonal to Y. We may and will assume $F(z) = 0$. The tangency of ∇F to Y implies that at each point of Y,

$$(2\text{-}7) \qquad \frac{\partial F}{\partial y_1}(0, y_2, \ldots, y_{n+1}) = 0.$$

The Hadamard lemma applied to F gives smooth functions $K_1, \ldots, K_{n+1}$ such that

$$(2\text{-}8) \qquad F = y_1 K_1(y_1, \ldots, y_{n+1}) + \sum_{j=2}^{n+1} y_j K_j(y_1, y_2, \ldots, y_{n+1}).$$

We can assume that for $j > 1$, K_j does not depend on y_1. Indeed, if it does, we write (again using the Hadamard lemma)

$$K_j(y_1, \ldots, y_{n+1}) = K_j(0, y_2, \ldots, y_{n+1}) + y_1 L_{1j}(y_1, \ldots, y_{n+1})$$

for smooth functions $L_{12}, \ldots, L_{1,n+1}$, and then replace K_j by $K_j(0, y_2, \ldots, y_{n+1})$ and K_1 by $K_1 + \sum y_j L_{1j}$. Condition (2-7) implies now that $K_1(0, y_2, \ldots, y_{n+1}) = 0$, hence

$$K_1(y_1, \ldots, y_{n+1}) = y_1 H_{11}(y_1, \ldots, y_{n+1})$$

for some function H_{11}. By the Hadamard lemma applied to $K_2, \ldots, K_{n+1}$ we infer that there exist functions H_{jk} for $j, k = 2, \ldots, n + 1$ such that

$$(2\text{-}9) \qquad F = y_1^2 H_{11}(y_1, \ldots, y_{n+1}) + \sum_{j,k=2}^{n} y_j y_k H_{jk}(y_2, \ldots, y_{n+1}).$$

Notice that the functions H_{11} and H_{jk} for $j, k = 2, \ldots, n+1$ evaluated at z correspond to the second derivatives of F at z. The nondegeneracy of $D^2 F(z)$ implies that $H_{11}(z) \neq 0$; by continuity H_{11} does not vanish in a neighborhood of z. After replacing $y_1 \sqrt{\pm H_{11}}$ by x_1, we can assume that $H_{11} = \pm 1$. Finally, the sum in (2-9) can be written as $\sum_{j \geqslant 2} \epsilon_j x_j^2$ ($\epsilon_j = \pm 1$) by the classical Morse lemma [11, Lemma 2.2]. $\square$

The next result is completely standard by now.

Lemma 2-10 *Assume that F is a Morse function on a cobordism (Ω, Y) between (Σ_0, M_0) and (Σ_1, M_1). If F has no critical points, then $(\Omega, Y) \cong (\Sigma_0, M_0) \times [0, 1]$. Furthermore, we can choose the diffeomorphism to map the level set $F^{-1}(t)$ to the set $\Sigma_0 \times \{t\}$.*

Proof The proof is identical to the classical case; see eg [12, Theorem 3.4]. $\square$

2.2 Half-handles

For any k we consider the k–dimensional disk $D^k = \{x_1^2 + \cdots + x_k^2 \leqslant 1\}$. In the classical theory, an n–dimensional handle of index k is the n–dimensional manifold $H = D^k \times D^{n-k}$ with boundary

$$\partial H = \left(\partial D^k \times D^{n-k} \right) \cup \left(D^k \times \partial D^{n-k} \right) = B_0 \cup B_0'.$$

Given an n–manifold with boundary $(\Sigma, \partial \Sigma)$ and a distinguished embedding $\phi \colon B_0 \to \partial \Sigma$, the effect of a classical handle attachment is the n–dimensional manifold with boundary

$$(\Sigma', \partial \Sigma') = (\Sigma \cup H, (\partial \Sigma \setminus B_0) \cup B_0'),$$

where we glue along $\phi(B_0)$ identified with B_0. The boundary $\partial \Sigma'$ is the effect of surgery on $\phi(B_0) \subset \partial \Sigma$. We now extend this construction to relative cobordisms of manifolds with boundary, using "half-handles". Since our ambient space Ω is $(n+1)$–dimensional, $(n+1)$ is the dimension of the handles, and they induce n–dimensional handle attachments on Y.

In order to do this, for any $k \geqslant 1$ we distinguish the following subsets of D^k: the "half-disk" $D_+^k := D^k \cap \{x_1 \geqslant 0\}$, and its boundary subsets $S_+^{k-1} := \partial D^k \cap \{x_1 \geqslant 0\}$, $S_0^{k-2} := \partial D^k \cap \{x_1 = 0\}$ and $D_0^{k-1} := D^k \cap \{x_1 = 0\}$. Clearly, S_0^{k-2} is a boundary of the two $(k-1)$–disks S_+^{k-1} and D_0^{k-1}; see Figure 2. We will call x_1 the *cutting coordinate*.

Definition 2-11 Let $0 \leqslant k \leqslant n$. An $(n+1)$–*dimensional right half-handle of index* k is the $(n+1)$–dimensional manifold $H_{right} = D^k \times D_+^{n+1-k}$, with boundary subdivided into three pieces $\partial H_{right} = B \cup C \cup N$, where

$$B := \partial D^k \times D_+^{n+1-k}, \quad C := D^k \times D_0^{n-k}, \quad N := D^k \times S_+^{n-k}.$$

One also has the intersections

$$B_0 := C \cap B = \partial D^k \times D_0^{n-k}, \quad N_0 := C \cap N = D^k \times S_0^{n-k-1}.$$

Hence the handle H is cut along C into two pieces, one of them is the half-handle H_{right}. Note that (C, B_0) is a n–dimensional handle of index k. See Figure 3 for an example of a right half-handle.
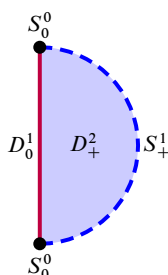

Figure 2: Various parts of a "half-disk" explaining the notation introduced before Definition 2-11.
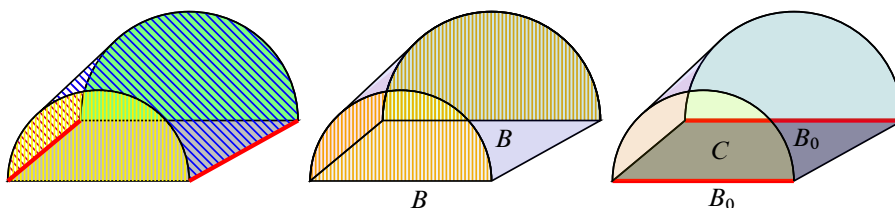

Figure 3: A right half-handle of index 1. The two half-disks form B, while C is the bottom rectangle.

Symmetrically, we define the left half-handles by cutting the handle H along the left–component disk D^k; see Figure 4.

Definition 2-12 Fix k with $1 \leqslant k \leqslant n+1$. An $(n+1)$–*dimensional left half-handle of index* k is the $(n+1)$–dimensional disk $H_{left} := D_+^k \times D^{n+1-k}$ with boundary subdivided into three pieces $\partial H_{left} = B \cup C \cup N$, where

$$B := S_+^{k-1} \times D^{n+1-k}, \quad C := D_0^{k-1} \times D^{n+1-k}, \quad N := D_+^k \times \partial D^{n+1-k}.$$

We also set $B_0 := C \cap B = S_0^{k-2} \times D^{n+1-k}$ and $N_0 := N \cap C = D_0^{k-1} \times \partial D^{n+1-k}$.
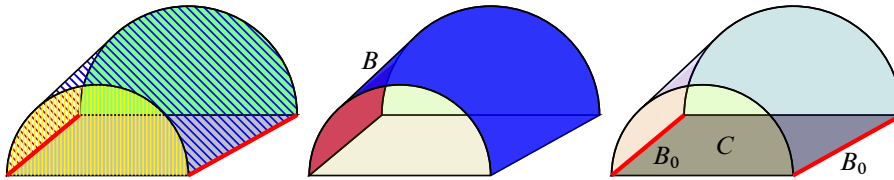
Figure 4: A left half-handle of dimension 3 and index $k = 2$. The two lines are B_0, the bottom rectangle is C. B is the surface between the two half circles in the picture.

Remark 2-13 The right half-handle and left-half handle are abstractly diffeomorphic to an $(n + 1)$–dimensional disk. The right half-handle and left-half handle are each abstractly diffeomorphic to an $(n + 1)$–dimensional disk. The difference is that the boundary is split into several components and this splitting is different for right half-handles and left half-handles.

A *half-handle* will from now on refer to either a right half-handle or left half-handle. We pass to half-handle attachments. We will attach a half-handle along B. The definitions of the right half-handle attachment and the left half-handle attachment are formally very similar, but there are significant differences in the properties of the two operations.

Definition 2-14 Let $(\Omega, Y; \Sigma_0, M_0, \Sigma_1, M_1)$ be an $(n + 1)$–dimensional relative cobordism. Given an embedding $\Phi\colon (B, B_0) \hookrightarrow (\Sigma_1, M_1)$ define the relative cobordism $(\Omega', Y'; \Sigma_0, M_0, \Sigma_1', M_1')$ obtained from $(\Omega, Y; \Sigma_0, M_0, \Sigma_1, M_1)$ by *attaching a (right or left) half-handle of index k* by

$$\Omega' = \Omega \cup_B H, \qquad Y' = Y \cup_{B_0} C,$$
$$\Sigma_1' = (\Sigma_1 \setminus B) \cup N, \quad M_1' = (M_1 \setminus B_0) \cup N_0.$$

See Figure 5 and Figure 6 for right, respectively left half-handle attachments.

We point out that in the case of the right half-handle attachment, any embedding of B_0 into M_1 determines (up to an isotopy) an embedding of pairs $(B, B_0) \hookrightarrow (\Sigma_1, M_1)$. Indeed, as $(B, B_0) = \partial D^k \times (D_+^{n+1-k}, D_0^{n-k})$, a map $\phi\colon B_0 \hookrightarrow M_1$ extends to a map $\Phi\colon B \hookrightarrow \Sigma_1$ in a collar neighbourhood of M_1 in Σ_1. (This is not the case in the left half-handle attachment.)

In particular, *in the case of right attachments, we specify only the embedding $B_0 \hookrightarrow M_1$*.

Example 2-15 (a) The right half-handle attachment of index 0 is the disconnected sum $\Omega \sqcup D^{n+1}_+$ with boundary $\partial D^{n+1}_+ = S^n_+ \cup D^n_0$. We think of the first disk S^n_+ as a part of Σ_1', while the second disk as a part of Y', and $M_1' = M_1 \cup S^{n-1}_0$.
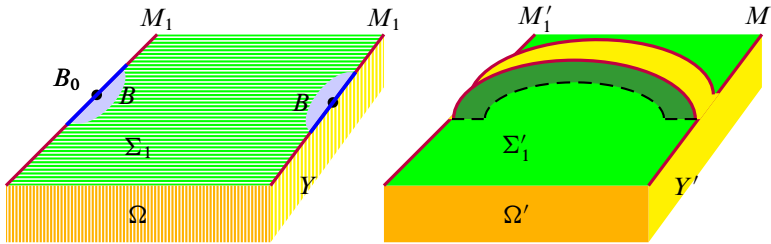
Figure 5: Right half-handle attachment. Here $k = 1$, $n = 2$. On the right, the two black points represent a sphere S^0 with a neighborhood B_0 in M_1 and B in Σ_1. In the picture on the right the dark green colored part of the handle belongs to Σ_1, the dashed lines belong to Σ_1 and are drawn only to make the picture look more "three-dimensional".
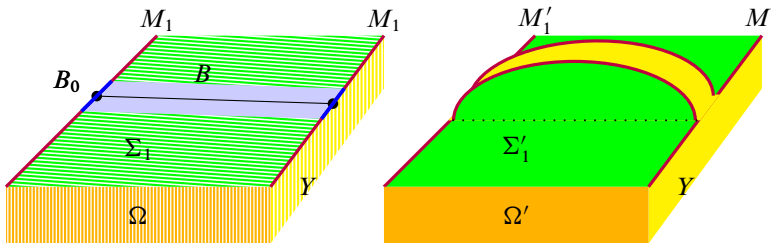

Figure 6: Left half-handle attachment with $k = 2$ and $n = 2$. This time the sphere on the left (denoted by two points) bounds a disk in Σ_1.

(b) We describe the left half-handle attachment for $k = 1$. In this case B_0 is empty. If we are given an embedding of $B \cong \{1\} \times D^n$ into $\Sigma_1 \setminus M_1$, we glue $[0, 1] \times D^n$ to Ω along B. Then we set $Y' = Y \sqcup \{0\} \times D^n$, $\Sigma_1' = (\Sigma_1 \setminus B) \cup [0, 1] \times \partial B$ and $M' = M \sqcup \{0\} \times \partial B$.

Example 2-16 There is another way of looking at left half-handle attachments. Suppose we are given a model of Ω (in Figure 7) made of clay. The height function is the Morse function F. On the top of Ω, that is on Σ_1, we specify an arc γ with boundary in M_1 (the arc is a 1–dimensional disk, that is, in our situation $k = 1 + 1 = 2$). We press down slightly a tubular neighborhood of the arc as on the right side of Figure 7. The resulting manifold is a result of a left half–handle attachment of index 2. Notice that for index-1 left half–handle attachment we should have specified a disk inside Σ_1 (with boundary disjoint from M_1).

Remark 2-17 In the next subsection we shall see that crossing a boundary stable critical point corresponds to a left half-handle attachment, while a boundary unstable
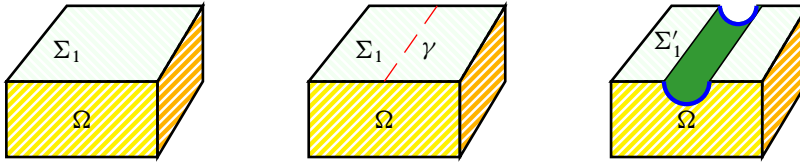
Figure 7: A "clay" variant of a left half-handle attachment as explained in Example 2-16. There are only small differences between this picture and Figure 6. Here the handle is "pushed down inside Ω", in the formal definition it is glued on top of Σ_1.

critical point corresponds to a right half-handle attachment. Theorem 3-1 can be interpreted informally as splitting a handle into a right half-handle and a left half-handle. This also motivates the name "half-handle".

2.3 Elementary properties of half-handle attachments

The following results are trivial consequences of the definitions.

Lemma 2-18 *Let Ω' be the result of a right half-handle attachment to Ω along $(B, B_0) \hookrightarrow (\Sigma_1, M_1)$. Let B' be B pushed slightly off M_1 into the interior of Σ_1. Let $\widetilde{\Omega}$ be the result of attaching a (standard) handle of index k to Ω along B'. Then Ω' and $\widetilde{\Omega}$ are diffeomorphic.*

Proof When we forget about C and B_0, the pair (H_{right}, B) is a standard $(n+1)$–dimensional handle of index k. ☐

For instance, the effect of a right half-handle attachment on Ω is the same as the effect of a standard handle attachment of the same index.

The situation is completely different in the case of left half-handle attachments.

Lemma 2-19 *If Ω' is the result of a left half-handle attachment, then Ω' is diffeomorphic to Ω.*

Proof By definition the pair (H_{left}, B) is diffeomorphic to the pair $(D^n \times [0, 1], D^n \times \{0\})$. Attaching H_{left} along B to Ω does not change the diffeomorphism type of Ω. ☐

The effect on Y of right and left half-handle attachments are almost the same, the only difference is the index shift by 1.

Lemma 2-20 *If $(\Omega', Y'; \Sigma_0, M_0, \Sigma'_1, M'_1)$ is the result of a left (respectively, right) half-handle attachment to $(\Omega, Y; \Sigma_0, M_0, \Sigma_1, M_1)$ along $(B, B_0) \hookrightarrow (\Sigma_1, M_1)$, then Y' is the result of a classical handle attachment of index $k - 1$ (respectively k) along B_0.*

Proof This follows immediately from Definition 2-14. □

The effects of half handle attachment on Σ are also easily described. The next lemma is a direct consequence of the definitions; its proof is omitted. We refer to Figures 5 and 6.

Lemma 2-21 (a) *If $(\Omega', Y'; \Sigma_0, M_0, \Sigma'_1, M'_1)$ is the result of index k right half-handle attachment to $(\Omega, Y; \Sigma_0, M_0, \Sigma_1, M_1)$ along $B_0 \hookrightarrow M_1$, then Σ'_1 is diffeomorphic to $\Sigma_1 \cup_{B_0} N$, where N is an n–dimensional disk $D^k \times D^{n-k}$ and $B_0 = S^{k-1} \times D^{n-k}$.*

(b) *Suppose $(\Omega', Y'; \Sigma_0, M_0, \Sigma'_1, M'_1)$ is the result of left half-handle attachment to $(\Omega, Y; \Sigma_0, M_0, \Sigma_1, M_1)$ along $(B, B_0) \hookrightarrow (\Sigma_1, M_1)$. Then Σ'_1 is diffeomorphic to $\Sigma_1 \setminus B$.*

Example 2-22 Suppose $n = 3$, so Σ_1 and Σ'_1 are three-dimensional manifolds with boundary. The effects on Σ_1 of left and right half-handle attachments to Ω are the following (the numbers $0, 1, 2, 3, 4$ are the index of a handle, "L" and "R" stand for "left" and "right"):

(0R) Σ'_1 is a disjoint union of Σ_1 and a 3–ball.

(1L) A 3–ball is removed from the interior of Σ_1.

(1R) A 1–handle (that is a thickened arc) is added to Σ_1. The attaching region is formed by thickening two point on $\partial \Sigma_1$.

(2L) An arc γ is chosen inside Σ_1 such that $\partial \gamma \subset \partial \Sigma_1$. The manifold Σ'_1 is then Σ_1 with a tubular neighborhood of γ removed.

(2R) A 2–handle (a thickened two-disk) is added to Σ_1. The attaching region is formed by thickening a circle belonging to $\partial \Sigma_1$.

(3L) A disk D is specified inside Σ_1 such that $\partial D \subset \partial \Sigma_1$. Then a tubular neighborhood of D is removed.

(3R) A 3–handle (a ball) is added to Σ_1. Notice that adding a 3–ball destroys one component of the boundary.

(4L) A connected component of Σ_1 that is a ball, is removed from Σ_1. This is the opposite of the (0R) move.

Lemma 2-21 and Example 2-22 emphasize that right half-handle attachments and left half-handle attachments are somehow dual operations on Σ. This can be seen also at the Morse function level: changing a Morse function F to $-F$ changes all right half-handles to left-half handles and conversely; see Section 2.4 and 2.5 below. But the above lemma shows another aspect as well: a right half handle attachment consists of gluing a disk, a left half-handle attachment consists of removing a disk. Indeed, in the case of right attachment, $(\Sigma_1', M_1') = (\Sigma_1 \cup D^k \times D^{n-k}, \partial\Sigma_1')$ is associated with an embedding $\Phi \colon \partial D^k \times D^{n-k} \to M_1$. On the other hand, by definition, for an embedding $\Phi' \colon (D^{k-1} \times D^{n+1-k}, \partial D^{k-1} \times D^{n+1-k}) \to (\Sigma_1, M_1)$, the pair

$$(\Sigma_1', M_1') = (\text{closure of } (\Sigma_1 \setminus D^{k-1} \times D^{n+1-k}), \partial\Sigma_1')$$

is obtained from (Σ_1, M_1) by a *handle detachment* of index $k-1$. We formulate this observation as a rephrasing of Lemma 2-21.

Corollary 2-23 *The effect on (Σ_1, M_1) of a right half–handle attachment of index k is a handle attachment of index k to (Σ_1, M_1). Likewise, the effect on (Σ_1, M_1) of a left half-handle attachment of index k is a handle detachment of index $k-1$. In particular, M_1' is obtained from M_1 as the result of a k surgery in the first case, and $k-1$ surgery in the second.*

The duality can also be seen as follows: we can cancel any handle attachment by a suitably defined handle detachment, and conversely.

The following definition introduces a terminology which is rather self-explanatory. We include it for completeness of the exposition.

Definition 2-24 A cobordism (Ω', Y') between (Σ, M) and (Σ', M') is a *right (respectively left) half-handle attachment* of index k if $(\Omega', Y', \Sigma', M')$ is a result of right (respectively left) half-handle attachments of index k (in the sense of Definition 2-14) to $(\Sigma \times [0,1], M \times [0,1], \Sigma \times \{0\}, M \times \{0\}, \Sigma \times \{1\}, M \times \{1\})$.

We conclude this section by studying homological properties of handle attachment. These properties will be used in [3]. The proofs are standard and are left to the reader.

Let (H_+, C, B, N) be a half-handle of index k.

Lemma 2-25 *If (H_{right}, C, B, N) is a right half-handle, then the pair (C, B_0) is a strong deformation retract of (H_+^r, B), while $(D^k, \partial D^k)$ is a strong deformation retract of (C, B_0). In particular, $H_j(H_+^r, B) \cong H_j(C, B_0) = \mathbb{Z}$ for $j = k$, and it is zero otherwise.*

The situation is completely different for left half-handles.

Lemma 2-26 *If (H_{left}, C, B, N) is a left half-handle, then the pair (H_{left}, B) retracts onto the trivial pair* (point, point). *In particular, all the relative homologies $H_*(H_{left}, B)$ vanish. On the other hand, (D_0^{k-1}, S_0^{k-2}) is a strong deformation retract of (C, B_0), hence $H_j(C, B_0) = \mathbb{Z}$ for $j = k - 1$, and it is zero otherwise. Therefore, the inclusion $(C, B_0) \hookrightarrow (H_{left}, B)$ induces a surjection on homologies.*

2.4 Boundary critical points and half-handles

Consider a Morse function F on a cobordism (Ω, Y) and assume that it has a single boundary critical point z of index k with critical value c and no interior critical points.

Theorem 2-27 *If z is boundary stable (unstable), then the cobordism is a left (right) half-handle attachment of index k respectively.*

Proof We can assume that $c = F(z) = 0$. Let us chose a neighborhood U of z in Ω. Shrinking U if necessary, we can assume that there are Morse coordinates $x_1, \ldots, x_{n+1}$ on U (see Lemma 2-6) and in these coordinates U is a half-ball of radius 2ρ for some positive number ρ:

$$U = \{x_1^2 + \cdots + x_{n+1}^2 \leqslant 4\rho^2\} \cap \{x_1 \geqslant 0\}.$$

The intersection $Y \cap U$ defined by $\{x_1 = 0\}$, and

$$F(x_1, \ldots, x_{n+1}) = -a^2 + b^2,$$

where if z is boundary stable we set

$$(2\text{-}28) \qquad a^2 = x_1^2 + x_2^2 + \cdots + x_k^2, \quad b^2 = x_{k+1}^2 + \cdots + x_{n+1}^2 \quad (k \geqslant 1),$$

and if z is boundary unstable

$$(2\text{-}29) \qquad a^2 = x_2^2 + \cdots + x_{k+1}^2, \quad b^2 = x_1^2 + x_{k+2}^2 + \cdots + x_{n+1}^2 \quad (k \geqslant 0).$$

We also assume that $x_1, \ldots, x_{n+1}$ is an orthonormal Euclidean coordinate system.

Next, we consider $\varepsilon > 0$ such that $\varepsilon \ll \rho$, and we define a space $\tilde{H}$ bounded by the following conditions (see Figure 8):

$$\tilde{H} := \{-a^2 + b^2 \in [-\varepsilon^2, \varepsilon^2], \, a^2 b^2 \leqslant \rho^4 - \varepsilon^4, \, x_1 \geqslant 0\}.$$

Observe that

$$\tilde{H} \subset U.$$

Let us now define the following parts of the boundary of $\tilde{H}$:

$$
\begin{aligned}
(2\text{-}30) \quad
&\tilde{B} = \partial \tilde{H} \cap \{-a^2 + b^2 = -\varepsilon^2\} \subset F^{-1}(-\varepsilon^2), \\
&\tilde{P} = \partial \tilde{H} \cap \{-a^2 + b^2 = \varepsilon^2\} \subset F^{-1}(\varepsilon^2), \\
&\tilde{K} = \partial \tilde{H} \cap \{a^2 b^2 = \rho^4 - \varepsilon^4\}, \\
&\tilde{C} = \partial \tilde{H} \cap \{x_1 = 0\} \subset Y.
\end{aligned}
$$

We have $\tilde{B} \cup \tilde{P} \cup \tilde{K} \cup \tilde{C} = \partial \tilde{H}$ (in Figure 8 we do not see $\tilde{C}$, because this would require one more dimension). If z is boundary unstable and $k = 0$ in (2-29) then the term a^2 is missing and $\tilde{B} = \varnothing$. Otherwise $\tilde{B} \neq \varnothing$.
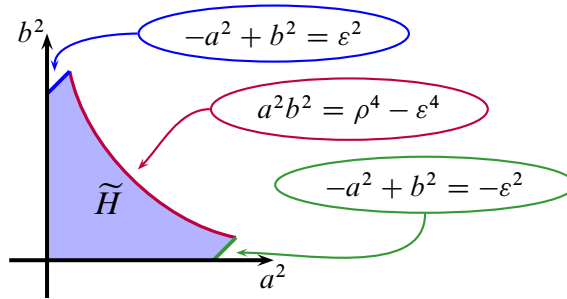
Figure 8: A schematic presentation of $\tilde{H}$, $\tilde{B}$, $\tilde{P}$, $\tilde{K}$ from the proof of Theorem 2-27. To each point (a^2, b^2) in $\tilde{H}$ on the picture correspond all those points $(x_1, \dots, x_{n+1})$ for which (2-28) or (2-29) holds and $x_1 \geqslant 0$.

Lemma 2-31 *The flow of ∇F is tangent to $\tilde{K}$.*

Proof Assume the critical point is boundary stable. The differential equation

$$\frac{d\boldsymbol{x}}{dt} = \nabla F = (-2x_1, \dots, -2x_k, 2x_{k+1}, \dots, 2x_{n+1})$$

has a solution

$$(x_1, \dots, x_{n+1}) \to (e^{-2t} x_1, \dots, e^{-2t} x_k, e^{2t} x_{k+1}, \dots, e^{2t} x_{n+1}).$$

It follows that $a^2 \to e^{-4t} a^2$ and $b^2 \to e^{4t} b^2$, and the hypersurface $a^2 b^2 = $ constant is preserved by the flow of ∇F. $\qquad \square$

Lemma 2-32 *The inclusion of pairs of spaces*

$$\left(F^{-1}(-\varepsilon^2) \cup_{\widetilde{B}} \widetilde{H}, Y \cap \left(F^{-1}(-\varepsilon^2) \cup_{\widetilde{B}} \widetilde{H}\right)\right) \subset (\Omega, Y)$$

admits a strong deformation retract.

Proof By Lemma 2-10 we can assume that

$$(\Omega, Y) = (F^{-1}([-\varepsilon^2, \varepsilon^2]), Y \cap F^{-1}([-\varepsilon^2, \varepsilon^2])).$$

First we assume that $\widetilde{B}$ is not empty, and it is given by the Equation (2-30) in U. Set $\Sigma_- = $ closure of $(F^{-1}(-\varepsilon^2) \setminus \widetilde{B})$ and let T_- be the part of the boundary of Σ_- given by

$$T_- = \text{closure of } (\partial \Sigma_- \setminus \partial F^{-1}(-\varepsilon^2)).$$

We have $T_- \subset \widetilde{B}$; see Figure 9. Let us choose a collar of T_- in Σ_-, that is a subspace $U_- \subset \Sigma_-$ diffeomorphic to $T_- \times [0, 1]$, T_- identified with $T_- \times \{0\}$ and $\partial T_- \times [0, 1] \subset \partial \Sigma_- \cap \partial F^{-1}(-\varepsilon^2)$. Let T'_- be the space identified with $T_- \times \{1\}$ by this diffeomorphism.

Similarly, let $\Sigma_+ = $ closure of $(F^{-1}(\varepsilon^2) \setminus \widetilde{P})$, and $T_+ = $ closure of $(\partial \Sigma_+ \setminus \partial F^{-1}(\varepsilon^2))$. We also define Ω_0 as the closure of $\Omega \setminus \widetilde{H}$. Clearly F has no critical points in Ω_0 and ∇F is everywhere tangent to $\partial \Omega_0 \setminus (\Sigma_- \cup \Sigma_+) = (Y \cap \Omega_0) \cup \widetilde{K}$ by Lemma 2-31. In particular, by Lemma 2-10, the flow of ∇F on Ω_0 yields a diffeomorphism between Σ_- and Σ_+, mapping T_- to T_+. We define $V \subset \Omega$ as the closure of the set of points v such that a trajectory going through v hits U_-. Lemma 2-10 implies that there is a diffeomorphism $V \cong T_- \times [0, 1] \times [-\varepsilon^2, \varepsilon^2]$ such that for $(x, t, s) \in V$ we have $F(x, t, s) = s$. Finally, we also define $V^* := \{(x, t, s) \in V : s \leqslant \varepsilon^2(1 - 2t)\}$.
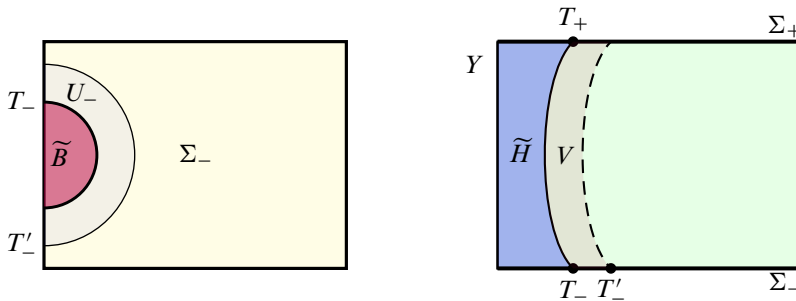

Figure 9: Notation used in Lemma 2-32. Note that the left picture is drawn on Σ_-, while the right one is on Ω.

We define the contraction in two steps: vertical and horizontal. The vertical contraction is defined as follows. For $v \in \widetilde{H} \cup V^*$ we define $\Pi_V(v) = v$. For a point $v \in \Omega_0 \setminus V$

we take for $\Pi_V(v)$ the unique point $s \in \Sigma_-$ such that a trajectory of ∇F goes from s to v. Finally if $v = (x,t,s) \in V \setminus V^*$ we define $\Pi_V(v) = (x,t,\varepsilon^2(1-2t))$.

By construction, the image of Π_V is $\tilde{H} \cup V^* \cup F^{-1}(-\varepsilon^2)$. Next, we define Π_H. Note that Π_H will be defined only on the image of Π_V.

$\Pi_H = \mathrm{id}$ on $\tilde{H} \cup F^{-1}(-\varepsilon^2)$, and maps $(x,t,s) \in V^*$ to $(x, t - (\varepsilon^2 + s)/(2\varepsilon^2), -\varepsilon^2)$ if $s \leqslant \varepsilon^2(2t-1)$, and to $(x,0,s-2\varepsilon^2 t)$ otherwise. Note that the expressions agree for any (x,t,s) with $s = \varepsilon^2(2t-1)$ and these points are sent to $(x,0,-\varepsilon^2)$. Both Π_H and Π_V are continuous retractions, by smoothing corners we can modify them into smooth retractions; also they can be extended in a natural way to strong deformation retracts. By construction, the retracts preserve Y too. See also Figure 10.

If $\tilde{B}$ is empty, then $\tilde{H}$ is necessarily a unstable (right) half-handle of index 0, $F^{-1}([-\varepsilon^2, \varepsilon^2])$ is a disconnected sum of $\tilde{H}$ and the manifold $F^{-1}(-\varepsilon^2) \times [-\varepsilon^2, \varepsilon^2]$. $\square$
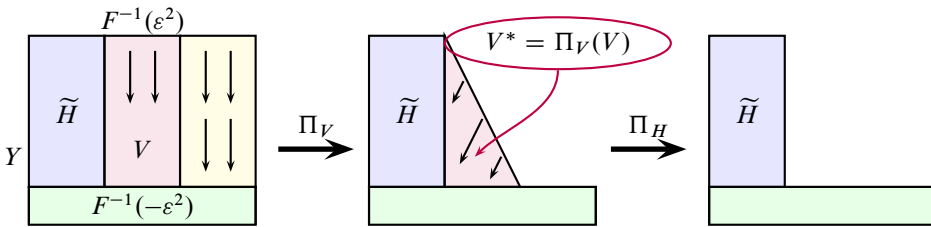
Figure 10: Contractions Π_H and Π_V from the proof of Lemma 2-32. The set V is now drawn as a rectangle.

Continuation of the proof of Theorem 2-27 We want to show that $\tilde{H}$ is a half-handle.

By Section 2.2 we have the following description in local coordinates of the left half-handle (2-33) and right half-handle (2-34) with cutting coordinate x_1:

(2-33) $\quad H_{left} = \{x_1^2 + \cdots + x_k^2 \leqslant 1\} \cap \{x_{k+1}^2 + \cdots + x_{n+1}^2 \leqslant 1\} \cap \{x_1 \geqslant 0\},$

(2-34) $\quad H_{right} = \{x_2^2 + \ldots + x_{k+1}^2 \leqslant 1\} \cap \{x_1^2 + x_{k+2}^2 + \ldots + x_{n+1}^2 \leqslant 1\} \cap \{x_1 \geqslant 0\}.$

We consider the subsets R and S of $\mathbb{R}^2$ given by

$$R = \{(u,v) \in \mathbb{R}^2 : u \geqslant 0, \, v \geqslant 0, \, uv \leqslant \rho^4 - \varepsilon^4, \, -u + v \in [-\varepsilon^2, \varepsilon^2]\},$$
$$S = \{(u,v) \in \mathbb{R}^2 : u \in [0,\varepsilon], \, v \in [0,\varepsilon]\}.$$

(Note that R can be seen in Figure 8 if we replace a^2 by u and b^2 by v.) These subsets are clearly diffeomorphic. We choose a diffeomorphism ψ that maps the edge

of R given by $\{-u + v = -\varepsilon^2\}$ to the edge $\{u = \varepsilon\}$ of S and the images of coordinate axes are the corresponding coordinate axes.

We use ψ to construct a diffeomorphism Ψ between $\widetilde{H}$ and H_{right} (respectively H_{left}) as follows. First let us write $\psi(u, v) = (\psi_1(u, v), \psi_2(u, v))$. As ψ maps axes to axes, we have $\psi_1(0, v) = 0$ and $\psi_2(u, 0) = 0$. Furthermore $\psi_1, \psi_2 \geqslant 0$. By Hadamard's lemma there exist smooth functions ξ and η such that

$$\psi(u, v) = (u\xi(u, v)^2, v\eta(u, v)^2).$$

We define now

$$\Psi(x_1, \ldots, x_{n+1}) = (\xi(a, b)x_1, \ldots, \xi(a, b)x_k, \eta(a, b)x_{k+1}, \ldots, \eta(a, b)x_{n+1})$$

if z is boundary stable, and

$$\Psi(x_1, \ldots, x_{n+1})$$
$$= (\eta(a, b)x_1, \xi(a, b)x_2, \ldots, \xi(a, b)x_k, \eta(a, b)x_{k+1}, \ldots, \eta(a, b)x_{n+1})$$

if z is boundary unstable. Here a and b are given by (2-28) or (2-29). By construction, Ψ maps $(\widetilde{H}, \widetilde{B}, \widetilde{C})$ diffeomorphically to the triple (H, B, C), where

$$H = \{a^2 \in [0, \varepsilon^2],\ b^2 \in [0, \varepsilon^2],\ x_1 \geqslant 0\},$$
$$B = \{a^2 = -\varepsilon^2,\ b^2 \in [0, \varepsilon^2],\ x_1 \geqslant 0\},$$
$$C = \{a^2 \in [0, \varepsilon^2],\ b^2 \in [0, \varepsilon^2],\ x_1 = 0\}.$$

After substituting for a and b the values from (2-28) or (2-29) (depending on whether z is boundary stable or unstable), we recover the model (2-34) of a right half-handle if z is boundary unstable; or the model (2-33) of a left half-handle (both of index k). $\square$

The fact that each half-handle can be presented in a left or right model will be now used to show the following converse to Theorem 2-27. The result for nonboundary case can be found in [12, Theorem 3.12].

Proposition 2-35 *Let $(\Omega, Y) = (\Sigma_0 \times [0, 1], M_0 \times [0, 1])$ be a product cobordism between (Σ_0, M_0) and $(\Sigma_1, M_1) \cong (\Sigma_0, M_0)$. Let us be given a half-handle (H, C, B) of index k and an embedding of $B_0 = C \cap B$ into M_1 (respectively an embedding of (B, B_0) into (Σ_1, M_1)), and let (Ω', Y') be the result of a right half-handle attachment along B_0 (respectively, a left half-handle attachment along (B, B_0)) of index k. Then, there exists a Morse function $F\colon (\Omega', Y') \to \mathbb{R}$, which has a single boundary unstable critical point (respectively, a single boundary stable critical point) of index k on H and no other critical points. In particular, F is a Morse function on a cobordism (Ω', Y').*

Proof We shall prove the result for right half-handle attachment, the other case is completely analogous. The proof consists mostly on reading "back to front" the proof of Theorem 2-27; we shall use notation from this theorem, with $\varepsilon = 1$ and $\rho = 2$.

In the case of a right half-handle B_0 is embedded into M_1 and we extend this embedding to an embedding of B into Σ_1 (see Definition 2-14 and the remark just after it).

The manifold Ω' is constructed in two steps. First, we glue a handle $\widetilde{H}$ to $\Omega = \Sigma_0 \times [0, 1]$ along B obtaining a manifold Ω''. The result is as in Figure 10 (the figure on the right). After this gluing, a vertical component of $\partial\widetilde{H}$ appears (in notation of (2-30) this vertical component is $\widetilde{K}$).

We glue now $\Sigma_- \times [-1, 1]$ to Ω'' so as to obtain Ω' as in Figure 10 on the left and in the way that Ω'' is diffeomorphic to Ω'. The way we do that is the following. The boundary of $\Sigma_- \times [-1, 1]$ decomposes into three parts. The first part is $\Sigma_- \times \{-1\}$; we glue it to $\Sigma_0 \times \{1\}$ (notice that Σ_- is Σ_0 with B removed). The second part is $\partial\Sigma_- \times [-1, 1]$. This part is identified with $\widetilde{K}$, in fact, the flow of ∇F studied in Lemma 2-31 induces a diffeomorphism of $\widetilde{K}$ with $\partial\Sigma_- \times [-1, 1]$. We glue together $\widetilde{K}$ and $\partial\Sigma_- \times [-1, 1]$ using this identification. The third part of the boundary, that is, $\Sigma_- \times \{1\}$, is not glued. It follows from an argument as in Lemma 2-32 that Ω' is diffeomorphic to Ω''.

The manifold Ω' consists of three components: $\Sigma_0 \times [0, 1]$, H and $\Sigma_- \times [-1, 1]$. We define a function F on each component separately, namely.

$$F(x) = \begin{cases} t & \text{if } x = (v, t) \in \Sigma_0 \times \{t\} \subset \Sigma_0 \times [0, 1] = \Omega, \\ 2 + t & \text{if } x = (v, t) \in \Sigma_- \times \{t\} \subset \Sigma_- \times [-1, 1], \\ 2 - \sum_{j=1}^{k} x_j^2 + \sum_{j=k+1}^{n+1} x_j^2 & \text{if } x = (x_1, \ldots, x_{n+1}) \in H. \end{cases}$$

As defined, F is smooth on each of the three components. It is also globally continuous. In fact, the identification of $\widetilde{K}$ with $\Sigma_- \times [-1, 1]$ can be done so that F is continuous on $\widetilde{K}$. On the part $\Sigma_0 \times \{1\} \subset \Omega'$, all the three components give the same value, that is 1.

Given the construction of F, it remains to perturb F (that is, to approximate it uniformly near $\widetilde{K} \cup \Sigma_0 \times \{1\}$) to a smooth function and in the way that F does not get any new critical points. For a general piecewise smooth function this is impossible, we can consider the real valued function $x \mapsto |x|$: any smooth approximation must have a critical point near $x = 0$. The reason for this is that near the nonsmooth point the topology of level sets of $|x|$ changes. This is essentially the main obstruction.

In our situation, the topology of the level sets of F does not change near $\widetilde{K}$, nor near $\Sigma_0 \times \{1\}$, that is, near any gluing region. This is enough to show that F can be approximated near its nonsmooth locus by a smooth function without introducing additional critical points. The proof of this fact is standard, but technical. Instead of giving all the details, we sketch a proof of a weaker result, Lemma 2-36. This result takes care of approximating the function F near $\Sigma_0 \times \{1\}$. Approximation near the whole of $\widetilde{K} \cup \Sigma_0 \times \{1\}$ follows essentially the same pattern and is left to the reader.

Given the approximation result, the proof of Proposition 2-35 is finished. $\qquad\square$

Lemma 2-36 (Approximating piecewise smooth functions by smooth functions) *Suppose that N is a smooth, compact manifold. Let $\pi\colon N \times [-1, 1] \to [-1, 1]$ be the projection onto the second factor. Let $N_0 = N \times \{0\}$, $N_+ = N \times [0, 1]$ and $N_- = N \times [-1, 0]$. Let $f\colon N \times [-1, 1]$ be a continuous function. Let f_+ and f_- be the restrictions to N_+ and N_- respectively. Suppose that*

  (a) *f_+ and f_- are smooth and have no critical points on $N \times [-1, 1]$;*

  (b) *$f_+^{-1}(0) = f_-^{-1}(0) = N_0$;*

  (c) *the image of f_+ is contained in $\mathbb{R}_{\geq 0}$ and the image of f_- is contained in $\mathbb{R}_{\leq 0}$;*

  (d) *the scalar product $\langle f_\pm, \nabla\pi \rangle$ is positive on $N_\pm \setminus N_0$.*

Then for any $\theta > 0$ there exist $\varepsilon, \delta \in (0, \theta)$ and a smooth function $g\colon N \times [-1, 1] \to \mathbb{R}$ such that

  (i) *g agrees with f_- on $N \times [-1, -\delta]$ and with f_+ on $N \times [\delta, 1]$;*

  (ii) *g takes values in $[-\varepsilon, \varepsilon]$ on $N \times [-\delta, \delta]$;*

  (iii) *g has no critical points on $N \times [-1, 1]$.*

Sketch of proof By compactness, the continuity of $f_\pm$ and assumptions (b), (c) there exists $\delta' > 0$ such that $f_+(N \times [0, \delta']) \subset [0, \theta/2]$ and $f_-(N \times [-\delta', 0]) \subset [-\theta/2, 0]$. We set $\varepsilon = \theta$ and $\delta = \min(\delta', \theta/2)$.

Choose a partition of unity subordinate to the covering $[-1, 1] = [-1, -\delta/2) \cup (-\delta, \delta) \cup (\delta/2, 1]$. The three functions corresponding to this partition are denoted by ϕ_-, ϕ_0 and ϕ_+ respectively.

Define $\Phi_\bullet\colon N \times [-1, 1] \to [0, 1]$ as compositions $\phi_\bullet \circ \pi$, where $\bullet$ is any of "+", "−" and "0". Consider the vector field

$$ v = \Phi_+ \nabla f_+ + \Phi_0 \nabla\pi + \Phi_- \nabla f_-. $$

By point (a) of the assumptions v is a smooth vector field. Assumption (d) implies that $\langle \nabla \pi, v \rangle > 0$ everywhere on $N \times [-1, 1]$, that is, v is a gradient–like vector field for π. In particular, v does not vanish on $N \times [-1, 1]$.

Let h be a positive C^∞ function. Set $v_h = hv$. Then v_h is also a gradient–like vector field for π. The trajectories of v coincide with those of v_h: multiplication by h changes only the speed of going along a trajectory.

We integrate the vector field v_h to a function $g_h \colon N \times [-1, 1] \to \mathbb{R}$. This means we first set $g_h \equiv f_-$ on $N \times \{-1\}$. Next, suppose $x \in N \times (-1, 1]$. Let $\gamma \colon U \to N \times [-1, 1]$ (here U is a closed interval) be a trajectory of v_h such that $\gamma(0) = x$. Since v_h is gradient–like for π, γ must have come from $N \times \{-1\}$ in the past, more precisely, there exist $t_x < 0$ and $y \in N \times \{-1\}$ such that $\gamma(t_x) = y$. We set

$$g_h(x) = g_h(y) - t_x.$$

Since v_h is smooth, by the implicit function theorem g_h is a smooth function.

Choosing the normalizing function h appropriately we can guarantee that $g := g_h$ satisfies (i). Namely, we set $h = \|\nabla f_-\|^{-2}$ so that the directional derivative $\langle v_h, \nabla f_- \rangle \equiv 1$ on $N \times [-1, -\delta]$. This implies that $g = f_-$ on $N \times [-1, -\delta]$. The choice of h on $N \times [-\delta, \delta]$ is such that the time the trajectory goes from a point $x_- \in N \times \{-\delta\}$ to some $x_+ \in N \times \{\delta\}$ is equal to $f_+(x_+) - f_-(x_-)$. The latter expression is positive by assumption (b). This implies that $g = f_+$ on $N \times \{\delta\}$ and condition (ii) is satisfied automatically. Finally we set $h = \|\nabla f_+\|^{-2}$ on $N \times [\delta, 1]$. The verification of condition (i) is straightforward.

As g_h is strictly increasing on trajectories of v_h, it cannot have any critical points. $\quad \square$

2.5 Left and right product cobordisms and traces of handle attachments

In this subsection we create a dictionary between surgery theoretical notions (traces of handle attachments and detachments) and Morse theoretical (additions of half-handles). The main result of this subsection, Proposition 2-38, is a direct consequence of the results proved earlier in the article.

To begin with, let (Ω, Y) be a cobordism between (Σ_0, M_0) and (Σ_1, M_1).

Definition 2-37 We shall say that Ω is a *left product cobordism* if $\Omega \cong \Sigma_0 \times [0, 1]$. Similarly, if $\Omega \cong \Sigma_1 \times [0, 1]$, then we shall say that Ω is a *right product cobordism*.

Proposition 2-38 (a) *If (Ω, Y) is a cobordism between (Σ_0, M_0) and (Σ_1, M_1) consisting only of left half-handle attachments, then it is a left-product cobordism. Likewise, if it consists only of right half-handle attachments, then it is a right product cobordism.*

(b) *Let $F\colon \Omega \to [0,1]$ be a Morse function in the sense of Definition 1-4. Assume that F has no critical points in the interior of Ω. If all critical points on the boundary are boundary stable, then F is a left-product cobordism. If all critical points are boundary unstable, then F is a right product cobordism.*

Proof Statements (a) and (b) are equivalent via Theorem 2-27 and Proposition 2-35. The stable-unstable (right-left) statements are also equivalent by replacing the Morse function F by $-F$. The stable case follows from Lemma 2-19. □

The next results of this subsection will be not used in this paper, but we add them because they bridge surgery techniques and applications, eg with [14] or [3].

In order to clarify what we wish, let us recall that by Theorem 2-27 if a Morse function F defined on a cobordism (Ω, Y) has only one critical point of boundary type then (Ω, Y) is a half-handle attachment. Proposition 2-35 is the converse of this; the (total) space of a half-handle attachment can be thought as a cobordism with a Morse function on it with only one critical point.

We wish to establish the analogues of these statements "at the level of Σ". In Section 2.3 we proved that the output of a right/left half-handle attachment induces a handle attachment/detachment at the level of Σ. The next lemma is the converse of this statement. (In fact, the output cobordism provided by it can be identified with the cobordism constructed in Proposition 2-35.)

Lemma 2-39 *Assume that (Σ_1, M_1) is the result of a handle attachment (respectively detachment) to (Σ_0, M_0). Then, there exists a cobordism $(\Omega, Y; \Sigma_0, M_0, \Sigma_1, M_1)$ such that $\Omega \cong \Sigma_1 \times [0, 1]$ (respectively $\Omega \cong \Sigma_0 \times [0, 1]$).*

Proof Assume that (Σ_1, M_1) arises from a handle attachment to (Σ_0, M_0), ie $\Sigma_1 = \Sigma_0 \cup D^k \times D^{n-k}$. Let us define $\Omega = \Sigma_1 \times [0, 1]$. The boundary $\partial \Omega$ can be split as

$$\partial \Omega = \left(\Sigma_0 \cup D^k \times D^{n-k}\right) \times \{0\} \cup \left(M_1 \times [0, 1]\right) \cup \left(\Sigma_1 \times \{1\}\right)$$
$$= \Sigma_0 \times \{0\} \cup Y \cup \Sigma_1 \times \{1\},$$

where $Y = D^k \times D^{n-k} \cup (M_1 \times [0, 1])$. Its $D^k \times D^{n-k}$ part can be "pushed inside" Ω transforming (diffeomorphically) Ω into a cobordism; see Figure 11.

An analogous construction can be used in the case of a handle detachment. If (Σ_1', M_1') is the result of a handle detachment from (Σ_0, M_0), then the trace of the handle detachment is the cobordism between (Σ_0, M_0) and (Σ_1', M_1') such that

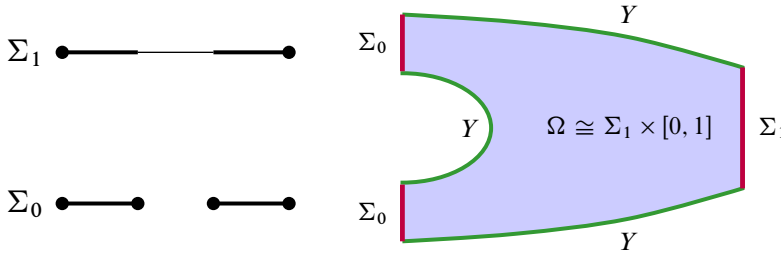$$(\Omega', Y') = (\Sigma_0 \times [0, 1], M_0 \times [0, 1] \cup D^k \times D^{n-k}). \qquad \square$$

Figure 11: Lemma 2-39. On the left a 1–handle is attached to Σ_0. On the right there is a cobordism between Σ_0 and Σ_1, which is a right product cobordism.

Definition 2-40 The cobordism $(\Omega, Y; \Sigma_0, M_0, \Sigma_1, M_1)$ determined by Lemma 2-39 is called the *trace* of a handle attachment of (Σ_0, M_0) (respectively the *trace* of a handle detachment).

3 Splitting interior handles

We prove here the theorem about moving critical points to the boundary.

Theorem 3-1 *Assume that on a cobordism* (Ω, Y) *between* (Σ_0, M_0) *and* (Σ_1, M_1) *we have a Morse function* F *with a single critical point* z *of index* $k \in \{1, \dots, n\}$ *in the interior of* Ω *situated on the level set* $\Sigma_{1/2} = F^{-1}(F(z))$. *If*

(3-2)   *the connected component of* $\Sigma_{1/2}$ *containing* z *has nonempty intersection with* Y,

then there exists a function $G \colon \Omega \to [0, 1]$, *such that:*

- G *agrees with* F *in a neighborhood of* $\Sigma_0 \cup \Sigma_1$.
- ∇G *is everywhere tangent to* Y.
- G *has exactly two critical points* z^s *and* z^u, *which are both on the boundary and of index* k. *The point* z^s *is boundary stable and* z^u *is boundary unstable.*
- *There exists a Riemannian metric such that there is a single trajectory of* ∇G *from* z^s *to* z^u *inside* Y.

Remark 3-3 A careful reading of the proof shows that we can in fact construct a smooth homotopy G_t such that $F = G_0$, $G = G_1$ and there exists $t_0 \in (0, 1)$ such that G_t has a single interior critical point for $t < t_0$, two boundary critical points for $t > t_0$ and a degenerate critical point on the boundary for $t = t_0$. See Remark 3-15.

The proof of Theorem 3-1 occupies Sections 3.2 to 3.4. We make a detailed discussion of Condition (3-2) in Section 3.5.

3.1 About the proof

The argument is based on the following two-dimensional picture. Consider the set $Z = \{(x, y) \in \mathbb{R}^2 : x \geqslant 0\}$ and the function $D \colon Z \to \mathbb{R}$ given by

$$D(x, y) = y^3 - yx^2 + ay,$$

where $a \in \mathbb{R}$ is a parameter. Observe that the boundary of Z given by $\{x = 0\}$ is invariant under the gradient flow of D (see Figure 12).

Lemma 3-4 *For $a > 0$, D has a single Morse critical point in the interior of Z. For $a < 0$, D has two Morse critical points on the boundary of Z.*

Proof Critical points of D are given by $\frac{\partial D}{\partial x} = \frac{\partial D}{\partial y} = 0$, that is, $xy = 0$ and $3y^2 - x^2 + a = 0$. The first equation means that $y = 0$ or $x = 0$ and then we get solutions $(\pm\sqrt{a}, 0)$ and $(0, \pm\sqrt{-a/3})$. In the case $a > 0$ we consider only first two solutions (and only one of them belongs to Z), while if $a < 0$, only the last two solutions are real and they correspond to boundary critical points. Checking that these critical points are Morse is straightforward and is left to the reader. $\qquad\square$

For $a = 0$, D acquires a D_4^- singularity at the origin (see eg [1, Section 17.1]).

In the proof of Theorem 3-1, we start by introducing "local/global" coordinates $(x, y, u_1, \dots, u_{n-1})$ at z, in which F has the form $D(x, y) \pm u_1^2 \pm \cdots \pm u_{n-1}^2$, hence it also parametrizes a neighborhood of a path connecting z with a point of Y. Then we change the parameter a (which we originally assume to be equal to 1) to $-\delta$, where δ is very small positive number (which corresponds to moving the critical point to the boundary along the chosen path).

3.2 Proof of Theorem 3-1 under an additional assumption

We first give the proof assuming the existence of such coordinate system as in 3.1, described explicitly in the next proposition (which is proved in Section 3.4). We use the hypotheses and notation of Theorem 3-1.
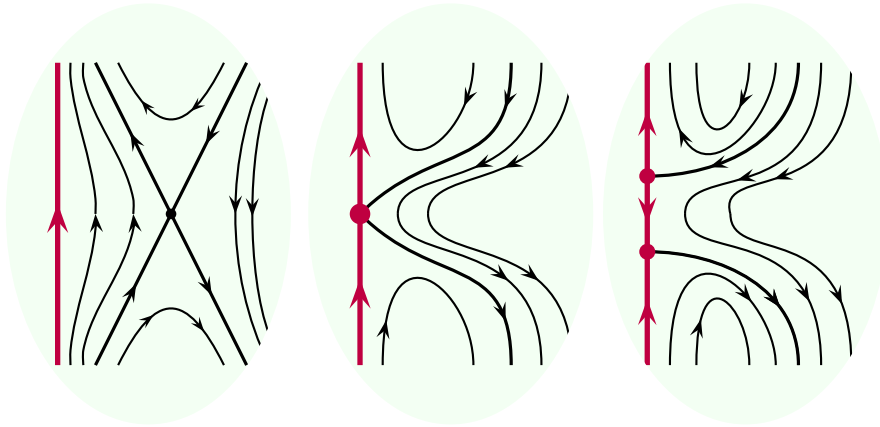
Figure 12: The trajectories of the gradient vector field of D for values of $a > 0$, $a = 0$ and $a < 0$.

Proposition 3-5 *There exists $\eta > 0$, $\eta \ll 1$ and an open "half-disk" $U \subset \Omega$, intersecting Y along a disk, and coordinates $x, y, u_1, \ldots, u_{n-1}$ such that in these coordinates U is given by*

$$0 \leqslant x < 3 + \eta, \quad |y| < \eta, \quad \sum_{j=1}^{n-1} u_j^2 < \eta^2,$$

$U \cap Y$ *is given by $\{x = 0\}$, and in these coordinates F is given by*

$$y^3 - yx^2 + y + \tfrac{1}{2} + \sum_{j=1}^{n-1} \epsilon_j u_j^2,$$

where $\epsilon_1, \ldots, \epsilon_{n-1} \in \{\pm 1\}$ are choices of signs. In particular $\#\{ j \colon \epsilon_j = -1\} = k - 1$, where $k = \operatorname{ind}_z F$.

Assuming the proposition, we prove Theorem 3-1. Let us introduce some abbreviations:

$$(3\text{-}6) \qquad \vec{u} = (u_1, \ldots, u_{n-1}), \quad \vec{u}^2 = \sum_{j=1}^{n-1} \epsilon_j u_j^2, \quad \|\vec{u}\|^2 = \sum_{j=1}^{n-1} u_j^2.$$

We fix a small real number $\varepsilon > 0$ such that $\varepsilon \ll \eta$ and two subsets $U_1 \subset U_2$ of U by

$$U_1 = \{|y| \leqslant \varepsilon, \, x \leqslant 3\} \cup \{(x-3)^2 + y^2 \leqslant \varepsilon^2\},$$
$$U_2 = \{|y| \leqslant 2\varepsilon, \, x \leqslant 3\} \cup \{(x-3)^2 + y^2 \leqslant 4\varepsilon^2\}.$$

The difference $U_{21} := \overline{U_2 \setminus U_1}$ splits into two subsets $S_1 \cup S_2$ (see Figure 13), where

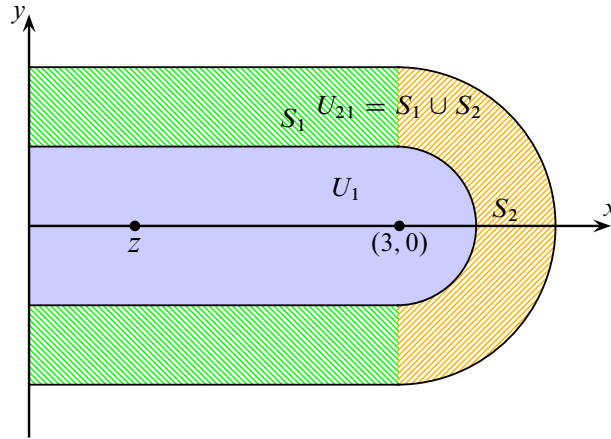$$S_1 = U_{21} \cap \{x \leqslant 3\}, \quad S_2 = U_{21} \cap \{x \geqslant 3\}.$$

Figure 13: Sets U_1, U_{21}, S_1 and S_2 in two dimensions (coordinates x and y).

For a point $v = (x, y, u_1, \ldots, u_{n-1}) \in U$, let us define

$$\widetilde{s}(v) = \begin{cases} 1 & \text{if } v \in U_1, \\ 0 & \text{if } v \in \overline{U \setminus U_2}, \\ 2 - \frac{|y|}{\varepsilon} & \text{if } v \in S_1, \\ 2 - \frac{\sqrt{(x-3)^2 + y^2}}{\varepsilon} & \text{if } v \in S_2. \end{cases}$$

The above formula defines a continuous function $\widetilde{s} \colon U_2 \to [0, 1]$. It is smooth away of $\partial S_1 \cup \partial S_2$. We can perturb it to a C^∞ function $s \colon U_2 \to [0, 1]$, with the following properties:

(S1) $s^{-1}(1) = U_1$, $s^{-1}(0) = \{|y| \geqslant 2\varepsilon - \varepsilon^2\} \cup \{(x - 3)^2 + y^2 \geqslant 4\varepsilon^2 - \varepsilon^3, x \geq 3\}$ (this is a thin region near the boundary of U_2).

(S2) $\frac{\partial s}{\partial u_j} = 0$ for any $j = 1, \ldots, n - 1$.

(S3) $\frac{\partial s}{\partial x} = 0$, and $|\frac{\partial s}{\partial y}| < 2/\varepsilon$ at all points of S_1. Furthermore $y \frac{\partial s}{\partial y} < 0$ at all points of S_1.

(S4) If $v \in S_2$ and we choose radial coordinates $x = 3 + r \cos \theta$, $y = r \sin \theta$ (where $r \in [\varepsilon, 2\varepsilon]$ and $\theta \in [-\pi/2, \pi/2]$), then $|\frac{\partial s}{\partial r}| < \frac{2}{\varepsilon}$ and $|\frac{\partial s}{\partial \theta}| < \varepsilon$

Observe that $\widetilde{s}$ satisfies (S1)–(S4) at every point for which it is smooth; the only issue is that on $S_1 \cap S_2$, $\widetilde{s}$ fails to be C^2.

Now let us choose a smooth decreasing function $\phi \colon [0, \eta^2] \to [0, 1]$, which is equal to 0 on $[\frac{3}{4}\eta^2, \eta^2]$ and $\phi(0) = 1$. We define now a new function $b \colon U_2 \to [0, 1]$ by the

formula

(3-7) $b(x, y, \vec{u}) = s(x, y, \vec{u}) \cdot \phi(\|\vec{u}\|^2).$

Let us finally define the function $G \colon \Omega \to [0, 1]$ by

(3-8) $G(w) = \begin{cases} F(w) & \text{if } w \notin U_2 \\ y^3 - yx^2 + y - (\delta+1)b(x, y, \vec{u})y + \frac{1}{2} + \vec{u}^2 & \text{if } w = (x, y, \vec{u}) \in U_2, \end{cases}$

where $\delta > 0$ is a very small number. Later we shall show that it is enough to take $\delta < \varepsilon^2/2$. In the following lemmas we shall prove that G satisfies the conditions of Theorem 3-1.

Lemma 3-9 *The function G is smooth.*

Proof It is a routine checking and we leave it for the reader.   □

In the next two lemmas we show that G has no critical points in U_{21}.

Lemma 3-10 *G has no critical points on $U_{21} \cap \{y = 0\}$.*

Proof If $(x, 0, u_1, \ldots, u_{n-1}) \in U_{21}$ then $x > 3$. Consider the derivative over y of G:

(3-11) $\dfrac{\partial G}{\partial y} = 3y^2 - x^2 + 1 - (\delta+1)b - (\delta+1)\phi(u_1^2 + \cdots + u_{n-1}^2)\dfrac{\partial s}{\partial y}y.$

Taking $y = 0$ we get $-x^2 + 1 - (\delta+1)b$. Since b takes values in $[0, 1]$ and $x > 3$, one gets $\frac{\partial G}{\partial y} < 0$.   □

Lemma 3-12 *If $\delta < 3\varepsilon^2$, then G has no critical points on $U_{21} \cap \{y \neq 0\}$.*

Proof Assume that $\frac{\partial G}{\partial x} = 0$ for some $(x, y, \vec{u})$. Then

$$y\left(-2x - (\delta+1)\frac{\partial s}{\partial x}\phi\right) = 0.$$

As $y \neq 0$, the expression in parentheses should be zero. If $0 < x \leq 3$, then by (S3) we have $\frac{\partial s}{\partial x} = 0$. Hence the above equality can not hold. Assume that $x = 0$. In the derivative over y (see Equation (3-11)), the expression $-(\delta+1)\phi\frac{\partial s}{\partial y}y$ is nonnegative by (S3). Furthermore $b < 1$, hence

$$\frac{\partial G}{\partial y} \geq 3y^2 - \delta.$$

Now if $\delta < 3\varepsilon^2$ then there are no critical points with $x = 0$. It remains to deal with the case $(x, y, u_1, \ldots, u_{n-1}) \in S_2$. Consider the derivative $\frac{\partial G}{\partial y}$. By (S4) and the chain rule we have

$$\left| y \frac{\partial s}{\partial y} \right| = \left| y \frac{\partial r}{\partial y} \frac{\partial s}{\partial r} + y \frac{\partial \theta}{\partial y} \frac{\partial s}{\partial \theta} \right| \leq \left| \frac{y^2}{r} \cdot \frac{\partial s}{\partial r} \right| + \left| \frac{(x-3)y}{r^2} \cdot \frac{\partial s}{\partial \theta} \right| < r \frac{2}{\varepsilon} + \varepsilon < 5.$$

Furthermore $|1 - (\delta + 1)b| \leq 1$, and $|3y^2| < 1$ because ε is small. As $x \geq 3$, we have $\frac{\partial G}{\partial y} < 0$ on S_2. $\qquad\square$

On U_1 the function G is given by

(3-13) $$G(x, y, \vec{u}) = y^3 - yx^2 - \delta y + \vec{u}^2 + \tfrac{1}{2}.$$

As in Section 3.1 we study the critical points in U_1.

Lemma 3-14 *G has two critical points on U_1 at*

$$z^s := (0, \sqrt{\delta/3}, 0, \ldots, 0),$$
$$z^u := (0, -\sqrt{\delta/3}, 0, \ldots, 0).$$

Both critical points are boundary, both of Morse index k, z^s is stable, while z^u is unstable.

Proof The derivative of G vanishes only at z^s and z^u. Indices are immediately computed from (3-13). The point z^s is boundary stable, because for z^s the expression $-yx^2$ is negative and the boundary is given by $x = 0$, hence it is attracting in the normal direction. Similarly we prove for z^u. See also Figure 12 for the two-dimensional picture. $\qquad\square$

Remark 3-15 If we define $G_t = y^3 - yx^2 + y - t(\delta + 1)b \cdot y + \tfrac{1}{2} + \vec{u}^2$ for $t \in [0, 1]$, then the same argument as in Lemmas 3-10 and 3-12 shows that G_t has no critical points in $U_2 \setminus U_1$. As for critical points in U_1, observe that on U_1 we have

$$G_t = y^3 - yx^2 + (1 - t(1 + \delta))y + \tfrac{1}{2} + \vec{u}^2.$$

Let $t_0 = 1/(1 + \delta)$. If $t > t_0$, the function G_t has two critical points on the boundary Y, while for $t < t_0$, G_t has a single critical point in the interior $U_1 \setminus Y$. If $t = t_0$, G_t has a single degenerate critical point on Y. In this way we construct an "isotopy" between F and G.

Let us now choose a Riemannian metric g' on

$$U_1' := U_1 \cap \{\|\vec{u}\| < \varepsilon\}$$

by the condition that $(x, y, u_1, \ldots, u_{n-1})$ be orthonormal coordinates (cf Remark 3-18 below). Clearly, any metric g on Ω can be changed near U_1 so as to agree with g' on U_1'. In this metric the gradient of G is

(3-16) $(-2xy, 3y^2 - x^2 - \delta, 2\epsilon_1 u_1, \ldots, 2\epsilon_{n-1} u_{n-1}).$

We want to show that there is a single trajectory starting from z^s and terminating at z^u. Clearly, there is one trajectory from z^s to z^u which stays in U_1' (having $y = 0$ and $\vec{u} = 0$). In order to eliminate the others, we need the following lemma.

Lemma 3-17 *Let γ be a trajectory of ∇G starting from z^s. Let w be the point, where γ hits $\partial U_1'$ for the first time. If δ is sufficiently small, then $G(w) > G(z^u)$.*

Proof Assume that $\gamma(t)$ is such trajectory. Assume that among numbers ϵ_i, we have $\epsilon_i = -1$ for $i \leqslant k - 1$ and $\epsilon_i = 1$ otherwise. As z^s is a critical point of the vector field ∇G with a nondegenerate linear part, we conclude that the limit

$$\lim_{t \to -\infty} \frac{\gamma'(t)}{\|\gamma'(t)\|} =: v = (x_0, y_0, u_{01}, \ldots, u_{0,n-1})$$

exists. The vector v is the tangent vector to the curve γ at the point z^s, and it lies in the unstable space. Hence $x_0 = 0$ as $(1, 0, \ldots, 0)$ is a stable direction; similarly $u_{01} = \cdots = u_{0,k-1} = 0$. Therefore, until γ hits the boundary of U_1' for the first time, we have

$$x = u_1 = \cdots = u_{k-1} = 0.$$

Set also $g(y) = y^3 - \delta y$. One has the following cases, depending the position of w, where γ hits $\partial U_1'$ for the first time: (a) $y = -\varepsilon$, (b) $y = \varepsilon$, or (c) $\|\vec{u}\|^2 = \eta^2$. The case (a) cannot happen since G is increasing along the trajectory, hence $G(w) > G(z^s)$, a fact which contradicts $g(-\varepsilon) < g(\sqrt{\delta/3})$ valid for $2\delta < \varepsilon^2$. In case (b), $G(w) > G(z^u)$ follows from $g(\varepsilon) > g(-\sqrt{\delta/3})$. Finally, assume the case (c). Then, as $u_{01} = \cdots = u_{0,k-1} = 0$, we obtain $\vec{u}^2 = \|\vec{u}\|^2 = \eta^2$. Then $G(w) - G(z^s) \geqslant \eta^2$, because the contribution to G from $y^3 - \delta y$ increases along γ. Hence $G(w) > G(z^u)$ follows again since $\varepsilon \ll \eta$. $\square$

Given the above lemma it is clear that if a trajectory γ leaves U_1', then G becomes bigger than $G(z^u)$. As G increases along any trajectory, it is impossible that such trajectory limits in z^u. The proof of Theorem 3-1, up to Proposition 3-5, is accomplished.

Remark 3-18 The metric g' defined below Remark 3-15 can be chosen so that $(x, y, \vec{u})$ forms an *orthogonal*, but not necessarily *orthonormal* coordinate system. Each component of the vector field (3-16) is then multiplied by a positive constant, the statement of Lemma 3-17 still holds with essentially the same proof. However, g' cannot be just any metric; we can choose a metric g' in a way that there is an arbitrary number of trajectories from z^s to z^u (topologically changing the metric can produce a pair of mutually canceling intersection points between the unstable manifold of z^s and the stable manifold of z^u).

3.3 An auxiliary construction.

The following construction is a crucial ingredient in the proof of Proposition 3-5; see the next section. Set
$$Z = \{(x, y) \in \mathbb{R}^2 : x \geqslant 0\},$$

and define the two functions

$$(3\text{-}19) \qquad A(x, y) = \frac{x^3}{3\sqrt{3}} - \sqrt{3}xy^2 - \frac{x}{\sqrt{3}} + \frac{2}{3\sqrt{3}}, \quad B(x, y) = y^3 - yx^2 + y.$$

Observe that

$$A + iB = \left(\frac{x}{\sqrt{3}} - iy\right)^3 - \left(\frac{x}{\sqrt{3}} - iy\right) + \frac{2}{3\sqrt{3}}.$$

Up to a linear transformation, the map $(x, y) \mapsto A + iB$ is a holomorphic map. Thus it shares several geometric properties of a holomorphic map. For example, it is an open map, and the singular points are precisely the points where the gradient of B vanishes.

Let us choose $\delta > 0$ smaller than $2/(3\sqrt{3})$. Consider two sets

$$(3\text{-}20) \qquad \begin{aligned} Z_1 &= \{(x, y) \in Z, \ x < 1, \ A(x, y) \geq \delta\}, \\ Z_2 &= \{(x, y) \in Z, \ x > 1, \ A(x, y) \geq \delta\}. \end{aligned}$$

We have the following result.

Lemma 3-21 *The map $\psi(x, y) = (A(x, y), B(x, y))$ maps Z_1 and Z_2 diffeomorphically onto E_1 and E_2 respectively, where*

$$E_1 = \left\{(a, b) \in \mathbb{R}^2 : a \in \left[\delta, \frac{2}{3\sqrt{3}}\right]\right\}, \quad E_2 = \{(a, b) \in \mathbb{R}^2 : a \geq \delta\}.$$

Proof One readily checks that $\psi \colon Z_1 \to V_1$ and $\psi \colon Z_2 \to V_2$ are bijections. As the derivative $D\psi$ is nondegenerate on $Z_1 \cup Z_2$, ψ is a diffeomorphism between the two pairs of sets. □
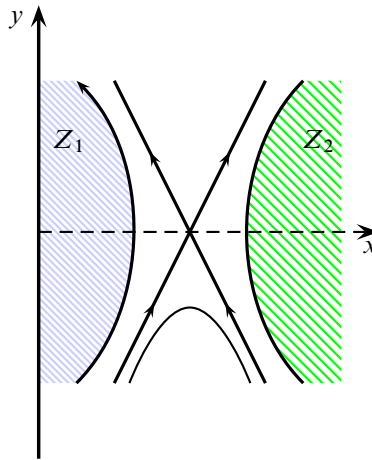
Figure 14: Sets Z_1 and Z_2 from Section 3.3. There is also drawn the singular level set $A^{-1}(0)$.

3.4 Proof of Proposition 3-5

First, as z is a critical point of index $k \in \{1, \dots, n\}$, by the Morse Lemma 2-6 we can find a neighborhood $\widetilde{V}$ of z and a chart $h_1 \colon \widetilde{V} \to \mathbb{R}^{n+1}$, with coordinates $(x', y, \vec{u})$ such that

$$F \circ h_1^{-1}(x', y, \vec{u}) = x'y + \vec{u}^2 + \tfrac{1}{2}.$$

Remark 3-22 The term $x'y$ (corresponding to a hyperbolic quadratic form) is the moment when the assumption that $k \neq 0, n+1$ is used.

Let us define a map $h_2(x, y, \vec{u}) = (x', y, \vec{u})$, where $x' = y^2 + 1 - x^2$. By the inverse function theorem, h_2 is a local diffeomorphism near $(1, 0, \dots, 0)$. Shrinking $\widetilde{V}$ if needed, and considering $h_3 = h_2^{-1} \circ h_1$, we obtain $h_3(z) = (1, 0, \dots, 0)$ and

(3-23) $\qquad F \circ h_3^{-1}(x, y, \vec{u}) = y^3 - yx^2 + y + \vec{u}^2 + \tfrac{1}{2} = B(x, y) + \vec{u}^2 + \tfrac{1}{2}.$

Let us pick now $\xi > 0$ such that the cylinder

$$V = \{|x - 1| < \xi, \ |y| < \xi, \ \|\vec{u}\| < \xi\} \subset \mathbb{R}^{n+1}$$

lies entirely in $h_3(\widetilde{V})$. By shrinking $\widetilde{V}$ we may in fact assume that $h_3(\widetilde{V}) = V$. If $0 < \delta \ll 2/(3\sqrt{3})$ is sufficiently small then $A(x, 0) < \delta$ implies $|x - 1| < \xi$. Choose such a δ, and set

$$V_1 := V \cap \{x < 1, \ A(x, y) \geq \delta\}$$

(compare (3-20)). By Lemma 3-21 the map

(3-24) $$\Psi_1(x, y, \vec{u}) = (A(x, y), B(x, y) + \vec{u}^2, \vec{u}),$$

is a diffeomorphism (being the composition of $\psi \oplus \mathrm{Id}_{\mathbb{R}^{n+1}}$ and a "triangular" map). Set $C_1 := \Psi_1(V_1)$ and $\tilde{V}_1 := h_3^{-1}(V_1)$. Finally, let

$$h = \Psi_1 \circ h_3.$$

Using (3-23) we obtain that

$$F \circ h^{-1}(a, b, \vec{u}) = b + \tfrac{1}{2}.$$

Let $\theta > \delta$ be sufficiently close to δ satisfying the inclusion

$$D_1 := [\delta, \theta] \times (-\theta, \theta) \times (-\theta, \theta)^{n-1} \subset C_1.$$

Let $\tilde{D}_1 = h^{-1}(D_1) \subset \tilde{V}_1$; see Figure 15.

Lemma 3-25 *If θ and δ are small enough, there is an closed ball $\tilde{W}$ in Ω, containing $\tilde{D}_1$, such that h extends to a diffeomorphism between $\tilde{W}$ and $[\delta, 2/(3\sqrt{3})] \times [-\theta, \theta] \times [-\theta, \theta]^{n-1}$ with $F \circ h^{-1}(a, b, \vec{u}) = b + \tfrac{1}{2}$, sending points with $a = 2/(3\sqrt{3})$ to Y.*

In the proof we shall use the following result.

Lemma 3-26 *There exists a smooth curve $\gamma \colon [\delta, 2/(3\sqrt{3})] \to \Omega$, such that:*

- *$\gamma(2/(3\sqrt{3})) \in Y$.*
- *$\gamma(t) \in \Sigma_{1/2}$.*
- *$\gamma(t) \in \tilde{D}_1$ if and only if $t \in [\delta, \theta]$.*
- *$h(\gamma(t)) = (t, 0, \dots, 0)$.*
- *γ omits $\tilde{V} \setminus \tilde{V}_1$.*
- *γ is transverse to Y.*

Proof of Lemma 3-26 Let $p = h^{-1}(\theta, 0, \dots, 0) \in \Sigma_{1/2}$. Let $B \subset \Sigma_{1/2}$ be an open ball with center z and $p \in \partial B$. Let Σ' be the connected component of $\Sigma_{1/2}$ containing p. We consider two cases.

Case 1 If $\Sigma' \setminus B$ is connected, it is also path connected. By (3-2), there exists a path $\tilde{\gamma} \subset \Sigma' \setminus B$ joining p with a point on the boundary. We can assume that $\tilde{\gamma}$ is transverse to Y. We choose $\gamma = h^{-1}([\delta, \theta] \times \{0, \dots, 0\}) \cup \tilde{\gamma}$ (and we smooth a possible corner at p). It is clear that γ omits $\tilde{V} \setminus \tilde{V}_1$ and that we can find a parametrization of γ by the interval $[\delta, 2/(3\sqrt{3})]$.
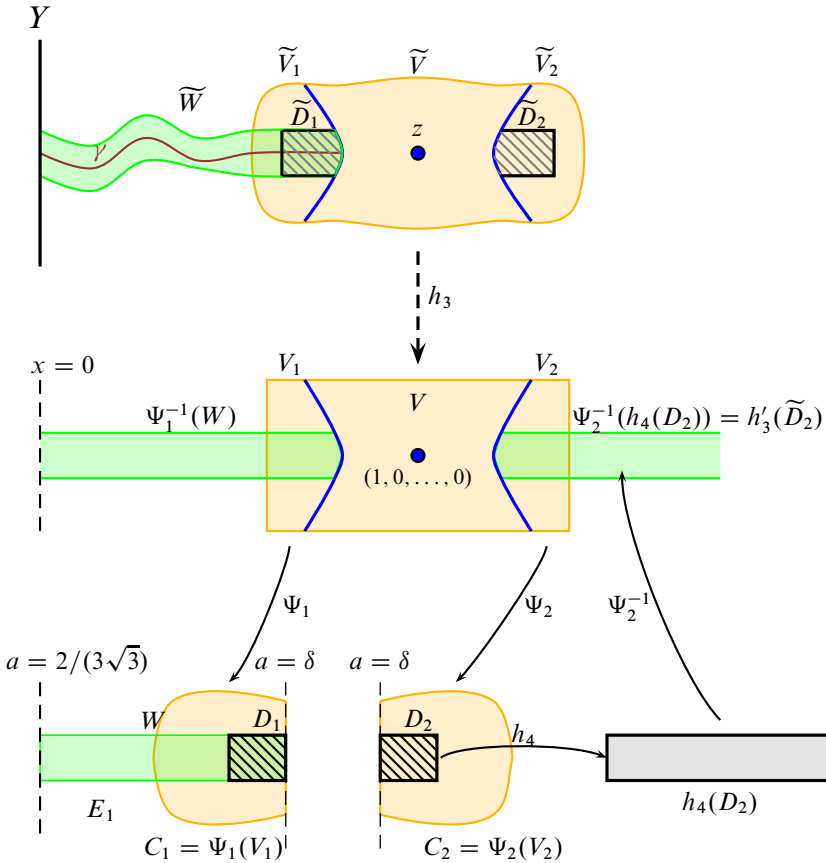
Figure 15: Notation used in Section 3.4. The top line is the picture on Ω, the middle line is in coordinates such that F is equal to $y^3 - yx^2 + y + \frac{1}{2} + \vec{u}^2$. The bottom line is in coordinates such that $F = b + \frac{1}{2}$. There is no mistake, the line $a = \delta$ appears twice on the picture, in coordinates on C_1 and on C_2.

Case 2 If $\Sigma' \setminus B$ is not connected, then as Σ' is connected, by a homological argument we have $n = 1$ and $k = 1$. Since Σ' is connected and has boundary, then Σ' is an interval and B is an interval too. Then $\Sigma' \setminus B$ consists of two intervals, each intersecting Y. One of these intervals contains p. So p is connected to Y by an interval, which omits B. We conclude the proof by the same argument as in the above case, when $\Sigma' \setminus B$ was connected. $\qquad\square$

Proof of Lemma 3-25 Given Lemma 3-26, let us choose a tubular neighborhood X of γ in $F^{-1}(\frac{1}{2}) \setminus (\widetilde{V} \setminus \widetilde{V}_1)$. Shrinking X if needed we can assume that it is a disk and $X_1 := X \cap \widetilde{V} = \widetilde{D}_1 \cap F^{-1}(\frac{1}{2})$. Now let τ be the vector field on $\widetilde{D}_1$ given
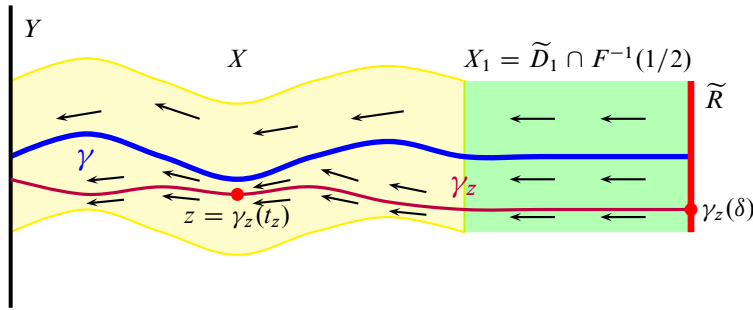
Figure 16: Proof of Lemma 3-25. Construction of the vector field τ. Picture on $F^{-1}(\frac{1}{2})$. The parallel vector field from the region on the right is extended to the whole X so that it is tangent to γ.

by $(Dh)^{-1}(1, 0, \ldots, 0)$, where Dh denotes the derivative of h. This vector field is everywhere tangent to X_1 and

$$(3\text{-}27) \qquad \tau|_{\gamma \cap \widetilde{D}_1} = \frac{d}{dt}\gamma(t)$$

by definition of γ. We extend τ to a smooth vector field on the of whole X, such that (3-27) holds on the whole of γ. For any point $z \in \gamma$, the trajectory of τ (which is γ) eventually hits Y and, on the other end, it hits the "right wall"

$$\widetilde{R} = h^{-1}(\{\delta\} \times \{0\} \times (-\theta, \theta)^{n-1}).$$

(compare Figure 16; note that the horizontal coordinate there increases from right to left for consistency with Figure 15). Since γ is transverse to $\widetilde{R}$ and to Y, by the implicit function theorem trajectories close to τ also start at $\widetilde{R}$ and end up at Y. Shrinking X if necessary we may assume that each point of X lies on the trajectory of τ which connects a point of $\widetilde{R}$ to some point of Y, and all the trajectories are transverse to both Y and $\widetilde{R}$.

We can now rescale τ (that is multiply by a suitable smooth function constant on trajectories) so that all the trajectories go from $\widetilde{R}$ to Y in time $2/(3\sqrt{3}) - \delta$, ie the same time as γ does. The rescaled vector field allows us to introduce coordinates on X in the following way. For any point $z \in X$, let γ_z be the trajectory of τ, going through z. We can assume that $\gamma_z(\delta) \in \widetilde{R}$. Let $t_z = \gamma_z^{-1}(z)$, ie the moment when γ_z passes through z. Since we normalized γ_z, we know that $t_z \in [\delta, 2/(3\sqrt{3})]$ and $t_z = 2/(3\sqrt{3})$ if and only if $z \in Y \cap X$.

Let $\vec{u}_z$ be such that $h(\gamma_z(\delta)) = (\delta, 0, \vec{u}_z)$. The vector $\vec{u}_z$ might be thought of as a coordinate on $\widetilde{R}$. We define now

$$h(z) = (t_z, 0, \vec{u}_z).$$

This maps clearly extends h to the whole of X.

Now let $\widetilde{W}$ be a tubular neighborhood of X in $\Omega \setminus (\widetilde{V} \setminus \widetilde{V}_1)$. We use the flow of ∇F to extend coordinates from X to $\widetilde{W}$. More precisely, shrinking $\widetilde{W}$ if needed we may assume that for each $w \in \widetilde{W}$ the trajectory of ∇F intersects X. This intersection is necessarily transverse and it is in one point, which we denote by $z_w \in X$. We define now

$$h(w) = (t_{z_w}, F(w) - F(z_w), \vec{u}_{z_w}).$$

As h is a local diffeomorphism on X (because ∇F is transverse to X), it is also a local diffeomorphism near X. We put $W = h(\widetilde{W})$. Clearly both definitions of h on $\widetilde{V}$ and $\widetilde{W}$ agree. We may now decrease θ and shrink W so that

$$W = [\delta, 2/(3\sqrt{3})] \times (-\theta, \theta) \times (-\theta, \theta)^{n-1}.$$

We have $F \circ h^{-1}(a, b, \vec{u}) = b + \frac{1}{2}$. We now extend h_3 over $\widetilde{W}$ by the formula $h_3 = \Psi_1^{-1} \circ h$. $\qquad\square$

Consider now

$$V_2 := V \cap \{x > 1, \, A(x, y) \geq \delta\}.$$

Let $\Psi_2 \colon V \to \mathbb{R}^{n+1}$ be given by $\Psi_2(x, y, \vec{u}) = (a, b, \vec{u}) = (A(x, y), B(x, y) + \vec{u}^2, \vec{u})$, provided by the same formula as Ψ_1 in (3-24) but the image now satisfies $a \geq \delta$, cf Lemma 3-21.

Let $C_2 = \Psi_2(V_2)$, and let us choose θ' sufficiently small such that

$$D_2 := [\delta, \theta'] \times (-\theta', \theta') \times (-\theta', \theta')^{n-1} \subset C_2.$$

We shall denote $h = \Psi_2 \circ h_3$ and $\widetilde{D}_2 = h^{-1}(D_2)$.

Let us now fix $M > 0$ large enough and consider a map $h_4 \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ of the form

$$h_4(a, b, \vec{u}) = (\phi(a), b, \vec{u}),$$

where $\phi \colon [\delta, \theta'] \cong [\delta, M]$ is a strictly increasing smooth function, which is an identity near δ. Consider the map $h_3' \colon \Psi_2^{-1} \circ h_4 \circ h \colon \widetilde{D}_2 \to \mathbb{R}^{n+1}$. Since h is an identity for a close to δ, this map agrees with h_3 for a close to δ. Furthermore $F \circ h_4^{-1}(a, b, \vec{u}) = F \circ h^{-1} \circ h_4^{-1}(a, b, \vec{u}) = b + \frac{1}{2}$ by a straightforward computation. On the other hand,

the point $h^{-1}(\theta', 0, \ldots, 0) \in \widetilde{D}_2$ is mapped by h'_3 to $(M, 0, \ldots, 0) \in \mathbb{R}^{n+1}$, where M can be arbitrary large, eg $M > 3$.

Having gathered all the necessary maps, we now conclude the proof. Let

$$\widetilde{U} = \widetilde{W} \cup (\widetilde{V} \setminus h_3^{-1}(V_1 \cup V_2) \cup \widetilde{D}_2.$$

The map $h_3 \colon \widetilde{U} \to [0, \infty) \times \mathbb{R}^n$ is given by h_3 on $\widetilde{W}$ and on $\widetilde{V} \setminus h_3^{-1}(V_2)$, and by h'_3 on $\widetilde{D}_2$. This map is a diffeomorphism onto its image, so it is a chart near z. By construction $F \circ h_3^{-1}$ is equal to $y^3 - yx^2 + y + \vec{u}^2 + \frac{1}{2}$ and $h_3(\widetilde{W})$ contains the segment with endpoints $(0, 0, \ldots, 0)$ and $(3, 0, \ldots, 0)$. Since it is an open subset, it contains $[0, 3 + \eta) \times (-\eta, \eta) \times (-\eta, \eta)^{n-1}$ for $\eta > 0$ small enough. The inverse image of this cube gives the required chart.

This ends the proof of Theorem 3-1 which moves a single interior critical point to the boundary. Section 4 generalizes this fact for multiple points; one of the needed tools will be the rearrangements of the critical values/points.

3.5 Condition (3-2) revisited

We will provide two sufficient conditions which imply Condition (3-2). One is valid for arbitrary $n \geqslant 1$, the other one holds only in the case $n = 1$. We shall keep the notation from previous subsections, in particular (Ω, Y) is a cobordism between (Σ_0, M_0) and (Σ_1, M_1), $F \colon \Omega \to [0, 1]$ is a Morse function with a single critical point z in the interior of Ω, and $F(z) = \frac{1}{2}$. Let $\Sigma_{1/2} = F^{-1}(\frac{1}{2})$ and Σ' be the connected component of $\Sigma_{1/2}$ such that $z \in \Sigma'$.

Proposition 3-28 *If Σ_0, Σ_1 and Ω have no closed connected components, then $\Sigma' \cap Y \neq \varnothing$. In particular, in Theorem 3-1 we can assume that Σ_0, Σ_1 and Ω have no closed connected components instead of (3-2).*

Proof Let $p = h^{-1}(\theta, 0, \ldots, 0) \in \widetilde{D}_1 \subset \Omega$ and let B be an open ball in Σ' near z, such that $p \in \partial B$. It is enough to show that p can be connected to Y by a path in $\Sigma_{1/2}$, which misses B (compare Lemma 3-26).

Let us choose a Riemannian metric on Ω. Let W_z^s be the stable manifold of z and let T be the intersection of W_z^s and Σ_0. This is a $(k-1)$–dimensional sphere. The flow of ∇F induces a diffeomorphism $\Phi \colon \Sigma_{1/2} \setminus B \cong \Sigma_0 \setminus B_0$, where B_0 is a tubular neighborhood of T in Σ_0 (here we tacitly use the fact that δ and θ are small enough); see Figure 17. Let $p_0 = \Phi(p)$. Let Σ'_0 be the connected component of Σ_0 which contains B_0.

Now we will analyze several cases. Recall that $k = \mathrm{ind}_z F \in \{1, \dots, n\}$. First we assume that $k < n$. Then $\Sigma_0' \setminus T$ is connected, so p_0 can be connected to the boundary of Σ_0' — which is nonempty by the assumptions of the proposition — by a path γ_0. Now the inverse image $\Phi^{-1}(\gamma_0)$ is the required path.
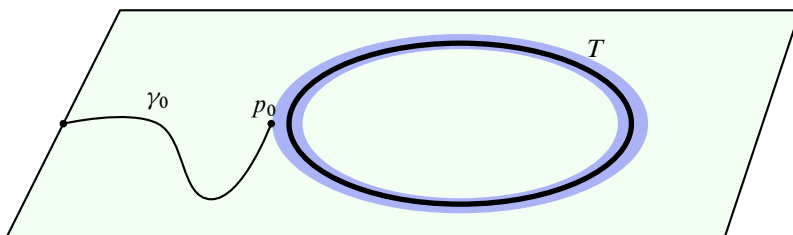

Figure 17: Notation on Σ_0.

If $k = n > 1$ then we reverse the cobordism and look at $-F$, hence this case is covered by the previous one (since $k = n$ will be replaced by $k = 1 < n$).

Finally, it remains to deal with the situation $k = n = 1$. Then $\dim \Sigma_0 = 1$. T consists of two points. Assume first that they lie in a single connected component Σ_0' of Σ_0. We shall show that this is impossible. As Σ_0' is connected with nontrivial boundary, it is an interval. The situation is like on Figure 18. Now as F has precisely one Morse critical point of index 1, Σ_1 is the result of a surgery on Σ_0. This surgery consists of removing two inner segments from Σ_0 and gluing back two other segments, which in Figure 18 are drawn as dashed arc. But then Σ_1 has a closed connected component, which contradicts assumptions of Theorem 3-1.

Therefore, T lies in two connected components of Σ_0. The situation is drawn in Figure 19, and it is straightforward to see that p_0 (either p_0' or p_0'' in Figure 19) can be connected to M_0 by a segment omitting B_0.   □

The proof of Proposition 3-28 suggests that the case $n = 1$ is different from case $n > 1$. We shall provide now a full characterization of the failure to (3-2).

Proposition 3-29 *Assume that $k = n = 1$ and Ω is connected. If (3-2) does not hold, then Ω is a pair of pants, Σ_0 is a circle and Σ_1 is a disjoint union of two circles; or vice versa: $\Sigma_1 = S^1$ and Σ_0 is a disjoint union of two circles. In particular, $Y = \varnothing$.*

Proof A one-handle attached to a surface changes the number of boundary components by ± 1. Let us assume that Σ_1 has fewer components than Σ_0, if not we can reverse the cobordism. As Ω is connected, Σ_0 has two components and Σ_1 only one. Let $A_0 \subset \Sigma_0$ be the attaching region, ie the union of two closed intervals to which the one-handle is
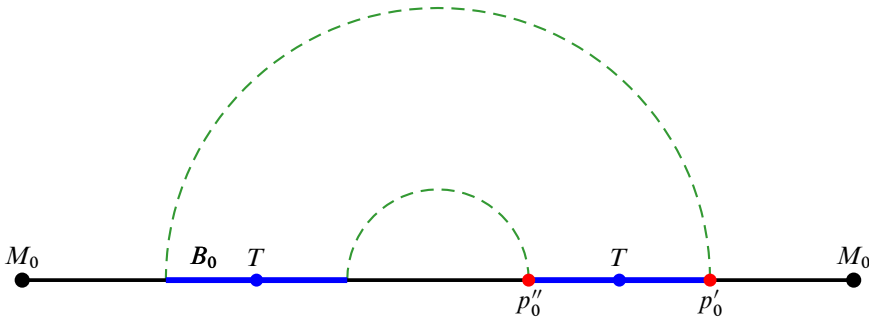
Figure 18: Proof of Proposition 3-28. Case $k = 1$ and $n = 1$ and T lies in two components of Σ. Σ_0 is the horizontal segment. The points p_0' and p_0'' are the two possible positions of the point p_0.
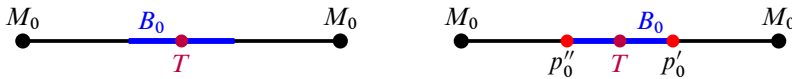

Figure 19: Proof of Proposition 3-28. Case $k = 1$ and $n = 1$ and T lies in two components of Σ_0. The points p_0' and p_0'' are the two possible positions of the point p_0. Both can be connected to the boundary M_0.

attached. With the notation of Section 3.5 we have $(\Sigma', z) \cong (\Sigma_0/A_0, A_0/A_0)$, where the quotient denotes collapsing a space to a point. In particular z cannot be joined to Y by a path in Σ' if and only if Σ_0 is disjoint from Y. Hence Σ_0 is closed, that is, it is a union of two circles. □

4 Rearrangements of boundary handles

4.1 Preliminaries

Let (Ω, Y) be a cobordism between two n–dimensional manifolds with boundary (Σ_0, M_0) and (Σ_1, M_1). Let F be a Morse function, with critical points $w_1, \ldots, w_k \in \operatorname{Int} \Omega$ and $y_1, \ldots, y_l \in Y$. In the classical theory (that is, when $Y = \varnothing$), the Thom–Milnor–Smale theorem (see [12, Section 4]) says that we can alter F without introducing new critical points such that if $\operatorname{ind} w_i < \operatorname{ind} w_j$, then $F(w_i) < F(w_j)$ as well. We want to prove similar results in our more general case.

In this section we rely very strongly on [12, Section 4].

4.2 Elementary rearrangement theorems

We shall begin with the case $k + l = 2$, ie F has two critical points. For a critical point p we shall denote by K_p the union $W_p^s \cup \{p\} \cup W_p^u$, ie the set of all points $x \in \Omega$, such that the trajectory $\phi_t(x)$ ($t \in \mathbb{R}$), of the gradient vector field ∇F contains p in its limit set. Elementary rearrangement theorems deal with the case when the two sets K_{p_1} and K_{p_2} for the two critical points are disjoint.

Proposition 4-1 (Rearrangement of critical points) *Let p_1 and p_2 be two critical points, and assume that $K_1 := K_{p_1}$ and $K_2 := K_{p_2}$ are disjoint. Let us choose $a_1, a_2 \in (0, 1)$. Then, there exist a Morse function $G \colon \Omega \to [0, 1]$, with critical points exactly at p_1 and p_2, such that $G(p_i) = a_i$, $i = 1, 2$; furthermore, near p_1 and p_2, the difference $F - G$ is a locally constant function.*

Remark 4-2 If both critical points are on the boundary, in order to guarantee the above existence, we need even to change the Riemannian metric away from K_1 and K_2.

Proof Similarly to [12, Section 4] we will use an auxiliary result. Its proof is postponed after the end of proof of Proposition 4-1.

Lemma 4-3 *There exists a smooth function $\mu \colon \Omega \to [0, 1]$ with the following properties:*

(M1) $\mu \equiv 0$ *in a neighborhood of K_1.*

(M2) $\mu \equiv 1$ *in a neighborhood of K_2.*

(M3) μ *is constant on trajectories of ∇F.*

Furthermore, if at least one of the critical points is interior, we have

(M4) μ *is constant on Y.*

We continue with the proof of Proposition 4-1. We choose a smooth function $\Psi \colon [0, 1] \times [0, 1] \to [0, 1]$ with:

(PS1) $\frac{\partial \Psi}{\partial x}(x, y) > 0$ for all $(x, y) \in [0, 1] \times [0, 1]$.

(PS2) There exists $\delta > 0$ such that $\Psi(x, y) = x$ for all $x \in [0, \delta] \cup [1 - \delta, 1]$ and $y \in [0, 1]$.

(PS3) For any $s \in (-\delta, \delta)$ we have $\Psi(F(p_1) + s, 0) = a_1 + s$ and $\Psi(F(p_2) + s, 1) = a_2 + s$.

For any $\eta \in \Omega$ we define $G(\eta) = \Psi(F(\eta), \mu(\eta))$. From the properties (PS3), (M1) and (M2) we see that near p_i, G differs from F by a constant. The property (PS2) ensures that G agrees with F in a neighborhood of Σ_0 and Σ_1. Let us show that ∇G does not vanish away from p_i. By the chain rule we have

$$(4\text{-}4) \qquad\qquad \nabla G = \frac{\partial \Psi}{\partial x} \nabla F + \frac{\partial \Psi}{\partial y} \nabla \mu.$$

Since μ is constant on all trajectories of ∇F, the scalar product $\langle \nabla F, \nabla \mu \rangle = 0$. Then the property (PS1) guarantees that $\langle \nabla G, \nabla F \rangle > 0$ away from p_1 and p_2.

We need to show that ∇G is everywhere tangent to Y. If one of the points is interior, by (M4) $\nabla \mu$ vanishes on Y, hence ∇G is parallel on Y to ∇F and we are done. Next assume that both critical points are on the boundary. Let us choose an open subset U of Y such that $\nabla \mu|_U = 0$ and $K_1 \cup K_2 \subset U$. This is possible, because of the properties (M1) and (M2). Then let us choose a neighborhood W in Ω of $Y \setminus U$, disjoint from K_1 and K_2. Observe that $dG(\nabla F) = \langle \nabla G, \nabla F \rangle > 0$. As $\nabla F \in TY$ one has $TY \not\subset \ker dG$, so by Lemma 1-7 we can change the metric in W so that ∇G is everywhere tangent to Y. $\qquad\square$

Proof of Lemma 4-3 Let us define $T_1 = K_1 \cap \Sigma_0$ and $T_2 = K_2 \cap \Sigma_0$. Assume that T_1 and T_2 are not empty. For each $\eta \in \Omega \setminus K_1 \cup K_2$, let $\pi(\eta)$ be the intersection of the trajectory of η under ∇F with Σ_0. This gives a map $\pi \colon \Omega \setminus (K_1 \cup K_2) \to \Sigma_0 \setminus (T_1 \cup T_2)$.

Let us define μ first on Σ_0 by the following conditions: $\mu \equiv 1$ in a neighborhood of T_2, $\mu \equiv 0$ in a neighborhood of T_1. Furthermore, if either T_1 or T_2 is disjoint from the boundary M_0 we extend μ to a constant function on M_0. Finally, we extend μ to the whole Ω by picking $\mu(\eta) = \mu(\pi(\eta))$ if $\eta \notin K_1 \cup K_2$, and $\mu|_{K_i}(\eta) = i - 1$, $i = 1, 2$.

If $T_1 = \varnothing$, then $\mathrm{ind}_F\, p_1 = 0$ and the proof of the rearrangement theorem is completely straightforward. $\qquad\square$

4.3 Morse–Smale condition on manifolds with boundary

In the classical theory, the Morse–Smale condition imposed on a Morse function $F \colon M \to \mathbb{R}$ means that for each pair of two critical points p_1, p_2 of M the intersection of stable manifold $W^s_{p_1}$ with the unstable manifold of $W^u_{p_2}$ is transverse. (Note that this Morse–Smale condition also depends on the choice of Riemannian metric on M.) Following [9, Definition 2.4.2], we reformulate the Morse–Smale condition in the following way.

Definition 4-5 The function F is called *Morse–Smale* if for any two critical points p_1 and p_2, the intersection of $\operatorname{Int}\Omega \cap W^s_{p_1}$ with $\operatorname{Int}\Omega \cap W^u_{p_2}$ is transverse (as the intersection in the $(n+1)$–dimensional manifold Ω) and the intersection of $Y \cap W^s_{p_1}$ with $Y \cap W^u_{p_2}$ is transverse (as an intersection in the n–dimensional manifold Y).

The Morse–Smale functions form an open-dense subset of all C^2 smooth functions satisfying the condition (1-8). The proof is the same as in the case of the Morse functions on manifolds without boundary; see for example [13, Theorem 2.27].

Assume now that F is Morse–Smale. Given two critical points of F, p_1 and p_2, we want to check whether $W^s_{p_1} \cap W^u_{p_2} = \varnothing$. This depends not only on the indices, but also on whether either of the two points are boundary stable. We show this in a tabulated form in Table 1, where $\operatorname{ind} p_1 = k$ and $\operatorname{ind} p_2 = l$. In studying the intersection, we remark that $W^s_{p_1} \cap W^u_{p_2}$ is formed from trajectories, so if for dimensional reasons we have $\dim W^s_{p_1} \cap W^u_{p_2} < 1$, it immediately follows that this intersection is empty.

type of p_1	type of p_2	$\dim_\Omega W^s_{p_1}$	$\dim_Y W^s_{p_1}$	$\dim_\Omega W^u_{p_2}$	$\dim_Y W^u_{p_2}$	empty if
interior	interior	k	$\varnothing$	$n+1-l$	$\varnothing$	$k \leqslant l$
interior	b. stable	k	$\varnothing$	$\varnothing$	$n+1-l$	always
interior	b. unstable	k	$\varnothing$	$n+1-l$	$n-l$	$k \leqslant l$
b. stable	interior	k	$k-1$	$n+1-l$	$\varnothing$	$k \leqslant l$
b. stable	b. stable	k	$k-1$	$\varnothing$	$n-l$	$k \leqslant l$
b. stable	b. unstable	k	$k-1$	$n+1-l$	$n-l$	$k \leqslant l$
b. unstable	interior	$\varnothing$	k	$n+1-l$	$\varnothing$	always
b. unstable	b. stable	$\varnothing$	k	$\varnothing$	$n+1-l$	$\mathbf{k < l}$
b. unstable	b. unstable	$\varnothing$	k	$n+1-l$	$n-l$	$k \leqslant l.$

Table 1: The last column shows whether there might exist trajectories from z_2 to z_1 under the Morse–Smale condition. We write $\dim_\Omega W^s = \dim(W^s \cap \operatorname{Int}\Omega)$ and $\dim_Y W^s = \dim(W^s \cap Y)$. $\varnothing$ means the manifold is empty.

4.4 Global rearrangement theorem

Let us combine the rearrangement theorems from Section 4.2 with the computations in Table 1.

Proposition 4-6 *Let F be a Morse function on a cobordism (Ω, Y) between (Σ_0, M_0) and (Σ_1, M_1). Let $w_1, \dots, w_m$ be the interior critical points of F and let $y_1, \dots, y_{k+l}$ be the boundary critical points, where the points $y_1, \dots, y_k$ are boundary stable and*

$y_{k+1}, \ldots, y_{k+l}$ are boundary unstable. Let us choose real numbers $0 < c_0 < c_1 < \cdots < c_{n+1} < 1$, $0 < c_1^s < \cdots < c_{n+1}^s < 1$, $0 < c_0^u < \cdots < c_n^u < 1$ satisfying

$$
\begin{aligned}
c_{i-1}^s &< c_i < c_{i+1}^s, \\
c_{i-1}^u &< c_i < c_{i+1}^u, \\
c_{i-1}^u &< c_i^s < c_i^u
\end{aligned}
$$

(4-7)

for all $i \in \{0, \ldots, n+1\}$ (we can assume that $c_{-1} = c_0^s = c_{-1}^u = 0$, $c_{n+2} = c_{n+2}^s = c_{n+1}^u = 1$ so that (4-7) makes sense for all i).

Then there exists another Morse function G on the cobordism (Ω, Y) with critical points $w_1, \ldots, w_m$ in the interior, $y_1, \ldots, y_{k+l}$ on the boundary, such that if ind $w_j = l$, then $G(w_j) = c_l$, if ind $y_i = l$ and y_i is boundary stable then $G(y_i) = c_l^s$, and if ind $y_i = l$ and y_i is boundary unstable, then $G(y_i) = c_l^u$.

Proof Given the elementary rearrangement result (Proposition 4-1), the proof is completely standard (see the proof of Theorem 4.8 in [12]). Note only that we need to have $c_i^s < c_i^u$ in the statement, because there might be a trajectory from a boundary stable critical point to a boundary unstable of the same index. However, we are free to choose $c_i < c_i^s$ or $c_i \in (c_i^s, c_i^u)$ or $c_i > c_i^u$. □

4.5 Moving more handles to the boundary at once

Before we formulate Theorem 4-10, let us introduce the following technical notion.

Definition 4-8 The Morse function F on the cobordism (Ω, Y) is called *technically good* if it has the following properties.

(TG1) If p_1, p_2 are (interior or boundary) critical points of F then ind $p_1 <$ ind p_2 implies $F(p_1) < F(p_2)$.

(TG2) There exist regular values of F, say $c, d \in [0, 1]$, with $c < d$ such that $F^{-1}[0, c]$ contains those and only those critical points which have index 0 or which are boundary stable critical points of index 1, and $F^{-1}[d, 1]$ contains those and only those critical points which have index $n + 1$ and boundary unstable critical points of index n.

(TG3) There are no pairs of 0 and 1 (interior) handles of F that can be canceled (in the sense of Section 5).

(TG4) There are no pairs of n and $n+1$ (interior) handles of F that can be canceled.

Lemma 4-9 *Each function F can be made technically good without introducing new critical points.*

Proof   By Proposition 4-6 we can rearrange the critical points of F, proving (TG1) and (TG2). The properties (TG3) and (TG4) can be guaranteed, using the handle cancellation theorem (eg [12, Theorem 5.4]). We refer to the beginning of Section 5 for an explanation that one can use the handle cancellation theorem if the manifold in question has boundary.   □

Theorem 4-10   *Let (Ω, Y) be a cobordism between (Σ_0, M_0) and (Σ_1, M_1). Let F be a technically good Morse function on that cobordism, which has critical points $y_1, \ldots, y_k$ on the boundary Y, points $z_1, \ldots, z_{l+m}$ in the interior $\mathrm{Int}\,\Omega$, of which $z_{m+1}, \ldots, z_{l+m}$ have index 0 or $n+1$ and the indices of $z_1, \ldots, z_m$ are in $\{1, \ldots, n\}$. Suppose furthermore the following properties are satisfied:*

(I1)   *Σ_0 and Σ_1 have no closed connected components.*

(I2)   *Ω has no closed connected components.*

Then there exists a Morse function $G\colon \Omega \to [0, 1]$, on the cobordism (Ω, Y), with critical points $y_1, \ldots, y_k \in Y$, $z_{m+1}, \ldots, z_{l+m}$ and $z_1^s, z_1^u, \ldots, z_m^s, z_m^u$ such that:

- *$\mathrm{ind}_G\, y_i = \mathrm{ind}_F\, y_i$ for $i = 1, \ldots, k$ and for $j = m+1, \ldots, m+l$ we have $\mathrm{ind}_G\, z_j = \mathrm{ind}_F\, z_j$.*

- *For $j = 1, \ldots, m$, $\mathrm{ind}_G\, z_j^s = \mathrm{ind}_G\, z_j^u = \mathrm{ind}_F\, z_j$.*

- *For $j = 1, \ldots, m$, z_j^s and z_j^u are on the boundary Y, furthermore z_j^s is boundary stable, z_j^u is boundary unstable and $G(z_j^s) < G(z_j^u)$.*

In other words, we can move all critical points to the boundary at once. To prove Theorem 4-10 we use Theorem 3-1 independently for each critical point $z_1, \ldots, z_m$. We need to ensure that Condition (3-2) holds. This is done in Proposition 4-11 stated below. Given these two ingredients the proof is straightforward.

4.6   Topological ingredients needed in the proof of Theorem 4-10

Proposition 4-11   *Let F be a technically good Morse function on the cobordism (Ω, Y). Assume that Σ_0, Σ_1 and Ω have no closed connected components. Let c, d be as in Definition 4-8.*

(a)   *If $n > 1$, then for any $y \in [c, d]$, the inverse image $F^{-1}(y)$ has no closed connected component.*

(b)   *If $n = 1$, then after possibly rearranging the critical values of the interior critical points of index 1, for any interior critical point $z \in \Omega$ of F of index 1, z can be connected with Y by a curve lying entirely in $F^{-1}(F(z))$; and furthermore all the critical points are on different levels.*
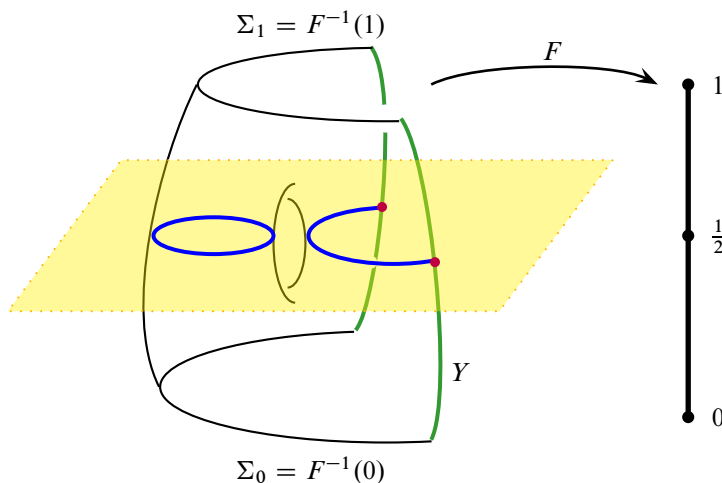
Figure 20: The statement of Proposition 4-11(a) does not hold if $n = 1$. Here F is the height function. The level set $F^{-1}(\frac{1}{2})$, drawn on the picture, has two connected components, one of which is closed.

Remark 4-12 The distinction between the cases $n > 1$ and $n = 1$ is necessary. The conclusion of point (a) of Proposition 4-11 is not necessarily valid if $n = 1$; see Figure 20 for a simple counterexample.

First let us prove several lemmas, which are simple consequences of the assumptions of Proposition 4-11. We use assumptions and notation of Proposition 4-11.

Lemma 4-13 *Let* $x, y \in [0, 1]$ *with* $x < y$. *If* Ω' *is a connected component of* $F^{-1}[x, y]$ *then either* $\Omega' \cap Y = \varnothing$, *or for any* $u \in [x, y] \cap [c, d]$ *we have* $F^{-1}(u) \cap \Omega' \cap Y \neq \varnothing$.

Proof Assume that for some $u \in [x, y] \cap [c, d]$ the intersection $F^{-1}(u) \cap \Omega' \cap Y = \varnothing$ and $\Omega' \cap Y \neq \varnothing$. Then either $\Omega' \cap Y \cap F^{-1}[0, u]$ or $\Omega' \cap Y \cap F^{-1}[u, 1]$ is not empty. Assume the first possibility (the other one is symmetric) and let $Y' = \Omega' \cap Y \cap F^{-1}[0, u]$. Let $f = F|_{Y'}$ be the restriction. Then Y' is compact and f has a local maximum on Y'. This maximum corresponds to a critical point of f of index n, so either a boundary stable critical point of F of index $n + 1$, or a boundary unstable critical point of index n. But the corresponding critical value is smaller than u, so smaller than d, which contradicts property (TG2). □

Lemma 4-14 *For any* $x \in [c, 1]$ *the set* $F^{-1}[0, x]$ *cannot have a connected component disjoint from* Y.

Proof Assume the contrary, and let Ω' be a connected component of $F^{-1}[0, x]$ disjoint from Y. Let Ω_1 be the connected component of Ω containing Ω'. Suppose that $\Omega_1 \cap Y = \varnothing$. Then either $\Omega_1 \cap (\Sigma_0 \cup \Sigma_1) = \varnothing$ or $\Omega_1 \cap (\Sigma_0 \cup \Sigma_1) \neq \varnothing$. In the first case Ω_1 is a closed connected component of Ω, in the second either $\Omega_1 \cap \Sigma_0$, or $\Omega_1 \cap \Sigma_1$ is not empty, so either Σ_0 or Σ_1 has a closed connected component. The contradiction implies that $\Omega_1 \cap Y \neq \varnothing$.

By Lemma 4-13 we have then $F^{-1}(x) \cap \Omega_1 \cap Y \neq \varnothing$, hence $\Omega'' := (F^{-1}[0, x] \cap \Omega_1) \backslash \Omega'$ is not empty and is disjoint from Ω'. As Ω' and Ω'' both belong to Ω_1 which is connected, there must be a critical point $z \in \Omega_1$ of index 1, which joins Ω' to Ω''. We have $F(z) > x$ and Ω', Ω'' belong to two different connected components of $F^{-1}[0, F(z))$ and to a single connected component of $F^{-1}[0, F(z)]$. The connected component of $F^{-1}[0, F(z))$ containing Ω' has empty intersection with Y (by Lemma 4-13) hence z must be an interior critical point of index 1. We also remind the reader that all critical points of F on Ω' are interior critical points, because $\Omega' \cap Y = \varnothing$.

Let W^s be the stable manifold of z of the vector field ∇F. Then $W^s \cap \Omega_1$ must be a connected curve, with nonempty intersection with Ω'. One of its boundaries is either on $\Sigma_0 \cap \Omega'$ or it is a critical point of F in Ω', necessarily interior and by the Morse–Smale condition, its index is 0. In the first case, $\Sigma_0 \cap \Omega'$ is not empty and since it is disjoint from Y, Σ_0 has a closed connected component. In the other case, we have in Ω_1 a single trajectory between a critical point of index 0 and a critical point of index 1. This contradicts (TG3). $\square$

Lemma 4-15 *Assume that for some $y \in [c, d]$ Σ_1 and Σ_2 are two disjoint connected components of $F^{-1}(y)$. If F has no interior or boundary unstable critical points of index n with critical value in $[c, y)$, then Σ_1 and Σ_2 belong to two different connected components of $F^{-1}[0, y]$.*

Proof For $x < y$ and close to y the sets Σ_1 and Σ_2 lie in two different connected components of $F^{-1}(x, y]$. Let $x_0 \geqslant 0$ be the smallest x with that property. Assume $x_0 > 0$. Then x_0 is a critical value of F. The number of connected components of $F^{-1}(x)$ increases as x crosses x_0. Thus the corresponding critical point is either an interior critical point of index n, or a boundary unstable critical point of index n. But then $x_0 \geqslant c$ because of (TG2), so we have $x_0 \in [c, y)$ which contradicts the assumptions of the lemma.

It follows that $x_0 = 0$. As F has no critical points on Σ_0, it follows that Σ_1 and Σ_2 belong to different components of $F^{-1}[0, y]$. $\square$

Lemma 4-16 Let $y \in [c, d]$ be chosen so that there are no interior or boundary unstable critical points of index n with critical values in $[c, y)$. Then $F^{-1}(y)$ has no closed connected components.

Proof Assume that Σ' is a closed connected component of $F^{-1}(y)$. Let $\Sigma'' = F^{-1}(y) \setminus \Sigma'$, it is not empty by Lemma 4-13 (applied for $x = 0$, $y = 1$), for otherwise $\Omega \cap Y = \varnothing$. Let Ω' be the connected component of $F^{-1}[0, y]$ containing Σ'. By Lemma 4-15, Ω' and Σ'' are disjoint, in particular $\Omega' \cap F^{-1}(y) \cap Y = \Sigma' \cap Y = \varnothing$. By Lemma 4-13, $\Omega' \cap Y = \varnothing$. But this contradicts Lemma 4-14. $\qquad\square$

Remark 4-17 There exists a symmetric formulation of the last three lemmas, which can be obtained by considering the function $1 - F$ instead of F. For instance, in Lemma 4-16, the symmetric assumption is that there are no interior or boundary stable critical points of index 1 in $(y, d]$. The statement is the same.

Proof of Proposition 4-11 Case $n > 1$ Let $x \in [c, d]$ be a noncritical value such that all the critical points of F with index n have critical value greater than x, and all critical points with index smaller than n have critical values smaller than x. Such x exists because of (TG1). If $y \leqslant x$, then Lemma 4-16 guarantees that $F^{-1}(y)$ has no closed connected components. If $y > x$, then $F^{-1}[y, d]$ has no critical points of index 1 (as $n > 1$), so we apply the symmetric counterpart of Lemma 4-16.

Case $n = 1$ The property (TG2) implies that the only critical points of $F|_{[c,d]}$ are the interior critical points of index 1. Let us call them $z_1, \ldots, z_m$. Since they are all of the same index, by Proposition 4-6 we are able to rearrange the values $F(z_1), \ldots, F(z_m)$ at will. Let us fix $c_1, \ldots, c_m$ with the property that $c < c_1 < \cdots < c_m < d$. Let us first rearrange the points $z_1, \ldots, z_m$ so that $F(z_1) = \cdots = F(z_m) = c_1$.

The singular level set $F^{-1}(c_1)$ is a singular manifold with m singular points $z_1, \ldots, z_m$, which are double points. By Lemma 4-16, $F^{-1}(c_1)$ has no closed connected components. In particular each of the points $z_1, \ldots, z_m$, can be connected to Y by a curve lying in $F^{-1}(c_1)$. At least one of those points can be connected to Y by a curve γ, which misses all the other critical points. We relabel the critical points so that this point is z_m. We rearrange the critical points so that $F(z_m) = c_m$ and the value $F(z_1) = \cdots = F(z_{m-1}) = c_1$. By construction, z_m can be connected to Y by a curve lying in $F^{-1}(c_m)$.

The procedure now is repeated, ie assume that we have already moved $z_{k+1}, \ldots, z_m$ to levels $c_{k+1}, \ldots, c_m$ respectively. Then $F^{-1}(c_1)$ still has no closed connected components by Lemma 4-16. We assume that z_k can be connected to Y by a curve in $F^{-1}(c_1)$ omitting all the other critical points. Then we rearrange the critical values so that $F(z_k) = c_k$. The proof is accomplished by an inductive argument. $\qquad\square$

4.7 Splitting of cobordisms

We have now all the ingredients needed to prove our main theorem about splitting cobordisms. We slightly change the notation in this subsection; the cobordism will be between (Σ, M) and (Σ', M').

Theorem 4-18 *Let (Ω, Y) be a cobordism between (Σ, M) and (Σ', M'). If*

- *Σ and Σ' have no closed connected components,*
- *Ω has no closed connected component,*

then the relative cobordism can be expressed as a union

$$\Omega = \Omega_0 \cup \Omega_{1/2} \cup \Omega_1 \cup \Omega_{3/2} \cup \cdots \cup \Omega_{n+1/2} \cup \Omega_{n+1}$$

such that $\partial \Omega_s = \Sigma_s \cup \Sigma_{s+1/2} \cup Y_s$ with $\Sigma_0 = \Sigma$, $\Sigma_{n+3/2} = \Sigma'$, $Y = Y_0 \cup \cdots \cup Y_{n+1}$. In other words (Ω_s, Y_s) is a cobordism between (Σ_s, M_s) and $(\Sigma_{s+1/2}, M_{s+1/2})$, where $M_s = \partial \Sigma_s = \Sigma_s \cap Y_s$. Furthermore:

- *(Ω_0, Y_0) is a cobordism given by a sequence of index-0 handle attachments.*
- *For $k = 1, \ldots, n+1$, $(\Omega_{k-1/2}, Y_{k-1/2})$ is a left product cobordism, given by a sequence of index-k left half-handle attachments.*
- *For $k = 1, \ldots, n$, (Ω_k, Y_k) is a right product cobordism, given by a sequence of index-k right half-handle attachments.*
- *(Ω_{n+1}, Y_{n+1}) is a cobordism provided by a sequence of index-$(n+1)$ handle attachments.*

Proof Let us begin with a Morse function F on the cobordism which has only boundary stable critical points (see Remark 2-3). Assume that $w_1, \ldots, w_m$ are the interior critical points and $y_1, \ldots, y_k$ are the boundary critical points. After a rearrangement of critical points and the cancellation of pairs of critical points as in Lemma 4-9 we can make F technically good. After applying Theorem 4-10 we get that F can have only 0 handles and $n+1$ handles as interior handles. Let us write $\theta = 1/(4n+6)$ and choose $c_0 = \theta$, $c_1^s = 3\theta$, $c_1^u = 5\theta$, $\ldots$, $c_k^s = (4k-1)\theta$, $c_k^u = (4k+1)\theta$, $\ldots$, $c_{n+1}^s = 1-3\theta$, $c_{n+1} = 1-\theta$. We rearrange the function F according to Proposition 4-6. Then we define for $k = 0, \frac{1}{2}, 1, \ldots, n+1$ the manifold $\Omega_k = F^{-1}[4k\theta, (4k+2)\theta]$, $Y_k = \Omega_k \cap Y$ and $\Sigma_k = F^{-1}(4k\theta)$.

By construction, each part (Ω_k, Y_k) contains critical points only of one type: for $k = 0$ and $n+1$ they are interior critical points, for $k = 1, \ldots, n$ they are boundary unstable of index k and for $k = \frac{1}{2}, \ldots, n+\frac{1}{2}$, they are boundary stable of index $k+\frac{1}{2}$ and we conclude the proof by Proposition 2-38. □

Remark 4-19 If the cobordism is a product on the boundary, ie $Y = M \times [0, 1]$, we can choose the initial Morse function to have no critical points on the boundary. Then all the critical points of F come in pairs, z_j^s and z_j^u with z_j^s boundary stable, z_j^u boundary unstable and there is a single trajectory of ∇F going from z_j^s do z_j^u.

The strength of Theorem 4-18 is that it is much easier to study the difference between the intersection forms on (Σ_k, M_k) and on $(\Sigma_{k \pm 1/2}, M_{k \pm 1/2})$. We refer to [3] for an application of this fact.

5 The cancellation of boundary handles

In this section we assume that F is a Morse function on the cobordism (Ω, Y) satisfying the Kronheimer–Mrowka–Morse–Smale regularity condition (Definition 4-5). We assume that F has precisely two critical points z and w, with ind $z = k$ and ind $w = k + 1$ and that there exists a single trajectory γ of ∇F going from z to w. If z and w are both interior critical points, then [12, Theorem 5.4] implies that (Ω, Y) is a product cobordism. In fact, Milnor's proof modifies F only in a small neighborhood of γ, which avoids the boundary Y. Hence it does not matter that in our case the cobordism has a boundary.

We want to extend this result to the case of boundary critical points. In some cases an analogue of the Milnor's theorem holds, in other cases we can show that it cannot hold.

5.1 Elementary cancellation theorems

The following generalization of Milnor's theorem was first obtained in [7, Theorem 1].

Theorem 5-1 *Let z and w be a boundary critical points of index k and $k + 1$, respectively. Assume that γ is a single trajectory joining z and w. Furthermore, assume that both z and w are boundary stable, or both boundary unstable. Then (Ω, Y) is a product cobordism.*

As usual, it is enough to prove the result for boundary unstable critical points, the other case is covered if we change F to $1 - F$. Note also, see Section 5.2, the assumption that both critical points are boundary stable, or both boundary unstable is essential.

A careful reading of Milnor [12, pages 46–66] shows that the proof there applies to this situation with only small modifications. Below we present only three steps of that proof, adjusted to our situation. We refer to [12] for all the missing details.

Let ξ be the gradient vector field of F. The proof relies on the following proposition (see the Preliminary hypothesis 5.5 in [12], proved on pages 55–66).

Proposition 5-2 *There exist an open neighborhood U of γ and a coordinate map $g\colon U \to \mathbb{R}_{\geqslant 0} \times \mathbb{R}^n$ and a gradient-like vector field ξ' agreeing with ξ away from U such that:*

- $g(Y) \subset \{x_1 = 0\}$, *and* $g(U) \subset \{x_1 \geqslant 0\}$.
- $g(z) = (0, 0, \ldots, 0)$.
- $g(w) = (0, 1, \ldots, 0)$.
- $g_* \xi' = \eta = (x_1, v(x_2), -x_3, \ldots, -x_k, x_{k+1}, \ldots, x_{n+1})$, *where v is a smooth function positive in $(0, 1)$, zero at 0 and 1 and negative elsewhere. Moreover $|\frac{dv}{dx_2}| = 1$ near 0 and 1.*

Furthermore, U can be made arbitrary small (around γ).

Given the proposition, we argue in the same way as in the classical case, cf [12, pages 50–55]: we improve the vector field ξ' in U so that it becomes a gradient like vector field of a function F', which has no critical points at all. Then the cobordism is a product cobordism.

The proof of Proposition 5-2 is a natural modification of Milnor's proof. After applying arguments as in [12, pages 55–58] the proof boils down to the following result.

Proposition 5-3 (cf [12, Theorem 5.6]) *Let $a + b = n$, $a \geq 1$ and $b \geq 0$ and write a point $x \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}^{a-1} \times \mathbb{R}^b$ as (x_a, x_b) with $x_a \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}^{a-1}$ and $x_b \in \mathbb{R}^b$. Assume that $h\colon (\mathbb{R}_{\geqslant 0} \times \mathbb{R}^{n-1}, \{0\} \times \mathbb{R}^{n-1}) \to (\mathbb{R}_{\geqslant 0} \times \mathbb{R}^{n-1}, \{0\} \times \mathbb{R}^{n-1})$ is an orientation preserving diffeomorphism such that $h(0) = 0$. Suppose that $h(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^{a-1} \times \{0\})$ intersects $\{0\} \times \{0\} \times \mathbb{R}^b$ only at the origin and the intersection is transverse and the intersection index is $+1$. Then, given any neighborhood N of $0 \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}^{n-1}$, there exists a smooth isotopy h'_t for $t \in [0, 1]$ of diffeomorphisms from $(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^{n-1}, \{0\} \times \mathbb{R}^{n-1})$ to itself with $h'_0 = h$ such that:*

(I) $h'_t(x) = h(x)$ *away from N.*

(II) $h'_1(x) = x$ *in some small neighborhood N_1 of 0 such that $\overline{N}_1 \subset N$.*

(III) $h'_1(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^{a-1} \times \{0\}) \cap \{0\} \times \{0\} \times \mathbb{R}^b = \{0\} \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}^n$.

Remark 5-4 The transversality assumption from the assumption of Proposition 5-3 is equivalent to the flow of ξ being Morse–Smale.

The proof of Proposition 5-3 in Milnor's book is given on pages 59–66. We prove here only the analogue of [12, Lemma 5.7]. For all other results we refer to Milnor's book.

Lemma 5-5 *Let h be as in the hypothesis of Proposition 5-3. Then there exists a smooth isotopy $h_t \colon \mathbb{R}_{\geqslant 0} \times \mathbb{R}^{n-1} \to \mathbb{R}_{\geqslant 0} \times \mathbb{R}^{n-1}$, with h_0 the identity map and $h_1 = h$, such that for each t we have $h_t(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^{a-1}) \cap \mathbb{R}^b = 0$.*

Proof We follow the proof of [12, Lemma 5.7]. We shall construct the required isotopy in two steps. First we isotope h by $h_t(x) = \frac{1}{t} h(tx)$. Then $h_1 = h$ and h_0 (defined to be the limit as $t \to 0$) is a linear map, the derivative of h at 0. If this linear map is an identity, we are done. Otherwise h_0 is just a nondegenerate linear map and clearly it maps $\mathbb{R}_{\geqslant 0} \times \mathbb{R}^{a-1} \times \mathbb{R}^b$ diffeomorphically onto itself. It means that under the decomposition $\mathbb{R}^n = \mathbb{R} \oplus \mathbb{R}^{a-1} \oplus \mathbb{R}^b$, h_0 has the block structure

$$h_0 = \begin{pmatrix} a_{11} & 0 & 0 \\ * & A & B \\ * & C & D \end{pmatrix},$$

where $a_{11} > 0$, and stars denote unimportant terms. As h_0 is orientation-preserving, $\det h_0 > 0$. We can apply a homotopy of linear maps which changes the first column of h_0 to $(a_{11}, 0, \ldots, 0)$ and preserves all the other entries of h_0. We do not change the determinant and the condition $h_0(\mathbb{R}_{\geqslant 0} \oplus \mathbb{R}^{a-1}) \cap \mathbb{R}^b = 0$ is preserved (it means that $a_{11} \det A > 0$). Let

$$h_{00} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

We have $\det h_0 = a_{11} \det h_{00}$, so $\det h_{00} > 0$. We use the same reasoning as in Milnor's proof to find a homotopy of h_{00} to the identity matrix, finishing the proof. $\qquad\square$

5.2 Noncancellation results

The two results below have completely obvious proofs, we state them to contrast with Theorem 5-1.

Lemma 5-6 *Assume that a Morse function F on the cobordism (Ω, Y) between (Σ_0, M_0) and (Σ_1, M_1) has two critical points z and w. Suppose z is an interior critical point and w is a boundary critical point. Then (Ω, Y) is not a product cobordism.*

Proof F restricted to Y has a single critical point, so the cobordism between M_0 and M_1 cannot be trivial. $\qquad\square$

Lemma 5-7 *Suppose that F has two critical points z and w. Assume that z is boundary stable and w is boundary unstable. Then (Ω, Y) is not a product.*

Proof If it were a product, we would have $H_*(\Omega, \Sigma_0) = 0$. We shall show that this is not the case.

If $F(z) = F(w)$, then there are no trajectories between z and w, so by Proposition 4-1 we can ensure that $F(z) < F(w)$. So we can always assume that $F(z) \neq F(w)$. For simplicity assume that $F(z) < F(w)$. Let c be a regular value such that $F(z) < c < F(w)$.

By Lemma 2-19 $F^{-1}[0, c] \sim \Sigma_0 \times [0, c]$. Then $H_*(\Omega, \Sigma_0) \cong H_*(\Omega, F^{-1}[0, c])$. Now Ω arises from $F^{-1}[0, c]$ by a right half-handle addition, hence $H_*(\Omega, F^{-1}[0, c]) \cong H_*(H, B)$, where (H, B) is the corresponding right half-handle. But $H_*(H, B)$ is not trivial by Lemma 2-25 (or Lemma 2-18). □

References

[1]   **V I Arnol′d**, **S M Guseĭn-Zade**, **A N Varchenko**, *Singularities of differentiable maps, I: The classification of critical points, caustics and wave fronts*, Monographs in Mathematics 82, Birkhäuser, Boston, MA (1985)   MR777682

[2]   **J M Bloom**, *The combinatorics of Morse theory with boundary*, from: "Proceedings of the Gökova Geometry–Topology Conference 2012", (S Akbulut, D Auroux, T Önder, editors), Int. Press, Somerville, MA (2013) 43–88   MR3203357

[3]   **M Borodzik**, **A Némethi**, **A Ranicki**, *Codimension 2 embeddings, algebraic surgery and Seifert forms*, preprint (2012)   `arXiv:1211.5964`

[4]   **M Borodzik**, **A Némethi**, **A Ranicki**, *On the semicontinuity of the mod 2 spectrum of hypersurface singularities*, J. Algebraic Geom. 24 (2015) 379–398   MR3311588

[5]   **D Braess**, *Morse-Theorie für berandete Mannigfaltigkeiten*, Math. Ann. 208 (1974) 133–148   MR0348790

[6]   **M Goresky**, **R MacPherson**, *Stratified Morse theory*, Ergeb. Math. Grenzgeb. 14, Springer, Berlin (1988)

[7]   **B Hajduk**, *Minimal m–functions*, Fund. Math. 111 (1981) 179–200   MR611759

[8]   **A Jankowski**, **R Rubinsztein**, *Functions with non-degenerate critical points on manifolds with boundary*, Comment. Math. Prace Mat. 16 (1972) 99–112   MR0331409

[9]   **P Kronheimer**, **T Mrowka**, *Monopoles and three-manifolds*, New Mathematical Monographs 10, Cambridge Univ. Press (2007)   MR2388043

[10]   **F Laudenbach**, *A Morse complex on manifolds with boundary*, Geom. Dedicata 153 (2011) 47–57   MR2819662

[11]   **J Milnor**, *Morse theory*, Ann. Math. Studies 51, Princeton Univ. Press (1963) MR0163331

[12] **J Milnor**, *Lectures on the h–cobordism theorem*, Princeton Univ. Press (1965) MR0190942

[13] **L I Nicolaescu**, *An invitation to Morse theory*, Springer, Berlin (2007)   MR2298610

[14] **A Ranicki**, *High-dimensional knot theory: Algebraic surgery in codimension* 2, Springer, New York (1998)   MR1713074

[15] **D Salamon**, *Morse theory, the Conley index and Floer homology*, Bull. London Math. Soc. 22 (1990) 113–140   MR1045282

[16] **S Smale**, *On gradient dynamical systems*, Ann. Math. 74 (1961) 199–206   MR0133139

[17] **S Smale**, *On the structure of manifolds*, Amer. J. Math. 84 (1962) 387–399 MR0153022

[18] **E Witten**, *Supersymmetry and Morse theory*, J. Differential Geom. 17 (1982) 661–692 MR683171

Institute of Mathematics, University of Warsaw
ul. Banacha 2, 02-097 Warszawa, Poland

Alfréd Rényi Institute of Mathematics
Reáltanoda u. 13-15., 1053 Budapest, Hungary

School of Mathematics, University of Edinburgh
Edinburgh EH9 3JZ, UK

mcboro@mimuw.edu.pl,   nemethi.andras@renyi.mta.hu,   nemethi@renyi.hu,
a.ranicki@ed.ac.uk