LINK CONCORDANCE IMPLIES LINK HOMOTOPY

MACIEJ BORODZIK, MARK POWELL, AND PETER TEICHNER

ABSTRACT. We show that link concordance implies link homotopy for immersions of codimension at least two. As a consequence, for $n \geq 2$ every embedding $\sqcup^r S^n \hookrightarrow S^{n+2}$ is link homotopically trivial. This means there is a homotopy to the trivial link, such that, for each time parameter, distinct S^n components are mapped disjointly into S^{n+2}. In other words, beyond the classical case of embedded circles in S^3, there are no 'linking modulo knotting' phenomena in codimension two. To date, this was only known for $n = 2$.

In our proofs we follow, expand, and complete unpublished notes of the third author developing stratified Morse theory for generic immersions, where the d-th stratum is given by points that have d preimages under the generic immersion. We include a discussion of gradient like vector fields, their strata preserving flows, and Cerf theory. This generalizes the case of embeddings, having only two strata, as studied by Perron, Sharpe, and Rourke, and expanded by the first two authors.

Two vital operations in Cerf theory are the rearrangement and cancellation of critical points. In the setting of stratified theory, there are additional rearrangement and cancellation obstructions arising from intersections of ascending and descending membranes for critical points of the Morse function restricted to various strata.

We show that both additional obstructions vanish in codimension at least three, implying a smooth proof of Hudson's result that embedded concordance implies isotopy. In codimension at least two, we show that only the rearrangement obstruction vanishes and we introduce finger moves that eliminate the cancellation obstruction. This is done carefully and only at the expense of introducing new self-intersection points into the components of the immersion. Therefore, our moves keep distinct components disjoint and hence preserve the link homotopy class.

2020 *Mathematics Subject Classification.* Primary: 57K45, 57R42, 57R70.

Key words and phrases. Immersed Morse theory, immersed Cerf theory, stratified Morse theory, concordance, isotopy, link homotopy.

Contents

Part 1. Main results and outlines of proofs.

1.1. Link maps, link homotopy, and link concordance

For a closed manifold X and a connected manifold Y, a continuous map $c: X \to Y$ is called a *link map* if it keeps distinct connected components of X disjoint in Y. In other words, if and only if the induced map $\pi_0(X) \to \pi_0(\mathrm{Image}(c))$ is a bijection. A *link homotopy* is a homotopy through link maps. A *link concordance* is a link map $C: X \times [0,1] \to Y \times [0,1]$ such that $C^{-1}(Y \times \{i\}) = X \times \{i\}$ for $i = 0, 1$.

This notion of link concordance is an analogue for link maps of the usual notion of embedded link concordance of embedded links. However note that when restricted to embedded links, our notion of link concordance does not give the notion of embedded link concordance one finds in the literature, since C need not be an embedding, even if $C|_{X \times \{0,1\}}$ is an embedding.

We focus initially on the case that Y is a sphere S^N and X is a disjoint union of spheres $\sqcup_{i=1}^r S^{n_i}$, with $n_i < N$ for each i.

Milnor [Mil54, Mil57] introduced link homotopies for classical links $L: \sqcup^r S^1 \hookrightarrow S^3$, to distinguish the phenomenon of 'knotting' from that of 'linking'. The Hopf link (formed by two unknots) is not link homotopic to an unlink, as detected by its linking number. Milnor's invariant $\overline{\mu}_{123}$ shows that the Borromean rings are also link homotopically essential.

The fact that Milnor's invariants are also concordance invariants led to the question of whether link concordance implies link homotopy. This was proven in 1979 for 1-dimensional links in S^3 by Giffen [Gif79] and Goldsmith [Gol79] (there is also a later account by Habegger [Hab92]). We extend this result to all dimensions, provided the codimension is at least two.

Theorem 1.1.1. *For any dimensions $n_i \leq N - 2$, two link maps $L, L': \sqcup_{i=1}^r S^{n_i} \to S^N$ are link homotopic if and only if they are link concordant.*

In his PhD thesis [Bar99], Bartels showed that in codimension two and for $n \geq 2$, every embedded link $L: \sqcup^r S^n \hookrightarrow S^{n+2}$ is link concordant to the unlink. Prior to this, Massey and Rolfsen [MR85] had introduced some high dimensional link homotopy invariants and asked whether link concordance implies link homotopy, and also whether every embedding $S^2 \sqcup S^2 \hookrightarrow S^4$ is link homotopically trivial. Our paper shows that both questions have an affirmative answer. Indeed the latter holds much more generally: by combining Theorem 1.1.1 with [Bar99] we obtain the following.

Corollary 1.1.2. *Every smooth embedding $\sqcup^r S^n \hookrightarrow S^{n+2}, n \geq 2$, is link homotopically trivial.*

Corollary 1.1.2 was proven in [BT99] for $n = 2$, where some of the techniques of this paper were introduced. Later [ST19] showed by different methods that a link map $S^2 \sqcup S^2 \to S^4$ with one (topologically) embedded component is link homotopically trivial. Fenn and Rolfsen constructed the first nontrivial link map, exhibited in [FR86, Figure 4]. It has one self-intersection of each component. The main result of [ST19] showed that the group (under connected sum) of link maps $S^2 \sqcup S^2 \to S^4$ is a free module over the commutative ring $\mathbb{Z}[x_1, x_2]/(x_1 \cdot x_2)$, freely generated by the Fenn-Rolfsen link map.

We end our link homotopy discussion for embedded spherical links by looking at the codimension one case. The Schoenflies theorem holds for $n \neq 3$; see [Mil65, Proposition D, p. 112] for $n \geq 4$, implying that a smooth codimension one submanifold $K \subseteq \mathbb{R}^{n+1}$ that is diffeomorphic to S^n is ambiently isotopic to the round n-sphere $S^n \subseteq \mathbb{R}^{n+1}$. We shall refer to K as a (unparametrized) *knot* and define a codimension one link $L \subseteq \mathbb{R}^{n+1}$ to be an *ordered* sequence

$(K_1, \dots, K_r)$ of pairwise disjoint codimension one knots. From the Schoenflies theorem it follows that such links are classified up to isotopy by their dual tree.

Definition 1.1.3. The *dual tree* $t(L)$ of a codimension one link $(K_1, \dots, K_r) = L \subseteq \mathbb{R}^{n+1}$ has $\pi_0(\mathbb{R}^{n+1} \smallsetminus L)$ as set of vertices, with a specified *root vertex* for the unbounded component. For each component K_i of L there is a unique edge e_i in $t(L)$ whose boundary consists of the two connected components of $\mathbb{R}^{n+1} \smallsetminus L$ whose closures meet K_i.

Proposition 1.1.4. *For $n \neq 3$, two smooth codimension one links $L, L' \subseteq \mathbb{R}^{n+1}$ are ambiently isotopic if and only if they are link concordant. Moreover, a generically immersed link concordance induces an isomorphism on dual trees, and if there is an isomorphism $t(L) \cong t(L')$ of rooted, edge-ordered trees, then it is unique and is induced by an ambient isotopy from L to L'.*

The notion of a link concordance between unparametrized links is discussed in Definition 6.4.1. Remark 6.4.2 contains our result on parametrised links modulo link concordance, classified by $t(L)$ with oriented internal edges. We could not find results on link concordance in codimension one in the literature, so we prove them in Section 6.4, where we also discuss *parametrised* knots $S^n \hookrightarrow \mathbb{R}^{n+1}$ and links, and their relation to exotic smooth structures on S^{n+1}.

The classification of *unordered* codimension one links follows from Proposition 1.1.4, by allowing non-order preserving isomorphisms of rooted tree. Proposition 1.1.4 is stronger, namely it implies that there is an ambient self-isotopy of L that permutes its components if and only if the rooted tree $t(L)$ admits a self-isomorphisms with the given permutation of edges. To see this, apply Proposition 1.1.4 with L' the same link as L but with a different ordering, so that the self-isomorphism of the rooted tree becomes an edge-order preserving isomorphism $t(L) \cong t(L')$.

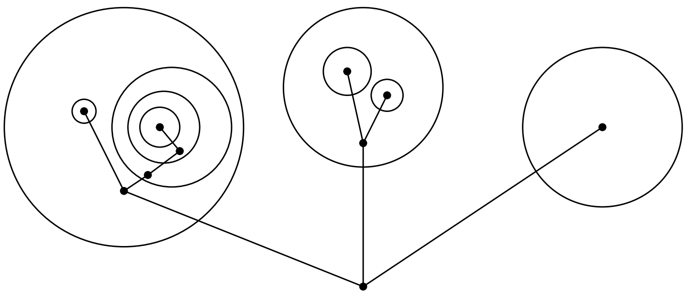

FIGURE 1. Any rooted tree arises as dual tree in dimensions $n \geq 1$.

Note the subtle distinction in Proposition 1.1.4 between the first statement 'link concordance implies isotopy', the most surprising part, and the second statement "a *generically immersed* link concordance *induces* an isomorphism". Figure 62 shows an example where a (non-generically immersed) link concordance does not induce a tree isomorphism (even though the trees are isomorphic).

It follows that approximating immersions are a very useful tool in the study of link homotopy, even in codimension one. Their existence is guaranteed by the following well-known result, which holds in all codimensions at least one, and that we shall apply in the rest of the paper in codimension at least two.

Proposition 1.1.5. *Assuming $n_i \le N-1$ for all i, a link map $L: \sqcup_{i=1}^r S^{n_i} \to S^N$ is link homotopic to an immersion. Similarly, a link concordance $(\sqcup_{i=1}^r S^{n_i}) \times [0,1] \to S^N \times [0,1]$ that restricts to an immersion on the boundary is link homotopic (rel. boundary) to an immersion.*

Proof. Note that for compact domains, link maps are open in the space of all continuous maps. Hence a smooth approximation (in the compact-open topology) will stay a link map (and not map onto S^N, i.e. lie in the parallelisable manifold $\mathbb{R}^N$). Then one shows that L is covered by a monomorphism of tangent bundles, and so by the h-principle is thus homotopic to an immersion. To see that L is covered by a bundle map, fix an inclusion map $\iota_i: \mathbb{R}^{n_i+1} \to \mathbb{R}^N$, and let $L_i: S^{n_i} \to \mathbb{R}^N$ be the ith component of L. Then the composition

$$TS^{n_i} \to TS^{n_i} \oplus \varepsilon \cong S^{n_i} \times \mathbb{R}^{n_i+1} \xrightarrow{(L_i,\iota_i)} \mathbb{R}^N \times \mathbb{R}^N \cong T\mathbb{R}^N$$

covers L_i and is fibrewise injective, as desired. By [Hir59, Thm.5.10], this codimension ≥ 1 immersion can be chosen arbitrarily close to L and is thus link homotopic to L. All arguments have relative versions. □

1.2. Link concordance implies link homotopy for immersions

The proof of Theorem 1.1.1 consists of a number of steps. In the first step, we use Proposition 1.1.5 to improve our link concordance to an immersion. In the second step, we apply the following main result of this paper that holds for all immersions. It was announced in [Tei96] where the strategy of our current proof was outlined but not completed.

Theorem 1.2.1. *Fix a closed $(n-1)$-manifold A and a compact $(n-1+k)$-manifold Y, both smooth. Consider a link concordance $G: A \times [0,1] \to Y \times [0,1]$ that is also an immersion, such that $G|_{A\times\{0\}} = g_0 \times \{0\}$, $G|_{A\times\{1\}} = g_1 \times \{1\}$ with (link) immersions $g_i: A \to Y$.*

If the codimension $k \ge 2$, then there is a regular link homotopy G_τ (rel. boundary) with $G_0 = G$ and such that G_1 is a level-preserving link immersion. In particular g_0 and g_1 are regularly link homotopic.

Moreover, if the codimension $k \ge 3$ and G is an embedding, then G_τ may be chosen to be an embedding for all $\tau \in [0,1]$. In particular g_0 and g_1 are ambiently isotopic.

Remark 1.2.2. Let $V_n(n+k)$ be the Stiefel manifold of orthonormal n-frames in $\mathbb{R}^{n+k}$. It follows from the h-principle and the fibre bundle (with projection given by the last vector)

$$V_{n-1}(n-1+k) \longrightarrow V_n(n+k) \longrightarrow S^{n-1+k}$$

that for $k \ge 2$, a concordance immersion is regularly homotopic (rel. boundary) to a level preserving immersion, i.e. to the track of a regular homotopy. However, we see no argument along these lines that would imply that the outcome is a link homotopy, i.e. that disjoint components would be kept disjoint during the regular homotopy that comes from the h-principle. To obtain the essential output in Theorem 1.2.1, namely a link homotopy, we have to add more geometric techniques.

Figures 2 and 3 imply that our result is the best possible in various ways. The example in Figure 2 shows that Theorem 1.2.1 is not true in codimension one, even with only one component. We use that the winding number in $\mathbb{Z}$ is a regular homotopy invariant of immersions $S^n \looparrowright \mathbb{R}^{n+1}$, which for $n=1$ is the Whitney-Graustein Theorem [Whi37]. The example shows that the winding number can change along an immersed concordance (even though it is invariant modulo 2), and hence immmersed concordance does not imply regular homotopy in general. We do not know whether link concordance implies (non-regular) link homotopy for codimension one (non-spherical) link maps.

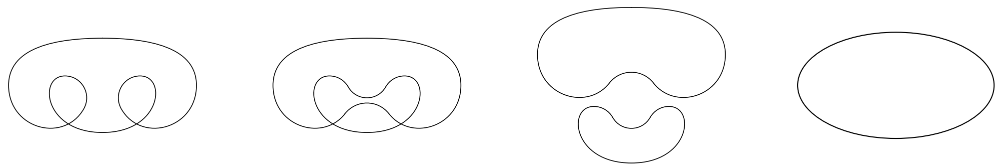

FIGURE 2. Immersed concordance does not imply regular homotopy in codimension one. Going from left to right: the singular link $S^1 \looparrowright \mathbb{R}^2$ to the left is transformed by a saddle move to obtain a singular 2-component link. Then, a Whitney move transforms the singular link to a 2-component unlink. Finally, a minimum cancels one of the components. The trace of these moves provides a singular concordance between the singular link on the left and the unknot on the right. The winding numbers of the immersions are different so they are not regularly homotopic.

Figure 3 shows that in codimension two there exist concordances that are not isotopic to the trace of any isotopy. Even though the minimum and maximum of the trefoil knot cancel in the domain 1-manifold, they cannot be cancelled ambiently, otherwise the trefoil knot would be trivial (as obstructed for example by the Alexander polynomial or Arf invariant). In this case the points at the top and bottom are still isotopic, via a different, simpler isotopy. However one dimension higher there are codimension two examples showing that concordance does not imply isotopy. In Figure 5 below we recall that the stevedore knot is concordant to the unknot, but is not isotopic to it. Again, the last claim can be seen using the Alexander polynomial.

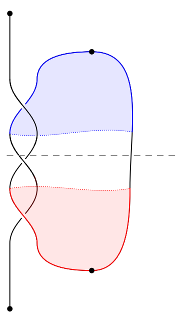

FIGURE 3. A nontrivial cancellation obstruction in codimension two. The intersection of membranes with the planar level presented as a horizontal dashed line is shown in Figure 4.

Theorem 1.2.1 nevertheless has the following consequence in dimension three.

Proposition 1.2.3. *Every classical link $\sqcup^r S^1 \hookrightarrow S^3$ is link homotopic to the closure of a braid with r strands.*

This result was first proven in [HL90] by an inductive procedure using Milnor's classification of almost trivial links in terms of a certain quotient of the fundamental group [Mil54]. Our proof is an alternative direct geometric argument. We prove Proposition 1.2.3 in Section 6.2. It may help the reader of the proof of Theorem 6.1.1 to first understand the low dimensional case, which can be more easily visualised.

For the proof of Theorem 1.2.1 we develop Morse theory on stratifications that come from generic immersions, including gradient-like vector fields, their flows and membranes. This technical core of our paper is described in the next section.

1.3. Morse immersions – an outline of the proof of Theorem 1.2.1

Let N and Ω be smooth manifolds, $F\colon \Omega \to \mathbb{R}$ a Morse function and $G\colon N \looparrowright \Omega$ a generic immersion. A *Morse immersion* is a pair (F, G) that satisfies suitable compatibility conditions between F and G, given in detail in Definition 2.4.1. Roughly speaking, we require that the restriction of F to each G-multiple point stratum

$$\Omega[d] := \{x \in \Omega : |G^{-1}(x)| = d\}, \quad d \geq 0$$

is again a Morse function. In case the generic immersion G is fixed throughout some argument, we also refer to F as an *immersed Morse function* (for G).

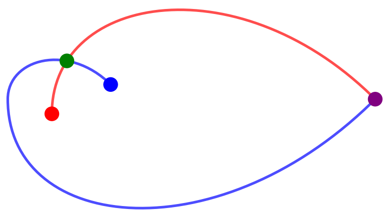

Figure 4. Intersecting membranes in a planar level for the minimum and maximum in Figure 3.

As with classical Morse theory, the key is to understand the behaviour near critical points, how critical points confer topological information, and how Morse immersions can be manipulated in order to cancel or rearrange critical points. This sometimes means changing F alone, sometimes G alone and in some cases, both maps have to be changed.

The main tool for studying Morse immersions is a suitable generalisation of a gradient-like vector field, a so-called gradient-like immersed vector field, in short a *grim* vector field. Roughly, a vector field ξ on Ω is *grim* for a Morse immersion (F, G) if $\partial_\xi F \geq 0$, with equality only at critical points of a restriction $F|_{\Omega[d]}$, and ξ is tangent to all the strata $\Omega[d]$. A more precise definition involves local coordinates for the pair (F, ξ) near critical points of the restrictions $F|_{\Omega[d]}$; see Definition 2.5.1. This characterisation was introduced in [Tei96] and is a direct generalisation of the definition of an embedded gradient like vector field due to Sharpe [Sha88]; see also [BP16].

Grim vector fields are not linear near critical points of $F|_{\Omega[d]}$ for $d > 0$. Therefore the notions of stable (descending) and unstable (ascending) manifolds of a critical point require revision. This leads to the construction of descending and ascending *membranes* in Section 2.6, which are manifolds with corners, described via a product formalism in [Tei96] but not here. These membranes were used in [BT99] if $\dim(Q) = 5$ and generalise those studied by Rourke [Rou70], Perron [Per75], Sharpe [Sha88], and the first two authors [BP16] in the context of embedded manifolds.

One of the key properties of membranes is as follows. Suppose $p \in \Omega[d]$ is a critical point of $F|_{\Omega[d]}$. The stable and unstable manifolds of ξ for p can be defined on $\Omega[d]$. The ascending and descending membranes have strictly higher dimension and correspond to trajectories limiting to p as $t \to \mp\infty$. The membranes associated with a point on $\Omega[d]$ do not lie entirely within $\Omega[d]$ unless $d = 0$. The intersections of membranes of different critical points away from $\Omega[d]$ give rise to an obstruction to rearrangement or cancellation that is not seen in classical Morse theory (see [BP16] for a detailed discussion of the embedded case). From this it follows that it might not be possible to rearrange (respectively cancel) two critical points p, q of $F|_{\Omega[d]}$, that could be rearranged (respectively cancelled) in the classical setting. This obstruction gives a Morse–theoretical explanation of the existence of non-ambiently isotopic embeddings and immersions. For example, Figure 4 explains from this point of view why the trefoil knot is nontrivial.

To be more precise, a careful dimension counting argument combined with a generalised Morse–Smale condition shows that these extra obstructions for rearrangements of critical points are void if the codimension is at least two. In codimension at least three, there are also no obstructions for cancelling a pair of immersed critical points that would be cancellable in the non-immersed case. In the present paper, our primary focus is on codimension two, where rearrangement is possible, but cancellation can be obstructed.

The main result of this paper implements the idea that these cancellation obstructions can be removed at the expense of adding self-intersections into the components of the domain of $G\colon N \looparrowright \Omega$. Suppose the codimension is 2 and p, q are critical points of $F|_{\Omega[1]}$. Assume they can be cancelled in this stratum, i.e. there is precisely one trajectory on $\Omega[1]$ of a Morse–Smale gradient-like vector field ξ that connects p and q. As mentioned above, there might be other trajectories of ξ in Ω connecting p and q. A dimension counting argument and compactness shows that there can be only finitely many such trajectories. Our main technical tool is the introduction of (a sequence of) *finger moves* that decrease the number of trajectories from p to q that stay outside $\Omega[1]$. The price to pay is the introduction of an $(n-2)$–dimensional sphere of self-intersections in a component of N for each finger move. The Morse function F will have some critical points on this sphere, so we introduced new critical points on $\Omega[2]$. After a finite number of such moves, the original critical points p and q can be cancelled. An example of this is illustrated in Figure 5.

Now let us describe an outline of the entire proof. After a perturbing $G\colon A \times [0,1] \looparrowright Y \times [0,1]$ we may assume that the composition

$$f_0\colon A \times [0,1] \xrightarrow{G} Y \times [0,1] \xrightarrow{\mathrm{pr}_2} [0,1]$$

is a Morse function. It is connected to the projection $f_1 := \mathrm{pr}_2\colon A \times [0,1] \to [0,1]$ by a path $f_\tau\colon A \times [0,1] \to [0,1]$ of *generalised Morse functions*, namely functions that are either Morse or have isolated birth or death points. Moreover, this path can be chosen to be independent of τ in a neighbourhood of $A \times \{0,1\}$.

The question is whether we can *lift* such a path f_τ to a regular homotopy G_τ of G. More precisely, we seek a regular homotopy G_τ and a homotopy $\mathrm{pr}_2 \circ G_\tau \simeq f_\tau$ through generic paths of generalised Morse functions that agree with f_0 and f_1 at $\tau = 0, 1$ respectively. Here generic means that there are only finitely many births and deaths, and that they occur at different τ values. If such a regular homotopy G_τ exists, then at the end of the regular homotopy there are no critical points, and so G_1 is level preserving, and is therefore the trace of a regular homotopy.

More generally, suppose (F_0, G_0) is a Morse immersion with $G_0\colon N \looparrowright \Omega$ a generic immersion and $F_0\colon \Omega \to \mathbb{R}$ a Morse function. Consider $f_0 := F_0 \circ G_0$ and assume that f_τ is a smooth path of functions on N. Assume that f_τ is Morse for all but finitely many parameters of τ, where a single birth or death occurs. We aim to *lift* the path of functions f_τ to a path (F_τ, G_τ) of Morse immersions in the sense that $F_\tau \circ G_\tau$ and f_τ are homotopic (relative their endpoints)

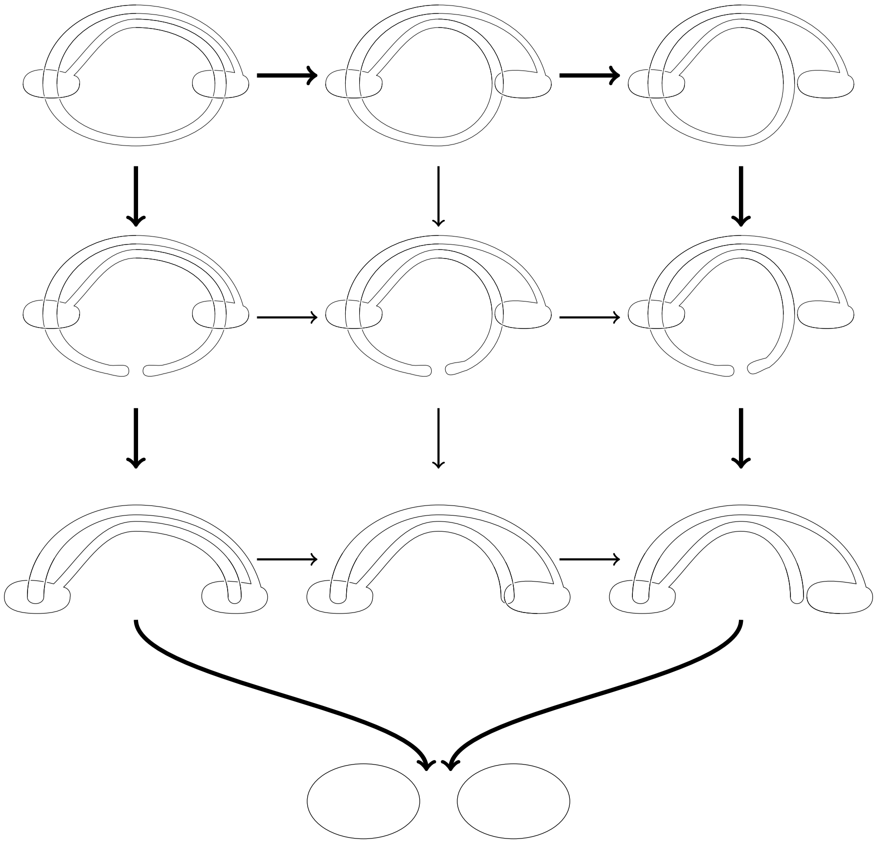

FIGURE 5. A an embedded concordance and a link homotopy (rel. boundary). The arrows from the top-left corner to the bottom represent an embedded concordance from the stevedore knot to the unknot (the obvious deletion of a circle is not presented). The arrows from top-left to the top-right, and then down, represent an immersed concordance with one minimum and one saddle point. The middle route downwards represents the midpoint of a regular homotopy between these two concordances. Going only along the top row from the left to the right gives an immersed concordance from the stevedore knot to the unknot (the knot in the top-right corner is already an unknot). In the right-then-down route, the saddle point and the minimum can be cancelled. A detailed explanation is given in Section 6.3.

through generic paths of generalised Morse functions on N. The obstructions to lifting only arise if:

- there is a τ with a death in f_τ, that is a pair of critical points p, q on N is cancelled, but there is an obstruction (from intersecting membranes) for cancellation of the pair p, q in Ω, or
- there are two critical points p, q on N that get rearranged, see Figure 1.3. In general, such rearrangement need not be possible in Ω but we show that in codimension at least two it is possible, provided the critical point of higher index of f_τ goes below the critical point of lower index.

We solve the first problem in codimension at least two by performing a regular homotopy that introduces self-intersections in the components of N, but that removes the cancellation obstruction. To see that a regular homotopy is necessary in general to obtain a level-preserving map G_1, and that this cannot be done via embeddings, recall that there are many pairs of knots $S^n \subseteq S^{n+2}$ that are concordant but not isotopic.

As an example for the reordering obstruction, let $G: N = S^1 \sqcup S^1 \hookrightarrow \Omega = \mathbb{R}^3$ be the Hopf link. Let f_τ be a path of functions with $f_0 = \mathrm{pr}_3 \circ g$ that pulls the circles apart as in Figure 1.3, i.e. there is a $c \in \mathbb{R}$ such that f_1 maps one circle above c and the other circle below c. Note that f_τ moves a critical point of index 1 below a critical point of index 0 and that this can be realized by a regular homotopy G_τ in the sense that $f_\tau = \mathrm{pr}_3 \circ G_\tau$.

However, the obvious G_τ is not a link homotopy because for some τ, the image $G_\tau(N)$ is connected. In fact, there is no way we can lift the path f_τ by a link homotopy G_τ because the Hopf link has nontrivial linking number.

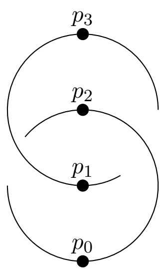

FIGURE 6. A nontrivial reordering obstruction. The critical point p_2 can be abstractly moved below p_1, but it cannot be moved in the embedded case, without creating self-intersections or adding extra critical points.

One of the main technical results in this paper is the following Path Lifting Theorem, which gives a criterion for such lifts of f_τ to exist. We use the notation $\mathrm{bd}\, N$ for the boundary of a manifold N. We fix compact manifolds N and Ω of dimensions n and $n+k$ respectively.

Theorem 1.3.1 (Path Lifting). *Let $(F_0\colon \Omega \to \mathbb{R}, G_0\colon N \looparrowright \Omega)$ be a Morse immersion such that $G_0^{-1}\,\mathrm{bd}(\Omega) = \mathrm{bd}\, N$ and F_0 is constant on connected components of $\mathrm{bd}\,\Omega$. Starting with $f_0 := F_0 \circ G_0$, let $f_\tau\colon N \to \mathbb{R}$ be a generic path of functions that contains no rearrangements where a critical point of higher index goes below a critical point of a lower index.*

If the codimension $k \geq 2$ then f_τ lifts, i.e. there exist:

- *a regular homotopy $G_\tau\colon N \to \Omega$ (rel. boundary) such that G_τ is a generic immersion for all but finitely many τ. Moreover, if G_0 is a link map then so is G_τ for all τ.*
- *a generic path of Morse functions $F_\tau\colon \Omega \to \mathbb{R}$ with (F_1, G_1) a Morse immersion and such that $F_\tau \circ G_\tau$ and f_τ are homotopic in the space of generic paths of functions.*

Furthermore, if G_0 is an embedding, then G_τ can be chosen to be an ambient isotopy if $k \geq 3$, or $k = 2$ and the path f_τ has no deaths.

Remark. In the statement of the Path Lifting Theorem 4.5.1, we will make the assumptions on *genericity* of the path f_τ precise.

From the Path Lifting Theorem, there is a direct way to deduce the Link Concordance Implies Link Homotopy Theorem 1.1.1. Apply the Path Lifting Theorem to $N = A \times [0,1], \Omega = Y \times [0,1]$ and let f_τ be a generic path connecting the given $\mathrm{pr}_2 \circ G_0$ to pr_2. Then we get a path of Morse functions F_τ with $F_0 = \mathrm{pr}_2$ and hence no F_τ has a critical point on the top dimensional stratum $\Omega[0]$.

We construct the regular homotopy by identifying the level sets of the composition $F_1 \circ G_1$. Each such level set is shown to be diffeomorphic to A; the restriction of G_1 to these level sets yields a regular homotopy. We refer to Proposition 2.5.10 for more details.

1.4. A guide through the paper

The paper is split into six parts. After the introduction (Part 1), in Part 2 we develop Immersed Morse Theory. We give the definition of an immersed Morse function and a gradient-like immersed vector field (called a *grim vector field* for short). We prove many technical results about the existence of grim vector fields. If $G: N \looparrowright \Omega$ is a generic immersion, we show how to pass from a grim vector field on Ω to a gradient-like vector field on N and vice versa. These technical constructions are important while proving results on path lifting. A key term from Part 2 is that of *membranes* (Section 2.6), which are analogues of stable and unstable manifolds from classical Morse theory.

Part 3 deals with families of Morse functions. Theorems on rearrangement and cancellation of critical points are proved for immersed manifolds and interpreted as creating a one parameter family of functions such that rearrangement, respectively cancellation, occurs along the family. We also recall the rudiments of classical Cerf theory. Our focus is on one-parameter phenomena, that is: birth, rearrangement, and death. We recast Cerf's results on uniqueness (up to homotopy) of paths of rearrangements in the language of gradient vector fields.

In Part 4, we introduce the concept of *path lifting*. This means that we are given an immersion $G: N \looparrowright \Omega$ and a Morse function $F_0: \Omega \to \mathbb{R}$. We suppose that there is a path of functions $f_\tau: N \to \mathbb{R}$, $\tau \in [0,1]$, such that $F_0 \circ G = f_0$. We want to find a family of functions $F_\tau: \Omega \to \mathbb{R}$ and a regular homotopy $G_\tau: N \to \Omega$ such that $f_\tau = F_\tau \circ G_\tau$ and such that F_τ has no births away from N. The Path Lifting Theorem 4.5.1 gives precise criteria under which such a lift can be found.

The proof of the Path Lifting Theorem in codimension two involves the finger move, which is stated at the end of Part 4 but proven in Part 5. The finger move is one of the central notions of the paper. It allows us to trade intersections of membranes for extra double points of $G(N)$. The whole of Part 5 comprises the construction together with the proof that it has the desired effect.

Part 6 contains examples and applications. The first example is the proof of Proposition 1.2.3. The second example gives the details of the example in Figure 5. It describes a finger move on a concordance in $S^3 \times [0,1]$ bounded by the stevedore knot 6_1 and the unknot, converting that cylinder $S^1 \times [0,1] \subseteq S^3 \times [0,1]$, via a regular homotopy, to the trace of a regular homotopy between 6_1 and the unknot.

Both examples serve as an introduction to the finger move in low dimensional situations where it can be more easily visualised. The reader who wishes to skip the technical construction of the finger move can read this part first. The prerequisite knowledge is the notion of a membrane and the statement of the Cancellation Theorem 3.4.1.

We prove Proposition 1.1.4 on codimension one links in Section 6.4. Section 6.5 shows applications of our results to regular homotopies of surfaces.

Remark. We take the opportunity to point out that the proof of Vector Field Integration Lemma in [BP16] contains a mistake. The function constructed in [BP16, Lemma 3.10] is smooth on trajectories, but it need not be continuous in general: it can have jumps when a trajectory hits a neighbourhood of a critical point, while a nearby trajectory does not hit this neighbourhood. We give a rigorous proof of that result in this paper in a more general setting, that is for immersed manifolds.

1.4.1. **Assumptions and conventions.** Throughout, all manifolds will be compact and smooth, and all immersions and embeddings will be smooth. Manifolds can be orientable or nonorientable, and need not be connected.

The notion of a 'boundary' can refer to the boundary of a manifold, or to the point-set boundary of a subset $U \subseteq \Omega$. In the latter case we write ∂U for $\overline{U} \setminus \mathring{U}$, the closure minus the interior. To avoid potential confusion we shall use $\operatorname{bd} M$ when we mean the manifold boundary of a manifold M. The notation $\operatorname{bd} M$ is somewhat unusual, but we prefer to avoid ambiguity.

Whenever a manifold has nonempty boundary, we will assume that the Morse function F is such that for each connected component $\operatorname{bd}_i \Omega$ of the boundary $\operatorname{bd} \Omega$, there is an $a_i \in \mathbb{R}$ such that $F(\operatorname{bd}_i \Omega) = \{a_i\}$, i.e. F is constant on each connected component of the boundary of Ω. For product manifolds $\Omega = Y \times I$, we will usually assume that Y has no boundary, $F(Y \times \{a\}) = a$ for $a = 0, 1$ and $F(Y \times [0,1]) = [0,1]$.

Acknowledgements. The first and second authors thank the Max Planck Institute for Mathematics in Bonn, the University of Regensburg, and the Banach Center in Warsaw, for hospitality during the time that part of this paper was written. MB was supported by OPUS 2019/B/35/ST1/01120 grant. MP was partially supported by EPSRC New Investigator grant EP/T028335/2 and EPSRC New Horizons grant EP/V04821X/2.

Part 2. **Morse immersions**

We are now going to develop *Morse immersions*, i.e. stratified Morse theory on the multiple-point sets of a generic immersion. The key notion is Definition 2.4.1, a specialisation of the more general notion of a stratified Morse function; see Definition 2.1.3, and compare [GM88].

The second definition, and actually the most important one, is that of a gradient-like immersed Morse vector field, or *grim vector field* for short. We study the ascending and descending submanifolds of critical points, which following Perron [Per75] and Sharpe [Sha88] we call *membranes*, as in the embedded case [BP16].

The results on grim vector fields that we prove are mostly technical, but they are needed in the following parts of the paper. First, we prove existence results, which, given the specific local form of a grim vector field, are not completely trivial. A related result extends a suitable vector field from a submanifold to a grim vector field.

Grim vector fields are used to study the change of level sets of an immersed Morse function. In particular, we show that if a Morse function has no critical points on the zeroth and first stratum, then it induces regular homotopy between level sets. This observation is used in the proof that immersed link concordance implies regular link homotopy.

2.1. Stratified manifolds

The reader acquainted with work of Goresky and McPherson might ask about the relation of our immersed Morse theory with the stratified Morse theory developed in [GM88]. The aim of this section is to review their approach and show how it relates to ours. In short, Goresky–McPherson develop Morse theory in much broader context, but they do not spend much time studying trajectories of vector fields on general stratified spaces. Their motivations were somewhat different to ours.

Recall that a *stratification* $\mathcal{S}$ of a topological space Ω is a set partition of Ω into smooth manifolds $\{\Omega[d]\}_{d\in\mathcal{I}}$, called the *strata*. In the present paper the indexing set $\mathcal{I}$ will always be a subset of the nonnegative integers $\mathbb{N}_0$. We will assume the following two conditions.

(S-1) *Local finiteness:* every point in Ω has a neighbourhood that meets only finitely many strata.

(S-2) *Local triviality:* for every $p \in \Omega[d]$, there is an open set $U \ni p$ in Ω with a stratum preserving diffeomorphism $\varphi\colon U \xrightarrow{\cong} S \times T$ where S is an open set in $\mathbb{R}^k$ for some k, trivially stratified, and T is a stratified manifold with a point stratum $\{q\} \subseteq T$, and $U \cap \Omega[d]$ is identified with $S \times \{q\}$ under φ.

The definition of a stratification might be too general for various results to hold. Often one imposes regularity conditions on the stratification, which is easiest to do when Ω is a smooth manifold and the strata $\Omega[d]$ are smooth submanifolds. We will assume the following regularity conditions.

Definition 2.1.1. A stratification of Ω satisfies *Whitney's conditions* if for every pair of strata, represented locally by disjoint submanifolds $X, Y \subseteq \Omega$, the following hold.

(W-1) Whenever a sequence of points $x_1, x_2, \dots \in X$ converges in Ω to a point $y \in Y$, and the sequence of tangent planes $T_{x_m}X$ converges to a plane $T \subseteq T_y\Omega$, then T contains the tangent plane to Y at y, i.e. $T \supseteq T_yY$.

(W-2) For each sequence $x_1, x_2, \dots$ of points in X and each sequence $y_1, y_2, \dots$ of points in Y, both converging to the same point y in Y, such that the sequence of secant lines L_m between x_m and y_m converges to a line L, and the sequence of tangent planes $T_{x_m}X$ converges to a plane T, we have that $L \subseteq T$.

The limits of secant lines and of tangent planes are taken in the topology on the Grassmannian bundle of $T\Omega$ associated with 1-dimensional (respectively $(\dim X)$-dimensional) subspaces. The Grassmannian space at each fibre inherits a topology as a quotient of the space of $\dim\Omega \times 1$ (respectively $\dim\Omega \times \dim S'$) matrices with rank 1 (respectively $\dim X$).

Given a generic immersion $G\colon N \looparrowright \Omega$, it defines a stratification of Ω by declaring

$$\Omega[d] := \{w \in \Omega : |G^{-1}(w)| = d\}, \quad d \geq 0.$$

Another important example of a stratified manifold is given by manifolds-with-corners. Recall that N is an n-dimensional manifold-with-corners if N is a second countable Hausdorff topological space locally modelled on $\mathbb{R}^{n-d} \times \mathbb{R}^d_{\geq 0}$, for some $d \geq 0$ that depends on the point in N. We define a stratification of N by declaring that $w \in N[d]$ if a neighbourhood of w in $N[d]$ is diffeomorphic to an open subset of $\mathbb{R}^{n-d} \times \mathbb{R}^d_{\geq 0}$, via a diffeomorphism sending w to 0. This type of stratification occurs in this article when discussing ascending and descending membranes of vector fields.

The following result will be used in Section 2.6.2 to guarantee the analogue of the Morse-Smale condition in immersed Morse theory. Its proof can be found in [GM88, Section I.1.3].

Lemma 2.1.2 (Stratified General Position)**.** *Let $\mathcal{S}$ be a locally trivial stratification of a closed manifold Ω satisfying Whitney's conditions, i.e. (S-1),(S-2),(W-1),(W-2) all hold for $\mathcal{S}$. Let $Y_1, Y_2 \subseteq \Omega$. Assume that for every stratum $S \in \mathcal{S}$ the intersections $Y_i \cap S$ are compact submanifolds of S.*

Then there exists a stratum preserving diffeomorphism h of Ω, stratum preserving isotopic to the identity, such that $h(Y_1 \cap S)$ intersects $Y_2 \cap S$ transversely for all strata $S \in \mathcal{S}$. In particular, if

$$\dim(Y_1 \cap S) + \dim(Y_2 \cap S) < \dim S$$

then $h(Y_1 \cap S)$ is disjoint from $Y_2 \cap S$.

Having defined stratifications we can now introduce stratified Morse functions. The definition makes sense for any stratification, and so we give it in this generality, however studying Morse functions on stratifications violating Whitney's conditions is much more involved, and we will not do so.

As usual, given a smooth function $f: L \to \mathbb{R}$, where L is a manifold of dimension ℓ, a point $p \in L$ is a *critical point* if there is a chart $\varphi: U \to L$ with $0 \in U \subseteq \mathbb{R}^\ell$ and $\varphi(0) = p$, such that $\left(\frac{\partial(f\circ\varphi)}{\partial x_i}(0)\right)_{i=1}^{\ell}$ vanishes. The critical point p is *nondegenerate* if the Hessian matrix of second partial derivatives $\left(\frac{\partial^2(f\circ\varphi)}{\partial x_i \partial x_j}(0)\right)_{i,j=1}^{\ell}$ is nondegenerate.

Definition 2.1.3 (Stratified Morse function)**.** A *Morse function* on a stratification $\mathcal{S}$ of Ω is a smooth function $F: \Omega \to \mathbb{R}$ such that:

(M-1) for every stratum S, the restriction $F|_S$ has no degenerate critical points;
(M-2) for every stratum S and for every $p \in S$ and every limit plane $Q = \lim_{p_i \to p} T_{p_i} S'$ we have that $dF_p(Q) \neq 0$, where the $p_i \in S' \neq S$ converge to p.

As above the limit of tangent planes is taken in the topology on the Grassmannian bundle of $T\Omega$ associated with ($\dim S'$-dimensional) subspaces.

A *critical point* of F is by definition an ordinary critical point of one of the restrictions $F|_S$ and the *index* of such a critical point is the ordinary index of the Morse function $F|_S$.

If, in addition to (M-1) and (M-2), all critical values are distinct, then F is called an *excellent Morse function* on $\mathcal{S}$.

Condition (M-2) might sound technical, but its intuitive meaning is that a critical point of the restriction $F|_S$ does not lie on deeper strata. See [GM88, Section I.1.4] for a more detailed discussion. We return to (M-2) in (IM-2) of Definition 2.4.1, which gives a further specialisation of Definition 2.1.3 to immersed Morse functions.

2.2. Generic immersions

Linear subspaces $V_1, \dots, V_\ell$ of a vector space V 'are in general position', 'meet generically', 'intersect transversely', or 'are transverse' if the diagonal map

$$V \longrightarrow \bigoplus_{i=1}^{\ell} (V/V_i)$$

is an epimorphism. If V is equipped with an inner product then this condition translates into the orthogonal complements of the V_i being linearly independent in V. If $M_1, \dots, M_\ell$ are smooth submanifolds of a smooth manifold Ω, one says that they 'meet, or intersect, transversely' at a point $x \in \Omega$ if the tangent spaces $T_x M_1, \dots, T_x M_\ell$ are transverse in $T_x\Omega$.

A smooth map $G: N \to \Omega$ is an *immersion* if $dG: T_pN \to T_{G(p)}\Omega$ is injective for every $p \in N$. Our immersions will be assumed *neat*, which means that $G^{-1}(\operatorname{bd}\Omega) = \operatorname{bd} N$ and G is transverse to $\operatorname{bd}\Omega$.

For a map $G: N \to \Omega$ and $d \in \mathbb{N}_0$, recall that $\Omega[d] \subseteq \Omega$ is the set of points $x \in \Omega$ such that $G^{-1}(x)$ has cardinality d, i.e. there exist exactly d points $u_1, \dots, u_d \in N$ such that $G(u_i) = x$.

Definition 2.2.1. For closed N, an immersion $G: N \to \Omega$ is *generic*, denoted $N \looparrowright \Omega$, if for all $x \in \Omega[d]$, the linear subspaces $dG(T_{u_1}), \dots, dG(T_{u_d})$, for $G(u_i) = x$, meet generically in $T_x\Omega$. If $\operatorname{bd} N$ is nonempty, an immersion G is *generic* if it is neat and the two restrictions $G|_{\operatorname{bd} N}: \operatorname{bd} N \to \operatorname{bd}\Omega$ and $G|_{N \setminus \operatorname{bd} N}: N \setminus \operatorname{bd} N \to \Omega \setminus \operatorname{bd}\Omega$ are both generic immersions.

Note that the genericity condition implies that $\Omega[d]$ is empty whenever $dk > n + k$.

A subspace of Ω will be called a *generically immersed manifold* if it is the image of a manifold N under a generic immersion $G: N \looparrowright \Omega$. From now on N and Ω will be assumed to be compact and immersions will be usually assumed generic. When we study paths of immersions in later parts, we will have to consider the minimal generic failure to a generic immersion at isolated points.

Proposition 2.2.2. *For a generic immersion $G: N \looparrowright \Omega$ the following hold:*

- *$\Omega[d]$ is a submanifold of codimension $(d-1)k$ in Ω for all $d \in \mathbb{N}_0$,*
- *The closure of $\Omega[d]$ is the union $\bigcup_{d' \geq d} O[d']$,*
- *$\{\Omega[d], d \in \mathbb{N}_0\}$ forms a stratification of Ω satisfying conditions (S-1), (S-2), (W-1), and (W-2) from Section 2.1.*

Sketch of proof. The stratification has only finitely many nonempty strata and hence it is also locally finite, i.e. (S-1) holds. For local triviality, we let $x \in \Omega[d]$. Then we take S to be an open neighbourhood of x in $\Omega[d]$. The final transversality condition of Definition 2.2.1 implies that we have a local description of the normal bundle to $\Omega[d]$ at x as a stratified manifold $T \cong \mathbb{R}^{(d-1)k}$ with the origin as a stratum. Thus (S-2) holds.

In order for a sequence of points to occur as in (W-1) and (W-2), there must be an open set $U \cong \mathbb{R}^n$ in N containing the limit point y and the tail of the sequence (x_m), such that $G|_U$ is a smooth embedding and y lies in a deeper stratum than the x_m. In this case the Whitney conditions (W-1) and (W-2) are automatic. □

For example, the fact that $\Omega[d]$ has codimension $(d-1)k$ has the following consequences in low dimensions. A surface N generically immersed in a 4-manifold Ω cannot have triple points, because in that case $d = 3$ and $k = 2$, so triple points would have codimension 4 in N. For generic immersions of 3-manifolds in 5-manifolds, the double points are arcs and circles, and the triple points are again empty.

The next result is a standard consequence of the multijet transversality theorem; see for instance [GG73, Section III.3]. We will prove some generalisations of this result while studying paths of immersions in Subsection 3.1.2.

Lemma 2.2.3. *Generic immersions form an open set and dense subspace of the space of all immersion under the Whitney C^∞-topology. Moreover, generic immersions are* stable *in the sense that every $G: N \looparrowright \Omega$ has an open neighbourhood in $C^\infty(N, \Omega)$ in which each map differs from G only by ambient isotopies of N and Ω.*

Definition 2.2.4. Let $G: N \looparrowright \Omega$ be a generic immersion with $x \in \Omega[d]$. A *branch* of G through x is a neighbourhood $V \subseteq N$ of some point $v \in G^{-1}(x)$ such that the restriction $G|_V$ is one-to-one.

Note that G and $v \in G^{-1}(x)$ determine the germ of a branch V and hence there are exactly d distinct germs of branches through $w \in \Omega[d]$; see Figure 7.

2.3. Jet extensions and the Thom-Boardman stratification

2.3.1. **Multijet transversality.** For two smooth manifolds X and Y, and an integer r, the *jet space* $J^r(X, Y)$ is the space of Taylor expansions up to order r of maps from X to Y (a formal definition can be found in [GG73, Section II.2]). For a C^r-smooth map $G: X \to Y$, we consider its *jet extension* $\mathrm{j}^r G: X \to J^r(X, Y)$. The *source map*, assigning to a Taylor expansion of a map at a point $x \in X$, the point x itself, is denoted $\alpha: J^r(X, Y) \to X$.

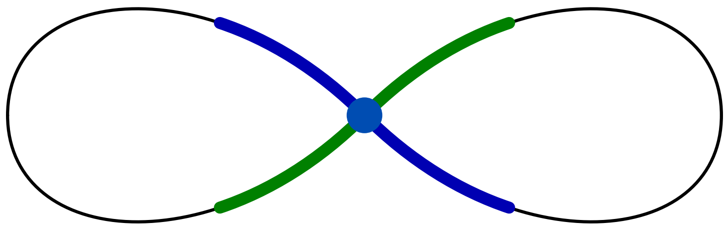

FIGURE 7. Two local branches passing through a double point of a figure 8 in the plane.

We now recall the construction of multijet spaces, referring to [GG73, Section II.4] for more details. Choose an integer $s \geq 1$, where the case $s = 1$ recovers standard jet spaces. The space $X^{(s)} \subseteq X^{\times s} = X \times \cdots \times X$ is the space of distinct s-tuples of points in X, i.e. the ordered configuration space of s points in X. We define $J^r_s(X,Y)$ to be the preimage of $X^{(s)}$ under the s-fold product of the source map $\alpha^{\times s}\colon J^r(X,Y)^{\times s} \to X^{\times s}$. The space $J^r_s(X,Y)$ is called the *multijet space*. For a smooth map $G\colon X \to Y$, there is the notion of a *multijet extension* $\mathrm{j}^r_s G\colon X^{(s)} \to J^r_s(X,Y)$, assigning to an s-tuple $x_1,\dots,x_s$ of pairwise distinct points in X, the Taylor expansions up to order r at these points. The multijet extension has the property that $\alpha^{\times s} \circ \mathrm{j}^r_s G = \mathrm{Id}|_{X^{(s)}}$.

Before we state the Multijet Transversality Theorem we briefly introduce residual subsets and their basic properties.

Definition 2.3.1 (Residual subset, compare [GG73, Definition III.3.2])**.** A subset $A \subseteq T$ is *residual* in T if it is a countable intersection of subsets of X, each of whose interiors is dense in X. A space T is called a *Baire space* if any residual subset of T is dense.

Any Banach space is complete, hence a Baire space. In particular, if X, Y are smooth manifolds, then the space of r-times differentiable maps $C^r(X,Y)$ is a Baire space. The next result has a more complex proof, because C^∞ is only a Fréchet space.

Proposition 2.3.2 (see [GG73, Proposition III.3.3])**.** *Suppose X, Y are two smooth manifolds. Then $C^\infty(X,Y)$ is a Baire space.*

Knowing that a subset is residual is often more useful than knowing it is dense. One reason is that countable intersections of dense subsets are not necessarily dense (let $(q_i)_{i=1}^\infty$ be an enumeration of the elements of $\mathbb{Q}$ and consider $A_i = \mathbb{Q} \smallsetminus \{q_i\}$), but the analogous property is true for residual subsets. We record this as a lemma, whose proof is immediate from the definitions.

Lemma 2.3.3. *Suppose that $A_i \subseteq X$ is residual for $i = 1, 2, \dots$. Then $\bigcap_{i=1}^\infty A_i$ is residual.*

The following result [GG73, Theorem II.4.13] will be extensively used in the future.

Theorem 2.3.4 (Multijet Transversality Theorem)**.** *For any submanifold $S \subseteq J^r_s(X,Y)$ the set of maps $G \in C^\infty(X,Y)$ whose s-fold multijet extension $\mathrm{j}^r_s G$ is transverse to S is residual in $C^\infty(X,Y)$.*

A residual set need not be open, in general. The following corollary translates Theorem 2.3.4 into a ready-to-use criterion. To prove it, use Theorem 2.3.4 in the case S is not a manifold, but a stratified space. This does not pose problems: one simply inducts on the strata (note that we strive to prove that the set of transverse maps is residual, not open, which requires extra conditions). We refer to the discussion in [Arn83, Section 29].

Corollary 2.3.5. *Suppose S is a Whitney stratified subset of $J^r_s(X,Y)$.*

(i) If $\operatorname{codim} S > s\dim X$, the space of maps from $H: X \to Y$, whose s-fold r-th multijet extension misses S, is residual in $C^\infty(X,Y)$.

(ii) If $\operatorname{codim} S > s\dim X + 1$, then the space of paths $H: X \times [0,1] \to Y$ such that for all $\tau \in [0,1]$, the s-fold r-th jet extension of H_τ misses S, is residual in $C^\infty(X\times[0,1],Y)$.

(iii) More generally, if $\operatorname{codim} S > s\dim X + k$, then the space k parameter families $H: X \times D^k \to Y$ such that for all $\tau \in D^k$, the s-fold r-th jet extension of H_τ misses S, is residual in $C^\infty(X \times D^k, Y)$.

The third item implies both the first and second, but since the first and second are the main statements we will need, we state these explicitly too.

Theorem 2.3.4 generalises the standard Thom transversality theorem (case $s = 1$). For $s = 1$, the space of maps missing a subset S satisfying the hypothesis of Corollary 2.3.5 is not only residual, but open-dense. Openness can be sometimes achieved by explicit methods.

Remark. We will use the terminology that some subsets of spaces of smooth functions, such as $\mathcal{I}$, are defined using strata in the corresponding jet spaces. The strata we talk about are subsets of a finitely dimensional jet space, or a multijet space. These sets, like $\mathcal{I}$, are defined as sets of functions whose jet extension misses some given strata. Whereas the actual subsets of smooth functions are infinite dimensional, as subsets of an infinite dimensional Fréchet space, e.g. $C^\infty(N,\Omega)$, the multijet transversality theorem allows us to analyse finite dimensionally, and hence conclude that the original subsets in the Fréchet space are residual.

2.3.2. **Thom–Boardman stratification.** Suppose Z, Y be two smooth compact manifolds of dimension m and m'. For a map $\phi: Z \to Y$, and an integer $r \geq 0$, we define the set $\beta^r\phi$, to be the set of points $x \in Z$ such that $\dim D\phi(T_xZ) = \min(\dim Z, \dim Y) - r$, compare [GG73, Secion VI.1]. This space need not be a manifold, but assume for the moment it is. Then, setting $r_1 = r$ and assuming $r > 0$, we can iterate this procedure. Namely, for $r_2 \geq 0$ define $\beta^{r_1,r_2}\phi$ to be the set of points $x \in \beta^{r_1}\phi$ such that $D(\phi|_{\beta^{r_1}\phi})$ has rank $\min(\dim \beta^{r_1}\phi, \dim Y) - r_2$. Thom–Boardman theory deals with these spaces. In particular, there is a formula for the expected dimension of the sets $\beta^{r_1,\dots,r_\ell}\phi$. The sets $\beta^{r_1,\dots,r_\ell}\phi$ can be defined, using jet extensions, even if β^{r_1} is not smooth, compare [GG73, Section VI.5]. We require that $r_1, \dots, r_{\ell-1} > 0$ and $r_\ell \geq 0$.

Definition 2.3.6 (Thom–Boardman map)**.** A smooth map $\phi: Z \to Y$ is called *Thom-Boardman* if all the $\beta^{r_1,\dots,r_\ell}Z$ are smooth of expected dimension.

The jet transversality theorem implies the following result.

Proposition 2.3.7 (compare [GG73, Theorem VI.5.2])**.** *The set of Thom–Boardman maps is residual in $C^\infty(Z,Y)$.*

Golubitsky-Guillemin [GG73, Theorem VI.5.2] proved residuality of maps satisfying an extra condition, called the NC-condition. We will use the extended version below. Our discussion is specialised to the following situation. Let X, Y be two smooth compact manifolds of dimensions $m+1$ and m. Define $\mathcal{P}(X,Y) \subseteq C^\infty(X,Y)$ by the conditions:

- Π is onto;
- for any $y \in Y$, $\Pi^{-1}(y) \cong [0,1]$,
- for any $x \in X$, $D\Pi(x): T_xX \to T_{\Pi(x)}Y$ is onto.

In the applications, such Π arises naturally from a non-vanishing vector field on X and contraction along trajectories. It is clear from the definition, that $\mathcal{P}(X,Y)$ is open. Note that if $\Pi \in \mathcal{P}$, $Z \subseteq X$ is a submanifold, and $\phi = \Pi|_Z$, then $\beta^{r_1,\dots,r_\ell}\phi = 0$ if at least any of the r_i is greater than 1. Moreover, if $r_1 = \dots = r_\ell$ and $\dim Z = \dim X - k$, then the expected dimension of $\beta^{r_1,\dots,r_\ell}\phi$ is $\dim Z - k\ell$. While latter observation follows from the general formula presented in [GG73, Section VI.5] or [AGZV12, Section I.2], it is quite instructive to understand case $\ell = 1$. Then, for $x \in Z$, we know that $x \in \beta^1\phi$, if and only if $\ker D\Phi \subseteq T_x Z$. Now $\dim\ker D\Phi = 1$, and the codimension of $T_x Z$ in $T_x X$ is k. Hence, the condition $\ker D\Phi \subseteq T_x Z$ is of codimension k.

For this situation, we use the notation $\beta^\ell\phi$ for $\beta^{1,\dots,1}\phi$, and $\beta^{\ell,0}$ for $\beta^{1,\dots,1,0}\phi$, where the sequence of 1's in both definitions has length 1. We adopt a specific variant of Definition 2.3.8

Definition 2.3.8. The map $\Pi\colon X \to Y$ is the *Thom–Boardman projection*, if $\Pi \in \mathcal{P}(X,Y)$, $\phi := \Pi|_Z$ is Thom–Boardman, and $\ell \geq 0$, $\beta^{\ell,0}\phi$ is an immersion with normal crossings.

Lemma 2.3.9. *The set of Thom–Boardman projections is residual in $\mathcal{P}(X,Y)$.*

Proof. Suppose Z and X are compact. The restriction map $C^\infty(X,Y) \to C^\infty(Z,Y)$ is open. As the preimage of a dense subset under an open map is dense, the preimage of a residual set under open map is open. By [GG73, Theorem VI.5.2] (see Proposition 2.3.7), the set of maps ϕ that are Thom–Boardman and satisfy the NC-condition is residual. Hence, the set of map $\Pi \in C^\infty(X,Y)$ whose restriction shares these properties, also is. Next, $\mathcal{P}(X,Y)$ is open in $C^\infty(X,Y)$, meaning that the set of maps $\Pi \in \mathcal{P}(X,Y)$ satisfying that are Thom–Boardman and satisfy the NC-condition, is open in $\mathcal{P}(X,Y)$. But this is exactly the statement of the lemma. □

2.4. Immersed Morse functions

The general notion of a stratified Morse function (Definition 2.1.3) can be specialised to the following definition.

Definition 2.4.1 (Immersed Morse function)**.** For a generic immersion $G: N \looparrowright \Omega$, consider its d-fold point startification $\{\Omega[d], d \geq 0\}$ of Ω. A function $F\colon\Omega \to \mathbb{R}$ is called an *immersed Morse function for G* if for each d and each critical point p of the restriction $F|_{\Omega[d]}$.

(IM-1) the Hessian of $F|_{\Omega[d]}$ is nondegenerate at p and

(IM-2) p is not a critical point of F restricted to the $(d-1)$-fold intersection $Y_1 \cap \dots \cap \widehat{Y_i} \cap \dots \cap Y_d$, for every $i \in \{1,\dots,d\}$, where $Y_1,\dots,Y_d$ are branches of $\Omega[d]$ at p.

If additionally for any two $p \neq p'$ such that p is a critical point of $F|_{\Omega[d]}$ and p' is a critical point of $F|_{\Omega[d']}$ we have $F(p) \neq F(p')$, we call F an *excellent* immersed Morse function.

Remark. Item (IM-1) in Definition 2.4.1 corresponds to condition (M-1) of Definition 2.1.3. Item (IM-2) corresponds to condition (M-2). In the Introduction, we referred to the pair (F,G) as a Morse immersion but since G is fixed in many constructions, it is important now to focus on the properties of F relative to the fixed stratification $\{\Omega[d]\}$ given by G.

Convention 2.4.2. If it is clear from the context, we will call an immersed Morse function simply a Morse function. A point p is called a *critical point of F* if p is a critical point of $F|_{\Omega[d]}$ for some stratum $\Omega[d]$.

We can characterise the local behaviour of a Morse function.

Lemma 2.4.3 (Immersed Morse Lemma)**.** *Suppose $F\colon\Omega \to \mathbb{R}$ is an immersed Morse function and let $p \in \Omega[d]$ be a critical point of $F|_{\Omega[d]}$ of index h. There exist local coordinates at p, denoted*

$$(x_1,\dots,x_m,y_{11},\dots,y_{1k},\dots,y_{d1},\dots,y_{dk}),$$

where $m = n + k - dk$, such that the jth branch of M at p is given, for $j = 1, \dots, d$, by $\{y_{j1} = \dots = y_{jk} = 0\}$, and F in these coordinates has the form

$$F(x_1, \dots, y_{dk}) = F(p) - x_1^2 - \dots - x_h^2 + x_{h+1}^2 + \dots + x_m^2 + \sum_{j=1}^{d} y_{j1}. \tag{2.4.1}$$

In particular, the Immersed Morse Lemma implies that all critical points of an immersed Morse function are isolated.

Proof. We begin with a general result, from which we shall deduce the Immersed Morse Lemma.

Lemma 2.4.4. *Suppose $M^n \subseteq \Omega^{n+k}$ is a generically immersed manifold, let $p \in \Omega[d]$, and assume there are local coordinates $(x_1, \dots, x_m, y_{11}, \dots, y_{dk})$ near p (with $m = n + k - dk = \dim \Omega[d]$) such that the branches of M through p are given by $Y_j = \{y_{j1} = \dots = y_{jk} = 0\}$. Assume that there is a smooth function $F\colon \Omega \to \mathbb{R}$ (not necessarily Morse) such that for each $j = 1, \dots, d$ there exists an index i_j with $\frac{\partial F}{\partial y_{ji_j}}(p) \neq 0$.*

Then there exist local coordinates $(\widetilde{x}_1, \dots, \widetilde{y}_{dk})$ near p such that $Y_j = \{\widetilde{y}_{j1} = \dots = \widetilde{y}_{jk} = 0\}$ for $j = 1, \dots, d$ and the function F has the form

$$F(\widetilde{x}_1, \dots, \widetilde{x}_m, \widetilde{y}_{11}, \dots, \widetilde{y}_{dk}) = F_1(\widetilde{x}_1, \dots, \widetilde{x}_m) + \sum_{j=1}^{d} \widetilde{y}_{j1}$$

for some function F_1.

First we show how to prove the Immersed Morse Lemma from Lemma 2.4.4. Take coordinates $(x_1, \dots, y_{dk})$ near p as in Lemma 2.4.4. Note that $\frac{\partial F}{\partial x_i}(p) = 0$ for all $i = 1, \dots, m$ for otherwise p is not a critical point of $\Omega[d]$ (the coordinates $(x_1, \dots, x_m)$ are local coordinates on $\Omega[d]$). By item (IM-2) of Definition 2.4.1 we claim that for any $j = 1, \dots, d$, there exists an index i_j such that $\frac{\partial F}{\partial y_{ji_j}}(p) \neq 0$. Indeed, if such an index does not exist for given j, the function F restricted to $Y_1 \cap \dots \cap Y_{j-1} \cap Y_{j+1} \cap \dots \cap Y_d$ would have a critical point at p.

By applying Lemma 2.4.4 we obtain that, after changing coordinates, $F(x_1, \dots, y_{dk}) = F_1(x_1, \dots, x_m) + \sum_{j=1}^{d} y_{d1}$. Clearly $F_1 = F|_{\Omega[d]}$ so by item (IM-1) of Definition 2.4.1, F_1 has a nondegenerate critical point at p. We apply the standard Morse Lemma to F_1 (see e.g. [Mil63]) to obtain a change of coordinates of $x_1, \dots, x_m$ to turn F_1 into a sum of quadratic terms. This concludes the proof of the Immersed Morse Lemma, modulo the proof of Lemma 2.4.4, which give next. □

Proof of Lemma 2.4.4. The proof uses induction in the next claim. The precise statement we prove is the following.

Claim. For each $j = 1, \dots, d$, there are local coordinates $(\widetilde{x}_1, \dots, \widetilde{y}_{dk})$ near p such that $Y_j = \{\widetilde{y}_{j1} = \dots = \widetilde{y}_{jk} = 0\}$ and the function F has the form

$$F(\widetilde{x}_1, \dots, \widetilde{x}_m, \widetilde{y}_{11}, \dots, \widetilde{y}_{dk}) = F_1(\widetilde{x}_1, \dots, \widetilde{x}_m, \widetilde{y}_{11}, \dots, \widetilde{y}_{jk}) + \sum_{j'=j+1}^{d} \widetilde{y}_{j'1}.$$

for some function F_1.

Note that for $j = d$ the claim is trivial (take $x_i = \widetilde{x}_i$, $y_i = \widetilde{y}_i$ and $F_1 = F$), while for $j = 0$ the claim corresponds precisely to the statement of Lemma 2.4.4. Now to prove the claim for $j - 1$, assuming the claim for j, we use the inverse function theorem. For simplicity assume that variables for which the claim for j holds are without tildes and the variables that we strive to define are going to have the tilde sign.

First number the variables in such a way that $\frac{\partial F}{\partial y_{j1}}(p) \neq 0$. Define

$$\widetilde{y}_{j1} = F_1(x_1, \dots, x_m, y_{11}, \dots, y_{jk}) - F_1(x_1, \dots, x_m, y_{11}, \dots, y_{j-1,k}, 0, \dots, 0).$$

With this variable and $\widetilde{F}_1 = F_1(x_1, \dots, x_m, y_{11}, y_{j-1,k}, 0, \dots, 0)$ we have

$$F_1(x_1, \dots, x_m, y_{11}, \dots, y_{jk}) = \widetilde{F}_1(x_1, \dots, y_{j-1,k}) + \widetilde{y}_{j1}.$$

Therefore we need to make sure that the change of variables

$$(x_1, \dots, x_m, y_{11}, \dots, y_{jk}) \mapsto (x_1, \dots, y_{j-1,k}, \widetilde{y}_{j1}, y_{j2}, \dots, y_{jk}) \tag{2.4.2}$$

is a local diffeomorphism that preserves strata.

Verification that the change (2.4.2) is a local diffeomorphism is a straightforward application of the inverse function theorem together with the observation that $\frac{\partial F}{\partial y_{j1}}(p) \neq 0$. Clearly the image of the set $\{y_{i1} = \dots = y_{ik} = 0\}$ for $i < j$ is again the set $\{y_{i1} = \dots = y_{ik} = 0\}$ after the change (2.4.2). Therefore the change of variables preserves strata $Y_1, \dots, Y_{j-1}$. The strata $Y_{j+1}, \dots, Y_d$ are unaffected. To see that Y_j is preserved, note that locally near p we have $y_{j1} = \dots = y_{jk} = 0$ if and only if $\widetilde{y}_{j1} = y_{j2} = \dots = y_{jk} = 0$. This uses again that $\frac{\partial F}{\partial y_{j1}}(p) \neq 0$.

The change (2.4.2) provides the necessary induction step. $\square$

Theorem 2.4.5 (Density of Morse functions)**.** *For every generic immersion $G \colon N \looparrowright \Omega$, there exists an immersed Morse function for G. Furthermore the set of immersed Morse functions is open and dense in $C^2(\Omega; \mathbb{R})$.*

While the theorem admits a direct proof using the transversality theorem, we will deduce it from Corollary 3.1.18, which will be proven in Subsection 3.1.3. The proof does not use material from Sections 2.5 – 2.8.

2.5. Gradient like immersed (grim) vector fields

The gradient of a Morse function is a vector field with specific properties. A gradient-like vector field with respect to an immersed Morse function F has many of the key properties of the gradient of a Morse function, but is modified in the neighbourhood of the critical points so that the critical points of F on all the strata are also critical points of the vector field. It is technically extremely useful to have the notion of gradient-like vector fields. They perform the same function as gradients, but are much more flexible.

Definition 2.5.1 (Grim vector field)**.** A vector field ξ on Ω is called a *gradient-like immersed vector field* with respect to an immersed Morse function F, for short a *grim* vector field, if ξ has the following properties.

(G-1) (Positivity) The directional derivative $\partial_\xi F$ is nonnegative and vanishes precisely at the critical points of F.

(G-2) (Tangency) If $w \in \Omega[d]$, then $\xi(w) \in T_w\Omega[d]$. That is, ξ is tangent to the stratum.

(G-3) (Local behaviour at singular points) If $p \in \Omega[d]$ is a critical point of ξ, then there are local coordinates centred at p

$$(x_1, \dots, x_m, y_{11}, \dots, y_{1k}, \dots, y_{d1}, \dots, y_{dk})$$

(with $m = n + k - dk$), such that $\Omega[d]$ is given locally by

$$\{y_{11} = \dots = y_{1k} = 0\} \cup \dots \cup \{y_{d1} = \dots = y_{dk} = 0\},$$

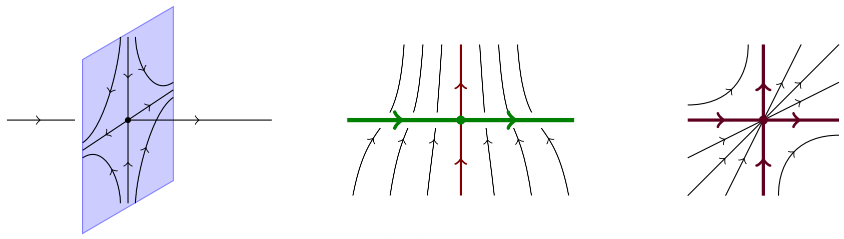

FIGURE 8. Trajectories of a grim vector field near a critical point. Left picture: a critical point of depth one and codimension two. Middle picture: a critical point of depth one and codimension two (the manifold Y is the horizontal line, the 'vertical' trajectories that are drawn go below the line). Right picture: a critical point of depth two and codimension one (the stratified manifold is the union of the central horizontal and vertical lines).

the function F is given by (2.4.1), that is:

$$F(x_1,\dots,y_{dk}) = F(p) - x_1^2 - \dots - x_h^2 + x_{h+1}^2 + \dots + x_m^2 + \sum_{j=1}^{d} y_{j1}$$

for some h with $0 \le h \le m$, and in these coordinates ξ has the form

$$\xi = (-x_1,\dots,-x_h,x_{h+1},\dots,x_m,\sum_{i=1}^{k} y_{1i}^2,0,\dots,0,\sum_{i=1}^{k} y_{2i}^2,0,\dots,0,\sum_{i=1}^{k} y_{di}^2,0,\dots,0). \tag{2.5.1}$$

In particular the $\frac{\partial}{\partial y_{ji}}$-coordinate of ξ is $\sum_{\ell=1}^k y_{j\ell}^2$ for $i=1$ and is zero for $i=2,\dots,k$.

Note that in the local coordinates around the critical points, the difference between ∇F and ξ are the terms $\sum_{i=1}^k y_{ji}^2$ in ξ, whereas the corresponding term in ∇F is 1. The rôle in ξ of the term $\sum_{i=1}^k y_{ji}^2$ is to arrange that ξ vanishes at each critical point. We do not have a cubic term in F, since if we did that, at critical points on deeper strata, F would also have a critical point considered as a function on Ω. The linear terms we use mean that critical points on deeper strata are not also critical points of $F\colon\Omega\to\mathbb{R}$.

2.5.1. **Existence of grim vector fields.**

Theorem 2.5.2 (Existence of grim vector fields)**.** *Every immersed Morse function F admits a grim vector field. Moreover, for every immersed Morse function F and every grim vector field ξ for it, there exists a Riemannian metric on Ω such that $\xi = \nabla F$ away from an arbitrarily small neighbourhood of the set of critical points of F.*

Remark. The conditions (2.4.1) and (2.5.1) exclude the possibility that $\xi = \nabla F$ near a critical point of F restricted to each stratum.

Proof. The proof is similar to the corresponding proof in [BP16, Section 3]. Let $p_1,\dots,p_s$ be the critical points of F. For each $i\in\{1,\dots,s\}$, choose an open neighbourhood U_i' of p_i such that F has the required special form of Definition 2.4.1. Choose U_i to be a neighbourhood of p_i such that $\overline{U}_i \subseteq U_i'$, and let V be an open subset of Ω such that $V\cup\bigcup U_i' = \Omega$ and $U_j\cap V=\varnothing$ for any $i\in\{1,\dots,s\}$. On each of the U_j' we choose a metric g_i in which the local coordinate system of (2.4.1) is orthonormal. Define ξ_i on U_i by (2.5.1).

Let us now define the metric on V. Let $u\in V$ be a point in the d–th stratum. Choose a coordinate system $\{x_1,\dots,x_m,y_{11},\dots,y_{dk}\}$ (with $m=n+k-dk$ as usual) on an open set U_u

containing u and contained in V such that the j–th branch of $\Omega[d]$ is given by $y_{j1} = \cdots = y_{jk} = 0$. As $u \in V$, u is not a critical point of $F|_{\Omega[d]}$. Therefore, there exists an index $i \in \{1,\dots,m\}$ such that $\frac{\partial F}{\partial x_i}(u) \neq 0$. Without loss of generality we suppose it is x_1 and $\epsilon \in \{-1,1\}$ is such that $\epsilon\frac{\partial F}{\partial x_1}(u) > 0$. Shrink U_u if necessary in order to guarantee that $\epsilon\frac{\partial F}{\partial x_1} > \delta_u > 0$ for some $\delta_u > 0$ at all points in U_u. Now for each $u' \in U_u$ there exists a positive definite (and in particular invertible) symmetric matrix $A(u')$ such that the first row, and by symmetry also the first column, is

$$\left(\epsilon\frac{\partial F}{\partial x_1}(u'),\dots,\epsilon\frac{\partial F}{\partial y_{dk}}(u')\right).$$

The sign of the entries other than the first is not controlled, but by choosing the remaining entries of the matrix judiciously we can arrange for the entire matrix to be positive definite. We can also assume that the coefficients of $A(u')$ depend smoothly on u'. We define the metric g_u on U_u such that $A(u')$ is the matrix of the scalar product of vectors in $T_{u'}\Omega$ in the chosen coordinate system.

The gradient of F in the metric g_u satisfies by definition that $\nabla F(u')^T A(u')\nu = dF(\nu)(u')$ for all $\nu \in T_{u'}\Omega$ and for all $u' \in U_u$. As the first row and column of $A(u')$ is ϵ times the vector of partial derivatives of F, it follows that $\nabla F(u')^T = (\epsilon,0,\dots,0) = \epsilon\frac{\partial}{\partial x_1}$. Hence, $\nabla F(u')$ is tangent to all branches of M passing through u'. Define $\xi_u := \nabla F$ in U_u. Note that $\partial_{\xi_u}\nabla F > 0$ on the whole of U_u.

Cover Ω by the sets U_i' for $i = 1,\dots,s$ and U_u for $u \in V$. Let $\phi_1,\dots,\phi_s,\{\phi_u\}_{u\in V}$ be a partition of unity subordinate to this covering. We define $\xi = \sum\phi_i\xi_i + \sum\phi_u\xi_u$ and $g = \sum\phi_i g_i + \sum\phi_u g_u$. From the construction we see that

- ξ has the local form (2.5.1) in each of the U_i, because $\xi = \xi_i$ in U_i;
- if $w \in \Omega[d]$, then $\xi \in T_w\Omega[d]$, because all the ξ_i and ξ_u had this property and belonging to $T_w\Omega[d]$ is preserved under the linear combination operation;
- $\partial_\xi F \geq 0$ with equalities precisely at the critical points of F. This follows from the fact that $\partial_{\xi_i}F \geq 0$ and $\partial_{\xi_u}F > 0$.

This shows that ξ is a grim vector field for F. Moreover, away from the neighbourhood $\bigcup_i U_i'$ of the critical points of F, we have $\nabla F = \xi$. □

An argument analogous to the proof of Theorem 2.5.2 can be used to prove the following result. Its rough meaning is that for a given Morse function F we can extend a vector field η from the submanifold M to a grim vector field on the whole of Ω. The formulation is given in the form that best suits applications in Part 4 below.

Proposition 2.5.3. *Let $F\colon\Omega \to \mathbb{R}$ be an immersed Morse function. Suppose $U \subseteq M$ is an open subset of M contained entirely in the first stratum of M (if M is embedded we can take $U = M$). Let η be a vector field on U that is tangent to M and which is a gradient-like vector field for $F|_U$. Then for any open subset $V \subseteq U$ such that $\overline{V} \subseteq U$ there exists a grim vector field ξ for F such that $\xi|_V = \eta$. If M is embedded and $U = M$, we can take $V = M$.*

Remark 2.5.4. Proposition 2.5.3 is a generalisation of the following classical result. If $f\colon M \to \mathbb{R}$ is a Morse function, U is an open set and η is a vector field that is gradient-like for f on U, then for any open subset V such that $\overline{V} \subseteq U$, there exists a gradient-like vector field η' on M for f with $\eta'|_V = \eta|_V$. This follows from Proposition 2.5.3 by taking $M = \Omega$.

Proof of Proposition 2.5.3. The idea of the proof is first to extend the vector field η to an open subset of Ω containing V and glue it with some grim vector field on Ω.

Let $p_1, \dots, p_s$ be the critical points of F contained in U. For each such point p_i there exist an open set U_{p_i} and coordinates $x_1, \dots, x_n, y_1, \dots, y_k$ such that

- M is given by $\{y_1 = \dots = y_k = 0\}$;
- the vector field η on $U_{p_i} \cap M$ is equal to $(-x_1, \dots, -x_h, x_{h+1}, \dots, x_n)$ for some h. Note that η has n coordinates, and not $n + k$, because η is a vector field on M and not on Ω;
- with h from the previous item, we have $F = -x_1^2 - \dots - x_h^2 + x_{h+1}^2 + \dots + x_n^2 + y_1 + F(p_i)$.

The second item follows from the fact that η is gradient like on Ω. We extend the vector field to the whole of U_{p_i} via

$$\xi_{p_i} = (-x_1, \dots, -x_h, x_{h+1}, \dots, x_n, \sum_{\ell=1}^{k} y_\ell^2, 0, \dots, 0).$$

Suppose $w \in U$ is not a critical point of F. Choose a neighbourhood U_w of w in Ω such that it does not contain any critical points of F. By the rectifiability lemma for vector fields, see e.g. [Arn06, Sections 7 and 32], on shrinking U_w if needed, we may assume that there are coordinates $(x_1, \dots, x_n, y_1, \dots, y_k)$ on U_w such that $M \cap U_w$ is given by $\{y_1 = \dots = y_k = 0\}$ and $\eta = \frac{\partial}{\partial x_1}$ on $M \cap U_w$.

We define $\xi_w = \frac{\partial}{\partial x_1}$. Note that $\partial_\eta F(w') > 0$ for $w' \in U_w \cap M$. Therefore, shrinking further U_w if necessary, we may and shall assume that $\partial_{\xi_w} F(w') > 0$ for all $w' \in U_w$. Finally, let ξ_Ω be any grim vector field for F.

The union of the sets $\{U_w\}$ and $\{U_{p_i}\}$ covers the closure $\overline{V} \subseteq M$. As $\overline{V}$ is compact, there exists an open set $U_V \subseteq \Omega$ containing $\overline{V}$ and whose closure is contained in the union of U_w and U_{p_i}. Define $U_\Omega = \Omega \setminus \overline{U}_V$. Then U_Ω is disjoint from $\overline{V}$. Also U_Ω, together with the sets U_w and U_{p_i}, cover the whole of Ω. We use a partition of unity to glue the vector fields ξ_w on U_w, ξ_{p_i} on U_{p_i}, and ξ_Ω on U_Ω, to a vector field ξ.

Every critical point of F belongs to precisely one subset of the cover. Hence, the vector field ξ satisfies (G-3) near each critical point of F (note that (G-3) is usually not preserved under taking a convex combination of two vector fields, so an extra argument was needed). $\square$

In a similar spirit we can prove the following result.

Proposition 2.5.5. *Suppose $F : \Omega \to \mathbb{R}$ is an immersed Morse function and ξ is a vector field on Ω satisfying $\partial_\xi F \geq 0$, with equality precisely at critical points of F (including critical points of F restricted to deeper strata), and such that ξ is tangent to all the strata.*

Then there exists a grim vector field $\widehat{\xi}$ agreeing with ξ everywhere except perhaps small neighbourhoods of the critical points of F.

Proof. We sketch a quick argument. By the assumptions, ξ already satisfies (G-1) and (G-2) from Definition 2.5.1. We need to take care of (G-3); more specifically, we need to make sure that ξ satisfies (2.5.1) in some local coordinates near each of the critical points of F.

To guarantee this, take $p \in \Omega$ to be a critical point of F and assume p lies on the d-th stratum. By the Immersed Morse Lemma 2.4.3, we may choose a neighbourhood U_p of p such that there are coordinates $\{x_1, \dots, x_n, y_{11}, \dots, y_{dk}\}$ with the property that the j-th branch is given by $\{y_{j1} = \dots = y_{jk} = 0\}$ and F has the form $-x_1^2 - \dots - x_h^2 + x_{h+1}^2 + \dots + x_m^2 + y_{11} + \dots + y_{d1} + F(p)$. Here h is the index of p and $m = n + k - dk$. Keeping in mind (2.5.1), define a vector field on U_p:

$$\xi_p = (-x_1, \dots, -x_h, x_{h+1}, \dots, x_m, \sum_i y_{1i}^2, 0, \dots, 0, \sum_i y_{2i}^2, 0, \dots, 0).$$

Since both F and ξ_p are given by explicit formulae, we may promptly verify that $\partial_{\xi_p} F \geq 0$ everywhere on U_p with equality only at p.

We will now glue ξ with ξ_p. Let $\phi: \Omega \to [0,1]$ be a bump function equal to 1 in a small neighbourhood of p and vanishing away from U. Replace ξ by $\widehat{\xi} = (1-\phi)\xi + \phi\xi_p$. With this definition, $\widehat{\xi}$ still satisfies (G-1) and (G-2), because both conditions are preserved under taking convex linear combination of vector fields. Moreover, $\widehat{\xi}$ satisfies (G-3) near p. We perform the same procedure for all critical points. The new vector field satisfies axiom (G-3) at all critical points. □

2.5.2. **Pull-backs.** Let $G: N \to \Omega$ be a generic immersion, let $M = G(N)$ and let $F: \Omega \to \mathbb{R}$ be an immersed Morse function for (Ω, M). Choose a grim vector field ξ on Ω for F. Define $f: N \to \mathbb{R}$ by $f(u) = F \circ G(u)$. This is a Morse function on N; we refer to it as the 'pull-back' of F. We ask whether we can find a vector field η on N that is a 'pull-back' of ξ, and which is a gradient-like vector field for N.

As our first approximation we could try $\eta_0(u) = (DG(u))^{-1}(\xi)$, where DG denotes the derivative. Since ξ is tangent to M and $DG(u)$ is injective, the vector field η_0 is well-defined. Moreover, unless $p = G(q)$ is a critical point of F, $\eta_0(u) \neq 0$ and $\partial_{\eta_0} f > 0$. However, if p is a critical point of F on the stratum $\Omega[d]$ with $d \geq 2$, then ξ vanishes at p and so η_0 vanishes at each point of the preimage of p, whereas $G^{-1}(p)$ consists of precisely d points, none of them being a critical point of f by item (IM-2) in Definition 2.4.1. Therefore, η_0 is not a gradient-like vector field for f and some alterations are needed. These are described in the next lemma.

Lemma 2.5.6 (Pull-back Lemma)**.** *For any open subset U of N containing all the preimages $G^{-1}(p)$ of critical points of F at depth d at least 2, there exists a vector field η on N such that:*

(P-1) $\eta = \eta_0$ *away from* U*;*
(P-2) η *is gradient-like for* f*.*

Proof. Suppose $p \in \Omega[d]$, $d \geq 2$, is a critical point of F. There are local coordinates $(x_1, \dots, y_{dk})$ in Ω near p such that ξ has the form (2.5.1). The point p has precisely d preimages under G. Let q be one such preimage. Renumbering coordinates if needed, we may and will assume that q corresponds to the branch of M through p given by $y_{d1} = \cdots = y_{dk} = 0$.

The local coordinates near p induce local coordinates in a neighbourhood U_q of q: these are

$$(x_1, \dots, x_m, y_{11}, \dots, y_{d-1,k}).$$

In these coordinates we have

$$\begin{aligned} f &= -x_1^2 - \cdots - x_h^2 + x_{h+1}^2 + \cdots + x_m^2 + y_{11} + \cdots + y_{d-1,1} \\ \eta_0 &= (-x_1, \dots, -x_h, x_{h+1}, \dots, x_m, \sum y_{1i}^2, 0, \dots, \sum y_{d-1,i}^2, 0, \dots). \end{aligned}$$

Here, as usual, h denotes the index of p and $m = n + k - dk$.

Shrink U_q if needed so that the closure of U_q is contained in U and the sets U_q corresponding to different critical points are pairwise disjoint. Choose another neighbourhood U_q' of q such that $\overline{U'}_q \subseteq U_q$. Consider a bump function $\phi_q: N \to [0,1]$ supported on U_q and equal to 1 on U_q'. For $u \in U_q$ we define

$$\eta(u) = \eta_0(u) + \phi \cdot \sum_{j=1}^{d-1} \frac{\partial}{\partial y_{j1}}.$$

For $u \in N \setminus \cup U_q$ define $\eta(u) = \eta_0(u)$. We have $\frac{\partial}{\partial y_{j1}} f = 1$ for all $j = 1, \dots, d-1$, so the new vector field satisfies $\partial_\eta f > 0$ everywhere on U_q.

By construction, $\partial_\eta f > 0$ in a neighbourhood of preimages of all critical points of F at depth 2 or more. As mentioned above, this condition was the only obstruction for η to be a gradient-like vector field. □

2.5.3. **Handle attachments.** Now we study the changes in the topology of the level sets while crossing a critical level of an immersed Morse function $F\colon \Omega \to \mathbb{R}$. Recall that $\dim N = n$ and $\dim \Omega = n + k$. Therefore $\Omega[d]$ has dimension $n + k - dk$.

Proposition 2.5.7. *Suppose $p \in F$ is an index h immersed critical point at the d-th stratum, and suppose that $F^{-1}(F(p))$ has no other critical points. Choose $\varepsilon > 0$ small enough so that F restricted to $F^{-1}(F(p) - \varepsilon, F(p) + \varepsilon)$ has p as its only critical point. For $j = 0, 1, 2, \dots$ denote by $\Omega[j]^\pm$ the intersection of the stratum $\Omega[j]$ with $F^{-1}(F(p) \pm \varepsilon)$.*

- *If $j > d$, then $\Omega[j]^+$ is diffeomorphic to $\Omega[j]^-$;*
- *$\Omega[d]^+$ arises from $\Omega[d]^-$ by a surgery on an embedded $S^{h-1} \times D^{m-h+1}$, that is a surgery of index equal to the index of p, where $m = \dim \Omega[d] = n + k - dk$.*

Proof. The first part follows because on $\Omega[j] \cap F^{-1}(F(p) - \varepsilon, F(p) + \varepsilon)$ the vector field ξ has no critical points. The second part is completely standard. □

Remark. If $j < d$, then $\Omega[j]^+$ usually differs from $\Omega[j]^-$. A precise description of the change uses a generalisation of the 'rising water principle'; see [GS99, Section 6.2]. Since we do not need it in the paper, we shall not dive further into this problem. We aim, however, to return to the rising water principle in a subsequent paper.

We will need a generalisation of [Mil65, Lemma 4.7] and [BP16, Lemma 3.10], enabling us to realise an isotopy from the identity map of a level set to some other self-diffeomorphism of that level set as the flow of a gradient-like immersed vector field. Recall that the *flow* of a vector field ξ on Ω is the 1-parameter family of diffeomorphisms of Ω generated by, or induced by, ξ. See [Mil63, pages 10–11] or [Arn06, Section 7].

Suppose that M is an immersed image in Ω, that $F\colon \Omega \to \mathbb{R}$ is a Morse function and that ξ is a grim vector field inducing a flow Ξ_s. Assume that $a, b \in \mathbb{R}$ are such that there are no critical points of F in $F^{-1}[a, b]$. Use the flow Ξ_s of ξ to identify $F^{-1}[a, b]$ with $F^{-1}(a) \times [a, b]$, as follows. For a point $z \in F^{-1}[a, b]$ we let s_- and s_+ be the real numbers such that $F(\Xi_{s_-}(z)) = a$ and $F(\Xi_{s_+}(z)) = b$. We map $z \in F^{-1}[a, b]$ to the pair $(\Xi_{s_-}(z), \frac{s_+ b - s_- a}{s_+ - s_-}) \in F^{-1}(a) \times [a, b]$. The inverse of this identification induces a diffeomorphism

$$\varphi\colon F^{-1}(a) = F^{-1}(a) \times \{b\} \to F^{-1}(b).$$

Lemma 2.5.8 (Isotopy Insertion)**.** *Let $F\colon \Omega \to \mathbb{R}$, ξ, φ, and $a, b \in \mathbb{R}$ be as above. Let Ψ be a strata preserving self-diffeomorphism of $F^{-1}(a)$ and suppose Ψ is isotopic to the identity through strata-preserving diffeomorphisms.*

There exists a new grim vector field $\widetilde{\xi}$, agreeing with ξ away from $F^{-1}(a, b)$, such that the flow of $\widetilde{\xi}$ induces a strata-preserving diffeomorphism

$$\widetilde{h}\colon (\Omega, M) \cap F^{-1}(a) \to (\Omega, M) \cap F^{-1}(b)$$

(as described above) such that $\widetilde{\varphi} = \varphi \circ \Psi$.

Proof. We follow the proof of [Mil65, Lemma 4.7]. As described above, the flow Ξ_s of ξ yields an identification of $F^{-1}(a) \times [a, b]$ with $F^{-1}[a, b]$. Denote it by $\Phi\colon F^{-1}(a) \times [a, b] \to F^{-1}[a, b]$, so that $\Phi|_{F^{-1}(a)\times\{b\}} = \varphi\colon F^{-1}(a)\times\{b\} = F^{-1}(a) \to F^{-1}(b)$. Let $\Psi_s\colon F^{-1}(a)\to F^{-1}(a)$, $s\in[a,b]$ be an isotopy such that Ψ_a is the identity of $F^{-1}(a)$, and $\Psi_b = \Psi$. We require that Ψ_s does not depend on s for s close to a and b and that Ψ_s preserves the strata of M. Then Ψ_s defines a map $\overline{\Psi}\colon F^{-1}(a)\times[a,b]\to F^{-1}(a)\times[a,b]$ by the formula $(w,s)\mapsto(\Psi_s(w),s)$.

Define the new vector field $\widetilde{\xi} = D(\Phi\circ\overline{\Psi}\circ\Phi^{-1})\xi$ (here D denotes the derivative of the self-diffeomorphism of $F^{-1}[a,b]$). Note that since Φ and $\overline{\Psi}$ are both strata-preserving, and ξ is tangent to the strata, the vector field $\widetilde{\xi}$ is also tangent to the strata of M. The assumption that Ψ_s does not depend on t for t close to $\{a,b\}$ implies that ξ agrees with $\widetilde{\xi}$ near $F^{-1}(\{a,b\})$, so we can extend the vector field $\widetilde{\xi}$ by ξ to the whole of Ω.

By the definition of the flow, the flow of $\widetilde{\xi}$ induces a diffeomorphism $\widetilde{\varphi}\colon F^{-1}(a)\to F^{-1}(b)$ with $\widetilde{\varphi} = \varphi\circ\Psi$, as required. □

The following simple result is a very useful feature of grim vector fields. It is a generalisation of Proposition 2.5.7. It says that if we have an immersed Morse function with no critical points on the zeroth and first strata in the inverse image of an interval, then we obtain a regular homotopy. We will use this as part of the proof that immersed link concordance implies regular link homotopy.

Theorem 2.5.9 (Crossing Deeper Strata)**.** *Suppose $F\colon\Omega\to[0,1]$ is an immersed Morse function and that F does not have any critical points on the zeroth and first stratum. Then $\Omega\cong\Omega_0\times[0,1]$ where $\Omega_0 = F^{-1}(0)$, and there exists a regular homotopy $H_t\colon M_0\to\Omega_0$, where $M_0 := M\cap\Omega_0$, such that M is the trace of the regular homotopy H_t. That is, $M = \overline{H}(M_0\times[0,1])$ where $\overline{H}(w,t) = (H_t(w),t)$.*

Proof. Since F does not have critical points on the zeroth stratum, F regarded as a Morse function on Ω has no critical points at all, so $\Omega\cong\Omega_0\times[0,1]$. As M is the image of a generic immersion, we have $M = G(N)$ for some N. Define $f\colon N\to[0,1]$ by $f(w) = F(G(w))$. As F does not have critical points on the first stratum, f is a Morse function without critical points, hence we can identify N with $M_0\times[0,1]$ in such a way that f is the projection onto the second factor. The homotopy H_t is constructed as follows: for $x\in M_0$, let

$$H_t(x) = G(x,t)\in\Omega_0\times\{t\}\cong\Omega_0.$$

We claim that H_t is a regular homotopy. To see this, we need to show that $DH_t(x)$ is a monomorphism for all $(x,t)\in N\times[0,1]$. To see this, we take $w = (x,t)\in N\times[0,1]$. The map G being an immersion means that $DG(w)$ is a monomorphism. As $f(w) = F(G(w))$, DG takes $\ker Df$ to $\ker DF$. Since Df is non-degenerate, $\ker Df$ is of codimension one in $T_w(N\times[0,1])$. If $DG|_{\ker Df}$ has nontrivial kernel at w, then $\dim DG(\ker Df(w)) < \dim N$. And then $\dim DG(T_w(N\times[0,1])) < \dim N + 1$, contradicting the fact that DG is non-degenerate. But $DG|_{\ker Df} = DH_t$. So DH_t is a monomorphism. □

The homotopy constructed in Theorem 2.5.9 might be not sufficient to prove that concordance implies regular link homotopy. In that setting, we will be given an identification of Ω with $\Omega_0\times[0,1]$ and N as $N_0\times[0,1]$. Theorem 2.5.9 gives (possibly) different such product structures, and so does not give the desired regular homotopy. We have the following improvement of Theorem 2.5.9, under stronger assumptions.

Proposition 2.5.10. *Suppose $N = N_0\times[0,1]$, $\Omega = \Omega_0\times[0,1]$, and $G_\tau\colon N\to\Omega$ is a family of immersions such that $G_0(N_0\times\{0\})\subseteq\Omega_0\times\{0\}$, $G_0(N_0\times\{1\})\subseteq\Omega_0\times\{1\}$, and G_τ does not depend on τ on $N_0\times\{0,1\}$. Assume that there exists a path of functions $F_\tau\colon\Omega\to[0,1]$, $\tau\in[0,1]$, such that F_τ is Morse (as a function on Ω), $F_\tau^{-1}(0) = \Omega_0\times\{0\}$, $F_\tau^{-1}(1) = \Omega_0\times\{1\}$,*

F_0 is projection onto the second factor, and F_1 is an immersed Morse function for $G_1(N)$ with no critical points on the first stratum.

Then the maps

$$H_i = N_0 \to N_0 \times \{i\} \xrightarrow{G_0|_{N_0\times\{i\}}} \Omega \times \{i\} \to \Omega,$$

$i = 0,1$, are regularly homotopic. Moreover, if N is disconnected and for each $\tau \in [0,1]$, we have that G_τ does not create intersections between different components of N, then neither does the regular homotopy between H_0 and H_1.

Proof. As F_0 has no critical points (as a Morse function on Ω), none of the F_τ have critical points either. Stability of Morse functions (see [GG73, Proposition III.2.2]), lets us find a family $\Psi_\tau\colon \Omega \to \Omega$, preserving $\Omega_0 \times \{0,1\}$, and a family of strictly increasing maps $\lambda_\tau\colon [0,1] \to [0,1]$, such that

$$F_\tau = \lambda_\tau \circ F_0 \circ \Psi_\tau.$$

The family Ψ_τ can be assumed to be the identity on $\Omega_0 \times \{0\}$ and to preserve $\Omega_0 \times \{1\}$. For simplicity of the argument we may and will replace the path F_τ by $\lambda_\tau^{-1} \circ F_\tau$ so that we have λ_τ being the identity for all τ.

Now define a new family of immersions, $G'_\tau := \Psi_\tau \circ G_\tau$. Then, $F_0 \circ G'_\tau = F_\tau \circ G_\tau$. In particular, since by hypothesis F_1 has no critical points of depth one, with respect to the stratification determined by $G_1(N)$, it follows that F_0 has no critical points at depth 1 with respect to the stratification determined by $G'_1(N)$. Thus by Theorem 2.5.9, the composition $F_0 \circ G'_1$ induces a product structure on N, which we denote by $\psi\colon N \xrightarrow{\cong} N_0 \times [0,1]$. Define a family of maps as

$$H'_t\colon N_0 \to N_0 \times \{t\} \subseteq N \xrightarrow{G'_1} \Omega_0 \times \{t\} = \Omega_0.$$

By hypothesis, G_τ is independent of τ on $N_0 \times \{0,1\}$, and moreover we arranged for Ψ_τ to act as the identity near $\Omega_0 \times \{0,1\}$. Hence $H'_0 = H_0$ and $H'_1 = H_1$, and so H'_t gives a regular homotopy between H_0 and H_1..

It remains to prove that if G_τ keeps connected components of N separate, then so does the regular homotopy. Note that G_τ separates connected components if and only if G'_τ does. Hence in particular G'_1 separates connected components. But the homotopy H_t in the proof of Theorem 2.5.9 is constructed by tracing along what is called $G(N)$ in the notation of that theorem, the role of which in the present proposition is played by $G'_1(N) = G'_1(N_0 \times [0,1]$. Hence no intersections between different components are created. □

2.6. Membranes

2.6.1. **Definitions.** Recall that $M^n \subseteq \Omega^{n+k}$ is a generically immersed manifold, $F\colon \Omega \to \mathbb{R}$ is an immersed Morse function and ξ is a grim vector field for F. In the case of a critical point on $\Omega \setminus M$, as in classical Morse theory, the stable and unstable manifolds are discs (at least near the critical point), whose intersections with level sets $F^{-1}(c)$ are spheres whose dimension depends on the index of the critical point. If $p \in \Omega[d]$ is a critical point of ξ and $d > 1$, the linear part of ξ is degenerate at p.

Since we know the form of ξ in a neighbourhood of p, we can explicitly describe the dynamics of the vector field near p. The objects of interest to us are the set of points that are attracted to the critical point (in classical Morse theory, the stable or descending manifold) and the set of points that are repelled (in classical Morse theory, this corresponds to the unstable or ascending manifold).

Definition 2.6.1 (Membranes)**.** For a critical point p of ξ, the ascending (respectively descending) membrane $\mathbb{M}_{\mathrm{a}}(p)$ (respectively $\mathbb{M}_{\mathrm{d}}(p)$) the set of points w such that a trajectory of ξ through w reaches p in the infinite past (respectively in the infinite future).

The notion is a straightforward generalisation of the membranes of Perron [Per75] and Sharpe [Sha88], which were defined for embedded submanifolds; see also [BP16].

In local coordinates (2.5.1) the descending membrane is given by

$$\{x_{h+1} = \cdots = x_r = 0 = y_{12} = \cdots = y_{1k} = y_{22} = \cdots = y_{2k} = \cdots = y_{dk}\} \cap \{y_{11} \le 0,\, y_{21} \le 0, \dots, y_{d1} \le 0\}.$$

Similarly, the ascending membrane is given by

$$\{x_1 = \cdots = x_h = 0 = y_{12} = \cdots = y_{1k} = y_{22} = \cdots = y_{2k} = \cdots = y_{dk}\} \cap \{y_{11} \ge 0,\, y_{21} \ge 0, \dots, y_{d1} \ge 0\}.$$

The ascending and descending membranes $\mathbb{M}_{\mathrm{a}}(p)$ and $\mathbb{M}_{\mathrm{d}}(p)$ are manifolds-with-corners. In particular they are stratified manifolds in the sense of Section 2.1: the d-th stratum of $\mathbb{M}_{\mathrm{a}}(p)$ is the intersection of $\mathbb{M}_{\mathrm{a}}(p)$ with $\Omega[d]$. In particular, we will be able to use Stratified General Position Lemma 2.1.2 to study intersections of membranes of different critical points. As a first step in this direction, we prove the following result.

Lemma 2.6.2. *Given a critical point p of ξ of index h and depth d, for any $j \le d$ we have* $\dim \mathbb{M}_{\mathrm{d}}(p) \cap \Omega[j] = h + d - j$*, and* $\dim \mathbb{M}_{\mathrm{a}}(p) \cap \Omega[j] = n - k(d-1) - h + d - j$*.*

Proof. The stratum at depth d has dimension $n - k(d-1)$. The stable manifold at the d-th stratum has dimension h and the unstable manifold has the complementary dimension $n-k(d-1)-h$. This gives the claimed answer when $d = j$. Now the dimension of the membrane at each stratum increases by one as we go to a more shallow level (i.e. decreasing j by one). ☐

2.6.2. **The Morse-Smale condition.** In classical Morse theory, the Morse–Smale condition means that stable and unstable manifolds of different critical points of a vector field intersect transversely. The generalisation to the case of grim vector fields is a straightforward extension of the approach of [BP16].

Definition 2.6.3. Let $M \subseteq \Omega$ be an immersed manifold, $F\colon \Omega \to \mathbb{R}$ a Morse function and ξ a grim vector field for F. We will say that ξ satisfies the *Morse–Smale condition* if for any two critical points p, p' of ξ, the ascending membrane of p intersects the descending membrane of p' transversely, in the sense of the intersection of stratified manifolds in the stratified manifold Ω.

More specifically, for every $j = 0, 1, \dots$, the intersection of membranes $\mathbb{M}_{\mathrm{a}}(p) \cap \Omega[j]$ and $\mathbb{M}_{\mathrm{d}}(p') \cap \Omega[j]$ is required to be transverse in $\Omega[j]$.

Proposition 2.6.4. *Every grim vector field can be perturbed outside of critical level sets to a vector field that satisfies the Morse–Smale condition.*

Sketch of proof. The proof is essentially a repetition of the proof given in [Mil65, Theorem 5.2]. The only difference is that we use the fact that we need to use the analogue of [Mil65, Lemma 5.3] in the stratified category, but this is essentially the statement of Lemma 2.1.2. ☐

For Morse–Smale grim vector fields we can calculate the dimension of the intersections of membranes.

Lemma 2.6.5. *Suppose a grim vector field ξ is Morse–Smale and p, p' are two critical points with $F(p) < F(p')$, such that the index of p is h and the index of p' is h'. Suppose that p lies* *on the d–th stratum and p' lies on the d'–th stratum. Then for every $c \in (F(p), F(p'))$ and for depth $j \geq 0$ we have*

$$\dim \mathbb{M}_{\mathrm{a}}(p) \cap \mathbb{M}_{\mathrm{d}}(p') \cap \Omega[j] \cap F^{-1}(c) = (h' + d') - (h + d) + (k - 2)(j - d) - 1.$$

Proof. According to Lemma 2.6.2 we have

$$\dim \mathbb{M}_{\mathrm{a}}(p) \cap \Omega[j] = n - k(d-1) - h + d - j \text{ and } \dim \mathbb{M}_{\mathrm{d}}(p') \cap \Omega[j] = h' + d' - j.$$

We also have $\dim \Omega[j] = n - k(j-1)$. By transversality, the dimension of the intersection in $\Omega[j]$ is equal to

$$\begin{aligned}\left(h' + d' - j\right) + (n - k(d-1) - h + d - j) - (n - k(j-1)) \\ = (h' + d') - (h + d) + (k-2)(j-d).\end{aligned}$$

Intersecting $\mathbb{M}_{\mathrm{a}}(p) \cap \mathbb{M}_{\mathrm{d}}(p') \cap \Omega[j]$ with the level set $F^{-1}(c)$ drops the dimension by 1. □

The following corollary will be extremely useful.

Corollary 2.6.6. *Suppose the codimension $k \geq 2$. With the notation of Lemma 2.6.5, if $\mathbb{M}_{\mathrm{a}}(p) \cap \mathbb{M}_{\mathrm{d}}(p')$ is nonempty, then $h' + d' > h + d$.*

Proof. By Lemma 2.6.5,

$$\dim \mathbb{M}_{\mathrm{a}}(p) \cap \mathbb{M}_{\mathrm{d}}(p') \cap \Omega[j] \cap F^{-1}(c) = (h' + d') - (h + d) + (k - 2)(j - d) - 1.$$

If $\mathbb{M}_{\mathrm{a}}(p) \cap \mathbb{M}_{\mathrm{d}}(p') \neq \varnothing$, then this number is nonnegative for some j. Clearly $j \leq d$, because by the tangency condition (G-2) the depth of a limit point of a trajectory is never smaller than the depth of a generic point of the trajectory.

Since $k \geq 2$, it follows that $(k-2)(j-d) \leq 0$. Thus

$$(h' + d') - (h + d) - 1 \geq (h' + d') - (h + d) + (k - 2)(j - d) - 1 \geq 0,$$

and therefore $h' + d' > h + d$ as claimed. □

Intersections of membranes give rise to obstructions to performing rearrangement and cancellation of critical points. The contrapositive of Corollary 2.6.6 can be used to show that such intersections do not occur.

2.7. Grim neighbourhoods and broken trajectories

2.7.1. **Broken trajectories.** Let $F\colon \Omega \to \mathbb{R}$ be a Morse function and let ξ be a grim vector field for M. Recall that a *trajectory* of ξ is a curve $\gamma\colon \mathbb{R} \to \Omega$ such that $\frac{d}{dt}\gamma(t) = \xi(\gamma(t))$. The *limit points* of the trajectory are $z := \lim_{t\to\infty} \gamma(t) \in \Omega$ (forward limit) and $w := \lim_{t\to-\infty} \gamma(t) \in \Omega$ (backward limit). We say that γ reaches z in the *infinite future* and w in the *infinite past.*

Definition 2.7.1. Let ξ be a smooth vector field in Ω.

(a) A *broken trajectory* of ξ is a finite collection $\gamma_1, \dots, \gamma_s$ of trajectories of ξ such that the forward limit of each of the γ_i is the backward limit of γ_{i+1} fo $i = 1, \dots, s-1$.

(b) If Ω' is the closure of an open set of Ω, a *zigzag trajectory* in Ω' is a finite collection of trajectories $\gamma_1, \dots, \gamma_s$ all contained in Ω' and such that one of the limit points of γ_i coincides with one of the limit points of γ_{i+1} for every $i = 1, \dots, s-1$; see Figure 9.

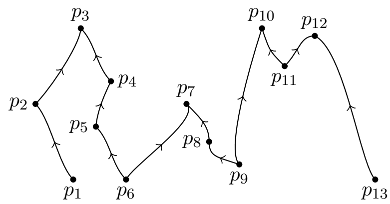

FIGURE 9. An example of a zigzag trajectory. All points $p_1, \dots, p_s$ are critical points of the vector field.

Definition 2.7.2. Let $F: \Omega \to \mathbb{R}$ be a Morse function and let ξ be a grim vector field for F. Let $a < b$ be two non-critical values of F. For two points $w, z \in F^{-1}([a,b])$ we say that $w \sim_0 z$ if there exists a (subset of a) trajectory of ξ union its limit points, contained in $F^{-1}([a,b])$, connecting w and z. We define the relation $\sim_{\xi,a,b}$ as the transitive closure of the relation $\sim_0$.

For any point $w \in F^{-1}[a,b]$ we define $\mathcal{K}_{\xi,a,b}(w)$ to be closure of the set of points z in $F^{-1}[a,b]$ such that $w \sim_{\xi,a,b} z$.

Remark. We have that $w \sim_{\xi,a,b} z$ if and only if there is a zigzag trajectory of ξ in $F^{-1}[a,b]$ connecting w with z.

2.7.2. **Grim neighbourhoods.** In the study of dynamics of vector fields, there is the notion of an index pair; see e.g. [CZ84, Sal90]. If ξ is a vector field, an *index pair* is a pair of sets (X, L) such that $L \subseteq X$ is the exit set for ξ. That is, if γ is a trajectory of ξ, $\gamma(0) = x \in X$, then there exists $t_0 > 0$ such that for $t \in (0, t_0)$, if $x \in X \setminus L$ then $\gamma(t) \in X$, and if $x \in L$ then $\gamma(t) \notin X$. Index pairs are used to study the behaviour of ξ near critical points.

Definition 2.7.3 (Index triple)**.** A triple $(X, L_{\text{in}}, L_{\text{out}})$ forms an *index triple* for ξ, if (X, L_{out}) forms an index pair for ξ, and (X, L_{in}) forms an index pair for $-\xi$.

The definition means that any trajectory of ξ that intersects X enters X through L_{in} and exits through L_{out}.

Example 2.7.4. Suppose $p \in \Omega$ is an isolated critical point of ξ. Then any neighbourhood U of p contains an index triple that contains p (see [Sal85, Theorem 4.3] for a more general statement).

In our applications, we will impose smoothness conditions on the index triple. This leads to a notion of a regular index triple.

Definition 2.7.5 (Regular index triple)**.** A *regular index triple* is a triple $(X, \partial_{\text{in}} X, \partial_{\text{out}} X)$ such that X is a manifold-with-corners of codimension 0 in Ω, $\partial_{\text{in}} X, \partial_{\text{out}} X \subseteq \partial X$ are manifolds with boundary (of codimension 0 in ∂X), $\partial_{\tan} X := \partial X \setminus (\partial_{\text{in}} X \cup \partial_{\text{out}} X)$ is a submanifold, and the corners of ∂X are at $\partial_{\text{in}} X \cap \partial_{\tan} X$ and $\partial_{\text{out}} X \cap \partial_{\tan} X$. We also require that $\partial_{\text{in}} X$ and $\partial_{\text{out}} X$ be disjoint.

We point out that there is no ambiguity with writing $\operatorname{bd} X$ or ∂X in Definition 2.7.5, since X is a codimension zero submanifold of Ω, and so $\operatorname{bd} X = \partial X$, i.e. the manifold boundary and the point-set boundary/frontier agree.

It follows from the definition that ξ is everywhere tangent to $\partial_{\tan} X$: if not, the trajectory of ξ through a point in $\partial_{\tan} X$ would leave X immediately either in future or in the past, so that point would belong to either $\partial_{\text{in}} X$ or $\partial_{\text{out}} X$.

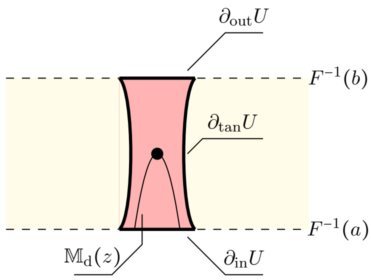

FIGURE 10. A grim neighbourhood.

Remark 2.7.6. For an arbitrary vector field, given an isolated critical point p, and a neighbourhood V containing p but no other critical points, there need not exist a regular index triple X with $p \in X \subseteq V$. However, if p is a critical point of a grim vector field ξ, then an index triple exists; see Lemma 2.7.9 below.

Definition 2.7.7 (Grim neighbourhood)**.** Suppose F is an immersed Morse function and ξ is a grim vector field for F. A *grim neighbourhood* of a set C is a regular triple index $(X, \partial_{\text{in}}X, \partial_{\text{out}}X)$ such that $C \subseteq X$, $X \setminus C$ has no critical points, and each of $\partial_{\text{in}}X, \partial_{\text{out}}X$ belong to a single level set of F.

In most cases, we will consider a grim neighbourhood for an isolated critical point of ξ. We make the following observation.

Lemma 2.7.8. *Let U_C be a grim neighbourhood of a closed subset C with $\partial_{\text{out}}U_C \subseteq F^{-1}(b)$ and $\partial_{\text{in}}U_C \subseteq F^{-1}(a)$ for some a, b (necessarily $a < b$).*

Let $a', b' \in \mathbb{R}$ be such that $C \subseteq F^{-1}(a', b')$ and $a < a'$, $b' < b$. Then $U_C \cap F^{-1}(a', b')$ is also a grim neighbourhood of C.

The proof of Lemma 2.7.8 is completely straightforward and will be omitted. We pass to the construction of a grim neighbourhood.

Lemma 2.7.9. *Suppose $p \in \Omega$ is a critical point of F and a, b are such that there is no other critical point in $F^{-1}[a, b]$. Then for any open subset $V \subseteq F^{-1}(a)$ containing $\mathbb{M}_{\text{d}}(p) \cap F^{-1}(a)$ there exists a grim neighbourhood $U \subseteq \Omega$ of p such that $\overline{V} = \partial_{\text{in}}U \subseteq F^{-1}(a)$ and $\partial_{\text{out}}(U) \subseteq F^{-1}(b)$.*

Proof. Let W be the complement of V in $F^{-1}(a)$. Define $U = F^{-1}(a, b) \setminus \Xi_{(a,b)}(W)$, where we recall that $\Xi_{(a,b)}(W)$ is the set of points in $F^{-1}(a, b)$ that are reached from W by the flow of ξ. The verification that U satisfies the desired conditions is straightforward. □

Remark. The same argument shows that if $V \subseteq F^{-1}(b)$ contains $\mathbb{M}_{\text{a}}(z) \cap F^{-1}(b)$, then there is a grim neighbourhood U of z such that $\partial_{\text{out}}U = \overline{V}$.

We can now state a result on grim neighbourhoods.

Lemma 2.7.10. *For any open subset $V \subseteq F^{-1}[a, b]$ with $\mathcal{K}_{\xi,a,b}(w) \subseteq V$, there exists a grim neighbourhood V' of $\mathcal{K}_{\xi,a,b}(w)$ in $F^{-1}[a, b]$ such that $V' \subseteq V$.*

Proof. In the proof we will use a weak version of the Broken Trajectory Lemma 2.7.14, which appears below i.e. in the presence of the function F. This formulation of the Broken Trajectory Lemma is well-known, and is commonly used in the proof that $\partial^2 = 0$ in Morse homology; compare e.g. [Sal90, Proof of Theorem 3.1]. It does not depend on Lemma 2.7.10, so we do not have a circular dependence of lemmas.

Suppose $\mathcal{K}_{\xi,a,b}(w)$ contains critical points $p_1,\dots,p_r$. If $r = 0$, that is, there are no critical points, then $\mathcal{K}_{\xi,a,b}(w)$ is just the intersection of $F^{-1}[a,b]$ with the trajectory through w. In that case, we let w_0 be the intersection of that trajectory with $F^{-1}(a)$. For any $m > 0$, we choose W_m to be the ball in $F^{-1}(a)$, with center w_0 and radius $1/m$, and V_m to be $\Xi_{[a,b]}W_m$, the set of points in $F^{-1}([a,b])$ that can be reached from W_m by the flow of ξ. Then $\bigcap_m V_m = \mathcal{K}_{\xi,a,b}(w)$, so for large m, by compactness, $V_m \subseteq V$.

For the rest of the proof, assume $r > 0$. Choose a sequence of open sets $V_m \subseteq F^{-1}[a,b]$ containing $p_1,\dots,p_r$ such that $V_m \supset V_{m+1}$ and $\bigcap V_m = \{p_1,\dots,p_r\}$. For instance, one could take for V_m the union of open balls of radius $1/m$ centred at the p_i in some metric. For each V_m set

$$U_m = \bigcup_{z\in V_m} \mathcal{K}_{\xi,a,b}(z).$$

As w is connected to at least one critical point by a trajectory of ξ, $w \in U_m$ for all m. From the construction of U_m, we also have $U_m \supseteq U_{m+1}$. We claim that U_m is open, and that $\bigcap U_m = \mathcal{K}_{\xi,a,b}(w)$. Before we prove openness, we make an observation.

Lemma 2.7.11. *For m sufficiently large, U_m does not contain any critical point other than $p_1,\dots,p_r$.*

Proof. Suppose for contradiction that $x_m \in U_m$ is a critical point of ξ different than $p_1,\dots,p_r$. Suppose a trajectory of ξ from (or to) x_m hits a point $y_m \in V_m$. In fact, x_m could be connected to yet another critical point of ξ, in that case, we choose x_m to be a point connected by a trajectory (not a zigzag trajectory) to a point in V_m. On passing to a subsequence, the points y_m converge to a point y, which is necessarily one of the points $p_1,\dots,p_r$. On the other hand, the points x_m are critical points of ξ different than $p_1,\dots,p_r$. The number of critical points is finite, so the sequence x_m, up to passing to a subsequence, is constant. Call it x. By the Broken Trajectory Lemma 2.7.14, there is a broken trajectory from x to y. But this means that there is a zigzag trajectory from w to x, contradicting the assumption that $p_1,\dots,p_r$ were all of the critical points in $\mathcal{K}_{\xi,a,b}(w)$. □

Next, we prove openness.

Lemma 2.7.12. *For sufficiently large m, the subset U_m is open.*

Proof. Let $x \in U_m$. We want to show that an open subset of x is contained in U_m. The easy case is that x is a critical point of ξ. Then x is one of the $p_1,\dots,p_r$ (by Lemma 2.7.11), and the open subset V_m, which contains x, is in U_m.

Suppose x is not a critical point of ξ. Then x is connected by a zigzag trajectory to a point $v \in V_m$. If this trajectory is a zigzag trajectory, and not just a trajectory, choose a part of it γ that passes through x and hits a critical point, say p_1. Either way, γ must enter V_m, so γ connects x to some point $v \in V_m$. This means that there is a non-critical point $v \in V$ and a trajectory of ξ connecting x to v. The time to travel from x to v is finite, so by continuous dependence of solutions on the initial conditions, any point sufficiently close to x can be connected to a point in V_m by a trajectory of ξ, and hence also lies in U_m. Hence U_m is open, as desired. □

Lemma 2.7.13. *We have that $\bigcap U_m = \mathcal{K}_{\xi,a,b}(w)$.*

Proof. The inclusion $\mathcal{K}_{\xi,a,b}(w) \subseteq U_m$ follows from the definition of U_m. Therefore, we have to prove that $\bigcap U_m \subseteq \mathcal{K}_{\xi,a,b}(w)$.

To this end, take $w' \in \bigcap U_m$. By definition, it is connected by a zigzag trajectory to a sequence of points $v_m \in V_m$. Arguing as in the proof of Lemma 2.7.12, we show that w' is connected by an actual trajectory of ξ with a point v_m. The points v_m have a converging subsequence, let v be the limit. It is necessarily one of the critical points $p_1, \dots, p_r$. The trajectories connecting w' to v_m limit into a broken trajectory connecting w' to v. This means that w' is connected to one of the points $p_1, \dots, p_r$. But these points belong to $\mathcal{K}_{\xi,a,b}(w)$. Hence w' also belongs to $\mathcal{K}_{\xi,a,b}(w)$. We have shown that $\bigcap U_m = \mathcal{K}_{\xi,a,b}(w)$. □

Continuing the proof of Lemma 2.7.10, we aim to deduce that one of the U_m must be contained in V. This follows from compactness of $\mathcal{K}_{\xi,a,b}$ by the following classical argument. Consider the cover of $F^{-1}[a,b]$ by V and $X_m := F^{-1}[a,b] \setminus \overline{U}_m$. This is an open cover of a compact set, so it has a finite subcover. The subcover must contain some sets $X_{m_1}, \dots, X_{m_\ell}$ and possibly V. Suppose $m_1 < \dots < m_\ell$. Then $X_{m_1} \subseteq \dots \subseteq X_{m_\ell}$, therefore we may assume that the cover contains only one set X_m. It necessarily contains V, then. That is to say, for some m, $X_m \cup V = F^{-1}[a,b]$. Suppose $y \in F^{-1}[a,b]$ is such that $y \in \overline{U}_m$ but $y \notin V$. Then $y \notin X_m \cup V$, which is a contradiction. This means that $\overline{U}_m \subseteq V$, so we just take $V' = U_m$. This completes the proof of Lemma 2.7.10. □

2.7.3. **The Broken Trajectory Lemma.** The notion of a broken trajectory is connected with the slogan that the closure of the space of trajectories is obtained by adding the broken trajectories. For gradient-like vector fields of Morse functions, the result is standard; see [Sal90]. As we aim to use the result for the proof of Vector Field Integration Lemma below, we do not use Morse functions, only vector fields. The assumptions are therefore somewhat more complicated than usual.

Lemma 2.7.14 (Broken Trajectory). *Assume ξ is a vector field on Ω, and Ω is presented as a disjoint union of three sets $\Omega_- \cup \Omega_0 \cup \Omega_+$, where Ω_- and Ω_+ are open, and Ω_0 is compact. Suppose that*

(BT-1) ξ has finitely many critical points in Ω_0 and they are all isolated. Moreover, there are no critical points on $\overline{\Omega}_0 \setminus \Omega_0$;
(BT-2) There is a pairwise disjoint family of subsets $\{X_p\}$, with $X_p \subseteq \Omega_0$, of pairwise disjoint regular index triples for all the critical points $p \in \Omega_0$;
(BT-3) Any trajectory of ξ through a point $z \in \Omega_0$ has either a forward limit at a critical point $p \in \Omega_0$, or it leaves in finite time to Ω_+;
(BT-4) Any trajectory of ξ through a point $z \in \Omega_0$ has either a backward limit in a critical point $p \in \Omega_0$, or it leaves in finite time in the past to Ω_-;
(BT-5) The set Ω_+ (respectively Ω_-) is forward invariant, respectively backward invariant. That is, a trajectory staying in Ω_+, respectively Ω_-, stays forever in the future in Ω_+, respectively forever in the past in Ω_-..

Suppose $\{z_\ell\}$ and $\{w_\ell\}$ are sequences of points in Ω_0 and that for every ℓ there is a trajectory of ξ passing through z_ℓ and then through w_ℓ. Assume $\lim_{\ell\to\infty} z_\ell = z$ and $\lim_{\ell\to\infty} w_\ell = w$ and $z, w \in \operatorname{Int}\Omega$. Then either there exists a (possibly broken) trajectory of ξ passing through z and then through w, or a (non-constant) broken trajectory connecting a critical point of ξ in Ω_0 with itself.

Proof. Assume first that z, w are not critical points of ξ. The case that at least one of z, w is a critical point requires only minor adjustments to the argument below, and is left to the reader.

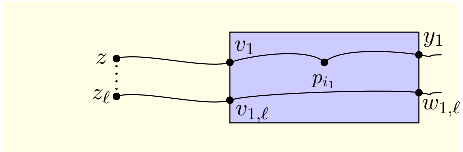

FIGURE 11. Notation of the proof of the Broken Trajectory Lemma 2.7.14.

By (BT-1), we let $p_1, \dots, p_m$ be critical points of ξ in Ω_0. Choose a pairwise disjoint regular index triples $(X_1, \partial_{\text{out}}X_1, \partial_{\text{in}}X_1)$, ..., $(X_m, \partial_{\text{out}}X_m, \partial_{\text{in}}X_m)$ for $p_1, \dots, p_m$, which exist by (BT-2). Note that each of $X_1, \dots, X_m$ is open and has compact closure. We assume also that z, w are disjoint from the closures of $X_1, \dots, X_m$; in particular, so are the points z_ℓ and w_ℓ.

We consider two simple cases.

First, we let T_ℓ be the time it takes to go from z_ℓ to w_ℓ. If T_ℓ has a bounded subsequence, on passing to a subsequence we have that $T_\ell \to T$ for some $T \in \mathbb{R}$. By continuous dependence of solutions on the initial conditions, we conclude that the trajectory of ξ starting from z reaches w after time T. This concludes the proof.

The second simple case is that the trajectory through z hits Ω_+. Then nearby trajectories also hit Ω_+ in finite time. Once a trajectory enters Ω_+, it stays there forever by (BT-5). Thus, we get a contradiction with contradicting the fact that the trajectories connect z_ℓ to w_ℓ.

The last case is that $T \to \infty$, and the trajectory through z stays forever in Ω_0. By (BT-3), it has to hit a critical point, say p_{i_1}. That is to say, at some moment, say t_1, it enters X_{i_1} through, say $v_1 \in \partial_{\text{in}}X_{i_1}$ and stays in X_{i_1} forever on. Note that the trajectory hits the interior of X_{i_1} after finite time. That is to say, there is a sequence of points $v_{1,\ell} \in \partial_{\text{in}}X_{i_1}$, converging to v_1, such that the trajectory through z_ℓ hits $v_{1,\ell}$.

We claim that the trajectory through $v_{1,\ell}$ hits $\partial_{\text{out}}X_{i_1}$ for ℓ sufficiently large. Otherwise, this trajectory stays forever in X_{i_1}, hence $w_\ell \in X_{i_1}$, so $w \in \overline{X}_{i_1}$, which is a contradiction. Let $w_{1,\ell} \in \partial_{\text{out}}X_{i_1}$ be the point of the first exit of the trajectory of ξ through $v_{1,\ell}$. Up to passing to a subsequence, we may and will assume that $w_{1,\ell}$ converge to a point y_1.

The backward limit of the trajectory through y_1 has to be p_{i_1}. Otherwise, the trajectory through y_1 exits (in the past) X_{i_1} through a point v_1'. As y_1 is the limit of $w_{1,\ell}$, and ξ connects $v_{1,\ell}$ to $w_{1,\ell}$, we infer that v_1' is the limit of $v_{1,\ell}$. But the limit of $v_{1,\ell}$ was v_1, the trajectory through v_1 hits p_{i_1} (does not reach y_1, in particular); the contradiction shows that the backward limit of y_1 is p_{i_1}. Compare Figure 11.

We start the inductive process, by looking at the forward behaviour of the trajectory through y_1. To see this, recall that $w_{1,\ell}$ converges to y_1. The trajectory through $w_{1,\ell}$ hits z_ℓ in the past by construction, so hits w_ℓ in the future. If the time required to reach w_ℓ from $w_{1,\ell}$ is bounded, then the trajectory through y_1 reaches w in finite time. We declare the limit trajectory to be the union of a trajectory from z through v_1 to p_{i_1}, and the trajectory from p_{i_1} through y_1 to w.

If the time is infinite, we repeat the procedure, i.e. we find the index i_2 such that the trajectory through y_1 hits p_{i_2}. We assume that v_2 is the point of last entry of this trajectory through y_1, and we let $v_{2,\ell} \in \partial_{\text{in}}X_{i_2}$ be the sequence of points on the trajectory through $w_{1,\ell}$ converging to v_2. We let $w_{2,\ell} \in \partial_{\text{out}}X_{i_2}$ be the points on the same trajectory, when they exit X_{i_2}, and y_2 be the limit of a subsequence of $w_{2,\ell}$. As above, we show that the backward trajectory through y_2 hits p_{i_2}.

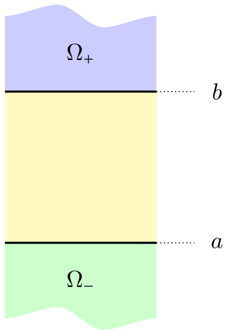

FIGURE 12. Notation of Vector Field Integration Theorem 2.8.1.

We repeat the process to obtain a sequence of indices $i_1, \ldots,$, points $v_1, v_2, \dots \in \partial_{\text{in}} X_{i_1}, \partial_{\text{in}} X_{i_2}, \dots$ and points $y_1, \dots \in \partial_{\text{out}} X_{i_1}, \partial_{\text{out}} X_{i_2}, \dots$. The point y_i is the limit of points $w_{i,\ell}$ as $\ell \to \infty$, and the trajectory through the original point z_ℓ passes through all the points $w_{1,\ell}, w_{2,\ell}, \dots$.

There are two possibilities: either the process stops after a finite time, that is, the trajectory through the point y_m reaches w in finite time. Then there is a trajectory from z through v_1 to p_{i_1}, from p_{i_1} through y_1 and v_2 to p_{i_2}, and so on, and the trajectory from p_{i_m} through y_m to w. This is the desired broken trajectory.

The other option is that the process does not stop. As the number of singular points is finite, there have to be repetitions among the indices $i_1, i_2, \dots$. Suppose $i_s = i_{s'}$ and $s' > s$. Then there is a broken trajectory that:

- starts at p_{i_s} and leaves X_{i_s} through y_s;
- enters $X_{i_{s+1}}$ through v_{s+1} and hits $p_{i_{s+1}}$;
- resumes from $p_{i_{s+1}}$ and leaves $X_{i_{s+1}}$ through y_{s+1};
- continues this behaviour until it enters $X_{i_{s'}}$ through $v_{s'}$ and hits $p_{i_{s'}}$.

This means that we have created a circular trajectory, as in the last clause of the lemma's statement. □

2.8. INTEGRATING GRIM VECTOR FIELDS

The purpose of this section is to prove the following result, called the Vector Field Integration Theorem. It says that under some natural assumptions, a vector field can be turned into a Morse function. In the upcoming theorem and proof, $\Omega_\pm$ are open subsets, and $\partial\Omega_\pm$ denotes the point-set boundary, or frontier, namely the closure minus the interior.

Theorem 2.8.1 (Vector Field Integration)**.** *Assume there are two (possibly empty) disjoint open subsets Ω_- and Ω_+ of Ω and suppose $F\colon \overline{\Omega}_- \sqcup \overline{\Omega}_+ \to \mathbb{R}$ is an immersed Morse function such that $F|_{\partial\Omega_-} = a$ and $F_{\partial\Omega_+} = b$ for some real numbers $a < b$. Moreover suppose F restricts to $F\colon \Omega_- \to (-\infty, a)$ and $F\colon \Omega_+ \to (b, \infty)$; see Figure 12. Suppose ξ is a vector field on Ω that is a grim vector field for F on $\Omega_- \sqcup \Omega_+$ satisfying the following conditions (reminiscent of (G-2) and (G-3) of Definition 2.5.1):*

- *(compare (G-2)) If $w \in \Omega[d]$, then $\xi(w) \in T_w\Omega[d]$. That is, ξ is tangent to all the strata.*
- *(compare (G-3)) If $p \in \Omega[d]$ is a critical point of ξ, then there are local coordinates centred at p*

$$(x_1, \dots, x_m, y_{11}, \dots, y_{1k}, \dots, y_{d1}, \dots, y_{dk})$$

(with $m = n + k - kd$), such that $\Omega[d]$ is given locally by the intersection of d strata:

$$\{y_{11} = \cdots = y_{1k} = 0\} \cap \ldots \cap \{y_{d1} = \cdots = y_{dk} = 0\},$$

such that ξ has the form

$$\xi = (-x_1, \ldots, -x_h, x_{h+1}, \ldots, x_m, \sum_{j=1}^{k} y_{1i}^2, 0, \ldots, 0, \sum_{j=1}^{k} y_{2i}^2, 0, \ldots, 0, \sum_{j=1}^{k} y_{di}^2, 0, \ldots, 0).$$

Suppose additionally that ξ satisfies the three conditions below.

- *(G-4) For each trajectory $\gamma\colon\mathbb{R} \to \Omega$ of ξ, either the limit $\lim_{s\to\infty} \gamma(s)$ (respectively $\lim_{s\to-\infty} \gamma(s)$) exists and belongs to $\Omega \smallsetminus \overline{\Omega_- \cup \Omega_+}$, or γ enters Ω_+ in the future (respectively, enters Ω_- in the past) and stays there forever.*
- *(G-5) There are no broken circular trajectories, that is there is no collection of trajectories $\{\gamma_1, \ldots, \gamma_{n+1}\}$ with $\gamma_{n+1} = \gamma_1$ such that $\lim_{s\to\infty} \gamma_i(s) = \lim_{s\to-\infty} \gamma_{i+1}(s)$ for $i = 1, \ldots, n$.*
- *(G-6) For each point $w \in \partial\Omega_-$ (respectively $w \in \partial\Omega_+$) the vector field ξ points out of Ω_- (respectively into Ω_+). In particular, there are no critical points on $\partial\Omega_- \cup \partial_+\Omega$.*

Then $F\colon\overline{\Omega}_- \sqcup \overline{\Omega}_+ \to \mathbb{R}$ extends to an immersed Morse function $F\colon\Omega \to \mathbb{R}$ such that ξ is a grim vector field for F.

This generalises the vector field integration lemma in [BP16, Section 3.1]. We remark that there was a mistake in that proof, because the function constructed in [BP16] was not necessarily continuous: it was smooth along trajectories, but could have jumps at points such that a trajectory through that point hits the boundary of V_i. This is a rather technical mistake and can be fixed by carefully adding a smoothing function. We provide a new proof here, both extending to the immersed case and fixing the previous error.

Proof. If there are no critical points of ξ in $\Omega\smallsetminus(\Omega_-\cup\Omega_+)$, the definition of F is straightforward and follows [Mil65, Proof of Assertion 5, page 54]. Namely, take $w \in \Omega \smallsetminus (\Omega_- \cup \Omega_+)$. By (G-4) the trajectory of ξ through w enters Ω_+ in the future and Ω_- in the past. Let T_+ be the time it takes to go from w to Ω_+ and T_- the time it takes to go from Ω_- to w. We set

$$F(w) = a + \frac{T_-}{T_- + T_+}(b - a).$$

The function F is smooth except possibly at $\partial\Omega_-$ and $\partial\Omega_+$, but using a standard argument (similar to the argument used in [Mil65, Proof of Assertion 5, page 54]) we can make it smooth on Ω.

Next, suppose there are critical points of ξ in $\Omega \smallsetminus (\Omega_+ \cup \Omega_-)$. Choose a partial order of all the critical points of ξ in $\Omega \smallsetminus (\Omega_+ \cup \Omega_-)$ by requiring that $p \le p'$ if there exists a (possibly broken) trajectory of ξ starting from p and ending up in p'. The condition (G-5) says $p \le p'$ is indeed a partial order. Every partial order has an associated strict partial order, where we write $p < p'$ if $p \le p'$ and $p \ne p'$.

We will first construct a function F near critical points and then use the flow of ξ to extend it over the whole of Ω. Define a function F first on critical points of ξ. The only condition on the values of F are that $F(p) \le F(p')$ whenever $p \le p'$ and $a < F(p) < b$. Next, each critical point p of ξ has a neighbourhood in which the vector field has the form (2.5.1). We choose pairwise disjoint neighbourhoods U_p and impose the form of F on U_p to be given by (2.4.1). Shrink U_p in such a way that if $p < p'$, then $\sup_{w\in U_p} F(w) < \inf_{w'\in U_{p'}} F(w')$.

Having defined F on $\bigcup U_p$, we may speak of grim neighbourhoods of p (as long as they are in U_p).

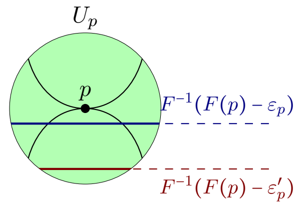

FIGURE 13. Choice of ε_p in the proof of Lemma 2.8.2. While ε_p satisfies the conditions spelled out in the proof, ε'_p is too large.

Lemma 2.8.2. *Suppose p is a critical point of ξ. There exists a grim neighbourhood $V_p \subseteq U_p$ such that if a trajectory of ξ leaves V_p then it never returns to V_p again.*

Proof. A grim neighbourhood V_p exists by Lemma 2.7.9 applied to $F|_{U_p}$. We shrink it until it satisfies the required conditions. First, fix $\varepsilon_p > 0$ to be such that $F^{-1}(F(p)-\varepsilon_p) \cap \mathbb{M}_{\mathrm{d}}(p)$ and $F^{-1}(F(p)+\varepsilon_p) \cap \mathbb{M}_{\mathrm{a}}(p)$ are separated from the boundary of U_p and each trajectory of ξ that hits p in the future (respectively in the past) hits $F^{-1}(F(p)-\varepsilon_p) \cap \mathbb{M}_{\mathrm{d}}(p)$ (respectively hits $F^{-1}(F(p)+\varepsilon_p)$); see Figure 13. This condition is nontrivial, because for the moment F is only defined inside U_p. With this choice of ε_p we let $\mathcal{V}_p$ be the family of all grim neighbourhoods of p whose 'in' boundary is on $F^{-1}(F(p)-\varepsilon_p)$ and whose 'out' boundary is on $F^{-1}(F(p)+\varepsilon_p)$. Note that $\bigcap_{V_p \in \mathcal{V}_p} V_p$ is the union of the ascending and descending membrane of p intersected with $F^{-1}(F(p)-\varepsilon_p, F(p)+\varepsilon_p)$. Observe that if $V, V' \in \mathcal{V}_p$, then $V \cap V' \in \mathcal{V}_p$. As Ω is second countable, a standard argument shows that there exists a nested family $V_1 \supset V_2 \supset \dots$ of grim neighbourhoods in $\mathcal{V}_p$ such that $\bigcap_\ell V_\ell = (\mathbb{M}_{\mathrm{a}}(p) \cup \mathbb{M}_{\mathrm{d}}(p)) \cap F^{-1}(F(p)-\varepsilon_p, F(p)+\varepsilon_p)$

Choose such a nested family. We will show that in this nested family, there is a grim neighbourhood V_p such that if a trajectory of ξ leaves V_p then it never returns to V_p again. Suppose for a contradiction that for every $\ell = 1, 2, \dots,$ there exists a trajectory of ξ leaving V_ℓ and entering V_ℓ again. Let w_ℓ be the first exit point and let z_ℓ be the next entrance point; see Figure 14. We have that $w_\ell \in F^{-1}(F(p)+\varepsilon_p)$ and $z_\ell \in F^{-1}(F(p)-\varepsilon_p)$. By passing to a subsequence we may and will assume that $z_\ell \to z$ and $w_\ell \to w$ for some z, w i.e. the subsequences converge and their limits are z and w respectively. Now by assumption $z \in \mathbb{M}_{\mathrm{d}}(p)$ and $w \in \mathbb{M}_{\mathrm{a}}(p)$. In particular, there is a broken trajectory of ξ starting at z and ending at w. On the other hand, by the Broken Trajectory Lemma 2.7.14 there exists a possibly broken trajectory of ξ starting at w and ending at z, or a circular trajectory. In the second case we get an immediate contradiction with (G-5). For the first case, taking the union of these two broken trajectories we obtain a circular broken trajectory, contradicting (G-5). This completes the proof of the lemma. □

As a next step towards the proof of the Vector Field Integration Theorem 2.8.1, we prove the following result.

Lemma 2.8.3. *Suppose $p < p'$. Then, on shrinking neighbourhoods V_p and $V_{p'}$, we may assume that no trajectory starting from $V_{p'}$ hits V_p in the future.*

Proof. Suppose such neighbourhoods do not exist. Acting as in the proof of Lemma 2.8.2, we create a trajectory starting at p' and hitting p in the infinite future (the reader might imagine that the points w_i of the proof of Lemma 2.8.2 belong $\partial_{\mathrm{out}} V_{p'}$, while points z_j belong to $\partial_{\mathrm{in}} V_p$). As $p < p'$, there exists a possibly broken trajectory from p to p'. Connecting these two broken trajectories, we obtain a circular broken trajectory, contradicting (G-5). □

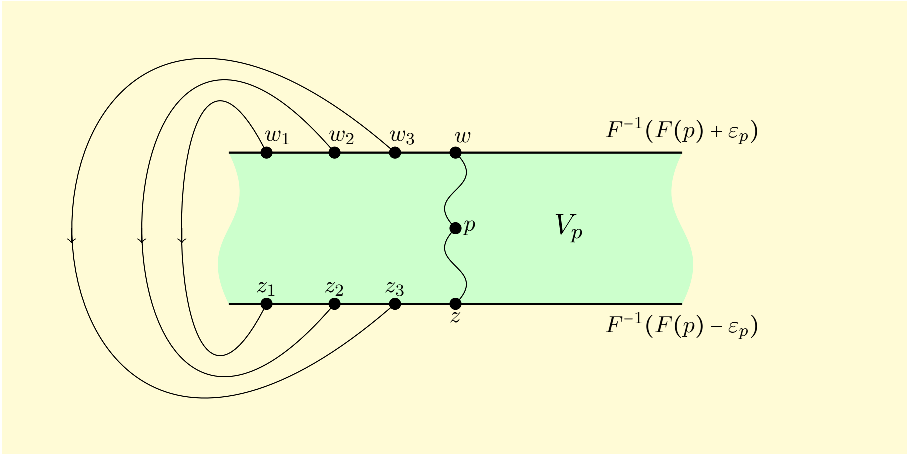

FIGURE 14. Existence of points z_ℓ, w_ℓ and their limit in the proof of Lemma 2.8.2.

Continuing the proof of the Vector Field Integration Theorem 2.8.1, we enumerate the critical points of F by $p_1, \dots, p_S$ and denote by c_i^- and c_i^+ the minimum and maximum of F on V_{p_i}. We have $c_i^- < c_i^+ \in \mathbb{R}$. We may and will assume that the critical points are ordered in such a way that $c_i^+ < c_{i+1}^- \in \mathbb{R}$ for all i.

We need to set up some more terminology. We will say that F is defined *up to the level set* a' if there is a set $\Omega'_- \subseteq \Omega$ such that F is defined on Ω'_-, $\partial\Omega'_- = F^{-1}(a')$ is the exit set of Ω'_-, and ξ is grim for F on Ω'_-. If F is defined up to a level set a' such that ξ has no critical points on $\Omega \smallsetminus (\Omega_+ \cup \Omega_-)$, we can extend F to the whole of Ω using the argument from the beginning of the proof.

Suppose F is defined up to the level set a', where $a' \in [c_i^+, c_{i+1}^-]$. This assumption means in particular that $V_{p_i} \subseteq \Omega'_-$. We have two cases. Either $a' = c_{i+1}^-$ or $a' < c_{i+1}^-$. Assume first $a' < c_{i+1}^-$. Let $A = F^{-1}(a')$. Define subsets $A_{i+1}, \dots, A_S \subseteq A$ by the condition that a trajectory of ξ starting from A_ℓ hits $\overline{V}_{p_\ell}$ before hitting other neighbourhoods. Let C_ℓ be the union of the trajectories of ξ starting from A_ℓ and before they hit $\overline{V}_{z_\ell}$; see Figure 15. Also let $T_\ell \colon A_\ell \to (0, \infty)$ be the function measuring the shortest time taken to get from the given point in A_ℓ to some point in V_ℓ.

First, observe that A_{i+1} is closed, because a trajectory starting from A and hitting $\overline{V}_{i+1}$ does not hit any other critical points by the ordering of critical points. In particular C_{i+1} is closed. Rescale the vector field ξ by a function supported in a neighbourhood of C_{i+1} (but disjoint from A and V_{i+1}) in such a way that T_{i+1} is constant on A_{i+1} and equal to $c_{i+1}^- - a$. Redefine the functions T_ℓ, for each $\ell \geq i+1$, with respect to the new vector field.

Our goal is to define the value $F(w)$ for a point w using the time required to go from A to w. The problem with doing precisely this is that some neighbourhoods $\overline{V}_p$ may be reached too early. To remedy this we need to slow down the vector field ξ along some trajectories. To this end, for $\ell > i+1$ we let A_ℓ^{bad} be the set of points $w \in A_\ell$ such that $T_\ell(w) \leq c_{i+1}^+$. Suppose ℓ_0 is the smallest index for which A_ℓ^{bad} is nonempty. Note that the closure of A_ℓ^{bad} is disjoint from the union $A_{i+1} \cup \dots \cup A_{\ell_0 - 1}$ (this union is a closed set). Define the cylinder C_ℓ^{bad} to be the union of all trajectories of ξ starting from A_ℓ^{bad} before they reach $\overline{V}_\ell$. Rescale ξ in the cylinder C_ℓ^{bad} in such a way that ξ runs slower in C_ℓ^{bad} and so making A_ℓ^{bad} empty. By induction we arrive at the situation in which all the sets A_ℓ^{bad} are empty.

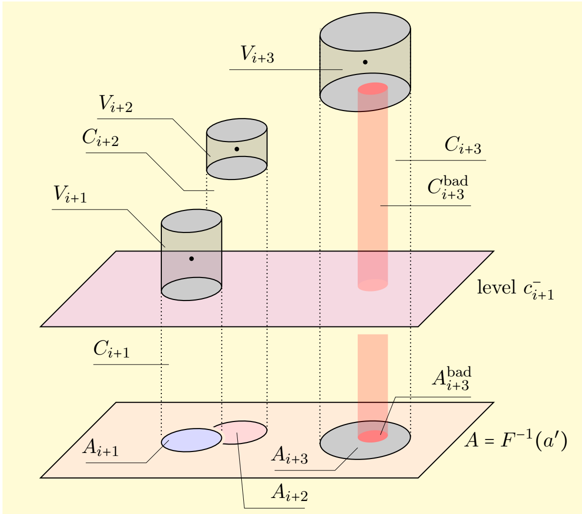

FIGURE 15. Proof of the Vector Field Integration Theorem 2.8.1.

Now let Ω_{i+1} be the set of points reached by ξ in time less than or equal to $c^-_{i+1} - a'$. By construction, ξ is nonvanishing on Ω_{i+1}. For a point $w \in \Omega_{i+1}$ that is reached from A in time t_w we set

$$F(w) = a' + t_w.$$

It is clear that F is a continuous function, and it can be made smooth near A by carefully rescaling the vector field ξ. In this way, we have defined F up to the level set c^-_{i+1}.

Now suppose that F is defined up to the level set $a' = c^-_{i+1}$. Define sets A_ℓ, C_ℓ analogously, as well as the function T_ℓ. Alter the vector field ξ near the boundary of $V_{p_{i+1}}$ to guarantee that $\partial_\xi F = 1$ near the boundary of $V_{p_{i+1}}$. This is a technical step, which simplifies the construction later.

For $\ell = i+2, \dots$ we define A^{bad}_ℓ as the set of points $w \in A_\ell$ such that $T_\ell(w) \le c^+_{i+1} - c^-_{i+1}$ and we perform an analogous rescaling as above to make A^{bad}_ℓ empty.

Define Ω^0_{i+1} to be the union of the set of points in Ω reached from $A \smallsetminus V_{p_{i+1}}$ by ξ in time less or equal to $c^+_{i+1} - c^-_{i+1}$. For a point $w \in \Omega_{i+1}$ that is reached from $A \smallsetminus V_{p_{i+1}}$ in time t_w we set

$$F(w) = a' + t_w.$$

As $\partial_\xi F = 1$ near the boundary of $V_{p_{i+1}}$, the function F is well-defined and continuous, and we can make it smooth by carefully altering F near $\partial V_{p_{i+1}}$. An argument, following the ideas presented in [Mil65, Proof of Assertion 5, page 54], is presented as Lemma 2.8.4 below. In this way we have defined F up to the level set c^+_{i+1}.

The above procedure in the proof of Vector Field Integration Theorem 2.8.1 was to increase gradually the level sets up to which F is defined. This was the key technical difficulty of the proof. At the final stage of the proof, we need to specify the starting level set. If Ω_- is nonempty, the assumptions of the proposition tell us that F is defined up to the level set a. If Ω_- is empty, then the critical point p_1 is necessarily an index 0 critical point. We set $\Omega_- = V_{p_1}$, and then F is well-defined up to the level set c^+_0.

A small technical problem arises: the extension we construct has not been guaranteed yet to be smooth near the level sets of F corresponding to the values $c^\pm_i$: it is a priori only continuous and piecewise smooth. The proof of the Vector Field Integration Theorem 2.8.1 is completed using the following general result, which upgrades the continuous function we have just defined to a smooth function.

Lemma 2.8.4. *Assume that $F\colon\Omega\to\mathbb{R}$ is a continuous function with a smooth level set $Z:=F^{-1}(c)$. Set $\Omega_-=F^{-1}(-\infty,c]$ and $\Omega_+=F^{-1}[c,\infty)$. Suppose F is smooth on Ω_- and Ω_+ and has no critical points near Ω_-.*

Assume that ξ is a smooth vector field such that $\partial_\xi F>0$ in a neighbourhood of Z in Ω_- and $\partial_\xi F>0$ in a neighbourhood of Z in Ω_+.

Then, for any $\varepsilon>0$, there exists a smooth function $F_\varepsilon\colon\Omega\to\mathbb{R}$ such that $\partial_\xi F_\varepsilon>0$, $\|F-F_\varepsilon\|<\varepsilon$, and such that we have $F(z)=F_\varepsilon(z)$ whenever $|F(z)-c|>\varepsilon$.

Applying Lemma 2.8.4 to the function F constructed above with c ranging over $c_i^\pm$ and ε choosen in such a way that the intervals $[c_i^\bullet-\varepsilon,c_i^\bullet+\varepsilon]$, $\bullet=\pm$, do not overlap, completes the proof of the Vector Field Integration Theorem 2.8.1. □

Sketch of proof of Lemma 2.8.4. Choose $\delta>0$ sufficiently small, so that F has no critical points on $F^{-1}[c-\delta,c+\delta]$. Let $Z_-=F^{-1}(c-\delta)$, $Z_+=F^{-1}(c+\delta)$. From the assumptions we deduce that the flow of ξ flows from Z_- to Z_+.

Multiply the vector field ξ by a smooth positive function in such a way that $\partial_\xi F\equiv1$ near Z_- and Z_+ and for any point $z\in Z_-$ the time to reach Z_+ is equal to 2δ. We set

$$F_\delta(z)=\begin{cases}F(z) & :F(z)\notin[c-\delta,c+\delta]\\ c-\delta+t(z) & :F(z)\in[c-\delta,c+\delta],\end{cases}$$

where $t(z)$ is the time to reach z starting from the unique point $z_-\in Z_-$ such that the trajectory of ξ through z_- hits z.

The function F is smooth on $F^{-1}(c-\delta,c+\delta)$. As the time to reach Z_+ from Z_+ is equal to 2δ, we infer that F_δ is continuous on $F^{-1}(c\pm\delta)$, in particular, it is continuous everywhere. Next, by construction, $\partial_\xi F_\delta\equiv1$ on $F^{-1}[c-\delta,c+\delta]$. Since we assumed that $\partial_\xi F\equiv1$ near $Z_\pm$, the function F_δ agrees with F on a neighbourhood of $Z_\pm$. That is to say, F_δ is smooth.

Next, for any $z\in F^{-1}[c-\delta,c+\delta]$ we have $F_\delta(z),F(z)\in[c-\delta,c+\delta]$. Hence $|F_\delta(z)-F(z)|<2\delta$. The statement is proved by setting $\varepsilon=\delta/2$. □

Remark. During the proof of the Vector Field Integration Theorem 2.8.1 we were multiplying the vector field ξ by a positive function. Note that this does not affect the veracity of the final statement. If $\widehat{\xi}$ is the 'corrected' vector field, $\widehat{\xi}=\phi\xi$ for some $\phi\colon\Omega\to\mathbb{R}_{>0}$ such that $\phi\equiv1$ in a neighbourhood of each of the critical points of ξ, then $\widehat{\xi}$ is a grim vector field for F if and only if ξ is.

Part 3. Families of Morse functions.

Before we prove the Path Lifting Theorem 4.5.1, which is the main technical tool of the present paper, we need to investigate paths of Morse functions in detail. In this part, Part 3, we study one-parameter families of Morse functions. Classical theorems on rearrangement and cancellations of critical points of a Morse function are generalised to the case of immersed Morse functions and phrased in the language of a one-parameter family of functions (see Sections 3.2 and 3.4). In the case of cancellations, these function cease to be Morse functions for one value of the parameter. This motivates the study of one-parameter families in the spirit of Cerf. We recall his approach in Section 3.1.

Now we introduce a few conventions used throughout Part 3 and 4. We consider a smooth compact manifold N of dimension n is immersed into an ambient manifold Ω of dimension $n+k$. The immersion is usually denoted by G, possibly with some decorations (e.g. subscripts). Both N and Ω are allowed to have boundary, in which case we require that G maps the boundary, and only the boundary, to the boundary. Moreover, any Morse function is assumed to be locally constant on the boundary, while vector fields are transverse to the boundary.

Immersed Morse functions on Ω are denoted by capital F, whereas Morse functions on N are denoted by lowercase f. Usually we have $f = F \circ G$, but the precise formula can vary, and it is always specified explicitly. A grim vector field on Ω is denoted by ξ. A vector field on N is denoted by η.

A path of functions is usually denoted by F_τ. The moments on the path are usually denoted by Greek letters, like F_τ for $\tau \in [\alpha, \beta]$ as opposed to level sets of functions, which are denoted by Latin letters, like $F^{-1}[a, b]$.

The path F_τ is *supported* on a subset $U \subseteq \Omega$, if $F_\tau|_{\Omega \setminus U}$ does not depend on τ. The *concatenation* of paths F^1_τ and F^2_τ is denoted by $F^1_\tau \cdot F^2_\tau$.

In Part 3, we focus on immersed Morse functions on Ω, where $G\colon N \to \Omega$ is an immersion and N has all connected components of the same dimension. Necessary changes if this is not the case are given in Section 4.6.

3.1. Regular paths

The aim of this section is to introduce regular paths. We specify generic conditions that guarantee nice behaviour of a function or a map, as well as conditions on paths of objects in which the mildest possible degeneracies occur. We begin by recalling the stratification of the space $\mathcal{F} := C^\infty(N, \mathbb{R})$ in the spirit of [Cer70]. In particular, we introduce the notion of a *regular path of functions*, also called an $\mathcal{F}^1$-path.

Next, we use analogous methods to stratify the space of functions $C^\infty(\Omega, \mathbb{R})$ (with a fixed immersed submanifold in Ω) and of smooth maps $N \to \Omega$, in particular to study regular (in a suitable sense) paths of immersions. Next, we combine the latter two notions. First, in the presence of a function $F\colon \Omega \to \mathbb{R}$, we study the set of maps $G\colon N \to \Omega$, such that F is an immersed Morse function with respect to $M := G(N)$. Finally, we study pairs (F, G), where $F\colon \Omega \to \mathbb{R}$ and $G\colon N \to \Omega$. We specify what it means for a path of such pairs (F_τ, G_τ) to be regular. In Subsection 3.1.6, we discuss the necessary changes in the theory when Ω and N have boundary.

3.1.1. **Ambient Cerf theory.** Cerf's theory [Cer70] deals with paths of functions, that is with smooth maps $f_\bullet\colon [0,1] \to \mathcal{F}$, where

$$\mathcal{F} := C^\infty(N, \mathbb{R}).$$

Here N is either a smooth, closed manifold, or a smooth compact cobordism $(N; \mathrm{bd}_0 N, \mathrm{bd}_1 N)$ with a decomposition of the boundary $\mathrm{bd}\, N = \mathrm{bd}_0 N \sqcup \mathrm{bd}_1 N$. In the latter case we impose neatness conditions, specified in Subsection 3.1.6 below. Now we recall the parts of this theory that are needed in the proof of Path Lifting Theorem 4.5.1.

Recall that a Morse function is called *excellent* if the function mapping critical points to critical values is injective. The subspace of excellent Morse functions is open and dense in $\mathcal{F}$, which follows from Theorem 2.4.5 and an easy argument to see that the excellent condition is open and dense.

We will analyse the function space $\mathcal{F}$ with a special emphasis on paths in $\mathcal{F}$. The space $\mathcal{F}$ has a decomposition [Cer70, Mat69], of which we need only a part. We describe subsets $\mathcal{F}^0$, $\mathcal{F}^1_\alpha$, and $\mathcal{F}^1_\beta$ in $\mathcal{F}$.

- $\mathcal{F}^0$ consists of all excellent Morse functions, meaning that the critical values of distinct critical points are distinct.
- $\mathcal{F}^1_\alpha$ are functions which in addition to nondegenerate critical points have exactly one cubic singularity (that is, an A_2 singularity) with local normal form
$$x_1^3 - x_2^2 - \dots - x_j^2 + x_{j+1}^2 + \dots + x_n^2.$$
All critical values are assumed to be distinct.
- $\mathcal{F}^1_\beta$ consists of all Morse functions with exactly two critical points having the same value.

Let $\mathcal{F}^1 := \mathcal{F}^1_\alpha \cup \mathcal{F}^1_\beta$. Then $\mathcal{F}^0$ and $\mathcal{F}^0 \cup \mathcal{F}^1$ are open in $\mathcal{F}$ by [Cer70, I,2.1]. When we want to emphasise the domain we write $\mathcal{F}(N)$, $\mathcal{F}^0(N)$, and $\mathcal{F}^1(N)$. We make the following definition.

Definition 3.1.1. A path $f_\bullet\colon [0,1] \to \mathcal{F}^0 \cup \mathcal{F}^1$ is called an $\mathcal{F}^1$*-path* or a *regular path* if $f_\tau \in \mathcal{F}^1$ for only finitely many time parameters τ_i, $f_0, f_1 \in \mathcal{F}^0$, and if $f_\bullet$ intersects $\mathcal{F}^1$ transversely.

The next result follows from the fact that $\mathcal{F}^1$ is a codimension one subspace of $\mathcal{F}$; its proof can be found in [Cer70, II.3.2, III.1.3, III.2.3].

Lemma 3.1.2. *Each path with endpoints in $\mathcal{F}^0$ can be perturbed rel. boundary to an $\mathcal{F}^1$-path. In particular, if $f_0, f_1 \in \mathcal{F}^0$, then there is an $\mathcal{F}^1$-path f_τ, $\tau \in [0,1]$ connecting f_0 and f_1.*

The study of $\mathcal{F}^1$-paths can be divided into the study of paths that intersect $\mathcal{F}^1$ at precisely one time value.

Definition 3.1.3. An $\mathcal{F}^1$-path f_τ, $\tau \in [0,1]$ is called a *path of birth*, respectively a *path of death*, or a *path of rearrangement*, if $f_\tau \in \mathcal{F}^1$ at precisely one time value τ_0, and:

- *birth:* $f_{\tau_0} \in \mathcal{F}^1_\alpha$ and f_τ has two more critical points for $\tau > \tau_0$ than for $\tau < \tau_0$;
- *death:* $f_{\tau_0} \in \mathcal{F}^1_\alpha$ and f_τ has two fewer critical points for $\tau > \tau_0$ than for $\tau < \tau_0$;
- *rearrangement:* $f_{\tau_0} \in \mathcal{F}^1_\beta$.

There is a local normal form for each of these three paths. That is, any path of birth, death or rearrangement can be changed by homotopy keeping the endpoints fixed to a path which crosses the stratum $\mathcal{F}^1$ in a specific way. These forms are called *elementary paths of birth, death, or rearrangement*; see [Cer70, Part III]. The description of normal forms will be useful when discussing lifting of elementary paths. These normal forms will be discussed in Section 3.5 for the birth case, and Section 3.6 for the death case. For the case of paths of rearrangements, we will reformulate Cerf's method using the notion of ξ-paths in Section 3.3, because vector fields allow us greater control when constructing regular homotopies.

While phrasing and proving Path Lifting Theorem 4.5.1, it will be much easier (and enough for all applications so far) to lift a path up to suitably defined homotopy. This motivates the following definition.

Definition 3.1.4. Let f_τ and $\widetilde{f}_\tau$ be $\mathcal{F}^1$-paths.

(i) Suppose that $f_0 = \widetilde{f}_0$ and $f_1 = \widetilde{f}_1$. Then f_τ and $\widetilde{f}_\tau$ are $\mathcal{F}^1$*-homotopic* if they are homotopic rel. endpoints ($\tau = 0,1$) through $\mathcal{F}^1$-paths in $\mathcal{F}_0 \cup \mathcal{F}_1$.

(ii) We say that f_τ and $\widetilde{f}_\tau$ are *lax homotopic* if there is a homotopy $h_{\sigma,\tau}$ of paths such that for every σ the path $\tau \mapsto h_{\sigma,\tau}$ is an $\mathcal{F}^1$-path, $h_{0,\tau} = f_\tau$, $h_{1,\tau} = \widetilde{f}_\tau$ and the paths $\sigma \mapsto h_{\sigma,0}$, $\sigma \mapsto h_{\sigma,1}$ belong to $\mathcal{F}^0$.

(iii) Suppose $f_\tau, \widetilde{f}_\tau$ are $\mathcal{F}^1$-paths and g_σ, $\sigma \in [0,1]$ is a path in $\mathcal{F}^0$ (i.e. a path of excellent Morse functions) such that $g_0 = f_0$ and $g_1 = \widetilde{f}_0$. We say that f_τ and $\widetilde{f}_\tau$ are *lax-homotopic over* g_σ if there exists a lax homotopy $h_{\sigma,\tau}$ as in the previous item, with the additional property that $h_{\sigma,0} = g_\sigma$.

(iv) If f_τ and $\widetilde{f}_\tau$ are lax homotopic over a constant path $g_\sigma \equiv f_0 = \widetilde{f}_0$, then the paths f_τ and $\widetilde{f}_\tau$ are called *left-homotopic*.

Remark 3.1.5. Consider a pair of $\mathcal{F}^1$-paths f_τ and $\widetilde{f}_\tau$ with the same endpoints $f_0 = \widetilde{f}_0$ and $f_1 = \widetilde{f}_1$. Then f_τ and $\widetilde{f}_\tau$ are lax homotopic if and only if they are $\mathcal{F}^1$-homotopic.

We describe the proof of the harder, forward direction. If $f_\tau, \widetilde{f}_\tau$ are lax homotopic, and $h_{\sigma,\tau}$ is the corresponding homotopy, then replacing $\tau \mapsto h_{\sigma,\tau}$ by the concatenation of paths $\sigma' \mapsto h_{\sigma',0}$, $\sigma' \in [0,\sigma]$, $\tau \mapsto h_{\sigma,\tau}$ and the inverse of the path $\sigma' \mapsto h_{\sigma',1}$, $\sigma' \in [0,\sigma]$, yields an $\mathcal{F}^1$-path for all $\sigma \in [0,1]$, with endpoints f_0 and f_1. Hence it is straightforward to construct an $\mathcal{F}^1$-homotopy.

3.1.2. **Paths of immersions.** We now begin the discussion of the immersed case, which will stretch over the next few subsections. We will consider an n-dimensional compact manifold N and a compact manifold Ω. The aim of this subsection is to study singularities that may occur on a generic path of immersions from N to Ω. Most of the results are not new; we refer e.g. to [Ekh01, Section 3.2].

Definition 3.1.6. The space $\mathcal{I}(N,\Omega)$, in short $\mathcal{I}$, is by definition the space of C^∞ immersions from N to Ω. If N and Ω have boundary, we assume the immersion to be neat and generic on the boundary (Definition 3.1.32 below).

We do not discuss the case where N and Ω have boundary in detail until Subsection 3.1.6. Note that $\mathcal{I}(N,\Omega)$ is open in $C^\infty(N,\Omega)$; see [GG73, Lemma II.5.5]. Define

$$\mathcal{I}^0 = \mathcal{I}^0(N,\Omega)$$

to be the space of generic immersions (see Definition 2.2.1). By Lemma 2.2.3, $\mathcal{I}^0$ is open and dense in $\mathcal{I}$.

Our aim is to study one-parameter deformations of a generic immersion. Our approach is essentially that of Ekholm [Ekh01, Section 3], except that we do not have restrictions on dimensions: these restrictions are not needed for [Ekh01, Section 3], rather they are used in later sections of that work.

In the next definition we define a preliminary subspace of immersions, containing all immersions where genericity fails at depth d for at least one point. After the definition, we will proceed to consider the points where genericity fails more than the minimal possible amount.

Definition 3.1.7. Let $d \geq 2$. The space $\widetilde{\mathcal{I}}^1_d$ is by definition the subspace of immersions $G \in \mathcal{I}$ for which there exists a d-tuple $u_1, \dots, u_d$ of distinct points of N that are all mapped by the immersion to the same point $w := G(u_1) = \cdots = G(u_d)$, such that there are pairwise disjoint open subsets $U_1, \dots, U_d$ of N, with U_j containing u_j, with the properties:

(CI-1) $G(U_1), \dots, G(U_{d-1})$ intersect transversely at w (compare the beginning of Section 2.2);

(CI-2) with $M_1 = G(U_1) \cap \cdots \cap G(U_{d-1})$, $M_2 = G(U_d)$, $T_w M_1 + T_w M_2$ is a positive codimension subspace in $T_w\Omega$.

For $d \geq 2$ we will consider subspaces $\mathcal{I}^2_{\alpha,d}$, $\mathcal{I}^2_{\beta,d}$, and $\mathcal{I}^2_{\gamma,d,d'}$ of $\widetilde{\mathcal{I}}^1_d$, whose definition is given below. We will then set

$$\mathcal{I}^2_\alpha := \bigcup_d \mathcal{I}^2_{\alpha,d},\ \mathcal{I}^2_\beta = \bigcup_d \mathcal{I}^2_{\beta,d},\ \mathcal{I}^2_\gamma = \bigcup_{d,d'} \mathcal{I}^2_{\gamma,d,d'}.$$

Then with

$$\mathcal{I}^2 := \mathcal{I}^2_\alpha \cup \mathcal{I}^2_\beta \cup \mathcal{I}^2_\gamma,$$

we note that $\mathcal{I}^2 \subseteq \widetilde{\mathcal{I}}^1$ and define

$$\mathcal{I}^1 = \widetilde{\mathcal{I}}^1 \smallsetminus \mathcal{I}^2.$$

The idea of the definition of $\mathcal{I}^1$ is that we remove from $\widetilde{\mathcal{I}}^1$ all the immersions, namely those in $\mathcal{I}^2$ for which the failure of genericity is worse than the minimal possible failure, so that a generic path lies in $\mathcal{I}^1$.

Now we give the definitions of $\mathcal{I}^2_{\alpha,d}$, $\mathcal{I}^2_{\beta,d}$, and $\mathcal{I}^2_{\gamma,d,d'}$.

- (*extra branch passing through tangency*) The subspace $\mathcal{I}^2_{\alpha,d} \subseteq \widetilde{\mathcal{I}}^1_d$ consists of maps such that there exists a set of points $u_1,\dots,u_d$ such that $G(u_1) = \cdots = G(u_d)$, satisfying items (CI-1) and (CI-2) of Definition 3.1.7, but there is another point u_{d+1} mapped to $G(u_1)$.
- (*higher tangency*) The subspace $\mathcal{I}^2_{\beta,d} \subseteq \widetilde{\mathcal{I}}^1_d$ consists of maps for which there is a set of points $u_1,\dots,u_d$ satisfying items (CI-1), and (CI-2) of Definition 3.1.7, and $T_wM_1 + T_wM_2$ is of codimension at least 2 in $T_w\Omega$.
- (*two tangencies at the same time*) For $d' \geq 2$, the subspace $\mathcal{I}^2_{\gamma,d,d'} \subseteq \widetilde{\mathcal{I}}^1_d$ is the set of maps for which there is a set of points $u_1,\dots,u_d$ satisfying (CI-1) and (CI-2), as well as another set of points $u'_1,\dots,u'_{d'}$ with $G(u'_1) = \cdots = G(u'_{d'}) \neq G(u_1)$ satisfying (CI-1) and (CI-2).

The following result echoes [Ekh01, Lemma 3.4].

Proposition 3.1.8. *Suppose $G_\tau\colon N \looparrowright \Omega$, $\tau \in [0,1]$, is a path in $\mathcal{I}$ such that $G_0, G_1 \in \mathcal{I}^0$. Then there exists a path G'_τ, arbitrarily close (in the C^∞ topology) to G_τ, with $G'_0 = G_0$ and $G'_1 = G_1$, such that G'_τ lies in $\mathcal{I}^0$ for all but finitely many values of τ, and for these values we have that $G_\tau \in \mathcal{I}^1$.*

In the proof of Proposition 3.1.8, we will heavily use multijet transversality from Subsection 2.3.1. As a step towards Proposition 3.1.8, we give the following result.

Lemma 3.1.9. *Suppose $r > 0$, $i_1, i_2 \geq 1$ are such that $(i_1 - 1)k \leq n$ and $(i_2 - 1)k \leq n$. We consider the subspace of $J^1_{i_1+i_2}(N,\Omega)$ determined by the condition corresponding to maps $G\colon N \to \Omega$ such that there are pairwise distinct points $u_1,\dots,u_{i_1+i_2} \in N$ such that $G(u_1) = G(u_2) = \cdots = G(u_{i_1+i_2})$, and there exist open subsets $U_1,\dots,U_{i_1+i_2}$ of N, for which the images $G(U_1),\dots,G(U_{i_1})$ are transverse at $w := G(u_1)$, the images $G(U_{i_1+1}),\dots,G(U_{i_1+i_2})$ are transverse at w, and*

$$\dim(T_w(G(U_1) \cap \cdots \cap G(U_{i_1})) + T_w(G(U_{i_1+1}) \cap \cdots \cap G(U_{i_1+i_2}))) \leq \dim\Omega - r.$$

The subspace of the (i_1+i_2)-fold 1-multijet space $J^1_{i_1+i_2}(N,\Omega)$ has codimension

$$(i_1 + i_2 - 1)(n + k) + r(n - (i_1 + i_2 - 1)k + r)$$

in $J^1_{i_1+i_2}(N,\Omega)$.

Proof. The condition that $G(u_1) = G(u_2) = \cdots = G(u_{i_1+i_2})$ is of codimension $(i_1+i_2-1)\dim\Omega = (i_1+i_2-1)(n+k)$, which accounts for the first summand. The condition that $G(U_1), \dots, G(U_{i_1})$ are transverse at w is open. Define

$$M_1 \coloneqq G(U_1) \cap \cdots \cap G(U_{i_1}).$$

By transversality, near w, this is a manifold of dimension $s_1 \coloneqq n - (i_1 - 1)k$.

Analogously, transversality of $G(U_{i_1+1}), \dots, G(U_{i_1+i_2})$ is an open condition; we set

$$M_2 \coloneqq G(U_{i_1+1}) \cap \cdots \cap G(U_{i_1+i_2}).$$

This is a manifold (near w) of dimension $s_2 \coloneqq n - (i_2 - 1)k$. Set $V_1 = T_w M_1$ and $V_2 = T_w M_2$. The condition that $V_1 + V_2$ forms a codimension r subspace of $T_w\Omega$ can be computed as follows. Choose a basis $e_{i1}, \dots, e_{is_i}$ of V_i. We have two cases.

- if $\ell = n + k - r - s_1 \geq 0$, then we choose the vectors $e_{11}, \dots, e_{1s_1}, e_{21}, \dots, e_{2\ell}$ generically. They span a subspace of dimension $n+k-r$. The remaining $s_2-\ell$ vectors, $e_{2,\ell+1}, \dots e_{2s_2}$ are all compelled to stay the codimension r subspace, which gives $r(s_1+s_2-(n+k-r)) = r(n - (i_1 + i_2 - 1)k + r)$ conditions.
- if $\ell < 0$, we can choose only the vectors $e_{11}, \dots, e_{1,n+k-r}$ points generically so that they span a subspace of dimension $n+k-r$. There remains $s_1-(n+k-r)+s_2 = s_2-\ell$ vectors $e_{1,n+k-r+1}, \dots, e_{1s_1}$, $e_{21}, \dots, e_{2s_2}$, which are all forced to belong to the codimension r subspace. This gives the same amount of conditions as the first case.

This accounts for the second summand. Altogether, we have $(i_1 + i_2 - 1)(n + k) + r(n - (i_1 + i_2 - 1)k + r)$ conditions as desired. □

Proof of Proposition 3.1.8. Choose $d \geq 2$. If $kd > n + 1$, then the condition that there exist d points $u_1, \dots, u_d$ such that $G(u_1) = \cdots = G(u_d)$ is of codimension $(d-1)(n+k) > nd + 1$ in the multijet space $J^0_d(N, \Omega)$. Hence the stratum $\mathcal{I}^1_d$ is missed by a (one-dimensional) path of immersions by Corollary 2.3.5 (ii).

Hence, for the rest of the proof, we suppose $d \geq 2$ and $kd \leq n + 1$. The strategy is as follows. Every path of immersions can be assumed to lie in $\mathcal{I}^0 \cup \widetilde{\mathcal{I}}^1$. We then consider the codimension of the subspaces $\mathcal{I}^2_{\alpha,d}$, $\mathcal{I}^2_{\beta,d}$, and $\mathcal{I}^2_{\gamma,d,d'}$, and we check that the complements are each residual, so each subspace can be avoided by a generic path of immersions. Then we use that a countable intersection of residual sets (given by these three sets, for all values of d and d') is residual (Lemma 2.3.3).

Lemma 3.1.10. *The conditions on $\mathcal{I}^2_{\alpha,d}$ define a subspace of codimension equal to $n(d+1) + k + 1$ in $J^1_{d+1}(N, \Omega)$. In particular, $\mathcal{I}^2_{\alpha,d}$ is missed by a generic k-parameter family of immersions from N to Ω.*

Proof. By Lemma 3.1.9 with $r = 1$, $i_1 = d-1$, $i_2 = 1$, we get $dn+1$ conditions. The extra point u_{d+1} mapped to $G(u_1)$ gives $(n + k)$ more conditions. This gives $n(d+1) + k + 1$ conditions. These conditions involve $i_1 + i_2 + 1 = d + 1$ points of N, so we end up with a codimension $n(d+1)+k+1$ subspace of $J^1_{d+1}(N, \Omega)$. Since $\dim N^{(d+1)} = n(d+1)$, the space $\mathcal{I}^2_{\alpha,d}$ is avoided by a generic k-parameter family of immersions by Corollary 2.3.5 (iii). □

Lemma 3.1.11. *The conditions on $\mathcal{I}^2_{\beta,d}$ define a subspace of codimension greater than or equal to $nd+4$ in $J^1_d(N, \Omega)$. A generic 3-parameter family of immersions from N to Ω avoids $\mathcal{I}^2_{\beta,d}$.*

Proof. We apply Lemma 3.1.9 for $i_1 = d - 1$ and $i_2 = 1$. As $r \geq 2$, this gives codimension $nd + (r-1)(n - (d-1)k) + r^2 > nd + r^2 \geq nd + 4$ by Corollary 2.3.5 (iii). □

Next, we deal with the case of independent tangencies. For this, we assume that $d, d' \geq 2$ are such that both $kd \leq n+1$ and $kd' \leq n+1$.

Lemma 3.1.12. *The conditions on $\mathcal{I}^2_{\gamma,d,d'}$ define a subspace of codimension greater than $n(d+d')+2$ in $J^1_{d+d'}(N,\Omega)$. In particular, a generic one-parameter family avoids $\mathcal{I}^2_{\gamma,d,d'}$.*

Proof. Conditions (CI-1), and (CI-2) give $nd+1$ conditions involving points $u_1,\dots,u_d$ and, independently, $nd'+1$ conditions involving points $u'_1,\dots,u'_{d'}$. Altogether, we get $n(d+d')+2$ conditions in $J^1_{d+d'}(N,\Omega)$. We conclude by Corollary 2.3.5 (ii). □

We can now finish the proof of Proposition 3.1.8. Lemmas 3.1.10, 3.1.11, and 3.1.12 show that the sets of one-parameter families missing $\mathcal{I}^2_{\alpha,d}$, $\mathcal{I}^2_{\beta,d}$ and $\mathcal{I}^2_{\gamma,d,d'}$ form a residual set in the set of all smooth immersions from N to Ω, so a residual set in the space of all smooth immersions from N to Ω by openness of $\mathcal{I}(N,\Omega)$ in $C^\infty(N,\Omega)$. The intersection of finitely many residual sets is residual by Lemma 2.3.3, which gives the statement of Proposition 3.1.8. □

Definition 3.1.13. A path G_τ, $\tau \in [0,1]$ of immersions from N to Ω is called *regular* if $G_0, G_1 \in \mathcal{I}^0$ and $G_\tau \in \mathcal{I}^0$ for all but finitely many values of τ, and if whenever G_τ is not in $\mathcal{I}^0$, then it lies in $\mathcal{I}^1$.

Proposition 3.1.8 can be rephrased in the following way.

Corollary 3.1.14. *Every path G_τ with endpoints generic immersions can be perturbed rel. endpoints to a regular path of immersions.*

We conclude this subsection with the following observation. Suppose G_τ is a regular path such that for $\tau \neq \frac{1}{2}$, G_τ is a generic immersion, and for $\tau = \frac{1}{2}$, G_τ has a tangency at the d-th stratum. The topology of $G_\tau(N)$ changes: in particular the d-th stratum undergoes a surgery while crossing the value $\tau = \frac{1}{2}$. A precise description can be deduced from [Ekh01, Lemma 3.5], which can be easily generalised to the case when $\dim N = n$ and $\dim \Omega = n+k$ (as mentioned above, Ekholm's paper [Ekh01] has specific dimension constraints, but the proofs in his Section 3 work without these restrictions).

Remark. Note that a regular homotopy is a path of immersions, but it is not the case that every regular homotopy is a regular path of immersions.

3.1.3. **Paths of immersed Morse functions.** In this subsection, we consider a fixed map $G\colon N \to \Omega$, which we assume to be a generic immersion. We set $M = G(N)$ and define the strata $\Omega[0], \Omega[1], \dots$ as in Subsection 1.3. Consider the space of maps $\mathcal{A} := C^\infty(\Omega;\mathbb{R})$.

Definition 3.1.15. Let $d \geq 0$.

- (*non Morse critical point*) The subspace $\mathcal{A}_{\alpha,d} \subseteq \mathcal{A}$ is by definition the space of functions $F \in \mathcal{A}$ such that there exists a point $p \in \Omega[d]$ with $Df_d(p) = 0$ (with $f_d := F|_{\Omega[d]}$) and such that $D^2 f_d(p)$ has nontrivial kernel. Inside $\mathcal{A}_{\alpha,d}$ we specify the following subsets.
  - (*non-degenerate non-Morse singularity*) Define the subspace $\mathcal{A}^1_{\alpha,d} \subseteq \mathcal{A}_{\alpha,d}$ by the conditions that $\dim\ker D^2 f_d(p) = 1$, and for $v \in \ker D^2 f_d(p)$, $v \neq 0$, we have $D^3 f_d(p)(v,v,v) \neq 0$.
  - (*degenerate non-Morse singularity*) Define the subspace $\mathcal{A}^2_{\alpha,d} \subseteq \mathcal{A}_{\alpha,d}$ by the condition that $\dim\ker D^2 f_d(p) \geq 2$, or for all $v \in \ker D^2 f_d(p)$, we have $D^3 f_d(p)(v,v,v) = 0$.

- (*two critical points at the same level*) The subspace $\mathcal{A}^1_{\beta,d_1,d_2} \subseteq \mathcal{A}$ is the set of maps such that there are two critical points p_1 of $F|_{\Omega[d_1]}$ and p_2 of $F|_{\Omega[d_2]}$ such that $F(p_1) = F(p_2)$, but no third critical point has this property. The subspace $\mathcal{A}^2_{\beta,d_1,d_2,d_3}$ is the set of maps such that there exists p_i, critical points of $F|_{\Omega[d_i]}$, with $F(p_1) = F(p_2) = F(p_3)$.
- (*extra branch passing through a critical point*) For each s with $1 \leq s < d$, define the subspace $\mathcal{A}_{\gamma,d,s} \subseteq \mathcal{A}$ to be the subspace of smooth functions such that there is a critical point $p \in \Omega[d]$ of $F|_{\Omega[d]}$ and there are branches $Y_1, \dots, Y_s$, through p such that F restricted to $Y_1 \cap \dots \cap Y_s$ has a critical point at p.
- (*two events at the same time*) The subspace $\mathcal{A}_\omega$ is the subspace of maps that satisfy independently two conditions from the above list, e.g. two non-Morse critical points, two critical points at the same level etc.

We let

$$\mathcal{A}_\alpha = \bigcup_d \mathcal{A}_{\alpha,d},\ \mathcal{A}_\beta = \bigcup_{d_1,d_2} \mathcal{A}_{\beta,d_1,d_2},\ \text{and } \mathcal{A}_\gamma = \bigcup_{d,s} \mathcal{A}_{\gamma,d,s}.$$

Furthermore we set

$$\mathcal{A}^1_\alpha = \bigcup_d \mathcal{A}^i_{\alpha,d};\ \ \mathcal{A}^2_\alpha = \bigcup_d \mathcal{A}^i_{\alpha,d};\ \ \mathcal{A}^1_\beta = \bigcup_{d_1,d_2} \mathcal{A}^1_{\beta,d_1,d_2} \text{ and } \mathcal{A}^2_\beta = \bigcup_{d_1,d_2,d_3} \mathcal{A}^i_{\beta,d_1,d_2,d_3},$$

so that $\mathcal{A}_\alpha = \mathcal{A}^1_\alpha \cup \mathcal{A}^2_\alpha$, and $\mathcal{A}_\beta = \mathcal{A}^1_\beta \cup \mathcal{A}^2_\beta$. The set $\mathcal{A}^0$ is the complement

$$\mathcal{A}^0 := \mathcal{A} \smallsetminus (\mathcal{A}_\alpha \cup \mathcal{A}_\beta \cup \mathcal{A}_\gamma).$$

Note that the set $\mathcal{A}^0$ consists precisely of the excellent immersed Morse functions in the sense of Definition 2.4.1. The subspace $\mathcal{A}_\alpha$ corresponds to the failure of condition (IM-1). The subspace $\mathcal{A}_\beta$ corresponds to the non-excellent Morse functions, while $\mathcal{A}_\gamma$ is the failure of the (IM-2) condition.

Lemma 3.1.16.

(i) The set $\mathcal{A}^1_{\alpha,d}$ is defined, via the jet extension map, using a union of (finitely many) strata of codimension at least $n+k+1$ in $J^3(\Omega;\mathbb{R})$.
(ii) The set $\mathcal{A}^2_{\alpha,d}$ is defined using a union of strata of codimension at least $n+k+2$ in $J^3(\Omega;\mathbb{R})$.
(iii) The set $\mathcal{A}^1_{\beta,d_1,d_2}$ is defined by a stratum of codimension $2(n+k)+1$ in $J^1_2(\Omega;\mathbb{R})$, while $\mathcal{A}^2_{\beta,d_1,d_2,d_3}$ is defined by a stratum of codimension $3(n+k)+2$ in $J^1_3(\Omega;\mathbb{R})$,
(iv) The set $\mathcal{A}_{\gamma,d,s}$ is defined by a stratum of codimension at least $n+2k$ in $J^1(\Omega;\mathbb{R})$.

Proof. The stratum $\mathcal{A}_{\alpha,d}$ is defined by the following conditions: a point p belongs to $\Omega[d]$ (codimension equal to $\operatorname{codim}\Omega[d] \subseteq \Omega$), the derivative $DF(p)$ vanishes at $T_p\Omega[d]$ (codimension equal to $\dim\Omega[d]$). Now $\det D^2 f_d(p) = 0$ is one more condition. Altogether, there are $1+\dim\Omega = n+k+1$ conditions. Next, within $\mathcal{A}_{\alpha,d}$ we specify $\mathcal{A}^1_{\alpha,d}$ by open conditions (codimension zero). The set $\mathcal{A}^2_{\alpha,d}$ is the complement of $\mathcal{A}^1_{\alpha,d}$ in $\mathcal{A}_{\alpha,d}$ and it is a union of two strata (one is when $\dim\ker D^2 f_d(p) > 1$, the other is when $D^3 f_d(p)(v,v,v) = 0$). Each of these two strata is defined using one extra condition, so the codimension is one higher, namely $n+k+2$.

The stratum $\mathcal{A}^1_{\beta,d_1,d_2}$ is defined by $p_1 \in \Omega[d_1]$ and $DF(p_1)$ vanishes on $T_{p_1}\Omega[d_1]$ (codimension equal to $\dim\Omega$ as above). The same codimension applies for $p_2 \in \Omega[d_2]$ being a critical point of $F|_{\Omega[d_2]}$. Next, there is an additional condition $F(p_1) = F(p_2)$, so we get

$2\dim\Omega + 1 = 2(n+k)+1$ conditions. The same reasoning gives the codimension of the stratum $\mathcal{A}^2_{\beta,d_1,d_2,d_3}$.

Finally, consider $\mathcal{A}_{\gamma,d,s}$. For $p \in \Omega[d]$, the condition that DF vanishes on $T_p(Y_1 \cap \cdots \cap Y_s)$ specifies a stratum of codimension $\dim(Y_1 \cap \cdots \cap Y_s) = \dim\Omega[d] + k(d-s) \geq \dim\Omega[d] + k$. Together with $p \in \Omega[d]$ ($\dim\Omega - \dim\Omega[d]$ conditions), we have codimension at least $n+2k$. □

We deal with $\mathcal{A}_\omega$. We use the following principle.

Lemma 3.1.17 (Principle of independent singularities)**.** *Let X, Y be two smooth manifolds. Suppose two properties of a function define subspaces $\mathcal{A}_1 \subseteq J^{a_1}_{b_1}(X,Y)$, and $\mathcal{A}_2 \subseteq J^{a_2}_{b_2}(X,Y)$. A simultaneous occurrence of these two properties defines a subspace $\mathcal{A}_{12}$ of $J^{\max(a_1,a_2)}_{b_1+b_2}(X,Y)$ of codimension* $\operatorname{codim}\mathcal{A}_1 + \operatorname{codim}\mathcal{A}_2$*. In particular, if two properties $\mathcal{A}_1, \mathcal{A}_2$ have codimension at least $b_1\dim\Omega + 1$ and $b_2\dim\Omega+1$ respectively, then a generic 1-parameter family avoids the situation when two properties appear simultaneously.*

Proof. Say the first property depends on the values of functions and differentials (up to ordeer a_1) at points $p_1, \ldots, p_{b_1}$, and the total number of conditions is $\operatorname{codim}\mathcal{A}^1$ conditions on them. Suppose analogously, for the second property, that it involves $\operatorname{codim}\mathcal{A}^2$ conditions on the values of functions and differentials at another b_2-tuple of points, $p'_1, \ldots, p'_{b_2}$. The simultaneous occurrence means that we simply merge the conditions, so that we have $\operatorname{codim}\mathcal{A}^1 + \operatorname{codim}\mathcal{A}^2$ conditions on a (b_1+b_2)-tuple of points $p_1, \ldots, p_{b_1}, p'_1, \ldots, p'_{b_2}$.

The second part follows from Corollary 2.3.5(ii). □

We point out that multijet transversality is adapted to the situation, where the two tuples of points $p_1, \ldots, p_{b_1}$ and $p'_1, \ldots, p'_{b_2}$ in the proof of Lemma 3.1.17 do not coincide. That is, the two properties are really considered as independent.

Corollary 3.1.18. *The set of functions $F \in \mathcal{A}$ avoiding $\mathcal{A}_\alpha \cup \mathcal{A}_\beta \cup \mathcal{A}_\gamma$ is open and dense. The set of paths of functions $F_\tau \in \mathcal{A}$, $\tau \in [0,1]$ avoiding $\mathcal{A}^2_\alpha$, $\mathcal{A}^2_\beta$, $\mathcal{A}_\gamma$ and $\mathcal{A}_\omega$, and intersecting $\mathcal{A}^1_\alpha$ and $\mathcal{A}^1_\beta$ in finitely many values of $\tau \in (0,1)$, is open and dense.*

In particular, any two functions $F_0, F_1 \in \mathcal{A} \setminus \mathcal{A}_\alpha \cup \mathcal{A}_\beta \cup \mathcal{A}_\gamma$ can be connected by such a path.

Proof. The parameter counting argument of Lemma 3.1.16 and the Multijet Transversality Theorem 2.3.4, used via Corollary 2.3.5, provide us with density both of functions and of paths, where for $\mathcal{A}_\omega$ we use the independence principle enunciated in Lemma 3.1.17. As $\mathcal{A}_\alpha$ and $\mathcal{A}_\gamma$ involve only a single point (so can be handled by the Jet Transversality Theorem), we have openness. However, the set of functions avoiding $\mathcal{A}_\beta$ is clearly open, as well as the set of paths having finitely many rearrangements. □

Definition 3.1.19. We call a path of functions as in the corollary an $\mathcal{A}^1$*-regular path* of functions on Ω.

We remark that Corollary 3.1.18 implies Theorem 2.4.5, so we have now provided the proof of the latter theorem, as promised.

3.1.4. **Paths of immersions relative to a function.** Now we consider the situation where $F\colon \Omega \to \mathbb{R}$ is fixed and F is Morse (in the classical sense) as a function on Ω. Suppose $G\colon N \looparrowright \Omega$ is an immersion.

Definition 3.1.20. The map G is called *F-regular* if G is a generic immersion (i.e. $G \in \mathcal{I}^0$), and F is immersed Morse with respect to $M = G(N)$.

We let $\mathcal{R}^0$ denote the set of all F-regular immersions. We now consider the complement of $\mathcal{R}^0$ in $\mathcal{R} := C^\infty(N, \Omega)$. The following definition is similar in spirit to Definition 3.1.15, except that in Definition 3.1.15 we discussed the subspaces of $C^\infty(\Omega, \mathbb{R})$, while here we consider the analogous conditions, but on the space $\mathcal{R}$ of smooth maps from N to Ω.

Definition 3.1.21. Let $\mathcal{R} := C^\infty(N, \Omega)$ and fix an excellent Morse function $F\colon \Omega \to \mathbb{R}$.

- (*a non-Morse singularity at the stratum of depth d*) The subspace $\mathcal{R}_{\alpha,d}(F) \subseteq \mathcal{R}$ of smooth maps is by definition the space of maps $G \in \mathcal{I}^0$ such that, with $M := G(N)$ and $f_d = F|_{\Omega[d]}$, there is a point $p \in \Omega[d]$ with $Df_d(p) = 0$, $f_d = F|_{\Omega[d]}$, and with $D^2 f_d(p)$ degenerate. Inside $\mathcal{R}_{\alpha,d}(F)$ we specify the following subspaces.
  - (*non-degenerate case*) The subspace $\mathcal{R}^1_{\alpha,d}(F) \subseteq \mathcal{I}^0 \subseteq \mathcal{R}_{\alpha,d}(F)$ consists of those maps that additionally satisfy that $D^2 f_d(p)$ has one-dimensional kernel spanned by v and $D^3 f_d(p)(v,v,v) \neq 0$.
  - (*degenerate case*) The subspace $\mathcal{R}^2_{\alpha,d}(F) \subseteq \mathcal{R}_{\alpha,d}(F)$ consists of those maps that additionally satisfy that $D^2 f_d(p)$ has either at least two-dimensional kernel, or it is a one dimensional kernel and $D^3 f_d(p)(v,v,v) = 0$ for all $v \in \ker D^2 f_d(p)$.
- (*two critical points on the same level set*) The subspace $\mathcal{R}^1_{\beta,d_1,d_2}(F) \subseteq \mathcal{I}^0 \subseteq \mathcal{R}$ is the set of maps $G \in \mathcal{I}^0$ such that with $M = G(N)$, there are two critical points p_1 of $F|_{\Omega[d_1]}$ and p_2 of $F|_{\Omega[d_2]}$ such that $F(p_1) = F(p_2)$, but no other critical points have critical value $F(p_1)$. The subspace $\mathcal{R}^2_{\beta,d_1,d_2,d_3}(F)$ is the set of maps such that $p_i \in \Omega[d_i]$ is a critical point of $F|_{\Omega[d_i]}$, $i = 1,2,3$ and $F(p_1) = F(p_2) = F(p_3)$;
- (*extra branch passing through critical points*) The subspace $\mathcal{R}_{\gamma,d,s}(F) \subseteq \mathcal{R}$ is the space of maps $G \in \mathcal{I}^0$ such that there is a critical point $p \in \Omega[d]$ of $F|_{\Omega[d]}$ and there are branches $Y_1, \dots, Y_s$, $s < d$, through p such that F restricted to $Y_1 \cap \dots \cap Y_s$ has a critical point at p.
- (*failure to be a generic immersion*) The subspace $\mathcal{R}_{\delta,d} \subseteq \mathcal{R}$ is the space of maps having a d-fold non-transversality point (similar to $\widetilde{\mathcal{I}}^1_d$). We let $w \in \Omega$ be a non-transversality point of G and let $u_1, \dots, u_d$ be the preimages of W and let $U_1, \dots, U_d$ be pairwise disjoint open neighbourhoods in N of $u_1, \dots, u_d$. We specify two subspaces of $\mathcal{R}_{\delta,d}$.
  - (*simple non-transversality point*) The subspace $\mathcal{R}^1_{\delta,d}(F) \subseteq \mathcal{R}_{\delta,d}$ is the space of maps in $\mathcal{I}^1$, where for any subset $(i_1, \dots, i_r)$ of $(1, \dots, d)$ for which $M' := G(U_{i_1}) \cap \dots \cap G(U_{i_r})$ is a transverse intersection, and $\dim M' > 0$, the restriction $F|_{M'}$ has nonvanishing derivative at w.
  - (*degenerate non-transversality point*) The subspace $\mathcal{R}^2_{\delta,d}(F) \subseteq \mathcal{R}_{\delta,d}$ is the space of maps G where either $G \in \mathcal{I}^2$, or $F|_{M'}$ has a critical point at w for some transverse intersection M' defined as in the previous item.
  - (*two events happening at the same time*) The subspace $\mathcal{R}^2_\omega(F)$ is the subspace of maps that satisfy independently two conditions from the above list, e.g. two non-Morse critical points, two critical points at the same level etc.

We let

$$\mathcal{R}^i_\alpha(F) := \bigcup_d \mathcal{R}^i_{\alpha,d}(F),\ \mathcal{R}^1_\beta(F) := \bigcup_{d_1,d_2} \mathcal{R}^1_{\beta,d_1,d_2}(F),\ \text{and } \mathcal{R}_\gamma(F) = \bigcup_{d,s} \mathcal{R}_{\gamma,d,s}(F).$$

We also define $\mathcal{R}^i_\delta(F) = \bigcup_d \mathcal{R}^i_{\delta,d}(F)$ for $i = 1,2$, and $\mathcal{R}^2_\beta(F) = \bigcup_{d_1,d_2,d_3} \mathcal{R}^2_{\beta,d_1,d_2,d_3}(F)$. Set

$$\mathcal{R}^1(F) = \mathcal{R}^1_\alpha(F) \cup \mathcal{R}^1_\beta(F) \cup \mathcal{R}^1_\delta(F) \text{ and } \mathcal{R}^2(F) = \mathcal{R}^2_\alpha(F) \cup \mathcal{R}^2_\beta(F) \cup \mathcal{R}_\gamma(F) \cup \mathcal{R}^2_\delta(F).$$

Similarly to the case of the $\mathcal{A}$ spaces, the subspace $\mathcal{R}_\alpha(F)$ corresponds to the failure of condition (IM-1) of F as an immersed Morse function on $M = G(N)$. The subspace $\mathcal{R}_\beta$ corresponds to non-excellent Morse functions, while $\mathcal{R}_\gamma$ is the failure of the (IM-2) condition. Finally, $\mathcal{R}_\delta$ corresponds to maps G that are not immersions. The following statement has essentially the same proof as for the $\mathcal{A}$ spaces.

Lemma 3.1.22. *The set of maps G that avoid $\mathcal{R}^1(F) \cup \mathcal{R}^2(F)$ is open-dense.*

The 1-parameter variant of Lemma 3.1.22 holds as well, but the $\mathcal{R}^1$ stratum cannot be avoided in general.

Lemma 3.1.23. *Suppose $k > 1$. Suppose G_τ is a path in the space of immersions $\mathcal{I} = \mathcal{I}(N,\Omega) \subseteq C^\infty(N,\Omega)$. Then G_τ can be perturbed in such a way that G_τ avoids $\mathcal{R}^2(F)$, avoids $\mathcal{I}^2$, and hits $\mathcal{R}^1(F) \cup \mathcal{I}^1$ only at finitely many points. Moreover, if the original path was such that F is immersed Morse with respect to $G_0(N)$ and $G_1(N)$, then we may assume that the perturbation fixes G_0 and G_1.*

Proof. The parameter counting argument for $\mathcal{R}_\alpha, \mathcal{R}_\beta$, and $\mathcal{R}_\gamma$ is the same as in the case of the corresponding spaces $\mathcal{A}$, so the set of maps missing $\mathcal{R}^2_\alpha \cup \mathcal{R}_\gamma$ is residual. For $\mathcal{R}_\delta$ note that the conditions defining $\mathcal{R}^1_\delta$ form an open-dense subset in the space $\mathcal{I}^1$: indeed, non-vanishing of the derivative of F at each of M' is an open-dense condition. The case of $\mathcal{R}_\omega$ is handled by Lemma 3.1.17, because the principle remains the same. The intersection of a residual set with an open-dense subset is residual, hence dense. □

Motivated by Lemma 3.1.23, we introduce the following definition.

Definition 3.1.24. Suppose F is an excellent Morse function on Ω. A path $G_\tau \in \mathcal{I}$, $\tau \in [0,1]$, of immersions from N to Ω is called an *F-regular path* if G_0, G_1 are generic immersions, F is immersed Morse with respect to $G_0(N), G_1(N)$, and the path G_τ is transverse to $\mathcal{R}^1(F)$ avoiding $\mathcal{R}^2(F)$.

With this definition, Lemma 3.1.23 allows us to state that any path of immersions can be perturbed to an F-regular path. Note that a function belonging to an F-regular path of functions need not be itself an F-regular function.

3.1.5. **Double paths.** We now merge the notion of a path of Morse functions and the path of immersions into the notion of a double path. A *double path* is a pair of paths (F_τ, G_τ), $\tau \in [0,1]$, such that $F_\tau\colon \Omega \to \mathbb{R}$ is an $\mathcal{F}^1$-path (Definition 3.1.1), and for each τ, $G_\tau\colon N \to \Omega$ is an immersion.

We will specify regularity conditions on a double path, which guarantee that for all but finitely many $\tau \in [0,1]$, F_τ is an immersed Morse function with respect to the pair $(\Omega, G_\tau(N))$, while for those finitely many parameter values where it is not, F_τ has mildest possible singularities.

The logic behind our construction is that we will usually fix an $\mathcal{F}^1$-regular path $F_\tau\colon \Omega \to \mathbb{R}$, and perturb G_τ as in Subsection 3.1.4. We will make precise what it means for G_τ to be an F_τ-regular path of functions, where F_τ changes with τ too.

To begin the plan sketched above, fix an $\mathcal{F}^1$-path of functions $F_\tau\colon \Omega \to \mathbb{R}$ on Ω. Let $\mathcal{D} := C^\infty(N \times [0,1], \Omega)$. We think of a path G_τ as an element of $\mathcal{D}$ (by currying). To specify generic regularity conditions, one defines subspaces $\mathcal{D}_\alpha$, $\mathcal{D}_\beta$, $\mathcal{D}_\gamma$, and $\mathcal{D}_\delta$ of $\mathcal{D}$, with respect to the fixed path F_τ, by analogy with the spaces $\mathcal{R}_\alpha$, $\mathcal{R}_\beta$, $\mathcal{R}_\gamma$, and $\mathcal{R}_\delta$. There is one extra subspace needed, $\mathcal{D}^2_\varepsilon$, which takes care of points where F_τ fails to be Morse at the same time as G_τ fails to be generic. Here are the precise definitions.

Definition 3.1.25. Fix an $\mathcal{F}^1$-path F_τ.

- (*non-Morse singularity at higher depth*) The subspace $\mathcal{D}_{\alpha,d} \subseteq \mathcal{D}$ of smooth paths is by definition the space of paths G_τ such that there exists τ_1 with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}_{\alpha,d}(F_{\tau_1})$. Inside $\mathcal{D}_{\alpha,d}$ we specify the following subspaces.
  - (*non-degenerate non-Morse critical point*) The subspace $\mathcal{D}^1_{\alpha,d} \subseteq \mathcal{D}_{\alpha,d}$ consists of those paths such that there exists τ_1 with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}^1_{\alpha,d}(F_{\tau_1})$.
  - (*degenerate non-Morse critical point*) The subspace $\mathcal{D}^2_{\alpha,d} \subseteq \mathcal{D}_{\alpha,d}$ consists of those paths G_τ such that there exists τ_1 with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}^2_{\alpha,d}(F_{\tau_1})$.
- (*two critical points on the same level set*) The subspace $\mathcal{D}^1_{\beta,d_1,d_2} \subseteq \mathcal{D}$ is the set of paths G_τ such that there exists $\tau_1 \in [0,1]$ with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}^1_{\beta,d_1,d_2}(F_{\tau_1})$.
- (*three or more critical points on the same level set*) The subspace $\mathcal{D}^2_{\beta,d_1,d_2,d_3} \subseteq \mathcal{D}$ is the set of paths G_τ such that there exists $\tau_1 \in [0,1]$ with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}^2_{\beta,d_1,d_2,d_3}(F_{\tau_1})$.
- (*extra branch passing through a critical point*) For $s < d$, the subspace $\mathcal{D}_{\gamma,d,s} \subseteq \mathcal{D}$ is the space of paths G_τ such that there exists $\tau_1 \in [0,1]$ with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}_{\gamma,d,s}(F_{\tau_1})$.
- (*failure to be a generic immersion*) The subspace $\mathcal{D}_{\delta,d} \subseteq \mathcal{D}$ is the space of paths G_τ such that there exists $\tau_1 \in [0,1]$ with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}_{\delta,d}(F_{\tau_1})$. We specify two subspaces of $\mathcal{D}_{\delta,d}$.
  - The subspace $\mathcal{D}^1_{\delta,d} \subseteq \mathcal{D}_{\delta,d}$ is the space of paths G_τ such that there exists $\tau_1 \in [0,1]$ with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}^1_{\delta,d}(F_{\tau_1})$.
  - The subspace $\mathcal{D}^2_{\delta,d} \subseteq \mathcal{D}_{\delta,d}$ is the space of paths G_τ such that there exists $\tau_1 \in [0,1]$ with F_{τ_1} a Morse function and $G_{\tau_1} \in \mathcal{R}^1_{\delta,d}(F_{\tau_1})$.
- (*failure of F_τ to be excellent Morse*) The subspace $\mathcal{D}^2_\varepsilon$ is the set of paths G_τ for which there exists $\tau_1 \in [0,1]$ such that F_{τ_1} is not Morse and G_{τ_1} is not a generic immersion.
- (*two events at the same time*) The subspace $\mathcal{D}_\omega$ is the subspace of paths such that for some $\tau_1 \in [0,1]$, $G_{\tau_1} \in \mathcal{R}_\omega(F_{\tau_1})$.

We let

$$\mathcal{D}^i_\alpha := \bigcup_d \mathcal{D}^i_{\alpha,d},\ \mathcal{D}^1_\beta := \bigcup_{d_1,d_2} \mathcal{D}^1_{\beta,d_1,d_2}, \mathcal{D}_\gamma = \bigcup_{d,s} \mathcal{D}_{\gamma,d,s},\ \text{and}\ \mathcal{D}^i_\delta = \bigcup_d \mathcal{D}^i_{\delta,d}.$$

Set also $\mathcal{D}^2_\beta := \bigcup_{d_1,d_2,d_3} \mathcal{D}^2_{\beta,d_1,d_2,d_3}$.

$$\mathcal{D}^1 = \mathcal{D}^1_\alpha \cup \mathcal{D}^1_\beta \cup \mathcal{D}^1_\delta \text{ and } \mathcal{D}^2 = \mathcal{D}^2_\alpha \cup \mathcal{D}^1_\beta \cup \mathcal{D}_\gamma \cup \mathcal{D}^2_\delta \cup \mathcal{D}^2_\varepsilon \cup \mathcal{D}_\omega.$$

Proposition 3.1.26. *Fix an $\mathcal{F}^1$-path of functions $F_\tau\colon \Omega \to \mathbb{R}$. For any path G_τ of immersions, there is a C^∞-close regular path of immersions G'_τ such that G'_τ misses $\mathcal{D}^2$, and hits $\mathcal{D}^1$ only for finitely many values of τ.*

Sketch of proof. The dimension counting argument for the spaces $\mathcal{D}^i_\bullet$, for $\bullet \in \{\alpha,\beta,\gamma,\delta\}$, is analogous to the dimension counting argument for the $\mathcal{A}$ spaces, given in Lemma 3.1.16 and Corollary 3.1.18; as usual this relies upon the Multijet Transverality Theorem 2.3.4. It follows that the path G_τ can be perturbed to avoid $\mathcal{D}^2_\alpha \cup \mathcal{D}^2_\gamma \cup \mathcal{D}^2_\delta$ and to intersect $\mathcal{D}^1$ at finitely many values. Lemma 3.1.17 can be applied to show that the path G_τ can also be perturbed to avoid D_ω. It remains to deal with $\mathcal{D}^2_\varepsilon$. Suppose $\tau_1,\dots,\tau_r$ are precisely the values for which F_{τ_i} is not Morse. The condition $\mathcal{D}^2_\varepsilon$ means that G_{τ_i} is not a generic immersion for one of these values. Since being a generic immersion is open and dense, we can perturb G_τ to arrange that G_{τ_i} is a generic immersion for $i = 1, \dots, r$. After this perturbation, G_τ avoids $\mathcal{D}^2_\varepsilon$. □

Definition 3.1.27. A double path is called *regular* if:

- F_0 is immersed Morse with respect to $G_0(N)$, and F_1 is immersed Morse with respect to $G_1(N)$.
- F_τ is an $\mathcal{F}^1$-path of Morse functions, when regarded as a path of functions on Ω;
- G_τ is a regular path of immersions;
- G_τ intersects $\mathcal{D}^1$ at finitely many points and misses $\mathcal{D}^2$.

Proposition 3.1.26 can be rephrased in the following manner.

Corollary 3.1.28. *If (F_τ, G_τ) is a double path, then there exists a regular double path (F'_τ, G'_τ) arbitrarily close to it. Moreover, if G_0 is a generic immersion and F_0 is immersed Morse with respect to $G_0(N)$, then we can take $F'_0 = F_0$, and $G'_0 = G_0$.*

Proof. First perturb F_τ to be an $\mathcal{F}^1$-path (Definition 3.1.1), and then use Proposition 3.1.26 to perturb G_τ. □

Example 3.1.29. Suppose (F_τ, G_τ) is such that F_τ is a $\mathcal{F}^1$-path and $G_\tau = G_0$ is an regular immersion. Then (F_τ, G_τ) is a regular double path.

Example 3.1.30. Assume (F_τ, G_τ) is such that $F_\tau = F_0$ and G_τ is a regular F_0-path in the sense of Definition 3.1.20. Then (F_τ, G_τ) is a regular double path.

We need the following statement.

Lemma 3.1.31. *Suppose (F_τ, G_τ) is a regular double path. Then $F_\tau \circ G_\tau$ is an $\mathcal{F}^1$-path of functions on N.*

Proof. Take $u \in N$ and choose $U \subseteq N$, a neighbourhood of u, such that $G_\tau|_U$ is an embedding. With this choice, it is clear that u is a critical point of $F_\tau \circ G_\tau$ if and only if $G_\tau(u)$ is a critical point of F_τ restricted to U. As G_τ omits $\mathcal{D}_\gamma$, if $F_\tau|_{G_\tau(U)}$ has a critical point on $G_\tau(U)$, $G_\tau(u)$ must belong to the first stratum. This shows that the critical points of $F_\tau \circ G_\tau$ correspond to critical points of F_τ on the first stratum.

The correspondence goes further. Morse critical points of $F_\tau \circ G_\tau$ correspond to Morse critical points of F_τ on the first stratum, and births/deaths of pairs of critical points of $F_\tau \circ G_\tau$ correspond to births/deaths of pairs of critical points of F_τ on the first stratum. That is, the double path missing $\mathcal{D}^2_{\alpha,1}$ implies that the non-Morse critical points of $F_\tau \circ G_\tau$ are precisely births and deaths. A refinement of the argument, showing that $F_\tau \circ G_\tau$ is actually transverse to $\mathcal{F}^1$ if G_τ is transverse to $\mathcal{D}^1_{\alpha,1}$, is left to the reader.

Transversality of G_τ to $\mathcal{D}_{\beta,1,1}$ means that there are finitely many rearrangements along the path $F_\tau \circ G_\tau$. In other words, $F_\tau \circ G_\tau$ is a path of functions with finitely many births, deaths and rearrangements. For all but finitely many values of τ, $F_\tau \circ G_\tau$ is excellent Morse, for the remaining values it belongs to $\mathcal{F}^1$. □

3.1.6. **Neat paths for manifolds with boundary.** So far, the discussion of path of maps from Ω to $\mathbb{R}$ and from N to Ω was done for N and Ω closed. In our applications, we allow for N and Ω to be manifolds with boundary, provided the maps behave nicely near the boundary. In this short subsection, we make this intuition precise. The bottom line is that all the results that are proved about paths of functions on closed manifolds carry over to the case of compact manifolds with boundary, as long as suitable neatness conditions are preserved.

We recall the definition of neat immersions and extend the definition to *very neat immersions.*

Definition 3.1.32. An immersion $G: N \looparrowright \Omega$ is *neat* if $G^{-1}(\operatorname{bd}\Omega) = \operatorname{bd} N$, and there exist collar neighbourhoods of $\operatorname{bd} N$ and $\operatorname{bd}\Omega$ such that in these coordinates, for $x \in \operatorname{bd} N$ and $t \in [0,1]$ we have that $G(x,t) = (G(x,0),t)$.

Definition 3.1.33. A path G_τ, $\tau \in [0,1]$ of immersions is *very neat* if G_0 is neat, the restriction $G|_{\operatorname{bd} N}: \operatorname{bd} N \looparrowright \operatorname{bd}\Omega$ is a generic immersion, and if there is a neighbourhood U of $\operatorname{bd} N$ in N, and a neighbourhood W of $\operatorname{bd}\Omega$ in Ω, such that $U \cong \operatorname{bd} N \times [0,1)$, $W \cong \operatorname{bd}\Omega \times [0,1)$, and for each $\tau \in [0,1]$ and $(x,t) \in N \times [0,1]$, we have that $G_\tau(x,t) = (G_0(x,0),t)$. That is, we require the path to be independent of τ near the boundary of N.

Any neat path of immersions can be perturbed to a regular, neat path of immersions. In fact, being neat implies that G_τ is a regular immersion on U for each τ. Next, regularity is local, so we can choose a smaller product neighbourhood U' of $\operatorname{bd} N$ and a perturbation of G_τ to a regular path of immersion, such that the perturbation is supported away from U'. Then the newly created path is still a neat path, with U' replacing U and $G_0(U')$ replacing W.

Suppose Ω is a manifold whose boundary is a union $\operatorname{bd}_-\Omega \sqcup \operatorname{bd}_+\Omega$. A function $F: \Omega \to [0,1]$ is called *neat*, if $F^{-1}(0) = \operatorname{bd}_-\Omega$, $F^{-1}(1) = \operatorname{bd}_+\Omega$ and F has no critical points in a neighbourhood W of $\operatorname{bd}\Omega$. In the immersed case we assume moreover that M is a neatly immersed manifold (i.e. the image of a neat immersion) and that F has no critical points near $\operatorname{bd}\Omega$ on each stratum.

Definition 3.1.34. A path of functions $F_\tau: \Omega \to \mathbb{R}$, $\tau \in [0,1]$, is called *neat* if, for any τ, $F_\tau^{-1}(0) = \operatorname{bd}_-\Omega$, $F_\tau^{-1}(1) = \operatorname{bd}_+\Omega$, and there is an open subset $W_\tau \subseteq \Omega$ containing $\operatorname{bd}\Omega$ such $F_\tau|_W = F_0$ and F_0 has no critical points on W_τ.

Proposition 3.1.35. *Any neat path of functions can be perturbed to a neat $\mathcal{F}^1$-path, possibly at the expense of shrinking the open subset W.*

This follows as above, because the property of being an $\mathcal{F}^1$-path is local. We now pass to double paths of functions.

Definition 3.1.36. A double path (F_τ, G_τ) is called *neat* if F_τ is neat and G_τ is very neat.

Proposition 3.1.37. *If (F_τ, G_τ) is a neat double path, there exists a perturbation, fixing the endpoints at $\tau = 0, 1$, to a neat and regular double path.*

Proof. The proof combines the arguments in this subsection. First, by Proposition 3.1.35 we can perturb F_τ to a neat $\mathcal{F}^1$-path at the expense of shrinking a neighbourhood of $\operatorname{bd}\Omega$ on which F_τ is independent of τ. Next, we perturb G_τ to a very neat path of immersions such that (F_τ, G_τ) is a regular double path, as in Proposition 3.1.26, using the local argument once again to avoid changing G_τ near $\operatorname{bd} N$. □

3.2. Rearrangement for immersed Morse functions

Motivated by Cerf theory, we state the rearrangement theorem under conditions which guarantee disjointness of the appropriate ascending and descending membranes. The rearrangement theorem creates a suitable path of functions in which a rearrangement occurs. Despite this formulation, the proof of the rearrangement theorem is a relatively straightforward generalisation of an analogous result for the embedded case [BP16]. This in turn was proven in a similar way to the corresponding result in Milnor's book [Mil65]. We begin by introducing some notation.

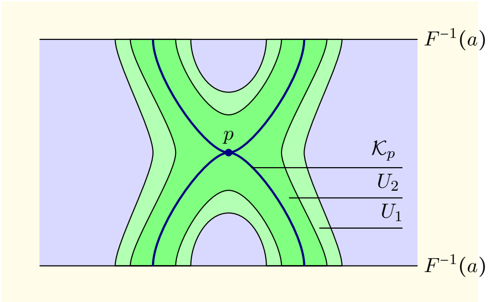

FIGURE 16. Proof of the rearrangement theorem.

Theorem 3.2.1 (Rearrangement)**.** *Let M be an n-dimensional immersed manifold of codimension k in Ω. Let $F\colon\Omega\to\mathbb{R}$ be a Morse function and let p be a critical point of F. Let ξ be a grim vector field for F. Let $[a,b]$ be an interval in $\mathbb{R}$ containing $F(p)$.*

For every $c\in[a,b]$ and for any open subset V of $F^{-1}[a,b]$ containing $\mathcal{K}_{\xi,a,b}(p)$, there exists a path of Morse functions F_τ, $\tau\in[0,1]$, supported in V, that moves the critical point p to the level set c. More specifically, the following hold for F_τ.

- *$F_0 = F$ i.e. the path starts at F.*
- *$F_1(p) = c$.*
- *For every $\tau\in[0,1]$ the critical points of F_τ are the same as the critical points of F_0.*
- *There exists a neighbourhood U of p such that $F_\tau(w) = F_0(w)+F_\tau(p)-F_0(p)$ for each $w\in U$. In particular the index of p is unchanged.*
- *For every $w\notin V$ and for every $w\in F_0^{-1}\{a,b\}$ we have $F_\tau(w)=F_0(w)$ for all τ. That is, $F_\tau(c)\neq F_0(x)$ implies that $x\in V$.*
- *The vector field ξ is a grim vector field for F_τ for every $\tau\in[0,1]$.*

The theorem implies that we can move the value of p under F_1 to any given $c\in[a,b]$ by only changing F within a pre-determined neighbourhood of $\mathcal{K}_{\xi,a,b}(p)$.

Proof. The proof is an extension of the argument of Milnor in [Mil65, Section 4]. Set

$$\Omega' := F^{-1}[a,b] \text{ and } \mathcal{K}_p := \mathcal{K}_{\xi,a,b}(p).$$

We claim that there exists a function $\mu\colon\Omega'\to[0,1]$ such that μ is constant on trajectories of ξ, and moreover $\mu\equiv 0$ outside V and $\mu\equiv 1$ in a neighbourhood of $\mathcal{K}_p$. To see this, first construct two grim neighbourhoods U_1, U_2 of $\mathcal{K}_p$ with the property that $U_2\subseteq U_1\subseteq V$ and the boundary of U_2 intersects the boundary of U_1 only at the level sets $F^{-1}(a)$ and $F^{-1}(b)$ and

$$\partial(\overline{U}_1\cap F^{-1}(a))\cap\partial(\overline{U}_2\cap F^{-1}(a)) = \partial(\overline{U}_1\cap F^{-1}(b))\cap\partial(\overline{U}_2\cap F^{-1}(b)) = \varnothing; \tag{3.2.1}$$

see Figure 16. The existence of a grim neighbourhood of $\mathcal{K}_p$ follows directly from Lemma 2.7.10. Note that the assumption of Lemma 2.7.10 is satisfied by the properties of ξ. The function μ is constructed as follows.

- Away from $\overline{U}_1$ we set $\mu = 0$;
- On $\overline{U}_2$ we set $\mu = 1$;
- We first extend μ on $F^{-1}(a)$ across $F^{-1}(a)\cap(\overline{U}_1\smallsetminus\overline{U}_2)$ to a smooth function.

Finally, take a point $w\in U_1\smallsetminus\overline{U}_2$. As there are no critical points of ξ in $U_1\smallsetminus U_2$, and both U_1 and U_2 are ξ-invariant, the trajectory of ξ through w hits $F^{-1}(a)$ in the past at a point

$z \in F^{-1}(a) \cap (\overline{U}_1 \setminus \overline{U}_2)$. We declare $\mu(w) := \mu(z)$. This completes the definition of the function $\mu: \Omega' \to [0,1]$.

Now choose an auxiliary smooth function $\Psi: \mathbb{R} \times [0,1] \times [0,1] \to \mathbb{R}$, written as $\Psi(t,\tau,s)$, such that:

(O-1) We have $\Psi(t,\tau,s) = t$, whenever at least one of the following holds:
- $t \le a$;
- $t \ge b$;
- $\tau = 0$;
- $s = 0$.

(O-2) The map $t \mapsto \Psi(t,\tau,s)$ is monotone increasing for all τ and s.

(O-3) For all τ, and for sufficiently small $\delta > 0$ the function $t \mapsto \Psi(t,\tau,1)$ is linear with derivative 1 if $t \in [F(p)-\delta, F(p)+\delta]$.

(O-4) $\Psi(F(p),1,1) = c$.

We define

$$F_\tau(w) = \Psi(F(w), \tau, \mu(w)).$$

We have $\partial_\xi F_\tau(w) = \frac{\partial \Psi}{\partial t} \partial_\xi F + \frac{\partial \Psi}{\partial s} \partial_\xi \mu$. Now by (O-2) we have $\frac{\partial \Psi}{\partial t} > 0$, and since ξ is grim we know that $\partial_\xi F \ge 0$ (with equality only at the critical points of F). By the construction of μ we have that $\partial_\xi \mu = 0$. It follows that $\partial_\xi F_\tau(w) \ge 0$ with equality only at the critical points of F.

Verifying that all the required properties of F_τ are satisfied is routine. For example, by (O-1) we see that $F_\tau(w) = F(w)$ whenever $F(w) \notin [a,b]$ or $\tau = 0$. The behaviour of the function F_τ near p is governed by (O-3) and the rearrangement is guaranteed by (O-4). The last condition in (O-1), that is $\Psi(t,\tau,0) = t$ together with the fact that $\mu = 0$ outside of U_2 implies that we do not change the function F except near $\mathcal{K}_p$. □

Corollary 3.2.2. *Let $F: \Omega \to \mathbb{R}$ be a Morse function, let p, p' be two critical points of F, and let $a < b$ be two non-critical values such that $a < F(p) < F(p') < b$. Suppose there is a grim vector field ξ such that $\mathcal{K}_{\xi,a,b}(p)$ and $\mathcal{K}_{\xi,a,b}(p')$ are disjoint. Then there exists a path of Morse functions F_τ, $\tau \in [0,1]$, $F_0 = F$, such that ξ is a grim vector field for F_τ for every τ, and such that $F_1(p) > F_1(p')$. The path is supported on an arbitrarily small open neighbourhood $V \subseteq F^{-1}([a,b])$ of $\mathcal{K}_{\xi,a,b}(p)$ and $\mathcal{K}_{\xi,a,b}(p')$.*

Proof. Apply the Rearrangement Theorem 3.2.1 to p with $c = \frac{1}{2}(F(p') + b)$. □

Remark. If F_τ is a path constructed in Corollary 3.2.2, then for every τ the critical points of F_τ are the same as the critical points of F_0. Furthermore, there exists a neighbourhood U of p such that $F_\tau(w) = F_0(w) + F_\tau(p) - F_0(p)$ for each $w \in U$. In particular the index of p is the same for every $\tau \in [0,1]$.

As explained in [BP16, Section 4] in the embedded case, in codimension $k > 1$ we can always use the local rearrangement theorem to order critical points in such a way that higher index critical points have higher critical value. In the immersed case there is an analogous result, once we have replaced the index h with $h + d$, where d is the depth. The statement and the proof are essentially the same as in Milnor's book [Mil65, Section 4] with a necessary technical modification for the immersed case. We begin with the following definition.

Definition 3.2.3. A Morse function $F: \Omega \to \mathbb{R}$ is called an *ordered Morse function* if for any two critical points p and p' of indices h and h' and at depth d and d', respectively, the condition $h' + d' > h + d$ implies that $F(p') > F(p)$.

Remark. An ordered Morse function is not required to be excellent. Two critical points are allowed to have the same critical value, as long as the sum $h+d$ is equal for both points.

Theorem 3.2.4 (Global Rearrangement)**.** *Suppose $k \geq 2$. Let $M^n \subseteq \Omega^{n+k}$ be an immersed manifold and let $F: \Omega \to \mathbb{R}$ be an immersed Morse function. Then there exists a path F_τ of Morse functions such that $F_0 = F$ and F_1 is ordered. The path can be chosen in such a way that ξ is a grim vector field for F_τ for all τ.*

Furthermore, if p and p' are critical points of F and $h+d = h'+d'$ (where h, h' are the indices and d, d' are the depths), then any prescribed order of $F_1(p)$ and $F_1(p')$ can be achieved. In particular, we can arrange, if desired, for F_1 to be excellent.

Proof. By Corollary 2.6.6 if $h+d \geq h'+d'$ and ξ is Morse–Smale, then $\mathbb{M}_{\mathrm{a}}(p) \cap \mathbb{M}_{\mathrm{d}}(p') = \varnothing$. We conclude the proof in the standard way, e.g. as in [Mil65, Section 4]. □

3.3. ξ-paths

A ξ-path is, roughly speaking, a family of functions with a common grim vector field ξ. The vector field ξ gives us a precise control on the behaviour of the family of functions. We describe paths on Ω, i.e. in the immersed setting. The case of paths on N, i.e. on a smooth manifold (where the paths are called η-paths) is then given as a specialisation. Both notions of a ξ-path and of an η-path are used throughout the rest of Parts 3 and 4.

3.3.1. ξ-paths for immersed Morse functions.

Definition 3.3.1. Let $F = F_0: \Omega \to \mathbb{R}$ be an immersed Morse function and let ξ be a grim vector field with respect to F_0. A path $F_\tau: \Omega \to \mathbb{R}$ of immersed Morse functions is called a *ξ-path* if ξ is a grim vector field for F_τ for every τ.

Example 3.3.2. The path constructed in Rearrangement Theorem 3.2.4 is a ξ-path.

We make the following observation.

Lemma 3.3.3. *If F_τ is a ξ-path, then a point $p \in \Omega$ is a critical point of F_τ if and only if p is a critical point of F_0. The index of p as a critical point of F_τ does not depend on τ.*

Proof. By definition ξ vanishes only at critical points of F_τ. Therefore, critical points of F_τ are the same for all τ.

Suppose p is a critical point of F_τ at depth d. By the definition, the index of p is the dimension of the stable manifold of ξ restricted to the stratum $\Omega[d]$. As such, it does not depend on τ. □

Definition 3.3.4 (A special ξ-path)**.** A ξ-path $F_\tau: \Omega \to \mathbb{R}$ is *special* if for every critical point p of ξ, there exists a neighbourhood U of p and a real valued function $c_p: [0,1] \to \mathbb{R}$ such that $F_\tau|_U = F_0|_U + c_p(\tau)$ for all $\tau \in [0,1]$.

Lemma 3.3.5. *Suppose F_0 and F_1 are two immersed Morse functions with a common grim vector field ξ. Assume that $F_0 - F_1$ is locally constant in some neighbourhood of the critical points of ξ. Then there exists a special ξ-path F_τ connecting F_0 and F_1.*

Proof. Define $F_\tau = \tau F_1 + (1-\tau)F_0$. At $\tau = 0,1$ we get F_0, F_1 respectively. As $F_0 - F_1$ is constant in some neighbourhood of each of the critical points, for each $\tau \in [0,1]$, in these neighbourhoods we can write $F_\tau = F_0 + \tau K_p$ for some $K_p \in \mathbb{R}$ that depends on the critical point p. Therefore F_τ satisfies the conditions (IM-1) and (IM-2) of Definition 2.4.1 for each $\tau \in [0,1]$. Thus F_τ is immersed Morse.

Obviously, $\partial_\xi F_\tau > 0$ except at critical points of F_τ. Hence ξ is grim for F_τ. □

A special case of a ξ-path is a path of rearrangement as constructed in the proof of the Rearrangement Theorem 3.2.1.

Definition 3.3.6 (Elementary ξ-path of rearrangement)**.** A special ξ-path $F_\tau\colon\Omega \to \mathbb{R}$ is called an *elementary ξ-path of rearrangement* if there is a unique critical point p such that $c_p(\tau)$ is not identically zero, and there exists a unique critical point q such that the function $\tau \mapsto F_\tau(p) - F_\tau(q)$ crosses zero as τ goes from 0 to 1. Moreover, we assume that for this pair $F_\tau(p) - F_\tau(q)$ has precisely one zero for $\tau \in [0,1]$ and this is a simple zero, i.e. at the zero, the derivative with respect to τ does not vanish.

The following result is an analogue of Cerf's uniqueness of rearrangement.

Proposition 3.3.7. *Suppose F_τ and $\widetilde{F}_\tau$ are two elementary ξ-paths of rearrangement that move the critical value of p_- above the critical value of another critical point p_+ and $F_0 = \widetilde{F}_0$. Assume that F_0, $\widetilde{F}_0$, F_1, and $\widetilde{F}_1$ are excellent immersed Morse functions. Then the paths F_τ and $\widetilde{F}_\tau$ are left-homotopic* (*Definition 3.1.4*)*.*

Proof. In a neighbourhood of p_+, the functions F_τ and $\widetilde{F}_\tau$ are independent of τ. In particular, $F_\tau(p_+) = \widetilde{F}_\tau(p_+) = F_0(p_+)$.

Reparametrise $\widetilde{F}_\tau$ and F_τ if needed (which can be achieved by a left-homotopy) so that $F_\tau(p_-), \widetilde{F}_\tau(p_-) < F_0(p_+)$ for $\tau < 1/2$ and $F_\tau(p_-), \widetilde{F}_\tau(p_-) > F_0(p_+)$ for $\tau > 1/2$. For $\sigma \in [0,1]$, define

$$H_{\sigma,\tau}(z) = (1-\sigma)F_\tau(z) + \sigma\widetilde{F}_\tau(z).$$

We first show that $H_{\sigma,\tau}$ is Morse. We clearly have $\partial_\xi H_{\sigma,\tau} \geq 0$ with equality only at critical points of F_0. That is to say, the only critical points of $H_{\sigma,\tau}$ are those of F_0. We need to show that $H_{\sigma,\tau}$ satisfies both items of Definition 2.4.1 at these points. To this end, recall that F_τ and $\widetilde{F}_\tau$ are special ξ-paths. Hence, for each critical point p' there exists a neighbourhood U such that F_τ and $\widetilde{F}_\tau$ restricted to U differ from F_0 by a constant (depending on τ and on U). Hence, $H_{\sigma,\tau}$ on U differs from F_0 by a constant. The Morse condition on $H_{\sigma,\tau}$ follows from the Morse condition on F_0. Hence, $H_{\sigma,\tau}$ is Morse.

Clearly, for all $\sigma \in [0,1]$, $H_{\sigma,\tau}(p_-) < H_{\sigma,\tau}(p_+)$ for $\tau < 1/2$ and $H_{\sigma,\tau} > H_{\sigma,\tau}(p_+)$ for $\tau > 1/2$. Therefore, a rearrangement occurs at $\tau = 1/2$ for each path $\tau \mapsto H_{\sigma,\tau}$ for a fixed value of σ. Moreover, the path F_τ being an $\mathcal{F}^1$-path means that the stratum $\mathcal{F}^1_\beta$ is intersected transversely, which amounts to saying that

$$\frac{d}{d\tau}(F_\tau(p_-) - F_\tau(p_+))|_{\tau=1/2} > 0.$$

An analogous condition is satisfied by $\widetilde{F}_\tau$, hence it is also satisfied by $H_{\sigma,\tau}$ for all σ. That is, $\tau \mapsto H_{\sigma,\tau}$ intersects the $\mathcal{F}^1$-stratum of $C^\infty(\Omega,\mathbb{R})$ transversely, as required by Definition 3.1.4. □

Now we discuss the case of two paths of rearrangements relative to different vector fields.

Proposition 3.3.8. *Let $F\colon\Omega \to \mathbb{R}$ be an immersed Morse function, let p be a critical point of F with $F(p) = c$. Assume that ξ and $\widehat{\xi}$ are grim vector fields for F. Let F_τ* (*respectively $\widehat{F}_\tau$*)*, be an elementary ξ-path* (*respectively, an elementary $\widehat{\xi}$-path*) *of rearrangement that moves the point p to the level $c' = F_1(p)$, with $F_0 = \widehat{F}_0$.*

Assume also that there exists an interval $[a,b]$ containing c and c', such that ξ and $\widehat{\xi}$ are both tangent to $\mathcal{K}_{\xi,a,b}(p)$ and agree in a neighbourhood of $\mathcal{K}_{\xi,a,b}(p)$. Then F_τ and $\widehat{F}_\tau$ are left-homotopic.

Proof. Let U be a neighbourhood of $\mathcal{K}_{\xi,a,b}(p)$ such that $\xi = \widehat{\xi}$ in U. Define $\xi_\sigma = (1-\sigma)\xi + \sigma\widehat{\xi}$. This need not be a grim vector field in general (a convex combination of grim vector fields is not necessarily grim), but it is certainly grim in U because there $\xi_\sigma = \xi = \widehat{\xi}$.

For each σ, let $\tau \mapsto F_{\sigma,\tau}$ be the path as in the proof of Theorem 3.2.1. To construct these paths, we use the same functions μ and Ψ for all σ: note that ξ_σ does not depend on σ on U, so we can indeed use the same μ. Note that for this construction we do not need ξ_σ to be grim everywhere on Ω, only on U, because the path constructed in Theorem 3.2.1 is supported on U. By construction, all the paths $\tau \mapsto F_{\sigma,\tau}$ are $\mathcal{F}^1$-paths, so $F_{0,\tau}$ and $F_{1,\tau}$ are left-homotopic.

The paths F_τ and $F_{0,\tau}$ need not be equal, but they are left-homotopic by virtue of Proposition 3.3.7. Likewise, the path $\widehat{F}_\tau$ and $F_{1,\tau}$ are left-homotopic for the same reason. Hence, F_τ and $\widehat{F}_\tau$ are left-homotopic as well. □

3.3.2. **Paths on a smooth manifold.** We pass to studying paths on the manifold N. We have two possibilities: either we assume that N is closed, or that N is compact with boundary and the functions on N behave nicely near the boundary. The precise formulation of nice behaviour was given in Subsection 3.1.6. For simplicity, in this subsection we assume that N is closed, leaving necessary adaptations to the reader.

Definition 3.3.9. An *η-path* is a path of Morse functions $f_\tau\colon N \to \mathbb{R}$ with a common gradient-like vector field η. An *η-path* of rearrangement is an η-path f_τ such that precisely two critical points are rearranged along f_τ.

Recall that vector fields on N are denoted using letter η, therefore what was said for ξ-paths for Ω, will be now put into the context of η-paths on N.

Lemma 3.3.10. *Suppose $f_0, f_1\colon N \to \mathbb{R}$ are two Morse functions. Assume η is a gradient-like vector field for both f_0 and f_1. Then the path $f_\tau = (1-\tau)f_0 + \tau f_1$ is an η-path of Morse functions.*

Proof. The same argument that was used in the proof of Lemma 3.3.3 shows that f_0 and f_1 have the same critical points. As $\partial_\eta f_0, \partial_\eta f_1 > 0$ away from the set of critical points, we infer that $\partial_\eta f_\tau > 0$ away from the set of critical points of f_0. In particular, critical points of f_τ are the same as the critical points of f_0.

It remains to show that f_τ is Morse. Take one critical point $q \in N$. Let W^s and W^u be the tangent spaces at q to the stable and unstable manifolds of η respectively. Since f_0 and f_1 are Morse, $D^2 f_0(q)$ and $D^2 f_1(q)$ are both negative (respectively: positive) definite forms on W^s (respectively: on W^u). This shows that $D^2 f_\tau(q) = (1-\tau)D^2 f_0(q) + \tau D^2 f_1(q)$ is negative definite on W^s (respectively: positive definite on W^u). This means that $D^2 f_\tau(q)$ is of full rank. That is, q is a Morse critical point of f_τ. This implies that f_τ has only Morse critical points and that $\partial_\eta f_\tau \geq 0$ with equality only at these points. □

As a consequence we prove a variant of Proposition 3.3.8 for paths starting from different points.

Lemma 3.3.11. *Suppose η is a common gradient-like vector field for two functions $f\colon N \to \mathbb{R}$ and $\widetilde{f}\colon N \to \mathbb{R}$ belonging to the same connected component of $\mathcal{F}^0$. Then any two η-paths of rearrangements starting from f and $\widetilde{f}$ that move a critical point q_- above the critical point q_+ are lax homotopic.*

Moreover, if $f = \widetilde{f}$, then the paths are left-homotopic.

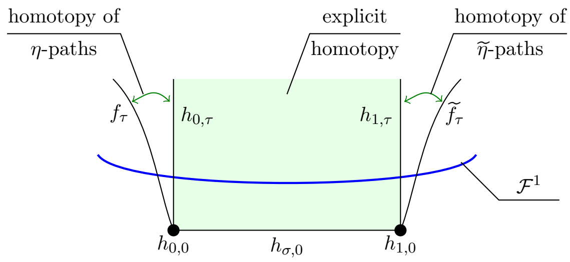

FIGURE 17. Proof of Lemma 3.3.12, schematic of paths.

Proof. Choose two η-paths f_τ and $\widetilde{f}_\tau$ of rearrangements that intersect the $\mathcal{F}^1_\beta$ stratum at $\tau = 1/2$, with $f_0 = f$ and $\widetilde{f}_0 = \widetilde{f}$. Set $h_{\sigma,\tau} = (1-\sigma)f_\tau + \sigma\widetilde{f}_\tau$. We have proved in Lemma 3.3.10 that η is gradient-like for $h_{\sigma,\tau}$. We need to show that the paths $\tau \mapsto h_{\sigma,\tau}$ rearrange precisely the critical points q_- and q_+ for all σ.

As f and $\widetilde{f}$ belong to the same connected component of $\mathcal{F}^0$, for any two critical points q_1, q_2 of f (which are also critical points of $\widetilde{f}$), we have $f(q_1) < f(q_2)$ if and only if $\widetilde{f}(q_1) < \widetilde{f}(q_2)$. It is then easy to check that if q_1, q_2 are two critical points, then:

- if $q_1 = q_-$ and $q_2 = q_+$, then $h_{\sigma,\tau}(q_1) < h_{\sigma,\tau}(q_2)$ for $\tau < 1/2$. There is equality for $\tau = 1/2$ and $h_{\sigma,\tau}(q_1) > h_{\sigma,\tau}(q_2)$ for $\tau > 1/2$.
- if $\{q_1, q_2\} \neq \{q_-, q_+\}$, then $h_{\sigma,\tau}(q_1) < h_{\sigma,\tau}(q_2)$ if and only if $f(q_1) < f(q_2)$.

That is to say, for any $\sigma \in [0,1]$, the path $\sigma \mapsto h_{\sigma,\tau}$ is an η-path of rearrangement intersecting $\mathcal{F}^1_\beta$ at $\tau = 1/2$. As in the proof of Proposition 3.3.7, we check that if the intersections of f_τ and $\widetilde{f}_\tau$ with $\mathcal{F}^1_\beta$ are transverse, then so is the intersection of $\sigma \mapsto h_{\sigma,\tau}$. It follows that $h_{\sigma,\tau}$ is a lax homotopy between the paths f_τ and $\widetilde{f}_\tau$.

Note that the lax homotopy is over the path $h_{\sigma,0} = (1-\sigma)f_0 + \sigma\widetilde{f}_0$. That is, if $f_0 = \widetilde{f}_0$, then, $h_{\sigma,0} = f_0$, proving that $h_{\sigma,\tau}$ is in fact a left-homotopy. $\square$

Another way of creating lax homotopic paths is the following.

Lemma 3.3.12. *Suppose $f\colon N \to \mathbb{R}$ is a Morse function and η is a gradient-like vector field for f. Assume that q_- and q_+ are two consecutive (with respect to the value of f) critical points of f. Suppose also U is an open subset containing a neighbourhood of the unstable manifold of q_- up to a level set above q_+, as well as a neighbourhood of q_+.*

Assume $h_{\sigma,0}$, $\sigma \in [0,1]$, $h_{00} = f$, is a path of excellent Morse functions whose support is disjoint from U. Let $\widetilde{\eta}$ be a gradient-like vector field for h_{10} agreeing with η on U.

If f_τ, $\tau \in [0,1]$ is an η-path of rearrangement starting at f and lifting q_- above q_+ and $\widetilde{f}_\tau$ is an $\widetilde{\eta}$-path of rearrangement starting at h_{10} and lifting q_- above q_+, then the paths $\tau \mapsto f_\tau$ and $\tau \mapsto \widetilde{f}_\tau$ are lax homotopic over $h_{\sigma,0}$.

Proof. The aim of the proof is to create a homotopy of paths that is guided by $h_{\sigma,0}$ away from U and and an η-path inside U. Such a homotopy will be constructed explicitly. In the second step, we will use this homotopy to show that f_τ and $\widetilde{f}_\tau$ are lax homotopic; compare Figure 17.

To begin with, choose $h_{0,\tau}$ to be an η-path of rearrangement starting at h_{00}, lifting q_- above q_+ and supported on U. Assume that rearrangement occurs at $\tau = 1/2$ and $h_{0,\tau}$ intersects $\mathcal{F}^1_\beta$

transversely. For $\tau \in [0,1]$ and $\sigma \in [0,1]$ define

$$h_{\sigma,\tau}(u) = \begin{cases} h_{0,\tau}(u) & u \in U \\ h_{\sigma,0}(u) & u \notin U. \end{cases} \tag{3.3.1}$$

The definition of $h_{\sigma,\tau}$ coincides with previously defined $h_{0,\tau}$ if $\sigma = 0$, and agrees with given $h_{\sigma,0}$ if $\tau = 0$. Indeed, if $\tau = 0$, the right-hand side of (3.3.1) is $h_{\sigma,0}$ away from U, and $h_{0,0}$ in U, but in U, $h_{0,\sigma} = h_{0,0}$. Hence, the right-hand side is equal to $h_{\sigma,0}$. Furthermore, note that the right-hand side of (3.3.1) if $\sigma = 0$ is equal to $h_{0,\tau}$ in U, and $h_{0,0}$ away from U, but $h_{0,\tau}$ is supported on U. Hence, (3.3.1) indeed extends the definition of $h_{\sigma,\tau}$ from the set $\{\tau\sigma = 0\}$.

The paths $\tau \mapsto h_{0,\tau}$ and $\sigma \mapsto h_{\sigma,0}$ have disjoint supports. This means that $h_{\sigma,\tau}$ is smooth and depends smoothly on σ, τ. Next, since $\sigma \mapsto h_{\sigma,0}$ is a path of excellent Morse functions, no critical point of f other than $q_\pm$ has critical value in the interval $[f(q_-), f(q_+)]$. That is to say, the only rearrangements that occur along $\tau \mapsto h_{\sigma,\tau}$ are at $\tau = 1/2$, where the critical values of q_- and q_+ agree. It follows that $\tau \mapsto h_{\sigma,\tau}$ is a path of rearrangements for each σ, and $\sigma \mapsto (\tau \mapsto h_{\sigma,\tau})$ is a lax homotopy between $\tau \mapsto h_{\tau,0}$ and $\tau \mapsto h_{\tau,1}$, and this is clearly a lax homotopy over $h_{\sigma,0}$.

In the remainder of the proof, we will show that $\tau \mapsto h_{1,\tau}$ and $\widetilde{f}_\tau$ are left-homotopic and that $\tau \mapsto h_{0,\tau}$ and f_τ are left-homotopic. Concatenating these homotopies will give us a lax homotopy over $h_{\sigma,0}$ between f_τ and $\widetilde{f}_\tau$.

Proving homotopy between $h_{\tau,1}$ and $\widetilde{f}_\tau$. We aim to show that $h_{1,\tau}$ and $\widetilde{f}_\tau$ are both $\widetilde{\eta}$-paths of rearrangement. For $\widetilde{f}_\tau$, this is one of the assumptions of the lemma. For $h_{1,\tau}$, we note that $\widetilde{\eta}$ is gradient-like for $h_{0,1}$, that is for the starting point. The path $\tau \mapsto h_{1,\tau}$ is supported on U, and η is gradient-like for $h_{1,\tau}$ in U. But on U, $\eta = \widetilde{\eta}$. That is, $\widetilde{\eta}$ is gradient-like for $h_{1,\tau}$ for all τ, hence it is a $\widetilde{\eta}$-path. Any two $\widetilde{\eta}$-paths of rearrangement with the same starting point and rearranging the same pair of critical points are left-homotopic by Lemma 3.3.11.

Proving homotopy between $f_{0,\tau}$ and f_τ. By construction, $\tau \mapsto h_{0,\tau}$ is an η-path of rearrangement. By the assumptions, f_τ also is an η-path. Two η-paths of rearrangement rearranging the same pair of critical points are left-homotopic by Lemma 3.3.11. □

Remark 3.3.13. The proof of Lemma 3.3.12 implies also that the paths f_τ and the concatenation of $\widetilde{f}_\tau$ with $h_{\sigma,0}$ are left-homotopic. Given the two left homotopies at the end of the proof, the statement is equivalent to saying that the paths $\tau \mapsto h_{0,1}$ and the concatenation of paths $h_{\sigma,0}$ and $h_{1,\tau}$ are left-homotopic. But this left homotopy is easily constructed using $h_{\sigma,\tau}$.

3.4. The Cancellation Theorem

The cancellation theorem is about the possibility of simplifying a Morse function by cancelling a pair of critical points. The assumptions are similar to those in the ambient case [Mil65, Section 5]: the two critical points must be of consecutive indices and there should be precisely one trajectory of a Morse–Smale gradient-like vector field connecting the two critical points. In the immersed case, there is one more assumption, namely the two critical points and the unique trajectory between them should all be at the same depth.

In the proof of Path Lifting Theorem 4.5.1, we will need only the case of critical points at depth 1, which was already done in [BP16, Theorem 5.1]. However, we will need a more detailed statement than in [BP16] for the purpose of using uniqueness of death. Therefore, we give a more detailed statement and a more detailed proof.

Theorem 3.4.1 (Cancellation). *Let $F:\Omega \to \mathbb{R}$ be an immersed Morse function with respect to $M = G(N)$, for $G:N \to \Omega$ a generic immersion, with a grim vector field ξ satisfying the Morse-Smale condition. Let p_- and p_+ be two critical points of F that are at depth d, the index of p_- is equal to h, and suppose that the index of p_+ is equal to $h+1$.*

Suppose there is a single trajectory γ of ξ connecting p_- and p_+, and that this trajectory belongs to the d-th stratum. Finally, assume that there are no critical points of F in $F^{-1}[p_-, p_+]$ other than p_- and p_+.

Then the critical points p_- and p_+ can be cancelled. That is, there exists a smooth family ξ_σ, $\sigma \in [0,1]$ of grim vector fields for F, and a path F_τ, $\tau \in [0,2]$ such that:

(C-1) $F_0 = F$, $\xi_0 = \xi$;

(C-2) The vector fields ξ_σ are grim vector fields for F_0, they all satisfy the Morse–Smale conditions, and γ is the single trajectory of ξ_σ connecting p_- to p_+;

(C-3) The vector field ξ_1 is grim for all functions F_τ, $\tau \in [0,1]$, in particular these functions have the same critical points;

(C-4) The path F_τ, $\tau \in [0,1]$ can be chosen to be supported on a predefined grim neighbourhood of γ;

(C-5) In a smaller neighbourhood U of γ, there exist coordinates $x_1, \dots, x_m$, $y_{11}, \dots, y_{dk}$ (with $m = n + k - dk$) and a strictly increasing function $\Upsilon:\mathbb{R} \to \mathbb{R}$ such that for $\tau \in [1,2]$, $F_\tau = \Upsilon \circ H^1_\tau$, where

$$H^1_\tau(x_1, \dots, y_{dk}) = x_1^3 - 3x_1 + 6(\tau - 1)x_1 - x_2^2 - \dots - x_h^2 + x_{h+1}^2 + \dots + x_m^2 + y_{11} + \dots + y_{d1}.$$

In particular, F_2 has the critical points p_- and p_+ removed.

In the statement of Theorem 3.4.1 we specify a path F_τ, $\tau \in [0,2]$. The part $\tau \in [0,1]$ is technical and serves as a preparation for the second part. The actual cancellation occurs along the part of the path with $\tau \in [1,2]$. Recall that $n+k := \dim \Omega$, $n = \dim N$, $m = n+k-dk$, h is the index, and d is the depth.

The path F_τ, $\tau \in [0,2]$ constructed in Theorem 3.4.1 (or more generally, the double path (F_τ, G) with G constant) depends on many choices. In the absolute case, Proposition 3.6.11 shows that different choices lead to left-homotopic paths. Compare Remark 3.6.12.

We also note that with $G_\tau := G$ for $\tau \in [0,2]$, G_τ is a regular immersion for all τ, and F_τ crosses the $\mathcal{F}^1$ stratum only at one point, when the two critical points p_- and p_+ are cancelled. That is (F_τ, G_τ) is a regular double path.

Proof of Theorem 3.4.1. The proof relies on several technical lemmas. We first give the proof of the theorem, and then prove the lemmas. The first lemma is an analogue of [Mil65, Assertion 6]. It specifies a local coordinate system. It also takes care of the property (C-2).

Lemma 3.4.2. *There is a neighbourhood U of γ, and a path ξ_σ of grim vector fields for F, with $\xi_0 = \xi$, satisfying the Morse-Smale condition, such that $\xi_\sigma = \xi$ away from γ, and there is a coordinate system $(x_1, \dots, x_r, y_{11}, \dots, y_{dk})$ on the whole of U, such that the branches are given by $\{y_{i1} = \dots = y_{ik} = 0\}$, and*

$$\xi_1 = (v(x_1), -x_2, \dots, -x_h, x_{h+1}, \dots, x_r, \sum y_{1i}^2, 0, \dots, \sum y_{2i}^2, 0, \dots, \sum y_{di}^2, 0, \dots). \tag{3.4.1}$$

Here v vanishes for precisely two values: 0 and 1, is positive for $v \in (0,1)$, negative elsewhere and $\frac{\partial v}{\partial x_1}$ is 1 near $x_1 = 0$ and -1 at $x_1 = 1$.

Remark. The function v is not specified here, as it is not specified in [Mil65] either. Careful analysis of the proof of [Mil65, Assertion 6, page 55] reveals that we can find the coordinate system for any predefined function v satisfying the above restrictions. The same applies here.

As mentioned above, we defer the proof of the lemma until the end of the proof of the theorem. Note that in the coordinates of the lemma, $p_- = (0,\dots,0)$ and $p_+ = (1,0,\dots,0)$. The coordinate system constructed in Lemma 3.4.2 is defined in a neighbourhood of γ, that is, locally. In the next lemma, we extend it to a grim vector field, that is, to a semi-local coordinate system.

Lemma 3.4.3. *There is a grim neighbourhood U' of γ, containing U, such that the coordinate system of Lemma 3.4.2 can be extended over U', and such that ξ_1 has still the form (3.4.1) in U'.*

Given the lemma, keeping in mind (C-5), we introduce an auxiliary function on U'.

$$H(x_1,\dots,y_{dk}) = x_1^3 - 3x_1 - x_2^2 - \dots - x_h^2 + x_{h+1}^2 + \dots + x_m^2 + y_{11} + \dots + y_{d1}. \tag{3.4.2}$$

Lemma 3.4.4. *On U' we have $\partial_{\xi_1} H \geq 0$ with equality only at p_- and p_+.*

Proof. The result is immediate from (3.4.1). □

As cancelling the critical points of H is straightforward, we aim to replace locally F by H. The result constructs a path F_τ, $\tau \in [0,1]$ with the properties (C-4).

Lemma 3.4.5. *There is a strictly increasing function $\Upsilon\colon \mathbb{R} \to \mathbb{R}$, a neighbourhood U_1 of γ, $U_1 \subseteq U$ and a ξ_1-path of functions F_τ, $\tau \in [0,1]$ such that $F_0 = F$ and $F_1|_{U_1} = \Upsilon \circ H$.*

We defer the proof of the lemma until the end of the proof of the theorem. The rest of the proof of Theorem 3.4.1 is now straightforward. Choose a cut-off function ω supported in U_1, equal to 1 on a neighbourhood $U_2 \subseteq U_1$ of γ. For $\tau \in [1,2]$ we set

$$H_\tau^\omega(x_1,\dots,y_{dk}) = x_1^3 - 3x_1 + 6(\tau-1)\omega x_1 - x_2^2 - \dots - x_h^2 + x_{h+1}^2 + \dots + x_m^2 + y_{11} + \dots + y_{d1},$$

and then

$$F_\tau(z) = \begin{cases} F_1(z) & z \notin U_1 \\ \Upsilon \circ H_\tau^\omega(z) & z \in U_1. \end{cases}$$

A straightforward check using the chain rule and the fact that Υ is strictly increasing shows that F_2 has no critical points in U_2 (which gives U in the statement of Theorem 3.4.1, where $\omega \equiv 1$. In U_2, we also have that $F_\tau = \Upsilon \circ H_\tau^1$. The proof of Theorem 3.4.1 is complete, modulo Lemmas 3.4.2, 3.4.3 and 3.4.5, whose proofs follow. □

Proof of Lemma 3.4.2. Choose local neighbourhoods U_-, U_+ of critical points with local Morse coordinates as in Definition 2.5.1. In particular, ξ has the form (2.5.1). Assume that U_- and U_+ are balls in these coordinates, and that γ intersects ∂U_- only at one point, which we call w_-. Moreover, assume that w_+ is the single point of intersection of γ with ∂U_+. Let $V_0 \subseteq \partial U_-$ be a neighbourhood of w_- in ∂U_- such that any trajectory starting from V_0 eventually hits U_+. Define U_0 to be the set of points on these trajectories, that is the set of x such that the trajectory through x hits V_0 in the past and U_+ in the future.

Extend the coordinate system from U_- to U_0 in the following way. For a point $x \in U_0$, a trajectory of ξ through x hits a point $w_x \in V_0$. Near w_x we already have coordinates from U_-. Then we assign the coordinates to x using the flow ξ in such a way that ξ has the form (2.5.1) on the whole of $U_- \cup U_0$. For simplicity, we will assume that $\overline{U}_0 \cap U_-$, respectively $\overline{U}_0 \cap U_+$ lies on one level set of F_τ. Denote this level set c_-, respectively c_+.

On the intersection $\overline{U}_0 \cap \overline{U}_+$ we have two coordinate systems. One comes from U_0 and the other comes from U_+. If the two coordinate systems match (up to possibly shifting the value of x_1), then the lemma is proved with $\xi_\sigma \equiv \xi$. If not, we need to find a suitable isotopy

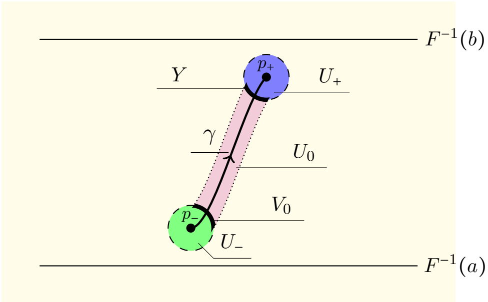

FIGURE 18. Notation of proof of Cancellation Theorem 3.4.1.

between the two coordinate systems and alter the vector field ξ using the Isotopy Insertion Lemma 2.5.8 to induce this isotopy by the flow. The proof builds on the proof of [Mil65, Assertion 6], and uses it in the last part.

Write $Y = \overline{U}_0 \cap \overline{U}_+$. The two coordinate systems specify maps $h_0{:}Y \to Z_0$, $h_1{:}Y \to Z_0'$, where $Z_0, Z_0' \subseteq \mathbb{R}^{n+k-1} = \mathbb{R}^{h-1} \oplus \mathbb{R}^{m-h} \oplus \mathbb{R}^k \oplus \dots \oplus \mathbb{R}^k$, with d copies of $\mathbb{R}^k$. Both h_0 and h_1 take the unique intersection point of γ with Y to $(0, \dots, 0) \in \mathbb{R}^{n+k-1}$.

We aim to show that there is a compactly supported homotopy of $h_1 \circ h_0^{-1}$ to a map that is the identity on a neighbourhood of $(0, \dots, 0) \in Z_0$. To be more precise, denote the first summand of the space $\mathbb{R}^{n+k-1}$ by L_0^s ('s' for 'stable'), the second by L_0^u ('u' for 'unstable') and all the others by $L_1, \dots, L_d$. Set $L_0 = L_0^s \oplus L_0^u$. Note that the choice of coordinate systems on U_- and U_+ involves choosing a numbering of branches. We choose these two numberings to be consistent, that is, the flow of ξ maps the i-th branch in U_- to the i-th branch in U_+.

We can summarise the properties of $h_1 \circ h_0^{-1}$ in the following lemma.

Lemma 3.4.6. *For any subset $\mathfrak{p}$ of $\{1, \dots, d\}$, let $L_{\mathfrak{p}} = \bigoplus_{i \in \mathfrak{p}} L_i$. Then for any $\mathfrak{p}$ the map $h_1 \circ h_0^{-1}$:*

(H-1) maps $L_0 \oplus L_{\mathfrak{p}} \to L_0 \oplus L_{\mathfrak{p}}$;
(H-2) Moreover, the image of L_0^u in L_0 intersects L_0^s transversely and only at 0.

Proof. Item (H-1) follows from the fact that $L_0 \oplus L_{\mathfrak{p}}$ is the image under h_0 of the intersection of branches $\bigcap_{i \notin \mathfrak{p}} Y_i$. Item (H-2) follows from the fact that the flow is Morse–Smale. □

The proof of Lemma 3.4.2 consists of subsequent simplifications of the composition $h_1 \circ h_0^{-1}$, done on smaller and smaller neighbourhoods of $(0, \dots, 0) \in Z_0$. The first result in this direction is reduction to the linear case.

Lemma 3.4.7 (Reduction to the Linear Case)**.** *There exists an open subset $Z_1 \ni \{0\}$, $Z_1 \subseteq Z_0$, and an isotopy h_τ^1 with $h_0^1 = h_1 \circ h_0^{-1}$ such that h_τ^1 is independent of τ in a neighbourhood of the boundary of $\overline{Z}_0$ and $h_1^1 = D(h_1 \circ h_0^{-1})(0)$ identically on Z_1. Moreover for each τ, h_τ^1 satisfies items (H-1) and (H-2).*

Proof of Lemma 3.4.7. To simplify the proof, endow $\mathbb{R}^{n+k-1}$ with an inner product. Suppose V_ℓ is a ball with radius $1/\ell$ and centre 0. For sufficiently large ℓ, the closure of the ball V_ℓ is contained in Z_0. The isotopy will be independent of τ away from V_ℓ.

Let ϕ_ℓ be a cut-off function supported on V_ℓ, equal to 1 on $V_{2\ell}$ and whose derivative is bounded by $\|D\phi_\ell\| \le 4\ell$. With $A := D(h_1 \circ h_0^{-1})(0)$, define

$$h_{\ell,\tau}^1 = A\tau\phi_\ell + (1 - \tau\phi_\ell) h_1 \circ h_0^{-1}.$$

Note that for all choices of ℓ, item (H-1) is satisfied because it is satisfied by A and $h_1 \circ h_0^{-1}$. Condition (H-2) is open in the C^1-norm. Thus, we need to show that $h^1_{\ell,\tau}$ is close to $h_1 \circ h_0^{-1}$ in the C^1-norm. This is equivalent to saying that $\kappa_\ell := \phi_\ell(A - h_1 \circ h_0^{-1})$ converges to 0 in the C^1-norm. By the Taylor formula, there is a constant $c_2 < \infty$ such that $\|Ax - h_1 \circ h_0^{-1}(x)\| \leq c_2\|x\|^2$. Thus, if ϕ_ℓ is supported on the ball of radius $1/\ell$, we know that the C^0-norm of κ_ℓ is bounded by c_2/ℓ^2. Looking at the derivatives, we note

$$\|D\kappa_\ell(x)\| = \|D\phi_\ell \cdot (A - h_1 \circ h_0^{-1}) + \phi_\ell D(A - h_1 \circ h_0^{-1})\| \leq \|D\phi_\ell \cdot (A - h_1 \circ h_0^{-1})\| + \|\phi_\ell D(A - h_1 \circ h_0^{-1})\|.$$

The first term is bounded by $4\ell c_2/\ell^2 = 4c_2/\ell$, because we control the derivative of ϕ_ℓ via $\|D\phi_\ell\| \leq 4\ell$, and $\|A - h_1 \circ h_0^{-1}\| \leq c_2/\ell^2$. The second term is bounded by $2c_2/\ell$, because $|\phi_\ell| \leq 1$, and by differentiating the Taylor approximation $\|D(A - h_1 \circ h_0^{-1})\| \leq 2c_2\|x\|$, which on the ball of radius $1/\ell$ is bounded by $2c_2/\ell$. Thus altogether $\|D\kappa_\ell(x)\| \leq 6c_2/\ell$.

Thus, for sufficiently large ℓ, $h^1_{\ell,\tau}$ satisfies (H-2). We then declare $Z_1 := V_{2\ell}$ and $h^1_\tau := h^1_{\ell,\tau}$. □

The next step reduces the linear map to a linear map preserving the decomposition of $\mathbb{R}^{n+k-1}$ into $L_0 \oplus L_1 \oplus \cdots \oplus L_d$.

Lemma 3.4.8 (Reduction to block structure)**.** *There is an open neighbourhood $Z_2 \subseteq Z_1$ of 0, a path h^2_τ of isotopies, supported on a compact set contained in Z_1 and containing Z_2 such that $h^2_0 = h^1_1$, and h^2_τ satisfies items (H-1) and (H-2). Moreover, h^2_1 preserves the decomposition $\mathbb{R}^{n+k-1} = L_0 \oplus L_1 \oplus \cdots \oplus L_d$.*

Proof. With the decomposition $\mathbb{R}^{n+k-1} = L_0 \oplus \cdots \oplus L_d$, we can write A_{ij}, $i, j = 0, \dots, d$ for the submatrix giving the map from L_i to L_j. With this notation, we have

- A_{ij} is zero except if $i = j$ or $j = 0$;
- A_{00} maps L^u_0 to a subspace transverse to L^s_0.

In fact, the first part follows from (H-1), the second from (H-2). As A is linear, these two conditions are actually equivalent to the conditions (H-1) and (H-2). This form of A is almost what is required by Lemma 3.4.8, except for the terms A_{i0} that spoil the block structure of A. Write $A = A_0 + A_1$, where A_0 is the sum of diagonal blocks A_{ii} and A_1 is the sum of the off-diagonal blocks A_{i0} $(i > 0)$.

Define $h^2_\tau(x) = A_0 x + (1 - \tau\phi(x))A_1 x$ for a cut-off function ϕ. We require that ϕ be supported on Z_5, and that it is equal to 1 on a smaller neighbourhood Z_6 of $0 \in \mathbb{R}^{n+k-1}$ Note that h^2_τ is no longer linear (the off-diagonal terms in a matrix representing h^2_τ depend on x. Nevertheless, conditions (H-1) and (H-2) are easily seen to be satisfied. Indeed, condition (H-1) follows from the block structure. The transversality condition (H-2) involves only the A_{00} block, which is the same for h^2_τ and A. □

We can now modify the map A to the identity acting block by block. Recall that Z_5 is the neighbourhood of $0 \in \mathbb{R}^{n+k}$ such that $h^2_1(x)$ is the linear map given by the block matrix. Shrink Z_5 if needed to make sure it has a product structure, i.e. $Z_5 = (Z_5 \cap L_0) \times \cdots \times (Z_5 \cap L_d)$.

Lemma 3.4.9 (Blockwise modification)**.** *For any $i = 0, \dots, d$, there is an isotopy $h^3_{i,\tau}$ of restrictions of $Z_5 \cap L_i$, supported such that $h^3_{i,0} = A_{ii}$ and $h^3_{i,1}$ is the identity on an open subset $Z_{7i} \subseteq Z_5 \cap L_i$.*

The isotopy preserves conditions (H-1) and (H-2).

Proof. The case $i = 0$ corresponds to the absolute case (classical cancellation theorem). This is the statement of [Mil65, Theorem 5.6]. Suppose $i > 0$. We need to find an isotopy $h^3_{i,\tau}$ such that $h^3_{i,0}$ is A_{ii}, $h^3_{i,1}$ on a smaller subset is the identity. This is a straightforward problem in linear algebra, which we leave to the reader. □

By taking the direct sum of the maps $h^3_{i,\tau}$, we obtain the following result.

Corollary 3.4.10. *There is an isotopy h^3_τ supported on Z_5 such that $h^3_0 = h^2_1$ and h^3_1 is the identity on $Z_6 \coloneqq Z_{60} \times \dots \times Z_{6d}$. The isotopy preserves conditions (H-1) and (H-2).*

Summarizing the proof of Lemma 3.4.2, we concatenate the isotopies h^1_τ (reducing to the linear part), h^2_τ (reducing to the block part), and h^3_τ (reducing to the identity), to obtain an isotopy whose starting map is $h_1 \circ h_0^{-1}$, and whose end map is the identity on Z_6. Denote this isotopy by h_σ. Note that h_σ is the identity near the boundary of Z_0. We inject this isotopy into U_0 by modifying the vector field ξ. As a result, we obtain a family of vector fields ξ_σ, such that the map induced by ξ_σ is h_σ. The vector field ξ_σ is tangent to the branches (because of (H-1)), by (H-2) it does not create any new trajectories between p_- and p_+, and it is Morse–Smale, as desired.

The end of the proof of Lemma 3.4.2 follows the same lines as the proof of Milnor's cancellation theorem [Mil65, Page 57]. Let U_6 be the subset of U_0 made out of trajectories that pass through $h_0(Z_6)$. The flow of ξ_1 induces the identity in the coordinates on U_- and U_+. Hence, we can glue these coordinates in such a way that ξ_1 has the form (3.4.1) in these coordinates. □

Remark 3.4.11. The proof of Lemma 3.4.2 required a modification of the vector field ξ by injecting an isotopy h_σ that is a concatenation of the isotopies h^1_τ, h^2_τ, and h^3_τ. Suppose ξ_σ is a vector field obtained by injecting only an isotopy h_τ, for $\tau \in [0,\sigma]$. By the property (H-2), which is preserved by h_τ, there is precisely one trajectory of ξ_σ on the first stratum that connects p_- with p_+. This property will be important when discussing the uniqueness of death.

We are now in position to prove Lemma 3.4.3.

Proof of Lemma 3.4.3. Recall that $\overline{U}_0 \cap \overline{U}_+$ belongs to a single level set $F^{-1}(c_+)$. Choose a pair of grim neighbourhoods X_- and X_+ of $\mathbb{M}_{\mathrm{a}}(p_-)$ and $\mathbb{M}_{\mathrm{d}}(p_+)$ respectively, in $F^{-1}[a,b]$, such that $X_- \cap F^{-1}(c_+)$ intersects $X_+ \cap F^{-1}(c_+)$ along an open subset in $F^{-1}(c_+)$ whose closure is contained in Z_7. By shrinking X_- and X_+ if needed, we may assume that there are coordinates in both X_- and X_+, on which ξ_1 has the form (3.4.1). The flow of ξ_1 induces the identity in these coordinates by Lemma 3.4.2, hence the two coordinate systems match together to produce a coordinate system on $X_- \cup X_+$. □

We can now prove Lemma 3.4.5, and by this conclude the proof of Cancellation Theorem 3.4.1.

Proof of Lemma 3.4.5. We use the notation of the proof of Lemma 3.4.3. In particular, $X_- \cup X_+$ is a grim neighbourhood of $\mathbb{M}_{\mathrm{a}}(p_-) \cup \mathbb{M}_{\mathrm{d}}(p_+)$ in $F^{-1}[a,b]$. Choose a', b' in such a way that $F(a) < F(a') < F(p_-) < F(p_+) < F(b') < F(b)$. Let $X_0 = (X_+ \cup X_-) \cap F^{-1}[a',b']$. Let X_1 be a grim neighbourhood of $\mathbb{M}_{\mathrm{a}}(p_-) \cup \mathbb{M}_{\mathrm{d}}(p_+)$ in $F^{-1}[a',b']$ such that $\overline{X_1 \cap F^{-1}(a')} \subseteq X_0$. Such an X_1 exists because of Lemma 2.7.10. Define $\mu\colon F^{-1}(a') \to [0,1]$ to be a cut-off function equal to 1 on $X_1 \cap F^{-1}(a')$ and to 0 away from $X_0 \cap F^{-1}(a')$. Extend μ to a ξ invariant function on the whole of $F^{-1}[a,b]$.

Set ζ_- to be the infimum of H on $X_0 \times [0,1]$ (the last coordinates being for the τ variable), and set ζ_+ to be the supremum of H. Choose $\Upsilon\colon \mathbb{R} \to \mathbb{R}$ in such a way that $a < \Upsilon(\zeta_-) <$

$\Upsilon(\zeta_+) < b$, and Υ is linear near $F(p_-)$ and $F(p_+)$ with derivative 1. On $F^{-1}[a',b']$ we define the function F_τ via:

$$H_\tau = (1-\tau\mu)F + \tau\mu\Upsilon\circ H. \tag{3.4.3}$$

We have $\partial_{\xi_1} H_\tau = (1-\tau\mu)\partial_{\xi_1}F + \tau\mu\partial_{\xi_1}(\Upsilon\circ H)$. This is clearly a non-negative number, because $\partial_{\xi_1}F$ and $\partial_{\xi_1}(\Upsilon\circ H)$ are nonnegative, and positive except at p_- and p_+. Moreover, $\mu\equiv 1$ in a neighbourhood of p_- and p_+, so the path H_τ is actually a special ξ_1-path (see Definition 3.3.4). Furthermore, H_τ on $F^{-1}(a')$ is strictly greater than a, and on $F^{-1}(b')$ is strictly smaller that b. This allows us to define a function F_τ in the following way.

We set $F_\tau = F$ away from $F^{-1}[a,b]$ and $F_\tau = H_\tau$ on $F^{-1}[a',b']$. Suppose $z\in F^{-1}(a,a')$ (the case $z\in F^{-1}(b',b)$ is analogous). As F has no critical points in $(a,a')\cup(b',b)$, there are unique points $z_0\in F^{-1}(a)$ and $z_1\in F^{-1}(a')$ such that ξ_1 flows from z_0 through z to z_1. Suppose t_1 is the time taken to get from z_0 to z_1 and t_0 is the time to get from z to z_1. We set

$$F_\tau(z) \coloneqq \frac{t_0}{t_1}F(z_0) + \frac{t_1-t_0}{t_1}H_\tau(z_1).$$

This interpolation is continuous. There might be non-smooth points at the level sets a,a',b',b, but in a presence of a vector field, we can smooth them in a standard way; see Lemma 2.8.4. □

3.5. Paths of birth

We defined paths of birth in Definition 3.1.3. Now we are going to specify a concrete class of paths of birth, so-called elementary paths of birth. The construction is due to Cerf [Cer70, Section III.2.1, p. 66]. Fix $h=0,\dots,n$, $\varepsilon>0$, and $\tau\in[0,1]$. Define the function $\mathfrak{r}_\tau\colon\mathbb{R}^n\to\mathbb{R}$ by

$$\mathfrak{r}_\tau(x_1,\dots,x_n) = -x_1^2-\dots-x_h^2+x_{h+1}^2+\dots+x_{n-1}^2+x_n^3-(2\tau-1)\varepsilon x_n.$$

The function $\mathfrak{r}_\tau$ is a standard bifurcation of an A_2 singularity: it has no critical points for $\tau<1/2$, it has a single critical point for $\tau=1/2$ and two Morse critical points of indices h and $h+1$, respectively, for $\tau>1/2$.

Choose a bump function $\omega\colon\mathbb{R}^n\to[0,1]$, identically equal to 1 in a neighbourhood of 0 and having compact support. Define

$$\mathfrak{r}_{\omega,\tau}(x_1,\dots,x_n) = -x_1^2-\dots-x_h^2+x_{h+1}^2+\dots+x_{n-1}^2+x_n^3-(2\tau\omega-1)\varepsilon x_n. \tag{3.5.1}$$

Choose sufficiently small $\varepsilon>0$. Then $\mathfrak{r}_{\omega,\tau}$ has no critical points in $\omega^{-1}(0,1)$. In particular, $\mathfrak{r}_\tau$ and $\mathfrak{r}_{\omega,\tau}$ have the same critical points, and they agree in the neighbourhood of these critical points; see [Cer70, Section III.1.1].

Define $\psi\colon\mathbb{R}^n\to\mathbb{R}^n$ by

$$\psi(x_1,\dots,x_n) = (x_1,\dots,x_{n-1},\mathfrak{r}_0(x_1,\dots,x_n)),$$

so that $\mathfrak{r}_0\circ\psi^{-1} = x_n$.

Definition 3.5.1. The path $\mathfrak{u}_\tau = \mathfrak{r}_{\omega,\tau}\circ\psi^{-1}\colon\mathbb{R}^n\to\mathbb{R}$ is called the *elementary path of birth on* $\mathbb{R}^n$ (of index h).

The path $\mathfrak{u}_\tau$ is a path such that $\mathfrak{u}_0 = x_n$ and $\mathfrak{u}_\tau$ acquires a pair of critical points. Moreover, as ω has compact support, $\mathfrak{u}_\tau = \mathfrak{u}_0$ away from a compact set in $\mathbb{R}^n$. This motivates the following definition; compare [Cer70, Section III.1.2].

Definition 3.5.2 (Elementary path of birth)**.** A path $f_\tau\colon N \to \mathbb{R}$ is an *elementary path of birth* at q_0 (of index h), if there exists an open set U containing q_0, and local coordinates $x_1, \dots, x_n$ in U such that

- $q_0 = (0, \dots, 0)$;
- $f_\tau = f_0$ away from U;
- $f_\tau(x_1, \dots, x_n) = \Upsilon \circ \mathfrak{u}_\tau(x_1, \dots, x_n)$

for some smooth, increasing function $\Upsilon\colon \mathbb{R} \to \mathbb{R}$.

Remark 3.5.3. By choosing appropriate ω (with small support) we can assume that $\mathfrak{u}_\tau = \mathfrak{u}_0$ away from an arbitrarily small neighbourhood of $(0, \dots, 0)$.

We quote two results of Cerf.

Proposition 3.5.4 (see [Cer70, Proposition III.1.3.1])**.** *Any path of birth f_τ can be given an arbitrarily small perturbation to another path of birth $\widetilde{f}_\tau$ such that for some $\delta_0 > 0$, $\widetilde{f}_\tau$ restricted to $\tau \in [1/2 - \delta_0, 1/2 + \delta_0]$ is an elementary path of birth.*

In the proof of Proposition 3.5.4, Cerf uses the fact that through each point of $\mathcal{F}^1_\alpha$ one can pass an elementary path of birth. It follows promptly that we can assume that the paths $\widetilde{f}_\tau$ and f_τ can be assumed to cross $\mathcal{F}^1_\alpha$ at the same point. This leads to the following statement.

Corollary 3.5.5. *Any path of birth is $\mathcal{F}^1$-homotopic to an elementary path of birth.*

The second result shows that paths of birth are, up to lax homotopy, determined by the index of the critical points that are created, and the connected component of the level set, in which they are constructed.

Proposition 3.5.6 (see [Cer70, Corollaire III.1.3.2, page 67])**.** *Suppose two paths of birth $\widetilde{f}_\tau$ and f_τ have f_0 and $\widetilde{f}_0$ in the same connected component of $\mathcal{F}^0$ and f_1 and $\widetilde{f}_1$ in the same connected component of $\mathcal{F}^0$. Suppose the two paths create critical points q_0, respectively $\widetilde{q}_0$, on the same component of the fixed level set $f^{-1}_{1/2}(c)$, and q_0 and $\widetilde{q}_0$ are of the same index. Then f_τ and $\widetilde{f}_\tau$ are lax homotopic.*

Our statement slightly differs from Cerf's result, hence we indicate how to use Cerf's result to prove Proposition 3.5.6.

Proof. Suppose g_0 and g_1 are the two points of intersection of f_τ and $\widetilde{f}_\tau$ with $\mathcal{F}^1$. By Corollary 3.5.5, there exists elementary paths of birth $f_{0,\tau}$ and $f_{1,\tau}$, passing through g_0, respectively g_1, and such that $f_{0,\tau}$ and f_τ are $\mathcal{F}^1$-homotopic and $f_{1,\tau}$ and $\widetilde{f}_\tau$ are $\mathcal{F}^1$-homotopic. It is enough to show that $f_{0,\tau}$ and $f_{1,\tau}$ are lax homotopic. This follows promptly from [Cer70, Corollaire III.1.3.2]. □

3.6. Paths of death

Throughout Section 3.6 we will assume that $f\colon N \to \mathbb{R}$ is a Morse function, and that q_- and q_+ are critical points of indices h_- and $h_+ = h_- + 1$ respectively. Set $a = f(q_-)$, $b = f(q_+)$. Assume that $a < b$ and there are no critical points of f with critical values in (a, b). Fix $c \in (a, b)$.

First, we recall the definition of Cerf's ascending and descending caps (what he calls, *en français, nappes ascendantes et descendantes*). These are an abstract form of (parts of) the ascending and descending manifolds of critical points. Next, we recast the statements of Cerf in the language of vector fields.

3.6.1. **Paths of death via ascending and descending caps.** This subsection is based on [Cer70, Section III.2].

Definition 3.6.1 (Pair of caps)**.** A *pair of caps* is a pair of smooth embeddings $\varphi_D: D^{h_+} \to N$, $\varphi_A: D^{n-h_-} \to N$, where D^{h_+} and D^{n-h_-} are closed discs with $\varphi_A(0) = q_-$, $\varphi_D(0) = q_+$, and the composition $f \circ \varphi_D$ (respectively $f \circ \varphi_A$) is a quadratic function equal to c at the boundary $\operatorname{bd} D^{h_+}$ (respectively $\operatorname{bd} D^{n-h_-}$) and having a global maximum b at zero (respectively, having a global minimum a at zero).

A pair of caps shall usually be denoted by (D, A) with $D := \varphi_D(D^{h_+})$, $A := \varphi_A(D^{n-h_-})$.

Definition 3.6.2 (Caps in good position)**.** Pair of caps (D, A) is called a *pair of caps in good position*, in short a *pair of good caps*, if the boundaries of D and A intersect transversely at a single point in $f^{-1}(c)$.

Example 3.6.3. Let η be a gradient-like vector field for f. We set $D = W^s_{q_+}(\eta) \cap f^{-1}[c, b]$, $A = W^u_{q_-}(\eta) \cap f^{-1}[a, c]$, that is, D and A are parts of the stable (respectively, unstable) manifolds of q_+ (respectively q_-). Then (D, A) form a pair of caps. Moreover, (D, A) are in a good position if and only if η satisfies the assumptions of Milnor's Cancellation Theorem, that is, if $W^s_{q_+}(\eta)$ intersects $W^u_{q_-}(\eta)$ transversely along a single trajectory connecting q_- and q_+; compare [Mil65, Theorem 5.4].

Definition 3.6.4. If η satisfies the assumptions of Milnor's Cancellation Theorem, the pair (D, A) of Example 3.6.3 is called the pair of good caps *associated with* η and denoted (D_η, A_η).

Remark. A gradient vector field ∇f (for some metric) gives rise to a pair of caps (D, A) and any pair of caps arises from a gradient vector field; see [Cer70, Section III.2.1]. As any gradient-like vector field is an actual gradient for some choice of metric, we can rephrase Cerf's statement as: 'any pair of caps (D, A) can be constructed via a suitable gradient-like vector field η'.

Suppose (D, A) is a pair of good caps. The following is a key notion for constructing elementary paths of death; see [Cer70, page 70].

Definition 3.6.5 (Double neighbourhood adapted to (D, A))**.** A map $\varphi: D^n \to N$ is a *double neighbourhood adapted to a pair of good caps* (D, A), if:

- $U = \varphi(D^n)$ contains $D \cup A$;
- with coordinates $x_1, \dots, x_n$ on U induced from coordinates on D^n, we have $f \circ \varphi = \Upsilon \circ \mathfrak{u}_1$ for some increasing function $\Upsilon: \mathbb{R} \to \mathbb{R}$;
- if ∇ denotes the gradient in U in the metric where $x_1, \dots, x_n$ are orthonormal coordinates, then the caps D and A are given by the ascending and descending manifolds of ∇f respectively.

Remark. Throughout the paper, a double neighbourhood adapted to (D, A) is always chosen with respect to the chosen standard model as described in [Cer70, Section III.2.2].

Suppose $\varphi: U \to N$ is a double neighbourhood adapted to (D, A).

Definition 3.6.6 (compare [Cer70, Section III.2.3])**.** The path of functions $f_\tau(u)$ given by $f_\tau(u) = f(u)$ if $u \neq U$ and $f_\tau(u) = \Upsilon \circ \mathfrak{u}_{1-\tau}(\varphi^{-1}(u))$ is called the *elementary path of death associated with* φ.

Lemma 3.6.7 ([Cer70, Corollary, page 72]). *Any path of death f_τ can be written as a composition $f^1_\bullet \cdot \widetilde{f}_\bullet \cdot f^0_\bullet$ such that f^1 and f^0 are paths of excellent Morse functions and $\widetilde{f_\tau}$ is an elementary path of death supported on an arbitrarily small neighbourhood of some pair (D, A) of caps in good position.*

Lemma 3.6.8 (Homotopy of paths of death). *Suppose φ_σ, $\sigma \in [0,1]$ is a smooth family of maps from $D^n \to N$ satisfying the conditions of Definition 3.6.5. Let $\tau \mapsto h_{\sigma,\tau}$ be the path of Definition 3.6.6 for φ_σ starting from a given path $h_{\sigma,0} = g_\sigma$ in $\mathcal{F}^0$. Then $h_{\sigma,\tau}$ is a left homotopy between paths $h_{0,\tau}$ and $h_{1,\tau}$.*

Proof. First, it is routine to see that $h_{\sigma,\tau}$ depend smoothly on the parameters σ and τ. By construction, any path $\tau \mapsto h_{\sigma,\tau}$ crosses $\mathcal{F}^1$ transversely, and at precisely one parameter value of τ (not equal to 0 and 1). This indicates that $\tau \mapsto h_{\sigma,\tau}$ is an $\mathcal{F}^1$-path. Moreover, the endpoints $h_{\sigma,1}$ all belong to $\mathcal{F}^0$, so they are in the same connected component to $\mathcal{F}^0$. This shows that $h_{\sigma,\tau}$ is a homotopy between $h_{0,\tau}$ and $h_{1,\tau}$ as desired. □

Let $\mathcal{P}$ denote the subspace of all double neighbourhoods adapted to a pair of good caps (D, A) with the topology given by uniform convergence. Let $\mathcal{P}(D, A)$ be the subspace of all double neighbourhoods adapted to a fixed pair of good caps (D, A) (Cerf denotes this space by $\mathcal{P}'$). The following result of Cerf is essential to prove uniqueness of deaths.

Lemma 3.6.9 (see [Cer70, page 70, item 3c]). *The space $\mathcal{P}(D, A)$ is nonempty and path-connected. If $\mathcal{N}$ denotes the space of all pairs of good caps (D, A), then the assignment $\mathcal{P} \to \mathcal{N}$ is a locally trivial fibration with fibres $\mathcal{P}(D, A)$.*

Combining Lemma 3.6.8 with Lemma 3.6.9, we obtain the following result.

Theorem 3.6.10 (Uniqueness of Death, geometric version). *Assume g_σ is a path of excellent Morse functions. Suppose $\varphi_{D,\sigma}: D^{h_+} \to N$ and $\varphi_{A,\sigma}: D^{n-h_-} \to N$ are smooth families of functions (with $\sigma \in [0,1]$) such that $D_\sigma := \varphi_{D,\sigma}(D^{h_+})$ and $A_\sigma := \varphi_{A,\sigma}(D^{n-h_0})$ form a pair of good caps for g_σ all $\sigma \in [0,1]$. Suppose φ_0 and φ_1 are two double neighbourhoods adapted to (D_0, A_0) and (D_1, A_1), respectively.*

Let $h_{0,\tau}$ and $h_{1,\tau}$ be the paths of death associated with φ_0 and φ_1 as in Definition 3.6.6, where $h_{0,0} = g_0$ and $h_{1,0} = g_1$. Then the paths $\tau \mapsto h_{0,\tau}$ and $\tau \mapsto h_{1,\tau}$ are lax homotopic over g_σ.

Proof. The assumptions imply that there is a path ζ_σ in the space $\mathcal{N}$, connecting pairs of good caps (D_σ, A_σ), the path being given by $(\varphi_{D,\sigma}, \varphi_{A,\sigma})$. By Lemma 3.6.9, there is a lift of ζ_σ to a path $\widetilde{\zeta_\sigma}$ in $\mathcal{P}$, whose endpoints are φ_0 and φ_1. By Lemma 3.6.8 the paths constructed using φ_0 and φ_1 are lax homotopic over g_σ. □

Remark. Cerf's statement on unicity of death [Cer70, Proposition III.2.4.4] has slightly different statement, because Cerf uses extra assumptions on dimension and simple connectivity of the spaces to guarantee that the rather complicated assumptions of Theorem 3.6.10 are satisfied. But our proof follows his proof very closely.

3.6.2. **Paths of death via vector fields.** Our Cancellation Theorem 3.4.1 (where the absolute case is obtained by setting $d = 0$) creates a certain path of functions along which two critical points are destroyed. The initial piece of data is a choice of a gradient-like vector field. We aim to combine the two approaches.

Assume $f: N \to \mathbb{R}$ is a Morse function, q_- and q_+ are two critical points of indices h and $h+1$ respectively, with no critical points in $f^{-1}[f(q_-), f(q_+)]$ apart from q_-, q_+, and η is a

Morse–Smale vector field, gradient-like for f, such that there exists a unique trajectory of η connecting q_- and q_+. Let f_τ, $\tau \in [0,2]$, be the path of functions from Theorem 3.4.1, and let η_σ, for $\sigma \in [0,1]$, be the path of vector fields from that theorem: that is η_0 is the original vector field, and η_1 is the vector field obtained after applying the modifications required by Lemma 3.4.2; compare Remark 3.4.11. Recall that in the absolute case we use the notation f, η, instead of F, ξ.

Proposition 3.6.11. *Let (D_σ, A_σ) be the good pair of caps constructed via η_σ as in Example 3.6.3. Let $\widetilde{f}_\tau$ be an elementary path of death with respect to the good pair of caps (D_0, A_0) with $\widetilde{f}_0 = f_0$. Then $\widetilde{f}_\tau$ is left-homotopic to f_τ.*

Proof. Let $\tau \mapsto h_{\sigma,\tau}$ be the elementary path of death starting from f_0 with respect to the good pair of caps (D_σ, A_σ). By Lemma 3.4.2 (and Remark 3.4.11), the assumptions of Theorem 3.6.10 are satisfied, so $h_{0,\tau}$ and $h_{1,\tau}$ are lax homotopic over $g_\sigma := h_{\sigma,0}$. As $h_{\sigma,0} = f_0$, we conclude that $h_{0,\tau}$ and $h_{1,\tau}$ are left-homotopic. Also note that $\widetilde{f}_\tau = h_{0,\tau}$.

We aim to prove that f_τ and $h_{1,\tau}$ are left-homotopic. Note that Lemma 3.4.2 in the proof of Theorem 3.4.1 constructs an explicit coordinate system and the sets X_- and X_+ that are a double neighbourhood adapted to (D_1, A_1).

In this neighbourhood, the path $f_{1+\tau}$, $\tau \in [0,1]$, is – by construction – an elementary path of death; compare item (C-5) of Theorem 3.4.1. By Theorem 3.6.10, $f_{1+\tau}$, for $\tau \in [0,1]$, is lax homotopic to $h_{0,\tau}$ over f_τ, $\tau \in [0,1]$.

This implies that the paths $\widetilde{f}_\tau$ and $f_{1+\tau}$ are lax homotopic over f_τ. That is, there exists a homotopy $\widetilde{h}_{\sigma,\tau}$ such that $\widetilde{h}_{\sigma,0} = f_\sigma$, $\widetilde{h}_{0,\tau} = \widetilde{f}_\tau$, and $\widetilde{h}_{1,\tau} = f_{1+\tau}$. Reparametrizing this homotopy to

$$h'_{\sigma,\tau} = \begin{cases} \widetilde{h}_{2\sigma\tau,0} & \tau \le \frac{1}{2} \\ \widetilde{h}_{\sigma,2\tau-1} & \tau \ge \frac{1}{2} \end{cases}$$

yields a left-homotopy between $\widetilde{f}_\tau$ and f_τ (the latter path being for $\tau \in [0,2]$). □

Remark 3.6.12. The construction of the path in the proof of Theorem 3.4.1 involved many choices, like the choice of isotopy h_τ, of neighbourhoods X_-, X_+, etc. Proposition 3.6.11 shows in particular that (at least in the absolute case) the choices do not matter, up to left-homotopy of paths.

From Proposition 3.6.11 we give another variant of Uniqueness of Death Theorem 3.6.10, which is phrased in terms of vector fields and their stable/unstable manifolds only.

Lemma 3.6.13 (Uniqueness of Death, Vector Field version, common starting function)**.** *Suppose η_σ, $\sigma \in [0,1]$ is a family of gradient-like vector fields for f satisfying the Morse–Smale conditions. Assume that for all σ, η_σ has a single trajectory connecting critical points q_- and q_+.*

The paths of death from Theorem 3.4.1, constructed using η_0 and η_1 and cancelling q_- and q_+, are left-homotopic.

Proof. Let $\tau \mapsto h_{\sigma,\tau}$ denote the path of death from Theorem 3.4.1 constructed with η_σ. Let (D_σ, A_σ) be the pair of good caps constructed from η_σ. Let $\tau \mapsto \widetilde{h}_{\sigma,\tau}$ be Cerf's path of death constructed with (D_σ, A_σ). By Proposition 3.6.11 the paths $h_{0,\tau}$ and $\widetilde{h}_{0,\tau}$, as well as the paths $h_{1,\tau}$ and $\widetilde{h}_{1,\tau}$, are left-homotopic. By the Uniqueness of Death Theorem 3.6.10, the paths $\widetilde{h}_{0,\tau}$ and $\widetilde{h}_{1,\tau}$ are left-homotopic. □

Lemma 3.6.13 admits a refinement if η_0 and η_1 are gradient-like vector fields for two distinct Morse functions belonging to the same connected component of $\mathcal{F}^0$. The precise formulation might sound artificial, and it is far from the most general statement, but it gives us a ready-to-use criterion to be applied in later sections. The argument resembles the proof of Lemma 3.3.12. We are again using a cut-and-paste argument to create a homotopy.

Lemma 3.6.14. *Suppose g_σ, $\sigma \in [0,1]$ is a path of excellent Morse functions, q_-, q_+ are two critical points of each of g_σ, and η is a gradient-like vector field for f_0, such that q_- and q_+ are in cancelling position. Write D for the unstable manifold of q_+ intersected with $g_0^{-1}[g_0(q_-), g_0(q_+)]$, and A for the stable manifold of q_- intersected with $f_0^{-1}[f_0(q_-), f_0(q_+)]$.*

Suppose there exists a neighbourhood U of $D \cup A$ such that $g_\sigma|_U = g_0|_U$ for all σ. Assume that η' is a gradient-like vector field for g_1 such that $\eta'|_U = \eta|_U$.

Then any two Milnor paths of death starting from g_0 with η and starting from g_1 with η' are lax homotopic over g_σ.

Proof. By Proposition 3.6.11 the paths from Theorem 3.4.1 are left-homotopic to Cerf paths of death associated to good caps for η and η', respectively. Therefore, it is enough to prove that the Cerf paths of death associated with good pair of caps constructed with η and η' are lax homotopic over g_σ.

Now, the good pair of caps associated with η and η' are isotopic: this is precisely the part of D, A cut out by a level set between $g_0(q_-)$ and $g_0(q_+)$. We conclude by Theorem 3.6.10. □

Finally, we have a statement on uniqueness of death for a family of functions with a common vector field.

Lemma 3.6.15. *Suppose $h_{\sigma,0}$, $\sigma \in [0,1]$ is a path of excellent Morse functions and η is a Morse–Smale gradient-like vector field for all $h_{\sigma,0}$. Assume there exists precisely one trajectory of η connecting two critical points q_- and q_+ and that $h_{\sigma,0}$ is independent of σ near $\mathcal{K}_- \cap \mathcal{K}_+$, where $\mathcal{K}_-$, $\mathcal{K}_+$ are the intersections of the unstable manifold of q_- (respectively, the stable manifold of q_+) with the level sets $h_{0,0}^{-1}[h_{0,0}(q_-), h_{0,0}(q_+)]$. The paths of death starting with $h_{0,0}$ and $h_{1,0}$ and constructed with η, cancelling this pair, are lax-homotopic over $h_{\sigma,0}$.*

Proof. Let these two paths of death, starting from $h_{0,0}$ and $h_{1,0}$ be called g_τ and $\widetilde{g}_\tau$ respectively. Let U be an open set containing $\mathcal{K}_- \cup \mathcal{K}_+$ on which $h_{\sigma,0}$ is equal to $f_{0,0}$. Let U_1 be another neighbourhood of $\mathcal{K}_- \cup \mathcal{K}_+$ with $\overline{U}_1 \subseteq U$. Construct a path of death $h_{0,\tau}$ starting from $h_{0,0}$, cancelling q_- with q_+, with guiding vector field η, and such that the path is supported on U_1. Also let $\phi\colon N \to \mathbb{R}$ be a cut-off function, supported on U and equal to 1 on U_1. For $\sigma \in [0,1]$ define:

$$h_{\sigma,\tau} = h_{0,\tau}\phi + (1-\phi)h_{\sigma,0}.$$

Away from U, $h_{\sigma,\tau} = h_{\sigma,0}$. On U_1, $h_{\sigma,\tau} = h_{0,\tau}$. On $U \smallsetminus U_1$, $h_{0,\tau} = h_{0,0} = h_{\sigma,0}$, so $h_{\sigma,\tau} = h_{0,0}$. In particular $\sigma \mapsto h_{\sigma,\tau}$ is an $\mathcal{F}^1$-homotopy of paths of death.

The path $h_{1,\tau}$ is a path of death, supported on U_1, with starting point $h_{1,0}$ and constructed using η. Any two η paths with the same starting point are left-homotopic. Thus, $h_{1,\tau}$ is left-homotopic to $\widetilde{g}_\tau$. Likewise, $h_{0,\tau}$ is left-homotopic to g_τ. Hence, g_τ and $\widetilde{g}_\tau$ are lax homotopic over $h_{\sigma,0}$. See Figure 19 for the notation. □

Part 4. Path Lifting

This part gives the main technical tool needed to prove the Concordance Implies Regular Homotopy Theorem 6.1.1, namely the Path Lifting Theorem 4.5.1. First we need some more advanced setup, namely the lifting lemmas from Sections 4.1, 4.2, 4.3, and 4.4.

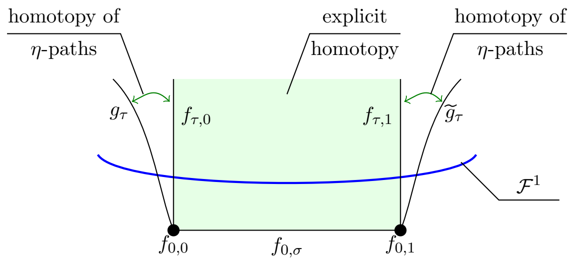

FIGURE 19. Schematic of the proof of Lemma 3.6.15.

The Path Lifting Theorem is one of the key novelties of this paper. To state it, suppose that N is a compact manifold, and let $f_\tau\colon N \to \mathbb{R}$, $\tau \in [0,1]$ be an $\mathcal{F}^1$-path of functions. The precise definition of an $\mathcal{F}^1$-path was given in Definition 3.1.1. Let $G\colon N \to \Omega$ be a generic immersion and suppose $F_0\colon \Omega \to \mathbb{R}$ is such that $F_0 \circ G = f_0$. We ask whether there exists a regular path of immersions $G_\tau\colon N \to \Omega$ with $G_0 = G$, and an $\mathcal{F}^1$-path functions $F_\tau\colon \Omega \to \mathbb{R}$ such that $F_\tau \circ G_\tau = f_\tau$ and F_τ has no new critical points away from $G_\tau(N)$. The goal of this part is to prove Path Lifting Theorem 4.5.1, which states that this is possible. An important step in the proof is the *finger move*. We delay a detailed description of the finger move until Part 5, since this is a significant detour which takes around 30 pages.

Throughout Part 3, we assume that N has all connected components of the same dimension. Necessary changes if this is not the case are given in Section 4.6 at the end of this part.

4.1. Generalities on path lifting

In this section we make precise the notion of lifting a path of functions, and address the problem of lifting elementary paths.

Definition 4.1.1. Suppose $f_\tau\colon N \to \mathbb{R}$, $\tau \in [0,1]$ is an $\mathcal{F}^1$-path of functions on N, as in Definition 3.1.1. Let $G\colon N \looparrowright \Omega$ be a generic immersion with $M = G(N)$. Finally, let $F\colon \Omega \to \mathbb{R}$ be an immersed Morse function such that $F \circ G = f_0$.

Suppose we are given a path (F_τ, G_τ) with $\tau \in [0,1]$, $F_\tau\colon \Omega \to \mathbb{R}$ and $G_\tau\colon N \to \Omega$ such that $F_0 = F$, $G_0 = G$, and (F_τ, G_τ) is a regular double path (see Definition 3.1.27) with the property that $F_\tau\colon \Omega \to \mathbb{R}$ is Morse for all $\tau \in [0,1]$ (not necessarily immersed Morse). Set $\widetilde{f}_\tau = F_\tau \circ G_\tau$.

(1) We say that the pair (F_τ, G_τ) *lifts the family* f_τ if $f_\tau = \widetilde{f}_\tau$.
(2) We say that (F_τ, G_τ) *weakly lifts the family* f_τ if $f_1 = \widetilde{f}_1$ and $\widetilde{f}_\tau$ is $\mathcal{F}^1$-homotopic to f_τ.

It is a key property of the lift that the function F_τ is Morse for all $\tau \in [0,1]$, that is, it does not acquire any critical points in the top stratum. It is not hard to lift a path f_τ at the cost of creating critical points away from N: take $G_\tau = G$ for all τ, choose a path of functions on Ω extending the given function on N, and perturb it to be generic using the methods of Subsection 3.1.3. But lifting this way, we would not have any control on the critical points which might appear on the zeroth stratum, so this would not be a very useful approach.

We begin with Lemma 4.1.2, which lifts a path of Morse functions with no rearrangements. Then we prove lemmas on lifting rearrangements, births, and deaths. These lemmas give us an inductive argument to prove the main theorem, namely the Path Lifting Theorem 4.5.1.

Lemma 4.1.2 (Lifting paths of excellent Morse functions)**.**

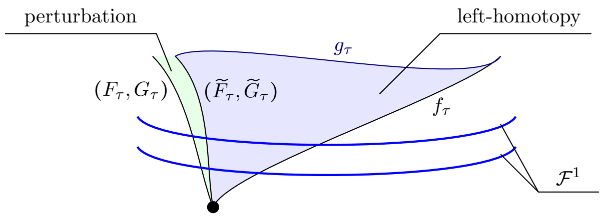

FIGURE 20. Schematic of the proof of Corollary 4.1.4.

(a) *Assume N and Ω are closed. Suppose f_τ is a path of excellent Morse functions on N, $G_0: N \to \Omega$ is a generic immersion, and $F_0: \Omega \to \mathbb{R}$ is an immersed Morse function for $M = G_0(N)$, such that $F_0 \circ G_0 = f_0$.*

Then there exists a family Ψ_τ of self-diffeomorphisms of N, and a family Υ_τ of orientation preserving self-diffeomorphisms of $\mathbb{R}$, such that the composition

$$F_\tau = \Upsilon_\tau \circ F_0 \circ G_0 \circ \Psi_\tau$$

defines a lift of the path f_τ, that is, for $G_\tau = G_0 \circ \Psi_\tau$ we have $f_\tau = F_\tau \circ G_\tau$.

(b) *If N and Ω are compact, f_τ is a neat path, F_0 is neat and G_0 is neat, then the statement holds, with (F_τ, G_τ) also being neat.*

Proof. Assume first that N and Ω are closed. By stability of Morse functions [GG73, Proposition III.2.2], there is a family Ψ_τ of self-diffeomorphisms of N and a family $\Upsilon_\tau: \mathbb{R} \to \mathbb{R}$ of strictly increasing functions such that $f_\tau = \Upsilon_\tau \circ f_0 \circ \Psi_\tau$. We define $G_\tau = G_0 \circ \Psi_\tau$ and $F_\tau = \Upsilon_\tau \circ F_0$. It is elementary to check that F_τ satisfies all the needed assumptions.

If N and Ω have boundary, the argument of [GG73] provides us with Ψ_τ being the identity near the boundary of N and Υ_τ being the identity near 0. This means that the path G_τ is independent of τ near the boundary of N, while F_τ is independent of τ near the boundary of Ω. That is to say, the double path (F_τ, G_τ) is neat. □

Remark 4.1.3. Lemma 4.1.2 not only says that the path f_τ can be lifted, but also gives a very precise recipe for a lift. For example, it follows from the form of F_τ and G_τ that the image $M = G_\tau(N)$ is independent of τ, only the parametrisation changes. In particular, G_τ is a generic immersion for all τ and F_τ is a path of immersed Morse functions. In particular (F_τ, G_τ) is a regular double path. Moreover, if ξ is a grim vector field for F_0, then one readily checks that ξ is a grim vector field for the whole family F_τ.

The following corollary will be useful.

Corollary 4.1.4. *Suppose f_τ is an $\mathcal{F}^1$-path of functions, and (F_τ, G_τ) are such that f_τ and $\widetilde{f}_\tau := F_\tau \circ G_\tau$ are left-homotopic. Then there exists a regular double path (F'_τ, G'_τ) that weakly lifts f_τ.*

Proof. For technical reasons, we apply a small perturbation to (F_τ, G_τ), denoted by $(\widetilde{F}_\tau, \widetilde{G}_\tau)$ so that the latter forms a regular double path; see Corollary 3.1.28. As G_0 is a generic immersion, and F_0 is an immersed Morse function with respect to $G_0(N)$, we may assume that the perturbation fixes (F_0, G_0). The condition that $(\widetilde{F}_\tau, \widetilde{G}_\tau)$ is a *small perturbation* means that we insist that $\widetilde{F}_\tau \circ \widetilde{G}_\tau$ and $F_\tau \circ G_\tau$ be left-homotopic. From this it follows that $\widetilde{F}_\tau \circ \widetilde{G}_\tau$ and f_τ are left-homotopic. Let $\widetilde{h}_{\sigma,\tau}$ be such left-homotopy. That is to say, $\widetilde{h}_{0,\tau} = \widetilde{F}_\tau \circ \widetilde{G}_\tau$, while $\widetilde{h}_{1,\tau} = f_\tau$. Denote $g_\tau = \widetilde{h}_{\tau,1}$; see Figure 20. Since $(\widetilde{F}_\tau, \widetilde{G}_\tau)$ is a regular double path, $\widetilde{G}_1$ is a generic immersion and $\widetilde{F}_1$ is an immersed Morse function with respect to $\widetilde{G}_1$. Lemma 4.1.2 allows us to find a lift of g_τ starting from $(\widetilde{F}_1, \widetilde{G}_1)$. We denote this lift by $(\widetilde{F}_{1+\tau}, \widetilde{G}_{1+\tau})$, $\tau \in [0,1]$. By construction, $\widetilde{G}_{1+\tau}$ is a generic immersion for all τ, and $\widetilde{F}_{1+\tau}$ is an immersed Morse function. That is, $(\widetilde{F}_{1+\tau}, \widetilde{G}_{1+\tau})$ is a regular double path. We concatenate now the paths $(\widetilde{F}_\tau, \widetilde{G}_\tau)$, $\tau \in [0,1]$ and $(\widetilde{F}_{1+\tau}, \widetilde{G}_{1+\tau})$, to obtain a path $(\widetilde{F}_\tau, \widetilde{G}_\tau)$, $\tau \in [0,2]$. The path is a regular double path (as a concatenation of two regular double paths), with $\widetilde{F}_2 \circ \widetilde{G}_2 = g_1 = f_1$.

After rescaling the time, we obtain a path (F'_τ, G'_τ), with $F'_\tau = \widetilde{F}_{2\tau}$, $G'_\tau = \widetilde{G}_{2\tau}$. We claim that $f'_\tau := F'_\tau \circ G'_\tau$ is $\mathcal{F}^1$-homotopic to f_τ. The homotopy is quite easy to construct explicitly. We define it as a concatenation of the homotopy $\widetilde{h}_{\sigma,\tau}$ and the path g_σ.

$$h_{\sigma,\tau} = \begin{cases} \widetilde{h}_{\sigma,2\tau} & \tau \le \frac{1}{2} \\ g_{\sigma+(2\tau-1)(1-\sigma)} & \tau \ge \frac{1}{2}. \end{cases}$$

It is routine to check that $\widetilde{h}_{\sigma,\tau}$ being a left-homotopy implies that $h_{\sigma,\tau}$ is an $\mathcal{F}^1$-homotopy. □

Motivated by Corollary 4.1.4 we introduce a piece of terminology.

Definition 4.1.5. A path (F_τ, G_τ), $\tau \in [0,\ell]$ can be *promoted to a weak lift of* f_τ, if (up to a possible perturbation of (F_τ, G_τ) rel. starting point, there exists a regular double path $(F_{\ell+\tau}, G_{\ell+\tau})$, $\tau \in [0,1]$ such that the whole path $(F_{(\ell+1)\tau}, G_{(\ell+1)\tau})$ is a weak lift of f_τ.

Corollary 4.1.4 admits the following generalisation.

Lemma 4.1.6. *Suppose f_τ, $\tau \in [0,1]$ is an $\mathcal{F}^1$ path, while (F_τ, G_τ), for $\tau \in [0,\ell]$ and $\ell > 1$, is a regular double path such that:*

- $g_\sigma := F_\sigma \circ G_\sigma$, $\sigma \in [0, \ell-1]$ *is an $\mathcal{F}^0$-path;*
- $\widetilde{f}_\tau := F_{\ell-1+\tau} \circ G_{\ell-1+\tau}$ *is lax homotopic to f_τ over g_τ.*

Then there exists a path $(F_{\ell+\tau}, G_{\ell+\tau})$, $\tau \in [0,1]$, such that f_τ and $F_{(\ell+1)\tau} \circ G_{(\ell+1)\tau}$ are homotopic rel. endpoints as $\mathcal{F}^1$-paths. That is, (F_τ, G_τ) can be promoted to a weak lift of f_τ.

Proof. For simplicity, assume that $\ell = 2$. Let $h_{\sigma,\tau}$ be a lax homotopy over g_τ between f_τ and $\widetilde{f}_\tau$. Set $\widetilde{g}_\sigma = h_{1-\sigma,1}$. This is an $\mathcal{F}^0$-path, moreover, $\widetilde{g}_0 = \widetilde{f}_1 = F_2 \circ G_2$ and $\widetilde{g}_1 = f_1$. Lift this path using Lemma 4.1.2 to a regular double path $(F_{2+\tau}, G_{2+\tau})$ with $F_{2+\tau} \circ G_{2+\tau} = \widetilde{f}_\tau$. We claim that $F_{3\tau} \circ G_{3\tau}$ is homotopic to f_τ. This is routine. We set

$$\widetilde{h}_{\sigma,\tau} = \begin{cases} g_{3\sigma\cdot\tau} & \tau \in [0,1/3] \\ h_{\sigma,3\tau-1} & \tau \in [1/3,2/3] \\ \widetilde{g}_{3\sigma\tau-3\tau+1} & \tau \in [2/3,1]. \end{cases}$$

With this choice $h_{0,\tau}$ is a reparametrisation of f_τ, while $h_{1,\tau}$ is a concatenation of g_τ, $\widetilde{f}_\tau$ and $\widetilde{g}_\tau$. That is, $h_{1,\tau} = F_{3\tau} \circ G_{3\tau}$ up to reparametrisation. □

4.2. Lifting paths of birth

Throughout Section 4.2, we assume that N and Ω are closed. The proof can be easily adapted to the case that N and Ω have nonempty boundary, provided that f_τ is a neat path, F_0 is neat, and G_0 is neat. The goal of this section is to prove the following result.

Lemma 4.2.1 (Lifting paths of birth)**.** *Suppose $f_\tau \colon N \to \mathbb{R}$ is a path of birth of index h. Let $G_0 \colon N \to \Omega$ be a generic immersion and let $F_0 \colon \Omega \to \mathbb{R}$ be an immersed Morse function such that $F_0 \circ G_0 = f_0$.*

Then there exists a path $F_\tau \colon \Omega \to \mathbb{R}$ and a path G_τ of generic immersions, such that:

- *the path F_τ is excellent Morse except for $\tau = 1/2$;*
- *F_τ has a birth at the first stratum for $\tau = 1/2$;*
- *(F_τ, G_τ) weakly lifts the path f_τ;*
- *G_τ is as close to G_0 as we please (for positive distance).*

Proof. The idea of the proof is to perturb the map G, if needed, so that the point at which the birth occurs is on the first stratum. This constructs the path G_τ. The next part creates the birth for the immersed Morse function F. Finally, we invoke Uniqueness of Birth Proposition 3.5.6 to show that the path we construct indeed lifts the original path. Except for Lemma 4.2.2 below, the result is straightforward. The details follow.

Without loss of generality we may assume that the birth occurs at $\tau = 1/2$. Let $q \in N$ be the point at which birth occurs and let $c = f_{1/2}(q)$ be the level set of q. Suppose $\delta > 0$ is such that, for $\tau \in [1/2-\delta, 1/2)$, no critical point of f_τ has critical value c. Such a δ exists, because $f_{1/2}$ has q as the only critical point at the level set c. The path f_τ for $\tau \in [0, 1/2-\delta]$ is a path of excellent Morse functions, which we can lift by Lemma 4.1.2. It is enough to lift the restricted path f_τ for $\tau \geq 1/2-\delta$. The interval $[1/2-\delta, 1/2]$ can be stretched to $[0, 1/2]$, so the path f_τ for $\tau \in [1/2-\delta, 1]$ can be reparametrised to a path on $[0,1]$ agreeing with the old path on $[1/2, 1]$. Therefore, without losing generality, we can assume that f_τ has no critical points at the level set c for $\tau < 1/2$.

The subset of $C^\infty(N, \Omega)$ consisting of maps that take q to the first stratum is clearly open-dense. Hence, we perturb G_0 in such a way that q is mapped to the first stratum. This perturbation can be done via a family of generic immersions. That is, let G_τ, $\tau \in [0, 1/4]$, be a path of generic immersions, supported near p and such that $p := G_{1/4}(q)$ belongs to the first stratum. By Lemma 3.1.23, we may and will assume that G_τ is F_0-regular (see Definition 3.1.20). Extend the path G_τ by setting $G_\tau = G_{1/4}$ for $\tau > 1/4$.

Choose a neighbourhood U of p with coordinates $x_1, \dots, x_n, y_1, \dots, y_k$ such that $G_{1/4}(N) \cap U = \{y_1 = \dots = y_k = 0\}$. Note that F_0 restricted to $G_{1/4}(N)$ does not have a critical point at p. Therefore there is an index $i = 1, \dots, n$, such that $\frac{\partial F_0}{\partial x_i}(p) \neq 0$. Without loss of generality we assume that $i = n$.

Consider the map

$$\psi(x_1, \dots, x_n, y_1, \dots, y_k) = (x_1, \dots, x_{n-1}, F_0(x_1, \dots, x_n, y_1, \dots, y_k), y_1, \dots, y_k).$$

This map is a local diffeomorphism since the derivative is invertible. By composing the coordinate chart with the local inverse map, and shrinking U if necessary to where the inverse is defined, we obtain new coordinates on U. By a mild abuse of notation we still denote these coordinates by $(x_1, \dots, y_k)$, but we have that $F_0 \equiv x_n$.

Set $F_\tau = F_0$ for $\tau \leq 1/4$, so that for the time $\tau \in [0, 1/4]$ we modify (perturb) G and for $\tau > 1/4$ we modify F. Choose a bump function ϕ supported on U, with $\phi \equiv 1$ in a smaller neighbourhood U' of p. For $\tau \in [1/4, 3/4]$, define

$$F_\tau = F_0 + \mathfrak{u}_{2\tau-1/2}(x_1, \dots, x_n) - x_n + (\tau - 1/4)\varepsilon\phi y_1,$$

where $\mathfrak{u}_\tau$ is the elementary path of birth of index h as in Definition 3.5.2 such that $\mathfrak{u}_\tau = \mathfrak{u}_0$ away from U'. The rôle of the last term in the definition of F_τ is to make sure that the function F_τ has no new critical points when regarded as a function on Ω. The real number $\varepsilon > 0$ is sufficiently small and will be determined shortly.

Lemma 4.2.2. *For sufficiently small $\varepsilon > 0$ and $\delta > 0$, the function F_τ for $\tau \in [0, 1/2+\delta]$ has no critical points on the zeroth stratum in U'.*

Proof. Let $U'' \subseteq U'$ be an open subset containing p such that $\phi \equiv 1$ on U''. We choose $\delta > 0$ by the condition that all the critical points of the function $\mathfrak{u}_{2\tau-1/2}(x_1,\dots,x_n)$ belong to U'' for $1/4 \le \tau \le 1/2+\delta$. Such δ exists, because $\mathfrak{u}_{2\tau-1/2}$ creates a critical point at $p \in U''$ for $\tau = 1/2$ and the critical points p_-, p_+ of $\mathfrak{u}_\tau$ are at points $(0,\dots,0,\pm\alpha(\tau))$, where $\alpha(\tau)$ depends smoothly on τ.

This means that for $1/4 \le \tau \le 1/2+\delta$, $\mathfrak{u}_{2\tau-1/2}$ has no critical points in $\overline{U'} \smallsetminus U''$. Choose a Riemannian metric. Then there is a constant $c > 0$ such that $\|D\mathfrak{u}_\tau(z)\| > c$ for all $z \in \overline{U'} \smallsetminus U''$. Set C to be the upper bound on $\overline{U'} \smallsetminus U''$ for $\|D\phi y_1\|$. If $\varepsilon < \frac{c}{2C}$, then

$$\|DF_\tau\| > c - \frac{3}{4}\varepsilon C > c - \frac{3cC}{8C} = \frac{5}{8}c > \frac{1}{2}c$$

on $\overline{U'} \smallsetminus U''$, so the derivative is nonzero there. Next, on U'', the derivative with respect to y_1 is positive, so there are no critical points on U''. Thus there are no critical points on the zeroth stratum of U', as desired. □

Corollary 4.2.3. *The function F_τ is immersed Morse for $\tau \in [0,1/2+\delta] \smallsetminus \{1/2\}$. For $\tau > 1/2$ it has two more critical points on the first stratum than for $\tau < 1/2$.*

We continue with the proof of Lemma 4.2.1. Reparametrise F_τ to $F_{\vartheta(\tau)}$ where the diffeomorphism $\vartheta\colon [0,1/2+\delta] \to [0,1]$ has positive derivative and maps $1/4$ to $1/4$ and $1/2$ to $1/2$. The reparametrized path still acquires the new critical points at $\tau = 1/2$, but the path is extended up to $[0,1]$ (and not on $[0,1/2+\delta]$).

Define $\widetilde{f}_\tau = F_{\vartheta(\tau)} \circ G_\tau$. Note that $\widetilde{f}_\tau$, for $\tau \in [0,1/4]$, is arbitrary close to f_0, in particular $\widetilde{f}_\tau$ belongs to the same stratum of $\mathcal{F}^0$ as f_0. Next, $\widetilde{f}_\tau$ for $\tau \in [1/4,1]$ acquires a critical point at $\tau = 1/2$, that is, it crosses the stratum $\mathcal{F}^1$ at $\tau = 1/2$. The crossing is transverse, the path $\mathfrak{u}_\tau$ crosses the stratum transversely and the reparametrisation ϑ we used has non-vanishing derivative at $\tau = 1/2$. In particular, $\tau \mapsto \widetilde{f}_\tau$ is a path of birth.

Now the birth on $\widetilde{f}_\tau$ occurs at the same point of N as the birth for f_τ. Moreover, the indices of the critical points created by $\widetilde{f}_\tau$ and f_τ are the same. We invoke Proposition 3.5.6 to see that $\widetilde{f}_\tau$ and f_τ are left-homotopic. By Corollary 4.1.4, the double path $(F_{\vartheta(\tau)}, G_\tau)$ can be promoted to a weak lift of f_τ. □

4.3. Lifting paths of rearrangement

Next, we lift paths of rearrangement. Throughout Section 4.3 we will assume that $k = \dim\Omega - \dim N > 1$. Remark 4.3.3 explains the obstruction to lifting if $k = 1$, so the assumption is necessary.

Lemma 4.3.1 (Lifting paths of rearrangement)**.** *Let $f_0\colon N \to \mathbb{R}$ be a Morse function and let η be a gradient-like vector field for f_0. Assume η is Morse–Smale. Let q_-, q_+ be two critical points of f_0 with $\operatorname{ind} q_- \ge \operatorname{ind} q_+$ and $f_0(q_-) = a < b = f_0(q_+)$. Suppose there are no other critical points of f_0 in $f_0^{-1}[a,b]$. Assume that there exists a generic immersion $G_0\colon N \to \Omega$ and an immersed excellent Morse function $F_0\colon \Omega \to \mathbb{R}$ such that $f_0 = F_0 \circ G_0$.*

If $k > 1$, then an elementary η-path f_τ, $\tau \in [0,1]$, moving the critical point q_- above the critical point q_+, can be weakly lifted to a regular double path (F_τ, G_τ) such that F_τ is Morse on the zeroth and first stratum for all τ, and $F_\tau \circ G_\tau$ is an η-path (Definition 3.3.9).

The path F_τ is supported on $F_0^{-1}[a-\varepsilon, b+\varepsilon]$, while G_τ is supported on $f_0^{-1}[a-\varepsilon, b+\varepsilon]$, where $\varepsilon > 0$ can be chosen as small as we please. Moreover, if f_τ is neat, G_0 is very neat, and F_0 is neat, then (F_τ, G_τ) is neat.

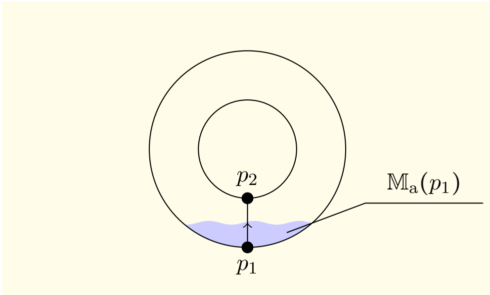

FIGURE 21. Lifting paths of rearrangements. Here M is a union of two circles in $\mathbb{R}^2$. Two critical points p_1 and p_2 are connected by a trajectory staying in the zeroth stratum. The point p_2 cannot be moved below p_1.

We recall that the condition that f_τ is elementary means that $\tau \mapsto f_\tau(q_-) - f_\tau(q_+)$ has precisely one zero in the interval $[0,1]$ and this is a simple zero; compare Definition 3.3.6. This amounts to saying that f_τ intersects $\mathcal{F}^1_\beta$ transversely at a single point.

In general, not all maps G_τ will be generic immersions, in particular, some new self-intersections of the components of N can be created. In the follow-up to Lemma 4.3.1, we control the double points created by G_τ.

Addendum 4.3.2. *The path G_τ of Lemma 4.3.1 can be constructed to have the following properties.*

(IR-1) if G_0 is an embedding, then G_τ is an embedding for all τ;
(IR-2) if $N = N_1 \sqcup \cdots \sqcup N_\ell$ and the images N_i and N_j (for all $i \neq j$) under G_0 are disjoint, then $G_\tau(N_i) \cap G_\tau(N_j) = \varnothing$ for all τ whenever $i \neq j$.
(IR-3) if $k = 2$ and $n \leq 3$, then G_τ is a generic immersion for all τ.

The proof of Lemma 4.3.1 takes the remainder of Section 4.3. Addendum 4.3.2 is proved alongside Lemma 4.3.1. We first introduce the notation, and list two extra conditions (LR-1) and (LR-2). Then we prove the lemma assuming (LR-2). Next, we arrange the function F_0 and the embedding G_0 in such a way that (LR-1) holds. Then we show how to improve a pair (F_0, G_0) satisfying (LR-1) so that (LR-2) is satisfied. A summary of the entire proof is given in Subsection 4.3.8.

4.3.1. **Notation of the proof of Lemma 4.3.1.** Let $h_- := \operatorname{ind} q_-$, $h_+ := \operatorname{ind} q_+$. Set $p_- := G_0(q_-)$ and $p_+ := G_0(q_+)$. Then p_-, p_+ are critical points of F_0 of indices h_-, h_+, respectively. Both of p_- and p_+ are at depth 1. In particular, there are neighbourhoods of q_- and q_+, such that G_0 maps these neighbourhoods to the first stratum of $M = G_0(N)$. For $\varepsilon > 0$ sufficiently small, G_0 maps $\mathcal{K}_{\eta,a-\varepsilon,a+\varepsilon}(q_-)$ and $\mathcal{K}_{\eta,b-\varepsilon,b+\varepsilon}(q_+)$ (refer to Definition 2.7.2 for the construction of $\mathcal{K}$) to the first stratum. Shrink ε if necessary to ensure that F_0 has no critical points with critical values in $[a-\varepsilon, a+\varepsilon] \cup [b-\varepsilon, b+\varepsilon]$ other than p_- and p_+.

Denote $\mathcal{K}_- := \mathcal{K}_{\eta,a,b+\varepsilon}(q_-)$ and $\mathcal{K}_+ := \mathcal{K}_{\eta,a-\varepsilon,b}(q_+)$. With this notation, $\mathcal{K}_-$ is the unstable manifold of q_- and $\mathcal{K}_+$ is the stable manifold of q_+. The dimensions are $\dim \mathcal{K}_- = n - h_-$, $\dim \mathcal{K}_+ = h_+$. By the Morse–Smale condition on η, and since $h_- \geq h_+$, it follows that $\mathcal{K}_-$ and $\mathcal{K}_+$ are disjoint. Set

$$\mathcal{L}_- = G_0(\mathcal{K}_-) \text{ and } \mathcal{L}_+ = G_0(\mathcal{K}_+). \tag{4.3.1}$$

Remark 4.3.3. If $k = 1$ and $\operatorname{ind} p_+ = \operatorname{ind} p_-$, there can be a trajectory of ξ from p_- to p_+ in the zeroth stratum, even though there are no trajectories in the first. Such a zero-stratum trajectory is an obstruction to lifting rearrangements in codimension one. Hence our assumption that $k \geq 2$; compare Figure 21.

4.3.2. **Proof of Lemma 4.3.1 under extra conditions.** We first list extra conditions which simplify the proof of Lemma 4.3.1.

(LR-1) The points $p_- = G_0(q_-)$ and $p_+ = G_0(q_+)$ are the only critical points of F_0 at depth 0 or 1 in $F_0^{-1}[a-\varepsilon, b+\varepsilon]$;

(LR-2) Either $\mathcal{L}_-$ or $\mathcal{L}_+$ belong to the first stratum of $G_0(N)$, i.e. at least one of $\mathcal{L}_-$ or $\mathcal{L}_+$ misses the singular image of G_0.

Condition (LR-2) will be used to construct a lift, while Condition (LR-1) is used to guarantee Condition (LR-2).

Lemma 4.3.4. *Suppose (F_0, G_0) satisfies (LR-2). Then there exists a path (F_τ, G_τ) with $G_\tau = G_0$ for all $\tau \in [0,1]$ that weakly lifts f_τ.*

Proof. Suppose $\mathcal{L}_-$ belongs to the first stratum. As the first stratum is open, it follows that there exists an open set $V \subseteq N$ containing $\mathcal{K}_-$ such that $G_0(V)$ belongs to the first stratum. Choose a grim vector field ξ on Ω for F_0, which is equal to $DG_0(\eta)$ in a neighbourhood of $\mathcal{L}_-$ in $M = G_0(N)$; compare Proposition 2.5.3. Assume that ξ is immersed Morse–Smale. Let $\widetilde{\eta}$ be the pull-back of ξ.

There might be critical points of F_0 in $F_0^{-1}[a,b]$ other than p_- and p_+. If that is the case, take a critical point $p \in F_0^{-1}[a,b]$ distinct from p_- and p_+. Assume it has $h_p + d_p \leq h_- + 1$ and it has the smallest value of F_0 among critical points with this property. The Morse–Smale condition implies that there are no broken trajectories in $F_0^{-1}[a, F_0(p)]$ connecting p_- with p, so we rearrange using ξ, applying Theorem 3.2.1, so that the critical value of p is smaller than a. Using this procedure, we move all critical points with $h_p + d_p \leq h_- + 1$ below the level set a. We denote by (F_σ, G_σ), $\sigma \in [0,1]$, the corresponding ξ-path, noting that it is supported away from $\mathcal{L}_-$ and away from all critical points of $F_0 \circ G_0$. By construction, $G_\sigma = G_0$. On shrinking V if needed, we may and will assume that the path is supported away from $G_0(V)$.

We claim that for F_1, $\mathcal{K}_{\xi,a,b}(p_-)$ and $\mathcal{K}_{\xi,a,b}(p_+)$ are disjoint. Clearly, By Corollary 2.6.6, as $h_{p_+} \leq h_{p_-}$, implies that there are no broken trajectories from p_- to p_+. But also, if there exists a critical point p in $F_1^{-1}[a,b]$ and a trajectory from p to p_+, then $h_p + d_p \leq h_{p_+} + 1$, also by Corollary 2.6.6. But $h_{p_+} + 1 \leq h_{p_-} + 1$ and all critical points with $h_p + d_p \leq h_{p_-} + 1$ have been moved away from $F_1^{-1}[a,b]$. Thus, no critical point is connected with p_+, so indeed, $\mathcal{K}_{\xi,a,b}(p_-) \cap \mathcal{K}_{\xi,a,b}(p_+) = \varnothing$.

We use the Rearrangement Theorem 3.2.1 to find an elementary ξ-path $F_{1+\tau}$, $\tau \in [0,1]$ of Morse functions that lifts p_- above p_+. This path can be chosen to be supported on an open set $U \subseteq \Omega$ such that $U \cap M \subseteq G_0(V)$. In particular, $\widetilde{f}_\tau := F_{1+\tau} \circ G_{1+\tau}$, $\tau \in [0,1]$ can be supported on V.

The path $\sigma \mapsto g_\sigma := F_\sigma \circ G_\sigma$, $\sigma \in [0,1]$ is a path of excellent Morse functions. We are in position to use Lemma 3.3.12. As an outcome, $\widetilde{f}_\tau$ and f_τ, $\tau \in [0,1]$, are lax homotopic over $h_{\sigma,0}$. This is precisely the situation of Lemma 4.1.6 with $\ell = 2$. That is, we can promote (F_τ, G_τ) to a weak lift of f_τ.

If $\mathcal{L}_+$ belongs to the first stratum, the proof is analogous. The vector field ξ is assumed to extend $DG_0(\eta)$ from an open subset of M containing $G_0(\mathcal{K}_+)$ and we construct a ξ-path that moves the critical point p_+ below p_-. This completes the proof of Lemma 4.3.1. □

Suppose G_0 satisfies (LR-2). The lift (F_τ, G_τ) we have constructed has the following properties:

- G_τ is a generic immersion for all τ;
- the image $G_\tau(N)$ does not change; in fact, the only moment we change G_τ is when we complete the path $\widetilde{f}_\tau$ to a weak lift. Corollary 4.1.4 constructs this completion in such a way that $G_\tau(N)$ changes by reparametrisation of N;
- the function F_τ is immersed Morse for all τ. The vector field ξ used in the construction of the rearrangement path is grim for all F_τ; compare Remark 4.1.3.

4.3.3. **Enforcing condition (LR-1).** We return to assuming the hypotheses of Lemma 4.3.1, and we arrange for (LR-1) to hold.

Lemma 4.3.5. *There exists a path $F_\tau\colon\Omega \to \mathbb{R}$ of immersed Morse functions such that:*

- $F_\tau(z) = F_0(z)$ *if* $F_0(z) < a - 2\varepsilon$ *or* $F_0(z) > b + 2\varepsilon$*;*
- $F_\tau = F_0$ *in a neighbourhood of* $\mathcal{L}_-$ *and* $\mathcal{L}_+$*;*
- *the only critical points of* F_1 *with critical values in* $[a-\varepsilon, b+\varepsilon]$ *are* p_- *and* p_+*, except if* $k = 2$ *and* $h_- = h_+$*, in which case there is also potentially a finite number of critical points at depth* $d > 1$*, for which the sum of the index and the depth is equal to* $h_- + 1$*;*
- *Suppose that* $N = N_1 \sqcup \cdots \sqcup N_\ell$ *and the images* $G_0(N_1), \dots, G_0(N_\ell)$ *are pairwise disjoint. If* p_- *and* p_+ *belong to different connected components, then all the other critical points can be moved so that their critical values do not lie in* $[a-\varepsilon, b+\varepsilon]$*. If* p_+ *and* p_- *belong to the same component of* $G_0(N)$*, then the only critical points from the previous item that cannot in general be moved away from* $[a-\varepsilon, b+\varepsilon]$ *are those that lie on the same component of* $G_0(N)$ *as* p_- *and* p_+*;*
- *The composition* $F_\tau \circ G_0$ *is a path of excellent Morse functions on* N*.*

Note that this lemma does indeed arrange for (LR-1) to hold. In the case $k = 2$ and $h_- = h_+$, the critical points that can occur with critical values in $[a-\varepsilon, b+\varepsilon]$ have depth at least 2, and so are not relevant to (LR-1).

Proof of Lemma 4.3.5. Note that $\mathcal{L}_-$ and $\mathcal{L}_+$ (defined in (4.3.1)) are stratified manifolds of dimensions $n - h_-$ and h_+, respectively. Assuming that G_0 is generic, $\mathcal{L}_- \cap \mathcal{L}_+ = \varnothing$. Choose a grim vector field ξ for F_0 which is Morse–Smale and also satisfies the following extra condition.

Condition 4.3.6. The membranes of all the critical points of ξ in $F_0^{-1}[a,b]$ different from p_- and p_+ are transverse to $\mathcal{L}_-$ and $\mathcal{L}_+$.

This can be achieved in the same way as the Morse–Smale condition; see Subsection 2.6.2. Using this condition we prove the following result.

Lemma 4.3.7. *Let* $p \neq p_-, p_+$ *be a critical point of* F_0 *in* $F_0^{-1}[a-\varepsilon, b+\varepsilon]$ *of index* h_p *and depth* $d_p > 0$*. If* $h_p + d_p < h_- + 1$*, then the descending membrane of* p *is disjoint from* $\mathcal{L}_-$*. If* $h_p + (k-1)(d_p - 1) > h_+$*, then the ascending membrane is disjoint from* $\mathcal{L}_+$*.*

Remark. If $d_p = 0$, the membranes of p belong to the zeroth stratum, so they are automatically disjoint from $\mathcal{L}_-$ and $\mathcal{L}_+$.

Proof. By Lemma 2.6.2, $\dim \mathbb{M}_{\mathrm{d}}(p) \cap \Omega[j] = h_p + d_p - j$ and $\dim \mathbb{M}_{\mathrm{a}}(p) \cap \Omega[j] = n - k(d_p - 1) - h_p + d_p - j$. For each $j > 0$, ($\mathcal{L}_- \cap \Omega[0] = \varnothing$ by construction), we have $\dim \mathcal{L}_- \cap \Omega[j] = n - h_- - k(j-1)$. Then $\mathbb{M}_{\mathrm{d}}(p) \cap \mathcal{L}_- \cap \Omega[j]$ is empty if:

$$h_p + d_p - j + n - h_- - k(j-1) < n - k(j-1) = \dim \Omega[j]. \tag{4.3.2}$$

Transforming this formula yields
$$h_p + d_p < h_- + j.$$
This holds for all $j > 0$, since we assume that $h_p + d_p < h_- + 1$. Thus $\mathbb{M}_{\mathrm{d}}(p) \cap \mathcal{L}_- \cap \Omega[j] = \varnothing$ for all $j \geq 0$. Hence, since $M = \cup_{j \geq 0}\Omega[j]$, we have that $\mathbb{M}_{\mathrm{d}}(p) \cap \mathcal{L}_- = \varnothing$, as required.

The argument for the $\mathcal{L}_+$ is analogous. Again, $\mathcal{L}_+ \cap \Omega[0] = \varnothing$. As $\dim \mathcal{L}_+ \cap \Omega[j] = h_+ - k(j-1)$, $\mathbb{M}_{\mathrm{a}}(p) \cap \mathcal{L}_+ \cap \Omega[j] = \varnothing$ if

$$n - k(d_p - 1) - h_p + d_p - j + h_+ - k(j-1) < n - k(j-1). \tag{4.3.3}$$

This translates into $h_+ < h_p + k(d_p - 1) - (d_p - j)$. This holds for all $j > 0$ if $h_+ < h_p + (k-1)(d_p - 1)$. Thus under this condition $\mathbb{M}_{\mathrm{a}}(p) \cap \mathcal{L}_+ \cap \Omega[j] = \varnothing$ for all $j \geq 0$, and hence $\mathbb{M}_{\mathrm{a}}(p) \cap \mathcal{L}_+ = \varnothing$. □

Remark 4.3.8. The intersection of $\mathcal{L}_-$ with the descending membrane $\mathbb{M}_{\mathrm{d}}(p)$ need not consist of trajectories of ξ, so it can be nonempty even if the expected dimension is zero. Therefore, we get strict inequalities in Lemma 4.3.7.

Continuing the proof of Lemma 4.3.5, we introduce a useful piece of terminology.

Definition 4.3.9. A ξ-path of rearrangement F_τ is *safe* if the support of F_τ is disjoint from $\mathcal{L}_- \cup \mathcal{L}_+$.

Continuing the proof of Lemma 4.3.5, we state and prove the following result.

Lemma 4.3.10. *There exists an F_τ path of rearrangements that is a concatenation of safe ξ-path of rearrangements, supported on $F_0^{-1}[a-\varepsilon, b+\varepsilon]$ such that for any two pairs of critical points of F_τ, p, p' with $p, p' \in F_0^{-1}[a+\varepsilon, b-\varepsilon]$, we have $h_p + d_p \leq h_{p'} + d_{p'}$ whenever $F_1(p) < F_1(p')$.*

Furthermore, the only critical points between p_- and p_+ necessarily have $h_- + d_- \leq h_p + d_p \leq h_+ + d_+$.

Remark. If $h_- + d_- > h_+ + d_+$, then the statement of the lemma means that after the rearrangements are made, there are no critical points between p_- and p_+

Proof. We present an algorithm for performing the necessary rearrangements. We take two consecutive critical points p, p', that is points such that $F_0(p) < F_0(p')$ and such that there are no critical points between p and p'. We claim that at least one of the situations occur.

(SR-1) $h_p + d_p \leq h_{p'} + d_{p'}$;
(SR-2) $h_p + d_p > h_{p'} + d_{p'}$ and at least one of $d_p, d_{p'}$ is zero;
(SR-3) $h_p + d_p > h_{p'} + d_{p'}$ and $h_{p'} + d_{p'} < h_- + 1$;
(SR-4) $h_p + d_p > h_{p'} + d_{p'}$ and $h_p + (k-1)(d_p - 1) > h_+$.

The claim is proved by contradiction. If none of the four cases occurs, we have (i) $h_p + d_p > h_{p'} + d_{p'}$, (ii) $h_{p'} + d_{p'} \geq h_- + 1$, and (iii) $h_p + (k-1)(d_p - 1) \leq h_+$. We also have that (iv) $h_- \geq h_+$, by the assumptions of Lemma 4.3.1, and $d_p > 0$, $d_{p'} > 0$ by (SR-2). Combining these yields:

$$h_p + d_p = h_p + (d_p - 1) + 1 \overset{(k \geq 2)}{\leq} h_p + (k-1)(d_p - 1) + 1 \overset{(iii)}{\leq} h_+ + 1 \overset{(iv)}{\leq} h_- + 1 \overset{(ii)}{\leq} h_{p'} + d_{p'} \overset{(i)}{<} h_p + d_p.$$

The first inequality used that $d_p \geq 1$. Since this is absurd, we deduce that the above list is indeed exhaustive, as asserted.

We describe now a local algorithm deciding on a case-by-case basis which type of rearrangement should be performed for two consecutive points.

In case (SR-1), there is no need to rearrange. In case (SR-2), the membranes of the critical point at depth zero are disjoint from $\mathcal{L}_-$ and $\mathcal{L}_+$, so we can rearrange by moving p above p' (if $d_p = 0$), or p' below p (if $d_{p'} = 0$). If $d_p = d_{p'} = 0$, we can perform either of the two moves.

Note that in this move, we rearrange only a depth zero critical point, so that if p or p' happen to be either of p_- or p_+, then they are not moved.

In case (SR-3), the dimension counting argument of Lemma 4.3.7 ensures that the descending membrane of p' is disjoint from the set $\mathcal{L}_-$. Therefore, we can perform safe rearrangements moving p' below p, unless $p' = p_-$ or $p' = p_+$.

If $p' = p_-$, then there is a critical point below p_-, with $h_p + d_p > h_{p_-} + d_{p_-}$. Originally, the function F_0 has no critical points in $F_0^{-1}[a - \varepsilon, a]$. If during the procedure we are now describing, some critical point lands below p, then it must have $h_p + d_p < h_{p_-} + d_{p_-}$. Therefore, the only situation where the move is forbidden is if $p' = p_+$ is the top critical point.

Similarly, in case (SR-4), the ascending membrane of p is disjoint from $\mathcal{L}_+$, so we can safely move p above p'. The discussion of special cases is analogous: we cannot perform the move of case (SR-4) only if $p = p_-$.

Having specified the local algorithm, we apply it first for all the critical points between p_- and p_+. In this way, using a bubble sort, we arrange critical points in ascending order of $h_p + d_p$. Next, we move to the case when $p = p_-$. If the critical point p' directly above p_- has $h_p + d_p > h_{p'} + d_{p'}$, then, as $h_p + d_p = h_- + 1$, we land in case (SR-3). Therefore, we can move p' below p_-. Inductively, we move all critical points below p_-, for which $h_{p'} + d_{p'} < h_- + 1$.

Next, suppose that $p' = p_+$ and p is immediately below p_+ with $h_p + d_p > h_{p'} + d_{p'}$, and $p \neq p_-$. If $d_p = 0$, we fit into case (SR-2), we lift p above p'. If $d_p \neq 0$, then, as $k \geq 2$, $h_p + (k-1)(d_p - 1) \geq h_p + d_p - 1$, so we are in case (SR-4). We can lift the critical point p above p_+.

Therefore, the only critical points that can remain between p_- and p_+ are those with $h_p + d_p \geq h_- + d_-$ (because all others can be safely pushed below p_- or above p_+) and, simultaneously, $h_p + d_p \leq h_+ + d_+$. □

We study the critical points that remain between p_- and p_+ in greater detail. These are precisely the points at which the local algorithm fails to move away from the region between p_- and p_+. The only situation when this can happen is if $h_- + d_- = h_+ + d_+$.

Corollary 4.3.11. *Suppose $h_- + d_- = h_+ + d_+$. Then all the critical points p with $d_p = 0$ can be moved above the critical point p_+. Moreover, if $k \geq 3$, all the critical points can be moved.*

Proof. We refine the part of the proof of Lemma 4.3.10. First take two critical points p and p' between p_- and p_+ such that p' is right above p and p' is allowed to be p_+. Necessarily, $h_p + d_p = h_{p'} + d_{p'}$ (now all critical points between p and p' satisfy this). If $d_p = 0$ and $d_{p'} > 0$, then we safely move p above p' as in case (SR-2) of the proof of Lemma 4.3.10. Iterating this process, we safely rearrange all critical points with $d_p = 0$ by moving them above p_+. This completes the first part of the corollary.

Next, suppose $k \geq 3$ and that we have a critical point p right below p_+. Suppose it is not p_-. By the assumptions, $h_p + d_p = h_+ + d_+$. As all critical points in the zeroth stratum have been moved above p_+, we necessarily have $d_p > 0$. Since $p \neq p_-$, we have $d_p \neq 1$ too, so $d_p \geq 2$. Then

$$h_p + (k-1)(d_p - 1) = h_p + (k-2)(d_p - 1) + d_p - 1 > h_p + d_p - 1 = h_+ + d_+ - 1 > h_+.$$

Thus, we are in case (SR-4), so we can safely lift p above p_+. □

Corollary 4.3.11 takes care of the first three items of Lemma 4.3.5.

We now refine the above procedure to prove the fourth item of Lemma 4.3.5. Suppose $G_0(N_1), \dots, G_0(N_\ell)$ are disjoint, and $p_- := G_0(q_-)$ belongs to $G_0(N_1)$. This is relevant only if $h_- = h_+$, because this condition is part of the statement of the fourth item. Suppose p is a critical point in $F_0^{-1}[a-\varepsilon, b+\varepsilon]$ at depth $d_p \geq 2$ and with $h_p + d_p = h_- + 1$, and suppose there are no critical points with critical values in $(F(p_-), F(p))$. Assume that p belongs, say, to $G_0(N_2)$. By the dimension condition, there are no trajectories between p_- and p that belong to the zeroth stratum. On the other hand, $\mathcal{L}_-$ belongs to a different component of $G_0(N)$ than the part of the membrane of p belonging to the first and deeper strata (here we use the fact that $\mathcal{L}_- \subseteq G_0(N_1)$ and $p \in G_0(N_2)$). Hence, the descending membrane of p is disjoint from $\mathcal{L}_-$, so we can safely move p below p_-. This proves the fourth item.

Concatenating all the ξ-paths of rearrangement used in these two procedures (first organizing all the critical points between p_- and p_+ and then second moving most of their critical values out of the region between the values of p_- and p_+) creates a ξ-path of rearrangement F_τ with support disjoint from $\mathcal{L}_- \cup \mathcal{L}_+$, i.e. a safe path. The only critical points of F_1 in $F_1^{-1}(a,b)$ are those with $h_p + d_p = h_- + 1 = h_+ + 1$ and $d_p > 0$. If $\varepsilon > 0$ is small, $F_1^{-1}[a-\varepsilon, b+\varepsilon]$ also contains only these points. We have not moved any of the critical points at depth 1, and so $F_\tau \circ G_0$ is a path of excellent Morse functions, as required by the last item of the lemma. This concludes the proof of Lemma 4.3.5. □

The path constructed in Lemma 4.3.5 does not change G_0, so in particular no new double points of $M = G_0(N)$ are created. Moreover, F_τ is a path of immersed Morse functions, without births or deaths on the deeper strata.

We have done enough to obtain (IR-1) of Addendum 4.3.2. If G_0 is an embedding then (LR-2) is automatically satisfied, and so no further operations are necessary. The procedure described in Subsection 4.3.8 will give rise to a regular double path (F_τ, G_τ) with G_τ an embedding for all τ.

4.3.4. **Enforcing condition (LR-2): the regular homotopy.** The next result, ensuring that it is possible to change the map G_0 by a path of immersions in such a way that G_1 maps $\mathcal{K}_-$ to the first stratum, is the most difficult step in the proof of Lifting Rearrangement Lemma 4.3.1. The precise construction of the isotopy G_τ stretches across Subsections 4.3.4, 4.3.5, 4.3.6, and 4.3.7. The main part of the proof occurs in this subsection. In Subsection 4.3.5 we show that we can arrange for G_τ to be an immersion for all τ. In Subsection 4.3.6 we show how to modify the construction to arrange for (IR-2) to hold. In Subsection 4.3.7 we address (IR-3). Then in Subsection 4.3.8 we summarise the entire proof.

Proposition 4.3.12. *Suppose (F_0, G_0) satisfies condition (LR-1), and η is gradient-like for $F_0 \circ G_0$. There exists a generic path of immersions $G_\tau : N \to \mathbb{R}$, $\tau \in [0,1]$, such that G_τ is a regular F_0-path (see Definition 3.1.24), $F_0 \circ G_\tau$ is an η-path (Definition 3.3.9) of excellent Morse functions, and $G_1(\mathcal{K}_-) \cap F_0^{-1}([a,b])$ is contained in the first stratum of $G_1(N)$.*

Proof of Proposition 4.3.12 modulo Subsection 4.3.5. We will prove the lemma modulo a key technical lemma, that will be proven in the following subsection. Define

$$\Omega' = F_0^{-1}[a-\varepsilon, b+\varepsilon] \text{ and } N' = G_0^{-1}(\Omega') = f_0^{-1}[a-\varepsilon, b+\varepsilon]. \tag{4.3.4}$$

The idea is to push the set $G_0(\mathcal{K}_-) \cap \Omega[2]$ above the level set of $F_0^{-1}(b)$, so that $G_0(\mathcal{K}_-)$ does not intersect $\Omega[2]$ below $F_0^{-1}(b)$. Hence also $G_0(\mathcal{K}_-)$ will not intersect $\Omega[d]$ for $d \geq 3$, because $\Omega[d] \subseteq \overline{\Omega[2]}$. The push will be guided by a vector field which is not necessarily tangent to $G_0(N)$. Choose a vector field on Ω', which is a gradient vector field for F_0 as *an ordinary Morse function*. This means that the vector field need not be tangent to M. To emphasise the difference between grim vector fields and this vector field, we denote the latter by ∇F_0. Note that F_0, regarded as an ordinary Morse function, has no critical points on Ω' (by Condition (LR-1)), therefore ∇F_0 does not vanish in Ω'.

We will assume that ∇F_0 is generic. The precise meaning of the genericity condition will be specified in few places, especially in Subsection 4.3.6. The next result, Lemma 4.3.13, gives a suitable framework for genericity: instead of perturbing the Riemannian metric defining ∇F_0, we can perturb the induced projection. This allows us to use transversality arguments in a more flexible way.

A gradient-like vector field ∇F_0 induces a projection from Ω' onto $Y := F_0^{-1}(a+\varepsilon)$: a point $z \in \Omega'$ is mapped to the unique point $y \in Y$ lying on the same trajectory of ∇F_0. We denote this projection by $\Pi: \Omega' \to Y$. For future use we give the following standard result.

Lemma 4.3.13. *Suppose $\Pi': \Omega' \to Y$ is another projection, and Π' is sufficiently close (in the C^1-norm) to Π. Then there is a gradient-like vector field for F_0 inducing Π'.*

Proof. For any $c \in [a+\varepsilon, b+\varepsilon]$ the restriction $\Pi|_{F_0^{-1}(c)}$ is a local diffeomorphism, because ∇F_0 is transverse to the level sets of F_0. This property is open, hence if $\Pi' \in C^\infty(\Omega', Y)$ is sufficiently close to Π, then Π' restricted to $F_0^{-1}(c)$, for $c \in [a+\varepsilon, b+\varepsilon]$, is a local diffeomorphism as well.

Choose a metric on Ω'. For each point $z \in \Omega'$, let $\xi(z) \in \ker \Pi'$ be the unit vector such that $\partial_{\xi(z)} F_0(z) > 0$. (We know that $\partial_{\xi(z)} F_0(z) \neq 0$ because $\xi(z)$ is transverse to the level set $F_0^{-1}(F_0(z))$.) We choose the orientation of $\xi(z)$ in such a way that $\partial_{\xi(z)} F_0(z) > 0$. By construction, $\xi(z)$ is a gradient-like vector field for F_0. Moreover, since $\xi(z) \in \ker D\Pi'(z)$ for all z, the trajectories of $\xi(z)$ are contained in the fibres of Π'. □

We use this lemma to specify the first of several genericity conditions imposed on ∇F_0.

Lemma 4.3.14. *The set of maps Π such that for each sign $\pm$, $(\Pi \circ G_0)^{-1}(\Pi \circ G_0(q_\pm)) = \{q_\pm\}$, is open dense in the set of all smooth maps from Ω' to Y.*

Proof. The double point set of $\Pi \circ G_0$ is of codimension $k-1$, hence it may be avoided generically by the zero-dimensional manifold $\{q_-, q_+\}$. It follows from transversality that the given set of maps is open-dense. □

It follows that there exists a nonempty open neighbourhood $U \subseteq N$ of q_- with the property that a trajectory of ∇F_0 through any point in $G_0(u)$ for $u \in U$, does not hit $G_0(N)$. Decrease ε, if necessary to ensure that

$$\mathcal{K}_- \cap f^{-1}[a, a+2\varepsilon] \subseteq U. \tag{4.3.5}$$

Define $Z \subseteq f^{-1}(a+\varepsilon)$ to be that subset of $z \in f^{-1}(a+\varepsilon)$ that satisfy Condition 4.3.15 below.

Condition 4.3.15. Whenever $w \in N$ lies on a trajectory of η through z and $a+\varepsilon \le f(w) < a+2\varepsilon$, then the trajectory of ∇F_0 through $G_0(w)$ only intersects $G_0(N)$ in Ω' at $G_0(w)$.

We refer to the schematic of Z in Figure 22. The conditions defining Z are open, hence Z is an open subset of $f^{-1}(a+\varepsilon)$. Note that (4.3.5) implies also that $\mathcal{K}_- \cap f^{-1}[a+\varepsilon] \subseteq Z$. In particular Z is nonempty.

As $\mathcal{K}_-$ is disjoint from $\mathcal{K}_+$, there exists an open subset $Z_0 \subseteq Z$ ('open' meaning open in $f^{-1}(a+\varepsilon)$) containing $\mathcal{K}_- \cap f^{-1}(a+\varepsilon)$ and disjoint from $\mathcal{K}_+$.

Define $R: \mathbb{R} \to \mathbb{R}$ to be a smooth increasing function such that $R(x) = x$ for $x \notin [a+\varepsilon, b+\varepsilon]$ and such that R maps:

- $[a+\varepsilon, a+2\varepsilon]$ onto $[a+\varepsilon, b]$;
- $[a+2\varepsilon, b]$ onto $[b, b+\varepsilon/2]$;
- $[b, b+\varepsilon]$ onto $[b+\varepsilon/2, b+\varepsilon]$.

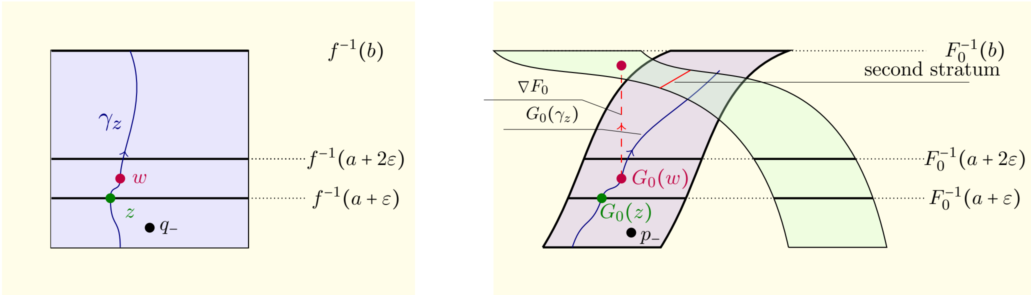

FIGURE 22. The point z does not belong to Z. In fact, the trajectory γ_z of η, in N passes through w (left picture), and the trajectory of ∇F_0 through $G_0(w)$, in $G_0(N) \subseteq \Omega$, reaches a point on $G_0(N)$ (right picture).

The function R will be our *rescaling function*, denoting the height of the lift. Note that, by construction

$$R(x) \geq x \text{ for all } x \in \mathbb{R}. \tag{4.3.6}$$

We define a new function, which is called the *lift function* $P\colon \Omega' \times [0,1] \to \Omega'$. For $\theta \in [0,1]$, and $z \in \Omega'$, we define $P(z,\theta)$ to be the unique point in Ω' belonging to the same trajectory of ∇F_0 as z, such that

$$F_0(P(z,\theta)) = F_0(z) + \theta(R(F_0(z)) - F_0(z)). \tag{4.3.7}$$

Note that the function $P\colon \Omega' \times [0,1] \to \Omega'$ is a smooth function, because of the smooth dependence of solutions of an ODE on initial conditions.

We aim to define G_τ in such a way that a point $u \in \mathcal{K}_-$ is mapped to $P(G_0(u),\tau)$, i.e. it is moved up from the level set $f(u)$ to the level set of $(1-\tau)f(u) + \tau R(f(u))$. For the map to be smooth, we need to also lift points near $\mathcal{K}_-$. To this end, we need to define a cut-off function

$$\mu\colon f^{-1}[a+\varepsilon, b+\varepsilon] \to [0,1].$$

In the construction, we will adjust the function μ in such a way that

(i) η is a gradient-like vector field for $F_0 \circ G_\tau$ for all $\tau \in [0,1]$; and
(ii) G_τ is an immersion for all $\tau \in [0,1]$.

We address the first property next. The second property will be recorded below as Lemma 4.3.19, the proof of which we will defer to Subsection 4.3.5.

To prove that η is gradient-like for $F_0 \circ G_\tau$, we will need the following properties of μ:

(μ-1) $\partial_\eta \mu \geq 0$, and equality holds only at critical points of f and at places where the value of μ lies in $\{0,1\}$;
(μ-2) the support of μ intersected with $f^{-1}(a+\varepsilon)$ is contained in Z_0;
(μ-3) μ is equal to 1 on $\mathcal{K}_- \cap f^{-1}[a+2\varepsilon, b+\varepsilon]$ and vanishes on $\mathcal{K}_+$.

The next lemma shows that such a function exists.

Lemma 4.3.16. *There exists a smooth function $\mu\colon f^{-1}[a+\varepsilon, b+\varepsilon] \to [0,1]$ satisfying (μ-1), (μ-2), and (μ-3).*

Proof. Take a smooth function $\mu_0\colon f_0^{-1}(a+\varepsilon) \to [0,1]$ supported on Z_0 and equal to 1 on $\mathcal{K}_- \cap f^{-1}(a+\varepsilon) \subseteq Z_0$. Define $\Omega_0' := F_0^{-1}[a+2\varepsilon, b+\varepsilon]$ and $N_0' := G_0^{-1}(\Omega_0')$.

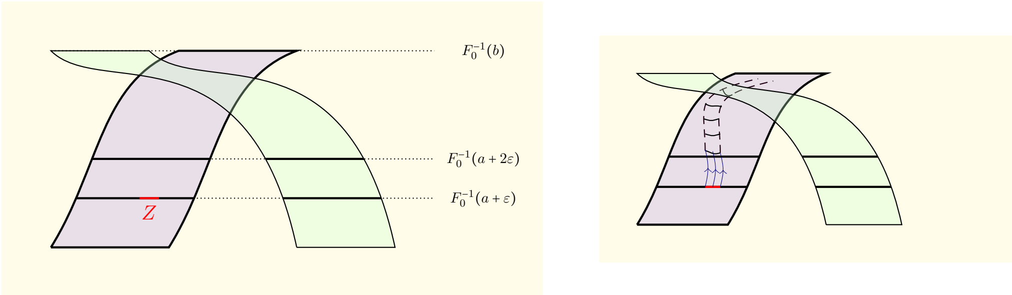

FIGURE 23. A schematic of G_τ. Points on a trajectory through Z are lifted up by flowing along ∇F_0.

Let ν be a smooth non-decreasing function from $\mathbb{R}_{\geq 0}$ to $[0,1]$ equal to 0 for $t \in \mathbb{R}$ close to 0 and equal to 1 for $t > t_0$, where t_0 is the minimal time needed to reach $f_0^{-1}(a+2\varepsilon)$ by going from Z_0 via η.

Then define $\mu\colon N_0' \to \mathbb{R}_{\geq 0}$ via $\mu(z) = \nu(t)\mu_0(z_0)$ for any $z \in N_0'$ if the trajectory of η through z hits the level set $a+\varepsilon$ at the point z_0 and the time to go from z_0 to z is equal to t. With this construction, μ has all the required properties. □

We will impose one more condition on μ in Subsection 4.3.5, and show that μ can be assumed to also satisfy this.

Now we define a path of maps $G_\tau\colon N \to \Omega$. The definition is as follows. Take $u \in N$. Set $G_\tau(u) = G_0(u)$ if $f_0(u) \notin N_0'$. If $f_0(u) \in N_0'$, we set

$$G_\tau(u) = P(G_0(u), \tau\mu(u)). \tag{4.3.8}$$

Compare Figure 23. We list the properties of G_τ.

- $G_\tau(u)$ depends smoothly on τ and u, in particular G_τ is a homotopy between the maps G_0 and G_1;
- $G_\tau(u) = G_0(u)$ on $\mathcal{K}_+$ by (μ-3). Moreover, if a trajectory of η through u misses Z, then $G_\tau(u) = G_0(u)$;
- $G_\tau = G_0$ above the level set $f_0^{-1}(b+\varepsilon)$ and below the level set $f_0^{-1}(a+\varepsilon)$.

Set

$$\widetilde{f}_\tau = F_0 \circ G_\tau\colon N \to \mathbb{R}.$$

Lemma 4.3.17. *The family $\widetilde{f}_\tau\colon N \to \mathbb{R}$ is an η-path of excellent Morse functions.*

Proof. It is clear from the construction that $\widetilde{f}_\tau$ does not depend on τ for any of the critical points of f_0. In fact, the only critical point of f_0 in N_0' is q_+, and we have that $\mu(q_+) = 0$.

We prove that $\partial_\eta f_\tau \geq 0$ with equality only at critical points of f_τ. Given a point $u \in N$, (4.3.7) implies that $P(G_0(u), \tau\mu(u))$ is at the level set $F_0(G_0(u)) + \tau\mu(u)(R(F_0(G_0(u))) - F_0(G_0(u)))$. Noting that $F_0(G_0(u)) = f_0(u)$ and $f_\tau(u) = F_0(G_\tau(u))$ we see that (4.3.7) and (4.3.8) imply:

$$f_\tau(u) = f_0(u) + \tau\mu(u)(R(f_0(u)) - f_0(u)). \tag{4.3.9}$$

Applying ∂_η to both sides of (4.3.9) we obtain.

$$
\begin{aligned}
\partial_\eta f_\tau(u) =&\partial_\eta f_0(u)+ \\
&+\tau(R(f_0(u))-f_0(u))\partial_\eta\mu+(R'(f_0(u))\partial_\eta f_0(u)-\partial_\eta f_0(u))\tau\mu= \\
=&\underbrace{(1-\tau\mu)\partial_\eta f_0(u)}_{(1)}+\underbrace{\tau(R(f_0(u))-f_0(u))\partial_\eta\mu}_{(2)}+\underbrace{\tau\mu R'(f_0(u))\partial_\eta f_0(u)}_{(3)}.
\end{aligned}
\tag{4.3.10}
$$

We study the terms (1), (2), and (3).

(1) As η is a gradient-like vector field and $\mu(u)\in[0,1]$, term (1) is non-negative. It is zero only if $\mu(u)=\tau=1$ or u is a critical point of f.

(2) By construction, $\partial_\eta\mu\geq 0$, with equality at the critical points, and also $R(f_0(u))-f_0(u)\geq 0$ (compare (4.3.6)), so (2) is non-negative.

(3) The derivative R' is positive. The term (3) is non-negative. It is positive unless $\tau\mu=0$ or u is a critical point of f_0.

From the list, since $\tau\mu$ cannot be 0 and 1 simultaneously, it is clear that $\partial f_\tau(u)\geq 0$ with equality exactly at the critical points of f_0. □

Continuing the proof of Proposition 4.3.12, we have the following result.

Lemma 4.3.18. *$G_1(\mathcal{K}_-)\cap F_0^{-1}[a,b]$ is disjoint from the set of the double points of $G_1(N)$.*

Proof. Take $w\in\mathcal{K}_-\cap f_0^{-1}[a,b]$. Consider the following three cases.

- If $f_0(w)\leq a+\varepsilon$, then $G_1(w)=G_0(w)$ belongs to the first stratum, because $\mathcal{K}_-\cap f_0^{-1}[a,a+\varepsilon]\subseteq U$.
- Suppose $f_0(w)\in[a+\varepsilon,a+2\varepsilon]$. Let $u\in f_0^{-1}(a+\varepsilon)$ belong to the same trajectory of η as w. Then $u\in Z$ and by Condition 4.3.15, $G_1(w)$ belongs to the first stratum.
- If $f_0(w)>a+2\varepsilon$, then $\mu(w)=1$, so by (4.3.9) we have $f_\tau(w)=(1-\tau)f_0(w)+R(f_0(w))>b$. So $G_1(w)\notin F_0^{-1}[a,b]$. □

We will need one more result, whose proof is deferred to later subsections.

Lemma 4.3.19. *There exists a choice of μ such that G_τ is an immersion for all $\tau\in[0,1]$.*

We will prove Lemma 4.3.19 in Subsection 4.3.5. Modulo this, Lemmas 4.3.17, 4.3.18, and 4.3.19 conclude the proof of Proposition 4.3.12. □

In the next subsection, we will check that G_τ is an immersion for all τ, proving Lemma 4.3.19. This done, we will perturb G_τ in such a way that (F_τ,G_τ) becomes a regular double path (see Definition 3.1.27). The construction will be further tweaked in Subsections 4.3.6 and 4.3.7, in order to arrange for (IR-2) and (IR-3) to hold. A summary of the entire proof of Lemma 4.3.1 will be given in Subsection 4.3.8.

4.3.5. **Enforcing condition (LR-2): proving that G_τ is an immersion.** We will now prove that if ∇F_0 is sufficiently generic, and for a careful choice of μ, then G_τ is an immersion for all τ.

Recall that $\Omega'=F_0^{-1}[a-\varepsilon,b+\varepsilon]$. Set $Y=F_0^{-1}(a+\varepsilon)$. As explained in Subsection 4.3.4, the vector field ∇F_0 induces a projection $\Pi\colon\Omega'\to Y$. We begin with the following observation.

Lemma 4.3.20. *The map $\Pi\circ G_\tau\colon N\to Y$ does not depend on τ.*

Proof. For any point $u\in N$, by the construction of G_τ, the points $G_\tau(u)$, $\tau\in[0,1]$ belong to the same trajectory of ∇F_0. Therefore, they are mapped by Π to the same point. □

Corollary 4.3.21. *Let $u \in N$ and let $z = G_0(u)$. Suppose that ∇F_0 is not tangent at z to the branch of $G_0(N)$ containing $G_0(u)$. Then G_τ is an immersion near u for all τ.*

Proof. By the chain rule, $D(\Pi \circ G_0) = D\Pi(G_0(u)) \circ DG_0(u)$. As G_0 is an immersion, $DG_0(u)$ has trivial kernel. The differential $D\Pi(z)$ has kernel spanned by $\nabla F_0(z)$. By assumption, it follows that the kernel of $D\Pi(z)$ intersects the image of $DG_0(u)$ only at 0. Hence $D(\Pi \circ G_0)(u)$ has trivial kernel. By Lemma 4.3.20, $D(\Pi \circ G_\tau)(u)$ has trivial kernel for all τ. Since $D(\Pi \circ G_\tau) = D\Pi(G_\tau(u)) \circ DG_\tau(u)$, it follows that $DG_\tau(u)$ has trivial kernel for all τ, and hence by continuity of the derivative, G_τ is an immersion near u. □

Thus we only need to study points at which ∇F_0 is tangent to $G_0(N)$ at z. We show that generically, the set of such points $z \in G_0(N)$ is itself a manifold. In proving genericity, it is most convenient to perturb the immersion G_0.

Lemma 4.3.22. *There is an open-dense subset of maps G_0 such that the set of points $u \in N$, with the property that ∇F_0 is tangent to the branch of M with tangent vectors $DG_0(u)(T_uN) \subseteq T_{G_0(u)}\Omega$, forms a smooth manifold of dimension $n-k$, missing q_- and q_+.*

Proof. The condition $\nabla F_0(G_0(u)) \subseteq \operatorname{Im} DG_0(u)$ defines a smooth submanifold in the 1-fold jet space $J^1(N,\Omega)$ of maps from N to Ω; compare Subsection 3.1.2 for the relevant definitions. The codimension is equal to k. This can be seen as follows: $DG_0(u)$ is a codimension k subspace and $\nabla F_0(G_0(u))$ is a 1-dimensional vector. The condition that a vector fits into a codimension k space is of codimension k. By the Thom transversality theorem (compare Theorem 2.3.4 for $s = 1$) the set of maps G_0 transverse to this stratum is open-dense. If, possibly after a perturbation, G_0 is transverse to this stratum, then the set of points in N such that the jet extension of G_0 hits the stratum, is a smooth submanifold of the same codimension, that is, k. □

Remark. We can understand Lemma 4.3.22 in the broader context of the Thom-Boardman stratification of Subsection 2.3.2. The set of points u where $\nabla F_0(G_0(u))$ is tangent to $G_0(N)$ at $G_0(u)$ actually corresponds to the Thom-Boardman stratum β^1 of the map $\Pi \circ G_0$. In particular, the dimension of this set can be alternatively computed using e.g. [AGZV12, Section I.2].

Perturb G_0 so that it satisfies the genericity condition spelled out in Lemma 4.3.22. Let $\Sigma \subseteq N$ be the set of points u such that the image of $DG_0(u)$ contains $\nabla F_0(G_0(u))$; see Figure 24. As critical points of f_0 form a finite set, by further perturbing G_0, we may and shall assume that Σ misses all critical points of f_0. For each point u in Σ, there is a unique vector $\upsilon_u \in T_uN$ such that $DG_0(u)(\upsilon_u) = \nabla F_0$. Hence we obtain a section $\upsilon\colon \Sigma \to TN$ of $TN|_\Sigma$. We make a trivial observation for future use.

Lemma 4.3.23. *If υ_u is parallel to $\eta(u)$, then υ_u is not a negative multiple of $\eta(u)$.*

Proof. We have $f_0 = F_0 \circ G_0$, $\upsilon_u \in T_uN$, and $DG_0(\upsilon_u) = \nabla F_0$. Thus

$$\partial_\upsilon f_0(u) \overset{(1)}{=} \partial_{DG_0(\upsilon)} F_0(G_0(u)) \overset{(2)}{=} \partial_{\nabla F_0} F_0(G_0(u)) \overset{(3)}{=} \|(\nabla F_0)(G_0(u))\|^2 \geq 0. \tag{4.3.11}$$

Here, (1) is functoriality of the differential. The second equality (2) is the definition of υ_u. The third equality (3) is the definition of the directional derivative as the scalar product with the gradient. Hence $\partial_\upsilon f_0 \geq 0$. Moreover, $\partial_\eta f_0 \geq 0$, because the vector field η is gradient-like for f_0. Notice that we assumed that Σ does not pass through any of the critical points. Hence, $\partial_\eta f_0$ is actually positive. If υ_u is a negative multiple of η, we must have $\partial_\upsilon f_0 = \partial_\eta f_0 = 0$, which can hold only if u is a critical point of f_0. But $u \in \Sigma$ is not a critical point. □

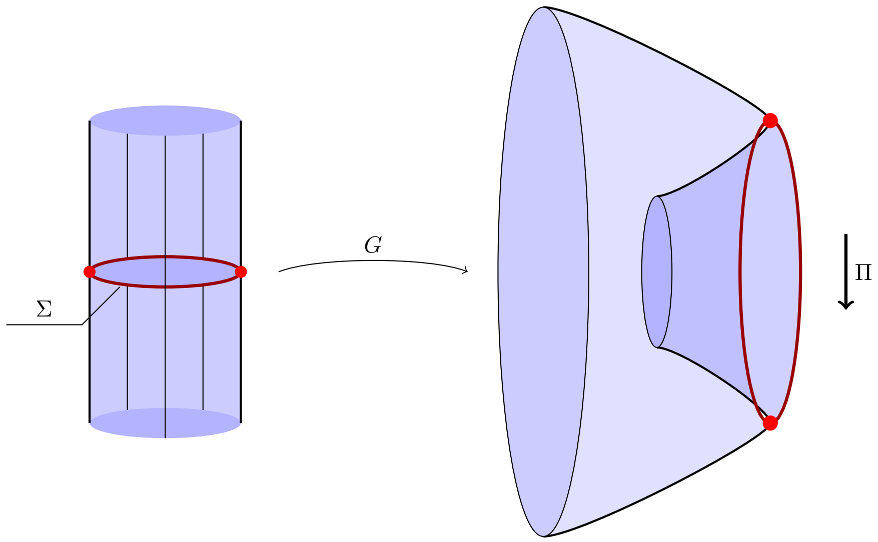

FIGURE 24. The set Σ and its stratification. The thickened points on Σ represent the second level of the Thom-Boardman stratification with respect to the map $\Pi \circ G$.

The key property of υ_u is that we can translate the condition on G_τ being an immersion into a condition involving μ and υ.

Lemma 4.3.24. *Suppose $u \in \Sigma$. Assume that the function μ satisfies*

$$\partial_\upsilon \mu(u) \geq 0. \tag{4.3.12}$$

Then G_τ is an immersion near u for all τ.

Proof. We know that $\dim \ker D(\Pi \circ G_0)(u) = 1$ is spanned by υ_u. Therefore, the kernel of $DG_\tau(u)$ can either be trivial, or it can be a one-dimensional subspace spanned by υ_u. To prove the lemma, we need to make sure that υ_u is not in the kernel of $DG_0(u)$. Since $f_\tau = F_0 \circ G_\tau$, if $\upsilon_u \in \ker DG_\tau(u)$, then $\partial_\upsilon f_\tau(u) = 0$. Hence, it is enough to show that (4.3.12) implies $\partial_\upsilon f_\tau(u) > 0$.

To this end, we use (4.3.9). Differentiating both sides of (4.3.9) with respect to υ, we obtain a formula analogous to (4.3.10).

$$\partial_\upsilon f_\tau(u) = (1 - \tau\mu)\partial_\upsilon f_0(u) + \tau(R(f_0(u)) - f_0(u))\partial_\upsilon \mu + \tau\mu R'(f_0(u))\partial_\upsilon f_0(u).$$

Note that $\partial_\upsilon f_0$ is positive, since it is $\partial_{\nabla F_0} F_0$ by (4.3.11). An analysis akin to the one following (4.3.9) leads to the statement that $\partial_\upsilon \mu(u) \geq 0$ implies that $\partial_\upsilon f_\tau > 0$. □

In light of Lemma 4.3.24 (and keeping in mind Corollary 4.3.21), we proceed to the main technical result of this subsection.

Proposition 4.3.25. *There exists a smooth function $\mu \colon f_0^{-1}[a + \varepsilon, b + \varepsilon] \to [0, 1]$ satisfying items (μ-1), (μ-2), and (μ-3) above and such that for any $u \in \Sigma$, $\partial_\upsilon \mu(u) \geq 0$ with equality only at points where $\mu(u) = 0, 1$.*

Proof. One might expect this proof to start with a function μ satisfying items (μ-1), (μ-2), and (μ-3), and then improving it to satisfy the statement of Proposition 4.3.25. However, there are technical problems near $\mu^{-1}(\{0, 1\})$. To avoid them, we will proceed in a slightly indirect way.

Let $\kappa > 0$ be an integer: we will describe how it is determined at the end of the proof. Choose a family $S_1, S_2, \dots, S_\kappa$ of open subsets of $f_0^{-1}(a+\varepsilon)$ with the properties that $\overline{S}_1 \subseteq Z_0$, $\overline{S}_{i+1} \subseteq S_i$ and $\mathcal{K}_- \cap f^{-1}(a+\varepsilon) \subseteq S_\kappa$. For $i = 1, \dots, \kappa$, let T_i be the subset of $f^{-1}[a+\varepsilon(1+\frac{i-1}{4\kappa}), b+\varepsilon(1+\frac{\kappa-i-1}{4\kappa})]$ consisting of the points that are reached from S_i by a trajectory of η. By construction

$$\overline{T}_{i+1} \subseteq T_i.$$

We note that here we assume η has no critical points other than q_+ in $f^{-1}[a+\varepsilon, b+\frac{5}{4}\varepsilon]$, and not just in $f^{-1}[a+\varepsilon, b+\varepsilon]$. But this does not pose any problems to arrange, e.g. by decreasing ε.

By Condition 4.3.15, no trajectory of η starting from Z hits Σ before reaching the level set $f^{-1}(a+2\varepsilon)$. This is because of Condition 4.3.15. We assume that the set S_1 is chosen in such a way that no trajectory of η starting from S_1 hits Σ before reaching the level set $f^{-1}(a+\frac{3}{2}\varepsilon)$.

We construct a function $\zeta: T_1 \to [0, \frac{5}{4}]$ with the following properties:

(ζ-1) $\partial_\eta \zeta > 0$ on T_1;
(ζ-2) Suppose $u \in T_1$ is connected to $S_1 \setminus S_\kappa$ by a trajectory of η. Then $\zeta(u) \in [0, \frac{1}{4}]$;
(ζ-3) ζ is greater or equal to 1 on $\mathcal{K}_- \cap f^{-1}[a+2\varepsilon, b+\varepsilon]$.

There is a correspondence between (μ-1) and (ζ-1), (μ-2) and (ζ-2), as well as between (μ-3) and (ζ-3). The precise connection will be made precise in due course, at the end of the proof; see (4.3.14) for the final definition of μ and the proof, using the conditions (ζ-1), (ζ-2), and (ζ-3), that the conditions (μ-1), (μ-2), and (μ-3) are satisfied for this definition (recall that we proved in Lemma 4.3.16 that a function μ satisfying these criteria exists, but we need a refined version to guarantee that G_τ is an immersion for all τ).

Lemma 4.3.26. *There exists ζ exists satisfying (ζ-1), (ζ-2), and (ζ-3).*

Sketch of proof. As these conditions resemble strongly the conditions for μ, the proof that ζ exists satisfying these properties is a straightforward adaptation of the proof of Lemma 4.3.16; we therefore leave the details to the reader. □

Our goal is to improve the function ζ so that $\partial_\upsilon \zeta(u) > 0$ for all $u \in \Sigma \cap T_\kappa$, and hence since the T_i are nested this proves that $\partial_\upsilon \zeta(u) > 0$ for all points u in $\Sigma \cap T_i$, for all i. We will need the following technical result.

Lemma 4.3.27. *Suppose X is a closed submanifold of Σ. Suppose U is an open subset of X such that neither υ, nor η is tangent to X at any point of $\overline{U}$. Then there exists a regular index triple $V \subseteq N$ (see Definition 2.7.5) such that $V \cap X = U$ and if a trajectory of η exits $\overline{V}$ at v_+ and then re-enters $\overline{V}$ at v_-, then $\zeta(v_-) > \zeta(v_+)$.*

Proof. Choose a Riemannian metric on Ω. Let E be the normal bundle for X in N. Since neither υ nor η is tangent to $\overline{U}$, we may and will assume that η and υ are orthogonal to $T\overline{U}$. As υ is not parallel to η either, we choose a metric in which η is orthogonal to υ. In particular, η and υ can be regarded as sections of the normal bundle to $\overline{U}$. Let L be the rank 1 subbundle of $E|_{\overline{U}}$ generated by η. Let $L^\perp$ be the orthogonal complement to L in $E|_{\overline{U}}$. We have that $\upsilon_u \in L^\perp_u$, a fact which we will not use until Lemma 4.3.28. For $\rho > 0$, we let $\widetilde{D}(\rho)$ denote the disc bundle of $L^\perp$ consisting of closed discs of radius ρ. By compactness of $\overline{U}$, we may and will choose ρ sufficiently small so that $\exp: \widetilde{D}(\rho) \to N$ is a diffeomorphism onto its image. Let $\overline{D}(\rho) \subseteq N$ be the image of $\widetilde{D}(\rho)$ under the exp map. Clearly, $\overline{D}(\rho) \cap X = \overline{U}$. We define $D(\rho)$ as the interior of $\overline{D}(\rho)$. Note that $\overline{D}(\rho)$ is a codimension one submanifold of N.

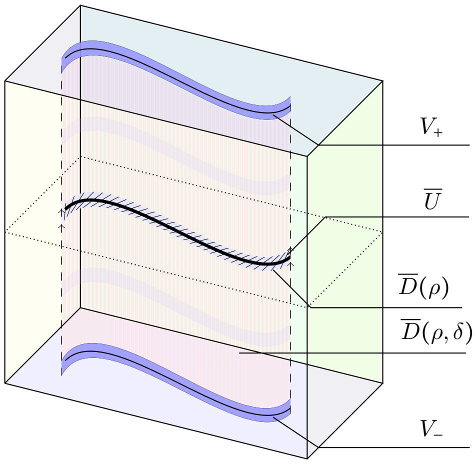

FIGURE 25. Proof of Lemma 4.3.28. Construction of $\overline{D}(\rho)$.

We have chosen η to be orthogonal to $L^\perp$. This means that η is transverse to $\overline{D}(\rho)$ at least at the points in $\overline{U}$. By openness of the transversality condition, we may and will decrease ρ in such a way that η is transverse to $\overline{D}(\rho)$ at every point of $\overline{D}(\rho)$.

We now thicken $\overline{D}(\rho)$ in the direction of η. Choose $\delta > 0$. By compactness of $\overline{D}(\rho)$ we may and will assume that any trajectory of η connecting two distinct points of $\overline{U}$ takes at least 3δ to travel. Define $\overline{D}(\rho,\delta)$ as the set of points on trajectories of η that can be reached from a point $u \in \overline{D}(\rho)$ in time in $[-\delta,\delta]$. With that choice of δ, the flow of η induces a diffeomorphism $\overline{D}(\rho,\delta) \cong \overline{D}(\rho) \times [-\delta,\delta]$.

We declare V to be the interior of $\overline{D}(\rho,\delta)$. The entry set $\partial_{\text{in}}V$ is the set of points that reach $\overline{D}(\rho)$ in time δ. The exit set $\partial_{\text{out}}V$ is the set of points that are reached from $\overline{D}(\rho)$ in time δ.

Note that if a trajectory of η exits $\partial_{\text{out}}V$ and then hits $\partial_{\text{in}}V$, it has to travel for time at least δ (otherwise there would be a trajectory connecting two points of $\overline{D}(\rho)$ in time less than 3δ). As ζ increases along the trajectories, if a trajectory leaves $\overline{V}$ at v_+ and then hits $\overline{V}$ at v_-, we have $\zeta(v_-) > \zeta(v_+)$. □

Remark. By compactness of N', the condition $\partial_\eta\zeta > 0$ can quickly lead to a universal constant $C_{\eta\zeta}$ such that $\zeta(v_-) > \zeta(v_+) + C_{\eta\zeta}$.

The next lemma is the key step in the proof of Proposition 4.3.25.

Lemma 4.3.28. *Suppose X is a closed submanifold of Σ, and let U be an open subset of X such that neither v nor η is tangent to X at any point of $\overline{U}$. Let V be as in Lemma 4.3.27. Let $\varepsilon > 0$ and U' be an open subset of U such that $\overline{U}' \subseteq U$. Set $V' \subseteq V$ to be the set obtained by thickening $\overline{U}'$ as in the proof of Lemma 4.3.27 with the same parameters ρ and δ.*

Then ζ can be altered to a function $\widetilde{\zeta}$ such that

- $\partial_v\widetilde{\zeta}(u) > 0$ *for all* $u \in \overline{U'}$;
- $|\widetilde{\zeta}(u) - \zeta(u)| < \varepsilon$;
- $\widetilde{\zeta}(u) = \zeta(u)$ *away from* V.

Proof. In the proof of Lemma 4.3.27, we chose a Riemannian metric on N. We will use the same metric in this proof. We begin with yet another thickening of $\overline{U}$, this time in the

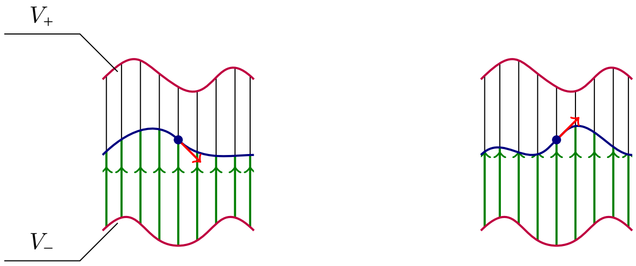

FIGURE 26. Proof of Lemma 4.3.28. The height function is ζ on the left and $\widetilde{\zeta}$ on the right. The diffeomorphisms λ preserve the trajectories of η, which are represented by vertical lines.

direction of ν_u. To this end, we choose $\theta_0 > 0$ with $\theta_0 < \rho$, and we allow ourselves to decrease θ_0 further, if need be. Consider the line bundle E over $\overline{U}$ spanned by the vector field ν_u. Note that $E \subseteq L^\perp$. The disc bundle $DE(\theta_0)$ is associated with the line bundle E, and the fibres are assumed to be closed and of radius θ_0. In particular, $DE(\theta_0)$ can be regarded as a subbundle of $\widetilde{D}(\rho)$. The exp map takes $DE(\theta_0)$ diffeomorphically onto its image. We denote this image by $E_v(\theta_0)$.

We need to introduce some notation. A point z in $E_v(\theta_0)$ is the image of a point $(u, \theta(z)) \in DE(\theta_0)$, where u is the base and $\theta(z) \in [-\theta_0, \theta_0]$ is the coordinate in the fibre. We also let $E_u \subseteq D_v(\theta_0)$, for $u \in \overline{U}$, be the image of the fibre of $DE(\theta_0)$ over u under the exp map.

For a point $v \in \overline{D}(\rho)$, we let v_-, v_+ denote the entry point and the exit point, respectively, of the trajectory of η. We choose δ to be small enough so that $\zeta(v_-) > \zeta(v_+) + \varepsilon$. In the rest of the proof, we will be assuming that η inside V is rescaled in such a way that a trajectory of η hitting v_- at $t = 0$, hits $\overline{D}(\rho)$ at time $t = 1/2$ and v_+ at time $t = 1$. This procedure is needed only to make the formulae more concise.

Our aim is to construct a family of orientation preserving self-diffeomorphisms λ_{v_-} of $[0,1]$, smoothly depending on a parameter $v_- \in \partial_{\mathrm{in}} V$. We require that λ_{v_-} be the identity:

- if v_- belongs to the boundary of $\partial_{\mathrm{in}} V$, that is, if the trajectory of η through v_- hits $\partial D_v(\theta_0)$ before leaving V;
- if the trajectory from v_- hits $\overline{U}$ before leaving V.

Next, suppose v_- passes through a point v in E_u, with $u \in \overline{U}'$. Suppose $\theta(v) < \theta_0/2$. We set λ_{v_-} to be the quadratic (or linear in the degenerate case) map such that $\lambda_{v_-}(0) = 0$, $\lambda_{v^-}(1) = 1$ and the condition on $\lambda_{v_-}(1/2)$ that is going to be specified momentarily. For $t \in [0,1]$, we set $\kappa(v_-, t) \in N$ to be point on the trajectory of η through v_- at time t, so that $\kappa(v_-, 0) = v_-$ and $\kappa(v_-, 1) = v_+$. As the trajectory of η hits $\overline{D}(\rho)$ at time value $1/2$, we have $\kappa(v_-, 1/2) \in \overline{D}(\rho)$. Note that the path $t \mapsto \kappa(v_-, \lambda_{v_-}(t))$ specifies the same trajectory of η as $t \mapsto \kappa(v_-, t)$, but after reparametrisation.

The last condition specifying λ_{v_-} is then

$$\zeta(\kappa(v_-, \lambda_{v_-}(1/2))) = \zeta(\kappa(u_-, 1/2)) + \theta(v) = \frac{1}{2} + \theta(v), \tag{4.3.13}$$

where the last equality follows from the second item of the bullet list above. The values at 0, 1/2, and 1 specify λ_{v_-} uniquely. Furthermore, it is easy to see that a quadratic function p attaining values 0 at 0 and 1 at 1 restricts to a diffeomorphism of $[0,1]$ if $p(1/2) \in (1/4, 1/2) \cup (1/2, 3/4)$, where $p(1/2) = 1/2$ is the case corresponding to the linear function. To see this, we write the general form of p as $p(x) = \frac{x(x-a)}{1-a}$, with $a \in \mathbb{R} \smallsetminus \{1\}$. Such a function is a diffeomorphism of the interval $[0,1]$ if and only if $a \notin [0,2]$. With $p(1/2) = q$, we have $a = \frac{4q-1}{4q-2}$. We have $a > 2$ if $q \in (1/2, 3/4)$, and $a < 0$ if $q \in (1/4, 1/2)$.

Moreover, λ_{v_-} depends smoothly on points $v_- \in \partial_{\text{in}} V$ such that a trajectory through v_- hits $E_v(\theta_0/2)$, the image under exp of the subbundle of the disc bundle of E whose fibres are discs of radius $\theta_0/2$. Note that the condition that $\overline{U}' \subseteq U$ guarantees that the definition of λ_{v_-} does not contradict the points in the itemized list.

If θ is sufficiently close to 0 (that is, if θ_0 is sufficiently small), then

$$\frac{1}{2} + \theta(v) = \zeta(\kappa(u_-, \lambda_{u_-}(1/2))) + \theta(v) \in (\zeta(\kappa(v_-, 0)), \zeta(\kappa(v_-, 1))),$$

If $|\theta(v)| < \frac{1}{4}$, which can be guaranteed by choosing $\theta_0 < \frac{1}{4}$, then λ_{v_-} is indeed a diffeomorphism.

We have defined λ_{v_-} on the boundary of V_- and on all the set of points v_- such that a trajectory of η through v_- hits $E_v(\theta_0)$. We extend λ_{v_-} smoothly through all points on V_-, on which it has not been defined yet. A possible extension is to specify a smooth map $v_- \mapsto \lambda_{v_-}(1/2)$, taking values in $[1/4, 3/4]$ since, as mentioned above, this gives rise to a smooth family of diffeomorphisms of $[0,1]$.

Now for a point $v \in V$ we set

$$\widetilde{\zeta}(v) = \zeta(\lambda_{v_-}(t_v)),$$

where t_v is the time to get from v_- to v; see Figure 26. With this definition, by (4.3.13), for $v \in L_u$ we have $\widetilde{\zeta}(v) = \widetilde{\zeta}(u) + \theta_u$. In particular, $\partial_v \widetilde{\zeta}(u) = 1$. Furthermore, as $\zeta(v_-) > \zeta(v_+) - \varepsilon$, and $\zeta(v_-) < \zeta(v), \widetilde{\zeta}(v) < \zeta(v_+)$, we obtain that $|\widetilde{\zeta}(u) - \zeta(u)| < \varepsilon$ as well. □

We continue proving Proposition 4.3.25. First, consider the map $\Pi \circ G_0 \colon T_1 \to Y$ (here Π is the projection along ∇F_0). As $\dim \ker D\Pi = 1$, there is a Thom-Boardman stratification of $\Pi \circ G_0$; see Subsection 2.3.2. We denote by Σ_m the part of the stratum $\beta^{\mathbf{m}}\Pi$ of this map contained in T_1. This is a smooth manifold with boundary. There is an index m_0 such that Σ_{m_0} is nonempty and all subsequent Σ_i are empty (there is a universal bound of m_0 in terms of the dimensions n and k, but we do not need it).

By definition, $\Pi \circ G_0$ restricted to any $\Sigma_m \setminus \Sigma_{m+1}$ is an immersion. That is, v_u is not tangent to $\Sigma_m \setminus \Sigma_{m+1}$. However, η might be tangent to $\Sigma_m \setminus \Sigma_{m+1}$. Therefore, consider the second projection $\Pi_\eta \colon N' \to f_0^{-1}(a + \varepsilon)$ along the trajectories of η. Let $\Sigma_{m,s}$ stand for the sth Thom–Boardman stratum $\beta^{\mathbf{s}}\Pi_\eta|_{\Sigma_m \setminus \Sigma_{m+1}}$. By Lemma 2.3.9, $\Sigma_{m,s}$ is a smooth manifold. Again, for any m, there is $s_m \in \mathbb{N}$ such that Σ_{m,s_m} is nonempty and $\Sigma_{m,i}$ is empty for all $i > s_m$. We can now finally define

$$\kappa = 1 + \sum_m s_m.$$

We proceed to prove that $\partial_v \zeta(u) > 0$ by induction on i. Consider $\Sigma_{m_0, s_{m_0}}$. This is a closed submanifold of T_1. Let $R \subseteq \Sigma_{m_0, s_{m_0}}$ be the set of points for which η is parallel to v_u. By Lemma 4.3.23, on R, η is a positive multiple of v_u. In particular, there is a neighbourhood U_R of R such that $\partial_v \zeta > 0$ everywhere on $\overline{U}_R$. We consider $\overline{U} \coloneqq \Sigma_{m_0, s_{m_0}} \setminus \operatorname{Int} U_R$, letting U be the interior of $\overline{U}$. We let $U' = U \cap T_2$ and $\varepsilon < \frac{1}{4\kappa}$. Using Lemma 4.3.28 we improve ζ on U in such a way that $\partial_v \zeta(u) > 0$ for all $u \in \overline{U}'$. This means that $\partial_v \zeta(u) > 0$ everywhere on $\Sigma_{m_0, s_{m_0}} \cap \overline{T}_2$.

We proceed with the induction step, inducting on i in T_i. Suppose that $\partial_v \zeta(u) > 0$ for $u \in \Sigma_{m,s+1} \cap \overline{T_i}$. We aim to improve ζ so that this also holds on $\Sigma_{m,s} \cap \overline{T}_{i+1}$. The procedure is analogous to the first step. Let R be the subset of $\Sigma_{m,s}$ such that η is parallel to v. For u in the compact set $(R \cup \Sigma_{m,s+1}) \cap \overline{T_i}$, we already know that $\partial_v \zeta(u) > 0$. In fact, on $R \cap \overline{T_i}$ this follows from Lemma 4.3.23, while on $\Sigma_{m,s+1} \cap \overline{T_i}$ this is the induction assumption. There

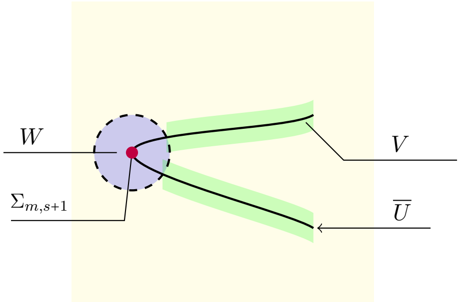

FIGURE 27. Proof of Proposition 4.3.25. Constructing $\overline{U}$.

exists an open subset $W \subseteq \Sigma_{m,s}$ containing $R \cup \Sigma_{m,s+1}$ such that $\partial_v\zeta(u) > 0$ on the whole of $\overline{W}$. We set $U = \Sigma_{m,s} \setminus W$ and $U' = U \cap T_{i+1}$; see Figure 27. We apply Lemma 4.3.28 to obtain that $\partial_v\zeta(u) > 0$ for $u \in \overline{U'}$. This means that $\partial_v\zeta(u) > 0$ on $\overline{W} \cup \overline{U'}$ which contains $\Sigma_{m,s} \cap \overline{T_{i+1}}$. This completes the induction step.

The inductive argument proves the statement until we reach $\Sigma_{m,0}$, that is, the statement holds for the whole of Σ_m. We declare Σ_{m-1,s_m+1} to be Σ_m, and we proceed with induction. The procedure continues until we conclude that $\partial_v\zeta > 0$ everywhere on $\Sigma \cap T_\kappa$.

The final function ζ has been constructed. Now, as the final step in the proof, we construct the function μ using the function ζ. We then show that μ satisfies the required conditions. Recall from Lemma 4.3.16 that constructing a function satisfying (μ-1), (μ-2), and (μ-3) was relatively straightforward. The key here is that we also need $\partial_v\mu(u) \geq 0$.

To this end, choose a smooth function $\chi\colon [0, \frac{5}{4}] \to [0,1]$ with the following properties.

(CH-1) χ maps $[0, \frac{1}{2}]$ to 0;
(CH-2) χ maps $[\frac{3}{4}, \frac{5}{4}]$ to 1;
(CH-3) $\chi'(x) > 0$ for $x \in (\frac{1}{2}, \frac{3}{4})$.

We set

$$\mu(u) = \chi(\zeta(u)). \tag{4.3.14}$$

To prove (μ-1) we use the property of (ζ-1) together with (CH-3). The chain rule yields that $\partial_\eta\mu > 0$ except at the locus where $\chi = \{0, 1\}$. At these latter points, $\partial_\eta\mu = 0$.

The same argument applies for the derivative along v_u, proving the last property for μ in the statement of Proposition 4.3.25.

Next, note that subsequent applications of Lemma 4.3.28 will not increase the value of ζ by more than $\frac{1}{4}$ (we apply this lemma κ times, and we choose $\varepsilon < \frac{1}{4\kappa}$). It follows from ($\zeta$-2) that ζ is less than $\frac{1}{2}$ at all points that are connected with $S_1 \setminus S_\kappa$ by a trajectory of η. Therefore, $\mu = 0$ on that set by (CH-1). Hence we can extend μ to the whole of $f^{-1}(a+\varepsilon, b+\varepsilon)$ by declaring it to be zero away from S_1. This proves (μ-2) and the second part of (μ-3). The fact (ζ-3) that initially $\zeta \geq 1$ on $\mathcal{K}_- \cap f^{-1}[a+2\varepsilon, b+\varepsilon]$ implies that after all the alterations we have $\zeta \geq \frac{3}{4}$ on that set. Hence by (CH-2), we have that $\chi \circ \zeta = 1$, proving (μ-3). We have proven that (μ-1), (μ-2), and (μ-3) hold. Also, we argued that $\partial_v(\mu) \geq 0$. So μ satisfies all the required conditions. ☐

4.3.6. **Enforcing condition (LR-2): arranging (IR-2).** We have shown that for any ∇F_0, the function μ can be changed in such a way that G_τ is a regular homotopy. Now, we will alter ∇F_0 to make sure that item (IR-2) of Addendum 4.3.2 is satisfied. Recall that this says that if $N = N_1 \sqcup \cdots \sqcup N_\ell$ and the images N_i and N_j (for all $i \neq j$) under G_0 are disjoint, then $G_\tau(N_i) \cap G_\tau(N_j) = \varnothing$ for all τ whenever $i \neq j$.

While pushing a point on $G_0(N)$ via the lift map P of Subsection 4.3.4 we might create new double points. In fact, we create a self-intersection of G_τ each time that $P(G_0(u), \tau\mu(u))$ hits another component of $G_0(N)$. In particular, self-intersections are created only among points that are identified under the map Π. We will use this principle to prove the next result.

Proposition 4.3.29. *Suppose $N = N_1 \sqcup \cdots \sqcup N_\ell$ and $G_0(N_1), \dots, G_0(N_\ell)$ are pairwise disjoint. Then there exists a vector field ∇F_0 such that $G_\tau(N_1), \dots, G_\tau(N_\ell)$ are pairwise disjoint for every $\tau \in [0,1]$.*

Proof. Choose a grim vector field ξ for F_0. It is usually not generic, in the sense that the map Π induced by ξ is not Thom-Boardman (see Definition 2.3.6), in the sense that the Thom-Boardman strata are not smooth of the expected dimension. However, ξ has the following important property.

Lemma 4.3.30. *No two points z_i, z_j with $z_i \in G_0(N_i) \cap \Omega'$ and $z_j \in G_0(N_j) \cap \Omega'$, $i \neq j$, are connected by a trajectory of ξ.*

Recall that $\Omega' = F_0^{-1}(a - \varepsilon, b + \varepsilon)$ as in (4.3.4).

Proof. The result relies on property (LR-1). If z_i and z_j are in distinct connected components of $G_0(N)$, the only trajectory of ξ that could possibly connect them has to lie in the zeroth stratum. As z_i, z_j are both on the first stratum or deeper, the only possibility that they can be connected by a trajectory in a shallower stratum is that both z_i and z_j are critical points of F_0 (this follows from the definition of the grim vector field; see Definition 2.5.1).

However, the fourth item of Lemma 4.3.5 implies that all critical points of F_0 in Ω' belong to the same component $G_0(N_i)$. This means that even if there is a trajectory in the zeroth stratum, it connects only critical points on the same connected component. ☐

Remark. We could also argue that there are no trajectories on the zeroth stratum for dimensional reasons. However, the present argument will also be used in proving an analogous result when lifting paths of death, where a dimension counting argument does not work.

Now suppose that $p_1, \dots, p_r$ are critical points of F_0 in Ω' (p_- and p_+ being in this set). The critical point p_i is assumed to belong to $G_0(N_{j_i})$. Choose balls $B_1, \dots, B_r$ around these points.

Lemma 4.3.31. *For sufficiently small balls $B_1, \dots, B_r$, there is no trajectory of ξ that passes through B_i and a component $G_0(N_j)$ with $j \neq j_i$. Also, there are no trajectories that connect two balls B_i and $B_{i'}$ whenever $j_i \neq j_{i'}$.*

Sketch of proof. If for every choice of balls there were a trajectory violating one of these conditions, on taking smaller and smaller balls, and passing to a limit we would conclude that there is a broken trajectory between two different components of $G_0(N)$. But such a broken trajectory would be composed of actual trajectories connecting components of $G_0(N)$ (no such trajectory can terminate at the zeroth stratum, because there are no critical points at depth 0). Therefore, there would be at least one unbroken trajectory connecting components of $G_0(N)$, contradicting Lemma 4.3.30. ☐

Now modify the vector field ξ inside $B_1, \dots, B_p$ so that ξ has no critical points in Ω'. That is, replace the coordinate of y_{11} in (2.5.1) from $\sum y_{1j}^2$ with $\phi + (1-\phi)\sum y_{1j}^2$, where ϕ is a bump function supported at B_- and equal to 1 at p_- (respectively supported on B_+ and equal to 1 at p_+). Let ξ' be the replaced vector field. It clearly satisfies the statement of Lemma 4.3.31, so it also satisfies the statement of Lemma 4.3.30.

Now note that the condition in the statement of Lemma 4.3.30 is open. Hence, we may perturb ξ' to ξ'', in such a way that ξ'' satisfies the genericity condition of Lemma 4.3.14 and such that no trajectory of ξ connects two different components of $G_0(N)$. Now ξ'' is gradient-like for F_0 in Ω'. We choose a metric in such a way that $\nabla F_0 = \xi''$ (compare Theorem 2.5.2). □

By Proposition 4.3.29, the previous proof with this choice of ∇F_0 gives the desired statement, with the addition that (IR-2) holds: if $G_0(N_i) \cap G_0(N_j)$ for $i \neq j$, then no new intersections are created between the different components.

4.3.7. **Condition (LR-2) for small dimensions: arranging (IR-3).** In this subsection, we focus on the case $k = 2$. The first lemma holds for $k > 2$ as well, however, so we state it in general.

Lemma 4.3.32. *Let $k \geq 2$. Suppose $h_- > n - k$ or $h_+ < k$. Then, for generic G_0, Condition (LR-2) is automatically satisfied.*

Proof. If $h_- > n - k$, the manifold $\mathcal{K}_-$ has dimension strictly less than k. The set of points in N that are mapped to the second stratum by G_0 has codimension k, so if $h_- > n - k$, this set is generically avoided by $\mathcal{K}_-$. If $h_+ < k$, then $\dim \mathcal{K}_+ < k$, and we repeat the argument. □

Corollary 4.3.33. *Suppose $k = 2$ and $n \in \{1, 2, 3\}$. Then condition (LR-2) is satisfied for generic G_0.*

Proof. Suppose $n = 3$. The only situation when $\dim \mathcal{K}_+ \geq k$ is that $h_+ = 2, 3$. In both cases, we know that $h_- \geq 2, 3$, because $h_- \geq h_+$. Then $h_- > n - k = 1$, so we apply Lemma 4.3.32. The argument for $n = 1, 2$ is analogous. □

Remark. If $n = 4$, Condition (LR-2) is satisfied for generic G_0 unless $h_- = h_+ = 2$. Indeed, for $n = 4$, $\dim \mathcal{K}_+ \geq 2$, $\dim \mathcal{K}_- \leq 2$ and $\dim \mathcal{K}_+ = \dim \mathcal{K}_-$ can hold only when $h_- = h_+$.

4.3.8. **Proof of Lifting Rearrangement Lemma 4.3.1.** We are now in position to prove Lifting of Rearrangement Lemma without assuming (LR-1) and (LR-2) at the beginning. The proof consists of improving the initial pair (F_0, G_0) so as to assert (LR-2). Then, we lift rearrangement. The key difficulty we address in this subsection is to show that the path (F_τ, G_τ) we construct is indeed a weak lift, that is $F_\tau \circ G_\tau$ is left-homotopic to the original path. This relies on patiently controlling each step improving (F_0, G_0) to satisfy conditions (LR-1) and then (LR-2). A schematic of the proof is given in Figure 28.

Consider an immersed Morse function $F_0 \colon \Omega \to \mathbb{R}$ and a vector field η on N that is gradient-like for $f_0 = F_0 \circ G_0$. We denote the original η-path of rearrangement by f_τ.

At first, we guarantee condition (LR-1). Lemma 4.3.5 creates a double path (F_τ, G_τ), $\tau \in [0, 1]$, such that F_τ is a ξ-path of rearrangement, F_τ is supported away from $\mathcal{L}_- \cup \mathcal{L}_+$. Set $\widetilde{f}_\tau := F_\tau \circ G_\tau$. Note that F_τ is not moving any critical points at depth 1, hence $\widetilde{f}_\tau$ is actually a path of excellent Morse functions. As ξ was constructed independently from η, we cannot expect that η is gradient-like for $\widetilde{f}_1$. However, it is so near $\mathcal{K}_- \cup \mathcal{K}_+$, which is enough for our purpose, as we now explain.

Lemma 4.3.34. *There is a vector field $\widetilde{\eta}$ on N that agrees with η near $\mathcal{K}_-$ and $\mathcal{K}_+$ and such that $\widetilde{f}_\tau$ is a $\widetilde{\eta}$-path of excellent Morse functions.*

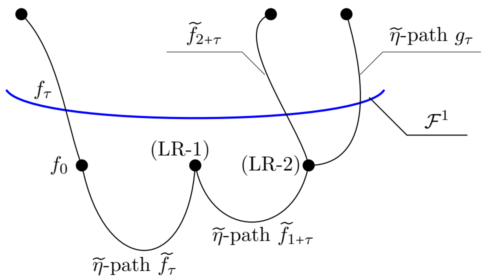

FIGURE 28. Paths created in the proof of Lifting Rearrangment Lemma 4.3.1.

Proof. Choose $U_1, U_2 \subseteq N$ to be two open subsets with the property that $U_1 \supset \overline{U_2} \supset (\mathcal{K}_- \cup \mathcal{K}_+)$ and such that $\widetilde{f}_\tau$ is supported on $N \setminus U_1$. This is possible because Lemma 4.3.5 uses only safe rearrangements. Shrink U_1 so that U_1 contains only q_-, q_+ as critical points of f_0.

Let ϕ be a bump function equal to 1 on U_2 and supported on U_1. Let η_ξ be the pull-back of ξ via G_1, as in Subsection 2.5.2. Then η_ξ is a gradient-like vector field for g_τ. Also, restricted to U_1, η is gradient-like vector field for $\widetilde{f}_\tau$, because $\widetilde{f}_\tau|_{U_1} = \widetilde{f}_0|_{U_1}$. Set $\widetilde{\eta} = \phi\eta + (1-\phi)\eta_\xi$. Then $\widetilde{\eta}$ is gradient-like for $\widetilde{f}_1$ and agrees with η on U_2. □

Next, Proposition 4.3.12 applied to $(F_1, G_1, \widetilde{\eta})$ creates another double path, which we now denote $(F_{1+\tau}, G_{1+\tau})$, $\tau \in [0,1]$, such that $\widetilde{f}_{1+\tau} := F_{1+\tau} \circ G_{1+\tau}$ is a path of excellent Morse functions, $\widetilde{\eta}$ is gradient-like for all $\widetilde{f}_{1+\tau}$, and $\widetilde{f}_{1+\tau}$ does not depend on τ near $\mathcal{K}_- \cup \mathcal{K}_+$. The main outcome of Proposition 4.3.12 is that G_2 satisfies (LR-2).

Let $\widetilde{f}_{2+\tau}$ be the $\widetilde{\eta}$-path of rearrangement starting from $\widetilde{f}_2 = F_2 \circ G_2$. The pair of functions (F_2, G_2) satisfies (LR-2). Consider the $\widetilde{\eta}$-path g_τ, $\tau \in [0,1]$ that starts at $g_0 := \widetilde{f}_2$ and lifts the critical point q_- above the critical point q_+. By Lemma 4.3.4, the path can be weakly lifted. That is, there exists a double path $(F_{2+\tau}, G_{2+\tau})$, $\tau \in [0,1]$, such that $\widetilde{f}_{2+\tau} := F_{2+\tau} \circ G_{2+\tau}$ is weakly homotopic to $g_{2+\tau}$.

Lemma 4.3.35. *The path $\widetilde{f}_{2+\tau}$ is lax homotopic to f_τ over $\widetilde{f}_\tau$, $\tau \in [0,2]$.*

Proof. As $\widetilde{f}_{2+\tau}$ is left-homotopic to g_τ, it is enough to show that g_τ is lax homotopic to f_τ over $\widetilde{f}_\tau$. This is precisely the statement of the lemma. We have set $h_{\sigma,0} = \widetilde{f}_{2\sigma,0}$ is a path of excellent Morse functions supported away from a neighbourhood of $\mathcal{K}_-$ and $\mathcal{K}_+$. The path f_τ is an η-path, while the path g_τ is a $\widetilde{\eta}$-path, and the two vector fields agree in a neighbourhood of $\mathcal{K}_- \cup \mathcal{K}_+$. □

Corollary 4.3.36. *The path (F_τ, G_τ), $\tau \in [0,3]$ can be promoted to a weak lift of f_τ.*

Proof. Lemma 4.3.35 implies that the paths (F_τ, G_τ), $\tau \in [0,3]$ satisfy the assumptions of Lemma 4.1.6. □

To conclude the proof of Lifting Rearrangment Lemma 4.3.1, we need to show that G_τ has the properties stated in Addendum 4.3.2. This is done with a case-by-case analysis.

- If G_0 is an embedding, then (LR-2) is always satisfied. Therefore, we do not create self-intersections by changing G_0 in such a way that (IR-1) holds;
- If N is a disjoint union $N_1 \sqcup \cdots \sqcup N_\ell$ with $G_0(N_i)$ and $G_0(N_j)$ disjoint for each $i \neq j$, then we use the methods from Subsection 4.3.6 to choose an appropriate vector field ξ so that condition (IR-2) is guaranteed.

- If $k = 2$ and $\dim N \leq 3$, then we use Subsection 4.3.7 to show that (LR-2) is automatically satisfied.

We have constructed a weak lift of the rearrangement by a regular double path, that does not introduce critical points on the zeroth or first stratum, and we have shown that the conditions in Addendum 4.3.2 can be arranged, when appropriate. This completes the proof of Lemma 4.3.1.

4.4. Lifting paths of death

We now pass to the most important, and technically most difficult question, namely lifting paths of death. In the proof of Lemma 4.4.1, we are going to use the finger move, described in detail in Part 5.

Lemma 4.4.1 (Lifting paths of death)**.** *Let $G\colon N \looparrowright \Omega$ be a generic immersion with $\dim N = n$, $\dim \Omega = n+k$, and $k \geq 2$. Suppose $F\colon \Omega \to \mathbb{R}$ is an immersed Morse function and $f_\tau\colon N \to \mathbb{R}$, for $\tau \in [0,1]$, is an elementary path of death such that $f_0 = F \circ G$.*

Then there exists a regular double path (F_τ, G_τ) that weakly lifts f_τ.

Suppose moreover that the level sets of the critical points that are cancelled are a, b. The path F_τ is supported on $F_0^{-1}[a-\varepsilon, b+\varepsilon]$, while G_τ is supported on $f_0^{-1}[a-\varepsilon, b+\varepsilon]$, where $\varepsilon > 0$ can be chosen as small as we please. In particular, if f_τ is neat, G_0 is neat and F_0 is neat, then (F_τ, G_τ) is neat.

As in the case of the Lifting Paths of Rearrangement Lemma 4.3.1, we give the follow-up result, which specifies additional properties of G_τ under certain conditions;

Addendum 4.4.2. *The map G_τ has the following properties.*

- *If $k > 2$, and G_0 is an embedding, then G_τ is an embedding for all τ.*
- *If $k = 2$, and G_0 is an embedding, then any new self-intersections are only added via a finite number of finger moves.*
- *If $k = 2$ and $n \in \{1, 2\}$, then any new self-intersections are only added via a finite number of finger moves. If $n = 1$ self-intersections only occur at finitely many values of $\tau \in (0,1)$.*
- *If $N = N_1 \sqcup \cdots \sqcup N_\ell$ and $G_0(N_1), \ldots, G_0(N_\ell)$ are pairwise disjoint, then for all $\tau \in [0,1]$, we also have that $G_\tau(N_1), \ldots, G_\tau(N_\ell)$ are pairwise disjoint.*

The proof of Lemma 4.4.1 occupies the whole of Section 4.4, except for the finger move, which occupies the whole of Part 5. Schematically, the proof is similar to the proof of the Lifting Paths of Rearrangement Lemma 4.3.1. First we set up notation, then we prove Lemma 4.4.1 under the conditions (LD-1)–(LD-5) specified in Subsection 4.4.2. The next subsections, 4.4.3, 4.4.4, 4.4.5, 4.4.6, and 4.4.7, show how to arrange the function F_0 and the embedding G_0 in such a way that the conditions (LD-1)–(LD-5) (from Subsection 4.4.2) are satisfied. Finally, in Subsection 4.4.9, we give a precise summary of all steps in the proof of Lemma 4.4.1.

4.4.1. **Notation of the proof of Lemma 4.4.1.** We let q_-, q_+ denote the critical points of f that are cancelled along f_τ. The indices are h_- and h_+ ($h_+ = h_- + 1$). By Assertion 3.6.7, it is enough to consider f_τ that is an elementary path of death from Definition 3.6.6. We choose a double neighbourhood for f and we let η be a gradient-like vector field for f such that this neighbourhood is adapted to the pair of good caps (D_η, A_η). (It may help to refer back to Section 3.6 for the terminology.)

Let γ be the unique trajectory of η connecting q_- and q_+. As $\dim\gamma = 1$ and the union of strata $\bigcup_{d\geq 2}\Omega[d]$ is of codimension 2, we may perturb G_0 to a generic immersion such that $G_0(\gamma)$ is disjoint from the second stratum and all higher strata. Note that a perturbation can be realised as a path G_τ of F_0-regular generic immersions such that $F_0 \circ G_\tau$ is a path of an excellent Morse functions. Hence, from now on, we will be assuming that $G_0(\gamma)$ is disjoint from $\bigcup_{d\geq 2}\Omega[d]$.

Set $a = f(q_-)$, $b = f(q_+)$. Choose a neighbourhood U_0 of γ such that G_0 maps U_0 to the first stratum. Write $\mathcal{K}_- \subseteq N$, respectively $\mathcal{K}_+ \subseteq N$ for the unstable manifold of q_-, respectively for the stable manifold of q_+. We choose $\varepsilon > 0$ such that $\mathcal{K}_- \cap f^{-1}[a, a+\varepsilon]$ and $\mathcal{K}_+ \cap f^{-1}[b-\varepsilon, b]$ are contained in U_0. Note that $\gamma \subseteq \mathcal{K}_- \cap \mathcal{K}_+ \cap f^{-1}[a,b]$. Define $\mathcal{L}_- = G_0(\mathcal{K}_-)$, $\mathcal{L}_+ = G_0(\mathcal{K}_+)$. As usual, we set $p_+ = G_0(q_+)$ and $p_- = G_0(q_-)$.

4.4.2. **Lifting paths of death under extra conditions.** We will structure the proof using the following conditions on η, f, and G_0.

(LD-1) There are no critical points of F at depth 0 with critical values in $[a-\varepsilon, b+\varepsilon]$;
(LD-2) The unique trajectory γ on N, between q_- and q_+, is mapped to the first stratum;
(LD-3) The subsets $\mathcal{K}_-$ and $\mathcal{K}_+$ of N are mapped to the first stratum of $G_0(N)$;
(LD-4) There exists a grim vector field ξ for F such that the pull-back $\widetilde{\eta}$ of ξ via G_0 gives rise to a path of death that is left-homotopic to a path of death constructed with η, and moreover there are no critical points of F in $F^{-1}[a-\varepsilon, b+\varepsilon]$ other than p_- and p_+;
(LD-5) There are no trajectories of ξ from p_- to p_+ other than the image of the trajectory γ.

Remark. (LD-1) and (LD-3) are the analogues of (LR-1) and (LR-2) respectively; however (LD-3) requires that *both* $\mathcal{K}_-$ and $\mathcal{K}_+$ are mapped to the first stratum, whereas (LR-2) required only one of them.

A condition on the lack of other critical points appears twice, in (LD-1) and in (LD-4). First, before we construct an appropriate vector field, we are allowed to perform only safe rearrangements (see the discussion in the proof of Lemma 4.3.7). The lack of critical points at depth 0 is important in ensuring (LD-3). Once we have a vector field extending η, we have more control when we perform rearrangements, and this enables us to move all critical points away from between p_- and p_+. Arranging for (LD-5) to hold is rather difficult: it will require the finger move, which takes the whole of Part 5 to construct. Removing the other critical points and arranging that there are no trajectories other than γ are both needed to apply Cancellation Theorem 3.4.1.

Lemma 4.4.3 (Conditional lifting of paths of death)**.** *Suppose conditions (LD-4) and (LD-5) are satisfied. Then there exists a path F_τ such that (F_τ, G_0) weakly lifts f_τ.*

Proof. The Cancellation Theorem 3.4.1 uses a vector field ξ to create a regular double path (F_τ, G_τ), $\tau \in [0,2]$ that cancels the pair of critical points p_- and p_+. This is possible since (LD-5) is satisfied. Then the composition $F_\tau \circ G_\tau$ is an $\widetilde{\eta}$ path of death. By (LD-4), it is left-homotopic to the original path of death constructed with η. By Corollary 4.1.4, the path (F_τ, G_τ) can be promoted to a weak lift of the original η-path. □

Remark. In the proof of Lemma 4.4.3 we did not use properties (LD-1), (LD-2), and (LD-3). However these properties will be needed to construct a vector field ξ satisfying (LD-4) and (LD-5).

4.4.3. **Removing critical points from** $[a-\varepsilon, b+\varepsilon]$**.** In this section we arrange that condition (LD-1) holds. The procedure uses the rearrangement theorem and the approach is similar to the approach of Subsection 4.3.3. In particular, we might need to replace the original vector field η by another gradient-like vector field, $\widetilde{\eta}$, which agrees with the original one near $\mathcal{K}_- \cup \mathcal{K}_+$.

Lemma 4.4.4. *There exists a path of Morse functions $F_\tau : \Omega \to \mathbb{R}$, such that $F_\tau = F_0$ on U_0, F_1 has no critical points in $[a-\varepsilon, b+\varepsilon]$ at depth 0 and $F_\tau \circ G_0$ is a path of Morse functions on N with constant critical values, and such that the support of $F_\tau \circ G_0$ is disjoint from $\mathcal{K}_- \cup \mathcal{K}_+$.*

Proof. The proof follows the argument of Lemma 4.3.5 and discovers what are the depths and indices of the critical points that can and cannot be moved out of the region $F_0^{-1}[a-\varepsilon, b+\varepsilon]$. The only difference with Lemma 4.3.5 is that $h_+ = h_- + 1$, where we recall that $h_\pm$ is the index of $q_\pm$.

We begin with the case $k = 2$, which is the most rigid. Choose a grim vector field ξ for F. By Lemma 4.3.7, a critical point p can be safely (see Definition 4.3.9) moved up by a ξ-path if $h_p + d_p > h_+ + 1 = h_- + 2$ (or, if $d_p = 0$), and p can be moved down by a ξ-path if $h_p + d_p < h_- + 1$. Acting as in the proof of Lemma 4.3.5 we can move all critical points with $h_p + d_p \leq h_-$ below p_- and all critical points with $h_p + d_p \geq h_- + 3$ above p_+. There remain critical points with $h_p + d_p = h_- + 1$ or $h_p + d_p = h_- + 2$. These points cannot be safely rearranged in general.

Suppose a point p' is between p_- and p_+ and $d_{p'} = 0$. If $h_{p'} = h_- + 1$, we can safely move it below p_-; here any rearrangement with other critical points is safe and possible, because the only critical points have $h_p + d_p \geq h_- + 1$. Compare the algorithm for handling (SR-2) in the proof of Lemma 4.3.10 above. On the other hand, if $h_{p'} = h_- + 2$, we move it up.

Concatenating the paths of rearrangement, we construct a ξ-path F_τ supported away from $\mathcal{L}_- \cup \mathcal{L}_+$. The only critical points of F_1 in $F_1^{-1}(a,b)$ are at depth 2 or more (recall that the critical points at depth 1 correspond to critical points of $F_\tau \circ G_\tau$, and since we consider a path of death there are none; see Section 3.6) and are such that $h_p + d_p = h_- + 1$ or $h_p + d_p = h_- + 2$. This concludes the case $k = 2$.

Suppose $k > 2$. Lemma 4.3.7 allows us to safely move a critical point p below p_- as long as $h_p + d_p < h_- + 1$. If $h_p + (k-1)(d_p - 1) > h_+ = h_- + 1$, we can safely move a point p above p_+. The only possibility that a critical point cann be safely moved neither above p_+ nor below p_- is if $h_p + d_p \geq h_- + 1$ and $h_p + (k-1)(d_p - 1) \leq h_+ = h_- + 1$ simultaneously. These two conditions imply that $h_p + d_p \geq h_p + (k-1)(d_p - 1)$, that is, $d_p \leq \frac{k-1}{k-2}$. Since $d_p \geq 2$ and $k \geq 3$ by assumptions, the only possibility is that $d_p = 2$ and $k = 3$. In that case, we must have $h_p + d_p = h_- + 1$. ☐

We remark that the rearrangements performed during the proof of Lemma 4.4.4, might result in η not being a gradient-like vector field for $f_1 := F_1 \circ G_0$. However, $f_1 = f_0$ near $\mathcal{K}_- \cup \mathcal{K}_+$. Therefore, η is a gradient-like vector field for f_1 near $\mathcal{K}_- \cup \mathcal{K}_+$. By Proposition 2.5.3, there exists a gradient-like vector field $\widetilde{\eta}$ for f_1 agreeing with η near $\mathcal{K}_- \cup \mathcal{K}_+$.

We can refine the result if N is a union of disjoint components, and G_0 maps all these components to pairwise disjoint stratified manifolds.

Lemma 4.4.5. *Suppose $N = N_1 \sqcup \cdots \sqcup N_\ell$, the images $G_0(N_1), \ldots, G_0(N_\ell)$ are pairwise disjoint, and p_+, p_- belong to the component $G_0(N_1)$. Then the path of rearrangement from Lemma 4.4.4 can be chosen to additionally push all critical points away from $F_0^{-1}(a,b)$, except for those that belong to N_1.*

Remark. Unlike in Lemma 4.3.5, we need not consider the case where p_- and p_+ are in separate components; the existence of a trajectory in $G_0(N)$ connecting q_- to q_+ implies that p_- and p_+ belong to the same connected component of $G_0(N)$.

Proof of Lemma 4.4.5. Lemma 4.4.4 allows us to move critical points away from $F_0^{-1}(a,b)$ under certain conditions on $h_p + d_p$ depending on the codimension k. If $k > 3$, we are able to move all critical points, so the statement of Lemma 4.4.5 follows from Lemma 4.4.4.

Assume $k = 2$. By Lemma 4.4.4, only critical points that cannot be safely moved away from $F_0^{-1}(a,b)$ are those for which $h_p + d_p = h_- + 1$ or $h_p + d_p = h_- + 2$, and $d_p \geq 2$.

The proof in that case follows the same lines as the proof of the fourth item of Lemma 4.3.5, so here we only give a quick sketch. If a critical point p has $h_p + d_p = h_- + 1$, its descending membrane is disjoint from the ascending membrane of p_- for dimensional reasons. Next, if p and p_- belong to different components of $G_0(N)$, the membrane of p belongs to a different component of $G_0(N)$ than $\mathcal{L}_- = G_0(\mathcal{K}_-)$. Moreover, for dimensional reasons, the descending membrane of p is disjoint from all the ascending membranes of critical points between p_- and p. Hence, p can be safely moved below the level set of p_-.

In this way, we can push all critical points with $h_p + d_p = h_- + 1$ below the level set of p_-, except for those that belong to the same component of $G_0(N)$ as p_-. Reversing this construction, we push all critical points with $h_p + d_p = h_- + 2$ above the level set of p_+, with the exception of critical points belonging to the same component as p_+.

For $k = 3$ we essentially repeat the argument, except that we only need to move critical points with $h_p + d_p = h_- + 1$ and $d_p = 2$. This completes the proof that (LD-1) can be arranged, and moreover if the $G_0(N_i)$ are pairwise disjoint, the only critical points remaining in $F^{-1}(a - \varepsilon, b + \varepsilon)$ are on the connected component $G_0(N_1)$. □

4.4.4. **Proving (LD-2).** In order to prove (LD-3) (the analogue of (LR-2)), we need an intermediate step. This is the place where the present situation differs from the proof of Proposition 4.3.12 used for lifting rearrangements in the previous section. As in Subsection 4.3.4, we will be working with a gradient-like vector field for F_0 as an ordinary Morse function. Again, to distinguish these gradient-like vector fields from grim vector fields, we use the notation ∇F_0. Note that any gradient-like vector field is a gradient for some choice of Riemannian metric, so this is not a significant abuse of notation.

Lemma 4.4.6. *There exists a vector field ∇F_0, gradient-like for F_0, such that with $\Omega' = F_0^{-1}[a - \varepsilon, b + \varepsilon]$ we have:*

- *the projection Π from Ω' to the level set $F^{-1}(a - \varepsilon)$ is a Thom-Boardman map (see Definition 2.3.6);*
- *suppose $N = N_1 \sqcup \cdots \sqcup N_\ell$ and that $G_0(N_i)$ are pairwise disjoint for $i = 1\dots,\ell$. Then there are open balls near critical points of F_0, which we can choose to be as small as we please, such that there are no trajectories of ∇F_0 that connect two distinct components $G_0(N_i)$ and $G_0(N_j)$.*

Proof. The first item follows from the arguments that have already been used in the proof of Proposition 4.3.12. Namely, the map Π is defned as the projection onto $F^{-1}(a - \varepsilon)$ along the trajectories of ∇F_0. By Lemma 4.3.13, a small perturbation of Π can be realized by a perturbation of ∇F_0 within the family of gradient-like vector fields for F_0. Furthermore, Proposition 2.3.7 guarantees, that any Π can be perturbed to a Thom–Boardman map. That is, any gradient-like vector field ∇F_0 can be perturbed to a gradient-like vector field satisfying the first condition. The second part is proved in the same way as Lemmas 4.3.30 and 4.3.31. □

We need to add an extra condition regarding the vector field ∇F_0.

Definition 4.4.7. We say that γ is in *good position* with respect to G_0 if for every $w \in G_0(\gamma)$ the trajectory of ∇F_0 through w does not hit $G_0(N) \cap \Omega'$, except at w.

Lemma 4.4.8.

- *If $k > 2$, then for a generic gradient-like vector field ∇F_0, γ is in a good position with respect to G_0.*
- *If $k = 2$, and F_0 satisfies (LD-1), then there exists a regular F_0-path of immersions G_τ such that G_τ is fixed on γ, η is a gradient-like vector field for $F_0 \circ G_\tau$, for all $\tau \in [0,1]$, and γ is in good position with respect to G_1.*
- *If $N = N_1 \sqcup \cdots \sqcup N_\ell$ and $G_0(N_1), \dots, G_0(N_\ell)$ are pairwise disjoint, then $G_\tau(N_1), \dots, G_\tau(N_\ell)$ are also pairwise disjoint for all $\tau \in [0,1]$.*

Proof. For each point z in γ, let ν_z denote the set of points in Ω' that lie on the trajectory of ∇F_0 through $G_0(z)$. The set $C = \bigcup_{z \in \gamma} \nu_z$ is two-dimensional and contains $G_0(\gamma)$ by construction. This set can be reinterpreted as follows: the flow of ∇F_0 induces a projection Π of Ω' onto the level set $F_0^{-1}(a - \varepsilon)$; see Lemma 4.3.13 above and the discussion in Subsection 4.3.4. Suppose Π is generic with respect to γ, that is, it satisfies the following two conditions.

(1) First $\beta^\ell \Pi|_\gamma$ is smooth of expected dimension (where $\beta^{\mathbf{i}}\gamma$ denotes the Thom–Boardman stratification, see Subsection 2.3.2). As $\dim \gamma = 1$, the expected dimension of $\beta^1\gamma$ is negative. This means that $\beta^{\mathbf{1}}\gamma$ is empty, so that $\Pi|_\gamma$ is a generic immersion.
(2) Secondly, we assume that the set of double points of the image is of expected dimension. This set is of codimension $(n + k - 3)$ in $\Pi(\gamma)$. As we assumed that $k \geq 2$, the set of double points is empty unless $n = 1$ and $k = 2$. This means that $\Pi^{-1}\Pi(\gamma)$ is either a smooth surface (if $n > 1$ or $k > 3$) or a surface with 1-dimensional set of singularities: each such singularity is a product of an ordinary double point of a curve and a segment.

These two conditions imply that in both cases, $\Pi^{-1}\Pi(\gamma)$ is Whitney stratified (Definition 2.1.1). The set C is an open subset of $\Pi^{-1}\Pi(\gamma)$, so it inherits Whitney stratification. In particular, stratified general position arguments can be applied to C as follows.

If $k > 2$, the set C can be made disjoint from $G_0(N) \smallsetminus G_0(\gamma)$ (which is of codimension $k > \dim C$) by perturbing Π. This proves the first item.

For the remaining part of the proof, we will assume that $k = 2$. As C is stratified of dimension 2, a general position argument guarantees that C intersects $G_0(N) \smallsetminus G_0(\gamma)$ at finitely many points, and each such point is a smooth point of C (because the singular set of C has dimension $1 < \operatorname{codim} G_0(N) \smallsetminus G_0(\gamma)$). Moreover, C avoids the second stratum of $G_0(N)$. Let $w_1, \dots, w_t \in N$ be such that $C \cap (G_0(N) \smallsetminus G_0(\gamma)) = \{G_0(w_1), \dots, G_0(w_t)\}$; see Figure 29.

Similarly to Subsection 4.3.5, define $\Sigma \subseteq N$ to be the subset of points $u \in N$ such that ∇F_0 is tangent at $G_0(u)$ to the branch of $G_0(N)$. Denote the collection of trajectories of η through points on $w_1, \dots, w_t$ by $\gamma_1, \dots, \gamma_t$. The curves $\gamma_1, \dots, \gamma_t$ are one-dimensional, while $\Sigma \subseteq N$ is of codimension k. By general position arguments (choosing appropriate Π) we may and will assume that all the curves γ_i are disjoint from Σ.

The original trajectory γ is not any of the $\gamma_1, \dots, \gamma_t$, therefore, each of the γ_i either does not terminate at q_+, or does not start at q_- (or both). Suppose γ_1 does not terminate at q_+. As γ_1 is a trajectory of η, the fact that γ_1 does not terminate at q_+ implies that γ_1 is disjoint from $\mathcal{K}_+$.

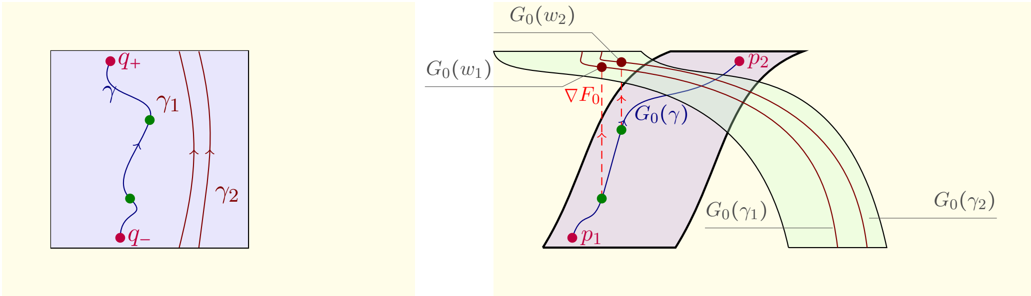

FIGURE 29. Notation of proof of Lemma 4.4.8.

Our aim is to lift γ_1 so that the point w_1 lands above the level $f^{-1}(b+\varepsilon)$, and then to apply induction. To accomplish the first goal, we use the procedure of the proof of Proposition 4.3.12. Set $c_1 = f(w_1)$ and $\varepsilon' > 0$ such that $c_1 - \varepsilon' > a$.

Choose a neighbourhood U_1 of $\gamma_1 \cap f^{-1}(c_1 - \varepsilon', b+\varepsilon)$ disjoint from $\mathcal{K}_+$ and all the other curves $\gamma_2, \dots, \gamma_t$. We assume that all trajectories of η passing through U_1 avoid Σ in $f^{-1}([c_1 - \varepsilon', b+\varepsilon])$. Such a choice of U_1 is possible, since γ_1 avoids Σ itself. Pick a function

$$\mu\colon f^{-1}[a+\varepsilon, b+\varepsilon] \to [0,1]$$

satisfying analogues of items (μ-1), (μ-2), and (μ-3). Namely:

- $\partial_\eta \mu \geq 0$;
- μ is supported on U_1;
- $\mu = 1$ on γ_1.

Existence of such μ follows from the same arguments as in Lemma 4.3.16. Note that in this specific case, we need not care about the condition $\partial_\upsilon \mu \geq 0$ as in Lemma 4.3.24. In fact, μ is supported on U_1 and we have chosen U_1 so as to avoid Σ. The vector field υ was defined only at points on Σ.

This function, together with the choice of the rescaling function R and the lift function P as in the proof of Proposition 4.3.12, allows us to create a regular homotopy lifting the trajectory γ_1 and its neighbourhood along the trajectories of ∇F_0, to arrange that no point of γ_1 belongs to the same trajectory of ∇F_0 as a point on γ. We have the following properties.

- As we already mentioned, the procedure is needed only for $k = 2$.
- If N is presented as a disjoint sum of N_i with $G_0(N_i)$ pairwise disjoint, then γ_1 belongs to the same component as γ. While we might create some new self-intersections, the only new self-intersections we create are within $G(N_1)$.

If γ_1 terminates at q_+, then it cannot start at q_-. We push it down instead of up, by the same procedure. An inductive argument constructs the required regular homotopy. □

4.4.5. **From (LD-1) and (LD-2) to (LD-3).**

Lemma 4.4.9. *Suppose (F_0, G_0, η) satisfies (LD-1) and (LD-2). There exists a regular homotopy G_τ such that η is a gradient-like vector field for $F_0 \circ G_\tau$ for all τ and $G_1(\mathcal{K}_-)$ and such that $G_1(\mathcal{K}_+)$ belong to the first stratum of $G_1(N)$, i.e. such that (LD-3) is satisfied. The path (F_τ, G_τ) with $F_\tau = F_0$ is a regular double path.*

Moreover:

- *if $G_0(N)$ is an embedding, then $G_\tau = G_0$.*

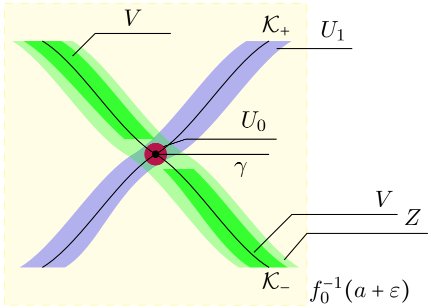

FIGURE 30. Notation of the proof of Lemma 4.4.9. The figure represents the level set $f_0^{-1}(a+\varepsilon)$.

- *if $G_0(N)$ is a disjoint union of $G_0(N_1),\dots,G_0(N_\ell)$ (with $N = N_1 \sqcup \dots \sqcup N_\ell$), then $G_\tau(N)$ is a disjoint union of $G_\tau(N_1),\dots,G_\tau(N_\ell)$ for all $\tau \in [0,1]$.*

Remark. The condition (LD-3) is stronger than (LD-2), that is G_1 satisfies (LD-2) as well. However, the procedure of arranging for (LD-3) might create extra self-intersections, hence we do not control critical points at deeper strata. That is, (F_1, G_1) need not satisfy (LD-1). These critical points will be dealt later, with (LD-4), which in turn implies (LD-1).

Proof. If $G_0(N)$ is an embedding then it automatically satisfies (LD-3).

The proof of Lemma 4.4.9 uses the whole machinery of the proof of Proposition 4.3.12. The presence of the trajectory γ prevents us from applying that proposition directly, because in the proof of Proposition 4.3.12, we were assuming that $\mathcal{K}_+$ and $\mathcal{K}_-$ are disjoint. However, thanks to (LD-2), we can apply these arguments with only minor modifications.

Choose $\varepsilon > 0$. As in the proof of Proposition 4.3.12, define Z by Condition 4.3.15. We note that $\mathcal{K}_- \cap f_0^{-1}(a+\varepsilon)$ belongs to Z. Recall that Z is an open subset of the level set $f_0^{-1}(a+\varepsilon)$.

Let $U_0 \subseteq f_0^{-1}(a+\varepsilon)$ be a neighbourhood of $\gamma \cap f_0^{-1}(a+\varepsilon)$ such that if $z \in U_0$, if λ_z is a trajectory of η through z (its part in $N' := G_0^{-1}(\Omega')$; see (4.3.4)), and if $w \in G_0(\lambda_z)$, then the trajectory of ∇F_0 through w does not hit $G_0(N)$ at any point in Ω' except at w itself. This is a condition analogous to Definition 4.4.7. Note that γ is in good position, so $\gamma \cap f_0^{-1}(a+\varepsilon)$ indeed belongs to U_0. In particular U_0 is not empty. The condition defining U_0 is open, so we may indeed choose U_0 to be open.

Let V be an open set of $f_0^{-1}(a+\varepsilon)$ such that $\overline{V} \subseteq Z$, V contains $(\mathcal{K}_- \cap f_0^{-1}(a+\varepsilon)) \setminus U_0$, and V is disjoint from some open subset U_1 containing $\mathcal{K}_+$. Notice that the only point on $f_0^{-1}(a+\varepsilon)$ that belongs to $\mathcal{K}_- \cap \mathcal{K}_+$ is the intersection point of γ with the level set (and that belongs to U_0). Therefore, such a V exists. Let Z_η and V_η be the sets of points in $f^{-1}[a+\varepsilon, b+\varepsilon]$ that are reached from Z, respectively V, by a trajectory of η; compare Figure 30.

To ensure property (LD-3) holds, we construct yet another cut-off function μ, with similar properties to the function with the same name from Section 4.3, but different from it. Let

$$\mu\colon f_0^{-1}[a+\varepsilon, b+\varepsilon] \to [0,1]$$

be a smooth function such that

- μ is supported on Z_η;
- $\mu \equiv 1$ on V_η;
- $\partial_\eta \mu \geq 0$ with equality only at points where μ is zero or one;
- $\partial_\upsilon \mu(u) \geq 0$ for all $u \in \Sigma \cap Z_\eta$ with equality only for when μ is zero or one, here υ is as in Subsection 4.3.5.

The construction of μ is analogous to the construction in Lemma 4.3.16. Using μ, and constructing an appropriate scaling function R and a lift function P as in Subsection 4.3.4, we define the isotopy G_τ, $\tau \in [1,2]$, that lifts all the points w with $f_1(w) \in [a+2\varepsilon, b]$ whose trajectory passes through V above the level set $F_0^{-1}(b)$. The construction of G_τ is via the same formula as in the proof of Proposition 4.3.12; see especially (4.3.8).

The definition of G_τ implies that the arguments of Lemma 4.3.17 apply to show that $F_0 \circ G_\tau$ is an η-path of functions. To show that G_1 maps $\mathcal{K}_-$ to the first stratum, we use arguments analogous to the proof of Lemma 4.3.18. Take $w \in \mathcal{K}_-$, and let $w' \in f_0^{-1}(a+\varepsilon)$ be the point on the same trajectory of η as $\mathcal{K}_-$. Consider the following two cases.

- If $w' \in V$, then
  - either $f_0(w) > a+\varepsilon$ and so $F_0 \circ G_1(w) > b + \varepsilon/2$, and we are done;
  - or $f_0(w) < a+\varepsilon$ and then $G_\tau(w)$ belongs to the first stratum by construction.
- If $w' \notin V$, then, as $w' \in \mathcal{K}_-$, we infer that $w' \in U_0$ (by construction, $V \cup U_0$ contains $\mathcal{K}_- \cap f_0^{-1}(a+\varepsilon)$). Then, by definition of U_0, the trajectory of ∇F_0 through $G_0(w)$ does not hit $G_0(N')$, so $G_\tau(w')$ belongs to the first stratum.

After this isotopy $\mathcal{K}_-$ is mapped to the first stratum.

An analogous operation improves G_0 in such a way that $\mathcal{K}_+$ is eventually mapped to the first stratum. The concatenation of the two paths (one making $\mathcal{K}_-$ map to the first stratum, the other making $\mathcal{K}_+$ map to the first stratum) is the desired path of immersions G_τ. We perturb it rel G_0 to be a regular F_0-path in the sense of Definition 3.1.20, so that (F_τ, G_τ), with $F_\tau = F_0$ is a regular double path.

To prove the second item of the lemma, we use the same method as in the proof of Lemma 4.4.5. The conditions on ∇F_0, guaranteed by Lemma 4.4.6, ensure that we do not create intersections between different components of N. □

4.4.6. **Condition (LD-4) and the vector field ξ.** Throughout Subsection 4.4.6 we assume that (F_0, G_0) satisfies (LD-3). Recall that the condition (LD-4) tells about the existence of a suitable vector field ξ and the lack of critical points in $F^{-1}[a-\varepsilon, b+\varepsilon]$ other than p_- and p_+. The next lemma constructs the vector field. That vector field will be used to move all the unnecessary critical points away from the interval $F^{-1}[a-\varepsilon, b+\varepsilon]$.

Lemma 4.4.10. *Suppose F is immersed Morse, $G\colon N \looparrowright \Omega$ is a generic immersion, $f = F \circ G$ and η is gradient-like for f. Assume G maps $\mathcal{K}_-$ and $\mathcal{K}_+$ to the first stratum, i.e. (LD-3) holds. There exists a grim vector field ξ for F such that the pull-back $\widetilde{\eta}$ of ξ to N agrees with η in an open subset of N containing $\mathcal{K}_- \cup \mathcal{K}_+$. If G is an embedding, we can guarantee that $\eta' = \eta$.*

As the pull-back $\widetilde{\eta}$ of ξ agrees with η in an open subset of $\mathcal{K}_- \cup \mathcal{K}_+$, Lemma 3.6.15 gives us a left-homotopy of the η-path and the $\widetilde{\eta}$-path of death.

Proof of Lemma 4.4.10. As (F,G) satisfies (LD-3), the map G takes a neighbourhood $U_1 \subseteq N$ of $\mathcal{K}_- \cup \mathcal{K}_+$ to the first stratum. Take $U_2 \subseteq N$ open such that $\mathcal{K}_- \cup \mathcal{K}_+ \subseteq U_2 \subseteq \overline{U}_2 \subseteq U_1$. By Proposition 2.5.3, applied to $U = G(U_1)$ there exists a grim vector field ξ for F, which is equal to $DG(\eta)$ on the whole of $G(U_2)$. We perturb ξ so that it is Morse–Smale.

By definition, the pull-back $\widetilde{\eta}$ of ξ agrees with η on U. □

To guarantee (LD-4), we need to rearrange critical points to move them away from between $F(p_-) = a$ and $F(p_+) = b$. This includes the critical points that have not been rearranged before and the critical points that might have been created while enforcing (LD-3).

Lemma 4.4.11. *Suppose (F,G) satisfies (LD-3) and ξ is as constructed in Lemma 4.4.10. Then, there exists a path of functions (F_τ, G_τ) such that F_τ is a ξ-path, $F_\tau \circ G_\tau$ is a path of excellent Morse functions supported away from $\mathcal{K}_- \cup \mathcal{K}_+$, $G_\tau = G$ (in particular, (F_τ, G_τ) is a regular double path) and moreover, F_1 has no critical points in $F_1^{-1}[a-\varepsilon, b+\varepsilon]$ other than $p-$ and p_+.*

Proof. We proceed in a standard way. That is, we take the first critical point above p_- with $h_p + d_p \leq h_- + 1$. We move it below p using Rearrangement Theorem 3.2.1 (compare the proof of Lemma 4.3.4). Inductively, we construct a special ξ-path that moves all critical points with $h_p + d_p \leq h_- + 1$ below p_-. We can do it in such a way that the path is supported away from $\mathcal{L}_-$. Next, the same argument allows us to move all critical points p with $h_p + d_p \geq h_+ + 1$ above the level set p_+. This done, since $h_+ = h_- + 1$, there are no more critical points between p_- and p_+. We let (F_τ, G_τ), for $\tau \in [0,1]$, denote the resulting path. This is a ξ-path, a concatenation of ξ-paths of rearrangments. Therefore $F_\tau \circ G_\tau$ is an $\widetilde{\eta}$-path. Notice that we do not rearrange pairs of critical points at depth 1, that is, $F_\tau \circ G_\tau$ is a path of excellent functions. □

Remark. Unlike in item (LD-1), our vector field ξ extends $\widetilde{\eta}$. That is, the condition that the membranes of points being rearranged miss $\mathcal{L}_-$ and $\mathcal{L}_+$ is included in the Morse–Smale condition, and need not be additionally guaranteed by the condition on a safe rearrangement.

For the reader's convenience we summarise the properties of the functions we have constructed.

Corollary 4.4.12. *The pair (F_1, G_1) that is the outcome of Lemma 4.4.11 satisfies conditions (LD-1), (LD-2), (LD-3), and (LD-4).*

Proof. The vector field ξ from Lemma 4.4.10 is grim for F_1, because Lemma 4.4.11 constructed F_1 as the end of a ξ-path. Next, Lemma 4.4.11 does not change the immersion G, in particular (LD-3) is preserved and (LD-4) is arranged to hold. Next, condition (LD-3) generalises (and implies) the condition (LD-2): the former one tells about the behaviour of G on $\mathcal{K}_- \cup \mathcal{K}_+$, the latter controls G on $\gamma = \mathcal{K}_- \cap \mathcal{K}_+$. Finally, condition (LD-4) controls all critical points between p_- and p_+, while condition (LD-1) controls only some of them. That is, (LD-4) implies (LD-1). □

4.4.7. **The Finger Move Theorem.** Before we complete the proof of Lifting Paths of Death Lemma 4.4.1, we need the most important ingredient, namely the Finger Move Theorem, which we now state. The proof is deferred to Part 5.

Theorem 4.4.13 (Finger Move)**.** *Let $G\colon N \looparrowright \Omega$ be a generic immersion with* $\dim N = n$ *and* $\dim \Omega = n+2$. *Let $F\colon \Omega \to \mathbb{R}$ be an immersed Morse function with respect to $M = G(N)$ and let ξ be a grim vector field for F satisfying the Morse–Smale condition. Set $f = F \circ G\colon N \to \mathbb{R}$ and let η be the pull-back of ξ as in Section 2.5.2, a gradient-like vector field on N. Assume that ξ and η are Morse–Smale.*

Suppose that there exist critical points $q_-, q_+ \in N$ of indices h and $h+1$ respectively, such that there is a single trajectory γ of η connecting q_- and q_+, and G maps the following subsets to the first stratum of $M = G(N)$, where $a := F(p_-)$, $b := F(p_+)$: (i) γ; (ii) the part of the stable manifold of q_+ in $F^{-1}(a,b)$; and (iii) the part of the unstable manifold of q_- in $F^{-1}(a,b)$. Suppose also there exists $\varepsilon > 0$ such that ξ has no critical points in the region $F^{-1}(a-\varepsilon, b+\varepsilon)$ other than p_- and p_+. Let $r > 0$ be the number of trajectories of ξ connecting $p_- = G(q_-)$ to $p_+ = G(q_+)$ that do not lie on $M = G(N)$.

Then there exists a path (F_τ, G_τ), for $\tau \in [0,2]$, and a grim vector field $\overline{\xi}$ for F_2, such that F_τ for $\tau \in [0,1]$ is an arbitrarily small perturbation of F_0, $F_{1+\tau} = F_1$ for $\tau \in [0,1]$, and $G_\tau = G_0$ for $\tau \in [0,1]$. The path (F_τ, G_τ) and $\overline{\xi}$ satisfy the following properties.

- *(FM-1) The Morse function F_2 has at most four more critical points than F_0 and these critical points belong to the second stratum of $G_2(N)$. Also they lie in the image of a single connected component of N under G_2.*
- *(FM-2) $F_2 = F_0$ near all critical points of F_0.*
- *(FM-3) The trajectories of $\overline{\xi}$ connecting p_- to p_+ are precisely the trajectories of ξ connecting p_- to p_+, except that one trajectory outside of $G_0(N)$ (and outside of $G_2(N)$) is removed.*
- *(FM-4) Set $f_\tau = F_\tau \circ G_\tau$. The path $f_{1+\tau}$ is a constant path for $\tau \in [0,1]$. Moreover, η is a gradient-like vector field for all f_τ.*
- *(FM-5) There is a pull-back $\overline{\eta}$ of $\overline{\xi}$ via G_2 such that the paths of death starting from $F_2 \circ G_2$ and constructed with η and $\overline{\eta}$ are left-homotopic.*
- *(FM-6) The map $G_\tau(N)$ fails to be a generic immersion only for one time value of τ (but still it is an immersion). The double points that are created are within the same connected component of $G_0(N)$.*

4.4.8. **Condition (LD-5).** The next result uses the finger move for its proof, which will be the topic of Part 5.

Lemma 4.4.14. *Suppose (F, G) satisfy (LD-1)–(LD-4). Then.*

- *If $k > 2$, then ξ automatically satisfies (LD-5);*
- *If $k = 2$, then there exists a regular homotopy G_τ and a family of functions F_τ, with (F_τ, G_τ) a regular double path, as well as a grim vector field $\widetilde{\xi}$ for the Morse function F_2 such that*
  - *ξ satisfies (LD-3), (LD-4), and (LD-5);*
  - *with $\widetilde{\eta}$ the pullback of $\widetilde{\xi}$, the paths of death starting from $F \circ G$ with η and the path of death starting from $F_2 \circ G_2$ with $\widetilde{\eta}$ are lax homotopic over $F_{2\tau} \circ G_{2\tau}$.*

Proof. Suppose first $k > 2$. Then, by Lemma 2.6.5 applied to $p_- = G_0(q_-)$ and $p_+ = G_0(q_+)$, it follows that $\dim \mathbb{M}_{\mathrm{a}}(p_-) \cap \mathbb{M}_{\mathrm{d}}(p_+) \cap \Omega[0] \cap F^{-1}(c) < 0$. In particular, the only trajectory of ξ connecting p_- to p_+ is the image of the trajectory γ of η connecting G_0. This means that (LD-5) is automatically satisfied.

From now on, suppose that $k = 2$. Lemma 2.6.5 implies that the number of trajectories of ξ from p_- to p_+ that lie in $\Omega \smallsetminus G_0(N)$ is finite. Let ℓ be this number. If $\ell = 0$, nothing needs to be done.

Assume $\ell > 0$. We first check that the properties (LD-1)—(LD-4) imply that the assumption of the Finger Move Theorem 4.4.13 are satisfied. The vector field ξ is Morse–Smale as well as its pull-back $\widetilde{\eta}$. The curve γ is mapped to the first stratum by (LD-2), and the parts of the stable/unstable manifolds are of q_- and q_+, that is, $\mathcal{K}_-$ and $\mathcal{K}_+$, are mapped to the first stratum by (LD-3). The assumption of the Finger Move Theorem 4.4.13 on the lack of critical points is guaranteed by (LD-4).

Apply the finger move. That is, by Theorem 4.4.13, we can create a path (F_τ, G_τ), $\tau \in [0,2]$, and a new grim vector field $\overline{\xi}$ for F_2. The path has all the properties (FM-1) – (FM-6). By Lemma 5.5.3, η is gradient-like for $F_2 \circ G_2$. By Lemma 3.6.15 with $h_{\sigma,0} = F_{2\tau} \circ G_{2\tau}$, the η-paths of death constructed from $F_2 \circ G_2$ and from $F \circ G$ are lax homotopic over $F_{2\tau} \circ G_{2\tau}$. By (FM-5), the pull-back $\overline{\eta}$ of $\overline{\xi}$ via G_2 has the property that the $\overline{\eta}$-path of death starting from $F_2 \circ G_2$ and the η-path of death starting from $F_2 \circ G_2$ are left-homotopic. That is, to say, the $\overline{\eta}$ path of death starting from $F_2 \circ G_2$ and the η-path of death starting from $F \circ G$ are lax homotopic over $F_{2\tau} \circ G_{2\tau}$. Moreover, by (FM-3), $\overline{\xi}$ has $\ell - 1$ trajectories outside of $G_2(N)$ connecting p_- with p_+. All this means that (F, G) satisfies condition (LD-3).

Now, by (FM-1), F_2 acquires four critical points on the second stratum. We perform rearrangements, using $\overline{\xi}$ as the guiding vector field, to move the four new critical points away from the region $F_2^{-1}[a, b]$. The rearrangement is done as in Lemma 4.4.11. Call (F_3, G_3) the resulting function. As the rearrangement is done using the vector field $\overline{\xi}$, and $\overline{\eta}$ is a pull-back, the path $F_{2+\tau} \circ G_{2+\tau}$ is an $\overline{\eta}$-path of excellent Morse functions. In particular, by Lemma 3.6.15, the $\overline{\eta}$-paths of death starting with $F_3 \circ G_3$ and starting with $F_2 \circ G_2$ are lax homotopic over $F_{2+\tau} \circ G_{2+\tau}$. That is to say, the $\overline{\eta}$-path of death starting from $F_3 \circ G_3$ is lax homotopic over $F_{3\tau} \circ G_{3\tau}$ to the original path of death. Note that the pair (F_3, G_3) satisfies:

- the condition (LD-3), with $\overline{\xi}$ being the corresponding grim vector field. This is because (F_2, G_2) already satisfied that condition;
- the condition (LD-4), because we just move all critical points that were created by the finger move away from $F_2^{-1}[a, b]$;
- the conditions (LD-1) and (LD-2) as consequences of conditions (LD-3) and (LD-4).

This places us in a position to use induction. Starting with (F_3, G_3), we create further paths by a subsequent use of the finger move. More precisely, the j-th path denoted by $(F_{3j-3+\tau}, G_{3j-3+\tau})$, $\tau \in [0, 3]$, is constructed using the finger move applied to (F_{3j-3}, G_{3j-3}) and the vector field $\widetilde{\xi}_j$, grim for F_{3j-3}, with $\widetilde{\xi}_0 = \xi$, $\widetilde{\xi}_1 = \overline{\xi}$. The vector field $\widetilde{\xi}_j$ has $\ell - j$ trajectories connecting p_- with p_+. Moreover, the induction argument shows that the grim path of death starting from $F_{3j-3} \circ G_{3j-3}$, and using the pull-back of $\widetilde{\xi}_j$, is lax homotopic to the original one over $F_{(3j-3)\tau} \circ G_{(3j-3)\tau}$.

Each of the vector fields $\widetilde{\xi}_j$ satisfies (LD-4). Eventually, $(F_{3\ell}, G_{3\ell})$ and $\widetilde{\xi}_\ell$ satisfy (LD-3), (LD-4), and (LD-5). We conclude the proof by reparametrizing the path to the interval $[0, 1]$. □

4.4.9. **The proof of the Death Lifting Lemma 4.4.1.** Let $F_0 = F$ and $G_0 = G$. We set $f_0 = F_0 \circ G_0$, η a gradient-like vector field for f_0, and f_τ an η-path of death cancelling q_- and q_+. The proof follows the pattern of the proof of the Rearrangement Lifting Lemma 4.3.1, but is technically more involved. The reader should keep in mind that throughout the proof a few new vector fields are constructed, like $\widetilde{\eta}$, η', and $\overline{\eta}$. The key property is that except for $\overline{\eta}$, all these vector fields agree with the original vector field η near $\mathcal{K}_- \cup \mathcal{K}_+$, so that all paths of death created with these vector fields are lax-homotopic to the original one. The case of $\overline{\eta}$ is a bit special, as it arises from the finger move, but its behaviour is controlled by (FM-5) of Finger Move Theorem 4.4.13. A schematic of the proof is presented in Figure 31. Lemma 4.4.4 takes care of (LD-1). It creates a path (F_τ, G_τ) that replaces (F_0, G_0) by a (F_1, G_1) (with $G_\tau = G_0$) and (F_1, G_1) satisfying (LD-1). Set $\widetilde{f}_\tau = F_\tau \circ G_\tau$. Clearly η need not be gradient-like for $\widetilde{f}_1$. However, the arguments of Lemma 4.3.34 from Subsection 4.3.8 allow us to construct a vector field $\widetilde{\eta}$, gradient like for $\widetilde{f}_\tau$ and such that $\widetilde{\eta} = \eta$ near $\mathcal{K}_- \cup \mathcal{K}_+$.

If G_1 is an embedding (that is if G_0 is), (LD-3) is already satisfied. Otherwise, we create a path whose destination satisfies that condition. We choose a generic gradient-like vector field ∇F_0 and, if necessary, replace G_1 by a path of immersions (denoted G_τ in Subsection 4.4.4, but here we call it $G_{1+\tau}$, for $\tau \in [0, 1]$, to avoid a notation clash), with $F_{1+\tau} \equiv F_1$ such that the trajectory γ is in good position with respect to G_2 and $F_{1+\tau} \circ G_{1+\tau}$ is an $\widetilde{\eta}$-path of Morse functions. This is done in Lemma 4.4.8.

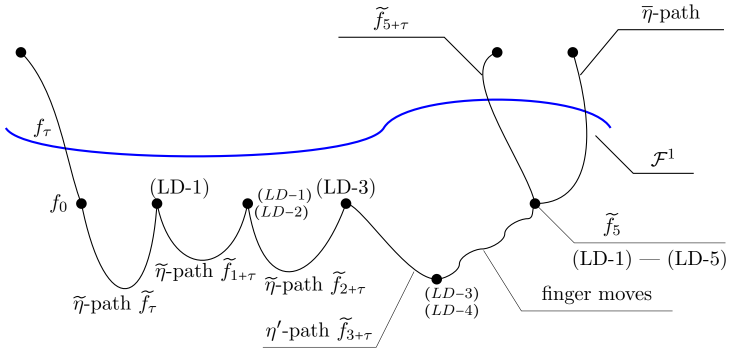

FIGURE 31. A schematic of the proof of Death Lifting Lemma 4.4.1.

With this choice, (F_2, G_2) satisfies (LD-2). However, (F_2, G_2) might acquire some critical points on deeper strata, as the path $(F_{1+\tau}, G_{1+\tau})$ might potentially create extra self-intersections. The key point is that no critical points at depth 0 are created, because $F_{1+\tau}$ is constant. That is, (F_2, G_2) again satisfies (LD-1).

Once (LD-1) and (LD-2) are satisfied, Lemma 4.4.9 constructs a path of functions, which we here denote $(F_{2+\tau}, G_{2+\tau})$, such that $F_{2+\tau} \circ G_{2+\tau}$ is a $\widetilde{\eta}$-path of Morse functions. Moreover (F_3, G_3) satisfies (LD-3).

Lemma 4.4.10 applied to $(F_3, G_3, \widetilde{\eta})$ constructs a vector field ξ, grim for F_3, such that the pull-back, η' agrees with $\widetilde{\eta}$ near $\mathcal{K}_- \cup \mathcal{K}_+$. Note that η' agrees also with the original vector field η in an open subset containing $\mathcal{K}_- \cup \mathcal{K}_+$. In Lemma 4.4.11, we construct a path, which we now denote $(F_{3+\tau}, G_{3+\tau})$, that moves all critical points away from between p_- and p_+. This is a path of rearrangement, such that $F_{3+\tau} \circ G_{3+\tau}$ is an η' path.

The resulting functions (F_4, G_4) satisfy conditions (LD-3) and (LD-4), so they also satisfy (LD-1) (which is weaker than (LD-4)) and (LD-2) (which is weaker than (LD-3)).

We continue with the proof of Death Lifting Lemma 4.4.1 by studying ξ in detail. The vector field ξ might have more trajectories connecting p_- and p_+. We apply finger moves through Lemma 4.4.14 applied to $(F, G) = (F_4, G_4)$, the vector field ξ and its pull-back η'. We obtain a path, which we denote $(F_{4+\tau}, G_{4+\tau})$, $\tau \in [0,1]$, and a vector field $\overline{\xi}$, grim for F_5.

Lemma 4.4.14 implies that $(F_5, G_5, \overline{\xi})$ satisfies (LD-4) and (LD-5), as well as all the previous conditions. Lemma 4.4.3 applied to that system of functions constructs a weak lift of the $\overline{\eta}$ path of death starting from $F_5 \circ G_5$. Denote it by $(F_{5+\tau}, G_{5+\tau})$, $\tau \in [0,1]$. Set $\widetilde{f}_{5+\tau} := F_{5+\tau} \circ G_{5+\tau}$.

Lemma 4.4.15. *The path $\widetilde{f}_{5+\tau}$ is lax-homotopic to the path f_τ over $\widetilde{f}_\tau$, $\tau \in [0,5]$.*

Proof. We make this proof in several steps.

- An $\widetilde{\eta}$-path of death starting from $\widetilde{f}_1$ is lax homotopic to an η-path of death starting from $f_0 = \widetilde{f}_0$ over $\widetilde{f}_\tau$ (which is f_τ) by Lemma 3.6.14.
- An $\widetilde{\eta}$-path of death starting from $\widetilde{f}_2$ is lax homotopic to an $\widetilde{\eta}$-path of death starting from $\widetilde{f}_1$ over $\widetilde{f}_{1+\tau}$ by Lemma 3.6.15.
- An $\widetilde{\eta}$-path of death starting from $\widetilde{f}_3$ is lax homotopic to an $\widetilde{\eta}$-path of death starting from $\widetilde{f}_2$ over $\widetilde{f}_{2+\tau}$ by Lemma 3.6.15.
- An η'-path of death starting from $\widetilde{f}_4$ is lax homotopic to an $\widetilde{\eta}$-path of death starting from $\widetilde{f}_3$ over $\widetilde{f}_{3+\tau}$. Here we use Lemma 3.6.14, because we change the vector field, however, we do not change it near $\mathcal{K}_- \cup \mathcal{K}_+$.

- An $\overline{\eta}$-path of death starting from $\widetilde{f}_5$ is lax homotopic to an η'-path of death starting from $\widetilde{f}_4$ over $\widetilde{f}_{4+\tau}$. This is a consequence of Lemma 4.4.14, which in turn relies on the property (FM-5) of the Finger Move Theorem 4.4.13.
- By construction $\widetilde{f}_{5+\tau}$ is a weak lift of an $\overline{\eta}$-path of death starting from $\widetilde{f}_5$. As such, $\widetilde{f}_{5+\tau}$ is left-homotopic to an $\overline{\eta}$-path of death starting from $\widetilde{f}_5$.

We now concatenate the homotopies. Eventually we get that $\widetilde{f}_{5+\tau}$ is lax homotopic to f_τ over $\widetilde{f}_{5\tau}$. □

Lemma 4.4.15 will put us in a position to use Lemma 4.1.6. The following result is needed to check all the assumptions.

Lemma 4.4.16. *The path $(F_{6\tau}, G_{6\tau})$ is a regular double path.*

Proof. The proof requires a case-by-case analysis of the components of the path.

- The path (F_τ, G_τ) changes only F_τ, fixing $G_\tau = G_0$ being a regular immersion. Moreover F_τ is an $\mathcal{F}^1$-path, compare Lemma 4.4.4, so (F_τ, G_τ) is a regular double path.
- The path $(F_{1+\tau}, G_{1+\tau})$ created in Lemma 4.4.8 is such that $F_{1+\tau}$ is fixed, and $G_{1+\tau}$ is a path of immersions (an F_1-regular path in the sense of Definition 3.1.20), in particular $(F_{1+\tau}, G_{1+\tau})$ is a regular double path.
- Likewise, Lemma 4.4.9 creates a regular double path $(F_{2+\tau}, G_{2+\tau})$.
- The path $(F_{3+\tau}, G_{3+\tau})$ is a path of rearrangments with $G_{3+\tau} = G_3$, see Lemma 4.4.11. That is, it is a regular double path.
- Lemma 4.4.14 creating the path $(F_{4+\tau}, G_{4+\tau})$ creates a regular double path.
- The weak lift $(F_{5+\tau}, G_{5+\tau})$ is created using Lemma 4.4.3 creating a regular double path, actually, that lemma does not change $G_{5+\tau}$, only $F_{5+\tau}$.

Concatenation of regular double paths is a regular double path. □

Give Lemmata 4.4.15 and 4.4.16, we promote $(F_{6\tau}, G_{6\tau})$ to a weak lift of f_τ using Lemma 4.1.6. This completes the proof of the Death Lifting Lemma 4.4.1.

4.5. The Path Lifting Theorem

We can now combine the results from Section 4.1 on lifting $\mathcal{F}^0$-paths and paths of birth, rearrangement and death, to obtain the following result.

Theorem 4.5.1 (Path Lifting)**.** *Let N be a closed manifold of dimension n and let f_τ, for $\tau \in [0,1]$, be an $\mathcal{F}^1$-path of functions such that there are no rearrangements for which a critical point of higher index goes below a critical point of a lower index. Let $G_0\colon N \to \Omega$ be a generic immersion, where Ω is a compact manifold of dimension $n+k$. Suppose $F_0\colon \Omega \to \mathbb{R}$ is an immersed Morse function such that $F_0 \circ G_0 = f_0$.*

If $k \geq 2$, then there exists a weak lift (Definition 4.1.1) of f_τ. That is, there exist:

- *a regular path (an $\mathcal{F}^1$-path) of functions $\widetilde{f}_\tau\colon N \to \mathbb{R}$ such that $\widetilde{f}_0 = f_0$ and $\widetilde{f}_1 = f_1$ and the paths $\widetilde{f}_\tau$ and f_τ are $\mathcal{F}^1$-homotopic;*
- *a regular double path (F_τ, G_τ) such that $F_\tau \circ G_\tau = \widetilde{f}_\tau$ with the following properties for G_τ.*
  - *(i) If connected components of N have disjoint image under G, then this also holds for G_τ for each $\tau \in [0,1]$.*
  - *(ii) if G_0 is an embedding and either $k \geq 3$ or f_τ has no deaths, then G_τ is an ambient isotopy.*

- *The family F_τ has no births or deaths on the zeroth stratum, that is to say, for each $\tau \in [0,1]$, F_τ is a classical Morse function when regarded as a function from Ω to $\mathbb{R}$, forgetting the stratification.*

If N and Ω are compact with nonempty boundary, G_0 is a neat immersion and f_τ is neat as a path of functions on N, while F_0 is a neat function, then there is a neat regular double path (F_τ, G_τ) satisfying all the required properties.

Proof. Let T be the number of 'events' on f_τ, where by an event we mean a birth, a rearrangement, or a death. As f_τ is an $\mathcal{F}^1$-path we make sure that at each of these events, the path f_τ is an elementary path, respectively of birth, of rearrangement, or of death. We may and shall assume that these events occur at distinct times.

We perform induction over T. If $T = 0$, then f_τ is a path of excellent Morse functions. We use Lemma 4.1.2.

Otherwise suppose that at τ_0 there appears one of the events: birth, death or rearrangement, and that $\tau_0 > 0$ is minimal with this property. Let $\delta > 0$ be such that the path f_τ, $\tau \in [\tau_0 - \delta, \tau_0 + \delta]$ is an elementary path. We split our path f_τ into three parts: the first part, for $\tau \in [0, \tau_0 - \delta]$, is a $\mathcal{F}^1$-path of Morse functions, which can be lifted using Lemma 4.1.2. For $\tau \in [\tau_0 - \delta, \tau_0 + \delta]$ we lift the path using Lemma 4.2.1 (if there is a birth at τ_0), Lemma 4.3.1 (if there is a rearrangement at τ_0), or Lemma 4.4.1 (if there is a death at τ_0). Note that if there is a rearrangement at τ_0, a critical point with a smaller index can never go above a critical point of higher index by the assumptions of the theorem. Therefore, the assumptions of Lemma 4.3.1 are satisfied.

The path for $\tau \in [\tau_0 + \delta, 1]$ has $T - 1$ events and we use the inductive hypothesis.

Now we provide an argument that the result works for N and Ω with nonempty boundary. First, we perturb f_τ to a $\mathcal{F}^1$-path that is neat. This is done using methods from Subsection 3.1.6. Therefore, we only need to prove that we can lift a neat path of Morse functions to a neat path, and that we can lift an elementary path of neat functions to a neat path on Ω.

The first step, that is, lifting a neat path of Morse functions to a neat path, is done by the last part of Lemma 4.1.2. Next, we note that each kind of elementary path (births, rearrangements and deaths) is lifted in such a way that the initial Morse function F_0 and the initial embedding G_0 are unchanged away from a set of the form $F_0^{-1}(c,d)$: for example, for rearrangements, if a and b are critical values of points that are going to be rearranged, we set $c = a - \varepsilon$, $d = b + \varepsilon$ for $\varepsilon \ll 1$. This means that each time we lift an elementary path, we do not alter the function F_0 near the boundary of Ω, nor do we alter the function G_0 near the boundary of N. Each time we lift an elementary path, we might need to shrink the neighbourhoods of $\operatorname{bd} N$ and $\operatorname{bd}\Omega$, on which G_τ, respectively F_τ are independent of τ, but we perform only finitely many such operations. Hence, the resulting regular double path (F_τ, G_τ) is neat. □

4.6. Path lifting for N having components of varying dimensions

Now we discuss the case that $N = N_1 \sqcup \cdots \sqcup N_r$ is a union of components of dimensions $n_1, \dots, n_r$, with $n_i \le \dim \Omega - 2$ for all i. We suppose that $G_0(N_i) \cap G_0(N_j) = \varnothing$ unless $i = j$. Set $k_i = \dim \Omega - \dim N_i$ to be the codimension of N_i.

First, the notions of immersed Morse functions and of grim vector fields can be generalised in an obvious way. With the stratification $\Omega[d]$ given by points at depth d (even if the connected components of $\Omega[d]$ might have different codimensions), the local form of a grim vector field is still given by (2.5.1), with the necessary adjustments of the range of indices of the x and y variables. The ascending and descending membranes of a critical point $p \in G_0(N_i)$ have dimensions governed by the n_i, the depth of p, and the codimension k_i. Existence of Morse functions and of grim vector fields is proved in this situation without changes.

In the Cancellation Theorem 3.4.1 we work with the image of a single connected component of N. Indeed, the two critical points that are to be cancelled are connected by a trajectory of the grim vector field that lies in the first stratum. That is, everything happens within a single connected component of N. The dimension restriction is not used.

The situation is different with the rearrangement theorem, where we might rearrange critical points belonging to different components. The key property we use is dimension counting. It is routine to check that if $k_i \geq 2$ for all i, then critical points p_- and p_+ with $\operatorname{ind} p_- \geq \operatorname{ind} p_+$ can be rearranged regardless of their depth and of the dimension of the manifolds they belong to. One way to see this (and to avoid lengthy calculations) is that rearrangement is possible in codimension two, and the dimension of membranes at the same depth does not increase if the codimension k_i increases. Another way is to check that if $p_- \in N_i$ and $p_+ \in N_j$, $i \neq j$, then the only trajectories connecting p_- to p_+ are necessarily in the zeroth stratum, since $G_0(N_i) \cap G_0(N_j) = \emptyset$. The calculation of dimensions of the membranes is straightforward.

We pass to a discussion of path lifting. Lifting of births is purely local. In general, we lift a path governed by a vector field η on N (an η-path of rearrangment or an η-path of death), this terminology is valid even if the components of N have different dimensions. Going into technical details, the partial rearrangement, i.e. ensuring (LR-1) and (LD-1), relies on dimension counting arguments. These arguments go through if the components have various dimensions (the same argument as in the previous paragraph applies). The most difficult case is still if p_- and p_+ both belong to a single codimension two component.

Arranging for properties (LR-2) and (LD-3) to hold is one of the most sophisticated technical part of the proof of the Path Lifting Theorem. Recall that we use an extra gradient-like vector field, which we call ∇F_0, and a special cut-off function μ supported in a neighbourhood of the ascending manifold $\mathcal{K}_-$ of q_-. The function μ is nonzero only on the component of N containing q_-, that is, the whole discussion constructing μ is independent of the dimensions of the components of N not containing q_-. The vector field ∇F_0 is constructed so as to avoid self-intersections of different components of $G_0(N)$. This holds even if the N_i have different dimensions.

Finally, we need to discuss the finger move. It is used if the critical points q_- and q_+ belonging to the *same* component of N are connected by a trajectory of a vector field ξ that lies in the zero stratum. As q_- and q_+ belong to the same component, and the component has codimension two, the finger move concerns only that single connected component of N. Other components do not interfere, regardless of their dimension.

Part 5. The finger move

Throughout Part 5 we assume that the codimension $k = 2$. Our work is focused on the image of N. Therefore, unlike in Part 4, we will mostly work on $M := G_0(N)$.

Plan of proof of Theorem 4.4.13. The proof of Theorem 4.4.13 requires Sections 5.1 through 5.5. As a guide for the reader through the proof, we present the plan.

The first challenge is to find a suitable coordinate system. The construction of this coordinate system is done in Sections 5.1 and 5.2 and consists of several steps. In fact, the coordinate system we define is not merely a local coordinate system near a critical point, but it is suitably extended along a "guiding curve." In a sense, this a semi-local coordinate system.

The actual finger move is described in Section 5.3. This section begins with an explicit formula for the finger move in Subsection 5.3.1. Formally, this induces an isotopy $G_{1+\tau}$ of Theorem 4.4.13. One of the key properties is that $G_{1+\tau}$ is level-set-preserving, that is $F \circ G_{1+\tau} = F \circ G_1 = F \circ G_0$ for all $\tau \in [0,1]$. Condition (FM-6) follows from the construction of G_τ, and it is addressed at the end of Subsection 5.3.1.

In Subsection 5.3.2 we construct a suitable vector field $\widehat{\xi}$, which is tangent to $G_2(M)$. In Subsection 5.3.3 we prove that $\partial_{\widehat{\xi}} F \geq 0$. Already the vector field $\widehat{\xi}$ satisfies the most important property of Theorem 4.4.13, namely (FM-3). However, $\widehat{\xi}$ is not a grim vector field on F; moreover, F is not an immersed Morse function for $G_2(N)$. To remedy this, in Subsection 5.3.4 we construct an explicit perturbation F_τ, $\tau \in [0,1]$, of F and an explicit perturbation of the vector field ξ_τ and of the function F_τ.

The key property (FM-3) is proved in Section 5.4 as Theorem 5.4.1 with two key lemmas (Lemma 5.2.3 and 5.2.5) proved in Section 5.2. Since F_1 is a perturbation of F, the proof of Theorem 5.4.1 holds both for F and $\widehat{\xi}$ and for F_1 and $\overline{\xi}$.

Finally, we study the pull-back of $\overline{\xi}$, denoted by $\overline{\eta}$, which is a vector field on N, gradient-like for $f_2 = F_2 \circ G_2$. Properties (FM-4) and (FM-5) are proved in Section 5.5. By a careful analysis of the ascending and descending manifolds of $\overline{\eta}$, in the same section we show that the grim paths of death constructed via $\overline{\eta}$ and via η are homotopic. □

5.1. The outer shell

As noted above, the proof of Finger Move Theorem 4.4.13 begins with a construction of a suitable open subset with concrete coordinates, in which the finger move will be performed. In fact, we will construct three nested neighbourhoods of one of the critical points.

Subsections 5.1.1 through 5.1.5 construct the outer shell, V_{out}. The construction is rather technical. For the reader who wants to skip the details, we summarise the properties of V_{out} in Subsection 5.1.6.

5.1.1. **The set V_{out}.** In this section we construct the biggest (outer) neighbourhood of one of the critical points, which we will denote V_{out}.

We set

$$c_- = F(p_-),\ \ c_+ = F(p_+)$$

and we denote by Ξ_s the flow of the vector field ξ.

Suppose that there are no other critical points in $F^{-1}([c_-, c_+])$. Note that this can always be arranged by the Global Rearrangement Theorem 3.2.4. Choose local coordinates around p_+ given by $(x_1, \dots, x_n, y_1, y_2)$, so that p_+ is given by $(0, \dots, 0)$, M is locally given by $\{y_1 = y_2 = 0\}$, and the grim vector field around p_+ is given by $(-x_1, \dots, -x_{h+1}, x_{h+2}, \dots, x_n, y_1^2 + y_2^2, 0)$. Let U_{coor} be an open set containing p_+ on which these coordinates are defined. In these local coordinates $\mathbb{M}_{\text{d}}(p_+)$ is given by

$$\mathbb{M}_{\text{d}}(p_+) = \{x_{h+2} = \dots = x_n = y_2 = 0\} \cap \{y_1 \leq 0\}.$$

Choose $c_{\text{bot}} \in (c_-, c_+)$ such that for any $c \in [c_{\text{bot}}, c_+]$, the intersection of the descending membrane $\mathbb{M}_{\text{d}}(p_+)$ with the level set $F^{-1}(c)$ is contained in the coordinate neighbourhood U_{coor}. Likewise, choose a regular value $c_{\text{top}} > c_+$ in such a way that $\mathbb{M}_{\text{a}}(p_+) \cap F^{-1}(c) \subseteq U_{\text{coor}}$ for any $c \in [c_+, c_{\text{top}}]$; see Figure 32. It is clear that such c_{bot} and c_{top} exist. The set U_{out} will be contained inside $F^{-1}[c_{\text{bot}}, c_{\text{top}}]$. To define it choose $\varepsilon_{\text{out}} > 0$ and let Y_{out} be an open

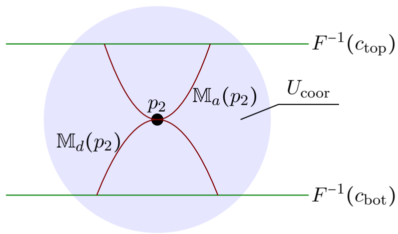

FIGURE 32. The set U_{coor} and the level sets c_{bot}, c_{top}.

tubular neighbourhood of $\mathbb{M}_{\mathrm{d}}(p_+) \cap F^{-1}(c_{\mathrm{bot}})$ in the level set $F^{-1}(c_{\mathrm{bot}})$ consisting of points at distance less than $\varepsilon_{\mathrm{out}}$ from $\mathbb{M}_{\mathrm{d}}(p_+) \cap F^{-1}(c_{\mathrm{bot}})$. We will assume that $\varepsilon_{\mathrm{out}} > 0$ is small, and the precise meaning of 'small' in this context will be clarified soon. For now define U_{out} as

$$\overline{U}_{\mathrm{out}} := \overline{\bigcup_{s \in \mathbb{R}} \Xi_s(Y_{\mathrm{out}})} \cap F^{-1}[c_{\mathrm{bot}}, c_{\mathrm{top}}]. \quad U_{\mathrm{out}} = \operatorname{Int} \overline{U}_{\mathrm{out}}. \tag{5.1.1}$$

That is, U_{out} is the grim neighbourhood of p_+ in the sense of Definition 2.7.7. Clearly, $\overline{U}_{\mathrm{out}}$ is the closure of U_{out}.

Lemma 5.1.1. *If Y_{out} is sufficiently thin (that is, if $\varepsilon_{\mathrm{out}}$ is sufficiently small), then U_{out} is contained in the coordinate neighbourhood U_{coor}.*

Proof. Let Y_n be the set Y_{out} defined with $\varepsilon_{\mathrm{out}} = \frac{1}{n}$ and let U_n be the corresponding set U_{out} (that is, U_n is constructed via (5.1.1) by taking $Y_{\mathrm{out}} = Y_n$). By construction, $\overline{U}_{n+1} \cap F^{-1}(c_{\mathrm{bot}}, c_{\mathrm{top}}) \subseteq U_n$. It follows from the proof of Lemma 2.7.9 that $\bigcap U_n = (\mathbb{M}_{\mathrm{a}}(p_+) \cup \mathbb{M}_{\mathrm{d}}(p_+)) \cap F^{-1}(c_{\mathrm{bot}}, c_{\mathrm{top}})$, so that $\bigcap U_n \subseteq U_{\mathrm{coor}}$. Since U_{coor} is open, it follows that for large n we must have $U_n \subseteq U_{\mathrm{coor}}$ as well. ☐

We have constructed the outer neighbourhood U_{out} starting from the neighbourhood Y_{out}. We leave ourselves the possibility of shrinking Y_{out} further by decreasing $\varepsilon_{\mathrm{out}}$, but henceforth the levels c_{bot} and c_{top} will remain fixed.

5.1.2. **The position of $\mathbb{M}_{\mathrm{a}}(p_-)$.** By the Morse–Smale condition, the membranes $\mathbb{M}_{\mathrm{a}}(p_-)$ and $\mathbb{M}_{\mathrm{d}}(p_+)$ intersect transversely along a finite number of trajectories. Intersecting $\mathbb{M}_{\mathrm{a}}(p_-) \cap \mathbb{M}_{\mathrm{d}}(p_+)$ with the level set $F^{-1}(c_{\mathrm{bot}})$ yields a finite number of points $z_0, z_1, \dots, z_r$. Here z_0 is assumed to lie on M, that is z_0 corresponds to the trajectory of ξ on M that connects p_- and p_+.

If Y_{out} is small enough, the intersection $\mathbb{M}_{\mathrm{a}}(p_-) \cap Y_{\mathrm{out}}$ consists of $r+1$ discs of dimension $n-h$. In fact, each connected component of $\mathbb{M}_{\mathrm{a}}(p_-) \cap Y_{\mathrm{out}}$ that does not contain $\mathbb{M}_{\mathrm{d}}(p_+) \cap Y_{\mathrm{out}}$ is separated from it by some distance, and thus shrinking Y_{out} (decreasing $\varepsilon_{\mathrm{out}}$) we can make this component disjoint from U_{out}; see Figure 33. The $r+1$ discs $D_0, D_1, \dots, D_r$ forming $\mathbb{M}_{\mathrm{a}}(p_-) \cap F^{-1}(c_{\mathrm{bot}}) \cap U_{\mathrm{out}}$ are transverse to $\mathbb{M}_{\mathrm{d}}(p_+)$, because $\mathbb{M}_{\mathrm{a}}(p_-)$ intersects $\mathbb{M}_{\mathrm{d}}(p_+)$ transversely (this is the Morse–Smale condition) and they intersect $\mathbb{M}_{\mathrm{d}}(p_+)$ at the points $z_0, z_1, \dots, z_r$. These discs can be isotoped in such a way that for $k > 0$ the disc D_k around z_k has the form

$$x_1 = s_{1,k}, \dots, x_{h+1} = s_{h+1,k} \text{ for } k = 1, \dots, r, \tag{5.1.2}$$

for some constants $s_{i,\ell} \in \mathbb{R}$. For one of the points z_m, after applying the aforementioned isotopy, we can and will assume that all of the constants $s_{i,m} = 0$, for $i = 1, \dots, h+1$ and that $s_{1,j} \neq 0$ for all $j \neq m$. The choice of m is arbitrary if $\dim \mathbb{M}_{\mathrm{a}}(p_-) \cap F^{-1}(c_{\mathrm{bot}}) > 1$ (that is, if

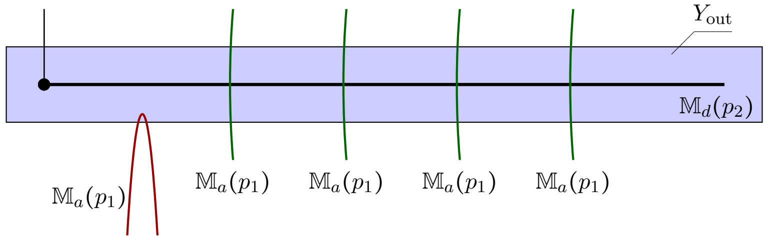

FIGURE 33. The membranes $\mathbb{M}_{\text{a}}(p_-)$ and $\mathbb{M}_{\text{d}}(p_+)$ at the level set c_{bot}. The set Y_{out} is shrunk to avoid the component on the left.

$h < n-1$); if $\dim \mathbb{M}_{\text{a}}(p_-) \cap F^{-1}(c_{\text{bot}})$ is 1-dimensional, then we take m to be the index of the point *nearest* to M on the guiding curve, see Subsection 5.1.2 below.

Given (5.1.2), the coordinates of the point z_ℓ are

$$(s_{1,\ell}, \dots, s_{h+1,\ell}, 0, \dots, 0, c_{\text{bot}} - s_{1,\ell}^2 - \dots - s_{h+1,\ell}^2, 0). \tag{5.1.3}$$

We remark that the y_1 coordinate of z_ℓ is determined by the fact that $F(z_\ell) = c_{\text{bot}}$. We also isotope the disc D_0 within M in such a way that it is given by (5.1.2), but with the extra conditions that the y_1 coordinate of z_0 is equal to 0, and the y_2 coordinate on D_0 is either non-positive or non-negative.

According to the Isotopy Insertion Lemma 2.5.8, the isotopy of the discs required to arrange the local description above of the intersection of $\mathbb{M}_{\text{a}}(p_-)$ with $\mathbb{M}_{\text{d}}(p_+)$, can be obtained by locally altering the vector field ξ in the set $F^{-1}(c', c_{\text{bot}})$, for some $c' < c_{\text{bot}}$ close to c_{bot}. From now on we will assume that such an alteration has been made, and that the intersection $\mathbb{M}_{\text{a}}(p_-) \cap Y_{\text{out}}$ is given by (5.1.2).

5.1.3. **The guiding curve.** Consider the point z_m on the intersection $\mathbb{M}_{\text{a}}(p_-) \cap \mathbb{M}_{\text{d}}(p_+) \cap F^{-1}(c_{\text{bot}})$, for which we have set $s_{1,m} = \dots = s_{h+1,m} = 0$. Take an embedded curve $\varpi\colon [0, \varepsilon_{\text{out}}] \to \mathbb{M}_{\text{a}}(p_-) \cap F^{-1}(c_{\text{bot}})$ given by

$$\varpi(s) = (0, \dots, 0, 0, \dots, 0, 0, s); \tag{5.1.4}$$

see Figure 34, in which $m = 4$. Since $x_1, \dots, x_n, y_1, y_2$ is an orthonormal coordinate system and Y_{out} is the set of points at distance less than or equal to ε_{out} from $\mathbb{M}_{\text{d}}(p_+) \cap F^{-1}(c_{\text{bot}})$, we have that $\varpi(\varepsilon_{\text{out}}) \in \partial Y_{\text{out}}$. Next, we want to extend ϖ to a map $\varpi\colon [0,1] \to \mathbb{M}_{\text{a}}(p_-) \cap F^{-1}(c_{\text{bot}})$ such that

$$\begin{aligned} \varpi((\varepsilon_{\text{out}}, 1]) \cap Y_{\text{out}} &= \emptyset, \\ \varpi(1) &\in M \cap F^{-1}(c_{\text{bot}}), \text{ and} \\ \varpi(1) &\notin \mathbb{M}_{\text{d}}(p_+). \end{aligned}$$

That is, ϖ is a path in $\mathbb{M}_{\text{a}}(p_-)$ and in the level set $F^{-1}(c_{\text{bot}})$, from an intersection point in $\mathbb{M}_{\text{a}}(p_-) \cap \mathbb{M}_{\text{d}}(p_+)$, to a point on M in the boundary of $\mathbb{M}_{\text{a}}(p_-) \cap F^{-1}(c_{\text{bot}})$; see Figure 35. The curve ϖ will correspond to the guiding curve in Section 6.3.

The construction of ϖ is as follows. The set $\mathbb{M}_{\text{a}}(p_-) \cap F^{-1}(c_{\text{bot}})$ is path connected. Its boundary, the unstable manifold of p_-, consists of at least two points. Precisely one point of the boundary belongs to $\mathbb{M}_{\text{d}}(p_+)$: this point corresponds to the unique trajectory on M connecting p_- with p_+. There exists a point z_{end} on $\text{bd}(\mathbb{M}_{\text{a}}(p_-) \cap F^{-1}(c_{\text{bot}}))$ disjoint from $\mathbb{M}_{\text{d}}(p_+)$. By path connectedness, we extend the curve ϖ to a curve in $\mathbb{M}_{\text{a}}(p_-) \cap F^{-1}(c_{\text{bot}})$ in such a way that $\varpi(1) = z_{\text{end}}$ and ϖ is transverse to M at z_{end}.

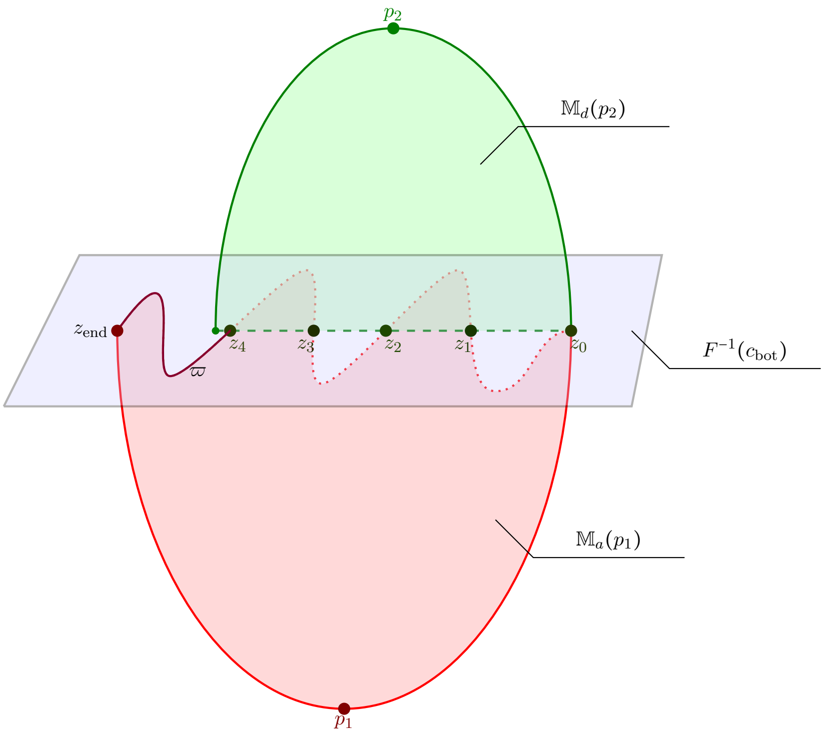

FIGURE 34. The notation of Section 5.1.2. For clarity only the part of the membrane $\mathbb{M}_\mathrm{d}(p_+)$ in $\{F \geq c_{\text{bot}}\}$ is drawn, as well as the part of the membrane $\mathbb{M}_\mathrm{a}(p_-)$ below the level set $F^{-1}(c_{\text{bot}})$.

It remains to show that ϖ can be chosen in such a way that $\varpi((\varepsilon_{\text{out}}, 1]) \cap Y_{\text{out}} = \varnothing$. Note that $\mathbb{M}_\mathrm{d}(p_+)$ intersects $\mathbb{M}_\mathrm{a}(p_-) \cap F^{-1}(c_{\text{bot}})$ along points $z_0, \dots, z_r$ and so Y_{out} intersects $\mathbb{M}_\mathrm{a}(p_-) \cap F^{-1}(c_{\text{bot}})$ along open balls near these points. If $\dim \mathbb{M}_\mathrm{a}(p_-) \cap F^{-1}(c_{\text{bot}}) > 1$, the complement $\mathbb{M}_\mathrm{a}(p_-) \cap F^{-1}(c_{\text{bot}}) \smallsetminus Y_{\text{out}}$ is still path connected so the choice of ϖ does not pose any problems.

If $\dim \mathbb{M}_\mathrm{a}(p_-) \cap F^{-1}(c_{\text{bot}}) = 1$, though, $\mathbb{M}_\mathrm{a}(p_-) \cap F^{-1}(c_{\text{bot}})$ is an arc with end points z_0 and z_{end} (there is only one possibility of chosing z_{end}). The set Y_{out} separates this arc. In order that ϖ exists, we need to choose the point z_m, from which ϖ emerges, to be the *nearest* point (among $z_1, \dots, z_r$) to z_{end} on $\mathbb{M}_\mathrm{a}(p_-) \cap F^{-1}(c_{\text{bot}})$, compare Figure 34, where $m = 4$ is the only possible choice. It might also happen that (5.1.4) defining ϖ points in the other direction, i.e. in Figure 34, the curve from z_4 goes towards z_3 instead of z_{end} as the parameter y_2 increases. If this happens, we globally replace y_2 by $-y_2$. Note that F does not depend on y_2 and ξ depends on y_2 only via y_2^2, so this change takes one coordinate system to another one, in which both F and ξ have the same shape.

We have constructed a guiding curve ϖ. For technical reasons we need to extend ϖ slightly across the point $\varpi(1)$, so we will assume, for some $\varepsilon_\gamma > 0$, that $\varpi\colon [0, 1 + \varepsilon_\gamma] \to F^{-1}(c_{\text{bot}})$ is defined as a small extension of $\varpi\colon [0,1] \to F^{-1}(c_{\text{bot}})$, and that ϖ is transverse to M at $\varpi(1)$. This is why ϖ extends slightly beyond M in Figures 35 and 36.

5.1.4. **Coordinates on a neighbourhood of the guiding curve.** Now we are going to extend the coordinate system from U_{out} to a neighbourhood of ϖ. First the coordinate system

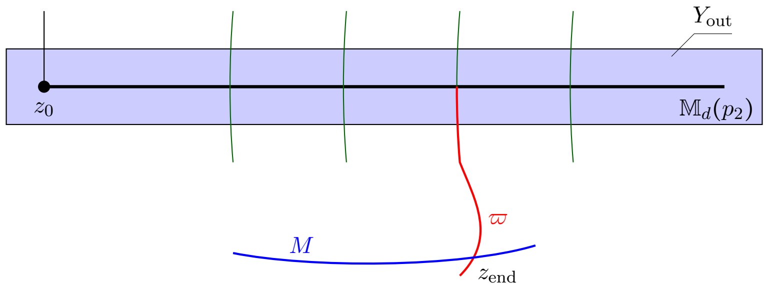

FIGURE 35. The guiding curve ϖ.

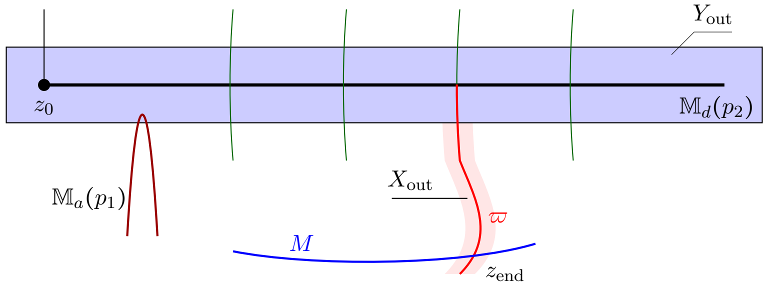

FIGURE 36. The set X_{out}.

will be constructed in the level set $F^{-1}(c_{\text{bot}})$, then in later subsections it will be extended to $F^{-1}[c_{\text{bot}}, c_{\text{top}}]$.

Choose a small neighbourhood X_{out} of $\varpi(\varepsilon_{\text{out}}/2, 1+\varepsilon_\gamma)$ in $F^{-1}(c_{\text{bot}})$, shown in Figure 36. We will assume that X_{out} has a product structure, that is $X_{\text{out}} \cong \varpi(\varepsilon_{\text{out}}/2, 1+\varepsilon_\gamma) \times D$ for some n-dimensional disc D (Ω is $(n+2)$-dimensional so level sets are $(n+1)$-dimensional) and that X_{out} intersects the manifold M away from the descending manifold of p_+. In the local coordinate system in $Y_{\text{out}} \cap X_{\text{out}}$, we will assume that X_{out} is given by $\{x_1^2 + \dots + x_n^2 \le \widetilde{\varepsilon}_{\text{out}}^2, y_2 \in (\varepsilon_{\text{out}}/2, \varepsilon_{\text{out}})\} \cap Y_{\text{out}}$, where $\widetilde{\varepsilon}_{\text{out}} \in (0, \varepsilon_{\text{out}}/2)$ is a real number. Recall that y_1 is fixed by being on the level set $F^{-1}(c_{\text{bot}})$, and y_2 changes along ϖ.

With respect to the product structure $X_{\text{out}} \cong \varpi \times D$, we write $\partial X_{\text{out}} = D_- \cup R_{\text{tang}} \cup D_+$, where $D_- = \varpi(\varepsilon_{\text{out}}/2) \times D$, $D_+ = \varpi(1+\varepsilon_\gamma) \times D$ and $R_{\text{tang}} = \varpi \times \operatorname{bd} D$. Here D_- is given by $X_{\text{out}} \cap \{y_2 = \varepsilon_{\text{out}}/2\} \cap \{x_1^2 + \dots + x_n^2 \le \widetilde{\varepsilon}_{\text{out}}^2\}$. As $\widetilde{\varepsilon}_{\text{out}} < \varepsilon_{\text{out}}/2$ we have that $D_- \subseteq Y_{\text{out}}$.

Choose an auxiliary nonvanishing vector field v on X_{out} such that the following hold:

(L-1) The vector field v is tangent to ϖ and to $\mathbb{M}_a(p_-) \cap F^{-1}(c_{\text{bot}})$.
(L-2) The vector field v is transverse to D_- and to D_+. Moreover v points into X_{out} on D_- and out of X_{out} on D_+.
(L-3) The vector field v is tangent to R_{tang}.
(L-4) For any $w \in X_{\text{out}} \setminus (D_+ \cup D_-)$, the trajectory of v through w hits D_+ in the future and hits D_- in the past.
(L-5) The time it takes the trajectory of v to go from D_- to D_+ is equal to $1+\varepsilon_\gamma$.
(L-6) In the local coordinates on $U_{\text{out}} \cap X_{\text{out}}$, the vector field v is given by $\frac{\partial}{\partial y_2}$.

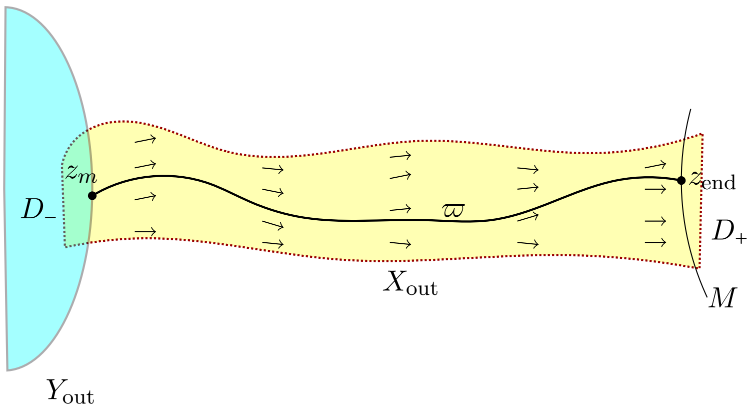

FIGURE 37. The auxiliary vector field υ and its flow.

The existence of such a υ is clear because X_{out} is homeomorphic to a disc. Notice that (L-6) is compatible with (L-1)—(L-5). Condition (L-5) can be obtained after all the other conditions by rescaling. The flow of υ allows us to extend the coordinate system from $U_{\mathrm{out}} \cap X_{\mathrm{out}}$ to the whole of X_{out}.

Take a point $u \in X_{\mathrm{out}}$ and let θ be the trajectory of υ through u. Parameterise θ as $\theta(s)$ so that $\theta(\varepsilon_{\mathrm{out}}/2) = w \in D_-$ and $\theta'(s) = \upsilon_{\theta(s)}$ for all $s \in [\varepsilon_{\mathrm{out}}/2, 1+\varepsilon_\gamma)$. Then, identifying w with its expression in coordinates in U_{out}, define the coordinates of u to be $w + (0, \dots, 0, 0, s_u)$, where s_u is the unique value of $s \in [\varepsilon_{\mathrm{out}}, 1+\varepsilon_\gamma)$ such that $\theta(s_u) = u$.

We conclude this subsection by making an assumption, whose rôle is to simplify the formulae that will appear in the sequel.

Lemma 5.1.2. *By possibly altering the vector field ξ below the level set $F^{-1}(c_{\mathrm{bot}})$, we may assume that the connected component of $M \cap F^{-1}(c_{\mathrm{bot}}) \cap X_{\mathrm{out}}$ containing $\varpi(1)$ is given by the set of equations $x_1 = 0$, $y_2 = 1$ and $y_1 = x_1^2 + \dots + x_{h+1}^2 - x_{h+2}^2 - \dots - x_n^2$. It follows in particular that $\mathbb{M}_{\mathrm{a}}(p_-) \cap X_{\mathrm{out}}$ is contained in $\{y_2 \le 1\}$.*

Sketch of proof. We know that $M \cap X_{\mathrm{out}} \cap F^{-1}(c_{\mathrm{bot}})$ is an $(n-1)$–dimensional disc containing $(0, \dots, 0, 0, 1)$ and also the equations $x_1 = 0, y_2 = 1$, $y_1 = x_1^2 + \dots + x_{h+1}^2 - x_{h+2}^2 - \dots - x_n^2$ define an $(n-1)$–dimensional disc in X_1 containing $(0, \dots, 0, 0, 1)$. These two discs are isotopic, so by an isotopy of M we can achieve that they coincide. Such an isotopy forces a change in ξ below the level set of $F^{-1}(c_{\mathrm{bot}})$, by the Isotopy Insertion Lemma 2.5.8; compare Section 5.1.2. ☐

5.1.5. **Flowing the guiding curve to different level sets.** Recall that we defined Ξ_s to be the flow of ξ and define

$$\overline{Z}_{\mathrm{out}} := \bigcup_{s \in \mathbb{R}} \Xi_s(\overline{X}_{\mathrm{out}}) \cap F^{-1}[c_{\mathrm{bot}}, c_{\mathrm{top}}], \quad Z_{\mathrm{out}} = \operatorname{Int} \overline{Z}_{\mathrm{out}}.$$

Then set

$$V_{\mathrm{out}} := U_{\mathrm{out}} \cup Z_{\mathrm{out}}.$$

Using the flow of ξ, we can extend the coordinate system from $U_{\mathrm{out}} \cup X_{\mathrm{out}}$ to V_{out}. Note that the fact that $\varpi(1)$ does not belong to the descending membrane of p_+ implies that Z_{out} is a product $X_{\mathrm{out}} \times (c_{\mathrm{bot}}, c_{\mathrm{top}})$ and the vector field ξ does not vanish on Z_{out}.

Remark. This is a place where we assume that there are no critical points of F other than p_- and p_+. Otherwise, we should strive to show that there are no trajectories connecting $\overline{X}_{\mathrm{out}}$ with a critical point.

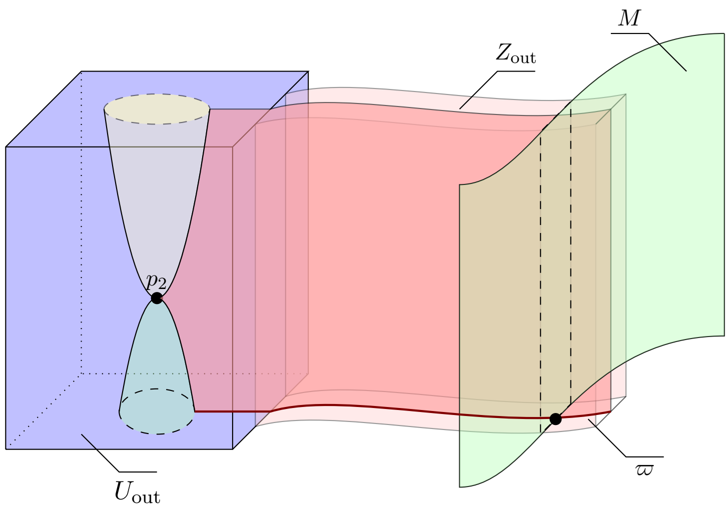

FIGURE 38. The sets Z_{out} and U_{out}.

Recall that U_{out} is a neighbourhood of the critical point p_+, whereas X_{out}, a neighbourhood of the part of ϖ that is outside U_{out}, is a subset of the level set $F^{-1}(c_{\mathrm{bot}})$; see Figure 38.

We want to extend our coordinate system to Z_{out}. The extension on Z_{out} is given by the requirement that the vector field ξ is given by

$$(-x_1, \dots, -x_{h+1}, x_{h+2}, \dots, x_n, y_1^2 + y_2^2, 0). \tag{5.1.5}$$

The procedure for extending the coordinate system is analogous to extending the coordinate system to X_{out} using the auxiliary vector field υ, which was described above.

Our aim is to show that in the new coordinate system F is given by $-x_1^2 - \dots - x_{h+1}^2 + x_{h+2}^2 + \dots + x_n^2 + y_1$ (up to a constant). In order to obtain this condition and (5.1.5) at the same time, we need to rescale ξ by a positive factor. The procedure is as follows.

Suppose $\phi\colon Z_{\mathrm{out}} \to (0, \infty)$ is a smooth function and consider the vector field $\phi\xi$. Take a point $u \in Z_{\mathrm{out}}$. Suppose there is a trajectory θ_ϕ of $\phi \cdot \xi$ such that $\theta_\phi(0) = u' \in X_{\mathrm{out}}$. Suppose u' has coordinates $(x_1', \dots, x_n', y_1', y_2')$. For $s \geq 0$, we set the coordinates of $\theta_\phi(s)$ to be $(x_1^\phi, \dots, y_2^\phi)$, where

$$\begin{aligned} &x_i^\phi = e^{-s} x_i' \text{ for } i \leq h+1, \quad x_i^\phi = e^s x_i' \text{ for } i > h+1, \\ &y_1^\phi = y_2' \tan(s y_2' + \arctan(y_1'/y_2')), \quad y_2^\phi = y_2'. \end{aligned} \tag{5.1.6}$$

For simplicity, we do not write explicitly the dependence of $x_1^\phi, \dots, y_2^\phi$ on the variable s. In these coordinates, the vector field $\phi\xi$ has the form

$$(-x_1^\phi, \dots, -x_{h+1}^\phi, x_{h+2}^\phi, \dots, x_n^\phi, (y_1^\phi)^2 + (y_2^\phi)^2, 0). \tag{5.1.7}$$

To justify this, one must think of $(x'_1, \dots, x'_n, y'_1, y'_2)$ as constants, and differentiate x_i^ϕ and y_i^ϕ with respect to the variable s. This yields:

$$\begin{aligned}
\frac{dx_i^\phi}{ds} &= \frac{d}{ds}(-e^{-s}x'_i) = -x_i^\phi, \text{ for } i \le h+1,\\
\frac{dx_i^\phi}{ds} &= \frac{d}{ds}(e^{s}x'_i) = x_i^\phi \text{ for } i > h+1,\\
\frac{dy_1^\phi}{ds} &= (y'_2)^2 \sec^2\left(sy'_2 + \arctan(y'_1/y'_2)\right) =\\
&=(y'_2)^2\left(1+\tan^2\left(sy'_2 + \arctan(y'_1/y'_2)\right)\right) = (y'_2)^2 + (y'_1)^2.
\end{aligned}$$

Also note that at $s = 0$ we recover $x_i^\phi = x'_i$ and $y_i^\phi = y'_i$. Therefore the path $s \mapsto (x_1^\phi, \dots, x_n^\phi, y_1^\phi, y_2^\phi)$ is the trajectory of the vector field (5.1.5), but it is also, by the definition, the trajectory $\theta_\phi(s)$ of $\phi\xi$. Therefore, indeed, $\phi\xi$ is given by (5.1.5).

The next result adjusts the function ϕ in such a way that F has the desired form in the new coordinates.

Lemma 5.1.3. *There exists a unique choice of smooth function $\phi\colon Z_{\mathrm{out}} \to (0,\infty)$ such that, in the coordinate system $(x_1^\phi, \dots, y_2^\phi)$, F has the form*

$$F = -(x_1^\phi)^2 - \ldots - (x_{h+1}^\phi)^2 + (x_{h+2}^\phi)^2 + \dots + (x_n^\phi)^2 + (y_1^\phi) + c_+, \tag{5.1.8}$$

where we recall that $c_+ = F(p_+)$.

Proof. Throughout the proof, whenever we talk of a trajectory of a vector field (usually ξ or $\phi\xi$), we mean the part of the trajectory that stays in Z_{out}.

Choose a point $u' \in X_{\mathrm{out}}$, whose coordinates are $x'_1, \dots, x'_n, y'_1, y'_2$. Notice that $y'_2 \ge \varepsilon_{\mathrm{out}}/2 > 0$. Define an abstract function $\nu\colon \mathbb{R}_{\ge 0} \to \mathbb{R}$ by

$$\nu(s) = -e^{-2s}({x'_1}^2 + \ldots + {x'_{h+1}}^2) + e^{2s}({x'_{h+2}}^2 + \ldots + {x'_n}^2) + y'_2 \tan(sy'_2 + \arctan(y'_1/y'_2)).$$

Observe that this is always an increasing function.

Let θ be the trajectory of ξ such that $\theta(0) = u'$. The functions $s \mapsto F(\theta(s))$ and $s \mapsto \nu(s)$ are both smooth, strictly increasing functions of a real positive argument that agree at $s = 0$ and ν is unbounded. By one-dimensional inverse function theorem, there exists an increasing function $\psi\colon \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ with $\psi(0) = 0$ such that

$$F(\theta(\psi(s))) = \nu(s). \tag{5.1.9}$$

Note that the function ψ depends implicitly on the choice of u'. The formula

$$\phi(\theta(\psi(s))) = \frac{d}{ds}\psi(s) \tag{5.1.10}$$

defines the value of ϕ on the trajectory of ξ through u'. Different choices of initial points u' allow us to define ϕ on the whole of Z_{out}. With this definition, a straightforward argument involving smooth dependence of solutions on intial conditions, shows that ϕ is a smooth function on the whole of Z_{out}.

We claim that with the choice of ϕ from (5.1.10)

$$\theta_\phi(s) = \theta(\psi(s)). \tag{5.1.11}$$

To see this, notice that the two functions are equal at $s = 0$. We calculate the differentials of both sides. By the definition of θ_ϕ:

$$\frac{d}{ds}\theta_\phi(s) = \phi(\theta_\phi(s))\xi(\theta_\phi(s)) \tag{5.1.12}$$

On the other hand:

$$\begin{aligned}\frac{d}{ds}\theta(\psi(s)) &= \frac{d}{ds}\psi(s)\theta'_s(\psi(s)) \overset{(*)}{=} \phi(\theta(\psi(s)))\theta'(\psi(s)) \overset{(**)}{=} \\ &= \phi(\theta(\psi(s)))\xi(\theta(\psi(s))).\end{aligned} \tag{5.1.13}$$

Here (*) is precisely (5.1.10), whereas (**) holds because $s \mapsto \theta$ is the trajectory of ξ. Therefore the two functions, $\theta_\phi(s)$ and $\theta(\psi(s))$ satisfy the same differental equation

$$\frac{d}{ds}y(s) = \phi(y(s))\xi(y(s)) \tag{5.1.14}$$

with the same initial condition. Hence, they are equal. Note that the differential equation (5.1.14) is exactly the flow of the vector field $\phi\xi$.

Combining (5.1.9) with (5.1.11) we see that

$$F(\theta_\phi(s)) = \nu(s).$$

With the choice of coordinate system as in (5.1.6) we note that F has the desired form. □

By construction, the function ϕ is uniquely determined on Z_{out}, in particular $\phi \equiv 1$ on $U_{\mathrm{out}} \cap X_{\mathrm{out}}$. We extend ϕ by 1 to the whole of X_{out} and to a smooth positive function on the whole of Ω (which we take equal to 1 away from a small neighbourhood of $Z_{\mathrm{out}} \cap X_{\mathrm{out}}$). We rescale now ξ to $\phi\xi$; this preserves all the membranes. We also drop superscript ϕ from the coordinates $x_1, \dots, y_2$. By Lemma 5.1.3 the function F has the form (5.1.8) on the whole of Z_{out}.

We have now completed the definition of the set V_{out} and the coordinate system on it. In what follows, we will refer to V_{out} as the *outer shell.*

5.1.6. **Properties of the outer shell.** Before we pass to the construction of V_{mid} and V_{inn}, we sum up the construction of V_{out}. It is an open subset of Ω, endowed with a coordinate system $(x_1, \dots, x_n, y_1, y_2)$, such that the following holds.

- (V-1) The vector field ξ is given by $(-x_1, \dots, -x_{h+1}, x_{h+2}, \dots, x_n, y_1^2 + y_2^2, 0)$. In particular $\mathbb{M}_{\mathrm{d}}(p_+)$ is given by $\{x_{h+2} = \dots = x_n = 0, y_2 = 0, y_1 \le 0\}$.
- (V-2) The manifold M is given by $\{x_1 = 0, y_2 = 1\} \cup \{y_1 = 0, y_2 = 0\}$. The second component contains the critical point p_+ with coordinates $(0, \dots, 0)$.
- (V-3) The function F is given by $-x_1^2 - \dots - x_{h+1}^2 + x_{h+2}^2 + \dots + x_n^2 + y_1 + c_+$.
- (V-4) The set V_{out} is a grim neighbourhood of p_+; in particular the entrance set of V_{out} is contained in $F^{-1}(c_{\mathrm{bot}})$ and the exit set belongs to $F^{-1}(c_{\mathrm{top}})$.

5.2. The middle and the inner shells

5.2.1. **Definining the middle shell.** The middle and inner shells V_{mid} and V_{inn} are constructed as open subsets satisfying $\overline{V}_{\mathrm{inn}} \subseteq V_{\mathrm{mid}}$, $V_{\mathrm{mid}} \subseteq V_{\mathrm{out}}$. The objective of defining V_{inn} and V_{mid} is that the finger move will be done inside V_{inn}, and the vector field ξ will be changed inside V_{mid}. The coordinate system was defined on a larger set V_{out} so that we can control all the trajectories of ξ reaching p_+.

Set

$$V_{\mathrm{mid}} := \{(x, y) \in V_{\mathrm{out}} \colon |x_1| < \varepsilon_{\mathrm{fng}}\};$$

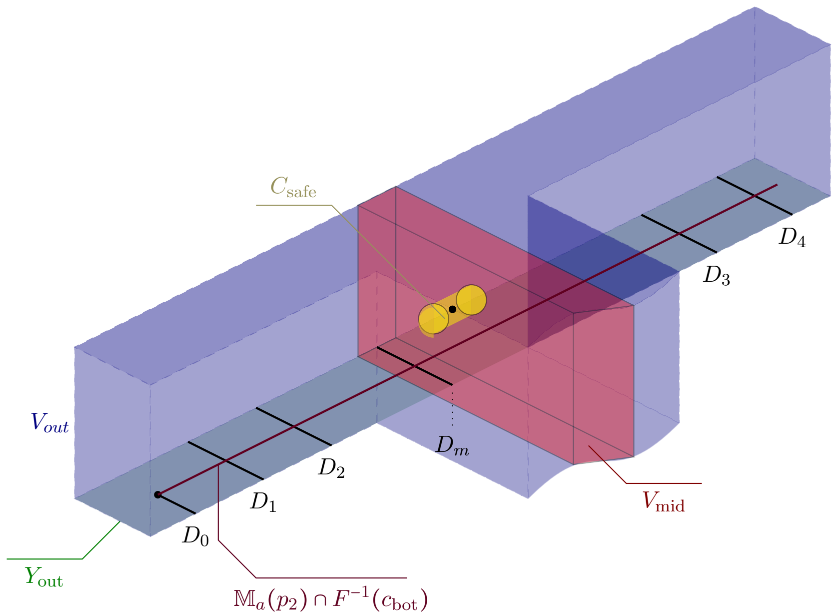

FIGURE 39. The middle shell.

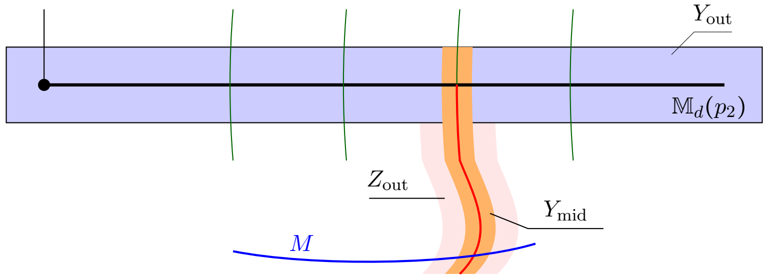

FIGURE 40. The set $Y_{\text{mid}} \coloneqq V_{\text{mid}} \cap F^{-1}(c_{\text{bot}})$.

see Figure 39. The main objective will be to fix the quantity ε_{fng}. Before we do this, let us denote.

$$\overline{Y}_{\text{mid}} \coloneqq \overline{V}_{\text{mid}} \cap F^{-1}(c_{\text{bot}}), \quad Y_{\text{mid}} = \operatorname{Int} Y_{\text{mid}},$$

where the interior is taken in the level set $F^{-1}(c_{\text{bot}})$; see Figure 40.

As a first step towards adjusting the value of ε_{fng}, we prove the following result.

Lemma 5.2.1. *There exists $\varepsilon^0_{\text{fng}} > 0$ such that if $\varepsilon_{\text{fng}} < \varepsilon^0_{\text{fng}}$, the middle set the set Y_{mid} is disjoint from the discs D_i if $i \neq m$; see Figure 40.*

Proof. By (5.1.2), the disc D_i is contained in $\{x_1 = s_{i,1}\}$, and for $i \neq m$, $s_{i,1} \neq 0$. We take $\varepsilon_{\text{fng}} < |s_{i,1}|$ for all $i \neq m$. □

The set V_{mid} is definitely not a grim neighbourhood of $V_{\text{mid}} \cap F^{-1}(c_{\text{bot}})$, because it does not contain the whole of $\mathbb{M}_{\text{d}}(p_+) \cap F^{-1}(c_{\text{bot}})$. However, if we restrict to the subspace $\{x_1 = 0\}$, then V_{mid} turns out to be a grim neighbourhood of p_+. We record this property for future use.

Lemma 5.2.2. *The set* $V_{\text{mid}} \cap \{x_1 = 0\}$ *is a grim neighbourhood of* p_+ *in* $V_{\text{out}} \cap \{x_1 = 0\}$.

Proof. By definition we have $V_{\text{mid}} \cap \{x_1 = 0\} = V_{\text{out}} \cap \{x_1 = 0\}$. The set V_{out} is the grim neighbourhood of p_+, see (V-4). The hypersurface $\{x_1 = 0\}$ is preserved by ξ. The statement follows quickly. $\square$

5.2.2. **Trajectories starting from D_i.** The choice of ε_{fng} is determined by the control of trajectories starting at D_i and coming close to p_+. We will consider two subsets C_{safe} and C_{vs} ('vs' standing for 'very safe') and shrink ε_{fng} in such a way that any trajectory from D_i enters V_{mid} through C_{vs} and, if it exits C_{safe} later on, it does it above the level set c_+. The finger move will not alter the vector field outside of $V_{\text{mid}} \smallsetminus C_{\text{safe}}$. This means that the finger move will not create any new trajectories from D_i to p_+; compare Theorem 5.4.1.

To begin with, let $\varepsilon_{\text{safe}}, \varepsilon_{\text{vs}}$ be two positive numbers with $\varepsilon_{\text{safe}} > \varepsilon_{\text{vs}}$. Define

$$C_{\text{safe}} = \{x_2^2 + \dots + x_n^2 + y_1^2 + y_2^2 \le \varepsilon_{\text{safe}}^2, x_1 \in [-\varepsilon_{\text{fng}}, \varepsilon_{\text{fng}}]\} \tag{5.2.1}$$

$$C_{\text{vs}} = \{x_2^2 + \dots + x_n^2 + y_1^2 + y_2^2 \le \varepsilon_{\text{vs}}^2, x_1 \in [-\varepsilon_{\text{fng}}, \varepsilon_{\text{fng}}]\}. \tag{5.2.2}$$

The purpose of Subsection 5.2.2 is to prove the following two results. The first one tells, that we can find a sufficiently thin neighbourhood of $\{x_1 = 0\}$ (controlled by the parameter ε_{fng}) a neighbourhood C_{vs} of the critical point, such that if a trajectory starts from a point D_i and hits V_{mid}, then either it hits C_{vs} (which is unavoidable, since there are trajectories of ξ from D_i to p_+), or it enters V_{mid} at the level set, where it has no chance to hit p_+ afterwards.

Lemma 5.2.3. *There exists* $\varepsilon_{\text{fng}}^0 > 0$ *such that for any* $\varepsilon_{\text{fng}} < \varepsilon_{\text{fng}}^0$ *there exists* $\varepsilon_{\text{vs}}^0 > 0$ *with the property that if* $\varepsilon_{\text{vs}} < \varepsilon_{\text{vs}}^0$, *then the following holds.*

If $w_1 \in D_i$, $i \neq m$ *and the trajectory of* ξ *through* w_1 *enters* V_{mid} *through* $w_2 \in \partial V_{\text{mid}}$ *(we require that* w_2 *is the point of the first entrance of the trajectory to* V_{mid} *after it hits* w_1*), then:*

(a) either $w_2 \in C_{\text{vs}}$*;*
(b) or $F(w_2) > c_+$.

Proof. Choose a point $w_1 \in D_i$. Write its coordinates as

$$w_1 = (s_{i,1}, \dots, s_{i,h+1}, x_{h+2}, \dots, x_n, y_{1,1}, y_2),$$

where

$$y_{1,1} = c_{\text{bot}} - c_+ + s_{i,1}^2 + \dots + s_{i,h+1}^2 - x_{h+2}^2 - \cdots x_n^2; \tag{5.2.3}$$

see (5.1.2) and (5.1.3). Denote

$$\Delta_- = s_{i,1}^2 + \dots + s_{i,h+1}^2, \quad \Delta_+ = x_{h+2}^2 + \dots + x_n^2.$$

Note that Δ_- is fixed for all the points in D_i, while Δ_+ can vary. There is a uniform upper bound on Δ_- and Δ_+ on the whole of $D_0 \cup \dots \cup D_r \smallsetminus D_m$.

Let $\gamma(s)$ be the trajectory of ξ such that $\gamma(0) = w_1$. If γ leads out of V_{out} without hitting V_{mid}, then, by (V-4), it leads out at the level set c_{top}, so it never reaches p_+; in that case there is nothing to prove. So assume γ reaches V_{mid} and let t_{hit} be the time of the first entrance of γ into V_{mid}. Set $w_2 = \gamma(t_{\text{hit}}) \in \partial \overline{V_{\text{mid}}}$.

Given (V-1), the coordinates of w_2 are calculated as follows,

$$w_2 = \gamma(t_{\text{hit}}) = (e^{-t_{\text{hit}}} s_{i,1}, \dots, e^{-t_{\text{hit}}} s_{i,h+1}, e^{t_{\text{hit}}} x_{i,h+2}, \dots, e^{t_{\text{hit}}} x_{i,n}, y_1(t_{\text{hit}}), y_2). \tag{5.2.4}$$

where $y_1(t_{\text{hit}})$ can also be determined, compare (5.1.6), but we do not use the explicit formula here.

By the definition of V_{mid}, the x_1-coordinate of w_2 is $\pm\varepsilon_{\mathrm{fng}}$. Therefore

$$t_{\mathrm{hit}} = \ln \frac{|s_{i,1}|}{\varepsilon_{\mathrm{fng}}}.$$

Note that by making $\varepsilon_{\mathrm{fng}}$ sufficiently small, we can make t_{hit} as large as we please.

Using (5.2.4) and (V-3), we compute

$$F(w_2) = -e^{-2t_{\mathrm{hit}}}\Delta_- + e^{2t_{\mathrm{hit}}}\Delta_+ + y_1(t_{\mathrm{hit}}) + c_+. \tag{5.2.5}$$

In the following estimate we recall t_{hit} is a function of w_1, but t_{hit} can be made uniformly large for all $w_1 \in D_i$ (and for all indices $i \neq m$) by choosing sufficiently small $\varepsilon_{\mathrm{fng}}$

Lemma 5.2.4. *For any $\delta > 0$ there exists $t_0 > 0$ such that for every $w_1 \in D_i$, if $t_{\mathrm{hit}} = t_{\mathrm{hit}}(w_1) > t_0$, then $y_1(t_{\mathrm{hit}}) > -\delta - \Delta_+$.*

Proof. The proof relies on explicit calculations. While we could use the formula as in (5.1.6) for y_1, we prefer to use more legible, but more rough, estimates. Let $\upsilon(s)$ be the y_1-coordinate of the point $\gamma(s)$. Note that $\upsilon(s)$ satisfies:

$$\begin{aligned} \frac{d}{ds}\upsilon &= \upsilon^2 + y_2^2 \\ \upsilon(0) &= c_{\mathrm{bot}} - c_+ + \Delta_- - \Delta_+. \end{aligned} \tag{5.2.6}$$

The first equation holds because the y_1 coordinate of ξ is $y_1^2 + y_2^2$. The second holds by (5.2.3). We have $\frac{d}{ds}\upsilon(s) \geq \upsilon^2$. Let $\upsilon_0(s)$ be the solution to the ODE $\frac{d}{ds}\upsilon_0(s) = \upsilon_0^2$ with initial condition $\upsilon_0(0) = \upsilon(0)$. That is,

$$\upsilon_0(s) = \frac{1}{\frac{1}{\upsilon(0)} - s}.$$

Gronwall's inequality from the classical theory of ODEs implies that $\upsilon(s) \geq \upsilon_0(s)$ for $s \geq 0$, so

$$\upsilon(t_{\mathrm{hit}}) \geq \frac{1}{\frac{1}{\upsilon(0)} - t_{\mathrm{hit}}}.$$

Now $\frac{1}{\frac{1}{c_{\mathrm{bot}} - c_+ + \Delta_- - \Delta_+} - t_{\mathrm{hit}}}$ tends to zero as $t_{\mathrm{hit}} \to \infty$. Hence, for sufficiently large t_{hit} we have $\upsilon(t_{\mathrm{hit}}) \geq -\delta - \Delta_+$. ∎

Resuming the proof of Lemma 5.2.3, choose $\delta < \varepsilon_{\mathrm{vs}}^2/10$ and adjust $\varepsilon_{\mathrm{fng}}$ according to Lemma 5.2.4. Recall that Δ_- is constant on D_i and, making $\varepsilon_{\mathrm{fng}}$ smaller if necessary, choose $\varepsilon_{\mathrm{fng}}$ in such a way that $e^{-2t_{\mathrm{hit}}}\Delta_- < \varepsilon_{\mathrm{vs}}^2/10$ as well. Then (5.2.5) implies that

$$\begin{aligned} F(w_2) = -e^{-2t_{\mathrm{hit}}}\Delta_- + e^{2t_{\mathrm{hit}}}\Delta_+ + y_1(t_{\mathrm{hit}}) + c_+ &\geq -e^{-2t_{\mathrm{hit}}}\Delta_- + e^{2t_{\mathrm{hit}}}\Delta_+ - \delta - \Delta_+ + c_+ \\ &\geq (e^{2t_{\mathrm{hit}}} - 1)\Delta_+ + c_+ - \varepsilon_{\mathrm{vs}}^2/5. \end{aligned} \tag{5.2.7}$$

Suppose $w_2 \notin C_{\mathrm{vs}}$. Recall that C_{vs} was given by (5.2.2), that is $|x_1| \leq \varepsilon_{\mathrm{fng}}$ and $x_2^2 + \dots + y_2^2 \leq \varepsilon_{\mathrm{vs}}^2$. If $w_2 \notin C_{\mathrm{vs}}$, by (5.2.4), we obtain the following conditions:

$$\left|e^{-t_{\mathrm{hit}}} s_{i,1}\right| > \varepsilon_{\mathrm{fng}}, \text{ or } e^{-2t_{\mathrm{hit}}}\Delta_- + e^{2t_{\mathrm{hit}}}\Delta_+ + y_1(t_{\mathrm{hit}})^2 + y_2^2 > \varepsilon_{\mathrm{vs}}^2. \tag{5.2.8}$$

However, we know that w_2 is on the boundary of V_{mid}, so $|e^{-t_{\mathrm{hit}}} s_{i,1}| = \varepsilon_{\mathrm{fng}}$. Therefore, the second inequality in (5.2.8) has to hold. There are four terms on the left-hand side of the second inequality, so at least one of them is greater than one fourth of the right-hand side. We will list relevant cases and try to show that either there is an immediate contradiction (using the fact that t_{hit} is large), or that $F(w_2) > c_+$. These cases are:

- $e^{-2t_{\mathrm{hit}}}\Delta_- > \varepsilon_{\mathrm{vs}}^2/4$. This is impossible if $t_{\mathrm{hit}} > s_0$ for some fixed s_0, which can be made uniform for all w_1.
- $e^{2t_{\mathrm{hit}}}\Delta_+ > \varepsilon_{\mathrm{vs}}^2/4$. Then $(e^{2t_{\mathrm{hit}}}-1)\Delta_+ > \frac{2}{9}\varepsilon_{\mathrm{vs}}^2$ if t_{hit} is large, and so by (5.2.7) we conclude that $F(w_2) > c_+$.
- $y_1(t_{\mathrm{hit}}) < -\sqrt{\varepsilon_{\mathrm{vs}}^2/4} = -\varepsilon_{\mathrm{vs}}/2$. But then by Lemma 5.2.4 we infer that $-\varepsilon_{\mathrm{vs}}/2 > -\delta - \Delta_+ > -\varepsilon_{\mathrm{vs}}^2/10 - \Delta_+$, which implies that $\Delta_+ \geq \varepsilon_{\mathrm{vs}}/2 - \varepsilon_{\mathrm{vs}}^2/10$. This is greater than $\varepsilon_{\mathrm{vs}}^2/4$ if $\varepsilon_{\mathrm{vs}}$ is small. We reduce to the previous case.
- $y_1(t_{\mathrm{hit}}) > \sqrt{\varepsilon_{\mathrm{vs}}^2/4} = \varepsilon_{\mathrm{vs}}/2$ and $e^{-2t_{\mathrm{hit}}}\Delta_- < \varepsilon_{\mathrm{vs}}^2/4$. Then by (5.2.5) we obtain that $F(w_2) \geq c_+ + \varepsilon_{\mathrm{vs}}/2 - \varepsilon_{\mathrm{vs}}^2/10 > c_+$, again provided $\varepsilon_{\mathrm{vs}}$ is small.
- $y_2^2(t_{\mathrm{hit}}) \geq \varepsilon_{\mathrm{vs}}^2/4$. The y_2-coordinate is constant along the trajectory of ξ, hence the y_1 coordinate of $\gamma(s)$ satisfies the inequality $\frac{d}{ds}v(s) \geq \varepsilon_{\mathrm{vs}}^2/4$ (compare (5.2.6)) hence $v(t_{\mathrm{hit}}) \geq v(0) + t_{\mathrm{hit}}\varepsilon_{\mathrm{vs}}^2/4$. With t_{hit} sufficiently large we arrive at the previous case, that is $y_1(t_{\mathrm{hit}}) > \varepsilon_{\mathrm{vs}}/2$.

We have assumed that w_2 is not contained in C_{vs}, and have shown that $F(w_2) > c_+$, as desired. □

The second result after Lemma 5.2.3 also adjusts the value of $\varepsilon_{\mathrm{fng}}$, but with respect to another parameter. We fix a neighbourhood C_{safe} as above, whose size is controlled by a parameter $\varepsilon_{\mathrm{safe}}$. We prove we can choose $\varepsilon_{\mathrm{fng}}$ small enough that if a trajectory enters C_{vs} then on exiting the larger set C_{safe}, it is at a level set above p_+, so that it has no chances to reach p_+ in the future.

Lemma 5.2.5. *For any $\varepsilon_{\mathrm{safe}} > 0$, there exists $\varepsilon_{\mathrm{vs}}^0 > 0$ such that if $\varepsilon_{\mathrm{vs}} < \varepsilon_{\mathrm{vs}}^0$, then if a trajectory of ξ enters V_{mid} through $w_2 \in C_{\mathrm{vs}}$, and exits C_{safe} at a point w_3, then $F(w_3) > c_+$.*

Proof. We let the coordinates of w_2 be

$$w_2 = (x_1, \dots, x_n, y_1, y_2).$$

We also set

$$\Delta_- = x_1^2 + \dots + x_{h+1}^2,\ \ \Delta_+ = x_{h+2}^2 + \dots + x_n^2.$$

As $w_2 \in C_{\mathrm{vs}}$ we have that

$$\Delta_-,\ \Delta_+,\ y_1^2,\ y_2^2 \leq \varepsilon_{\mathrm{vs}}^2, \quad |x_1| < \varepsilon_{\mathrm{fng}}. \tag{5.2.9}$$

The trajectory of ξ through w_2 leaves C_{safe}. We let s_0 be the time after which this trajectory leaves C_{safe} for the first time and let w_3 be the point at which this trajectory leaves C_{safe}. The coordinates of w_3 are clearly

$$w_3 = (e^{-s_0}x_1, \dots, e^{-s_0}x_{h+1}, e^{s_0}x_{h+2}, \dots, e^{s_0}x_n, \widetilde{y}_1, y_2). \tag{5.2.10}$$

Here, $\widetilde{y}_1$ is the y_1-coordinate of w_3. We have $\widetilde{y}_1 \geq y_1$ and $y_1 \geq -\varepsilon_{\mathrm{vs}}$ by (5.2.9), so

$$\widetilde{y}_1 \geq -\varepsilon_{\mathrm{vs}}. \tag{5.2.11}$$

By property (V-3) combined with (5.2.10):

$$F(w_3) - c_+ = e^{2s_0}\Delta_+ - e^{-2s_0}\Delta_- + \widetilde{y}_1. \tag{5.2.12}$$

In (5.2.1) C_{safe} is given by two conditions: $|x_1| \leq \varepsilon_{\mathrm{fng}}$, $x_2^2 + \dots + \dots + y_2^2 \leq \varepsilon_{\mathrm{safe}}^2$. As $w_3 \in \partial C_{\mathrm{safe}}$, one of the two inequalities must be equality. We argue as in the proof of Lemma 5.2.3: either the first inequality is an equality, or one of the four terms $x_2^2 + \dots + x_{h+1}^2$, $x_{h+2}^2 + \dots + x_n^2$, y_1^2, y_2^2 at w_3 should be greater or equal than $\varepsilon_{\mathrm{safe}}^2/4$. The coordinates of w_3 are given by (5.2.10), so $w_3 \in \partial C_{\mathrm{safe}}$ implies that one of the following holds.

(a) $e^{-2s_0}\Delta_- \geq \varepsilon_{\text{safe}}^2/4$;
(b) $e^{2s_0}\Delta_+ \geq \varepsilon_{\text{safe}}^2/4$;
(c) $\widetilde{y}_1 \geq \varepsilon_{\text{safe}}/2$;
(d) $\widetilde{y}_1 \geq -\varepsilon_{\text{safe}}/2$;
(e) $|y_2| \geq \varepsilon_{\text{safe}}/2$;
(f) $|e^{-2s_0}x_1| \geq \varepsilon_{\text{fng}}$.

We now show that among these possibilities some are self-contradictory (given (5.2.9)), the others lead to $F(w_3) > c_+$.

- Case (a) is impossible, because $\Delta_- < \varepsilon_{\text{vs}}^2$; compare (5.2.9).
- If (b) holds, then in (5.2.12) we use that $\Delta_- \leq \varepsilon_{\text{vs}}^2$, $e^{s_0} > 1$ and (5.2.11) to get $F(w_3) - c_+ \geq \varepsilon_{\text{safe}}^2 - \varepsilon_{\text{vs}}^2 - \varepsilon_{\text{vs}}$. If ε_{vs} is small enough compared with $\varepsilon_{\text{safe}}$, we have $F(w_3) - c_+ > 0$.
- In case (c) we have $F(w_3) - c_+ \geq \varepsilon_{\text{safe}} - e^{-2s_0}\Delta_- \geq \varepsilon_{\text{safe}} - \varepsilon_{\text{vs}}^2$. This difference is positive if ε_{vs} is sufficiently small.
- Case (d) contradicts (5.2.11).
- Case (e) cannot occur, because y_2 is constant on the trajectory, so we must have $|y_2| \leq \varepsilon_{\text{vs}}$.
- Finally, we discuss case (f). Note that $|x_1| < \varepsilon_{\text{fng}}$ by (5.2.9). As $s_0 > 0$, we have that $|e^{-2s_0}x_1| < \varepsilon_{\text{fng}}$.

We have shown that $F(w_3) > c_+$, so the lemma is proved. □

Condition 5.2.6. The value of ε_{fng} is chosen in such a way that the statements of Lemmas 5.2.1 and 5.2.3 hold.

The values of $\varepsilon_{\text{safe}}$ and ε_{vs} will be adjusted in later subsections according to Lemmas 5.2.3 and 5.2.5.

5.2.3. **The inner shell V_{inn}.** We will also define the set V_{inn}. First introduce the following piece of notation. For every $\alpha \in [0,1]$ we denote

$$c_{-\alpha} = c_+ + (c_{\text{bot}} - c_+)\alpha \text{ and } c_{+\alpha} = c_+ + (c_{\text{top}} - c_+)\alpha. \tag{5.2.13}$$

For $\varepsilon_R > 0$ consider a cylinder

$$V_{\text{inn}}(\varepsilon_R) = \{x_1^2 + \cdots + x_n^2 \leq \varepsilon_R^2, y_1 \in [c_{-1/2} - c_+, c_{+1/2} - c_+], y_2 \in [-\varepsilon_R, 1 + \varepsilon_R]\} \subseteq V_{\text{out}}. \tag{5.2.14}$$

Lemma 5.2.7. *For ε_R small enough, the interior of $V_{\text{inn}}(\varepsilon_R)$ is a subset of V_{mid}.*

Proof. The intersection of all of the $V_{\text{inn}}(\varepsilon_R)$ is the 2-dimensional rectangle $(0,\dots,0,t,s)$, $t \in [c_{-1/2} - c_+, c_{+1/2} - c_+]$, $s \in [0,1]$. This rectangle belongs to V_{mid} and we can choose a neighbourhood U_{fng} of the rectangle contained in V_{mid}. For ε_{R_0} sufficiently small, the interior of $V_{\text{inn}}(\varepsilon_{R_0})$ (interior in $F^{-1}[c_{-1/2}, c_{+1/2}]$) is contained in U_{fng}. Hence, for $\varepsilon_R = \varepsilon_{R_0}/2$, the whole of $V_{\text{inn}}(\varepsilon_R)$ belongs to V_{mid}. □

Condition 5.2.8. From now on set ε_R in such a way that

$$V_{\text{inn}} := V_{\text{inn}}(\varepsilon_R)$$

is contained in V_{mid}; compare Figure 41.

The purpose of the sets $V_{\text{inn}}, V_{\text{mid}}$ and V_{out} will be as follows.

- V_{inn} is the set where the finger move is conducted.
- V_{mid} is the set where the vector field ξ is changed after the finger move, so that it is still a grim vector field of some function.

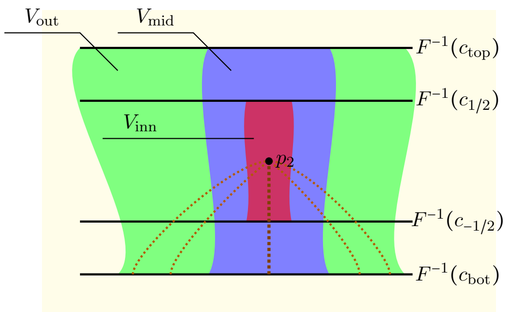

FIGURE 41. The position of the sets V_{inn}, V_{mid} and V_{out} presented in a two-dimensional schematic. The dotted lines represent the trajectories of ξ connecting p_- and p_+.

- V_{out} is the set where we control the trajectories of the vector field ξ.

5.3. The finger move

Having defined a suitable neighbourhood of the guiding curve in Subsection 5.1, we pass to an explicit description of the finger move in Subsection 5.3.1, that is, of the isotopy $G_{1+\tau}$. Note that we assume that $G_\tau \equiv G$ for $\tau \in [0,1]$. In Subsection 5.3.2 we show how to change the vector field ξ to a new vector field $\widehat{\xi}$. In Subsection 5.3.3 we study the critical points of $\widehat{\xi}$ and show that $\partial_{\widehat{\xi}} F \geq 0$, which shows that $\widehat{\xi}$ is a good candidate for a grim vector field. It turns out that $\widehat{\xi}$ vanishes on the newly created sphere of double points. In fact, the function F is not an immersed Morse function on $G_2(M)$ (it has non-isolated critical points on the second stratum after the finger move). Therefore, we define an explicit perturbation of both F and $\widehat{\xi}$ near that sphere of double points. This is done in Subsection 5.3.4, where we define the path F_τ for $\tau \in [0,1]$ and a perturbation of $\widehat{\xi}$, denoted by $\overline{\xi}$. Property (FM-1) of Finger Move Theorem is proved in Subsection 5.3.4, and (FM-2) follows immediately from the construction of F_τ. The key property of $\widehat{\xi}$, namely (FM-3), is proved in Section 5.4.

5.3.1. **Explicit description of the finger move.** As described in Section 6.3 via an explicit example, we will start at the level $c_{+1/4}$. We will push M along the guiding curve ϖ. The length of the push will vary as the height varies.

By (V-2), the intersection of M with V_{out} consists of two connected components. The component containing p_+ is

$$M_{p_+} := M \cap U_{\text{out}}.$$

The second component belongs to $\{y_2 = 1\}$. We consider its part contained in V_{inn}.

$$M_{\text{orig}} = M \cap V_{\text{inn}} \cap \{y_2 = 1\}.$$

Now we define the finger move, that is, the path $G_{1+\tau}$ of immersions such that $G_1 = G$. Choose $u \in N$. If $G(u) \notin M_{\text{orig}}$, we set $G_{1+\tau}(u) = G(u)$ for all $\tau \in [0,1]$. If $G(u) = (x_1, \dots, y_2) \in M_{\text{orig}}$, we set

$$G_{1+\tau}(u) := (x_1, \dots, y_1, (1-\tau)y_2 + \tau\mathfrak{s}(x_1, \dots, x_n, y_1)) \in V_{\text{inn}}, \tag{5.3.1}$$

$\tau \in [0,1]$, for a function $\mathfrak{s}\colon \mathbb{R}^{n+1} \to \mathbb{R}$ which we will define shortly. Note that as F does not depend on the y_2 coordinate, see property (V-3), we immediately obtain.

Lemma 5.3.1. *The composition $F \circ G_{1+\tau}$ is equal to $F \circ G$.*

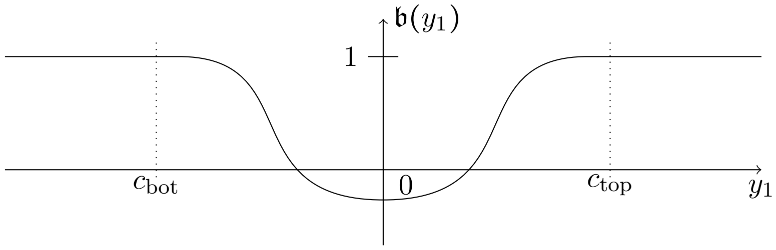

FIGURE 42. A graph of $\mathfrak{b}\colon \mathbb{R} \to \mathbb{R}$.

By definition, the finger move replaces the original part of M, M_{orig} by

$$M_{\mathrm{fng}} := \{x_1 = 0, y_2 = \mathfrak{s}(x_1, \dots, x_n, y_1)\}, \tag{5.3.2}$$

that is to say $G_2(N) = (M \smallsetminus M_{\mathrm{orig}}) \cup M_{\mathrm{fng}}$. We define $M_{\mathrm{new}} = G_2(N)$. To complete the construction, we need to define $\mathfrak{s}$. Set

$$\mathfrak{s}(x_1, \dots, y_1) = \mathfrak{a}(y_1) \min\{x_1^2 + x_2^2 + \dots + x_n^2, \varepsilon_R^2\} + \mathfrak{b}(y_1), \tag{5.3.3}$$

where $\mathfrak{a}\colon \mathbb{R} \to [0, \infty)$ and $\mathfrak{b}\colon \mathbb{R} \to \mathbb{R}$ are two functions such that

$$\mathfrak{a}(y_1)\varepsilon_R^2 + \mathfrak{b}(y_1) \equiv 1. \tag{5.3.4}$$

The function $\mathfrak{a}$ is nonnegative. When all $x_i = 0$, the function is $\mathfrak{b}(y_1)$. Moreover we require that:

Condition 5.3.2.

(A-1) $\mathfrak{b} \equiv 1$ (and so $\mathfrak{a} \equiv 0$) for $y_1 \notin [c_{-1/4} - c_+, c_{+1/4} - c_+]$;
(A-2) $\mathfrak{b}$ has a minimum at $y_1 = 0$, it is non-increasing for $y_1 > 0$ and non-decreasing for $y_1 < 0$;
(A-3) $\mathfrak{b}(0) < 0$ and $|\mathfrak{b}(0)| < \varepsilon_R$, so that $M_{\mathrm{fng}} \subseteq V_{\mathrm{inn}}$;
(A-4) $\mathfrak{b}'(y) \neq 0$ whenever $\mathfrak{b}(y) = 0$.

A graph of an exemplary function for $\mathfrak{b}$ is presented in Figure 42. A schematic of the finger move is depicted in Figure 43.

Remark. As the reader may have noticed, M_{new} has corners along the common boundary of M_{fng} and M_{orig}. These corners can be avoided by taking a smooth approximation of $\min(x_1^2 + x_2^2 + \dots + x_n^2, \varepsilon_R^2)$ in (5.3.2). We will not do that explicitly because it would make the exposition even more complicated.

Note that M_{new} has extra double points. In fact, we have the following result towards proving property (FM-6)

Lemma 5.3.3. *The map $G_{1+\tau}$ is an immersion and it is regular for all values of τ except for a single value of τ.*

Proof. Since G_1 is an immersion, and G_τ changes only the coordinate y_2, which was constant on M_{orig}, it is clear that G_τ is an immersion. In fact, an easy calculation shows that the double point set of $G_\tau(M_{\mathrm{orig}}) \cap M_{p_+}$ intersect along the sphere $x_1 = y_1 = y_2 = 0$ and $\{x_2^2 + \dots + x_n^2 = \varepsilon_R^2 - \frac{1}{\tau\mathfrak{a}(0)}\}$. The only τ for which the immersion is not regular, is when the radius of the sphere is zero, that is $\tau = \varepsilon_R^2/\mathfrak{a}(0)$. ∎

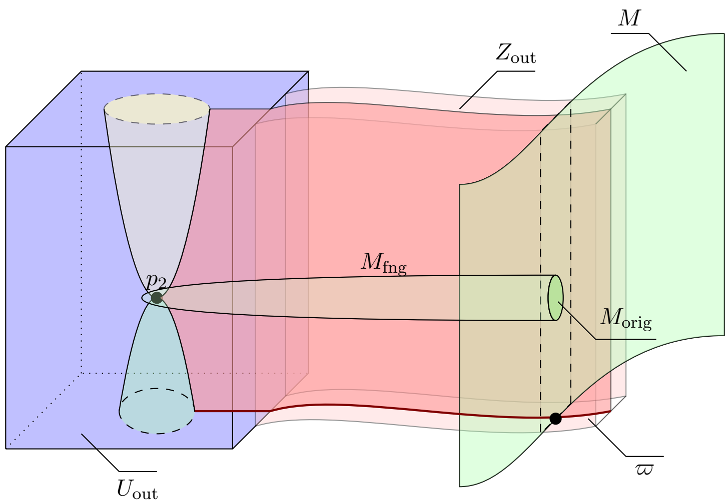

FIGURE 43. A schematic of the finger move.

From this lemma we deduce that $M_{p_+} \cap M_{\text{fng}}$ is given by

$$\Sigma = \{x_2^2 + \cdots + x_n^2 = \varepsilon_\Sigma^2\} \cap \{x_1 = y_1 = y_2 = 0\}, \tag{5.3.5}$$

where

$$\varepsilon_\Sigma^2 = -\frac{\mathfrak{b}(0)}{\mathfrak{a}(0)}. \tag{5.3.6}$$

The double point set is an $(n-2)$-dimensional sphere. Note that the two branches of M intersect transversely.

The last observation in this subsection prove (FM-6).

Lemma 5.3.4. *The components M_{orig} and M_{p_+} of $M \cap V_{\text{out}}$ belong to the same connected component of M.*

Proof. The part M_{p_+} is a subset of the connected component of M containing p_+. Note that p_+ and p_- are connected by a single trajectory of ξ, hence M_{p_+} belongs to the same connected component of M as p_-. The intersection of the ascending membrane $\mathbb{M}_a(p_-)$ with M belongs to the same connected component of M as p_-. The endpoint of the guiding curve $\varpi(1)$ belongs to the same connected component of M as p_-, because it belongs to the ascending membrane. But M_{orig} contains $\varpi(1)$ by construction. □

5.3.2. **Construction of $\widehat{\xi}$.** The next goal is to modify the vector field ξ to obtain a vector field $\widehat{\xi}$, which will be perturbed to $\overline{\xi}$ later on. The definition of $\widehat{\xi}$ is more important, because it will already have the property (FM-3) from Theorem 4.4.13 – one fewer trajectory between p_- and p_+. Note that

$$\partial_\xi \mathfrak{s}(x_1, \dots, y_2) = 0 \text{ if } x_1^2 + x_2^2 + \dots + x_n^2 \geq \varepsilon_R^2, \tag{5.3.7}$$

which follows immediately from (5.3.3), since then $\mathfrak{s}$ is constant. It is easy to see that the vector field

$$\widetilde{\xi} := (-x_1, \dots, -x_{h+1}, x_{h+2}, \dots, x_n, y_1^2 + y_2^2, \partial_\xi \mathfrak{s}) = \xi + \partial_\xi \mathfrak{s} \frac{\partial}{\partial y_2} \tag{5.3.8}$$

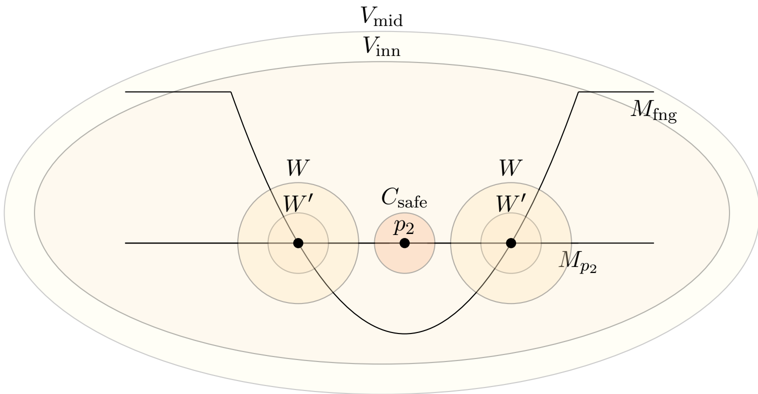

FIGURE 44. The subsets V_{inn}, V_{mid}, W and W' that we use to study the vector field $\widehat{\xi}$. We draw the intersection with the hyperplane $\{x_1 = 0\}$.

is tangent to M_{fng}. Moreover, by (V-3) we have $\frac{\partial}{\partial y_2} F = 0$, hence

$$\partial_{\widetilde{\xi}} F = \partial_\xi F \tag{5.3.9}$$

everywhere on V_{out}.

The two vector fields ξ and $\widetilde{\xi}$ do not match on $M_{\text{fng}} \cap M_{p_+}$. To rectify this, we will interpolate between ξ and $\widetilde{\xi}$, to obtain a third vector field, which will be $\widehat{\xi}$. The most naive approach, that is, to define the new vector field $\widehat{\xi}$ as a linear combination of ξ and $\widetilde{\xi}$ multiplied by two functions vanishing on M_{fng} and M_{p_+} respectively, does not quite work, because a nontrivial linear combination of ξ and $\widetilde{\xi}$ does not have to be tangent to the stratum $M_{\text{fng}} \cap M_{p_+}$.

Let e_1 be the vector field

$$e_1 := (0, x_2, \dots, x_n, 0, 0). \tag{5.3.10}$$

We claim that

$$\partial_{e_1} \mathfrak{s} > 0 \tag{5.3.11}$$

everywhere on Σ. To see this, note that if $x_1^2 + \dots + x_n^2 \le \varepsilon_R^2$ (which holds on Σ), then

$$\partial_{e_1} \mathfrak{s} = 2(x_2^2 + \dots + x_n^2)\mathfrak{a}(y_1) + \mathfrak{b}(y_1).$$

On Σ, $x_2^2 + \dots + x_n^2 = \varepsilon_\Sigma^2$, so $\partial_{e_1} \mathfrak{s} = 2\varepsilon_\Sigma^2 \mathfrak{a}(0) + \mathfrak{b}(0)$. By (5.3.6), we obtain $\partial_{e_1} \mathfrak{s} = \varepsilon_\Sigma^2 \mathfrak{a}(0) > 0$.

Choose $W \subseteq V_{\text{inn}}$ to be a neighbourhood of the double point set Σ satsfying the following two conditions; see Figure 44.

Condition 5.3.5. The set W is disjoint from p_+, and $\partial_{e_1} \mathfrak{s} > 0$ everywhere on W.

The latter is possible because of (5.3.11). Let W' be a neighbourhood of Σ in Ω such that the closure of W' is contained in W.

Now we adjust $\varepsilon_{\text{safe}}$ to satisfy the assumptions of Lemma 5.2.5 and by the following condition:

Condition 5.3.6. C_{safe} is a subset of V_{inn}, the closures of C_{safe} and W are disjoint, and C_{safe} is disjoint from M_{fng}.

Set $\rho_{\text{fng}}\colon V_{\text{mid}} \to \mathbb{R}_{\ge 0}$ to be the square of the distance to M_{fng}, and set $\rho_+\colon V_{\text{mid}} \to \mathbb{R}_{\ge 0}$ to be the square of the distance to M_{p_+}. Finally, choose two bump functions $\phi_1, \phi_2\colon \Omega \to [0,1]$,

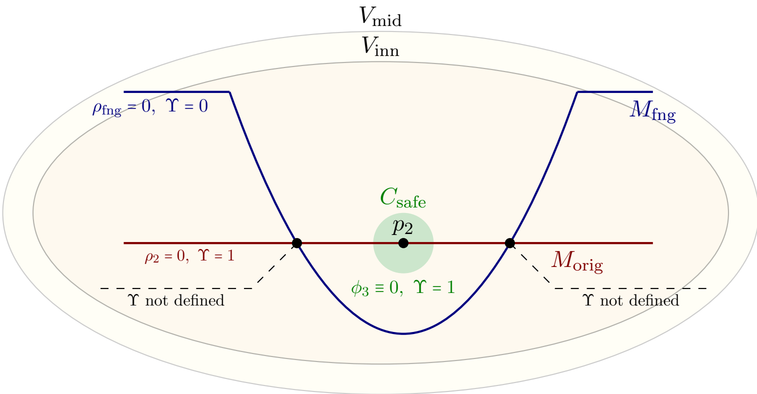

FIGURE 45. Figure 44 revisited. We indicate the values of ρ_{fng}, ρ_+ and Υ.

where ϕ_1 is a function supported on V_{mid}, equal to 1 on W, and ϕ_2 is a function supported in W and equal to 1 on W'

In W, define the decomposition of the tangent bundle $T\Omega$ into $E_{\tan}$, $E_1 = \operatorname{span}\{\frac{\partial}{\partial y_1}, (\partial_{e_1}\mathfrak{s})\frac{\partial}{\partial y_2} + e_1\}$ and $E_2 = \operatorname{span}\{e_1, \frac{\partial}{\partial x_1}\}$. By Condition 5.3.5, $E_1 \cap E_2 = 0$. The subbundle $E_{\tan}$ can be defined as the orthogonal complement of E_1 in $\operatorname{span}\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\}$. Recall that we are still using the metric for which the coordinate vectors $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial y_2}$ are orthonormal.

The bundle $E_{\tan}$ is orthogonal to $E_1 \oplus E_2$, but E_1 is not necessarily orthogonal to E_2. By definition $(E_{\tan} \oplus E_2)|_{M_{p_+} \cap W}$ is the tangent bundle to $M_{p_+} \cap W$, and $(E_{\tan} \oplus E_1)|_{M_{\mathrm{fng}} \cap W}$ is the tangent bundle to $M_{\mathrm{fng}} \cap W$. Moreover, $E_{\tan}$ restricted to $M_{p_+} \cap M_{\mathrm{fng}}$ is the tangent bundle to $M_{p_+} \cap M_{\mathrm{fng}}$.

According to the decomposition $T\Omega = E_{\tan} \oplus E_1 \oplus E_2$ we write $\xi = \xi_{\tan} + \xi_1 + \xi_2$ and $\widetilde{\xi} = \widetilde{\xi}_{\tan} + \widetilde{\xi}_1 + \widetilde{\xi}_2$ on W. Notice that $\widetilde{\xi} - \xi$ is parallel to $\frac{\partial}{\partial y_2}$, which belongs to $E_1 \oplus E_2$, hence $\widetilde{\xi}_{\tan} = \xi_{\tan}$ in W. Set

$$\Upsilon = \left(\frac{\rho_{\mathrm{fng}}}{\rho_{\mathrm{fng}} + \rho_+}\right)^{\phi_3}.$$

Then $\Upsilon \equiv 1$ on C_{safe}. Also $\Upsilon \equiv 0$ on the whole of M_{fng} and $\Upsilon \equiv 1$ on the whole of M_{p_+} ($\rho_+ = 0$ there); see Figure 45. The function Υ is not well-defined on $M_{\mathrm{fng}} \cap M_{p_+}$, but, as we will see below, this does not affect the correctness of the definition of $\widehat{\xi}$. The interpolated vector field is given by:

(5.3.12) $$\widehat{\xi} = \xi^a + \xi^b + \xi^c + \xi^d,$$

where:

$$\xi^a = (1 - \phi_1)\xi \qquad \xi^b = \phi_1\left(\Upsilon\xi_{\tan} + (1 - \Upsilon)\widetilde{\xi}_{\tan}\right)$$
$$\xi^c = \phi_1(1 - \phi_2)\left(\Upsilon(\xi_1 + \xi_2) + (1 - \Upsilon)(\widetilde{\xi}_1 + \widetilde{\xi}_2)\right) \qquad \xi^d = \phi_1\phi_2\left(\rho_+\widetilde{\xi} + \rho_{\mathrm{fng}}\xi\right).$$

The form of $\widehat{\xi}$ in different regions is given in Lemma 5.3.8 and Lemma 5.3.10.

Remark. Notice that $\xi_{\tan}$, $\widetilde{\xi}_{\tan}$, and $\xi_i, \widetilde{\xi}_i$ for $i = 1, 2$, are defined only on W. However, outside W the vector field $\widehat{\xi}$ depends only on the sum $\xi_{\tan} + \xi_1 + \xi_2$ and $\widetilde{\xi}_{\tan} + \widetilde{\xi}_1 + \widetilde{\xi}_2$. Therefore, to define $\widehat{\xi}$ everywhere on V_{mid}, extend ξ_i, $\widetilde{\xi}_i$ over V_{mid} in such a way that $\xi_{\tan} + \xi_1 + \xi_2 = \xi$ and $\widetilde{\xi}_{\tan} + \widetilde{\xi}_1 + \widetilde{\xi}_2 = \widetilde{\xi}$, and then $\widehat{\xi}$ will not depend on the extension.

One can think of $\widehat{\xi}$ on M_{p_+} as taking the vector field ξ and slowing it in the transverse direction to M_{p_+} as the trajectory draws near M_{p_+}.

In the following series of lemmas we explain the formulae for $\widehat{\xi}$ and show that $\widehat{\xi}$ is tangent to M_{fng} and to M_{p_+}.

Lemma 5.3.7. *If $w \notin V_{\text{mid}}$, then $\widehat{\xi}(w) = \xi(w)$.*

Proof. We have $\phi_1(w) = 0$, so $\xi^a = \xi$ and $\xi^b = \xi^c = \xi^d = 0$. □

Lemma 5.3.8. *If $w \in V_{\text{mid}} \setminus V_{\text{inn}}$, then $\widehat{\xi}$ is a linear combination of ξ and $\widetilde{\xi}$; more precisely*

$$\widehat{\xi} = ((1 - \phi_1) + \phi_1 \Upsilon)\xi + \phi_1(1 - \Upsilon)\widetilde{\xi}. \tag{5.3.13}$$

Proof. This follows from the fact that $\phi_2 \equiv 0$ on $V_{\text{mid}} \setminus V_{\text{inn}}$. □

Lemma 5.3.9. *If $w \in V_{\text{mid}} \setminus V_{\text{inn}}$ and $w \in M_{\text{fng}}$, then $\widehat{\xi}$ is tangent to M_{fng}; more precisely $\widehat{\xi} = \xi$ on $M_{\text{fng}} \cap (V_{\text{mid}} \setminus V_{\text{inn}})$.*

Proof. On $M_{\text{fng}} \setminus V_{\text{inn}}$ we have $x_2^2 + \dots + x_n^2 \geq \varepsilon_R^2$. Therefore, by the definition of $\widetilde{\xi}$ we have $\widetilde{\xi} = \xi$. We conclude by (5.3.13). □

Lemma 5.3.10. *Suppose $w \in V_{\text{inn}} \setminus W$. Then $\widehat{\xi}(w) = \Upsilon\xi(w) + (1 - \Upsilon)\widetilde{\xi}(w)$. In particular, $\widehat{\xi}(w) = \xi(w)$ if $w \in M_{p_+}$ and $\widehat{\xi}(w) = \widetilde{\xi}(w)$ if $w \in M_{\text{fng}}$.*

Proof. We have $\phi_1 \equiv 1$ and $\phi_2 \equiv 0$, hence $\xi^a = \xi^d = 0$ and so

$$\widehat{\xi} = \xi^b + \xi^c = \Upsilon\xi + (1 - \Upsilon)\widetilde{\xi}.$$

The statement follows, since $\Upsilon(w) = 0$ if $w \in M_{\text{fng}}$ and $\Upsilon(w) = 1$ if $w \in M_{p_+}$. □

As $C_{\text{safe}} \subseteq V_{\text{inn}} \setminus W$ and $\Upsilon \equiv 1$ on C_{safe} we obtain the following corollary.

Corollary 5.3.11. *On C_{safe} we have $\widehat{\xi} = \xi$.*

Lemma 5.3.12. *Suppose $w \in W \setminus W'$. Then $\widehat{\xi}$ is a nonnegative linear combination of ξ, $\widetilde{\xi}$ and $\xi_{\tan}$, more precisely*

$$\widehat{\xi} = (1 - \phi_2)(\Upsilon\xi + (1 - \Upsilon)\widetilde{\xi}) + \phi_2(\rho_+\widetilde{\xi} + \rho_{\text{fng}}\xi + \xi_{\tan}). \tag{5.3.14}$$

Proof. We use the fact that $\xi_{\tan} = \widetilde{\xi}_{\tan}$ in W. □

Lemma 5.3.13. *Suppose $w \in W \setminus W'$. If $w \in M_{p_+}$, then $\widehat{\xi}(w)$ is tangent to M_{p_+}, and if $w \in M_{\text{fng}}$, then $\widehat{\xi}(w)$ is tangent to M_{fng}.*

Proof. This follows from (5.3.14). If $w \in M_{p_+}$, then $\Upsilon(w) = 1$ and $\rho_+(w) = 0$, while if $w \in M_{\text{fng}}$, then $\Upsilon(w) = 0$ and $\rho_{\text{fng}}(w) = 0$. The vector field $\xi_{\tan}$ is tangent both to M_{p_+} and to M_{fng}. □

Lemma 5.3.14. *If $w \in W'$ then $\xi(w)$ is still given by (5.3.14), but the formula simplifies further to*

$$\widehat{\xi}(w) = \xi_{\tan}(w) + \rho_+\widetilde{\xi}_1(w) + \rho_{\text{fng}}\xi_2(w). \tag{5.3.15}$$

In particular $\widehat{\xi}$ is well-defined (even if Υ is not) and $\widehat{\xi}$ is tangent to M_{fng}, M_{p_+} and $M_{\text{fng}} \cap M_{p_+}$.

Proof. We have $\phi_1(w) = \phi_2(w) = 1$, so $\xi^a = \xi^c = 0$. Moreover $\xi_{\tan} = \widetilde{\xi}_{\tan}$, hence $\xi^b = \xi_{\tan}$. The formula for ξ^d simplifies to $\xi^d = \rho_+\widetilde{\xi}_1 + \rho_{\text{fng}}\xi_2$. Hence we obtain (5.3.15) and from this formula the tangency of $\widehat{\xi}$ follows immediately. □

Combining the above lemmas we obtain the following corollary.

Corollary 5.3.15. *The vector field $\widehat{\xi}$ is tangent to M_{p_+}, M_{fng} and $M_{p_+} \cap M_{\text{fng}}$.*

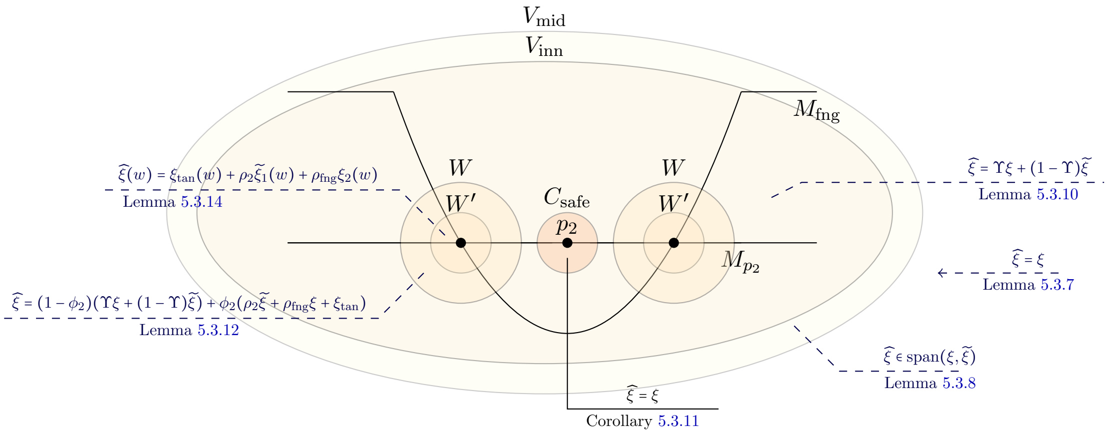

FIGURE 46. Graphical presentation of the statements of Lemmas 5.3.7 through 5.3.14.

5.3.3. **Positivity of $\widehat{\xi}$.** We pass to the study of the critical points of $\widehat{\xi}$.

Proposition 5.3.16.

(a) *The vector field $\widehat{\xi}$ satisfies $\partial_{\widehat{\xi}} F \geq 0$.*

(b) *Away from $M_{\mathrm{fng}} \cap M_{p_+}$, the vector field $\widehat{\xi}$ has the same critical points as ξ and they have the same local behaviour.*

Moreover on $M_{\mathrm{fng}} \cap M_{p_+}$, $\widehat{\xi}$ vanishes at points where $x_2 = \cdots = x_{h+1} = 0$ or $x_{h+2} = \cdots = x_n = 0$.

Proof. By (5.3.9), $\partial_{\widehat{\xi}} F \geq 0$. Inside W, $\widehat{\xi}$ is a linear combination of ξ, $\widetilde{\xi}$ and $\xi_{\tan}$ (see Lemmas 5.3.12 and 5.3.14) with nonnegative coefficients. Therefore, in order to show that $\partial_{\widehat{\xi}} F \geq 0$, it is enough to show that $\partial_{\xi_{\tan}} F \geq 0$.

Recall from (5.3.10) that e_1 is the vector field $(0, x_2, \dots, x_n, 0, 0)$ in V_{inn}. The projection of ξ to the space orthogonal to $\frac{\partial}{\partial x_1}$, $\frac{\partial}{\partial y_1}$ and $\frac{\partial}{\partial y_2}$ is equal to the vector field $\xi_{2:n}$ defined as

$$\xi_{2:n} := (0, -x_2, \dots, -x_{h+1}, x_{h+2}, \dots, x_n, 0, 0). \tag{5.3.16}$$

Then

$$\xi_{\tan} = \xi_{2:n} - \frac{\langle \xi_{2:n}, e_1 \rangle}{\langle e_1, e_1 \rangle} e_1. \tag{5.3.17}$$

Set $\Delta = -x_2^2 - \cdots - x_{h+1}^2 + x_{h+2}^2 + \cdots + x_n^2$. Then the following holds.

$$\langle \xi_{2:n}, e_1 \rangle = \Delta \qquad \langle e_1, e_1 \rangle = x_2^2 + \cdots + x_n^2 \qquad \frac{1}{2}\partial_{e_1} F = \Delta \qquad \frac{1}{2}\partial_{\xi_{2:n}} F = x_2^2 + \cdots + x_n^2.$$

Therefore

$$\frac{1}{2}\partial_{\xi_{\tan}} F = x_2^2 + \cdots + x_n^2 - \frac{\Delta^2}{x_2^2 + \cdots + x_n^2} \geq 0. \tag{5.3.18}$$

The last inequality holds because $|\Delta| \leq x_2^2 + \dots + x_n^2$. This concludes the proof that $\partial_{\widehat{\xi}} F \geq 0$ in W, and so proves point (a) of the proposition.

We pass to the study of the critical points of $\widehat{\xi}$. Firstly, outside W the vector field $\widehat{\xi}$ is a linear combination of ξ and $\widetilde{\xi}$ with coefficients summing to 1. Now $\widetilde{\xi} - \xi$ is parallel to $\frac{\partial}{\partial y_2}$, so it is linearly independent of ξ. Thus, for $w \notin W$, $\widehat{\xi}(w) = 0$ only if $\xi(w) = 0$, that is, only if w

is a critical point of ξ. By construction, $\widehat{\xi} = \xi$ near each critical point of ξ. This shows point (b) outside W.

Now we analyse the critical points of $\widehat{\xi}$ inside W. First, by (5.3.9), $\partial_\xi F(w) = \partial_{\widetilde{\xi}} F(w)$ on $V_{\text{out}} \supseteq W$. If $w \notin M_{p_+} \cap M_{\text{fng}}$, then in (5.3.14) (which holds also inside W' by Lemma 5.3.14), either the coefficient at ξ or the coefficient at $\widetilde{\xi}$ is nonzero. Moreover, $\partial_\xi F(w) = \partial_{\widetilde{\xi}} F(w) > 0$, because within $F^{-1}([c_{\text{bot}}, c_{\text{top}}])$, we have $\partial_\xi F = 0$ only at the point p_+, which does not belong to W. Since we have already shown that $\partial_{\xi_{\tan}} F \geq 0$, we conclude, using that all coefficients in (5.3.12) are positive, that $\partial_{\widehat{\xi}} F(w) > 0$, and so $\widehat{\xi}(w) \neq 0$. This concludes the proof of point (b) inside W. □

5.3.4. **The path F_τ.** In this subsection we construct the path F_τ and the vector field $\overline{\xi}$. We also prove property (FM-1) of Finger Move Theorem 4.4.13.

The function F is – in general – not an immersed Morse function for M_{new}. More precisely we have the following result.

Lemma 5.3.17. *Recall that $\Sigma = M_{p_+} \cap M_{\text{fng}}$. The function F restricted to Σ has critical points on the whole subset $\Sigma \cap \{x_2 = \dots = x_{h+1} = 0\}$ and $\Sigma \cap \{x_{h+2} = \dots = x_n = 0\}$.*

Proof. By (5.3.5) the set Σ is given as $x_2^2 + \dots + x_n^2 = \varepsilon_\Sigma^2$, $x_1 = y_1 = 0$. On this set, by property (V-3), the function F is given by $-x_2^2 - \dots - x_{h+1}^2 + x_{h+2}^2 + \dots + x_n^2 + c_+$. Then F is constant on the intersection of Σ with $\{x_2 = \dots = x_{h+1} = 0\}$ and $\{x_{h+2} = \dots = x_n = 0\}$. □

Our goal is to construct an explicit perturbation of F and of $\widehat{\xi}$ inside $W' \subseteq W$. To this end choose a real vector $(\mu_2, \dots, \mu_n)$ with $|\mu_i| < \frac{1}{n}$ for each i. On our coordinate chart in V_{out}, define $L\colon V_{\text{out}} \to \mathbb{R}$ and a vector field ξ^μ by:

$$L := 2\sum_{j=2}^n \mu_j x_j, \quad \xi^\mu := \sum_{j=2}^n \mu_j \frac{\partial}{\partial x_j}.$$

Let $\xi^\mu_{\tan}$ be the projection of the vector field ξ^μ to the sub-bundle orthogonal to e_1. Decrease the coefficients μ_j so that $|L(z)| < 1$ for all $z \in W$. Recall that $\rho_2\colon V_{\text{mid}} \to \mathbb{R}_{\geq 0}$ is the square of the distance to M_{p_+} and $\rho_{\text{fng}}\colon V_{\text{mid}} \to \mathbb{R}_{\geq 0}$ is the square of the distance to M_{fng}. For $\varepsilon_\lambda > 0$ sufficiently small define the set U_λ by the inequalities $\{\rho_{\text{fng}} \leq \varepsilon_\lambda, \rho_2 \leq \varepsilon_\lambda\}$. If $\varepsilon_\lambda \ll 1$, we have $U_\lambda \subseteq W'$. From now on we assume that this is the case. Let also $U'_\lambda \subseteq U_\lambda$ be given by $\{\rho_{\text{fng}} \leq \frac{1}{2}\varepsilon_\lambda, \rho_2 \leq \frac{1}{2}\varepsilon_\lambda\}$. The choice of the factor $\frac{1}{2}$ in the definition of U'_λ is arbitrary.

Theorem 5.3.18. *There exists $\varepsilon_\lambda > 0$, a smooth function $\phi_\lambda\colon \Omega \to [0,1]$ vanishing outside U_λ and equal to 1 in U'_λ, and a parameter $\sigma_0 > 0$ such that for any $\sigma \in (0, \sigma_0]$ the function*

$$F - \sigma\phi_\lambda L$$

is an immersed Morse function for $G_2(N)$. It has at most four critical points when restricted to Σ. Moreover, the vector field

$$\widehat{\xi}_\sigma := \widehat{\xi} - \sigma\phi_\lambda \xi^\mu_{\tan}$$

is a grim vector field for $F - \sigma\phi_\lambda L$.

Proof. Fix $\varepsilon_\lambda > 0$ such that $U_\lambda \subseteq W'$. Choose $\phi_\lambda\colon \Omega \to [0,1]$ in such a way that

$$\|\nabla \phi_\lambda\| \leq 4/\varepsilon_\lambda. \tag{5.3.19}$$

That this can be done is easy to believe if one notices that it is essentially equivalent to approximating a piecewise linear function $f\colon [0,2] \to [0,1]$ with $f(x) = 1$ for $x \leq 1/2$ and

$f(x) = 0$ for $x \geq 1$ by a smooth function with derivative $|f'(x)| < 4$ for all $x \in [0,2]$. Choose c, K_ξ, and K_F to be real positive numbers satisfying

$$\partial_\xi F(z) > c, \ \|\widehat{\xi}(z)\| < K_\xi, \text{ and } \|\nabla F(z)\| < K_F,$$

for all $z \in \overline{W'}$. Such c exists, because $\partial_\xi F$ is positive on the whole of $\overline{W'}$.

Our discussion takes place entirely in W'. By (5.3.15) we have $\widehat{\xi} = \xi_{\tan} + \rho_2 \xi + \rho_{\text{fng}} \widetilde{\xi}$. By (5.3.18) and (5.3.9), $\partial_{\widehat{\xi}} F \geq (\rho_2 + \rho_{\text{fng}} \partial_\xi) F$. Hence,

$$\partial_{\widehat{\xi}_\sigma}(F - \sigma\phi_\lambda L) \geq (\rho_{\text{fng}} + \rho_2)\partial_\xi F - \sigma \partial_{\widehat{\xi}}(\phi_\lambda L) - \sigma\phi_\lambda \partial_{\xi^\mu_{\tan}}(F - \sigma\phi_\lambda L).$$

We first show that this quantity is nonzero on $U_\lambda \smallsetminus U_{\lambda'}$.

In $U_\lambda \smallsetminus U'_\lambda$ we have $\rho_2 > \frac{1}{2}\varepsilon_\lambda$ or $\rho_{\text{fng}} > \frac{1}{2}\varepsilon_\lambda$, $\partial_\xi F = \partial_{\widetilde{\xi}} F > c$, and $\partial_{\xi_{\tan}} F \geq 0$ by (5.3.18). Therefore

$$\partial_{\widehat{\xi}} F > \frac{1}{2} c \varepsilon_\lambda. \tag{5.3.20}$$

Now observe that $\partial_{\widehat{\xi}} \phi_\lambda \leq \|\widehat{\xi}\| \, \|\nabla \phi_\lambda\| \leq 4K_\xi \varepsilon_\lambda^{-1}$. Hence, using that $|L| < 1$, $\|\nabla L\| < 1$, and that ϕ_λ is valued in $[0,1]$, we have:

$$|\partial_{\widehat{\xi}} \phi_\lambda L| = |\phi_\lambda \partial_{\widehat{\xi}} L + L \partial_{\widehat{\xi}} \phi_\lambda| \leq K_\xi (1 + 4\varepsilon_\lambda^{-1}). \tag{5.3.21}$$

The vector field ξ^μ has length at most 1 since $|\mu_i| < \frac{1}{n}$, so $\xi^\mu_{\tan}$ also has length at most 1 as well, being an orthogonal projection. Therefore we have

$$|\partial_{\xi^\mu_{\tan}} F| \leq \|\nabla F\| \cdot \|\xi^\mu_{\tan}\| \leq K_F. \tag{5.3.22}$$

In addition, also using that $|L| < 1$, $\|\nabla L\| < 1$ on U_λ, and that ϕ_λ is valued in $[0,1]$:

$$|\partial_{\xi^\mu_{\tan}} \phi_\lambda L| = |L \partial_{\xi^\mu_{\tan}} \phi_\lambda + \phi_\lambda \partial_{\xi^\mu_{\tan}} L| \leq 4\varepsilon_\lambda^{-1} + 1. \tag{5.3.23}$$

Combining (5.3.20), (5.3.21), (5.3.22) and (5.3.23) we obtain

$$\partial_{\widehat{\xi}_\sigma}(F - \sigma\phi_\lambda L) \geq \frac{1}{2} c\varepsilon_\lambda - \sigma K_\xi (1 + 4\varepsilon_\lambda^{-1}) - \sigma K_F - \sigma^2(1 + 4\varepsilon^{-1} + 1))$$

In the above inequality, only the first term does not depend on σ. Therefore, there exists $\sigma_0 > 0$ such that if $|\sigma| < \sigma_0$, then $\partial_{\widehat{\xi}_\sigma}(F - \sigma\phi_\lambda L)$ is positive everywhere on $U_\lambda \smallsetminus U'_\lambda$.

We now estimate $\partial_{\xi_\sigma}(F - \sigma\phi_\lambda L)$ inside $U_{\lambda'}$. Then $\phi_\lambda \equiv 1$ so we have

$$\widehat{\xi}_\sigma = \widehat{\xi} - \sigma \xi^\mu_{\tan}.$$

Let $\xi'_\sigma = \xi_{\tan} - \sigma\xi^\mu_{\tan}$ so that $\widehat{\xi}_\sigma = \xi'_\sigma + \rho_2\xi + \rho_{\text{fng}}\widetilde{\xi}$. As $\partial_\xi F > c$, for σ sufficiently small, we have $\partial_\xi(F - \sigma L), \partial_{\widetilde{\xi}}(F - \sigma L) > 0$. We need to show that $\partial_{\xi'_\sigma}(F - \sigma L) \geq 0$.

Set $H_\sigma = (F - \sigma L) + x_1^2 - y_1$. By definition, H_σ does not depend on y_1 and x_1, because these terms in the definition of H_σ cancel with the corresponding terms in F; compare (V-3). By (V-1), $\partial_\xi(F - \sigma L - H_\sigma) = \partial_{\widetilde{\xi}}(F - \sigma L - H_\sigma) \geq 0$ and $\partial_{\xi_{\tan}}(F - \sigma L - H_\sigma) = \partial_{\xi^\mu_{\tan}}(F - \sigma L - H_\sigma) = 0$. Therefore we strive to show that $\partial_{\xi'_\sigma} H_\sigma \geq 0$, and this will imply that $\partial_{\xi'_\sigma}(F - \sigma L) \geq 0$.

Note that $\nabla H_\sigma = 2\xi_{2:n} - 2\sigma\xi_\mu$. The orthogonal projection of $\xi_{2:n} - \sigma\xi_\mu$ along e_1 is equal precisely $\xi_{\tan} - \sigma\xi^\mu_{\tan} = \xi'_\sigma$. This shows that

$$\partial_{\xi'_\sigma} H_\sigma \geq 0. \tag{5.3.24}$$

As alluded to above, this implies that $\partial_{\widehat{\xi}_\sigma}(F - \sigma L) \geq 0$ on U'_λ.

The remaining part of the proof is devoted to checking when $\partial_{\widehat{\xi}_\sigma}(F - \sigma L) = 0$. First of all, if $\rho_{\mathrm{fng}} > 0$ or $\rho_2 > 0$, then $\partial_{\widehat{\xi}_\sigma}(F - \sigma L) > 0$, because $\partial_{\xi'_\sigma}(F - \sigma L) \geq 0$. Therefore, if $\partial_{\widehat{\xi}_\sigma}(F - \sigma L)(z) = 0$, then $\rho_{\mathrm{fng}}(z) = \rho_2(z) = 0$, that is, $z \in \Sigma$.

Next, if $\partial_{\widehat{\xi}_\sigma}(F - \sigma L)(z) = 0$, we need to have equality in (5.3.24). This is possible only if $2\xi_{2:n} - 2\sigma\xi_\mu$ is parallel to e_1. For $2\xi_{2:n} - 2\sigma\xi_\mu$ to be parallel to e_1 we have, for some $\kappa \in \mathbb{R}$, that $(-1-\kappa)x_i = \mu_i$ for $2 \leq i \leq h$ and $(1-\kappa)x_i = \mu_i$ for $h+1 \leq i \leq n$. Solving for x_i and substituting into $x_2^2 + \cdots + x_n^2 = \varepsilon_\Sigma^2$ yields a quartic equation for κ. Thus there are at most 4 solutions. □

Given Theorem 5.3.18 we choose $\sigma_1 < \sigma_0$. We define

$$\overline{\xi} = \widehat{\xi}_{\sigma_1}. \tag{5.3.25}$$

We also define the path F_τ that will satisfy the requirements of Finger Move Theorem 4.4.13. For each $\tau \in [0,1]$ we set

$$F_\tau := F - \tau\sigma_1\phi_\lambda L.$$

Remark. We leave ourselves the possibility of decreasing the value of σ_1 later on. The precise order of choosing different variables is explained in Remark 5.4.2.

5.4. Property (FM-3)

Our next result proves property (FM-3) from the Finger Move Theorem 4.4.13. Our result is valid both for F_1 and the grim vector field $\overline{\xi}$, as well as for the original function F and the unperturbed vector field $\widehat{\xi}$.

Recall that $\mathbb{M}_{\mathrm{a}}(p_-) \cap F^{-1}(c_{\mathrm{bot}}) \cap U_{\mathrm{out}}$ consists of discs $D_0, D_1, \ldots, D_r$ containing points $z_0, z_1, \ldots, z_r \in \mathbb{M}_{\mathrm{a}}(p_-) \cap \mathbb{M}_{\mathrm{d}}(p_+) \cap F^{-1}(c_{\mathrm{bot}})$, and that we seek to remove $z_m \in D_m$. Recall that Y_{mid} is the interior of $\overline{V}_{\mathrm{mid}} \cap F^{-1}(c_{\mathrm{bot}})$ and that $V_{\mathrm{mid}} \subseteq F^{-1}(c_{\mathrm{bot}}, c_{\mathrm{top}})$.

Theorem 5.4.1. *There are no trajectories of $\overline{\xi}$ from p_- to p_+ that go through Y_{mid}. If $\varepsilon_{\mathrm{fng}}$ is very small, then the only trajectories from p_- to p_+ are those trajectories of ξ that go through discs $D_0, D_1, \ldots, D_r$ (see (5.1.2)) for $r \neq m$.*

In other words, the number of trajectories of ξ from p_- to p_+ is one less than the number of trajectories of $\overline{\xi}$ from p_- to p_+.

Proof. Suppose γ is a trajectory of $\overline{\xi}$ connecting p_- with p_+. Assume γ does not go through Y_{mid}. As below the level set c_{bot}, $\xi = \overline{\xi}$, γ is a trajectory of ξ at least until it reaches the level set c_{bot}. Let w_1 be the intersection of γ with $F^{-1}(c_{\mathrm{bot}})$. Clearly $w_1 \in \mathbb{M}_{\mathrm{a}}(p_-)$.

By (V-4), and since $\xi = \overline{\xi}$ on $\partial V_{\mathrm{out}}$, a trajectory of $\overline{\xi}$ can enter V_{out} only through Y_{out}. Hence, $w_1 \in Y_{\mathrm{out}}$. As $\mathbb{M}_{\mathrm{a}}(p_-) \cap Y_{\mathrm{out}}$ is the union of discs $D_0, \ldots, D_r$, we note that $w_1 \in D_i$ for some i.

We consider now two cases. Either $i = m$, or $i \neq m$.

We begin with the first case, that is, $i = m$. Our aim is to show that this case is not possible. The set $P := \{x_1 = 0\}$ is easily seen to be invariant under the flow of $\overline{\xi}$, moreover $D_m \subseteq \{x_1 = 0\}$. As the trajectory hits the point $w_1 \in P$, the trajectory stays on P until reaching p_+, compare Lemma 5.2.2.

Consider the subsets

$$P_\pm = \{\pm\big(y_2 - \mathfrak{a}(y_1)\min\{x_2^2 + \cdots + x_n^2, \varepsilon_R^2\} + \mathfrak{b}(y_1)\big) \geq 0\}.$$

Note that we can also write $P_\pm$ as $\{\pm(y_2 - \mathfrak{s}) \geq 0\}$, where $\mathfrak{s}$ is the function from (5.3.3). We have $P_+ \cup P_- = P$ and $P_+ \cap P_- \subseteq M_{\mathrm{fng}}$, see Figure 47. As $\overline{\xi}$ is tangent to M_{fng}, there can be

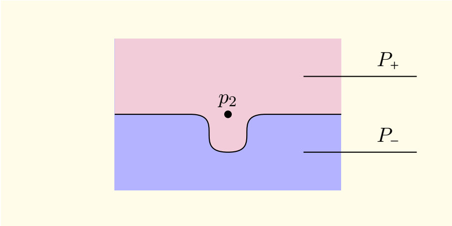

FIGURE 47. The sets P_+ and P_- from the proof of Theorem 5.4.1. The trajectories of $\overline{\xi}$ that enter through P_- will never reach p_+.

no trajectory of $\overline{\xi}$ that goes from P_- to P_+ and does not belong to M_{fng}. By construction, $p_+ \in P_+$.

Recall that $w_1 \in D_m$ is the point of entrance of γ to V_{mid}. Denote by $(x_1, x_2, \dots, x_n, y_1, y_2)$ its coordinates. We know that $y_2 < 1$ and $x_1 = \dots = x_{h+1} = 0$, compare (5.1.2). On the other hand, we have $F(w_1) = c_{\mathrm{bot}}$. By (V-3), $y_1 + x_{h+2}^2 + \dots + x_n^2 = c_{\mathrm{bot}} - c_+$, so $y_1 \le c_{\mathrm{bot}} - c_+$. Here, recall that $c_+ = F(p_+)$. It follows that $y_1 \notin [c_{-1/4} - c_+, c_{+1/4} - c_+]$. Property (A-1) implies that $\mathfrak{b}(y_1) = 1$, so it follows from (5.3.3) that $\mathfrak{s}(w_1) = 1$. As $y_2 < 1$, we infer that $w_1 \in P_-$. Hence, γ does not reach p_+.

We pass to the proof of the second part, that is, $w_1 \in D_i$ for $i \neq m$. As V_{mid} is disjoint from D_i by Lemma 5.2.1, the trajectory γ must enter V_{mid} (the point p_+ belongs to the interior of V_{mid}). Let $w_2 \in \partial V_{\mathrm{mid}}$ be the point of the entrance. The trajectory γ is a trajectory of ξ until it reaches w_2. Hence, we can use Lemmas 5.2.3 and 5.2.5. We have the following cases.

- The point w_2 is outside C_{vs}. By Lemma 5.2.3, $F(w_2) > c_+ = F(p_+)$. As $\partial_{\overline{\xi}} F \ge 0$, such a trajectory will never reach p_+.
- The trajectory hits V_{mid} at $w_2 \in C_{\mathrm{vs}}$ and exits C_{safe} at a point w_3. By Lemma 5.2.5 we have that $F(w_3) > c_+$, so the trajectory will never reach p_+ again.
- The trajectory enters C_{vs} and does not lead out of C_{safe}. Then for all the time it is a trajectory of ξ, because $\xi = \overline{\xi}$ in C_{safe}.

All this means that every trajectory of $\overline{\xi}$ that connects p_- and p_+ and passes through D_i with $i \neq m$ is actually a trajectory of ξ.

Thus we have indeed reduced the number of trajectories by 1, namely, by eliminating the trajectory passing through D_m. □

We conclude the section with an important remark.

Remark 5.4.2. The order of choosing and adjusting various constants, after V_{mid} has been constructed, is the following.

- Adjust $\varepsilon_{\mathrm{fng}}$ according to Lemma 5.2.3; see Condition 5.2.6.
- Choose ε_R in such a way that Lemma 5.2.7 is satisfied; see Condition 5.2.8;
- The functions $\mathfrak{a}$, $\mathfrak{b}$ are adjusted to ε_R as in Condition 5.3.2.
- The subsets W and W' are defined as neighbourhoods of the double point set (5.3.5). They are chosen to be disjoint from p_+ and contained in V_{mid}.
- Take a $\varepsilon_{\mathrm{safe}} > 0$ such that $C_{\mathrm{safe}} \subseteq V_{\mathrm{mid}}$ is disjoint from W. This choice is made in Condition 5.3.6.
- Adjust $\varepsilon_{\mathrm{vs}}$ according to Lemmas 5.2.3 and 5.2.5.
- Adjust the value $\sigma_1 < \sigma_0$ in the definition of $\overline{\xi}$ and F_τ in (5.3.25). The adjustment is used in Section 5.5 while proving property (FM-4) of Finger Move Theorem 4.4.13.

5.5. Proof of properties (FM-4) and (FM-5)

We recall the statements of (FM-4) and (FM-5).

(FM-4) Set $f_\tau = F_\tau \circ G_\tau$. The path $f_{1+\tau}$ is a constant path for $\tau \in [0,1]$. Moreover, η is a gradient-like vector field for all f_τ.

(FM-5) There is a pull-back $\overline{\eta}$ of $\overline{\xi}$ via G_2 such that the grim paths of death starting from $F_2 \circ G_2$ and constructed with η and $\overline{\eta}$ are left-homotopic.

We begin with the following result.

Lemma 5.5.1. *For any $u \in N$ and for any $\tau \in [0,2]$ we have $F \circ G_\tau(u) = F \circ G_0(u)$.*

Proof. We use (5.3.1). Note that if $u \notin G^{-1}(M_{\mathrm{orig}})$, then $G_\tau(u) = G(u)$, so there is nothing to prove. If $G(u) = (x_1, \dots, y_2) \in M_{\mathrm{orig}}$, the finger move affects only the y_2 coordinate of $G(u)$, but F does not depend on the y_2 coordinate of a point in V_{out}. □

Corollary 5.5.2. *The vector field η is a gradient-like vector field for $F \circ G_\tau$ for all τ.*

Next we pass to the function $F_2 \circ G_2$. Recall that the path F_τ was constructed in Theorem 5.3.18. The size of perturbation was controlled by a parameter $\sigma_1 < \sigma_0$, compare (5.3.25).

Lemma 5.5.3. *Suppose σ_1 is sufficiently small. Then η is a gradient-like vector field for $F_\tau \circ G_2$, for all $\tau \in [0,2]$.*

Proof. We use notation from Subsection 5.3.4. We have $F_\tau = F - \tau\sigma_1\phi_\lambda L$. Let $Z = G_2^{-1}(U_\lambda)$. As ϕ_λ has support on U_λ, it is enough to prove the statement on Z. Note that Z is separated from the critical set of $F \circ G_2$. As η is gradient-like for $F \circ G_2$, there exists $c > 0$ such that $\partial_\eta(F \circ G_2)(u) > c$ for all $u \in Z$. Denote by K_ϕ the supremum of $|\partial_\eta(\phi_\lambda L \circ G)|$ on Z. Choose $\sigma_1 < c/K_\phi$. For all $\sigma < \sigma_1$ we have $\partial_\eta F_\tau \circ G_2(u) > 0$, if $u \in Z$. □

Now define $\overline{\eta}$ to be a pull-back of $\overline{\xi}$ via G_2 in the sense of Section 2.5.2. As η itself is a pull-back of ξ via G, we may and will assume that $\overline{\eta}(u) = \eta(u)$ if $G_2(u) = G(u)$ and $\xi(G(u)) = \overline{\xi}(G(u))$. This means that $\overline{\eta}(u) = \eta(u)$ away from $G_2^{-1}(V_{\mathrm{out}} \smallsetminus C_{\mathrm{safe}})$.

By construction, $\overline{\eta}$ is a gradient-like vector field for $f_2 = F_2 \circ G_2$. For $\tau \in [0,1]$, set:

$$\eta_\tau = (1-\tau)\eta + \tau\overline{\eta}.$$

Lemma 5.5.4. *The vector field η_τ is a gradient-like vector field for $f_2 \colon N \to \mathbb{R}$.*

Proof. By construction, η and $\overline{\eta}$ agree away from the preimage of $G_2^{-1}(V_{\mathrm{out}} \smallsetminus C_{\mathrm{safe}})$. In particular, η and $\overline{\eta}$ agree near all critical point of f_2. Away from these points, $\partial_\eta f_2 > 0$ by Corollary 5.5.2, and $\partial_{\overline{\eta}} f_2 > 0$ as $\overline{\eta}$ is gradient-like for f_2. The statement promptly follows. □

Theorem 5.5.5. *The grim paths of death starting at f_2 constructed by η and η_τ are left-homotopic.*

Proof. Our goal is to eventually use the Uniqueness of Death Lemma 3.6.13. To this end, we need to study the stable and unstable manifolds of q_- and q_+ with respect to η_τ, where $q_- := G^{-1}(p_-)$ and $q_+ := G^{-1}(p_+)$, i.e. the preimages in N of $p_-, p_+ \in \Omega$.

Consider the level set $F^{-1}(c_{\mathrm{bot}})$. Since this is the level set below the finger move is performed, we have that for any $w \in F^{-1}(c_{\mathrm{bot}})$, $F(w) = F_\tau(w)$. Moreover, if $u \in f^{-1}(c_{\mathrm{bot}})$, then $G_\tau(u) = G(u)$ for all τ. Our first step toward the proof of Theorem 5.5.5 is the following result, which relies on Theorem 5.4.1. The subset $W \subseteq V_{\mathrm{inn}}$ was chosen in Section 5.3.2, as a small neighbourhood of the double point set $M_{\mathrm{fng}} \cap M_{p_+}$, with $p_+ \notin W$.

Lemma 5.5.6. *Suppose a trajectory of ξ hits $\overline{W}$ below the level set c_+. If it hits p_- in the infinite past, then it crosses $F^{-1}(c_{\text{bot}})$ at a point of D_m.*

Proof. Any trajectory of ξ entering V_{out} and hitting p_- in the infinite past must cross one of the discs $D_0, \dots, D_r$. Consider such a trajectory and let $w_1 \in D_0 \cup \dots D_r$ be the point of its entrance to V_{out}.

Assume the trajectory reaches W before the level set c_+. Suppose towards contradiction that $w_1 \in D_j$, $j \neq m$. As $W \subseteq V_{\text{mid}}$, the trajectory must enter V_{mid}. Let w_2 be the point it enters V_{mid}. By Lemma 5.2.3, either $F(w_2) > c_+$ and we are done, or $w_2 \in C_{\text{vs}}$. In the latter case, the trajectory has to leave C_{safe} before entering W. By Lemma 5.2.5, such a trajectory leaves C_{safe} above the level set c_+. The contradiction shows that $j = m$. □

The above result has implications for trajectories of η_τ on N. The following corollary explains the behaviour of the unstable manifold of q_- with respect to η_τ.

Corollary 5.5.7. *Suppose a trajectory of η_τ for $\tau \in [0,1]$ hits $G^{-1}(W \cap M_{p_+})$. Then it does not reach q_- in the infinite past.*

Proof. Assume for a contradiction that γ_u is a trajectory of η_τ that starts at q_- and reaches $G^{-1}(W \cap M_{p_+})$. Let $t_0 > 0$ be such that $\gamma(t_0)$ hits $G^{-1}(V_{\text{mid}})$. Then, for $t < t_0$, γ_u is a trajectory of η, because η_τ agrees with η away from $G^{-1}(V_{\text{mid}})$. Next, on $G^{-1}(V_{\text{mid}} \smallsetminus V_{\text{inn}})$, $\eta = \overline{\eta}$, so also $\eta = \eta_\tau$, by Lemma 5.3.9. Similarly, on $G^{-1}(V_{\text{mid}} \smallsetminus W)$ we use Lemma 5.3.10 to show that $\eta = \overline{\eta} = \eta_\tau$. Note that Lemma 5.3.10 cannot be used inside $G^{-1}(W)$, because we use a perturbation of $\widehat{\xi}$ to $\overline{\xi}$. Anyway, we proved that until it reaches $G^{-1}(W)$, the trajectory γ_u is a trajectory of η. The trajectory γ_u between q_- and $G^{-1}(W)$ does not hit any preimage of the second stratum. By Pull-back Lemma 2.5.6, the image $G(\gamma_u)$ is a trajectory of ξ starting at p_- and reaching W. By Lemma 5.5.6, we conclude that $u_0 \in G^{-1}(D_m)$. But D_m is disjoint from M. The contradiction shows that γ_u cannot reach q_- in the infinite past. □

Lemma 5.5.8. *Suppose a trajectory of η_τ passes through $u \in f^{-1}(c_{\text{bot}})$. If this trajectory hits $G^{-1}(M_{\text{orig}})$, then it does not reach q_+ in the infinite future.*

Proof. First we prove this result for $\tau = 0$, that is for $\eta_\tau = \eta$. Since M_{orig} belongs to the subset $\{y_2 \equiv 1\}$, a trajectory of ξ passing through this subset will have $y_2 = 1$, at least until it leaves V_{out}. But any trajectory leaving V_{out} leaves it above the level set $F(p_+)$, so the trajectory will not reach p_+. From this it follows that if a trajectory of η leaves $G^{-1}(M_{\text{orig}})$, it will not reach q_+.

Also note that if a trajectory of ξ leaves M_{orig}, then it will not reach M_{p_+}. This is because M_{orig} has y_2–coordinate equal to 1 and the y_2 coordinate near M_{p_+} is close to zero.

Suppose a trajectory $\gamma_u(t)$ with $\gamma_u(0) = u$ of η_τ passes through $G^{-1}(M_{\text{orig}})$ and reaches q_+. In general, such a trajectory can hit $G^{-1}(M_{\text{orig}})$ infinitely many times, but there has to be $t_{\max}$ such that $\gamma_u(t_{\max}) \in \partial G^{-1}(M_{\text{orig}})$ and $\gamma_u(t) \notin G^{-1}(M_{\text{orig}})$ for $t > t_{max}$. Now from the time moment $t_{\max}$ onwards, γ_u is a trajectory of η: it does not return to $G^{-1}(M_{\text{orig}})$, it does not hit $G^{-1}(M_{p_+})$, and away from $G^{-1}(M_{\text{orig}} \cup M_{p_+})$ we have $\eta = \eta_\tau$, because η was not modified there. We have already shown that no trajectory of η can leave M_{orig} and hit q_+. □

For a fixed $\tau \in [0,1]$, let A_τ be the intersection of the unstable (ascending) manifold of q_- (with respect to η_τ) with the level set $f^{-1}(c_{\text{bot}})$. Let B_τ be the intersection of the stable (descending) manifold of q_+ (with respect to η_τ) with the level set $f^{-1}(c_{\text{bot}})$.

Lemma 5.5.9. *For any τ the submanifolds A_τ and B_τ intersect at a single point. Near this point $A_\tau = A_0$ and $B_\tau = B_0$.*

Proof. Below the level set c_{bot} we have $\eta = \eta_\tau$, hence $A_\tau = A_0$. On B_τ we specify the subset B^0_τ, as the set of points $u \in B_\tau$ such that the trajectory of η_τ through u reaches $G^{-1}(M_{p_+} \cap W)$ in the future. As W is open, B^0_τ is an open subset of B_τ. Let B^1_τ be the complement of B^0_τ in B_τ. By Lemma 5.5.8, if $u \in B^1_\tau$, then the trajectory of u is in fact a trajectory of η. In particular $B^1_\tau = B^1_0$.

By Corollary 5.5.7 the whole intersection of A_τ and B_τ occurs in the interior of B^1_τ. That is, $A_\tau \cap B_\tau = A_\tau \cap B^1_\tau = A_0 \cap B^1_0 = A_0 \cap B_0$ as desired. □

Now we finish the proof of Theorem 5.5.5. Lemma 5.5.9 implies that the assumptions of the Uniqueness of Death Lemma 3.6.13 are satisfied. This concludes the proof of Theorem 5.5.5 and therefore of the last property of Theorem 4.4.13. □

Part 6. Examples and applications

This part gives examples and applications of our theory. We already explained some applications to link homotopy in the introduction.

First we prove the Singular Concordance implies Regular Homotopy Theorem 6.1.1 in Section 6.1. Next, we illustrate the Singular Concordance implies Regular Homotopy Theorem 6.1.1 using some (previously known) low dimensional cases. We also give an application to homotopies between two surfaces in a 4-manifold.

In Section 6.2 we prove Proposition 1.2.3, which states that any classical link with n components is link homotopic to the closure of a braid with n strands. While the result is well-known, our proof illustrates the methods in the proof of the Singular Concordance implies Regular Homotopy Theorem 6.1.1.

Section 6.3 deals with another example. The stevedore knot $6_1 \subseteq S^3$ is slice, that is it bounds a smoothly embedded disc in D^4. For one such disc, the function 'distance to origin' restricted to this disc has two minima and one saddle. It is not possible to cancel a minimum–saddle pair in the embedded case, even though such cancellation is possible restricting to the disc, forgetting about the embedding. To see this, observe that if this cancellation could be achieved, then the stevedore knot would be unknotted. However, after a finger move, the embedded cancellation becomes possible. This example serves as an alternative geometric explanation of the finger move. It is our interpretation of the method sketched by Habegger in [Hab92].

Section 6.5 investigates regular homotopies between surfaces in 4-manifolds, and shows that every regular homotopy can be expressed as a combination of finger moves and Whitney moves, and that all the finger moves can be assumed to happen first. This fact is often quoted in the literature.

6.1. Immersed link concordance implies regular link homotopy

Having the Path Lifting Theorem 4.5.1 at our disposal, we can prove the main result of the article.

Theorem 6.1.1 (Immersed link concordance implies regular link homotopy)**.** *Suppose A is a closed $(n-1)$-dimensional manifold and let Y be a compact $(n+k-1)$-dimensional manifold. Assume that $g_0, g_1 \colon A \to Y$ are concordant generic immersions, via a generic immersion $G \colon A \times [0,1] \to Y \times [0,1]$ such that $G|_{A\times\{0\}} = g_0 \times \{0\}$, $G|_{A\times\{1\}} = g_1 \times \{1\}$.*

(1) *If $k > 2$, and G is an embedding, then there are rel. boundary isotopies of diffeomorphisms H_t of $Y \times [0,1]$ and K_t of $A \times [0,1]$, such that*

$$G' := H_1 \circ G \circ K_1 : A \times [0,1] \to Y \times [0,1]$$

is a level-preserving embedding. In particular g_0 and g_1 are isotopic.

(2) *If $k \geq 2$, then there is a rel. boundary regular homotopy $H_t: A \times [0,1] \to Y \times [0,1]$ with $H_0 = G$ such that H_1 is a level-preserving generic immersion. In particular g_0 and g_1 are regularly homotopic. Moreover, for every pair of connected components of $A \times [0,1]$ that has disjoint image under G, the same holds for H_t, for all $t \in [0,1]$. In particular, if H_0 is an immersed link concordance then H_1 is a regular link homotopy.*

Before we prove Theorem 6.1.1 we recall one more result of Cerf.

Proposition 6.1.2 (see [Cer70, Theorem V.1.1]). *Suppose $f_0, f_1: N \to \mathbb{R}$ are two Morse functions with the property that whenever q_-, q_+ are two critical points of f_0 (respectively f_1), with $\operatorname{ind} q_- < \operatorname{ind} q_+$, then $f_0(q_+) > f_0(q_-)$ respectively $f_1(q_+) > f_1(q_-)$.*

Then there exists a $\mathcal{F}^1$-path of functions f_τ such that for every rearrangement occurring along f_τ, no critical point of higher index goes below a critical point of lower index.

Proof of Theorem 6.1.1. Set $\Omega = Y \times [0,1]$ and let $F: \Omega \to [0,1]$ be the projection onto the first factor. Let $N = A \times [0,1]$. Perturb F if necessary to make it an immersed Morse function with respect to $M = G(N)$, without introducing critical points of F on $\Omega \setminus M$. Set $f: N \to [0,1]$ to be $f = F \circ G$. Then f is a Morse function on N.

If F has no critical points (and no critical points on all strata of M), we conclude that $G(A \times \{0\})$ is isotopic to $G(A \times \{1\})$. In fact, we can identify $F^{-1}(t)$ with Y and under this identification, $M \cap F^{-1}(t)$ is a set A_t isotopic to A_0. If F has no critical points on the zeroth and first stratum of M, we conclude that $A \times \{0\}$ is regularly homotopic to $A \times \{1\}$ by Theorem 2.5.10. The aim is to achieve these conditions on F in cases (1) and (2) respectively.

Now, f is a Morse function on $A \times [0,1]$, equal to 0 on $A \times \{0\}$ and equal to 1 on $A \times \{1\}$. Let f_1 be the Morse function on $A \times [0,1]$ that is a projection onto the second factor. Note that f_1 and f can be made equal on a neighbourhood of $A \times \{0,1\}$. Then we connect f and f_1 by a neat path of functions. By Lemma 3.1.2, we may perturb this path to a neat $\mathcal{F}^1$-path of functions f_τ connecting f to a function f_1.

By Proposition 6.1.2, we may and shall assume that:

- f_τ is an elementary path near each event (birth, death and rearrangement);
- f_τ has no rearrangements such that a critical point of higher index goes below a critical point of lower index.

These are precisely the assumptions for the Path Lifting Theorem 4.5.1. This means that there exists a regular double path (F_τ, G_τ) that is a weak lift of f_τ. Moreover F_τ is a constant path near $F^{-1}(0)$ and $F^{-1}(1)$. The path of immersions G_τ changes the map G by a regular homotopy, but only introduces intersection points within each connected component. If $k \geq 3$ and G is an embedding then G_τ is an ambient isotopy.

The functions F_τ, regarded as ordinary Morse functions on Ω, have no critical points. We invoke Proposition 2.5.10 to conclude that $G(A \times \{0\})$ is regularly homotopic to $G(A \times \{1\})$. The regular homotopy does not create any intersections between different connected components of A. Moreover, if $k > 2$ and G was initially an embedding, then so is G_τ for all τ. Thus M is the trace of an ambient isotopy. ☐

6.2. A 3-dimensional example: the proof of Proposition 1.2.3

We recall the statement of Proposition 1.2.3 for the convenience of the reader.

Proposition 6.2.1. *Every classical link in S^3 with n components is link homotopic to the closure of a braid with n strands.*

Proof. Consider a link in S^3 with n components. Choose a disc D^2 intersecting each of the n components of the link transversely in a single point, and thicken it to obtain $D^2 \times I$ intersecting the link in $\{q_i\} \times I$ for some points $q_i \in D^2$. Remove the interior of $D^2 \times I$ from S^3, and identify the remainder also with $D^2 \times I$. We obtain a concordance $G: \sqcup^n I \hookrightarrow D^2 \times I$ from $\{(q_i, 0)\}_{i=1}^n$ to $\{(q_i, 1)\}_{i=1}^n$, which is called a *string link* in [HL90]. An example is shown in Figure 48. Write G for this string link.

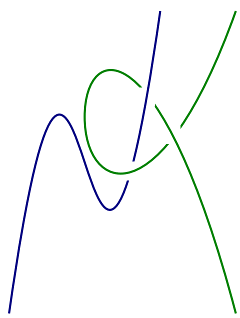

Figure 48. A string link.

After a perturbation, we may assume that the composition $\pi_2 \circ G: n \cdot I \to I = [0,1]$ (where π_2 is a projection onto the second factor) is a Morse function. If this Morse function has no critical points then the string link is a braid, and we are done. We will explain how to perform a link homotopy to reduce the number of local minima by one, and thus simultaneously the number of local maxima by one (which must occur for Euler characteristic reasons). An induction on the number of critical points then completes the proof.

By Theorem 3.2.4, by an isotopy we may arrange that all the local minima occur below the local maxima, as is already the case in Figure 48. Assume that the local minima occur in $D^2 \times [0, 1/2]$, and the local maxima are in $D^2 \times [1/2, 1]$, with $1/2 \in I$ a regular value. Then the disc $L := D^2 \times \{1/2\} \subseteq D^2 \times I$ is an intermediate level set of the Morse function π_2. The image $G(\sqcup^n I)$ intersects L transversely in $k \geq n+2$ points $p_1, \ldots, p_k$; compare Figure 49.

Take two such points p_i and p_j and suppose they are connected by an arc in $G(\sqcup^n I)$ that does not intersect L. In Figure 49 such pairs of points are (p_1, p_2), (p_2, p_4), (p_3, p_5) and (p_3, p_6). Each such pair corresponds either to a local minimum or to a local maximum of the function $\pi_2 \circ G$. For two such points, write $a_{ij} = a_{ji}$ for the projection of the corresponding arc to the disc L. Such arcs are also drawn in Figure 49. In the language of Section 2.6, these arcs are the intersections of the ascending and descending membranes of the critical points with the level set L. We assume that the arcs a_{ij} intersect one another transversely. Moreover, the arcs corresponding to maxima (respectively minima) are mutually disjoint, while the projections of minima and maxima arcs can intersect each other. The transverse intersections assumption is equivalent to the Morse–Smale condition (Definition 2.6.3).

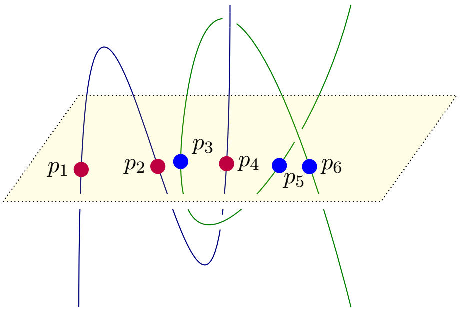

FIGURE 49. Intersecting the string link of Figure 48 with an intermediate level set L.

If we consider the projection of $G(\sqcup^n I)$ to the disc L, in addition to the arcs just discussed we also see two arcs per component of $G(\sqcup^n I)$, running from one of the p_k to $(q_\ell, 1/2) \in L = D^2 \times \{1/2\}$, for some ℓ. These correspond to the projections of monotone arcs with one endpoint p_k and one endpoint either $q_\ell^+ := q_\ell \times \{1\}$ or $q_\ell^- = q_\ell \times \{0\}$. (Recall that $q_\ell^\pm$ are the endpoints of the ℓ-th component of $G(\sqcup^n I)$.) These arcs in L can intersect anything, including themselves; we do not need to control these intersections. They represent a portion of the link components where the embedding is already monotone increasing with respect to the I coordinate. Denote the arc in L by $b_\ell^\pm$, using b_ℓ^+ if one endpoint of the preimage in $D^2 \times I$ is $q_\ell^+ \in D^2 \times \{1\}$, and b_ℓ^- if the endpoint is $q_\ell^- \in D^2 \times \{0\}$. Figure 50 shows the configuration in L coming from Figure 48.

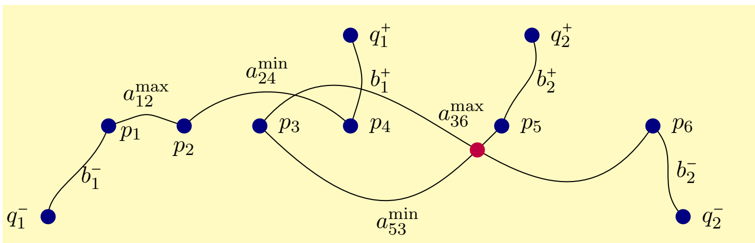

FIGURE 50. The projection of the arcs containing the critical points, and the arcs $b_i^\pm$ to the level set L. In the language of Definition 2.6.1, the arcs between two p_i are the intersections of the ascending and descending membranes of the local minima and maxima with L. The points $q_i \times \{0\}$ and $q_i \times \{1\}$ have been displaced from one another so their projections can be shown as distinct points $q_i^\pm$. Arcs labelled with min represent the projections of minima, while arcs labelled with max represent the projections of maxima.

Let $a_{\min}$ and $a_{\max}$ be two arcs among the $\{a_{ij}\}$ that meet at one of their end points, with $a_{\min}$ (respectively $a_{\max}$) the projection of an arc containing a local minimum $w_{\min}$ of $\pi_2 \circ G$ (respectively, a local maximum $w_{\max}$). Choose these to be the first such pair that arises while travelling along a link component starting at q_ℓ^- for some ℓ, i.e. $a_{\max}$ must be adjacent to b_ℓ^-. If this is their only intersection point then isotope the arc $a_{\min}$ that is below the intermediate level to slightly above it, and isotope the arc $a_{\max}$ that is above the intermediate level to

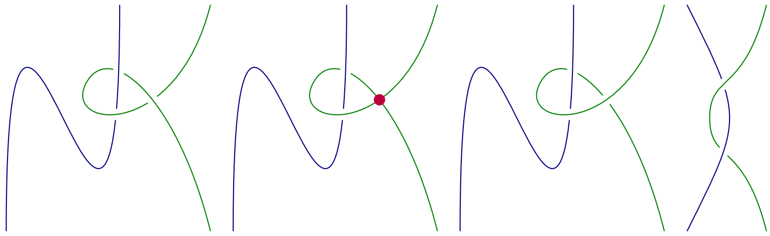

Figure 51. A homotopy of the string link that cancels critical points, even when there are intersections between membranes.

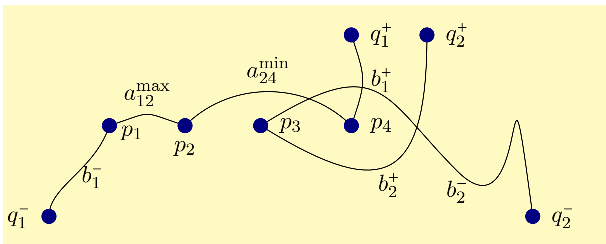

Figure 52. The effect of the homotopy in the first three frames of Figure 51 on the configuration of arcs in L.

below it. Then further isotope both arcs downwards, without creating maxima nor minima, so that the arc $a_{\min}$, apart from the endpoint that is furthest along the ℓth component, is moved back slightly below the intermediate level; move this endpoint so as to lie on L. The result is to cancel the critical points $w_{\min}$ and $w_{\max}$. In the model case presented in Figure 50, we can perform this operation for the arcs a_{12} and a_{24} connecting (p_1, p_2) and (p_2, p_4). The effect on the projection to ℓ is that the arcs $a_{\max}$ and $a_{\min}$ are both added to b_ℓ^-.

To fix notation, let i, j, and k be defined by setting $a_{\min} = a_{ij}$ and $a_{\max} = a_{jk}$. Even if the arcs $a_{\min}$ and $a_{\max}$ meet in their interior, that is if $(a_{ij} \cap a_{jk}) \smallsetminus \{p_j\}$ is nonempty (this is the case in Figure 50 for arcs a_{53} and a_{36} connecting (p_3, p_5) and (p_3, p_6)), we can still perform the same cancellation motion — push a_{ij} above L, push a_{jk} below L, and then move all apart from the endpoint p_i of a_{ij} to be below L. Since the maxima (minima) arcs are mutually disjoint, this only creates intersections between the two pushed arcs. Since they share a common point p_j, the arcs a_{ij} and a_{jk} are projections of arcs of the same component of the string link $G(\sqcup^n I)$, and so we have performed a link homotopy. An example is drawn in Figure 51, and the effect of the homotopy in Figure 51 on the configuration of arcs is shown in Figures 52 and 53.

The effect of the cancellation on the configuration of arcs diagram, in L, is to remove the points of intersection p_j and p_k of $G(\sqcup^n I)$ with L. The arcs a_{ij} and a_{jk} are joined with b_ℓ^-. The point $p_i \in L$ at the far end of a_{ij} remains. The number of maxima and the number of minima of $\pi_2 \circ G$ have both been reduced by one. Since $\sqcup^n I$ is compact, a finite induction suffices to remove all maxima and minima. Once all that remains are the monotone arcs $b_i^\pm$, we have a braid, and it has been obtained by a link homotopy from our original string link G. The closure of this braid is therefore link homotopic to the original link in S^3 we started with. $\square$

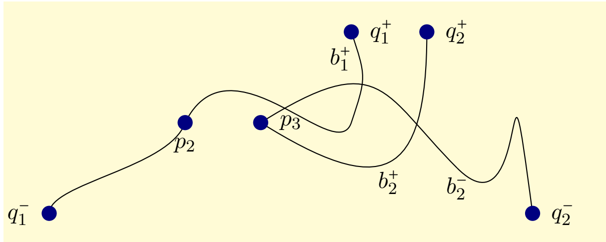

FIGURE 53. The effect of the homotopy in the final two frames of Figure 51 on the configuration of arcs in L.

6.3. A 4-DIMENSIONAL EXAMPLE

We provide another example, this time involving surfaces in 4-dimensional space. We explain Figure 5 from the introduction in detail. This example enables us to illustrate the finger move, and the idea of the proof of Theorem 6.1.1. We use the simplest nontrivial slice knot 6_1, the stevedore knot, which is shown in Figure 54. This example motivates the more general construction of the finger move from Part 5.

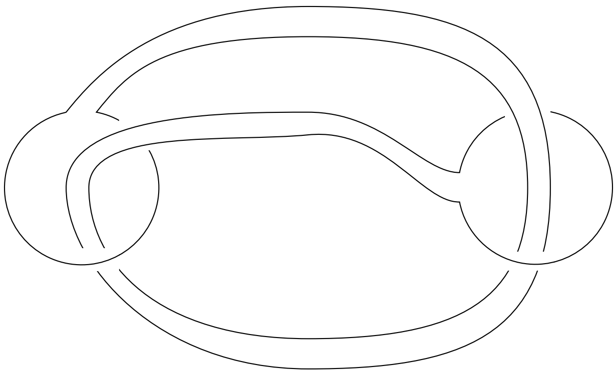

FIGURE 54. The stevedore knot 6_1.

The stevedore knot 6_1 bounds a disc Δ in D^4, but no embedding of the disc can be built from a single 0-handle and no 1-handles, in such a way that the handle decomposition comes from a Morse function on D^4 with a single index 0 critical point; if it could, the knot would be trivial. However, the knot bounds a disc admitting a handle decomposition with two 0-handles and one 1-handle. This can be seen in Figure 55. Our aim is to cancel a 0–1 handle pair at the expense of a regular homotopy of Δ. This will yield a link null-homotopy of 6_1. Of course we know this is possible, because every knot in S^3 is null homotopic. Our aim is to give an explicit demonstration of the finger move.

To fix notation, consider $F\colon D^4 \to [0,1]$ given by the radius, and assume F is an immersed Morse function with respect to Δ. Denote the critical points corresponding to 0-handles by p_-, p'_-, and the critical point corresponding to the 1-handle by p_+. Choose a grim vector field ξ on D^4 for F.

In the absolute setting, i.e. if we were to ignore the fact that they lie on an embedded submanifold and just consider trajectories of ξ on Δ, either pair of 0- and 1-handles could be cancelled. This is not possible when we consider the ambient trajectories of ξ as well. The obstruction to cancellation can be expressed in terms of the membranes $\mathbb{M}^0_{\mathrm{d}}(p_+)$ and

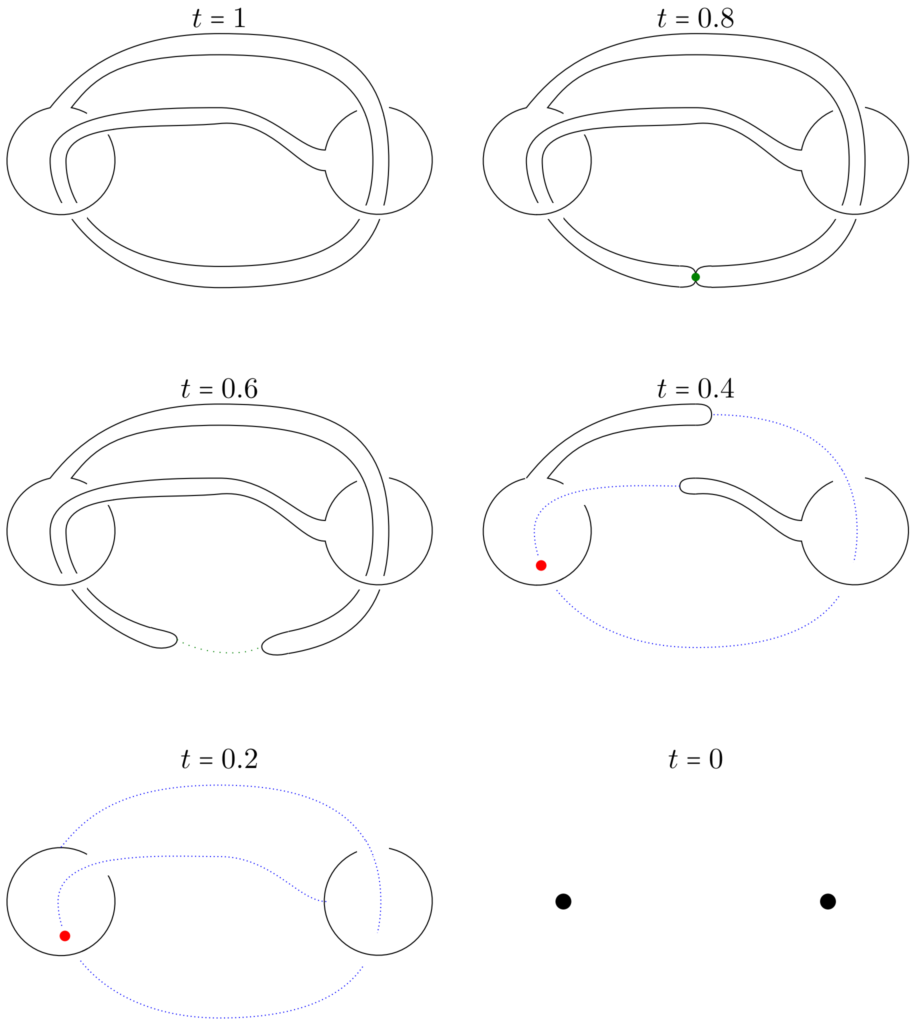

FIGURE 55. A movie indicating that the stevedore knot 6_1 bounds a disc Δ in D^4 such that the distance-to-origin function t on D^4 restricts to a Morse function on Δ with precisely two minima and one saddle.

$\mathbb{M}^0_{\mathrm{a}}(p_-)$ or $\mathbb{M}^0_{\mathrm{a}}(p'_-)$. When the membranes of the critical points have nontrivial intersection in their interior, this prevents one from cancelling the critical points, even when there is a unique trajectory on Δ between the critical points. As mentioned before, this was studied in [Per75, Sha88, BP16]. Each such intersection between membranes corresponds to a trajectory that starts at the index zero critical point, leaves the embedded disc immediately and then later arrives at the index 1 critical point. The hypothesis of the Cancellation Theorem 3.4.1 that there be only one trajectory of a gradient-like immersed vector field between the critical points, and that this must lie on Δ, is not satisfied. The ascending membrane of one of the index zero critical points p_-, p'_-, and the descending membrane of the index one critical point p_+, are depicted in Figure 56. We can see that they intersect.

However, we can cancel the intersection points in the interiors of membranes after a suitable regular homotopy of Δ. This homotopy can be constructed as follows. Choose a level set c in between the two critical level sets c_- and c_+ for p_- and p_+ respectively. Choose a curve γ

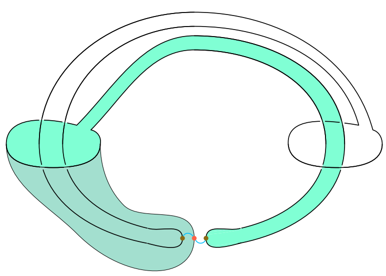

FIGURE 56. The ascending membrane (shaded) of one of the two index 0 critical points p_- of the disc Δ, below the index 1 critical level set. The small arc represents the descending membrane of the index 1 critical point p_+ of Δ. The intersection of the two membranes is one point in the middle of the arc.

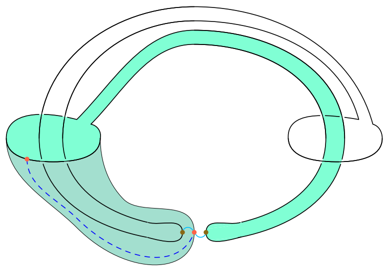

FIGURE 57. The guiding curve γ represented as a dotted arc connecting the intersection point of the membranes with the link.

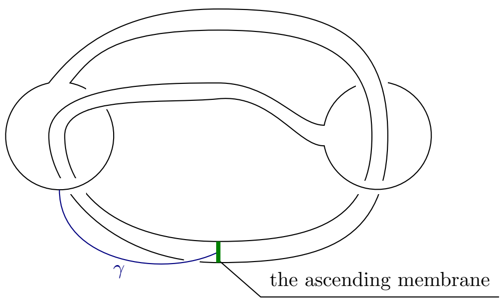

FIGURE 58. The guiding curve γ above the level set c_+.

that lies in this level set, that lies on the ascending membrane of p_-, and which connects the intersection point of the ascending and descending membranes $\mathbb{M}^0_{\mathrm{a}}(p_-) \cap \mathbb{M}^0_{\mathrm{d}}(p_+) \cap F^{-1}(c)$ to a point on the ascending sphere $\mathbb{M}^1_{\mathrm{a}}(p_-)$ of p_-. See Figure 57. Now take the curve γ and flow it upwards, to a level in between the level set c and a level set just above the index 1 critical point. That is, to the level $c_+ + \delta$, where $\delta > 0$ is a small real number. Let γ_t be the result of flowing γ to the level set t. So we are considering $\gamma_{c_+ + \delta}$.

Above c_+, there is a problem with flowing at the endpoint $\gamma(0)$ of γ that lies on $\mathbb{M}^0_{\mathrm{d}}(p_+)$. We therefore consider, in the level sets $F^{-1}(\widetilde{c})$, for $\widetilde{c}$ in $(c_+, c_+ + \delta]$, the closure of $F^{-1}(\widetilde{c}) \cap \bigcup_{t \in \mathbb{R}} \gamma_t$; see Figure 58.

Now we describe the homotopy for the finger move. It is supported in the level sets $[c_+ - \delta, c_+ + \delta]$. We describe the intersection of the finger-moved Δ with $F^{-1}([c_+ - \delta, c_+ + \delta])$,

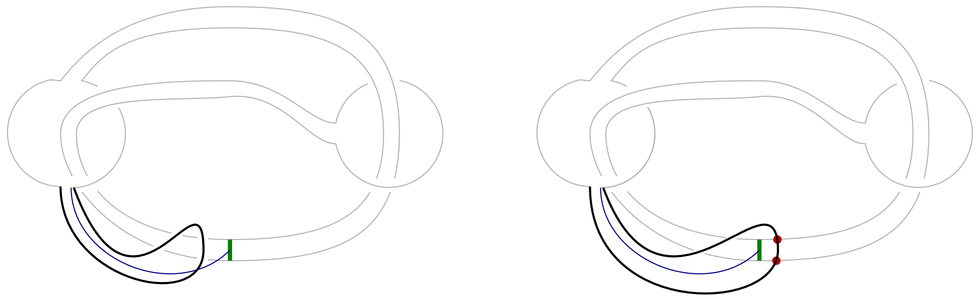

FIGURE 59. The outcome of dragging part of the knot along the guiding curve, shown at the level sets $F^{-1}(c_+ + 3\delta/4)$ on the left, and at $F^{-1}(c_+ + \delta/2)$ on the right.

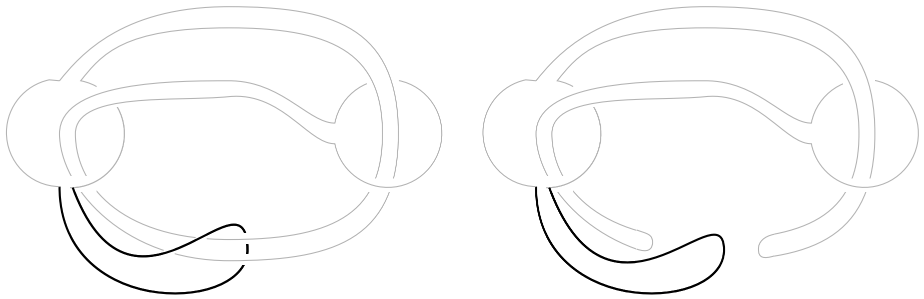

FIGURE 60. Left: the maximum extent of the finger move, in $F^{-1}(c + \delta/4)$. The finger is pushed past the band. The resulting knot is the unknot. Right: the finger is dragged back to where it started, below the level set $F^{-1}(c_+)$. The right figure shows $F^{-1}(c_+ - \delta/4)$.

starting at the level set $c_+ + \delta$, *above* the critical level $F(p_+) = c_+$ for the index 1 critical point p_+. We will work downwards. We push a neighbourhood of the endpoint of γ that lies on the ascending sphere of p_- along γ by a 'finger move' as in Figures 59.

The length of the finger depends on the level $c_+ + t\delta$, $t \in [-1, 1]$. For $c_+ + \delta = c_{\text{top}}$, the length is zero. As t decreases from 1, the finger gets longer. At some level set $c_{\text{int}} \in (c_+, c_+ + \delta)$, the finger crosses the two lower strands of the knot, creating a pair of self-intersections; see Figure 59(left). This happens at the two points of $\mathbb{M}^1_{\text{a}}(p_+) \cap F^{-1}(c_{\text{int}})$; see Figure 59(right). Moving the finger slightly further in the level sets below c_{int}, we no longer have self-intersections. The knot after this move, which is an unknot, is depicted on the left of Figure 60.

Next, we have to undo the finger, in the reverse of the above process. We will do this within the level sets $[c_+ - \delta, c_+)$. We keep the finger constant between c_{int} and c_+, then below c_+ we begin to retract it. We do this in such a way that the disc Δ is unaltered by the move at the level set $c_+ - \delta$, and the finger reduces in length gradually as we move down from c_+ to $c_+ - \delta$. Below the index one critical point p_+ i.e. below the level set c_+, the band is cut by the saddle point, so the finger can shorten without introducing any more self intersections of Δ. Compare the right of Figure 60.

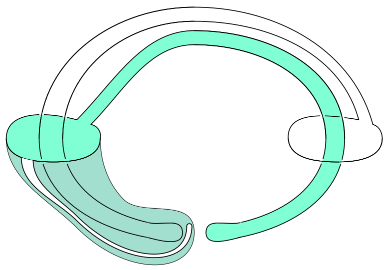

FIGURE 61. After drilling a neighbourhood of the guiding curve, the membranes do not intersect any more.

The finger move transforms the disc Δ into an immersed disc Δ'. The two discs only differ in the region $F^{-1}([c_+ - \delta, c_+ + \delta])$. The Morse function on D^4 induces an immersed Morse function F on Δ'. Apart from the two index zero critical points p_- and p'_-, and one index 1 critical point p_+, F now has two critical points on the double point stratum (recall that any isolated double point is automatically a critical point). There is a single trajectory on Δ' connecting p_- to p_+, and a single trajectory on Δ' connecting p'_- to p_+. But now the trajectory connecting p_- to p_+, that used to run outside Δ, has no corresponding trajectory in the complement of Δ'. See Figure 61.

So we are ready to apply Cancellation Theorem 3.4.1 to cancel the pair of critical points p_- and p_+. Note that the disc below the level set c_{int} is a slice disc for the unknot. Some care must be taken to show that no new trajectories from p_- to p_+ appear as a result of the finger move. It turns out that broken trajectories, that is trajectories going from p_-, through a critical point on a deeper stratum, and ending up in p_+, can be excluded by a dimension counting argument. Ordinary trajectories do not appear since we have just removed them with the finger move. Thus we can cancel p_- and p_+, leaving just a single index 0 critical point p'_-. The new disc Δ' therefore determines a link homotopy from the original knot 6_1 to the unknot. The above construction is described precisely, and in generality, in carefully chosen coordinates, in Section 5.3.

In this dimension, the finger move introduced a 0-sphere S^0 of self intersections. In higher dimensions, the finger move introduces a higher dimensional sphere S^k of new self-intersections.

6.4. CODIMENSION ONE LINKS

The Schoenflies theorem holds for $n \neq 3$, proven by elementary methods for $n < 3$ and using the h-cobordism theorem for $n > 3$ [Mil65, p. 112]. More precisely, given a smooth embedding $\varphi\colon S^n \hookrightarrow \mathbb{R}^{n+1}$, the h-cobordism theorem implies that for $n > 3$ the (closure of the) bounded part of its complement is diffeomorphic to D^{n+1}. Since embeddings of codimension 0 discs (in any connected manifold) are all isotopic, up to a reflection, it follows that the image $K := \varphi(S^n)$ is ambiently isotopic to the unknot $U := S^n \subseteq \mathbb{R}^{n+1}$, the round n-sphere.

We note that precomposing U by a self-diffeomorphism of S^n can change the isotopy class of the embedding $\varphi : S^n \hookrightarrow \mathbb{R}^{n+1}$ and in fact, Cerf's pseudo-isotopy theorem implies that this precomposition inducess bijections

$$\pi_0 \operatorname{Diff}(S^n) \cong \pi_0 \operatorname{Emb}(S^n, \mathbb{R}^{n+1}) \quad \text{for } n > 4.$$

By Kervaire-Milnor [KM63] the (orientation preserving part of the) left hand mapping class group is isomorphic to the group Θ_{n+1} of exotic $(n+1)$-spheres, a highly nontrivial (finite) group.

That is why we do not work with *parametrized knots* φ in this codimension one discussion but with their images, i.e. smooth codimension one submanifolds $K = \varphi(S^n)$ (that happen to be diffeomorphic to S^n). In that spirit, we defined a codimension one link $L \subseteq \mathbb{R}^{n+1}$ to be an *ordered* sequence $(K_1, \dots, K_r)$ of disjoint codimension one knots.

Definition 6.4.1. Given unparametrized knots $K_0, K_1 \subseteq \mathbb{R}^{n+1}$, we say that a concordance $g: S^n \times [0,1] \to \mathbb{R}^{n+1} \times [0,1]$ connects K_0 and K_1 if $g|_{S^n \times \{i\}}$ is a parametrisation of K_i for $i = 0, 1$. Similarly for homotopy between K_0 and K_1 and unparametrized link concordance respectively link homotopy. Alternatively, we could require maps of manifolds whose domain happens to be diffeomorphic to $S^n \times [0,1]$, but that notion is equivalent.

We note that two embeddings $\varphi_0, \varphi_1: S^n \hookrightarrow \mathbb{R}^{n+1}$ that differ by an *orientation preserving* diffeomorphism of S^n are homotopic without changing their image, because the diffeomorphism is homotopic to the identity. So changing the parametrisations of the components of an embedding $\varphi: \sqcup^r S^n \hookrightarrow \mathbb{R}^{n+1}$ in this way preserves its link homotopy and link concordance class.

However, if we precompose φ by a reflection on one component then its link concordance class can change, depending on whether that component was innermost or not. In the former case, the component is null homotopic in the complement of the other components and hence homotopic to the precomposition with a reflection. In the latter case, there is at least one other component on the inside and hence the winding number around that component detects the orientation.

Remark 6.4.2. This discussion leads naturally to the notion of *oriented* unparametrized links $L = (K_1, \dots, K_r)$ for which the notion of link concordance is slightly more interesting. Recalling the *dual tree* from Definition 1.1.3, it turns out that the orientations of the non-innermost submanifolds K_i lead to *orientations* of those edges in $t(L)$ that are "internal", i.e. not adjacent to a leave, and so they correspond to innermost components K_i, see the discussion in the next paragraph. Then a link concordance preserves these orientations if and only if the trees are isomorphic as rooted, edge-ordered and internally oriented trees.

Furthermore, it follows that isomorphism classes of such trees also classify the set of parametrised links $\mathrm{Emb}(\sqcup^r S^n, \mathbb{R}^{n+1})$, modulo link concordance.

Now we give the proof of Proposition 1.1.4, whose statement we recall here.

Proposition 6.4.3. *For $n \neq 3$, two smooth codimension one links $L, L' \subseteq \mathbb{R}^{n+1}$ are ambiently isotopic if and only if they are link concordant. Moreover, a generically immersed link concordance induces an isomorphism on dual trees, and if there is an isomorphism $t(L) \cong t(L')$ of rooted, edge-ordered trees, then it is unique and is induced by an ambient isotopy from L to L'.*

Proof. The arguments will all use the following notion of an *innermost* component of L. Start with the first component K_1 and consider its "inside" ball J_1, the closure of the bounded part of $\mathbb{R}^{n+1} \smallsetminus K_1$. If there is another component K_j inside K_1, i.e. in J_1, then consider its inside ball J_j etc. By finiteness, there must be a component K_i of L that has no other component inside and hence it bounds a ball in $\mathbb{R}^{n+1} \smallsetminus L$ - we call K_i innermost and note that there may be several innermost components in L. By induction on the number of components r of L, where we only need to add an innermost component, one can easily show the following.

(1) If L has r components then $\mathbb{R}^{n+1} \smallsetminus L$ has $r+1$ connected components, exactly one of which is unbounded.
(2) For every component K of L, there are exactly two connected components of $\mathbb{R}^{n+1} \smallsetminus L$ whose closures meet K.
(3) We have that $t(L)$ is a well-defined tree, with an ordering of the edges induced from the ordering of the components of L.
(4) If C is a connected component of $\mathbb{R}^{n+1} \smallsetminus L$ representing a vertex v_C in $t(L)$ then the edges adjacent to v_C correspond to the components of L that lie in the closure (or equivalently, in the point set boundary) of C.
(5) The component C is innermost if and only if v_C is a univalent vertex.
(6) If there is an isomorphism $t(L) \cong t(L')$ of rooted edge-ordered trees, it is unique, and it is realised by an ambient isotopy from L to L'.

The subtle part of Proposition 1.1.4 is to show that link concordance implies ambient isotopy. By Proposition 1.1.5 we can approximate the link concordance by an immersed link concordance

$$G\colon (\sqcup^r S^n) \times [0,1] \looparrowright \mathbb{R}^{n+1} \times [0,1].$$

from L to L', that we may assume is in general position as in Definition 2.2.1

Then we show that G induces an isomorphism $t(L) \cong t(L')$ which is implied by the following result, Lemma 6.4.4, which proves that the obvious inclusion-induced maps on vertices and edges are an isomorphism of rooted, edge-ordered trees. $\square$

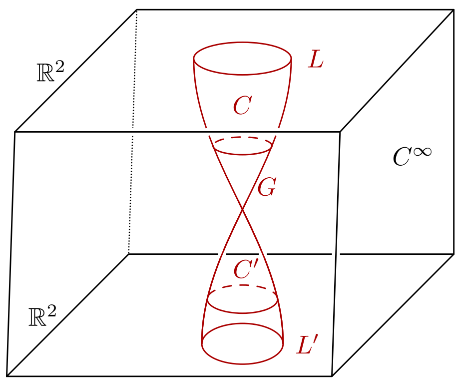

FIGURE 62. An 'hourglass' knot concordance G between two copies L and L' of a standard $S^1 \subseteq \mathbb{R}^2$. This G is not generically immersed. The connected components of $\mathbb{R}^2 \times [0,1] \smallsetminus G$ are C, C', and the unbounded region C^∞. With $j\colon \pi_0(\mathbb{R}^2 \smallsetminus L) \to \pi_0(\mathbb{R}^2 \times [0,1] \smallsetminus G)$ and $j'\colon \pi_0(\mathbb{R}^2 \smallsetminus L') \to \pi_0(\mathbb{R}^2 \times [0,1] \smallsetminus G)$ the inclusion-induced maps, we have $\{C_\infty, C\} = \operatorname{Image}(j) \neq \operatorname{Image}(j') = \{C_\infty, C'\}$. Although L and L' are isotopic, and $t(L) \cong t(L')$, this isomorphism is not induced by the concordance G. Since G is not generically immersed, this is consistent with Proposition 6.4.3.

Lemma 6.4.4. *The two inclusions from the boundary induce injective maps j, j'*

$$\pi_0(\mathbb{R}^{n+1} \smallsetminus L) \xrightarrow{j} \pi_0(\mathbb{R}^{n+1} \times [0,1] \smallsetminus G) \xleftarrow{j'} \pi_0(\mathbb{R}^{n+1} \smallsetminus L')$$

with equal images, $\operatorname{Image}(j) = \operatorname{Image}(j')$.

Remark 6.4.5. The injectivity of the maps j, j' also holds for an arbitrary (non-immersed) link concordance by Alexander duality, as will become apparent from the proof of Lemma 6.4.4. However, equality of their images does not hold in general, as illustrated by the example in Figure 62.

Proof of Lemma 6.4.4. The first step can be proven using Alexander duality, but we use a transversality argument instead.

Step 1. Generically, a smooth circle $c: S^1 \to \mathbb{R}^{n+1} \times [0,1]$ intersects a smooth concordance $g: S^n \times [0,1] \to \mathbb{R}^{n+1} \times [0,1]$ in an even number of points. To see this, note that the circle bounds a smooth map $D^2 \to \mathbb{R}^{n+1} \times [0,1]$ which generically intersects g in a compact 1-manifold with boundary $g \pitchfork c$, so this consists of an even number of points.

Step 2. Injectivity of $j: \pi_0(\mathbb{R}^{n+1} \smallsetminus L) \hookrightarrow \pi_0(\mathbb{R}^{n+1} \times [0,1] \smallsetminus G)$. Assume that there are two components C_1, C_2 of $\mathbb{R}^{n+1} \smallsetminus L$ that map to the same component, $j(C_1) = j(C_2)$. That means that there is a path γ_{12} in $\mathbb{R}^{n+1} \times [0,1] \smallsetminus G$ from a point $p_1 \in C_1$ to $p_2 \in C_2$. Consider the geodesic in $t(L)$ from the vertex v_{C_1} to v_{C_2} and assume that K_i is a component of L that corresponds to an edge on that geodesic. This means that there is a path γ in $\mathbb{R}^{n+1}$ from p_1 to p_2 that intersects K_i generically an odd number of times. So $\gamma_G \cup \gamma$ is a circle in $\mathbb{R}^{n+1} \times [0,1]$ that intersects the component G_i of G generically an odd number of times, a contradiction to the observation in Step 1.

Step 3. Setting up an induction on r, the number of components of L to show the equality $\mathrm{Image}(j) = \mathrm{Image}(j')$. Let K_i be an innermost component of L and let $\widehat{L} := L \smallsetminus K_i$. Then $\widehat{G} := G \smallsetminus G_i$ is an immersed concordance from $\widehat{L}$ to $\widehat{L}' := L' \smallsetminus K_i'$ and so the statement holds by induction. The base case that G is empty is easily seen to hold. To show it for G, it suffices to prove that there is a path γ in $\mathbb{R}^{n+1} \times [0,1] \smallsetminus G$ that starts on the inside of K_i' and ends on the inside of K_i. This is one of the places where we use that G is a generic immersion. Namely, we show that such a path γ exists arbitrarily close to G_i, and hence γ automatically can be assumed to miss the other components of G. So we are left with showing the inductive step purely in terms of the innermost component we removed.

Step 4. To prove the inductive step, we need to show that for a generically immersed concordance g between knots K to K', there is a path γ in $\mathbb{R}^{n+1} \times [0,1] \smallsetminus g$ that starts on the inside of K', ends on the inside of K, and stays arbitrarily close to g.

To prove this statement, consider the normal bundle νg of $g: S^n \times [0,1] \looparrowright \mathbb{R}^{n+1} \times [0,1]$, a 1-dimensional vector bundle on $S^n \times [0,1]$. This is a trivial bundle since its restriction to $S^n \times \{0\}$ can be trivialized by the normal vector that points into the inside of K. As a consequence, we can extend that section to a non-vanishing section σ of νg. Using the exponential map locally, where g is an embedding, we get a map $s: S^n \times [0,1] \to \mathbb{R}^{n+1} \times [0,1]$ that we think of as a 'parallel' of g. Here we used again that g is a generic immersion; this step is not possible with the 'hourglass' concordance in Figure 62.

If g were an embedding then s would also be an embedding (and actually be an honest parallel), lying in the complement of g. However, the double point manifold of g has codimension one in $S^n \times [0,1]$, the triple point manifold has codimension 2, etc. Therefore, we can at least pick a smooth arc $\alpha: [0,1] \to S^n \times [0,1]$ from $(x,0)$ to $(x,1), x \in S^n$, that misses the triple, quadruple and higher singular points of g and intersects the double point set generically in finitely many points $\{t_1, \dots, t_m\} \subseteq [0,1]$.

By taking $s \circ \alpha: [0,1] \to \mathbb{R}^{n+1} \times [0,1]$, we obtain an embedded arc β in $\mathbb{R}^{n+1} \times [0,1]$ that starts on the inside of K', ends on the inside of K, and meets g transversely in $2m$ points. One pair of these intersections comes from each point $\alpha(t_i)$ lying in the double point set of g. So β follows the normal vector field along α. At each double point of α, we have the local picture in Figure 63. The easiest way to describe it is to start with α given locally in $\mathbb{R}^2$ by the coordinate axes, intersecting in the origin $(0,0)$. Observe that $\mathbb{R}^2 \smallsetminus \alpha$ has four connected components, namely the four (open) quadrants. By transversality, g is given locally by the product of this $\mathbb{R}^2$ with $\mathbb{R}^n$, with $\{0\} \times \mathbb{R}^n$ becoming the local double point manifold of g. Note that g itself may twist around α from one double point of α to the next and that β just twists along as its normal vector field. So we have no control over how β moves between these double points, it just continues to stay normal to g and glues together to a smoothly embedded arc (since we are assuming $n > 0$).

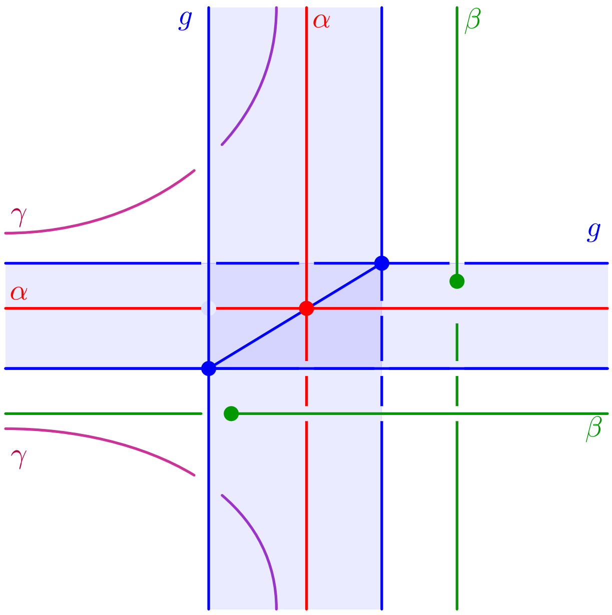

FIGURE 63. Constructing the paths α, its parallel β, and ultimately γ.

From the local description we see, at each double point of α, a unique turn that we can introduce just before β hits g. We can keep the new local arc γ normal to both sheets of g and in addition disjoint from g as in Figure 63. This comes from the fact that in a coordinate ball around the double point, the complement of g has exactly four connected components, aforementioned the quadrants in $\mathbb{R}^2$ times $\mathbb{R}^n$. The required global path γ in the complement of g is then constructed by the following algorithm.

(1) Start with $\gamma(0) := \beta(0)$ lying on the inside of K' and let $\gamma(t) := \beta(t)$ for all $t < t_1$, the first double point of α.
(2) At that double point, let γ make the unique smooth turn that keeps it disjoint from g and then continue normally to g along α. This continuation may be either following β, or may be on the other side, i.e. following the negative $-\sigma$ of the section σ used to define s and β, applied to α. We denote this curve by $-\beta$. The submanifold γ may also have a different local orientation to $\pm\beta$.
(3) This process of building an embedding γ of a compact connected 1-manifold continues uniquely, up to isotopy, turning at all double points $\{t_1, \ldots, t_m\}$ of α in the prescribed way, and staying normal to g along α otherwise.
(4) The resulting 1-manifold can be parametrised to become an embedding of $[0,1]$ with a second boundary point $\gamma(1)$. By construction, $\gamma(1)$ is equal to either σ or $-\sigma$

applied to a boundary point of α. So $\gamma(1)$ equals $\sigma(\alpha(0)) = \beta(0)$, $-\sigma(\alpha(0)) = -\beta(0)$, $\sigma(\alpha(1)) = \beta(1)$, or $-\sigma(\alpha(1)) = -\beta(1)$. However $\gamma(1)$ cannot be equal to $\gamma(0) = \beta(0)$ since the manifold γ has two boundary points.

The last step is to argue that $\gamma(1) = \pm\beta(1)$ and lies on the inside of K, because the other two possibilities contradict Step 1 above. If $\gamma(1) = -\beta(0)$ still lies in $\mathbb{R}^{n+1} \smallsetminus K'$, then we can glue γ with the short path $\delta\colon [0,1] \to \mathbb{R}^{n+1}$ sending $t \mapsto (\frac{1}{2}(t+1) \cdot \beta(0)$ through K' to get a circle that intersects g once, which is a contradiction. Similarly, if $\gamma(1)$ lies in $\mathbb{R}^{n+1} \smallsetminus K$, but outside of K, then we can connect $\gamma(1)$ by a path in the unbounded component, far away from g, to $-\beta(0)$. Using δ again we obtain a circle that intersects g once, which is another contradiction. Hence $\gamma(1)$ is equal to whichever of $\pm\beta(1)$ lies inside K. This completes the proof of Step 3, and hence completes the proof of Lemma 6.4.4. □

6.5. Regular homotopies of surfaces

A regular homotopy between immersed compact surfaces $G_i\colon \Sigma \looparrowright M$, $i = 0,1$, in a closed 4-manifold M, is a smooth homotopy $G\colon \Sigma \times I \to M$ through immersions, with $G_t\colon \Sigma \looparrowright M$ an immersion, and with $G_t^{-1}(\operatorname{bd} M) = \operatorname{bd}\Sigma$ for all $t \in I$. We consider the trace $G'\colon \Sigma \times I \to M \times I$ that sends $G'(s,t) = (G(s,t),t)$. This is a level-preserving generic immersion. The following result is often used in the 4-manifolds literature.

Theorem 6.5.1. *Let $G\colon \Sigma \times I \looparrowright M \times I$ be a generic immersion between generically immersed surfaces, e.g. the trace of a regular homotopy between $G_0(\Sigma)$ and $G_1(\Sigma)$. Then G is regularly homotopic to a concatenation of finger moves, Whitney moves, and ambient isotopies leading from G_0 to G_1. Moreover, we may assume that all the finger moves occur before all the Whitney moves.*

Proof. By general position (Lemma 2.2.3), we may and shall assume that G induces a stratification on $M \times I$ with:

- $\Omega[d] = \varnothing$ for $d \geq 3$,
- $\Omega[2]$ the double point set, a collection of disjointly embedded 1-manifolds,
- $\Omega[1]$ the complement in $G(\Sigma \times I)$ of the double point set, and
- $\Omega[0]$ the complement $M \times I \smallsetminus G(\Sigma \times I)$.

By Lemma 3.1.22, we may assume after a perturbation of G that the projection $F\colon M \times I \to I$ is an immersed Morse function with respect to $G(\Sigma \times I)$. By Theorem 6.1.1, G is smoothly regularly homotopic rel. boundary to the trace of a regular homotopy.

Since G is level preserving, the only critical points occur on $\Omega[2]$. Since $\Omega[2]$ is a 1-manifold, the critical points are maxima and minima, i.e. index 0 and index 1 critical points of the Morse function restricted to the stratum $\Omega[2]$. Minima correspond to finger moves of $G_t(\Sigma)$ and maxima correspond to Whitney moves. By the Rearrangement Theorem 3.2.4, we may rearrange the critical points so that maxima are above minima. Consider a single pair of a Whitney move and a finger move, where the Whitney move occurs first. To see that the rearrangement theorem applies, note that the descending membrane of a finger move is 3-dimensional, and thus 2-dimensional in each level set, while the ascending membrane of a Whitney move is 2-dimensional, and thus 1-dimensional in each level set. The level sets are 4-dimensional, so by the Morse-Smale condition the membranes may be assumed disjoint. More precisely, Theorem 3.2.4 alters the Morse function by a 1-parameter family to achieve the rearrangement, with a corresponding 1-parameter family of grim vector fields. It does so without introducing any new critical points. We perform an ambient isotopy that returns the Morse function to the original projection; see Lemma 4.1.2. The rearrangement switches the order of our given pair of a Whitney move and finger move. By applying this as many times as necessary, it follows that by a regular homotopy G can be made into the trace of a regular homotopy, and all the finger moves i.e. minima on $\Omega[2]$, can be placed before the Whitney moves i.e. the maxima on $\Omega[2]$. □

References

[AGZV12] V. Arnold, S. Gusein-Zade, and A. Varchenko. *Singularities of differentiable maps. Volume 1.* Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2012.

[Arn83] V. I. Arnold. *Geometrical methods in the theory of ordinary differential equations*, volume 250 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, New York, 1983. Translated from the Russian by Joseph Szücs, Translation edited by Mark Levi.

[Arn06] Vladimir I. Arnold. *Ordinary differential equations.* Universitext. Springer-Verlag, Berlin, 2006. Translated from the Russian by Roger Cooke, Second printing of the 1992 edition.

[Bar99] Arthur Clemens Bartels. *Link homotopy in codimension two.* ProQuest LLC, Ann Arbor, MI, 1999. Thesis (Ph.D.)–University of California, San Diego.

[BP16] M. Borodzik and M. Powell. Embedded Morse Theory and Relative Splitting of Cobordisms of Manifolds. *J. Geom. Anal.*, 26(1):57–87, 2016.

[BT99] A. Bartels and P. Teichner. All two-dimensional links are null homotopic. *Geom. Topol.*, 3:235–252, 1999.

[Cer70] J. Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. *Inst. Hautes Études Sci. Publ. Math.*, (39):5–173, 1970.

[CZ84] Charles Conley and Eduard Zehnder. Morse-type index theory for flows and periodic solutions for Hamiltonian equations. *Comm. Pure Appl. Math.*, 37(2):207–253, 1984.

[Ekh01] T. Ekholm. Invariants of generic immersions. *Pacific J. Math.*, 199(2):321–345, 2001.

[FR86] Roger Fenn and Dale Rolfsen. Spheres may link homotopically in 4-space. *J. London Math. Soc. (2)*, 34(1):177–184, 1986.

[GG73] M. Golubitsky and V. Guillemin. *Stable mappings and their singularities.* Springer-Verlag, New York-Heidelberg, 1973. Graduate Texts in Mathematics, Vol. 14.

[Gif79] C. H. Giffen. Link concordance implies link homotopy. *Math. Scand.*, 45(2):243–254, 1979.

[GM88] M. Goresky and R. MacPherson. *Stratified Morse theory*, volume 14 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3).* Springer-Verlag, Berlin, 1988.

[Gol79] D. L. Goldsmith. Concordance implies homotopy for classical links in M^3. *Comment. Math. Helv.*, 54(3):347–355, 1979.

[GS99] R. E. Gompf and A. I. Stipsicz. *4-manifolds and Kirby calculus*, volume 20 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1999.

[Hab92] N. Habegger. Applications of Morse theory to link theory. In *Knots 90 (Osaka, 1990)*, pages 389–394. de Gruyter, Berlin, 1992.

[Hir59] M. W. Hirsch. Immersions of manifolds. *Trans. Amer. Math. Soc.*, 93:242–276, 1959.

[HL90] N. Habegger and X. S. Lin. The classification of links up to link-homotopy. *J. Amer. Math. Soc.*, 3(2):389–419, 1990.

[KM63] M. A. Kervaire and J. W. Milnor. Groups of homotopy spheres. I. *Ann. of Math. (2)*, 77:504–537, 1963.

[Mat69] J. N. Mather. Stability of C^∞ mappings. II. Infinitesimal stability implies stability. *Ann. of Math. (2)*, 89:254–291, 1969.

[Mil54] J. W. Milnor. Link groups. *Ann. of Math. (2)*, 59:177–195, 1954.

[Mil57] J. W. Milnor. Isotopy of links. Algebraic geometry and topology. In *A symposium in honor of S. Lefschetz*, pages 280–306. Princeton University Press, Princeton, N. J., 1957.

[Mil63] J. W. Milnor. *Morse theory.* Princeton University Press, Princeton, N.J., 1963. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51.

[Mil65] J. W. Milnor. *Lectures on the h-cobordism theorem.* Princeton University Press, Princeton, N.J., 1965. Notes by L. Siebenmann and J. Sondow.

[MR85] W. S. Massey and D. Rolfsen. Homotopy classification of higher-dimensional links. *Indiana Univ. Math. J.*, 34(2):375–391, 1985.

[Per75] B. Perron. Pseudo-isotopies de plongements en codimension 2. *Bull. Soc. Math. France*, 103(3):289–339, 1975.

[Rou70] C. P. Rourke. Embedded handle theory, concordance and isotopy. In *Topology of Manifolds (Proc. Inst., Univ. of Georgia, Athens, Ga., 1969)*, pages 431–438. Markham, Chicago, Ill., 1970.

[Sal85] Dietmar Salamon. Connected simple systems and the Conley index of isolated invariant sets. *Trans. Amer. Math. Soc.*, 291(1):1–41, 1985.

[Sal90] D. Salamon. Morse theory, the Conley index and Floer homology. *Bull. London Math. Soc.*, 22(2):113–140, 1990.

[Sha88] R. W. Sharpe. Total absolute curvature and embedded Morse numbers. *J. Differential Geom.*, 28(1):59–92, 1988.

[ST19] R. Schneiderman and P. Teichner. The group of disjoint 2-spheres in 4-space. *Annals of Mathematics*, 190(3):669–750, 2019.

[Tei96] Peter Teichner. Stratified morse theory and link homotopy. 1996. Unpublished notes.

[Whi37] H. Whitney. On regular closed curves in the plane. *Compositio Math.*, 4:276–284, 1937.

Institute of Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland
Email address: mcboro@mimuw.edu.pl

School of Mathematics and Statistics, University of Glasgow, University Place, Glasgow, G12 8QQ, United Kingdom
Email address: mark.powell@glasgow.ac.uk

Max Planck Institut für Mathematik, Vivatsgasse 7, 53111 Bonn, Germany
Email address: teichner@mpim-bonn.mpg.de