Families of Morse Functions Parameterized by Maxima

Michael D. Hirsch & Morris W. Hirsch

0. Introduction

A C^2 function $f: M \to \mathbf{R}$ on a Riemannian manifold M is called a *Morse function* if, at every critical point p, in local coordinates the Hessian matrix Hf_p is nonsingular. The classical Morse's lemma states that, in suitable local coordinates centered at a critical point, f is expressed as a sum and difference of squares of coordinates.

Morse functions have been of great importance in differential topology and dynamical systems. Möbius's proof of the classification of surfaces decomposed a surface into elementary pieces bounded by noncritical level sets of a Morse function (Möbius [9]; see Hirsch [3] for a modern presentation). In a far-reaching development of Möbius's idea, Smale [12] used Morse functions in an essential way to prove his generalized Poincaré conjecture: he showed that, for a generic set of Morse functions f, the stable and unstable manifolds of singular points of the gradient vector field ∇f provide dual cell decompositions of M. For discussion of Smale's work see Hirsch [4].

Stable and unstable manifolds of ∇f also played a key role in Smale's work [13] on structural stability of generic gradient flows. One-parameter families of Morse-like functions were utilized by Cerf [2] for analyzing the diffeomorphism group of S^3.

In a different type of application, which motivated this paper, Rimon and Koditschek [11] used gradients of Morse functions to attack a problem in robotics. They gave a formula for analytic Morse functions on domains consisting of a closed n-ball D^n with the the interiors of smaller n-balls deleted, having prescribed critical points and boundary values. It is noteworthy that the Rimon–Koditschek formula depends discontinuously on a certain parameter derived from the data. This raises the problem of determining whether continuous formulas exist.

We simplify the problem by considering only Morse functions $f: D^2 \to \mathbf{R}$ on the closed unit disk $D = D^2$ having the following properties.

Received April 8, 1997. Revision received October 20, 1997.

The first author was partially supported by National Science Foundation grants DMS-9404261 and ASC-92527186 and the University Research Council of Emory University. The second author was partially supported by National Science Foundation grant DMS-9113250.

Michigan Math. J. 45 (1998).

HYPOTHESIS 0.1.

(a) f has precisely r local maxima;
(b) f has no local minima;
(c) the gradient vector field ∇f of f points inward at boundary points of D;
(d) f is of differentiability class C^2.

The space $\mathcal{M}_r$ of such functions is given the C^2 topology. Thus $\mathcal{M}_r$ is an open set in the Banach space of C^2 functions on the disk.

We are chiefly concerned with $r = 3$. The Morse inequalities (or older results going back to Poincaré and even Maxwell) imply that a function in $\mathcal{M}_3$ has precisely two saddle points in addition to the three local maxima.

Let $\mathcal{C}_r$ be the space of unordered sets of r distinct points in the interior Int D of the disk. Define the continuous map $\mu_r : \mathcal{M}_r \to \mathcal{C}_r$, which assigns to each f its set of maxima. In Section 3 we show that this is a locally trivial fibre space.

Our main result is that μ_3 does not have a section.

THEOREM 0.2. *There is no continuous map $\sigma : \mathcal{C}_3 \to \mathcal{M}_3$ such that $\mu_3 \circ \sigma$ is the identity map of $\mathcal{C}_3$.*

Note that it is easy to construct a section of $\sigma : \mathcal{C}_2 \to \mathcal{M}_2$ of μ_2; for example, $\sigma(a, b) = \sigma_{a,b}$ where

$$\sigma_{a,b}(x) = -\|x - a\|^2 \|x - b\|^2. \tag{1}$$

The same formula works in any dimension. The analogous formula for three or more critical points does not always define a Morse function, however; this is discussed in Section 3.

We consider also the space $\hat{\mathcal{M}}_r$ of *labeled* Morse functions in $\mathcal{M}_r$: An element of $\hat{\mathcal{M}}_r$ is an ordered pair $(f, (m_1, \ldots, m_r))$, where $f \in \mathcal{M}_r$ and $(m_1, \ldots, m_r)$ is a list of its maxima. $\hat{\mathcal{M}}_r$ is topologized so that the natural "forgetful" map $\Phi_{\mathcal{M}} : \hat{\mathcal{M}}_r \to \mathcal{M}_r$ is a covering space.

Denote by $\hat{\mathcal{C}}_r \subset D \times \cdots \times D$ the manifold of lists of r distinct points of the disk. The forgetful map $\Phi_{\mathcal{V}} : \hat{\mathcal{C}}_r \to \mathcal{C}_r$ is a covering space. The map $\hat{\mu} = \hat{\mu}_r : \hat{\mathcal{M}}_r \to \hat{\mathcal{C}}_r$, which assigns to each labeled Morse function its list of maxima, is continuous.

The question of whether $\hat{\mu}_r$ has a section is a natural one. It is easy to see that $\hat{\mu}_2$ has a section. We prove in Section 3 that, in contrast to Theorem 0.2, $\hat{\mu}_3 : \hat{\mathcal{C}}_3 \to \hat{\mathcal{M}}_3$ *does* have a section.

Theorem 0.2 is an immediate corollary of a more general result, Theorem 0.4, about families of C^1 vector fields on D^2 whose flows have dynamical properties similar to the gradients of the Morse functions considered in Theorem 0.2, but which are not required to be gradient fields.

Our main tools are unstable manifolds, and especially their saddle connections. A key computation makes use of the structure of the braid group on three strands.

Analogous questions can be raised for higher-dimensional disks, and indeed for Morse functions on arbitrary manifolds. We suspect that, except for very special cases, there do not exist sections.

All maps are assumed or easily proved to be continuous unless the contrary is mentioned. Sections, in particular, are continuous.

It will be more convenient to work with vector fields than with functions. Let $X\colon D \to \mathbf{R}^2$ be a C^1 vector field. An *equilibrium* $p \in D$ is a point for which $X(p) = 0$, or equivalently, a fixed point for the flow Ψ generated by X. Recall that p is *hyperbolic* if the Jacobian matrix $DX(p)$ is invertible and has no pure imaginary eigenvalue. A hyperbolic equilibrium is a *sink* if the real parts of the eigenvalues are negative, and a *saddle* if one eigenvalue is positive and the other is negative. A sink is an asymptotically stable fixed point for Ψ.

Let r denote a positive integer. We work in the space $\mathcal{V}_r$ of C^1 vector fields $X\colon D \to \mathbf{R}^2$ on the disk having the following properties.

HYPOTHESIS 0.3.

(a) all equilibria are hyperbolic with real eigenvalues;
(b) there are r sinks, $r-1$ saddles, and no other equilibria;
(c) X is transverse inward on ∂D;
(d) every complete semitrajectory in D converges to an equilibrium;
(e) there are no heteroclinic or homoclinic loops.

These properties are equivalent to Smale's Axiom A, and they imply the existence of a smooth Liapunov function for the corresponding flow (see Smale [13]).

It is easy to see that $\mathcal{V}_r$ is an open set of the Banach space of C^1 vector fields on the disk (compare Abraham and Robbins [1]). In particular, we may consider $\mathcal{V}_r$ as an infinite-dimensional analytic manifold.

It is clear that the gradient vector field of a function in $\mathcal{M}_r$ lies in $\mathcal{V}_r$, and that the resulting map $\mathcal{M}_r \to \mathcal{V}_r$ is a homeomorphism onto its image. By assigning to each vector field its set of sinks, we define a continuous map $\mathcal{V}_r \to \mathcal{C}_r$. By a convenient abuse of notation we denote this by μ_r. In Theorem 3.2 we prove that μ_r is a fibre bundle projection.

Theorem 0.2 is an immediate consequence of the following.

THEOREM 0.4. *The map* $\mu_3\colon \mathcal{V}_3 \to \mathcal{C}_3$ *has no section.*

As the spaces $\mathcal{V}_r$ for $r > 3$ will play little role in this paper, we simplify notation by setting $\mathcal{V}_3 = \mathcal{V}$ and $\mu_3 = \mu$.

Later we shall need to consider the covering space $\Phi_{\mathcal{V}}\colon \hat{\mathcal{V}} \to \mathcal{V}$, where $\hat{\mathcal{V}}$ is the space of pairs comprising a vector field in $\mathcal{V}$ together with an ordering of its sinks, and $\Phi_{\mathcal{V}}$ is the forgetful map.

Outline of Proof. The proof is by contradiction. A section $\sigma : \mathcal{C}_3 \to \mathcal{V}$ of μ induces the homomorphism of fundamental groups $\sigma_*\colon \pi_1(\mathcal{C}_3) \to \pi_1(\mathcal{V})$. By Hypothesis 0.3(b), every $X \in \mathcal{V}$ has two saddles. These vary continuously with X. The resulting continuous map $\nu\colon \mathcal{V} \to \mathcal{C}_2$ induces the homomorphism $\nu_*\colon \pi_1(\mathcal{V}) \to \pi_1(\mathcal{C}_2)$ of fundamental groups.

Since $\pi_1(\mathcal{C}_r) = B_r$, the braid group on r strands, σ determines in this way a homomorphism of braid groups, $\nu_* \circ \sigma_*\colon B_3 \to B_2$. This homomorphism may, of course, depend on σ; however, by analyzing σ on a particular toral subset of

$\mathcal{C}_3$, we can make an a priori computation of the value of $\nu_* \circ \sigma_*$ on a certain element of B_3. Finally, it will be easy to show algebraically, from the structure of braid groups, that no homomorphism $B_3 \to B_2$ can realize this computation. This contradiction will complete the proof.

Section 1 sets up the basic definitions and constructions needed for Section 2, in which Theorem 0.4 is proven. Section 3 further explores the bundle of Morse functions and has some positive results. The final section contains some open problems and discussion.

1. Loops, Isotopies, and Winding Numbers

In this section we develop the topological tools needed for the proof of the main theorem. Smale [13] or Hirsch and Pugh [5] can be consulted for the theory of stable and unstable manifolds.

The Unstable Manifold Complex

Let Ψ denote the flow generated by a vector field $X \in \mathcal{V}$. Let $p \in \operatorname{Int} D$ be a saddle equilibrium for X. The *stable manifold* $W^s(p, X) = W^s(p)$ is the set of $x \in D$ such that $\lim_{t\to\infty} \Psi_t(x) = p$. There is an injective C^1 immersion $h\colon J \to D$, where $J \subset \mathbf{R}$ is a closed infinite interval, such that $h(J) = W^s(p, X)$. If $W^s(p, X)$ is disjoint from the boundary circle ∂D, then $J = \mathbf{R}$; otherwise, the unique endpoint of J is mapped by h to $W^s(p, X) \cap \partial D$.

The *unstable manifold* of p is the set $W^u(p) = W^u(p, X)$ comprising all points $x \in D$ such that $\lim_{t\to-\infty} \Psi_t(x) = p$. It is the injective image of a C^1 immersion of $\mathbf{R}$.

Deleting p from its stable manifold leaves two *branches* of the stable manifold, homeomorphic to open or half-open intervals. The unstable manifold similarly has two branches. Each branch is invariant under the flow and is therefore a C^2 curve; its closure is a C^1 arc. Only a stable branch (i.e., a branch of a stable manifold) can be a half-open interval, by Hypothesis 0.3(c). This occurs when a point of the stable branch is on the disk's boundary.

A *saddle connection* is a branch that connects two saddles, or connects one saddle to itself; however, the latter situation (a homoclinic loop) is precluded by Hypothesis 0.3(e). Nor can two branches reciprocally connect a pair of saddles (producing a heteroclinic loop). By 0.3(e), there is at most one saddle connection.

Each unstable branch is a complete orbit, which by 0.3(d) connects two equilibria. These equilibria are distinct by 0.3(e); one or both may be saddles.

Given $X \in \mathcal{V}$, let K_X be the *unstable manifold complex* of the saddles of X, that is, the closure of the union of the unstable manifolds of the saddles. We show in Proposition 1.1 that K_X is a contractible graph having as vertices the three sinks and two saddles of X. These graphs divide into two homeomorphism types—"T-shaped" and "V-shaped"—according to whether X has a saddle connection or not. By labeling the equilibria we make K_X into a labeled graph.

Figure 1 is a picture of the two types of complexes. Small parts of the stable manifolds have been drawn to indicate the saddles, and arrows on the unstable manifolds indicate the direction of the flow.

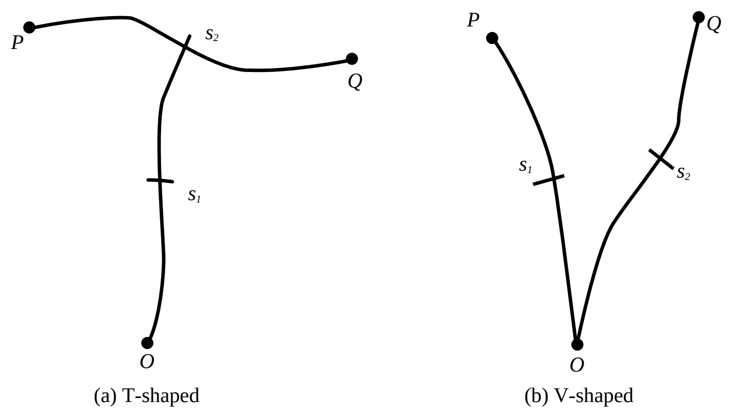

Figure 1 The two combinatorial types of unstable manifold complexes

PROPOSITION 1.1. *For each vector field $X \in \mathcal{C}_3$, one of the following conditions holds for the complex K_X:*

(i) *there are no saddle connections and K_X is homeomorphic to a compact interval;*
(ii) *there is just one saddle connection and K_X is homeomorphic to the triod*

$$\mathsf{T} = \{(x, y) \in \mathbf{R}^2 : xy = 0, \ -1 \le x \le 1, \ 0 \le y \le 1\}.$$

Proof. Suppose there is no saddle connection; then each unstable branch limits at a sink. Since there are four unstable branches and three sinks, two unstable branches must limit at the same sink.

These two branches can not be from the same saddle, say s. For otherwise the closure of their union would form a Jordan curve bounding a simply connected compact invariant set $R \subset \operatorname{Int} D$. For any point $x \in W^s(s) \cap R$, there exists an equilibrium $\lim_{t\to\infty} \phi_{-t} x = p$. But p must be a saddle, violating the assumption of no saddle connections.

A similar argument shows that of the four unstable branches, precisely two limit at a common sink. This suffices to prove (i). Similarly, if there is a saddle connection it is easy to see that (ii) holds. □

Let $\mathcal{V}_\mathsf{T} \subset \mathcal{V}$ denote the closed set of vector fields in $\mathcal{V}$ having a saddle connection; that is, $\mathcal{V}_\mathsf{T}$ is the set of vector fields with T-shaped unstable manifold complexes. It is known that $\mathcal{V}_\mathsf{T}$ is a closed set which is a C^1 submanifold (Abraham and Robbins [1]).

Isotopies of Unstable Manifold Complexes

A *path* in a space E is a map $I \to E$ defined on the closed unit interval $I = [0, 1]$. A path which identifies the endpoints of I is a *loop*.

Let $Y : I \to \mathcal{V}$ be path, that is, a 1-parameter family of vector fields. Denote the unstable manifold complex $K_{Y(t)}$ by K_t.

A *labeling* for Y is an indexed set of five paths $\{\lambda_i : I \to D : i = 1, \ldots, 5\}$ in D such that, for each $t \in I$, the set $\{\lambda_i(t),\ i = 1, \ldots, 5\}$ consists of the five equilibria of $Y(t)$. Even if Y is a loop, it is possible that some of the paths λ_i are not loops. If all the λ_i *are* loops then we say that the labeling $\{\lambda_i\}$ is *consistent*.

Because the equilibria occurring in a 1-parameter family of vector fields in V cannot coalesce (owing to hyperbolicity), every labeling $\{\lambda_i\}$ has the important property that, for all t,

$$\lambda_i(t) \neq \lambda_j(t) \quad \text{if } i \neq j.$$

By continuity, the labeling is completely determined by its initial values $\{\lambda_i(0)\}$. Moreover, for each i, the type (sink or saddle) of the equilibrium $\lambda_i(t)$ remains constant as t varies.

We often denote a labeling by $\{O(t), P(t), Q(t), s_1(t), s_2(t)\}$, with O, P, Q labeling sinks and s_1, s_2 labeling saddles.

LEMMA 1.2. *Every path Y in $\mathcal{V}$ has a labeling.*

Proof. It suffices to prove this locally in I, that is, to show that every point of I has a closed interval neighborhood $J \subset I$ such that $Y|J$ has a labeling. For this it suffices to construct the paths $\lambda_i(t)$ separately for each i.

Fix $t_0 \in I$ and let $p_0 \in \operatorname{Int} D$ be an equilibrium for the vector field $Y(t_0) : D \to \mathbf{R}^2$. We seek a map $\lambda : J \to \mathbf{R}^2$ in a subinterval J containing t_0 such that

$$Y(t)(\lambda(t)) = 0, \tag{2}$$

$$\lambda(t_0) = p_0. \tag{3}$$

Hyperbolicity of p_0 validates application of the implicit function theorem to obtain the desired map λ. □

The path $Y : I \to \mathcal{V}$ is *pure* if either $Y(I) \subset \mathcal{V}_{\mathrm{T}}$ or $Y(I) \subset \mathcal{V} \setminus \mathcal{V}_{\mathrm{T}}$. Thus either all the K_t are T-shaped or all are V-shaped. In the first case we say Y has *type* T and in the second case *type* V.

An *isotopy* of unstable manifold complexes over Y is a map

$$H : K_0 \times I \to D, \quad H(x, t) = H_t(x), \tag{4}$$

with the following properties:

(a) for each $t \in I$, H_t maps K_0 homeomorphically onto K_t,
(b) H_0 is the identity map of K_0.

Each vector field $Y(t)$ has nondegenerate equilibria (Hypothesis 0.3(a)); thus, given a labeling $\{\lambda_i\}$, both $\{\lambda_i\}$ and $\{H_t(\lambda_i(0))\}$ are labelings with the same initial values. Hence, $H_t(\lambda_i(0)) = \lambda_i(t)$ for all $i \in \{1, \ldots, 5\}$ with $t \in I$.

PROPOSITION 1.3. *If $Y : I \to \mathcal{V}$ is a pure path, then there is an isotopy of unstable manifold complexes over Y. In fact, every labeling extends to an isotopy.*

The first thing we need to establish is that the way unstable branches connect equilibria cannot change discontinuously in t. Let $\{\lambda_i\}$ be a labeling for a pure path Y. We abbreviate $W^s(\lambda_i(t), Y(t)) = W^s(i, t)$, and similarly for W^u.

An i, j *connection at* $t \in I$ is a branch of $W^u(i, t)$ going from a saddle $\lambda_i(t)$ to a sink or another saddle $\lambda_j(t)$. Equivalently, it is $W^u(i, t) \cap W^s(j, t)$.

LEMMA 1.4. *If $Y : I \to \mathcal{V}$ is a pure path and there is an i, j connection at some $t_0 \in I$, then there is an i, j connection at each $t \in I$.*

Proof. Fix a compact arc A in the branch $W^u(i, t_0) \cap W^s(j, t_0)$ having $\lambda_i(t_0)$ as an endpoint. Continuity of unstable manifolds (Hirsch and Pugh [5]) means that, for any $\varepsilon > 0$, there exists $\delta > 0$ such that if $|t - t_0| < \delta$ then $W^u(i, t)$ contains a compact arc A_t having $\lambda_i(t)$ as endpoint, and there is diffeomorphism $g_t : A \to A_t$ within ε of the identity map of A in the C^1 topology.

Consider the case where $\lambda_j(t_0)$ is a sink for $Y(t_0)$, so that $W^s(j, t_0)$ is an open set in D. It is well known (see e.g. Hirsch and Smale [6, Chap. 16]) that, for any compact set $Q \subset W^s(j, t_0)$, there is a neighborhood $\mathcal{N}$ of $Y(t_0)$ in space of C^1 vector fields on D such that if $X \in \mathcal{N}$ then X has a unique hyperbolic sink q with the property that $Q \subset W^s(q, X)$. This implies that for t sufficiently near t_0 we have $A \cap W^s(j, t) \neq \emptyset$. It follows that there exists $\delta > 0$ such that $|t - t_0| < \delta$ implies $g_t(A) \cap W^s(j, t) \neq \emptyset$. Since $g_t(A) \subset W^u(\lambda_i(t))$, this exhibits an i, j connection for t near t_0.

Suppose now that $\lambda_j(t_0)$ is a saddle and that $\lambda_i(t_0)$ is also. Then K_{t_0} is T-shaped, so K_t is T-shaped for all t. Of the four unstable branches, three go to sinks and one goes from $\lambda_i(t_0)$ to $\lambda_j(t_0)$. Under sufficiently small perturbations, the corresponding unstable branches go to corresponding sinks, by the preceding continuity considerations. Therefore it must be that, for all t, a branch of $W^u(\lambda_i, t)$ connects to a saddle. Since there are no homoclinic loops, it must connect to $W^u(\lambda_j, t)$. $\square$

Proof of Proposition 1.3. For each i we define H_t on equilibria in K_0 by

$$H_t(\lambda_i(0)) = \lambda_i(t).$$

Let i, j be such that there is an i, j connection at some $t \in I$. Hence, by Lemma 1.4 there is an i, j connection at t for all $t \in I$. We denote this branch by $B_t(i, j) = W^u(i, t) \cap W^s(j, t)$; its closure is a C^1 compact arc $C_t(i, j)$. For each $t \in I$, the union of these arcs is K_t.

We define $H_t(i, j) : C_0(i, j) \to C_t(i, j)$ so that it multiplies length by a constant, as follows. For any $s \in I$ and $z \in C_s(i, j)$, let $A_s(z)$ denote the arc length along $C_s(i, j)$ from $\lambda_i(s)$ to z. For $x \in C_0(i, j)$ we define $H_t(x)$ to be the point $y \in C_t(i, j)$ such that

$$\frac{A_t(y)}{A_t(1)} = \frac{A_0(x)}{A_0(1)}.$$

Continuity of unstable manifolds implies that the resulting map $H(x, t) : K_0 \to K_t$ is continuous and thus a homeomorphism. $\square$

It is easy to see, using purity of Y and continuity of stable manifolds, that for any $t_1, t_2 \in S^1$ the unstable manifold complexes of $X(t_1)$ and $X(t_2)$ are isomorphic as labeled graphs.

COROLLARY 1.5. *All the unstable manifold complexes K_t for a pure path are isomorphic as labeled graphs.*

Winding Numbers

Let $S^1 = \{x \in \mathbf{R}^2 : \|x\| = 1\}$ denote the unit circle. Recall that to each loop $\lambda\colon I \to S^1$ there corresponds an integer $\deg_S(\lambda)$ called its *degree*, defined as follows. Express $\lambda(t)$ continuously as $\lambda(t) = (\cos\theta(t), \sin\theta(t))$. As $\theta(t)$ is well defined modulo 2π, we can define $\deg_S(\lambda) = (\theta(1) - \theta(0))/2\pi$. Loops that are homotopic (through loops) have the same degree.

Let P^1 denote the (real) projective line, that is, the space of lines through the origin in $\mathbf{R}^2$. An element $L \in P^1$ is represented in *homogeneous coordinates* by $[a, b]$, where $(a, b) \neq (0, 0)$ is any point on $L \setminus (0, 0)$. The formula

$$h(\cos\theta, \sin\theta) = \left[\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\right]$$

defines a homeomorphism $h\colon S^1 \to P^1$.

To a loop $\lambda\colon I \to P^1$ we also assign a degree:

$$\deg_P(\lambda) = \deg_S(h^{-1} \circ \lambda),$$

noting that $h^{-1} \circ \lambda$ is a loop in the circle.

Consider now the canonical double covering projection

$$\pi : S^1 \to P^1, \quad (x, y) \mapsto [x, y].$$

The following lemma is well known.

LEMMA 1.6.

(a) *If* $\lambda\colon I \to S^1$ *is a loop, then* $\deg_P(\pi \circ \lambda) = 2\deg_S \lambda$.
(b) *A loop in* P^1 *lifts through* π *to a loop in* S^1 *if and only if its degree is even*.

Recall that $\mathcal{C}_2$ is the space of unordered pairs of distinct points in the disk and that $\hat{\mathcal{C}}_2$ is the space of ordered pairs of such points. The canonical map $\psi : \mathcal{C}_2 \to P^1$ into the (real) projective line P^1 is defined by sending $\{u, v\} \in \mathcal{C}_2$ to the point of P^1 with homogeneous coordinates $[u_1 - v_1, u_2 - v_2]$. Thus $\psi(\{u, v\}) = [u_1 - v_1, u_2 - v_2]$. We denote by $\hat{\psi} : \mathcal{C}_2 \to S^1$ the unique continuous map covering ψ so that the following diagram commutes:

$$\begin{array}{ccc} \hat{\mathcal{C}}_2 & \xrightarrow{\hat{\psi}} & S^1 \\ \downarrow & & \downarrow \\ \mathcal{C}_2 & \xrightarrow{\psi} & P^1, \end{array}$$

where the vertical maps are the canonical projections.

Let $\rho\colon I \to \mathcal{C}_2$ be a loop. To ρ we associate an integer, called its *projective relative winding number* $\mathrm{PRWN}(\rho)$, defined as the degree of the composition

$$I \xrightarrow{\lambda} \mathcal{C}_2 \xrightarrow{\psi} P^1.$$

This is invariant under homotopies of loops in $\mathcal{C}_2$.

LEMMA 1.7. *A loop $\rho\colon I \to \mathcal{C}_2$ lifts to a loop in $\hat{\mathcal{C}}_2$ if and only if* PRWN(ρ) *is even.*

Proof. Consider the commutative diagram

$$\begin{array}{ccccc} & & \hat{\mathcal{C}}_2 & \xrightarrow{\hat{\psi}} & S^1 \\ & & \downarrow & & \downarrow \\ I & \xrightarrow{\rho} & \mathcal{C}_2 & \xrightarrow{\psi} & P^1; \end{array}$$

PRWN(ρ) is even if and only if $\psi \circ \rho$ lifts to a loop $I \to S^1$, by Lemma 1.6. By properties of covering spaces, this is equivalent to ρ lifting to a loop in $\hat{\mathcal{C}}_2$. □

In the rest of this section we assume that $Y\colon I \to \mathcal{V}$ is a pure loop with labeling

$$\{\lambda_i(t)\} = \{O(t), P(t), Q(t), s_1(t), s_2(t)\}$$

in the notation of Section 1. Without loss of generality we assume that the labeling of K_0 is as in Figure 1.

Then for all t, the equilibria $O(t)$, $P(t)$, and $Q(t)$ are sinks, while $s_1(t)$ and $s_2(t)$ are saddles. Since Y is a loop, each sink (or saddle) for $Y(1)$ coincides with a sink (or saddle, respectively) for $Y(0)$, but perhaps with a different label. It turns out (Proposition 1.9) that if Y takes values in $\mathcal{V}_{\mathrm{T}}$ then the labeling is in fact consistent: $\lambda_i(0) = \lambda_i(1)$ for $i = 1, \dots, 5$.

A glance at Figure 1 shows that $P(t)$ and $Q(t)$ can be topologically distinguished from $O(t)$ in K_t (but not from each other). That is, no homeomorphism of K_t can send $P(t)$ or $Q(t)$ to $O(t)$. Moreover, the saddles are topologically distinguishable from the sinks.

Consider a pair of distinct indices $1 \le i, j \le 5$ such that the map

$$g_{ij}\colon I \to \mathcal{C}_2, \quad t \mapsto \{\lambda_i(t), \lambda_j(t)\}$$

is a loop; that is, $\{\lambda_i(0), \lambda_j(0)\} = \{\lambda_i(1), \lambda_j(1)\}$. Denote the projective relative winding number of g_{ij} by $d_{ij}^Y = d_{ij}$. Notice that g_{ij} and g_{ji} are the same map, whence $d_{ij} = d_{ji}$.

PROPOSITION 1.8. *Let $Y\colon [0,1] \to \mathcal{V}$ be a pure loop. Then d_{ij}^Y is the same for all pairs (i, j), where $i, j = 1, \dots, 5$ and $i \neq j$, such that g_{ij} is a loop.*

Proof. Consider two such pairs (i, j) and (k, l). We look for paths $u, v\colon I \to K_0$ such that

(a) $\{u(0), v(0)\} = \{\lambda_i(0), \lambda_j(0)\}$,
(b) $\{u(1), v(1)\} = \{\lambda_k(0), \lambda_l(0)\}$, and
(c) $u(s) \neq v(s)$, $s \in I$.

Given such u and v, we obtain a homotopy from g_{ij} to g_{kl} as follows. Define

$$G_s\colon I \to P^1,$$

$$G_s(t) = \psi(\{H_t(u(s)), H_t(v(s))\}), \quad s, t \in I,$$

where H is an isotopy of unstable manifold complexes covering the path Y (see (4)). Then H_1 is a homeomorphism of K_0 to itself that permutes the equilibria among themselves, and the saddles among themselves, and fixes $O(0)$. Moreover H_1 maps each unstable branch of $Y(0)$ to an unstable branch of $Y(0)$.

Properties (a) and (b) ensure that $G_0 = g_{ij}$ and $G_1 = g_{kl}$, while (c) implies that $\{H_t(u(s)), H_t(v(s))\} \in \mathcal{C}_2$. Thus the existence of such u, v will imply $d_{ij} = d_{kl}$.

If such a pair u, v exists then we write $i, j \simeq k, l$. Standard homotopy theory implies that this is an equivalence relation. It therefore suffices to prove that, for any $i, j, k, l \in \{1, \ldots, 5\}$ with $i \neq j$ and $k \neq l$, we have either $i, j \simeq k, l$ or $i, j \simeq l, k$.

Consider first the case that Y takes values in $\mathcal{V}_\mathsf{T}$. Let the equilibria in K_0 be labeled O, P, Q, s_1, s_2, as in the left-hand side of Figure 1. We extend these labels to a labeling of Y by Lemma 1.2.

In order to prove $O, s_1 \simeq O, s_2$, we define $u\colon I \to K_0$ to be the constant path at O and $v\colon I \to K_0$ to be the path from s_1 to s_2 whose image is the closure of $W^u(s_1) \cap W^s(s_2)$—in other words, that branch of $W^u(s_1)$ which leads to s_2 (the upper branch in Figure 1).

A similar construction shows that $O, s_2 \simeq s_1, s_2$. Now, by transitivity, $O, s_1 \simeq s_1, s_2$.

Proceeding in this way reveals that $i, j \simeq k, l$ for $i \neq j$ and $k \neq l$. This proves the proposition for the T-shaped case. The proof for the V-shaped case of unstable manifold complexes is similar, except that for each such i, j, k, l we can only prove that either $i, j \simeq k, l$ or else $i, j \simeq l, k$. This, however, suffices to complete the proof. $\square$

PROPOSITION 1.9. *Suppose Y is a pure loop in $\mathcal{V}_\mathsf{T}$. Then every labeling is consistent.*

Proof. We assume the labeling is

$$\{\lambda_1(t), \lambda_2(t), \lambda_3(t), \lambda_4(t), \lambda_5(t)\} = \{O(t), P(t), Q(t), s_1(t), s_2(t)\}$$

as usual.

By Lemma 1.7 it suffices to prove that d_{ij} is even when $i \neq j$. By Proposition 1.8 it is enough to prove that $d_{4,5}$ is even or, equivalently, that the saddles cannot be interchanged by an isotopy of T-shaped unstable manifold complexes. This is obvious because in a T-shaped complex the saddles are topologically distinct. $\square$

The analogous statement is false for V-shaped pure loops, as there are loops in $\mathcal{V} \setminus \mathcal{V}_\mathsf{T}$ whose induced isotopy of unstable manifold complexes starts and ends with the right hand diagram in Figure 1 (ignoring labels) but exchanges P and Q.

2. Proof of the Main Theorem

In proving the main theorem we make use of a certain torus $T \subset \mathcal{C}_3$ of configurations. If $\sigma : \mathcal{C}_3 \to \mathcal{V}$ is a section then $\sigma(T)$ is a torus of vector fields. From loops in T we obtain, via σ, isotopies of stable manifold complexes and hence winding numbers.

Tori

The *standard torus* T^2 is the Cartesian product $S^1 \times S^1$ of unit circles. The two projections $T^2 \to S^1$ induce an isomorphism of the fundamental group of T^2 with $\mathbf{Z} \times \mathbf{Z}$. Thus to each loop in T^2 we assign an ordered pair of integers called its *type.* To an oriented embedded circle $\Sigma \subset T^2$ we assign the type of any characteristic loop for Σ. The type of an unoriented embedded circle means the type for some convenient orientation.

LEMMA 2.1. *Let $\Sigma \subset T^2$ be an oriented embedded circle of type (m, n). Then one of the following holds*:

(a) $m = n = 0$; *or*
(b) $|m|$ *and* $|n|$ *are relatively prime.*

Proof. This well-known topological fact (due to Hopf?) can be proved as follows from the theory of covering spaces and the classification of surfaces. Suppose (a) does not hold. Then Σ represents a nontrivial homotopy class $\alpha \in \pi_1(T^2) = \mathbf{Z} \times \mathbf{Z}$. There is a unique maximal positive integer k such that $\alpha = k\beta$ for some nonzero class $\beta \in \pi_1(T^2)$, namely, the greatest common divisor of $|m|, |n|$. If $k = 1$ then the proof is complete.

There is a covering space $p\colon E \to T^2$ such that $\pi_1(E) = \mathbf{Z}$, $p_*(1) = \beta$, and $p_*(k) = \alpha$; we may take $E = \mathbf{R}^2 \setminus 0$. The covering homotopy theorem yields a map $f\colon \Sigma \to E$ such that $p \circ f$ is the identity map of Σ. Therefore $f(\Sigma) \subset E$ is a Jordan curve representing k times the generator of $\pi_1(E)$.

In other words, the winding number of the Jordan curve $f(\Sigma)$ about the origin in $\mathbf{R}^2$ is k, and a classical theorem says $k = 1$. $\square$

We now describe an important torus of configurations (see Figure 2). Define the subset $T \subset \mathcal{C}_3$ to be the set of those unordered triples in the interior of D that can be labeled (O, P, Q), where O is the origin and P, Q are antipodal points on the circle of radius $\frac{1}{3}$ centered at a point at distance $\frac{1}{2}$ from O.

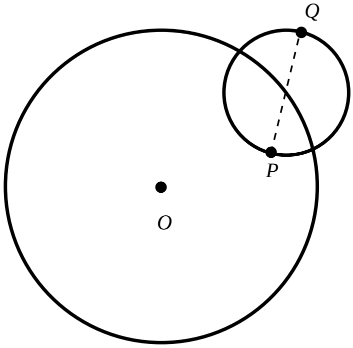

Figure 2 The torus T of configurations in $\mathcal{C}_3$

The torus T is an analytic submanifold of $\mathcal{C}_3$, diffeomorphic to the (nonstandard) torus $S^1 \times P^1$ by the map

$$r = (r_1, r_2)\colon T \to S^1 \times P^1, \quad \{O, P, Q\} \mapsto \left(\frac{P}{\|P\|}, \ \psi(\{P, Q\}) \right),$$

where $\psi\colon \mathcal{C}_2 \to P^1$ is the canonical map defined in Section 1. Notice that r induces an isomorphism r_* of fundamental groups.

A loop $\lambda\colon I \to T$ determines an element in the fundamental group of T, which is isomorphic to $\mathbf{Z} \times \mathbf{Z}$ under r_*. Thus to each loop we associate an ordered pair of integers (m, n), called its *type*: m is the degree of the loop $r_1 \circ \lambda\colon I \to S^1$, while n is the degree of $r_2 \circ \lambda\colon I \to P^1$. From this and the definition of the projective relative winding number (PRWN), the following computation follows easily.

LEMMA 2.2. *Let $\lambda\colon I \to T$, $t \mapsto (O(t), P(t), Q(t))$ be a loop such that $O(t)$ is the origin, $P(0) = P(1)$, and $Q(0) = Q(1)$. Let $a \in \mathbf{Z}$ be the PRWN of the loop in $\mathcal{C}_2$ defined by $t \mapsto (O(t), P(t))$. Similarly, let $b \in \mathbf{Z}$ denote the PRWN of $t \mapsto (P(t), Q(t))$. Then λ has type $(a/2, b)$.*

The following results will be used to show that certain kinds of loops in $\mathcal{V}$ must contain vector fields having T-shaped unstable manifold complexes.

LEMMA 2.3. *Let $\Sigma \subset T$ be an embedded circle. Assume there exists an embedded circle $\Lambda \subset \mathcal{V} \setminus \mathcal{V}_{\mathrm{T}}$ that maps homeomorphically onto Σ under the canonical map $\mu\colon \mathcal{V} \to \mathcal{C}_3$. Then, with suitable orientation, Σ has type $(1, 2)$ or $(0, 0)$.*

Proof. Fix a characteristic map $\nu\colon I \to \Lambda$ for Λ. Then $\nu\colon I \to \mathcal{V}$ is a pure loop in the sense of the preceding sections, with V-shaped unstable manifold complexes.

Consider a labeling $\{\lambda_i\}$ for $\mu \circ \nu$ with

$$\{\lambda_1(0), \lambda_2(0), \lambda_3(0), \lambda_4(0), \lambda_5(0)\} = \{O, P, Q, s_1, s_2\}$$

and $O(t)$ at the origin. The labeling may not be consistent, in which case the loop interchanges P with Q and s_1 with s_2. In any case, $\nu\#\nu$, the composition of the path ν with itself, is consistently labeled.

It follows from Lemma 1.7 that the projective relative winding number d_{12} of $\nu\#\nu$ is even. Therefore, from Proposition 1.8, $d_{ij} = 2n$ for all $i \neq j$. Thus $\nu\#\nu$ has type $(n, 2n)$ by Lemma 2.2, so the degree of Σ is $(n/2, n)$. Since $n/2$ must be integral, Σ has degree $(k, 2k)$. By Lemma 2.1, $|k| = 1$ or 0. □

COROLLARY 2.4. *Let $\Lambda \subset \mathcal{V}$ be a set mapped homeomorphically by $\mu\colon \mathcal{V} \to \mathcal{C}_3$ to an embedded circle $\Sigma \subset T$. If Σ has type $(1, 0)$, then Λ meets $\mathcal{V}_{\mathrm{T}}$ in at least one point.*

The next proposition concerns a circle of type $(1, 2)$, embedded in the torus T, that is covered by a pure type T loop Y in $\mathcal{V}$. Figure 3 suggests such a circle, together with the unstable manifold complex K_t for $Y(t)$.

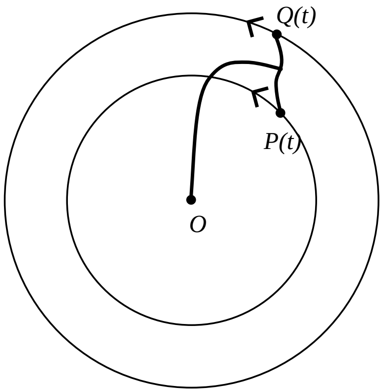

Figure 3 A (1, 2)-circle in T lifting to a type-T loop

PROPOSITION 2.5. *Assume $Q^2 \subset \mathcal{V}$ is a C^1 2-dimensional submanifold transverse to $\mathcal{V}_\mathsf{T}$ that maps homeomorphically onto T under μ. Then $Q^2 \cap \mathcal{V}_\mathsf{T}$ contains an oriented embedded circle Λ such that $\mu(\Lambda)$ has type $(1, 2)$.*

Proof. We first show that $Q^2 \cap \mathcal{V}_\mathsf{T}$ is nonempty. In fact, from Corollary 2.4 we see that, if $\Sigma \subset T$ is any embedded circle of type $(1, 0)$, then the circle $\Lambda = (\mu|Q^2)^{-1}(\Sigma)$ must meet $\mathcal{V}_\mathsf{T}$. Therefore, by the transversality assumption, $Q^2 \cap \mathcal{V}_\mathsf{T}$ is a nonempty compact submanifold of Q^2 of dimension 1 and hence a finite union of disjoint embedded circles $\Lambda_j \subset T$ $(j = 1, \ldots, N,\ N \geq 1)$.

Set $\mu(\Lambda_j) = \Sigma_j \subset T$. As a loop in T, Σ_j has type $(n_j, 2n_j)$ with $|n_j| \leq 1$, by Lemma 2.3. We orient Σ_j so that $n_j \geq 0$.

It cannot be that all $n_j = 0$. To see this, suppose otherwise. Then each Σ_j bounds disk D_j in T and, since the boundaries are disjoint, each pair of these disks is either disjoint or nested. We may therefore change their indexing so that, for some $1 \leq m \leq N$, the maximal disks are disjoint disks $D_1, \ldots, D_m$ and their union contains all the $D_j,\ j = 1, \ldots, N$.

Now, by shrinking $D_1, \ldots, D_m$ isotopically, one can see that there exists a type-$(1, 0)$ embedded circle $\Sigma \subset T$ disjoint from the D_j. Thus Σ is disjoint from $\mu(Q^2 \cap \mathcal{T})$, contradicting the second sentence of the proof. Therefore $n_j = 1$ for some j. $\square$

Completion of Proof of the Main Theorem

The following notation will be used. If β is a loop in a space E, then $[\beta]$ denotes the element of the fundamental group $\pi_1(E)$ represented by β. When β is a characteristic loop for an oriented embedded circle $\Gamma \subset E$, we also denote this element by $[\Gamma]$.

Suppose now there is a section $\sigma : \mathcal{C}_3 \to \mathcal{V}$. Invoking standard approximation techniques, we assume the restriction $\sigma|T$ is smooth and transverse to $\mathcal{V}_\mathsf{T}$.

LEMMA 2.6. *There is an oriented embedded circle* $\Lambda \subset T$ *such that*:

(i) $\sigma(\Lambda) \subset \mathcal{V}_{\mathrm{T}}$, *and*
(ii) Λ *has type* $(1, 2)$.

Proof. Apply Proposition 2.5 to $Q^2 = \sigma(T)$. □

Let $\gamma : \mathcal{V} \to \mathcal{C}_2$ assign to each vector field its set of saddles. We obtain a continuous map $\tau = \gamma \circ \sigma : \mathcal{C}_3 \to \mathcal{C}_2$ and the corresponding homomorphism τ_* of fundamental groups, which in this case are braid groups. Thus a section gives us a homomorphism of braid groups $\tau_* \colon B_3 \to B_2$. Because $\mathcal{C}_2$ admits the projective line as deformation retract, we make the identification $B_2 = \mathbf{Z}$.

The standard presentation of B_3 has generators a_1, a_2 and the relation $a_1a_2a_1 = a_2a_1a_2$; see Figure 4.

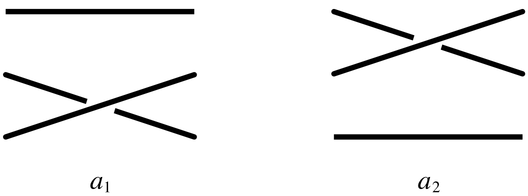

Figure 4 The two generators of B_3

Let $\iota \colon T \to \mathcal{C}_3$ denote the inclusion. Consider the following commutative diagram of continuous maps,

$$\begin{tikzcd} T \arrow[r, "\sigma \circ \iota"] \arrow[d, "\iota"'] & \mathcal{V} \arrow[d, "\nu"] \\ \mathcal{C}_3 \arrow[ur, "\sigma"] \arrow[r, "\tau"'] & \mathcal{C}_2, \end{tikzcd}$$

as well as the corresponding commutative diagram of homomorphisms of fundamental groups,

$$\begin{tikzcd} \pi_1(T) \arrow[r, "\sigma_* \iota_*"] \arrow[d, "\iota_*"'] & \pi_1(\mathcal{V}) \arrow[d, "\nu_*"] \\ B_3 \arrow[ur, "\sigma_*"] \arrow[r, "\tau_*"'] & B_2 = \mathbf{Z}. \end{tikzcd}$$

We will show that the section σ cannot exist by computing $\tau_*\iota_*[\Lambda] \in \mathbf{Z}$ in two different ways. On the one hand, $\tau_*\iota_*[\Lambda] = 2$. But on the other hand,

$$\iota_*[\Lambda] = (a_1a_2)^3 \in B_3,$$

so the fact that τ_* is a homomorphism makes $\tau_*\iota_*[\Lambda]*$ divisible by 3.

To carry out these calculations, fix a characteristic loop for Λ, denoted by

$$\alpha\colon I \to T, \quad t \mapsto \{O, P(t), Q(t)\}.$$

Set $X(t) = \sigma(\alpha(t)) \in \mathcal{V}$. Then $X\colon I \to \mathcal{V}$ is a loop with values in $\mathcal{V}_{\mathsf{T}}$. Let $s_1(t), s_2(t)$ be a continuous labeling of the saddles of $X(t)$. Then Proposition 1.9 implies that we have a consistent labeling of the equilibria of $X(t)$ as

$$\{\lambda_i(t)\} = \{O, P(t), Q(t), s_1(t), s_2(t)\}.$$

Consistency means that each path $\lambda_i\colon I \to D$ is a loop.

With this notation $\gamma(X(t)) = \{s_1(t), s_2(t)\}$, and from the definition of τ we see that $\tau_* \iota_*[\Lambda]$ is the PRWN of the loop $t \mapsto \{s_1(t), s_2(t)\}$ in $\mathcal{C}_2$.

By Proposition 1.8, we can therefore calculate $\tau_* \iota_*[\Lambda]$ as the PRWN of the loop $t \mapsto \{P(t), Q(t)\}$; this loop can also be expressed as

$$g \circ \alpha\colon I \to \mathcal{C}_2,$$

where

$$g\colon T \to \mathcal{C}_2, \quad \{O, P, Q\} \mapsto \{P, Q\}.$$

It follows that $\tau_* \iota_*[\Lambda]$ is the homotopy class of $g \circ \kappa\colon I \to \mathcal{C}_2$, where κ is any loop in the torus homotopic to α; that is, κ can be any $(1, 2)$-loop. For this purpose we choose the *standard* $(1, 2)$*-loop,*

$$\kappa\colon I \to T, \quad t \mapsto \{O, P(t), Q(t)\},$$

for which $P(t)$ and $Q(t)$ are both collinear with O and each wind once around the origin in concentric circles, centered at O, of radii $\frac{1}{6}$ and $\frac{5}{6}$ respectively (see Figure 3). Because the PRWN of $(g \circ \kappa) = 2$, this gives

$$\tau_* \iota_*[\Lambda] = 2.$$

It is interesting that this is independent of the choice of σ.

On the other hand, we can also calculate $\tau_* \iota_*[\Lambda]$ as $\tau_*(\iota_*[\kappa])$. Figure 5 suggests the proof of the fact that

$$\iota_*[\kappa] = (a_1 a_2)^3 \in B_3.$$

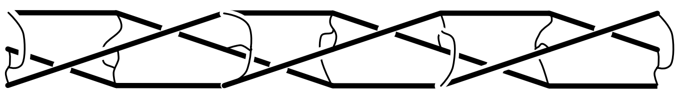

Figure 5 A $(1, 2)$-loop in T yields $(\sigma_1 \sigma_2)^3$ in B_3

Therefore, using additive notation in $B_2 = \mathbf{Z}$, we have

$$\tau_* \iota_*[\Lambda] = 3\tau_*(a_1 a_2).$$

This yields the contradiction that 3 divides 2, completing the proof of Theorem 0.2. $\square$

3. Sections of Morse Bundles

In this section we discuss some settings in which there is a section of the Morse bundle for low values of r, the number of maxima. Our first task is to show that $\mu_r\colon \mathcal{M}_r \to \mathcal{C}_r$ is in fact a fibre bundle. Some of the material on equivariant maps is adapted from Palais [10].

Group Actions, Fibre Bundles, and Sections

A map $p\colon E \to B$ will also be denoted by (p, E, B). For a subset $U \subset B$ we set $E|U = p^{-1}U$.

We say (p, E, B) is *trivial* over an open set U if there is a space F (called the *standard fibre*) and a commutative diagram

$$\begin{tikzcd} U \times F \arrow[r, "h"] \arrow[d, "p_1"'] & E|U \arrow[d, "p"] \\ U \arrow[r, "="'] & U, \end{tikzcd}$$

where h is a homeomorphism and p_1 is projection on the first factor. We call h a *local trivialization* of (p, E, B) over U. If $U = B$ then we call h a *global trivialization* and say that (p, E, B) is globally trivial. Notice that a globally trivial map (p, E, B) has a cross-section.

If there is a local trivialization over each open set of some open cover of B, then (p, E, B) is *locally trivial* map, also called a *fibre bundle.* For the theory of fibre bundles see the books by Hu [7], Husemoller [8] or Steenrod [14].

Let G be a topological group with identity element e, acting continuously on a topological space Z through homeomorphisms. Thus there is given a map $\rho\colon G \times Z \to Z$, written

$$\rho(g, x) = \rho_g(x) = g \cdot x,$$

such that the map $g \mapsto \rho_g$ is a homomorphism into the group of homeomorphisms of Z.

Let $U \subset Z$ be an open set and let $z \in U$. A *local transsection* at (U, z) of the action ρ is a map $\sigma\colon (U, z) \to (G, e)$ such that $\sigma(x) \cdot z = x$ for all $x \in U$. If $U = Z$ then we call σ a *global transsection*. We say that G acts *equivariantly* on (p, E, B) if we have actions of G on E and B such that $g \cdot p(x) = p(g \cdot x)$ for all $x \in E$.

PROPOSITION 3.1. *Let G act equivariantly on (p, E, B).*

(a) *If the action of G on B has a global transsection, then (p, E, B) is globally trivial.*

(b) *If every $b \in B$ belongs to an open set $U \subset B$ such that the action of G on B has a local transsection at (U, b), then (p, E, B) is locally trivial.*

(c) *Suppose B is a manifold and G is a Lie group, and suppose the action of G on B is simply transitive. Then (p, E, B) is globally trivial and therefore has a cross-section.*

Proof. Suppose $\rho\colon (B, b) \to (G, e)$ is a global section for ρ. Set $F = E_b$. Then the map

$$h\colon B \times E_b \to E, \quad (x, y) \mapsto \rho(x) \cdot y$$

is a global trivialization. This proves (a), and (b) follows.

To prove (c), fix any $b \in B$. For $x \in B$ define $\sigma(x) \in G$ to be the unique group element g such that $g \cdot b = x$. The resulting map $\sigma\colon B \to G$ is the inverse to the evaluation map at b: $ev_b\colon G \to B$, $g \mapsto g \cdot b$. As G and B are locally Euclidean and ev_b is continuous and bijective, it follows from invariance of domain that its inverse σ is continuous. □

THEOREM 3.2. *For every r, the following maps are fibre bundles*: $(\mu_r, \mathcal{M}_r, \mathcal{C}_r)$, $(\hat{\mu}_r, \hat{\mathcal{M}}_r, \hat{\mathcal{C}}_r)$, *and* $(\mu_r, \mathcal{V}_r, \mathcal{C}_r)$.

Proof. It is easy to see that the group of C^2 diffeomorphisms of the disk acts equivariantly on each of these three maps with local transsections, so that Proposition 3.1(b) applies. □

A classical topological method of obtaining sections is the following.

PROPOSITION 3.3. *Let (p, E, B) be a locally trivial map of metric spaces. Let $B_1 \subset B$ be a closed subspace that is a deformation retract of B. Then every section of $E\,|\,B_1$ extends to a global section of (p, E, B).*

Proof. This follows from the homotopy lifting property of fibre bundles (Hu [7]). □

A Canonical Morse Function

Let $C = \{c_1, \ldots, c_r\} \in \mathcal{C}_r$ be any set of $r \geq 1$ distinct points in $\mathbf{R}^2$. Define a map $f_C\colon \mathbf{R}^2 \to \mathbf{R}$ by

$$f_C(z) = \prod_{j=1}^{r} \|z - c_j\|^2. \tag{5}$$

Denote by $p_C\colon \mathcal{C}_C \to \mathcal{C}_C$ the monic polynomial

$$p_C(z) = \prod_j (z - c_j) \tag{6}$$

whose set of roots is C, so that $f_C(z) = |p_C(z)|^2$.

The goal of this subsection is to prove the following theorem.

THEOREM 3.4.

(a) *f_C is a Morse function if and only if the polynomials p_C and p'_C have only simple roots.*

(b) *Suppose $r = 3$. Then f_C is a Morse function if and only if C is not the set of vertices of an equilateral triangle.*

For the proof it is convenient to identify complex numbers with vectors in $\mathbf{R}^2$. The *real* inner product of $z, w \in \mathbf{C} = \mathbf{R}^2$ is denoted by $\langle z, w\rangle$.

Multiplication by the complex number $z = a + ib$ determines a linear operator in $\mathbf{R}^2$ whose matrix in the standard basis is

$$M(z) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

With this notation we have $wz = M(w)z$. Note that $M(\bar{z}) = M(z)^{\mathrm{T}} =$ the transpose of $M(z)$.

Let $p(z)$ be a complex analytic function defined in some open subset W of the plane. We consider p as a map $\mathbf{R}^2 \to \mathbf{R}^2$; then its real derivative DF_z is a linear operator on $\mathbf{R}^2$ whose matrix is $M(p'(z))$.

The conjugation $z \mapsto \bar{z}$ defines a real (not complex) linear transformation of $\mathbf{R}^2$ whose matrix is denoted by

$$R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Define a real-valued function f on the plane by

$$f(z) = \tfrac{1}{2}|p(z)|^2 = \tfrac{1}{2}\langle p(z), p(z)\rangle.$$

Denote its gradient vector at z by ∇f_z and its Hessian matrix by Hf_z.

LEMMA 3.5.

$$\nabla f_z = \overline{p'(z)}\, p(z), \tag{7}$$

$$Hf_z = M(\overline{p''(z)}\,\overline{p(z)})R + |p'(z)|^2 I. \tag{8}$$

Proof. From the chain rule we have that, for every $\xi \in \mathbf{R}^2$,

$$Df_z\xi = \langle p(z), M(p'(z))\xi\rangle \tag{9}$$

$$= \langle M(p'(z))^T p(z), \xi\rangle, \tag{10}$$

$$Df_z\xi = \langle \overline{p'(z)}\, p(z), \xi\rangle, \tag{11}$$

which is equivalent to (7). Differentiating (11) with respect to z along ξ, we obtain

$$Hf_z(\xi, \xi) = \langle \overline{p''(z)}\,\bar{\xi}\, p(z), \xi\rangle + \langle \overline{p'(z)}\, p'(z)\xi, \xi\rangle$$
$$= \langle \overline{p''(z)}\, p(z)R\xi, \xi\rangle + |p'(x)|^2\langle \xi, \xi\rangle,$$

which yields (8). □

COROLLARY 3.6. *The function $|p(z)|^2$is a Morse function if and only if both p and p' have only simple zeroes.*

Proof. From (7) we see that z_0 is a critical point of $|p(z)|^2$ if and only if z_0 is a zero of p' or p.

Suppose $p(z_0) = 0$. Then equation (8) shows that z is a degenerate critical point of $|p(z)|^2$ if and only if $p'(z_0) = 0$, which is equivalent to z_0 being a multiple root of p.

If $p'(z_0) = 0$, then (8) shows that z_0 is a degenerate critical point of $|p(z)|^2$ if and only if $p''(z_0) = 0$; that is, z_0 is a multiple root of p'. $\square$

COROLLARY 3.7. *The gradient vector at z of the function $f_C(z) = |p_C(z)|^2$ is*

$$\nabla_z f_C = |p_C(x)|^2 \sum_{j=1}^{r} \frac{(z - c_j)}{|z - c_j|^2}.$$

Proof. Follows from (6) and (7). $\square$

COROLLARY 3.8. *Let $C = \{c_1, \ldots, c_r\}$ be a set of r distinct points in the open unit disk. Then $\nabla_z f_C$ points outward at every boundary point z of the disk.*

Proof. Follows from Corollary 3.7, because $\langle z, z - c_j \rangle > 0$ if $|z| = 1$ and $|c_j| < 1$. $\square$

Proof of Theorem 3.4. Part (a) follows from Corollary 3.6. To prove (b), we set $C = \{a, b, c\}$ and apply Corollary 3.6; thus we need to show that a, b, c are equidistant from each other if and only if p'_C has a double root. Now

$$p'_C(z) = 3z^2 - 2(a + b + c)z + ab + bc + ca,$$

and its discriminant is

$$(a + b + c)^2 - 3(ab + bc + ca) = a^2 + b^2 + c^2 - (ab + bc + ca).$$

Therefore f_C is not a Morse function if and only if

$$a^2 + b^2 + c^2 - ab + bc + ca = 0.$$

Since the hypotheses and conclusion are unchanged if we replace f_C by $f \circ H$ where $H : \mathbf{R}^2 \to \mathbf{R}^2$ is a Euclidean isometry or homothety, we may assume that $c = 0$ and $b = 1$. Then f_C is a non-Morse function if and only if $a^2 + 1 - a = 0$. But this equation holds if and only if a, 1, and 0 form an equilateral triangle. $\square$

Sections of Some Morse Bundles

Let $\mathcal{S}_r \subset \mathcal{C}_r$ denote the set of configurations C of r points in the plane for which the polynomial p'_C (see Equation (6)) does not have simple roots. Let $\hat{\mathcal{S}}_r \subset \hat{\mathcal{C}}_r$ denote the set of labeled sets of r points whose unlabeled configurations lie in $\mathcal{S}_r$. It is easy to see that $\mathcal{S}_r \subset \mathcal{C}_r$ and $\hat{\mathcal{S}}_r \subset \hat{\mathcal{C}}_r$ are algebraic varieties of codimension 2. If $C \in \mathcal{C}_r \setminus \mathcal{S}_r$ then $-f_C$ is a Morse function satisfying Hypothesis 0.1 by Theorem 3.4(b) and Corollary 3.8.

Let $\upsilon : \hat{\mathcal{C}}_r \to \mathcal{C}_r$ denote the natural covering space projection. We define the *canonical sections*

$$\kappa_r : \mathcal{C}_r \setminus \mathcal{S}_r \to \mathcal{M}_r, \quad C \mapsto -f_C, \tag{12}$$

and

$$\hat{\kappa}_r : \hat{\mathcal{C}}_r \setminus \hat{\mathcal{S}}_r \to \hat{\mathcal{M}}_r, \quad \hat{C} \mapsto -f_{\upsilon}(C). \tag{13}$$

In Section 2 we defined a certain torus $T \subset \mathcal{C}_3$ and proved that no section of $\mathcal{V}_3|T$, and a fortiori of $\mathcal{M}_3|T$, extends to a global section. We did not, however, construct a section of $\mathcal{M}_3|T$. We now see from Corollary 3.8 and Theorem 3.4(b) that κ_3 defines such a section, as it is clear that no equilateral triangle has its set of vertices in T.

From Proposition 3.1(b) we derive the following.

PROPOSITION 3.9. *Let $\mu: \mathcal{M} \to \mathcal{C}$ be the bundle of Morse functions over a submanifold $\mathcal{C} \subset \mathcal{C}_r$ of configurations of maxima. Let G be a Lie group of diffeomorphisms of the disk D such that the induced action of G on $\mathcal{C}_r$ is simply transitive on $\mathcal{C}$. Then there exists a section of $(\mu, \mathcal{M}, \mathcal{C})$. In fact, for any Morse function $F \in \mathcal{M}$ we obtain a section $\sigma : \mathcal{C} \to \mathcal{M}$ through F by defining*

$$\sigma(C) = F \circ g_C^{-1}, \tag{14}$$

where $g_C \in G$ is the unique group element such that $g_C(\mu(F)) = C$.

SCHOLIUM. We can choose the section σ such that, for every $C \in \mathcal{C}$, the Morse function $\sigma(C)$ takes the same constant value y on the boundary of D. This can be proved from Hypothesis 0.1 by deforming any section, but it follows instantly from equation (14) by choosing F to have constant boundary value y. Such boundary conditions are motivated by the original robotics formulation of the section problem by Rimon and Koditschek [11].

There are many other possibilities; for example, we could make all the Morse functions in the section have exactly two critical points when restricted to the boundary.

EXAMPLE 3.10. To obtain a section of the Morse bundle over $\mathcal{C}_1$, we can simply use the canonical section κ_1 defined previously (since $\mathcal{S}_1$ is empty). Alternatively, to specify boundary behavior we can use Proposition 3.9 and its Scholium, since the group H of hyperbolic translations of the Poincaré disk acts simply transitively on $\mathcal{C}_1$. Thus we have our last corollary.

COROLLARY 3.11. *The map $\mu_1: \mathcal{M}_1 \to \mathcal{C}_1$ has a section with constant boundary values.*

EXAMPLE 3.12. Consider $\hat{\mu}_2: \hat{\mathcal{M}}_2 \to \hat{\mathcal{C}}_2$, the bundle of Morse functions with two *labeled* maxima in the open disk, as well as the corresponding unlabeled bundle $\mu_2: \mathcal{M}_2 \to \mathcal{C}_2$. Clearly, $\hat{\kappa}_2$ and κ_2 are sections.

EXAMPLE 3.13. Consider $\hat{\mu}_3: \hat{\mathcal{M}}_3 \to \hat{\mathcal{C}}_3$, the Morse functions with three labeled maxima. Take $B_1 \subset \hat{\mathcal{C}}_3$ to be the set of labeled configurations (P, Q, R) such that the distance $\delta(P, Q, R)$ from P to the (Euclidean) line QR is less than $\|Q - R\|/2$, which prevents the triangle (P, Q, R) from being equilateral.

It is easy to see that B_1 *is a deformation retract of* $\hat{\mathcal{C}}_3$: We deform the ordered triple of distinct points $(P, Q, R) \in \hat{\mathcal{C}}_3$ along the path $t \mapsto (P_t, Q, R)$, $0 \le t \le 1$, defined as follows. Let $P' \in QR$ be the point nearest to P. Define

$$P_t = (1 - t)P + t(0.1P + 0.9P').$$

Because the standard section $\hat{\kappa}_3$ maps B_1 into $\hat{\mathcal{M}}_3$, application of Proposition 3.3 yields the following.

THEOREM 3.14. *The map* $\hat{\mu}_3 \colon \hat{\mathcal{M}}_3 \to \hat{\mathcal{C}}_3$ *has a section.*

4. Open Questions

There is an abundance of open and easily stated problems in this area.

QUESTION 4.1. What spaces of configurations on which manifolds admit sections of the bundle of Morse function?

This paper is a first attack on this question. The global nature of the question makes it difficult to use results on one configuration space for another. An obvious conjecture, of which Theorem 0.2 is just the case $r = 3$, is the following.

CONJECTURE 4.2. For $r \geq 3$, the map $\mu_r \colon \mathcal{M}_r \to \mathcal{C}_r$ does not have a section.

Many interesting questions arise concerning the topology of Morse bundles; we give one example.

QUESTION 4.3. What is the topology of the standard fibre $\mathcal{F}_r$ of the bundle of Morse functions $(\mu_r, \mathcal{M}_r, \mathcal{C}_r)$? What is the primary obstruction to a cross-section?

ACKNOWLEDGMENT. We thank the anonymous referee for several improvements.

References

[1] R. Abraham and J. Robbin, *Transversal mappings and flows,* Benjamin, New York, 1967.
[2] J. Cerf, *Isotopie et pseudo-isotopie,* Proceedings of International Congress of Mathematicians (Moscow), pp. 429–437, 1966.
[3] M. W. Hirsch, *Differential topology,* Graduate Texts in Mathematics, 33, Springer, New York, 1976.
[4] ———, *The work of Stephen Smale in differential topology,* From Topology to Computation: Proceedings of the Smalefest (New York) (M. W. Hirsch, J. E. Marsden, and M. Shub, eds.), pp. 83–106, Springer, New York, 1993.
[5] M. W. Hirsch and C. C. Pugh, *Stable manifolds and hyperbolic sets,* Global Analysis (Providence, RI) (S.-S. Chern and S. Smale, eds.), Proc. Sympos. Pure Math, 14, pp. 133–163, Amer. Math. Soc., Providence, RI, 1970.
[6] M. W. Hirsch and S. Smale, *Differential equations, dynamical sytems, and linear algebra,* Academic Press, New York, 1974.
[7] S.-T. Hu, *Homotopy theory,* Academic Press, New York, 1959.
[8] D. Husemoller, *Fibre bundles,* McGraw-Hill, New York, 1966.
[9] A. F. Möbius, *Theorie der elementaren Verwandschaft,* Leipziger Sitzungberichte math.-phys., Classe 15 (1869) [also in *Werke,* Bd. 2].
[10] R. S. Palais, *Local triviality of the restriction map for embeddings,* Comment. Math. Helv. 34 (1960), 305–312.

[11] E. Rimon and D. E. Koditschek, *Robot navigation functions on manifolds with boundary,* Adv. in Appl. Math. 11 (1990), 412–442.
[12] S. Smale, *Generalized Poincaré conjecture in dimensions greater than four,* Ann. of Math. (2) 74 (1961), 391–406.
[13] ———, *On gradient dynamical systems,* Ann. of Math. (2) 74 (1961), 199–206.
[14] N. Steenrod, *The topology of fibre bundles,* Princeton University Press, Princeton, NJ, 1951.

M. D. Hirsch
Department of Mathematics
Emory University
Atlanta, GA 30322

hirsch@mathcs.emory.edu

M. W. Hirsch
Department of Mathematics
University of California
Berkeley, CA 94720-3840

hirsch@math.berkeley.edu