Lecture IX: The Pseudo-Isoptoy Theorem: Statement, First Reductions, and Method of Proof

November 11, 2020

1 The Definitions and the Statement and Corollaries

Let W be a compact smooth oriented manifold. Diff(W) denotes the group of orientation-preserving diffeomorphisms of W with the C^∞-topology.

Definition 1.1. Let V be a compact smooth oriented manifold without boundary. A *pseudo-isotopy* of V is a diffeomoprhism $\Phi\colon V \times I \to V \times I$ whose restriction to $V \times \{0\}$ is the identity. The set of pseudo-isotopies of V is identified with the subgroup of Diff$(V \times I)$ of diffeomorphism whose restriction to $V \times \{0\}$ is the identity.. The group off pseudo-isotopies of V is denoted $\mathcal{P}(V)$, or simply $\mathcal{P}$ if V is clear from context. An *isotopy* of V is a level-preserving element of $\mathcal{P}(V)$. These form a subgroup $\mathcal{I}(V) \subset \mathcal{P}(V)$ called the *isotopy group* of V.

For Φ is a pseudo-isotopy, we denote by $\Phi_1\colon V \to V$ the diffeomorphism that is the restriction of Φ to $V \times \{1\}$. The group $\mathcal{P}(V)$ acts of Diff(V) (on the right) by $g \cdot \Phi = g \circ \Phi_1$. Diffeomorphisms g_1 and g_2 are said to be *pseudo-isotopic* if they are in the same orbit of the action of $\mathcal{P}(V)$. Restricting to $\mathcal{I}(V)$ gives an action of the group of isotopies of V on Diff(V). Two diffeomorphisms are said to be *isotopic* if they are in the same orbit of this action. Clearly, if two diffeomorphisms are isotopic then they are pseudo-isotopic.

Notice that since $\mathcal{I}(V)$ is the space of paths in Diff(V) based at the identity diffeomorphism, $\mathcal{I}(V)$ is contractible and in particular is connected.

Theorem 1.2. *(Cerf) Let V be a closed smooth connected manifold. Suppose $\pi_1(V) = 0$ and the dimension of V is ≥ 5. Then $\mathcal{P}(V)$ is connected.*

Corollary 1.3. *With V as in the theorem, two orientation-preserving diffeomorphisms of V are isotopic if and only if they are pseudo-isotopic.*

Proof. (that the theorem implies the corollary) Let diffeomorphisms h_0 and h_1 be pseudo-isotopic. Then there is a pseudo-isotopy Φ such that $h_1 = h_0 \circ \Phi_1$. According to the theorem here is a path $\Phi(s)$ of pseudo-isotopies from the identity to $\Phi = \Phi(1)$. Then $h_0 \cdot \Phi(s) = h_0 \circ \Phi(s)(1)$ is an isotopy of h_0 to h_1. □

Definition 1.4. An oriented smooth manifold homotopy equivalent to S^n is a *homotopy n-sphere.*

Suppose that Σ_0 and Σ_1 are homotopy n-spheres. Then their connected sum $\Sigma_0 \# \Sigma_1$ ia a homotopy n-sphere. This determinies a commutative semi-group structure on the set of orientation-preerving diffeomorphism classes of homotopy n-spheres.

Definition 1.5. The semi-group of orientation-preserving diffeomorphism classes of oriented smooth manifolds homotopy equivalent to S^n is denoted Γ_n.

Lemma 1.6. *For $n \geq 5$, Γ_n is an abelian group.*

Proof. For any homotopy n-sphere Σ, denote by $-\Sigma$ the same smooth manifold with the opposite orientation. The smooth manifold $\Sigma \#(-\Sigma)$ is diffeomorphic to the boundary of $W = (\Sigma \setminus B^n) \times I$. the manifold W is homotopy equivalent to B^{n+1}, removing a ball from the interior of W gives an h-cobordism from $\Sigma \#(-\Sigma)$ to S^n. Since $n \geq 5$, then by the h-cobordism theorem, $\Sigma \#(-\Sigma)$ is diffeomorphic to S^n and hence $(-\Sigma)$ is the inverse of Σ in Γ_n □

Corollary 1.7. *For $n \geq 6$*

- $\pi_0(\mathrm{Diff}(D^n) = 0$
- $\pi_0(\mathrm{Diff}(S^{n-1})) \equiv \Gamma_n$

Proof. Given a diffeomorphism of D^n there is an isotopopy of it to a diffeomorphism g that is the identity on the ball $D^n(1/2)$ of radius $1/2$. The restriction of g to $D^n \setminus \mathrm{int}(D^n(1/2)$ is identified with a pseudo-isotopy. Deforming this through pseudo isotopies to the identity produces an isotopy of g to the identitty.

By the h-cobordism theorem we know that any homotopy n-sphere ($n \geq 6$) is obtained by gluing two balls together by a diffeomorphism of their boundaries and the result is diffeomorphism to S^n if and only if the diffeomorphism extends on the ball. Thus, Γ_n is identified with the quotient of

$$\pi_0(\mathrm{Diff}(D^n)) \to \pi_0(\mathrm{Diff}(S^{n-1})),$$

thus, the second statement follows from the first. □

2 First Reductions: Connections to Morse Theory

We fix V a compact, connected smooth manifold without boundary.

Definition 2.1. Let $\mathcal{F}(V) = \mathcal{F}$ be the space of smooth functions

$$f : V \times I \to I$$

without critical points on the boundary and with $f^{-1}(i) = V \times \{i\}$ for $i = 0, 1$, endowed with the C^∞-topology. This is an open convex subset of the Frechet affine space $\mathcal{G}$ of C^∞ functions $g : V \times I \to \mathbb{R}$ with $g^{-1}(i) = V \times \{i\}$ for $i = 0, 1$. Let $\mathcal{E}(V) \subset \mathcal{F}(V)$ be the open subspace of functions without critical points.

The group $\mathcal{P}(V)$ acts on $\mathcal{F}(V)$ from the left by

$$\Phi \cdot f = f \circ \Phi^{-1}.$$

This action stabilizes $\mathcal{E}(V)$.

Define a map $p : \mathcal{P}(V) \to \mathcal{E}(V)$ as follows. Given a pseudo-isotopy $\Phi : V \times I \to V \times I$, we define $p(\Phi) = f_\Phi : V \times I \to I$ to be $\pi_2 \circ \Phi^{-1}$. Since Φ is a diffeomorphism, $f_\Phi \in \mathcal{E}(V)$; i.e., it has no critical points. The level sets of $f_\Phi^{-1}(t)$ are the images $\Phi(V \times \{t\})$ of the natural level sets of $V \times I$. Also, $\mathcal{I}(V)$ is the fiber $p^{-1}(\pi_2)$, where $\pi_2 : V \times I \to I$ is projection to the second factor..

Fix be a Riemannian metric G_V on V and let $G = G_V + dt^{\otimes 2}$ be the product of this metric with the Euclidean metric on the interval. We use this metric to construct a map $\mathcal{E}(V) \to \mathcal{P}(V)$ that is a cross section of p.

For any $g \in \mathcal{E}(V)$ construct the vector field ∇g using the metric G, and let $a_g : V \times I \to V$ be the map that associates to any $w \in V \times I$ the initial point in $V \times \{0\}$ of the flow line for ∇g through w. The map $w \mapsto (a_g(w), g(w))$ is a pseudo-isotopy $\Phi(g) : V \times I \to V \times I$. It is clear that for any $t \in I$, we have $\Phi(g)(V \times t) = g^{-1}(t)$. That is to say $\pi_2 \circ \Phi(g)^{-1} = g$. This means that $p \circ \Phi(g) = g$ so that $g \mapsto \Phi(g)$ is a cross section of $p : \mathcal{P}(V) \to \mathcal{E}(V)$.

Now let us consider the fibers of $p : \mathcal{P}(V) \to \mathcal{E}(V)$.

Claim 2.2. *Two pseudo-isotopies Φ and Ψ are in the same fiber of p if and only if there is an isotopy I such that $\Phi \circ I = \Psi$.*

Proof. Two elements in $\Phi, \Psi \in \mathcal{P}(V)$ are in the same fiber fi and only if

$$\pi_2 \circ \Phi^{-1} = \pi_2 \circ \Psi^{-1}$$

which holds if and only if

$$\pi_2 \circ (\Phi^{-1} \circ \Psi) = \pi_2.$$

The second equation is the statement that $\Phi^{-1} \circ \Psi$ is level-preserving and hence an isotopy. Clearly $\Phi \circ (\Phi^{-1} \circ \Psi) = \Psi$. □

Proposition 2.3. *Let $\Phi\colon V \times I \to V \times I$ be a pseudo-isotopy. Then there is a path of elements $g_s \in \mathcal{E}(V)$, $0 \leq s \leq 1$, connecting f_Φ to the projection $V \times I \to I$, if and only there is a one-parameter family of pseudo-isotopies connecting Φ to an isotopy or equivalently a one-parameter family of pseudo-isotopies connecting Φ to the identity.*

Proof. Since $\mathcal{I}(V)$ is contractible and hence path connected, the last equivalence is clear.

If there is a one-parameter path ω in $\mathcal{P}(V)$ from Φ to an isotopy, then $p(\omega)$ is a path in $\mathcal{E}(V)$ connecting f_Φ to π_2.

Conversely, if there is a path $\overline{\omega}$ in $\mathcal{E}(V)$ connecting f_Φ to π_2, then by the above claim $\Phi(\overline{\omega})$ is a one-parameter family of pseudo-isotopies from Ψ to an isotopy and $\Psi = \Phi \circ J$ for some isotopy J. But since $\mathcal{I}(V)$ is contractible, there is a one-parameter family of isotopies $J(s)$ connecting Id to J. The composition $\Phi \circ J(s)$ is then a one parameter family of pseudo-isotopies connecting Φ to $\Phi \circ J = \Psi$. The concatenation of these two paths gives the required path of pseudo-isotopies connecting Φ to an isotopy. □

We come to the main technical result that we shall estblbish.

Theorem 2.4. *If $\pi_1(\mathcal{F}(V), \mathcal{E}(V)) = \{*\}$, i.e., if every path in $\mathcal{F}(V)$ with end points in $\mathcal{E}(V)$ deforms relative to its end points to a path in $\mathcal{E}(V)$, then every pseudo-isotopy of V deforms through pseudo-isotopies to the identity. In particular, in this case any diffeomorphism $\varphi\colon V \to V$ pseudo-isotopic to the identity is isotopic to the identity.*

Proof. Since $\mathcal{F}(V)$ is a convex subset in an affine submanifold of a Frechet vector space, it is contractible. Thus, the statement that $\pi_1(\mathcal{F}(V), \mathcal{E}(V)) = \{*\}$ is equivalent to the statement that $\mathcal{E}(V)$ is path connected. The result then follows from Proposition 2.3. □

We have reduced the question of whether or not, for a closed manifold V, pseudo-isotopy implies isotopy to the question of whether or not the subspace of functions $g \in \mathcal{E}(V)$ is path connected, or equivalently, whether any path of function in $\mathcal{F}(V)$ with end points in $\mathcal{E}(V)$ is homotopic, relatie to the end points, to a path in $\mathcal{E}(V)$.

Theorem 2.5. *Suppose that $\pi_1(V) = 0$ and the dimension of V is at least 5. Then $\pi_1(\mathcal{F}(V), \mathcal{E}(V)) = \{*\}$, i.e., every path in $\mathcal{F}(V)$ with end points in $\mathcal{E}(V)$ deforms relative to its end points to a path in $\mathcal{E}(V)$.*

Corollary 2.6. *With V as in the statement of the theorem, then any pseudo-isotopy of V is connected by a path of pseudo-isotopies to the identity and in particular pseudo-isotopy implies isotopy for V.*

3 Method of Proof

3.1 The Stratification of $\mathcal{F}(V)$

There is a stratification of $\mathcal{F} = \mathcal{F}(V)$. We denote by $\mathcal{F}^i$ the strata of codimension i. The open set $\mathcal{F}^0$ is the space of Morse functions with distinct critical values. The space of $\mathcal{F}^1$ has two types of components: (i) those whose points are Morse functions with two critical points with the same critical value and all other critical points with distinct values and (ii) those with one degenerate critical point, a birth, and other than than all critical points non-degenerate and values of distinct critical points being different.

The first step in the argument is to show that any path in $(\mathcal{F}, \mathcal{E})$ is homotopic relative to its endpoints to a smooth path in $\mathcal{F}^0 \cup \mathcal{F}^1$ that meets $\mathcal{F}^1$ transversely. In particular, there are only finitely many parameter values at which the path meets $\mathcal{F}^1(V)$. We call such paths *good* paths.

We display the information about the family of critical points and their critical values in a graghic. The *graphic* of a good path $\{f_s\}_{0 \le s \le 1}$ in $\mathcal{F} \cup \mathcal{F}^1$ that crosses $\mathcal{F}^1$ transversally and starts and ends at a function without critical points is a graph in the unit square in the (s,t)-pane which is a smooth codimension-1 submanifold except at a finite set of points. The exceptional points are crossing and cusps. Each smooth arc in the graph projects diffomorpically onto a sub-infterval of the parameter interval. The graph is disjoint from the boundary of the square. The horizontal direction is the parameter s in the homotopy and the vertical direction is the parameter for the interval direction in $V \times I$. The intersection of the graph with the vertical line $\{s = s_0\}$ is the critical values of f_{s_0}. We can also label each

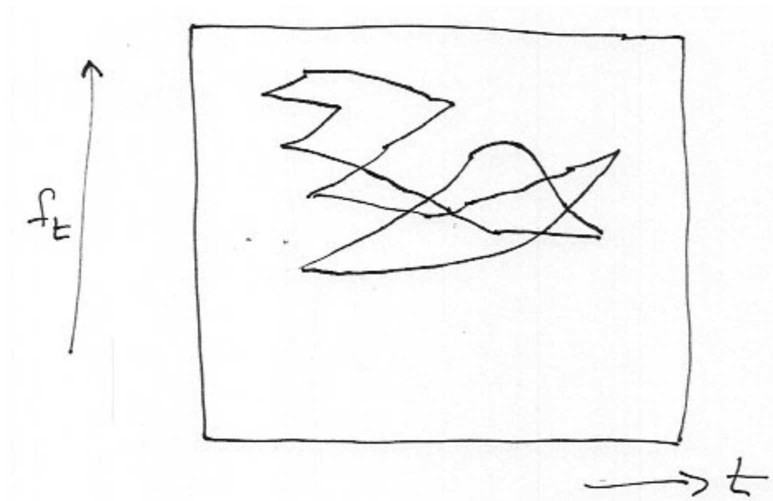

Figure 1: Graphic of a path f_t.

component of the smooth locus with the index of the critical point. See Figure 1

3.2 Analogue of Self-Indexing

Once we have a good family in $\mathcal{F}^0 \cup \mathcal{F}^1$ we perform moves so as to arrange that the critical points are ordered in the sense that for every s all non-degenerate critical points of f_s are ordered in the sense that if p and q are non-degenerate critical points with the index of f_s at p is less than its index at q the $f_s(p) < f_s(q)$. This implies that for every s, if there is a cusp for f_s at which a pair of critical points of index $i+1$ and i are created or annihilated then the value of f_s at this degenerate critical point lies above all critical values for critical points of index i and below all critical values for critical points of index $i+1$.

In order to arrange the critical points in this way, we need moves that remove a pair of crossings of two components of the smooth locus of the graph, a move that allows us the pass a component of the smooth locus across a crossing point, and a move that allows us to pass a component of the smooth locus across a cusp. The first move follows from the fact that crossings are unique up to homotopy, The second and third moves are accomplished by moving the path, relative to its endpoints, across a component of $\mathcal{F}^2$. The second move uses the components of $\mathcal{F}^2$ given by the condition that all critical points are Morse, one critical value has three

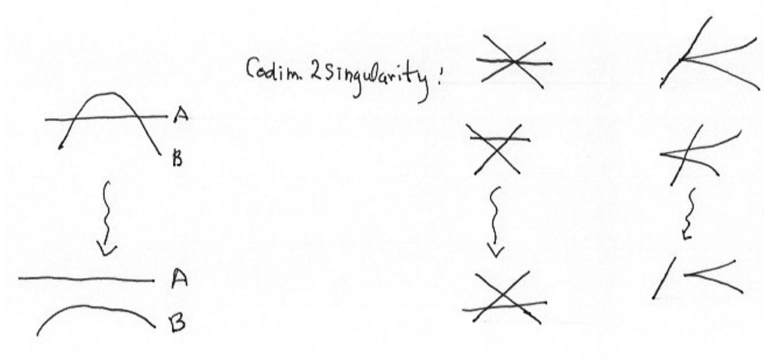

Figure 2: Moves required to order the critical points.

critical points in its pre-image, and all other critical values are distinct. The third move uses components of $\mathcal{F}^2$ given by the condition that a birth and a non-degenerate critical point have the same and all other critical points are non-degenerate with distinct critical values. See Figure 2

3.3 Concentrating the indices in two adjacent dimensions

Up to now in the argument, we have only rearranged the critical graphic; we have not removed any critical points of any of the f_s. To remove critical points we must use the remaining type of codimension-2 strata: the *swallow tall*. The model for a swallow tail singularity is

$$f(x, y, u) = -|x|^2 + |y|^2 - u^4$$

in coordinates (x, y, u) for $\mathbb{R}^i \times \mathbb{R}^{n-i-1} \times \mathbb{R}$. These lie in the codimension two stratum $\mathcal{F}^2$. Such singularities are a union of components of this stratum. Nearby Morse functions to the swallow tail given above have three critical points: two of index $i+1$ and one of index i. Furthermore, the intersection of a loop encicling $\mathcal{F}^2$ near a swallow tail singularity meets with $\mathcal{F}^1$ in three points: two births and a crossing point. Moving a path over a stratum of this type allows us to replace the three critical points with a single critical point of index $i+1$. See Figure 3

There is also a type of path called a Smale path that is used in conjunction with the swallow tail move. This path in $\mathcal{F}^0 \cup \mathcal{F}^1$ is the path of a birth of critical points p and q of index $i+1, i+2$ followed by a crossing of p below all other critical points of index $t+1$ and then a death where p annihilates a critical point of index i. This has the effect (which we used in

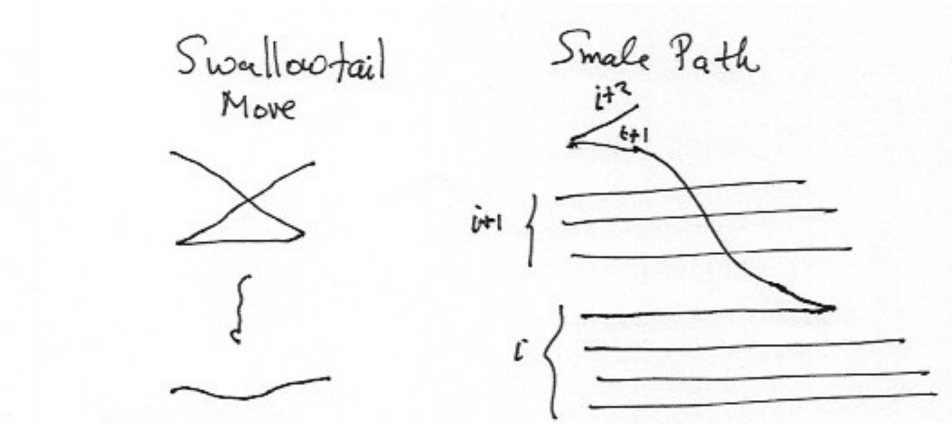

Figure 3: Swallow Tail and Smale path.

the proof of the h-cobordism theorem) of trading a critical point of index i for one of index $i+2$. Used in conjunction, the swallow tail move and the Smale paths are used to remove a critical point of the lowest index at the expense of adding a critical point of index two higher. Repeating this move allows us to remove all the critical points of the lowest index k as long as $k \leq n-4$.

Repeating these moves for critical points of the next higher index and also doing the moves to remove the lowest index critical points of $-f$, we continue to remove all critical points of extremal index until we arrive at a situation where are the critical points are of two consecutive indices k and $k+1$ where k is pre-determined, subject only to the condition that $2 \leq k \leq n-3$.

3.4 Removing the critical points of index k and $k+1$

Now we come to the algebraic K-theory part of the argument which we need to cancel the remaining critical points of index k and $k+1$, for some k satisfying $2 \leq k \leq n-3$.

Let $M \subset V \times I$ be a level set $f^{-1}(t)$ between the critical points of index k and those of index $k+1$. Let $W_+ = f^{-1}([t,1])$ and $W_- = f^{-1}([0,t])$ and set

$$L^+ \subset H_k(M) = \mathrm{Im}(H_{k+1}(W_+, M) \to H_k(M))$$

and

$$L^- \subset H_{n-k-1}(M) = \mathrm{Im}(H_{n-k}(W_-, M) \to H_{n-k-1}(M)).$$

Poincaré duality induces a perfect pairing

$$PD\colon L^+ \otimes L^- \to \mathbb{Z}.$$

Suppose that the boundaries of the descending disks for the critical points of index $k+1$ and the boundaries for the ascending disks for the critical points of index k give bases β^+ and β^- for L^+ and L^- that are dual under PD. Then, under the assumptions that the dimension of $V \times I$ is at least 6, the Whitney trick and the Morse lemma allow us to kill the critical points in dual pairs.

Smale's proof of the h-cobordism theorem then comes down to the fact that inside $G_q = GL(q, \mathbb{Z})$ the group T_q of lower triangular matrices and the symmetric group S_q generate all of G_q. For by changing the orientations of the descending disks, changing the gradient-like vector field to do handle slides we change change the basis α^+ of L^+ by any element T_q, and by reordering the critical points of index $k{+}1$ we can change α^+ by the action by any element of S_q. That is to say, given descending disks with corresponding basis α^+ of L^+, we can do handle slides, orientation reversals and reordering of the critical points to produce critical points and change the descending disks so that the homology classes of their boundary spheres form a dual basis β^+ to α^-.

Of course to study paths of Morse functions with non-degenerate critical points of index k and $k+1$ (and births and deaths), we need to study all possible ways to do moves on both bases α^+ and α^- that produce a combinatorial path from a given pair of bases α^+ for L^+ and α^- for L^- to dual bases, where we use elements of the symmetric group and lower triangular matrices to make the moves on both bases. It turns out that any algebraic move comes from a path of Morse functions with crossing critical points (and reordering the basis element).

To study pairs of moves of this nature we make a CW complex $\mathcal{S}(L^+)$ out of T_q and the symmetric group S_q. The zero cells of this complex are the maximal flags G_q/T_q.

$$F_1 \subset F_2 \subset \cdots \subset F_q$$

for L^+. The one-cells connect flags that differ by a single transposition in two adjacent basis vectors. There are two cells, rectangles, generated by pairs of transpositions that commute, and hexagons for pairs of transpositions of the form $s_a = (a, a+1)$ and $s_{a+1} = (a+1, a+2)$ coming from the relation in the symmetric group

$$s_a s_{a+1} s_a = s_{a+1} s_a a_{a+1}.$$

We consider $\mathcal{S}(L^+) \times \mathcal{S}(L^-)$ and let Δ be the the full subcomplex whose vertices are dual bases. Since every path in $\mathcal{S}(L^+) \times \mathcal{S}(L^-)$ with endpoints in Δ is the image of a path in $\mathcal{F}^1(V)$ and any one-parameter family of such paths (rel end points) is the image of a one-parameter family of paths in $\mathcal{F}^1(V)$, if $\pi_1(\mathcal{S}(L^+) \times \mathcal{S}(L^-), \Delta)$ were trivial then we could deform one parameter of functions relative the end points to a one-;parameter family where at every stage the bases from the descending and ascending pheres are dual and there are no births nor deaths. Then we could simply cancel the corresponding dual pairs in a continuous fashion throughout the entire family leaving a family without critical points.

But $\pi_1(\mathcal{S}(L^+) \times \mathcal{S}(L^-), \Delta)$ is not trivial. We can however find a simple set of generators for the relative homotopy semi-group. The last step is to see that for each generator in the list and any lift to a path of in $\mathcal{F}^1(V)$, using the the moves as before on one-parameter families, this family can be deformed relative its end points to a family whose graphic is constant. Thus we can using the basic moves on paths in $\mathcal{F}^1(V)$ we can deform the path relative to its endpoints to a path without any critical points.

This completes a outline of the proof that we can deform a path with only critical points in degrees k and $k+1$, relative to its endpoints, to one without critical points, hence proving that $\mathcal{E}(V)$ is path connected and establish Cerf's theorem