Recollections of Morse Theory and the h-Cobordism Theorem

Vincent Woltmann

We aim to give a quick recap of Morse theory and sketch a proof of the h-cobordism theorem.

Roughly, Morse theory studies smooth manifolds by looking at the critical points of functions from these manifolds to $\mathbb{R}$. The classical motivational example is the standing torus (i.e. with the "hole" towards the viewer) and its height function: This function has 4 critical points, one of index 0 at the bottom, two of index 1 where the hole "starts" and "ends" and one of index 2 at the top. These counts match the respective ranks of the homology of the torus and not by coincidence: It turns out that one can use these critical points to define a "Morse Homology" that matches the CW-homology.

More elementary one gets the Morse-Inequalities, which say that a smooth function (fulfilling certain technical conditions) has at least b_i critical points of index i where b_i is the i-th Betty number. As always when dealing with inequalities this makes the question arise whether they are tight i.e. whether there always exists a function $f : M \to \mathbb{R}$ with exactly b_i critical points of index i for every i. This is a quite hard question in general, which is why we are only going to deal with a quite special case (which is still hard enough) that already has far reaching implications, but we have to recall some definitions beforehand.

Let us start with the basics, just to make sure:

Definition 1. *A point $p \in M$ is called a critical point of a function $f : M \to \mathbb{R}$ if $(\mathrm{d}f)_p = 0$*

For technical reasons we are only interested in non-degenerate critical points, a definition which requires the Hessian. In general we need a connection to define the Hessian as $\mathcal{L}_X(\mathcal{L}_Y f) + \mathrm{d}f(\nabla_X Y)$, but at critical points this is independent of the choice connection, allowing us to define:

Definition 2. *Let $p \in M$ be a critical point of $f : M \to \mathbb{R}$. The Hessian of f at p is defined as*

$$\mathrm{Hess}(f)(X, Y) := \mathcal{L}_X(\mathcal{L}_Y f)$$

Now we can define non-degeneracy:

Definition 3. *A critical point $p \in M$ of $f : M \to \mathbb{R}$ is called non-degenerate if its Hessian has full rank.*

Using the Hessian we can now also define the index of a critical point:

Definition 4. *The (Morse-)Index of a critical point is defined as the dimension of the largest subspace on which the Hessian is negative definite.*

All of these concepts are so important because of the following Lemma:

Lemma 5 (Morse Lemma). *Let $p \in M$ be a critical point of $f : M \to \mathbb{R}$. Then there exist coordinates $x_1, ..., x_n$ such that*

$$f(x) = -x_1^2 - ... - x_i^2 + x_{i+1}^2 + ... + x_n^2$$

where i is the index of p.

This Lemma is truly the heart of classical morse theory, saying that every function looks "the same" around critical points of the same index. Since this breaks for degenerate critical points, let us also define what is means for a function to be "nice"

Definition 6. *We call a smooth function $f : M \to \mathbb{R}$ Morse if it has no degenerate critical points.*

As it turns out, being Morse is a generic condition, which is why we are going to assume that every function in this talk is Morse.

The statement we are going to proof today is formulated in the language of cobordism, so let us also recall:

Definition 7. *A cobordism $(W, \partial_\pm)$ is a smooth manifold W together with a splitting of its boundary into a positive part $\partial_+ W$ and negative part $\partial\_ W$.*

Now we can finally state the important theorem of today:

Theorem 8 (h-cobordism theorem). *Let W be a cobordism of dimension $\dim W \geq 6$ with $W, \partial_\pm W$ simply connected and $H_*(W, \partial\_ W; \mathbb{Z}) = 0$. Then W admits a Morse function without critical points that is constant on $\partial_\pm W$*

> **Remark**
>
> This theorem was first proven by Smale in 1962 and earned him a Fields Medal because it upheaved the understanding of manifolds at the time. Since up to dimension 4, increasing the dimension increases the difficulty of getting results it was assumed that higher dimensional manifolds must be even harder to understand. But it turns out that starting with dimension 5 it usually gets easier, e.g. the generalized Poincaré Conjecture almost immediately follows from this for dimension larger or equal to 5.

Proving this requires some helping lemmata, which are useful tools in general Morse theory, but to formulate them we need some additional definitions. A lot of Morse theory depends on the flow of the gradient of Morse functions. Defining these gradients requires a choice of Riemannian metric, but the morse theoretic consequences and definitions are independent of the choice of metric. Alternatively one can also choose a vector field that behaves like a gradient:

Definition 9. *Let $(W, \partial W_\pm)$ be a cobordism and $f : W \to \mathbb{R}$ a Morse function. A vector field X on W is called gradient-like for f if it satisfies*

$$\mathcal{L}_X f \geq \delta(|X|^2 + |\mathrm{d}f|^2)$$

for a $\delta > 0$, where $|X|$ is the norm with respect to some Riemannian metric and $|\mathrm{d}f|$ is the respective dual norm.

`(..) => ..` [Note that at critical points of f/zeros of X the left hand side is zero and thus $X/\mathrm{d}f$ has to be zero there and therefore critical points of f correspond to the zeros of X.] The most important consequence of these vector fields is the definition of stable/unstable manifolds:

Definition 10 (Stable/Unstable Manifold). *Let (W, f, X) be a Smale cobordism. The stable/unstable manifold W_p^-/W_p^+ of a critical point $p \in W$ of f is the set of all points converging to p under the forward/backward flow of X.*

> **Remark**
>
> These stable/unstable manifolds are one of the reasons why Morse theory is so powerful. The fact that f has only one critical point on them ensures that they are completely described by the index of the critical point, therefore any Morse function gives us a handle decomposition. (which also proves that every smooth manifold is a CW complex)

We can collect all of these fact together in one definition:

Definition 11 (Smale cobordism). *A cobordism $(W, \partial W_\pm)$ together with a Morse function $f : W \to \mathbb{R}$ that has no critical points on the boundary and a pseudo-gradient vector field X is called a Smale cobordism (W, f, X).*

We are sometimes interested in Smale cobordism that are "as simple as possible":

Definition 12 (Elementary Smale cobordism). *A Smale cobordism (W, f, X) is called elementary if there are no X-trajectories between different critical points of f.*

As a consequence, the stable manifolds of elementary Smale cobordisms are discs that intersect ∂_W along spheres, which we are going to call stable discs/spheres (respective unstable discs/spheres).

Elementary Smale cobordisms are interesting, because every Smale cobordism can be split into elementary ones (this essentially works because non-degenerate critical points behave nicely, in particular they are isolated). Also useful for is the fact that the critical values of a elementary Smale cobordism can be changed quite freely:

Lemma 13 (The elementary moving lemma). *Let (W, f, X) be an elementary Smale cobordism with $f_{\partial_\pm W}) = a_\pm$ and critical points $p_1, ..., p_n$ of values $f(p_i) = c_i$. For $i = 1, ..., n$ let $c_i : [0, 1] \to (a_-, a_+)$ be smooth functions with $c_i(0) = c_i$. Then there exists a smooth family $f_t, t \in [0, 1]$, of Morse functions for which X is a pseudo gradient vector field with $f_0 = f$ and $f_t = f$ on $\mathcal{O}p(\partial W)$ such that $f_t(p_i) = c_i(t)$*

Proof. There are no X-trajectories between critical points by assumption, turning the stable manifolds into disjoint discs, whose boundaries are stable spheres denoted by $S_i \subset \partial\_W$ for a critical point p_i. We are then going to extend these into disjoint tubular neighborhoods in ∂_W and denote the closure of their image under the forward flow of X as V_i and then again with smaller tubular neighborhoods contained in the previous ones, whose forward flow is denoted as $U_i \subset V_i$. Reflecting on this definition, one notes that $V_i \setminus \text{Int } U_i$ is diffeomorphic to $S^{k_1-1} \times S^{m-k_i-1} \times [0, 1] \times [a_-, a_+]$ where m is the dimension of W and k_i the index of p_i. This holds true because the stable manifold of p_i is a k_i-dimensional disc with a $(k_i - 1)$-sphere as boundary, with the tubular neighborhoods adding a factor of $S^{m-k_i-1} \times [0, 1]$. Also since the trajectories don't meet any critical points the factor of $[a_-, a_+]$ gets added by the forward flow. Taking this manifold as a choice of coordinates for W turns f into $f(u, v, x, y) = y$ and $X = \partial_y$.

Now fix a inversion function $\rho : [0, 1] \to [0, 1]$ which equals 1 near 0 and 0 near 1. Pick smooth families of diffeomorphisms $\sigma_{i,t}$ with $\sigma_{i,0} = \text{Id}$, $\sigma_{i,t}(c_i) = c_i(t)$ and $\sigma_{i,t}$ being the identity near $a_\pm$ for every i, t. Using the coordinates established above we can define

$$f_t := \begin{cases} (1 - \rho(x))y + \rho(x)\sigma_{i,t}(y) & \text{on } \; V_i \setminus \text{Int } U_i \\ \sigma_{i,t} \circ f & \text{on } \; U_i \\ f & \text{on } \; W \setminus \bigcup_i V_i \end{cases}$$

Critical values p_i lie in the interior of U_i and therefore $f_{t(p_i)} = \sigma_{i,t}(f(p_i)) = c_{i(t)}$. We also have on $V_i \setminus \text{Int } U_i$:

$$\mathcal{L}_X(f_t) = \frac{\partial f_t}{\partial y} = (1 - \rho(x)) + \rho(x)\sigma_i'(y) > 0$$

hence X is a pseudo-gradient vector field for f_t, making f_t the desired function. □

Moving critical values is an invaluable tool, but we still need to get rid of critical values in some way. The next lemma partially takes care of that:

Lemma 14 (Cancellation of critical points). *Let (W, X, f) be a Smale cobordism with two critical points of index $k-1$ and k connected by a unique trajectory of X along which the stable and unstable manifolds intersect transversely. Then there exists a family of Smale cobordisms (W, X_t, f_t) fixed on $\mathcal{O}p\ \partial W$ with $(X_0, f_0) = (X, f)$ such that (X_1, f_1) has no critical points.*

Sadly this Lemma is too difficult to prove in our limited time, which is why we have to point to [1, Theorem 5.4] for a proof.

Since it also doesn't allow arbitrary removal of critical points, but only cancellation, we ironically sometimes have to create critical points to cancel them out later, enabled by the following lemma:

Lemma 15 (Creation of critical points). *Let (W, X, f) be a Smale cobordism without critical points. Then for $1 \leq k \leq \dim W, p \in \operatorname{Int} W$ there exists a family of Smale cobordisms (W, X_t, f_t) that only changes near p with $(X_0, f_0) = (X, f)$ such that (X_1, f_1) has two critical points of index $k-1$ and k that are connected by a unique trajectory of X_1 such that the stable and unstable manifolds intersect transversely.*

For the proof we again refer to [1, Lemma 8.2]

The last lemma we need is a bit more technical:

Lemma 16 (Moving attaching spheres). *Let (W, X, f) be a Smale cobordism and $p \in W$ a critical point with stable manifold W_p^- that intersects ∂_W along a sphere $S \subset \partial\_W$. Given an isotopy $S_t \subset \partial\_W$ of S there exists a homotopy X_t fixed on ∂W of gradient-like vector fields from X such that $W_p^-(X_t)$ intersects ∂_W along S_t.*

We are going to need this to simplify certain intersections. Now we can finally start with our proof itself:

Proof. We start by using Lemma 13 to move every critical point of index λ to level λ (A morse function like this is also called self-indexing), the only obstacle to this is that this Lemma can only be applied to elementary Smale cobordisms, but using transversality we can move the stable/unstable manifolds away from each other, allowing us to apply our lemma to the general case.

Now for every $k \in \mathbb{N}$ consider the regular level set $\Sigma = f^{-1}\left(k - \frac{1}{2}\right)$ i.e. the level set between the critical points of index $k-1$ and k. Let $p_1, ..., p_s$ be the critical points of index k and $q_1, ..., q_t$ those of index $k-1$. Denote the stable spheres of p_i by $S_i^- \subset \Sigma$ and the unstable spheres of q_j as S_j^+ and let A be the matrix of homological intersections numbers $A_{ij} = S_i^- \cdot S_j^+$.

We can modify A using so called handle slides: Take two critical points p_i, p_j of level k, move p_i to a slightly higher level and using Lemma 16 modify X to "slide" the stable manifold of p_i over the stable manifold of p_j, replacing the homology class $[S_i^-]$ by $[S_i^-] + [S_j^-]$, effectively adding the j-th row of A to the i-th row. Since $H_*(W, \partial\_W; \mathbb{Z}) = 0$ we can use these elementary row operations to ensure that $S_i^- \cdot S_j^+ = 1$ for $i = 1, ..., r \leq \min(s, t)$ with all other intersection numbers being zero.

Now before we continue we are going to have to get rid of critical points of index $0, 1, m-1, m$ with $m = \dim W$. We can use Lemma 14 to get rid of the critical points of index 0 and m and then use the "Smale trick": Use Lemma 15 to create a pair of critical points of index 2 and 3 and then use the critical point of index 2 to cancel out a critical point of index 1. Similar with critical points of index $m-1$.

No we can remove all remaining "homologically unnecessary" intersections (i.e. intersections that are cancelled out by others) that still remain. To accomplish this, consider two transverse intersection points $z_\pm$ with local intersection indices ± 1. We can connect them by paths in S_i^- and S_j^+ to obtain an embedded loop γ in Σ. Per assumption W is simply connected, hence γ bounds a disk. We also have $\dim W \geq 6$, allowing us to apply a theorem by Whitney (see [2] for details) to choose this disk Δ such that it meets $S_i^- \cup S_j^+$ only transversely along the boundary. Pushing S_i^- over Δ using Lemma 16 then eliminates the intersection points $z_\pm$.

We end up with S_i^- and S_i^+ intersecting in a unique point for every $i = 1, ..., r$. In particular p_i and q_i are connected by a unique X-trajectory, allowing us to eliminate them using Lemma 14. Thus we end up with a Morse function with no critical points. □

Bibliography

[1] J. Milnor, *Lectures on the h-cobordism theorem.* Princeton, NJ: Princeton University Press, 2025.

[2] H. Whitney, “The self-intersections of a smooth n-manifold in $2n$-space”, *Annals of Mathematics*, vol. 45, pp. 220–246, 1944, doi: 10.2307/1969265.