Lecture I: Basics of Morse Theory

September 14, 2020

1 Basic Definitions and Results

Let M be a compact smooth manifold. We shall denote by $C^\infty(M;\mathbb{R})$ the space of C^∞-functions on M, with the C^∞-topology. This is Frechet vector space. More generally for N a smooth manifold we denote by $C^\infty(M,N)$ the space of C^∞ mappings with the C^∞ topology. This is Frechet manifold.

For any $f \in C^\infty(M;\mathbb{R})$ we denote by $Df\colon M \to T^*M$ the map.

$$x \mapsto Df_x\colon TM_x \to \mathbb{R},$$

where Df_x is viewed as an element of the dual space T^*M_x. This is a smooth section of T^*M. The correspondence $f \mapsto Df$ defines a continuous mapping $C^\infty(M,\mathbb{R}) \to C^\infty(M,T^*M)$.

Definition 1.1. Let $f \in C^\infty(M;\mathbb{R})$.

- Then $x \in M$ is a *critical point* of f if $Df_x\colon TM_x \to \mathbb{R}$ is zero; i.e., x is a critical point of f if and only if $Df(x)$ lies in the zero section of T^*M.

- A critical point $x \in M$ of f is *non-degenerate* if $Df\colon M \to T^*M$ is transverse to the zero section of T^*M at x.

- $t \in \mathbb{R}$ is a *critical value* of f if there is a critical point x of f with $f(x) = t$.

- Any real number that is not a critical value of f is called *a regular value* of f.

Suppose that W is a compact manifold with boundary and fix a decomposition $\partial W = \partial_- W \coprod \partial_+ W$, where each of $\partial_\pm W$ is a union of connected components. Then we define $C_0^\infty(W,\partial_- W,\partial_+ W)$ to be the set of C^∞ functions f on W satisfying:

- f is locally constant on ∂W,
- f has no critical points on the boundary
- Df is positive, resp. negative, on outward pointing normals at points of $\partial_+ W$, resp. $\partial_- W$.

If $\partial_\pm W$ are clear from context when we will use the notation $C_0^\infty(W)$ instead of $C_0^\infty(W, \partial_- W, \partial_+ W)$ to refer to this space of functions.

2 The Hessian and Local Models

2.1 The Hessian

At any point $(x, 0)$ of the zero section of T^*M there is a canonical decomposition

$$T(T^*M)_{(x,0)} = TM_x \oplus T^*M_x,$$

where the first factor is the tangent space to the zero section and the second factor is the space of vertical tangents. [The second subspace of $T(T^*M)$ is canonically defined at every point of T^*M, but the first is not. One way to produce the first at a general point of T^*M is to fix a connection on T^*M.]

Fix a smooth function $f\colon M \to \mathbb{R}$ and suppose that x is a critical point of f. Then we have the composition

$$TM_x \stackrel{D(Df)}{\longrightarrow} T(T^*M)_{(x,0)} = TM_x \oplus T^*M_x \stackrel{\pi_2}{\longrightarrow} T^*M_x.$$

We denote this composition $H_x(f)\colon TM_x \to T^*M_x$ and call it the *Hessian* of f at x.

Lemma 2.1. *The map $H_x(f)$ is self-adjoint. Any local coordinate system $(x^1, \ldots, x^n)$ near x determines a basis for TM consisting of $\{\partial/\partial_{x^i}\}_{i=1}^n$ and the dual basis $\{dx^i\}_{i=1}^n$ for T^*M defined throughout the neighborhood. With respect to these bases at x, we have*

$$H_x(f) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\Big|_x \right).$$

Proof. In the given local coordinates, Df is the map

$$Df(y) = (y, \sum_i \frac{\partial f}{\partial x^i}(y) dx^i),$$

and by definition

$$H_x(f) = D(\pi_2(Df(y))) = \sum_j \frac{\partial}{\partial x^j}(\pi_2 Df)(x)dx^j.$$

We have

$$\frac{\partial}{\partial x^j}(\pi_2 Df)(x) = \sum_{i,j} \frac{\partial^2 f}{\partial x^j \partial x^i}(x)dx^i.$$

Since the matrix of second partials is symmetric, this establishes that $H_x(f)$ is self-adjoint. □

Fix a finite dimensional real vector space V. For any self-adjoint mapping $H\colon V \to V^*$ there is a basis B for V such that the matrix for H is diagonal in the basis B for V and the dual basis B^* for V^*. The number of positive, negative, and zero eigenvalues are independent of the choice of such a basis. The numbers of zero eigenvalues is called the *nullity* of H, or its *co-rank*. The *rank* of H is $\dim(V)$ minus the co-rank of H. The number of (strictly) negative eigenvalues of H is called the *index* of H. It is the maximal dimension of a linear subspace of V on which H is negative definite.

Definition 2.2. Let x be a critical point of $f\colon M \to \mathbb{R}$. The *index* of the critical point x is the index of $H_x(f)$.

2.2 Local Models

Here is a local model for a non-degenerate critical point.

Lemma 2.3. *Let $U \subset \mathbb{R}^n$ be an open neighborhood of the origin Let $f\colon U \to \mathbb{R}$ have a non-degenerate critical point at the origin. Then there is an open ball $V \subset \mathbb{R}^n$ and a diffeomorphism onto an open neighborhood of the origin $\varphi\colon V \to \varphi(V) \subset U$ with the property that $f \circ \varphi = \sum_i \epsilon_i |x^i|^2$ with each ϵ_i equal to ± 1.*

Proof. We begin with a claim.

Claim 2.4. *If a smooth function g defined in a neighborhood of 0 in $\mathbb{R}^n$ vanishes at 0, then there is a smooth coordinate system $(x^1, \ldots, x^n)$ defined near 0 so that $g(x^1, \ldots, x^n) = \sum_i x^i h_i(x^1, \ldots, x^n)$ for smooth functions h_i with $h_i(0) = \partial g/\partial x^i(0)$.*

Proof. Since $g(0) = 0$, the fundamental theorem of calculus implies

$$g(x) = \int_0^1 \sum_i x^i (\partial g/\partial x^i)(tx)dt = \sum_i x^i \int_0^1 (\partial g/\partial x^i)(tx)dt.$$

Set $h_i = \int_0^1 (\partial g/\partial x^i)(tx)dt$. □

Now suppose that f has a non-degenerate critical point at the origin. Now we choose coordinates $(x^1, \ldots, x^n)$ near the origin so that the matrix of second partials of f at 0 is a diagonal matrix with all diagonal entries ± 1. Then the claim gives $(\partial f/\partial x^i)(x) = \sum_j x^j k_{ij}$. Applying the claim again with these expressions for the $\partial f/\partial x^i$ we see there are smooth functions $h_{i,j}$ such that

$$f(x) = \sum_{i,j} x^i x^j h_{i,j}.$$

By averaging we can assume that the $h_{i,j} = h_{j,i}$ for all i, j.

Denote the i^{th} diagonal entry by ϵ_i. Then $h_{i,i}(0) = \epsilon_i$. Restrict to a ball around the origin where the $h_{i,i}$ are all non-zero.

We set $y^1 = \sqrt{\epsilon_1 h_{1,1}} x^1 + \sum_{i=2}^n \frac{x^i h_{1,i}}{\sqrt{\epsilon_1 h_{11}}}$. Then $(y^1, x^2, \ldots, x^n)$ is a coordinate system near the origin and in this system

$$f = \epsilon_i (y^1)^2 + \sum_{i,j \geq 2} x^i x^j h_{i,j},$$

where the $h_{i,j}$ are new functions of the coordinates $(y^1, x^2, \ldots, x^n)$.. Continuing this way by induction we find a new set of coordinates $(y^1, \ldots, y^n)$ in which $f = \sum_i \epsilon_i (y^i)^2$. □

Definition 2.5. For a non-degenerate critical point p of a smooth function f, local coordinates in a neighborhood of p in which $f = f(p) + \sum_i \epsilon_i (x^i)^2$ with the $\epsilon_i = \pm 1$ are called *standard coordinates near* p.

Corollary 2.6. *If p is a non-degenerate critical point of f, then, in the ring of germs of functions at the origin, the ideal generated by the partial derivatives of f is the ideal of all germs vanishing at p.*

3 Morse Functions

Definition 3.1. Let M be a compact manifold, possibly with boundary with a decomposition $\partial M = \partial_- M \coprod \partial_+ M$ as before. A *Morse function* on M is an element $f \in C_0^\infty(M; \mathbb{R})$ all of whose critical points (automatically interior points) are non-degenerate.

One way that Morse functions on compact manifolds with boundary arise is shown in the following example.

Example. Let $f\colon M \to \mathbb{R}$ be a Morse function on a closed manifold. For regular values $b < c$ let $W = f^{-1}([b,c])$ and set $\partial_- W = f^{-1}(b)$ and $\partial_+ W = f^{-1}(c)$. The restriction of f to W is a Morse function.

Lemma 3.2. *Let M be a compact manifold, possibly with boundary.. The set of Morse functions is an open subset of set of $C_0^\infty(M;\mathbb{R})$.*

Proof. As we have seen, the condition that f be a Morse function is that $Df\colon M \to T^*M$ be transverse to the zero section (including the fact that Df is nowhere zero on ∂M). Clearly, $f \mapsto Df$ is a continuous map $C_0^\infty(M;\mathbb{R}) \to C^\infty(M, T^*M)$. The subset of maps $M \to T^*M$ transverse to the zero section form an open subset of all C^∞-mappings. The result follows. □

Notice that it follows that the number of critical points is a locally constant function. Also, as we vary f in a small neighborhood the critical points vary continuously and Hessians at the critical point also vary continuously. Since the index is a discrete invariant of the Hessian, it is a locally constant function on the set of non-degenerate quadratic forms. Thus, for all Morse functions sufficiently close to f we have a one-to-one correspondence of critical points preserving the indices.

Theorem 3.3. *The set of Morse functions is a dense subset of $C_0^\infty(M;\mathbb{R})$.*

Proof. The space $J^1(M,\mathbb{R})$ is naturally identified with the cotangent bundle of M and $J^1(f) = Df$. Apply Thom's Transversality Theorem to the map $f \mapsto J^1(f)$ and W equal to the zero section $Z \subset T^*M$. Since a function $f\colon M \to \mathbb{R}$ is a Morse function if and only if $J^1(f) = Df$ is transverse to Z, this gives the result. □

There is an analogue of this argument using a finite dimension family inside $C_0^\infty(M;\mathbb{R})$, see Theorem 4.9 on page 54 of Golubitsky-Guillemin.

A variant of this argument applies in the following relative situation.

Proposition 3.4. *Let W be a smooth manifold. Suppose that $f\colon W \to \mathbb{R}$ is an element of $C_0^\infty(W;\mathbb{R})$ whose critical C set of f is compact and contained in the interior of M . For any neighborhood U of C, there is an arbitrarily small C_0^∞-perturbation $\hat{f}$ of f, with the perturbation supported in U and which has only non-degenerate critical points.*

Proposition 3.5. *Let M be a smooth manifold and let $f\colon M \to \mathbb{R}$ be a C^∞-function. Suppose that the critical set $C \subset M$ of f with is compact.*

If the index of every critical point is at least λ, then there is an arbitrarily small perturbation $\hat{f}$ of f with $\hat{f} = f$ outside any fixed neighborhood U of C such that every critical point of $\hat{f}$ is non-degenerate and of index $\geq \lambda$.

Proof. Take a sequence g_n of Morse functions agreeing with f outside of U and converging to f in the C^∞-topology. We claim that eventually all the critical points of g_n have index $\geq \lambda$. Suppose not. Then pass to a subsequence and find points x_n in the critical set for g_n at which the index is less than λ. Passing to a further subsequence we arrange that the x_n converge to a critical point y of f, and indeed all x_n lie in a local coordinate system for y. Of course, the matrix of second partials of g_n at x_n in this local coordinate system converge to the matrix of second partials of f at y. Since the subset of symmetric matrices that have at least λ negative eigenvalues is an open condition, this implies that the matrix of second partials of g_n at x_n have at least λ negative eigenvalues. This contradicts the fact that x_n is a critical point for g_n whose Hessian has fewer than λ negative eigenvalues. □

We can also arrange that distinct critical points of a Morse function lie in different level sets of the function

Corollary 3.6. *Let M be a compact manifold. The set of Morse functions with the property that the pre-image of each critical value has exactly one critical point is an open dense set in $C_0^\infty(M;\mathbb{R})$.*

Proof. Clearly the condition that the restriction of a Morse function to its critical points gives an injective function from the set of critical points to $\mathbb{R}$ is an open condition on the set of Morse functions. Since the set of Morse functions is itself an open subset of $C_0^\infty(M;\mathbb{R})$, it follows that those with this extra condition form an open subset of $C_0^\infty(M;\mathbb{R})$.

Let us show that these Morse functions form a dense set. To do this we need only show that any Morse function is a limit of those with distinct critical values. Let f be a Morse function and $p_1, \ldots, p_N$ its critical points. Choose disjoint neighborhoods $U_j \subset \operatorname{int} M$ centered at p_j with standard local coordinates on the U_j. For each j let ψ_j be supported in U_j and equal to 1 in some neighborhood of p_j. For $t = (t_1, \ldots, t_N)$ is a sufficiently small neighborhood of the origin in $\mathbb{R}^N$, the function $f(t) = f + \sum_j t_j \psi_j$ is a Morse function. Again restricting to a smaller neighborhood of the origin, $f(t)$ has the same number of critical points as f. But each critical point of f is a critical point of $f(t)$. Thus, for t in this smaller neighborhood of the origin, the critical points of $f(t)$ are the $\{p_1, \ldots, p_N\}$. The critical value under $f(t)$ of p_j is $f(p_j) + t_j$. Thus, for generic t in this neighborhood, the values of the p_j under $f(t)$ are all distinct. □

4 A Finite Dimensional Analogue

Here is a purely finite dimensional analogue of the fact that the Morse functions are open dense.

Proposition 4.1. *Let $M \subset \mathbb{R}^N$ be a compact, smooth submanifold (without boundary). For each u in the unit sphere S^{N-1} we have a smooth function $f_u\colon M \to \mathbb{R}$ given $f_u(x) = \langle x, u\rangle$ where the inner is the Euclidean inner product. For an open dense set of $u \in S^{N-1}$, the function f_u is a Morse function.*

Proof. We define $P(M)$ to be the smooth locally trivial fiber bundle over M whose fiber at $x \in M$ is the Grassmannian of codimension-1 linear subspaces of the normal space at x to M. There is a smooth map

$$\varphi\colon P(M) \to Gr(N-1, N)$$

that assigns to $L \subset \nu_x(M)$ the codimension-1 hyperplane $L \oplus T_xM$. For any unit vector u let $V_u = u^\perp$. Then we have the corresponding point $\{V_u\} \in Gr(N-1, N)$. The critical points of f_u are the image in M under the projection $P(M) \to M$ of $\varphi^{-1}(\{V_u\})$. The map f_u is a Morse function if and only if $\{V_u\}$ is a regular value for $\varphi\colon P(M) \to Gr(N-1, N)$. The standard finite dimensional version of Sard's Theorem implies that the regular values of φ are dense. In this case since the domain is compact, they are an open dense. □

5 Bott-Morse Functions

There is a generalization of the notion of a Morse function which is often a more natural context in which to work, especially in geometric situations where there is a compact symmetry group.

Definition 5.1. A function $f\colon M \to \mathbb{R}$ is *Bott-Morse* function if the set of critical points is a submanifold (possibly with components of different dimensions) and if at every critical point $x \in M$ the null space of the Hessian $H_x(f)$ is the tangent space at x to the critical submanifold.

Remark 5.2. The null space of a symmetric bilinear form $H\colon V \otimes V \to \mathbb{R}$ is the set of $v \in V$ for which $H(v, \cdot)\colon V \to \mathbb{R}$ is trivial. It is easily seen to be a subspace indeed the maximal subspace on which H is identically zero. H induces a symmetric biinear form on the quotient of V by its null space and this form is non-degenerate.

Remark 5.3. (i) In passing a critical level, the change in the topology is achieved by attaching a disk bundle of dimension equal to the index of the critical points along each component of the critical manifold.

(ii) One way to perturb a Bott-Morse function f (at least in the case when the critical manifold is compact) is to choose a Morse function g on the critical manifold and then extend it to all of M, say by pulling back to the normal bundle and damping out away from the critical submanifold. The family $f + \epsilon = f + \epsilon g$ will be a Morse function for all $\epsilon > 0$ sufficiently small. Then the change in topology is given by attaching descending disks out of these critical points. This replaces the attachment of the disk bundle over the critical submanifold in the Bott-Morse picture but gives the same total change in the topology.

Remark 5.4. 1. If $G \times M \to M$ is an action of a compact Lie group on a compact manifold and $f\colon M \to \mathbb{R}$ is a G-invariant function, then for a critical point x of f, the orbit $G \cdot x$ consists of critical points. Thus, unless x is a fixed point for the action of the component of the identity of G, x cannot be an isolated fixed point, and hence cannot be a non-degenerate critical point.
2. Let $G \times M \to M$ be a free action of a compact group and denote by $\pi\colon M \to M/G$ the natural quotient map. if $f\colon M/G \to \mathbb{R}$ is a Morse function, then $f \circ \pi\colon M \to \mathbb{R}$ is Bott-Morse with each component of the critical set being a copy of G.