2 Morse Theory

2.1 Sard's lemma

In this subsection we recall some basic facts about Baire category theory and the Sard's lemma about the critical values of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ without proof.

Definition 2.1. Let X be a topological space. A *first category subset* of X is a countable union of closed subsets with empty interior. A *second category subset* is a subset which is not a first category subset.

Lemma 2.2 (Baire)**.** *In a complete metric space the complement of a first category subset is dense.*

The following terminology is commonly used in the literature about topology or symplectic geometry. If X is a complete metric space and P is some property, we say that "a generic element of X satisfies P" or, interchangeably, that "P is generic (in X)" if the elements of X which do not satisfy P form a first category subset. This implies not only that elements satisfying P are dense in X, but also that, given countably many generic properties $P_1, \ldots, P_n, \ldots$, a generic element of X satisfies all of them. In fact a countable union of first category subsets is a first category subset.

Let $F\colon \mathbb{R}^n \to \mathbb{R}^m$. A *critical point* of F is a point $\mathbf{x} \in \mathbb{R}^n$ such that $d_{\mathbf{x}}F$ is not surjective. A *critical value* of F is a point $\mathbf{y} \in \mathbb{R}^m$ such that $F(\mathbf{x}) = \mathbf{y}$ for some critical point $\mathbf{x}$.

Theorem 2.3 (Sard's lemma)**.** *Let $F\colon \mathbb{R}^n \to \mathbb{R}^m$ be a smooth function. Then the set of critical values of F is a first category subset of $\mathbb{R}^m$.*

From this we obtain immediately the following corollary.

Corollary 2.4. *Let M and N be smooth manifolds and $f\colon M \to N$ a smooth map between them. Then the set of critical values of f is a subset of first category of N.*

Often Sard's lemma is stated with measure zero subsets replacing first category subsets. The two statements cannot be derived from one another; in fact there are first category subsets of positive measure and zero measure subset of second category. However the two versions of Sard's lemma have similar proofs and are functionally equivalent. We have chosen the one with Baire category because it is the one which generalizes to the infinite dimensional setting which we will need later.

2.2 Morse functions

Let M be a smooth manifold and $f\colon M \to \mathbb{R}$ a smooth function. We recall that a critical point of f is a point $p \in M$ such that $d_p f = 0$. The set of critical points of f will be denoted by $Crit(f)$. It is easy to see that $Crit(f)$ is closed in M. A critical value of f is an element $x \in \mathbb{R}$ which is the image of a critical point and a *regular value* is an element $x \in \mathbb{R}$ which is not a critical value. We will recall here some definitions and results, mostly without proof.

Lemma 2.5. *Let M be a smooth manifold and $f\colon M \to \mathbb{R}$ a smooth function. If $p \in M$ is a critical point for f and X, Y are vector fields on a neighbourhood of p, then*

1. $(XYf)(p) = (YXf)(p)$, *and*
2. $XYf(p) = 0$ *if* $X_p = 0$ *or* $Y_p = 0$.

Proof. We have $(XYf)(p) - (YXf)(p) = d_p f([X,Y]) = 0$ because $d_p f = 0$. If $X_p = 0$ then $XYf(p) = 0$ by the definition of vector field. If $Y_p = 0$ then $XYf(p) = 0$ because of (1). □

Lemma 2.5 implies that $(XYf)(p)$ depends only on X_p and Y_p at a point $p \in Crit(f)$ and thus motivates the following definition.

Definition 2.6. Let M be a smooth manifold and $f\colon M \to \mathbb{R}$ a smooth function. The *Hessian* of f at the critical point p is the symmetric bilinear form

$$H_p f\colon T_pM \times T_pM \to \mathbb{R}$$

defined as follows. Given two tangent vectors $X_p, Y_p \in T_pM$, we extend them to vector fields X and Y in an arbitrary way and define $H_p f(X_p, Y_p) = (XYf)(p)$.

Definition 2.7. A function $f\colon M \to \mathbb{R}$ is a *Morse function* if $H_p f$ is non-degenerate at any point $p \in Crit(f)$. The *Morse index* of a critical point p, denoted $i(p)$, is the dimension of the negative space of $H_p f$.

For every Morse function f we will denote $Crit_k(f)$ the set of the critical points of f with Morse index k.

Lemma 2.8. *For any smooth manifold M, Morse functions are generic in $C^\infty(M)$.*

Morse functions have simple local models near their critical points.

Lemma 2.9 (Morse lemma)**.** *Let M be a smooth manifold of dimension n and $f\colon M \to \mathbb{R}$ a smooth function. If $p \in M$ is a Morse critical point of index k, then there is a neighbourhood U of p and a chart $\phi\colon U \to \mathbb{R}^n$ with $\phi(p) = \mathbf{0}$ such that*

$$(f \circ \phi^{-1})(x_1, \dots, x_n) = f(p) - x_1^2 - \dots - x_k^2 + x_{k+1}^2 + \dots + x_n^2$$

in a neighbourhood of $\mathbf{0}$ in $\mathbb{R}^n$.

The Morse lemma implies that the critical points of a Morse function are isolated. In particular a Morse function on a compact manifold has finitely many critical points.

2.3 Morse homology

If we fix a Riemannian metric g on M, we can define the gradient vector field ∇f of a function $f \in C^\infty$ by $g(\nabla f, X) = df(X)$ for all vector field X on M. The flow generated by $-\nabla f$ is called the *negative gradient flow* of f.

If f is a Morse function and φ is it negative gradient flow, for every critical point p we define the *stable submanifold* $W^s(p)$ as

$$W^s(p) = \{q \in M : \lim_{t \to +\infty} \varphi_t(q) = p\}$$

and the *unstable manifold* $W^u(p)$ as

$$W^u(p) = \{q \in M : \lim_{t \to -\infty} \varphi_t(q) = p\}.$$

The stable and unstable manifolds are open submanifolds of M and moreover $\dim W^u(p) = i(p)$ and $\dim W^s(p) = \dim M - i(p)$.

Definition 2.10. A pair (g, f), where g is a Riemannian metric and f is a Morse function on M, is called *Morse-Smale* if for any pair of critical points p, q, the manifolds $W^s(p)$ and $W^u(q)$ intersect transversely.

Lemma 2.11. *Given a Morse function f, the pair (f, g) is Morse-Smale for a generic Riemannian metric g.*

Given two critical points p, q, we define

$$\mathcal{M}(p,q) = W^u(p) \cap W^s(q).$$

If (f, g) is a Morse-Smale pair, then $\mathcal{M}(p,q)$ is a smooth manifold of dimension

$$\dim \mathcal{M}(p,q) = i(p) - i(q).$$

It is clear that $\mathcal{M}(p,q)$ depends on both f and g.

Since $\mathcal{M}(p,q)$ is a union of trajectories, the negative gradient flow φ induces an action of $\mathbb{R}$ on $\mathcal{M}(p,q)$. When $i(p) > i(q)$ we denote

$$\mathcal{M}^*(p,q) = \mathcal{M}(p,q)/\mathbb{R}$$

the quotient by the $\mathbb{R}$-action. Then $\mathcal{M}^*(p,q)$ is a manifold of dimension $i(p) - i(q) - 1$ which can be regarded as a space which parametrizes the negative gradient flow trajectories from p to q. The manifolds $\mathcal{M}(p,q)$ and $\mathcal{M}^*(p,q)$ are often called *(Morse) moduli spaces*[5].

Theorem 2.12. *Let (f, g) be a Morse-Smale pair on a closed manifold.*

- *If p, q are critical points of f with $i(p) \leq i(q)$, then $\mathcal{M}(p,q) = \emptyset$ if $p \neq q$, and $\mathcal{M}(p,q)$ consists of constant trajectories if $p = q$.*
- *If p, q are critical points of f with $i(p) - i(q) = 1$, then $\mathcal{M}^*(p,q)$ is a finite set.*
- *If $i(p) - i(q) = 2$, then $\mathcal{M}^*(p,q)$ is a 1-dimensional manifold with finitely many connected components. The noncompact connected components of $\mathcal{M}^*(p,q)$ admit natural compactifications which are homeomorphic to closed intervals. The boundary of the compactification $\overline{\mathcal{M}}^*(p,q)$ of $\mathcal{M}^*(p,q)$ is*

$$\partial \overline{\mathcal{M}}^*(p,q) = \bigcup_{i(r)=i(p)-1} \mathcal{M}^*(p,r) \times \mathcal{M}^*(r,q).$$

We will build a chain complex out of a Morse-Smale pair (f, g), the so-called *Morse complex*. In order to avoid the mild complications which are necessary to define the Morse complex over the integers, we will work over $\mathbb{Z}/2\mathbb{Z}$. If p, q are critical points of f with $i(p) - i(q) = 1$, we define $\#\mathcal{M}^*(p,q)$ as the count modulo 2 of the number of negative gradient flow trajectories from p to q. This count makes sense by Theorem 2.12.

[5] In differential geometry we call "moduli space" a topological space which parametrizes the solutions of a differential equation, possibly up to the action of some symmetry group.

Definition 2.13 (Morse complex)**.** Let (f, g) be a Morse-Smale pair on a closed manifold. We define the *Morse complex* $(C_*(f,g), \partial)$, where

$$C_i(f,h) = \bigoplus_{p \in Crit_i(f)} \mathbb{Z}/2\mathbb{Z}\, p$$

and the differential $\partial\colon C_i(f,g) \to C_{i-1}(f,g)$ is defined as

$$\partial(p) = \sum_{q \in Crit_{i-1}(f)} \#\mathcal{M}^*(p,q) q. \tag{9}$$

The identity $\partial^2 = 0$ is an algebraic way to encode the geometry of the compactification of 1-dimensional moduli spaces.

Theorem 2.14. *If (f, g) is a Morse-Bott pair, then $(C_*(f,g), \partial)$ is a chain complex.*

Proof. We need to prove that $\partial^2 = 0$. If we apply Equation (9) twice we obtain

$$\partial^2(p) = \sum_{q \in Crit_{i-2}(f)} \left(\sum_{r \in Crit_{i-1}(f)} \#\mathcal{M}^*(p,r)\#\mathcal{M}(r,q) \right) q.$$

The quantity in the big parentheses is the number of broken trajectories from p to q. From Theorem 2.12 we know that they form the boundary of the compactification of the 1-dimensional moduli space $\mathcal{M}^*(p,q)$. Since a compact 1-dimensional manifold with boundary is a disjoint union of finitely many closed intervals and circles, and closed intervals have two boundary points each, we conclude that

$$\sum_{r \in Crit_{i-1}(f)} \#\mathcal{M}^*(p,r)\#\mathcal{M}^*(r,q) = 0.$$

$\square$

The homology of the Morse complex $(C_*(f,g), \partial)$ is called *Morse homology* and is denoted by $H_*(f,g)$. It is possible to prove directly that Morse homology is independent of (f, g) by constructing chain homotopies between Morse complexes arising from different choices of Morse-Smale pairs. However we will derive the topological invariance of Morse homology from an isomorphism with singular homology that will be described in the next section.

2.4 Comparison with singular homology

In this section we will sketch the proof that Morse homology is isomorphic to singular homology. The isomorphism will be proved in two steps. First we will define cellular homology and prove that it is isomorphic to singular homology. Then we will show that cellular homology is isomorphic to Morse homology.

A different route would also be possible: we could prove first that Morse homology is independent of the the Morse-Smale pair up to canonical isomorphism, and then show that it satisfies the axioms of Eilenberg-Steenrod. Then a theorem in homological algebra would imply that it is isomorphic to singular homology.

A Morse-Smale pair (f, g) on a smooth manifold M of dimension n induces a filtration

$$M_0 \subset \ldots \subset M_n = M, \tag{10}$$

where M_0 is a union of balls around the local minima of f, the negative gradient vector field $-\nabla f$ points inside M_i along ∂M_i, and $Crit_i(f) \subset int(M_i) \setminus M_{i-1}$. This filtration is constructed inductively: M_0 is a small neighbourhood of the index zero critical points (i.e. the local minima of f) and M_i is constructed from M_{i-1} by adding small neighbourhoods of the unstable manifolds $W^u(p)$ for all critical points $p \in Crit_i(f)$. Then (M_i, M_{i-1}) is an excision pair and moreover M_i retracts on $M_{i-1} \cup \left(\bigcup_{p \in Crit_i(f)} W^u(p) \right)$.

Lemma 2.15. *$C_i(f, g)$ is canonically isomorphic to $H_i(M_i, M_{i-1})$.*

Proof. For each $p \in Crit_i(f)$ we have $W^u(p) \setminus int(M_{i-1}) \cong D^i$. Then applying excision and homotopy invariance to the pair (M_i, M_{i-1}) we obtain

$$H_i(M_i, M_{i-1}) \cong \bigoplus_{p \in Crit_i(f)} H_i(D^i, \partial D^i) \cong C_i(f, g)$$

because $H_i(D^i, \partial D^i) \cong \mathbb{Z}$. □

The correspondence between $C_i(f, g)$ and $H_i(M_i, M_{i-1})$ is obtained by associating the class $[W^u(p)] \in H_i(M_i, M_{i-1})$ to the critical point $p \in Crit_i(f)$.

Lemma 2.16. *$H_i(M_{i+1}) \cong H_i(M)$ for all $i < n$ and $H_i(M_k) = 0$ for $i > k$.*

It is evident that, for $i = n$, $H_n(M_n) = H_n(M)$ because $M_n = M$.

Proof of Lemma 2.16. The relative homology $H_*(M_{k+1}, M_k)$ is concentrated in degree $k+1$ because, by homotopy invariance and excision,

$$H_*(M_{k+1}, M_k) \cong \bigoplus_{p \in Crit_{k+1}(f)} H_*(D^{k+1}, \partial D^{k+1})$$

and $H_*(D^{k+1}, \partial D^{k+1})$ is concentrated in degree $k+1$. Then the relative homology long exact sequence for the pair (M_{k+1}, M_k) implies that $H_i(M_k) \cong H_i(M_{k+1})$ if $i \neq k, k+1$. From this the lemma follows by induction on k. $\square$

Definition 2.17. We define the map $\Delta_i \colon H_i(M_i, M_{i-1}) \to H_{i-1}(M_{i-1}, M_{i-2})$ as the composition

$$H_i(M_i, M_{i-1}) \to H_{i-1}(M_{i-1}) \to H_{i-1}(M_{i-1}, M_{i-2}),$$

where the first map is the connecting homomorphism and the second one is the projection.

Proposition 2.18. *The maps Δ_* satisfy the relation $\Delta_{i+1} \circ \Delta_i = 0$ and the homology of the complex*

$$\to H_{i+1}(M_{i+1}, M_i) \xrightarrow{\Delta_{i+1}} H_i(M_i, M_{i-1}) \xrightarrow{\Delta_i} H_{i-1}(M_{i-1}, M_{i-2}) \to \tag{11}$$

is isomorphic to the singular homology of M.

The homology of the complex (11) is called *cellular homology.*

Proof. $\Delta_{i+1} \circ \Delta_i$ is the composition of the maps

$$\begin{tikzcd}
H_{i+1}(M_{i+1}, M_i) \arrow[d] & & \\
H_i(M_i) \arrow[r] & H_i(M_i, M_{i-1}) \arrow[r] & H_{i-1}(M_{i-1}) \arrow[d] \\
& & H_{i-1}(M_{i-1}, M_{i-2})
\end{tikzcd}$$

The central row in the diagram is a piece of the homology exact sequence for the pair (M_i, M_{i-1}). This proves that $\Delta_{i+1} \circ \Delta_i = 0$.

In order to prove the isomorphism between cellular homology and singular homology, let us consider the following diagram.

$$
\begin{tikzcd}
 & & & H_i(M_{i+1}, M_i) = 0 & \\
H_i(M_{i-1}) = 0 \arrow[dr] & & H_i(M_{i+1}) \arrow[ur] & & \\
 & H_i(M_i) \arrow[ur] \arrow[dr] & & & \\
H_{i+1}(M_{i+1}, M_i) \arrow[ur] \arrow[rr, "\Delta_{i+1}"] & & H_i(M_i, M_{i-1}) \arrow[rr, "\Delta_i"] \arrow[dr] & & H_{i-1}(M_{i-1}, M_{i-2}) \\
 & & & H_{i-1}(M_{i-1}) \arrow[ur] & \\
 & & H_{i-1}(M_{i-2}) = 0 \arrow[ur] & &
\end{tikzcd}
$$

We can identify $\ker \Delta_i \cong H_i(M_i)$, so $\dfrac{\ker \Delta_i}{\operatorname{Im} \Delta_{i+1}} \cong H_i(M_{i+1})$ and therefore $\dfrac{\ker \Delta_i}{\operatorname{Im} \Delta_{i+1}} \cong H_i(M)$ by Lemma 2.16. $\square$

We consider the filtration $M_0 \subset \ldots \subset M_n = M$ coming from a Morse-Smale pair (f, g) on M and identify $H_i(M_i, M_{i-1})$ with $C_i(f, g)$. Given $p \in Crit_i(f)$ and $q \in Crit_{i-1}(f)$, we denote by $\langle \Delta_i(p), q \rangle$ the coefficient of q in $\Delta_i(p)$. We also define $n(p, q) = \#\mathcal{M}^*(p, q)$.

Proposition 2.19. $\langle \Delta_i(p), q \rangle = n(p, q)$

Proof. Let $W^u(p)$ be the unstable manifold of p. We denote $\Lambda_p = W^u(p) \cap \partial M_{i-1}$, so that $\Delta_i(p) = [\Lambda_p] \in H_{i-1}(M_{i-1}, M_{i-2})$. Let $W^s(q)$ be the stable manifold for any point $q \in Crit_{i-1}(f)$. Suppose for a moment that $\Lambda_p \cap W^s(q) = \emptyset$. Then, for T sufficiently large, $\varphi_T(\Lambda_p) \subset M_{i-2}$, which implies that $[\Lambda_p] = 0$ in $H_{i-1}(M_{i-1}, M_{i-2})$.

For the general case, let Λ'_p be obtained from Λ_p by removing small discs around the intersections $\Lambda_p \cap W^s(q)$. Then, as before, for T sufficiently large $\varphi_T(\Lambda'_p) \subset M_{i-2}$. Moreover, if we choose T and the discs appropriately, φ_T pushes every disc around an intersection point in $\Lambda_p \cap W^s(q)$ to a disc which is arbitrarily close to $W^u(q) \setminus int(M_{i-2})$. This proves the lemma because $\#(\Lambda_p \cap W^s(q)) = n(p, q)$. $\square$