MORSE THEORY NOTES

1. Local normal form of a function at a non-degenerate critical point and the Morse Lemma.

1.1. Linear algebra step. Any $B \in Sym_n$ non-degenerate (zero is not an eigenvalue) can be diagonalized by an orthogonal matrix:

$$Q^T BQ = diag(\lambda_1, \ldots, \lambda_n), \quad \lambda_i \neq 0, \quad Q \in O_n.$$

Without requiring Q to be orthogonal, it is easy to see we can assume all diagonal entries are ± 1:

$$Q^T BQ = A = diag(a_1, \ldots a_n), \quad a_i = \pm 1, \quad Q \in GL_n.$$

Claim. Fixing a standard form $A = diag(a_1, \ldots, a_n)$ with $a_i = \pm 1$, the diagonalizing matrix depends smoothly on B, for $B \in Sym_n$ close to A in the natural topology of Sym_n. More precisely, there exists a neighborhood $\mathcal{N}_A$ of A in Sym_n and a smooth map $P : \mathcal{N}_A \to GL_n$ so that:

$$P(B)^T BP(B) = A, \quad \forall B \in \mathcal{N}_A, \quad P(A) = \mathbb{I}_n.$$

We present the proof in the case $n = 2$ (for the inductive argument for all n, see [Hirsch, p.145], but beware the typo: the factor $|b|^{-1/2}$ multiplies all entries of Q). To avoid LateX coding, write 2×2 matrices in the form $T = [t_{11}, t_{12}, t_{21}, t_{22}]$.

Given $A = [a_1, 0, 0, a_2]$ with $a_i = \pm 1$ and $B = [b, c, c, d] \in Sym_2$ close to A (so $b \sim a_1, c \sim 0, d \sim a_2$), let $Q = |b|^{-1/2}[1, 0, -cb^{-1}, 1]$. By direct computation, and using $\frac{b}{|b|} = a_1$, we find:

$$Q^T BQ = [a_1, 0, 0, d - c^2 b^{-1}].$$

Note $d - c^2 b^{-1} \sim a_2 \neq 0$, so define $\alpha > 0$ by $\alpha^2(d - c^2 b^{-1}) = a_2$. Setting $P = Q[1, 0, 0, \alpha] = |b|^{-1/2}[1, 0, -cb^{-1}, \alpha]$, we find $P^T BP = [a_1, 0, 0, a_2] = A$. Clearly the function with values in GL_2:

$$B = [b, c, c, d] \mapsto P(B) = |b|^{-1/2}[1, 0, -cb^{-1}, |d - c^2 b^{-1}|^{-1/2}] \in GL_2$$

is smooth in a neighborhood $\mathcal{N} \subset Sym_2$ of A, and satisfies $P(A) = \mathbb{I}_2$.

1.2. Calculus step. Let $g : V \to \mathbb{R}$ be a smooth function, where $V \subset \mathbb{R}^n$ is a convex neighborhood of the origin. Suppose $g(0) = 0$ and $dg(0) = 0$ (thus, 0 is a critical point of g.) There exists a smooth function from V to Sym_n, $x \mapsto B_x$, so that:

$$g(x) = \sum_{i,j=1}^{n} b_{ij}(x) x_i x_j, \quad b_{ij}(0) = \partial^2_{x_1 x_j} g(0).$$

Proof. By the FTC, using $g(0) = 0$ and integrating along the ray in V from the origin to $x \in V$:

$$g(x) = \int_0^1 \frac{dg}{dt}(tx)dt = (\int_0^1 \sum_{i=1}^{n} \partial_{x_i} g(tx)dt)x_i.$$

Applying the same integration argument to the partial derivatives (using the fact they all vanish at the origin):

$$(\partial_{x_i} g)(tx) = \int_0^1 \frac{dg_{x_i}}{ds}(stx)ds = \sum_j (\int_0^1 (\partial^2_{x_j x_i} g)(stx)ds)x_j.$$

We conclude:

$$g(x) = \sum_{i,j=1}^n b_{ij}(x)x_i x_j, \quad b_{ij}(x) = \int_0^1 \int_0^1 (\partial^2_{x_i x_j} g)(stx)dsdt.$$

Remark: Note that by taking V small enough, we obtain that the image B_x of this map is contained in an arbitrarily small neighborhood of the Hessian matrix of g at 0, $H_g(0) = [\partial^2_{x_i x_j} g(0)]$. From the expression for B_x we have $B_0 = H_g(0)$.

1.3 Normal form lemma. Let $g : U \to \mathbb{R}$ be a smooth function in a neighborhood of $0 \in \mathbb{R}^n$. Suppose $g(0) = 0$ and 0 is a non-degenerate critical point of g. Let $A = diag(a_1, \dots, a_n), a_i = \pm 1$, be the standard diagonal form of the Hessian $H_g(0)$. Then there exist small neighborhoods $V, V_1 \subset U$ of the origin and a diffeomorphism $\varphi : V \to V_1$, $y = \varphi(x)$, so that $g \circ \varphi^{-1}$ has the form:

$$(g \circ \varphi^{-1})(y) = \sum_i a_i y_i^2.$$

Proof. By (1.2), we have $g(x) = \sum_{i,j} b_{ij}(x)x_i x_j$, where $B_x = (b_{ij}(x)) \in Sym_n$ is defined and smooth in a neighborhood V of 0, $B_0 = H_g(0)$. Taking V small enough, we find $B_x \in \mathcal{N}$ for all $x \in V$, where $\mathcal{N}$ is the neighborhhood of A in Sym_n found in (1.1), on which the smooth map $P : \mathcal{N} \to GL_n$ is defined. Let $Q_x = P(B_x) \in GL_n$ (defined by $Q_x^T B_x Q_x = A$) and define φ on V by:

$$\varphi(x) = Q_x^{-1}[x].$$

Then $\varphi(0) = 0$ and the differential at 0 $D\varphi(0) = \mathbb{I}_n$. By the inverse function theorem, φ is a diffeomorphism in a neighborhhood of the origin (which we still denote by V), with image equal to V_1, a second neighborhood of the origin.

Let $y = \varphi(x)$ Then $y = Q_x^{-1}x$, or $x = Q_x y$. Thus:

$$g(x) = x^T B_x x = y^T Q_x^T B_x Q_x y = y^T A y = \sum_i a_i y_i^2,$$

or $(g \circ \varphi^{-1})(y) = \sum_i a_i y_i^2$.

1.4 Morse Lemma. If $0 \in \mathbb{R}^n$ is a nondegenerate critical point of $g : V \to \mathbb{R}$ of index k, we may assume the first k entries a_i in the standard form A of $H_g(0)$ equal -1 and the last $n - k$ equal 1. Thus there exists a decomposition $\mathbb{R}^n = \mathbb{R}^k \oplus \mathbb{R}^{n-k}$, $y = (u, v)$, with respect to which we have:

$$g \circ \varphi^{-1}(y) = -|u|^2 + |v|^2.$$

Thus, if $f : M \to \mathbb{R}$ is a smooth function on a smooth manifold M and $p \in M$ is a nondegenerate critical point of f, with index k, there exists a local chart $\psi : U \to V \subset \mathbb{R}^n$ at p so that, in this chart:

$$f \circ \psi^{-1}(y) = f(p) - |u|^2 + |v|^2, \quad y = (u, v) \in V \subset \mathbb{R}^k \oplus \mathbb{R}^{n-k}.$$

4. Morse inequalities.

Let M be a compact n-dimensional manifold, $f : M :\to \mathbb{R}$ a Morse function. For $p = 0, \ldots, n$, let:

$\beta_p =$ rank $H_p(M)$ (Betti numbers), $\nu_p = \#\{$critical points of index $p\}$ ('type numbers').

Theorem. We have:

$$(a) \sum_{p=0}^{n} (-1)^p \nu_p = \sum_{p=0}^{n} (-1)^p \beta_p := \chi(M) \quad \text{(euler characteristic)}.$$

$$(b) \text{for each } 0 \leq m \leq n : \sum_{p=0}^{m} (-1)^{m+p} \nu_p \geq \sum_{p=0}^{m} (-1)^{m+p} \beta_p.$$

Note that adding the statements of (b) for $m-1$ and m, we find:

$$(c) \nu_m \geq \beta_m, 0 \leq m \leq n.$$

Statement (c) is striking: the number of critical points of a given index for *any* Morse function on M (a generic kind of function) is bounded below by the rank of homology in that dimension: topology gives a lower bound on critical points!

Also, comparing the statements (b) for $m = n$ and $m = n-1$ yields (a). For this reason, sometimes (a),(c) are referred to as "weak Morse inequalities", while (b) are the "Morse inequalities".

Preparation for proof. First, by perturbing f slightly (without changing the critical points z_i or their index), we may assume the $f(z_i)$ are all distinct. For instance, let U_i be a small open neighborhood of z_i, and consider a smooth function $\lambda_i : M \to [0, 1]$, supported in U_i and identically 1 in a smaller neighborhood of z_i. Then, for judiciously chosen small $\epsilon_i > 0$, the function $g(x) = f(x) + \sum_i \epsilon_i \lambda_i(x)$ achieves what we want.

Thus each critical value c_i has only one critical point z_i (of index k_i between 0 and n) in its preimage. Pick the indices so the critical values c_i are an increasing sequence, then choose regular values $a_i, i = 0, \ldots, N$ separating the critical values, with a_0 smaller than the minimum value c_0 of f and a_N larger than the maximum value c_N (so $f(z_0) = c_0$ with index 0 and $f(z_N) = c_N$ of index n):

$$a_0 < c_0 < a_1 < c_1 < \ldots < a_{N-1} < c_N < a_N.$$

For $j = 0, \dots N$, consider the sublevel set $M_j = f^{-1}[a_0, a_j] = \{x \in M; f(x) \leq a_j\}$, a manifold with boundary ∂M_j equal to the (regular) level set at a_j. Clearly $M = M_N$ is the increasing union of the M_j, with $M_0 = \emptyset$. Adopt the notation:

$$\beta(p, j) = \text{ rk } H_p(M_j), j = 0, \dots, n, \quad \text{so } \beta(p, 0) = 0, \beta(p, N) = \beta_p.$$

$$\alpha(p, j) = \text{ rk } H_p(M_j, M_{j-1}), \quad j = 1, \dots N.$$

In part (3) we proved that M_j differs from M_{j-1} (up to homotopy type) by attaching a cell of dimension k_j. More precisely, there exists a closed k_j-cell e^{k_j} contained in the interior of M_j, with boundary $\partial e^{k_j} \subset M_{j-1}$, so that $M_{j-1} \cup e^{k_j}$ is a deformation retract of M_j. So we have inclusions:

$$A_j := (M_{j-1} \setminus \partial e^{k_j}) \subset M_{j-1} \subset M_{j-1} \cup e^{k_j}, \quad \overline{A_j} \subset M_{j-1},$$

$$(M_{j-1} \cup e^{k_j}) \setminus A_j = e^{k_j}, \quad M_{j-1} \setminus A_j = \partial e^{k_j}.$$

Thus, by the excision theorem (excising A_j) and the fact $M_{j-1} \cup e^{k_j}$ is a deformation retract of M_j, we have, for each $p = 0, \dots, n$:

$$H_p(M_j, M_{j-1}) \approx H_p(M_{j-1} \cup e^{k_j}, M_{j-1}) \approx H_p(e^{k_j}, \partial e^{k_j}),$$

which has rank equal to 1 if $p = k_j$ and 0 otherwise.

Thus the terms in the sum $\sum_{j=0}^{N} \alpha(p, j)$ (for fixed p) are equal to 1 if $k_j = p$ and zero otherwise. In other words, the value of the sum is the number of j for which $k_j = p$, that is, the number of critical points of f of index p, which we have denoted by ν_p.

Proof of statement (a) in the theorem. Recall the fact (exercise 5.5 in [Rotman, p. 87]) that in any long exact sequence of finitely generated abelian groups:

$$0 \to A_n \to A_{n-1} \to \dots \to A_1 \to A_0 \to 0$$

the alternating sum of ranks vanishes: $\sum_{j=0}^{n} (-1)^j rk(A_j) = 0$. Apply this to the long exact sequence in homology of the pair (M_j, M_{j-1}), for fixed $j = 1, \dots, N$:

$$0 \to H_n(M_{j-1}) \to H_n(M_j) \to H_n(M_j, M_{j-1}) \to H_{n-1}(M_{j-1}) \to \dots$$

$$\dots \to H_0(M_{j-1}) \to H_0(M_j) \to H_0(M_j, M_{j-1}) \to 0.$$

Grouping the terms in the alternating sum of ranks by threes, we find:

$$\sum_{p=0}^{n} (-1)^p [\beta(p, j-1) - \beta(p, j) + \alpha(p, j)] = 0.$$

Adding this expression over $j = 1, \dots, N$, we find (noting the telescoping sum, and using $\beta(p, N) = \beta_p, \beta(p, 0) = 0$):

$$\sum_{p=0}^{n} \sum_{j=1}^{N} (-1)^p \alpha(p, j) = \sum_{p=0}^{n} (-1)^p \sum_{j=1}^{N} [\beta(p, j) - \beta(p, j-1)] = \sum_{p=0}^{n} (-1)^p \beta_p,$$

and recalling that the sum on the left equals $\sum_{p=0}^{n}(-1)^p \nu_p$, this proves statement (a).

Proof of statement (b) in the theorem. This is similar, considering instead the long exact sequence for the pair (M_j, M_{j-1}) starting at homology in dimension m:

$$0 \to K_{m,j} \to H_m(M_{j-1}) \to H_m(M_j) \to H_m(M_j, M_{j-1}) \to H_{m-1}(M_{j-1}) \to \dots$$

$$\dots \to H_0(M_{j-1}) \to H_0(M_j) \to H_0(M_j, M_{j-1}) \to 0.$$

Here $K_{m,j}$ is the kernel of the inclusion homomorphism $H_m(M_{j-1}) \to H_m(M_j)$; denote its dimension by $\kappa_{m,j}$. Again grouping the alternating sum of ranks by threes (and including the additional term), we find for fixed j:

$$\sum_{p=0}^{m}(-1)^p[\beta(p, j-1) - \beta(p, j) + \alpha(p, j)] + (-1)^{m+1}\kappa_{m,j} = 0,$$

or

$$0 \le \kappa_{m,j} = \sum_{p=0}^{m}(-1)^{p+m}[\beta(p, j-1) - \beta(p, j) + \alpha(p, j)].$$

Adding over $j = 1, \dots, N$ as before, we find:

$$\sum_{p=0}^{m}(-1)^{p+m}[\nu_p - \beta_p] \ge 0,$$

which is statement (b).

Problems. (Source: [Hirsch, *Differential Topology*])

1. Let $M \subset \mathbb{R}^q$ be a compact smooth submanifold. For each $v \in S^q$, let $f_v : M \to \mathbb{R}$ be the function $f_v(x) = \langle v, x \rangle$. Then the set of $v \in S^q$ such that f_v is a Morse function is open and dense.

2. Define the function $g : RP^n \to \mathbb{R}$ on real projective space by the formula:

$$f[x_0, x_1, \ldots, x_n] = \frac{\sum \lambda_j |x_j|^2}{\sum |x_j|^2},$$

where the λ_j are distinct real numbers. Show f is a Morse function of type $(1, 1, 1, 1, \ldots 1, 1, 1)$. Describe the CW structure on RP^n that follows from this.

3. Let $f : S^n \to \mathbb{R}$ be a Morse function invariant under the antipodal map $x \mapsto -x$. Then f has at least two critical points of each index $0, 1, \ldots, n$. [Consider the function induced on RP^n. The $\mathbb{Z}_2$ Betti numbers of RP^n are $1, 1, \ldots, 1$.]

4. Let S be a compact orientable surface of genus p. (a) Every Morse function on S has at least $2p + 2$ critical points. (b) Some Morse function on S has exactly this number of critical points.