NOTES ON MORSE THEORY BY J.MILNOR

ZIYANG QIN

Abstract. These are my notes (and also transcripted handwritten notes) on *Morse Theory* by J.Milnor.

Notations.

(1) $\mathbb{R}$: Set of real numbers.
(2) M: A Riemannian manifold.
(3) $h(x) : M \to \mathbb{R}$: Arbitrary height function.
(4) M^a: Intersection of M with $h(x) \leq a$.
(5) T_pM: Tangent space of M at a point p.
(6) $g_* : T_pM \to T_qM$: pushforward for $g(p) = q, g : M \to N$.
(7) f_{**} on T_pM: Hessian of f at p.
(8) $B_r(p)$ and $B(p, r)$: They all stand for a ball of radius r centered at p. (Unless specified, the dimension equals the ambient space)
(9) e^λ: λ-cell; ∂e^λ: boundary of a λ-cell.

1. Non-Degenerate Smooth Functions On a Manifold

1.1. **Definitions and Lemma.**

Lemma 1.1. *Let f be a smooth function in a convex neighborhood V of 0 in $\mathbb{R}^n$ with $f(0) = 0$. Then $f(x_1, \dots, x_n) = \sum_{i=1}^n x_i g_i(x_1, \dots, x_n)$ for some suitable C^∞ functions g_i defined in V, with $g_i(0) = \frac{\partial f}{\partial x_i}(0)$.*

Proof. Take a point $X = (x_1, x_2, \dots, x_n) \in V$. Since the neighborhood is convex, the path from 0 to X is in the neighborhood.

We parametrize this as $\gamma(t) = (tx_1, tx_2, \dots, tx_n)$, $t \in [0, 1]$. Note that $f(\gamma(0)) = 0$, $f(\gamma(1)) = f(x)$.

$$f(x) = f(\gamma(1)) - f(\gamma(0)) = \int_0^1 \frac{d}{dt}\left[f(\gamma(t))\right] dt$$
$$= \int_0^1 \frac{d}{dt}[f(tx_1, tx_2, \dots, tx_n)]dt$$

Let $u_i = tx_i$(dummy variables of f).

$$= \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial u_i}(tx_1, \dots, tx_n)\frac{du_i}{dt}$$
$$= \int_0^1 (\sum_{i=1}^n \frac{\partial f(tx_1, \dots, tx_n)}{\partial u_i} \cdot x_i)dt$$
$$= \sum_{i=1}^n x_i \cdot (\int_0^1 \frac{\partial f(tx_1, \dots, tx_n)}{\partial u_i} dt)$$
$$= \sum_{i=1}^n x_i g_i(x_1, \dots, x_n).$$

□

Definition 1.2 (Index)**.** The index of a bilinear functional H is defined to be the dimension of the subspace on which H is negative definite.

Lemma 1.3 (Lemma by Morse)**.** *Let p be a non-degenerate critical point for f. Then there exists a local coordinate system $(y^1, \dots, y^n)$ in a neighborhood U of p with $y^i(p) = 0$ for all i and:*

$$f = f(p) - (y^1)^2 - \cdots - (y^\lambda)^2 + (y^{\lambda+1})^2 + \cdots + (y^n)^2 \tag{1.1}$$

*throughout U. λ is called the **index** of p.*

Remark 1.4. This means that in a neighborhood around non-degenerate critical points, the smooth function f locally behaves like hyperbolic functions no matter which "direction" you look into.

We partition the proof of 1.3 to be two claims.

Claim 1. If there is some expression like 1.1, λ is the index defined in Definition 1.2.

Proof. For any coordinate system $(z^1, \ldots, z^n)$, if:

$$f(q) = f(p) - (z^1(q))^2 - \cdots - (z^\lambda(q))^2 + (z^{\lambda+1}(q))^2 + \cdots + (z^n(q))^2.$$

Then the hessian is:

$$\frac{\partial f}{\partial z^i \partial z^j} = \begin{cases} -2 \text{ if } i = j \leq \lambda \\ 2 \text{ if } i = j > \lambda \\ 0 \text{ otherwise.} \end{cases}$$

Then the matrix looks like:

$$\begin{bmatrix} -2 & & & & & \\ & \ddots & & & & \\ & & -2 & & & \\ & & & 2 & & \\ & & & & \ddots & \\ & & & & & 2 \end{bmatrix}$$

Then there is a subspace of T_pM with $\dim = \lambda$ where f_{**} is negative definite, and a subspace V of dimension $n - \lambda$ on hwich f_{**} is positive definite. Rank Nullity Theorem guarantees that there cannot be a subspace of $\dim > \lambda$ on which f_{**} is negative definite. So λ is indeed the index. □

Claim 2. A suitable coordinate system $(y^1, \ldots, y^n)$ exists.

Proof. We show existance by constructing an example. Choose a local coordinate system in which p is the origin in $\mathbb{R}^n$ and $f(p) = f(0) = 0$.

By Lemma 1.1, we have:

$$f(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j g_j(x_1, \ldots, x_n).$$

By assumption, origin (p) is a critical point. So $g_j = \frac{\partial f}{\partial x^j}(0) = 0$.

Remark 1.5. This construction is crucial because in Lemma 1.1, we need 0 at origin.

Then we apply Lemma 1.1 again to g_j:

$$g_j(x_1, \dots, x_n) = \sum_{i=1}^{n} x_j \cdot h_{ij}(x_1, \dots, x_n)$$

for some smooth h_{ij}. Substitute back:

$$f(x_1, \dots, x_n) = \sum_{i=1}^{n} x_i x_j h_{ij}(x_1, \dots, x_n).$$

But here, h_{ij} may not be symmetric. Define the symmetric part of h_{ij} as:

$$\tilde{h}_{ij}(x) = \frac{1}{2}(h_{ij}(x) + h_{ji}(x)).$$

Then:

$$\begin{aligned}\sum_{i,j} x_i x_j \tilde{h}_{ij}(x) &= \frac{1}{2}\sum_{i,j} x_i x_j (h_{ij}(x) + h_{ji}(x)) \\ &= \frac{1}{2}(\sum_{i,j} x_i x_j h_{ij} + \sum_{i,j} x_i x_j h_{ji}).\end{aligned}$$

Since $x_i x_j = x_j x_i$, swap $i \leftrightarrow i$ for second part:

$$\sum_{i,j} x_i x_j \tilde{h}_i j(x) = 2 \cdot \frac{1}{2} \cdot \sum_{i,j} = \sum_{i,j} x_i x_j h_{ij}(x) = f(x).$$

So replacing h_{ij} with the symmetric version $\tilde{h}_{ij}$ does not change value.

Remark 1.6. For further notation, we will assume h_{ij} is symmetric. If it is not, we will use the above $\tilde{h}_{ij}$ to replace h_{ij} and keep the name h_{ij}.

Now, when we have symmetric version of h_{ij}, replace h with $\tilde{h}$.
Then we compute the partial derivatives at origin:

$$\begin{aligned}\frac{\partial}{\partial x_k}[x_i \cdot x_j \cdot h_{ij}(x)] &= \frac{\partial x_i}{\partial x_k} x_j \cdot h_{ij}(x) \\ &\quad + \frac{\partial x_j}{\partial x_k} x_i \cdot h_{ij}(x) + x_i x_j \frac{h_{ij}}{\partial x_i}(x).\end{aligned}$$

Note that $\frac{\partial x_i}{\partial x_k} = 1$ only when $i = k$, 0 otherwise. So we have $\sum_{i,j} \delta x_j h_{ij}(x) = \sum_j x_j h_{kj}(x)$, and the sum turns to be:

$$\frac{\partial f}{\partial x_k} = 2 \cdot \sum_j \cdot x_j \cdot h_{kj}(x) + \sum_{i,j} x_i x_j \frac{\partial h_{ij}}{\partial x_k}(x)$$

At $x = 0$, $x_i = 0$, $x_j = 0$, so $\frac{\partial f}{\partial x_k} = 0$ for all k.

This is consistent with the fact that p is a critical point.

Then we calculate the second derivatives:

$$\begin{aligned}\frac{\partial^2 f}{\partial x_l \partial x_k} &= \frac{\partial}{\partial x_l}(\frac{\partial f}{\partial x_k}) \\ &= 2 \cdot h_{lk}(x) + \text{ terms with } x_i \text{ or } x_j.\end{aligned}$$

Recall that at $x = 0$, $x_i = x_j = 0$, so:

$$\frac{\partial f}{\partial x_l \partial x_k} = 2 \cdot h_{lk}(0) = 2h_{kl}(0).$$

By remark 1.6, h_{ij} is symmetric.

Since p is a non-degenerate critical point, by definition, the hessian of f at p is non-singular. So $2h_{kl}(0)$ is non-singular.

Next we use a proof by induction to diagonalize the function f to the form in 1.1.

Inductive Hypothesis: Assume that in a neighborhood U, the function has been partially diagonalized:

$$\begin{aligned}f &= \pm(u_1)^2 \pm \cdots \pm (u_{r-1})^2 + \sum_{i,j \geq r} u_i u_j H_{ij}(u_1, \ldots, u_n) \\ &= \sum_{k=1}^{r-1} (\pm(u_k)^2) + \sum_{i,j \geq r} u_i u_j H_{ij}(u_1, \ldots, u_n).\end{aligned}$$

Base Case: r=1. This is trivial.

Quick observation of the matrix, we have:

- The top-left blocks are already diagonalized.
- Bottom-right block is the Hessian of $\sum_{i,j \geq r} u_i u_j H_{ij}(u)$ at p, which is $2(H_{ij}(0))_{i,j \geq r}$.

Recall that previous calculations show that the entire Hessian is non singular (This includes the bottom-right block) and p is non-degenerate.

So we perform a linear transformation on right bottom blocks to diagonalize $\sum_{i,j \geq r} u_i u_j H_{ij}(0)$.

It is a common result in Linear Algebra that we can find a transformation that produces a new matrix that satisfies:

- $H_{ij} = H_{ji}$.
- $H_{ii}(0) \neq 0 \quad \forall i$.

We perform this transofrmation, and replace the current H_{ij} with the output of the linear transofrmation.

Define a smooth function $g(u) = \sqrt{|H_{rr}|(u)|}$. Note that this is nozero in a neighborhood around 0 since $H_{rr}(0) \neq 0$.

We introduce new coordinates:

$$v_i = u_i \text{ for } i \neq r$$
$$v_r = g(u) \cdot \left[u_r + \sum_{i>r} \frac{u_i H_{ir}(u)}{H_{rr}(u)} \right].$$

Remark 1.7. The terms in the bracket are designed so that when you compute $(v_r)^2$, the cross terms involving u_r, u_i for $i > r$ will cancel out. Other terms are unchanged for $i \neq r$.

Remark 1.8. **This is the most crucial part in the proof.**

The map $(u_1, \dots, u_n) \to (v_1, \dots, v_n)$ has a non singular Jacobian at the origin(by construction it is an id matrix + an invertible block). Then, *Inverse Function Theorem* A.5 implies that $(v_1, \dots, v_n)$ forms a valid smooth coordinate system in some small enough neighborhood or the origin.

Then we can proceed with Inductive Step calculations:

Recall from Inductive Hypothesis, we have:

$$f = \sum_{k=1}^{r-1} (\pm (u_k)^2) + \sum_{i,j \geq r} u_i u_j H_{ij}(u_1, \dots, u_n).$$

We expand the later term:

$$\sum_{i,j \geq r} u_i u_j H_{ij} = H_{rr} {u_r}^2 + 2 \sum_{i>r} u_r u_i H_{ri} + \sum_{i,j>r} u_i u_j H_{ij}$$
$$f = \sum_{k=1}^{r-1} (\pm (u_k)^2) + H_{rr} {u_r}^2 + 2 \sum_{i>r} u_r u_i H_{ri} + \sum_{i,j>r} u_i u_j H_{ij}$$

Complete square for $H_{rr} {u_r}^2 + 2 \sum_{i>r} u_r u_i H_{ri}$:

$$H_{rr}(u_r + \sum_{i>r} \frac{u_i H_{ri}}{H_{rr}})^2 = H_{rr} {u_r}^2 + 2 u_r \sum_{i>r} u_i H_{ri} + \frac{1}{H_{rr}} (\sum_{i>r} u_i H_{ri})^2.$$

So rewrite f as:

$$f = \sum_{k=1}^{r-1} (\pm(u_k)^2) + H_{rr}(u_r + \sum_{i>r} \frac{u_i H_{ri}}{H_{rr}})^2 \\ - \frac{1}{H_{rr}} (\sum_{i>r} u_i H_{ri})^2 + \sum_{i,j>r} u_i u_j H_{ij}.$$

Let ϵ be the sign function of H_{rr}(i.e. $\epsilon = \mathrm{sgn}(H_r r)$). Then $H_{rr} = \epsilon \cdot g^2(u)$:

$$H_{rr}(u_r + \sum_{i>r} \frac{u_i H_r i}{H_r r})^2 = \epsilon g^2(u) \cdot \left[u_r + \sum_{i>r} \frac{u_i H_{ri}}{H_{rr}} \right]^2 \\ = \epsilon {v_r}^2.$$

Seems good enough. But we also need to deal with the remainder terms:

$$-\frac{1}{H_{rr}} \left(\sum_{i>r} u_i H_{ri} \right)^2 + \sum_{i,j>r} u_i u_j H_{ij}. \tag{1.2}$$

Substituting $u_i = v_i$ for $i \neq r$:

$$-\frac{1}{H_{rr}} \left(\sum_{i>r} v_i H_{ri} \right)^2 + \sum_{i,j>r} v_i v_j H_{ij}.$$

Expand $\left(\sum_{i>r} v_i H_{ri}\right)^2$:

$$\left(\sum_{i>r} v_i H_{ri} \right)^2 = \sum_{i,j>r} v_i v_j H_{ri} H_{rj}.$$

Then equation 1.2 becomes:

$$\sum_{i,j>r} v_i v_j \left[H_{ij} - \frac{H_{ri} H_{rj}}{H_{rr}} \right].$$

Define $\left[H_{ij} - \frac{H_{ri}H_{rj}}{H_{rr}}\right]$ as H'_{ij}. Putting all things together, we have:

$$f = \sum_{k=1}^{r-1} \pm(v_k)^2 + \epsilon \cdot (v_r)^2 + \sum_{i,j>r} v_i v_j H'_{ij}. \\ = \sum_{k=1}^{r} \pm(v_k)^2 + \sum_{i,j>r} v_i v_j H'_{ij}.$$

Thus, proof done by induction. $\square$

Lemma 1.9. *A smooth vector field on M which vanishes outside of a compact set $K \subset M$ generates a unique 1-parameter group of diffeomorphisms of M.*

2. Homotopy Type in Terms of Critical Value

If f is a real valued function on a manifold M, let $M^a = \{p \in M : f(p) \leq a\} = f^{-1}(-\infty, a]$.

Theorem 2.1. *Let f be a smooth real valued function on M. Let $a < b$, $f^{-1}[a, b]$ is* ***compact*** *(2.3) and contains no critical value of f.*

Then M^a is diffeomorphic to M^b. Moreover, M^a is a deform contract of M^b so that inclusion map $M^a \to M^b$ is a homotopy equivalence.

Proof. We choose a Riemannian Metric on M. Let $\langle \cdot, \cdot \rangle$ denote the inner product. We have the identity:

$$\langle X, \operatorname{grad} f \rangle = X(f) \quad \text{(directional derivative of } f \text{ along } X).$$

Recall that $\operatorname{grad} f = 0$ if and only if p is a critical point of f. Since there are no critical points in $f^{-1}[a, b]$, $\operatorname{grad} f \neq 0$. To prove the theorem, we construct a vector field that will "push the points down" to a in a finite, uniform amount of time.

If we just use grad as flow, some points will be moved too quickly or too slowly. We need to define a new function to control this speed.

Also, if we consider a curve $C : \mathbb{R} \to M$ with velocity vector $\frac{dC}{dt}$, we have:

$$\langle \frac{dC}{dt}, \operatorname{grad} \rangle = \frac{dC}{dt}(f) = \frac{d(f \circ C)}{dt}. \tag{2.1}$$

So the rate of change of f along a curve C is just the inner product of velocity vector with gradient vector.

Define $X_q = \rho(q) \cdot (\operatorname{grad} f)_q$ with $\rho : M \to \mathbb{R}$ defined as:

$$\begin{cases} \rho(q) = \frac{1}{\langle \operatorname{grad} f, \operatorname{grad} f \rangle_q} \text{ for all } q \in f^{-1}[a, b], \\ \rho(q) = 0 \text{ outside.} \end{cases}$$

This ensures X is compactly supported. And 1.9 implies that for all $q \in M$, the curve $q \to \phi_t(q)$ is unique integral curve starting from q.

The family $\{\phi_t\}$ is the flow of vector field X. $\varphi_t : M \to M$ is the "displacements" of points.

Fix a point q. Consider $t \to f(\varphi_t, q)$. We take the derivative:

$$\frac{d}{dt} f(\varphi_t(q)) = \langle \frac{d\varphi_t(q)}{dt}, \operatorname{grad} f \rangle \text{ by 2.1.} \tag{2.2}$$

Since $q \to \varphi_t(q)$ is the integral curve induced by $X = \frac{d\varphi_t(q)}{dt}$, we can transform 2.2 to:

$$\begin{aligned}\frac{d}{dt}f(\varphi_t(q)) = \langle X, \operatorname{grad} f\rangle &= \langle \rho \operatorname{grad} f, \operatorname{grad} f\rangle \\ &= \rho\langle \operatorname{grad} f, \operatorname{grad} f\rangle \\ &= 1.\end{aligned}$$

- $a \to b$:

  Consider $\varphi_{b-a} : M \to M$.

  For any point $q \in M^a$, flowing forward for time $b - a$ increases f at rate 1. $f(\varphi_{b-a}(q) \leq a + (b - a) = b$. This implies $\varphi(b - a) \in M^b$.

  This by definition is onto and invertible, so diffeomorphic (first part of theorem 2.1) is proved.
- $b \to a$:

  This time we need to flow backwards. This is tricky since if a point is on $f(q) = b$, we need time $b - a$, but we don't need that long if it is not.

  We need to find a continuous family of maps $\gamma_t : M^b \to M^b$.

  Define:

$$\gamma_t = \begin{cases} q \text{ if } f(q) \leq a \\ \varphi_{t(a-f(q))}(q) \text{ if } a \leq f(q) \leq b. \end{cases}$$

  If $q \in M^a$, $\gamma_t(q) = q$ for all t.

  If $a \leq f(q) \leq b$ and $t = 0$, $\gamma_0 = \varphi_0(q) = q$, this is the identity.

  If $a \leq f(q) \leq b$ and $t = 1$, $\gamma_1(q) = \varphi_{a-f(q)}(q)$.

  Since $f(a) \geq a$, $a - f(q) \leq 0$.

  Recall that when the flow is between a and b:

$$\frac{d}{ds}f(\varphi_s(q)) = 1 \implies f(\varphi_s(q)) = f(q) + s.$$

  Hence $s = a - f(q) \leq 0$:

$$f(\gamma_1(q)) = f(\varphi_{a-f(q)}(q)) = f(q) + (a - f(q)) = a.$$

  So γ_1 is on the level set $f(q) = a$. So $\gamma_1(q) \in M^a$.

  If $f(q) = a$, $s = 0$, and it is identity. So:

$$\begin{cases} \text{if } q \in f^{-1}[a,b], \gamma_1(q) \in M^a. \\ \text{if } q \text{ on } f(q) = a, \gamma_1(q) \in M^a. \\ \text{if } f(q) \leq a, \gamma_1(q) \in M^a. \end{cases}$$

So γ_1 is a retraction onto M^a. γ is also continuous since the flow φ_t is continuous.

□

Remark 2.2. The theorem is of course false when we remove "compact" condition. Trivial counter example inclusdes "removing a point from the walls of a cylinder". See *Diagram 3* in Milnor's book.

Remark 2.3. Furthermore, the theorem is not generally true if we replace "compact" with "complete" in the argument.

A counter-example is a shape that looks like the surface of Torricelli's Trumpet. This is non compact and non closed.

Theorem 2.4. *Let $f : M \to \mathbb{R}$ be a smooth function. Let p be a non-degenerate critical point with index λ. Let $f(p) = c$. Assume for some $\epsilon > 0$, $f^{-1}[c-\epsilon, c+\epsilon]$ is compact and contains no critical points of f other than p.*

Then for all sufficiently small ϵ, $M^{c+\epsilon} = f^{-1}(-\infty, c+\epsilon]$ has homotopy type of $M^{c-\epsilon} = f^{-1}(-\infty, c-\epsilon]$ with a λ-cell attached.

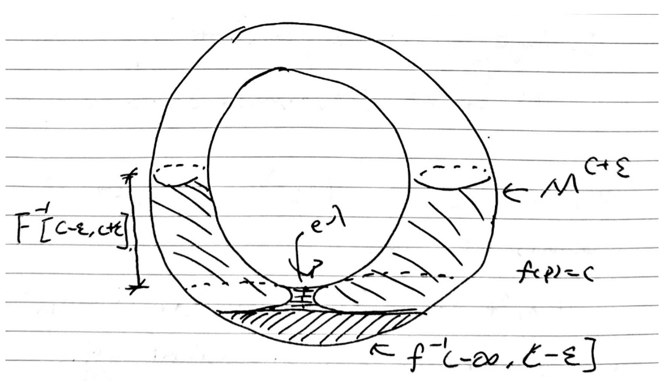

FIGURE 1

Proof. Since p is a non-degenerate critical point, we can apply Morse's Lemma 1.3.

Choose local coordinates $u_1, \ldots, u_n$ in a neighborhood U of p such that:

$$f = c - (u_1^2 + \cdots + u_\lambda^2) + (u_{\lambda+1}^2 + \cdots + u_n^2).$$

(As always, λ is the index: number of negative eigenvalues of f at p.)
Choose $\epsilon > 0$ small enough such that:

(1) $f^{-1}[c-\epsilon, c+\epsilon]$ is compact and contains no other critical points than p.
(2) Image of U under coordinate map contains the closed ball:

$$\{(u_1, \dots, u_n) : \sum u_i^2 \leq 2\epsilon\}.$$

Define the λ-cell as the set of points U where:

$$(u_1^2 + \cdots + u_\lambda^2) \leq \epsilon \text{ and } u_{\lambda+1} = \cdots = u_n = 0.$$

(In other words, we are just defining a λ dimensional disk.)
Then on the disk e^λ, f takes values:

$$f = c - (u_1^2 + \cdots + u_\lambda^2) \leq c,$$

with $f = c$ only at p. The boundary is where $u_1^2 + \cdots + u_\lambda^2 = \epsilon$ and $u_{\lambda+1} = \cdots = u_n = 0$.

So $f = c - \epsilon$ on the boundary, this is attached as we like. e^λ intersect $M^{c-\epsilon}$ only at $f = c - \epsilon$.

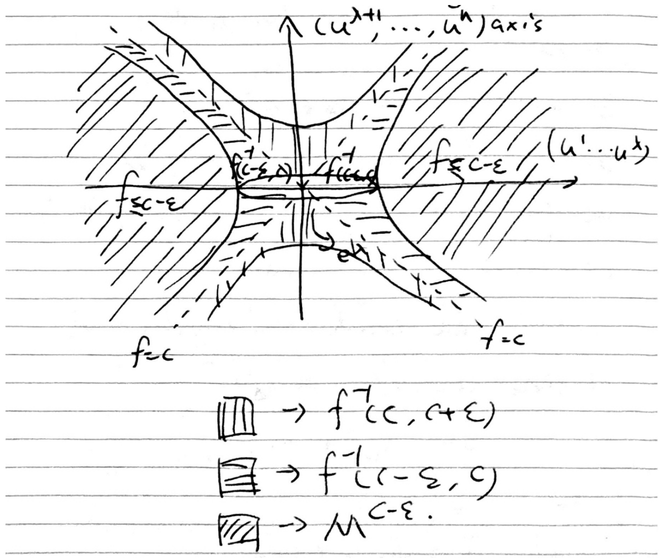

FIGURE 2

Because of the critical point, f is not useful. Define a new function F that coincides with f outside a neighborhood of p but is lower than f at p.

Choose a smooth function $\mu : \mathbb{R} \to \mathbb{R}$ with the following properties:

(1) $\mu(r) > \epsilon$.
(2) $\mu(r) = 0$ for $r \geq 2\epsilon$.
(3) $-1 < \mu'(r) \leq 0$ for all r, $\mu' = \frac{du}{dr}$.

(Why we require properties like this? See remark 2.5)

It is non increasing and non negative. We can always find such a function.

Within the coordinate neighborhood, define $a = (u_1^2, \cdots + u_\lambda^2)$, $b = (u_{\lambda+1}^2 + \cdots + u_n^2)$. So under this definition, $f = c - a + b$. Define:

$$\begin{aligned} F &= f - \mu(a + 2b) \\ &= c - a + b + \mu(a + 2b). \end{aligned}$$

Outside U, $F = f$. Since $\mu(a+2b) = 0$ when $a+2b \geq 2\epsilon$, and ball $\sum u_i^2 \leq 2\epsilon$. F must be well defined and smooth on all of M.

Claim 3.

$$F^{-1}(-\infty, c + \epsilon]$$

Proof. If outside $a + 2b \leq 2\epsilon$:

$$\mu(a + 2b) = 0 \implies F = f \implies F(q) \leq c + \epsilon \text{ if and only if } f \leq c + \epsilon.$$

If inside $a + 2b \leq 2\epsilon$:

Since $\mu \geq 0$, $F \leq f$. Also note that $f = c - a + b \leq c + b$. Also, since $a \geq 0$, $a + 2b \leq 2\epsilon \implies b \leq \epsilon$.

This means for all q in the ellipsoid $a + 2b \leq 2\epsilon$, $f(q) \leq c + \epsilon$. Now, if $F \leq c + \epsilon$, since $F \leq f$, we have $f \leq c + \epsilon$ automatically.

Thus $F^{-1}(-\infty, c + \epsilon] \subseteq M^{c+\epsilon}$.

On the other hand, if $f \leq c + \epsilon$, then:

$$F(q) \leq f(q) \leq c + \epsilon \implies M^{c+\epsilon} \subseteq F(-\infty, c + \epsilon].$$

This proves the claim. ∎

Remark 2.5. Why are we requiring these properties in 2?

Consider $a = (u_1)^2 + (u_2)^2 + \cdots + (u_\lambda)^2$, $b = (u_{\lambda+1})^2 + \cdots + (u_n)^2$, $F = f - \mu(a + 2b) = c - a + b - \mu(a + 2b)$.

Then $dF = \frac{\partial F}{\partial a} da + \frac{\partial F}{\partial b} db$.

$$\frac{\partial F}{da} = -1 - \mu'(a+2b), \quad \frac{\partial F}{\partial b} = 1 - 2\mu'(a+2b).$$

Property (3) of μ previously gurantees that $\frac{\partial F}{\partial a} < 0$ everywhere, $\frac{\partial F}{\partial b} > 0$ everywhere.

The one forms da and db linear independent except vanish at p.

At any other point $q \neq p$, at least one of da and db is non-zero. So dF cannot vanish at q.

Define $H = F^{-1}(-\infty, c-\epsilon] \setminus M^{c-\epsilon}$. $F^{-1}(-\infty, c-\epsilon]$ include points which:

(1) All points of $M^{c-\epsilon}$.
(2) Some points near $f(q) > c - \epsilon$ but $F(q) < c - \epsilon$.

Claim 4. $M^{c-\epsilon} \cup e^{\lambda}$ is a deformation retract of $M^{c-\epsilon} \cup H$.

Proof. To prove this, we construct a deformation retraction $\gamma_t : M^{c-\epsilon} \cup H \to M^{c-\epsilon} \cup H$, $t \in [0,1]$.

Use locall coordinates $u_1, \dots, u_n$. The construction needs to take care of 3 different regions.

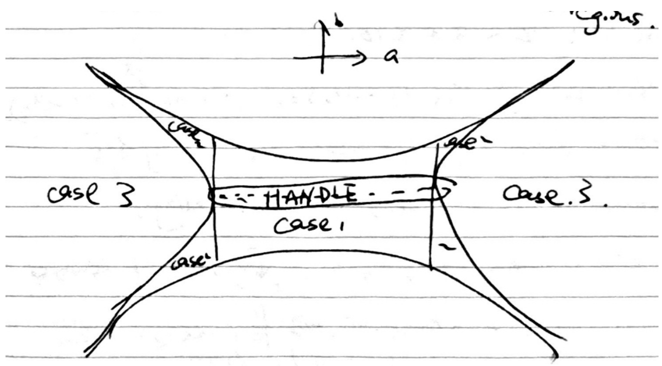

FIGURE 3. The middle part corresponds to Case 1. The four corner spikes corredpond to Case 2. The left and right parts correspond to Case 3. They will be addressed on by one below.

Case 1: $a < \epsilon$. Define $\gamma_t(u_1, \dots, u_n) = (u_1, \dots, u_\lambda, tu_{\lambda+1}, \dots, tu_n)$. If $t = 0$, γ_0 maps points to $(u_1, \dots, u_\lambda, 0, \dots, 0)$, which is a point in t_λ.

If $t = 1$, $\gamma_1 =$ identity. Recall that $\frac{\partial F}{\partial b} > 0$. F is increasing on principal direction of b. And b is positive by previous. So $t^2 b \leq b \implies F$ decreases under γ_t.

If $F(q) \leq c - \epsilon$ initially, then $F(\gamma_t(q)) \leq F(q) \leq c - \epsilon$. So $\gamma_t(q) \in F^{-1}(-\infty, c-\epsilon)$.

Case 2: $\epsilon < a < \epsilon + b$.

Define $\gamma_t(u_1, \ldots, u_n) = (u_1, \ldots, u_\lambda, s_t u_{\lambda+1}, \ldots, s_t u_n)$, where $s_t = t + (1-t)\sqrt{\frac{a-\epsilon}{b}}$.

When $t = 1$, $s_1 = 1$, the identity.

When $t = 0$, $s_0 = \sqrt{\frac{a-\epsilon}{b}}$, new b becomes $s^2 \cdot b = a - \epsilon \implies f = c - a + b = c - \epsilon$.

So $\gamma_0(q) \in f^{-1}(c-\epsilon) \subset M^{c-\epsilon}$.

Continuing on boundary:

(a) if $a = \epsilon$, $s_t = t$, same as **Case 1:**.
(b) if $a = \epsilon + b$, $s_t = 1$. We need to ensure this during construction of next case.

Note that s_t stays in $F^{-1}(-\infty, c-\epsilon]$ since $s_t \leq 1$, $s_t^2 b \leq b$. Recall F is increasing on b, $F(\gamma_t(q)) \leq F(q) \leq c - \epsilon$.

Case 3: $a > b + \epsilon$.

Define γ_t as he identity.

$$f(q) = c - a + b \leq c - (b + \epsilon) + b = c - \epsilon.$$

So $q \in M^{c-\epsilon}$. So $\gamma_t(q) = q$ remains in $M^{c-\epsilon}$.

Thus, take $s_t = \gamma_t$, γ_t is continuous across the regions and defines a deformation retraction $M^{c-\epsilon} \cup H$ onto $M^{c-\epsilon} \cup e^\lambda$.

□

Combine with Claim 3, we have $M^{c-\epsilon} \cup H$ is a deformation retraction of $M^{c+\epsilon}$. □

Theorem 2.6. *If f if a differentiable function on a manifold M with no degenerate critical points, and if each M^a is compact, then M has the homotopy type of a CW-complex, with one cell of dimension λ for each critical point of index λ.*

Lemma 2.7. *Let φ_0 and φ_1 be homotopic maps from the sphere $S^{\lambda-1}$ to a space X. Then the identity map on X extends to a homotopy equivalence:*

$$k : X \cup_{\varphi_0} e^\lambda \to X \cup_{\varphi_1} e^\lambda.$$

Proof. We try to find a continuous deformation k inverse l such that:

$$k \circ l \cong \mathrm{Id},$$
$$l \circ k \cong \mathrm{Id}\,.$$

On X, $k(x) = x$. On e^λ, use polar coordinates. Points are denoted as tu where $u \in S^{\lambda-1}$, $t \in [0,1]$. Define:

$$k(tu) = \begin{cases} 2tu \text{ if } 0 \le t \le \frac{1}{2} \\ \varphi_{2-2t}(u) \text{ if } \frac{1}{2} \le t \le \frac{1}{2}. \end{cases}$$

(where φ_t is the homotopy between φ_0 and φ_1.)

Continuity:

- At $t = \frac{1}{2}$:
  Inner ball: $k(tu) = u$. Outer ball: $k(tu) = \varphi_1(u)$.
  Recall that $u \in S^{\lambda-1}$, in $X \cup_{\varphi_1} e^\lambda$, u on boundary is just identified as $\varphi_1(u) \in X$. So these are the same point.
- At $t = 1$:
  outer ball: $k(tu) = \varphi_0(u)$. Similarly, when $t = 1$, u is identified $\varphi_0(u) \in X$. Recall k is identity on X, k maps φ_0 to itself.

Similarly define l:

$$l(x) = x \text{ for } x \in X$$
$$l(tu) = \begin{cases} 2tu \text{ for } 0 \le t \le \frac{1}{2} \\ \varphi_{2t-1}(u) \text{ for } \frac{1}{2} \le t \le 1. \end{cases}$$

When $t = \frac{1}{2}$, $l(tu) = \varphi_0(u) \implies u$ in $X \cup_{\varphi_1} e^\lambda$.
When $t = 1$, $l(tu) = \varphi_1(u)$. Continuity follows similarly.

Homotopy Equivalence: Check the requirements of $k \circ l$:
On X: $k \circ l(x) = x$. Constant.
On the cell:

$$0 \le t \le \frac{1}{4} : k \circ l(tu) = 4tu.$$
$$\frac{1}{4} \le t \le \frac{1}{2} : k \circ l(tu) = \varphi_{2-4t}(u).$$
$$\frac{1}{2} \le t \le 1 : k \circ l(tu) = \varphi_{2t-1}(u).$$

Define a homotopy $H : (X \cup_{\varphi_1} e^\lambda) \times [0,1] \to X \cup_{\varphi_1} e^\lambda$.

- On X: $H(x, \tau) = x$ for τ.

- On e^λ: define $H(tu, \tau)$ as:
  (1) $0 \le t \le \frac{1}{4}$: $H(tu, \tau) = t(4 - 3\tau)u$.
    - $\tau = 0$: $H(tu, 0) = 4tu$.
    - $\tau = 1$: $H(tu, 1) = tu = \mathrm{Id}$.
  (2) $\frac{1}{4} \le t \le \frac{1}{2}$:
    - $0 \le \tau \le \frac{1}{2}$: $H(tu, \tau) = \varphi_{2-4t+2\tau(4t-1)}u$.
    - $\frac{1}{2} \le \tau \le 1$: $H(tu, \tau) = [2 - 2\tau + (2\tau - 1)t]u$.
      * $\tau = 0$: $H = \varphi_{2-4t}(u) = k \circ l(tu)$.
      * $\tau = \frac{1}{2}$: $H = \varphi_1(u)$.
      * $\tau = 1$: $H = tu = \mathrm{Id}$.
  (3) $\frac{1}{2} \le t \le 1$:
    - $0 \le \tau \le \frac{1}{2}$: $H(tu, \tau) = \varphi_{2t-1+4\tau(1-t)}u$.
    - $\frac{1}{2} \le \tau \le 1$: $H(tu, \tau) = [2 - 2\tau + (2\tau - 1)t]u$
      * $\tau = 0$: $H = \varphi_{2t-1}(u) = k \circ l(tu)$.
      * $\tau = \frac{1}{2}$: $H = \varphi_1 u$.
      * $\tau = 1$: $H = tu = \mathrm{Id}$.

Thus H is continuous and we have shown $k \circ l \cong \mathrm{Id}$. The reverse direction is similar. □

Lemma 2.8. *Let $\varphi : S^{\lambda-1} \to X$ be an attaching map, and let $f : X \to Y$ be homotopy equivalence. then f extends to a homotopy equivalence:*

$$F : X \cup_\varphi e^\lambda \to Y \cup_{f\circ\varphi} e^\lambda.$$

Proof. Let $f : X \to Y$ be a homotopy equivalence with homotopy inverse $g : Y \to X$.

Similar to previous proof, there exist homotopies:

$$h : X \times I \to X \text{ with } h(x, 0) = g(f(x)) \text{ and } h(x, 1) = x,$$
$$k : X \times I \to Y \text{ with } g(y, 0) = f(g(y)) \text{ and } k(y, 1) = y.$$

Define $F : X \cup_\varphi e^\lambda \to Y \cup_{f\circ\varphi} e^\lambda$ by:

$$F(x) = f(x) \text{ for } x \in X,$$
$$F(tu) = tu \text{ for } tu \in e^\lambda \qquad (u \in S^{\lambda-1}, \quad t \in [0, 1]).$$

Define $G : Y \cup_{f\circ\varphi} e^\lambda \to X \cup_\varphi e^\lambda$ by:

$$G(y) = g(y) \text{ for } y \in X,$$
$$G(tu) = tu \text{ for } tu \in e^\lambda.$$

$G \circ F$:
On x, $(G \circ F)(x) = g(f(x)) = h(x, 0)$. On e^λ, $(G \circ F)(tu) = tu$. $\square$

3. The Morse Inequalities

Definition 3.1. Let S be a function thqt assigns integers to pairs of topological spaces. S is subadditive if whenever we have nested spaces $X \supset Y \supset Z$, the inequalities hold:

$$S(X, Z) \leq S(X, Y) + S(Y, Z).$$

If equality always hold, S is additive.

Example 3.2. *Betti Numbers* are subadditive. For a pair of topological spaces we are having above, Betti Bumber is defined as:

$$R_n(X, Y) = \operatorname{rank}_F H_n(X, Y; F).$$

(Where $R_n(X, Y)$ is the Betti Number, F is the coefficient vector field, and H_n is the nth homology group.)

Or informally, it is the number of "n dimensional holes". For example, for a torus:

$$R_0 = 1 \text{ one connected component,}$$
$$R_1 = 2 \text{ two independent loops,}$$
$$R_2 = 1 \text{ one cavaty.}$$

There is a beautiful image illustrating "$R_1 = 2$" for a torus on Wikipedia:

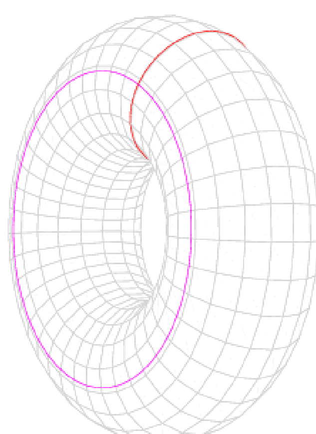

Figure 4

Example 3.3. Euler characterstic $x(X, Y)$ is additive.

Lemma 3.4. *Let S be subadditive and consider a filtration:*

$$X_0 \subset X_1 \subset \cdots \subset X_n.$$

Then $S(X_n, X_0) \leq \sum_{i=1}^{n} S(X_i, X_{i-1})$.

Theorem 3.5 (Weak Morse Inequalities)**.** *Let M be a compact manifold and $f : M \to \mathbb{R}$ is a smooth function with no degenerate critical points. If C_λ denotes the number of critical points of index λ on the compact manifold, then:*

$$R_\lambda(M) \leq C_\lambda,$$
$$\sum (-1)^\lambda R_\lambda(M) = \sum (-1)^\lambda C_\lambda.$$

Proof. Since M is compact and has only non-degenerate critical points, there are finitely many critical points. Order the critical values:

$$c_1 < c_2 < \cdots < c_k.$$

Choose a filtration:

$$\begin{cases} a_0 < c_1, \\ a_i \in (c_i, c_{i+1}), i \in [1, k-1], \\ a_k > c_k. \end{cases} \implies M^{a_0} \subset \cdots \subset M^{a_k} = M,$$

(where each interval contains i critical points.)

We are very familiar with this: when we pass through a critical point, Theorem 2.4 and Lemma 2.7 guarantees that $M^{a_i} \cong M^{a_{i-1}} \cup e^{\lambda_i}$. Now consider relative homotopy group:

$$\begin{aligned} H_h(M^{a_i}, M^{a_{i-1}}) &\cong H_h(M^{a_{i-1}} \cup e^{\lambda_i}, M^{a_{i-1}}), \\ &\cong H_h(e^{\lambda_i}, \partial e^{\lambda_i}). \end{aligned}$$

But $H_h(e^{\lambda_i}, \partial e^{\lambda_i})$ is $\mathcal{Z}$ if $h = \lambda_i$ and 0 otherwise. So:

$$R_h(M^{a_i}, M^{a_{i-1}}) = \begin{cases} 1 \text{ if } h = \lambda_i, \\ 0 \text{ otherwise.} \end{cases}$$

Recall R_h is subadditive. Fix h. For the filtration $M^{a_0} \subset M^{a_1} \subset \cdots \subset M^{a_k}$. We have by Lemma 3.4:

$$R_h(M^{a_k}, M^{a_0}) \leq \sum_{i=1}^{k} R_h(M^{a_i}, M^{a_{i-1}}).$$

- LHS: By construction, $a^0 <$ smallest critical point. So $R_h(M^{a_k}, M^{a_0}) = R_h(M, \emptyset) = R_h(M)$. (also, $a^k >$ greatest critical point).
- RHS: Previously we have $R_h(M^{a_i}, M^{a_{i-1}}) = 1$ only when $h = \lambda_i$. This means the critical point at level i has index h, so:

$$\begin{aligned}\sum_{i=1}^{k} R_h(M^{a_i}, M^{a_{i-1}}) &= \#\text{ of critical points with index } h,\\ &= C_h.\end{aligned}$$

This proves the first inequality. The second inequality follows similarly. We use an additive function χ.(euler characterstic) z □

Appendix A. An Interesting proof to Inverse Function Theorem

Theorem A.1 (Implicit Function Theorem)**.** *$F : \Omega \to \mathbb{R}^m$ be C^1 function on an open set $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$. Denote points in Ω by (x, y), where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. Suppose (a, b) satisfies:*

(1) $F(a, b) = 0$.
(2) The $m \times m$ matrix is invertible.

Then there exists open sets $X \subseteq \mathbb{R}^n$ containing a and $Y \subseteq \mathbb{R}^m$ containing b, and a unique C^1 function $f : X \to Y$ such that:

(1) $F(x, f(x)) = 0$ for all $x \in X$.
(2) $f(a) = b$.

Moreover, the Jacobian matrix of f is:

$$Jf(x) = -\left(\frac{\partial F(x, f(x))}{\partial y}\right)^{-1} \cdot \frac{\partial F(x, f(x))}{\partial x}. \tag{A.1}$$

Remark A.2. Let $z(x) = (x, f(x))$. Then A.1 easily follows from A.1:

$$\frac{d}{dx}F(z(x)) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot Jf(x) = 0$$
$$\implies Jf(x) = -(\frac{\partial F}{\partial y})^{-1} \cdot \frac{\partial F}{\partial x}.$$

The proof is skipped as it is a standard result in undergraduate analysis course. We present this theorem just because we need it to prove Inverse Function Theorem.

We will also use the following lemma.

Lemma A.3. *Let $f \in C^1(\Omega, \mathbb{R}^n)$, with $\Omega \subseteq \mathbb{R}^n$ open, and let $p \in \Omega$ such that $\det(JF(p)) \neq 0$. Then there exists $r > 0$ such that F is injective on $B_r(p)$.*

Proof. Since F is C^1, its first derivative is continuous, and $\det(JF)$ is continuous. Then there exists $r > 0$ such that for all $x \in B_r(p)$, $\det(JF)(x) \neq 0$. Take any two distinct points $a, b \in B_r(p)$. Apply Mean Value Theorem to each F_i: there exists points c_i between a_i and b_i such that:

$$F_i(b) - F_i(a) = \nabla F_i(c_i) \cdot (b - a).$$

Let M denote the matrix with i-th rtow to be ∇F_i. Since each $c_i \in B_r(p)$, $\det(M) \neq 0$.

Remark A.4. $\det(M) \neq 0$ is possible since $\det(JF(p)) \neq 0$, and $\det(JF)$ is continuous, so we may choose a neighborhood small enough such that $\det(M) \neq 0$.

Then M is invertible. So we have:

$$M(b-a) \neq 0 \implies F(b) - F(a) \neq 0 \implies F(b) \neq F(a).$$

□

Theorem A.5 (Inverse Function Theorem)**.** *Let $F : \Omega \to \mathbb{R}^n$ be C^1 function on $\Omega \subset \mathbb{R}^n$ and let $p \in \Omega$ such that $JF(p)$ is invertible. Then there exist an open set $X \in \Omega$ containing p, an open set $Y \subset \mathbb{R}^n$ containing $F(p)$, and a C^1 function $G : Y \to X$ such that:*

(1) $F(G(y)) = y$ for all $y \in X$.
(2) $G(F(x)) = x$ for all $x \in X$.

Moreover, $JG(y) = JF(G(y))^{-1}$ for all $y \in X$.

Proof. Define $\psi : \mathbb{R}^n \times \Omega \to \mathbb{R}^n$ by:

$$\psi(y, x) = F(x) - y.$$

Then ψ is C^1 on open set $\mathbb{R}^n \times \Omega$. Note that by construction, $\psi(F(p), p) = 0$.

Also, $\frac{\partial \psi}{\partial x}(y, x) = JF(x)$. Since $JF(p)$ is invertible, $JF(x)$ at $(F(p), p)$ is invertible.

Rename $y \leftrightarrow x$ and apply Implicit Function Theorem, there exist open sets $Y \subset \mathbb{R}^n$ containing $F(p)$ and $x' \subset \Omega$ containing p, and a C^1 function $G : Y \to X'$ such that:

$$\psi(y, G(y)) = 0 \text{ for all } y \in Y.$$

This implies $F(G(y)) - y = 0 \implies F(G(y)) = y \forall y \in Y$. This implies surjective.

By Lemma A.3, F is injective(we can shrink Ω if necessary to ensure this). So F is bijective.

Also, differentiate $F(G(y)) = y$ with respect to y. By chain rule:

$$\begin{aligned} &JF(G(y)) \cdot J(G(y)) = I \\ &\implies J(G(y)) = JF(G(y))^{-1}. \end{aligned}$$

□