Anders Krøger Evensen

Morse Theory applied to the Unitary Group

Bachelor's project in Mathematics
Supervisor: Markus Szymik
May 2021

Bachelor's project

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Anders Krøger Evensen

Morse Theory applied to the Unitary Group

Bachelor's project in Mathematics
Supervisor: Markus Szymik
May 2021

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

Introduction

There are many tools in mathematics which we can use to study manifolds. For example we might consider how functions on the manifold behave. By considering the critical points of functions from a manifold to the reals we can, through Morse theory, construct the manifold up to homotopy and even diffeomorphism. In this paper we introduce the basics of Morse theory as well as applying it to the the special case of the unitary group $U(n)$.

To give an intuitive explanation I will borrow some metaphors from [6].

Anyone that has ever hiked over hilly or mountainous terrain will have an intuitive understanding of the basic concepts of Morse theory. Walking across some terrain results in some change in altitude. During our hike we will come across some critical point where, momentarily, our altitude does not increase or decrease. These points essentially take one of three forms:

1. The top of a hill, where movement in any direction will decrease our altitude
2. The bottom of a crater, where any movement leads to an increase in altitude
3. The lowest point of a ridge between to mountains, i.e a saddle point. Here movement in one direction increases altitude, while the direction orthogonal to that decreases it.

This change in altitude may be considered as a smooth function $f : \mathbb{R}^2 \to \mathbb{R}$. Let's say that we find ourselves at the top of some mountain with at some pair of coordinates p. That means that we can choose some coordinates around p such that our function will look like $f(x, y) = -x^2 - y^2$. If we instead find ourselves in the bottom of some crater, we could choose coordinates such that we get $f(x, y) = x^2 + y^2$. And lastly if at a saddle point we can get coordinates that give us $f(x, y) = x^2 - y^2$.

Now if some terrain has a smooth height-function such that any critical point falls under one of these 3 types, we can gain global information about the shape of the terrain by only studying the critical points. This is the essence of Morse theory. We consider some function on a manifold with only non-degenerate critical points. In this case our function is called a *Morse function*. Given some Morse function we obtain global information about the manifold from its critical points.

There are however a functions that we would like to study that don't fulfill the requirements to be a Morse function. Therefore we must further generalize the theory. We do this by way of so-called Morse–Bott functions. For such functions on a manifold M we get that the set of critical points form submanifolds of M. Continuing with the above hiking example, if we've hiked to the top of a volcano we can now continuously move along the rim of its crater without ever losing or gaining height. So the height-function on such a terrain would not be a Morse function.

At the end of this paper we will see that, given the standard matrix representation of the unitary groups, the function on $U(n)$ given by $f(A) = Re(tr(A))$ is a Morse–Bott function. The main result of this paper is then that the critical submanifolds of f on $U(n)$ are Grassmannian manifolds.

The paper is split into two sections. The first covers the basics of Morse theory. We start by introducing the motivation for why we are interested in Morse theory as well as essential definitions and notations. When that is done we proceed with some fundamental results. The final of these being that any Morse function on a manifold gives a CW-complex homotopic to the manifold. We will also give the equivalent results for Morse-Bott functions.

The second section of the paper follows [3] and [7]. Though Frankel writes about all the classical groups, we will primarily concern ourselves with only $U(n)$.

1 Morse theory

1.1 Definitions and notation

To introduce Morse functions we first require a rigorous definition of what it means for a point on a manifold to be a critical point of a function.

Definition 1.1. : Let M be a smooth n-dimensional manifold and let

$$f : M \to \mathbb{R}$$

be a smooth map. Some point $p \in M$ with some coordinates $(x_1, x_2, \ldots, x_n)$ is said to be a critical point of f if: $\frac{\partial f}{\partial x_i} = 0, i = 1, \ldots, n$

Definition 1.2. : Let M and f be as in Def 1.1. Then we define the *Hessian* matrix at a point p, $H_f(p)$, as the $n \times n$ matrix with entries $(H_f(p))_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. Further a critical point p is said to be non-degenerate if the Hessian matrix at p is invertible, i.e. $\det(H_f(p)) \neq 0$.

We remark that even though it might not be immediately obvious, this definition is invariant under change of coordinates. That is to say in some neighborhood of some non-degenerate critical point $p \in M$ it will remain a non-degenerate critical point. Say we have some other coordinates $(y_1, \ldots, y_n)$ with change of basis matrix P. Then we get that in $\frac{\partial^2 f}{\partial x_i \partial x_j} = P^{-1}(\frac{\partial^2 f}{\partial y_i \partial y_j})P$.
These two definitions together lead to the definition of a Morse function.

Definition 1.3. (Morse function) A smooth function $f : M \to \mathbb{R}$ is said to be a Morse function if all its critical points are non-degenerate.

Let's give some concrete examples to familiarize ourselves with the definition.

Example 1.4. Consider the manifold S^1 as a sub-manifold of $\mathbb{R}^2$ given by coordinates $(\cos(t), \sin(t))$. The function $f(t) = \sin(t)$ is a Morse function. The critical points are at $t = \frac{-\pi}{2}$ and $t = \frac{\pi}{2}$ with Hessian matrix $(-\sin(t))$, which has non-zero determinant at the critical points.

Example 1.5. : Let's consider the function $f(x, y) = x^3 - 3xy^2$ on $\mathbb{R}^2$, the real part of the complex function $(a + bi)^3$. This function has Jacobian $(3x^2 - 3y^2, -6xy)$, so we see that the function has a critical point at $(0, 0)$. The Hessian at (x, y) is $\begin{pmatrix} 6x & -6y \\ -6y & -6x \end{pmatrix}$ which is seen to be singular at $(0, 0)$. So the function is not a Morse function.

In our hiking example we mentioned how we get coordinates looking like $\pm x^2 \pm y^2$ at critical points. Our first theorem, the Morse Lemma, generalizes this idea.

Theorem 1.6. *(Morse Lemma) Let p be a non-degenerate point of a Morse function $f : M \to \mathbb{R}$. Then there are local coordinates $(y_1, \ldots, y_n)$ about p such that f in these coordinates is given by*

$$f = f(p) - {y_1}^2 - {y_2}^2 - \cdots - {y_i}^2 + {y_{i+1}}^2 + \cdots + {y_n}^2$$

Proof. Before we begin the proof it's worth noting that we can assume that $p = 0$ and that $f(p) = 0$. This is done by an affine change in coordinates and by considering the smooth function $g = f - f(p)$ instead of f. Now we have that $\frac{\partial f}{\partial x_i}(0, \ldots, 0) = 0$ since 0 is a critical point. Since f is smooth we have that

$$f = \sum_{i=1}^{n} x_i g_i(x_1, \ldots, x_n)$$

where each g_i is some smooth function satisfying

$$\frac{\partial f}{\partial x_i}(0, \ldots, 0) = g_i(0, \ldots, 0) = 0.$$

Simply choose g_i to be $\int_0^1 \frac{\partial f}{\partial x_i}(tx_1, tx_2, \ldots, tx_n)dt$. Since each g_i is smooth we can repeat this process for each of those. This gives us

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

Now if we let $H_{ij} = \frac{h_{ij} + h_{ji}}{2}$ we get

$$f = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j H_{ij}(x_1, \ldots, x_n)$$

with $H_{ij} = H_{ji}$ and therefore also $\frac{\partial^2 f}{\partial x_i \partial x_j}(0,\dots,0) = 2H_{ij}(0,\dots,0)$. From our assumption that the critical point is non-degenerate we have that the matrix with entries H_{ij} is invertible. We may even assume, after some linear change of coordinates, that

$$2H_{11} = \frac{\partial^2 f}{\partial^2 x_i} \neq 0$$

now define

$$y_1 = \sqrt{|H_{11|}}\left(x_1 + \sum_{i=1}^{n} x_i \frac{H_{1i}}{H_{11}}\right)$$

The Jacobian of coordinate change from $(x_1,\dots,x_n)$ to $(y_1,x_2,\dots,x_n)$ is invertible so we have a new coordinate system $(y_1,x_2,x_3,\dots,x_n)$. Some calculations show that

$$y_1^2 = \begin{cases} H_{11}x_1^2 + 2\sum_{i=2}^n x_1 x_i H_{1i} + \frac{(\sum_{i=2}^m x_i H_{1i})^2}{H_{11}} & H_{11} > 0 \\ -(H_{11}x_1^2 + 2\sum_{i=2}^n x_1 x_i H_{1i} + \frac{(\sum_{i=2}^m x_i H_{1i})^2}{H_{11}}) & H_{11} < 0 \end{cases}$$

and that f in these coordinates becomes

$$f = \begin{cases} y_1^2 + \sum_{i=2}^n \sum_{i=2}^n x_1 x_i H_{ij} - \frac{(\sum_{i=2}^n x_i H_{1i})^2}{H_{11}} & H_{11} > 0 \\ -y_1^2 + \sum_{i=2}^n \sum_{i=2}^n x_1 x_i H_{ij} - \frac{(\sum_{i=2}^n x_i H_{1i})^2}{H_{11}} & H_{11} < 0 \end{cases}$$

Since we see that all terms after y_1^2 contain only $x_2, x_3, \dots, x_n$ we can repeat this process for n steps to obtain the desired result of $f = -y_1^2 - y_2^2 - \dots - y_i^2 + y_{i+1}^2 + \dots + y_n^2$ □

From the theorem we get the immediately following corollary:

Corollary 1.7. *The critical points of a Morse function $f : M \to \mathbb{R}$ are isolated. Further if M is compact f has finitely many critical points.*

Proof. I will not go into much detail for this proof. The idea is that there is only one critical point for $f = f(p) \pm x_1^2 \cdots \pm x_n^2$ in a neighborhood p. The second part follows from compact manifolds being sequentially compact and the first part. □

Now observe the fact that once we have such a coordinate system satisfying $f = f(p) - {y_1}^2 - {y_2}^2 - \dots - {y_i}^2 + {y_{i+1}}^2 + \dots + {y_n}^2$ then the Hessian matrix of f at p becomes

$$\begin{pmatrix} -2 & & & & & \\ & \ddots & & & & \\ & & -2 & & & \\ & & & 2 & & \\ & & & & \ddots & \\ & & & & & 2 \end{pmatrix}.$$

That is, there is some i-dimensional subspace of the tangent space of M at p where the Hessian is negative definite and some $(n-i)$-dimensional subspace where the Hessian is positive-definite. And this notion gives us a new defintion.

Definition 1.8. Let $f : M \to \mathbb{R}$ be a Morse Function with critical point p. The *index* of p is the maximal dimension of a subspace of the tangent space T_pM on which $H_f(p)$ is negative definite.

Example 1.9. Continuing with our Example 1.4. The top and bottom were our critical points with Hessian ± 1. So the critical point at the bottom has index 0 and the top has index 1. Likewise if we consider the height-function on S^n we will get two critical points with index 0 and n for the "north-pole" and "south-pole".

Example 1.10. : Lets consider the canonical Morse theory example: the torus with the height function. That is if we consider $T^2 = S^1 \times S^1$ with

$$f(\theta, \phi) = (R + r\cos(\theta))\cos(\phi)$$

for some $R > r > 0$. The gradient of this becomes

$$(-r\sin(\theta)\cos(\phi), -(r\cos(\theta) + R)\sin(\phi)).$$

So we have four critical points (0,0), (π,0), (0,π) and (π,π). The Hessian of the function becomes

$$\begin{pmatrix} -r\cos(\theta)\cos(\phi) & r\sin(\theta)\sin(\phi) \\ r\sin(\theta)\sin(\phi) & -(r\cos(\theta) + R)\cos(\phi) \end{pmatrix}.$$

Calculating eigenvalues one can find that the index for corresponding to the critical points above is respectively 0, 1, 1, 2. We see that at these are respectively a local minimum, a saddle point, another saddle point and a local maximum. In Figures 1 and 2 we have first pictured the torus with its critical points, secondly we see how one might go about constructing the torus by gluing together such minima, saddle points and maxima.

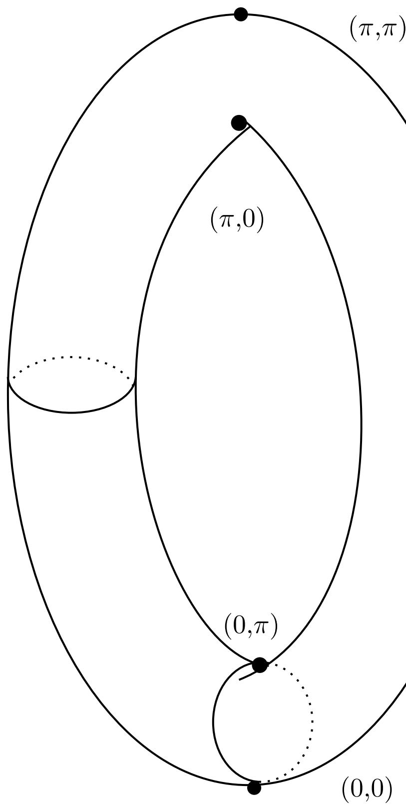

Figure 1: The torus with marked critical points of the height function

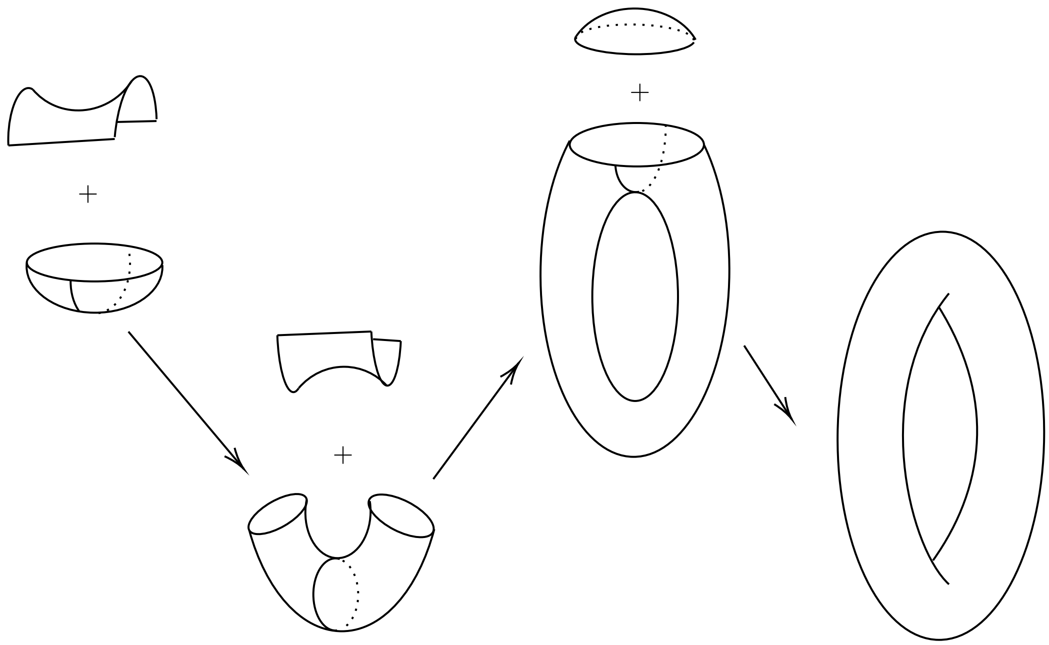

Figure 2: Constructing the torus

In Figure 2 we might see that before gluing we have some set that is just the points on the torus with height less than some number a. This motivates our next definition.

Definition 1.11. Let $f : M \to \mathbb{R}$ and $a, b \in \mathbb{R}$. Define the sublevel set M_a by

$$M_a = f^{-1}(-\infty, a] = \{m \in M \mid f(m) \leq a\},$$

and $M_{[a,b]} = \{m \in M \mid a \leq f(m) \leq b\}$

Example 1.12. Let $M = S^2$,the sphere centered at (0,0) in $\mathbb{R}^3$, and let f be the height function. i.e. $f(x, y, z) = z$ Then $M_c = \emptyset$, for all $c < -1$, M_{-1} is just the south pole, M_0 is the southern hemisphere and $M_d = S^2$, for all $d \geq 1$.

The next subsection is dedicated to explaining how these sublevel sets behave and how to construct any manifold M in the way we have done with the torus in Figure 2.

1.2 Fundamental results of Morse theory

The goal of this section will be to show that, through the lens of Morse theory, the points of interest on a manifold are the critical points. But before showing how the critical points affect the shape of our manifold lets first show that regular points leave our manifold unaffected.

Theorem 1.13. *Let $f : M \to \mathbb{R}$ be a smooth function and let $a, b \in \mathbb{R}$ such that the interval $[a, b]$ has no critical values and the set $M_{[a,b]}$ is compact. Then M_a is diffeomorphic to M_b. Further, M_a is a deformation retract of M_b*

The idea behind the proof of this theorem will be to follow the flow lines of some nice vector fields on M. We will follow the proof of Milnor [8] . And as such we will need to introduce the concept of a 1-parameter group of diffeomorphisms.

Definition 1.14. A ϕ is a smooth $\mathbb{R}$-action on M

$$\phi : \mathbb{R} \times M \to M$$

is called a *1-parameter group of diffeomorphisms.* That is:

1. For all $s, t \in \mathbb{R}$, the map $\phi_t : M \to M$ given by $\phi_t(m) = \phi(t, m)$ is a diffeomorphism

   and

2. $\phi_{t+s} = \phi_t \circ \phi_s$

Lets give some example before we proceed with the proof of Theorem 1.13.

Example 1.15. (Projection onto an axis) Lets start of with a somewhat trivial example. Lets consider the smooth map $\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ with $\phi(x, y) = (x, 0)$. Then we get that for any $t, s \in \mathbb{R}$

1. ϕ_t is a diffeomorphism as it is simply a translation of the line $y = t$ and
2. $\phi_{t+s} = (x, 0) = \phi_t \circ \phi_s$

Example 1.16. (Rotations of the circle) Let $M = S^1$ and let

$$\phi : \mathbb{R} \times S^1 \to S^1$$

be given by $\phi(t, (\cos(\theta), \sin(\theta))) = (\cos(\theta+t), \sin(\theta+t))$ for $t \in \mathbb{R}$ and $(\cos(\theta), \sin(\theta)) \in S^1$. Clearly this is a 1-parameter group of diffeomorphism.

Now given some vector field X on a manifold M and 1-parameter group ϕ we call X a generator of ϕ if

$$X_q(f) = \lim_{h \to 0} \frac{f(\phi_h(q)) - f(q)}{h}$$

for all continuous functions f and $q \in M$.

To prove Theorem 1.13 we will need the following lemma.

Lemma 1.17. *Let X be a smooth vector field on a manifold M such that X vanishes outside some compact $K \subset M$. Then X generates a unique 1-parameter group on M.*

This induced 1-parameter group is sometimes called the *flow* associated to the vector field X. It sends points along their integral curves on the vector field.

Proof. Let's start by assuming we have a 1-parameter group ϕ generated by some vector field X. Then if we fix some q let's consider the curve

$$t \mapsto \phi_t(q).$$

Such a curve will satisfy the ODE

$$\frac{d\phi_t(q)}{dt} = X_{\phi_t(q)}$$

as we get the

$$\frac{d\phi_t(q)}{dt}(f) = \lim_{h\to 0} \frac{f(\phi_{t+h}(q)) - f(\phi_t(q))}{h} = \lim_{h\to 0} \frac{f(\phi_h(p)) - f(p)}{h} = X_p(f).$$

It is a well known fact that such ODEs have unique solutions smoothly dependent on initial value. So since K is compact in M we have that we can find finitely many U_i with corresponding ε_i such that the U_i's cover K and

$$\frac{d\phi_t(q)}{dt} = X_{\phi_t(q)}$$

has a unique solution for $q \in U_i$ and for $|t| < \varepsilon_i$. Now let ε_0 be the smallest such ε. If we let $\phi_t = id$ for $q \notin K$ then it follows that for $|t| < \varepsilon_0$ and $q \in M$ we have a unique solution to our differential equation. Further the solutions can be considered smooth functions of both t and q and given $|t|, |s|, |t+s|$ all less than ε_0 we have $\phi_{t+s} = \phi_t \circ \phi_s$. Now if we have some r with $|r| > \varepsilon_0$ we can simply write $r = k\frac{\varepsilon_0}{2} + d$ for some integer k and some d with $|d| < \frac{\varepsilon_0}{2}$. Then we simply define

$$\phi_t = \underbrace{\phi_{\frac{\varepsilon}{2}} \circ \cdots \circ \phi_{\frac{\varepsilon}{2}}}_{\text{k times}} \circ \phi_d.$$

Now clearly ϕ satisfies our conditions and therefore completes our proof. $\square$

With this lemma we can proceed with proving the larger theorem at hand, i.e. Theorem 1.13.

Proof. The idea will be to define a nice vector field and to follow the integral curves (the solutions to the differential equations discussed in the previous proof) of that vector field. If we have our Morse function f and some Riemannian metric on M let the vector field ∇f be defined by the identity $\langle X, f\rangle = X(f)$. Where X is any vector field on M and $\langle -, -\rangle$ is the inner product defined by the Riemannian metric on M. Define the function $g : M \to \mathbb{R}$

given by $\frac{1}{\langle \nabla f, \nabla f \rangle}$ on the set $M_{[a,b]}$ and vanishes on some compact neighborhood of $M_{[a,b]}$. Now using this function let's define the vector field

$$X_q = g(q)\nabla f_q.$$

So since this vector field vanishes outside some compact set of M it satisfies the conditions of our lemma. That is we can find some 1-parameter group of diffeomorphisms

$$\phi_t : M \to M$$

generated by X. Now note the following property of the tangent vector for some curve c on M

$$\Big\langle \frac{dc}{dt}, \nabla f \Big\rangle = \frac{d(f \circ c)}{dt}$$

So define the curve on m given by

$$t \mapsto f(\phi_t(q))$$

for fixed $q \in M$. If q in addition lies in the set $M_{[a,b]}$ then we get that

$$\frac{d(f(\phi_t(q)))}{dt} = \Big\langle \frac{d(\phi_t(q))}{dt}, \nabla f \Big\rangle = \big\langle X, \nabla f \big\rangle = X(f) = +1$$

So we have a linear correspondence with derivative 1

$$t \mapsto f(\phi_t(q)).$$

With this construction its not hard to prove the first part of the theorem. So to find our diffeomorphism simply choose $\phi_{b-a} : M \to M$ maps M_a diffeomorphically onto M_b.

Finding a deformation retract is also quite easy now. Simply consider the following function

$$F : M_b \times I \to M_b$$

given by

$$F(q,t) = \begin{cases} id & , q \in M_a \\ \phi_{t-f(q)}(q) & , q \in M_{[a,b]} \end{cases}$$

This is clearly a deformation retract. □

So we have now shown that for some manifold M and Morse function f, the only change in the diffeomorphism type of the M happen at the critical points of f. The next step for us now will be to find some way to describe how these critical points affect M.

We have already described such a change in our example of the torus (see Figure 2), but we will have to generalize that process. Let $p_i \in M$ be the critical points of a Morse function f with corresponding critical values c_i such that $c_i \neq c_j$, for $i \neq j$. Order these such that $c_0 < c_2 < \cdots < c_n$. Now p_0 is a local minimum of f and thus f has the form $f = x_1^2 + x_2^2 + \cdots + x_m^2$ for some local coordinates at p_0. So the sublevel set $M_{c_0+\varepsilon} = \{(x_1, \ldots, x_m) | x_1^2 + x_2^2 + \cdots + x_m^2 \leq \varepsilon\}$ for some small $\varepsilon > 0$. That is to say that $M_{[c_0-\varepsilon, c_0+\varepsilon]}$ is an downwards facing m-dimensional disk D^m. Likewise for c_k we have a local we get a local maximum around p_k and so we have that $M_{[c_k-\varepsilon, c_k+\varepsilon]}$ looks like a downwards-facing disk. So when passing a critical point p_i that is either a local max or min we get the set $M_{c_i+\varepsilon}$ by adding either a downwards-facing or upwards-facing disk D^m to the set $M_{c_i-\varepsilon}$.

A slightly more difficult task is finding out how to describe the set $M_{[c_i-\varepsilon, c_i+\varepsilon]}$ if p_i is some critical point with index k that is neither maximal or minimal. To this end, let's once again consider the local coordinates around p_i given by the Morse lemma. So since p_i is of index k we get that we have coordinates $(x_1, \ldots, x_m)$ such that

$$f = c_i - x_1^2 - \cdots - x_k^2 + x_{k+1}^2 + \cdots + x_m^2$$

Now choose some $\varepsilon, \delta > 0$, s.t. $\delta \ll \varepsilon$ and that there are no other critical values in $M_{[c_i-\varepsilon, c_i+\varepsilon]}$. If we now consider the points in M satisfying

- $x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_m^2$
- $x_{k+1}^2 + \cdots + x_m^2$

This set is diffeomorphic to $D^k \times D^{m-k}$ and is called a *k-handle*. It contains within it the e^k-cell $D^k \times 0$

Figure 3 illustrates the situation.

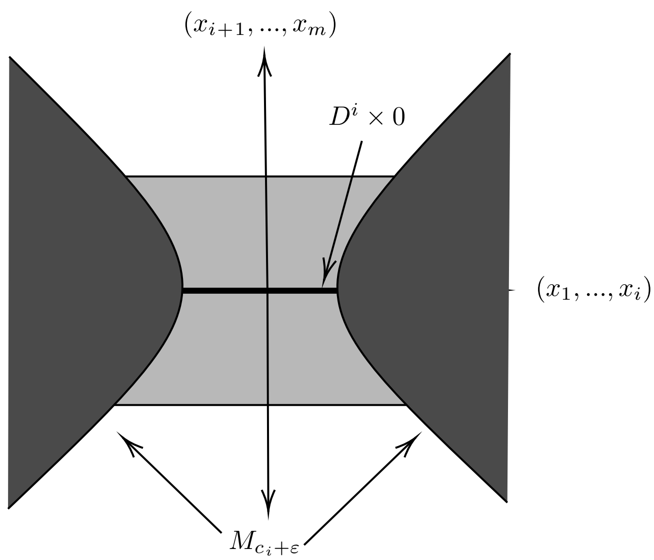

Figure 3

Considering the illustration above one might see that the set does not resemble a manifold, given the "corners" where the handle is attached. The next theorem seeks to find some manifold $\overline{M}$ that retracts into

$$M_{c_i-\varepsilon} \bigcup (D^i \times D^m - i).$$

Additionally we'll see that $M_{c_i+\varepsilon}$ retracts into $\overline{M}$ and $\overline{M}$ retracts into $M_{c_i-\varepsilon} \bigcup D^i \times 0$

Theorem 1.18. *Let f be a Morse function on the manifold M. Let p_i be a critical point of index i such that $M_{[c_i-\varepsilon,c_i+\varepsilon]}$ contains no other critical values. Then $M_{c_i+\varepsilon}$ is of the same homotopy type as $M_{c_i-\varepsilon}$ with an attached e^i-cell.*

Proof. The idea for this proof will be to construct some second function $F : M \to \mathbb{R}$ in such a way that F and f agree except for some small neighborhood of p_i where $F < f$. We will construct the aforementioned $\overline{M}$ from $F^{-1}(-\infty, c_i + \varepsilon)$. For this proof we will use M_a only for the set $f^{-1}(-\infty, a]$ so as not to confuse it with the set $F^{-1}(-\infty, a]$.

First let's assume we have coordinates $(u_1, \ldots, u_m)$ in a small neighborhood U of p_i such that the image of U under our coordinate chart contains the closed ball $\{(x_1, \ldots, x_m) \in \mathbb{R}^m | x_1^2 + x_2^2 + \cdots + x_n^2 \leq 2\varepsilon\}$. In the interest of making a suitable F let's first choose a function:

$$\rho : \mathbb{R} \to \mathbb{R}$$

such that

$$\rho(0) > \varepsilon$$
$$\rho(t) = 0 \text{ when } t \geq 2\varepsilon$$
$$\rho'(t) \in (-1, 0], \text{ for all } t$$

We also define the following two functions $\nu, \mu : U \to \mathbb{R}^+$, where

$$\nu(u_1, \ldots, u_m) = \sum_{n=1}^{i} x_n^2$$

$$\mu(u_1, \ldots, u_m) = \sum_{n=i+1}^{m} x_n^2$$

We now define our function F by letting

$$F = f - \rho(u_1^2 + \cdots + u_i^2 + 2u_{i+1}^2 + \cdots + 2u_m^2)$$

on U and $F = f$ outside of U. In terms of ν and μ we have

$$F = c_i - \nu + \mu + \rho(\nu + \mu)$$

on U. Now since outside the region bounded by the ellipsoid $\nu + 2\mu$ $F = f$. However in this region we have $F \leq f = c_i - \nu + \mu \leq c_i + \frac{1}{2}\nu + \mu \leq \varepsilon$. These two facts combined gives us $F^{-1}(-\infty, c_i + \varepsilon] = M_{c_i+\varepsilon}$.

If we're able to describe the critical points of F on U we might then also be able to use Theorem 1.13 to imply some information about M from F. Keep in mind that we don't require F to be a Morse function as the Theorem 1.13 only requires a function to be smooth. Now $dF = \frac{\partial T}{\nu} d\nu + \frac{\partial T}{\mu} d\mu$. But by our careful choice of ρ we get the two following inequalities

$$\frac{\partial F}{\partial \nu} d\nu = -1 - \rho'(\nu + 2\mu) < 0$$

and

$$\frac{\partial F}{\partial \mu} d\mu = 1 - 2(\nu + 2\mu) \geq 1$$

Now $d\nu$ and $d\mu$ only vanish at the origin so these facts combined gives us that the only critical point is at p_i. Since $F \leq f$ and $F^{-1}(-\infty, c_i + \varepsilon]$ we get $F^{-1}[c_i - \varepsilon, c_i + \varepsilon] \subset M_{[c_i - \varepsilon, c_i + \varepsilon]}$ and further $F^{-1}[c_i - \varepsilon, c_i + \varepsilon]$. But note that at the critical point $F(p) = c_i - \rho(0) < c_i - \varepsilon$. So $F^{-1}[c_i - \varepsilon, c_i + \varepsilon]$ contains no critical points and is therefore a deformation retract of $M_{c_i+\varepsilon}$. Now all that remains is to show that $F^{-1}[c_i - \varepsilon, c_i + \varepsilon]$ retracts into $M_{c_i - \varepsilon} \cup e^i$. We will create an explicit retraction for this. Firstly let the function

$$r : F^{-1}[c_i - \varepsilon, c_i + \varepsilon] \times I \to F^{-1}[c_i - \varepsilon, c_i + \varepsilon]$$

be the identity outside of U, but split the region on U into 3 cases. These are pictured in Figure 4

Case 1, where $\nu \leq \varepsilon$: Here we will let $r((u_1, \dots, u_m), t) = (x_1, \dots, x_i, tx_{i+1}, \dots, tx_m)$. We will let $D^i \times 0$ be our e^i-cell. As is noted in the figure. Then it's easy to see that the entire described set retracts into e^i.

Case 2, the region where $\varepsilon \leq \nu \leq \mu + \varepsilon$: First define the following function on the unit interval.

$$s(t) = t + (1-t)\sqrt{\frac{\nu - \varepsilon}{\mu}}$$

Then let r on this region be given by

$$r((x_1, \dots, x_m), t) = (x_1, \dots, x_i, s(t)x_{1+i}, \dots, s(t)x_m)$$

Note again that we have the identity for $t = 1$ and for $t = 0$ maps the entire region into $M_{c_i - \varepsilon}$. This part agrees with case 1 when the two cases meet, i.e $\nu = \varepsilon$.

Case 3. The region $M_{c_i - \varepsilon}$: Here we simply let r be the identity for all t. Again we see that this case agrees with the previous cases.

This completes the proof $\square$

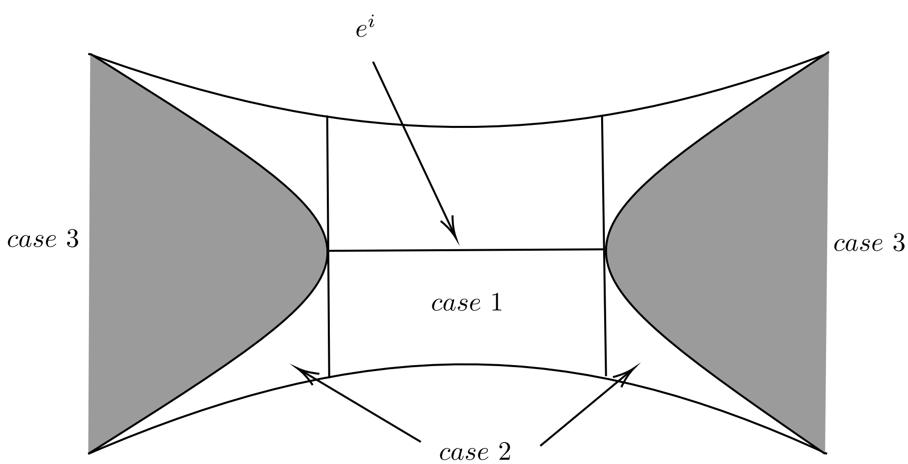

Figure 4

Corollary 1.19. *If the preimage of some critical value $c \in \mathbb{R}$ contains n critical points $\{p_i\}$ with corresponding indices λ_i, then $M_{c+\varepsilon}$ has homotopy type that of $M_{c-\varepsilon} \cup e^{\lambda_1} \cup \cdots \cup e^{\lambda_n}$*

Proof. We omit the proof as it is very similar to the preceding proof.

$\square$

1.3 Handlebody Decompositions

In addition to proving the theorem, we have done a lot of the work in showing that a Morse function defines a so called handlebody decomposition. Let's first define what we mean by this term.

Definition 1.20. Let X, Y be topological spaces and let $A \subset X$. Given some continuous map $f : A \to Y$ which we will call the *attaching map*. Define the attaching space

$$X \cup_f Y$$

as the quotient space

$$X \coprod Y /_{\sim}$$

Where the equivalence relation $\sim$ is given by $a \sim f(a)$. We say X is attached to Y along the attaching map f.

Example 1.21. Consider two copies of the unit interval. Let $A = \{0, 1\}$ and let $i : A$ be the inclusion of A into I. Then $I \cup_f I$ is a space homeomorphic to the circle S^1

Example 1.22. If A is some set with one point, then the attaching space $X \cup_f Y$ is $X \vee Y$, the wedge sum of X and Y.

The intuition for handlebody decompositions is that it is the smooth counterpart to CW-complexes, as the differing cells may be of differing dimensions and attaching maps are not smooth for CW-complexes. So whereas the constructions of CW-complexes often gives us information regarding the homotopy type of a manifold, constructing a handlebody decomposition might tell us something about the diffeomorphism type of said manifold. Our definition of a handlebody will be inductive through the attaching of spaces of the form $D^i \times D^{m-i}$. Not unlike that of CW-complexes with attaching of cells. The space $D^i \times D^{m-i}$ is called an *m-dimensional i-handle*, or just i-handle if the dimension m is implied through context.

We will assume all attaching maps in the definition are smooth embeddings, made possible by the tubular neighborhood theorem.

Definition 1.23. (Handlebody):

1. The m-disk D^m is an m-dimensional handlebody.
2. Let $\phi_1 : \partial D^{i_1} \times D^{m-i_1} \to \partial D^m$. The attaching space

   $$D^m \cup_{\phi_1} (D^{i_1} \times D^{m-i_1})$$

   is a m-dimensional handlebody. We will denote this space as $H(D^m; \phi_1)$

3. If we have a m-dimensional handlebody $\overline{M} = H(D^m; \phi_1, \ldots, \phi_k - 1)$ and let $\phi_k : (\partial D^{i_k} \times D^{m-i_k}) \to \partial\overline{M}$ be an attaching map. Then the attaching space
$$\overline{M} \cup \phi_k(D^{i_k} \times D^{m-i_k}) = H(D^m; \phi_1, \ldots, \phi_k)$$
is an m-dimensional handlebody.

Example 1.24. The possibly simplest example of a handlebody, outside of the disk D^n, is that of the n-sphere. We start as in the definition with an m-disk D^m. We will attach another D^m. Let the attaching map $\phi : \partial D^m \times 0 \to \partial D^m$ simply be given by the identity. The resulting handlebody $H(D;\phi)$ is the n-sphere S^n.

Example 1.25. Start with the 3-disk D^3. We attach a 1-handle somewhere onto the boundary of D^3 by some injective map ϕ. The resulting handlebody $H(D^3;\phi)$ is a space resembling a girya. I.e. $H(D;\phi)$ is diffeomorphic to a solid torus. This process could be repeated n times to construct a genus g handlebody.

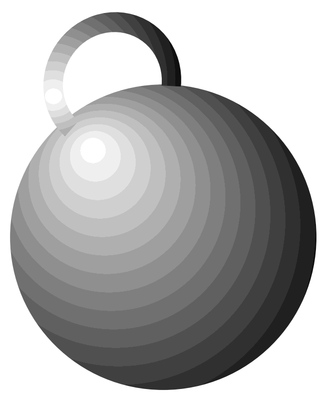

Figure 5: Girya

One might see that the structures in our definition of handlebodies very closely resembles a lot of those in our proof of Theorem 1.18. This observation motivates the following:

Theorem 1.26. *If M is a compact, smooth, m-dimensional manifold and $f : M \to \mathbb{R}$ is a Morse function. Then f defines a handlebody decomposition of M given by attaching an i-handle for each critical point of index i.*

Proof. The proof of will proceed inductively. We will assume that f has distinct critical values ordered such that $c_0 < c_1 < \cdots < c_k$ and with corresponding critical points $p_0, \ldots, p_k$. Our base case is that $M_{c_0+\varepsilon}$ is a handlebody. We quickly see that this is true as since we ordered our critical points the way we did p_0 is a local minimum. Then $M_{c_0+\varepsilon} \cong D^m$.

Now assume that $M_{c_{i-1}+\varepsilon}$ is a handlebody $H(D^m, \phi_1, \ldots, \phi_{i-1})$. The interval $[c_{i-1}+\varepsilon, c_i-\varepsilon]$ contains no critical values. So by Theorem 1.13 $M_{c_{i-1}+\varepsilon} \cong M_{c_i-\varepsilon}$. That is $M_{c_i-\varepsilon} \cong H(D^m, \phi_1, \ldots, \phi_{i-1})$. Proceeding as we did in our proof of Theorem 1.18 we get that if i_j is the index of c_i that

$$M_{c_i+\varepsilon} \cong M_{c_i-\varepsilon} \cup_{\phi_i} (D^{i_j} \times D^{m-i_j}).$$

Here ϕ_i is some attaching map determined by f. Thus through f we get $M_{c_i+\varepsilon} = H(D^m, \phi_1, \ldots, \phi_i)$. This completes our proof. □

An important remark to this theorem is that it does not give us a unique handlebody decomposition for the manifold M. As the theorem works for any given Morse function, another Morse function g would give another handlebody decompostion $H(D^m, \psi_1, \ldots, \psi_h)$. Indeed in the general case a given manifold X does not have a unique handlebody decompostition. As an example we have illustrated a handlebody construction of the 2-sphere S^2, differing from our previous example.

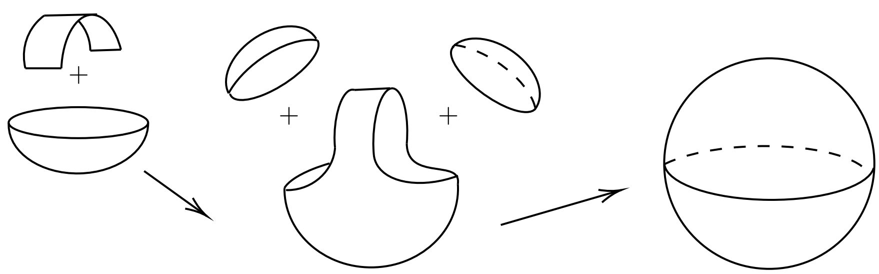

Figure 6

Now we used the constructions from our proof of Theorem 1.18 to make our handlebodies, but that theorem actually makes a statement about homotopy and attaching of e^i-cells. This may lead one to believe that a similar construction of CW-complexes might exist. The last theorem of this section, stated below, encapsulates this idea.

Theorem 1.27. *Given some smooth manifold M and a Morse function f on M such that each M_a is compact, then M is homotopy equivalent to a CW-complex with a cell of dimension i for each critical point of index i.*

Proof. We will state this theorem without proof. Those interested are referred to Milnor's proof [8, pp. 20-24]. □

1.4 Morse–Bott functions

Up until now we have only considered whether or not some given function *is* a Morse function or not. But a natural question to ask at this point is that of the existence of Morse functions. I.e. given some manifold M can we be sure that there is a Morse function $f : M \to \mathbb{R}$ and how is it given. The first part of this section will be dedicated to proving a stronger statement. Indeed given some smooth compact manifold M we will show that "almost all" smooth funcitons (in some sense) on M are Morse functions.

Theorem 1.28. *Given some smooth compact manifold M and some smooth function $g : M \to \mathbb{R}$, we may find another smooth function $f : M \to \mathbb{R}$ which is a Morse function.*

Proof. The outline of the proof is as follows. Since M is compact we may cover it with open subsets U_i and compact subsets K_i of U_i such that the K_i's also cover M. We will show that restricted to each of these U_i's we may construct a Morse function f_i from g. To complete the proof we will have to combine all these f_i's in such a way that we get a new function f which is a Morse function on all of M.

So if we are given some open subset $U \subset \mathbb{R}^n$ and a smooth function $h : U \to \mathbb{R}$ how do we make this function Morse? Our claim is that we may choose some point $(a_1, \ldots, a_n) \in \mathbb{R}$ that makes the function

$$\overline{h}(x_1, \ldots, x_n) = h(x_1, \ldots, x_n) - (a_1x_1 + \cdots + a_nx_n)$$

a Morse function. In fact we will see that most choices of $(a_1, \ldots, a_n)$ will work.

Our first step is to prove the assertion that we can create a Morse function for any open subset U of $\mathbb{R}^n$. To do this we will utilize Sard's Theorem. This theorem has become quite standard and as such we will not prove it. For a more detailed discussion we direct the reader to [5, pp. 205-207]. Simply put Sard's theorem states that given any smooth map between Euclidean spaces the set of critical values in the co-domain has Lebesgue measure 0. So if we consider the function

$$\nabla h = \left(\frac{\partial h}{\partial x_1}, \ldots, \frac{\partial h}{\partial x_n}\right)$$

This is a smooth map $\nabla h : \mathbb{R}^n \to \mathbb{R}^n$. Now observe that the Jacobian of ∇h is equal to that of the Hessian of h. So $p \in U$ being a critical point of ∇h is equivalent to $\det(H_h(p)) = 0$. Note also that $H_h(p) = H_{\overline{h}}(p)$.

Sard's theorem also directly implies a weaker statement. Since we know that the set of critical values of ∇h is 0 we at least know that there exists some $(a_1, \ldots, a_n) \in \mathbb{R}^n$ which is not a critical value of ∇h. This will be our candidate for the point described earlier. All that remains is to show that the function $\overline{h}$ is indeed Morse. Now $\overline{h}$ is obviously still smooth so pick some point $q \in U$ such that q is a critical point of $\overline{h}$. Then we have

$$\nabla h(q) = (a_1, \ldots, a_n)$$

But since $(a_1, \dots, a_n)$ is a regular value of ∇h we get

$$H_h(q) = H_{\overline{h}}(q) \neq 0$$

and therefore q is a non-degenerate critical point of $\overline{h}$. This holds for any critical point of $\overline{h}$ so it is a Morse function.

Now since M is compact we cover M with open coordinate neighborhoods $U_1, \dots, U_n$. We can also choose some compact $K_1, \dots, K_n$ lying in their respecting U_i such that $\bigcup_i^n K_i = M$. We will construct our Morse function f as above in such a way that it agrees with g outside some compact sets L_i contained in U_i and containing K_i, for $i = 1, \dots, n$. If we choose such a L_i's we can construct a smooth bump function $b_i : U_i \to \mathbb{R}$ satisfying the following:

$b_i(x) = 1$ when $x \in K_i$

$b_i(x) = 0$ when $x \in U_i \setminus L_i$

$b_i(x) \in [0, 1]$ when $x \in L_i$

That we can indeed construct such functions is a basic fact of manifold theory and as such we will not bother to prove it. We now have all the building blocks required to construct f. First, set $f_0 = g$. Define

$$f_i = \begin{cases} f_{i-1}(x_1, \dots, x_n) - (a_{1i}x_1 + \cdots + a_{ni}x_n)b_i(x_1, \dots, x_n), & (x_1, \dots, x_n) \in U_i \\ f_{i-1} & \textit{otherwise} \end{cases}$$

where $(a_{1i}, \dots, a_{ni})$ is chosen as above. Note that $f_i = f_i$ outside of L_i. Repeating this process for all i's we get a function $f_n = f$ which is Morse on all of M. This completes our proof □

Despite it being possible to create very many Morse functions, our theory still has some shortcomings. Creating Morse functions as we did in our above proof will tend to be a bit tedious. Further, many of the "nice" functions that we would like to work with, such as the trace which we consider in the next section, admit some sort of symmetry that stops them from being Morse functions. The symmetry of the function might for example cause the critical points to not be isolated. The conclusion one might reach is that we will have to expand our theory. Motivated by this we give the following more general definition:

Definition 1.29. (Morse–Bott function): Let M be a smooth manifold and let f be a smooth real-valued function on M. Then f is called a Morse-Bott function if it satisfies:

1. The critical points of f form a submanifold C of M made of a union of connected submanifolds $C = \cup_i C_i$
2. The null-space of the Hessian H_f at every point $c \in C$ coincides with the tangent space of C at c.

The manifolds C_i is called a *non-degenerate critical submanifolds* or simply *critical manifolds* of M.

Let's look at some examples to make ourselves comfortable with this definition.

Example 1.30. Any Morse function is also automatically a Morse-Bott function where the critical submanifolds are 0-dimensional.

Example 1.31. We've used as an example throughout the text so far the height function on an upright torus. We see that if we instead let f be the height function of the torus oriented such that it lies flat on the plane, i.e. the way one might usually depict it, then it is no longer a Morse function. It is however a Morse-Bott function. The critical submanifolds then become two copies of S^1.

Example 1.32. As we did with the normal Morse functions let's see if we can find some function on a manifold which is not a Morse–Bott function. Consider the manifold U which is just some neighborhood of $(0, 0) \in \mathbb{R}^2$ and let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by x^2y^2. The critical set C of f is then just the union of the x-axis and the y-axis in U. As C is not even a manifold f can not be a Morse–Bott function.

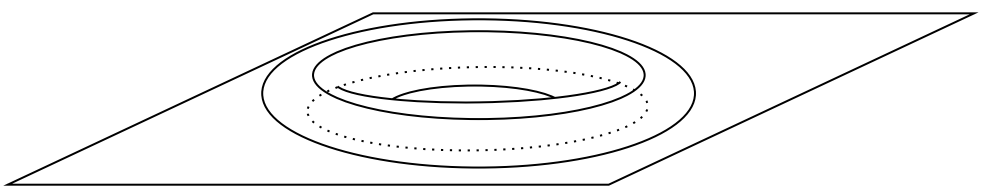

Figure 7: The torus, marked with critical submanifold

As for the standard Morse functions, the concept of indices of critical points is central to Morse–Bott functions. It will be defined in a similar fashion.

Definition 1.33. (Index of a critical submanifold): If C_i is a critical submanifold of M. The *index* of C_i is the dimension of the largest subspace of the tangent space at some $c \in C_i$ on which the Hessian H_f is negative definite.

Note that this definition is identical to that of Morse functions. Since the nullspace of H_f coincides with the tangent space on all of C_i, by our definition, and since each C_i is connected the index is independent on our choice of c. The following definition will also prove useful.

Definition 1.34. Let M be a smooth manifold and let $f : M \to \mathbb{R}$ be a Morse–Bott function. Then the normal bundle of a critical submanifold C_i splits into $P \oplus N$ such that the Hessian is positive definite when restricted to P and negative definite when restricted to N. The subbundle N is called the *negative bundle* of C_i.

Remark: The bundle P is simply the negative bundle of the function $-f$. Then the normal bundle of C_i is simply the direct product of the negative bundles of $-f$ and f

Though we do not prove this some of the theorems described in 2.2 have equivalent theorems for Morse–Bott functions. We state all of these in the following theorem. Proofs may be found in [2, pp. 324-325]

Theorem 1.35. *Let M be a smooth compact manifold and let f be a Morse-Bott function on M. Then:*

1. *If $[a, b]$ contains no critical values, then M_a is a deformation retract of M_b and M_a is diffeomorphic to M_b.*
2. *If $M_{[a,b]}$ contains a single critical submanifold C, then the space M_b has the same homotopy type as the attaching space*

$$M_a \cup N_C$$

   where N_C is the negative bundle of C. Attached along the sphere bundle of N_C.

3. *With the same setup of 2) $M_b = M_a \cup e^1 \cup \cdots \cup e^k$ where each e^l is a cell of dimension larger than or equal to the index of C*

2 The trace function and the unitary group

2.1 Introduction

Now that we have developed some theory, we should try to apply it in a setting that is a bit less trivial than the examples we have mentioned. Inspired by T. Frankel we will consider the unitary group $U(n) = \{U \in \mathcal{M}_{n\times n}(\mathbb{C}) \mid U^*U = I\}$. I.e. the group of unitary matrices with matrix-multiplication as binary operation. This group may also be endowed with the structure of a manifold. We will now quickly prove this fact.

Theorem 2.1. *The unitary groups, $U(n)$, are each endowed with a Lie-group structure.*

Proof. Lets start by proving that $U(n)$, with the binary operation of matrix multiplication, is indeed a group. $U(n)$, as a set, is a subset of the set of invertible matrices with complex entries. This is known to be a group so we only need to show that we have totality and inverses. If A and B are unitary then we have $(AB)^*(AB) = B^*A^*AB = B^*IB = B^*B = I$. So their product is again unitary. Now clearly for $A \in U(n)$ we have that $A^* = A^{-1} \in U(n)$.

So now we only need to show that $U(n)$ is a manifold. For this purpose we will utilize the Regular Value Theorem. As this is a basic theorem of manifold theory we refer the reader to [5, pp.21]. To this end let's consider the map $\Phi : M_{n\times n}(\mathbb{C}) \to H_{n\times n}(\mathbb{C})$ given by $A \mapsto A^*A$, where $M_{n\times n}(\mathbb{C})$ is the space of $n \times n$ matrices with complex entries and $H_{n\times n}(\mathbb{C})$ is the space hermitian matrices. Indeed we can see that ϕ maps into $H_{n\times n}(\mathbb{C})$ as

$$\Phi(A) = A^*A = A^*A^{**} = (A^*A)^* = (\Phi(A))^*$$

. Then at some $A \in M_{n\times n}(\mathbb{C})$ we get the tangent map

$$Df_A(V) = \frac{d}{dt} |_{t=0} (A + tV)^*(A + tV) = V^*A + A^*V$$

. So then we need to show that this map is surjective if A is unitary. In other words, if $W \in H_{n\times n}(\mathbb{C})$ is there some X such that $Df_A(X) = W$? We see that this is the case if we simply choose $X = \frac{1}{2}AW$ as then we have

$$Df_A(X) = \frac{1}{2}W^*A^*A + \frac{1}{2}A^*AW = \frac{1}{2}W^* + \frac{1}{2}W = W$$

. This concludes the proof.

□

Now for $n > 2$ the topology of $U(n)$ becomes rather intricate. For $n = 1$, the set $U(n)$ is simply all complex numbers with norm equal to 1. So $U(1) = S^1$. Already for $n = 2$ things become a bit hard to visualize. Though it is indeed known that $U(2)$ is diffeomorphic to $S^3 \times S^1$. So one might ask if we can find some Morse function that we might use to describe the shape of $U(n)$ in general. To this end let's try the function $f : U(n) \to \mathbb{R}$ with $A \mapsto Re(tr(A))$, the real part of the trace. However a problem with this arises straight away. The trace, and thus also the real part of the trace, is a class function. Meaning that for all $A \in U(n)$, we have $f(A) = f(gAg^{-1})$, for all $g \in U(n)$. Thus, if $f' = 0$ at some $A \in U(n)$ we also have that $f' = 0$ for all $B \in \Xi_A = \{gAg^{-1} \mid g \in U(n)\}$.

As we will see later, the subset Ξ_A is a (connected, positive dimensional) manifold. And therefore the critical points of the trace function are not isolated. By Corollary 1.7 we have that f can not be a Morse function. As we will see however, it is a Morse-Bott function.

Remark: Knudson actually gives an explicit Morse function on $U(n)$ in [6, pp. 38-39] by modifying the trace in the following way. If we choose some real numbers $1 < c_1 < c_2 < \cdots < c_n$, then Knudson shows that the function

$$f(A) = Re(c_1 A_{11} + c_2 A_{22} + \cdots + c_n A_{nn})$$

is a Morse function.

2.2 The trace and its critical submanifolds

Before we go about describing $U(n)$ we should discuss a fundamental Lie group property. Specifically we will talk about maximal tori and of a theorem regarding them. Though the theorem holds for any Lie-group, we will only give the motivational proof from $U(n)$.

Definition 2.2. A *torus* T in a compact Lie group G is any connected, compact and abelian Lie subgroup of G. If is T is maximal among such subgroups then it is called a *maximal torus* of G.

In $U(n)$ one maximal torus is given by the set of diagonal matrices in $U(n)$

$$D(n) = \{diag(e^{it_1}, \ldots, e^{it_n}) | t_1, \ldots, t_n \in [0, 2\pi)\}$$

Theorem 2.3. *Let $T \subset G$ be a maximal torus, then any element $g \in G$ is conjugate to some element in T. That is, there exists some $t \in T$ and some $h \in G$ s.t. $g = hth^{-1}$*

Proof. For $U(n)$ this is simply a restatement of the Spectral Theorem from linear algebra. For a proof see [4, pp. 401-402] $\square$

The first result towards describing $U(n)$ will be used to show that instead of finding all critical points of the trace we may instead only consider its critical points along some maximal torus. In this theorem and the rest of the section we will discuss notions of orthogonality of vectors and of vectors being tangent, as well as that of the gradient. This is with respect to the metric given by the inner-product inherited from the Riemannian metric on $U(n)$ given by the inner product

$$\langle A, B \rangle = tr(A^* B)$$

.

Theorem 2.4. *Let $f : U(n) \to \mathbb{R}$ be given by $x \mapsto Re(tr(x))$. Then ∇f is tangent to $D(n)$ at every point $t \in D(n)$*

Proof. First note that we may assume ∇f is non-zero as if it is zero then it is trivially tangent to all of $U(n)$. So if $\nabla f \neq 0$ at t then it is also non-zero on some neighborhood V of t. Consider the subset of G such that $f(g) = f(t)$. This is the level set of value of f at t. We will denote it

$$L_t(f) = \{g \in U(n) \mid f(g) = f(t)\} = f^{-1}(f(t))$$

.

Let's briefly consider an example similar to that in the beginning of the thesis. We will imagine we are hiking up a mountain. At some point we might come to a very steep section. We essentially have two options in proceeding with our goal of reaching the top. If we are experienced climbers or perhaps just a bit brave we might try continuing our climb straight upwards. That is choose the path where the height increases the most. If we are a bit more concerned for our safety we might want to find a less treacherous path. We proceed by walking along the level set of the height function at this height until we find a less steep part. These two paths will be perpendicular to each-other. In more technical terms: $L_t(f)$ is a $n^2 - 1$ dimensional submanifold which is tangent the gradient ∇f.

As we discussed in the introduction to this section the trace is a *class function*, i.e. $tr(gtg^{-1}) = tr(t)$. This means that the set $\{gtg^{-1} \mid g \in U(n)\}$ is a subset of $L_t(f)$. We will denote this subset Ξ_t. Let's try to describe this subset in more details as it is, as we will see, in fact a submanifold of $U(n)$. Define the map $\varrho : U(n) \to \Xi_t$ by $g \mapsto gtg^{-1}$. This map is by definition of Ξ_t onto. Let $C(t)$ be the centralizer of t, i.e. all elements of $U(n)$ commuting with t, i.e. $\{g \in U(n) \mid gtg^{-1} = t\}$. Note that for some element $c \in C(t)$ we have that $\varrho(gc) = \varrho(g)$. So this defines a new 1-1 map

$$\overline{\varrho} : (U(n) \big/ C(t)) \to \Xi_t$$

. Then we simply have that the manifold Ξ_t is given by

$$\Xi_t = \overline{\varrho}(U(n) \big/ C(h))$$

. As an example let's consider the special case of $t = I$, the identity matrix. Then the subset Ξ_t is again just I. Likewise $C(t) = U(n)$ and so $U(n) \big/ C(t) = I$.

Now let's choose some $\phi \in D(n)$ and some β be in the tangent space $T_\phi U(n)$ of $U(n)$ at ϕ. Then as $U(n)$ is just some submanifold of $M_{n\times n}(\mathbb{C})$ then an element of the tangent space to $U(n)$ is also just such a matrix. Now the tangent space is just the kernel of the tangent map described in 2.1. So the defining equation for the tangent space at ϕ is

$$\beta^*\phi + \phi^*\beta = 0.$$

Further the tangent space $T_\phi D(n)$ of $D(n)$ at ϕ is given by the set of diagonal matrices satisfying the same defining equation.

Let's now assume that β is orthogonal to $D(n)$. Then we have that for some $\eta \in T_\phi D(n)$ we have $\beta \perp \eta$. Now let δ be the matrix obtained by replacing all non-diagonal with zero. Then $\delta \in T_\phi D(n)$ and so $\delta \perp \beta$. Specifically we get that

$$df(\beta) = 0$$

as

$$Re(tr(\beta)) = 0.$$

This completes our proof. $\square$

The reason this proof is useful is that we have the following consequence *The set of critical values of f on $D(n)$ are the same as the critical values of* $f|_{D(n)}$. Then computing critical values is far easier. For some $\phi = (\phi_1, \ldots, \phi_n)$,

$$f(\phi) = \sum_i^n cos(\phi_i)$$

further making

$$df(\phi) = \sum_i^n -sin(\phi_i)$$

Then the critical values, call them α, along $D(n)$ are simply the matrices of the form

$$\begin{pmatrix} \pm 1 & & & 0 \\ & \pm 1 & & \\ & & \ddots & \\ 0 & & & \pm 1 \end{pmatrix}$$

Two such α, α' are conjugate if they have the same amount of negative signs. For convenience we will denote the set of block matrices given by $\begin{pmatrix} U(k) & 0 \\ 0 & U(n-k) \end{pmatrix}$ simply by $U(k) \times U(n-k)$. Likewise denote matrices of the form

$$\begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ 0 & & & & & -1 \end{pmatrix}$$

with k positive signs by $I(k) \times I(n-k)$. So since having the same amount of negative signs makes α, α' conjugate our search for the critical submanifolds reduces to finding the Ξ_α's where each α is of the form $I(k) \times -I(n-k)$ for some k. Now its not too hard to see that for such an α the centralizer is $C(\alpha) = U(k) \times U(n-k)$. So the our critical manifolds become

$$\Xi_\alpha = {U(n)} \Big/ {U(k) \times U(n-k)} = Gr_k(\mathbb{C}^n), \text{ for } k = 1, \dots, n$$

the Grassmannian manifold of complex k-planes in complex n-space.

What remains is to verify that f is indeed a Morse–Bott function as well as determining the index of each Ξ_α. For this part we will require another definition.

Definition 2.5. (Stable and unstable manifolds) Let F be some vector field on a manifold M vanishing on some set of points C and let $c \in C$. Then let ψ_t be the induced 1-parameter group. The *stable submanifold of* c is then

$$S_F(c) = \{m \in M | \lim_{t \to \infty} (m)\psi_t = c\}.$$

It is a submanifold(by the Stable Manifold Theorem) of M made from all the integral curves that end at c. Dually we can define the submanifold of integral curves with trajectories that diverge from c. We call it the *unstable manifold* and it is defined as the stable manifold of $-F$ at c:

$$U_F(c) = S_{-F}(c)$$

The stable manifold of C is

$$S_F(C) = \cup_i S_F(c_i), c_i \in C.$$

It maps diffeomorphically to the negative normal bundle of C. [1] We are now ready to state the main theorem of this section

Theorem 2.6. *Given a critical point $\alpha = I(k) \times I(n-k)$ of f on $D(n)$, α has as stable submanifold $U(k) \times -I(n-k)$ and therefore the index of Ξ_α is $dim(U(k) \times -I(n-k)) = k^2$.*

Proof. First, given α=$I(k) \times -I(n-k)$, consider the subgroup $\overline{C(\alpha)}$ of the centralizer given by matrices of the form $= U(k) \times I(n-k)$. Now the left translate of this subgroup $\alpha\overline{C(\alpha)}$ is $U(n) \times -I(n-k)$, namely our candidate for the stable submanifold wf α. As with the torus $D(n)$ we will want to show that it is tangent to ∇f.

Choose some $c \in \alpha\overline{C(\alpha)}$ and and let β be a vector of the tangent space of $U(n)$ at c. Now at any point $c \in U(n)$ the tangent space is given by $T_cU(n) = \{X \in M_{n\times n}(\mathbb{C}) | X^*c + c^*X = 0\}$. Since $c \in \alpha\overline{C(\alpha)}$ it is some block matrix of the form

$$\begin{pmatrix} c' & 0 \\ 0 & -I \end{pmatrix}.$$

Let us write β as a block matrix of same dimensions

$$\beta = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Written in this way a simple matrix calculation shows that the defining equation for the tangent space gives:

$$b_{11}^* c' + c'^* b_{11} = 0 \tag{2.7}$$

$$b_{22}^* + b_{22} = 0 \tag{2.8}$$

Now if we assume that β is orthogonal to $\alpha\overline{C(\alpha)}$ at c then we are done if we can show that $df(\beta) = 0$; i.e.

$$Re(tr(b_{11})) + Re(tr(b_{22})) = 0.$$

From 2.8 it is easy to see that we have $Re(tr(b_{22})) = 0$. Then what remains is to show that the first term is also zero. Choose some vector γ tangent to $\alpha\overline{C(\alpha)}$ at c. It is not to hard to see that for γ to be tangent it will have to be of the form

$$\begin{pmatrix} \gamma_{11} & 0 \\ 0 & 0 \end{pmatrix}.$$

But under our assumption that β is orthogonal to any such γ we get that β is orthogonal to $\begin{pmatrix} b_{11} & 0 \\ 0 & 0 \end{pmatrix}$ it follows that $b_{11} = 0$

The next part of the proof will be to confirm that $\alpha\overline{C(\alpha)}$ is indeed the stable manifold of α. First we will need some 1-parameter group on $\overline{C(\alpha)}$. Call this ψ_t. Then ψ_t takes on the form $\psi_t \times I(n-k)$. In order to get a curve tangent to $\alpha\overline{C(\alpha)}$ we make $\alpha\psi_t$. This is clearly tangent to $\alpha\overline{C(\alpha)}$ and has the property that $\alpha\psi_0 = \alpha$. We can now choose some $k \in \overline{C(\alpha)}$ such that the curve $k\psi_t k^{-1}$ travels along a maximal torus of $\overline{C(\alpha)}$. That is $k\psi_t k^{-1} = e^{it\theta_1} \times \cdots \times e^{it\theta_k} \times I(n-k)$. The corresponding curve $\alpha k\psi_t k^{-1}$ becomes, for some constant non-zero $(\theta_1, \dots, \theta_k)$

$$e^{it\theta_1} \times \cdots \times e^{it\theta_k} \times -I(n-k)$$

Evaluating our function along $\alpha\psi_t$ we get

$$f(\alpha k\psi_t k^{-1}) = (\sum_i^k cos(t_i r_i)) - (n-k) \leq f(\alpha)$$

and

$$\frac{d^2}{d^2t} f(\alpha k\psi_0 k^{-1}) = -\sum_i^k r_i^2.$$

So we have that on $\alpha\overline{C(\alpha)}$, the function f obtains its maximum at α. Also, more importantly, the stable manifold of α in $C(\alpha)$ is $\alpha\overline{C(\alpha)}$. This concludes the proof. $\square$

Remark:As shown in [3] this theorem may be generalized to all the classical groups. The simplest of which, given our proof for $U(n)$, is probably the symplectic group $Sp(n)$ as $Sp(n) \subset U(2n)$. One choice of maximal torus in $Sp(n)$ is that of the diagonal matrices in $U(2n)$ of the form

$$diag(e^{it_1} \times e^{-it_1} \times \cdots \times e^{it_n} \times e^{-it_n})$$

. The critical points are of the form $\alpha = I(2k) \times I(2n-2k)$ with

$$C(\alpha) = Sp(k) \times Sp(n-k)$$

and therefore again we get the critical submanifolds

$$\Xi_\alpha = {}^{Sp(n)}\Big/_{Sp(k) \times Sp(n-k)} = Gr_k(\mathbb{H}^n)$$

the Grassmanian manifold of quaternionic k-planes in quaternionic n-space. The result is similar for the orthogonal group with real Grassmannians, but to show that requires a bit more work.

References

[1] A. Banayaga. D. E. Hurtubise (2004). A Proof of the Morse-Bott Lemma *Expositiones Mathimaticae* , Volume 22, Issue 4 ,pp.365-373

[2] R. Bott.(1957). The Stable Homotopy of the Classical Groups *Annals of Mathematics* , Volume 70, Issue 2, pp. 313-337

[3] T. Frankel. (1965). Critical submanifolds of the Classical Groups and Stiefel Manifolds *Differential and Combinatorial Topology, A Symposium in Honor of Marston Morse* , (PMS-27), pp. 37-53

[4] S. H. Friedberg, A. J. Insel, L. E. Spence. (1989) *Linear Algebra* Harlow: Pearson

[5] V. Guillemin A. Pollack. (1974). *Differential Topology* New York: AMS Chelsea Publishing

[6] K. P. Knudson. (2015) *Morse Theory: Smooth and Discrete* Singapore: World Scientific

[7] H. Miller. (1985). Stable Splittings of Stiefel Manifolds *Topology* , Volume 24, Issue 4, pp.411-419

[8] J. Milnor. (1969) *Morse Theory* Princeton: Princeton University Press

NTNU
Kunnskap for en bedre verden