Morse Theory

Al Momin

Monday, October 17, 2005

1 THE example

Let M be a 2-torus embedded in $\mathbb{R}^3$, laying on it's side and tangent to the two planes $z = 0$ and $z = 1$. Let $f : M \to \mathbb{R}$ be the function which takes a point $x \in M$ to its z-coordinate in this embedding (that is, it's "height" function). Let's study what this function can tell us about the topology of the manifold M.
Define the sets

$$M^a := \{x \in M : f(x) \leq a\}$$

and investigate M^a for various a.

- $a < 0$, then $M^a = \phi$.
- $0 < a < \frac{2}{3}$, then M^a is a disc.
- $\frac{1}{3} < a < \frac{2}{3}$, then M^a is a cylinder.
- $\frac{2}{3} < a < 1$, then M^a is a genus-1 surface with a single boundary component.
- $a > 1$, then M^a is M is the torus.

Notice how the topology changes as you pass through a critical point (and conversely, how it doesn't change when you don't!). To describe this change, look at the homotopy type of M^a.

- $a < 0$, then $M^a = \phi$.
- $0 < a < \frac{1}{3}$, then M^a has the homotopy type of a point.
- $\frac{1}{3} < a < \frac{2}{3}$, then M^a, which is a cylinder, has the homotopy type of a circle. Notice how this involves attaching a 1-cell to a point.
- $\frac{2}{3} < a < 1$. A surface of geunus 1 with one boundary component deformation retracts onto a figure-8 (see diagram), and so has the homotopy type of a figure-8, which we can obtain from the circle by attaching a 1-cell.

- $a > 1$ Then M is just the torus, which we can obtain from the above by attaching a 2-cell (this is most obvious when you attach it onto the genus-1 surface with one boundary component first, and then deformation retract.

Do we recall defn of CW-complex here? Do we at least describe how to attach a k-cell to a CW-complex?

So, at least in this case, passing through a critical point has the effect of adding a cell to our CW-complex. The relationship between the dimension of the cell we add and the critical point is through the "index" of the critical point, as in the index of a quadratic form. We'll explain this in general in the next section. For this case, notice how the torus looks (more or less) locally like the graph of:

- $z = x^2 + y^2$ at the critical point corresponding to $a = 0$.
- $z = x^2 - y^2$ at the critical point corresponding to $a = \frac{1}{3}$
- $z = y^2 - x^2$ at the critical point corresponding to $a = \frac{2}{3}$
- $z = -x^2 - y^2$ at the critical point corresponding to $a = 1$.

In each case, notice how the number of negative signs (the "index") corresponds to the dimension of the cell we attach. This observation will generalize.

2 Definitions and Lemmas

Definition 1
A critical point of a function $F : M \to \mathbb{R}$ is a point p where $dF(p) = 0$. A critical point is said to be **non-degenerate if it's Hessian is non-degenerate.**

We can define the Hessian in two ways. The first is to take local coordinates and define it to be the symmetric bilinear form with matrix in those coordinates given by

$$\frac{\partial^2 f}{\partial x^i \partial x^j}$$

The other is an invariant definition. Given two vectors v, w in T_pM, define $f_{**}(v, w) = \tilde{v}_p(\tilde{w}(f))$, where $\tilde{v}$ and $\tilde{w}$ are vector fields extending v and w. Supposing it is well defined, we see it is symmetric because $[\tilde{v}, \tilde{w}]_p(f) = df(p)[\tilde{v}, \tilde{w}] = 0$ since p is a critical point for f. To see it is well defined, note that it is independent of the choice of $\tilde{v}$ by it's definition, then use the symmetry above (which holds for vector fields at least) to show it is also independent of the choice of $\tilde{w}$. One then sees that in local coordinates this definition coincides with the previous definition.

The Hessian, being a symmetric bilinear form which we can identify with the quadratic form $f_{**}(v, v)$, decomposes the tangent space at p into a direct sum of 3 subspaces: one where it is positive definite, one where it is negative definite, and one where it vanishes identically. Then non-degeneracy means that the nullity = dim(subspace on which it vanishes) is 0. We define the **index** to be the dimension of this maximal subspace on which the Hessian is negative definite.

Lemma 1 (Lemma of Morse)
Let p be a non-degenerate criticla point of f of index λ. Then there exists a coordinate system $(y^1, .., y^n)$ in a neighborhood U aournd p such that $y^i(p) = 0$ for each i and $f = f(p) - (y^1)^2 - ... - (y^\lambda)^2 + (y^{\lambda+1})^2 + ... + (y^n)^2$ throughout U.

Proof: Assume p is the origin and $f(p) = 0$ WOLOG, and write

$$f(x^1, ..., x^n) = \sum_{j=1}^{j=n} x^j \cdot g_j(x^1, ..., x^n)$$

(this can always be done), where $\frac{\partial g_j}{\partial x^j} = 0$ because p is a critical point. Then do the same thing to each g_j, writing it as

$$g_j(x^1, ..., x^n) = \sum_{i=1}^{n} x^i \cdot h_{ij}(x^1, ..., x^n)$$

Thus,

$$f(x^1, ..., x^n) = \sum_{i,j=1}^{n} x^i x^j \cdot h_{ij}(x^1, ..., x^n)$$

Without loss of generality, $h_{ij} = h_{ji}$, and the matrix $2 \cdot (h_{ij}(0))_{i,j}$ is equal to the Hessian $\frac{\partial^2 f}{\partial x^i \partial x^j}$ and so is non-degenerate. By completing the square in the usual way, this can be diagonalized with all entries on the diagonal ± 1. ■

This lemma tells us that the only information contained at non-degenerate critical points is the index, so we know exactly what the manifold looks like locally at such points. This is why we are able to describe the change in homotopy type as we cross such critical points. In contrast, degenerate critical points have many possible forms.

An immediate corollary is that non-degenerate critical points are isolated (among the set of all critical points). There is a particular class of cases where, though the critical points are degenerate, we can still say much about the local behaviour because the degeneracy is "nice". This is the Morse-Bott case which we will discuss at the end (time permitting).

Examples of degenerate critical points:

- (cubic) $f(x) = x^3$
- $f(x) = e^{-\frac{1}{x^2}} \sin^2(\frac{1}{x})$
- (monkey saddle) $\Re(x + iy)^3$

- $f(x, y) = x^2$
- $f(x, y) = x^2y^2$

3 The change of homotopy type across a critical point

State the main theorems, and give a brief indication of the proofs.

Theorem 1
Let $f : M \to \mathbb{R}$ be smooth. Suppose $a < b$ and that $f^{-1}[a, b]$ is compact and contains no critical points. Then M^a is diffeomorphic to M^b. In fact, M^a is a deformation retract of M^b (hence the inclusion map $M^a \hookrightarrow M^b$ is a homotopy equivalence).

Proof: Sketch only First we take a Riemannian metric $\langle \cdot, \cdot \rangle$ on M, and consider the vector field which is zero on $M^{a-\epsilon}$ and on $f^{-1}[a, b]$ is equal to

$$X(x) = \frac{-1}{\|\nabla f(x)\|^2} \nabla f(x)$$

(∇f is characterized by $\langle \nabla f, Y \rangle = Y(f)$). This vector field is easily constructed with bump-functions. It defines ϕ_t a flow on $f^{-1}[a, b]$ since it has compact support. Flowing for time $b - a$ takes all of M^b to M^a, since $\frac{d}{dt} f(\phi_t(x)) = 1$.
This can be made into a retract by running the flow on a point until the precise moment it enters M^a. ■
See Milnor for an easy example where compactness is needed

Theorem 2
Let $f : M \to \mathbb{R}$ be a smooth function and let p be a non-degenerate critical point of index λ. Let $c = f(p)$ and suppose that $\epsilon > 0$ is such that $f^{-1}[c - \epsilon, c + \epsilon]$ is compact and contains no other critical points. Then $M^{c+\epsilon}$ has the homotopy type of $M^{c-\epsilon}$ with a λ cell attached.

Proof: Sketch: Need diagram... ■
In the case that M is a compact manifold, and f is a function with only non-degenerate critical points, we can use this to construct a cell structure on M. Since "most" functions have only non-degenerate critical points we can always do this. Conversely, knowing the manifold's topology, we can make assertions on the number of critical points such a Morse function must have (since we must have at least as many critical points as we have cells).

As an example, consider the 2-torus as we did initially. At each stage we attach a cell of dimension 1 greater than the structure we had previously, and in each case the boundary map (which is described by the attaching map) is actually the zero map. Thus, H^i is the same as C^i (the free group generated by the i-cells), and so we see $H^0 = \mathbb{Z}$, $H^1 = \mathbb{Z}^2$, $H^2 = \mathbb{Z}$.

4 Some examples and applications

Theorem 3
Reeb *If M is a compact manifold and f is a differentiable function on M with only two critical points, both of which are non-degenerate, then M is homeomorphic to a sphere.*

Proof: The critical points must be the maximum and minimum of f on M. Suppose WOLOG the max is 1 and the min is 0. We see directly from the Morse Lemma that M^ϵ and $f^{-1}[1-\epsilon, 1]$ are both closed n-discs. However, by the first theorem above, we know that M^ϵ is diffeomorphic (hence homeomorphic) to $M^{1-\epsilon}$. Thus we realize M as a union of two closed discs meeting along their common boundary $f^{-1}(1-\epsilon)$. The homeomorphism between the sphere and M is now obvious. ■

Remark: Notice that this asserts only a homeomorphism, not necessarily a diffeomorphism (since there are several differentiable structures on certain spheres). ■

Example: We take the example of $\mathbb{C}P^n$. Consider the following function on $\mathbb{C}P^n$:

$$f[z_0 : ... : z_n] = \frac{\sum_{i=0}^n i \cdot |z_i|^2}{\sum_{i=0}^n |z_i|^2}$$

Note that it is well defined in homogeneous coordinates. $\mathbb{C}P^n$ is covered by the coordinate charts $(U_i, (w_0^i, ..., \hat{w_i^i}, ..., w_n^i))$ where $U_i = \{[z_0 : ... : z_n] : z_i \neq 0\}$, and $w_j^i = \frac{z_j/z_i}{\sqrt{\sum_{j=0}^n |z_j/z_i|^2}}$. In these coordinates, we have

$$f(w_0^i, ..., \hat{w_i^i}, ..., w_n^i) = \sum_{j=0}^n j \cdot |w_j^i|^2 = i + \sum_{j=0}^n (j-i) \cdot |w_j^i|^2$$

where we've used the fact $\sum_{j=0}^n |w_j^i|^2 = 1$. Notice that in U_i there is a unique critical point at the origin, which corresponds to the point $p_i = [0 : ... : 0 : 1 : 0 : ... : 0] \in \mathbb{C}P^n$ (with a 1 in the i^{th} position). Moreover, the U_i are coordinate systems as in the Morse Lemma, and we can read off the index to be $2i$. Since the U_i cover $\mathbb{C}P^n$, these are precisely the critical points, and they are all non-degenerate. So we can use the theorems of Morse to see that $\mathbb{C}P^n$ has the homotopy type of a CW-complex with a cell in each even dimension between 0 and $2n$. We even see explicitly that the $2i$-dimensional cell is just the set $\{[z_0 : ... : z_i : 0 : 0 : ... : 0] : z_i \neq 0\} = A_i = \mathbb{C}P^i \backslash \mathbb{C}P^{i-1} \subset \mathbb{C}P^n$ (for $f|_{A_i}$ takes values between 0 and i). ■

5 Elements of the Morse-Bott Case

Definition 2
*A smooth function $f : M \to \mathbb{R}$ is **clean** if each connected component of its critical point set C_f is a submanifold of M, and, in addition, the Hessian df^2 is non-degenerate in directions normal to C_f. The **index** of a component of C_f is the dimension of the subspace on which df^2 is positive definite (note that this is opposite how we had it before, to stay consistent with Guillemin Sternberg).*

Remark: Thus, in case each component of C_f is a point, the function is Morse. ■

In the Morse case, we saw how to decompose M into a disjoint union of cells (more or less anyways). In Morse-Bott theory, instead of a decomposition into cells, we have a decomposition in sets W^i, each of which is a fibre bundle with fibre a k-cell (where k is the index of the Hessian in the normal direction to C_f) over the components C_i of C_f (note that this agrees with the Morse case if f is Morse, for then each C_i is just a point). Explicitly, choosing a metric and letting ϕ be the flow of the positive gradient flow:

$$W^i = \{x \in M \| \phi(t) \to C_i \quad as \quad t \to \infty\}$$

The following two theorems will be useful for us:

Theorem 4
Suppose f is a Morse-Bott function with all critical manifolds C_i even dimensional and with corresponding index even. Then f has a unique local maximum (which is thus also the maximum).

Proof: Let $C_1, ..., C_k$ be the critical manifolds, each with index equal to 0, and let $C_{k+1}, .., C_l$ be the remaining critical manifolds (f is constant along each C_i, and so the values $a_1, ..., a_k$ of f along $C_1, ..., C_k$ are precisely the local maxima of f). Then $M = \bigcup W_i$ according to the discussion above, and each W_i is open for $i \leq k$, and each W_i for $i > k$ is of codim at least 2. But then the W_i for $i > k$ cannot disconnect M, so $M \setminus \bigcup_{i=k+1}^{l} W_i$ is connected and a union of open disjoint sets W_i, $i \leq k$. So $k = 1$ necessarily. ■

Corollary 1
The function Φ^ξ has a unique local maximum.

Proof: For, one can show that it is Morse-Bott, and that its critical manifolds are even dimensional with even index. ■