

1

Morse Theory in the 1990's

Martin Guest

Introduction

Since the publication of Milnor's book [47] in 1963, Morse theory has been a standard topic in the education of geometers and topologists. This book established such high standards for clarity of exposition and mathematical influence that it has been reprinted several times, and it is still the most popular introductory reference for the subject.

Morse theory is not merely a useful technique. It embodies a far-reaching *idea*, which relates analysis, topology and (most recently) physics. This no doubt is responsible for the resilience of Morse theory over the past several decades: despite the essential simplicity of the idea, it seems to re-emerge every few years to play a crucial role in some major new mathematical development. The title of Bott's article [13] was no doubt inspired by this resilience – here is a subject which appears to be completely 'worked out', yet which time after time has come back to yield something unexpected. Characteristically, a few years after the appearance of [13], Morse theory is again at the forefront of mathematics, as a motivating example of a 'topological field theory'.

This article is based loosely on four lectures given at a Graduate School in Differential Geometry, held at the University of Durham in September 1996. The purpose of the lectures was to give a topical introduction to Morse theory, for postgraduate students in the general area of geometry. In order to save time and yet provide a concrete focus, I omitted most of the standard proofs, and used Morse functions on *Grassmannian manifolds* as a fundamental collection of examples to illustrate the theorems. As well as introducing some basic aspects (such as Schubert varieties) of these important manifolds, this gave the opportunity of discussing a link between Morse theory and Lie theory. A second feature was that I emphasized from the start the fundamental role played by the *gradient flow lines* of a Morse function. With the benefit of hindsight – see §4 – this is a very natural point of view to take. It also fits well with the Grassmannian examples, where the flow lines are known explicitly.

In late 1997 I gave a more leisurely series of lectures for advanced undergraduate students at Tokyo Metropolitan University, and I took this opportunity to expand greatly my original notes. As in the earlier lectures, I emphasized the Grassmannians and the role of the gradient flow lines, but this time I went

‘beyond homology groups’ in order to illustrate the real power of Morse theory, and to give some idea of the developments since [13]. I have also tried to give an account of the ‘toric’ point of view, the full significance of which is only just beginning to be appreciated.

Before getting started, a few historical comments are appropriate. After the pioneering work of Morse, the ‘modern’ period of Morse theory began with Bott’s work in the 1950’s on the homology and homotopy groups of compact symmetric spaces. One of the main achievements of the Morse-theoretic approach was the extension of this work to the loop space of a symmetric space; this led to the discovery of the (Bott) Periodicity Theorem and ultimately to K-theory. Although the role of Morse theory in this area was quickly taken over by the new machinery of algebraic topology, the geometrical nature of Bott’s proof of the Periodicity Theorem still retains great appeal. During the 1960’s, Morse theory was used most prominently to investigate the topology of manifolds, and most prominently of all in the work of Smale, which led to a proof of the Poincaré Conjecture in dimensions greater than four. Following several dormant years in the early 1970’s, Morse theory returned as a guiding force in the development of mathematical Yang-Mills theory, in which the critical points of the Yang-Mills functional are studied. In the 1980’s, having been pushed out of the limelight by the rapidly developing analytic and algebraic geometrical aspects of gauge theory, Morse theory found a dramatic new role. This time the primary motivation was a new approach to Morse theory due to Witten, together with an extension of these ideas by Floer. In this approach the gradient flow lines play the central role, and this laid the groundwork for the ‘field-theoretic’ point of view pioneered by Cohen-Jones-Segal, Betz-Cohen, and Fukaya.

The main emphasis of these notes is the Morse theory of compact – in particular, finite-dimensional – manifolds. I have tried to give at least some references for each aspect of finite-dimensional Morse theory, but not for the infinite-dimensional theory where the subject is much more diverse.

Acknowledgements

I would like to thank John Bolton and Lyndon Woodward for their invitation to the Graduate School at the University of Durham in 1996, and Simon Salamon for encouraging me to contribute the lecture notes to this book. I also wish to acknowledge financial support from the US National Science Foundation, as well as the financial support and extraordinary hospitality of Tokyo Metropolitan University.

My own Morse-theoretical education began at Oxford with Brian Steer in the 1970’s, and I learned a lot from my fellow students Elias Micha, Andrew Pressley, Simon Salamon, Pepe Seade and Socorro Soberon. Later on I benefited greatly from discussions about Morse theory with Haynes Miller and Bill Richter. Most of all, however, I have Brian to thank for suggesting Morse theory as a suitable direction of study. At that time I sometimes worried that Morse theory was ‘too easy’ a subject for serious mathematical research, and only much later did I understand the (often repeated, but usually ignored) advice

that the simplest ideas are the best ones. Even then, with the applications in gauge theory still on the horizon, Brian predicted that in due course there would be tremendous developments in the subject. He was right, and for his guidance and encouragement I dedicate these notes to him, with gratitude.

1 Morse functions

In this section and the next we give a brief explanation of Morse theory, referring mainly to the first 40 pages of [47] for proofs. We begin by defining Morse functions and by mentioning several nontrivial examples. We introduce the flow lines and the stable and unstable manifolds, and give some examples to illustrate these fundamental concepts.

1.1 A basic question

Let us agree that it is important to study manifolds. One way to study a manifold might be to study all possible real-valued functions on it. Presumably, different types of manifolds will possess different types of functions.

Every manifold admits real-valued functions, e.g. constant functions. But it is not immediately obvious how to write down explicit formulae for nontrivial real-valued functions on a given manifold. For example, consider the Grassmannians

$$\begin{aligned} Gr_k(\mathbb{R}^n) &= \{\text{real } k\text{-dimensional linear subspaces of } \mathbb{R}^n\} \\ Gr_k(\mathbb{C}^n) &= \{\text{complex } k\text{-dimensional linear subspaces of } \mathbb{C}^n\}, \end{aligned}$$

which are compact manifolds of (real) dimensions $k(n-k)$, $2k(n-k)$ respectively. To give a function $f: M \rightarrow \mathbb{R}$, where $M = Gr_k(\mathbb{R}^n)$ or $Gr_k(\mathbb{C}^n)$, we must associate to each k -plane a real number. How can we do this in a natural and nontrivial way?

As a much easier example, consider the manifold S^1 consisting of complex numbers $e^{2\pi i\theta}$ of unit length (i.e. the circle). The angle θ defines a function $S^1 \rightarrow \mathbb{R}/\mathbb{Z}$, which is not quite what we want, but we can obtain a real-valued function by using $\cos 2\pi\theta$. It seems reasonable to regard this as the 'simplest' kind of nontrivial real-valued function on S^1 . Note that this function may be interpreted as the first coordinate of the embedding $S^1 \rightarrow \mathbb{R}^2$, $e^{2\pi i\theta} \mapsto (\cos 2\pi\theta, \sin 2\pi\theta)$. This suggests a useful source of functions on a general manifold M : first embed M into a euclidean space, then take a coordinate function. (But in order to find nice functions, one has to find a nice embedding.)

As a slight modification of the previous example, we could take the torus: $T = S^1 \times S^1$. For a fixed choice of $(a, b) \in \mathbb{R}^2$, we have a 'coordinate function' $(e^{2\pi ix}, e^{2\pi iy}) \mapsto a \cos 2\pi x + b \cos 2\pi y$. This is a coordinate function for an embedding of T in \mathbb{R}^4 . We could instead use the familiar embedding of T in \mathbb{R}^3 , and take a coordinate function there (page 1 of [47]). But we shall see later that this function is slightly less satisfactory, for the purposes of present-day Morse theory.

The basic question which Morse theory addresses is: *what is the relation between the properties of a manifold and the properties of its real-valued functions?*

By ‘properties’ we mean global properties, as all manifolds of the same dimension have the same local properties. Thus, Morse theory aims to relate topological properties of M with analytical properties of real-functions on M .

1.2 Morse functions

First we recall a standard definition:

Definition 1.2.1 *Let M be a (smooth)¹ manifold, and let $f: M \rightarrow \mathbb{R}$ be a (smooth) function. A point $m \in M$ is called a critical point of f if $Df_m = 0$.*

The derivative Df_m at m is a linear functional on the tangent space $T_m M$; thus m is critical if and only if this derivative is the zero linear functional.

In terms of a local coordinate chart $\phi: U \rightarrow \mathbb{R}^n$, where U is an open neighbourhood of m in M and $\phi(m) = 0$, f corresponds to the function

$$f \circ \phi^{-1}: U \rightarrow \mathbb{R},$$

and Df_m is represented by the $1 \times n$ matrix

$$\left(\frac{\partial(f \circ \phi^{-1})}{\partial x_1}(0), \dots, \frac{\partial(f \circ \phi^{-1})}{\partial x_n}(0) \right).$$

It will simplify notation if we just write f instead of $f \circ \phi^{-1}$. Using this convention, Taylor’s theorem may be written as

$$f(x) - f(0) = \sum a_i x_i + \frac{1}{2} \sum a_{ij} x_i x_j + \text{remainder}$$

where $a_i = \frac{\partial f}{\partial x_i}|_0$ and $a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}|_0$.

When m is a critical point, i.e. $\sum a_i x_i = 0$, it can be shown that the matrix (a_{ij}) defines a symmetric bilinear form on the tangent space $T_m M$. (This bilinear form is called the Hessian of f .) Hence it is diagonalizable, and the rank and nullity do not depend on the choice of ϕ . This leads to two more definitions:

Definition 1.2.2 *Let m be a critical point of $f: M \rightarrow \mathbb{R}$. The index of m is defined to be the index of (a_{ij}) , i.e. the number of negative eigenvalues of (a_{ij}) .*

Definition 1.2.3 *Let m be a critical point of $f: M \rightarrow \mathbb{R}$. We say that m is a nondegenerate critical point if and only if the nullity of (a_{ij}) , i.e. the dimension of the 0-eigenspace of (a_{ij}) , is zero.*

Since any function on a compact manifold has critical points (e.g. maxima and minima), we cannot get very far by considering functions without critical points. In other words, it is unreasonable to insist that the first term in the Taylor series be a nondegenerate linear functional at every point. The next most favourable condition is that a function has no degenerate critical points, i.e. that (at each critical point) the quadratic term in the Taylor series be nondegenerate:

¹ The word ‘smooth’, meaning ‘infinitely differentiable’, will usually be omitted in future.

Definition 1.2.4 Let M be a (smooth) manifold, and let $f: M \rightarrow \mathbb{R}$ be a (smooth) function. We say that f is a Morse function if and only if every critical point of f is nondegenerate.

To be a Morse function is in some sense a weak condition, as it can be shown that the space of Morse functions is dense in the space of functions. But in another sense² it is a strong condition, as Morse functions have a very special local canonical form:

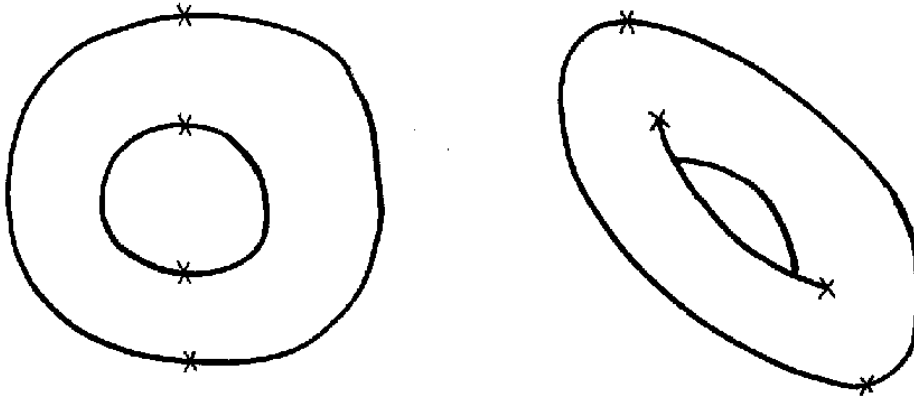
Lemma 1.2.5 (The Morse lemma) Let $f: M \rightarrow \mathbb{R}$ be a Morse function. Then, for any $m \in M$, there exists a local chart ϕ at m such that

$$f(x) - f(0) = - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^n x_i^2.$$

□

Note that the remainder term has disappeared. It follows from this formula that the index of any local maximum point is n , and the index of any local minimum point is 0.

Example 1.2.6 (Height functions on the torus) In the diagram below we have two embeddings of the torus T^2 in \mathbb{R}^3 . By taking the z -coordinate function, we obtain two real-valued functions on T^2 . The critical points are marked with crosses.



It is instructive to write down explicit formulae for these functions on T^2 , and to verify that the critical points are all nondegenerate, with indices 2, 1, 1, 0. □

²This somewhat paradoxical situation is indicative of a 'good' definition, perhaps.

Example 1.2.7 (General theory of height functions) Let M be a compact submanifold of \mathbb{R}^N . For any $v \in \mathbb{R}^N$ we may define the ‘height function’ h^v and the ‘distance function’ L^v on M by

$$h^v(m) = \langle m, v \rangle, \quad L^v(m) = \langle m - v, m - v \rangle$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product. (Note that these functions are essentially the same if M is embedded in the sphere $S^{N-1} \subseteq \mathbb{R}^N$.) It can be shown that, for almost all values of v , the functions h^v and L^v are Morse functions. The critical point theory of these functions is intimately related to the Riemannian geometry of M (with its induced metric). A brief treatment of L^v is given in [47]; the general theory is developed in [53]. \square

1.3 Geometry

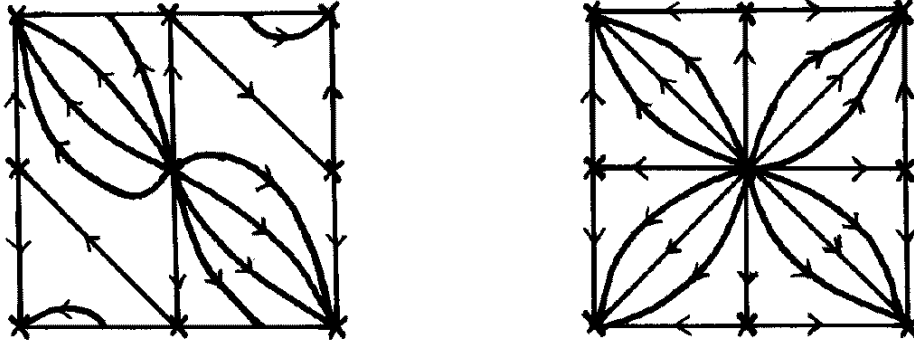
To obtain geometrical information from a Morse function $f: M \rightarrow \mathbb{R}$ it is useful to consider the gradient vector field ∇f of f . To define the gradient vector field we assume that a Riemannian metric $\langle \cdot, \cdot \rangle$ has been chosen on M , and we define $(\nabla f)_m$ by $\langle (\nabla f)_m, X \rangle = (Df)_m(X)$ (for all $X \in T_m M$). For many purposes (although not all!) the particular choice of Riemannian metric is unimportant.

By the theorem of local existence of solutions to first-order ordinary differential equations, there exists an integral curve γ of the vector field $-\nabla f$ through any point of M . (We introduce the minus sign because we want to consider γ as flowing ‘downwards’.) It should be noted that *explicit formulae* for these integral curves are not readily available in general.

If M is compact, then the domain of any such integral curve is \mathbb{R} , and \mathbb{R} acts on M as a group of diffeomorphisms. We shall denote this action by $m \mapsto t \cdot m$, for $m \in M$, $t \in \mathbb{R}$. In other words, $t \cdot m = \gamma(t)$, where γ is the solution of the o.d.e. $\gamma'(t) = -(\nabla f)_{\gamma(t)}$ such that $\gamma(0) = m$. There are only a finite number of critical points in this case (as critical points are isolated, by the Morse lemma). Each integral curve γ ‘begins’ and ‘ends’ at critical points, i.e. $\lim_{t \rightarrow \pm\infty} \gamma(t)$ are critical points. Evidently these integral curves are constrained by the global nature of M , and we shall see that they are a very useful tool for investigating M .

Example 1.3.1 Let $M = S^2$, and embed S^2 in \mathbb{R}^3 as the unit sphere. The z -coordinate function $f: S^2 \rightarrow \mathbb{R}$, $f(x, y, z) = z$ is a Morse function and it has precisely two critical points, the north pole and the south pole. It is the restriction of $F: \mathbb{R}^3 \rightarrow \mathbb{R}$, $F(x, y, z) = z$, and $-\nabla f$ is the component of $-\nabla F = (0, 0, -1)$ which is tangential to the sphere. Hence each integral curve of $-\nabla f$ is a line of longitude running from the north pole to the south pole (with a certain parametrization). \square

Example 1.3.2 (Height functions on the torus, continued) Consider the two height functions on T^2 from Example 1.2.6. The integral curves are illustrated below, with respect to the induced Riemannian metric from \mathbb{R}^3 . In each case the torus is represented by $\mathbb{R}^2/\mathbb{Z}^2$, and a fundamental domain is shown. The maximum point is at the centre of the square; the minimum point is represented by the four corners.



Example 1.3.3 It seems obvious that the least number of isolated critical points of any function on S^2 is 2 (proof?). Example 1.3.1 provides such a function. But what is least number of critical points of any function on the torus T^2 ? In the case of a Morse function, we shall see later that the answer is 4. But if arbitrary (smooth) functions are allowed, the answer is 3. (For an example, see [56], Example 1, page 19.) \square

1.4 Stable and unstable manifolds

Integral curves, or flow lines, may be assembled into the following important objects:

Definition 1.4.1 Let M be a compact manifold, let $f: M \rightarrow \mathbb{R}$ be a Morse function, and let m be a critical point of f . The stable manifold $S(m)$ of m is the set of points which flow ‘down’ to m , i.e.

$$S(m) = \{x \in M \mid \lim_{t \rightarrow \infty} t \cdot x = m\}.$$

The unstable manifold $U(m)$ of m is the set of points which flow ‘up’ to m , i.e.

$$U(m) = \{x \in M \mid \lim_{t \rightarrow -\infty} t \cdot x = m\}.$$

The next result can be proved from the Morse lemma.

Proposition 1.4.2 *Let the index of m be λ . Then $S(m), U(m)$ are homeomorphic (respectively) to $\mathbb{R}^{n-\lambda}, \mathbb{R}^\lambda$. \square*

It follows that a Morse function f on M provides two decompositions of M into disjoint ‘cells’:

$$M = \bigcup_{m \text{ critical}} S(m) = \bigcup_{m \text{ critical}} U(m).$$

We shall refer to these as the stable manifold decomposition and the unstable manifold decomposition associated to f . Observe that the unstable manifold decomposition associated to f is the same as the stable manifold decomposition associated to $-f$.

By intersecting these two decompositions, we obtain a finer one, namely

$$M = \bigcup_{m_1, m_2 \text{ critical}} F(m_1, m_2), \quad F(m_1, m_2) = U(m_1) \cap S(m_2).$$

This collects together the integral curves according to origin and destination: $F(m_1, m_2)$ consists of all integral curves which go from m_1 to m_2 .

Example 1.4.3 Consider the Morse function of Example 1.3.1, on $M = S^2$. The stable manifold decomposition has two pieces, i.e. one copy of each of \mathbb{R}^0 and \mathbb{R}^2 . The unstable manifold decomposition is similar. The intersection of these decompositions has three pieces – two points and one copy of $\mathbb{R}^2 - \{0\}$. \square

Example 1.4.4 (Height functions on the torus, continued) Consider again the two height functions on T^2 from Examples 1.2.6 and 1.3.2. In each case there are four unstable manifolds: a 0-cell, two 1-cells, and a 2-cell. The same is true for the stable manifolds. However, the decompositions $T^2 = \bigcup_{m_1, m_2 \text{ critical}} F(m_1, m_2)$ are quite different. \square

The behaviour of the stable and unstable manifolds is particularly nice if we impose the condition that they intersect transversely:

Definition 1.4.5 *A Morse function $f: M \rightarrow \mathbb{R}$ on a Riemannian manifold M is said to be a Morse-Smale function if $U(m_1)$ is transverse to $S(m_2)$ for all critical points m_1, m_2 of f .*

For the meaning of transversality, we refer to Chapter 3 of [41] or Chapter 3 of [40], or other texts on differential topology. This concept was introduced into Morse theory by Smale – see [60], and the historical discussion in [12].

The transversality condition implies that $U(m_1) \cap S(m_2)$ is a manifold, and that

$$\text{codim } U(m_1) \cap S(m_2) = \text{codim } U(m_1) + \text{codim } S(m_2)$$

whenever $U(m_1) \cap S(m_2)$ is nonempty. Since $F(m_1, m_2) = U(m_1) \cap S(m_2)$, and $\dim U(m_i) = \lambda_i$, we have

$$n - \dim F(m_1, m_2) = (n - \lambda_1) + \lambda_2$$

and hence

$$\dim F(m_1, m_2) = \lambda_1 - \lambda_2.$$

In particular, if there exists a flow line from m_1 to m_2 , then we must have $\lambda_1 > \lambda_2$.

Example 1.4.6 Of the two Morse functions on the torus in Example 1.2.6 (see also Example 1.3.2), only one satisfies the Morse-Smale condition. (Which one? Note that, by the previous paragraph, the existence of a flow line connecting two critical points of the same index is not possible for a Morse-Smale function.) \square

The concepts introduced so far will help us to address the question ‘what configurations of critical points and flow lines are possible?’ for a Morse (or Morse-Smale) function on a given manifold M . This is a more precise version of our original question ‘what kind of smooth functions are possible?’ on M .

Finally, we should mention that the behaviour of a smooth function is much less predictable without the Morse condition, i.e. in the presence of degenerate critical points. For example, critical points are not necessarily isolated, and flow lines do not necessarily converge to critical points. For an example of the latter phenomenon, see page 14 of [54].

2 Topology

Morse theory gives a fundamental relation between topology and analysis. This is traditionally expressed by the ‘Morse inequalities’. We describe various forms of this relation, and its generalizations.

2.1 Topology and analysis

The examples in §1 clearly suggest that there is a relation between the topology of M and the critical point data of a function $f: M \rightarrow \mathbb{R}$. We list four specific examples below, following [12]. In each case, topological information predicts the existence of critical points of f .

(i) It is an elementary fact of topology that, if M is compact, *then f has maximum and minimum points, and these are critical points.*

(ii) Assume that f has a finite number k of critical points (but is not necessarily a Morse function). Then we still have $M = \cup_{m \text{ critical}} S(m)$, and the $S(m)$ are a finite number of contractible sets (but not necessarily cells). It can be shown that this implies the following cohomological condition: if $m > k$, then the product of any m cohomology classes (of positive dimension) on M must be zero. We shall not make any use of this idea, which is part of the theory of Lusternik-Schnirelmann category, so we refer to Lecture 2 of [12] for further

information. However, we note that it gives a stronger version of the prediction of (i): *if M has i cohomology classes (of positive dimension) whose product is nonzero, then any function on M must have at least $i + 1$ critical points.*

For example, consider the torus T^2 . There are two one-dimensional cohomology classes whose product is nonzero, so any function on T^2 must have at least 3 critical points. It is not possible to find three cohomology classes (of positive dimension) whose product is nonzero, so we cannot improve the estimate beyond 3. This is just as well, in view of Example 1.3.3.

(iii) The ‘Minimax Principle’ – see [12].

(iv) Finally we come to our main example, the Morse inequalities. We shall consider this matter in detail later on, but the basic fact is easily stated: if f is a Morse function on a compact manifold M , then the number of critical points of index k is greater than or equal to the k th Betti number b_k of M . This is a considerable generalization of (i), but in fact it is only a hint of the power of Morse theory, as we shall see.

2.2 The main theorem of Morse theory

Recall from §1.4 that a Morse function f gives a decomposition

$$M = \bigcup_{m \text{ critical}} U(m)$$

of a compact manifold M , where $U(m)$ is homeomorphic to \mathbb{R}^{λ_m} , λ_m being the index of m . The main theorem of Morse theory gives information about how these pieces fit together:

Theorem 2.2.1 *Let M be a compact manifold, and let $f: M \rightarrow \mathbb{R}$ be a Morse function on M . Then M has the homotopy type of a cell complex, with one cell of dimension λ for each critical point of index λ .*

To be precise, this theorem means that M is homotopy equivalent to a topological space of the form

$$X_r = ((D^{\lambda_1} \cup_{f_1} D^{\lambda_2}) \cup_{f_2} D^{\lambda_3}) \cup_{f_3} \dots$$

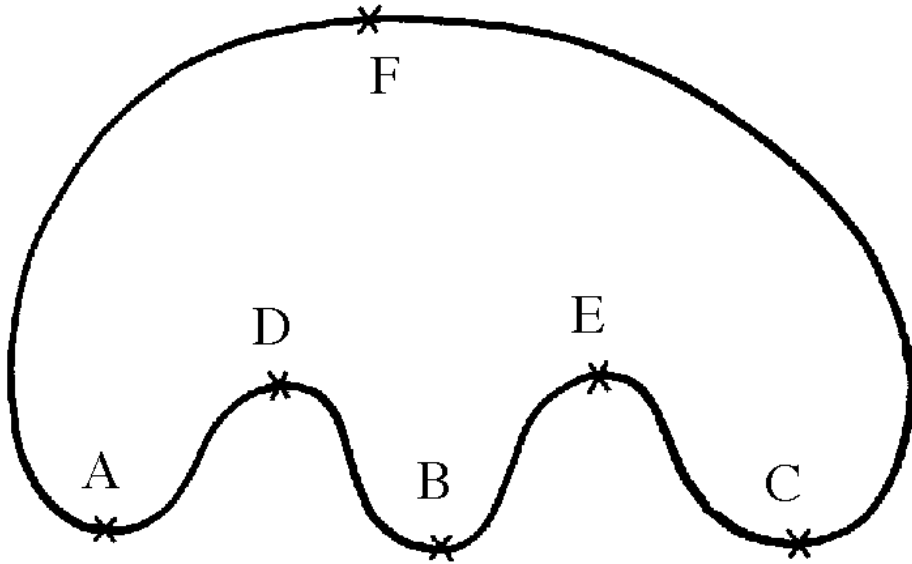
where $0 = \lambda_1, \lambda_2, \lambda_3, \dots, \lambda_r = m$ are the indices of the critical points of f (listed in increasing order, with repetitions where necessary), and $f_1, f_2, f_3, \dots, f_{r-1}$ are certain continuous ‘attaching maps’, with

$$f_i: \partial D^{\lambda_i+1} \rightarrow X_i.$$

To simplify the notation, we shall drop the parentheses in future and simply write $X_r = D^{\lambda_1} \cup_{f_1} D^{\lambda_2} \cup_{f_2} \dots \cup_{f_{r-1}} D^{\lambda_r}$.

In section 2.4 we sketch the main steps in the proof of this theorem. First, however, we give some simple examples.

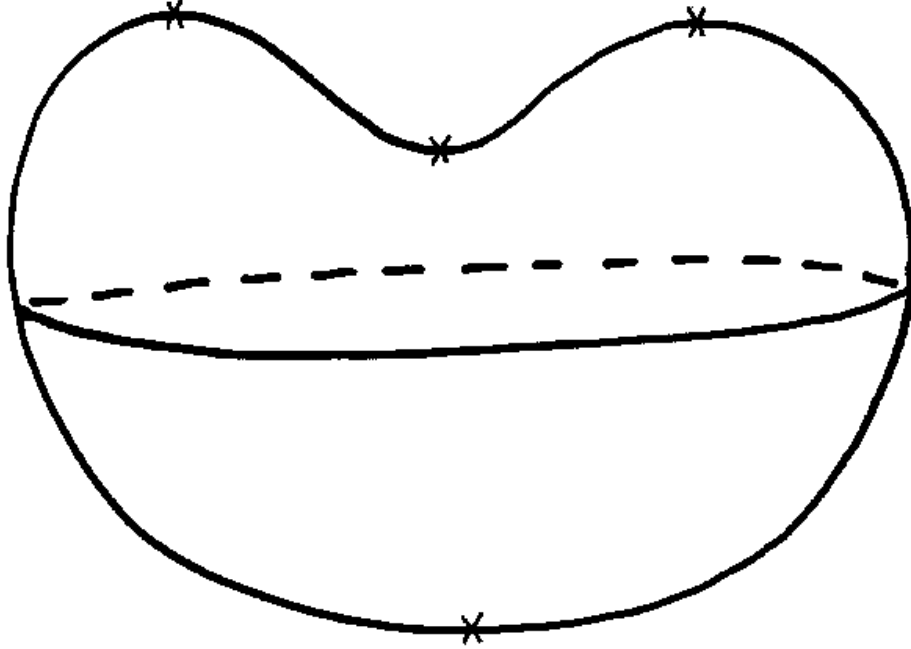
Example 2.2.2 The height function on S^1 defined by the illustrated embedding of S^1 in \mathbb{R}^2 has three local minima A, B, C and three local maxima D, E, F .



We have $S^1 \simeq \{A, B, C\} \cup_f D^1 \cup_g D^1 \cup_h D^1$. Each of the attaching maps f, g, h is an injective map from $\{-1, 1\}$ to $\{A, B, C\}$. Contemplation of this example suggests that one should be able to say more than ‘there is at least one minimum point and at least one maximum point’. It seems plausible, for example, that the number of local maxima must always be *equal* to the number of local minima for this manifold. We shall soon see that this is correct (Corollary 2.3.2). \square

Example 2.2.3 The height function on S^2 defined by the illustrated embedding of S^2 in \mathbb{R}^3 has one local minimum, one critical point of index 1, and two local

maxima.



We have $S^2 \simeq D^0 \cup_f D^1 \cup_g D^2 \cup_h D^2$. Here, $D^0 \cup_f D^1$ is a copy of S^1 , and the maps g, h serve to attach two hemispheres to this circle. \square

Example 2.2.4 (Height functions on the torus, continued) The two height functions on T^2 in Example 1.2.6 give two cell decompositions $T^2 \simeq D^0 \cup_f D^1 \cup_g D^1 \cup_h D^2$. However, the attaching maps behave quite differently in each case, as is clear from Examples 1.3.2 and 1.4.4. \square

Theorem 2.2.1 does not tell us anything about the attaching maps, other than the fact that they exist. This appears to be a serious deficiency. However, we can deduce quite a lot of information on the topology of M – for example, the Morse inequalities, which we discuss next – without knowing anything further. In the above examples, it is intuitively obvious what the attaching maps are.

2.3 The Morse inequalities

Let $f: M \rightarrow \mathbb{R}$ be a Morse function, on a compact manifold M . The Morse inequalities say that $m_i \geq b_i$, where m_i is the number of critical points of index i , and b_i is the i th Betti number of M , i.e. $b_i = \dim H_i(M)$. (We use any homology theory with coefficients in a field.) In the case of T^2 , we have $b_0 = b_2 = 1$ and $b_1 = 2$, so a Morse function on T^2 must have at least four critical points. We have already given examples where this happens.

In terms of the *Morse polynomial*

$$M(t) = \sum_{i=0}^n m_i t^i$$

and the *Poincaré polynomial*

$$P(t) = \sum_{i=0}^n b_i t^i$$

we may express the Morse inequalities symbolically as $M(t) \geq P(t)$. Using this convenient notation, there is a stronger form of the Morse inequalities:

Theorem 2.3.1 *Let $f: M \rightarrow \mathbb{R}$ be a Morse function, on a compact manifold M . Then $M(t) - P(t) = (1+t)Q(t)$, for some polynomial $Q(t)$ such that $Q(t) \geq 0$.*

We shall sketch the proof in the next section. Before that, we give some simple consequences and some examples. To begin with, we put $t = -1$ in the theorem to obtain an expression for the Euler characteristic of M :

Corollary 2.3.2 *Let $f: M \rightarrow \mathbb{R}$ be a Morse function, on a compact manifold M . Then $\sum_{i=0}^n (-1)^i m_i = \sum_{i=0}^n (-1)^i b_i$. \square*

(This is a special case of the Hopf Index Theorem on vector fields, namely for vector fields of the particular form ∇f .)

Corollary 2.3.3 *Let $f: M \rightarrow \mathbb{R}$ be a Morse function, on a compact manifold M . If $M(t)$ contains only even powers of t , then $M(t) = P(t)$.*

Proof Assume that $M(t)$ contains only even powers of t . Then $M(t) \geq P(t) + (1+t)Q(t)$, so neither $P(t)$ nor $(1+t)Q(t)$ can contain odd powers of t . But this is possible only if $Q(t)$ is the zero polynomial. \square

(Various generalizations of this result can be obtained by similar reasoning – the basic principle is that a ‘gap’ in the sequence m_0, m_1, m_2, \dots forces a relation between $M(t)$ and $Q(t)$. This is the ‘lacunary principle’ of Morse.)

Example 2.3.4 For our height functions on the torus T^2 , we have $M(t) = P(t) = 1 + 2t + t^2$. \square

Example 2.3.5 Let $M = \mathbb{C}\mathbb{P}^n$, i.e. n -dimensional complex projective space. This may be defined as $\mathbb{C}^{n+1} - \{0\}/\mathbb{C}^*$, where \mathbb{C}^* acts by multiplication on each coordinate of \mathbb{C}^{n+1} . It may be identified with the space of all complex lines in \mathbb{C}^{n+1} . The equivalence class of (z_0, \dots, z_n) will be denoted by the standard

notation $[z_0; \dots; z_n]$. A Morse function $f: \mathbb{C}\mathbb{P}^n \rightarrow \mathbb{R}$ is given on page 26 of [47]. It may be defined by the formula

$$f([z_0, \dots, z_n]) = \frac{\sum_{i=0}^n c_i |z_i|^2}{\sum_{i=0}^n |z_i|^2}$$

where $c_0 < \dots < c_n$ are fixed real numbers. The critical points of f are the coordinate axes $L_0 = [1; 0; \dots; 0]$, $L_1 = [0; 1; \dots; 0]$, \dots , $L_n = [0; 0; \dots; 1]$. The index of L_i is $2i$. Since all indices are even, Corollary 2.3.4 applies, and we deduce that the Poincaré polynomial of $\mathbb{C}\mathbb{P}^n$ is $1 + t^2 + t^4 + \dots + t^{2n}$. \square

It is already clear from these results and examples that Morse theory works ‘both ways’: (1) the homology groups of a manifold impose conditions on the critical points of any Morse function, and (2) the critical point data of a Morse function sometimes permits the computation of the homology groups.

If $f: M \rightarrow \mathbb{R}$ is a Morse function such that $M(t) = P(t)$, we say that f is *perfect*. The question naturally arises: does every compact manifold possess a perfect Morse function? This matter is discussed in detail in [56]; the answer, essentially, is negative. One reason is illustrated by the next example.

Example 2.3.6 Let $\mathbb{R}\mathbb{P}^n$ be n -dimensional real projective space, with $n \geq 2$. Consider the function $f: \mathbb{R}\mathbb{P}^n \rightarrow \mathbb{R}$ defined by the formula of Example 2.3.5. Is this a Morse function? (Answer: yes.) Is it a perfect Morse function? (Answer: no, if the coefficient field is \mathbb{R} ; yes, if the coefficient field is $\mathbb{Z}/2\mathbb{Z}$.) \square

2.4 Sketch proofs of the main theorems

Theorem 2.2.1 is a consequence of two fundamental results. Following [12] we call these Theorem A and Theorem B. They describe how the structure of the space

$$M^t = \{m \in M \mid f(m) \leq t\}$$

changes when t changes.

Theorem A If $f^{-1}([a, b])$ contains no critical point of f , then M^b is diffeomorphic to M^a .

The integral curves of (a slight modification of) $-\nabla f$ give the required diffeomorphism. See [47], pages 12–13.

Theorem B If $f^{-1}([a, b])$ contains a single critical point of f , then M^b is homotopy equivalent to $M^a \cup_f D^\lambda$, for some $f: \partial D^\lambda \rightarrow M^a$, where λ is the index of the critical point.

The proof is given in [47], pages 14–19 (see also the ‘Proof by picture’ in [12], pages 339–340). By Theorem A, it suffices to consider the situation in a small neighbourhood of the critical point. The example $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = -x^2 + y^2$, in a neighbourhood of the critical point $(0, 0)$, is instructive.

To deduce Theorem 2.2.1 from Theorems A and B, an induction argument is needed (see [47] again, pages 20–24).

There are various ways of proving (and expressing) the Morse inequalities, but all of them are based on the fact that Theorem B allows us to compute the relative homology groups $H_*(M^b, M^a)$ (in the situation of Theorem B, $H_*(M^b, M^a)$ is isomorphic to the coefficient field when $* = \lambda$, and is zero otherwise). A rather formal argument is given in [47], pages 28–31; the proof in [56] is perhaps more transparent. Later expositions of the Morse inequalities and the necessary background material may be found in Chapter 8 of [41] and Chapter 6 of [40], for example.)

A more intuitive version of the proof of Theorem B appears in [12]; the idea is to consider how the Morse and Poincaré polynomials $M^a(t)$, $P^a(t)$ of M^a change when we pass from a to b . Clearly we have

$$M^b(t) = M^a(t) + t^\lambda.$$

For the Poincaré polynomial there are two possibilities, as the λ -cell D^λ either introduces a new homology class in dimension λ or bounds a homology class of M^a in dimension $\lambda - 1$. (More precisely, the connecting homomorphism $H_\lambda(M^b, M^a) \rightarrow H_{\lambda-1}(M^a)$ in the long exact homology sequence is either zero or injective.) Thus we have

$$P^b(t) = P^a(t) + (t^\lambda \text{ or } -t^{\lambda-1}).$$

We are interested in the difference $M^b(t) - P^b(t)$, and for this we have

$$M^b(t) - P^b(t) = M^a(t) - P^a(t) + (0 \text{ or } t^\lambda + t^{\lambda-1}).$$

By induction, this gives the Morse inequalities in the form $M^b(t) \geq P^b(t)$. It also gives immediately the stronger result $M^b(t) - P^b(t) = (1+t)Q(t)$ with $Q(t) \geq 0$.

2.5 Generalization: Morse-Bott functions

Since Morse functions necessarily have isolated critical points, Morse theory immediately disqualifies many ‘natural’ functions. Constant functions provide trivial examples of this phenomenon, but there are many nontrivial ones, such as functions which are equivariant with respect to the action of a Lie group – here the orbit of a critical point consists entirely of critical points. We begin with a simple example:

Example 2.5.1 Consider the sphere $S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$. Define $f: S^2 \rightarrow \mathbb{R}$ by $f(x, y, z) = -z^2$. It is easy to verify that the critical points are

(1) the north pole $(0, 0, 1)$, (2) every point of the equator $z = 0$, and (3) the south pole $(0, 0, -1)$. It is also easy to verify that (1) and (3) are nondegenerate critical points of index zero. To understand what is happening at the equator, let us choose local coordinates

$$(\sqrt{1 - u^2 - v^2}, u, v) \mapsto (u, v)$$

around the point $(1, 0, 0)$. Then, with the notational conventions of §1, we have

$$f(u, v) = -v^2, \quad f(0, 0) = 0.$$

Comparing this with the form of the Morse lemma, we see that degeneracy is indicated here by the absence of $\pm u^2$. Now, the u -direction is precisely the direction of the equator, and along the equator f is constant, so degeneracy ‘in the u -direction’ is inescapable. (If $f: M \rightarrow \mathbb{R}$ is a smooth function and N is a connected submanifold of M consisting of critical points of f , then f is constant on N .) In the v -direction, however, we are in the situation of the Morse lemma. The integral curves of $-\nabla f$ (with respect to the standard induced metric on the sphere) are easy to imagine in this situation, and we appear to have the generalized cell decomposition

$$S^2 \simeq (\text{north pole} \cup \text{south pole}) \cup_g (\text{equator} \times [-\frac{1}{2}, \frac{1}{2}])$$

where

$$g: \text{equator} \times \{-\frac{1}{2}, \frac{1}{2}\} \rightarrow \text{north pole} \cup \text{south pole}$$

is the map which sends ‘equator $\times \{-\frac{1}{2}\}$ ’ to the south pole and ‘equator $\times \{\frac{1}{2}\}$ ’ to the north pole. (If we use $-f$ instead of f , the decomposition would involve attaching the northern and southern hemispheres to the equator.) \square

The generalization of Morse theory to such examples was developed by Bott in [10]. This depends on the following definition:

Definition 2.5.2 *Let V be a connected submanifold of M , such that every point of V is a critical point of $f: M \rightarrow \mathbb{R}$. We say that V is a nondegenerate critical manifold of f (or NDCM, for short) if, for every $v \in V$, $T_v V$ is equal to the null-space N_v of the bilinear form $(\partial^2 f / \partial x_i \partial x_j)$ on $T_v M$.*

With the hypotheses of the definition, it is obvious that $T_v V$ is contained in N_v ; for an NDCM they are required to be equal. Since V is assumed to be connected, each critical point $v \in V$ has the same index, and we refer to this number as the index of V .

Definition 2.5.3 *A function $f: M \rightarrow \mathbb{R}$ is a Morse-Bott function if every critical point of M belongs to a nondegenerate critical manifold.*

Any Morse function is a Morse-Bott function, of course. A Morse-Bott function is a Morse function if and only if each NDCM is a single point.

Bott's generalization of the main theorem of Morse theory is based on the fact that each stable and unstable 'cell' has the structure of a vector bundle over an NDCM (see [10]). In the above example, the unstable manifold of the equator is in fact a trivial vector bundle (of rank 1), but in general the bundle could be nontrivial. This vector bundle is usually called the *negative bundle* of the NDCM.

Theorem 2.5.4 *Let M be a compact manifold, and let $f: M \rightarrow \mathbb{R}$ be a Morse-Bott function on M . Then M has the homotopy type of a 'cell-bundle complex'. Each nondegenerate critical manifold V of index λ contributes a cell-bundle $D^\lambda(V)$ of rank λ , i.e. a fibre bundle over V with fibre D^λ .* \square

This means

$$M \simeq D^{\lambda_1}(V_1) \cup_{f_1} D^{\lambda_2}(V_2) \cup_{f_2} D^{\lambda_3}(V_3) \cup_{f_3} \dots \cup_{f_{r-1}} D^{\lambda_r}(V_r)$$

where $0 = \lambda_1, \lambda_2, \dots, \lambda_r \leq m$ are the indices of the nondegenerate critical manifolds V_1, V_2, \dots, V_r of f , and $f_1, f_2, f_3, \dots, f_{r-1}$ are certain maps, with

$$f_i: \partial D^{\lambda_{i+1}}(V_{i+1}) \rightarrow D^{\lambda_1}(V_1) \cup_{f_1} D^{\lambda_2}(V_2) \cup_{f_2} \dots \cup_{f_{i-1}} D^{\lambda_i}(V_i).$$

(The boundary $\partial D^\lambda(V)$ of a cell-bundle $D^\lambda(V)$ is a sphere-bundle, i.e. a fibre bundle over V with fibre $S^{\lambda-1}$.)

What would be an appropriate generalization of the Morse inequalities? Well, the natural generalization of Theorem B of section 2.4 leads one to consider the relative homology groups $H_*(D^\lambda(V), \partial D^\lambda(V))$. By the Thom Isomorphism Theorem we have

$$H_*(D^\lambda(V), \partial D^\lambda(V)) \cong H_{*-\lambda}(V; \theta)$$

where θ is the 'orientation sheaf' of the negative bundle. There are two commonly occurring situations where this orientation sheaf is constant: (1) if we use homology with coefficients in the field $\mathbb{Z}/2\mathbb{Z}$, then $\theta = \mathbb{Z}/2\mathbb{Z}$; (2) if M and V are complex manifolds and the negative bundle is a complex vector bundle, then $\theta = F$ is constant for any coefficient field F (or even for integer coefficients). If the negative bundle is trivial (as in Example 2.5.1), then θ is certainly constant.

For notational simplicity we shall assume from now on that we are in the favourable situation where θ is constant, so that $H_*(D^\lambda(V), \partial D^\lambda(V)) \cong H_{*-\lambda}(V)$. Following the argument of section 2.4, we see that the contribution of V to the Poincaré polynomial is either $t^\lambda P^V(t)$ or $-t^{\lambda-1} P^V(t)$, where $P^V(t)$ is the Poincaré polynomial of V . Let us define the *Morse-Bott polynomial* of f to be

$$MB(t) = \sum_{i=1}^r P^{V_i}(t) t^{\lambda_i}.$$

Then the argument of section 2.4 gives the following analogue of Theorem 2.3.1, which we refer to as the Morse-Bott inequalities:

Theorem 2.5.5 *Let $f: M \rightarrow \mathbb{R}$ be a Morse-Bott function, on a compact manifold M . Then $MB(t) - P(t) = (1+t)Q(t)$, where $Q(t) \geq 0$. \square*

In particular we have $MB(t) \geq P(t)$. The comments in section 2.3 on ‘gaps’ apply equally to this situation: if $MB(t)$ contains no odd power of t , then $MB(t) = P(t)$, i.e. f is a perfect Morse-Bott function.

Constant functions are now (satisfyingly) incorporated into the theory, because a constant function on M has one NDCM, namely M itself, of index zero, and we have $MB(t) = P(t)$. Somewhat more interesting are functions of the form $\pi \circ f$, where $\pi: E \rightarrow M$ is a fibre bundle and $f: M \rightarrow \mathbb{R}$ is a Morse-Bott function. Any function of this form is a Morse-Bott function; the NDCM’s are of the form $\pi^{-1}(V)$, where V is an NDCM of f , and the index of $\pi^{-1}(V)$ is equal to the index of V .

Bott explains the following intuition behind Theorem 2.5.5: if a Morse-Bott function is deformed slightly to give a Morse function, each NDCM V of index λ breaks up into a finite set of isolated critical points, and for each Betti number b_i of V there are b_i critical points of index $i + \lambda$. In the case of Example 2.5.1, the equator contributes $tP^{S^1}(t) = t(1+t)$, which is equivalent to having two isolated critical points of indices 1, 2. The way in which the equator can break up into two isolated critical points is illustrated in the diagrams below:



Let us look at some further concrete examples.

Example 2.5.6 In Example 2.3.5 we examined the Morse function

$$f([z_0, \dots, z_n]) = \frac{\sum_{i=0}^n c_i |z_i|^2}{\sum_{i=0}^n |z_i|^2}$$

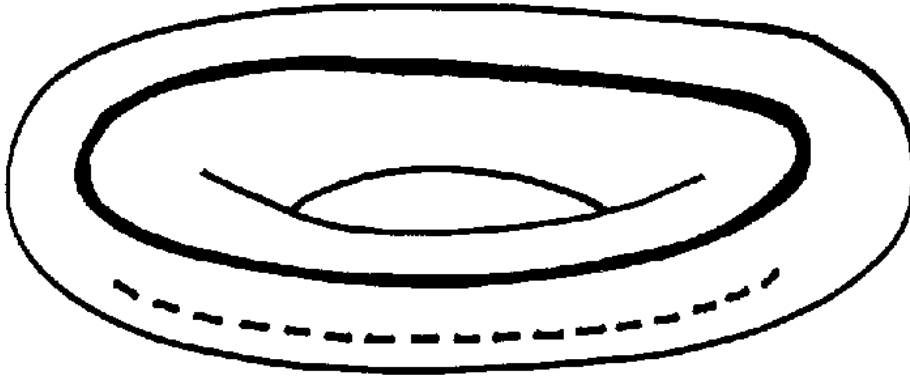
on $\mathbb{C}\mathbb{P}^n$, where $c_0 < \dots < c_n$. If we allow *arbitrary* real numbers $c_0 \leq \dots \leq c_n$ then the same formula defines a Morse-Bott function. Regarding c_0, \dots, c_n as the eigenvalues of a diagonal matrix $C = \text{diag}(c_0, \dots, c_n)$, we can describe the NDCM’s as the submanifolds $\mathbb{P}(V_1), \dots, \mathbb{P}(V_k)$ of $\mathbb{C}\mathbb{P}^n$, where V_1, \dots, V_k are the eigenspaces of C . For example, in the extreme case where $c_0 = 0, c_1 = \dots =$

$c_n = 1$, we have two NDCM's, namely

$$\begin{aligned} \{[1; 0; \dots; 0]\} &= \mathbb{C}\mathbb{P}^0 \text{ (an isolated critical point)} \\ \{[0; *; \dots; *] \mid * \in \mathbb{C}\} &\cong \mathbb{C}\mathbb{P}^{n-1}. \end{aligned}$$

Clearly the isolated point is an absolute minimum point, so it has index zero. The other NDCM consists of absolute maxima, so it has index 2 (i.e. $\dim \mathbb{C}\mathbb{P}^n - \dim \mathbb{C}\mathbb{P}^{n-1}$). Hence the Morse-Bott polynomial is $MB(t) = 1 + t^2(1 + t^2 + \dots + t^{2(n-1)})$. This is equal to the Poincaré polynomial of $\mathbb{C}\mathbb{P}^n$, so we have a perfect Morse-Bott function. (It is easy to see that, for any choice of $c_0 \leq \dots \leq c_n$, we obtain a perfect Morse-Bott function, in fact.) \square

Example 2.5.7 Let us consider a new height function on the torus T^2 , defined by the following embedding of T^2 in \mathbb{R}^3 :



There are two NDCM's, each being a copy of S^1 , with indices 0, 1. If we tilt this embedding slightly, we obtain one of our old Morse functions on T^2 (which one?). \square

2.6 Generalization: noncompactness and singularities

For a Morse function on a noncompact manifold M , there are two immediate difficulties. First, flow lines do not necessarily exist for all time (consider the manifold obtained by removing a non-critical point from a compact manifold). Second, even the flow lines which exist for all time do not necessarily converge to critical points (consider the result of removing a critical point!).

If the function $f: M \rightarrow \mathbb{R}$ is bounded below and proper, then the sets $f^{-1}(-\infty, a]$ are compact, and there is essentially no difficulty in doing Morse theory. For example, it is possible to construct Morse functions on $\mathbb{C}\mathbb{P}^\infty = \cup_{n \geq 0} \mathbb{C}\mathbb{P}^n$ by extending Example 2.3.5 in an obvious way. A more interesting example, perhaps, is the example which was fundamental for Morse himself, namely the space

of paths on a Riemannian manifold (see [47]). We shall say a little more about infinite-dimensional examples such as these in §4.

Morse theory can be extended in another direction, to functions on singular spaces. Different approaches can be found in [35], [42], and section 3.2 of [45].

2.7 Generalization: equivariant Morse theory

Let G be a compact Lie group acting (smoothly) on a compact manifold M . Assume that $f: M \rightarrow \mathbb{R}$ is G -equivariant, i.e. $f(gm) = f(m)$ for all $g \in G$, $m \in M$. Then, as we have already pointed out, f is unlikely to be a Morse function as the G -orbit of any critical point consists entirely of critical points. Morse-Bott theory is designed to cope with this kind of situation, but there is a further extension of Morse-Bott theory which applies to equivariant functions.

The basic idea (see [4] and [12]) is that one would like to relate the Morse theory of an equivariant Morse-Bott function $f: M \rightarrow \mathbb{R}$ to the ‘Morse theory’ of the induced map $f: M/G \rightarrow \mathbb{R}$. If G acts freely on M , then M/G is a compact manifold and $M \rightarrow M/G$ is a locally trivial bundle, so such a relation exists by the remarks following Theorem 2.5.4. But if the action of G is not free, one must find another way to ‘take account of the symmetry due to G ’. The method introduced by Atiyah and Bott is to replace M/G by the homotopy quotient

$$M//G = M \times_G EG$$

where $EG \rightarrow BG$ is a universal bundle for G . The (contractible) space EG is not a finite-dimensional manifold (if G is nontrivial), but it can be constructed as a limit of compact manifolds $(EG)_n$, and f extends naturally to a function on each compact manifold $M \times_G (EG)_n$. The following standard example ([4]) illustrates this construction.

Example 2.7.1 Let f be the height function on $M = S^2$ as in Example 1.3.1, and let $G = S^1$ act on S^2 by ‘rotation about the vertical axis’. Thus, the two critical points are isolated orbits, and all other orbits are circles. Clearly ordinary Morse theory does not work on the quotient space $S^2/S^1 \cong [-1, 1]$, so we consider instead

$$S^2//S^1 = S^1 \times_{S^1} S^\infty = \lim_{n \rightarrow \infty} S^2 \times_{S^1} S^{2n+1}.$$

Now, $S^2 \times_{S^1} S^{2n+1}$ is a manifold on which (the extension of) f has two non-degenerate critical manifolds, each homeomorphic to $\{\text{point}\} \times_{S^1} S^{2n+1} \cong \mathbb{C}\mathbb{P}^n$. The indices of these NDCM’s are 0 and 2, and so we have a perfect Morse-Bott function with Morse-Bott polynomial

$$t^0(1 + t^2 + \dots + t^{2n}) + t^2(1 + t^2 + \dots + t^{2n}).$$

Taking the limit $n \rightarrow \infty$, we obtain the Poincaré polynomial of $S^2//S^1$ as $(1 + t^2) \sum_{i \geq 0} t^{2i}$ (conveniently abbreviated as $(1 + t^2)/(1 - t^2)$). This agrees

with the well known fact that $S^2//S^1$ is homotopy equivalent to $\mathbb{C}\mathbb{P}^\infty \vee \mathbb{C}\mathbb{P}^\infty$. The Poincaré polynomial of $S^2 \times_{S^1} S^{2n+1}$ could also be obtained by using the fact that $S^2 \times_{S^1} S^{2n+1}$ is a bundle over $\mathbb{C}\mathbb{P}^n$ with fibre S^2 ; if we choose any perfect Morse function on $\mathbb{C}\mathbb{P}^n$ then we obtain a perfect Morse-Bott function on $S^2 \times_{S^1} S^{2n+1}$ by the remarks following Theorem 2.5.4. \square

The next instructive example is also taken from [4].

Example 2.7.2 Let $f: S^2 \rightarrow \mathbb{R}$ be the function $-z^2$ of Example 2.5.1. This Morse-Bott function is not perfect, and neither is the extended Morse-Bott function on $S^2 \times_{S^1} S^{2n+1}$. The NDCM's of the latter are

$$\begin{aligned} \{\text{point}\} \times_{S^1} S^{2n+1} &\cong \mathbb{C}\mathbb{P}^n \quad (\text{index } 2) \\ \{\text{equator}\} \times_{S^1} S^{2n+1} &\cong S^{2n+1} \quad (\text{index } 0) \\ \{\text{point}\} \times_{S^1} S^{2n+1} &\cong \mathbb{C}\mathbb{P}^n \quad (\text{index } 2) \end{aligned}$$

so the Morse-Bott polynomial is

$$t^0(1 + t^{2n+1}) + 2t^2(1 + t^2 + \dots + t^{2n}),$$

which has a superfluous term t^{2n+1} . But as $n \rightarrow \infty$, the effect of this term disappears, and we obtain a perfect 'Morse-Bott function' on $S^2//S^1$. \square

In general, a G -equivariant Morse-Bott function $f: M \rightarrow \mathbb{R}$ is converted by the above procedure into a 'Morse-Bott function' on $M//G$; each NDCM N in M of index λ is converted into an NDCM $N//G$ of index λ . Atiyah and Bott give a criterion for the new Morse-Bott function to be perfect (even when f is not), the 'Self-Completion Principle'.

It is important to understand what this extension of Morse theory does *not* do: it does not give additional information about the homology groups of M , but rather about those of the (much larger) auxiliary space $M//G$.

Sometimes it is useful to associate a *smaller* auxiliary space to the original Morse-theoretic data, especially when dealing with infinite-dimensional manifolds. One example of this is provided by the space of 'broken geodesics' in Morse's original application to the space $M = \Omega X$ of closed smooth paths on a manifold X , or (analogously) the space of 'algebraic loops' when X is a Lie group (see [57]). Another famous example is due to Floer ([24]).

In all of these situations, one may regard the homology groups of the auxiliary construction as generalized homology groups of the original space M ; from equivariant Morse theory one obtains equivariant homology groups, and from Floer's theory Floer homology groups.

In the case of equivariant homology, it is, roughly speaking, the *coefficient ring* that is being generalized, from $H^*(\text{point})$ to $H^*(BG)$. The equivariant (co)homology groups $H_G^*(M)$ are defined by $H_G^*(M) = H^*(M//G)$. This is a module over the ring $H_G^*(\text{point}) = H^*(BG)$, but it is not in general isomorphic to

the free module $H^*(M) \otimes H^*(BG)$; indeed, it is the $H^*(BG)$ -module structure (and the multiplicative structure) of $H_G^*(M)$ which reflects the nature of the action of G on M .

A second lesson from equivariant Morse theory is that one can sometimes ‘do Morse theory without a Morse function’: the space $M//G$ is not a manifold in general, yet we have constructed a Morse-Bott polynomial and the Morse-Bott inequalities are satisfied.

3 Morse functions on Grassmannians

We examine in detail a special but very important example, namely the Morse theory of a family of functions on the complex Grassmannian manifold $Gr_k(\mathbb{C}^n)$. Our main tool is an explicit formula for the integral curves. This will give information on the homology and cohomology of $Gr_k(\mathbb{C}^n)$, and on the behaviour of its Schubert subvarieties.

3.1 Morse functions on Grassmannians

The (complex) Grassmannian

$$Gr_k(\mathbb{C}^n) = \{ \text{complex } k\text{-dimensional linear subspaces of } \mathbb{C}^n \}$$

provides a good ‘test case’ for the theory of the previous two sections. In addition, $Gr_k(\mathbb{C}^n)$ is an important manifold which appears in many parts of mathematics. We shall study it in detail in this section.

To produce a ‘nice’ real-valued function on $Gr_k(\mathbb{C}^n)$, we use the embedding

$$\begin{aligned} Gr_k(\mathbb{C}^n) &\subseteq \text{SkewHerm}_n = \{ n \times n \text{ complex matrices } A \mid A^* = -A \} \\ V &\mapsto \sqrt{-1} \pi_V \end{aligned}$$

where $\pi_V: \mathbb{C}^n \rightarrow \mathbb{C}^n$ denotes ‘orthogonal projection on V ’ with respect to the standard Hermitian inner product on \mathbb{C}^n . The (real) vector space SkewHerm_n has an inner product $\langle\langle \cdot, \cdot \rangle\rangle$, defined by $\langle\langle A, B \rangle\rangle = \text{trace } AB^* = -\text{trace } AB$. We can obtain real-valued functions on $Gr_k(\mathbb{C}^n)$ by taking height functions with respect to this embedding. Specifically, we shall choose real numbers a_1, \dots, a_n and consider the function

$$f: Gr_k(\mathbb{C}^n) \rightarrow \mathbb{R}, \quad V \mapsto \langle\langle \sqrt{-1} \pi_V, \sqrt{-1} D \rangle\rangle = \text{trace } \pi_V D$$

where $D = \text{diag}(a_1, \dots, a_n)$ denotes the diagonal matrix whose diagonal entries are a_1, \dots, a_n .

In the special case $k = 1$ (so that $Gr_k(\mathbb{C}^n) = \mathbb{C}P^{n-1}$), we may write

$$V = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = [z] \in \mathbb{C}P^{n-1}$$

and we have

$$\pi_V = \frac{zz^*}{|z|^2}.$$

So our real-valued function $f: \mathbb{C}\mathbb{P}^{n-1} \rightarrow \mathbb{R}$ is given by

$$f([z]) = \frac{\sum_{i=0}^n a_i |z_i|^2}{\sum_{i=0}^n |z_i|^2}.$$

This is our old friend from Example 2.3.5.

Let $V_i = \mathbb{C}e_i$, i.e. the line spanned by the i th basis vector of \mathbb{C}^n . We shall prove:

Theorem 3.1.1 *Assume that $a_1 > \dots > a_n \geq 0$. Then f is a perfect Morse function. An element V of $Gr_k(\mathbb{C}^n)$ is a critical point of f if and only if $V = V_{i_1} \oplus \dots \oplus V_{i_k}$ for some i_1, \dots, i_k with $1 \leq i_1 < \dots < i_k \leq n$.*

One way (and indeed the conventional way) to prove this theorem would be by direct calculation, using a local chart. We shall take a slightly different point of view here, by concentrating on the integral curves of $-\nabla f$. Remarkably, there is an *explicit formula* for these integral curves, which greatly simplifies the Morse-theoretic analysis of f :

Lemma 3.1.2 *The integral curve γ of $-\nabla f$ through a point $V \in Gr_k(\mathbb{C}^n)$ is given by $\gamma(t) = \text{diag}(e^{-a_1 t}, \dots, e^{-a_n t})V = e^{-tD}V$.*

Remark It follows immediately from this lemma that the critical points of f are the special k -planes $V_{\mathbf{u}} = V_{u_1} \oplus \dots \oplus V_{u_k}$, as stated in the above theorem.

Proof It is easy to verify that the tangent space of $Gr_k(\mathbb{C}^n)$ is given as a subspace of SkewHerm_n by

$$T_V Gr_k(\mathbb{C}^n) = \{T - T^* \mid T \in \text{Hom}(V, V^\perp)\},$$

and that the orthogonal projection of any $X \in \text{SkewHerm}_n$ on this tangent space is $\pi_V X \pi_V^\perp + \pi_V^\perp X \pi_V$. Since f is the restriction of the linear function $X \mapsto \langle \langle X, \sqrt{-1} D \rangle \rangle$ on SkewHerm_n , the gradient ∇f_V is the projection of $\sqrt{-1} D$ on $T_V Gr_k(\mathbb{C}^n)$, so we have

$$-\nabla f_V = -\sqrt{-1} (\pi_V D \pi_V^\perp + \pi_V^\perp D \pi_V).$$

(This formula shows that V is a critical point of f if and only if $DV \subseteq V$, $DV^\perp \subseteq V^\perp$, i.e. if and only if V is of the form $V_{i_1} \oplus \dots \oplus V_{i_k}$.)

To find $\dot{\gamma}(t)$, we first write $e^{-tD} = U_t P_t$, where U_t is unitary and P_t is an invertible complex $n \times n$ matrix such that $P_t V = V$ (such a factorization may be accomplished by the Gram-Schmidt orthogonalization process). Using this we have

$$\begin{aligned} \dot{\gamma}(t) &= \frac{d}{dt} \sqrt{-1} \pi_{e^{-tD}V} = \frac{d}{dt} \sqrt{-1} \pi_{U_t V} = \frac{d}{dt} \sqrt{-1} U_t \pi_V U_t^{-1} \\ &= \sqrt{-1} \{ \dot{U}_t \pi_V U_t^{-1} - U_t \pi_V U_t^{-1} \dot{U}_t U_t^{-1} \} \\ &= \sqrt{-1} U_t \{ U_t^{-1} \dot{U}_t \pi_V - \pi_V U_t^{-1} \dot{U}_t \} U_t^{-1} \\ &= \sqrt{-1} U_t \{ \pi_V^\perp U_t^{-1} \dot{U}_t \pi_V - \pi_V U_t^{-1} \dot{U}_t \pi_V^\perp \} U_t^{-1}. \end{aligned}$$

As u varies from 1 to n , such points account for all of $\mathbb{C}\mathbb{P}^{n-1}$ – so we deduce that

$$S_u \cong \left\{ \left(\begin{array}{c} * \\ \vdots \\ * \\ 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) \text{ with } * \in \mathbb{C} \right\} \cong \mathbb{C}^{u-1}.$$

Similarly we have

$$U_u \cong \left\{ \left(\begin{array}{c} 0 \\ \vdots \\ 0 \\ 1 \\ * \\ \vdots \\ * \end{array} \right) \text{ with } * \in \mathbb{C} \right\} \cong \mathbb{C}^{n-u}.$$

It is clear from this description that S_u and U_u meet transversally at V_u , and that f must be a Morse function. Finally, since the stable and unstable manifolds are even-dimensional, f is perfect.

We use similar notation in the case of $Gr_k(\mathbb{C}^n)$ for general k . First, we represent the critical point $V_{\mathbf{u}} = V_{u_1} \oplus \dots \oplus V_{u_k}$ by an $n \times k$ matrix:

$$\begin{bmatrix} \vdots & & \vdots \\ V_{u_1} & \dots & V_{u_k} \\ \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}.$$

Then we observe that

$$\lim_{t \rightarrow \infty} \begin{pmatrix} e^{-a_1 t} & & \\ & \ddots & \\ & & e^{-a_n t} \end{pmatrix} \begin{bmatrix} * & * \\ \vdots & \vdots \\ * & \vdots \\ 1 & \vdots \\ & \ddots & * \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}.$$

But any $n \times k$ matrix may be brought into the ‘echelon form’

$$\begin{bmatrix} * & & * \\ \vdots & & \vdots \\ * & & \vdots \\ 1 & & \vdots \\ & \ddots & * \\ & & 1 \end{bmatrix}$$

by ‘column operations’, so we have found the stable manifold $S_{\mathbf{u}}$ of $V_{\mathbf{u}}$.

Now, an element of $S_{\mathbf{u}}$ may be represented by more than one matrix in echelon form. To obtain a unique representation, we should bring the matrix into ‘reduced echelon form’, so that it looks (for example) like this:

$$\begin{pmatrix} * & * & * \\ * & * & * \\ 1 & 0 & 0 \\ & 1 & 0 \\ & & * \\ & & 1 \\ & & 0 \\ & & 0 \end{pmatrix}.$$

It follows that

$$\begin{aligned} \dim_{\mathbb{C}} S_{\mathbf{u}} &= (u_1 - 1) + (u_2 - 2) + (u_3 - 3) + \dots + (u_k - k) \\ &= \sum_{i=1}^k u_i - \frac{1}{2}k(k+1) \end{aligned}$$

(and $S_{\mathbf{u}}$ may be identified with a complex vector space – or cell – of this dimension).

There is a similar description of the unstable manifold $U_{\mathbf{u}}$. We have

$$\lim_{t \rightarrow -\infty} \begin{pmatrix} e^{-a_1 t} & & \\ & \ddots & \\ & & e^{-a_n t} \end{pmatrix} \begin{bmatrix} 1 & & \\ * & \ddots & \\ \vdots & & 1 \\ \vdots & & * \\ \vdots & & \vdots \\ * & & * \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix},$$

The action of $A = \text{diag}(e^{-a_1 t}, \dots, e^{-a_n t})$ (which gives the integral curves of $-\nabla f$) is easy to express in term of the Lie algebra $\mathfrak{n}_{\mathbf{u}}$. For $X \in \mathfrak{n}_{\mathbf{u}}$, we have

$$\begin{aligned} A(\exp X)V_{\mathbf{u}} &= A(\exp X)A^{-1}AV_{\mathbf{u}} \\ &= A(\exp X)A^{-1}V_{\mathbf{u}} \text{ (as } V_{\mathbf{u}} \text{ is fixed by } A) \\ &= (\exp AXA^{-1})V_{\mathbf{u}}. \end{aligned}$$

The map $X \mapsto AXA^{-1}$ has the effect of multiplying the (i, j) th entry of X by $e^{-(a_i - a_j)t}$.

3.2 Morse-Bott functions on Grassmannians

By relaxing the condition $a_1 > \dots > a_n \geq 0$ to $a_1 \geq \dots \geq a_n \geq 0$, we obtain a Morse-Bott function on $Gr_k(\mathbb{C}^n)$. To investigate this, we introduce the notation

$$\mathbb{C}^n = E_1 \oplus \dots \oplus E_l$$

for the eigenspace decomposition of $D = \text{diag}(a_1, \dots, a_n)$. We write b_i for the eigenvalue on E_i . Thus, the *distinct* a_i 's are $b_1 > \dots > b_l$.

Theorem 3.2.1 *Assume that $a_1 \geq \dots \geq a_n \geq 0$. Then f is a perfect Morse-Bott function. An element V of $Gr_k(\mathbb{C}^n)$ is a critical point of f if and only if $V = V \cap E_1 \oplus \dots \oplus V \cap E_k$, i.e. if and only if V is spanned by eigenvectors of $D = \text{diag}(a_1, \dots, a_n)$. The NDCM containing such a V is diffeomorphic to $Gr_{c_1}(E_1) \times \dots \times Gr_{c_l}(E_l)$ where $c_i = \dim_{\mathbb{C}} V \cap E_i$.*

It may seem unnecessary to introduce Morse-Bott functions in a situation like this, where we already have a good supply of Morse functions. However, we shall see later that the special properties of Morse-Bott functions can be extremely useful.

To prove the theorem, and to identify the stable and unstable manifolds, we use the explicit formula for the integral curves of $-\nabla f$ which was given earlier. (The derivation of that formula is clearly valid for arbitrary real a_1, \dots, a_n .) We shall just sketch the main points here.

First, it is immediate from the form of the integral curves that the critical points are as stated in the theorem. Note that every $V_{\mathbf{u}}$ is certainly a critical point of f , but these are not the only critical points; the others are obtained by taking the orbits of the $V_{\mathbf{u}}$'s under the product of unitary groups $U(E_1) \times \dots \times U(E_l)$.

Let us now try to identify the stable manifold of a critical point $V_{\mathbf{u}}$. Each NDCM contains at least one point of this form. Moreover, since the negative bundles turn out to be homogeneous vector bundles on the NDCM $M_{\mathbf{u}} \cong Gr_{k_1}(E_1) \times \dots \times Gr_{k_l}(E_l)$, the stable manifold of $V_{\mathbf{u}}$ will determine the entire negative bundle.

Let g be the Morse function on $Gr_k(\mathbb{C}^n)$ corresponding to distinct real numbers $c_1 > \dots > c_n \geq 0$, as in the previous section. The stable manifold $S_{\mathbf{u}}^f$

of f . For $k = 1$ this is easy, but for general k it is a little harder. In this section we shall carry out the calculation, in two quite different ways.

Recall that f has $\binom{n}{k}$ critical points, namely the k -planes $V_{\mathbf{u}} = V_{u_1} \oplus \dots \oplus V_{u_k}$, where $1 \leq u_1 < \dots < u_k \leq n$. From 3.2 the (complex) dimension of the stable manifold $S_{\mathbf{u}}$ is $\sum_{i=1}^k (u_i - i)$, so the index of $V_{\mathbf{u}}$ as a critical point of $-f$ is $2 \sum_{i=1}^k (u_i - i)$.

If we define

$$a_d = |\{\mathbf{u} \mid \sum_{i=1}^k (u_i - i) = d\}|$$

then the Morse polynomial of $-f$ is $M(t) = \sum_{d \geq 0} a_d t^{2d}$. Our task, therefore, is to calculate a_d .

There is a one to one correspondence between

$$\{(u_1, \dots, u_k) \mid 1 \leq u_1 < \dots < u_k \leq n, \sum_{i=1}^k (u_i - i) = d\}$$

and

$$\{(p_1, \dots, p_k) = (u_1 - 1, \dots, u_k - k) \mid 0 \leq p_1 \leq \dots \leq p_k \leq n - k, \sum_{i=1}^k p_i = d\}.$$

Thus, a_d is equal to ‘the number of partitions of d into at most k integers, where each such integer is at most $n - k$ ’. Unfortunately there is no simple formula for a_d , but there is a formula for the generating function $G(t) = \sum_{d \geq 0} a_d t^d$ (and this is exactly what we want, since $M(t) = G(t^2)$).

Before stating this formula, we recall the well known fact that, if b_d denotes the number of (unrestricted) partitions of d , then

$$\sum_{d \geq 0} b_d t^d = \frac{1}{(1-t)(1-t^2)(1-t^3)\dots}$$

This means that b_d is equal to the coefficient of t^d in the formal expansion of the right hand side. The formula for $G(t)$ is:

Proposition 3.3.1 *The generating function $G(t) = \sum_{d \geq 0} a_d t^d$ is given by*

$$G(t) = \frac{(1 - t^{n-k+1})(1 - t^{n-k+2}) \dots (1 - t^n)}{(1-t)(1-t^2)\dots(1-t^k)}$$

This is a purely combinatorial statement, which may be proved directly. However, we shall give an indirect proof, by making use of a Morse-Bott function on $Gr_k(\mathbb{C}^n)$. In doing so, we shall find a simple inductive formula for $G(t)$ as well.

Let $g: Gr_k(\mathbb{C}^n) \rightarrow \mathbb{R}$ be the Morse-Bott function corresponding to the choice of real numbers

$$a_1 = 1, a_2 = \dots = a_n = 0.$$

Amongst the non-constant Morse-Bott functions of section 3.2, this is the ‘crudest’, in the sense that it has the least number of critical manifolds. (Morse functions are at the opposite extreme; they have the largest number of critical manifolds.)

From Theorem 3.2.1, the critical manifolds are as follows:

$$\begin{aligned} M_{max} &= \{V \in Gr_k(\mathbb{C}^n) \mid V = V_1 \oplus W, W \subseteq V_2 \oplus \dots \oplus V_n\} \cong Gr_{k-1}(\mathbb{C}^{n-1}) \\ M_{min} &= \{V \in Gr_k(\mathbb{C}^n) \mid V \subseteq V_2 \oplus \dots \oplus V_n\} \cong Gr_k(\mathbb{C}^{n-1}). \end{aligned}$$

The unstable manifold of M_{max} must be the complement of M_{min} , so the index of M_{max} for g is $2(\dim Gr_k(\mathbb{C}^n) - \dim Gr_{k-1}(\mathbb{C}^{n-1}))$, i.e. $2(n-k)$. The index of M_{min} is of course zero. So the Morse-Bott polynomial of g is

$$MB(t) = P_{k,n-1}(t) + t^{2(n-k)}P_{k-1,n-1}(t)$$

where $P_{i,j}(t)$ denotes the Poincaré polynomial of $Gr_i(\mathbb{C}^j)$.

Since g is a perfect Morse-Bott function, this gives an inductive formula for $P_{k,n}(t) = MB(t)$:

Proposition 3.3.2 *The Poincaré polynomial $P_{k,n}(t)$ of $Gr_k(\mathbb{C}^n)$ satisfies the relation*

$$P_{k,n}(t) = P_{k,n-1}(t) + t^{2(n-k)}P_{k-1,n-1}(t).$$

□

For example: $P_{2,4}(t) = P_{2,3}(t) + t^4P_{1,3}(t) = (1 + t^2 + t^4) + t^4(1 + t^2 + t^4) = 1 + t^2 + 2t^4 + t^6 + t^8$.

We can also use this to give a proof of Proposition 3.3.1, which is equivalent to the following slightly more symmetrical statement:

Proposition 3.3.3 *The Poincaré polynomial $P_{k,n}(t)$ of $Gr_k(\mathbb{C}^n)$ is given by*

$$P_{k,n}(t) = \frac{\prod_{i=1}^n (1 - t^{2i})}{\prod_{i=1}^k (1 - t^{2i}) \prod_{i=1}^{n-k} (1 - t^{2i})}.$$

Proof Denote the right hand side by $B_{k,n}(t)$. It is easy to verify that this satisfies the same recurrence relation as $P_{k,n}(t)$, i.e. $B_{k,n}(t) = B_{k,n-1}(t) + t^{2(n-k)}B_{k-1,n-1}(t)$. Since the recurrence relation determines $B_{k,n}(t)$ or $P_{k,n}(t)$ inductively, and these agree when $k = 1$, they must be equal. □

3.4 Schubert varieties

In section 3.1 we described explicitly the stable manifold $S_{\mathbf{u}}$ of a critical point $V_{\mathbf{u}} = V_{u_1} \oplus \dots \oplus V_{u_k}$ of a Morse function $f: Gr_k(\mathbb{C}^n) \rightarrow \mathbb{R}$. It has the form

$$S_{\mathbf{u}} = \left\{ V = \left[\begin{array}{ccc} * & * & * \\ * & * & * \\ 1 & 0 & 0 \\ & 1 & 0 \\ & & * \\ & & 1 \\ & & 0 \\ & & 0 \end{array} \right] \in Gr_k(\mathbb{C}^n) \mid * \in \mathbb{C} \right\}.$$

From 3.1 it is clear that such V are characterized *geometrically* by the following conditions:

$$\dim V \cap \mathbb{C}^i = \begin{cases} 0 & \text{if } 1 \leq i \leq u_1 - 1 \\ 1 & \text{if } u_1 \leq i \leq u_2 - 1 \\ \dots & \\ k & \text{if } u_k \leq i \leq n \end{cases}$$

In other words, we have

$$S_{\mathbf{u}} = \{V \in Gr_k(\mathbb{C}^n) \mid \dim V \cap \mathbb{C}^i = \dim V_{\mathbf{u}} \cap \mathbb{C}^i \text{ for all } i\}.$$

The condition on $\dim V \cap \mathbb{C}^i$ is called a ‘Schubert condition’. It can be specified either by listing $v_i = \dim V \cap \mathbb{C}^i$, i.e.

$$v_1 = \dots = v_{u_1-1} = 0, \quad v_{u_1} = \dots = v_{u_2-1} = 1, \quad \dots, \quad v_{u_k} = \dots = v_n = k$$

or, more economically, by listing those i such that $\dim V \cap \mathbb{C}^i = \dim V \cap \mathbb{C}^{i-1} + 1$, i.e.

$$u_1, u_2, \dots, u_k.$$

The k -tuple $\mathbf{u} = (u_1, \dots, u_k)$ is referred to as a ‘Schubert symbol’, and the set $S_{\mathbf{u}}$ is called the ‘Schubert cell’ associated to \mathbf{u} .

Similarly, the unstable manifold $U_{\mathbf{u}}$ is characterized by these conditions:

$$\dim V \cap (\mathbb{C}^{n-i})^{\perp} = \begin{cases} 0 & \text{if } 1 \leq i \leq n - u_k \\ 1 & \text{if } n - u_k + 1 \leq i \leq n - u_{k-1} \\ \dots & \\ k & \text{if } n - u_1 + 1 \leq i \leq n \end{cases}$$

Example 3.4.1 For the Morse function $f: Gr_2(\mathbb{C}^4) \rightarrow \mathbb{R}$ (of section 3.1), there are 6 Schubert cells. We give the matrix representations below, followed by the

sequence $\dim V \cap \mathbb{C}, \dots, \dim V \cap \mathbb{C}^4$, the Schubert symbol, and the dimension of the cell.

$$\begin{pmatrix} * & * \\ * & * \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad 0, 0, 1, 2 \quad (3, 4) \quad \dim_{\mathbb{C}} = 4$$

$$\begin{pmatrix} * & * \\ 1 & 0 \\ 0 & * \\ 0 & 1 \end{pmatrix} \quad 0, 1, 1, 2 \quad (2, 4) \quad \dim_{\mathbb{C}} = 3$$

$$\begin{pmatrix} * & * \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad 0, 1, 2, 2 \quad (2, 3) \quad \dim_{\mathbb{C}} = 2$$

$$\begin{pmatrix} 1 & 0 \\ 0 & * \\ 0 & * \\ 0 & 1 \end{pmatrix} \quad 1, 1, 1, 2 \quad (1, 4) \quad \dim_{\mathbb{C}} = 2$$

$$\begin{pmatrix} 1 & 0 \\ 0 & * \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad 1, 1, 2, 2 \quad (1, 3) \quad \dim_{\mathbb{C}} = 1$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad 1, 2, 2, 2 \quad (1, 2) \quad \dim_{\mathbb{C}} = 0$$

It follows that the Poincaré polynomial of $Gr_2(\mathbb{C}^4)$ is $1 + t^2 + 2t^4 + t^6 + t^8$. \square

Definition 3.4.2 *The Schubert variety $X_{\mathbf{u}}$ associated to the Schubert symbol \mathbf{u} is the closure of $S_{\mathbf{u}}$ (with respect to the usual topology of $Gr_k(\mathbb{C}^n)$), i.e.*

$$X_{\mathbf{u}} = \overline{S_{\mathbf{u}}} = \{V \in Gr_k(\mathbb{C}^n) \mid \dim V \cap \mathbb{C}^i \geq v_i \text{ for all } i\}.$$

It is easy to show that $X_{\mathbf{u}}$ is an algebraic subvariety of $Gr_k(\mathbb{C}^n)$ (see section 3.5). This subvariety may have singularities. For example, in the case of $Gr_2(\mathbb{C}^4)$, we have

$$X_{(2,4)} = \{V \in Gr_2(\mathbb{C}^4) \mid \dim V \cap \mathbb{C}^2 \geq 1\}$$

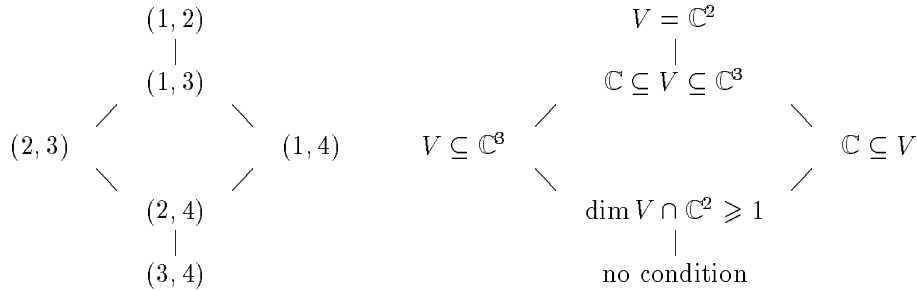
(the v_i 's are given by $(v_1, v_2, v_3, v_4) = (0, 1, 1, 2)$, but the conditions $\dim V \cap \mathbb{C} \geq 0$, $\dim V \cap \mathbb{C}^3 \geq 1$, $\dim V \cap \mathbb{C}^4 \geq 2$ are automatically³ satisfied). The point $V = \mathbb{C}^2$ is a singular point of $X_{(2,4)}$, but $X_{(2,4)} - \{\mathbb{C}^2\}$ is smooth, having the structure of a fibre bundle over $\mathbb{C}\mathbb{P}^1$ with fibre $\mathbb{C}\mathbb{P}^2 - \{\text{point}\}$.

Observe that

$$X_{(2,4)} = S_{(2,4)} \cup S_{(2,3)} \cup S_{(1,4)} \cup S_{(1,3)} \cup S_{(1,2)}.$$

It is clear from the definition that, in general, $X_{\mathbf{u}}$ is a disjoint union of Schubert cells. This gives rise to a partial order on the set of Schubert symbols: we define $\mathbf{u}_1 \leq \mathbf{u}_2$ if and only if $\overline{S_{\mathbf{u}_1}} \supseteq S_{\mathbf{u}_2}$.

In the case of $Gr_2(\mathbb{C}^4)$, this partial order is represented in the following diagrams. The second diagram indicates the conditions which define the Schubert varieties.



Although the partial order is a simple consequence of the definition of Schubert variety, it provides nontrivial information on the behaviour of the flow lines of $-\nabla f$. Namely, the condition $\mathbf{u}_1 \leq \mathbf{u}_2$ is equivalent to the existence of a flow line $\gamma(t)$ from \mathbf{u}_2 to \mathbf{u}_1 , i.e. such that $\lim_{t \rightarrow -\infty} \gamma(t) = V_{\mathbf{u}_2}$, $\lim_{t \rightarrow \infty} \gamma(t) = V_{\mathbf{u}_1}$. (This is not immediately obvious from the condition $\overline{S_{\mathbf{u}_1}} \supseteq S_{\mathbf{u}_2}$, but it does follow from the geometrical description of $S_{\mathbf{u}_1} \cap U_{\mathbf{u}_2}$.)

We shall return to Schubert varieties in section 3.6, when we discuss the cohomology ring of a Grassmannian, so we conclude this section with some further remarks.

First, since the flow lines of $-\nabla f$ ‘preserve’ the Schubert variety $X_{\mathbf{u}}$, we deduce that $X_{\mathbf{u}}$ inherits a decomposition into (possibly singular) ‘stable manifolds’ (or ‘unstable manifolds’).

Second, although the Morse function f depends on a choice of real numbers $a_1 > \dots > a_n$, the Schubert varieties are independent of this choice. In fact they

³If V, W are linear subspaces of \mathbb{C}^n of dimensions k, l respectively, then we have $W/W \cap V \cong W + V/V$, and hence $\dim W \cap V + \dim(W + V) = k + l$.

depend solely on the standard 'flag'

$$\mathbb{C} \subseteq \mathbb{C}^2 \subseteq \dots \subseteq \mathbb{C}^n,$$

or on the standard ordered basis e_1, \dots, e_n of \mathbb{C}^n . If a_1, \dots, a_n are arbitrary distinct real numbers, we obtain a similar Morse function, possibly corresponding to a re-ordering of e_1, \dots, e_n . More generally still, any choice of orthonormal basis, or equivalently any flag, corresponds to a similar Morse function. The formula for such a function is obtained by replacing the diagonal matrix A by UAU^{-1} , where U is a unitary matrix.

There is a geometrical description of the stable manifolds of the Morse-Bott functions considered in section 3.2, i.e. where the real numbers $a_1 \geq \dots \geq a_n$ are not necessarily distinct. We state the result without proof (as the easiest proof depends on the more general theory of §4). Recall that the critical manifolds are denoted $M_{\mathbf{u}}$, and that each such NDCM contains at least one point of the form $V_{\mathbf{u}}$. The stable manifold $S_{\mathbf{u}}$ of $M_{\mathbf{u}}$, i.e. the union of the stable manifolds of all points of $M_{\mathbf{u}}$, is then given by:

$$S_{\mathbf{u}} = \{V \in Gr_k(\mathbb{C}^n) \mid \dim V \cap (E_1 \oplus \dots \oplus E_i) = \dim V_{\mathbf{u}} \cap (E_1 \oplus \dots \oplus E_i) \text{ for all } i\},$$

where E_1, \dots, E_l are the eigenspaces of $\text{diag}(a_1, \dots, a_n)$ as in section 3.2. Thus, in this case, the Schubert 'cells' (or rather, Schubert cell-bundles) depend only on the 'partial flag'

$$E_1 \subseteq E_1 \oplus E_2 \subseteq \dots \subseteq E_1 \oplus \dots \oplus E_l = \mathbb{C}^n.$$

Conversely, as in the case of the Morse functions discussed earlier, *any* partial flag determines a Morse-Bott function on $Gr_k(\mathbb{C}^n)$.

The Schubert cell-bundles (or their NDCMs) are parametrized by l -tuples (c_1, \dots, c_l) of non-negative integers with $c_j \leq \dim E_j$ and $c_1 + \dots + c_l = k$; namely $c_j = \dim V_{\mathbf{u}} \cap E_j$. We consider (c_1, \dots, c_l) to be a 'generalized Schubert symbol'. The bundle projection map $S_{\mathbf{u}} \rightarrow M_{\mathbf{u}}$ is given explicitly by

$$V \mapsto (V_{(1)}, \dots, V_{(l)}) \in Gr_{c_1}(E_1) \times \dots \times Gr_{c_l}(E_l)$$

where

$$V_{(i)} = V \cap E_1 \oplus \dots \oplus E_i + E_1 \oplus \dots \oplus E_{i-1} / E_1 \oplus \dots \oplus E_{i-1}.$$

The integers $w_i = c_1 + \dots + c_i$, $i = 1, \dots, l$, are analogous to the integers v_i in the case of a Morse function. In terms of these integers, the Schubert cell-bundle $S_{\mathbf{u}}$ is

$$S_{\mathbf{u}} = \{V \in Gr_k(\mathbb{C}^n) \mid \dim V \cap (E_1 \oplus \dots \oplus E_i) = w_i \text{ for all } i\},$$

By taking the closure of a Schubert cell-bundle, we obtain a generalized Schubert variety, namely

$$X_{\mathbf{u}} = \overline{S_{\mathbf{u}}} = \{V \in Gr_k(\mathbb{C}^n) \mid \dim V \cap E_1 \oplus \dots \oplus E_i \geq w_i \text{ for all } i\}.$$

One of the advantages of having these explicit descriptions of Schubert cell-bundles is that we can compute easily the indices of the critical points; we shall give two examples below.

Example 3.4.3 Let us choose the partial flag $\mathbb{C} \subseteq \mathbb{C}^n$. This corresponds to a Morse-Bott function on $Gr_k(\mathbb{C}^n)$ with $a_1 > a_2 = \dots = a_n$. We have already considered such a function in section 3.3; there are two critical manifolds. The stable manifold of the maximum NDCM is just that NDCM, and the stable manifold of the minimum NDCM is the complement of the maximum NDCM.

The eigenspace decomposition of \mathbb{C}^n is given by $E_1 = \mathbb{C}$, $E_2 = \mathbb{C}^\perp$, and the generalized Schubert symbols are $(c_0, c_1) = (1, k-1)$ and $(c_0, c_1) = (0, k)$. \square

Example 3.4.4 Consider the Morse-Bott function on $\mathbb{C}\mathbb{P}^6 = Gr_1(\mathbb{C}^7)$ given by $a_1 = a_2 = 2$, $a_3 = a_4 = a_5 = 1$, $a_6 = a_7 = 0$. In this case the eigenspace decomposition of \mathbb{C}^7 is given by

$$\mathbb{C}^7 = E_1 \oplus E_2 \oplus E_3, \quad E_1 = V_1 \oplus V_2, E_2 = V_3 \oplus V_4 \oplus V_5, E_3 = V_6 \oplus V_7.$$

We list below the generalized Schubert symbols, followed by the NDCM, and then the Schubert cell-bundle.

$$\begin{aligned} (c_0, c_1, c_2) &= (1, 0, 0), & M_{\mathbf{u}} &= \mathbb{P}(E_1), & S_{\mathbf{u}} &= \mathbb{P}(E_1) \\ (c_0, c_1, c_2) &= (0, 1, 0), & M_{\mathbf{u}} &= \mathbb{P}(E_2), & S_{\mathbf{u}} &= \mathbb{P}(E_1 \oplus E_2) - \mathbb{P}(E_1) \\ (c_0, c_1, c_2) &= (0, 0, 1), & M_{\mathbf{u}} &= \mathbb{P}(E_3), & S_{\mathbf{u}} &= \mathbb{P}(E_1 \oplus E_2 \oplus E_3) - \mathbb{P}(E_1 \oplus E_2) \end{aligned}$$

In particular, the Morse indices (for the function $-f$) are 0, 4, 10 respectively, and the Morse-Bott polynomial is $t^0(1+t^2) + t^4(1+t^2+t^4) + t^{10}(1+t^2)$. This is equal to the Poincaré polynomial of $\mathbb{C}\mathbb{P}^6$, as expected. \square

3.5 Morse theory of the Plücker embedding

There is a well known embedding

$$Gr_k(\mathbb{C}^n) \rightarrow \mathbb{C}\mathbb{P}^N, \quad N = \binom{n}{k} - 1$$

called the Plücker embedding. It is defined by

$$V \mapsto \wedge^k V \subseteq \wedge^k \mathbb{C}^n \cong \mathbb{C}^{N+1}.$$

If e_1, \dots, e_n are the standard basis vectors of \mathbb{C}^n , then the vectors $e_{\mathbf{u}} = e_{u_1} \wedge \dots \wedge e_{u_k}$ with $1 \leq u_1 < \dots < u_k \leq n$ constitute a basis of $\wedge^k \mathbb{C}^n$.

Let $f: Gr_k(\mathbb{C}^n) \rightarrow \mathbb{R}$ be the Morse-Bott function defined by certain real numbers a_1, \dots, a_n , as in section 3.2. Then by Lemma 3.1.2 the one parameter diffeomorphism group of the vector field $-\nabla f$ is induced by the action

$$t \cdot \sum \lambda_i e_i = \sum \lambda_i e^{-ta_i} e_i$$

of \mathbb{R} on \mathbb{C}^n .

This action is the restriction of the action

$$t \cdot \sum \lambda_{\mathbf{u}} e_{\mathbf{u}} = \sum \lambda_{\mathbf{u}} e^{-t(a_{u_1} + \dots + a_{u_k})} e_{\mathbf{u}}$$

of \mathbb{R} on $\wedge^k \mathbb{C}^n$. But this action is the one parameter diffeomorphism group of the vector field $-\nabla F$, where $F: \mathbb{C}\mathbb{P}^N \rightarrow \mathbb{R}$ is the Morse-Bott function defined by the $\binom{n}{k}$ real numbers $a_{u_1} + \dots + a_{u_k}$. We conclude that the Morse-Bott theory of f on $Gr_k(\mathbb{C}^n)$ is just the 'restriction' of the (much simpler!) Morse-Bott theory of F on $\mathbb{C}\mathbb{P}^N$:

Proposition 3.5.1 *Let $M_{\mathbf{u}}, S_{\mathbf{u}}, X_{\mathbf{u}}$ denote the NDCM's, Schubert cell-bundles, and generalized Schubert varieties for the Morse-Bott function $f: Gr_k(\mathbb{C}^n) \rightarrow \mathbb{R}$. Let $M_{\mathbf{u}}^F, S_{\mathbf{u}}^F, X_{\mathbf{u}}^F$ denote the corresponding objects for the Morse-Bott function $F: \mathbb{C}\mathbb{P}^N \rightarrow \mathbb{R}$. Then we have $M_{\mathbf{u}} = Gr_k(\mathbb{C}^n) \cap M_{\mathbf{u}}^F$, $S_{\mathbf{u}} = Gr_k(\mathbb{C}^n) \cap S_{\mathbf{u}}^F$, and $X_{\mathbf{u}} \subseteq Gr_k(\mathbb{C}^n) \cap X_{\mathbf{u}}^F$. \square*

Observe that it is possible to choose the real numbers a_1, \dots, a_n so that both f and F are Morse functions. But it may happen that f is a Morse function even when F is not.

The spaces $M_{\mathbf{u}}^F, S_{\mathbf{u}}^F, X_{\mathbf{u}}^F$ are determined by an eigenspace decomposition

$$\mathbb{C}^{N+1} = \hat{E}_1 \oplus \dots \oplus \hat{E}_k$$

in the usual way. Using this notation, the NDCM's of F are the linear subspaces $\mathbb{P}(\hat{E}_i)$, and the stable manifold of $\mathbb{P}(\hat{E}_i)$ is given explicitly as a bundle over $\mathbb{P}(\hat{E}_i)$ by

$$\mathbb{P}(\hat{E}_0 \oplus \dots \oplus \hat{E}_i) - \mathbb{P}(\hat{E}_0 \oplus \dots \oplus \hat{E}_{i-1}).$$

The projection map to $\mathbb{P}(\hat{E}_i)$ sends a line L in $\hat{E}_0 \oplus \dots \oplus \hat{E}_i$ to the line $L + \hat{E}_0 \oplus \dots \oplus \hat{E}_{i-1}$ in $\hat{E}_0 \oplus \dots \oplus \hat{E}_i / \hat{E}_0 \oplus \dots \oplus \hat{E}_{i-1} \cong \hat{E}_i$.

The associated Schubert variety, i.e. the closure of this stable manifold, is just

$$\mathbb{P}(\hat{E}_0 \oplus \dots \oplus \hat{E}_i),$$

which is a *linear subspace* of $\mathbb{C}\mathbb{P}^N$. Although $X_{\mathbf{u}}$ is not necessarily *equal* to the intersection of this space with $Gr_k(\mathbb{C}^n)$, it is in fact true that $X_{\mathbf{u}}$ is given by the intersection of *some* linear subspace with $Gr_k(\mathbb{C}^n)$. This follows from the fact that

$$S_{\mathbf{u}} = Gr_k(\mathbb{C}^n) \cap \mathbb{P}(\hat{E}_0 \oplus \dots \oplus \hat{E}_i) - \mathbb{P}(\hat{E}_0 \oplus \dots \oplus \hat{E}_{i-1})$$

(as some of the linear equations defining $\mathbb{P}(\hat{E}_0 \oplus \dots \oplus \hat{E}_i)$ in \mathbb{C}^{n+1} may become dependent in the presence of the equations defining $Gr_k(\mathbb{C}^n)$).

The fact that the Plücker embedding is compatible with the natural Morse-Bott functions on $Gr_k(\mathbb{C}^n)$ and $\mathbb{C}\mathbb{P}^N$ may be explained group-theoretically. The key point is that the Plücker embedding is induced by an irreducible representation $U_n \rightarrow U_{N+1}$. However, it seems technically easier to work with the explicit formulae for the flow lines, as we have done in this section.

3.6 Cohomology of the Grassmannian, and the Schubert calculus

In this section we consider only Morse functions on $Gr_k(\mathbb{C}^n)$.

From a Schubert symbol $\mathbf{u} = (u_1, \dots, u_k)$ we obtain a cell $S_{\mathbf{u}}$ in $Gr_k(\mathbb{C}^n)$ of (real) dimension $2 \sum_{i=1}^k (u_i - i)$. By the Morse inequalities for the coefficient group \mathbb{Z} (or by standard theory of cellular homology) it follows that $H_*(Gr_k(\mathbb{C}^n); \mathbb{Z})$ is a free abelian group with one generator for each Schubert cell. Let $x_{\mathbf{u}}$ be the homology class represented by⁴ $X_{\mathbf{u}}$. Both $x_{\mathbf{u}}$ and $X_{\mathbf{u}}$ are referred to as ‘Schubert cycles’.

By Poincaré Duality we obtain a dual cohomology class $z_{\mathbf{u}}$ of dimension $2k(n-k) - 2 \sum_{i=1}^k (u_i - i)$. Thus, $\dim z_{\mathbf{u}} = \text{codim } x_{\mathbf{u}}$.

An example of particular interest is the generator of $H^2(Gr_k(\mathbb{C}^n); \mathbb{Z}) \cong \mathbb{Z}$; this corresponds to the unique codimension one Schubert cycle, i.e. to the Schubert symbol $(n-k, n-k+2, \dots, n-1, n)$. The Schubert conditions here are

$$\dim V \cap \mathbb{C} = \dots = \dim V \cap \mathbb{C}^{n-k-1} = 0, \quad \dim V \cap \mathbb{C}^{n-k} = 1, \quad \dots, \quad \dim V \cap \mathbb{C}^n = k,$$

and the Schubert variety is characterized by the single condition $\dim V \cap \mathbb{C}^{n-k} \geq 1$. In terms of the Plücker embedding (see section 3.5), we have

$$X_{(n-k, n-k+2, \dots, n-1, n)} = Gr_k(\mathbb{C}^n) \cap H$$

where $\mathbb{P}(H)$ is the Schubert variety for the critical point $V_{n-k} \wedge V_{n-k+2} \wedge \dots \wedge V_{n-1} \wedge V_n$ in $\mathbb{C}\mathbb{P}^N$. Now, since $a_1 > \dots > a_n$, we have $a_n + \dots + a_{n-k+1} > a_n + \dots + a_{n-k+2} + a_{n-k} > \dots$, so H is the *hyperplane* in \mathbb{C}^{N+1} orthogonal to $V_{n-k+1} \wedge \dots \wedge V_{n-1} \wedge V_n$. It is well known that the cohomology class dual to $\mathbb{P}(H)$ is a generator of $H^2(\mathbb{C}\mathbb{P}^N; \mathbb{Z})$, so we deduce that the induced homomorphism $H^2(\mathbb{C}\mathbb{P}^N; \mathbb{Z}) \rightarrow H^2(Gr_k(\mathbb{C}^n); \mathbb{Z})$ is an isomorphism.

The multiplicative behaviour of $H^*(Gr_k(\mathbb{C}^n); \mathbb{Z})$ is equivalent to the behaviour of the intersections of generic representatives of homology classes (we shall make a more precise statement shortly). As a first step towards describing this, we shall need a slight generalization of Schubert varieties.

In section 3.4 we pointed out that the definition of the Schubert varieties $X_{\mathbf{u}}$ depends only on the choice of the standard flag $\mathbb{C} \subseteq \mathbb{C}^2 \subseteq \dots \subseteq \mathbb{C}^n$. If we choose a new flag $F_1 \subseteq F_2 \subseteq \dots \subseteq F_n = \mathbb{C}^n$, denoted by F , then we obtain new objects $S_{\mathbf{u}}^F, X_{\mathbf{u}}^F, x_{\mathbf{u}}^F, z_{\mathbf{u}}^F$ defined in exactly the same way as $S_{\mathbf{u}}, X_{\mathbf{u}}, x_{\mathbf{u}}, z_{\mathbf{u}}$, but using the new flag instead of the standard flag. Since any two flags are related by an element of the unitary group U_n , however, we have $x_{\mathbf{u}}^F = x_{\mathbf{u}}$ and $z_{\mathbf{u}}^F = z_{\mathbf{u}}$. So we may regard the Schubert cycles $X_{\mathbf{u}}^F$ as a family of representatives for the homology class $x_{\mathbf{u}}$, parametrized by the space of all flags.

For example, consider the ‘opposite’ flag

$$(\mathbb{C}^{n-1})^\perp \subseteq (\mathbb{C}^{n-2})^\perp \subseteq \dots \subseteq \mathbb{C}^\perp \subseteq \mathbb{C}^n;$$

⁴The precise meaning of this is explained in Appendix B of [30].

let us denote the Schubert varieties with respect to this flag by $X_{\mathbf{u}}^c$. It is easy to check that

$$X_{\mathbf{u}}^c = \overline{U}_{\mathbf{u}^c}$$

where $\mathbf{u}^c = (n - u_k + 1, n - u_{k-1} + 1, \dots, n - u_1 + 1)$. Thus, both $\overline{S}_{\mathbf{u}}$ and $\overline{U}_{\mathbf{u}^c}$ are representatives of the same homology class $x_{\mathbf{u}}$.

We now state a special case of an important general principle (see Appendix B of [30]):

Theorem 3.6.1 *If $x_{\mathbf{u}}$ and $x_{\mathbf{v}}$ are Schubert cycles with $\dim z_{\mathbf{u}} + \dim z_{\mathbf{v}} = \dim Gr_k(\mathbb{C}^n)$, then the product $z_{\mathbf{u}}z_{\mathbf{v}} \in H^{\dim Gr_k(\mathbb{C}^n)}(Gr_k(\mathbb{C}^n); \mathbb{Z}) \cong \mathbb{Z}$ of the corresponding cohomology classes is equal to the intersection number of $X_{\mathbf{u}}$ and $X_{\mathbf{v}}$. \square*

This intersection number is equal to the number of points (counted with multiplicities) in $X_{\mathbf{u}}^{F_1} \cap X_{\mathbf{v}}^{F_2}$ whenever $X_{\mathbf{u}}^{F_1} \cap X_{\mathbf{v}}^{F_2}$ is a finite set.

Example 3.6.2 For any \mathbf{u} , we have $z_{\mathbf{u}}z_{\mathbf{u}^c} = 1$. This is because the dual homology classes are represented by $\overline{S}_{\mathbf{u}}$ and $\overline{U}_{\mathbf{u}}$ respectively, and these intersect at precisely one point, namely the critical point $V_{\mathbf{u}}$. (Of course, these homology classes may also be represented by $\overline{S}_{\mathbf{u}}$ and $\overline{S}_{\mathbf{u}^c}$. But this is of no interest to us as these cycles intersect at infinitely many points.) To see that the multiplicity of the intersection point is 1, one can use the Plücker embedding – it follows from our discussion in section 3.5 that the multiplicity of *any* isolated point of intersection of two Schubert varieties is precisely 1, since we are just taking the intersection of linear subspaces in $\mathbb{C}P^N$. \square

More generally, we have:

Proposition 3.6.3 *If $x_{\mathbf{u}}$ and $x_{\mathbf{v}}$ are Schubert cycles with $\dim z_{\mathbf{u}} + \dim z_{\mathbf{v}} = \dim Gr_k(\mathbb{C}^n)$, then*

$$z_{\mathbf{u}}z_{\mathbf{v}} = \begin{cases} 1 & \text{if } \mathbf{u} = \mathbf{v}^c \\ 0 & \text{otherwise} \end{cases}$$

Proof We have just seen that $z_{\mathbf{u}}z_{\mathbf{u}^c} = 1$. To show that $X_{\mathbf{u}} \cap X_{\mathbf{v}} = \emptyset$ if $\mathbf{v} \neq \mathbf{u}^c$, one may use the geometrical characterization of $X_{\mathbf{u}}$ and $X_{\mathbf{v}}$ – we omit the details. \square

This proposition is a manifestation of Poincaré Duality, and it allows us to determine the products $z_{\mathbf{u}}z_{\mathbf{v}}$ for *arbitrary* \mathbf{u}, \mathbf{v} . For we may express $z_{\mathbf{u}}z_{\mathbf{v}}$ in terms of the additive Schubert basis as

$$z_{\mathbf{u}}z_{\mathbf{v}} = a_1 z_{\mathbf{u}(1)} + \dots + a_r z_{\mathbf{u}(r)}$$

for some integers a_1, \dots, a_r , where $\dim z_{\mathbf{u}} + \dim z_{\mathbf{v}} = \dim z_{\mathbf{u}(i)}$ for each i . Then we obtain a_i by multiplying the above expression by $z_{\mathbf{u}(i)}^c$:

$$z_{\mathbf{u}} z_{\mathbf{v}} z_{\mathbf{u}(i)}^c = a_i$$

(all other products vanish, by Proposition 3.6.3). Thus, we have to calculate all triple products $z_{\mathbf{u}} z_{\mathbf{v}} z_{\mathbf{w}}$ such that $\dim z_{\mathbf{u}} + \dim z_{\mathbf{v}} + \dim z_{\mathbf{w}} = \dim Gr_k(\mathbb{C}^n)$. Theorem 3.6.1 generalizes to this situation, so we have to calculate the corresponding triple intersections of Schubert varieties.

Example 3.6.4 We shall carry out the calculation of some triple products for $Gr_2(\mathbb{C}^4)$, and hence determine the multiplicative structure of $H^*(Gr_2(\mathbb{C}^4); \mathbb{Z})$. First we list the additive generators:

$$\begin{aligned} z_{(3,4)} &\in H^0 \\ z_{(2,4)} &\in H^2 \\ z_{(1,4)}, z_{(2,3)} &\in H^4 \\ z_{(1,3)} &\in H^6 \\ 1 = z_{(1,2)} &\in H^8 \end{aligned}$$

Proposition 3.6.3 gives the following products:

$$z_{(2,4)} z_{(1,3)} = z_{(1,4)} z_{(1,4)} = z_{(2,3)} z_{(2,3)} = 1, \quad z_{(1,4)} z_{(2,3)} = 0.$$

Let us now try to compute $z_{(1,4)} z_{(2,4)} z_{(2,4)}$. We must find suitably generic representing cycles $X_{\mathbf{u}}^F$ for these classes, by choosing suitably generic flags F .

For $z_{(2,4)}$ we need two modifications of the standard representative

$$X_{(2,4)} = \{V \in Gr_2(\mathbb{C}^4) \mid \dim V \cap \mathbb{C}^2 \geq 1\}.$$

We shall choose the flags

$$\begin{aligned} F' : & V_1 \subseteq V_1 \oplus V_4 \subseteq V_1 \oplus V_2 \oplus V_4 \subseteq \mathbb{C}^4 \\ F'' : & V_2 \subseteq V_2 \oplus V_4 \subseteq V_1 \oplus V_2 \oplus V_4 \subseteq \mathbb{C}^4 \end{aligned}$$

The corresponding cycles are:

$$\begin{aligned} X'_{(2,4)} &= \{V \in Gr_2(\mathbb{C}^4) \mid \dim V \cap V_1 \oplus V_4 \geq 1\} \\ X''_{(2,4)} &= \{V \in Gr_2(\mathbb{C}^4) \mid \dim V \cap V_2 \oplus V_4 \geq 1\} \end{aligned}$$

For $z_{(1,4)}$ we shall choose the flag

$$F''' : V_3 \subseteq V_2 \oplus V_3 \subseteq V_1 \oplus V_2 \oplus V_3 \subseteq \mathbb{C}^4$$

i.e. we choose

$$X'''_{(1,4)} = \{V \in Gr_2(\mathbb{C}^4) \mid V_3 \subseteq V\}.$$

It may now be verified that

$$X'_{(2,4)} \cap X''_{(2,4)} \cap X'''_{(1,4)} = \{V_3 \oplus V_4\}$$

i.e. a single point. As in Example 3.6.2, we can see that the multiplicity of this point is 1. So we conclude that $z_{(1,4)}z_{(2,4)}z_{(2,4)} = 1$. Exactly the same argument gives $z_{(2,3)}z_{(2,4)}z_{(2,4)} = 1$.

The remaining (double) products in $H^*(Gr_2(\mathbb{C}^4); \mathbb{Z})$ are

$$\begin{aligned} z_{(2,4)}z_{(2,4)} &= az_{(2,3)} + bz_{(1,4)} \\ z_{(1,4)}z_{(2,4)} &= cz_{(1,3)} \\ z_{(2,3)}z_{(2,4)} &= dz_{(1,3)}. \end{aligned}$$

Using the two triple products which we have just calculated, we find that $a = b = c = d = 1$. \square

The same method works for $H^*(Gr_k(\mathbb{C}^n); \mathbb{Z})$, although this situation is of course more complicated. There are famous general formulae for the multiplicative structure, which constitute the ‘Schubert calculus’. An elementary approach to these formulae and their traditional applications can be found in [43]; other versions can be found in [36], [39], and [30].

From the theory of Chern classes, there is a well known ‘closed formula’ for the ring structure of $H^*(Gr_k(\mathbb{C}^n); \mathbb{Z})$, namely

$$\frac{\mathbb{Z}[c_1, \dots, c_{n-k}, d_1, \dots, d_k]}{(1 + c_1 + \dots + c_{n-k})(1 + d_1 + \dots + d_k)} = 1$$

where $c_i, d_i \in H^{2i}(Gr_k(\mathbb{C}^n); \mathbb{Z})$. (This is explained in detail in §23 of [15]; another good reference is [49].) It may be checked that this agrees with our description in the case of $Gr_2(\mathbb{C}^4)$. The dual of cohomology class d_i is represented by the Schubert variety $X_{\mathbf{u}}$ with $\mathbf{u} = (n - k, n - k + 1, \dots, n - k + i - 1, n - k + i + 1, \dots, n - 1, n)$; these are called ‘special Schubert varieties’. The cohomology class corresponding to a general Schubert variety is more complicated to describe (see [7] and [30]).

Viewing the cohomology ring of a manifold M as a collection of triple intersection numbers of (representatives of) homology classes is in fact a useful idea. The equivariant cohomology ring $H_G^*(M)$ (see section 2.7) may be described in terms of $H^*(BG)$ -valued triple intersections. Quantum cohomology (see section 4.4) appears in a similar way.

3.7 Next steps

We have now covered the ‘classical’ aspects of Morse theory, and in §4 we shall turn to more recent developments. As motivation for this, we mention here a couple of points which arise from our study of Grassmannians.

An immediate question is: when can the cohomology ring be determined directly from a Morse function? It is a well known limitation of classical Morse

theory that the Morse inequalities give information only about the *additive* structure of the cohomology ring. However, we have seen that the cohomology ring of $Gr_k(\mathbb{C}^n)$ can be found from explicit knowledge of the stable manifolds of a suitable Morse function. Was this just a special trick, or is there perhaps a more general theory which works for Morse functions on arbitrary compact manifolds?

A slightly more subtle (but related) question concerns the possible configurations of flow lines of a Morse function. We saw in §2 that the possible configurations of critical points of a Morse function $f : M \rightarrow \mathbb{R}$ are limited by the topology of M . For example, it is not possible to have a Morse function on $S^1 \times S^1$ with precisely three critical points whose indices are 0, 1 and 2. In the same way, it is not possible to have arbitrary configurations of flow lines connecting the critical points. The question arises as to how these configurations of flow lines are restricted.

These questions may be answered (for certain kinds of manifolds, at least) by generalizing Morse theory in a new way: instead of considering a single Morse function, we consider several. Indeed, when we computed triple products of cohomology classes in section 3.6, we were in fact making use of three ‘independent’ Morse functions on $Gr_k(\mathbb{C}^n)$. The general theory underlying this calculation is what we shall study next.

4 Morse theory in the 1990’s

We describe several recent applications of Morse theory, in which the gradient flow lines play a fundamental role. Although the level of discussion will be somewhat more advanced in this section, the case of a complex Grassmannian should be kept in mind as a typical example. We begin by discussing Morse functions which arise from torus actions on Kähler manifolds; these functions, which include the functions on Grassmannians in §3, have the crucial property that their gradient flow lines are explicitly identifiable. Then we describe the ‘new’ approach to Morse theory, due to Witten. After that we present the ‘field-theoretic’ Morse theory of Cohen-Jones-Segal, Betz-Cohen, and Fukaya.

4.1 Morse functions generated by torus actions

In §1 and §2 we gave a review of the ‘classical’ Morse theory, and then in §3 we illustrated this in detail for a particular example. We shall now focus on some contemporary aspects, which show that the power of Morse theory goes far beyond the computation of homology groups. Our starting point is an important family of examples which includes the Morse and Morse-Bott functions of §3.

Let $T \cong S^1 \times \dots \times S^1$ be a torus, and let M be a simply-connected compact Kähler manifold. Assume that the group T acts on M , and that this action preserves the complex structure J and the Kähler 2-form ω of M . It follows that the action also preserves the Kähler metric $\langle \cdot, \cdot \rangle$, as the latter is given by $\omega(A, B) = \langle A, JB \rangle$.

Let \mathfrak{t} denote the Lie algebra of T . Let $X \in \mathfrak{t}$ be any generator of the torus; this means that T is the closure of its (not necessarily closed) subgroup $\exp \mathbb{R}X$.

The formula

$$X_m^* = \frac{d}{dt} \exp tX \cdot m|_{t=0}$$

defines a vector field X^* on M .

Since M is simply connected, and $\omega(X^*, \cdot)$ is a closed 1-form, there is a function $f^X : M \rightarrow \mathbb{R}$ such that $df^X = \omega(X^*, \cdot)$. We shall see that f^X is a perfect Morse-Bott function, with particularly nice properties. Observe that we have

$$-\nabla f^X = JX^*,$$

from the formula $\langle \nabla f^X, A \rangle = df^X(A) = \omega(X^*, A) = -\langle JX^*, A \rangle$.

Theorem 4.1.1 ([25]) *For any generator X of \mathfrak{t} , the function f^X is a perfect Morse-Bott function on M . The critical points of f^X are the fixed points of the T -action, and they form a finite number of totally geodesic Kähler submanifolds of M .*

Sketch of the proof The fact that the critical points of f^X are the fixed points of T follows from the formula $-\nabla f^X = JX^*$.

Let m be a fixed point of T . Then there is an induced action of the Lie group T , and hence also of the Lie algebra \mathfrak{t} , on the vector space $T_m M$. This means that we have a Lie group homomorphism $\Theta_m : T \rightarrow Gl(T_m M)$ and a Lie algebra homomorphism $\theta_m : \mathfrak{t} \rightarrow \text{End}(T_m M)$.

Since T acts by isometries (i.e. the action of T preserves the metric), $\Theta_m(t)$ is an orthogonal transformation, and $\theta_m(X)$ a skew-symmetric transformation, for each $t \in T$, $X \in \mathfrak{t}$. As T is abelian, we may put these transformations simultaneously into canonical form. This means that there exists a decomposition

$$T_m M = V_0 \oplus V_1 \oplus \dots \oplus V_r$$

such that $\theta|_{V_0} = 0$ and

$$\theta|_{V_i} = \begin{pmatrix} 0 & w_i \\ -w_i & 0 \end{pmatrix}$$

for nontrivial linear functionals w_1, \dots, w_r on \mathfrak{t} . Note that the subspaces V_i for $i > 0$ are not uniquely determined in general, and that the linear functionals w_i are determined only up to sign.

By considering geodesics through m (see [44]), it can be shown that the connected component of the fixed point set of T containing m is a submanifold of M – the geodesics through m in the direction of V_0 give a local chart (via the exponential map). This argument also shows that the submanifold is totally geodesic, with tangent space V_0 at m .

Up to this point, we have used only the fact that the action of T preserves the metric. Since T preserves the complex structure, J commutes with the linear transformations $\Theta_m(t)$ and $\theta_m(t)$. We may therefore choose the decomposition so that J preserves each subspace V_i . It follows that each connected component

of the fixed point set of T is actually a Kähler manifold, and that (for $i > 0$) we may write

$$J|_{V_i} = \pm \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

An additional consequence of the existence of J is that each V_i acquires a natural orientation, and so the linear functionals ω_i are now determined uniquely (for a given choice of subspaces V_i).

We now turn to the computation of the Hessian H of f^X . In section 1.2 we gave a definition in terms of local coordinates, but this is equivalent to the following more invariant definition (see page 4 of [47]). For any vectors $V, W \in T_m M$, we have

$$H(V, W) = \tilde{V}(\tilde{W}(f^X))(m) = \tilde{V}df^X(\tilde{W})(m)$$

where \tilde{V}, \tilde{W} are any extensions of V, W to local vector fields on M . We therefore have $H(V, W) = \tilde{V}\omega(X^*, \tilde{W})(m) = \tilde{V}\alpha(X^*)(m)$, where $\alpha = -i_{\tilde{W}}\omega$ (and i denotes interior product). With a suitable choice of the extensions \tilde{V}, \tilde{W} , the well known formula for $d\alpha$ shows that $H(V, W) = -\omega(\tilde{W}, [\tilde{V}, X^*])$.

From the definition of θ , it follows that $\theta(X)V = [\tilde{V}, X^*]_m$. Hence we obtain the formula

$$H(V, W) = \langle W, -J\theta(X)V \rangle.$$

On each V_i with $i > 0$, it follows that H is equal to the inner product times the nontrivial linear functional ω_i . Hence the Hessian is nondegenerate on a space complementary to V_0 , i.e. f^X is a Morse-Bott function.

The index of the Hessian at a critical point m is even, being twice the number of ω_i 's such that $\omega_i(m) < 0$. This implies that the Morse-Bott function f^X is perfect, and so the proof of Theorem 4.1.1 is complete. \square

The linear functionals w_1, \dots, w_r on \mathfrak{t} which appear in this proof are called the (nonzero) weights of the action of T at the fixed point m . Theorem 4.1.1 generalizes to the case of a symplectic manifold, as was pointed out in [25]. We shall not need this extra generality, and in fact we shall make essential use of the complex structure J , so we shall only consider Kähler manifolds here.

The formula $-\nabla f^X = JX^*$ leads to a geometrical description of the flow lines of $-\nabla f^X$, as we shall explain next.

Lemma 4.1.2 *The action of $T = S^1 \times \dots \times S^1$ on M extends to an action of the complexified torus $T^{\mathbb{C}} = \mathbb{C}^* \times \dots \times \mathbb{C}^*$ on M . The vector $\sqrt{-1}X \in \mathfrak{t}^{\mathbb{C}}$ generates the vector field JX^* , i.e. $(\sqrt{-1}X)^* = JX^*$. The flow line γ of $-\nabla f^X$ passing through $m \in M$ is given by the action of the subgroup $\sqrt{-1}\mathbb{R}X$, i.e. $\gamma(t) = \exp \sqrt{-1}tX \cdot m$.*

Sketch of the proof The extension of the action is guaranteed by the fact that T preserves the complex structure of M (see [2]). A direct construction of the extension may be given by using the integral curves of JX^* . The fact

that $(\sqrt{-1} X)^* = JX^*$ follows (tautologically) from this, as does the required description of the flow line γ . \square

Thus, the flow lines of the gradient vector field associated to the action of T are given by the action of a subgroup of the larger group $T^{\mathbb{C}}$. *Conversely, whenever an 'algebraic torus' $\mathbb{C}^* \times \dots \times \mathbb{C}^*$ acts complex analytically on a compact Kähler manifold then we obtain both a Morse-Bott function and its gradient flow in the above manner.* Many manifolds do admit such actions, for example the Grassmannian $Gr_k(\mathbb{C}^n)$ (which is acted upon naturally by the group of diagonal matrices with nonzero complex diagonal entries). This provides a simple explanation of the rather tricky calculation of Lemma 3.1.2, by means of which we identified the gradient flow lines for the height functions. It suffices to assume that the action of $\mathbb{C}^* \times \dots \times \mathbb{C}^*$ preserves the complex structure, because an $S^1 \times \dots \times S^1$ -invariant Kähler metric can be obtained by averaging the given Kähler metric over the (compact) group $S^1 \times \dots \times S^1$.

We shall now address a question which lies at the heart of Morse theory: *how are the gradient flow lines of a Morse function $f: M \rightarrow \mathbb{R}$ arranged within M ?* For example, which pairs of critical points are connected by a flow line? And by how many flow lines? These questions can be answered by explicit computation of stable and unstable manifolds in the case of a height function on a Grassmannian, as we did in §3, but in general no such computation will be possible. Our main theme from now on will be to consider this question for Morse-Bott functions f^X associated to torus actions.

From the lemma it follows that the behaviour of the flow lines is related to the geometrical properties of the various orbits of the group $T^{\mathbb{C}}$. If M is an algebraic Kähler manifold, and the action of $T^{\mathbb{C}}$ is algebraic (as we shall assume from now on), the closures of such orbits are special algebraic varieties called toric varieties.⁵ In general toric varieties have singularities, but they are particularly amenable to study (see [50] and [31]) because they may be characterized by purely combinatorial objects, called 'fans'. Now, in the Kähler situation at least, the fan is equivalent to a more familiar combinatorial object, namely a convex polyhedron in Euclidean space. The next theorem describes this polyhedron.

Before stating the theorem, we need to introduce the moment map associated to the action of T on M . This is the map

$$\mu: M \rightarrow \mathfrak{t}^*$$

which is determined up to an additive constant by the condition

$$d\mu(\cdot)(Y) = \omega(Y^*, \cdot)$$

for all $Y \in \mathfrak{t}$. The definition of this map comes from symplectic geometry and classical mechanics, but it has a straightforward Morse-theoretic interpretation:

⁵This is essentially the definition of a toric variety.

$\mu(\cdot)(Y)$ is the Morse-Bott function f^Y associated to the sub-torus of T which is generated by Y .

The fact that our Morse-Bott function f^X is not alone, but is accompanied by a whole family of Morse-Bott functions f^Y parametrized by $Y \in \mathfrak{t}$, is significant. The moment map μ assembles these Morse-Bott functions into a single vector-valued function.

Theorem 4.1.3 ([2; 38]) *Let O_m denote the closure in M of the orbit of m under $T^{\mathbb{C}}$, i.e. $O_m = \overline{T^{\mathbb{C}} \cdot m}$. Then: (i) $\mu(O_m)$ is the convex hull of the finite set $\{\mu(m) \in \mathfrak{t}^* \mid m \text{ is a critical point of } f^X\}$, (ii) the inverse image (under μ) of each open face of $\mu(O_m)$ is a single $T^{\mathbb{C}}$ -orbit in O_m , and (iii) μ induces a homeomorphism $O_m/T \rightarrow \mu(O_m)$ (although the action of T on O_m is not necessarily free).*

Proof See Theorem 2 of [2]. □

The simplest nontrivial example of this theorem is given by the action of \mathbb{C}^* on $\mathbb{C}P^1 (\cong S^2)$ by $u \cdot [z_0; z_1] = [uz_0; z_1]$. The corresponding Morse-Bott function $f^X: S^2 \rightarrow \mathbb{R}$ is a height function, and there are two isolated critical points: the maximum point and the minimum point. We have $f^X = \mu$ in this situation, and $f^X(S^2)$ is obviously the line segment joining the maximum and minimum values.

We are specifically interested in the stable and unstable manifolds of f^X , and their intersections. Theorem 4.1.3 leads to the following information about these spaces. For simplicity we shall assume that M is actually a smooth projective variety, i.e. an algebraic submanifold of some complex projective space, with the induced Kähler structure. Furthermore we assume that $T^{\mathbb{C}}$ acts on M by projective transformations. From the method of section 3.5 we can then deduce that the closures of the stable and unstable manifolds of our Morse functions are irreducible algebraic subvarieties.

Theorem 4.1.4 ([46]) *Assume that M is a smooth projective variety, and that $T^{\mathbb{C}}$ acts on M by projective transformations. Let V be an irreducible algebraic subvariety of M which is preserved by the action of $T^{\mathbb{C}}$. Then $\mu(V)$ is the convex hull of the finite set $\{\mu(m) \in \mathfrak{t}^* \mid m \text{ is a critical point of } f^X \text{ in } V\}$.*

Sketch of the proof Let P_1, \dots, P_s be the distinct images under μ of the critical points of f^X which lie in V . (This is necessarily a finite set, as f^X is constant on any connected critical submanifold.) For any one-dimensional subalgebra $\mathbb{R}Y$ of \mathfrak{t} , the image of the continuous function $f^Y|_V$ is a closed finite interval in \mathbb{R} (as V is connected). Moreover, since $T^{\mathbb{C}}$ preserves V , the ends of this interval (i.e. the maximum and minimum values of f^Y) are of the form $f^Y(P_i), f^Y(P_j)$, for some i, j . But $f^Y = \pi_Y \circ \mu$, where $\pi_Y: \mathfrak{t}^* \rightarrow \mathbb{R}$ is given by evaluation at Y . It follows that $\mu(V)$ is contained in the convex hull of P_1, \dots, P_s .

Conversely, we must show that any point of this polyhedron is contained in $\mu(V)$. Let P_{i_1}, \dots, P_{i_t} denote the ‘exterior’ points of the polyhedron. For each

i_j , choose some $Y_j \in \mathfrak{t}$ such that the function $f^{Y_j}|_V$ has P_{i_j} as its absolute minimum value, occurring on a critical set V_j , where $V_j = V \cap M_j$ for some connected component M_j of the fixed point set of $T^{\mathbb{C}}$ on M . (This may be done by choosing a 'generic' Y_j such that P_{i_j} is the closest point of the polyhedron to the linear functional $\langle Y_j, \cdot \rangle$, where $\langle \cdot, \cdot \rangle$ is an invariant inner product on \mathfrak{t} .)

Let $S_j^V = V \cap S_j$, where S_j is the stable manifold of M_j (for f^{Y_j}). Since $T^{\mathbb{C}}$ preserves V , S_j^V is the stable 'manifold' of V_j (for $f^{Y_j}|_V$). Now, the stable manifold decomposition of M induces a decomposition of V . The closure (in V) of each piece of this decomposition is a subvariety of V , and precisely one such closure must be equal to V since V is irreducible. Denote this piece by $V \cap S$, where S is the corresponding stable manifold in M (for f^{Y_j}).

We claim that $S = S_j$. (This would be obvious if V were a smooth subvariety.) Assume that S and S_j are not equal, so that $S \cap S_j = \emptyset$. Then the closure of $V \cap S$ in V is disjoint from $V \cap S_j$, since the closure of S in M is disjoint from S_j . However this contradicts the defining property of $V \cap S$.

It follows that the complement of S_j in V is a subvariety of positive codimension. The complement of the intersection of all such S_j (for $j = 1, \dots, t$) is therefore also a subvariety of positive codimension; in particular this intersection is nonempty. For any point v of the intersection, we have $P_{i_1}, \dots, P_{i_t} \in \overline{T^{\mathbb{C}} \cdot v}$, because of the description of the flow lines as orbits of subgroups of $T^{\mathbb{C}}$. Hence $\mu(V)$, which contains $\mu(\overline{T^{\mathbb{C}} \cdot v})$, must contain the convex hull of P_1, \dots, P_s , by Theorem 4.1.3. \square

The case $V = M$ was stated and proved as one of the main theorems of [2],[38]. Although the proofs given there were direct, the possibility of deducing the result from Theorem 4.1.3 was in fact mentioned in [2]. Various special cases of this result (e.g. when V is a Schubert variety) were already known – we refer to [2],[38] for further information.

Theorem 4.1.4 suggests the possibility that the behaviour of the gradient flow lines of f may be encoded by combinatorial information within the polyhedron $\mu(M)$. We shall investigate this phenomenon, starting with the familiar case of height functions on Grassmannians from §3. Initial work in this direction was done in [33], where some of the results of [2], [38] were anticipated in the case of a Grassmannian.

Let $M = Gr_k(\mathbb{C}^n)$ and let T_n be the group of diagonal $n \times n$ matrices whose diagonal entries are complex numbers of unit length, and let $T_n^{\mathbb{C}}$ be its complexification, i.e. the group of diagonal $n \times n$ matrices with nonzero complex diagonal entries. We have a natural action of T_n (or $T_n^{\mathbb{C}}$) on M . The vector $X = \sqrt{-1}(a_1, \dots, a_n)$ in the Lie algebra \mathfrak{t}_n is a generator of T_n if and only if $\mathbb{R}(a_1, \dots, a_n) \cap \mathbb{Z}^n = \{0\}$, i.e. the line through (a_1, \dots, a_n) has 'irrational slope'.

From the above general theory, any choice of (a_1, \dots, a_n) gives rise to (1) an action of T (the sub-torus of T_n generated by $Y = \sqrt{-1}(a_1, \dots, a_n)$), and (2) a Morse-Bott function f^Y , whose critical points are the fixed points of T . From Lemma 3.1.2 we see that this function is (up to an additive constant) the

height function on $Gr_k(\mathbb{C}^n)$ defined in §3. The fixed points of T are of course the k -planes which can be spanned by eigenvalues of $\text{diag}(a_1, \dots, a_n)$.

From the formula for f^X in section 3.1, it follows that the moment map is given explicitly by

$$\mu(V) = \text{diagonal part of } \pi_V.$$

(We identify \mathfrak{t}^* with $\mathfrak{t} \cong \mathbb{R}^n$ by using the (restriction of the) standard Hermitian product on \mathbb{C}^n .) In particular,

$$\mu(V_{\mathbf{u}}) = e_{\mathbf{u}},$$

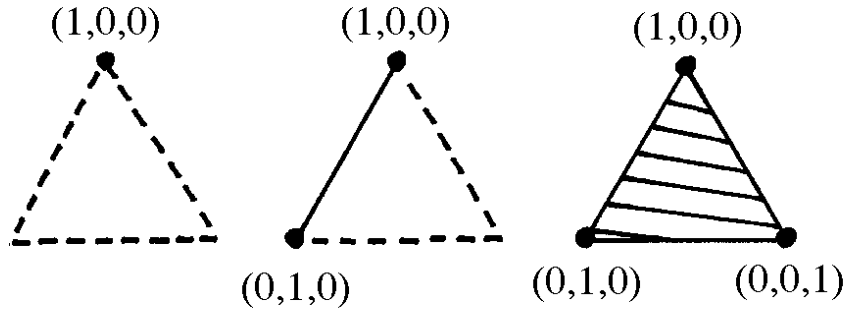
where $e_{\mathbf{u}} = e_{u_1} + \dots + e_{u_k}$ (and e_i denotes the i th basis vector of \mathbb{C}^n). Since every critical manifold of f^X contains at least one point of the form $V_{\mathbf{u}}$, the images of the critical points of f^X under μ are precisely the points $e_{\mathbf{u}}$. Hence the general theory gives

$$\mu(Gr_k(\mathbb{C}^n)) = \text{convex hull of } \{e_{\mathbf{u}} = e_{u_1} + \dots + e_{u_k} \mid 1 \leq u_1 < \dots < u_k \leq n\}.$$

Example 4.1.5 Let $k = 1$, so that $Gr_k(\mathbb{C}^n) = \mathbb{C}\mathbb{P}^{n-1}$. In this case the formula for f^X in section 3.1 shows that the moment map is given even more explicitly by

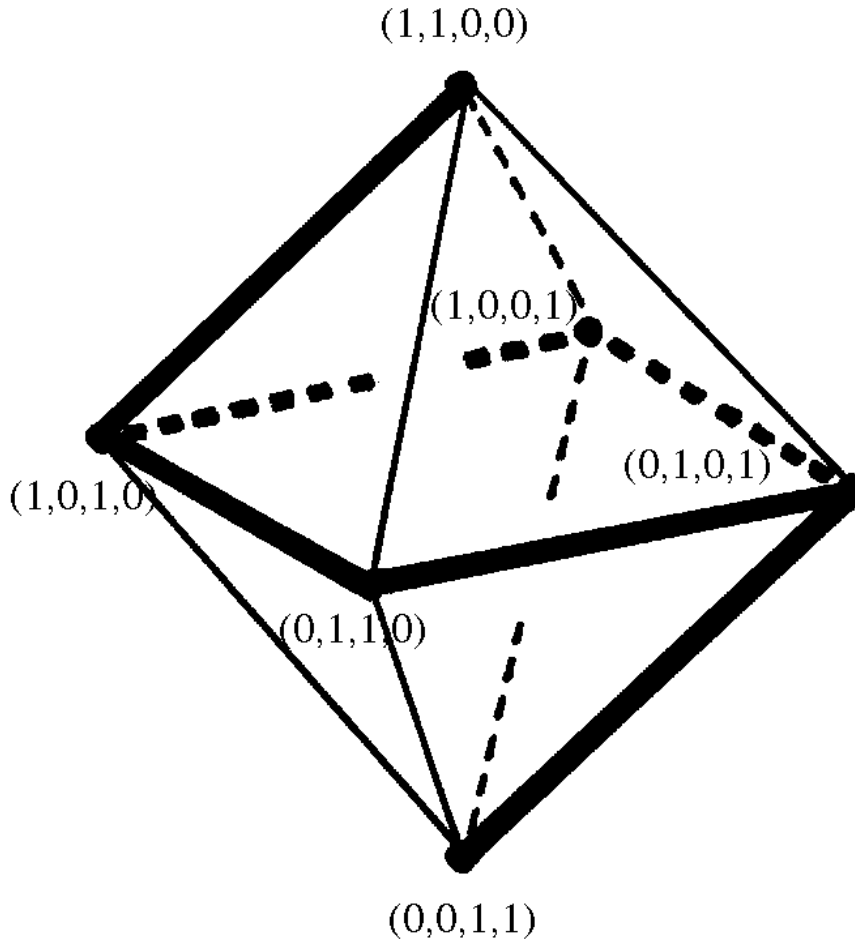
$$\mu([z]) = (|z_1|^2, \dots, |z_n|^2) / \sum_{i=1}^n |z_i|^2.$$

It follows immediately from this formula that $\mu(\mathbb{C}\mathbb{P}^{n-1})$ is the convex hull of the basis vectors e_1, \dots, e_n . Since the (closures of the) stable and unstable manifolds of f^X are respectively of the form $[*; \dots; *; 0; \dots; 0]$ and $[0; \dots; 0; *; \dots; *]$ (see section 3.1), their images under μ are also clear. For example, in the diagram below, we illustrate the images under μ of the closures of the stable manifolds of V_1, V_2, V_3 , for a Morse function $f^X: \mathbb{C}\mathbb{P}^2 \rightarrow \mathbb{R}$.



In fact, for $\mathbb{C}\mathbb{P}^{n-1}$, all assertions of Theorems 4.1.3 and 4.1.4 may be verified directly, from the formula for μ . \square

Example 4.1.6 Let us consider what the general theory says in the case of $Gr_2(\mathbb{C}^4)$. We have already investigated this space in some detail in §3 (starting with Example 3.4.1). The image of μ is the convex hull of the six points $e_i + e_j$ in \mathbb{R}^4 (with $1 \leq i < j \leq 4$). It may be verified that this polyhedron is a regular octahedron.



The heavy lines represent the partial order shown in the diagram following Definition 3.4.2. Combining the calculations of the stable manifolds in §3 with the statement of Theorem 4.1.4, we see that the image under μ of the closure of the stable manifold of a critical point $V_{\mathbf{u}}$ is the convex hull of those vertices which are greater than or equal to $e_{\mathbf{u}}$ in this partial order.

With a little more work, it is possible to verify the predictions of Theorem

4.1.3 in this case. First it is necessary to identify the various possible types of the closures of the orbits of the group $(\mathbb{C}^*)^4$, to which Theorem 4.1.3 will apply. Zero-dimensional orbits, i.e. points, correspond to the vertices of the octahedron. One-dimensional orbits (necessarily isomorphic to \mathbb{CP}^1) correspond to the edges. Two-dimensional orbits are of two types: copies of \mathbb{CP}^2 , which correspond to faces, and copies of $\mathbb{CP}^1 \times \mathbb{CP}^1$, which correspond to squares spanned by sets of four coplanar vertices. Finally, there are two types of three-dimensional orbits, whose images under μ are either half of the octahedron or the entire octahedron. The first type is represented by the Schubert variety $X_{(2,4)}$ (see section 3.4), and the second by the famous ‘tetrahedral complex’ (see [33]). \square

Returning to the general case of $Gr_k(\mathbb{C}^n)$, we mention that the partial order on the critical points of $f^X: Gr_k(\mathbb{C}^n) \rightarrow \mathbb{R}$ (or equivalently, on the vertices of the polyhedron) may be specified purely algebraically, in terms of the action of the symmetric group Σ_n (the Weyl group of $GL_n\mathbb{C}$). This is explained in [2].

In [34] and [32], a detailed study is made of the various sub-polyhedra of $\mu(Gr_k(\mathbb{C}^n))$ which arise from Schubert varieties associated to the Morse functions f^X and their intersections. These are characterized in terms of the combinatorial concept of a matroid.

By representing the stable and unstable manifolds of the function f^X as sub-polyhedra of the polyhedron $\mu(M)$, we have in principle solved the problem of understanding the behaviour of the flow lines of $-\nabla f^X$. The practical value of this solution depends on being able to extract this information in an efficient manner, and this may not always be possible. However, there are two general situations where reasonably explicit results may be expected, namely for generalized flag manifolds and for toric manifolds. The essential phenomenon in both these situations is that *there exists an orbit of the complex algebraic torus $T^{\mathbb{C}}$ whose closure contains all critical points of the Morse functions associated to the torus action*. We shall call this kind of $T^{\mathbb{C}}$ -action a *complete torus action*.

Example 4.1.7 A generalized flag manifold is by definition the quotient of a complex semisimple Lie group $G^{\mathbb{C}}$ by a parabolic subgroup P . Height functions on generalized flag manifolds – generalizing the Morse functions on Grassmannians in §3 – were first studied by Bott (see [11] and the article of Bott in [1]). Such functions are associated to torus actions (namely a maximal torus of $G^{\mathbb{C}}$) in the manner described at the beginning of this section. The index and, in the case of a Morse-Bott function, the nullity of a critical point may be computed in terms of the weights of the torus action (i.e. in terms of the roots of $G^{\mathbb{C}}$). Each stable or unstable manifold is an orbit of a certain subgroup of $G^{\mathbb{C}}$, as in the case $G^{\mathbb{C}}/P = Gr_k(\mathbb{C}^n)$. Indeed the decomposition of $G^{\mathbb{C}}/P$ into stable (or unstable) manifolds coincides with the well known ‘Bruhat decomposition’ of $G^{\mathbb{C}}/P$, a fact which was proved in [55] as well as in various later papers (e.g. [2]). A brief summary of this theory may be found in the Appendix of [37]. The image of the moment map for the action of a maximal torus on a generalized flag manifold was worked out in [2], generalizing earlier work of Kostant. The polyhedron

can be described (see [34]) as the convex hull of the *weights* of an irreducible representation of G ; it is well known that $G^{\mathbb{C}}/P$ is the projectivized orbit of the maximal weight vector of a suitable representation (the generalized Plücker embedding). Schubert varieties in generalized flag manifolds have been extensively studied from the point of view of Lie theory and algebraic geometry (see [39] for an introduction and further references). The sub-polyhedra obtained by taking the images of (various intersections of) Schubert varieties in generalized flag manifolds have been characterized in combinatorial terms in [34], generalizing the results mentioned earlier for Grassmannians. The homology classes represented by Schubert varieties have also been investigated thoroughly. In [7] these homology classes are related to the well known description of the cohomology ring due to Borel, by making use of the generalized Plücker embedding. A brief explanation of the latter work can be found in [59]. \square

Example 4.1.8 From the general theory of toric varieties, it is well known that a (smooth) toric variety with a Kähler metric is entirely determined by the image of its moment mapping. In particular, the behaviour of the flow lines of a Morse-Bott function associated to the given torus action is represented faithfully in the momentum polyhedron. Various geometrical and topological invariants of such manifolds have been computed explicitly in terms of this polyhedron; details can be found in [50] and [31] (see also [5] and [20]). There are relations between this and the previous example, as one may consider the toric varieties obtained as the closures of the orbits of a maximal torus of $G^{\mathbb{C}}$ acting on $G^{\mathbb{C}}/P$. These are singular varieties, in general; they have been studied in [34] and later in [23], [19], and [16]. \square

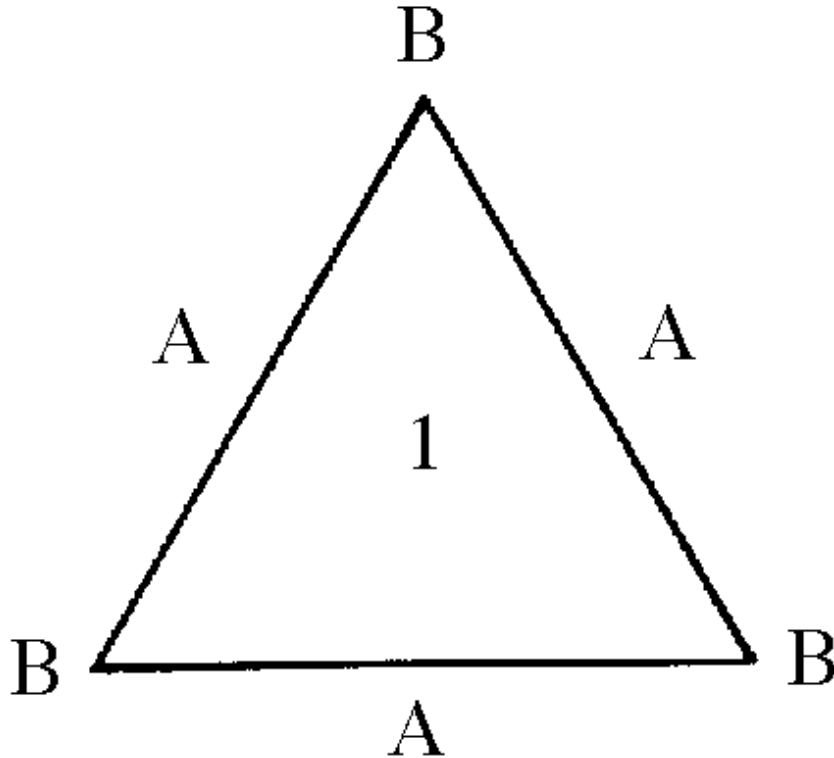
We shall now change our point of view slightly by focusing on the family of Morse-Bott functions f^Y (parametrized by $Y \in \mathfrak{t}$), rather than the single Morse function f^X (corresponding to a generator X of \mathfrak{t}). This theme will reach maturity in section 4.3, so as motivation for this we shall consider again the problem of computing the cohomology ring $H^*(Gr_2(\mathbb{C}^4))$ (cf. section 3.6).

Theorem 4.1.4 gives a representation of the (image under μ of the) Schubert variety $V = X_{\mathbf{u}}^Y$ for the Morse function $f^Y: Gr_k(\mathbb{C}^n) \rightarrow \mathbb{R}$. Namely, the image under μ of V is the convex hull of those points $e_{\mathbf{u}}$ such that $V_{\mathbf{u}} \in V$. This representation also applies to the (irreducible components of the) variety $V = X_{\mathbf{u}_1}^{Y_1} \cap \dots \cap X_{\mathbf{u}_k}^{Y_k}$. In section 3.6 we saw that the problem of calculating products in cohomology can in principle be reduced to the problem of calculating intersections of (pairs or) triples of Schubert varieties $X_{\mathbf{u}_1}^{Y_1} \cap X_{\mathbf{u}_2}^{Y_2} \cap X_{\mathbf{u}_3}^{Y_3}$. ‘In principle’ means⁶ ‘providing that all necessary triple intersections are transverse’.

⁶It can be shown that all intersections are indeed transverse, by using the fact that the stable manifolds are orbits of certain subgroups of $GL_n\mathbb{C}$; see [45]. This is not quite enough, for we must show in addition that there exist sufficiently many representatives of all cohomology classes. This will be obvious in the examples we consider, however.

To be more precise, we need to find all zero-dimensional triple intersections. By Theorem 4.1.4 (or by the Plücker embedding argument of section 3.6), such an intersection either consists of a single point or is empty. To determine which is the case, we just need to know which cohomology classes are represented by which sub-polyhedra of $\mu(Gr_k(\mathbb{C}^n))$. Let us consider two examples, $\mathbb{C}\mathbb{P}^2$ and $Gr_2(\mathbb{C}^4)$.

Example 4.1.9 The cohomology ring $H^*(\mathbb{C}\mathbb{P}^2)$ has additive generators in dimensions 0, 2, and 4; let us denote these respectively by 1, A , and B . They are dual to the fundamental homology classes of the Schubert varieties for a fixed Morse function f^Y . By varying Y in \mathfrak{t} , we arrive at the following representation of these cohomology classes on the triangular region $\mu(\mathbb{C}\mathbb{P}^2)$:

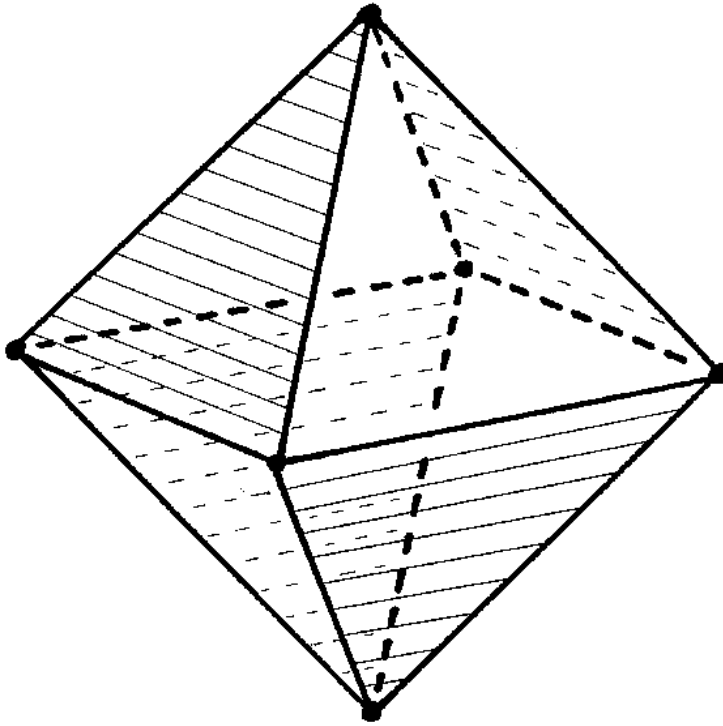


The structure of the cohomology ring $H^*(\mathbb{C}\mathbb{P}^2)$ is determined completely by the product A^2 , and from the diagram the intersection number of the dual Schubert varieties is 1. Hence $A^2 = 1B = B$, as expected. \square

Example 4.1.10 The cohomology ring $H^*(Gr_2(\mathbb{C}^4))$ has the six additive generators described in section 3.6. These are represented on the octahedron of Example 4.1.6 as follows:

- $z_{(3,4)} \in H^0$: the octahedron
- $z_{(2,4)} \in H^2$: the half octahedra
- $z_{(1,4)}, z_{(2,3)} \in H^4$: alternate faces
- $z_{(1,3)} \in H^6$: the edges
- $z_{(1,2)} \in H^8$: the vertices

In the diagram below, the (four) faces which represent the cohomology class $z_{(2,3)}$ are shaded.



It is now a simple matter to read off all zero-dimensional double and triple intersections. For example, $z_{(1,4)}z_{(2,4)}z_{(2,4)}$ is represented by the intersection of two half octahedra and a face. Any such intersection *giving a finite number*

of points gives precisely one point. So the product is equal to the generator of H^8 – as we found by a much more laborious calculation in Example 3.6.4. By contemplating the above diagram we can determine the entire cohomology ring of $Gr_2(\mathbb{C}^4)$! \square

To end this section, we emphasize two advantages of having a Morse-Bott function which is associated to a torus action. First, having a group action has computational advantages, as we have seen in the description of the gradient flow and the identification of the index (and nullity) of a critical point. Second, the torus action gives more than a single Morse-Bott function; it gives a whole family of related Morse-Bott functions, and this family can give more information than one of its members (as in the above calculation of the cohomology ring).

Additional comments (May, 2000): The theory of equivariant cohomology provides an algebraic explanation for the success of the above method of computing $H^*(M)$ from a (complete) torus action (see M. Goresky, R. Kottwitz, and R. MacPherson, *Equivariant cohomology, Koszul duality, and the localization theorem*, Invent. Math. 131 (1998), 25-83). The equivariant cohomology ring $H_T^*(M)$ (see section 2.7) is in this case isomorphic as a module over $H^*(BT)$ to the tensor product $H^*(M) \otimes H^*(BT)$. The localization theorem for equivariant cohomology expresses products of equivariant cohomology classes in terms of the fixed point data of the torus action (this is another manifestation of Morse theory). As a result, it is possible to describe $H^*(M)$ explicitly as a quotient of a polynomial ring, in terms of this data.

4.2 The Witten complex

In view of the importance of the gradient flow lines of a Morse function, it is perhaps not surprising that the basic theorems of Morse theory may be developed entirely from this point of view. In fact, much stronger results (than the traditional ‘Morse inequalities’) are possible, as we shall see in this section and the next one.

In this section we consider the goal of computing the homology groups of a manifold M . Traditionally, this is possible only for a *perfect* Morse function. However, if we assume that $f: M \rightarrow \mathbb{R}$ is a Morse-Smale function (i.e. the stable and unstable manifolds of f intersect transversely, as in Definition 1.4.5), then the homology may be calculated whether f is perfect or not. This method became widely understood only in the 1980’s, through the work of Witten and Floer (see [66] and [24]). It is easy to describe: one constructs a certain chain complex of abelian groups, the ‘Witten complex’, whose homology groups turn out to be the homology groups of M .

Let us assume first that M is oriented. The abelian groups are defined in terms of the critical points of the Morse-Smale function f by

$$C_i = \text{free abelian group on the set of critical points of index } i.$$

(Since f is a Morse function and M is compact, the groups C_i have finite rank.) The boundary operators $\partial_i: C_i \rightarrow C_{i-1}$ are defined in terms of the gradient flow

lines of $-\nabla f$ as follows. Let m be a critical point of f of index i (i.e. a generator of C_i). Then

$$\partial_i(m) = \sum_{\gamma} \epsilon(\gamma)m_{\gamma},$$

where (1) the sum is over all flow lines γ such that

$$\lim_{t \rightarrow -\infty} \gamma(t) = m, \quad \lim_{t \rightarrow \infty} \gamma(t) = (\text{a critical point}) m_{\gamma} \text{ of index } i - 1,$$

and (2) $\epsilon(\gamma)$ is either 1 or -1 , the choice depending on whether γ 'preserves or reverses orientation'.

Some explanation of (1) and (2) is necessary. First, since the Morse-Smale condition gives

$$\dim F(m, m_{\gamma}) = 1,$$

it follows that there are only finitely many such γ , hence the sum is finite. Second, to define the sign of $\epsilon(\gamma)$, we first choose arbitrary orientations of the unstable manifolds. Since M is oriented, and since the stable and unstable manifolds intersect transversely, we may then assign orientations to the stable manifolds in a consistent manner. The manifold $F(m, m_{\gamma})$ itself thus acquires an orientation. We define $\epsilon(\gamma)$ to be 1 if the natural orientation of γ agrees with its orientation as a component of $F(m, m_{\gamma})$; otherwise we define $\epsilon(\gamma)$ to be -1 .

It can be shown that (C_*, ∂_*) is a chain complex, i.e. that $\partial_{i-1} \circ \partial_i = 0$ for all i . For this, and for the proof of the next theorem, we refer to section 2 of [6], where a detailed discussion can be found.

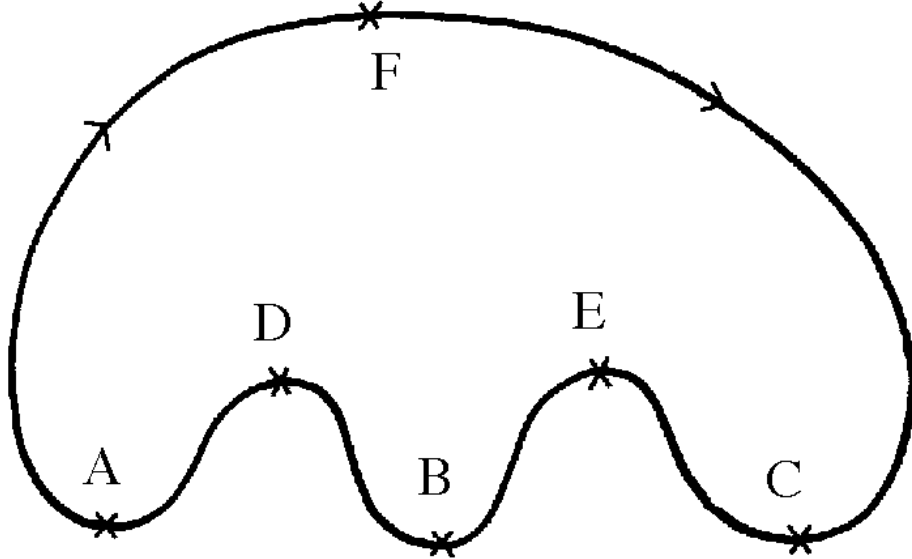
Theorem 4.2.1 *The homology groups of the chain complex (C_*, ∂_*) are isomorphic to the homology groups of M .* \square

If M is orientable, as in the above definition, then homology groups with coefficients in \mathbb{Z} are obtained. If M is not orientable, then some modifications to the definition are necessary. The simplest way to do this is to work over $\mathbb{Z}/2\mathbb{Z}$ instead of \mathbb{Z} , i.e. to replace C_i by $C_i \otimes \mathbb{Z}/2\mathbb{Z}$ and then to define $\epsilon(\gamma) = 1$ for all γ . In this case the theorem is true for homology groups with coefficients in $\mathbb{Z}/2\mathbb{Z}$.

It is possible to formulate and prove a similar theorem for the cohomology groups of M , using the dual chain complex (see [6]).

The Morse inequalities (Theorem 2.3.1) follow from the statement of the above theorem by a purely algebraic argument (using $\text{rank } H_i(M) = \text{rank Ker } \partial_i - \text{rank Im } \partial_{i+1}$ and $\text{rank } C_i = \text{rank Ker } \partial_i + \text{rank Im } \partial_i$). The 'lacunary principle', that f is necessarily perfect if all its critical points have even index, also follows immediately, since in this case we have $\partial_i = 0$ for all i .

Example 4.2.2 Consider the Morse function on the circle with three local maxima and three local minima depicted in Example 2.2.2.



This is of course not a perfect Morse function. But the homology groups may be calculated by using the Witten complex. If we choose the 'clockwise' orientation on S^1 and on all stable manifolds, then the map $\partial_1: C_1 \rightarrow C_0$ is given by:

$$\partial_1 F = -A + C, \quad \partial_1 D = A - B, \quad \partial_1 E = B - C.$$

The kernel and cokernel of ∂_1 are therefore both isomorphic to \mathbb{Z} . \square

Example 4.2.3 Consider the Morse function $f: \mathbb{R}\mathbb{P}^n \rightarrow \mathbb{R}$ defined by

$$f([x_0, \dots, x_n]) = \frac{\sum_{i=0}^n c_i |x_i|^2}{\sum_{i=0}^n |x_i|^2}$$

with $c_0 > \dots > c_n$, which we met in Example 2.3.6. Let us calculate the homology of $\mathbb{R}\mathbb{P}^n$ by using the Witten complex. First we take coefficients in $\mathbb{Z}/2\mathbb{Z}$, to avoid the problem of dealing with orientations.

The critical points are the coordinate axes V_0, \dots, V_n , and the indices are (respectively) $n, n-1, \dots, 0$. Thus $C_i = \mathbb{Z}/2\mathbb{Z} V_{n-i}$ for $0 \leq i \leq n$. As in the case of $\mathbb{C}\mathbb{P}^n$ in §3, we may identify the stable and unstable manifolds explicitly. In particular, we see that the space $F(V_i, V_{i+1})$ of points on flow lines from V_i to V_{i+1} is of the form

$$\{[0; \dots; 0; *, *, 0; \dots; 0] \in \mathbb{R}\mathbb{P}^n \mid * \in \mathbb{R}\}.$$

In other words, it is a copy of $\mathbb{R}\mathbb{P}^1 \cong S^1$, and so there are precisely two such flow lines. Thus, every homomorphism ∂_i is zero, and our Morse function is perfect.

With integer coefficients the situation is more complicated, particularly since we have not defined the Witten complex (over \mathbb{Z}) for a nonorientable manifold, and it is well known that $\mathbb{R}\mathbb{P}^n$ is orientable only when n is odd. However, the Witten complex can in fact be defined for any manifold (see section 2.1 of [45]), and it turns out that for $\mathbb{R}\mathbb{P}^n$ the maps ∂_i are given by

$$\partial_i(n) = \begin{cases} 2n & \text{if } i \text{ is even} \\ 0 & \text{if } i \text{ is odd} \end{cases}$$

This gives the integral homology groups of $\mathbb{R}\mathbb{P}^n$. □

Example 4.2.4 Any function associated to the standard torus action on a generalized flag manifold (see Example 4.1.7) has trivial Witten complex, since all critical points have even index. However, the real analogues of these complex manifolds (which include $\mathbb{R}\mathbb{P}^n$ and more generally the real Grassmannians) give rise to nontrivial Witten complexes, and it may be expected that these complexes are determined by the same combinatorial information that describes the behaviour of the flow lines. This is indeed the case; full details may be found in [45]. Extensive earlier work on the Morse theory of these spaces can be found in [14], [61], [62], [26], [22]. The convexity results of [2], [38] were extended to these spaces in [21]. □

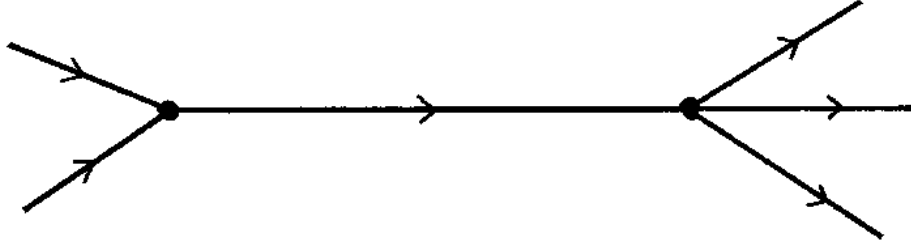
The above description of the Witten complex, phrased in the language of differential topology, is closely related to earlier work of Smale and Thom (see [27], [48], and [60]), in which the groups C_i appear as the relative homology groups of the pair (M_i, M_{i-1}) , for a suitable filtration $\{M_i\}$ of M , and the maps ∂_i appear as the connecting homomorphisms in the homology exact sequence. (The space M_i is obtained by taking all cells of dimension less than or equal to i in the usual Morse decomposition of M .) Witten's original motivation was actually quite different, as it arose from quantum theory. A brief description of Witten's point of view, together with further historical information, can be found in [13].

Finally we mention that an extension of the theory to the case of Morse-Bott functions is given in [6]. Another reference where full details of the material of this section can be found is the book [58].

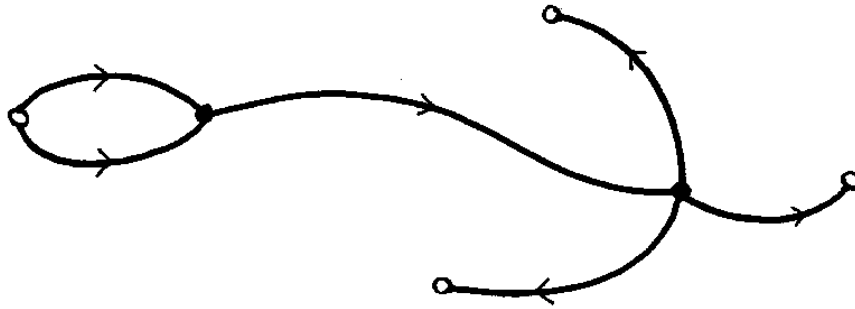
4.3 Morse theory as a topological field theory

In section 4.1 we have seen that it can be useful to study *families* of Morse functions on a given compact manifold M . This can be taken as motivation for the 'field-theoretic' approach to Morse theory of [8], [9], [17], [18], [28; 29], so called because it is based on similar constructions in gauge theory. As such constructions tend to involve rather elaborate preparations, we shall just give an informal description here.

The basic ingredient is a certain ‘moduli space’ $\mathcal{M}(\Gamma)$, which is a device for counting configurations of flow lines. The definition of this space depends on M and on a choice of an oriented connected graph Γ . We assume that Γ has n_1 edges parametrized by $(-\infty, 0]$ (‘incoming edges’), n_2 edges parametrized by $[0, 1]$ (‘internal edges’), and n_3 edges parametrized by $[0, \infty)$ (‘outgoing edges’). In the example below, we have $n_1 = 2$, $n_2 = 1$, and $n_3 = 3$.



An element of $\mathcal{M}(\Gamma)$ is a ‘configuration of flow lines of Morse-Smale functions on M , modelled on Γ ’, i.e. a continuous map $\mathcal{F}: \Gamma \rightarrow M$ such that, on each edge of Γ , \mathcal{F} is (part of) a flow line of a Morse-Smale function on M . Thus, for the graph illustrated above, $\mathcal{F}(\Gamma)$ might look like the diagram below, where the white dots are critical points approached by the incoming and outgoing flow lines.



Let $f_1^{\mathcal{F}}, \dots, f_{n_1+n_2+n_3}^{\mathcal{F}}$ be the Morse-Smale functions in the definition of \mathcal{F} , and let $a_1^{\mathcal{F}}, \dots, a_{n_1}^{\mathcal{F}}, a_{n_1+n_2+1}^{\mathcal{F}}, \dots, a_{n_1+n_2+n_3}^{\mathcal{F}}$ be the critical points which are approached by the incoming and outgoing flow lines in that definition (the white dots in the diagram). For any $(n_1 + n_2 + n_3)$ -tuple of functions $g = (g_1, \dots, g_{n_1+n_2+n_3})$ on M , we define

$$\mathcal{M}_g(\Gamma) = \{\mathcal{F} \in \mathcal{M}(\Gamma) \mid f_i^{\mathcal{F}} = g_i\}.$$

For any $(n_1 + n_3)$ -tuple of points $b = (b_1, \dots, b_{n_1}, b_{n_1+n_3+1}, \dots, b_{n_1+n_2+n_3})$ of M , we define

$$\mathcal{M}_g(\Gamma; b) = \{\mathcal{F} \in \mathcal{M}_g(\Gamma) \mid a_i^{\mathcal{F}} = b_i\}.$$

(These definitions are informal versions of the precise definitions in [9].)

We shall assume from now on that all stable and unstable manifolds of $f_1^{\mathcal{F}}, \dots, f_{n_1+n_2+n_3}^{\mathcal{F}}$ intersect transversely; in particular all these functions are Morse-Smale functions. Under this assumption, it can be shown that $\mathcal{M}_g(\Gamma; b)$ is a smooth manifold, that $\mathcal{M}_g(\Gamma; b)$ is oriented if M is oriented, and that

$$\dim \mathcal{M}_g(\Gamma; b) = \sum_{i=1}^{n_1} \text{index } b_i - \sum_{i=1}^{n_3} \text{index } b_{n_1+n_2+i} - (\dim M)(\dim H_1(\Gamma; \mathbb{R}) + n_1 - 1).$$

Example 4.3.1 Let Γ be the graph below with $n_1 = n_3 = 1$, $n_2 = 0$:

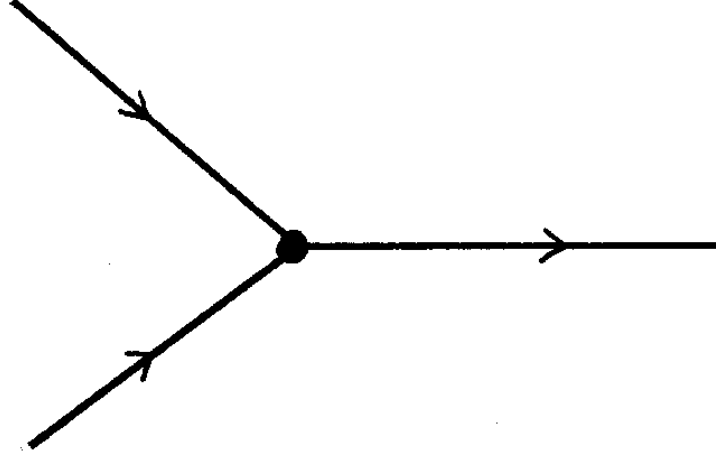


Let $g = (g_1, g_2)$, $b = (b_1, b_2)$. Then the points of $\mathcal{M}_g(\Gamma; b)$ are in one-to-one correspondence with the points of $U_{b_1}^{g_1} \cap S_{b_2}^{g_2}$, where $U_{b_i}^{g_i}$ is the unstable manifold of the critical point b_i of g_i , and $S_{b_i}^{g_i}$ is the stable manifold. From the transversality assumption we have (see section 1.4)

$$\text{codim } U_{b_1}^{g_1} \cap S_{b_2}^{g_2} = (\dim M - \text{index } b_1) + (\text{index } b_2).$$

This checks with the general formula above, i.e. $\dim \mathcal{M}_g(\Gamma; b) = \text{index } b_1 - \text{index } b_2$. □

Example 4.3.2 Let Γ be the graph below with $n_1 = 2$, $n_2 = 0$, and $n_3 = 1$:



Let $g = (g_1, g_2, g_3)$, $b = (b_1, b_2, b_3)$. In this situation the points of $\mathcal{M}_g(\Gamma; b)$ correspond to points of $U_{b_1}^{g_1} \cap U_{b_2}^{g_2} \cap S_{b_3}^{g_3}$. Transversality implies that

$$\text{codim } U_{b_1}^{g_1} \cap U_{b_2}^{g_2} \cap S_{b_3}^{g_3} = (\dim M - \text{index } b_1) + (\dim M - \text{index } b_2) + (\text{index } b_3).$$

Again this is consistent with the general formula. \square

We come now to the main part of the construction, which will associate to each graph Γ a topological invariant of M . This will make use of the Witten complex $(C_*(f), \partial_*)$ of a (Morse-Smale) function f , which was defined in the last section.

For a fixed choice of g (as above), we define

$$q_g(\Gamma) = \sum_{\mathcal{F}, b} e(\mathcal{F}) b_1 \otimes \dots \otimes b_{n_1} \otimes b_{n_1+n_2+1} \otimes \dots \otimes b_{n_1+n_2+n_3},$$

where the sum is over all \mathcal{F}, b such that $\mathcal{F} \in \mathcal{M}_g(\Gamma; b)$ and $\dim \mathcal{M}_g(\Gamma; b) = 0$. If M is oriented, so is the zero-dimensional manifold $\mathcal{M}_g(\Gamma; b)$, and $e(\mathcal{F})$ is plus or minus one, according to the orientation of \mathcal{F} as a point of $\mathcal{M}_g(\Gamma; b)$.

Thus, if M is oriented, we may regard $q_g(\Gamma)$ as an element of

$$\left(\bigotimes_{i=1}^{n_1} C_*(g_i) \right) \otimes \left(\bigotimes_{i=n_1+n_2+1}^{n_1+n_2+n_3} C^*(g_i) \right),$$

where $C^*(g_i)$ is the dual complex to $C_*(g_i)$. If M is not oriented, then the definition of $\epsilon(\mathcal{F})$ must be modified in the same way as the Witten complex.

The construction of $q_g(\Gamma)$ depends on the choice of the $(n_1 + n_2 + n_3)$ -tuple of Morse-Smale functions g , of course. Nevertheless, it turns out that this choice, and all other choices necessary for a rigorous definition of $q_g(\Gamma)$, are irrelevant:

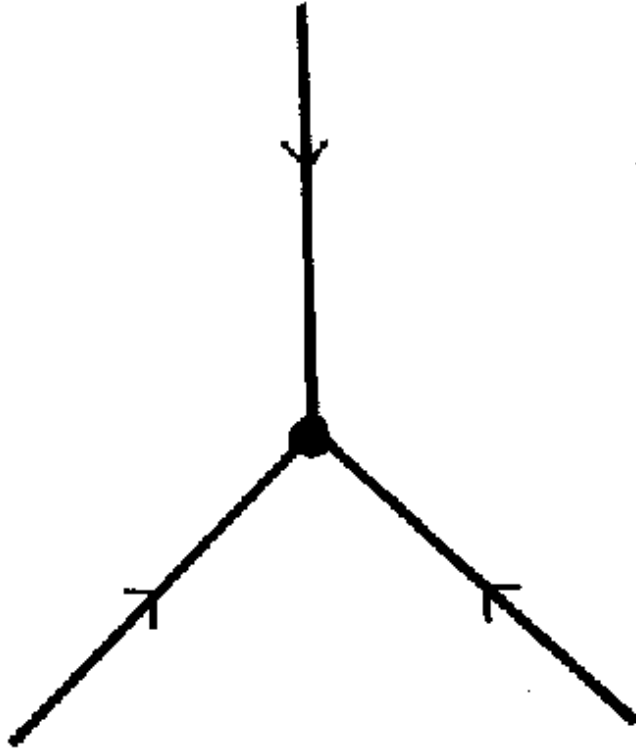
Theorem 4.3.3 ([8; 9; 28]) *The element $q_g(\Gamma)$ is annihilated by the (appropriate extension of) ∂_* , and so we obtain a class*

$$[q_g(\Gamma)] \in \left(\bigotimes_{i=1}^{n_1} H_*(M) \right) \otimes \left(\bigotimes_{i=n_1+n_2+1}^{n_1+n_2+n_3} H^*(M) \right).$$

This element depends only on the graph Γ . □

We denote the class $[q_g(\Gamma)]$ by $q(\Gamma)$. We may regard $q(\Gamma)$ as an element of the ring $\text{Hom}(\bigoplus H^*(M), \bigoplus H^*(M))$, i.e. as an *operation* on cohomology classes. Various well known operations can be obtained this way, by choosing suitable graphs. We shall discuss the case of the (cup) product operation in the following example.

Example 4.3.4 We choose the graph illustrated below:



If $\mathcal{M}_g(\Gamma; b) (\cong U_{b_1}^{g_1} \cap U_{b_2}^{g_2} \cap U_{b_3}^{g_3})$ is nonempty and zero-dimensional, then by transversality we must have $\sum_{i=1}^3 (\dim M - \text{index } b_i) = \dim M$. We obtain

$$q(\Gamma) \in \bigoplus_{i_1+i_2+i_3=\dim M} H_{\dim M-i_1}(M) \otimes H_{\dim M-i_2}(M) \otimes H_{\dim M-i_3}(M),$$

and hence also

$$q(\Gamma) \in \bigoplus_{i_1+i_2+i_3=\dim M} H^{i_1}(M)^* \otimes H^{i_2}(M)^* \otimes H^{i_3}(M)^*.$$

From our calculations of the triple product operation of $H^*(Gr_k(\mathbb{C}^n))$ in section 3.6 and the end of section 4.1, it is clear that $q(\Gamma)$ must be precisely that operation. This works for any orientable manifold M , because the main ingredient used in our calculation for $Gr_k(\mathbb{C}^n)$ was the existence of three functions, all of whose stable and unstable manifolds intersect transversely. In the general case, however, it is not easy to identify such ‘generic’ functions explicitly. The special feature of $Gr_k(\mathbb{C}^n)$, and indeed of any manifold with a complete torus action (as defined before Example 4.1.7), is that one only needs generic elements of the Lie algebra of the torus, and the existence of these is guaranteed by the convexity theorem. \square

The product operation has been described by other authors in terms of the Witten complex – see [6] and [65].

In all our examples so far we had $n_2 = 0$, and the space $\mathcal{M}_g(\Gamma; b)$ was identified with a subspace of M itself. In general, $\mathcal{M}_g(\Gamma; b)$ may be identified with a subspace of the $(n_2 + 1)$ -fold product $M \times \dots \times M$.

The theory described in this section makes use only of those $\mathcal{M}_g(\Gamma; b)$ which are zero-dimensional. In [17], [18] the higher dimensional spaces are used to construct a much more complicated algebraic object, which gives correspondingly more topological information.

4.4 Origins and other directions

The study of critical points of functions on infinite-dimensional spaces (e.g. on function spaces – the Calculus of Variations) has been a guiding principle right from the beginning of Morse theory. Rather surprisingly, perhaps, the development of almost all of Morse theory has been prompted by infinite-dimensional examples! Since the infinite-dimensional theory is much more complicated, it is usually presented as a generalization of traditional Morse theory, and we shall continue this tradition by giving only a brief list of examples, almost as an afterthought. Nevertheless, it is very likely that future directions in Morse theory will be strongly influenced by examples like these.

One of the earliest examples was the study of geodesics as critical points of the length or energy functional on the space of paths on a Riemannian manifold. Morse’s idea of using ‘broken geodesics’ to reduce the problem to a finite-dimensional problem is described in detail in [47]. Subsequently, a general theory of Morse functions on Hilbert manifolds was developed by Palais and Smale (see [51], [52]), under the assumption of ‘Condition (C)’. This condition, a substitute for compactness, is satisfied in the case of the geodesic problem, but unfortunately not in the case of many other important examples. For example, it is not satisfied in the case of the energy functional on the space of maps from a

Riemann surface into a Riemannian manifold. In this case the critical points are the harmonic maps, which are closely related to minimal surfaces.

For the Yang-Mills functional on the space of connections over a Riemann surface, a substitute for compactness was found by Atiyah and Bott ([4]). This led to new developments in finite-dimensional Morse theory, namely for the functional $|\mu|^2$ where μ is the moment map for the action of a (not necessarily abelian) Lie group on a manifold M . In [42], Kirwan showed that a version of Morse theory holds for this function, even though it is not a Morse-Bott function.

A general approach to Morse theory for the Yang-Mills functional (and other functionals, such as the energy functional on maps of Riemann surfaces) has been given by Taubes (see [63] and also the survey article [64] of Uhlenbeck).

All these examples focus on the critical points of a functional, but (as we have seen) it is possible to take a different point of view by focusing on the flow lines. Floer's idea of restricting attention to certain well-behaved flow lines is such a case, and this was also motivated by an infinite-dimensional example – the 'area functional' on closed paths in a symplectic manifold X . The flow lines in this case have particular geometrical significance: they may be regarded as 'holomorphic curves' in X , where X is given an almost complex structure compatible with its symplectic structure. Such holomorphic curves arose in earlier work of Gromov, and the homology theory which is computed by the Witten complex in this situation is known as Gromov-Floer theory. (It is also closely related to quantum cohomology theory.)

Floer applied his theory to an apparently quite different example, the Chern-Simons functional on the space of connections on certain three-dimensional manifolds. Again the flow lines have a geometrical meaning: they are Yang-Mills instantons. The homology theory arising here is called Floer homology.

References

1. M.F. Atiyah (ed.): 'Representation Theory of Lie Groups', Cambridge Univ. Press, Lond. Math. Soc. Lecture Notes 34, 1979.
2. M.F. Atiyah: *Convexity and commuting Hamiltonians*, Bull. Lond. Math. Soc. **16** (1982), 1–15.
3. M.F. Atiyah: *Topological quantum field theory*, Publ. Math. IHES **68** (1989), 175–186.
4. M.F. Atiyah and R. Bott: *The Yang-Mills equations over Riemann surfaces*, Phil. Trans. R. Soc. Lond. **A308** (1982), 523–615.
5. M. Audin : 'The Topology of Torus Actions on Symplectic Manifolds ', Birkhäuser, 1991.
6. D.M. Austin and P.J. Braam : *Morse-Bott theory and equivariant cohomology*, in 'The Floer Memorial Volume' (eds:) H. Hofer, C.H. Taubes, A. Weinstein, and E. Zehnder, Progress in Math. 133, Birkhäuser, 1995, pp. 123–183.
7. I.N. Bernstein, I.M. Gelfand and S.I. Gelfand: *Schubert cells and cohomology of the spaces G/P* , Russian Math. Surveys **28** (1973), 1–26.
8. M. Betz: *Categorical constructions in Morse theory and cohomology operations*, Ph.D. thesis, Stanford University, 1993.

9. M. Betz and R.L. Cohen: *Graph moduli spaces and cohomology operations*, Turkish J. of Math. **18** (1994), 23–41.
10. R. Bott: *Nondegenerate critical manifolds*, Annals of Math. **60** (1954), 248–261.
11. R. Bott: *An application of the Morse theory to the topology of Lie groups*, Bull. Soc. Math. France **84** (1956), 251–281.
12. R. Bott: *Lectures on Morse theory, old and new*, Bull. Amer. Math. Soc. **7** (1982), 331–358.
13. R. Bott: *Morse theory indomitable*, Publ. Math. IHES **68**, (1989), 99–114.
14. R. Bott and H. Samelson: *An application of the theory of Morse to symmetric spaces*, Amer. J. Math. **80** (1958), 964–1029.
15. R. Bott and L.W. Tu: ‘Differential Forms in Algebraic Topology’, Springer GTM **82**, 1982.
16. J.B. Carrell and A. Kurth: *Normality of torus orbits in G/P* , preprint.
17. R.L. Cohen, J.D.S. Jones, and G.B. Segal: *Morse theory and classifying spaces*, preprint, 1991.
18. R.L. Cohen, J.D.S. Jones, and G.B. Segal: *Floer’s infinite dimensional Morse theory and homotopy theory*, in ‘The Floer Memorial Volume’ (eds:) H. Hofer, C.H. Taubes, A. Weinstein, and E. Zehnder, Progress in Math. **133**, Birkhäuser, 1995, pp. 297–325 (also published in Surikaiseikikenkyujo Kokyuroku **883**, Kyoto, 1994, 68–96).
19. A. Dabrowski: *On normality of the closure of a generic torus orbit in G/P* , Pacific J. Math. **172** (1996), 321–330.
20. T. Delzant: *Hamiltoniens périodiques et image convexe de l’application moment*, Bull. Soc. Math. France **116** (1988), 315–339.
21. J.J. Duistermaat: *Convexity and tightness for restrictions of Hamiltonian functions to fixed point sets of an anti-symplectic involution*, Trans. Amer. Math. Soc. **275** (1983), 417–429.
22. J.J. Duistermaat, J.A.C. Kolk, and V.S. Varadarajan: *Functions, flows and oscillatory integrals on flag manifolds and conjugacy classes in real semisimple Lie groups*, Comp. Math. **49** (1983), 309–398.
23. H. Flaschka and L. Haine: *Torus orbits in G/P* , Pacific J. Math. **149** (1991), 251–292.
24. A. Floer: *Witten’s complex and infinite dimensional Morse theory*, J. Diff. Geom. **30** (1989), 207–221.
25. T. Frankel: *Fixed points and torsion on Kähler manifolds*, Annals of Math. **70** (1959), 1–8.
26. T. Frankel: *Critical submanifolds of the classical groups and Stiefel manifolds*, in ‘Differential and Combinatorial Topology’ (ed:) S.S. Cairns, Princeton, 1965, pp. 37–53.
27. J. Franks: *Morse-Smale flows and homotopy theory*, Topology **18** (1979), 199–215.
28. K. Fukaya: *Topological field theory and Morse theory*, Sugaku Expositions (translation from Suugaku, **46** (1994), 289–307) **10** (1997), 19–39.
29. K. Fukaya (notes by P. Seidel): *Floer homology, A_∞ -categories and topological field theory*, in ‘Geometry and Physics’ Lecture Notes in Pure and Applied Math. **184**, Marcel Dekker, (eds:) J.E. Anderson, J. Dupont, H. Pederson, A. Swann, 1997, pp. 9–32.
30. W. Fulton: ‘Young Tableaux’, Cambridge Univ. Press, 1997.
31. W. Fulton: ‘Introduction to Toric Varieties’, Annals of Math. Stud. **131**, Princeton

- Univ. Press, 1993.
32. I.M. Gelfand, M. Goresky, R.D. MacPherson and V.V. Serganova : *Combinatorial geometries, convex polyhedra, and Schubert cells*, Adv. Math. **63** (1987), 301–316.
 33. I.M. Gelfand and R.D. MacPherson: *Geometry in Grassmannians and a generalization of the dilogarithm*, Adv. Math. **44** (1982), 279–312.
 34. I.M. Gelfand and V.V. Serganova: *Combinatorial geometries and torus strata on homogeneous compact manifolds*, Usp. Mat. Nauk. **42**, (1987), 107–133 (Russian Math. Surveys 42 (1987), 133–168).
 35. M. Goresky and R. MacPherson: ‘Stratified Morse Theory’, Springer, 1988.
 36. P. Griffiths and J. Harris: ‘Principles of Algebraic Geometry’, Wiley, 1978.
 37. M.A. Guest and Y. Ohnita: *Group actions and deformations for harmonic maps*, J. Math. Soc. Japan. **45** (1993), 671–704.
 38. V. Guillemin and S. Sternberg: *Convexity properties of the moment mapping*, Invent. Math. **67** (1982), 491–513.
 39. H. Hiller: ‘Geometry of Coxeter Groups’, Research Notes in Math. 54, Pitman, 1982.
 40. M.W. Hirsch: ‘Differential Topology’, GTM 33, Springer, 1976.
 41. D.W. Kahn: ‘Introduction to Global Analysis’, Academic Press, 1980.
 42. F.C. Kirwan: ‘Cohomology of Quotients in Symplectic and Algebraic Geometry’, Math. Notes 31, Princeton Univ. Press, 1986.
 43. S.L. Kleiman and D. Laksov: *Schubert calculus*, Amer. Math. Monthly **79** (1972), 1061–1082.
 44. S. Kobayashi: *Fixed points of isometries*, Nagoya Math. J. **13** (1958), 63–68.
 45. R.R. Kocherlakota: *Integral homology of real flag manifolds and loop spaces of symmetric spaces*, Advances in Math. **110** (1995), 1–46.
 46. W. Liu, Ph.D. thesis, University of Rochester, 1998.
 47. J. Milnor: ‘Morse Theory’, Annals of Math. Studies 51, Princeton Univ. Press, 1963.
 48. J. Milnor: ‘Lectures on the h -cobordism Theorem’, Princeton Univ. Press, Math. Notes 1, 1965.
 49. J. Milnor and J.D. Stasheff: ‘Characteristic Classes’, Annals of Math. Studies 76, Princeton Univ. Press, 1974.
 50. T. Oda: ‘Convex Bodies and Algebraic Geometry: ‘An Introduction to the Theory of Toric Varieties’, Springer, 1988.
 51. R.S. Palais: *Morse theory on Hilbert manifolds*, Topology **2** (1963), 299–340.
 52. R.S. Palais: ‘Foundations of Global Non-linear Analysis’, Benjamin, 1968.
 53. R.S. Palais and C.-L. Terng: ‘Critical Point Theory and Submanifold Geometry’, Lecture Notes in Math. 1353, Springer, 1988.
 54. J. Palis and W. de Melo: ‘Geometric Theory of Dynamical Systems’, Springer, 1982.
 55. G.D. Parker: *Morse theory on Kähler homogeneous spaces*, Proc. Amer. Math. Soc. **34** (1972), 586–590.
 56. E. Pitcher: *Inequalities of critical point theory*, Bull. Amer. Math. Soc. **64** (1958), 1–30.
 57. A.N. Pressley: *The energy flow on the loop space of a compact Lie group*, J. Lond. Math. Soc. **26** (1982), 557–566.
 58. M. Schwarz: ‘Morse Homology’, Progress in Math. 111’, Birkhäuser, 1993.

59. G. Segal: *An introduction to the paper Schubert cells and cohomology of the spaces G/P* , in 'Representation Theory', Lond. Math. Soc. Lecture Notes 69 (eds:) I.M. Gelfand et al., Cambridge Univ. Press, 1982, pp. 111–114.
60. S. Smale: *Morse inequalities for dynamical systems*, Bull. Amer. Math. Soc. **66** (1960), 43–49.
61. M. Takeuchi: *Cell decompositions and Morse inequalities on certain symmetric spaces*, J. Fac. Sci. Univ. Tokyo **12** (1965), 81–192.
62. M. Takeuchi and S. Kobayashi: *Minimal embeddings of R -spaces*, J. Diff. Geom. **2** (1968), 203–215.
63. C.H. Taubes: *A framework for Morse theory for the Yang-Mills functional*, Invent. Math. **94** (1988), 327–402.
64. K. Uhlenbeck: *Applications of non-linear analysis in topology*, in 'Proc. Int. Congress of Math. I Math. Soc. Japan/Springer, 1991, pp. 261–279.
65. C. Viterbo: *The cup-product on the Thom-Smale-Witten complex, and Floer cohomology*, in 'The Floer Memorial Volume' (eds:) H. Hofer, C.H. Taubes, A. Weinstein, and E. Zehnder, Progress in Math. 133, Birkhäuser, 1995, pp. 609–625.
66. E. Witten: *Supersymmetry and Morse theory*, J. Diff. Geom. **17** (1982), 661–692.