Smooth and Discrete Morse Theory

Dominic Weiller

October 2015

A thesis submitted for the degree of Bachelor of Philosophy (Honours)
of the Australian National University

Acknowledgements

First I must thank my supervisor, Joan Licata. Her time spent explaining concepts has been invauable in helping me understand Morse theory, and her keen eye for, and patience with, grammatical errors has been truely appreciated. I'd like to thank Brett Parker for helping Joan first introduce me to Morse theory as well as David Smyth, Scott Morrison, Tony Licata, Vigliek Angeltveit, and Bryan Wang for teaching me so well over the last two years.

I'd like to thank those who shared the basement with me this year, particularly Jack, Suo Jun, Mark, Gino, and even the defector Eloise — who left our cosy basement for an office with a window — for the great company and conversation this year.

Finally I'd like the thank my family for always supporting me and leaving me dinner when I might arrive home from university after they'd already gone to bed.

Preface

Morse theory is a powerful tool that uses generic functions on a manifold to provide information about the manifold. Morse theory has been used to prove important results in the study of smooth manifolds such as the h-cobordism theorem and generalised Poincaré conjecture for dimensions greater than four [22, 26]. Recently, the ideas of Morse theory have been adapted for use with CW complexes [12]. This provides a tool for analysing these complexes, and in particular, a secondary way to study smooth manifolds, by first giving them a cell structure.

Chapter I reviews the basics of smooth Morse theory that will be needed for later chapters. We predominantly follow the approach of Milnor [22] and Nicolaescu [23], but with a few arguments and motivations shown more explicitly. The argument showing the weak Morse inequalities at the end of the chapter is original, though we suspect a similar argument must have existed previously.

Chapter II highlights the utility that being able to find Morse functions with specific properties can have, through a relatively simple classification of surfaces. We follow the approach of Champanerkar, Kumar, and Kumaresan [11] but fix an error in their proof; this is given by Lemma II.1.3. We also discuss the effect that modifying a Morse function on a surface has on the structure of sublevel sets; this discussion is inspired by the ideas in [11] but addresses issues not covered there.

The latter chapters are predominately focused on discussing discrete Morse theory and its relationship with smooth Morse theory. Chapter III introduces the theory discrete Morse functions developed by Robin Forman [12] in the 1990s. This theory extends many ideas from smooth Morse theory to the context of CW complexes. Theorem III.2.4 and the following corollary provide an original description of a process using a discrete Morse gradient to collapse a CW complex. This is used implicitly in the literature, though we could not find a similar explanation of this technique.

Section 4 of Chapter III discusses some similarities and differences in the collections of Morse or discrete Morse functions that an object supports. In particular we present an original argument that all discrete Morse functions on a CW complex are homotopic to one another, through a paths of discrete Morse functions; the analogous statement is not true in the smooth case.

Particularly interesting is the ability to take any smooth Morse function and induce a discrete Morse function with similar properties by triangulating the manifold; this is the content of Chapter IV. We discuss two methods for doing this, one due to Benedetti [5] and the other by Gallais [13]. While presenting Benedetti's approach we discuss shellable simplicial complexes and include a few original arguments concerning them. In particular we present original arguments that show the boundary of a simplex is shellable and that a shellable complex is endocollapsible; these results are mentioned in the literature, but we could not find proofs as accessible as the ones we provide.

Finally, Chapter V discusses the homology theory of discrete Morse theory introduced by

Forman [12], highlighting the striking similarity to the homology of smooth Morse theory. We present an original proof that the boundary map in fact gives us a chain complex; this extends the proof of Gallais [13] from the case of simplicial complexes to CW complexes. We also show that the discrete Morse homology (for CW complexes) is isomorphic to singular homology; this is a minor extension of the proof given by Gallais for simplicial complexes.

Contents

Chapter I

Smooth Morse theory

This chapter reviews the basics of Morse theory. We assume familiarity with the basics of differential geometry, particularly smooth manifolds, smooth maps, local coordinates, tangent bundles and vector fields. In addition we make use of some basic algebraic topology.

For a smooth manifold M, we will write M^m to denote that M has dimension m. In this section all maps are smooth functions. Given a smooth map $f : M \to N$ we will denote its differential by $df : TM \to TN$ and the differential at a point will be denoted by $d_p f : T_p M \to T_{f(p)} N$. Similarly, given a vector field X (a smooth section of the tangent bundle TM) on M we will write X_p to be the vector of X at p; that is $X_p = X(p) \in T_p M$.

Recall that a tangent vector $v \in T_p M$ is a derivation $C^\infty(M) \to \mathbb{R}$[1]. By $XY(f)$ we mean the composition of X and Y: $(XY)_p(f) = X_p(Y(f))$, note that given $f \in C^\infty(M)$, then $Y(f) \in C^\infty(M)$.

The Lie bracket of two vector fields X, Y will be denoted $[X, Y]$. This is the vector field given by $[X, Y](f) = XY(f) - YX(f)$.

1 Basic definitions

Definition I.1.1. Given a smooth function $f : M^m \to N^n$, a point $x \in M$ is *critical* if its differential $d_x f$ does not have full rank, i.e. $\dim \operatorname{im} d_x f < \min(m, n)$.

We will concern ourselves with real-valued functions in which case the condition that the differential has full rank is the same as saying the differential is non-zero.

We will want the critical points of our function to be 'not too flat'. Intuitively this will mean that second derivatives are non-zero. Recall that tangent vectors are directional derivative operators, then:

Definition I.1.2. The *Hessian* of $f : M \to \mathbb{R}$ at a critical point x is the bilinear map

$$\begin{aligned} H_{f,x} : T_x M \times T_x M &\to T_{f(x)} N \times T_{f(x)} N \\ (u, v) &\mapsto \tilde{u}_p(\tilde{v}(f)) \end{aligned}$$

where $\tilde{v}$ is a vector field so that $\tilde{v}_x = v$.

The Hessian then gives us an iterated second derivative where we first take the derivative with respect to $\tilde{v}$, giving us a new smooth function for which we take the derivative in the direction u at x. It is not immediate from this definition that the Hessian is well defined as

[1] $C^\infty(M)$ is the algebra of smooth functions $M \to \mathbb{R}$

it may depend on the chosen vector fields $\tilde{u}, \tilde{v}$. We will now see that it is well defined and that in addition it is symmetric.

First recall that the directional derivative $W(f)$ of f along a vector field W is given by $W_p(f) = d_p f(W_p)$. From this we can easily derive the following.

Lemma I.1.3. *Given a smooth function $f : M \to \mathbb{R}$, vector fields V, W on M and a critical point p of f we have $(VW)_p(f) - (WV)_p(f) = 0$.*

Proof. We have $(VW)_p(f) - (WV)_p(f) = [V, W]_p(f) = d_p([V, W]_p) = 0$ as p is critical. □

Lemma I.1.4. *The Hessian is independent of the chosen extension vector fields.*

Proof. Suppose p is a critical point of a smooth map f on M and $\tilde{u}, \tilde{u}'$ and $\tilde{v}, \tilde{v}'$ are vector fields on M so that $\tilde{u}_p = \tilde{u}'_p = u$ and $\tilde{v}_p = \tilde{v}'_p = v$. Then $(\tilde{u} - \tilde{u}')_p g = 0$ for all smooth functions g on M. In particular we have

$$\tilde{u}_p(\tilde{v}(f)) - \tilde{u}'_p(\tilde{v}(f)) = (\tilde{u} - \tilde{u}')_p(\tilde{v}(f)) = 0.$$

Similarly we have

$$\tilde{u}_p(\tilde{v}(f)) - \tilde{u}_p(\tilde{v}'(f)) = \tilde{u}_p((\tilde{v} - \tilde{v}')(f)) = 0$$

and hence the Hessian is independent of vector fields used. □

So we have that the Hessian is well defined; Lemma I.1.3 then tells us that the Hessian is symmetric.

Any symmetric bilinear map of vector spaces $H : V \times V \to W$ has an associated index and null space. The null space is the maximal subspace $U \subset V$ so that $H|_{U \times V} = 0$. The index of H is just the number of negative eigenvalues of H, equivalently this is the maximal dimension of a subspace on which H is negative definite.

Definition I.1.5. A critical point p of a smooth function $f : M \to \mathbb{R}$ is *non-degenerate* if the null space of $H_{f,p}$ is 0. If the null space of $H_{f,p}$ has positive dimension then p is *degenerate*.

Non-degenerate critical points are precisely the ones that are 'not too flat'. For a non-degenerate critical point p of f, we say that the *index* of p, denoted $\lambda(p)$, is the index of $H_{f,p}$. This tells us the number of linearly independent directions along which f decreases from p.

Definition I.1.6. A *Morse function* is a smooth function $M \to \mathbb{R}$ so that each critical point $p \in M$ is non-degenerate.

The set of critical points of a Morse function f will be denoted by $\mathrm{Cr}\, f$, and the set of index i critical points will be denoted $\mathrm{Cr}_i\, f$.

2 Morse functions and local coordinates

Morse functions can also be defined using local coordinates. Choosing local coordinates $(x^1, \dots, x^m)$ on $U \subset M^m$, a point $p \in M$ is critical for $f : M \to \mathbb{R}$ if $\frac{\partial f}{\partial x^i} = 0$ for all $i \in \{1, \dots, m\}$. The Hessian is then the familiar matrix of second derivatives

$$(H_{f,p})_{ij} = \frac{\partial^2 f}{\partial x^i \partial x^j}.$$

To see this, take coordinate expressions $u = \sum_i u^i \frac{\partial}{\partial x^i}|_p$ and $v = \sum_i v^i \frac{\partial}{\partial x^i}|_p$ for $u, v \in T_pM$. Then in U we have $\tilde{v} = \sum_i \tilde{v}^i \frac{\partial}{\partial x^i}$ where the $\tilde{v}^i$ are smooth functions with $\tilde{v}^i(p) = v^i$. Then we have

$$u(\tilde{v}(f)) = u\left(\sum_j \tilde{v}^j \frac{\partial f}{\partial x^j}\right) = \sum_i \sum_j u^i \tilde{v}^j \frac{\partial^2 f}{\partial x^i \partial x^j}.$$

Recall that any symmetric bilinear map has an eigenbasis. This means that we can choose coordinates where the bilinear map is represented by a diagonal matrix, and moreover its index will be the number of negative entries. We have a similar situation for Morse functions and their Hessians given to us by the Morse Lemma. Before proving the Morse Lemma we need a small tool that allows us to rewrite smooth functions.

Lemma I.2.1. *Let f be a smooth function on a convex neighbourhood of the origin in $\mathbb{R}^n$ with $f(0) = 0$, then we can write $f(x_1, \ldots, x_n) = \sum_{i=1}^n x_i g_i(x_1, \ldots, x_n)$ for some smooth functions g_i.*

Proof. As we have a convex subset of $\mathbb{R}^n$ we can write

$$f(x_1, \ldots, x_n) = f(x) - f(0) = \int_0^1 \frac{df(tx_1, \ldots, tx_n)}{dt} dt.$$

The chain rule then gives us

$$f(x_1, \ldots, x_n) = \int_0^1 \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) dt = \sum_{i=1}^n x_i \int_0^1 \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) dt,$$

so we take $g_i = \int_0^1 \frac{\partial f}{\partial x_i} dt$ □

Lemma I.2.2 (the Morse Lemma)**.** *For any non-degenerate critical point p of a smooth function $M^m \to \mathbb{R}$ there exists a neighbourhood U of p with local coordinates $(x^1, \ldots, x^m)$ centred[2] at p so that in these coordinates f is given by*

$$f(x) = f(p) - \sum_{i=1}^{\lambda(p)} (x^i)^2 + \sum_{i=\lambda(p)+1}^{m} (x^i)^2.$$

Proof. We follow the proof in [22], but show the diagonalisation process more explicitly.

Choose a chart centred at p with coordinates $(x^1, \ldots, x^n)$, and define in these coordinates $\tilde{f} = f - f(p)$. Then $\tilde{f}(0) = 0$. By Lemma I.2.1 we know that $\tilde{f}(x^1, \ldots, x^n) = \sum_{i=1}^n x^i g_i(x^1, \ldots, x^n)$. As 0 is a critical point of $\tilde{f}$ we have $g_i(0) = \frac{\partial f}{\partial x^i}(0) = 0$ and so we can apply Lemma I.2.1 to each g_i: $g_i(x^1, \ldots, x^n) = \sum_{j=1}^n x^j h_{ij}(x^1, \ldots, x^n)$ with $h_{ij}(0) = \frac{\partial^2 f}{\partial x^i \partial x^j}$. By replacing h_{ij} with $\frac{1}{2}(h_{ij} + h_{ji})$ we can assume that $h_{ij}(x^1, \ldots, x^n)$ is symmetric in i and j.

We can 'diagonalise' $\tilde{f}$ in a similar way to a quadratic form. Assume for induction that $\tilde{f} = \pm(x^1)^2 \pm \ldots \pm (x^{r-1})^2 + \sum_{i,j \geq r} x^i x^j h_{ij}(x^1, \ldots, x^n)$. If $h_{rr}(0) = 0$ then we can perform a linear change of coordinates to make $h_{rr}(0) \neq 0$. To do this note that there must be a non-zero term in $h_{ij}(0)$ for $i, j \geq r$ as otherwise the Hessian of f would be degenerate. Say $h_{kk'}(0) \neq 0$, with $k, k' \geq r$, then by setting $y^k = x^k + x^{k'}$ and $y^{k'} = x^k - x^{k'}$ we define coordinates u^i by $u^i = x^i$ for $i \neq k, k'$ and $u^k = y^k$ and $u^{k'} = y^{k'}$. Let $\tilde{h}_{ij}$ be h_{ij} in these new coordinates; we have $(y^k)^2 - (y^{k'})^2 = 4x^k x^{k'}$, so $\tilde{h}_{kk}(0)$ and $\tilde{h}_{k'k'}(0)$ are both non-zero;

[2] coordinates are said to be centred at a point x if x is the point 0 in these coordinates

by permuting coordinates we can assume that $k = r$. Let $g(u) = \sqrt{|\tilde{h}_{rr}(u)|}$, this is smooth in some neighbourhood of 0.

Define new coordinates v^i by $v^i = u^i$ for $i \neq r$ and $v^r(u) = g(u)\left(u^r + \sum_{i>r} u^i \frac{\tilde{h}_{ir}(u)}{\tilde{h}_{rr}(u)}\right)$. We have

$$(v^r)^2 = |\tilde{h}_{rr}|(u^r)^2 + \frac{|\tilde{h}_{rr}|}{\tilde{h}_{rr}^2}\left(\sum_{i>r} u^i \tilde{h}_{ir}\right)^2 + 2\frac{|\tilde{h}_{rr}|}{\tilde{h}_{rr}} \sum_{i>r} u^r u^i \tilde{h}_{ir}$$

where all $\tilde{h}_{ij}$ are evaluated at 0. One can check that this coordinate transformation has non-degenerate Jacobian, and thus gives us a coordinate system in some neighbourhood on 0. Note that the first ($|\tilde{h}_{rr}|(u^r)^2$) and final ($2\frac{|\tilde{h}_{rr}|}{\tilde{h}_{rr}} \sum_{i>r} u^r u^i \tilde{h}_{ir}$) terms or their negative appears in $\tilde{f}$ depending on whether $\tilde{h}_{rr}$ is positive or negative. We now have

$$\tilde{f}(v) = \left(\sum_{i<r} \pm (v^i)^2\right) \pm (v^r)^2 + \sum_{i,j>r} v^i v^j \tilde{h}_{ij} - \frac{1}{h_{rr}}\left(\sum_{i>r} u^i \tilde{h}_{ir}\right)^2$$

with the sign of the $(v^r)^2$ term corresponding to whether $\tilde{h}_{rr}$ is positive or negative. Notice that we have 'diagonalised' the contribution of the rth coordinate. By induction we can do this for all coordinates, and get a chart with coordinates v^i so that

$$\tilde{f}(v) = \sum \pm (v^i)^2$$

and so

$$f(v) = f(p) + \sum \pm (v^i)^2.$$

We can see that the number of negative terms (other than $f(p)$) must indeed be the index, as by computing the (diagonal) Hessian in these coordinates we get that many negative eigenvalues and the subspace of T_pM on which the Hessian is negative definite is independent of the coordinates. Thus by rearranging the order of the coordinates we get the desired form. $\square$

This is the key lemma that gives us information about the local structure of Morse functions. For example we immediately see that all critical points must be isolated, as locally (in the neighbourhood U of Morse's lemma) there are no critical points besides p. In turn this tells us that on compact manifolds a Morse function has finitely many critical points. We will call a neighbourhood of a critical point p with the properties given by the Morse Lemma a *Morse neighbourhood*.

3 Existence of Morse functions

Before presenting the basic results of Morse theory we pause briefly to discuss an important point: that Morse functions actually exists, and moreover they are generic in that every smooth function is 'almost' Morse. One simple way to think of Morse functions is as a type of 'height' function. Explicitly, if we are given a Morse function $f : M^m \to \mathbb{R}$ we can construct an embedding of M into some $\mathbb{R}^k$ so that f is given by projection to a coordinate.

To see this, we start by using the Whitney embedding theorem [29] to embed M^m into $\mathbb{R}^k$ for some $k \geq 2m$. Let $i_0 : M^m \to \mathbb{R}^k$ be such an embedding. We now construct an embedding $i : M \to \mathbb{R}^{k+1}, p \mapsto (i_0(p), f(p))$. We now see that f is given by the 'height' function $x \mapsto \langle e_{k+1}, x\rangle$ where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^{k+1}$.

We also have that almost any projection function on a submanifold of $\mathbb{R}^k$ is a Morse function. Let $\langle \cdot, \cdot \rangle$ denote the standard inner product on $\mathbb{R}^k$. Given a submanifold of $\mathbb{R}^k$ the function $\langle v, \cdot \rangle$ is Morse for almost all vectors $v \in \mathbb{R}^k$ [23]. These two results give us a very geometric way to think about Morse functions.

Definition I.3.1. A function $f : M \to \mathbb{R}$ is *exhaustive* if for all $a \in \mathbb{R}$ $f^{-1}(-\infty, a]$ is compact.

Exhaustive Morse functions have a useful property: for all $a < b$ we have $f^{-1}[a, b]$ is compact. This is a necessary hypothesis for the key structural theorems Theorems I.4.9 and I.4.10 that we will see shortly. The existence of exhaustive Morse functions is also guaranteed.

Lemma I.3.2. *Any smooth manifold admits exhaustive Morse functions.*

Proof. In addition to almost any height function of a submanifold $M^m \subset \mathbb{R}^k$ being Morse, Nicolaescu shows that almost any distance function is Morse. By a distance function we mean a function $f_x(p) = \frac{1}{2}|p - x|^2$ for some $x \in \mathbb{R}^k$; these functions are Morse for almost all $x \in \mathbb{R}^k$ [23]. In fact, if M is a closed subset of $\mathbb{R}^k$ and f_x is Morse, then it clear that f_x is exhaustive as the sets $f^{-1}(-\infty, b] = f^{-1}[0, b]$ are bounded (they are contained in the closed ball of radius b around x). By another version of the Whitney embedding theorem we can embed M as a closed subset of $\mathbb{R}^{2m+1}$ [20]; consider Morse distance functions on M under this embedding then proves that there are exhaustive Morse functions on M. $\square$

4 Structure from Morse functions

In this section we will present several key results of Morse theory that relate the structure of a manifold to Morse functions. Often we will require the use of the 'gradient' of a Morse function — in fact this vector field will be the key object used, rather than the Morse function itself. Without fixing a metric, we have no concept of a true gradient for a function. We have two solutions, for any smooth manifold we can always fix a metric and use the gradient given by the metric or we can make use of so-called 'gradient-like' vector fields.

4.1 Vector fields and flow lines

Vector fields on smooth manifolds can be thought of as describing some type of dynamics, carrying a point or submanifold to another, by realising that a vector field describes a family of differential equations. Given a vector field X on M, for any point $p \in M$ we have a curve $\gamma(t)$ given by the differential equation $\frac{d\gamma(t)}{dt} = X_{\gamma(t)}$ with initial solution $\gamma(0) = p$ where $\frac{d\gamma(t)}{dt}$ is the tangent to γ. Such an initial value problem has a unique solution for $|t| < \epsilon$ for some ϵ and in fact this solution depends smoothly on the initial point within some neighbourhood $U \ni p$. We will use vector fields to produce families of diffeomorphisms of a manifold which 'push' the manifold along these curves.

Definition I.4.1. A *flow* is a smooth group action of $(\mathbb{R}, +)$ on M.

Such a group action can be thought of as a family of diffeomorphisms $\phi_t : M \to M$ so that $\phi_{t+s} = \phi_t \circ \phi_s$ and the induced map $\mathbb{R} \times M \to M$ is smooth. Given such a family of diffeomorphisms we can define a vector field by

$$X_p(f) = \lim_{h \to 0} \frac{f(\phi_h(p)) - f(p)}{h}.$$

Such a vector field is said to *generate* the flow. The curves that are solutions to $\gamma'(t) = X_p(f)$ where X generates the flow are called *flow lines*.

Proposition I.4.2. *Given a vector field X on M vanishing outside a compact set $U \subset M$, there is a unique flow generated by X.*

Proof. The proof follows from the remarks at the beginning of this section, which give us local flows (i.e. smooth action of $(-\epsilon, \epsilon) \subset \mathbb{R}$ on M). These can then be patched together using a finite cover of $K = \bigcup_{i=1}^{n} U_i$ where we have a unique solution to $X_{\phi_t(p)} = \frac{d\phi_t(p)}{dt}$ on each $(-\epsilon_i, \epsilon_i) \times U_i$. Setting $\phi_t(p) = p$ for all $p \notin K$ (that is the derivative outside K is 0) and $\epsilon = \min_i \epsilon_i$ we then have a unique solution on $(-\epsilon, \epsilon) \times M$.

We can see that the homomorphism rule $\phi_{t+s} = \phi_t \circ \phi_s$ is respected as long as $|t|, |s|, |t+s|$ are all less than ϵ. This follows by considering the curves $t \mapsto \phi_{t+s}(p)$ and $t \mapsto \phi_t \circ \phi_s(p)$ and noting that both of these must be (local) solutions to $\frac{d\phi_t(q)}{dt} = X_{\phi_t(q)}$ with initial condition $\phi_0(q) = \phi_s(p)$. For larger t, we define ϕ_t by noting that for any t, we can write $t = k\frac{\epsilon}{2} + r$ where $|r| < \frac{\epsilon}{2}$, $k \in \mathbb{N}$. We define

$$\phi_t = \overbrace{\phi_{\frac{\epsilon}{2}} \circ \ldots \circ \phi_{\frac{\epsilon}{2}}}^{k \text{ times}} \circ \phi_r$$

if $k \geq 0$, and if $k < 0$ we use $\phi_{-\frac{\epsilon}{2}}$ instead. We can see that the homomorphism property is still satisfied; consider $\phi_t \circ \phi_s$, if both $t, s > \epsilon$, then we have

$$\begin{aligned} \phi_t \circ \phi_s &= \phi_{\frac{\epsilon}{2}} \circ \cdots \circ \phi_{\frac{\epsilon}{2}} \circ \phi_{r_t} \circ \phi_{\frac{\epsilon}{2}} \circ \cdots \phi_{\frac{\epsilon}{2}} \circ \phi_{r_s} \\ &= \phi_{\frac{\epsilon}{2}} \circ \cdots \circ \phi_{\frac{\epsilon}{2}} \circ \phi_{r_t} \circ \phi_{r_s} \\ &= \phi_{\frac{\epsilon}{2}} \circ \cdots \circ \phi_{\frac{\epsilon}{2}} \circ \phi_{r_t + r_s} \\ &= \phi_{t+s} \end{aligned}$$

where we get to the last line directly if $|r_t + r_s| < \frac{\epsilon}{2}$, or by pulling a $\phi_{\frac{\epsilon}{2}}$ out of $\phi_{r_t + r_s}$ if $r_t + r_s > \frac{\epsilon}{2}$ or absorbing the preceding $\frac{\epsilon}{2}$ if $r_t + r_s < -\frac{\epsilon}{2}$. This last step also shows the case of only one of t, s being greater than ϵ. The cases of one or both of t, s being less than $-\epsilon$ are similar computations. □

In particular if we have a compact manifold M then any vector field on M produces a flow.

Definition I.4.3. A *gradient-like* vector field for a Morse function $f : M^m \to \mathbb{R}$ is a vector field X so that $X(f) > 0$ and near a critical point p there are coordinates (x^i) so that in these coordinates $X(f) = -2\sum_{i=1}^{\lambda(p)} x^i \frac{\partial}{\partial x^i} + 2\sum_{i=\lambda(p)+1}^{m} x^i \frac{\partial}{\partial x^i}$.

Gradient-like vector fields can be obtained by choosing a Riemannian metric g that is given by $\sum_{i=1}^{m}(dx^i)^2$ in a Morse neighbourhood of each critical point [27]. Then the gradient with respect to this metric, $\nabla_g f$, is a gradient-like vector field.

Definition I.4.4. Given a Morse function $f : M^m \to \mathbb{R}$ and a gradient-like vector field ξ, let ϕ_t be the flow of $-\xi$. The *stable manifold* and *unstable manifold* of a critical point p are the subsets

$$M_s(p) = \{q \in M \mid \lim_{t\to\infty} \phi_t(q) = p\}$$

and

$$M_u(p) = \{q \in M \mid \lim_{t\to-\infty} \phi_t(q) = p\}$$

respectively.

The stable and unstable manifolds are submanifolds of M [23] consisting of the points that flow to p and away from p under the negative of the gradient-like vector field (note that $p = M_s(p) \cap M_u(p)$). This is the way we will normally think — in terms of flowing down in the direction (determined by ξ) of decrease in f. Intuitively, if one thinks of a Morse function as a height function, then the stable manifold is the set of points where water poured on the manifold would flow to p, and the unstable manifold tells us where water poured on p would flow. This explains an alternate terminology where the stable and unstable manifolds are called the ascending and descending manifolds respectively.

There is one property we will often want to require of a Morse function and a chosen gradient-like vector field: that the stable and unstable manifolds of different critical points intersect transversally.

Definition I.4.5. Two submanifolds $U, V \subset M$ have transverse intersection if for all $x \in U \cap V$, T_xM is spanned by T_xU and T_xV.

We note that the dimension of a transverse intersection has the expected dimension [8] (compare with the dimension of the transverse intersection of two vector subspaces).

Proposition I.4.6. *Suppose U^r, V^s are submanifolds of M^m with transverse intersection. Then $U \cap V$ is a submanifold of M with dimension $r + s - m$ (if $r + s - m < 0$ then we take this to mean $U \cap V = \varnothing$); equivalently* $\operatorname{codim} U + \operatorname{codim} V = \operatorname{codim} U \cap V$.

Definition I.4.7. The pair (f, ξ) where f is a Morse function and ξ is a gradient-like vector field for f is *Morse-Smale* if for any pair of critical points $p \neq q$, $M_s(p)$ and $M_u(q)$ intersect transversely.

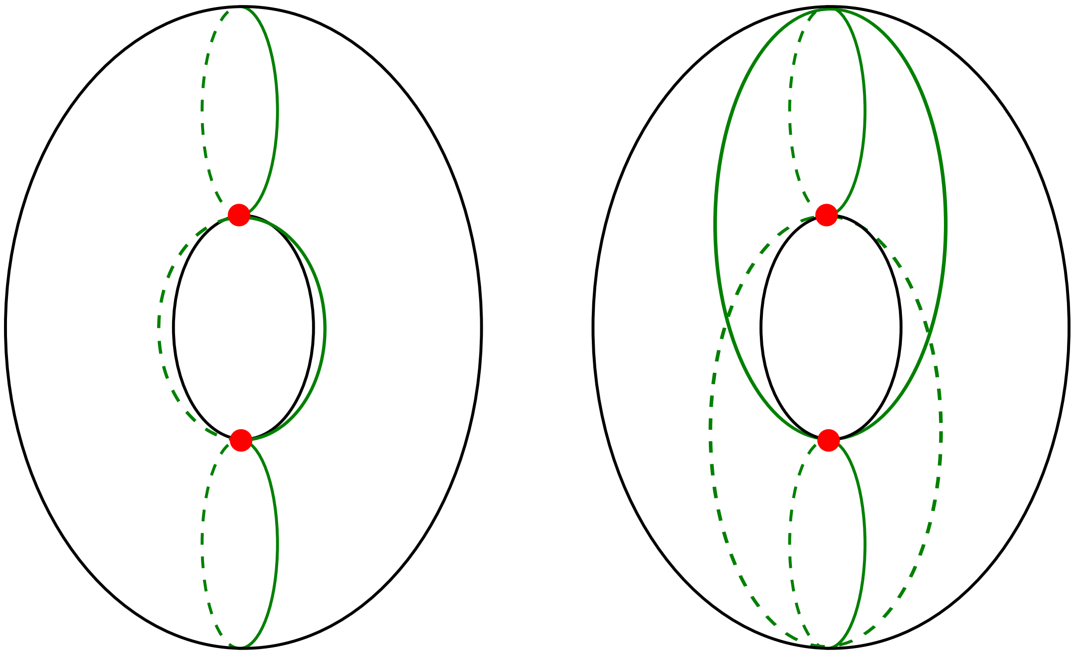
It turns out that for any Morse function, we can find a gradient-like vector field that is Morse-Smale and moreover the Morse-Smale property is generic (see [16, 23, 27]). Figure I.1(a) shows a torus with a Morse function given by height and gradient vector field induced from the metric on $\mathbb{R}^3$. The stable and unstable manifolds of the index one critical points are shown; note that these are not transverse as they intersect each other. Figure I.1(b) shows the situation if the torus is tilted slightly; note that this is Morse-Smale.

4.2 Handle decompositions

Morse theory is closely linked with the idea of decomposing manifolds into a sequence of balls (called handles) that are glued together. As discussed in Remark I.4.11 a Morse function provides a decomposition of a manifold into a sequence of handle attachments.

Definition I.4.8. Let M^m be a manifold with (non-empty) boundary and H a manifold diffeomorphic to D^m. Let $A \subset \partial H$ and $B \subset \partial M^m$ be $(m-1)$-submanifolds of the boundaries, both diffeomorphic to $S^{i-1} \times D^{m-i} \subset \mathbb{R}^m$ via diffeomorphisms that can be extended into collars of A and B. Let $M' = (M \sqcup H)/(A \equiv B)$ be the space obtained by gluing H to M using the identification provided by the diffeomorphisms. Then M' is the result of *attaching a smooth i-handle* to M.

A *handle decomposition* of a manifold M is given by a sequence of handles H_i so that $M = H_0 \cup H_1 \cup \cdots \cup H_n$ with the handles attached in order. If there are h_i handles of index i in a given handle decomposition of M^m, we call the sequence $(h_0, \ldots, h_m)$ the *handle vector* of this decomposition.



(a) Height function is not Morse-Smale for 'upright' torus; this is not Morse-Smale.

(b) Tilting the torus back a little gives Morse-Smale height function.

Figure I.1: Morse functions on a torus embedded in $\mathbb{R}^3$ given by height using gradient given by the standard inner product. Red dots are index 1 critical points and green lines are their stable and unstable manifolds.

4.3 Structural theorems

The two main structural theorems are Theorems I.4.9 and I.4.10. These are expressed in terms of *sublevel sets*: sets of the form $f^{-1}(\infty, a]$ which we denote ${}^{a}M$. Similarly we denote $f^{-1}[a, b]$ by ${}^{b}_{a}M$ and $f^{-1}[a, \infty)$ by ${}_{a}M$. Occasionally it will not be clear which function we are using to define a sublevel set (for example when we are relating the sublevel sets of two functions); when this is the case we will use superscripts to denote the function used, e.g. ${}^{b}_{a}M^{f}$. We can think of Theorems I.4.9 and I.4.10 as telling us when two sublevel sets are the same or different, with the presence of a critical point being the distinguishing feature.

These statements are often stated in forms of varying strength, specifically they can be stated in terms of diffeomorphism or homotopy type, the latter case usually given by the existence of a deformation retraction. Recall that a deformation retraction is a homotopy relative to a subspace between the identity map and a retraction to that subspace.

Theorem I.4.9. *Suppose $f : M^m \to \mathbb{R}$ is a Morse function and $a < b \in \mathbb{R}$ two distinct values so that ${}^{b}_{a}M$ is compact and contains no critical points of f. Then ${}^{b}M$ is diffeomorphic to ${}^{a}M$ and moreover ${}^{a}M$ is a deformation retract of ${}^{b}M$.*

Proof. We pick a Riemannian metric $\langle \cdot, \cdot \rangle$. Recall that the gradient with respect to this metric ∇f is characterised by $\langle \nabla f, X \rangle = X(f)$ for all vector fields X. We define a function

$$\rho = \frac{1}{\langle \nabla f, \nabla f \rangle}$$

on ${}_a^bM$ and so that ρ vanishes outside a compact neighbourhood of ${}_a^bM$. Define a vector field X by $X_p = \rho(p)(\nabla f)_p$; this vanishes outside a compact set of M and hence we have a flow $\phi_t : M \to M$ with $\frac{d\phi_t(p)}{dt} = X_{\phi_t(p)}$. Differentiating f along a curve $t \mapsto \phi_t(p)$ (that is the derivative of $t \mapsto f(\phi_t(p))$) we have if $\phi_t(p) \in {}_a^bM$ then

$$\frac{df(\phi_t(p))}{dt} = X(f) = \langle \nabla f, X\rangle = \frac{\langle \nabla f, \nabla f\rangle}{\langle \nabla f, \nabla f\rangle} = 1.$$

It follows that ϕ_{b-a} is a diffeomorphism ${}^aM \to {}^bM$. To see that there is a deformation retraction, we define a homotopy $r_t : {}^bM \to {}^aM$, by

$$r_t(p) = \begin{cases} p & \text{if } f(p) \leq a \\ \phi_{t(a-f(p))}(p) & \text{if } a < f(p) \leq b \end{cases}$$

It is clear that r_0 is the identity, r_1 is a retraction, and r_t is a homotopy relative to ${}^aM$ so ${}^aM$ is a deformation retract of ${}^bM$. □

Theorem I.4.10. *Let $f : M^m \to \mathbb{R}$ be a Morse function and p a critical point with $f(p) = c$ and index λ so that for some $\epsilon > 0$, the set ${}_{c-\epsilon}^{c+\epsilon}M$ is compact and contains no critical points other than p. Then ${}^{c+\epsilon}M$ has the same homotopy type as ${}^{c-\epsilon}M$ with a λ-cell e^λ attached; moreover ${}^{c-\epsilon}M \cup e^\lambda$ is a deformation retract of ${}^{c+\epsilon}M$.*

Proof. We will follow the proof in [22]. First choose a Morse neighbourhood U of p with coordinates (x^i) and some $\epsilon > 0$ so that ${}_{c-\epsilon}^{c+\epsilon}M$ is compact and the closed ball of radius 2ϵ is contained in U (using the coordinates of U). For convenience we will use x_- to mean the coordinates $(x^1, \ldots, x^\lambda)$ and x_+ for the coordinates $(x^{\lambda+1}, \ldots, x^m)$. We can then write $f|_U = -|x_-|^2 + |x_+|^2$.

We will now define a new Morse function F that takes lower values near p and agrees with f far enough away from p in such a way that the handle we attach to ${}^{c-\epsilon}M$ is given by ${}^{c-\epsilon}M^F \cap {}^{c-\epsilon}M^f$.

Let $\mu : \mathbb{R} \to \mathbb{R}$ be a smooth function so that $\mu(0) > \epsilon$, $\mu(t) = 0$ for $t \geq 2\epsilon$ and $-1 < \frac{d\mu}{dt} \leq 0$; this function will control the difference between F and f. Define F as

$$F(q) = \begin{cases} f(q) & q \notin U \\ f(q) - \mu(|x_-|^2 + 2|x_+|^2) & q \in U \end{cases}.$$

We now show some properties of F.

Claim. ${}^{c+\epsilon}M^F = {}^{c+\epsilon}M^f$

Proof. First note that $F|_{M\setminus U} = f|_{M\setminus U}$. We divide U into two regions: $A = \{|x_-|^2 + 2|x_+|^2 \geq 2\epsilon\}$ and $B = \{|x_-| + 2|x_+|^2 < 2\epsilon\}$. We have that $F|_A = f|_A$ as $\mu(|x_-|^2 + 2|x_+|^2) \geq 2\epsilon$ inside A, so $F|_A = f - 0 = f$ as μ is 0 on arguments $\geq 2\epsilon$. All that is left to show is that on B we have $F \leq c + \epsilon$. We have

$$F|_B \leq f = c - |x_-|^2 + |x_+|^2 \leq c - \frac{1}{2}|x_-|^2 + |x_+|^2 < c + \epsilon.$$

□

Claim. *${}^{c-\epsilon}M^F$ is diffeomorphic to ${}^{c+\epsilon}M^f$*

Proof. We only need to show that $\operatorname{Cr} F = \operatorname{Cr} f$ and that $p \notin {}^{c+\epsilon}_{c-\epsilon}M^F$. Then Theorem I.4.9 along with the first claim tells us that ${}^{c-\epsilon}M^F \cong {}^{c+\epsilon}M^F = {}^{c+\epsilon}M^f$.

Consider
$$\frac{\partial F}{\partial |x_-|^2} = -1 - \mu'(|x_-|^2 + 2|x_+|^2) < 0$$
and
$$\frac{\partial F}{\partial |x_+|^2} = 1 - 2\mu'(|x_-|^2 + 2|x_+|^2) \geq 1;$$
this tells us that F has no critical points in U as in the region B we have $dF = \frac{\partial F}{\partial |x_-|^2} d|x_-|^2 + \frac{\partial F}{\partial |x_+|^2} d|x_+|^2$. This is only 0 when both $d|x_-|^2$ and $d|x_+|^2$ are 0 (this only happens at the origin). So the critical points of F are the critical points of f. We now see that $p \notin {}^{c+\epsilon}_{c-\epsilon}M^F$ as $F(p) = c - \mu(0) < c - \epsilon$. $\square$

Let $H = \overline{{}^{c-\epsilon}M^F \setminus {}^{c-\epsilon}M^f}$ and let $e^\lambda = \{q \in U \mid |x_-|^2 < \epsilon, |x_+|^2 = 0\}$; e^λ is clearly a λ-dimensional closed disk. We have that $e^\lambda \subset H$ as $f(e^\lambda) = [c-\epsilon, c]$ and $F(e^\lambda) < c-\epsilon$. To finish the proof we need to show that ${}^{c-\epsilon}M^f \cup e^\lambda$ is a deformation retract of ${}^{c-\epsilon}M^f \cup H = {}^{c-\epsilon}M^F$; we will explicitly produce a deformation retraction r_t. We do this in pieces: begin by setting r_t equal to the identity outside U, we then define r_t on three other regions given by the inequalities $|x_-|^2 \leq \epsilon$, $\epsilon \leq |x_-|^2 \leq |x_+|^2 + \epsilon$ and $|x_+|^2 + \epsilon \leq |x_-|^2$.

- Consider the region given by $|x_-|^2 \leq \epsilon$. This region is in H and we see that projection onto the x_- coordinates sends any point to a point in e^λ, so we will define r_t to push along the x_+ coordinates. Define $r_t(x^1, \ldots, x^m) = (x^1, \ldots, x^\lambda, tx^{\lambda+1}, \ldots, tx^m)$; then r_0 maps this region of H onto e^λ and r_1 is the identity.

- Consider the region given by $\epsilon \leq |x_-|^2 \leq |x_+|^2 + \epsilon$. This region is the part of H where projection onto the x_- coordinates does not give a point in e^λ, so we need to do something a little more complicated than the previous case. Define $r_t(x^1, \ldots, x^m) = (x^1, \ldots, x^\lambda, s_t x^{\lambda+1}, \ldots, s_t x^m)$ where
$$s_t = t + (1-t)\sqrt{\frac{|x_-|^2 - \epsilon}{|x_+|^2}}.$$
We see that $s_t \in [0,1]$ with $s_1 = 1$ and $s_0 = \sqrt{\frac{|x_-|^2 - \epsilon}{|x_+|^2}}$. We then have that
$$\begin{aligned} f(r_0(x^1, \ldots, x^m)) &= c - |x_-|^2 + \frac{|x_-|^2 - \epsilon}{|x_+|^2}|x_+|^2 \\ &= c - \epsilon. \end{aligned}$$
When $|x_-|^2 = \epsilon$ we have that $s_t = t$ and so this coincides with the definition on r_t in the previous region along their common interface ($|x_-| = \epsilon$). In addition we have that when $|x_-|^2 = |x_+|^2 + \epsilon$ that $s_t = 1$ (hence r_t is the identity here).

- Consider the region $|x_+|^2 + \epsilon \leq |x_-|^2$. This region is in ${}^{c-\epsilon}M^f$ so we simply define r_t to be the identity here; this agrees with the previous region along their interface ($|x_+|^2 + \epsilon = |x_-|^2$). The intersection of the first region with this region is the points with $|x_-|^2 = \epsilon$ and $|x_+|^2 = 0$ which is the boundary of e^λ and so this definition coincides with that of the first region.

We now have that r_t is a deformation retraction from ${}^{c-\epsilon}M^f \cup H$ to ${}^{c-\epsilon}M^f \cup e^\lambda$. Notice that e^λ is attached to ${}^{c-\epsilon}M^f$ along its boundary; this completes the proof. □

Theorem I.4.10 uses a Morse function so that $f^{-1}(c)$ contains a single critical point. As the proof takes place locally, we can drop this hypothesis and have k critical points in $f^{-1}(c)$. We are then attaching a λ_i-cell for each critical point of index λ_i with value in $[c-\epsilon, c+\epsilon]$.

Remark I.4.11. There is a stronger form of this theorem that tells us the change in diffeomorphism type when passing a critical point. Passing a critical point of index λ actually corresponds to attaching a smooth λ-handle [1, 24]. Kosinski proves this stronger version by showing that the subset H considered above is actually a handle attached appropriately along ${}^{e-\epsilon}M^f$ [1]. The upgraded theorem states that given the same hypothesis as in Theorem I.4.10, ${}^bM$ is diffeomorphic to ${}^aM$ with a smooth λ-handle attached.

As a consequence of Theorems I.4.9 and I.4.10 we have

Theorem I.4.12. *Suppose $f : M \to \mathbb{R}$ is an exhaustive Morse function, then M is homotopy equivalent to a CW complex with one λ-cell for each index λ critical point.*

This follows quite simply if there are finitely many critical points, but is also true for functions with infinitely many critical points (recall that if there are infinitely many critical points the manifold must be non-compact) [22].

The Morse inequalities

Recall that the Euler characteristic of a CW complex X is given by the alternating sum of the number of i-cells: $\chi(X) = \sum_{i=0}^n (-1)^i n_i$ where n_i denotes the number of i-cells of X. This is the same as the alternating sum of Betti numbers b_i (see [14]), which are given by the ranks of the homology groups: $b_i(X) = \operatorname{rank} H_i(X)$. By considering the cellular homology, we can then see that Theorem I.4.12 implies that $\chi(M^m) = \sum_{i=0}^m (-1)^i |\mathrm{Cr}_i\, f|$.

Considering the cellular homology also shows that $b_i(X)$ is not greater than the number of i-cells in a CW complex X as the rank cellular homology groups must be less that the number of i in any given CW structure for the manifold. Hence we have a set of inequalities

$$b_l(M) \leq |\mathrm{Cr}_l\, f| \qquad \forall l \in \mathbb{N}$$

for a compact manifold M and any Morse function $f : M \to \mathbb{R}$; these are known as the weak Morse inequalities. This places a lower bound on the number of critical points that a Morse function can have, or equivalently, given a Morse function on a manifold we obtain an upper bound for the Betti numbers. The weak Morse inequalities can also be obtained from a stronger result, the (strong) Morse inequalities (see [22, 23]), which are

$$\sum_{i=0}^{l} (-1)^i b_{l-i}(M) \leq \sum_{i=0}^{l} (-1)^i |\mathrm{Cr}_l\, f|$$

for any $l \in \mathbb{N}$.

In later chapters we will be interested in the number of critical points of each index that a Morse function has. To simplify terminology we define the following.

Definition I.4.13. Suppose $f : M^m \to \mathbb{R}$ is a Morse function. The sequence of numbers $(|\mathrm{Cr}_0\, f|, |\mathrm{Cr}_1\, f|, \ldots, |\mathrm{Cr}_m\, f|)$ is called the *Morse vector* of f.

Definition I.4.14. A Morse function $f : M^m \to \mathbb{R}$ with Morse vector $(c_0, \ldots, c_m)$ is *perfect* if $c_i = b_i(M)$.

A perfect Morse function is the simplest we can hope to have, given the constraint of the weak Morse inequality. Note that there are manifolds that do not admit a perfect Morse function [23]. Given a smooth Morse function f, there is a simple criterion that detects when we can perturb f in order to 'cancel' two critical points; this is a useful way to simplify Morse functions. We will not do much with this process but include it here to allow comparison with the analogous statement in discrete Morse theory.

Theorem I.4.15. *Suppose p, q are critical points with consecutive indices ($\lambda(p) = \lambda(q) - 1$) of a Morse function $f : M \to \mathbb{R}$ and ξ is a gradient-like vector field so that $M_s(p)$ and $M_u(q)$ intersect transversely. Furthermore, suppose that for any level set between $f(p)$ and $f(q)$ this intersection consists of a single point. Then there is a Morse function g equal to f outside a neighbourhood U of $M_s(p) \cap M_u(q)$ so that g has no critical points in U.*

For a proof see [21, 19]. Note that the condition that any level set between p and q contains a single point of $M_s(p) \cap M_u(q)$ is satisfied if it is true for a single level set.

This process can also be reversed, to create two critical points of corresponding index. The cancellation of two critical points is sometimes referred to as death, with the operation that creates two critical points being referred to as birth.

Chapter II

Classifying surfaces by ordering functions

In this chapter we will examine a neat application of Morse theory. We will classify the smooth closed compact orientable 2-manifolds up to diffeomorphism by reordering the critical points of a given Morse function. This demonstrates the utility that local modifications of Morse functions bring to understanding the structure of manifolds. In this section we will refer to smooth closed compact orientable 2-manifolds as surfaces.

We will follow the approach of [11] but fix an error concerning the movement of index 1 critical points. The general strategy is to choose a Morse function that will give us a simple handle decomposition of a surface. We then show that we can understand a particular set of handle attachments and build our manifold by gluing these together.

Theorem I.4.10 tells us that a Morse function on M gives us a handle decomposition of M. We will see that we can change this decomposition so that handles are attached in a specific order.

Definition II.0.1. A Morse function is said to be *ordered* if $\lambda(p) < \lambda(q)$ implies $f(p) < f(q)$ and no critical points of the same index take the same value.

1 Surfaces admit ordered Morse functions

In this section we will prove that any surface admits an ordered Morse function. In fact a stronger statement is true; given a Morse function on any manifold, by performing certain local modifications we can produce an ordered Morse function with the same number of critical points of each index [21, 23].

We start by showing that any index 0 critical point can be lowered an arbitrary amount.

Lemma II.1.1. *Given a Morse function $f : M^m \to \mathbb{R}$ with an index 0 critical point p, there exists a neighbourhood U of p and a Morse function $g : M^m \to \mathbb{R}$ with the same critical points as f so that $g|_{M\setminus U} = f|_{M\setminus U}$ and $g(p) = f(p) - a$ for any $a \geq 0$.*

Proof. Let U be a Morse neighbourhood of p with local coordinates $(x^1, \ldots, x^m)$. Let $\epsilon > 0$ be such that U contains the closed ball of radius ϵ, $\overline{B_\epsilon(0)}$, and assume we have a smooth function $h : U \to \mathbb{R}$ so that $h(p) = a$, the mixed partial derivatives $\frac{\partial^2 h}{\partial x^1 \partial x^2}(p)$ are zero,

$h(x) = 0$ for $|x| > \epsilon$ and for all $|x| < \epsilon$ we have $\frac{\partial h}{\partial x^i}|_x \leq 0$ and $\frac{\partial^2 h}{\partial x^{i2}}(p) < 2$. Then

$$g(q) = \begin{cases} f(q) - h(q) & \text{if } q \in U \\ f(q) & \text{if } x \notin U \end{cases}$$

will be the desired function. To see this, first note that p is the only critical point of f in U and $g|_{M \setminus U} = f|_{M \setminus U}$. We have that $dg = df - dh$ in U, in particular $dg|_p = df|_p - dh|_p = 0$ so the critical points of g are the critical points of f as for any other $x \in U$ $dg \neq 0$ (all partial derivatives of g are positive in $U \setminus \{p\}$). The Hessian of g is given by

$$\begin{aligned} (H_{g,p})_{ij} &= \frac{\partial^2 g}{\partial x^i \partial x^j} \\ &= \frac{\partial^2 f}{\partial x^i \partial x^j} - \frac{\partial^2 h}{\partial x^i \partial x^j} \\ &= 2\delta_{ij} - a_i \delta_{ij} \end{aligned}$$

where $a_i < 2$, then we have that $H_{g,p}$ is diagonal in these coordinates and all its eigenvalues are positive; hence p is an index 0 critical point of g. So we have that g is a Morse function with the desired properties.

Recalling the bump functions with specific properties that are known to exist, we see that such a h exists. For an explicit example consider (in coordinates)

$$h(x^1, \dots, x^m) = a \prod_{i=1}^{m} \phi(x^i)$$

where

$$\phi(x) = \begin{cases} e^{-\frac{1}{1-(\frac{x^i}{\epsilon})^2}+1} & |x| < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

□

By considering the function $-f$ on M we have the following corollary.

Corollary II.1.2. *Given a Morse function $f : M^m \to \mathbb{R}$ with an index m critical point p there exists a neighbourhood U of p and a Morse function $g : M^m \to \mathbb{R}$ with the same critical points as f so that $g|_{M \setminus U} = f|_{M \setminus U}$ and $g(p) = f(p) + a$ for any $a \geq 0$.*

We now turn to the case of index 1 critical points on surfaces, which we will see can be moved up or down. There is a subtle difference when moving an intermediate index critical point p on M^m (not index 0 or m) that prevents us from naively pushing a neighbourhood of p up or down. The issue with just using any Morse neighbourhood of p is that extra critical points can be introduced, as is shown in Figure II.1.

Figure II.1 shows how naively lifting a critical point using a strategy like that in Lemma II.1.1 can produce new critical points: the flow lines and level sets can be perturbed when raising the index one critical point in such a way the new critical points are created. Note that flow lines must be transverse to level sets and can only intersect at critical points[1]. If we raise the critical point in only a small neighbourhood, then we can see that the part of the stable manifold that is not perturbed must meet the new unstable manifold, resulting in new critical points as we now have intersecting flow lines.

[1] This follows from the uniqueness of the solution to the associated differential equations

To remedy this we must modify f in a neighbourhood of the ascending (resp. descending) manifold of p if we are increasing (resp. decreasing) the value of p (see Figure II.2), and we must keep the value of the critical point below (resp. above) the value of the part of the ascending (resp. descending) manifold in the boundary of the neighbourhood. Otherwise we will encounter the situation depicted in Figure II.1.

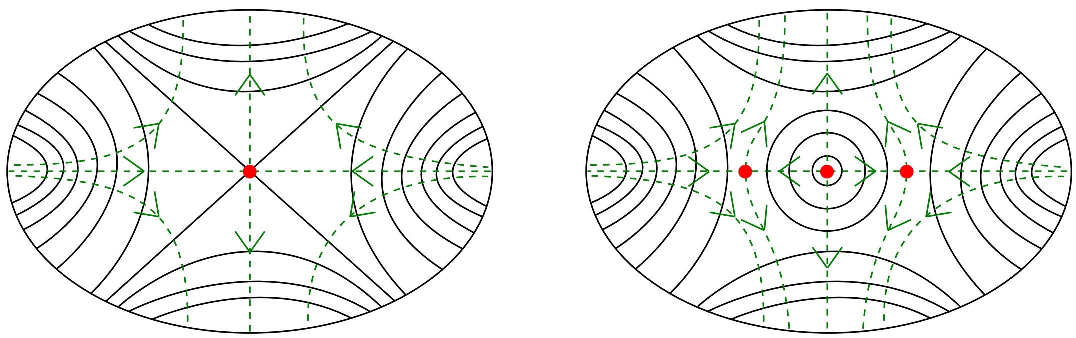


Figure II.1: Naively moving a critical point can produce extra critical points. Black lines represent level sets with dotted green (oriented) lines representing the flow and red dots the critical points. The sole index 1 critical point (left) is raised in a small neighbourhood; this alters the flow lines and produces new critical points (right).

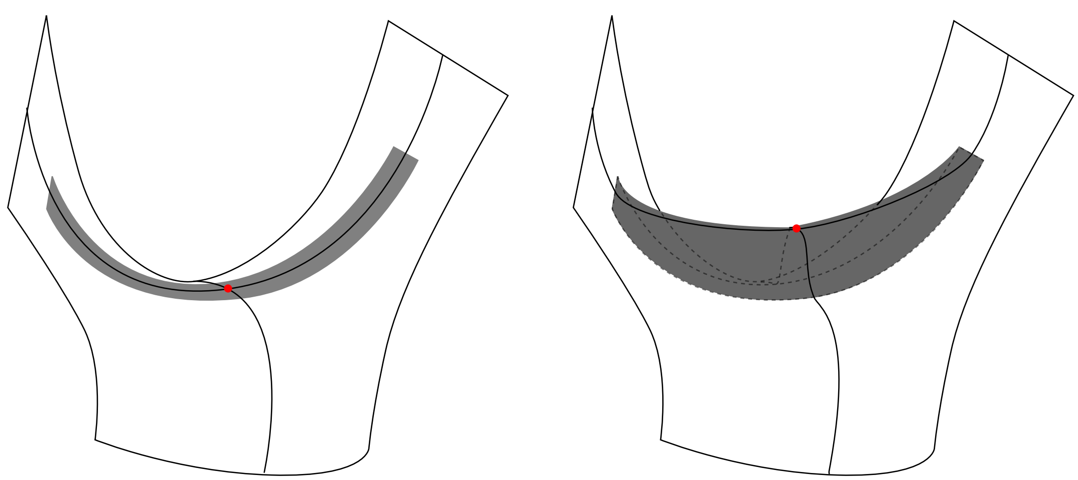


Figure II.2: To move an index 1 critical point up we must modify the Morse function in a neighbourhood of the ascending manifold; the red point is the critical point and the shaded region is the modified neighbourhood. Note the boundary of the shaded regions agree.

This is the error in [11]. The statement given there claims that the value of an index 1 critical point can be moved up or down an arbitrary amount but the proof provided has allowable values of the critical point dependent the size of Morse neighbourhoods that exist around the point. We potentially need to move the index 1 critical points further than just some $\epsilon > 0$ and it was not shown that there is a sufficiently large neighbourhood exists to accomplish this. Lemma II.1.3 is crucially used in Lemma II.2.3 to simplify an ordered Morse function to a desired form, but this may involve large movements of particular index 1 critical points so we correct the statement and proof in [11] to allow for this.

Lemma II.1.3. *Suppose f is an ordered Morse function on a surface M with some gradient-like vector field ξ and p is a critical point of index 1. Let a, b be two regular values with $a < f(p) = c < b$ so that $M_s(p) \cap {}_a^bM$ does not intersect the unstable manifold of any critical point in $\operatorname{Cr} f \cap {}_a^bM$. Then for any $c' \in (c, b)$ there is a Morse function g, equal to f outside an open neighbourhood of $M_s(p)$ so that $g(p) = c'$ and p is a critical point of index 1 for g; moreover no new critical points are introduced.*

Proof. We will construct a smooth function on some neighbourhood U of $M_s(p)$ so that p is an index 1 critical point and the function and its derivatives match up along the boundary of U. Consider a compact neighbourhood N of $M_s(p) \cap {}_a^{b-\epsilon}M$ (contained in ${}_a^bM$) so that N is disjoint from the unstable manifold of any critical point in ${}_c^bM$. We can then define a smooth function h on N with p an index 1 critical point of h, so that ξ is gradient-like for h along $M_s(p)$, $f(p) = c'$, and h is equal to f near the boundary of N. Then defining g by $g|_N = h$ and $g|_{M\setminus N} = f$ proves the theorem.

Define smooth path $\alpha : [-1, 1] \to M_s(p)$ so that $\alpha(\pm 1)$ gives the two points in $M_s(p) \cap f^{-1}(b-\epsilon)$ and $f \circ \alpha = \bar{h}$ is a Morse function on $[-1, 1]$. We can pick a tubular neighbourhood of N with coordinates (x, y) so that in these coordinates $f = \bar{h}(x) - y^2$ with p the origin and the line in $y = 0$ is $N \cap M_s(p)$ (this follows similarly to a statement in [19]).

Let $\tilde{h}$ be a smooth function $\tilde{h} : [-1, 1] \to \mathbb{R}$ so that $\bar{h} < \tilde{h}$, $\tilde{h}(0) = c'$, 0 is the only critical point of $\tilde{h}$ and $\tilde{h} = \bar{h}$ near $\{-1, 1\}$; it is clear that such a function exists. Let $\phi : \mathbb{R} \to \mathbb{R}$ be a smooth function with non-negative derivative so that $\phi(0) = 0$ and $\phi(r) = 1$ for $|r| > \epsilon'$ where $\epsilon' > 0$ is such that $f(\epsilon', 0) = b - \epsilon$. Then setting $h(x, y) = \tilde{h}(x) + (\bar{h}(x) - \tilde{h}(x))\phi(y^2) - y^2$ we have a suitable function h: we can see that h has a non-degenerate critical point at $(0, 0)$ of index 1 and this is the only critical point of h. □

Combining Lemmas II.1.1 and II.1.3 and corollary II.1.2 we have that surfaces admit ordered Morse functions.

Theorem II.1.4. *Any Morse function f on a surface can be perturbed in neighbourhoods of stable and unstable manifolds so that the resulting function is ordered.*

We also see that we can lower an index 1 critical point p in a similar way (consider the negative of the function and apply Lemma II.1.3). By choosing a Morse-Smale gradient, we can then move an index 1 critical point p as far up or down as we like so long as we do not meet the level of an index 0 (resp. 2) critical points whose stable (resp. unstable) manifolds intersect the unstable (resp. stable) manifold of p.

Remark II.1.5. Note that Lemma II.1.1 and corollary II.1.2 work in any dimension. There is also a generalisation of Lemma II.1.3 to any dimension, commonly known as the 'rearrangement lemma' [21, 23]. This allow for a Morse function, with a gradient-like vector field ξ, to be modified (locally) so as to change the value of any critical point p in ${}_a^bM$ to any $c \in [a, b]$ provided that $W(p) = M_s(p) \cup M_u(p)$ does not intersect $W(q)$ for any other critical point q in ${}_a^bM$. Moreover ξ is gradient-like for the new function.

1.1 Examining the movement of index 1 critical points

Using the flow given by a gradient like vector field we can follow any non-critical points through level sets of a manifold in either direction (up or down). Assuming that we have a non-critical point $x \in f^{-1}(t)$, then as long as the flow does not take x to a critical point there is a unique critical point in any level set $f^{-1}(t \pm \epsilon)$ that corresponds to x, and these points do not correspond to any other $y \neq x$ in $f^{-1}(t)$. In this section we will make use of these ideas to talk about how relationships between points changes in different level sets — what we really mean is the corresponding points under the flow.

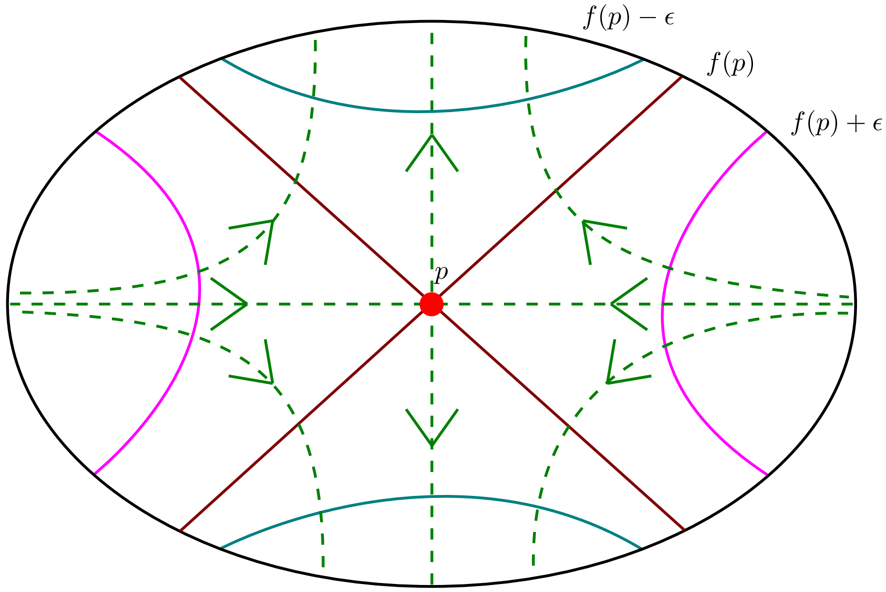


Figure II.3: Connectivity of level sets near a critical point p; level sets $f(p-\epsilon)$ (blue), $f(p)$ (brown), and $f(p+\epsilon)$ (pink) are shown with flow lines (dotted green). Notice how an interval in $f(p+\epsilon)$ than contains a point on a flow line heading to p is not mapped to an interval in $f(p-\epsilon)$ by the flow.

Let us restrict our focus to a surface M. The Morse lemma gives us an exact picture how boundary components of sublevel sets change as we pass over a critical point of index one. The key here is that if $f^{-1}(c)$ is a critical point p, then the connectivity of the components of $f^{-1}(c+\epsilon)$ and $f^{-1}(c-\epsilon)$ changes, see figure II.3.

Consider the points in Figure II.3 at the intersection of the level set $f(p)-\epsilon$ and the indicated flow lines that do not meet p; there are two pairs, one near the top of the diagram and one near the bottom. Suppose p is type 1, then both pairs are connected in $f^{-1}(f(p)-\epsilon)$ outside the shown neighbourhood. The points in $f^{-1}(f(p)+\epsilon)$ that are the image of the points in a pair under the flow are then not in the same component of $f^{-1}(f(p)-\epsilon)$.

Let p be a type 1 critical point, then $M_u(p)$ has both its intersections with $f^{-1}(c-\epsilon)$ in the same component. For clarity we assume p has type 1, though the same idea holds for type 2. We will consider what happens when p is moved just above another index 1 critical point q so that $c = f(p) < b < f(q) = d$ and p, q are the only critical points in ${}^{d+\epsilon}_{c-\epsilon}M$ for small enough ϵ; we will call the modified function g and let $g(p) \in (d, d+\epsilon)$ (we only move p past a single critical point). As we have seen, the handle attachment that corresponds to p for f takes place along a neighbourhood of $M_u(p) \cap f^{-1}(c-\epsilon)$. As p is type 1, $f^{-1}(b)$ has at least two components.

If $M_u(q)$ intersects $f^{-1}(b)$ in the same components, then we know that q has type 2. As we modify f in a neighbourhood of $M_s(p)$ that is disjoint from $M_u(q)$, the flow is unchanged for points in $M_u(q)$. This means that the handle corresponding to q is attached at the same points of the manifold after moving p above q. Notice however that after moving p we know that $M_u(q)$ intersects a single component of $g^{-1}(b)$ (as the handle of p is not there to separate it into two), and hence q has type 1 after moving p. This is shown schematically in Figure II.4; notice that the level set $g = b$ is a single component whereas $f = b$ is two components.

On the other hand, if the intersection of $M_u(q)$ and $f^{-1}(b)$ does not occur in exactly the same components as the intersection of $M_s(p)$ and $f^{-1}(b)$ then we see that $M_s(q)$ intersects

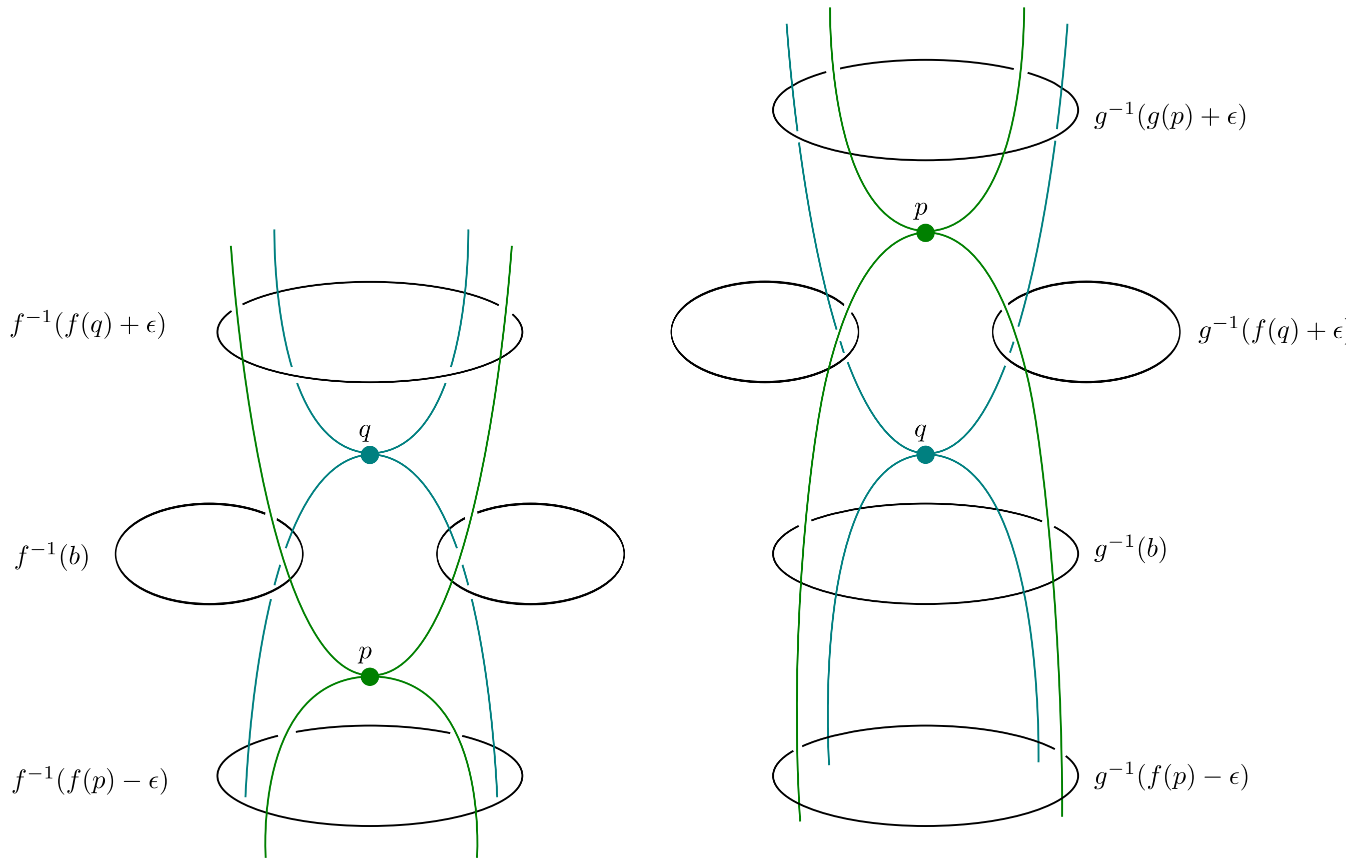


Figure II.4: Change in level sets when moving index 1 critical point p above critical point q (i.e., modifying f to g); as the unstable manifold of q intersects the same components of $f^{-1}(b)$ as the stable manifold of p, the types of p and q switch when moving p above q.

$g^{-1}(b)$ in the same number of components as it intersects $f^{-1}(b)$ and hence its type does not change.

To see that p must change type if q does, and similarly keep its type if q does, we count boundary components of sublevel sets. Noting that each critical point changes the number of boundary components by one, if q now adds one rather than subtracting one (i.e., q changes to type 1) p must subtract one (have type 2) after the move as the other critical points are unaffected due to the locality or the perturbation that produces g from f. Similarly, if q retains its type then similarly p must retain its type.

2 Classification

We now complete the classification of surfaces. The advantage of using ordered Morse functions is that we know that order in which different index handle attachments happen and so we know the index of handles in certain 'height' ranges. In particular Lemmas II.2.1 and II.2.3 will mean that we only need to worry about the interplay between a subset of the index 1 handles.

The following lemma reduces the number of critical points we will need to worry about under certain conditions; this lemma can be thought of as taking the place of cancellation (Theorem I.4.15), but is much simpler.

Lemma II.2.1. *Suppose $a < b$ are two regular values of an ordered Morse function f so that there are an equal number of index 0 critical points and index 1 critical points of type 2 in ${}_a^bM$ and these are the only critical points in ${}_a^bM$. Then ${}^bM$ is diffeomorphic to ${}^aM$.*

Proof. This follows immediately from considering the number of boundary components; say there are m index 0 critical points (and index 1 critical points of type 2) in ${}_a^bM$. As f is ordered, we know that ${}^aM$ is diffeomorphic to a disjoint union n of disks where n is the number of index 0 critical points in ${}^aM$. There is some regular value $a' \in (a, b)$ so that separates all index 0 and index 1 critical points. We then have that ${}^{a'}M$ is diffeomorphic to a disjoint union of $m + n$ disks. Attaching an index 1 handle between two different disks clearly gives us back a disk, so we have ${}^bM$ is a disjoint union of $m + n - m = n$ disks. $\square$

Proposition II.2.2. *Suppose M admits an ordered Morse function with a single index 0 critical point, two index 1 critical points and a single index 2 critical point. Then M is a torus.*

Proof. First note that the two index 1 critical points must be of differing type — the first must have type 1 as there is only one boundary component it can attach to, and the second must have type 2 as it must join the two boundary components produced by the type 1 handle in order for the index 2 handle attachment to result is a closed manifold. Attaching the first 1-handle gives us a cylinder. Then attaching a handle along sections of two different boundary components of the cylinder clearly gives us a torus minus a disk; this disk is the 2-handle that is attached last. $\square$

The following lemma, in conjunction with Lemma II.2.1, will reduce the number of index 1 critical points we need to worry about by showing that we can improve a given ordered Morse function to have a particularly nice form.

Lemma II.2.3. *Given an ordered Morse function $f : M \to \mathbb{R}$ there is an order Morse function g with the same critical points of f with regular values $a < b$ so that ${}^aM$ is diffeomorphic to a disk and ${}_bM$ is diffeomorphic to a disk. Moreover ${}_a^bM$ has an even number of index 1 critical points, half each of type 1 and type 2.*

Proof. We only need to show the claim for ${}^aM$ as the corresponding statement for ${}_bM$ follows by considering the function $-f$. Let α be a regular value that separates the index 0 and index 1 critical points and β a regular value that separates the index 1 and index 2 critical points. Then ${}^\alpha M^f$ is a disjoint union of m_0 disks, having m_0 boundary components and ${}^\beta M^f$ is connected, having m_2 boundary components. As ${}^\beta M^f$ is connected we know that there is at least $m_0 - 1$ index 1 critical points of type 2 whose flow lines travel down to distinct index 0 critical points.

We construct g by moving $m_0 - 1$ such critical points down one by one (using Lemma II.1.3) to lie between α and the lowest value of an index 1 critical point that has not been moved previously. Let a be a regular value for g separating the moved index 1 critical points from the rest of the index 1 critical points. Then we have that ${}^aM_g$ is diffeomorphic to a disk by Lemma II.2.1. Similarly we can do the same to get some b so that ${}_bM_g$ is a disk. We now have that there must be an equal number of type 1 and type 2 1-handles in between a and b as $g^{-1}(a)$ has the same number of components as $g^{-1}(b)$ (they each have a single component). $\square$

We will call a Morse function satisfying the properties of sublevel sets described in Lemma II.2.3 a *nice* ordered Morse function and we call the values a, b lower and upper *disk values* respectively. Note that upper and lower disk values are not unique; the set of disk values forms an open bounded set in $\mathbb{R}$ with two components, one for upper and one for lower disk values. We have the following immediate corollary.

Corollary II.2.4. *A surface admits a nice ordered Morse function.*

We are now set to prove our classification. We sate this in terms of connected sum of manifolds, recalling that smooth connected sums that respect orientation are well defined.

Theorem II.2.5. *Suppose $f : M \to \mathbb{R}$ is a nice ordered Morse function with lower and upper disk values $a < b$ and n index 1 critical points in ${}^{b}_{a}M$. Then M is the connected sum of $\frac{n}{2}$ tori.*

Proof. We being by modifying f on ${}^{b}_{a}M$ so that the critical points alternate in type as we move up the manifold (with respect to the new function). We can do this by altering f, moving type 2 critical points down as far as possible without changing their types. We now assume that the critical points of f in ${}^{b}_{a}M$ alternate.

We need to show that for a regular value δ, with p pairs of critical points (a type 1 followed by a type 2) in ${}^{\delta}_{a}M$, that ${}^{\delta}M$ is a connected sum of p tori with a disk removed. Assume for induction that this is true for the first $p-1$ pairs (the base case is given by Proposition II.2.2); then there is some $\gamma < \delta$ with ${}^{\gamma}M$ a connected sum of $p-1$ tori minus a disk. By Proposition II.2.2 we have that ${}^{\delta}_{\gamma}M$ is a torus minus two disks and gluing this to ${}^{\gamma}M$ together we have that ${}^{\delta}M$ is a connected sum of p tori, minus a disk. By induction, ${}^{b}M$ is a connected sum of $\frac{n}{2}$ tori minus a disk. This missing disk is given by ${}_{b}M$, and hence M is the connected sum of $\frac{n}{2}$ tori. $\square$

Chapter III

Discrete Morse theory

We will now review the basics of Forman's discrete Morse theory [12]. This was developed in the 1990s and allows us to generalise results and techniques of smooth Morse theory to CW complexes. This may be useful, for example, when trying to analyse discrete data sets [18]. Most of Forman's theory was initially developed with finite CW complexes in mind though parts of it have been subsequently generalised to infinite complexes [2]. We assume familiarity with CW complexes, particularly their construction via attaching maps, the degrees of these maps and the simplicial, cellular and singular homology theories of CW complexes.

1 Basic definitions and results

Definition III.1.1. A cell σ of a CW complex is a *regular face of* τ if the characteristic map of τ restricted to the inverse image of $\mathring{\sigma}$ is a homeomorphism and the closure of this inverse image is a closed p-ball. A CW complex is called *regular* if every face is regular.

Some common examples of CW complexes are not regular. For example consider the standard CW structure of the projective plane $\mathbb{R}P^2$ given by one 0-, 1- and 2-cell where the attaching map for the 2-cell has degree two. This is not regular as the attaching map for the 2-cell is a two-to-one map.

On the other hand, a simplicial complex is a regular CW complex. Though we will often deal with simplicial complexes (particularly when talking of PL manifolds), the basic definitions and theorems of discrete Morse theory hold for general CW complexes. Unless otherwise specified, all CW complexes are assumed to be finite.

To condense notation when first introducing a cell we will write $\sigma^{(p)}$ to mean a cell σ has $\dim \sigma = p$. We will write $\sigma < \tau$ for two cells σ, τ so that $\sigma \subset \overline{\tau}, \sigma \neq \tau$ and $\sigma \leq \tau$ if $\sigma < \tau$ or $\sigma = \tau$. For $\sigma < \tau$, we call σ a *face* of τ.

Lemma III.1.2. *Given cells $\nu < \sigma < \tau$ with ν a regular cell of σ and σ a regular cell of τ. Then there is a cell $\tilde{\sigma}^{(p)} \neq \sigma$ so that $\nu < \tilde{\sigma} < \tau$.*

Proof. This is checked easily by computing $\partial\tau$, $\partial\sigma$ and comparing with $\partial^2\tau = 0$ where ∂ is the cellular boundary map. □

Definition III.1.3. A *discrete Morse function* on a CW complex M is a function $f : M \to \mathbb{R}$ so that any irregular face σ of τ has $f(\sigma) < f(\tau)$. In addition, for any cell $\sigma^{(p)}$ we require

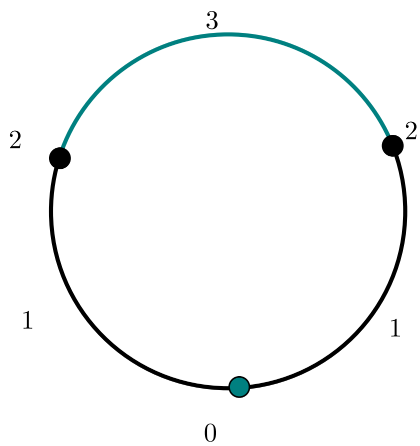
- $|\{\tau^{(p+1)} > \sigma \mid f(\tau) \leq f(\sigma)\}| \leq 1$

- $|\{\nu^{(p-1)} < \sigma \mid f(\nu) \geq f(\sigma)\}| \leq 1$

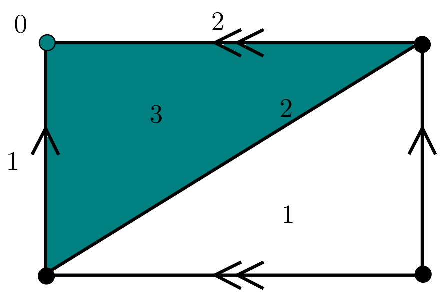
We will denote the sets above as $M_{f}{}^{+}_{\sigma}$ and $M_{f}{}^{-}_{\sigma}$ respectively, omitting the f where the chosen function is clear.

The definition says that given a discrete Morse function, the ‘general’ case is that the value of two cells $\sigma^{(p)} < \tau^{(p+1)}$ will decrease with the dimension of the cells, with possibly one exception.

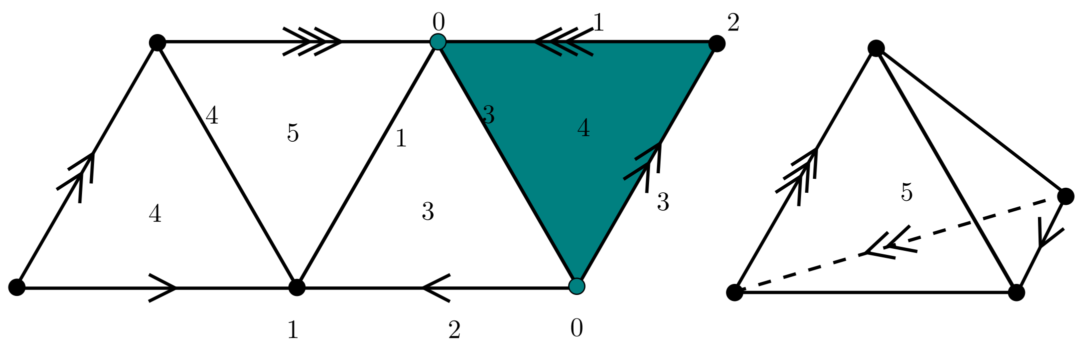
Figure III.1 gives some examples of discrete Morse functions in dimensions 1–3. Note that at most one of the sets $M^{\pm}_{\sigma}$ is non-empty for every cell σ, this is in fact be the general case as is shown by Lemma III.1.5.



(a) S^1



(b) The torus



(c) 3-simplex (right) and identification space for its boundary (left); the 5 in the right figure is the value of the 3-cell.

Figure III.1: Examples of discrete Morse functions on CW complexes; blue cells are critical.

When there are no exceptions we call a cell critical:

Definition III.1.4. A *critical cell* for a discrete Morse function is a cell $\sigma^{(p)}$ so that

- $|M^{+}_{\sigma}| = 0$

- $|M_\sigma^-| = 0$

The *index* of σ is defined to be its dimension.

Looking again at Figure III.1, we can see that the critical cells are given by the blue cells as for each blue cell the adjacent cells with dimension one higher or lower have higher and lower value respectively. In the same way that we defined the Morse vector of a smooth Morse function we define the *discrete Morse vector* of a discrete Morse function f on a d-complex to be the sequence $(c_0, \ldots, c_d)$ where c_i is the number of critical i cells.

Lemma III.1.5. *Suppose f is a discrete Morse function on M. For a non-critical cell $\sigma^{(p)}$ exactly one of M_σ^+ and M_σ^- contains a cell, the other is empty.*

Proof. If $\dim M = 0$ then we have M_σ^- is automatically empty, so we assume $\dim M > 0$. Suppose there is a $\tau^{(p+1)} > \sigma$ so that $f(\tau) \leq f(\sigma)$. By definition σ is a regular cell of τ and then any other $\tilde{\sigma}^{(p)} < \tau$ must have $f(\tilde{\sigma}) < f(\tau) \leq f(\sigma)$. If in addition there is some $\nu^{(p-1)} < \sigma$ with $f(\nu) \geq f(\sigma)$ then by Lemma III.1.2 there is a cell $\tilde{\sigma}^{(p)} \neq \sigma$ with $\nu < \tilde{\sigma} < \tau$. But then we must have $f(\sigma) < f(\nu) < f(\tilde{\sigma})$ giving us a contradiction (we already showed that $f(\tilde{\sigma}) < f(\sigma)$). □

Consider a pair of cells $\sigma^{(p)} < \tau^{(p+1)}$ in a regular CW complex with $f(\sigma) \geq f(\tau)$. By the above lemma we have that $f(\sigma') < f(\tau)$ for all other $\sigma'^{(p)} < \tau$. We then have that all $\nu^{(p-1)} < \tau$ have $f(\nu) < f(\tau)$ unless $\nu < \sigma$ and $f(\nu) < f(\sigma)$. This follows as Lemma III.1.2 tells us there must at least two p-faces of τ of which ν is a face. If ν has a higher value than τ, then by Lemma III.1.5 one of the p-faces of τ containing ν has a higher value than τ, and so must be σ. Note that ν must have lower value than σ by Lemma III.1.5. By inducting on dimension we see that for any $\alpha < \beta$ to have $f(\alpha) > f(\beta)$, M_β^- must contain a cell and α must be a face of this cell.

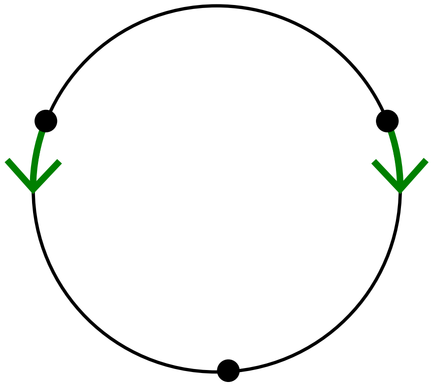
We now define the analogue of the descending gradient of a Morse function. Gradients (or gradient like vector fields) of Morse functions are often much easier to use than the functions themselves and generally contain all the information one needs when working with Morse theory. Given a discrete Morse function f we have seen by Lemma III.1.5 that non-critical cells come in (disjoint) pairs $\sigma^{(p)} < \tau^{(p+1)}$ where $f(\sigma) \geq f(\tau)$. Intuitively we can think of a decreasing gradient as given by arrows $\sigma \to \tau$ between these non-critical pairs. Then every non-critical cell will be either the initial or terminal point of an arrow (but not both) and critical cells will be neither. We will now make this definition rigorous. For a complex M we denote the set of p-cells by M_p.

Definition III.1.6. Given a discrete Morse function f on M, the *discrete Morse gradient* of f is a set of maps of cells $V_f : M_p \to M_{p+1} \cup \{0\}$ (one for each p) so that if $\tau^{(p+1)} > \sigma^{(p)}$ has $f(\sigma) \geq f(\tau)$ then $V(\sigma) = \tau$ and $V(\sigma) = 0$ otherwise.

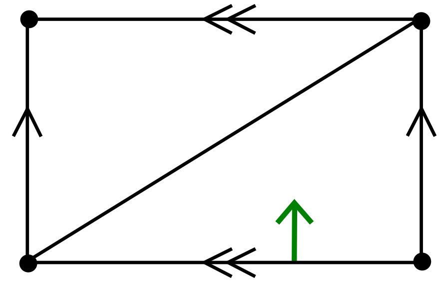
Considering that $M_{\dim M+1}$ is the empty set we see that top-dimensional cells must map to 0. We write $\sigma \prec \tau$ to say that σ and τ are two cells with $V_f(\sigma) = \tau$, we can think of $\sigma \prec \tau$ as meaning there is a 'vector' from σ to τ. This gives us a more compact way of writing the equivalent statement $\sigma^{(p)} < \tau^{(p+1)}$, $f(\sigma) \geq f(\tau)$.

Figure III.2 shows the discrete gradients of the functions in Figure III.1: each 'vector' (pair $\sigma \prec V(\sigma)$) is represented as an arrow from σ to $V_f(\sigma)$. Cells that are not the origin of an arrow are the cells with for which V_f is 0; note that these are either critical cells or in the image of V (i.e. the have an arrow pointing to them).

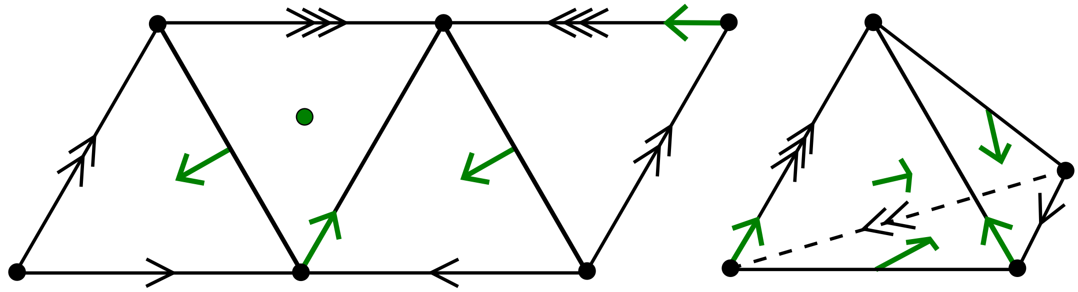
We say f, g are *equivalent discrete Morse functions* on M if they have the same discrete Morse gradient: $V_f = V_g$. This defines an equivalence relation on the collection of discrete Morse functions on M and clearly also respects restriction to subcomplexes of M.



(a) S^1



(b) The torus



(c) 3-simplex (right) and identification space for its boundary (left); the green dot (in the identification space) represent the arrow into the 3-cell (in the right figure).

Figure III.2: Examples of discrete Morse gradients on various cell complexes.

If we consider a cell $\sigma^{(p)}$ with $V(\sigma) = \tau^{(p+1)} \neq 0$ then we have by definition that $f(\tau) \leq f(\sigma)$, and moreover $f(\sigma') < f(\tau) \leq f(\sigma)$ for any $\sigma' < \tau$; in particular the other p-cells of τ have lower value than σ. This indicates how we can define a gradient path, but first we will introduce a slightly more general analogue of a vector field; gradient paths will then be special cases of paths for these vector fields.

The properties of discrete Morse functions can be reinterpreted in terms of their discrete gradients. The unique pairing property given by Lemma III.1.5 says that if $V_f(\sigma) = \tau$, then $V_f(\tau) = 0$ and moreover, $V_f^{-1}(\tau) = \{\sigma\}$. By definition we also have that σ is a regular face of τ if $V_f(\sigma) = \tau$. We enforce these properties to define the more general concept of a discrete vector field.

Definition III.1.7. A *discrete vector field* is a linear map $V : M \to M \cup \{0\}$ so that

- $V(M_p) \subset M_{p+1} \cup \{0\}$
- $V(\sigma) = 0$ if σ is an irregular face of $V(\sigma)$
- $V(\sigma) = 0$ if $\sigma \in \operatorname{im} V$
- Given $\sigma^{(p)}$, $|\{\nu^{(p-1)} \mid V(\nu) = \sigma\}| \leq 1$.

A discrete vector fields captures the local properties of a discrete gradient. We see that a discrete vector field has the unique pairing property of a discrete Morse gradient and a cell can only map to a cell of 1 dimension higher of which it is a regular face. In fact it is clear that any discrete Morse gradient is a discrete vector field. We would like to know when the converse holds: for which discrete vector fields do we actually have the gradient of some discrete Morse function? Knowing this will allow us to use vector fields without reference to a specific discrete Morse function. As we will see, it is really only the equivalence classes of discrete Morse functions (i.e. the discrete gradient) we care about, so knowing that some vector field is a gradient without explicitly writing down a discrete Morse function is advantageous.

Definition III.1.8. A *V-path* is a sequence of cells $\sigma_0^{(p)}, \sigma_1^{(p)}, \ldots, \sigma_n^{(p)}$ so that if $V(\sigma_i) = \tau \neq 0$, then $\sigma_i \neq \sigma_{i+1}$ and $\sigma_{i+1} < \tau$. If $V(\sigma_i) = 0$ then $i = n$. A path is called *closed* if $\sigma_n = \sigma_0$ and $n \neq 0$.

We say that the *length* of a V-path $\sigma_0, \ldots, \sigma_n$ is n. The key observation to make is that for a discrete gradient V_f the value of f must decrease along any V_f-path.

Lemma III.1.9. *Suppose f is a discrete Morse function and $\gamma : \sigma_0, \ldots, \sigma_n$ is a V_f-path. The $f(\sigma_{i+1}) < f(\sigma_i)$ for all $i \in \{0, \ldots, n\}$.*

Proof. We have $f(\sigma_i) \geq f(V_f(\sigma_i)) > f(\sigma')$ for any $\sigma' < V_f(\sigma_i), \sigma' \neq \sigma_i$. In particular $\sigma_{i+1} < V_f(\sigma_i)$ and hence $f(\sigma_{i+1}) < f(\sigma_i)$. □

This immediately implies the following.

Corollary III.1.10. *Given a discrete Morse function f there are no closed V_f-paths.*

This is in fact the only additional property required to ensure that a discrete vector field is a discrete Morse gradient.

Theorem III.1.11 (Forman)**.** *Given a discrete vector field V, $V = V_f$ for some discrete Morse function f if and only if V has no closed V-paths. Moreover f can be chosen to be self-indexing.*

Proof. Only one direction needs to be shown; that a discrete vector field is a discrete gradient for some function if there are no closed paths — the other direction is given by Corollary III.1.10. This is done by inductively constructing an explicit self-indexing discrete Morse function on the cells with dimension up to k, whose discrete Morse gradient is given V; see [12]. □

When V is a discrete gradient, we call a V-path a *gradient path* of V (or of any function that induces V). Gradient paths will in many ways play the role of the flow lines of Morse theory. From now on we will refer to discrete vector fields with no closed path as discrete Morse gradients even without reference to a discrete Morse function, highlighting the fact that we will not depend on a specific function.

The definition of V-paths is a little too restrictive for our purposes. We extend these definitions to allow us to talk of a V-path from $\tau^{(p+1)}$ to $\sigma^{(p)}$. This is in reality a V-path as defined previously starting at some $\sigma_0^{(p)} < \tau$ so that $V(\sigma_0) \neq \tau$. Notice that the value of a discrete Morse function that induces V will be higher on τ than on any cell of the path. This extension is predominately to simplify terminology: for example we can now talk of the set of all gradient paths from some critical $\tau^{(p+1)}$ to some $\sigma^{(p)}$, which is the set of all paths from faces of τ to σ.

Notation for gradient paths

We will pause briefly to introduce some notation that will make it easier to talk about gradient paths in later sections.

We will denote a gradient path from a cell $\alpha^{(p)}$ to a cell $\beta^{(p)}$ by $\alpha \rightsquigarrow \beta$. We will write $\alpha \overset{(\gamma)}{\rightsquigarrow} \beta$ to mean we have a gradient path $\alpha \rightsquigarrow \beta$ and moreover the path is the length one path (α, β) with $W(\alpha) = \gamma$, where W is the gradient we are considering. Recall that we defined a gradient path from a critical cell $\gamma^{(p+1)}$ to a cell $\beta^{(p)}$ to be a gradient path of p-cells from some $\alpha^{(p)} < \gamma$ $(\alpha \not\prec \gamma)$ to β; we will notate this as $\gamma \overset{(\alpha)}{\rightharpoondown} \beta$ and when we do not care about the specific α that begins the path (for example when talking about all paths from γ to β) we will write $\gamma \rightharpoondown \beta$.

By concatenating these notations we indicate a concatenation of gradient paths. For example $\gamma^{(p+1)} \overset{(\alpha)}{\rightharpoondown} \beta \overset{(\mu)}{\rightharpoondown} \nu$ indicates a gradient path of p-cells from the face $\alpha^{(p)} < \gamma$ to β followed by a gradient path of $(p-1)$-cells from $\mu^{(p-1)} < \beta$ to ν and moreover γ and β are critical. When the second path in a concatenation is a path of the form $\alpha \rightsquigarrow \beta$ then the concatenated paths form a new path where we simply append the sequence for $\alpha \rightsquigarrow \beta$ (without the initial α) to the end of the sequence for the preceding path. For example, $\gamma^{(p+1)} \rightharpoondown \alpha \rightsquigarrow \beta$ is a path of the form $\gamma \rightharpoondown \beta$ which will be given by a sequence of p-cells $\alpha_0, \dots, \alpha_{n-1}, \alpha, \beta_1, \dots, \beta_{n-1}, \beta$ where $\alpha_0, \dots, \alpha_{n-1}, \alpha$ is the sequence for $\gamma \rightharpoondown \alpha$ and $\alpha, \beta_1, \dots, \beta_{n-1}, \beta$ is the sequence for $\alpha \rightsquigarrow \beta$. In particular the expressions $\gamma \overset{(\alpha)}{\rightharpoondown} \beta$ and $\gamma \overset{(\alpha)}{\rightharpoondown} \alpha \rightsquigarrow \beta$ mean exactly the same thing as a gradient path of the form $\gamma \overset{(\alpha)}{\rightharpoondown} \alpha$ is necessarily given by the length 0 path α.

The main aim of these notations is to make it easier to recall the relative dimensions that cells of interest have, and describe the set of gradient paths with particular forms in a more compact way. This will be of particular use in Chapter V.

2 Cellular collapse

In this section we will introduce a tool that is used widely in discrete Morse theory, cellular collapse, and discuss the links between cellular collapse and discrete Morse gradients. In particular cellular collapse defines a discrete Morse gradient, which makes it a useful tool for discussing the existence of discrete Morse gradients on a complex.

Given a simplicial cell complex C, a *free face* of C is a cell σ such that σ is a face of only cell τ. Given a free face, an *elementary collapse* is the removal of a free face (which in turn requires the removal of τ) resulting in a subcomplex C'; we write $C \searrow_\sigma C'$ and say that C' is the result of collapsing the free face σ in C. Note in particular that $|C'| = |C| \setminus (\mathring{\sigma} \cup \mathring{\tau})$ and $C' = C \setminus \{\sigma, \tau\}$. We can also collapse free faces that are regular cells in CW complexes using the same approach.

More generally, we write $C \searrow C'$ and say C collapses to C' if there is a sequence of elementary collapses $C \searrow_{\sigma_1} C_1 \searrow_{\sigma_2} \cdots \searrow_{\sigma_n} C_n = C'$. Figure III.3 shows a sequence of collapses taking the 2-simplex to a vertex; notice that there is more than one way to do this. Cell collapse provides us with a way to detect when a subcomplex $C' \subset C$ is a deformation retract of C as if $C \searrow C'$ then C' is a deformation retract of C.

Definition III.2.1. A cell complex is *collapsible* if it collapses to a single vertex.

Collapsible complexes are contractible spaces. Note that the converse is not true: there are non-collapsible cell complexes homeomorphic to contractible spaces. A simple example of a non-collapsible cell complex that is contractible is the *dunce hat* [30]. This is obtained

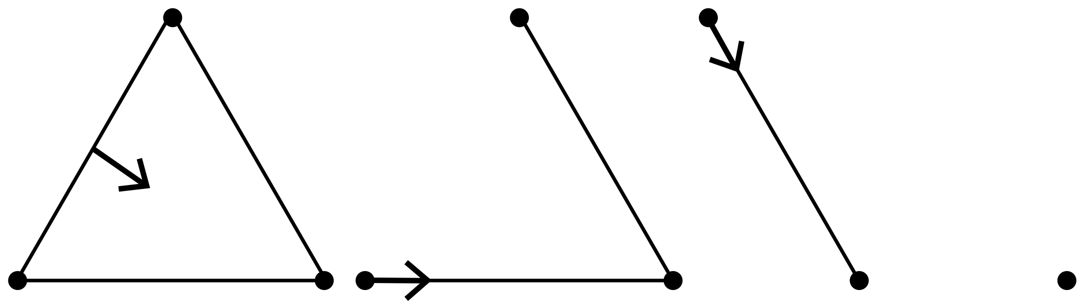


Figure III.3: A collapse sequence reducing a 2-simplex to a vertex; the arrows begin at the free face used for an elementary collapse.

from a 2-simplex given by vertices a, b, c by identifying the edges $ab \sim ac \sim bc$. This results in a space that is clearly not collapsible as there is no free face (the only 1-cell is not a regular face of the 2-cell). The dunce hat is contractible: as a CW complex it is a disk glued to a 1-cell e along eee^{-1}, which is homotopy equivalent to attaching along e, hence the dunce hat has trivial fundamental group.

A sequence of collapses $C \searrow_{\sigma_1} C^1 \ldots \searrow_{\sigma_n} C^n$ induces a discrete vector field V on C. First note that at each collapse step the unique cell for which σ_i is a face (denote it τ_i) must have dimension $\dim \sigma_i + 1$, so we set $V(\sigma_i) = \tau_i$ and $V(\sigma) = 0$ for all $\sigma \in C \setminus \{\sigma_i, \tau_i\}_{i=1}^n$.

Lemma III.2.2. *The vector field V defined by a sequence of collapses $C^0 \searrow_{\sigma_1} C^1 \searrow_{\sigma_2} \cdots \searrow_{\sigma_n} C_n$ is a discrete Morse gradient.*

Proof. By construction each cell σ satisfies exactly one of

1. σ is the image of one σ_i
2. $\sigma = \sigma_i$ for one i
3. σ is critical ($V(\sigma) = 0$ and $\sigma \notin \operatorname{im} V$).

This means we have a well defined vector field, and by Theorem III.1.11 we need only show that there are no non-stationary closed paths. This follows because we use free faces. Suppose $\sigma^{(p)}$ is a free face of C that appears in the sequence of collapses (note that it does not have to be C^0) so that its collapse removes $\tau^{(p+1)}$. This means that for any V-path in which σ appears, σ can only be the first cell of the path. This follows as for σ to appear anywhere else we would need some $\sigma'^{(p+1)}$ in the sequence to collapse τ as well (so that $V(\sigma') = \tau$), but this is impossible.

The converse to this says that if a cell is a non-initial cell of a V-path, then it is not a free face of C. For any closed path we can choose any cell to be the initial cell and hence no cell of a closed path can be a free face of C. Consider a closed path γ in C_i, the same argument says that no cell of this path can contain a free face of C_i. Consider the collapse $C_i \searrow_{\sigma_i} C_{i+1}$; we must have that γ is contained in C_{i+1}. By induction on the collapse sequence no element of the collapse sequence can be part of a closed path, but as every path arises from collapses (that is a V-path can only contain the cells σ_i) there cannot be any closed paths. $\square$

Remark III.2.3. The link between collapse and discrete gradients is deeper than that given by Lemma III.2.2. In addition to a sequence of collapses defining a discrete Morse gradient we also have that a discrete gradient gives us a recipe for reducing a cell complex M down to a subset of its 0-cells. While this process is implicitly suggested by Benedetti (e.g. in [5]) we could not find a proper description of this process. We provide our own here.

We begin by removing the top-dimensional critical cells from the complex to obtain M'. Then we perform the collapses given by gradient paths starting at free faces of M'. To see that we can collapse all of the remaining top-dimensional cells in this way consider a top-dimensional cell $\tau_0 \subset M'$; we will show that there is a gradient path from a free face of M' to the $(\dim M - 1)$-face of τ_0 with a higher value than τ_0.

There is some unique σ_0^τ so $\sigma_0^\tau \prec \tau$. There are two cases for σ; it is either free in M', in which case we are done, or there is some cell $\tau_{-1} \subset M'$, not τ_0, so that $\sigma_0^\tau < \tau_{-1}$ which by definition of M' is non-critical. We then have the same situation for τ_{-1} as for τ. It follows that there must be a path $\sigma_{-n}^\tau, \ldots, \sigma_0^\tau$ where σ_{-n}^τ is free a free face of τ_{-n} in M'. Note that τ_{-i} and hence the gradient path may not be uniquely determined for general complexes but for a triangulated manifold each non-free codimension 1 cell is a face of precisely two top-dimensional cells and hence the gradient path is unique for a triangulated manifold.

We want to perform the collapses given by this gradient path — $M_1 \searrow_{\sigma_{-n}^\tau} \cdots \searrow_{\sigma_0^\tau} M_1'$ — but after performing an elementary collapse, the next cell in the gradient path may not be free, see Figure III.4. To collapse along this gradient path we may need to collapse other cells first as shown in Figure III.4. The key point is that even if the next cell in a path σ_{-i} is not free after an elementary collapse, then there is only one way it 'wants' to collapse (into τ_{-i}) and the other cells that make it non-free can be collapsed earlier so as to make σ_{-i} free.

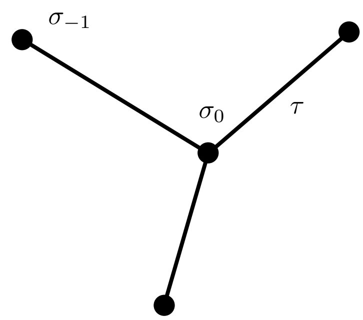


Figure III.4: After performing the collapse $\searrow_{\sigma_{-1}}$ we can see that σ_0 is not free.

We know in particular that there must be some codimension 1 free faces of M' that points to (has as its image under V a) top-dimensional faces. If we perform all these collapses, then either we have collapsed all top-dimensional faces or we obtain a new subcomplex that by the previous arguments must again have free faces that point to other top-dimensional faces. So by inductively collapsing all the free faces we can always reduce the number of top-dimensional faces, and hence remove all of them. In this way we have collapsed along all the gradient paths from free faces of M' to remove all the codimension 0 cells.

Theorem III.2.4. *Given a discrete gradient on an d-complex M, let M' be the subcomplex obtained by removing all critical d-cells from M. Then $M' \searrow M_1$ where M_1 is a subcomplex of M with strictly lower dimension (it has codimension 1 if there is a codimension 1 cell in M) and the elementary collapses used are given by the discrete gradient.*

Proof. This is proved by the preceding discussion. □

By induction on the dimension of M we obtain the following corollary.

Corollary III.2.5. *Given a discrete gradient on a d-complex M, the gradient describes a way to reduce M to its critical 0-cells by repeated application of Theorem III.2.4.*

Figure III.5 gives an example of using this procedure to reduce a simplicial complex onto its critical vertices. Each step is either given by the collapse of all free faces (indicated by $\searrow$) or is given by removing the top-dimensional critical cells of the preceding subcomplex (indicated by $\rightsquigarrow$).

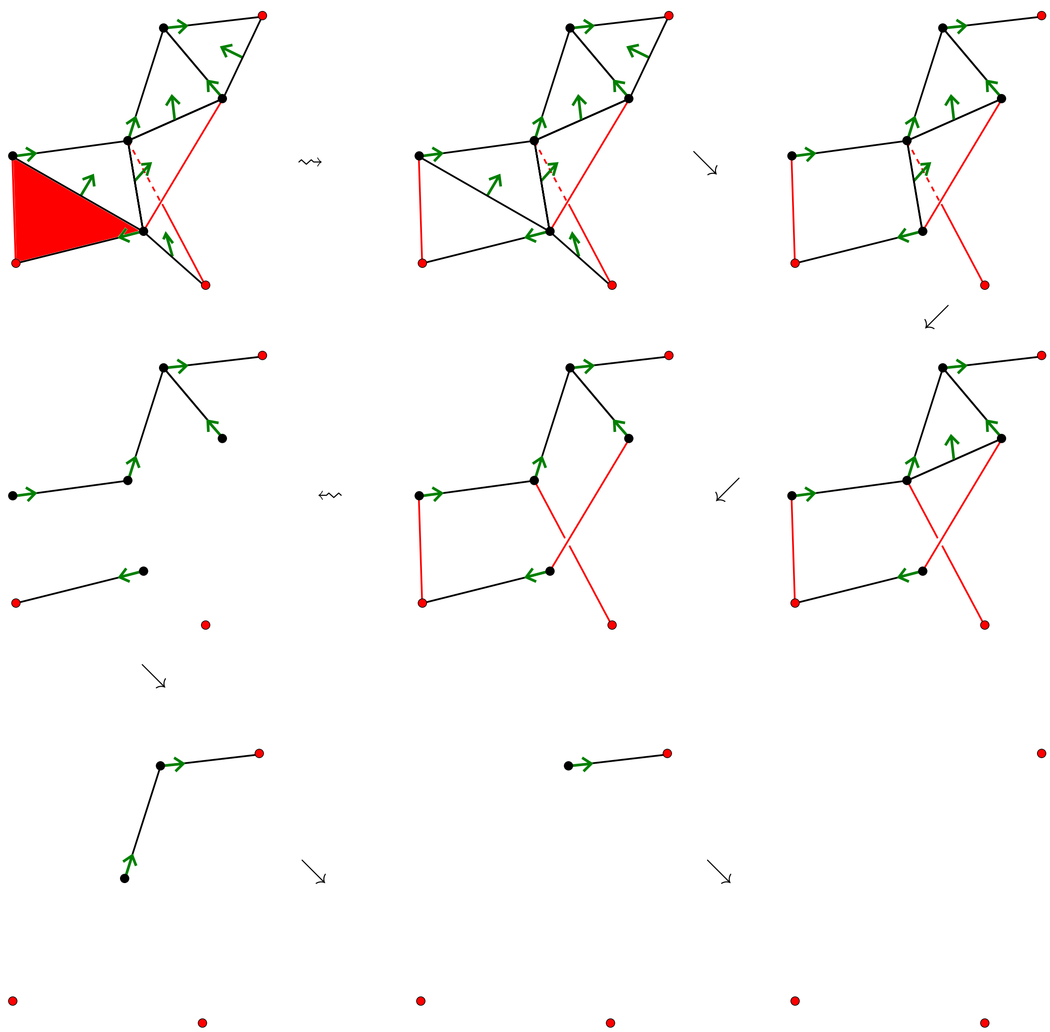


Figure III.5: Collapsing a simplicial complex using a discrete Morse gradient; critical cells are red and the steps given by $\rightsquigarrow$ are the removal of critical cells.

Discrete structural theorems

The structural theorems of smooth Morse theory (Theorems I.4.9 and I.4.10) have direct analogues in discrete Morse theory [12]. Let M be a CW complex with a discrete Morse function f; define $M(a)$ to be the set of cells σ so $f(\sigma) \leq a$ and all the faces of such cells.

Suppose M and f satisfy:

- For every pair of cells σ, τ, if τ is in the smallest subcomplex of M that contains σ and τ is not a face of σ, then $f(\tau) \leq f(\sigma)$.

- For $p, r \geq 0$, whenever $\tau^{(p+r+1)} > \sigma^{(p)}$ and $f(\tau) < f(\sigma)$ there is some $\tilde{\tau}^{(p+1)}$ so $\tilde{\tau} > \sigma$ and $f(\tilde{\tau}) \leq f(\tau)$.

Then we say M satisfies the *discrete Morse hypothesis.* The first of these conditions ensures that $M(a)$ is a subcomplex for any a, and the second condition tells us that to determine whether $\sigma^{(p)}$ is in $M(a)$ that we only need to look at the cells of dimension $p+1$.

We can now state the analogues of Theorems I.4.9 and I.4.10.

Theorem III.2.6. *Suppose M is a CW complex with a discrete Morse function f satisfying the discrete Morse hypothesis and $a < b$ are numbers so that no critical cell of f have value in $[a, b]$; then $M(b) \searrow M(a)$*

Theorem III.2.7. *Suppose M is a CW complex with a discrete Morse function satisfying the discrete Morse hypothesis and $a < b$ are such that there is a single critical point, $\sigma^{(p)}$, with value in $[a, b]$. Then $M(b)$ is homotopy equivalent to $M(a) \cup_{\partial e^p} e^p$, that is $M(a)$ with a p-cell attached along its boundary.*

For proofs see [12]. Note that any regular CW complex with a discrete Morse function will satisfy the discrete Morse hypothesis.

3 Cobordism

A *smooth cobordism* is a triple of manifolds (the first with boundary) (M, N^-, N^+) so that $\partial M = N^- \cup N^+$ and $N^- \cap N^+ = \varnothing$. We have already seen such objects in the context of smooth Morse theory — given a smooth Morse function the submanifolds ${}^aM$, ${}^b_aM$ are cobordisms $\big(f^{-1}(-\infty, a], \varnothing, f^{-1}(a)\big)$ and $\big(f^{-1}[a, b], f^{-1}(a), f^{-1}(b)\big)$ respectively. We define a Morse function on a cobordism (M, N^-, N^+) to be a Morse function on M so that there are no critical points on ∂M. As we have seen it can be useful to restrict focus to one such cobordism at a time.

In the discrete case we have an analogous object, for which we must modify the definition of a discrete Morse function slightly.

Definition III.3.1. A *cellular triad* is a triple (M, N^-, N^+) of cell complexes so that, $N^\pm$ is a subcomplex of M, $N^- \cap N^+ = \varnothing$ and every cell $\sigma^{(p)} \in N^\pm$ is a face of a unique cell $\tau^{(p+1)} \not\subset N^- \cup N^+$ and moreover σ is a regular face of τ. We will call N^- the *lower complex* and N^+ the *upper complex*

By taking $N^\pm = \varnothing$ we see that a cell complex is a special case of a cellular triad and so any results or definitions involving cellular triads will hold for complexes as well.

Definition III.3.2. A discrete Morse function on a cellular triad is a function $f : M \to [a, b]$ for some $a, b \in \mathbb{R}$ so that $f^{-1}(a) = N^-$, $f^{-1}(b) = N^+$ and for any cell not in $N^- \cup N^+$ f is a discrete Morse function in the sense of Definition III.1.3 (note that cells in $M^\pm_\sigma$ may be in $N^- \cup N^+$). We define the critical points of f to be the cells not in $N^- \cup N^+$ that satisfy the conditions of Definition III.1.4.

These definitions of a discrete Morse function and its critical points are the same as the definitions for a CW complex when we take $N^- \cup N^+ = \varnothing$. When $N^\pm$ is not empty we have a similar structure to smooth cobordisms defined by ranges of a Morse function, with 'top' and 'bottom' subcomplexes; the top, M^+, has non critical cells.

For a discrete gradient we must make a similar modification to the standard definition. We require that for a cell σ in N^+ the image of that cell is the unique cell τ given by Definition III.3.1. We use the standard definition on $M \setminus (N^- \cup N^+)$, and set the gradient to be constantly 0 on N^-. Note that no cell in N^+ maps to zero and that no cell in N^- is in the image of the gradient: adding these two properties to the definition of a discrete vector field we get the generalisation of Definition III.1.7 to cellular triads. Figure III.6 gives examples of these constructions for cellular triads. Notice that the restriction of a discrete Morse function on (M, N^-, N^+) to $N^\pm$ will usually not be a discrete Morse function on $N^\pm$. Theorem III.1.11 also holds for cellular triads.

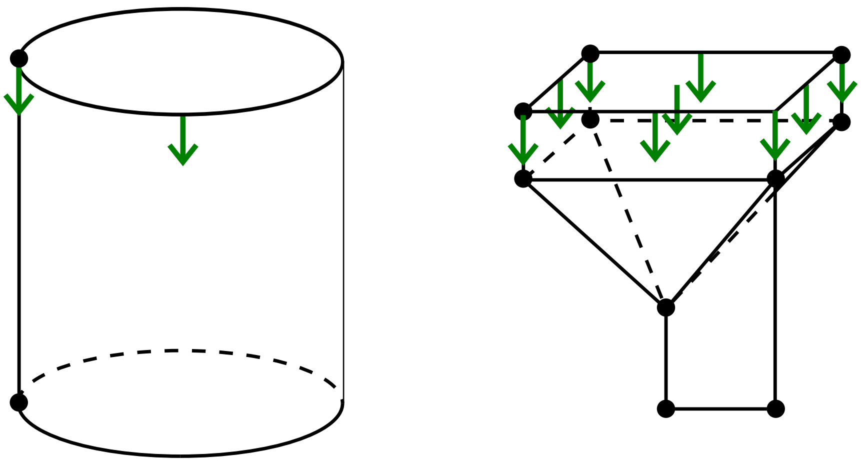


Figure III.6: Examples of two cellular triads, a cylinder (left) and 3-complex (right); the upper and lower complexes are the cells at the very top and bottom. The arrows represent the requirements of a discrete vector field on the upper complexes. Note that unlike a cobordism the upper and lower complexes need not have the same dimension (shown by the cellular triad on the right).

We can immediately notice that for a cellular triad all the cells of N^+ can be collapsed. This follows by definition because every face $\sigma^{(p)} \subset N^+$ a regular face of a unique $\tau^{(p+1)} \subset M \setminus (N^- \cup N^+)$. So we have that the $(\dim N^+)$-cells of N^+ are free faces and we can then collapse the cells of N^+ in order of decreasing dimension (after collapsing all $i+1$-cells, the i-cells in N^+ become free). It follows from Theorem III.2.4 that if there are no critical cells for a discrete Morse function on a cellular triad (M, N^-, N^+) then $M \searrow N^-$.

As the gradient on the bottom of a cellular triad is 0, we can stack two cellular triads $(M_1, N^-, N_1), (M_2, N_2, N^+)$ where the top of M_1 is glued to the bottom of M_2 (suppose N_1 and N_2 are combinatorially equivalent) resulting in the cellular triad $(M = M_1 \sqcup M_2 / N_1 \sim N_2, N^-, N^+)$. We will denote the subcomplex of M that is the gluing interface by N. If we have vector fields on both triads, then we can define a vector field on the new triad by taking the vector fields of either triad on the respective parts; on the interface we take the vector field from M_1.

We can see that if we had two discrete gradients on the triads then the gradient that we get after gluing is also a discrete gradient. To get a function showing the vector field is a gradient pick a function on each of the original triads f_1 on (M_1, N^-, N_1) and f_2 on (M_2, N_2, N^+) so that the vector fields on the two triads are the gradients of these function; we can assume that $f_1(N_1) = -\dim M_1$ and $f_2(N_2) = 0$. We define f on $((M_1 \sqcup M_2)/(N_1 \sim N_2), N^-, N^+)$ by $f|_{M_2 \setminus N} = f_2|_{M_2 \setminus N}$ and $f|_{M_1 \setminus N} = f_1|_{M_1 \setminus N}$. On N we then set $f|_N = -\operatorname{codim}$. We see immediately that this function's gradient is the vector field constructed above.

We conclude this section by introducing some terminology that will be helpful in later chapters.

Definition III.3.3. A *boundary critical discrete Morse function* is a discrete Morse function $M \to \mathbb{R}$ so that every cell in the boundary of M is critical.

We define the *interior Morse vector* for a boundary critical discrete Morse function on a d-complex to be $(c_0, \ldots, c_d)$ where c_i is the number of interior critical cells.

Remark III.3.4. Given a boundary critical discrete Morse function on M, the process of Corollary III.2.5 can be adapted to collapse M onto its boundary by not removing the (critical) boundary cells.

4 The set of discrete Morse functions on a complex

As we have already seen, in smooth Morse theory it is often useful to know that certain Morse functions exist, and to know how they can be related to each other. In the smooth case this is studied by Cerf theory. Ideally one would like to have a homotopy $F : M \times [0,1] \to \mathbb{R}$ between two Morse functions f_0 and f_1 so that $F|_{M\times\{t\}}$ is a Morse function for all $t \in [0,1]$, we will call this a *homotopy through Morse functions*. This is not always possible. As a manifold admits Morse functions with different numbers of critical points, we need to consider homotopies that include births and deaths of critical points. Births and deaths necessarily violate the non-degeneracy condition of Morse functions; at best we will have isolated cubic singularities. If a function would be Morse except for finitely many such singularities we say that it is a *generalised Morse function*. The space of generalised Morse functions is connected; in particular there is a homotopy through generalised Morse functions between any two Morse functions. The space of Morse functions is not connected on the other hand, as the number of critical points in a connected component is constant [17].

In this section we will explore the analogous question in the discrete case. We see that the discrete behaviour is much simpler: we will show that given two discrete Morse functions, there is a homotopy through discrete Morse functions between them. We will also examine how birth and death work in discrete Morse theory and show that the operations of birth and death can be modified to allow us to 'move' a critical point.

4.1 Moving between discrete Morse functions

It is often easier to consider the vector fields of smooth Morse functions or even the flow lines entering and leaving critical points, rather than the functions themselves. With this view we can ask whether there is a homotopy through Morse functions that get us from one of these vector fields to another. Note that if there is a gradient vector field X that is gradient-like for two (smooth) Morse functions f_0, f_1, we can see that the obvious homotopy $f_t = (1-t)f_0 + tf_1$ is a homotopy between f_0 and f_1 and that X is gradient-like for any f_t.

In the discrete case the question of the existence of homotopies through discrete Morse functions is much simpler. As we will see it is possible to view two Morse vector fields as being related through a 'straight line homotopy' of functions that induce those vector fields. To do this we use the notion of a 'flat' discrete Morse function.

Definition III.4.1. A discrete Morse function f is *flat* if for all $\sigma^{(k)} < \tau^{(k+1)}$ with $f(\sigma) \geq f(\tau)$ we have equality, i.e. $f(\sigma) = f(\tau)$ if $\sigma \prec \tau$.

Some authors build this requirement into their definition of a discrete Morse function. This can be done due to the following result.

Proposition III.4.2. *Any discrete Morse function is equivalent to a flat discrete Morse function.*

Proof. Let f be a discrete Morse function. We want to construct a discrete Morse function $\bar{f}$ so that for all cells $\sigma < \tau$, $f(\sigma) \geq f(\tau)$ if and only if $\bar{f}(\sigma) = \bar{f}(\tau)$. Consider two cells $\sigma^{(p)} \prec \tau^{(p+1)}$, then we have for all $\sigma'^{(p)}$ with $\sigma \neq \sigma' < \tau$ that $f(\sigma') < f(\tau)$ and for any other $\tau'^{(p)}$ with $\sigma < \tau'$, we have $f(\sigma) < f(\tau')$. It follows that σ could take any value in $[f(\tau), f(\sigma)]$ and we would still have a discrete Morse function with the same discrete gradient as the sets M^+_σ, M^-_τ, and $M^+_{\sigma'}, M^-_{\tau'}$ would remain the same. So by setting $f'(\sigma) = f(\tau)$ we have a discrete Morse function with the same gradient. As all pairs such pairs $\sigma \prec \tau$ are disjoint, we can do this for every pair of cells $\sigma \prec \tau$ to produce $\bar{f}$. □

Combining with Theorem III.1.11, this lemma show us that any discrete gradient is the discrete gradient of a flat discrete Morse function.

Lemma III.4.3. *If f, g are flat discrete Morse function on M then $f_t = (1-t)f + 1g$ defines a discrete Morse function for all $t \in [0,1]$. Moreover for $t \in (0,1)$, $f_t(\sigma) = f_t(\tau)$ if and only if $f(\sigma) = g(\sigma) = f(\tau)$ i.e. all f_t are equivalent for $t \in (0,1)$.*

Proof. This follows directly from the definitions; note that for $\sigma < \tau$ we have $f(\sigma) = f(\tau)$ if and only if $\dim \sigma = \dim \tau - 1$ and $V_f(\sigma) = \tau$, similarly $f(\sigma) < f(\tau)$ if and only if $V_f(\sigma) \neq \tau$. These properties follow from the fact that f and g are flat. We have three cases to consider, corresponding to whether there is a vector from σ to τ for both f and g, just one of them, or neither of them.

Suppose that we have $V_f(\sigma) = V_g(\sigma) = \tau$. Then

$$f_t(\sigma) = (1-t)f(\sigma) + tg(\sigma) = (1-t)f(\tau) + tg(\tau) = f_t(\tau)$$

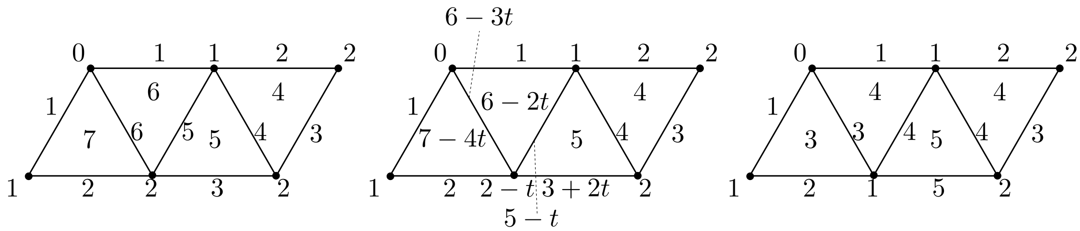
for all t. For the case where there is a vector for just one of the functions, assume without loss of generality that $V_f(\sigma) = \tau$ and $V_g(\sigma) \neq \tau$ then $f_t(\sigma) = (1-t)f(\tau) + tg(\sigma) < f_t(\tau)$ for all $t \in (0,1)$. Finally if $V_f(\sigma) \neq \tau$ and $V_g(\sigma) \neq \tau$ then $f_t(\sigma) < f_t(\tau)$ for all t. So we see that for all $t \in (0,1)$ $f_t(\sigma) = f_t(\tau)$ if and only if the same is true of both f and g. It follows that f_t is a discrete Morse function for all t and that for $t \in (0,1)$ all f_t are equivalent with the set of gradient paths of V_{f_t} being the intersection of the sets of gradient paths for V_f and V_g. □

With this proposition we see that the question of whether there is a homotopy of Morse functions that will get us from one discrete Morse gradient field to another is relatively trivial; there is always a homotopy through discrete Morse functions that will suffice. Figure III.7 shows a homotopy between two flat discrete Morse functions. We can see that the property of such a homotopy that f_t are all equivalent for $t \in (0,1)$ results in three discrete gradients appearing, the initial gradient of f, the final gradient of g and an intermediate gradient whose arrows are exactly the arrows common to both V_f and V_g.

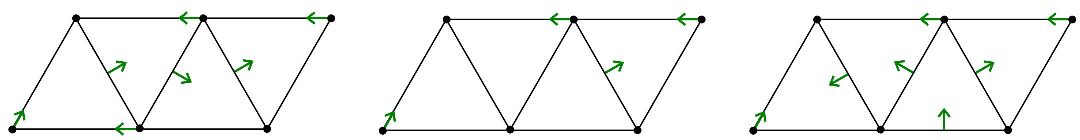
Similarly, the question for functions has a simple answer as well: there will be a homotopy between any two discrete Morse functions. This follows by 'flattening' both functions via homotopies through discrete Morse functions and then applying Lemma III.4.3. It is clear from the remarks made in the proof of Proposition III.4.2 and lemma III.4.3 that given any discrete Morse function f equivalent to a flat Morse function $\bar{f}$, the homotopy $(1-t)f + t\bar{f}$ will flatten f to $\bar{f}$ and the discrete Morse gradient will be constant with respect to t.

4.2 Birth and death

Similarly to cancellation in the smooth case (recall Theorem I.4.15), there is a simple process to cancel two adjacent-dimension critical cells and a simple criterion for detecting when this is possible.



(a) The function along the homotopy; (left) the original function f_0, (right) the final function f_1, (centre) the function f_t



(b) The gradient along the homotopy: (left) the gradient of f_0, (right) the gradient of f_1, (centre) the gradient of f_t for $t \in (0,1)$. Note that all functions f_t for $t \in (0,1)$ are equivalent (have the same gradient).

Figure III.7: A homotopy f_t from f_0 to f_1 given by $f_t = (1-t)f + tg$: (a) shows the functions f, f_t and f_1; (b) shows the corresponding discrete Morse gradients with $t \in (0,1)$.

Proposition III.4.4 (Forman)**.** *Suppose V is a discrete gradient on M so that for two critical cells $\sigma^{(p)}$, $\tau^{(p+1)}$ there is a unique gradient path $\gamma : \tau \rightharpoondown \sigma = \sigma_0, \ldots, \sigma_n = \sigma$ between them and σ is a regular face of σ_{n-1}, then the discrete vector field $\bar{V}$ defined by*

- $\bar{V}|_{M\setminus\gamma} = V|_{M\setminus\gamma}$
- $\bar{V}(\sigma_0) = \tau$
- $\bar{V}(\sigma_{i+1}) = V(\sigma_i)$ *for* $i \in \{0, \ldots, n-1\}$

is a discrete Morse gradient. Moreover there is a unique $\bar{V}$-path $\sigma \rightsquigarrow \sigma_0$.

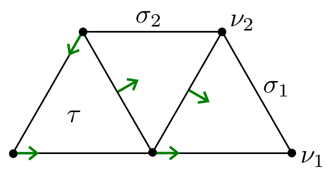
Proof. From the definition of $\bar{V}$ it is clear that the critical cells of $\bar{V}$ are exactly the critical cells of V minus $\{\tau, \sigma\}$.

Note that it is clear that $\bar{V}$ is in fact a discrete vector field. By Theorem III.1.11, to prove the proposition we just need to show that there are no closed $\bar{V}$-paths. As $\bar{V}$ differs from V only on γ we have that there is no closed $\bar{V}$-path in $M \setminus \gamma$ as then this would also be a closed V-path and clearly a closed $\bar{V}$-path cannot be contained in γ. So any closed $\bar{V}$ path must contain a segment $\sigma_i, \delta_0, \ldots, \delta_r, \sigma_j$ where $\delta_k \notin \gamma$. We have $V(\sigma_{i-1}) = \bar{V}(\sigma_i)$, so $\sigma_0, \ldots, \sigma_{i-1}, \delta_0, \ldots, \delta_r, \sigma_j, \ldots, \sigma_n$ would be a second V-path from τ to σ, contradicting the assumption that there is only one such gradient path.

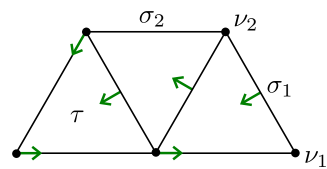
To see that the only one $\bar{V}$-path from σ_n to σ_0, is $\sigma_n, \ldots, \sigma_0$, note that any other gradient path must have a segment of the form $\sigma_i, \epsilon_0, \ldots, \epsilon_l, \sigma_j$ where $\epsilon_k \notin \gamma$ and $j < i$ (as otherwise we would have a closed $\bar{V}$-path), but then the V-path $\epsilon_0, \ldots, \epsilon_l, \sigma_j, \sigma_{j+1}, \ldots, \sigma_{i-1}, \epsilon_0$ is a closed V-path. $\square$

The relationship between $\bar{V}$ and V, $\bar{V}$ is intuitively given by 'flipping the arrows of V' along γ; this is demonstrated in figure Figure III.8. We call the effect of moving from V to $\bar{V}$ the 'death' of the critical cells σ, τ.

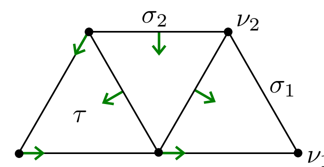
The following gives us the notion of birth for a discrete Morse function.



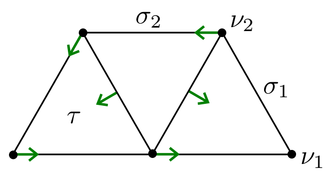
(a) The base discrete Morse gradient with critical cells $\tau, \sigma_1, \sigma_2, \nu_1, \nu_2$.



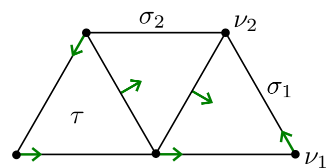
(b) Cancellation of τ and σ_1.



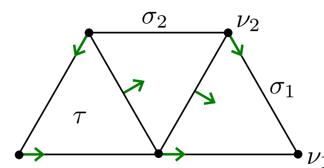
(c) Cancellation of τ and σ_2.



(d) Cancellation of σ_2 and ν_2.



(e) Cancellation of σ_1 and ν_1.



(f) Cancellation of σ_1 and ν_2.

Figure III.8: Examples of cancelling critical points; Figures III.8(b) to III.8(f) show the cancellation of all cancellable pairs.

Proposition III.4.5. *Given a discrete Morse function f with discrete gradient V on M, suppose there is a exactly one gradient p-path $\gamma : \sigma_0 \rightsquigarrow \sigma_n = \sigma_0, \sigma_1, \ldots, \sigma_n$ starting at σ_0 so that $V(\sigma_n) = \tau^{(p+1)}$. Define $\tilde{V}$ by*

- $\tilde{V}|_{M \setminus \gamma} = V$,
- $\tilde{V}(\sigma_{i+1}) = V(\sigma_i)$ *for* $i = 0, \ldots, n-1$, *and*
- $\tilde{V}(\sigma_0) = 0$.

Then $\tilde{V}$ is a discrete vector field with no closed $\tilde{V}$-paths. Moreover, the critical cells of $\tilde{V}$ are the critical cells of V along with $\{\sigma_0, \tau\}$ and there is a unique $\tilde{V}$ path $\tau \rightharpoondown \sigma_0$.

Proof. The proof essentially the same as Proposition III.4.4. □

The proofs for Propositions III.4.4 and III.4.5 work just as well for cellular triads. Combining Propositions III.4.4 and III.4.5 we have the following corollary.

Corollary III.4.6. *Critical cells created by birth can be cancelled, and two cancelled critical cells can be made critical again by birth.*

It is also possible to employ the same arrow flipping idea of birth and death to move a critical cell. Suppose $\tau^{(p+1)}$ is a critical cell and $\sigma^{(p)}$ a cell so that $V(\sigma) = \tau'$ and there is a unique gradient path from $\tau \rightharpoondown \sigma$, then reversing V along this path (as is done for birth and death) we can see that τ becomes non-critical and τ' becomes critical. Because there was only one gradient path from τ to σ after reversing the gradient path we still have a discrete gradient, and moreover the reversed path will be the only gradient path $\tau' \rightsquigarrow \tau$. We can think of this as the critical cell 'moving' from τ to τ'.

We can also move a critical cell by using a path in the same dimension. Suppose we have a unique gradient path from some non-critical $\sigma^{(p)}$ to a critical $\sigma'^{(p)}$. By reversing the gradient along this path σ becomes critical and σ' becomes non-critical in a similar fashion.

Chapter IV

Linking smooth and discrete Morse theory

As we have seen there are many objects in discrete Morse theory that resemble objects in smooth Morse theory. The data needed to described these discrete structures is much simpler however and as it is discrete can make it easier to do computations particularly with a computer. If we have some smooth manifold with a Morse function, it may then be useful to be able to describe this structure in terms of discrete Morse theory, to simplify the information we have. In this section see two different ways that smooth structure can be made discrete. The first is due to Benedetti [5] who showed that given a Morse function f on a smooth manifold, then any PL triangulation of the manifold can be subdivided (in a nice way) so that it supports a discrete Morse function so that the critical cells of a given dimension n are in bijection with the critical points of index n. The second approach we discuss is a construction by Gallais [13] that uses the original smooth Morse function (with a gradient) explicitly to produce a PL triangulation that follows along the unstable manifolds of f and a discrete Morse function where each critical cell of dimension n contains a single critical point of f, and the index of this critical point is n.

1 Polyhedral cell complexes

A *Euclidean cell* is the convex hull of finitely many points in some $\mathbb{R}^n$. Given a set of points $p_0, \ldots, p_m$, we denote their convex hull by $\operatorname{conv}(p_0, \ldots, p_m)$. If the points are linearly independent (that is the vectors $p_i - p_0$ are linearly independent for $i \neq 0$) then the convex hull is an m-dimensional *simplex*. A *face* of a Euclidean cell E is a Euclidean cell given by a subset of the vertices of E. Simplices are the generalisations of triangles to arbitrary dimensions: a 0-, 1-, 2-, 3-simplex is a point, line interval, triangle, tetrahedron respectively.

A *polyhedral complex* is a collection of Euclidean cells in some $\mathbb{R}^k$ so that each face of a cell is in the collection and the intersection of any two cells is a common face (or empty). We will mostly deal with *simplicial complexes* from here on which we define as polyhedral complexes where every cell is a simplex. Note that polyhedral complexes are CW complexes. We will also assume from now on that the polyhedral complexes we deal with are finite (have finitely many cells); in particular this implies that they are compact. Note that in a simplicial complex each n-cell is uniquely determined by its $n + 1$ vertices.

Given a cell complex C we denote by $|C|$ its *underlying set*, $|C| = \bigcup_{\sigma \in C} \sigma$. We will often want to consider a cell complex without a given cell (or perhaps without many cells). We define the *cell difference* of a complex C and cell σ to be $C - \sigma = C \setminus \{\tau \in C \mid \sigma \subset \tau\}$. That is, $C - \sigma$ is the maximal subcomplex of C that does not contain σ; this is how we generalise the cell difference to any set of cells. If we have a set of cell D then $C - D$ is the maximal subcomplex of $C \setminus D$, this is $C \setminus D'$ where $D' = \{\tau \in C \setminus D \mid \exists \sigma \in D', \sigma \subset \tau\}$. To see that this is indeed the maximal subcomplex, note that for any cell τ disjoint from the cells of D, every face of τ is also disjoint from D, so by only removing the cells that contain cells of D we get a subcomplex. It is clear that we need to remove all of D' and hence $C \setminus D'$ is the maximal subcomplex in $C \setminus D$.

Intuitively the cell difference takes a subset of cells and then removes all higher dimensional cells that contain anything in the subset. The number of cells removed correlates inversely with the dimension of the cells in the subset: removing a point removes all cells containing it, whereas if we remove a top-dimensional face nothing else is taken. See Figure IV.1 for examples of cell differences; the second column shows the cell difference of the green cells in the left column with the complex.

Definition IV.1.1. For a cell σ in a polyhedral cell complex C we define the

- *star* of σ, $\operatorname{star}(\sigma, C)$ to be the minimal subcomplex of C containing every cell that contains σ.
- *link* of σ to be $\operatorname{link}(\sigma, C) = (C - \sigma) \cap \operatorname{star}(\sigma, C)$
- *join* of σ and a cell τ to be $\sigma * \tau = \{\operatorname{conv}(\sigma', \tau') \mid \sigma' \in \sigma, \tau' \in \tau\}$.

When the complex C is implicit we will denote the star and link by $\operatorname{star} \sigma$ and $\operatorname{link} \sigma$ respectively.

Intuitively the star of σ is the smallest neighbourhood of σ that is a subcomplex and the link is the boundary of the star (other than any part of the boundary that contains σ). The star, link and join are all complexes. To construct the star we take the cells of C that contain σ as well as all faces of these cells (to make it a complex); alternatively we have $\operatorname{star} \sigma = \operatorname{link} \sigma * \bar{\sigma}$ where $\bar{\sigma}$ is defined to be subcomplex given by σ with all its faces.

Figure IV.1 gives examples of stars and links; each row depicts the cell difference, star and link of the green cell(s).

One key concept when dealing with polyhedral complexes is being able to subdivide the cells of a complex. If we think of a polyhedral complex as a way of partitioning the underlying space, then a subdivision gives us a finer partition.

Definition IV.1.2. Given two polyhedral complexes C, D with the same underlying space, we say that D is a *subdivision* of C if every cell of D is contained in a cell of C.

Definition IV.1.3. A *triangulation* of a manifold M is a homeomorphism $\phi : C \to M$ from a simplicial complex C. We say M is *triangulated* if we pick such a triangulation.

We may abuse notation and say that a subset of a triangulated manifold M is a simplex (resp. subcomplex) if its inverse image under ϕ is a simplex (resp. subcomplex) as the homeomorphism preserves the cell structure (the combinatorial data of how cells are glued together), only losing the guarantee that cells are convex hulls (it may not even make sense to speak of convex sets). Note that as we use finite simplicial complexes a manifold must be compact if it admits a triangulation.

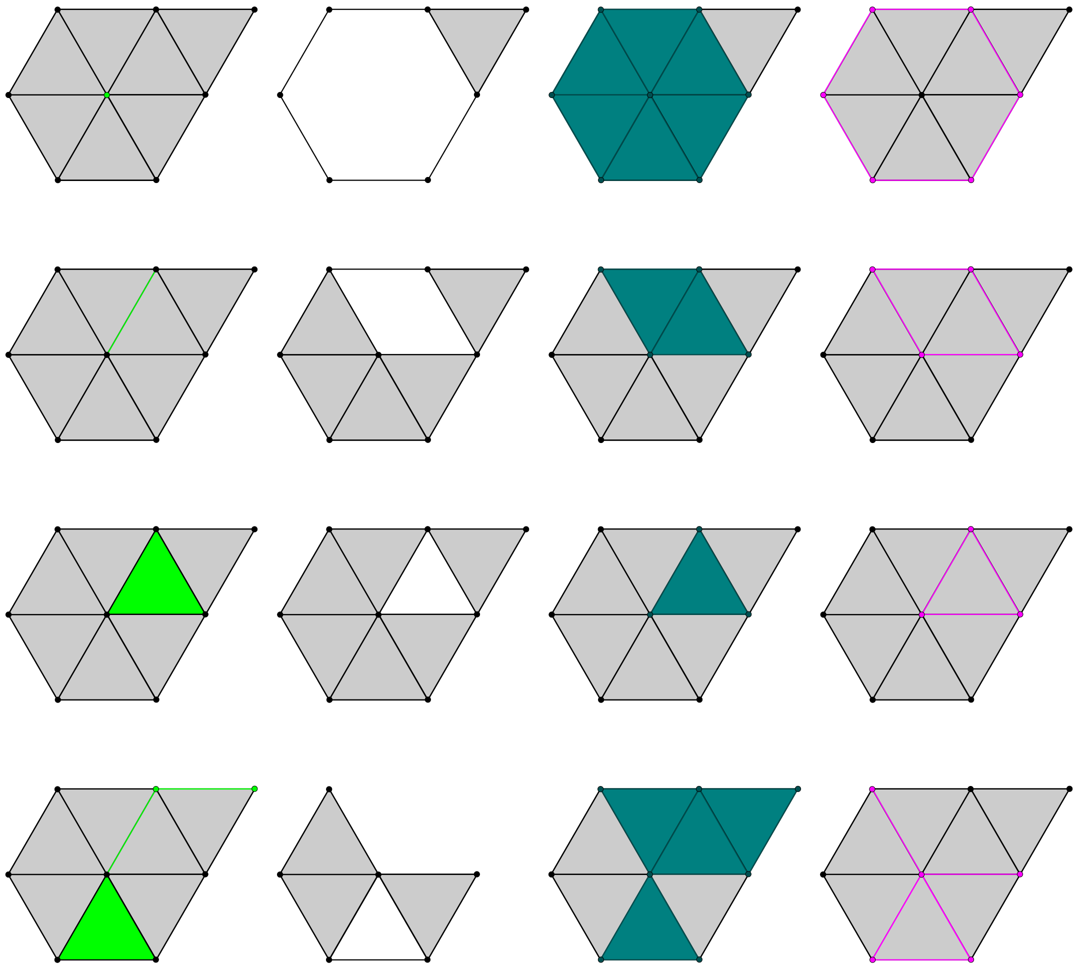


Figure IV.1: Examples of cell difference, stars and links in dimensions one and two: (left) green denotes the cell(s) we will take the cell difference, star and link of; (centre left) the cell difference; (centre right) the star is shown in teal; (right) the link is shown in purple.

2 PL manifolds

We will define PL manifolds in a combinatorial way without needing an atlas, as is common for definitions of smooth manifolds. There is a definition PL manifolds using charts with transition maps that are piecewise-linear (this is where the name comes from), but this definition is equivalent (up to PL-homeomorphism) to the one we give here [15]. The advantage of the combinatorial approach used here is that it will be easier to think in terms of the cells of a complex then in terms of piecewise-linear maps.

Given a polyhedral complex, we can view it as a poset (partially ordered set) of its cells, with the order given by inclusion. Moreover, we can think of a complex as a poset of subsets of vertices, as each face is uniquely determined by its vertices. Any face which is maximal in this poset is called a *facet*.

Definition IV.2.1. A polyhedral complex is *pure* if all its facets have the same dimension.

For example, consider Figure III.6. We can see that the complex on the right is not pure as it is a polyhedral 3-complex that has a 2-dimensional facet (the quadrilateral that intersects the lower complex). If we removed the 3-cells we would obtain a pure 2-complex. Note that the complex on the left of Figure III.6 does not fit our definition of a pure complex as it is a CW complex but not a polyhedral complex. To create a cylinder as a polyhedral complex requires at least 6 vertices (3 for each circle on the ends).

Definition IV.2.2. Two simplicial complexes are *combinatorially equivalent* if they are isomorphic as posets.

Combinatorial equivalence is an equivalence relation on simplicial complexes. It captures the idea that simplicial complexes are determined just by the combinatorial data of how faces relate to each other. Two simplicial complexes are *PL-homeomorphic* if there is some subdivision of both that are combinatorially equivalent.

A *PL n-ball* (resp. *PL n-sphere*) is a simplicial complex PL-homeomorphic to the n-simplex (resp. the boundary of the $(n+1)$-simplex).

Definition IV.2.3. A *PL manifold* is a simplicial complex so that the star of every vertex is a PL ball.

The requirement that stars of vertices are PL enforces that locally PL manifolds look like balls; note that this definition allows the possibility of a PL manifold having a non-empty boundary. We can characterise the boundary cells of a PL manifold as the cells σ for which $\operatorname{link}\sigma$ is a ball. This follows as if σ is in the interior of a complex, the link is the boundary sphere of the star, and if σ is in the boundary, then the link doesn't contain the cells of the star in the boundary (see Figure IV.2 and compare with the links of the vertex and 1-cell in Figure IV.1). A PL manifold without boundary is then a simplicial complex so that the link of every vertex is a PL sphere.

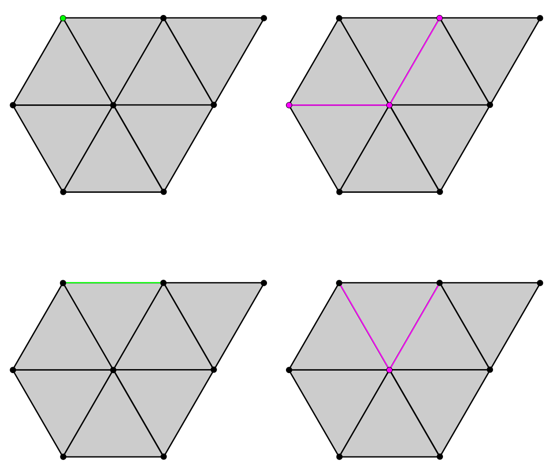


Figure IV.2: The links of cells detect the boundary of a PL manifold; the links of boundary cells are disks.

2.1 Triangulating smooth manifolds

Cairns [10] showed that there is a relatively simple process to triangulate any smooth manifold. Given a smooth compact m-manifold M, we use the Whitney embedding theorem to embed it in R^{2m} [29]. For a chosen $\epsilon > 0$ we can fix a finite number of points $x_1, \ldots, x_l$

so that the balls of radius ϵ centred at the x_i cover M. We then construct the Voronoi diagram given by the points x_i. The Voronoi diagram is a partition of $\mathbb{R}^{2m}$ into regions $U_i = \{x \in \mathbb{R}^{2m} \mid \|x - x_i\| \leq \|x - x_j\|, \forall j \neq i\}$ (note that these regions intersect along their boundaries). Cairns showed that for small enough ϵ the sets $V_i = U_i \cap M$ defines a polyhedral complex where the V_i are the facets (the lower dimensional cells are then just the intersections of these facets). A derived subdivision of this complex is then a simplicial complex, which gives us a PL triangulation for M. This triangulation is C^1 and so is unique up to PL-homeomorphism [28]. This construction can also be extended to the case of a non-compact manifold [9].

While any smooth structure gives us a PL triangulation, the converse is not true in general. Given a PL manifold there can be multiple smooth structures that can be placed on the manifold, or indeed none at all. However, in dimensions strictly less than 7 there is a unique smooth structure for each PL structure. See [5] for a review of the PL and smooth structures that may exist on manifolds.

3 From smooth to discrete Morse theory

Given a smooth Morse function f on a manifold M Gallais and Benedetti have both shown ways of constructing a PL triangulation that supports a discrete Morse function with analogous critical points, i.e. there is a discrete Morse function with a one-to-one correspondence between the (smooth) critical points of index i and the dimension i critical cells of the discrete Morse function.

Benedetti's approach transitions between the smooth and PL structure by using handle decompositions as an intermediary and shows that if we have a given PL handle decomposition then there is a discrete Morse function whose critical cells have dimension corresponding to the index of the handles.

A *PL handle attachment* $M \cup H$ works just as with in the smooth case, but in this case we only need that the subsets of the boundary of H that we glue along is a subcomplex homeomorphic to $S^{j-1} \times I^{m-j}$. We then identify a PL homeomorphic subcomplex in ∂M. Then we can glue them together (possibly after subdividing). A *PL handle decomposition* is then a sequence of handle attachments that form M.

We have already seen how to produce a PL triangulation for a given smooth manifold and how to obtain a smooth handle decomposition given a smooth Morse function. By triangulating each handle in a smooth handle decomposition separately, we get a PL handle decomposition with the same handle vector.

3.1 An induced discrete Morse function given any triangulation — Benedetti's approach

This section details Benedetti's approach to inducing a discrete Morse function on a manifold given a smooth Morse function. The main result is the following.

Theorem IV.3.1. *If M is an m-manifold with PL handle decomposition with c_i i-handles, then for large enough r the r-th derived subdivision has a discrete Morse function with c_i critical cells of index i.*

In particular this means that we can take any triangulation on the manifold and construct a suitable discrete Morse function on some derived subdivision. To prove this we will follow Benedetti [5]. We first introduce some technical tools.

Shellable and endocollapsible complexes

Definition IV.3.2. An m-dimensional simplicial complex M with non-empty boundary is *endocollapsible* if M with one m-face removed collapses onto the boundary of M.

We note that a complex is endocollapsible if and only if it admits a boundary critical discrete Morse function with exactly one interior critical point; this follows from the weak Morse inequalities of discrete Morse theory. If we consider attaching an m-handle H to an m-manifold M we can immediately see that if the handle is endocollapsible then any discrete Morse gradient on M will extend to a discrete Morse gradient on $M \cup H$ with one extra critical cell of index m. To see this take the gradient on the handle to be defined by the collapses — i.e., remove an m-cell of H and then define the gradient by the collapse of H onto its boundary — this gradient has as its critical cells the removed m-cell and the boundary of H. As the gradient is zero on the boundary, we can glue the gradients on M and H together by setting the gradient on the boundary of H to be given by the gradient on M.

Definition IV.3.3. A *shelling* for a simplicial complex is an ordering of its facets $F_1, \ldots, F_n$ so that $F_i \cap \bigcup_{j=1}^{i-1} F_j$ is pure of dimension $\dim F_i - 1$. A shelling for a 0-dimensional simplicial complex is then any ordering of its vertices. A complex is *shellable* if it has a shelling.

A shellable complex is one that can be formed by gluing facets together in a 'nice' way. Notice that given a shelling $F_1, \ldots, F_n$ for a shellable complex, the union of the first k of these faces gives us a shellable complex with shelling $F_1, \ldots, F_k$.

Proposition IV.3.4. *The boundary of the n-simplex is shellable. Moreover any ordering of facets is a shelling.*

Proof. This follows as any two codimension 1 faces of the n-simplex intersect in a codimension 2 face. Pick any ordering $F_0, \ldots, F_n$ of the $(n-1)$-faces of the n-simplex, then the intersection of F_i (this is an $(n-1)$-simplex) with the union of the previous faces in the shelling is a union of $(n-2)$-cells in ∂F_i. This is a pure subcomplex of ∂F_i and so we see that any ordering of the facets of $\partial \Delta^n$ is a shelling of $\partial \Delta^n$. □

This proposition immediately implies that given a shelling order $F_1, \ldots, F_n$ for a complex that the intersection of F_i with the union of the previous facets is shellable, as this is simply a subcomplex of the boundary of a simplex.

Lemma IV.3.5. *Suppose C is a shellable complex of dimension greater than 1 that is disconnected, then all cells of positive dimension must be in a single component, and all other components must be vertices. Moreover all these single-vertex components must make up the tail of any shelling.*

Proof. This follows as C has at least one facet in each component. Given an shelling order of facets $F_1, \ldots, F_n$ for C, if F_i, $i \neq 1$ is the first facet in a component then it must be 0-dimensional as it does not intersect the union of the previous facets. □

Note that a single simplex is shellable (it has only one facet) and endocollapsible. In fact, any shellable manifold is endocollapsible[1].

Lemma IV.3.6 (Benedetti)**.** *Shellable d-manifolds are endocollapsible.*

[1] By a shellable manifold we mean a manifold with a shellable triangulation, or equivalently a shellable complex homeomorphic to a (topological) manifold.

Proof. This follows by induction. Assume that all shellable manifolds with fewer than k facets or dimension lower that d are endocollapsible. Consider a shelling $F_1, \ldots, F_k$ of some triangulated d-manifold M. Then the subcomplex M' given by the shelling $F_1, \ldots, F_{k-1}$ is endocollapsible by the inductive assumption. As F_k is endocollapsible we have $M - F_k \searrow M' \cup \partial F_k = M' \cup \partial M$.

In a triangulated d-manifold, every $(d-1)$-cell is a face of at most two cells, so if $\sigma_1, \ldots, \sigma_l$ are the facets of $F_k \cap \bigcup_{i<k} F_i$ then it follows that each σ_i is free in M' and thus by the inductive assumption we have $M' \searrow \partial M' \setminus \sigma_i$. As $M' \cap F_k$ is $(d-1)$-dimensional and shellable we have that $\partial M' \setminus \sigma_i$ collapses onto $\partial M' \setminus (M' \mathring{\cap} F_k)$. As $(M' \mathring{\cap} F_k)$ is exactly the part of $\partial M'$ that is not in ∂M we see that $M \searrow \partial M$, that is M is endocollapsible. □

Stellar subdivisions

Given a polyhedral cell complex C, we define the *stellar subdivision* of a face $\sigma \in C$ by choosing some point $v \in \mathring{\sigma}$ and replacing σ with cells that are cones of cells $\tau < \sigma$ over v. More concretely we have the complex

$$\mathrm{St}(\sigma, C) = (C - \sigma) \cup \{\mathrm{conv}\,\{v\} \cup \tau \mid \tau \in \mathrm{star}(\tau, C)\}$$

which is the stellar subdivision of σ at v; we also call this 'starring σ'. Note that up to PL equivalence it does not matter which point v is chosen; this can be seen immediately by considering face posets of two subdivisions obtained by choosing different points. More generally, we call a subdivision of C a *stellar subdivision* if it is obtained by a sequence of such subdivisions.

Definition IV.3.7. A *derived subdivision* of a cell complex C is a subdivision $\mathrm{sd}\, M$ obtained by doing a stellar subdivision for every cell of C in order of decreasing dimension. The r-th derived subdivision is $\mathrm{sd}^r M$ defined inductively by $\mathrm{sd}^r M = \mathrm{sd}(\mathrm{sd}^{r-1} M)$.

Subdividing the cells in order of decreasing dimension is important as after subdividing a cell σ of dimension i subdividing a cell τ of dimension $i-1$ in σ will have a different stellar subdivision then if it had been subdivided first; for example see Figure IV.3. Conversely subdividing a lower dimensional cell early subdivides all cells it is a face of (see **??**) and hence prevents us from starring all original cells of the complex. If we take the barycentre of the cells of C for the stellar subdivisions we see that the well-known barycentric subdivision is a special case of a 1st derived subdivision.

In general stellar and derived subdivisions preserve many combinatorial properties of a simplicial complex. For our purposes we only need that derived subdivision preserves the shellability of a complex [7].

Proposition IV.3.8. *A derived subdivision of a shellable complex is shellable.*

As a simplex is shellable we have that a derived subdivision of a simplex is endocollapsible. This result can be strengthened to the case of a stellar subdivision. Jojić showed any stellar subdivision of a simplex is shellable [25]; combining this with Lemma IV.3.6 we have the have the following lemma.

Lemma IV.3.9. *A stellar subdivision of a simplex is endocollapsible.*

Benedetti and Adiprasito proved that PL balls becomes shellable after taking finitely many derived subdivisions [7].

Lemma IV.3.10. *For any PL ball B there is an r so that $\mathrm{sd}^r(B)$ is shellable.*

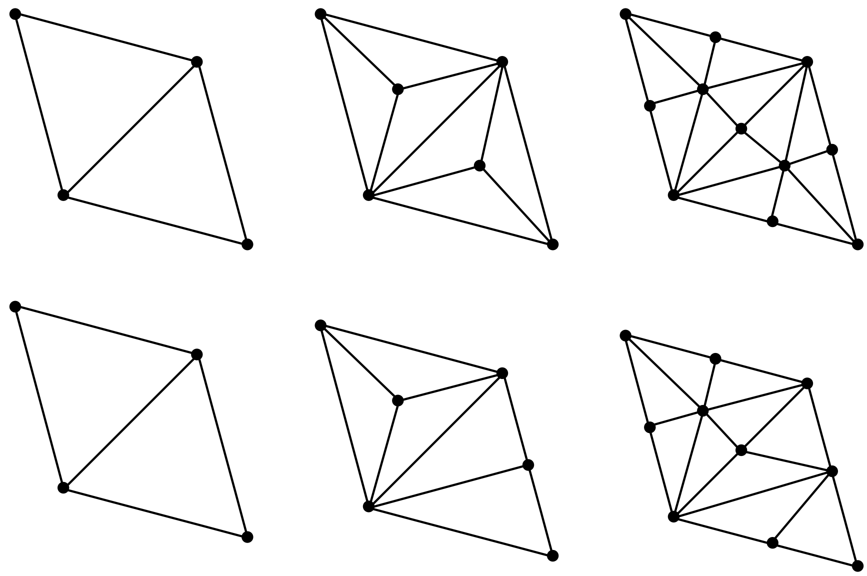


Figure IV.3: Subdividing a complex of two 2-simplices: the upper row shows a derived subdivision; the lower row shows subdividing one of the 1-cells before the 2-cell it is contained in, note that after doing so the original 2-cell can not be starred and hence we cannot subdivide to obtain the derived subdivision.

This also tells us that $\mathrm{sd}^r(B)$ will be endocollapsible via Lemma IV.3.6. In particular this means that if we subdivide enough (using derived subdivisions), the star of any vertex in a PL manifold becomes endocollapsible.

The next lemma is the key that let us take a discrete Morse function on a PL manifold, subdivide the manifold and induce a 'similar' discrete Morse function on the subdivision.

Lemma IV.3.11. *Let f be a discrete Morse function on a PL manifold M. Given a stellar subdivision M' there is a discrete Morse function f' so that*

- *There is a one-to-one correspondence between the critical i-cells of f and the critical i-cells of f';*
- *$\sigma' \subset M'$ is critical for f' if and only if $\sigma' \subset \sigma \subset M$ where σ is critical for f.*

Proof. We need only consider the case of subdividing a single face of M as any stellar subdivision is given by a sequence of such subdivisions. Suppose that M' is the subdivision $\mathrm{St}(\sigma, M)$ for some face $\sigma \subset M$ and let V be the discrete gradient for f, we will construct a discrete Morse vector field V' for M' satisfying the conditions.

Consider two cells $\gamma \prec \delta$ in M, we have three cases: neither γ or δ contains σ; both γ, δ contain σ; δ contains σ but γ does not. In the first case we set $V'(\gamma) = V(\gamma) = \delta$ as neither is subdivided so these are also cells of M'.

In the second case both γ and δ are subdivided in M'. Note that δ is the cone of γ over a single vertex v and after subdivision γ is partitioned into cells γ_i' which are cones of the faces in $\gamma - \sigma$ over the added vertex. Then we have that δ is subdivided into cells $\gamma_i' * v$; we set $V'(\gamma_i') = v * \gamma_i'$. This then defines V' on the interior cells of the subdivided γ and δ.

In the third case δ is partitioned into simplices $\delta_1, \ldots, \delta_l$ which are joins with v (the added vertex) of the simplices in δ. There is a unique such simplex — $v * \sigma$ — containing σ, so we set $V'(\sigma) = v * \sigma$. By Lemma IV.3.9 we know that the subdivision of δ is endocollapsible, so after performing the collapse $\delta \searrow \delta - \sigma$ we see that $\delta - \sigma$ collapses onto its boundary.

Now we consider what happens with a critical cell τ of f. If τ does not contain σ, then it is not subdivided so we set $V'(\tau) = 0$. If τ contains σ then it is partitioned into some set of $\dim \tau$-simplices $\tau_1, \ldots, \tau_l$. By Lemma IV.3.9 this subdivision of τ is endocollapsible, so we set $V'(\tau_i) = 0$ for some i and then we know subdivision minus τ_i collapses onto the boundary of τ. □

3.2 Proof of Theorem IV.3.1

Lemma IV.3.12 (Benedetti)**.** *If a PL d-manifold M has a discrete (resp. a boundary critical discrete) Morse function f with discrete (resp. interior) Morse vector $(c_0, \ldots, c_d)$ then M has an iterated derived subdivision that supports a boundary discrete (resp. discrete) Morse function with interior (resp. discrete) Morse vector $(c_d, \ldots, c_0)$.*

For the proof see [5, 4].

Lemma IV.3.13. *Suppose $M = B_1 \cup B_2$ is a union of two endocollapsible d-balls so that $B_1 \cap B_2$ is*

- *an $(d-1)$-dimensional subcomplex of ∂B_1 and ∂B_2, and*
- *homeomorphic to $S^{j-1} \times I^{d-j}$ for some $j \in \{1, \ldots, d\}$.*

Then any boundary-critical discrete Morse function f on $B_1 \cap B_2$ extends to a boundary critical discrete Morse function g on M whose critical interior cells are exactly the critical cells of f minus a $(d-1)$-cell, plus a d-cell.

Proof. We follow the proof in [5] but clarify one important point, that the $(d-1)$ cell chosen at the start can be a critical cell of f. This must be the case to have one less critical $(d-1)$ face in g than in f.

Let σ be a critical $(d-1)$-cell of f in $B_1 \cap B_2$ we know this must exists by the weak Morse inequalities for discrete Morse functions. We then have that there are exactly two d cells containing σ, $\tau_1 \in B_1$ and $\tau_2 \in B_2$. By the endocollapsibility of B_1 we have $B_1 - \tau_1 \searrow \partial B_1$, and this remains true when B_2 is glued to B_1 as the collapsed cells are in the interior of B_1, so $B_1 \cup B_2 - \tau_1 \searrow \partial B_1 \cup B_2$. We now have that σ is a free face of τ_2 in $\partial B_1 \cup B_2$, and then by the endocollapsibility of B_2 we have

$$\partial B_1 \cup B_2 \searrow_\sigma \partial B_1 \cup B_2 - \sigma \searrow \partial B_1 \cup \partial B_2 - \sigma.$$

Define a discrete Morse gradient V on $M \setminus (B_1 \cap B_2 - \sigma)$ by these collapses. The critical cells on this subset are exactly τ_1 and the boundary cells of M that are not in $\partial(B_1 \cap B_2)$. We then define $V = V_g$ on $B_1 \cap B_2$. Then V has as its interior critical faces τ_1 and all interior critical cells of g other than σ. □

The following lemma is an extension of this to the case where B_1 is not a ball.

Lemma IV.3.14. *Let $M = M' \cup B$ be a PL m-manifold where M' is a PL m-manifold with boundary and B is a PL endocollapsible m-ball. Suppose that*

- *$M' \cap B$ is a $(d-1)$ subcomplex of $\partial M \cap \partial B$;*
- *$M' \cap B$ is homeomorphic to $S^{j-1} \times I^{d-j}$ (i.e., we have a handle attachment); and*
- *M' has a boundary critical discrete Morse function f.*

Then a boundary critical discrete Morse function g on $M \cap B$ lifts to a boundary critical discrete Morse function h on $M \cup B$ whose critical interior cells are exactly those of g minus a $(d-1)$-face plus the critical interior cells of f.

Proof. The proof is the same as in Lemma IV.3.13 with f is used to collapse M' onto its boundary (see Remark III.3.4) in place of the endocollapsibility of B_1. □

Benedetti stated this with the additional hypothesis that the function f is perfect; we note that this is not necessary.

We now have the tools to prove Theorem IV.3.1. This will be done by proving the 'dual' problem, that for a PL manifold with PL handle vector $(c_0, \ldots, c_m)$ there is a boundary critical discrete Morse function with interior Morse vector $(c_m, \ldots, c_0)$. Then applying Lemma IV.3.12 gives us the result.

Proof of Theorem IV.3.1.
We will use induction on both the number of handles and the dimension of M, this will allow us to consider just the last handle in a decomposition with the inductive assumption that there is a suitable function for fewer handles and lower dimension. Then we will use Lemma IV.3.14 to extend the (inductively assumed) suitable functions.

First, suppose M decomposes into a single handle; this must be a 0 handle. Then M is a PL-ball, and hence by Lemmas IV.3.6 and IV.3.10 we have that for some r, $\operatorname{sd}^r M$ admits a boundary critical discrete Morse function with exactly one interior critical cell. In dimension one each 0-handle is clearly collapsible and each 1-handle is a partitioned interval, which is endocollapsible so the result is clear. These two cases provide the base case for our double induction.

Suppose that $m = \dim M > 1$ and we have more than one handle; write $M = M' \cup B$ where B is a j-handle attached to M'. We can assume that $j > 0$ as the case of attaching a 0-handle is addressed by the same argument as the single handle case with the function $g \sqcup g'$[2] where g' is a function induced by the collapsibility of a PL-ball after taking derived subdivisions. Note that M' has one handle fewer than M and $M' \cap B$ has dimension $\dim M - 1 = m - 1$, so the inductive assumption applies to these submanifolds.

Let $(h_0, \ldots, h_m)$ be the handle vector for the decomposition of M; we need to show that for some r there is a boundary critical discrete Morse function on $\operatorname{sd}^r M$ with exactly h_{m-i} critical interior i-cells. We have that $M' \cap B$ is $S^{j-1} \times I^{m-j}$. This is an $(m-1)$-dimensional manifold that has a handle decomposition into one 0-handle and one $(j-1)$-handle (note that these are $(d-1)$-dimensional handles). Taking the a-th derived subdivision for some a, we have $\operatorname{sd}^a M' \cap B$ has a PL handle decomposition into one 0-handle and one $(j-1)$-handle. The inductive assumption then says that there is some b and a boundary critical discrete Morse function g on $\operatorname{sd}^{a+b}(M' \cap B)$ whose only critical cells are exactly one critical $(m-1)$-cell and one critical $(m-j)$-cell (recall that $\dim M' \cap B = \dim M - 1$); we pick b large enough so that $\operatorname{sd}^b B$ is endocollapsible as well. Applying the inductive assumption to M' gives us a c and a boundary critical discrete Morse function f on $\operatorname{sd}^c M'$ with h_{m-i} interior critical i-cells for $i \neq j$ and $h_{m-j} - 1$ interior critical j-cells.

Define $r = \max\{a + b, c\}$. We can apply Lemma IV.3.14 to $\operatorname{sd}^r M = \operatorname{sd}^r M' \cup \operatorname{sd}^r B$. This gives us a boundary critical discrete Morse function k on $\operatorname{sd}^r M$ which has the critical interior cells of g and f minus an $(m-1)$-cell. Summing the interior critical cells of g and f, k has exactly h_{m-i} critical interior i-cells. □

[2] The disjoint union $g \sqcup g'$ of g and g' is the function on the disjoint union of their domains given by restriction to either g or g' on the two pieces.

By combining Theorem IV.3.1 with the results that a smooth Morse function induces a smooth handle decomposition of a manifold and any smooth handle decomposition induces a PL handle decomposition, we get the following corollary.

Corollary IV.3.15. *Any smooth Morse function f on M induces a PL structure on M and a discrete Morse function $\bar{f}$ such that the number of critical i-cells of the $\bar{f}$ is the same as the number of index i critical points of f. Moreover, we can see that the critical cells of $\bar{f}$ are contained in the same handle as its corresponding critical point of f.*

3.3 Triangulating along flow lines — Gallais approach

The similarities between the discrete and smooth Morse functions in Corollary IV.3.15 suggest the question of how similar one can make a discrete Morse function to a given smooth Morse function. This section will describe a construction by Gallais that provides even more topological similarities: the critical points are contained in the critical cells, there is a bijection between the set of gradient paths between critical cells and flow lines between critical points; in addition, the stable manifolds will subcomplexes of the triangulation near critical points.

First we recall a well known result about the product of simplicial complexes [13].

Proposition IV.3.16. *The product $C \times D$ of two simplicial complexes has a simplicial subdivision without adding any new vertices.*

An example for $\Delta^2 \times \Delta^1$ is shown in Figure IV.4.

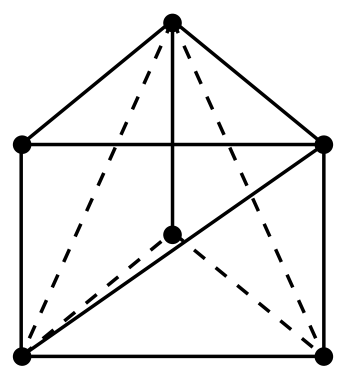


Figure IV.4: Simplicial subdivision of $\Delta^2 \times \Delta^1$ without adding new vertices.

The following lemma is the result of a fix explained by Benedetti [5] for the original statement given by Gallais [13]. Benedetti pointed out the it is not known whether any simplicial triangulation of a simplex is collapsible, something that Gallais proof relies on (Gallais proof states that X_0 is collapsible and uses this), however Benedetti and Adiprasito showed that the first derived subdivision of any simplicial subdivision is collapsible [6].

Lemma IV.3.17. *Suppose X_0 is a simplicial subdivision of Δ^k and $X_1 = \Delta^k$. Then there is a simplicial subdivision of $\Delta^k \times \Delta^1$ (we identify Δ^1 with $[0,1]$) so that $\Delta^k \times \{0\}$ is subdivided as $\operatorname{sd} X_0$ and $\Delta^k \times \{1\}$ is subdivided as $\operatorname{sd} X_1$; in addition $X \searrow \operatorname{sd} X_i \times \{i\}$ for $i \in \{0,1\}$. If X_0 is collapsible then we can instead find a triangulation so $X \searrow X_i \times \{i\}$.*

Proof. When $k = 0$ the result is clear as $\Delta^0 \times \Delta^1$ is Δ^1 and Δ^1 collapses onto either of its vertices. Suppose for induction that the claim holds up to dimension $k-1$. Then we can subdivide $\partial\Delta^k \times \Delta^1$ using the induction hypothesis so that each $\Delta^{(k-1)}$ in Δ^k and the corresponding subdivisions in X_0 make up the subdivision Y of $\partial\Delta^k \times \Delta^1$; we then have $Y \searrow \partial X_i$ for $i \in \{0,1\}$. We can then put X_1 and X_0 in along the corresponding ends to obtain a subdivision $Y \cup X_0 \times \{0\} \cup X_1 \times \{1\}$ of $\partial(\Delta^k \times \Delta^1)$.

We now subdivide $\Delta^k \times \Delta^1$ by taking the subdivision already constructed for its boundary and taking the join of each cell to some point v in the interior of $\Delta^k \times \Delta^1$; call this subdivision X'. Let X be the first derived subdivision of X'. We now show how to collapse X onto $X_0 \times \{0\}$. We will now see how to collapse $X \searrow \operatorname{sd} X_0$.

By collapsing in order of decreasing dimension, it is clear that we can collapse each face of $X'|_{X_1 \times \{1\}}$ with its join with v; by not collapsing the cells in $\partial X_1 \times \{1\}$ we have $X' \searrow X' - (\mathring{X_1} \times \{1\}) = X'|_{Y*v}$. It follows that we can collapse X onto the its corresponding subcomplex: $X \searrow X|_{Y*v}$.

We know we can collapse $X'|_{\partial X_1 \times \Delta^1} = Y$ onto $X'|_{\{\partial\Delta^k \times \{0\}\}}$, this collapsing also works with the join over v: for each elementary collapse $\searrow_\sigma$ in the collapse of Y to its base we perform the elementary collapse $\searrow_{\sigma*v}$. After collapsing the joins, we can perform the collapse of Y, giving us collapse $X'|_{Y*v} \searrow Y \cup X_0 * v \searrow X'|_{X_0*v}$. Again this means we can collapse $X|_{Y*v} \searrow X|_{Y \cup X_0 \times \{0\}*v} \searrow X|_{X_0 \times \{0\}*v}$.

We now collapse $X_0 * v$ to X_0; this is where we need that either X_0 is collapsible or that we have taken the first derived subdivision of X_0. As $\operatorname{sd} X_0$ is collapsible (see [6]) we know that we can collapse it to some point x. We can use the corresponding collapse on the join to collapse $\operatorname{sd}(X_0 * v) \searrow \operatorname{sd}(X_0 \cup x * v)$. Note that this is slightly different from collapsing joins seen before as we have the derived subdivision of the join; however, a similar process works: an elementary collapse $\searrow_\sigma$ in the sequence of collapses of $\operatorname{sd} X_0$ corresponds to the set of collapses $\searrow_{\sigma'}$ where σ' is a cell of $\operatorname{sd}(\sigma * v)$ not contained in the boundary of σ (again these must be performed in order of decreasing dimension). Finally we can collapse $\operatorname{sd}(X_0 \cup x * v) \searrow \operatorname{sd} X_0$ (this is just collapsing a subdivided 1-cell).

Combining the collapses above we have $X \searrow X|_{X_0 \times \{0\}} = \operatorname{sd} X_0 \times \{0\}$. The same process but working the other way shows that $X \searrow \operatorname{sd} X_1 \times \{1\}$. In the case that X_0 is collapsible then we can instead take $X = X'$ and essentially the same argument gives the result. $\square$

Lemma IV.3.18. *Let Δ^n be the simplex given by vertices $(\sigma_0, \dots, \sigma_n)$ and δ be the face of Δ^n given by $(\sigma_1, \dots, \sigma_n)$. Then for any $m \in \mathbb{N}$ there is a simplicial subdivision C of $\Delta^m \times \Delta^n$ so that*

- *$C|_{\Delta^m \times \delta}$ is the simplicial subdivision with no new vertices given by Proposition IV.3.16*
- *$C|_{\Delta^m \times \sigma_0} = \Delta^m$; and*
- *$C \searrow C|_{(\partial\Delta^m \times \Delta^n) \cup (\Delta^m \cup \sigma_0)}$.*

We omit the proof (see [13]), which is an inductive argument on the size of $m+n$, noting that when $m+n = 1$ the result is trivial; we will describe the triangulation. First a subdivision of $\partial(\Delta^m \times \Delta^n)$ is formed by decomposing this as $(\partial\Delta^m \times \Delta^n) \cup (\Delta^m \times (\sigma_0 * \partial\delta)) \cup (\Delta^m \times \delta)$. We subdivide $\Delta^m \times \delta$ without adding any new vertices by Proposition IV.3.16 and use the induction hypothesis (that a subdivision satisfying the lemma exists for $\Delta^i \times \Delta^j$ with $i + j < m + n$) to subdivide $(\partial\Delta^m \times (\sigma * \delta)) \cup (\Delta^m \times (\sigma * \partial))$. We then pick a point in the interior of the $n + m$ cell of $\Delta^m \times \Delta^n$ and subdivide $\Delta^m \times \Delta^n$ by taking the join of the subdivision of the boundary with this point.

The following lemma allows us to construct to triangulate a product cobordism (M, M^-, M^-)[3] given triangulations of M^- and M^+ which admit a common subdivision.

[3]A product cobordism is a cobordism (M, M^-, M^+) so that M is diffeomorphic to $M \times I$.

Lemma IV.3.19. *Let N be a submanifold (which can have boundary) of M (which may have corners). Suppose T_i^M and T_i^N are triangulations of M and N respectively so that T_i^N is a subcomplex of T_i^M for $i \in \{0,1\}$. If T_1^M and T_2^M admit a common subdivision then there is a subdivision T of $M \times [0,1]$ so that $(T|_{M\times\{i\}}, T|_{N\times\{i\}}) = (\operatorname{sd} T_i^M, \operatorname{sd} T_i^N)$ and moreover $T \searrow T_0^M \cup T|_{N\times[0,1]}$ and $T|_{N\times[0,1]} \searrow T_0^N$.*

Proof. Let $T^{1/2}$ be a common simplicial subdivision of T_1^M and $\frac{M}{2}$. We subdivide $M \times [0, \frac{1}{2}]$ and $M \times [\frac{1}{2}, 1]$ by using Lemma IV.3.17 so that along $M \times \{\frac{1}{2}\}$ we have the triangulation $T^{\frac{1}{2}}$. We can then take the union of these complexes to be the triangulation T on $M \times [0,1]$. Then we have $T|_{M\times\{i\}} = \operatorname{sd} T_i^M$ for $i \in \{0,1\}$ and we get the collapses $T \searrow \operatorname{sd} T_i^M$ from Lemma IV.3.17. □

Note that if both T_i^M are collapsible we remove the need to take the derived subdivision, and so have $T|_{M\times\{i\}} = T_i^M$.

We now construct a triangulation of cobordism with a Morse function that has just a single critical point. We follow Gallais [13], but expand some of the arguments in the construction; note that Gallais uses a different definition of the stable and unstable manifolds of a critical point (Gallais' definitions are reversed relative to those used here, using the flow of a gradient-like vector field whereas we use the negative of a gradient-like vector field); we have altered the wording of the statement to suit our definitions.

Theorem IV.3.20. *Suppose (f, ξ) is a Morse-Smale pair on a cobordism (M^m, M^-, M^+) with one critical point p of index i. Then there is a triangulation (T, T^-, T^+) of the cobordism with a discrete Morse gradient so that*

- *the unstable manifold of p is a subcomplex T_p^u of T and $T \searrow T_p^u \cup T^-$*
- *there is a i-cell $\sigma_p \subset T_p^u$ so that $p \in \sigma_p$ and $T_p^s - \sigma_p \searrow T_p^u \cap T^-$.*

Proof. The stable manifold of p is a diffeomorphic to D^{m-i} and we pick a tubular neighbourhood N of p so that ξ is transverse to ∂N; N can be thought of as the handle attached when passing over the level of p. We think of N in terms of a diffeomorphism to $D^i \times D^{m-i}$; the stable manifold is $\{0\} \times D^{m-i}$ and the unstable manifold is $D^i \times \{0\}$. We triangulate the unstable manifold with the i-simplex.

The triangulation of the stable manifold is formed by first triangulating its boundary $\{0\} \times S^{m-i-1}$ using any triangulation and then taking the triangulation of D^{m-i} to be the cone over $(0,0)$. We now want to triangulate all of $D^i \times D^{m-i}$. We use Lemma IV.3.18 to triangulate each $D^i \times ((0,0) * \sigma)$ for each simplex $\sigma \in \partial D^{m-i}$; here the δ of Lemma IV.3.18 is σ. This gives us a well defined triangulation of $D^i \times D^{m-i}$ as for any two simplices $\sigma_1, \sigma_2 \in \partial D^{m-i}$ the subdivisions produced will restrict to the same subdivisions of $D^i \times ((0,0) * (\sigma_1 \cap \sigma_2))$. So we have a triangulation of N, which we denote T^N. By taking the collapses given by Lemma IV.3.18 we have that $T^N \searrow T_p^u \cup T^N|_{\partial T_p^u \times D^{m-i}}$.

We now triangulate the rest of the cobordism. Notice that the flow gives us a diffeomorphism between $N \cup (M^- \setminus N)$ and every point in $M \setminus (N \cup M^-)$, which is the region that still requires triangulation.

Fix a triangulation T^- of M^- so that it agrees with T^N on $N \cap M^-$. For example, this can be done using any triangulation of $M^- \setminus (M^- \mathring{\cap} N)$, replacing the triangulation of $M^- \cap N$ with the triangulation given by N and re-triangulating the cells in $M \setminus N$ that share a face with T^N by stellar subdivision. Define $\partial_- N = M^- \cap N$ and $\partial_+ N = \partial N \setminus (\mathring{\partial_- N})$. These two submanifolds decompose ∂N into two pieces that with intersection $V = \partial_N \cap \partial_+ N$ which is $\partial D^i \times \partial D^{m-i} = S^{i-1} \times S^{m-i-1}$.

We now use the flow of ξ to carry the triangulation on $\partial_+ N$ up to M^+. As ξ is transverse to $\partial_+ N$ the flow is a diffeomorphism of $\partial_+ N$ (resp. V) onto some submanifold of M^+ which we denote $M^+_{\partial_+ N}$ (resp. M^+_V); we use the triangulation induced by the flow to on $M^+_{\partial_+ N}$ and M^+_V, in particular the triangulations of these triangulations are combinatorially equivalent to the triangulations of $\partial_+ N$ and V. The part of the manifold (along the flow) between $\partial_+ N$ and $M^+_{\partial_+ N}$ is clearly a product cobordism (diffeomorphic to $\partial_+ N \times I$) and so we can use Lemma IV.3.19 to triangulate it; this also gives us a triangulation of the product cobordism between V and M^+_V.

The same process will give us a triangulation between $M^- \setminus \mathring{\partial_-} N$ and its image under the flow, and on the cobordism between V and M^+_V, this will agree with the triangulation already constructed. Taking the part of this triangulation on M^+ together with the triangulation $M^+_{\partial_+ N}$, we have a triangulation T^+ of M^+.

Note that triangulating the cobordisms requires us the take the derived subdivision of the complex if the triangulations of $\partial_+ N$ and $M^- \setminus \mathring{\partial_-} N$ are not both collapsible, but recall that the collapses already discussed will still be valid, so we may assume (for notational convenience) that $\partial_+ N$ and $M^- \setminus \mathring{\partial_-} N$ were collapsible.

We have now completed the triangulation T of M, and in fact we have a cellular triad (T, T^-, T^+). By Lemma IV.3.19 we have that $T \searrow T^- \cup T^N$; if we then perform the collapse of $T^N \searrow T^u_p \cup T^N|_{\partial T^u_P \times D^{m-i}}$ already discussed we have $T \searrow T^u_p \cup T^-$.

As the i-simplex is endocollapsible we can then collapse $T^u_p \cup T^- \searrow T^-$ and so we have a discrete Morse gradient on the cellular triad (T, T^-, T^+) with only a single critical cell. □

Using Theorem IV.3.20 we can immediately see how to produce a triangulation for any manifold M with a Morse-Smale pair (f, ξ) where no two critical points have the same value of f.

Theorem IV.3.21. *Given a Morse-Smale pair (f, ξ) where $f : M \to \mathbb{R}$ is an ordered Morse function on a manifold M, there is a triangulation T of M with a discrete Morse function so that there is a bijection between critical cells of dimension i in T and the critical cells of index i; moreover the critical points are contained in their corresponding critical cell and the there is a neighbourhood of each critical point in its unstable manifold in which the unstable manifold is a subcomplex of T.*

Proof. To see this denote the critical points by p_i where $f(p_i) < f(p_{i+1})$ and cut M up into cobordisms ${}^{f(p_i)+\epsilon_i}_{f(p_i)-\epsilon_i}M$ only critical point is p_i, and product cobordisms ${}^{f(p_{i+1})-\epsilon_{i+1}}_{f(p_i)+\epsilon_i}M$ between them. We triangulate the non-trivial cobordisms using Theorem IV.3.20. Due to the fact that we use the flows of these cobordisms to triangulate the upper level sets $f^{-1}(f(p_i) + \epsilon_i)$, the triangulations on these will be C^1 (and we can choose a C^1 triangulation on the lower level sets $f^{-1}(f(p_i) - \epsilon_i)$. Any two C^1 triangulations of the same manifold admit a common simplicial subdivision [28] and so we can use Lemma IV.3.19 to triangulate the product cobordisms so that they agree on with the non-trivial cobordisms along their common level sets. Note that the product cobordisms admit discrete Morse gradient with no critical cells by Lemma IV.3.19 The rest of the statement follows immediately from Theorem IV.3.20 and the previously discussed method for defining a discrete Morse gradient when gluing cellular triads with given discrete Morse gradients together. □

It is also possible to drop the requirement that critical points have distinct values by choosing (for critical points of the same value) tubular neighbourhoods that are disjoint from each other.

Chapter V

Morse homology

One reason to be interested in Morse theory is that it provides a way to compute the homology of smooth manifolds. Using the flows between critical points of a Morse function one can define a chain complex whose homology is isomorphic to the singular homology of the manifold. Here we will briefly review the smooth Morse homology theory before showing the discrete analogue introduced by Forman [12].

1 Smooth Morse homology

We briefly recall the basic objects and structure of smooth Morse homology. Given a smooth Morse function on a compact manifold M and a Morse-Smale gradient ξ, the intersection of $M_s(p) \cap M_u(q)$ is a manifold with dimension $\dim M_s(p) + \dim M_u(q) - \dim m = (m - \lambda(p)) + \lambda(q) - m = \lambda(q) - \lambda(p)$ by Proposition I.4.6. This is the set of points on flow lines from q to p. Each flow line can be reparameterised by a translation of $\mathbb{R}$, changing the initial condition for the corresponding differential equation. Quotienting by this action of $\mathbb{R}$ gives the moduli space of flow lines from q to p.

$$\mathscr{M}(q,p) = (M_s(p) \cap M_u(q))/\mathbb{R}.$$

It follows that $\mathscr{M}(q,p)$ is a manifold with dimension $\lambda(q) - \lambda(p) - 1$ provided that $p \neq q$.

Of particular interest to us is the case where $\lambda(q) = \lambda(p) + 1$. In this case we have that $\mathscr{M}(q,p)$ is a collection of points. It can be shown that when $\lambda(q) = \lambda(p) + q$, $\mathscr{M}(q,p)$ is compact and hence consists of finitely many points [16]. By picking an orientation on the unstable manifold of q and the unstable manifold of p we get orientations on the flow lines can be used to induce an orientation on T_pM from the orientation of T_qM; note that reparameterising a flow line does not change this orientation and so the points $\mathscr{M}(q,p)$ induce orientations on T_pM. We define $n(q,p)$ as the signed count of the points in $\mathscr{M}(q,p)$ where the sign is positive if the induced orientation agrees with the orientation of T_pM given by the chosen orientation on the $M_u(p)$ [16, 23].

The Morse chain complex $(C_*^\infty(f,\xi), \partial^\infty)$ is these given by $C_i^\infty(f,\xi) = \mathbb{Z}\,\mathrm{Cr}_i f$ is the $\mathbb{Z}$-module generated by the index i critical points of f and the boundary map is

$$\begin{aligned} \partial^\infty : C_i &\to C_{i-1} \\ p &\mapsto \sum_{q \in \mathrm{Cr}_{i-1} f} n(p,q)q. \end{aligned}$$

This is a chain complex ($\partial^\infty \circ \partial^\infty = 0$) and its homology is isomorphic to singular homology [23, 16].

2 Discrete Morse homology

Given a discrete Morse function, Forman showed two different ways to construct a chain complex with homology isomorphic to singular homology. We will consider the one that is more closely analogous to Morse homology. For this section let M be a (finite) CW complex with a discrete gradient V. We define C_i^V to be the $\mathbb{Z}$-module generated by the critical i-cells of M; these are the chain groups for discrete Morse theory. As in the smooth case the boundary map will be defined as a linear combination of cells by counting the number of gradient paths between critical cells of adjacent dimension. We assume familiarity with the cellular homology of a CW complex and in this section ∂ will denote the cellular boundary map.

2.1 Induced orientations

In smooth Morse homology the orientations of flow lines between critical points are used to sign the contributions of ∂^∞. We will need to account for orientations induced by gradient paths between critical cells in a similar way. We use the notation $\langle \partial\tau, \sigma\rangle$ to denote the incidence number between τ and σ (this is symmetric with $\langle \partial\tau, \sigma\rangle = \langle \sigma, \partial\tau\rangle$). This is motivated by recognising $\langle \cdot, \cdot\rangle$ as an inner product on the $\mathbb{Z}$ module generated by the p-cells of M so that the p-cells form an orthonormal basis. By doing this we identify $-\sigma$ as being the cell σ but with the opposite orientation.

The orientation on the initial cell of a gradient path induces the orientation of the terminal cell of the path in a step-wise manner along the path. Given two p-cells $\sigma \neq \tilde{\sigma}$ and a $(p+1)$-cell τ so that $\sigma < \tau$ and $\tilde{\sigma} < \tau$, an orientation on σ induces an orientation on $\tilde{\sigma}$ given by the formula

$$\langle \partial\tau, \sigma\rangle\langle \partial\tau, \tilde{\sigma}\rangle = -1.$$

Notice that we can think of this in two equivalent ways: σ induces an orientation on τ so $\langle \partial\tau, \sigma\rangle = -1$ which then induces an orientation on $\tilde{\sigma}$ so $\langle \partial\tau, \tilde{\sigma}\rangle = 1$; or we can take the orientations for both σ and τ to be fixed and this pair of orientations will induce the orientation on $\tilde{\sigma}$ by the above formula. Notice that both points of view give the same result as taking the opposite orientation for τ gives us

$$\langle -\partial\tau, \sigma\rangle\langle -\partial\tau, \tilde{\sigma}\rangle = (-\langle \partial\tau, \sigma\rangle)(-\langle -\partial\tau, \tilde{\sigma}\rangle) = \langle \partial\tau, \sigma\rangle\langle \partial\tau, \tilde{\sigma}\rangle.$$

Figure V.1 shows the orientation induced by a 1-simplex on the other faces of a 2-simplex.

2.2 The boundary map

We now fix an orientation for all the cells in M. We can follow induced orientations along a gradient path $\gamma : \sigma_0, \ldots, \sigma_l$ by first taking the fixed orientation on σ_0 and $V(\sigma)$ to induce an orientation on σ_1; we then induce an orientation on σ_{i+1} by using the induced orientation on σ_i. We define the *multiplicity* $m(\gamma)$ of γ to be 1 if the orientation induced on σ_l is the fixed orientation of σ_l and -1 if the induced orientation is opposite. Algebraically this is given by

$$m(\gamma) = \prod_{i=0}^{l-1} -\langle \partial V(\sigma_i), \sigma_i\rangle\langle \partial V(\sigma_i), \sigma_{i+1}\rangle$$

for $l > 0$; if we have a length 0 path (the path is just a single cell) we define the multiplicity to be 1. This links the orientation of σ_0 to the orientation of σ_l, telling us whether these orientations agree or not.

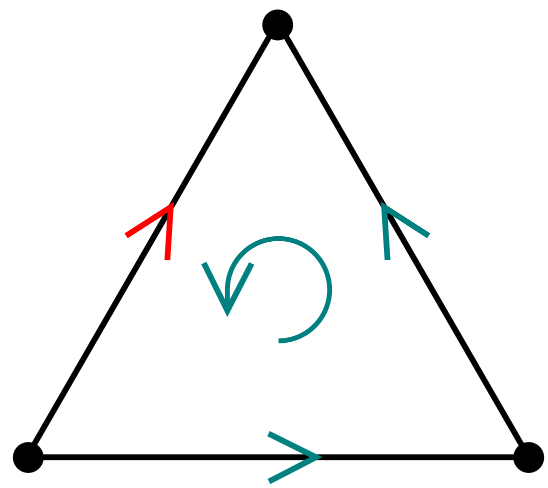


Figure V.1: The orientations induced (blue) by a given orientation (red) 1-simplex.

We now define the boundary map $\partial^V : C^V_{i+1} \to C^V_i$ for chains by

$$\partial^V(\tau) = \sum_{\sigma \in C_i} n(\tau, \sigma)\sigma$$

where

$$n(\tau, \sigma) = \sum_{\tilde{\sigma}^{(i)} < \tau} \langle \partial\tau, \tilde{\sigma} \rangle \sum_{\gamma \in \Gamma(\tilde{\sigma}, \sigma)} m(\gamma).$$

where $\Gamma(\alpha, \beta)$ denotes the set of gradient paths $\alpha \rightsquigarrow \beta$.

This is exactly analogous to the smooth case — $n(\tau, \sigma)$ counts all gradient paths from i-cells in $\tau^{(i+1)}$ to σ, just as it counted all flow lines between critical cells in the smooth case. The terms $\langle \partial\tau, \tilde{\sigma} \rangle$ account for the orientation in moving from τ to faces that begin a path and the multiplicities carry these to the end of the gradient paths.

The following two theorems (Theorems V.2.1 and V.2.3) were proved by Forman [12] using an indirect method. We present proofs following those of Gallais that are more direct [13]. Gallais' proofs of Theorems V.2.1 and V.2.3 are for simplicial complexes; we extend these to the case of CW complexes.

Theorem V.2.1. *(C^V_*, ∂^V) is a chain complex, i.e. $\partial^V \circ \partial^V = 0$.*

Proof. We only need to consider the contributions of sets $\{\gamma^{(p)} \overset{(\alpha)}{\rightharpoondown} \beta \overset{(\mu)}{\rightharpoondown} \nu\}$ of path juxtapositions where γ and ν are fixed and β is allowed to vary. This follows as we only need to show that the coefficient of each $(p-2)$-cell in $\partial^V \circ \partial^V(\gamma)$ is 0. We will define ${}^{\nu}(\partial^W \circ \partial^W)(\gamma)$ to be the contribution of $\partial^W \circ \partial^W(\gamma)$ that lies in $\mathbb{Z}\nu$. Then to prove that $\partial^V \circ \partial^V = 0$ we just need to show that ${}^{\nu}(\partial^V \circ \partial^V)(\gamma) = 0$ for any critical cells $\gamma^{(p)}$ and $\nu^{(p-2)}$.

We prove this by induction on the number of cells in the pairing of the vector field. If there are no paired cells, then we have that the boundary map is the cellular boundary map: $\partial^V = \partial$ and hence $\partial^V \circ \partial^V = 0$. Suppose for induction that for every vector field with $k - l$ paired edges (where $0 < l \leq k$) that we have the result. Given a vector field V with k paired edges, let V' be the vector field given by removing a single pair $\sigma^{(n)} \prec \tau^{(n+1)}$ from V (i.e. making σ and τ critical). By the induction hypothesis we have $\partial^{V'} \circ \partial^{V'} = 0$; moreover if $\dim \sigma = n$ then $C^V_i = C^{V'}_i$ for $i \neq n, n+1$ and $\partial^V|_{C^{V'}_i} = \partial^{V'}|_{C^{V'}_i}$ for $i \neq n+2, n+1, n$. It follows that we only have to check that $\partial^V \circ \partial^V(\sigma') = 0$ for these cases, where $\tau' \in C^V_n, C^V_{n+1}, C^V_{n+2}$.

We will write $\Gamma(\alpha,\beta)$ for the set of gradient V-paths $\alpha \rightsquigarrow \beta$ and $\Gamma'(\alpha,\beta)$ for the set gradient V'-paths $\alpha \rightsquigarrow \beta$.

Case 1 Suppose $\tau' \in C_n^V$. Notice that no path $\gamma : \tau' \rightharpoonup \sigma' \in C_{n-1}^V$ passes through σ (i.e., no path has some segment $\sigma_i \overset{(\sigma)}{\rightsquigarrow} \sigma_{i+1}$) since σ is paired with $\tau \in C_{n+1}$. It follows that making σ critical does not change $(n-1)$-dimensional gradient paths and so $\partial^V \circ \partial^V(\tau') = \partial^{V'} \circ \partial^{V'}(\tau') = 0$.

Case 2 Suppose that $\tau' \in C_{n+1}^V$ and $\nu \in C_{n-1}^V = C_{n-1}^{V'}$; we need to show that ${}^{\nu}(\partial^V \circ \partial^V)(\tau) = 0$. Let S be the set of juxtapositions of gradient V-paths $\tau' \rightharpoonup \sigma' \rightharpoonup \nu$ where $\sigma' \in C_n^V$. If there are no juxtapositions that go through σ (i.e., no path $\tau' \rightsquigarrow \sigma'$ in a juxtaposition has a segment $\sigma \overset{(\tau)}{\rightsquigarrow} \bar\sigma$), then we are done as $\partial^V \circ \partial^V(\tau') = \partial^{V'} \circ \partial^{V'}(\tau') = 0$ as S is also the set of juxtapositions for V'.

We will first explain the relationships between sets of juxtaposed paths and their contributions to ${}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau')$ and ${}^{\nu}(\partial^V \circ \partial^V)(\tau')$. This will show contributions cancel for V'. We will then see a direct computation drawn directly from this discussion that confirms the result. To simplify notation we will prescribe that any cell that is given by σ with some adornment (i.e. $\sigma', \tilde\sigma, \bar\sigma$) has dimension n and any cell given by ν with an adornment has dimension $n-1$.

Let S' be the set of juxtapositions of gradient V'-paths $\tau' \rightharpoonup \bar\sigma \rightharpoonup \nu$ for $\bar\sigma \in C_n^{V'} = C_n^V \cup \{\sigma\}$. Consider the difference between S and S'. When removing the pair $\sigma \prec \tau$ from V to get V' we lose the juxtapositions in S that go through σ, but gain juxtapositions of the form $\tau' \rightharpoonup \sigma \rightharpoonup \nu$. By the inductive assumption we have ${}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau) = 0$; this means that the set of juxtapositions $\tau \overset{(\sigma)}{\rightharpoonup} \sigma \rightharpoonup \nu$ (i.e., the ones through σ) cancels the contribution of the subset of juxtapositions $\tau \rightharpoonup \sigma' \rightharpoonup \nu$, where $\sigma' \neq \sigma$, for $\partial^{V'} \circ \partial^{V'}(\tau)$. The juxtapositions in S that do not go through σ are exactly the juxtapositions in S' that do not go through σ. We can also see that the juxtapositions in S that do go through σ have a specific form: $\tau' \rightharpoonup \sigma \overset{(\tau)}{\rightsquigarrow} \bar\sigma \rightsquigarrow \sigma' \rightharpoonup \nu$ where $\bar\sigma^{(p)} < \tau$, $\bar\sigma \neq \sigma$, note that we allow the possibility that $\bar\sigma = \sigma'$ (in which case $\bar\sigma \rightsquigarrow \sigma'$ is length 0).

The following schematics depict these gradient paths; each arrow represents the set of paths between cells with arrows to (resp. from) C_n^V representing the set of all paths to (resp. from) any critical n-cell. First we have the V'-paths:

$$\begin{tikzcd}
 & C_n^V \arrow[dr] & \\
\tau' \arrow[ur] \arrow[dr] & \tau \arrow[u, "(\bar\sigma)"'] \arrow[d, "(\sigma)"] & \nu \\
 & \sigma \arrow[ur] &
\end{tikzcd}$$

where $\sigma \neq \bar\sigma < \tau$.

When passing to V by adding pair $\sigma \prec \tau$ (i.e. setting $V(\sigma) = \tau$) we get the following

V-paths.

$$\begin{tikzcd}
 & C_n^V \arrow[dr, harpoon] & \\
\tau' \arrow[ur, harpoon] \arrow[dr, harpoon] & \bar{\sigma} \arrow[u, squiggly] & \nu \\
 & \sigma \arrow[u, squiggly, "(\tau)"] &
\end{tikzcd}$$

Note that definition of $n(\alpha^{(p+1)}, \beta^{(p)})$ still makes sense when β is not critical. Examining the contribution of juxtapositions $\tau' \rightharpoonup \sigma \overset{(\tau)}{\rightsquigarrow} \bar{\sigma} \rightsquigarrow \sigma' \rightharpoonup \nu$ to ${}^{\nu}(\partial^V \circ \partial^V)(\tau')$ we see that they contribute $-\langle \partial\tau, \sigma\rangle$ multiplied by both $n(\tau', \sigma)$ and the contribution of juxtapositions $\tau \rightharpoonup \bar{\sigma} \rightsquigarrow \sigma' \rightharpoonup \nu$ to ${}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau)$; this is exactly the contribution to ${}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau')$ of the juxtapositions in S' that go through σ. As the juxtapositions in S that do not go through σ are the same as the juxtapositions in S' that do not go through σ, we then have that ${}^{\nu}(\partial^V \circ \partial^V)(\tau') = {}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau') = 0$.

We now give an explicit computation that illustrates this. By the inductive assumption we have ${}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau') = 0$, in particular the contribution of juxtapositions in S' through σ cancels the contribution of juxtapositions in S' that do not go through σ:

$$\begin{aligned}
0 &= {}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau') \\
&= \sum_{\sigma' \in C_n^V} \sum_{\tilde{\sigma} < \tau'} \langle \partial\tau', \tilde{\sigma}\rangle \sum_{\gamma \in \Gamma'(\tilde{\sigma}, \sigma')} m(\gamma) \sum_{\tilde{\nu} < \sigma'} \langle \partial\sigma', \tilde{\nu}\rangle \sum_{\gamma \in \Gamma(\tilde{\nu}, \nu)} m(\gamma)\nu \\
&\quad + \sum_{\tilde{\sigma} < \tau'} \langle \partial\tau', \tilde{\sigma}\rangle \sum_{\gamma \in \Gamma'(\tilde{\sigma}, \sigma)} m(\gamma) \sum_{\tilde{\nu} < \sigma} \langle \partial\sigma, \tilde{\nu}\rangle \sum_{\gamma \in \Gamma(\tilde{\nu}, \nu)} m(\gamma)\nu. \qquad \mathbf{(1)}
\end{aligned}$$

where we have split up the contributions that do not go through σ (the part before the '+') and those that do (the part after '+').

Similarly, from ${}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau) = 0$ we have

$$\begin{aligned}
-\langle \partial\tau, \sigma\rangle \sum_{\tilde{\nu} < \sigma} \langle \partial\sigma, \tilde{\nu}\rangle \sum_{\gamma \in \Gamma'(\tilde{\nu}, \nu)} m(\gamma)\nu \\
= \sum_{\sigma' \in C_n^V} \sum_{\substack{\bar{\sigma} < \tau \\ \bar{\sigma} \neq \sigma}} \langle \partial\tau, \bar{\sigma}\rangle \sum_{\gamma \in \Gamma'(\bar{\sigma}, \sigma')} m(\gamma) \sum_{\tilde{\nu} < \sigma'} \langle \partial\sigma', \tilde{\nu}\rangle \sum_{\gamma \in \Gamma(\tilde{\nu}, \nu)} m(\gamma)\nu. \qquad \mathbf{(2)}
\end{aligned}$$

Denoting the set of gradient V-paths $\tilde{\sigma} \rightsquigarrow \sigma'$ that pass through σ by $\Gamma^{\sigma}(\tilde{\sigma}, \sigma')$ we have by definition

$$\sum_{\gamma \in \Gamma^{\sigma}(\tilde{\sigma}, \sigma')} m(\gamma) = -\langle \sigma, \partial\tau\rangle \sum_{\gamma \in \Gamma(\tilde{\sigma}, \sigma)} m(\gamma) \sum_{\substack{\bar{\sigma} < \tau \\ \bar{\sigma} \neq \sigma}} \langle \partial\tau, \bar{\sigma}\rangle \sum_{\gamma \in \Gamma(\bar{\sigma}, \sigma')} m(\gamma) \qquad \mathbf{(3)}$$

as

$$m(\tilde{\sigma} \rightsquigarrow \sigma \overset{(\tau)}{\rightsquigarrow} \bar{\sigma} \rightsquigarrow \sigma') = m(\tilde{\sigma} \rightsquigarrow \sigma) m(\sigma \overset{(\tau)}{\rightsquigarrow} \bar{\sigma}) m(\bar{\sigma} \rightsquigarrow \sigma')$$

and

$$m(\sigma \overset{(\tau)}{\rightsquigarrow} \bar{\sigma}) = -\langle \sigma, \partial\tau\rangle \langle \partial\tau, \bar{\sigma}\rangle.$$

We are now in a position to compute ${}^{\nu}(\partial^V \circ \partial^V)(\tau')$ and show it is 0. We start by noting that we can split up the set of paths from a $\tilde{\sigma} < \tau'$ to a (critical) cell $\sigma' \in C_n^V$ into those that pass through sigma, $\Gamma^\sigma(\tilde{\sigma}, \sigma')$, and those that do not, $\Gamma'(\tilde{\sigma}, \sigma')$. So we have

$$
\begin{aligned}
{}^{\nu}(\partial^V \circ \partial^V)(\tau') = \sum_{\sigma' \in C_n^V} \sum_{\tilde{\sigma} < \tau'} \langle \partial \tau', \tilde{\sigma} \rangle \left(\sum_{\gamma \in \Gamma^\sigma(\tilde{\sigma}, \sigma')} m(\gamma) + \sum_{\gamma \in \Gamma'(\tilde{\sigma}, \sigma)} m(\gamma) \right) \\
\times \left(\sum_{\tilde{\nu} < \sigma'} \langle \partial \sigma', \tilde{\nu} \rangle \sum_{\gamma \in \Gamma(\tilde{\nu}, \nu)} m(\gamma) \nu \right) \qquad (\mathbf{4})
\end{aligned}
$$

Let's consider just the terms with that use paths in the various Γ^σ. By expanding the sums over each $\Gamma^\sigma(\tilde{,}\sigma')$ (using (**3**)) and rearranging we get that these terms are

$$
\begin{aligned}
&\sum_{\tilde{\sigma} < \tau'} \langle \partial \tau', \tilde{\sigma} \rangle \sum_{\gamma \in \Gamma(\tilde{\sigma}, \sigma)} m(\gamma) \\
&\qquad \times -\langle \partial \tau, \sigma \rangle \sum_{\sigma' \in C_n^V} \sum_{\substack{\bar{\sigma} < \tau \\ \bar{\sigma} \neq \sigma}} \langle \partial \tau, \bar{\sigma} \rangle \sum_{\gamma \in \Gamma(\bar{\sigma}, \sigma')} m(\gamma) \sum_{\tilde{\nu} < \sigma'} \langle \partial \sigma', \tilde{\nu} \rangle \sum_{\gamma \in \Gamma(\tilde{\nu}, \nu)} m(\gamma) \nu. \qquad (\mathbf{5})
\end{aligned}
$$

Using (**2**) (noting that $\langle \partial \tau, \sigma \rangle = \pm 1$) we can simplify the factor on the second line and write (**5**) as

$$
\sum_{\tilde{\sigma} < \tau'} \langle \partial \tau', \tilde{\sigma} \rangle \sum_{\gamma \in \Gamma(\tilde{\sigma}, \sigma)} m(\gamma) \sum_{\tilde{\nu} < \sigma} \langle \partial \sigma, \tilde{\nu} \rangle \sum_{\gamma \in \Gamma'(\tilde{\nu}, \nu)} m(\gamma) \nu.
$$

Substituting this back into (**4**) and looking back at (**1**) we see that

$$
{}^{\nu}(\partial^V \circ \partial^V)(\tau') = {}^{\nu}(\partial^{V'} \circ \partial^{V'})(\tau') = 0.
$$

Case 3 The third case, with $\tau' \in C_{n+2}^V$ and $\nu \in V_n^V$ works similarly to the previous case; we will show the path schematics. The V'-paths to consider are

$$
\begin{tikzcd}
 & C_{n+1}^V \arrow[d] \arrow[dr] & \\
\tau' \arrow[ur] \arrow[dr] & \sigma & \nu \\
 & \tau \arrow[u, "(\sigma)"] \arrow[ur, "(\bar{\sigma})"'] &
\end{tikzcd}
$$

Where $\bar{\sigma} \neq \sigma$. Then passing to V we get

$$
\begin{tikzcd}
 & C_{n+1}^V \arrow[d] \arrow[dr] & \\
\tau' \arrow[ur] & \sigma \arrow[d, squiggly, "(\tau)"'] & \nu \\
 & \bar{\sigma} \arrow[ur, squiggly] &
\end{tikzcd}
$$

and similar considerations show that the contribution of these sets of paths cancel. $\square$

The key difference between our proof and Gallais' is that as Gallais uses simplicial complexes all incidence numbers are ± 1. This makes it clear that the contributions of gradient paths cancel over $\mathbb{Z}/2\mathbb{Z}$ and so Gallais shows that when passing from V'-paths to V-paths there is still a pairing between the contributions to $\partial^V \circ \partial^V$ and then that the signs of these are opposite. When using CW complexes it is less clear how to show that contributions still cancel in pairs as the incidence numbers between a critical cell and its faces can be any integer, so instead we work with the entire set of gradient paths at once.

We call the chain complex $(C^V_\bullet, \partial^V)$ the *discrete Morse chain complex.*

To prove that the homology of the discrete Morse chain complex is isomorphic to singular homology (Theorem V.2.3) we employ some general machinery for chain complexes known as 'Gaussian elimination' [13, 3].

Lemma V.2.2. *Suppose $\phi : A \to B$ is an isomorphism of $\mathbb{Z}$ modules, then the chain complex*

$$\to C_{n+1} \xrightarrow{\begin{pmatrix}\alpha\\ \beta\end{pmatrix}} A \oplus D \xrightarrow{\begin{pmatrix}\phi & \delta\\ \gamma & \epsilon\end{pmatrix}} B \oplus E \xrightarrow{\begin{pmatrix}\xi & \eta\end{pmatrix}} C_{n-2} \to$$

has homology isomorphic to the homology of

$$\to C_{n+1} \xrightarrow{\beta} D \xrightarrow{\epsilon - \gamma\phi^{-1}\delta} E \xrightarrow{\eta} C_{n-2} \to .$$

Proof. By applying the change of basis given by $T_1 = \begin{pmatrix}1 & \phi^{-1}\delta\\ 0 & 1\end{pmatrix}$ and $T_2 = \begin{pmatrix}1 & 0\\ -\phi^{-1}\gamma & 1\end{pmatrix}$ to $A \oplus D$ and $B \oplus E$ respectively, the first map transforms to

$$T_1\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}1 & \phi^{-1}\delta\\ 0 & 1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\alpha + \phi^{-1}\delta\beta\\ \beta\end{pmatrix} = \begin{pmatrix}0\\ \beta\end{pmatrix}$$

where the last step comes from the fact that $\phi\alpha + \delta\beta = 0$ as $d^2 = 0$. Similarly the final map transforms to

$$\begin{pmatrix}\xi & \eta\end{pmatrix} T_2^{-1} = \begin{pmatrix}\xi & \eta\end{pmatrix}\begin{pmatrix}1 & 0\\ \phi^{-1}\gamma & 1\end{pmatrix} = \begin{pmatrix}0 & \eta\end{pmatrix}.$$

The map in the middle transforms to

$$T_2\begin{pmatrix}\phi & \delta\\ \gamma & \epsilon\end{pmatrix} T_1^{-1} = \begin{pmatrix}\phi & 0\\ 0 & \epsilon - \gamma\phi^{-1}\delta\end{pmatrix}.$$

So now we have the chain complex

$$\to C_{n+1} \xrightarrow{\begin{pmatrix}0\\ \beta\end{pmatrix}} A \oplus D \xrightarrow{\begin{pmatrix}\phi & 0\\ 0 & \epsilon - \gamma\phi^{-1}\delta\end{pmatrix}} B \oplus E \xrightarrow{\begin{pmatrix}0 & \eta\end{pmatrix}} C_{n-2} \to .$$

This is a direct sum of two chain complexes, so we have that the homology groups split with this direct sum, but the homology of the summand $0 \to A \xrightarrow{\phi} B \to$ is trivial so we have that the chain complex

$$\to C_{n+1} \xrightarrow{\beta} D \xrightarrow{\epsilon - \gamma\phi^{-1}\delta} E \xrightarrow{\eta} C_{n-2} \to$$

has homology isomorphic to the original complex. □

Theorem V.2.3. *The homology of the discrete Morse chain complex is isomorphic to singular homology.*

Proof. We will prove this by showing that the discrete Morse chain complex has the same homology as the cellular chain complex. In the case where all cells are critical this is trivial as the chain complex is then equal to the cellular chain complex. Suppose for induction that the result is true for every discrete gradient with up to $k-1$ cell pairs. Let V be a discrete gradient with k cell pairing and let V' be the discrete gradient obtained by making one pair critical. By the remarks about the relationship between C_i^V and $C_i^{V'}$ in Theorem V.2.1 we can see that the chain complex for V' is the same as the chain complex for V except for the segment

$$\to C_{n+2}^V \xrightarrow{\begin{pmatrix}\alpha \\ \partial^V\end{pmatrix}} \mathbb{Z}\tau \oplus C_{n+1}^V \xrightarrow{\begin{pmatrix}\langle\partial\tau,\sigma\rangle & \delta \\ \gamma & \epsilon\end{pmatrix}} \mathbb{Z}\sigma \oplus C_n^V \xrightarrow{\begin{pmatrix}\xi & \partial^V\end{pmatrix}} C_{n-1}^V \to .$$

We know that $\langle\partial\tau,\sigma\rangle = \pm 1$, so it is an isomorphism $\mathbb{Z}\tau \to \mathbb{Z}\sigma$ and hence by Lemma V.2.2 it follows that the homology of this is isomorphic to the homology of

$$\to C_{n+2}^V \xrightarrow{\partial^V} C_{n+1}^V \xrightarrow{\epsilon-\gamma\langle\partial\tau,\sigma\rangle\delta} C_n^V \xrightarrow{\partial^V} C_{n-1}^V \to .$$

All that is left to show is that $\epsilon - \gamma\langle\partial\tau,\sigma\rangle\delta = \partial^V$. The contribution to ∂^V of paths that do not go through σ is the same as the contribution for V' — this is given by ϵ. For paths that do pass through σ, the segment of the paths that gets to σ contribute δ, then passing over the pair $\sigma \prec \tau$ gives $\langle\partial\tau,\sigma\rangle$ and then from faces of τ to the end of the paths is exactly the contribution of γ. To see that the minus sign is correct, note that in the formula for the multiplicity we have a factor $-\langle\partial\tau,\sigma\rangle$ and so the $\langle\partial\tau,\sigma\rangle$ must come with this minus sign. □

The induction makes use of the fact that the discrete Morse chain complex is actually equal to the cellular chain complex when every cell is critical. This observation lets us think of discrete Morse homology as a generalisation of cellular homology that allows us to simplify the chain complex by reducing the size of the chain groups. In cellular or simplicial homology, the number of cells increases the size of the chain groups and hence the complexity of calculating the homology groups. Discrete Morse homology can be seen as a way of allowing us to use more cells and using the gradient paths to reduce the size of the chain groups.

2.3 Example calculations

We give two quick homology computations using discrete Morse homology. The first of these (the projective plane) highlights the fact that when all cells are critical, discrete Morse homology reduces to cellular homology.

Projective plane

The usual CW complex structure of $\mathbb{R}P^2$ with one cell in each dimension is not regular and in fact each cell must be critical as no cell is a regular cell of another. Say $C_2 = \mathbb{Z}\tau$, $C_1 = \mathbb{Z}\sigma$ and $C_0 = \mathbb{Z}\nu$. As each cell is critical, the Morse chain complex is equal to the cellular chain complex. Explicitly, the map $\partial_2 : C_2 \to C_1$ maps $\tau \to 2\sigma$ as $\langle\partial\tau,\sigma\rangle = 2$ and $\partial_1 : C_1 \to \mathbb{C}_0$ is the zero map as $\langle\partial\sigma,\nu\rangle = 0$. It follows that $H_2(\mathbb{R}P^2) = 0$, $H_1(\mathbb{R}P^2) = \mathbb{Z}/2\mathbb{Z}$ and $H_0(\mathbb{R}P^2) = \mathbb{Z}$.

The torus

Consider the Δ-complex structure[1] for the torus depicted in Figure V.2, where the green arrows give a discrete Morse gradient V, the dashed blue indicate the gradient paths we need

[1]A Δ-complex can be thought of as a generalisation of a simplicial complex where a simplex need not be uniquely determined by its vertices; moreover the faces of a cell can coincide; see [14].

to compute the discrete Morse homology, and red indicates chosen orientations of the cells. The critical cells are $\tau^{(2)}, \sigma_1^{(1)}, \sigma_2^{(1)}, \nu^{(0)}$. The chain groups are $C_2^V = \mathbb{Z}\tau$, $C_1^V = \mathbb{Z}\sigma_1 \oplus \mathbb{Z}\sigma_2$ and $C_0^V = \mathbb{Z}\nu$.

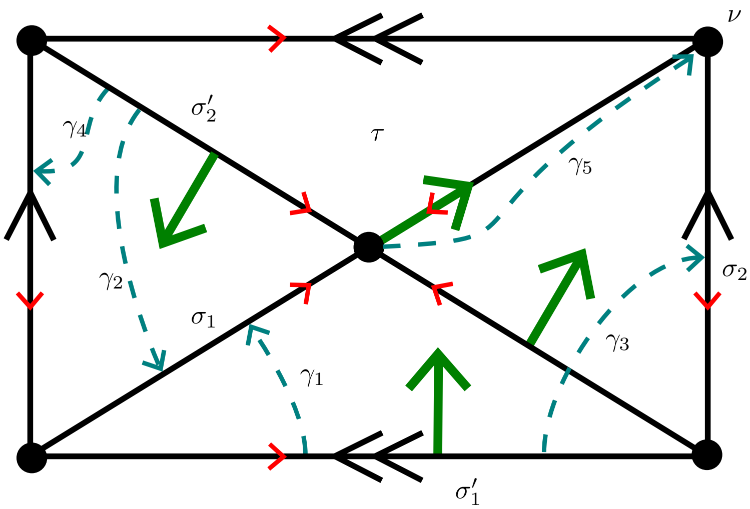


Figure V.2: Δ-complex structure for the torus; green arrows define a discrete Morse gradient and dashed blue indicates the gradient paths γ_i required for the computing the discrete Morse homology. Red indicates orientations on the 1-cells.

We have $m(\gamma_i) = 1$ for all $i \in \{1,2,3,4,5\}$. We can now compute the boundary map.

$$\begin{aligned}
\partial(\sigma_1) &= \langle \partial\sigma_1, \nu' \rangle \nu + \langle \partial\sigma_1, \nu \rangle \nu \\
&= \nu - \nu \\
&= 0 \\
\partial(\sigma_2) &= \langle \partial\sigma_2, \nu \rangle \nu \\
&= \langle \nu - \nu, \nu \rangle \\
&= 0
\end{aligned}$$

For each 2-cell pick the orientation so that two 1-cells in the boundary agree with this orientation. Then we have

$$\begin{aligned}
\partial^V(\tau) &= \langle \partial\tau, \sigma_1' \rangle(\sigma_1 + \sigma_1) + \langle \partial\tau, \sigma_2' \rangle(\sigma_1 + \sigma_2) \\
&= (\sigma_1 + \sigma_2) - (\sigma_1 + \sigma_2) \\
&= 0.
\end{aligned}$$

It follows that $H_2^V = C_2^V \cong \mathbb{Z}$, $H_1^V = C_1^V \cong \mathbb{Z} \times \mathbb{Z}$ and $H_0^V = C_1^V \cong \mathbb{Z}$ as all the boundary map is the zero map on each chain group.

Notice that even though we have more cells than needed to give the torus a cell structure, the chain groups are as simple as they can be (the discrete Morse gradient is perfect). This is in contrast with cellular homology, where the chain groups grow with the number of cells; we can think of the discrete Morse gradient as telling us how to simplify the chain groups of cellular homology.

Bibliography

[1] Antoni. *Differential manifolds.* Academic Press, Inc., Boston, MA, 1993.

[2] Rafael Ayala, Jose Antonio Vilches, Gregor Jerše, and Neža Mramor Kosta. Discrete gradient fields on infinite complexes. *Discrete Contin. Dyn. Syst.*, 30(3):623–639, 2011.

[3] Dror Bar-Natan. Fast Khovanov homology computations. *J. Knot Theory Ramifications*, 16(3):243–255, 2007.

[4] Bruno Benedetti. Discrete Morse theory for manifolds with boundary. *Trans. Amer. Math. Soc.*, 364(12):6631–6670, 2012.

[5] Bruno Benedetti. Smoothing discrete morse theory. *Ann. Sc. Norm. Super. Pisa, Classe di Scienze*, to appear. `http://arxiv.org/pdf/1212.0885v4`.

[6] Bruno Benedetti and Karim Adiprasito. Metric geometry, convexity and collapsibility. *preprint.* `http://arxiv.org/pdf/1107.5789v4.pdf`.

[7] Bruno Benedetti and Karim Adiprasito. Subdivisions, shellability and collapsibility of products. *preprint.* `http://arxiv.org/pdf/1202.6606.pdf`.

[8] Keith Burns and Gidea Marian. *Differential geometry and topology.* Chapman & Hall/CRC, Boca Raton, FL, 2005. With a view to dynamival systems.

[9] Stewart Cairns. Triangulation of the manifold of class one. *Bull. Amer. Math. Soc*, 41(8):549–522, 1935.

[10] Stewart Cairns. A simple triangulation method for smooth manifolds. *Bull. Amer. Math. Soc.*, 67:389–390, 1961.

[11] Abhijit Champanerkar, Ajit Kumar, and S. Kumaresan. Classification of surfaces via Morse theory. *Expo. Math.*, 18(1):37–73, 2000.

[12] Robin Forman. Morse theory for cell complexes. *Adv. Math.*, 134(1):90–145, 1998.

[13] Éttiene Gallais. Combinatorial realization of the Thom-Smale complex bia discrete morse theory. *Ann. Sc. Norm. Super. Pisa, Classe di Scienze*, 9(2):229–252, 2010.

[14] Allen Hatcher. *Algebraic topology.* Cambridge University Press, Cambridge, 2002.

[15] J. F. P. Hudson. *Piecewise linear topology.* University of Chicago lecture notes prepared with with assistance of J. L. Shaneson and J. Lees. W. A. Benjamin, Inc., 1969.

[16] Michael Hutchings. Lecture notes on Morse homology (with an eye towards Floer theory and pseudoholomorphic curves). *lecture notes.* `https://math.berkeley.edu/~hutching/teach/276-2010/mfp.ps`.

[17] Kiyoshi Igusa. On the homotopy type of the space of generalized Morse functions. *Topology*, 23(2):245–256, 1984.

[18] Henry King, Kevin Knudson, and Neža Mramor Kosta. Birth and death in discrete Morse theory. *preprint*. `http://arxiv.org/pdf/0808.0051.pdf`.

[19] François Laudenbach. A proof of Morse's theorem about the cancellation of critical points. *C. R. Math. Acad. Sci. Paris*, 351(11-12):483–488, 2013.

[20] John Lee. *Introduction to smooth manifolds.* Springer, New York, 2013.

[21] John Milnor. *Lectures on the h-cobordism theorem.* Notes by L. Siebenmann and J. Sondow. Princeton University Press, Princeton, N.J., 1965.

[22] John Milnor. *Morse theory.* Princeton University Press, 1973.

[23] Livui Nicolaescu. *Invititaion to Morse theory.* Springer, 2011.

[24] Richard Palais. Morse thoery on Hilbert manifolds. *Topology*, 2:299–340, 1963.

[25] Duško Jojić. Extendable shelling, simplicial and toric h-vector of some polytopes. *Publ. Inst. Math. (Beograd) (N.S.)*, 81(95):85–93, 2007.

[26] Stephen Smale. Generalized Poincaré's conjecture in dimensions greater than four. *Ann. of Math. (2)*, 74:391–406, 1961.

[27] Stephen Smale. On gradient dynamical systems. *Ann. Math. (2)*, 74(1):199–296, 1961.

[28] John Whitehead. On C^1-complexes. *Ann. of Math. (2)*, 41:809–824, 1940.

[29] Hassler Whitney. The self-intersections of a smooth n-manifold $2n$-space. *Ann. of Math. (2)*, 45:220–246, 1944.

[30] Erik Zeeman. On the dunce hat. *Topology*, 2:241–358, 1964.