Indian Institute of Technology Delhi

Department of Mathematics

A study on low dimensional manifold theory

Supervisor:

Prof. Biplab Basak

MSc Project by:

Rohin Garg

Contents

Introduction

This thesis is meant to be a study of low dimensional manifolds, in particular 4-manifolds. First we have studied introductory differential topology to get up to speed with manifolds and bundles *et al*, mainly using John Milnor's *Topology from the Differentiable Viewpoint*, after which we studied Morse Theory, mostly in the discrete setting. For this semester, we have been studying 0-efficient triangulations and the theory of shrinking surfaces. Having now gathered all these tools, we hope to use them to prove complexity bounds for certain classes of Lens spaces.

This project would not be possible without the helpful and constructive guidance of professor Biplab Basak, to whom I am grateful.

Chapter 1

Elementary Differential Topology[1]

1.1 Manifolds, smooth maps, & tangent spaces

We begin by introducing the basic machinery of smooth manifolds, starting by defining what a smooth manifold actually is. Consider $U \subset \mathbb{R}^k$ and $V \subset \mathbb{R}^l$ to be open sets. A map f from U to V is called smooth if all of the partial derivatives $\frac{\partial^n f}{\partial x_i \ldots \partial x_n}$ exist and are continuous.

More generally, let $X \subset \mathbb{R}^k$ and $Y \subset \mathbb{R}^l$ be arbitrary subsets of Euclidean spaces. A map $f : X \to Y$ is called smooth if for each $x \in X$ there exists an open set $U \subset \mathbb{R}^k$ containing x and a smooth mapping $F : U \to \mathbb{R}^l$ that coincides with f throughout $U \cap X$. If $f : X \to Y$ and $g : Y \to Z$ are smooth, note that the composition $g \circ f : X \to Z$ is also smooth. The identity map of any set X is automatically smooth.

Definition 1. *A map $f : X \to Y$ is called a diffeomorphism if f is a homeomorphism from X to Y and if both f and f^{-1} are smooth.*

Definition 2. *A subset $M \subset \mathbb{R}^k$ is called a smooth manifold of dimension m if each $x \in M$ has a neighborhood $W \cap M$ that is diffeomorphic to an open subset U of the Euclidean space $\mathbb{R}^m$. Any particular diffeomorphism $g : U \to W \cap M$ is called a parametrization of the region $W \cap M$. (The*

---

[1] This chapter is heavily based on Milnor's *Topology from the Differentiable Viewpoint*

inverse diffeomorphism $W \cap M \to U$ is called a system of coordinates on $W \cap M$.)

Example 1. *The unit sphere S^2, consisting of all $(x, y, z) \in \mathbb{R}^3$ with $x^2 + y^2 + z^2 = 1$ is a smooth manifold of dimension 2. In fact, the diffeomorphism $(x, y) \mapsto (x, y, \sqrt{1 - x^2 - y^2})$, for $x^2 + y^2 < 1$, parametrizes the region $z > 0$ of S^2. By interchanging the roles of x, y, z, and changing the signs of the variables, we obtain similar parametrizations for the regions $x > 0$, $y > 0$, $x < 0$, $y < 0$, and $z < 0$. Since these cover S^2, it follows that S^2 is a smooth manifold. More generally, the sphere $S^{n-1} \subset \mathbb{R}^n$ consisting of all $(x_1, \ldots, x_n)$ with $\sum x_i^2 = 1$ is a smooth manifold of dimension $n - 1$. For example, $S^0 \subset \mathbb{R}^1$ is a manifold consisting of just two points.*

Example 2. *A somewhat wilder example of a smooth manifold is given by the set of all $(x, y) \in \mathbb{R}^2$ with $x \neq 0$ and $y = \sin(1/x)$.*

1.2 Tangent Spaces and Derivatives

To define the notion of derivative df_x for a smooth map $f : M \to N$ of smooth manifolds, we first associate with each $x \in M \subset \mathbb{R}^k$ a linear subspace $T_xM \subset \mathbb{R}^k$ of dimension m called the tangent space of M at x. Then df_x will be a linear mapping from T_xM to T_yN, where $y = f(x)$. Elements of the vector space T_xM are called tangent vectors to M at x. Intuitively, one thinks of the m-dimensional hyperplane in $\mathbb{R}^k$ which best approximates M near x; then T_xM is the hyperplane through the origin that is parallel to this.

We first define the derivative for open sets. For any open set $U \subset \mathbb{R}^k$, the tangent space $T_{U,x}$ is defined to be the entire vector space $\mathbb{R}^k$. For any smooth map $f : U \to V$, the derivative $df_x : \mathbb{R}^k \to \mathbb{R}^l$ is defined by the formula

$$df_x(h) = \lim_{t \to 0} \frac{f(x + th) - f(x)}{t}$$

for $x \in U$, $h \in \mathbb{R}^k$. Clearly, $df_x(h)$ is a linear function of h.

$$df_x(h+l) = \lim_{t \to 0} \frac{f(x + th + tl) - f(x)}{t} = \lim_{t \to 0} \frac{f(x + th + tl) - f(x + th) + f(x + th) - f(x)}{t}$$

$$= \lim_{t \to 0} \frac{f(x + th + tl) - f(x + th)}{t} + df_x(h) = df_x(l) + df_x(h)$$

In fact, df_x is just that linear mapping which corresponds to the $l \times k$ Jacobian matrix $\left(\frac{\partial f}{\partial x_i}\right)$.

Here are two fundamental properties of the derivative operation:

1. **Chain rule:** If $f : U \to V$ and $g : V \to W$ are smooth maps, with $f(x) = y$, then
$$d(g \circ f)_x = dg_y \circ df_x.$$
In other words, to every commutative triangle of smooth maps between open subsets of $\mathbb{R}^k$, $\mathbb{R}^l$, $\mathbb{R}^m$,
$$\begin{tikzcd} & V \arrow[dr, "f"] & \\ U \arrow[ur, "g"] \arrow[rr, "g \circ f"] & & W \end{tikzcd}$$
there corresponds a commutative triangle of linear maps.
$$\begin{tikzcd} & V \arrow[dr, "df_x"] & \\ U \arrow[ur, "dg_y"] \arrow[rr, "d(g \circ f)_x"] & & W \end{tikzcd}$$

2. If I is the identity map of U, then dI_x is the identity map of $\mathbb{R}^k$. More generally, if $U \subset U'$ are open sets and $i : U \to U'$ is the inclusion map, then again di_x is the identity map of $\mathbb{R}^k$.

3. For a linear map $L : \mathbb{R}^k \to \mathbb{R}^l$, $dL_x = L$.

Proposition 1. *If f is a diffeomorphism between open sets $U \subset \mathbb{R}^k$ and $V \subset \mathbb{R}^l$, then k must equal l, and the linear mapping $df_x : \mathbb{R}^k \to \mathbb{R}^l$ must be nonsingular.*

Proof. The composition $f^{-1} \circ f$ is the identity map of U; hence $d(f^{-1})_y \circ df_x$ is the identity map of $\mathbb{R}^k$. Similarly, $df_x \circ d(f^{-1})_y$ is the identity map of $\mathbb{R}^l$. Thus df_x has a two-sided inverse, and it follows that $k = l$. $\square$

The above proposition has a partial converse which one may have seen in an advanced calculus course.

Theorem 1.2.1. ***Inverse Function Theorem****: Let $f : U \to \mathbb{R}^k$ be a smooth map, with U open in $\mathbb{R}^k$. If the derivative $df_x : \mathbb{R}^k \to \mathbb{R}^k$ is nonsingular, then f maps any sufficiently small open set U' about x diffeomorphically onto an open set $f(U')$.*

Remark 1. f may not be one-one in general, even if every df_x is nonsingular.

We now define the tangent space T_xM for an arbitrary smooth manifold $M \subset \mathbb{R}^k$. Choose a parametrization $g : U \to M \subset \mathbb{R}^k$ of a neighborhood $g(U)$ of x in M, with $g(u) = x$. Here U is an open subset of $\mathbb{R}^m$. Set T_xM equal to the image $dg_u(\mathbb{R}^m)$. We must prove that this construction does not depend on the particular choice of parametrization g. Let $h : V \to M \subset \mathbb{R}^k$ be another parametrization of a neighborhood $h(V)$ of x in M, and let $v = h^{-1}(x)$. Then $h^{-1} \circ g$ maps some neighborhood U_1 of u diffeomorphically onto a neighborhood V_1 of v. The commutative diagram of smooth maps between open sets

$$\begin{tikzcd}
 & \mathbb{R}^k & \\
U_1 \arrow[ur, "g"] \arrow[rr, "h^{-1}\circ g"] & & V_1 \arrow[ul, "h"']
\end{tikzcd}$$

gives rise to a commutative diagram of linear maps

$$\begin{tikzcd}
 & \mathbb{R}^k & \\
\mathbb{R}^m \arrow[ur, "dg_u"] \arrow[rr, "dh^{-1}\circ g_u"] & & \mathbb{R}^m \arrow[ul, "dh_v"']
\end{tikzcd}$$

It follows immediately that $\text{Image}(dg_u) = \text{Image}(dh_v)$. Thus, T_xM is well-defined.

Next, we need to show that T_xM is an m-dimensional vector space. Since $g^{-1} : g(U) \to U$ is a smooth mapping, we can choose an open set W containing x and a smooth map $F : W \to \mathbb{R}^m$ that coincides with g^{-1} on $W \cap g(U)$. Setting $U_0 = g^{-1}(W \cap g(U))$, we have the commutative diagram

$$\begin{tikzcd}
 & W \arrow[dr, "h"] & \\
U_0 \arrow[ur, "F"] \arrow[rr, "i_{U_0}"] & & \mathbb{R}^m
\end{tikzcd}$$

and so

$$\begin{tikzcd} & \mathbb{R}^k \arrow[dr, "dg_u"] & \\ \mathbb{R}^m \arrow[ur, "dF_x"] \arrow[rr, "Id_{\mathbb{R}^m}"] & & \mathbb{R}^m \end{tikzcd}$$

This diagram clearly implies that dg_u has rank m, and hence that its image T_xM has dimension m.

Now consider two smooth manifolds, $M \subset \mathbb{R}^k$ and $N \subset \mathbb{R}^l$, and a smooth map $f : M \to N$ with $f(x) = y$. The derivative $df_x : T_xM \to T_yN$ is defined as follows. Since f is smooth, there exists an open set W containing x and a smooth map $F : W \to \mathbb{R}^l$ that coincides with f on $W \cap M$. Define $df_x(v)$ to be equal to $dF_x(v)$ for all $v \in T_xM$. To justify this definition, we must prove that $dF_x(v)$ belongs to T_yN and that it does not depend on the particular choice of F. For this, choose parametrizations $g : U \to M \subset \mathbb{R}^k$ and $h : V \to N \subset \mathbb{R}^l$ for neighborhoods $g(U)$ of x and $h(y)$ of y. Replacing U by a smaller set if necessary, we may assume that $g(U) \subset W$ and that f maps $g(U)$ into $h(V)$. Then $h^{-1} \circ f \circ g$ from U to V is well defined. Now, we already have the following diagram:

$$\begin{tikzcd} W \arrow[r, "F"] & \mathbb{R}^l \\ U \arrow[u, "g"] \arrow[r, "h^{-1}\circ f\circ g"'] & V \arrow[u, "h"'] \end{tikzcd}$$

The corresponding diagram for the tangent spaces is

$$\begin{tikzcd} \mathbb{R}^k \arrow[r, "dF_x"] & \mathbb{R}^l \\ \mathbb{R}^m \arrow[u, "dg_u"] \arrow[r, "d(h^{-1}\circ f\circ g)_u"'] & \mathbb{R}^n \arrow[u, "dh_v"'] \end{tikzcd}$$

The second diagram implies that dF_x maps $TM_x = Im(dg_u)$ into $TN_y = Im(dh_v)$. Moreover, the map df_x does not depend on the particular choice of F, as we can define the derivative of f from the diagram itself by $df_x = dh_v \circ d(h^{-1} \circ f \circ g)_u \circ (dg_u)^{-1}$, and so $df_x : TM_x \to TN_y$ is a well defined linear map.

The derivative of a map between smooth manifolds shares the properties of the derivative between open sets:

1. **Chain rule:** If $f : M \to N$ and $g : N \to O$ are smooth maps, with $f(x) = y$, then
$$d(g \circ f)_x = dg_y \circ df_x.$$
2. If I is the identity map of M, then dI_x is the identity map of TM_x. More generally, if $M \subset N$ and $i : M \to N$ is the inclusion map, then $TM_x \subset TN_x$ with inclusion map dI_x.
3. If f is a diffeomorphism between M and N, then the linear mapping $df_x : TM_x \to TN_y$ is a vector isomorphism, and M and N have the same dimension.

Definition 3. *Let $f : M \to N$ be a smooth map between two smooth manifolds M & N with $dim(M) = dim(N)$. Then $x \in M$ is a regular point if df_x is invertible (and so by the Inverse Function Theorem maps a neighbourhood of x diffeomorphically onto a a neighbourhood of $y = f(x)$. $y \in N$ is a regular value if $f^{-1}(y)$ only has regular points.If df_x is not invertible, then x is a critical point of f and f(x) is a critical value.*

Now, if M is compact, we readily obtain some extra information regular values of smooth maps on M.

- **$f^{-1}(y)$ is finite (or even empty):** Clearly, since $f^{-1}(y)$ is a closed subset of a compact manifold M, it is compact. Additionally, $f^{-1}(y)$ is discrete: suppose $f^{-1}(y)$ was not discrete. Then there exists a $A \subset f^{-1}(y)$ such that $A = Im(\gamma)$, where γ is some continuous path. Without a loss of generality we assume that γ is injective. Pick $a_1 \in A$. Then since y is a regular value, df_{a_1} is invertible, and so f maps some neighbourhood U_{a_1} of a_1 diffeomorphically onto some neighbourhood V_{a_1} of y. Now, we pick $a_2 \in A$ such that $a_1 \neq a_2, |a_1 - a_2| < \epsilon$. Then $\exists t_1, t_2$ such that $|t_1 - t_2| < \delta$ and $\gamma(t_1) = a_1, \gamma(t_2) = a_2$. Furthermore, f maps some neighbourhood U_{a_2} of a_2 diffeomorphically onto some neighbourhood V_{a_2} of y. Then $y \in V_{a_1} \cap V_{a_2} \neq \emptyset$, and so $U_{a_1} \cap U_{a_2} \neq \emptyset$ $\implies a_1, a_2 \in U_{a_1} \cap U_{a_2}$ since $y \in V_{a_1} \cap V_{a_2}$. But since f is a diffeomorphism, $f(a_1) = y = f(a_2) \implies a_1 = a_2 \quad \perp$. Thus $f^{-1}(y)$ is discrete. And since a closed and discrete subset of a comapct space is finite, $f^{-1}(y)$ is finite.

- Let $\#f^{-1}(y)$ be the number of points in $f^{-1}(y)$. Then $\#f^{-1}(y)$ **is locally constant as a function of** y **(over regular values)**: Let $x_1, x_2, , , , x_k$ be the points in $f^{-1}(y)$ (finite since M is compact). Choose pairwise disjoint neighbourhoods $U_1, , , U_k$ that are mapped diffeomorphically onto $V_1, , , , V_k$ neighbourhoods of $y \in N$. Let $V = V_1 \cap V_2 \cap \cdots V_k \setminus f(M \setminus \cup_{i=1}^k U_i)$. Then $y' \in V, \#f^{-1}(y) = \#f^{-1}(y')$.

Corollary 1. *An immediate corollary of these basic definitions is the Fundamental Theorem of Algebra: every nonconstant complex polynomial has at least one zero. The proof uses the sterographic projection on S^2 and can be found in Milnor's book.*

1.3 Submanifolds and manifolds with boundary

While the set of critical values may not be finite, a seminal theorem by Arthur Sard in 1942, building on work by Anthony Morse, tells us that is in a certain sense 'small'.[2]

Theorem 1.3.1. *Let $f : U \to \mathbb{R}^n$ be a smooth defined on some open subset $U \subset \mathbb{R}^m$. If $C[\{x \in U | rank(df_x) < n\}$, then $f(C)$ has Lebesgue measure zero in $\mathbb{R}^n$.*

Note: If $m < n$, then by Sard's theorem $f(C)$ has measure zero. Thus, for the most part we shall focus on the case $m \geq n$.

We can extend this results of smooth manifolds in general: let $f : M \to N$ be a smooth from a manifold of dimension m to a manifold of dimension n. Let C is the set of $x \in M$ such that $rank(df_x) < n$ i.e. the set of critical points. Then since M can be covered by a countable collection of neighbourhoods diffeomorphic to open subsets of $\mathbb{R}^m$, then the set of regular values i.e. $\mathbb{R}^n - f(C)$ is everywhere dense ($f(C)$ has measure zero).

To use this result in the future, we need a few key lemmas.

[2] Sard, Arthur (1942), "The measure of the critical values of differentiable maps", Bulletin of the American Mathematical Society, 48 (12): 883–890, doi:10.1090/S0002-9904-1942-07811-6, MR 0007523, Zbl 0063.06720.

Lemma 1. *Let $f : M \to N$ be a smooth map between manifolds of dimension $m \geq n$ respectively, and y a regular value of f. Then $f^{-1}(y)$ is a smooth manifold of dimension $m - n$.*

Proof. Pick $x \in f^{-1}(y)$. y is regular, so df_x maps TM_x onto TN_y. Thus $ker(df_x)$ has dimension $m - n$. If $M \subset \mathbb{R}^k$, we choose a linear map $L : \mathbb{R}^k \to \mathbb{R}^{m-n}$ that is nonsingular on $ker(df_x)$. Next, define $F : M \to N \times \mathbb{R}^{m-n}$ by $F(z) = (f(z), L(z))$, then $dF_x = (df_x, L)$. By construction, dF_x is nonsingular. Thus F maps some neighbourhood U of x diffeomorphically onto some neighbourhood V of $(y, L(x))$. Now, F maps $f^{-1}(y)$ to the hyperplane $y \times \mathbb{R}^{m-n}$. Thus $f^{-1}(y) \cap U$ is mapped diffeomorphically $y \times \mathbb{R}^{m-n} \cap V$. Thus, $f^{-1}(y)$ is a smooth manifold of dimension $m - n$. □

We show that S^{m-1} is a smooth manifold. Consider $f : \mathbb{R}^n \to \mathbb{R}$ given by $f(x) = x_1^2 + x_2^2 + \ldots x_n^2$. Then $df_x(h) = 2(x_1 h_1 + \ldots x_n h_n)$. Then $\forall y \neq 0, y$ is a regular value. Then since 1 is a regular value, $f^{-1}(1) = S^{m-1}$ is a smooth manifold of dimension $m - (m - 1) = 1$.

If M' is a manifold is contained in M, we know $TM'_x \subseteq TM_x$. Then $(TM'_x)^{\perp}$, the orthogonal complement of TM'_x in TM_x, is the space of vectors normal to M' to M.

Lemma 2. *Let $M' = f^{-1}(y)$ for a regular value of $f : M \to N$. $Ker(df_x) : TM_x \to TN_y$ is precisely equal to tangent space $TM'_x \subset TM_x$ of the submanifold $M' = f^{-1}(y)$. Hence df_x maps the orthogonal complement of TM'_x isomorphically onto TN_y.*

Proof. We have the diagram

$$\begin{tikzcd} M' \arrow[r, "i_M"] \arrow[d, "f'"'] & M \arrow[d, "f"] \\ y \arrow[r, "i_y"'] & N \end{tikzcd}$$

we get the corresponding diagram

$$\begin{tikzcd} TM'_x \arrow[r, "id"] \arrow[d, "df'_x"'] & TM_x \arrow[d, "df_x"] \\ 0 \arrow[r, "0"'] & TN_y \end{tikzcd}$$

By going around the diagram, we can see that df_x maps TM'_x to zero in TN_y. Thus, $(TM'_x)^{\perp}$ is mapped isomorphically onto TN_y. □

Manifolds with Boundary

We can extend our existing definitions to define manifolds with a boundary. Consider
$$H^m = \{(x_1, \cdots, x_m)\,\varepsilon R^m \mid x_m \geq 0\}$$
The boundary ∂H^m is defined to be the hyperplane $R^{m-1} \times 0 \subset R^m$.

Definition 4. *A subset $X \subset R^k$ is called a smooth m-manifold with boundary if each $x \in X$ has a neighborhood $U \cap X$ diffeomorphic to an open subset $V \cap H^m$ of H^m. The boundary ∂X is the set of all points in X which correspond to points of ∂H^m under such a diffeomorphism. Then ∂X is a smooth manifold of dimension $m-1$, while the interior $X - \partial X$ is a smooth manifold of dimension m.*

We can generate a host of manifolds with boundary by looking at certain submanifolds of manifolds without boundary and with boundary.

Lemma 3. *Let M be a manifold without boundary and let $g : M \to \mathbb{R}$ be a smooth map with 0 as a regular value. The set of x in M with $g(x) \geq 0$ is a smooth manifold, with boundary equal to $g^{-1}(0)$.*

Proof. Since 0 is a regular value, $g^{-1}(0)$ is a smooth manifold with dimension $m-1$. Now, since $(0, \infty)$ is open and g is smooth, $g^{-1}(0, \infty)$ is open in M, and so is a smooth manifold of dimension m. Thus, $g^{-1}[0, \infty)$ is mapped to $R_{\geq 0}$ and so is a smooth manifold with boundary $g^{-1}(0)$. □

As an example, consider the unit disk D^n defined by $1 - \Sigma_{i=1}^n x_i^2 \geq 0$. Then D^n is a manifold with S^{n-1} as its boundary.

Lemma 4. *Now consider a smooth map $f : X \to N$ from an m-manifold with boundary to an n-manifold, where $m > n$. If $y \in N$ is a regular value, both for f and for the restriction $f\mid_{\partial X}$, then $f^{-1}(y) \subset X$ is a smooth $(m-n)$-manifold with boundary. The boundary of $\partial\,(f^{-1}(y))$ is the intersection of $f^{-1}(y)$ with ∂X.*

Proof. Since we have to prove a local property, using the coordinate charts on our manifold it suffices to consider the special case of a map $f : H^m \to R^n$, with regular value $y \in R^n$. Let $\bar{x} \in f^{-1}(y)$. If $\bar{x}$ is an interior point, then as before $f^{-1}(y)$ is a smooth manifold in the neighborhood of $\bar{x}$. Suppose that $\bar{x}$ is a boundary point. Choose a smooth map $g : U \to R^n$ that is defined throughout a neighborhood of $\bar{x}$ in R^m and coincides with f on $U \cap H^m$. Replacing U by a smaller neighborhood if necessary, we can assume that g has no critical points. Hence $g^{-1}(y)$ is a smooth manifold of dimension $m-n$.

Now, let $\pi_m : g^{-1}(y) \to R$ be the mth coordinate projection restricted to $g^{-1}(y)$,

$$\pi\left(x_1, \cdots, x_{\mathrm{m}}\right) = x_{\mathrm{m}}$$

Then π_m has 0 as a regular value: for $x \in \pi^{-1}(0)$, $T(g^{-1}(y))_x = ker(dg_x = df_z)$ at a point $x \in \pi^{-1}(0)$, and since $f\mid_{\partial H^m}$ is regular at x, $ker(dg_x)$ cannot be completely contained in $R^{m-1} \times 0$ (naturally since $m > m-1$).

Thus the set $g^{-1}(y) \cap H^m = f^{-1}(y) \cap U$, consisting of all $x \in g^{-1}(y)$ with $\pi(x) \geq 0$, is a smooth manifold, by previous lemma with boundary equal to $\pi^{-1}(0)$. □

For compact manifolds we have another important result.

Lemma 5. *For a compact manifold with boundary X, there does not exist a smooth map $f : X \to \partial X$ that leaves the boundary fixed pointwise i.e. no smooth retraction exists.*

Proof. Suppose such and f exists. Let $y \in \partial X$ be regular. Now, since by assumption $f\mid_{\partial X} = Id_{\partial X}$, y is also a regular value for $f\mid_{\partial X}$. Then by Lemma 4 $f^{-1}(y)$ is a smooth 1-manifold with boundary

$$f^{-1}(y) \cap \partial X = \{y\}$$

However, by the classification of compact manifolds, the only compact 1-manifolds are finite disjoint unions of S^1 (which have no boundary) and intervals (which have two boundary points). Thus, $\partial f^{-1}(y)$ must have an even number of boundary points, which is a contradiction. $\perp$ □

Corollary 2. *Since D^n is a compact manifold with boundary S^{n-1}, the identity map on S^{n-1} cannot be smoothly extended to a smooth map on D^n.*

Another important set of results that arise from these lemmas are Brouwer's Fixed Point Theorems.

Corollary 3. *Any smooth map $g : D^n \to D^n$ has a fixed point.*

Corollary 4. *Any continuous map $g : D^n \to D^n$ has a fixed point.*

The latter is a simple consequence of applying the Stone-Weierstrass approximation theorem to the former.

1.4 Degree mod 2 of a smooth map

A key tool for classification problems in topology and geometry is the use of 'invariants' (of a suitable category). Here, we introduce a certain type of weak homotopy invariant called the mod 2 degree of a map. From now on, for two smooth maps f, g we will say that $f \sim g$ to denote that the two maps are smoothly homotopic i.e. the two maps are homotopic and the homotopy map $F : X \times [0, 1] \to Y$ is itself smooth. If f, g are diffeomorphisms and for each t the map $F_t = F(x, t)$ is a diffeomorphism, we then say that f, g are smoothly isotopic.

Definition 5. *Let $f : M \to N$ be a smooth map with M a compact manifold without boundary and N connected, and dim(M) = dim(N). Then the mod 2 degree of f at y, denoted $deg_2(f, y) = \#f^{-1}(y)$.*

We will show that the mod 2 degree does not depend on the choice of regular value y and in fact depends only on the smooth homotopy class of f.

Lemma 6 (**Homotopy Lemma**). *Let $f, g : M \to N$ be smooth maps with M a compact manifold without boundary and N connected, dim(M) = dim(N), and $f \sim g$. If y is a regular value for both f and g, then $deg_2(f, y) = deg_2(g, y)$*

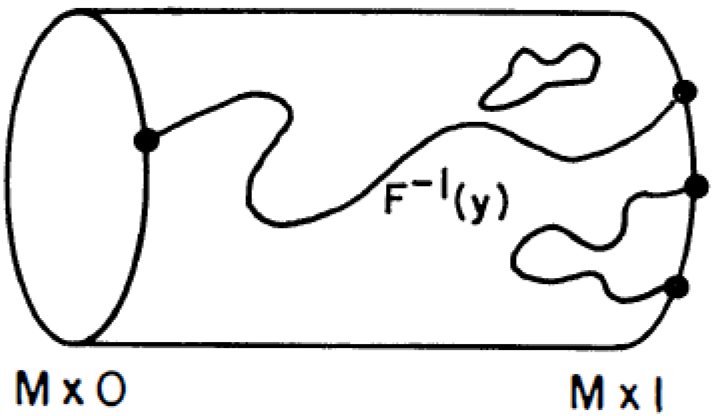
Proof. Let F be the smooth homotopy map between f and g. We then have the follwoing two cases.

The first case is when y is also a regular value of F. Then $F^{-1}(y)$ is a compact 1-manifold, with boundary equal to

$$F^{-1}(y) \cap (M \times 0 \cup M \times 1) = f^{-1}(y) \times 0 \cup g^{-1}(y) \times 1$$

which means that the total number of boundary points is $\#f^{-1}(y) + \#g^{-1}(y)$

Figure 1.1: The number of boundary points on the left is congruent to the number on the right modulo 2



Again, by the classification of compact 1-manifolds, $F^{-1}(y)$ must have an even number of boundary points. Thus $\#f^{-1}(y) + \#g^{-1}(y)$ is even, and therefore

$$\#f^{-1}(y) \equiv \#g^{-1}(y) \pmod 2$$

The second case is when y is not a regular value of F. Now, we had earlier showed that for a compact manifold and a smooth map h the function $\#h^{-1}(y)$ is locally constant over regular values. Thus there exist neighborhoods $V_1, V_2 \subset N$ of y, consisting of regular values of f and g respectively such that

$$\forall y' \in V_1 \#f^{-1}(y') = \#f^{-1}(y); \forall y' \in V_2 \#g^{-1}(y') = \#g^{-1}(y)$$

Next, we pick a regular value z of F in $V_1 \cap V_2$ (which always exists by Sard's theorem). Then

$$\#f^{-1}(y) = \#f^{-1}(z) \equiv \#g^{-1}(z) = \#g^{-1}(y)$$

□

We now show that the mod 2 degree is indeed a homotopy invariant of maps. We will use the following lemma, the proof of which can be found in Milnor's text.

Lemma 7 (Homogeneity Lemma)**.** *Let y and z be arbitrary interior points of the smooth, connected manifold N. Then there exists a diffeomorphism $h : N \to N$ that is smoothly isotopic to the identity and carries y into z.*

Theorem 1.4.1. *Assume that M is compact and without boundary, that N is connected, and that $f : M \to N$ is smooth. If y and z are regular values of f then*

$$deg_2(f) = \#f^{-1}(y) \equiv \#f^{-1}(z) \quad (\text{modulo}\, 2)$$

Indeed, $deg_2(f)$ depends only on the smooth homotopy class of f

Proof. Consider regular values y and z. Let h be a diffeomorphism of N which is isotopic to the identity and which carries y to z (this exists because of the Homogeneity Lemma). Then z is a regular value of the composition $h \circ f$. Since $h \circ f \sim f$, via the Homotopy Lemma we have that

$$\#(h \circ f)^{-1}(z) \equiv \#f^{-1}(z)(\text{mod}2)$$

However, we know that

$$(h \circ f)^{-1}(z) = f^{-1}h^{-1}(z) = f^{-1}(y)$$

and so

$$\#(h \circ f)^{-1}(z) = \#f^{-1}(y)$$

Therefore $\#f^{-1}(y) \equiv \#f^{-1}(z) \pmod{2}$ as required. Next, we consider a smooth map g such that $f \sim g$. Sard's theorem ensures that there exists an element $y \in N$ which is a regular value for both f and g. By the Homotopy Lemma, $\#f^{-1}(y) \equiv \#g^{-1}(y)$, and so

$$\deg_2 f \equiv \#f^{-1}(y) \equiv \#g^{-1}(y) \equiv \deg_2 g$$

□

We can now demonstrate some basic uses of the mod 2 degree:

1. A constant map $c : M \to M$ has even mod 2 degree. The identity map of M has odd degree. Hence the identity map of a compact manifold without boundary is not homotopic to a constant.

2. For the sphere, this result provides another proof of our earlier assertion that no smooth map $f : D^{n+1} \to S^n$ leaves the boundary sphere pointwise fixed: if such a map existed we would have a smooth homotopy $F : S^n \times [0,1] \to S^n, \quad F(x,t) = f(tx)$ between a constant map and the identity.

1.5 Oriented manifolds and the Brouwer degree of a smooth map

While the mod 2 degree is a powerful invariant, its power is 'limited' (we will show that it cannot distinguish between certain types of maps). To overcome its limitations, we will require the notion of an oriented manifold. We begin by defining the orientation of a vector space. An orientation for a finite dimensional real vector space is an equivalence class of ordered bases as follows: the ordered basis $(b_1, \cdots, b_n)$ determines the same orientation as the basis $(b'_1, \cdots, b'_n)$ if $b'_i = \sum a_{ij} b_i$ with $\det(a_{ij}) > 0$. It determines the opposite orientation if $\det(a_{ij}) < 0$. Thus each positive dimensional vector space has precisely two orientations. The vector space R^n has a standard orientation corresponding to the standard basis. For a zero dimensional vector space we simply say the orientation is positive or negative.

Definition 6. *An oriented smooth manifold consists of a manifold M together with a choice of orientation for each tangent space TM_x. If $m \geq 1$, these are required to fit together as follows: For each point of M there should exist a neighborhood $U \subset M$ and a diffeomorphism h mapping U onto an open subset of R^m or H^m which is orientation preserving, in the sense that for each $x \in U$ the isomorphism dh_x carries the specified orientation for TM_x into the standard orientation for R^m.*

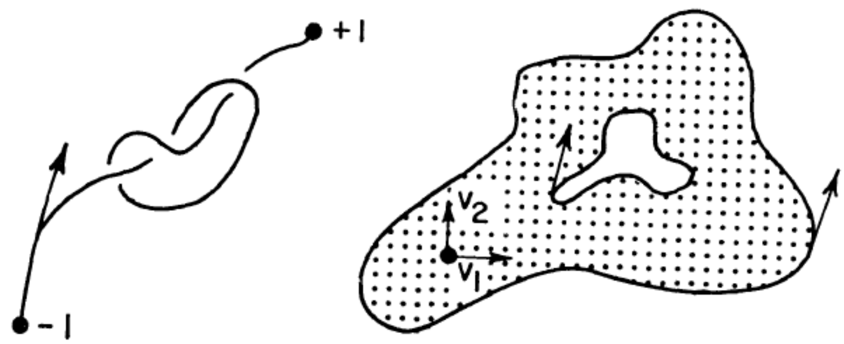
An orientable connected manifold has exactly two orientations. Now, if M has a boundary, then the tangent space at the boundary point has three kinds of vectors:

1. Vectors tangent to the boundary forming an $(m-1)$ dimensional subspace $T(\partial M)_x \subset TM_x$
2. outward facing vectors, forming an open half space bounded by $T(\partial M)_x$
3. inward facing vectors forming a complementary half space.

Then the orientation on M induces an orientation on ∂M in the following way: for $x \in \partial M$ choose a positively oriented basis $(v_1, v_2, \cdots, v_m)$ for TM_n in such a way that $v_2, \cdots, v_m$ are tangent to the boundary (assuming that $m \geq 2$) and that v_1 is a vector of the second kind i.e. an outward facing vector. Then the orientation on ∂M is the orientation induced by $(v_2, \cdots, v_m)$.

For a one dimensional manifold, each boundary point is assigned the orientation -1 or +1 based on whether a positively oriented vector at x points inward or outward.

Figure 1.2: Examples of boundary orientation



Brouwer Degree of a map

Definition 7. *Let M and N be oriented n-dimensional manifolds without boundary such that M is compact and N is connected and let f be a smooth map. The Brouwer degree (hereafter just degree) of f is defined as follows: let $x \in M$ be a regular point of f, so that $df_x : TM_x \to TN_{f(x)}$ is a linear isomorphism between oriented vector spaces. Define the sign of df_x to be +1 or -1 according as df_z preserves or reverses orientation. Then for any regular value $y \in N$ define* $\deg(f; y) = \sum_{x \in f^{-1}(y)} \operatorname{sign} df_x$

Proposition 2. $\deg(g \circ f) = \deg(g) \times \deg(f)$

Proof.

$$\deg(g \circ f) = \sum_{x \in (g \circ f)^{-1}(z)} \operatorname{sign} d(g \circ f)_x = \sum_{x \in (g \circ f)^{-1}(z)} \operatorname{sign} dg_{f(x)} \operatorname{sign} df_x$$

$$= \sum_{y \in g^{-1}(z)} \sum_{x \in f^{-1}(y)} \operatorname{sign} dg_y \operatorname{sign} df_x = \sum_{y \in g^{-1}(z)} \operatorname{sign} dg_y \sum_{x \in f^{-1}(y)} \operatorname{sign} df_x = \deg(g) \times \deg(f)$$

$\square$

Since M is compact, $deg(f;y)$ is a locally constant function over regular values. As with the mod 2 degree, we will show that the degree $deg(f)$ of a map does not depend on the choice of regular value and instead depends on the smooth homotopy class.

For this, we will need 2 key lemmas:

Lemma 8. *Let $M = \partial X$ for some compact oriented manifold X and that M has the boundary orientation induced from X. If $f : M \to N$ extends to a smooth map $F : X \to N$, then* $\deg(f;y) = 0$ *for every regular value y.*

Proof. First suppose that y is a regular value for F, as well as for $f = F\mid_M$. The compact 1-manifold $F^{-1}(y)$ is a finite union of arcs and circles, with only the boundary points of the arcs lying on $M = \partial X$. Let $\gamma \subset F^{-1}(y)$ be one of these arcs, with $\partial\gamma = \{a\} \cup \{b\}$. We want to show that for any such arc we have that $\operatorname{sign} df_a + \operatorname{sign} df_b = 0$ and so by taking the sum over all such arcs $\deg(f;y) = 0$.

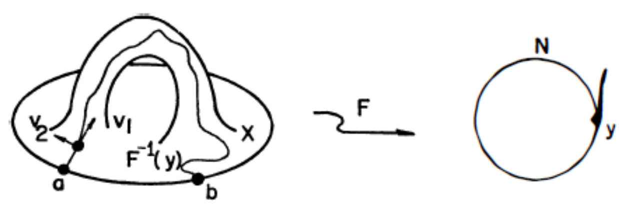


Figure 1.3: Orienting $F^{-1}(y)$

Now, the orientations for X and N determine an orientation for γ in the following manner: for any $x \in \gamma$, let $(v_1, \cdots, v_{n+1})$ be a positively oriented basis for TX_x with v_1 tangent to γ. Then v_1 determines the required orientation for $T\gamma_x$ if and only if dF_x carries $(v_2, \cdots, v_{n+1})$ into a positively oriented basis for TN_y. We define $v_1(x)$ to be the positively oriented unit vector tangent to γ at each $x \in \gamma$. Clearly v_1 is a smooth function, and $v_1(x)$ points outward at the 'ending' boundary point (say b) of γ and inward at the 'starting' boundary point a of γ. Then by definition of the sign function we have that $\operatorname{sign} df_a = -1, \quad \operatorname{sign} df_b = +1$ and so $\operatorname{sign} df_a + \operatorname{sign} df_b = 0$. Next, by taking the sum over all such arcs γ, we have that $\deg(f;y) = 0$.

More generally, suppose that y_0 is a regular value for f, but not for F. The function $\deg(f;y)$ is constant within some neighborhood U of y_0. Hence, by Sard's theorem, we can choose a regular value y for F within U and observe that $\deg(f;y_0) = \deg(f;y) = 0$. $\square$

Lemma 9. *Suppose* $f \sim g$. *Then degree* $\deg(g;y)$ *is equal to* $\deg(f;y)$ *for any common regular value* y.

Proof. Let f be smoothly homotopic via F. The manifold $[0,1] \times M$ can be oriented as a product, and will then have boundary consisting of $1 \times M$ (with the positive orientation) and $0 \times M$ (with the negative orientation). Thus the degree of $F \mid \partial\left([0,1] \times M^n\right)$ at a regular value y is equal to $\deg(g;y) - \deg(f;y)$. Now, note that the map $f+g$ defined on $\partial[0,1] \times M = 1 \times M - 0 \times M$ extends to the smooth map F defined on $[0,1] \times M$. Then by the previous theorem $\deg(g;y) - \deg(f;y) = 0$. Thus $\deg(g;y) = \deg(f;y)$ $\square$

We now prove our main theorems:

Theorem 1.5.1. *The integer* $\deg(f;y)$ *does not depend on the choice of regular value* y.

Proof. Let y, z be regular values for $f : M \to N$. Now, by the Homogeneity Lemma there exists a diffeomorphism $h : N \to N$ such that h is smoothly isotopic to the identity and carries y into z. So then z is a regular value of $h \circ f$, since $(h \circ f)^{-1}(z) = f^{-1} \circ h^{-1}(z) = f^{-1}(y)$. Using Lemma 2, since $h \circ f \sim f$, we get that $deg(h \circ f; z) = deg(f; z)$.
Since h is smoothly isotopic to the identity, h preserves orientation, and so $signdf_x = signd(h \circ f)_{y=f(x)}$. Thus, $deg(h \circ f; h(y)) = deg(f;y)$. Since $h(y) = z$, we get that $deg(f;y) = deg(f;z)$. $\square$

Thus from now on we shall simply use $deg(f)$.

Theorem 1.5.2. *If* $f \sim g$, *then* $\deg f = \deg g$.

Proof. Let $f \sim g$ and deg(f) = deg(f;y) for some regular value y of f. Suppose y is also a regular value of g, then by Lemma 2 $deg(f) = deg(g)$. If y is not a regular value, then by Sard's theorem we pick a common regular value y_0, and so by the previous theorem we get that deg(f) = deg(g). $\square$

We can now do some basic computations.

- The complex function $z \to z^k, z \neq 0$, maps the unit circle onto itself with degree k ($k \in \mathbb{Z}$)
- The constant mapping $f : M \to c \in N$ has degree zero, since $\forall y \in N$ such that $y \neq c$ we have that y is trivially a regular value.
- A diffeomorphism $f : M \to N$ has degree +1 or -1 depending on whether f preserves or reverses orientation. Thus an orientation reversing diffeomorphism of a compact manifold without boundary is not smoothly homotopic to the identity.
- One example of an orientation reversing diffeomorphism is provided by the reflection with respect to the ith coordinate $r_i : S^n \to S^n$, where $r_i(x_1, \cdots, x_{n+1}) = (x_1, \cdots, -x_i, \cdots, x_{n+1})$. Thus, the antipodal map of S^m has degree $(-1)^{n+1}$, since $-x = r_1 \circ r_2 \circ \cdots \circ r_{n+1}(x)$, and so if n is even then the antipodal map on S^n is not smoothly homotopic to the identity (something we could not deduce simply by using the mod 2 degree).

Using this last, we can prove a proposition related to vector fields on spheres.

Definition 8. *A smooth tangent vector field on $M \subset R^k$ is a smooth map $v : M \to R^k$ such that $v(x) \in TM_x$ for all $x \in M$.*

Thus for a sphere, a smooth tangent vector field satisfies $\forall x \in S^m, v(x) \cdot x = 0$

Proposition 3. *S^n admits a smooth field of nonzero tangent vectors if and only if n is odd.*

Proof. Since $v(x)$ is nonzero for all x, then without loss of generality we assume that $v(x) \cdot v(x) = 1$ for all $x \in S^n$, as we can always normalise the vector field as $\bar{v}(x) = \frac{v(x)}{\|v(x)\|}$. Thus we consider v to be a smooth function from S^m to itself. Next, we define a smooth homotopy $F : S^n \times [0, \pi] \to S^n$ by $F(x, y) = xcos(t) + v(x)sin(t)$. Clearly, F is smooth and $F(x, \theta) \cdot F(x, \theta) = 1$, while $F(x, 0) = x, \quad F(x, \pi) = -x$. Thus F is a smooth homotopy between the identity map and the antipodal map. But then as we have shown this means that n must be odd.

For the other implication, for $n = 2k - 1$, we define a tangent vector field $v(x_1, \cdots, x_{2k}) = (x_2, -x_1, x_4, -x_3, \cdots, x_{2k}, -x_{2k-1})$ defines a nonzero tangent vector field on S^m. Then $v(x) \cdot x = 0$ for all x and $v(x)$ is nonzero everywhere. $\square$

A further corollary incidentally, that the antipodal map of S^m is homotopic to the identity for n odd. Later on we will prove a result of Hopf which will demonstrate that two smooth maps from a connected n-manifold to the n-sphere are smoothly homotopic if and only if they have the same degree.

1.6 Vector fields, the Index sum, and Euler numbers

As the last proposition of our the previous example hinted at, the types of vector field that can be admitted on a manifold says something about the nature of the manifold. In this section, we will make that intuition concrete and show that in fact the types of smooth vector fields that can be defined on a smooth manifold actually tell us a lot of information. For this, we shall introduce a few more analytical concepts.

Definition 9. *Consider first an open set $U \subset R^m$ and a smooth vector field $v : U \to R^m$ with an isolated zero at the point $z \in U$. The function $\bar{v}(x) = v(x)/\|v(x)\|$ maps a small sphere centered at z into the unit sphere. The degree of this mapping is called the index ι of v at the zero z.*

Example 3. *Over $\mathbb{C}$, the polynomial z^k defines a smooth vector field with a zero of index k at the origin, while $\bar{z}^k$ defines a vector field with a zero of index $-k$.*

More visual examples of vector fields with indices $-1, 0, 1, 2$ are shown in Figure 1.4.

Definition 10. *Let f be a smooth map $M \to N$ with a vector field on each manifold. The vector fields v on M and v' on N correspond under f if the derivative df_z carries $v(x)$ into $v'(f(x))$ for each $x \in M$. If f is a diffeomorphism, then v' is uniquely determined by v, and we write $v' = df \circ v \circ f^{-1}$*

We will use the concept of the index of a vector field to build a topological invariant. As usual, we begin with some lemmas.

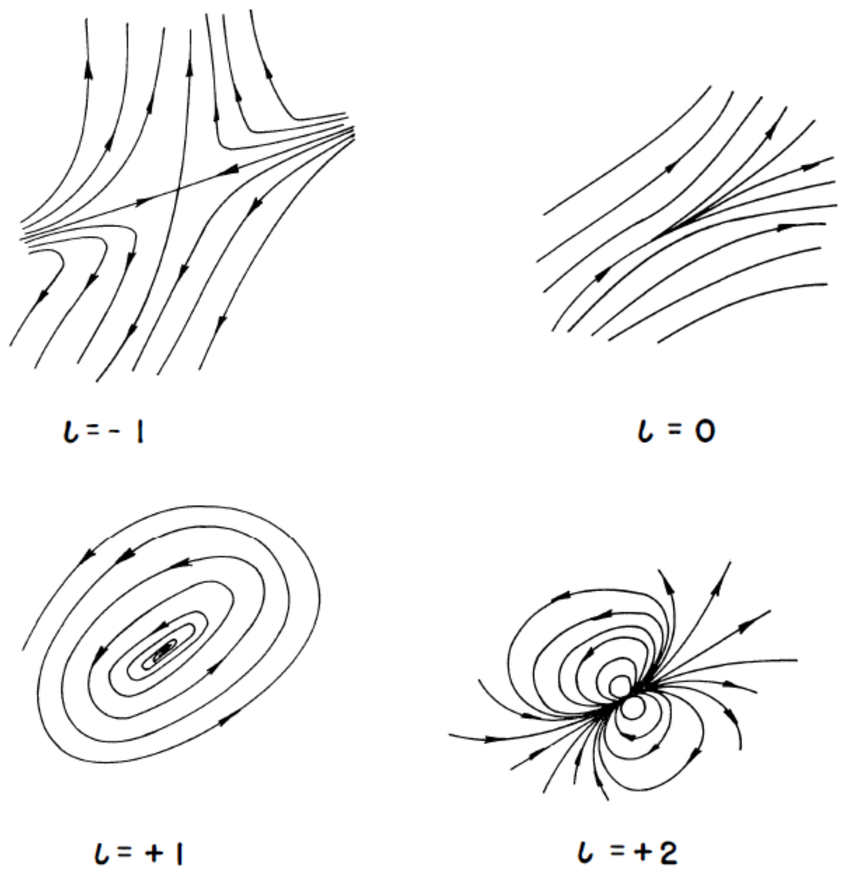


Figure 1.4: Examples of vectors fields with certain indices

Lemma 10. *Any orientation preserving diffeomorphism f of R^m is smoothly isotopic to the identity.*

Proof. Without a loss of generality we assume that $f(0) = 0$. A simple computation shows that we can define the derivative at 0 to be $df_0(x) = \lim_{t \to 0} \frac{f(tx)}{t}$, and so we can define an isotopy $F : R^m \times [0, 1] \to R^m$ by

$$F(x,t) = f(tx)/t \quad \text{for} \quad 0 < t \leq 1$$
$$F(x,0) = df_0(x)$$

To prove that F is smooth $t \to 0$, we write $f(x) = x_1 g_1(x) + \cdots + x_m g_m(x)$ where $g_1, \cdots, g_m$ are suitable smooth functions, and note that

$$F(x,t) = x_1 g_1(tx) + \cdots + x_{\mathrm{m}} g_{\mathrm{m}}(tx)$$

for all values of t. We can write f in the above manner as follows: $f(x) = \int_0^1 \frac{\partial}{\partial s} f(sx)ds = \sum_{i=1}^m \int_0^1 x_i \frac{\partial}{\partial x_i} f(sx)ds$, and so $g_i(x) = \int_0^1 \frac{\partial}{\partial x_i} f(sx)ds$. Thus f is smoothly isotopic to the linear mapping df_0, Now, since f is orientation preserving, $df_0 \in GL(n, \mathbb{R})_+$ which is the path component of $GL(n, \mathbb{R})$, and so df_0 is clearly isotopic to the identity. □

Lemma 11. *The index of a vector field is invariant under diffeomorphism: Suppose that the vector field v on U corresponds to $v' = df \circ v \circ f^{-1}$ on U' under a diffeomorphism $f : U \to U'$. Then the index of v at an isolated zero z is equal to the index of v' at $f(z)$.*

Proof. Without a loss of generality we assume that the $z = 0$ (i.e. 0 is the isolated zero) and that U is convex. Suppose f preserves orientation, then by Lemma 10 we can construct a one-parameter family of embeddings $f_t : U \to R^m$ with f_0 = identity, $f_1 = f$, and $f_t(0) = 0$ for all t. Let $v_t = df_t \circ v \circ f_t^{-1}$ on $f_t(U)$, which corresponds to v on U. The v_t are all well defined and nonzero on a sufficiently small sphere centered at 0, and so $\bar{v}(x) = v(x)/\|v(x)\|$ is well defined. Then via our constructed homotopy $\bar{v} = \bar{v_0} \sim \bar{v_t} \sim \bar{v_1} = \bar{v}'$. Clearly, smooth homotopies preserve degree, so the index of $v = v_0$ at 0 must be equal to the index of $v' = v_1$ at 0 . This proves Lemma 1 for orientation preserving diffeomorphisms.

In the case that f reverses orientation, it is enough to consider the special case of a reflection ρ. Then $v' = \rho \circ v \circ \rho^{-1}$ and so $\bar{v}'(x) = v'(x)/\|v'(x)\|$ on some ϵ-sphere satisfies $\bar{v}' = \rho \circ \bar{v} \circ \rho^{-1}$, which implies that the degree of $\bar{v}'$ equals the degree of $\bar{v}$. □

Using Lemma 11, we can define the index of a vector field v on an arbitrary manifold M as follows: If $g : U \to M$ is a parametrization of a neighborhood of z in M, then the index ι of v at z is defined to be equal to the index of the corresponding vector field $dg^{-1} \circ v \circ \mathrm{g}$ on U at the zero $g^{-1}(z)$.

By defining the index of a vector field on an arbitrary manifold, we can obtian an extremely powerful result, proven by Hopf in 1926 after earlier partial results by Poincaré, Brouwer, and Hadamard.

Theorem 1.6.1 (**Poincaré-Hopf Theorem**). *Let M be a compact manifold and w a smooth vector field on M with isolated zeros. If M has a boundary, then w is required to point outward at all boundary points.The index sum $\sum \iota$*

of the indices at the zeros of such a vector field is equal to the Euler number $\chi(M) = \sum_{i=0}^{m}(-1)^i \quad \mathrm{rank}(H_i(M))$, *where* $H_i(M)$ *is the ith Homology group of* M. *In particular this index sum is a topological invariant of* M : *it does not depend on the particular choice of vector field.*

For the complete proof of the theorem we refer the reader to Hopf's original paper.[3] We will prove some special cases of this theorem.

Lemma 12 (Hopf Lemma)**.** *Let* $X \subset R^m$ *be a compact* m*-manifold with boundary. The Gauss map* $g : \partial X \to S^{m-1}$ *assigns to each* $x \in \partial X$ *the outward unit normal vector at* x. *If* $v : X \to R^m$ *is a smooth vector field with isolated zeros, and if* v *points out of* X *along the boundary, then the index sum* $\sum \iota$ *is equal to the degree of the Gauss mapping from* ∂X *to* S^{m-1}. *In particular,* $\sum \iota$ *does not depend on the choice of* v.

Proof. We begin by removing an ϵ-ball around each isolated zero to obtain a new manifold x' with boundary. The function $\bar{v}(x) = v(x)/\|v(x)\|$ maps x' into S^{m-1}. Then by Lemma 8 the sum of the degrees of $\bar{v}$ restricted to the various boundary components is zero i.e. $\sum deg(\bar{v} \mid_{\partial X'})$. Now $\bar{v} \mid_{\partial X} \sim g$: define $u_t = t\bar{v} + (1-t)g$, and let $\bar{u} = \frac{u_t}{\|u_t\|}$. Furthermore, each removed ball was endowed with the opposite orientation as X: the orientation on each ∂B_j is given by the normal pointing into the sphere and away from the manifold. However, when we calculate the degree, we use the normal pointing outwards from the sphere. Thus, the degrees on the other boundary components add up to $-\sum \iota$, with the minus sign occurring since each B_j gets the wrong orientation. Hence $\deg(g) - \sum_\iota = 0$ □

Remark 2. The degree of g is also known as the "curvatura integra" of ∂X, since it can be expressed as a constant times the integral over ∂X of the Gaussian curvature. This was shown by Chern in 1945 [4], and the result is known as the Chern-Gauss-Bonnet Theorem (a generalisation of the Gauss-Bonnet Theorem for surfaces).

[3]Hopf, H.. "Vektorfelder in n-dimensionalen-Mannigfaltigkeiten." Mathematische Annalen 96 (1927): 225-250. ¡http://eudml.org/doc/159166¿.

[4]Chern, Shiing-shen (October 1945). "On the Curvatura Integra in a Riemannian Manifold". The Annals of Mathematics. 46 (4): 674–684. doi:10.2307/1969203. JSTOR 1969203. S2CID 123348816.

Definition 11. *Consider first a vector field v on an open set $U \subset R^m$ and think of v as a mapping $U \to R^m$, so that $dv_z : R^m \to R^m$ is defined. The vector field v is nondegenerate at z if the linear transformation dv_z is nonsingular.*

Lemma 13. *The index of v at a nondegenerate zero z is either +1 or -1 according as the determinant of dv_z is positive or negative.*

Proof. Consider v to be a diffeomorphism from some convex neighborhood U_0 of z into R^m. WLOG we assume that $z = 0$. If v preserves orientation, we have seen by Lemma 10 and 11 that $v \mid U_0$ can be deformed smoothly into the identity without introducing any new zeros, and so the index is equal to $+1$. If v reverses orientation, then v can similarly be deformed into a reflection; hence $\iota = -1$ □

Lemma 14. *z is a zero of a vector field w on a manifold $M \subset R^k$. Here, we consider w to be a map from M to R^k and so the derivative $dw_z : TM_z \to R^k$ is defined. In fact, the derivative dw_x carries TM_z into the subspace $TM_z \subset R^k$ and hence can be considered as a linear transformation from TM_z to itself. If this linear transformation has determinant $D \neq 0$ then z is an isolated zero of w with index equal to +1 or -1 according as D is positive or negative.*

Proof. Let $h : U \to M$ be a parametrization of some neighborhood of z. Let e^i denote the i-th basis vector of R^m and let $l^i = dh_{\mathrm{u}}\left(e^i\right) = \partial h / \partial u_i$ so that the vectors $t^1, \ldots, t^m$ form a basis for the tangent space $TM_{h(u)}$. First, we note that $dw_{h(w)}\left(t^i\right) = d(w \circ h)_u\left(e^i\right) = \partial w(h(u)) / \partial u_i$. Next, Let $v = \sum v_i e^i$ be the vector field on U which corresponds to the vector field w on M. By definition, since h is a local diffeomorphism $v = dh^{-1} \circ w \circ h$, so that $w(h(u)) = dh_u(v) = \sum v_i t^i$. Therefore $\partial w(h(u)) / \partial u_i = \sum_i \left(\partial v_i / \partial u_i\right) t^i + \sum_i v_i \left(\partial t^i / \partial u_i\right)$. Then, evaluating at the zero $h^{-1}(z)$ of v (so $v \equiv 0$), we get that

$$dw_z\left(t^i\right) = \sum_i \left(\partial v_i / \partial u_i\right) t^i + 0$$

This shows that dw_a maps TM_z into itself, and the determinant D of this linear transformation $TM_z \to TM_z$ is equal to the determinant of the matrix $\left(\partial v_i / \partial u_i\right)$, and so by the previous lemma has index +1 or -1. □

We can now prove another partial result for a compact manifold without boundary $M \subset R^k$. Let N_ϵ denote the closed ϵ-neighborhood of M (i.e., the set of all $x \in R^k$ with $\|x - y\| \leq \epsilon$ for some $y \in M$). For ϵ sufficiently small it can be shown that N_ϵ is also a smooth manifold with boundary.

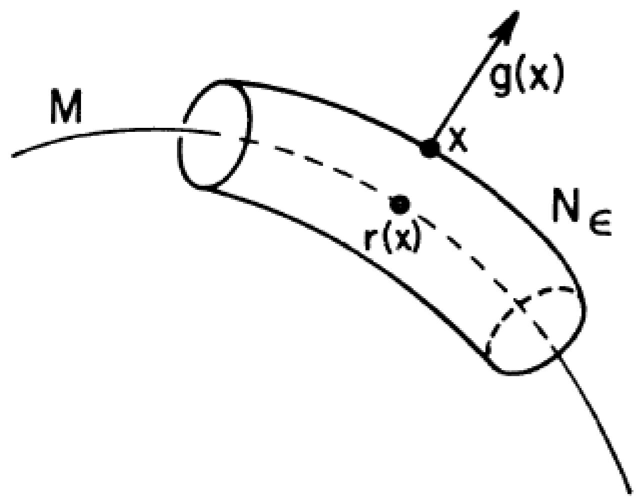


Figure 1.5: An example of an ϵ neighbourhood

Theorem 1.6.2. *For any vector field v on M with only nondegenerate zeros, the index sum $\sum$ is equal to the degree of the Gauss map*

$$g : \partial N_{\mathrm{e}} \to S^{k-1}$$

In particular this sum does not depend on the choice of vector field.

Proof. For $x \in N_{\mathrm{c}}$ we define $r(x) \in M$ to be the closest point of M. Note that the $x - r(x)$ is perpendicular to the tangent space of M at $r(x)$, for otherwise $r(x)$ would not be the closest point of M. For ϵ sufficiently small $r(x)$ is smooth and well defined. Consider also the squared distance function $\varphi(x) = \|x - r(x)\|^2$. A straightforward calculation shows that the gradient of φ is given by $\operatorname{grad} \varphi = 2(x - r(x))$. Then $\forall x$ of the level surface $\partial N_\epsilon = \varphi^{-1}(\epsilon^2)$, the outward unit normal vector is given by

$$g(x) = \frac{\operatorname{grad} \varphi}{\|\operatorname{grad} \varphi\|} = \frac{(x - r(x))}{\epsilon}$$

We extend v to a vector field w on N_ϵ by

$$w(x) = (x - r(x)) + v(r(x))$$

Then w points outward along the boundary, since the inner product $w(x) \cdot g(x)$ is equal to $\epsilon > 0$. Now, w only vanishes at the zeros of v in M since $(x - r(x))$ and $v(r(x))$ are mutually orthogonal. Computing the derivative of w at a zero z of w, we see that for $h \in TM_z$, since for $h \in TM_z$, $w(x) = x - r(x) + v(r(x)) = x - x + v(x) = v(x)$:

$$dw_z(h) = d(w \mid_{TM_z}) = dv_z(h)$$

and for $h \in TM_z^{\perp}$, since $v(z) = 0$ and for t small $r(z + th) = z$:

$$dw_z(h) = \lim_{t \to 0} \frac{w(z + th) - w(z)}{t} = \lim_{t \to 0} \frac{z + th - r(z + th) + v(r(z + th)) - z + r(z) - v(r(z))}{t}$$

$$= \lim_{t \to 0} \frac{z + th - z + v(z) - z + r(z) - v(z)}{t} = \lim_{t \to 0} \frac{th + 0 - z + z - 0}{t} = h$$

Additionally, $v(z) = 0 \implies r(z) = z$ and so $w(z) = z - z + 0 = 0$. Since w and v have the same determinant and the same zeroes, the index of w at z is equal to the index of v at z. Since w points outwards along the boundary, we apply the Hopf Lemma to w to get that

$$\sum \iota_z(v) = \sum \iota_z(w) = deg(g)$$

□

1.7 Framed manifolds and cobordisms

So far we have defined the degree of a smooth map $f : M \to M'$ for when the manifolds M and M' are oriented and have the same dimension. In this chapter, we will study a powerful generalisation of the degree of a map to smooth maps from an arbitrary compact manifold without boundary to a sphere. We begin with another round of definitions

Definition 12. *Let N and N' be compact n-submanifolds without boundary of a m-manifold M with. $m - n$ is called the codimension of the submanifolds. N is cobordant to N' within M if the subset $N \times [0, \epsilon) \cup N' \times (1 - \epsilon, 1]$ of $M \times [0, 1]$ can be extended to a compact manifold $X \subset M \times [0, 1]$ such that $\partial X = N \times 0 \cup N' \times 1$ and X does not intersect $M \times 0 \cup M \times 1$ except at the points of ∂X.*

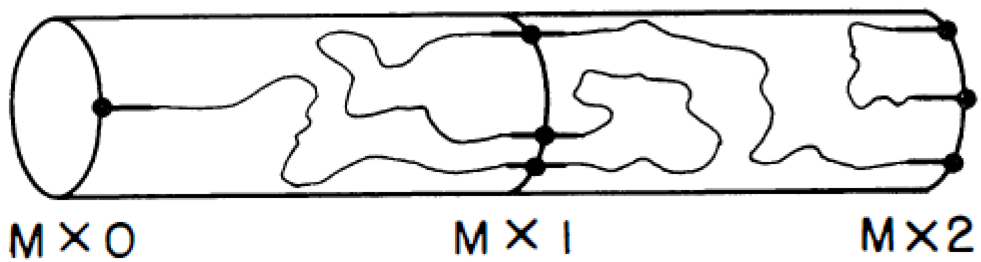


Figure 1.6: Transitvity of cobordisms

It can be easily shown that cobordisms form an equivalence relation on manifolds.

Definition 13. *A framing of the submanifold $N \subset M$ is a smooth function $\mathfrak{v}$ which assigns to each $x \in N$ a basis. $\mathfrak{v}(x) = (v^1(x), \cdots, v^{m-n}(x))$ for the space $TN_x^\perp \subset TM_x$ of normal vectors to N in M at x. The pair $(N, \mathfrak{v})$ is called a framed submanifold of M. Two framed submanifolds $(N, \mathfrak{v})$ and $(N', \mathfrak{w})$ are framed cobordant if there exists a cobordism $X \subset M \times [0,1]$ between N and N' and a framing u of X, so that*

$$u^i(x,t) = \big(v^i(x), 0\big) \text{ for } (x,t) \varepsilon N \times [0, \epsilon)$$
$$u^i(x,t) = \big(w^i(x), 0\big) \text{ for } (x,t) \varepsilon N' \times (1-\epsilon, 1]$$

Then framed cobordisms too form an equivalence relation.
Finally, we consider a smooth map $f : M \to S^p$ and a regular value $y \in S^p$. The map f induces a framing of the manifold $f^{-1}(y)$ in the following manner: Choose a positively oriented basis $\mathrm{v} = (v^1, \cdots, v^p)$ for the tangent space $T\left(S^p\right)_y$. For each $x \in f^{-1}(y)$ by Lemma 2 we have that $df_x : TM_x \to T\left(S^P\right)_y$ maps the subspace $Tf^{-1}(y)_x$ to zero and maps its orthogonal complement $Tf^{-1}(y)_x^\perp$ isomorphically onto $T\left(S^p\right)_y$. Hence $\forall i$ there is a unique vector $w^i(x) \in Tf^{-1}(y)_x^\perp \subset TM_x$ that maps into v^i under df_x. We denote $\mathfrak{w} = f^*\mathfrak{v}$ for the framing $w^1(x), \cdots, w^p(x)$ of $f^{-1}(y)$.

Definition 14. *The framed manifold $(f^{-1}(y), f^*\mathfrak{v})$ will be called the Pontryagin manifold associated with f.*

Now f admits many Pontryagin manifolds based on different choices of y and $\mathfrak{v}$, but as we shall show they all belong to the same framed cobordism class by proving three main theorems.

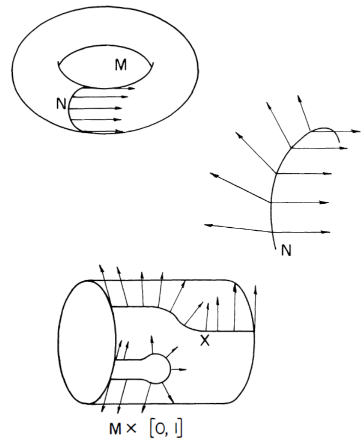


Figure 1.7: A framed cobordism

Theorem 1.7.1. *If y' is another regular value of f and $\mathfrak{v}'$ is a positively oriented basis for $T\left(S^p\right)_{y'}$, then the framed manifold $\left(f^{-1}\left(y'\right), f^*\mathfrak{v}'\right)$ is framed cobordant to $(f^{-1}(y), f^*\mathfrak{v})$.*

Theorem 1.7.2. *Two mappings from M to S^p are smoothly homotopic if and only if the associated Pontryagin manifolds are framed cobordant.*

Theorem 1.7.3. *Any compact framed submanifold $(N, \mathfrak{w})$ of codimension p in M occurs as Pontryagin manifold for some smooth mapping $f : M \to S^p$.*

Combining these three theorems, we get what is known as the Pontryagin-Thom correspondence: homotopy classes of maps from M to S^p are in one-one correspondence with the framed cobordism classes of submanifolds of codimension p!

To prove theorem A, we require three lemmas.

Lemma 15. *If v and $\forall$ are two different positively oriented bases at y, then the Pontryagin manifold $(f^{-1}(y), f^*\mathfrak{v})$ is framed cobordant to $(f^{-1}(y),\ f^*\mathfrak{v}')$*

Proof. Choose a smooth path from $\mathfrak{v}$ to $\mathfrak{v}'$ in the space of all positively oriented bases for $T\left(S^p\right)_y$ by identifying the latter with the space $GL^+(p, R)$ of matrices with positive determinant, which is path connected. Such a path gives rise to the required framing of the cobordism $f^{-1}(y) \times [0, 1]$. $\square$

Sometimes for convenience of notation we may omit $f^*\mathfrak{v}$ and simply talk of "the framed manifold $f^{-1}(y)$."

Lemma 16. *If y is a regular value of f, and z is sufficiently close to y, then $f^{-1}(z)$ is framed cobordant to $f^{-1}(y)$.*

Proof. We can assume that the set $f(C)$ of critical values is compact, we can choose $\epsilon > 0$ so that the ϵ-neighborhood of y contains only regular values. Given z with $\|z - y\| < \epsilon$, choose a smooth one-parameter family of rotations (i.e. an isotopy) $r_t : S^p \to S^p$ so that $r_1(y) = z$, and so that r_t is the identity for $0 \leq t < \epsilon'$, r_t equals r_1 for $1 - \epsilon' < t \leq 1$, and each $r_t^{-1}(z)$ lies on the great circle from y to z, and hence is a regular value of f.

Now, we define the homotopy $F : M \times [0,1] \to S^\Phi$ by $F(x,t) = r_t f(x)$. For each t z is a regular value of the composition $r_t \circ f : M \to S^D$
Then by construction z is a regular value for the mapping F. Hence $F^{-1}(z) \subset M \times [0,1]$ is a framed manifold and provides a framed cobordism between the framed manifolds $f^{-1}(z)$ and $(r_1 \circ f)^{-1}(z) = f^{-1} r_1^{-1}(z) = f^{-1}(y)$. □

Lemma 17. *If $f \sim g$ and y is a regular value for both, then $f^{-1}(y)$ is framed cobordant to $g^{-1}(y)$.*

Proof. We pick a homotopy F with

$$\begin{aligned} F(x,t) &= f(x) \quad 0 \leq t < \epsilon \\ F(x,t) &= g(x) \quad 1 - \epsilon < t \leq 1 \end{aligned}$$

By Sard's Theorem and the previous Lemma 16, we can choose a regular value z for F which is close enough to y so that $f^{-1}(z)$ is framed cobordant to $f^{-1}(y)$ and so that $g^{-1}(z)$ is framed cobordant to $g^{-1}(y)$. Then $F^{-1}(z)$ is a framed manifold and provides a framed cobordism between $f^{-1}(z)$ and $g^{-1}(z)$. □

We now prove Theorem 1.7.1.

Proof. Given any two regular values y and z for f, we can choose rotations

$$r_t : S^p \to S^p$$

so that r_0 is the identity and $r_1(y) = z$. Thus f is homotopic to $r_1 \circ f$; hence $f^{-1}(z)$ is framed cobordant to

$$(r_1 \circ f)^{-1}(z) = f^{-1} r_1^{-1}(z) = f^{-1}(y)$$

□

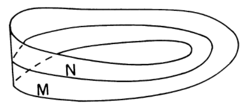


Figure 1.8: The Möbius strip is an example of an unframable manifold

The proof of Theorem 1.7.3 will require an important result, the proof of which can be found in Milnor's book or in *Differential Topology* by Guillemin and Pollack. Consider $N \subset M$ a framed submanifold of codimension p with framing $\mathfrak{v}$. Assume that N is compact and that M, N are without boundary.

Theorem 1.7.4 (Product Neighborhood Theorem). *Some neighborhood of N in M is diffeomorphic to the product $N \times R^*$. Furthermore the diffeomorphism can be chosen so that each $x \in N$ corresponds to $(x, 0) \in N \times R^p$ and so that each normal frame $\mathfrak{v}(x)$ corresponds to the standard basis for R^p.*

It must be noted that product neighborhoods do not exist for arbitrary submanifolds (for instance for unframable manifolds).
We now prove Theorem 1.7.3.

Proof. Consider $N \subset M$ be a compact framed manifold without boundary. We pick a product representation $g : N \times \mathbb{R}^p \to V \subset M$ for a neighbourhood V of N and then define the projection $\pi : V \to \mathbb{R}^p$ by $\pi(g(x, y)) = y$. Then 0 is a regular value of π and so the framed manifold $\pi^{-1}(0) = N$ can be framed as below.

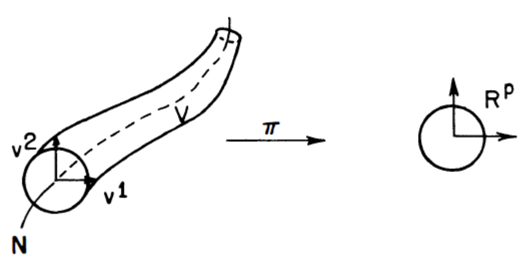


Figure 1.9: The framing of the constructed projection

Now choose a smooth map $\varphi : R^p \to S^p$ which maps every x with $\|x\| \geq 1$ into a base point s_0, and maps the open unit ball in R^p diffeomorphically onto $S^p - s_0$. Next, we define $f : M \to S^p$ by

$$\begin{array}{lll} f(x) = \varphi(\pi(x)) & \text{for} & x \in V \\ f(x) = s_0 & \text{for} & x \notin V \end{array}$$

Then f is smooth by construction, and the point $\varphi(0)$ is a regular value of f. Thus the corresponding Pontryagin manifold $f^{-1}(\varphi(0)) = \pi^{-1}(0)$ is equal to the framed manifold N. $\square$

Finally, to prove Theorem 1.7.2 we must first show that the Pontryagin manifold of a map determines its homotopy class.

Lemma 18. *Let $f, g : M \to S^p$ be smooth maps with a common regular value y. If the framed manifold $(f^{-1}(y), f^*\mathfrak{v})$ is equal to $(g^{-1}(y), g^*\mathfrak{v})$, then f is smoothly homotopic to g.*

Proof. Let $N = f^{-1}(y)$. Then $f^*\mathfrak{v} = g^*\mathfrak{v} \implies df_x = dg_x$. Suppose f coincides with g throughout an entire neighborhood V of N. Let $h : S^p - \{y\} \to R^p$ be the stereographic projection. Then the homotopy

$$\begin{array}{l} F(x,t) = f(x) \quad \text{for } x \in V \\ F(x,t) = h^{-1}[t \cdot h(f(x)) + (1-t) \cdot h(g(x))] \text{ for } x \in M - N \end{array}$$

shows that $f \sim g$.

Thus it is enough to show that we can deform f so that it coincides with g in some small neighborhood of N, being careful not to map any new points into y during the deformation. Choose a product representation $N \times R^p \to V \subset M$ for a neighborhood V of N, where V is small enough so that $f(V)$ and $g(V)$ do not contain the antipode $\bar{y}$ of y. Identifying V with $N \times R^p$ and identifying $S^p - \{\bar{y}\}$ with R^p, we obtain corresponding mappings $F, G : N \times R^p \to R^p$ with $F^{-1}(0) = G^{-1}(0) = N \times 0$ and with $dF_{(x,0)} = dG_{(x,0)} = \pi_{\mathbb{R}^p}$ for all $x \in N$. We will first find a constant c so that

$$F(x,u) \cdot u > 0, \quad G(x,u) \cdot u > 0$$

for $x \in N$ and $0 < \|u\| < c$. In such a case the points $F(x,u)$ and $G(x,u)$ belong to the same open half-space in R^p and so So the homotopy $(1 - t)F(x,u) + tG(x,u)$ between F and G will not map any new points into 0 for $\|u\| < c$. By Taylor's theorem

$$\|F(x,u) - u\| \leq c_1 \|u\|^2, \text{ for } \quad \|u\| \leq 1$$

Hence
$$|(F(x,u)-u)\cdot u| \leq c_1\|u\|^3$$
and
$$F(x,u)\cdot u \geq \|u\|^2 - c_1\|u\|^3 > 0$$
for $0 < \|u\| < \mathrm{Min}\left(c_1^{-1}, 1\right)$, with a similar inequality for G. To avoid moving distant points we select a smooth map $\lambda : R^p \to R$ with
$$\lambda(u) = 1 \text{ for } \|u\| \leq c/2$$
$$\lambda(u) = 0 \text{ for } \|u\| \geq c.$$
Then homotopy $F_t(x,u) = [1-\lambda(u)t]F(x,u) + \lambda(u)tG(x,u)$ deforms $F = F_0$ into a mapping F_1 that coincides with G in the region $\|u\| < c/2$, coincides with F for $\|u\| \geq c$, and has no new zeros. Then finally by making the corresponding deformation to f we get $f \sim g$. □

We finally prove Theorem 1.7.2

Proof. If $f \sim g$ are smoothly homotopic, then by Lemma 17 the Pontryagin manifolds $f^{-1}(y)$ and $g^{-1}(y)$ are framed cobordant. Conversely, given a framed cobordism $(X, \mathfrak{w})$ between $f^{-1}(y)$ and $g^{-1}(y)$, we construct a homotopy $F : M \times [0,1] \to S^p$ whose Pontryagin manifold $(F^{-1}(y), F^*\mathfrak{v})$ is precisely equal to $(X, \mathfrak{w})$. Setting $F_t(x) = F(x,t)$, note that the maps F_0 and f have exactly the same Pontryagin manifold. By Lemma 18 $F_0 \sim f$ and $F_1 \sim g$, and so $f \sim g$. □

Remark 3. We can easily generalise theorems 1.7.1,1.7.2, & 1.7.3 to apply to a manifold with boudnary by only considering mappings which carry the boundary into a fixed base point. The homotopy classes of such mappings are in one-one correspondence with the cobordism classes of framed submanifolds of codimension p. If $p \geq \frac{1}{2}m + 1$, then this set of homotopy classes can be given thestructure of an abelian group, called the p-th cohomotopy group $\pi^p(M, \partial M)$. The composition operation in $\pi^p(M, \partial M)$ corresponds to the union operation for disjoint framed submanifolds of Interior (M).

We end by sampling some theorems of Hopf. Let M be a connected and oriented manifold of dimension $m = p$. A framed submanifold of codimension p is just a finite set of points with a preferred basis at each. Let $\mathrm{sgn}(x)$ equal +1 or -1 according as the preferred basis determines the right or wrong orientation. Then $\sum \mathrm{sgn}(x)$ is clearly equal to the degree of the associated map $M \to S^m$. But it is not difficult to see that the framed cobordism class of the 0-manifold is uniquely determined by this integer $\sum \operatorname{sgn}(x)$. Thus we have proved the following.

Theorem 1.7.5. *If M is connected, oriented, and without boundary, then two maps $M \to S^m$ are smoothly homotopic if and only if they have the same degree.*

On the other hand, suppose that M is not orientable. Then given a basis for TM_z we can slide x around M in a closed loop so as to transform the given basis into one of opposite orientation and so we have:

Theorem 1.7.6. *If M is connected but nonorientable, then two maps $M \to S^m$ are homotopic if and only if they have the same mod 2 degree.*

Chapter 2

Discrete Morse Theory

2.1 CW complexes

Many main theorems of both smooth and discrete Morse Theory are best stated using the machinery of CW Complexes (as opposed to simplicial complexes). While there is sometimes a "loss of information" when going from simplicial or even Δ complexes to CW comples, CW complexes have a few advantages. For one, they are oftne easier to compute, since they are simples and 'smaller' - a simplicial complex decomposition of a topological space X will often have way more cells and simplices than the corresponding CW complex CW decomposition.

Let Bd denote the closed unit ball in d-dimensional Euclidean space. That is, $B^d = \{x \in E^d : |x| \leq 1\}$. The boundary of B^d is the unit $(d-1)$-sphere $S^{d-1} = \{x \in E^d : |x| = 1\}$. A d-cell is a topological space which is homeomorphic to B^d. If σ is a d-cell, then we denote by $\dot{\sigma}$ the subset of ξ corresponding to $S^{d-1} \subset Bd$ under any homeomorphism between B^d and ξ. A cell is a topological space which is a d-cell for some d.

The basic operation of CW complexes is the notion of attaching a cell. Let X be a topological space, σ a d-cell and $f : \dot{\sigma} \to X$ a continuous map. We let $X \cup_f \sigma$ denote the disjoint union of X and ξ quotiented out by the equivalence relation that each point $s \in \dot{\sigma}$ is identified with $f(s) \in X$. We refer to this operation by saying that $X \cup_f \sigma$ is the result of attaching the cell σ to X via f. The map f is called the attaching map. Note that the attaching map must be defined on all of $\dot{\sigma}$. That is, the entire boundary of σ must be "glued" to X.

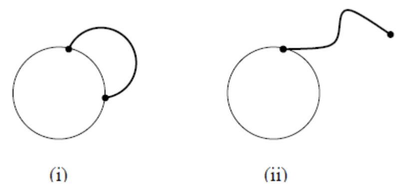


Figure 2.1: Correctly and incorrectly attached cells

Definition 15. *A finite CW complex is any topological space X such that there exists a finite nested sequence*

$$\emptyset \subset X_0 \subset X_1 \subset \cdots \subset X_n = X$$

such that for each $i = 0, 1, 2, \ldots, n, X_i$ is the result of attaching a cell to $X_{(i-1)}$. Note that this definition requires that X_0 be a 0 -cell. If X is a CW complex, we refer to any sequence of spaces as above as a CW decomposition of X. Suppose that in the CW decomposition, of the $n+1$ cells that are attached, exactly c_d are d-cells. Then we say that the CW complex X has a CW decomposition consisting of c_dd-cells for every d. We note that a (closed) d-simplex is a d-cell. Thus a finite simplicial complex is a CW complex, and has a CW decomposition in which the cells are precisely the closed simplices.

Example 4. *Below we demonstrate a CW decomposition of a 2-dimensional torus which, beginning with the 0 -cell, requires attaching two 1 -cells and then one 2-cell. Here we can see one of the most compelling reasons for considering CW complexes rather than just simplicial complexes. Every simplicial decomposition of the 2-torus has at least 7 vertices, 21 edges and 14 triangles.*

The homotopy type of $X \cup_f \sigma$ depends only on the homotopy type of X and the homotopy class of f.

Theorem 2.1.1. *Let $h : X \to X'$ denote a homotopy equivalence, σ a cell, and $f_1 : \dot{\sigma} \to X, f_2 : \dot{\sigma} \to X'$ two continuous maps. If $h \circ f_1$ is homotopic to f_2, then $X \cup_{f_1} \sigma$ and $X' \cup_{f_2} \sigma$ are homotopy equivalent.*

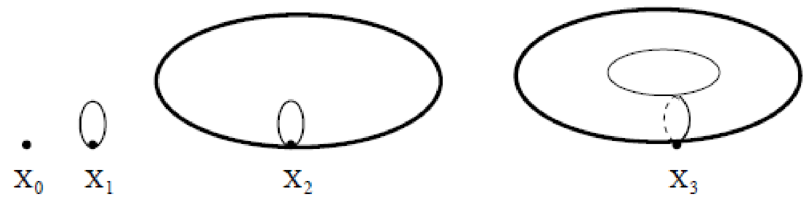


Figure 2.2: CW decomposition of the 2-Torus

Corollary 5. *Let X be a topological space, σ a cell, and $f_1, f_2 : \dot{\sigma} \to X$ two continuous maps. If f_1 and f_2 are homotopic, then $X \cup_{f_1} \sigma$ and $X \cup_{f_2} \sigma$ are homotopy equivalent.*

Let us now consider some basic examples of CW decompositions.

1. Suppose X is a topological space which has a CW decomposition consisting of exactly one 0 -cell and one d-cell. Then X has a CW decomposition $\emptyset \subset X_0 \subset X_1 = X$. The space X_0 must be the 0 -cell, and $X = X_1$ is the result of attaching the d-cell to X_0. Since X_0 consists of a single point, the only possible attaching map is the constant map. Thus X is constructed from taking a closed d-ball and identifying all of the points on its boundary. One can easily see that this implies that the resulting space is a d-sphere.

2. Let us next consider X is a topological space which has a CW decomposition consisting of exactly one 0 -cell and nd-cells. Then X has a CW decomposition such that X_0 is the 0 -cell, and for each $i = 1, 2, \ldots, n$ the space X_i is the result of attaching a d cell to $X_{(i-1)}$. From the previous example, we know that X_1 is a d-sphere. The space X_2 is constructed by attaching a d-cell to X_1. The attaching map is a continuous map from a $(d-1)$-sphere to X_1. Every map of the $(d-1)$-sphere into X_1 is homotopic to a constant map (since $\pi_{(d-1)}(X_1) \cong \pi_{(d-1)}\left(S^d\right) \cong 0$). If the attaching map is actually a constant map, then it is easy to see that the space X_2 is the wedge of two d-spheres, denoted by $S^d \wedge S^d$. (The wedge of a collection of topological spaces is the space resulting from choosing a point in each space, taking the disjoint union of the spaces, and identifying all of the chosen points.) Since the attaching map must be homotopic to a constant map, Corollary 5 implies that X_2 is homotopy equivalent to a wedge of two d-spheres.

When constructing X_3 by attaching a d-cell to X_2, the relevant information is a map from S^{d-1} to X_2, and the homotopy type of the resulting space is determined by the homotopy class of this map. All such maps are homotopic to a constant map (since $\pi_{d-1}(X_2) \cong \pi_{d-1}(S^d \wedge S^d) \cong 0$). Since X_2 is homotopy equivalent to a wedge of two d-spheres, and the attaching map is homotopic to a constant map, it follows from Theorem 1.3 that X_3 is homotopy equivalent to the space that would result from attaching a d-cell to $S^d \wedge S^d$ via a constant map, i.e., X_3 is homotopy equivalent to a wedge of three d-spheres. Continuing in this fashion, we can see that X must be homotopy equivalent to a wedge of nd-spheres.

3. Suppose that X is a CW complex which has a CW decomposition consisting of exactly one 0 -cell, one 1 -cell and one 2-cell. Let us consider a CW decomposition for X with these cells: $\emptyset \subset X_0 \subset X_1 \subset X_2 = X$. We know that X_0 is the 0 -cell. Suppose that X_1 is the result of attaching the 1-cell to X_0. Then X_1 must be a circle, and X_2 arises from attaching a 2-cell to X_1. The attaching map is a map from the boundary of the 2-cell, i.e., a circle, to X_1 which is also a circle. Up to homotopy, such a map is determined by its winding number, which can be taken to be a nonnegative integer. If the winding number is 0 , then without altering the homotopy type of X we may assume that the attaching map is a constant map, which yields that $X \sim S^1 \wedge S^2$ (where $\sim$ denotes homotopy equivalence). If the winding number 1 then without altering the homotopy type of X we may assume that the attaching map is a homeomorphism, in which case X is a 2-dimensional disc. If the winding number is 2 , then without altering the homotopy type of X we may assume that the attaching map is a standard degree 2 mapping (i.e., that wraps one circle around the other twice, with no backtracking), and in this case is$X$ is the 2-dimensional projective space $\mathbb{P}^2$. In fact, each winding number results in a homotopically distinct space. These spaces can be distinguished by their homology, since $H_1(X, \mathbb{Z})$ for the space X resulting from an attaching map with winding number n is isomorphic to $\mathbb{Z}/n\mathbb{Z}$.

The last example is particularly instructive: we assumed that the 1-cell was attached before the 2-cell, and we must consider the alternative order, in which X_1 is the result of attaching a 2-cell to X_0. In this case, X_1 is a 2-sphere, and $X = X_2$ is the result of attaching a 1-cell to X_1. The attaching map is a map of S^0 into S^2. Since S^2 is connected (i.e., $\pi_0\left(S^2\right)=0$) all such maps are homotopic to a constant map. Taking the attaching map to be a constant map yields that $X=S^1 \wedge S^2$. Thus adding the cells in this order merely resulted in fewer possibilities for the homotopy type of X. This is a general phenomenon. Generalizing ths argument and using the fact that $\pi_i\left(S^d\right)=0$ for $i<d$, we get

Theorem 2.1.2. *Let*

$$\emptyset \subset X_0 \subset X_2 \subset \cdots \subset X_n = X$$

be a CW decomposition of a finite CW complex X. Then X is homotopy equivalent to a finite CW decomposition with precisely the same number of cells of each dimension as above, and with the cells attached so that their dimensions form a nondecreasing sequence.

CW complexes also have a well defined theory of homology: the theory of cellular homology. Let X be a CW complex with a fixed CW decomposition. Suppose that in this decomposition X is constructed from exactly c_d cells of dimension d for each $d=0,1,2,\ldots,n=\dim(K)$, and let $C_d(X,\mathbb{Z})$ denote the space $\mathbb{Z}^{c_d}$ (more precisely, $C_d(X,\mathbb{Z})$ denotes the free abelian group generated by the d-cells of X, each endowed with an orientation).

Theorem 2.1.3. *There are boundary maps $\partial_d : C_d(X,\mathbb{Z}) \to C_{d-1}(X,\mathbb{Z})$, for each d, so that*

$$\partial_{d-1} \circ \partial_d = 0$$

and such that the resulting differential complex

$$0 \longrightarrow C_n(X,\mathbb{Z}) \xrightarrow{\theta_n} C_{n-1}(X,\mathbb{Z}) \xrightarrow{\theta_{n-1}} \cdots \xrightarrow{\theta_1} C_0(X,\mathbb{Z}) \longrightarrow 0$$

calculates the homology of X. That is, if we define

$$H_d(C,\partial) = \frac{\mathrm{Ker}\left(\partial_d\right)}{\mathrm{Im}\left(\partial_{d+1}\right)}$$

then for each d

$$H_d(C,\partial) \cong H_d(X,\mathbb{Z})$$

where $H_d(X,\mathbb{Z})$ denotes the singular homology of X.

The precise definition of the boundary map ∂ is slightly nontrivial and so we do not discuss it here.. However, much can be learned from just knowing of the existence of such a boundary map:choose a coefficient field $\mathbb{F}$, and tensor everything with $\mathbb{F}$ to get a differential complex

$$0 \longrightarrow C_n(X, \mathbb{F}) \xrightarrow{\theta_n} C_{n-1}(X, \mathbb{F}) \xrightarrow{\theta_{n-1}} \cdots \xrightarrow{\theta_1} C_0(X, \mathbb{F}) \longrightarrow 0$$

which calculates $H_*(X, \mathbb{F})$, where now $C_d(X, \mathbb{F}) \cong \mathbb{F}^{c_d}$. This leads us to the following deep inequalities.

Theorem 2.1.4. *Let X be a CW complex with a fixed CW decomposition with c_d cells of dimension d for each d. Fix a coefficient field $\mathbb{F}$ and let b_* denote the Betti numbers of X with respect to $\mathbb{F}$, i.e., $b_d = \dim(H_d(X, \mathbb{F}))$.*

1. *(The Weak Morse Inequalities) For each d*

$$c_d \geq b_d$$

2. *Let $\chi(X)$ denote the Euler characteristic of X, i.e.,*

$$\chi(X) = b_0 - b_1 + b_2 - \ldots$$

then we also have

$$\chi(X) = c_0 - c_1 + c_2 - \ldots$$

Theorem 2.1.5. *(The Strong Morse Inequalities). With all notation as in Theorem 1.7, for each $d = 0, 1, 2, \ldots$*

$$c_d - c_{d-1} + c_{d-2} - \cdots + (-1)^d c_0 \geq b_d - b_{d-1} + b_{d-2} - \cdots + (-1)^d b_0$$

As mentioned earlier using CW complexes over simplciial complexes allows us to use fewer cells. Consider the case where X is a two-dimensional torus, so that with respect to any coefficient field $b_0 = 1, b_1 = 2, b_2 = 1$. From the weak Morse inequalities, we have that for any CW decomposition,

$$\begin{aligned} c_0 &\geq b_0 = 1 \\ c_1 &\geq b_1 = 2 \\ c_2 &\geq b_2 = 1 \end{aligned}$$

A simplicial decomposition is a special case of a CW decomposition, so these inequalities are satisfied when c_d denotes the number of d-simplices in a fixed simplicial decomposition. However, every simplicial decomposition has at least 7 0-simplices, 21 1-simplices and 14 2-simplices, so these inequalities are far from equality. It is generally the case that for a simplicial decomposition these inequalities are very far from optimal, and hence are generally of little interest. On the other hand, earlier we demonstrated a CW decomposition of the two-torus with exactly one 0 -cell, two 1-cells and one 2-cell. The inequalities tell us, in particular, that one cannot build a two-torus using fewer cells.

2.2 The Basics of Morse Theory

Suppose we have a finite simplicial complex X. From the previous discussion, we can expect that X has a CW decomposition with many fewer cells than in the original simplicial decomposition. How can one go about finding such an "efficient" CW decomposition for X ? We begin by recalling that a finite simplicial complex is a finite set of vertices V, along with a set of subsets K of V. The set K satisfies two main properties: 1) $V \subseteq K$ 2) If $\alpha \in K$ and $\beta \subseteq \alpha$ then $\beta \in K$.

(We will freely refer to the simplicial complex simply as K). The elements of K are called simplices. If $\alpha \in K$, and α contains $p+1$ vertices, then we say that the dimension of α is p, and we will sometimes denote this by $\alpha^{(p)}$. For simplices α and β we will use the notation $\alpha < \beta$ or $\beta > \alpha$ to indicate that α is a proper subset of β (thinking of α and β as subsets of V), and say that α is a face of β.

Definition 16. *A function*

$$f : K \longrightarrow \mathbb{R}$$

is a discrete Morse function if for every $\alpha^{(p)} \in K$

1. $\#\left\{\beta^{(p+1)} > \alpha \mid f(\beta) \leq f(\alpha)\right\} \leq 1$
2. $\#\left\{\gamma^{(p-1)} < \alpha \mid f(\gamma) \geq f(\alpha)\right\} \leq 1$

A straightforward example here will help us understand this definition. Consider the two complexes shown in Figure 2.2. Here we indicate functions by writing next to each simplex the value of the function on that simplex. The function (i) is not a discrete Morse function as the edge $f^{-1}(0)$ violates rule (2), since it has 2 lower dimensional "neighbors" on which f takes on higher values, and the vertex $f^{-1}(5)$ violates rule (1), since it has 2 higher dimensional "neighbors" on which f takes on lower values. The function (ii) is a Morse function. Note that a discrete Morse function is not a continuous function on K. Rather, it is an assignment of a single number to each simplex.

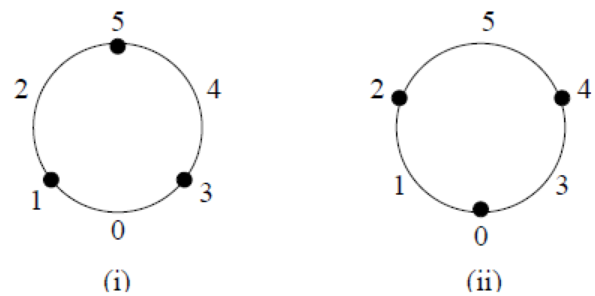


Figure 2.3: The first is not a discrete Morse function, the second one is

Definition 17. *A simplex* $\alpha^{(p)}$ *is critical if (1)* $\#\left\{\beta^{(p+1)} > \alpha \mid f(\beta) \leq f(\alpha)\right\} = 0$*, and (2)* $\#\left\{\gamma^{(p-1)} < \alpha \mid f(\gamma) \geq f(\alpha)\right\} = 0$*.*

Thus in our previous example the only critical simplices are, the vertex $f^{-1}(0)$ and the edge $f^{-1}(5)$ are critical, and there are no other critical simplices. Note here the following important lemma:

Lemma 19. *If K is a simplicial complex with a Morse function f, then for any simplex α, either*

1. $\#\left\{\beta^{(p+1)} > \alpha \mid f(\beta) \leq f(\alpha)\right\} = 0,$

   or

2. $\#\left\{\gamma^{(p-1)} < \alpha \mid f(\gamma) \geq f(\alpha)\right\} = 0$

Thus a simplex cannot simultaneously fail both tests for criticality.

Theorem 2.2.1. *Suppose K is a simplicial complex with a discrete Morse function. Then K is homotopy equivalent to a CW complex with exactly one cell of dimension p for each critical simplex of dimension p.*

Here we provide a sketch of this theorem. A discrete Morse function gives us a way to build the simplicial complex by attaching the simplices in the order prescribed by the function, i.e., adding first the simplices which are assigned the smallest values. More precisely, for any simplicial complex K with a discrete Morse function f, and any real number c, define the level subcomplex $K(c)$ by

$$K(c) = \cup_{f(\alpha)\leq c} \cup_{\beta\leq\alpha} \beta.$$

That is, $K(c)$ is the subcomplex consisting of all simplices α of K such that $f(\alpha) \leq c$ along with all of their faces. Thhen by using the following two lemmas the proof follows:

Lemma 20. *If there are no critical simplices α with $f(\alpha) \in (a, b]$, then $K(b)$ is homotopy equivalent to $K(a)$.*

Lemma 21. *If there is a single critical simplex α with $f(\alpha) \in (a, b]$ then there is a map $F : S^{(d-1)} \to K(a)$, where d is the dimension of α, such that $K(b)$ is homotopy equivalent to $K(a) \cup_F B^d$.*

Using our previous example we can see that these lemmas hold.

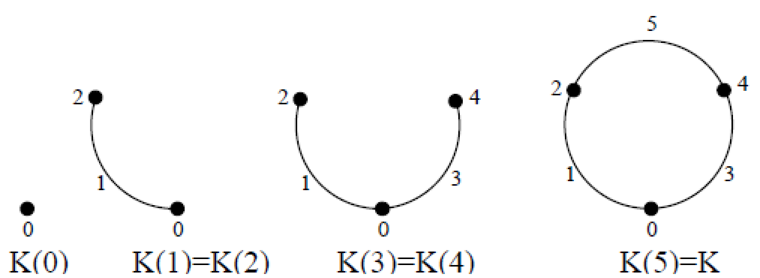


Figure 2.4: The level subcomplexes

Remark 4. For the first lemma, consider the transition from $K(0)$ to $K(1)$. We have not added any critical simplices, and, just as the lemma predicts,

$K(0)$ and $K(1)$ are homotopy equivalent. Let us try to understand why the homotopy type did not change. To construct $K(1)$ from $K(0)$, we first have to add the edge $f^{-1}(1)$. This edge is not critical because it has a codimension-one face which is assigned a higher value, namely the vertex $f^{-1}(2)$. In order to have $K(1)$ be a subcomplex, we must also add this vertex. Thus we see that the edge $f^{-1}(1)$ in $K(1)$ has a free face, i.e., a face which is not the face of any other simplex in $K(1)$. We can deformation retract $K(1)$ to $K(0)$ by "pushing in" the edge $f^{-1}(1)$ starting at the vertex $f^{-1}(2)$.

This is a very general phenomenon. That is, it follows from the axioms for a discrete Morse function that for any simplicial complex with any discrete Morse function, when passing from one level subcomplex to the next the noncritical simplices are added in pairs, each of which consists of a simplex and a free face. Suppose that $K_2 \subset K_1$ are simplicial complexes, and K_1 has exactly two simplices α and β that are not in K_2, where β is a free face of α. Then it is easy to see that K_2 is a deformation retract of K_1, and hence K_1 and K_2 are homotopy equivalent (see below). This special sort of combinatorial deformation retract is called a simplicial collapse. If one can transform a simplicial complex K_1 into a subcomplex K_2 by simplicial collapses, then we say that K_1 collapses to K_2, and we indicate this by $K_1 \searrow K_2$. the figure below shows a 2-dimensional simplex collapsing to one of its vertices.

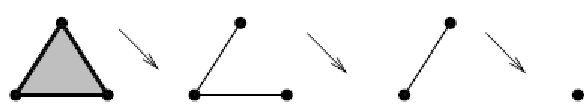


Figure 2.5: 2-simplex collapsing to a vertex

Remark 5. For the second lemma, let us see what happens when one adds a critical simplex, for example when making the transition from $K(4)$ to $K(5)$. In this case we are adding a critical edge. We can see clearly from the illustration that we pass from $K(4)$ to $K(5)$ by attaching a 1-cell. To see why this works in general, consider a critical d-simplex α. It follows from the definition of a critical simplex that each face of α is assigned a smaller value than α, which implies in turn that each face of α appears in a previous level subcomplex. Thus the entire boundary of α appears in an earlier level

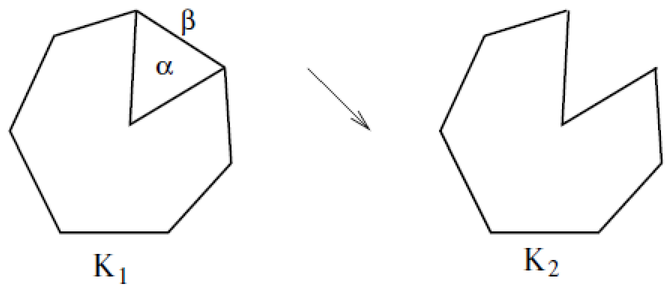


Figure 2.6: A general simplicial collapse

subcomplex, so that when it comes time to add α, we must "glue it in" along its entire boundary. This is precisely the process of attaching a d-cell.

Note here that we can always define a trivial discrete Morse function on any simplicial complex. Namely, one can simply let $f(\alpha) = \dim(\alpha)$ for each simplex α. In this case, every simplex is critical, and Theorem 2.2.1 is a tautology.

Furthermore, our new Theorem when combined with our earlier inequalities, using the fact that homotopy equivalent spaces have isomorphic homology, gives us a stronger version of the latter. Let K be a simplicial complex with a discrete Morse function. Let m_p denote the number of critical simplices of dimension p. Let $\mathbb{F}$ be any field, and $b_p = \dim H_p(K, \mathbb{F})$ the p^{th} Betti number with respect to $\mathbb{F}$.

Theorem 2.2.2 (The Weak Morse Inequalities).

- *For each* $p = 0, 1, 2, \ldots, n$ *(where* n *is the dimension of* K *)*

$$m_p \geq b_p$$

- $m_0 - m_1 + m_2 - \cdots + (-1)^n m_n = b_0 - b_1 + b_2 - \cdots + (-1)^n b_n \quad [= \chi(K)]$

Theorem 2.2.3 (The Strong Morse Inequalities).
For each $p = 0, 1, 2, \ldots, n, n+1$,

$$m_p - m_{p-1} + \cdots + (-1)^p m_0 \geq b_p - b_{p-1} + \cdots + (-1)^p b_0$$

2.3 Gradient Vector Fields

Let us now return to original example on the circle. Noncritical simplices occur in pairs. For example, the edge $f^{-1}(1)$ is not critical because it has a "lower dimensional neighbor" which is assigned a higher value, i.e., the vertex $f^{-1}(2)$. Similarly, the vertex $f^{-1}(2)$ is not critical because it has a "higher dimensional neighbor" which is assigned a lower value, i.e., the edge $f^{-1}(1)$. We indicate this pairing by drawing an arrow from the vertex $f^{-1}(2)$, pointing into the edge $f^{-1}(1)$. Similarly, we draw an arrow from the vertex $f^{-1}(4)$ pointing into the edge $f^{-1}(3)$. (see below) One can think of these arrows as pictorially indicating the simplicial collapse that is referred to in the proof of Lemma 20.

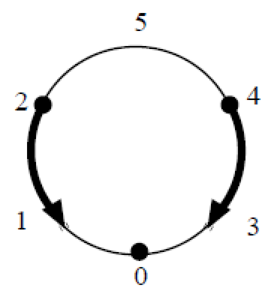


Figure 2.7: A gradient vector field for our previous example

This process can be applied to any simplicial complex with a discrete Morse function. The arrows are drawn as follows. Suppose $\alpha^{(p)}$ is a non-critical simplex with $\beta^{(p+1)} > \alpha$ satisfying $f(\beta) \leq f(\alpha)$. We then draw an arrow from α to β. Figure 3.2 illustrates a more complicated example. Note that the discrete Morse function drawn in this figure has one critical vertex, $f^{-1}(0)$, and one critical edge, $f^{-1}(11)$. Theorem 2.2.1 implies this simplicial complex is homotopy equivalent to a CW complex with exactly one 0 -cell and one 1-cell, i.e., a circle.

It follows from Lemma 19 that that every simplex α satisfies exactly one of the following:

1. α is the tail of exactly one arrow
2. α is the head of exactly one arrow

3. α is neither the head nor the tail of an arrow

Note that a simplex is critical if and only if it is neither the tail nor the head of any arrow. These arrows can be viewed as the discrete analogue of the gradient vector field of the Morse function.

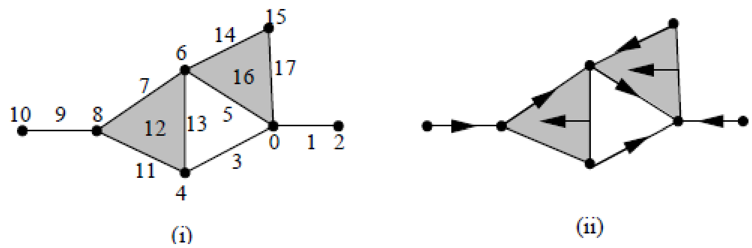


Figure 2.8: Another example fo a gradient vector field

As we will see in examples later, gradient vector fields contains all of the information that we will need to know about the original discrete Morse function for most applications. Thus, one "only" needs to find a gradient vector field. However, suppose we attach arrows to the simplices so that each simplex satisfies exactly one of properties (i),(ii),(iii) above. Then how do we know if that set of arrows is the gradient vector field of a discrete Morse function?

Let K be a simplicial complex with a discrete Morse function f. Then rather than thinking about the discrete gradient vector field V of f as a collection of arrows, we may equivalently describe V as a collection of pairs $\{\alpha^{(p)} < \beta^{(p+1)}\}$ of simplices of K, where $\{\alpha^{(p)} < \beta^{(p+1)}\}$ is in V if and only if $f(\beta) \leq f(\alpha)$. In other words, $\{\alpha^{(p)} < \beta^{(p+1)}\}$ is in V if and only if we have drawn an arrow that has α as its tail, and β as its head. The properties of a discrete Morse function imply that each simplex is in at most one pair of V.

Definition 18. *A discrete vector field V on K is a collection of pairs $\{\alpha^{(p)} < \beta^{(p+1)}\}$ of simplices of K such that each simplex is in at most one pair of V.*

Given a discrete vector field V on a simplicial complex K, a V-path is a sequence of simplices

$$\alpha_0^{(p)}, \beta_0^{(p+1)}, \alpha_1^{(p)}, \beta_1^{(p+1)}, \alpha_2^{(p)}, \dots, \beta_r^{(p+1)}, \alpha_{r+1}^{(p)}$$

such that for each $i = 0, \dots r, \{\alpha < \beta\} \in V$ and $\beta_i > \alpha_{i+1} \neq \alpha_i$. We say such a path is a *non-trivial closed path* if $r \geq 0$ and $\alpha_0 = \alpha_{r+1}$. If V is the gradient vector field of a discrete Morse function f, then we sometimes refer to a V-path as a *gradient path of* f.

Some powerful theorems follow form this definition:

Theorem 2.3.1. *Suppose V is the gradient vector field of a discrete Morse function f. Then a sequence of simplices is a V-path if and only if $\alpha_i < \beta_i > \alpha_{i+1}$ for each $i = 0, 1, \dots, r$, and*

$$f(\alpha_0) \geq f(\beta_0) > f(\alpha_1) \geq f(\beta_1) > \cdots \geq f(\beta_r) > f(\alpha_{r+1}).$$

The converse also holds.

Theorem 2.3.2. *A discrete vector field V is the gradient vector field of a discrete Morse function if and only if there are no non-trivial closed V-paths.*

The following theorem implies theorem 2.3.2, as we shall show later.

Theorem 2.3.3. *Let G be a directed graph. Then there is a real-valued function of the vertices that is strictly decreasing along each directed path if and only if there are no directed loops.*

2.4 Some Illustrative Examples

2.4.1 The Real Projective Plane

Our first example is the real projective plane $\mathbb{RP}^2$. The figure below shows a triangulation of the$\mathbb{RP}^2$. Note that the vertices along the boundary with the same labels are to be identified, as are the edges whose endpoints have the same labels. In the next figure we illustrate a discrete vector field V on this simplicial complex. One can easily see that there are no closed V-paths (since all V-paths go to the boundary of the figure and there are no closed V-paths on the boundary), and hence is a gradient vector field. The only simplices which are neither the head nor the tail of an arrow are the vertex labelled 1 , the edge e, and the triangle t. Thus, by Theorem 2.5 , the projective plane is homotopy equivalent to a CW complex with exactly one 0

-cell, one 1-cell and one 2-cell. (Of course, we already knew this from our discussion of Example 3 in Section 2.)

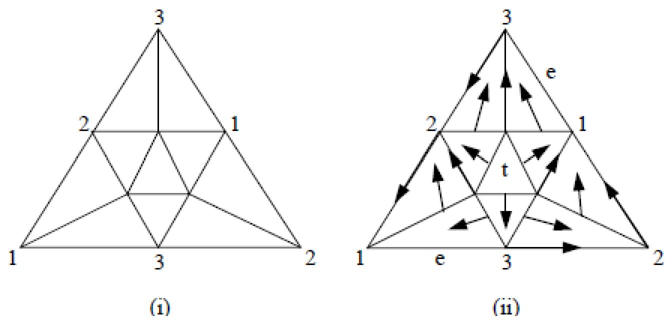


Figure 2.9: A triangulation of $\mathbb{RP}^2$ and a discrete gradient vector field on $\mathbb{RP}^2$

2.4.2 Sphere Theorems

The previous example gives rise to two potential concerns. The first is that from the main theorem we learn only a statement about "homotopy equivalence". This is sufficient if one is only interested in calculating homology or homotopy groups. However, one might be interested in determining the (PL-)homeomorphism type of the complex. This is possible, in some cases, using deep results of J.H.C. Whitehead. We shall remark on this later. The second potential point of concern is that as we saw earlier there are an infinite number of different homotopy types of CW complexes which can be built from exactly one 0 -cell, one 1 -cell and one 2-cell. One might wonder if Morse Theory can give us any additional information as to how the cells are attached. In fact, one can deduce much of this information if one has enough information about the gradient paths of the Morse function. Recall that a simplicial complex K is a combinatorial d-ball if K and the standard d-simplex Σ_d have isomorphic subdivisions. A simplicial complex K is a combinatorial $(d-1)$-sphere, if K and $\dot{\Sigma}_d$ have isomorphic subdivisions (where $\dot{\Sigma}_d$ denotes the boundary of Σ_d with its induced simplicial structure). A simplicial complex K is a combinatorial d-manifold with boundary if the link of every vertex is either a combinatorial $(d-1)$-sphere or a combinatorial $(d-1)$-ball.

Theorem 2.4.1. *Let K be a combinatorial d-manifold with boundary which simplicially collapses to a vertex. Then K is a combinatorial d-ball.*

It is with this theorem (and its generalizations) that one can strengthen the conclusion of Theorem 2.2.1 beyond homotopy equivalence. We present just one example.

Theorem 2.4.2. *Theorem 8.2. Let X be a combinatorial d-manifold with a discrete Morse function with exactly two critical simplices. Then X is a combinatorial d-sphere.*

The proof follows from the previous theorem. If X is a combinatorial d-manifold with a discrete Morse function f with exactly two critical simplices, then the critical simplices must be the minimum of f, which must occur at a vertex v, and the maximum of f, which must occur at a d-simplex α. Then $X - \alpha$ is a combinatorial d-manifold with boundary with a discrete Morse function with only a single critical simplex, namely the vertex v. It follows from Lemma 20 that $X - \alpha$ collapses to v Whitehead's theorem now implies that $X - \alpha$ is a combinatorial d-ball, which implies that X is a combinatorial d-sphere.

2.4.3 Complex of Not Connected Graphs

Our third example is the complex of not connected graphs. A number of fascinating simplicial complexes arise from the study of monotone graph properties. Let K_n denote the complete graph on n vertices, and suppose we have labelled the vertices $1, 2, \ldots, n$. Let $\mathcal{G}_n$ denote the spanning subgraphs of K_n, that is, the subgraphs of K_n that contain all n vertices. A subset $\mathcal{P} \subset \mathcal{G}_n$ is called a graph property of graphs with n vertices if inclusion in $\mathcal{P}$ only depends on the isomorphism type of the graph. That is, $\mathcal{P}$ is a graph property if for all pairs of graphs $G_1, G_2 \in \mathcal{G}_n$, if G_1 and G_2 are isomorphic (ignoring the labellings on the vertices) then $G_1 \in \mathcal{P}$ if and only if $G_2 \in \mathcal{P}$. A graph property $\mathcal{P}$ of graphs with n vertices is said to be monotone decreasing if for any graphs $G_1 \subset G_2 \in \mathcal{G}_n$, if $G_2 \in \mathcal{P}$ then $G_1 \in \mathcal{P}$.

Monotone decreasing properties abound in the study of graph theory. Here are some typical examples: graphs having no more than k edges (for any fixed k), graphs such that the degree of every vertex is less that δ (for any fixed δ), graphs which are not connected, graphs which are not i-connected (for any fixed i), graphs which do not have a Hamiltonian cycle, graphs which do not contain a minor isomorphic to H (for any fixed graph H), graphs which are r-colorable (for any fixed r), and bipartite graphs.

Any monotone decreasing graph property $\mathcal{P}$ gives rise to a simplicial complex $\mathcal{K}$ where the d-simplices of $\mathcal{K}$ are the graphs $G \in \mathcal{P}$ which have $d+1$ edges. In particular, if G is a d-simplex in $\mathcal{K}$, then the faces of G are all of the nontrivial spanning subgraphs of G (the monotonicity of $\mathcal{P}$ implies that each of these graphs is in $\mathcal{K}$). Said in another way, if $\mathcal{P}$ is nonempty, then the vertices of $\mathcal{K}$ are the edges of K_n, and a collection of vertices in $\mathcal{K}$ span a simplex if the spanning subgraph of K_n consisting of all edges which correspond to these vertices lies in $\mathcal{P}$.

For our purposes, we will try to use discrete Morse Theory to determine the topology of $\mathcal{N}_n$, the simplicial complex of not connected graphs on n vertices.Our goal is to construct a discrete gradient vector field V on $\mathcal{N}_n$, the simplicial complex of all not-connected graphs with the vertex set $\{1, 2, 3, \ldots, n\}$. The construction will be in steps. Let V_{12} denote the discrete vector field consisting of all pairs $\{G, G+(1,2)\}$, where G is any graph in $\mathcal{N}_n$ which does not contain the edge $(1,2)$ and such that $G+(1,2) \in \mathcal{N}_n$. Another way of describing V_{12} is that if G is any graph in $\mathcal{N}_n$ which contains the edge (1,2), then $G-(1,2)$ and G are paired in V_{12}. Actually, there is one exception to this rule. Let G^* denote the graph consisting of only the single edge $(1,2)$. Then $G^*-(1,2)$ is the empty graph, which corresponds to the empty simplex in $\mathcal{N}_n$, and may not be paired in a discrete vector field. Thus, G^* is unpaired in V_{12}.

The graphs in $\mathcal{N}_n$ other than G^* which are unpaired in V_{12} are those that do not contain the edge $(1,2)$ and have the property that $G+(1,2) \notin \mathcal{N}_n$. That is, those disconnected graphs G with the property that $G+(1,2)$ is connected. Such a graph must have exactly two connected components, one of which contains the vertex labelled 1, and one which contains the vertex labelled 2. We denote these connected components by G_1 and G_2, respectively as in the figure below.

Let G be a graph other than G^* which is unpaired in V_{12}, and consider vertex 3 . This vertex must either be in G_1 or G_2. Suppose that vertex 3 is in G_1. If G does not contain the edge $(1,3)$ then $G+(1,3)$ is also unpaired in V_{12}, so

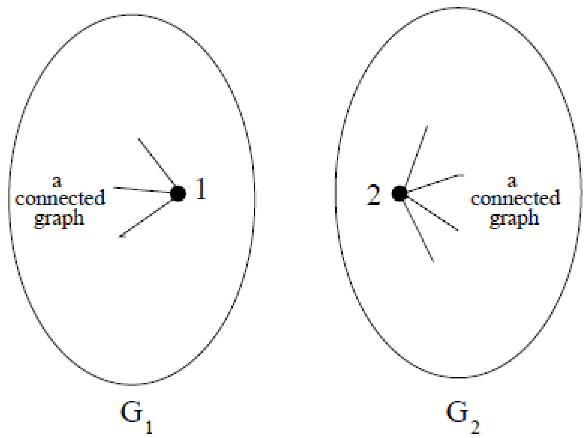


Figure 2.10: The graphs other than G^* which are unpaired in the vector field V_{12}

we can pair G with $G + (1, 3)$. If vertex 3 is in G_1, then the graph G is still unpaired if and only if G contains the edge $(1, 3)$ and $G - (1, 3)$ is the union of three connected components, one containing vertex 1 , one containing vertex 2 , and one containing vertex 3. Similarly, if vertex 3 is in G_2 and G does not contain the edge (2,3), then pair G with $G + (2, 3)$. Let V_3 denote the resulting discrete vector field.

The unpaired graphs in V_3 are G^* and those that either contain the edge $(1, 3)$ and have the property that $G - (1, 3)$ is the union of three connected components, one containing vertex 1 , one containing vertex 2 , and one containing vertex 3 , or contain the edge $(2, 3)$ and have the property that $G - (2, 3)$ is the union of three connected components, one containing vertex 1 , one containing vertex 2 , and one containing vertex 3. These can eb seen the figure below, in which the circles in this figure indicate connected graphs.

Now consider the location of the vertex labelled 4, and pair any graph G which is unpaired in V_3 with $G + (1, 4), G + (2, 4)$, or $G + (3, 4)$ if possible (at most one of these graphs is unpaired in V_3). Call the resulting discrete vector field V_4. We continue in this fashion, considering in turn the vertices labelled $5, 6, \ldots, n$. Let V_i denote the discrete vector field that has been constructed after the consideration of vertex i, and $V = V_n$ the final discrete

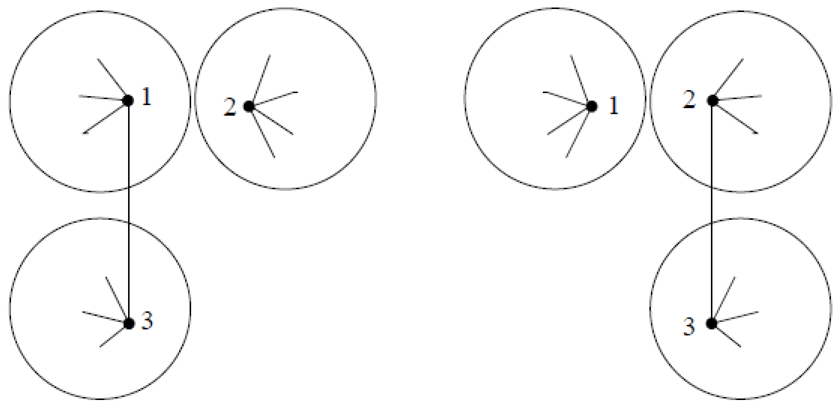


Figure 2.11: The graphs other than G^* which are unpaired in the vector field V_3

vector field. When we are done the only unpaired graphs in V will be G^* and those graphs that are the union of two connected trees, one containing the vertex 1 and one containing the vertex 2 . In addition, both trees have the property that the vertex labels are increasing along every ray starting from the vertex 1 or the vertex 2 . There are precisely ($n-1$)! such graphs, and they each have $n-2$ edges, and hence correspond to an $(n-3)$-simplex in $\mathcal{N}_n$.

It remains to see that the discrete vector field V is a gradient vector field, i.e., that there are no closed V-paths. We first check that V_{12} is a gradient vector field. Let $\gamma = \alpha_0^{(p)}, \beta_0^{(p+1)}, \alpha_1^{(p)}$ denote a V_{12}-path. Then α_0 must be the "tail of an arrow", i.e., the smaller graph of some pair in V_{12}, with β_0 being the head of the arrow, i.e., $\beta_0 = \alpha_0 + (1,2)$. The simplex α_1 is a codimension-one face of β_0 other than α_0. Thus, α_1 corresponds to a graph of the form $\alpha_0 + (1,2) - e$, where e is an edge of α_0 other than $(1,2)$. Since α_1 contains the edge $(1,2)$ it is the "head of an arrow" in V_{12}, i.e., the larger graph of some pair in V_{12}, which implies that γ cannot be continued to a longer V_{12}-path. This certainly implies that there are no closed V_{12}-paths.

A similar argument works for V. Recall that V is constructed in stages, by first considering the edge $(1,2)$ and then the vertices $3,4,5,\ldots$ in order. Let $\gamma = \alpha_0, \beta_0, \alpha_1$ denote a V-path. In particular, α_0 and β_0 must be paired in V. The reader can check that if α_0 and β_0 are first paired in $V_i, i \geq 3$, then either α_1 is the head of an arrow in V_i, in which case the V-path cannot be continued, or α_1 is paired in V_{i-1}. It follows by induction that there can be no closed V-paths. Thus we can finally say that V is a discrete gradient vector field on $\mathcal{N}_n$ with exactly one unpaired vertex, and $(n-1)$! unpaired $(n-3)$-simplices. Theorem 2.2.1 then allows us to conclude.

Definition 19. *The complex $\mathcal{N}_n$ of not connected graphs on n-vertices is homotopy equivalent to the wedge of $(n-1)$! spheres of dimension $(n-3)$.*

2.5 Some Combinatorial Aspects

Consider the Hasse diagram of K, that is, the partially ordered set of simplices of K ordered by the face relation. Consider the Hasse diagram as a directed graph. The vertices of the graph are in 1-1 correspondence with the simplices of K, and there is a directed edge from β to α if and only if α is a codimension-one face of β. Now let V be a combinatorial vector field. We modify the directed graph as follows. If $\{\alpha < \beta\} \in V$ then reverse the orientation of the edge between α and β, so that it now goes from α to β. A V-path can be thought of as a directed path in this modified graph. There are some directed paths in this modified Hasse diagram which are not V-paths as we have defined them. However, we can still obtain the following result

Theorem 2.5.1. *There are no nontrivial closed V-paths if and only if there are no nontrivial closed directed paths in the corresponding directed Hasse diagram.*

In the language of combinatorics a discrete vector field is a partial matching of the Hasse diagram, and a discrete vector field is a gradient vector field if the partial matching is acyclic in the above sense. We note here that this theorem allows us to conclude Theorem 3.3.2 from 3.3.3.

We can now restate some of our earlier theorems in this language. There is a very minor complication in that one usually includes the empty set as an element of the Hasse diagram (considered as a simplex of dimension -1) while we have not considered the empty set previously.

Theorem 2.5.2. *Let V be an acyclic partial matching of the Hasse diagram of K (of the sort described above - assume that the empty set is not paired*

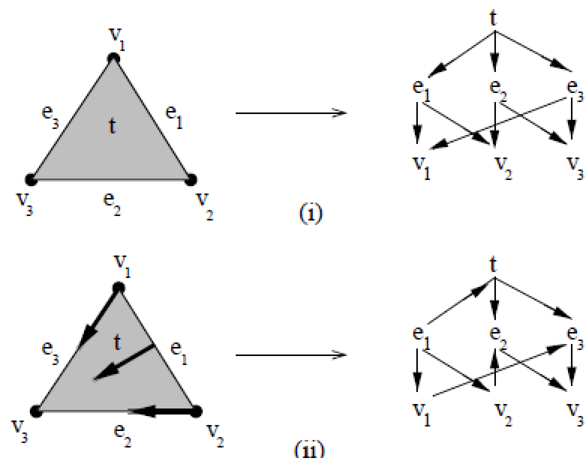


Figure 2.12: From a discrete vector field to a Directed Hesse diagram

with another simplex). Let u_p denote the number of unpaired p-simplices. Then K is homotopy equivalent to a CW-complex with exactly u_p cells of dimension p, for each $p \geq 0$.

If V is a complete matching, that is, every simplex (this time including the empty simplex) is paired with another simplex, we get the following result.

Theorem 2.5.3. *Let V be a complete acyclic matching of the Hasse diagram of K, then K collapses onto a vertex, so that, in particular, K is contractible.*

2.6 The Morse Complex

Let K be a simplicial complex with a Morse function f. Let $C_p(X, \mathbb{Z})$ denote the space of p-simplicial chains, and $\mathcal{M}_p \subseteq C_p(X, \mathbb{Z})$ the span of the critical p-simplices. We refer to $\mathcal{M}_*$ as the space of Morse chains. If we let m_p denote the number of critical p-simplices, then we obviously have

$$\mathcal{M}_p \cong \mathbb{Z}^{m_p}.$$

Since homotopy equivalent spaces have isomorphic homology, the following theorem (which is equivalent to the Strong Morse inequalities) follows from our previous results.

Theorem 2.6.1. *There are boundary maps $\tilde{\partial}_d : \mathcal{M}_p \to \mathcal{M}_{d-1}$, for each d, so that*

$$\tilde{\partial}_{d-1} \circ \tilde{\partial}_d = 0$$

and such that the resulting differential complex

$$0 \longrightarrow \mathcal{M}_n \xrightarrow{\bar{\theta}_n} \mathcal{M}_{n-1} \xrightarrow{\bar{\theta}_{n-1}} \cdots \xrightarrow{\bar{\theta}_1} \mathcal{M}_0 \longrightarrow 0$$

calculates the homology of X. That is, if we define

$$H_d(\mathcal{M}, \tilde{\partial}) = \frac{\operatorname{Ker}\left(\tilde{\partial}_d\right)}{\operatorname{Im}\left(\tilde{\partial}_{d+1}\right)}$$

then for each d

$$H_d(\mathcal{M}, \tilde{\partial}) \cong H_d(X, \mathbb{Z})$$

Let us now try and find an explicit formula for the boundary operator $\tilde{\partial}$. This requires a closer look at of the notion of a gradient path. Let α and $\tilde{\alpha}$ be p-simplices. Recall that a gradient path from $\tilde{\alpha}$ to α is a sequence of simplices

$$\tilde{\alpha} = \alpha_0^{(p)}, \beta_0^{(p+1)}, \alpha_1^{(p)}, \beta_1^{(p+1)}, \alpha_2^{(p)}, \ldots, \beta_r^{(p+1)}, \alpha_{r+1}^{(p)} = \alpha$$

such that $\alpha_i < \beta_i > \alpha_{i+1}$ for each $i = 0, 1, 2, \ldots, r$ and $f(\alpha_0) \geq f(\beta_0) > f(\alpha_1) \geq f(\beta_1) > \cdots \geq f(\beta_r) > f(\alpha_{r+1})$. Equivalently, if V is the gradient vector field of f, we require that for each i, α_i and β_i be paired in V and $\beta_i > \alpha_{i+1} \neq \alpha_i$. In Figure 7.2 we show a single gradient path from the boundary of a critical 2 -simplex β to a critical edge α, where the arrows indicate the gradient vector field V.

Given a gradient path as shown below, an orientation on β induces an orientation on α. We will not state the precise definition here. The idea is that one "slides" the orientation from β along the gradient path to α. For example, for the gradient path shown, the indicated orientation on β induces the indicated orientation on α.

Theorem 2.6.2. *Choose an orientation for each simplex. Then for any critical $(p+1)$ simplex β set*

$$\widetilde{\partial}\beta = \sum_{\text{critical } \alpha^{(p)}} c_{\alpha,\beta}\,\alpha$$

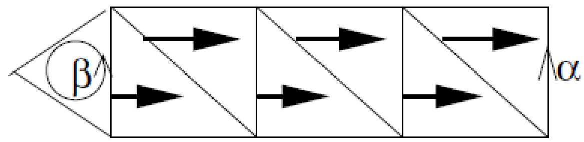


Figure 2.13: A gradient path from the boudnary of β to α

where

$$c_{\alpha,\beta} = \sum_{\gamma \in \Gamma(\beta,\alpha)} m(\gamma)$$

where $\Gamma(\beta, \alpha)$ *is the set of gradient paths which go from a maximal face of* β *to* α*. The multiplicity* $m(\gamma)$ *of any gradient path* γ *is equal to* ± 1*, depending on whether, given* γ*, the orientation on* β *induces the chosen orientation on* α*, or the opposite orientation. With this differential, the complex (7.1) computes the homology of* K*.*

Let us now try and use the Morse complex to compute the homology of $\mathbb{RP}^2$ using our earlier triangulation. We have already seen how discrete Morse Theory can help us see that $\mathbb{P}^2$ has a CW decomposition with exactly one 0 -cell, one 1-cell and one 2-cell. Here we will see how Morse Theory can distinguish between the spaces which have such a CW decomposition. In the figure below we redraw the gradient vector field, and indicate a chosen orientation on the critical edge e and the critical triangle t. Let us now calculate the boundary map in the Morse complex. To calculate $\tilde{\partial}(e)$, we must count all of the gradient paths from the boundary of e to v. There are precisely two such paths. Namely, following the unique gradient path beginning at each endpoint of e leads us to v. (The gradient path beginning at the head of e is the trivial path of 0 steps.) Since the orientation of e induces a+ orientation on the head of e, and a orientation on the tail of e, adding these two paths with their corresponding signs leads us to the formula that $\tilde{\partial}(e) = 0$. It can be seen from the illustration that there are precisely two gradient paths from the boundary of t to e, and, with the illustrated orientation for t, both induce the chosen orientation on e, so that $\tilde{\partial}(t) = 2e$. Therefore the homology of the real projective plane can be calculated from the following differential complex.

$$\mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \xrightarrow{0} \mathbb{Z} \longrightarrow 0.$$

Thus

$$H_0\left(\mathbb{P}^2, \mathbb{Z}\right) \cong \mathbb{Z}, \quad H_1\left(\mathbb{P}^2, \mathbb{Z}\right) \cong \mathbb{Z}/2\mathbb{Z}, \quad H_2\left(\mathbb{P}^2, \mathbb{Z}\right) \cong 0$$

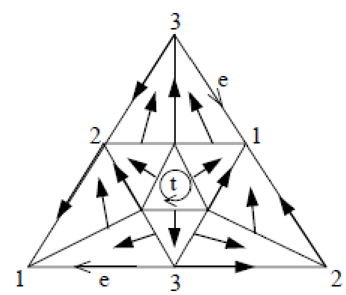


Figure 2.14: A gradient vector field on $\mathbb{RP}^2$

2.7 Cancelling Critical Points

One of the main problems in Morse Theory, whether in the combinatorial or smooth setting, is to find a Morse function for a given space with the fewest possible critical points. In general this is a very difficult problem, since, in particular, it contains the Poincaré conjecture - spheres can be recognized as those spaces which have a Morse function with precisely 2 critical points. In 1965, Milnor recast Smale's proof (1961) of the higher dimensional Poincaré conjecture (in fact, a proof is presented of the more general h-cobordism theorem) completely in the language of Morse Theory. Drastically oversimplifying matters, the proof of the higher Poincaré conjecture can be described as follows. Let M be a smooth manifold of dimension ≥ 5 which is homotopy equivalent to a sphere. Endow M with a (smooth) Morse function f. One then proceeds to show that the critical points of f can be cancelled out in pairs until one is left with a Morse function with exactly two critical points, which implies that M is a (topological) sphere. A key step in this proof is the "cancellation theorem" which provides a sufficient condition for two critical points to be cancelled. The analogous theorem also holds for discrete Morse functions. Moreover in the combinatorial setting the proof is much simpler.

Theorem 2.7.1. *Suppose f is a discrete Morse function on M such that $\beta^{(p+1)}$ and $\alpha^{(p)}$ are critical, and there is exactly one gradient path from the boundary β to α. Then there is another Morse function g on M with the* *same critical simplices except that α and β are no longer critical. Moreover, the gradient vector field associated to g is equal to the gradient vector field associated to f except along the unique gradient path from the boundary β to α.*

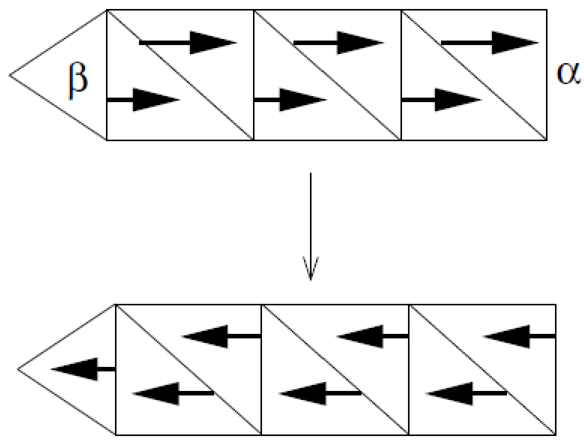


Figure 2.15: Cancelling critical points

We can sketch a proof of this. If, in the top drawing in the diagram above, the indicated gradient path is the only gradient path from the boundary of β to α, then we can reverse the gradient vector field along this path, replacing the figure by the vector field shown in the bottom. The uniqueness of the gradient path implies that the resulting discrete vector field has no closed orbits, and hence, is the gradient vector field of some Morse function. Moreover, α and β are not critical for this new Morse function, while the criticality of all other simplices is unchanged. This completes the proof.

The proof in the smooth case proceeds along the same lines. However, in addition to turning around those vectors along the unique gradient path from β to α, one must also adjust all nearby vectors so that the resulting vector field is smooth. Moreover, one must check that the new vector field is the gradient of a function, so that, in particular, modifying the vectors did not result in the creation of a closed orbit.

Chapter 3

3-manifolds and 0-efficient triangulations

3.1 Triangulations and Cell-Decompositions

We begin by defining a more general version of cell-decompositions, which some readers may have encountered as 'pseudo triangulations': in the usual triangulation, the tetrahedra are embedded and if two simplices meet at all, then they meet in a face of each, while for us cells are not necessarily embedded (though they are embedded on the interiors of each cell, face, and edge, and the intersection of two cells is a union of sub-cells of each).

Let $\mathbf{\Delta} = \{\tilde{\Delta}_1, \ldots, \tilde{\Delta}_t\}$ be a pairwise-disjoint collection of oriented, compact, convex, linear cells. Suppose Φ is a family of affine isomorphisms pairing faces of the cells in $\mathbf{\Delta}$ so that if $\phi \in \Phi$, then ϕ is an orientation-reversing affine isomorphism from a face $\sigma_i \in \tilde{\Delta}_i$ to a face $\sigma_j \in \tilde{\Delta}_j$, possibly $i = j$. We use $\mathbf{\Delta}/\Phi$ to denote the space obtained from the disjoint union of the $\tilde{\Delta}_i$ by setting $x \in \tilde{\sigma}_i$ equal to $\phi(x) \in \tilde{\sigma}_j$, with the identification topology. Then $\mathbf{\Delta}/\Phi$ is a 3-manifold, except possibly at the images of the vertices of the $\tilde{\Delta}_i$. (In a completely general setting, the identification space $\mathbf{\Delta}/\Phi$ may not be a 3-manifold at the image of the centers of some edges, as well as the images of the vertices; however, we have avoided this problem by orienting the $\tilde{\Delta}_i$ and choosing the affine isomorphisms $\phi \in \Phi$ orientation-reversing.) We collect all this information into a single symbol $\mathcal{T}$ and call $\mathcal{T}$ a *cell-decomposition* of $\mathbf{\Delta}/\Phi$; in this case, we also use just $|\mathcal{T}|$ to denote the space $\mathbf{\Delta}/\Phi$. A *cell* *(tetrahedron)*, *face*, *edge*, or *vertex* in this cell decomposition is, respectively, the image of a cell (tetrahedron), face, edge, or vertex from the collection $\mathbf{\Delta} = \{\tilde{\Delta}_1, \ldots, \tilde{\Delta}_t\}$. We will denote the image of the faces by $\mathcal{T}^{(2)}$, the image of the edges by $\mathcal{T}^{(1)}$, and the image of the vertices by $\mathcal{T}^{(0)}$. We call $\mathcal{T}^{(i)}$ the *i-skeleton of $\mathcal{T}$*; but, generally, we just refer to these as the faces, edges, or vertices of $\mathcal{T}$. We will denote the image of $\tilde{\Delta}_i$ by Δ_i and call $\tilde{\Delta}_i$ the *lift of* Δ_i. A cell is the quotient of a unique cell, and a face is the quotient of one or two faces; edges and vertices may be the quotient of a number of edges or vertices, respectively. While the cells are not necessarily embedded, the interior of each cell is embedded. We define the *order* or *valence of an edge* e of $\mathcal{T}$ to be the number of edges in the collection $\mathbf{\Delta} = \{\tilde{\Delta}_1, \ldots, \tilde{\Delta}_t\}$, which are identified to e. If the link (here we mean the boundary of a small regular neighborhood and not the combinatorial link) of each vertex is either a 2-sphere or a 2-cell, then the underlying point set is an oriented 3-manifold, possibly with boundary, and we say $\mathcal{T}$ is a *cell-decomposition* of the 3-manifold $M = |\mathcal{T}|$. If each cell in $\mathbf{\Delta}$ is a tetrahedron, then we say $\mathcal{T}$ is a *triangulation* of the 3-manifold $M = |\mathcal{T}|$. If the link of some vertex is a closed surface, distinct from the 2-sphere, we say $\mathcal{T}$ is an *ideal cell-decomposition* (or *ideal triangulation*) of the 3-manifold $M = |\mathcal{T}| \setminus |\mathcal{T}^{(0)}|$. In this case, the vertices of $\mathcal{T}$ are called *ideal vertices*, and the *index of an ideal vertex* is the genus of its linking surface.

We can define cell-decompositions of surfaces in a similar manner. If $\boldsymbol{\sigma} = \{\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n\}$ is a pairwise disjoint collection of compact, convex, planar polygons and Ψ is a family of linear isomorphisms pairing edges of the polygons in $\boldsymbol{\sigma}$ so that $\psi \in \Psi$, then ψ is a linear isomorphism of an edge e_i of $\tilde{\sigma}_i$ to an edge e_j of $\tilde{\sigma}_j$, possibly $i = j$. We use $\boldsymbol{\sigma}/\Psi$ to denote the space obtained from the disjoint union of the $\tilde{\sigma}_i$ by setting $x \in \tilde{e}_i$ equal to $\psi(x) \in \tilde{e}_j$, with the identification topology. We have that $\boldsymbol{\sigma}/\Psi$ is always a 2-manifold, possibly nonorientable. In this situation, we say we have a cell-decomposition of the 2-manifold $\boldsymbol{\sigma}/\Psi$. If each $\tilde{\sigma}_i$ is a triangle, we say we have a *triangulation* of the 2-manifold $\boldsymbol{\sigma}/\Psi$. Similar to the case for 3-manifolds, in a cell-decomposition of a 2-manifold, our cells are not embedded; however, the open cells are embedded.

A final key definition is the complexity of a manifold (in the sense proprosed by Matveev). In essence, it is the 'size' of a minimal triangulation. For our purposes, it is the minimum number of tetrahedra required for a triangulation of the manifold. A highly nontrivial theorem is that complexity is additive under the connected sum and boundary connected sum operations.

In our cell-decompositions, an edge can be a simple closed curve, an edge in a cell with endpoints (vertices) identified. Since we aim to investigate triangulatiosn with just one vertex, in our this case, every edge is a simple closed curve. Below, we give the possible configurations for triangles (2-cells). In Figure 3.1, parts (4) and (5), we have two edges identified to give faces which are *cones* (the latter is a *pinched cone*); in (6), we have a face which is a Möbius band; and in (7) and (8), we have all three edges identified, giving in (7) the classical *dunce hat* and in (8) a spine for $L(3,1)$.

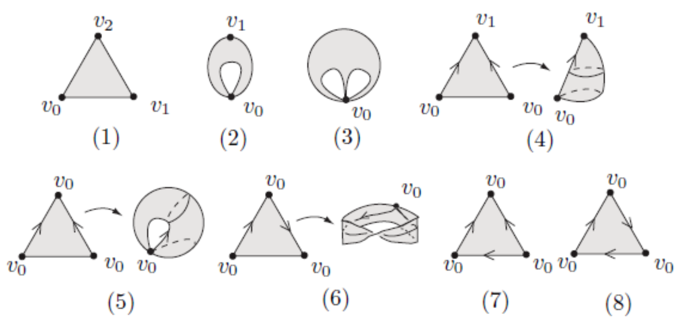


Figure 3.1: Possible configurations for a one-triangle triangulation

As for tetrahedra, we have in Figure 3.2 the seven distinct identifications of a single tetrahedron to give an orientable 3-manifold. Here, we wish to dwell on these diagrams. As we venture into the world of 3-manifolds, we may not have much of our usual visual intuition available to us to get a visual grasp of the objects we discuss. Thus, it becomes necessary to use a combination of topology and algebra to both help us visualise what these objects look like and to 'keep us honest'.

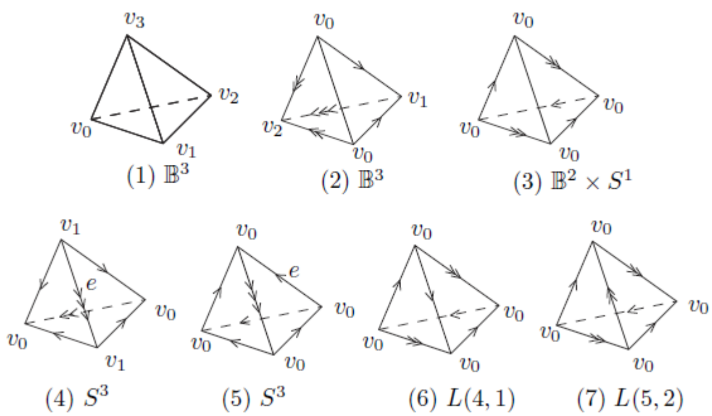
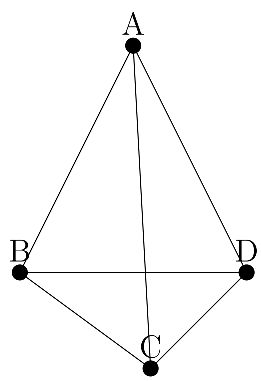


Figure 3.2: Possible configurations of one-tetrahedron triangulations

To help the reader acclimatise themselves with the 'dizzying heights' we are about to claim, we demonstrate here that the triangulations we have mentioned do indeed correspond to the 3-manifolds listed. (1) and even (2) can be quite trivially be seen to be the manifold $\mathbb{B}^3$, so we start with (3): the solid torus $\mathbb{B}^2 \times S^1$. There are many ways to visualise this, but for our purposes it helps to compute its homology. For the purposes of this section, we label the triangle in the following manner: the topmost vertex is A, the leftmost vertex is B, the bottom vertex is C, and the rightmost vertex is D.



Now, the given triangulation has 1 vertex, 3 edges, 3 faces, 1 tetrahedron, and no higher dimensional cells. We use the following rotated version of the triangulation: ABC $\sim$ CAD, with AB $\sim$ CA, BC $\sim$ AD, CA $\sim$ DA, and BA $\sim$ AC $\sim$ CD. Thus, we get the following chain complex:

- $C_4 = 0$

- $C_3 = \mathbb{Z}$
- $C_2 = \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$
- $C_1 = \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$
- $C_0 = \mathbb{Z}$

with

$$C_4 = 0 \xrightarrow{\partial_4} C_3 = \mathbb{Z} \xrightarrow{\partial_3} C_2 = \mathbb{Z}^3 \xrightarrow{\partial_2} C_1 = \mathbb{Z}^3 \xrightarrow{\partial_1} C_0 = \mathbb{Z} \xrightarrow{\partial_0} 0$$

Now, ∂_4 is the trivial map, so $Im(\partial_4) = 0$. Meanwhile, using the definition of the boundary map, $\partial_3[ABCD] = [BCD] - [ACD] + [ABD] - [ABC] = [BCD] + [ABD] + [CAD] - [ABC]$. By our identification, $[CAD] - [ABC] = 0$, so $\partial_3[ABCD] = [BCD] + [ABD]$ and $ker(\partial_3) = 0$. Thus

$$H_3 = \frac{ker(\partial_3)}{Im(\partial_4)} = ker(\partial_3) = 0$$

Next, we see that $\partial_2[ABC] = \partial_2[CAD] = [BC] - [AC] + [AB]; \partial_2[ABD] = [BD] - [AD] + [AB], \partial_3[BCD] = [CD] - [BD] + [BC]$. Thus, $Im(\partial_2(p, q, r)) = \langle (p - q + r)[BC], (2p - r + q)[AB], (q - r)[BD] \rangle$ after our identifications. Now $\partial_2(p, q, r) = 0 \implies q = r, p = 0$, which means that $ker(\partial_2) = \langle q[ABC], r[BCD] \rangle$. Thus

$$H_2 = \frac{ker(\partial_2)}{Im(\partial_3)} = \frac{\langle q[ABC], r[BCD] \rangle}{\langle p[ABD], r[BCD] \rangle} = \frac{\mathbb{Z}}{\mathbb{Z}} = 0$$

Since all the vertices are identified, $\partial_1[AB] = \partial_1[BC] = \partial_1[BD] = 0$, and so $Im(\partial_1) = 0, ker(\partial_1) = \mathbb{Z}^3$. Thus

$$H_1 = \frac{ker(\partial_1)}{Im(\partial_2)} = \frac{\langle p, q, r \rangle}{\langle p - q + r, 2p - r + q, q - r \rangle} = \frac{\langle p, q, r \rangle}{\langle p, 2p - r + q, q - r \rangle} = \frac{\langle p, q, r \rangle}{\langle p, -r + q, q - r \rangle}$$

$$= \frac{\langle p, q, r \rangle}{\langle p, 0, q - r \rangle} = \frac{\langle p, q - r, r \rangle}{\langle p, q - r, 0 \rangle} = \mathbb{Z}$$

Finally,

$$H_0 = \frac{ker(\partial_0)}{Im(\partial_1)} = ker(\partial_0) = \mathbb{Z}$$

The homology groups obtained are

$$H_i = \begin{cases} \mathbb{Z} & i = 0, 1 \\ 0 & i \geq 2 \end{cases}$$

which are indeed the homology groups for the solid torus. Essentially, what we are doing ehre is the following: we take two faces of a tetrahedron and glue them together in a orientation reversing manner by rotating an angle $\frac{2\pi}{3}$ clockwise or counterclockwise (indeed, if we glue the remaining 2 boudnary faces to each other we get a Lens space).

Since (4) and (5) are similar, it suffices to compute the homology for (4). The given triangulation has 2 vertices, 3 edges, 2 faces, 1 tetrahedron, and no higher dimensional cells. We use the identifications: ABC $\sim$ ADC, ABD $\sim$ CBD, AB $\sim$ AD, BC $\sim$ DC, CA $\sim$ CA, BA $\sim$ BC, DA $\sim$ DC, BD $\sim$ BD and A $\sim$ C, B $\sim$ D. Thus, we get the following chain complex:

- $C_4 = 0$
- $C_3 = \mathbb{Z}$
- $C_2 = \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$
- $C_1 = \mathbb{Z} \times \mathbb{Z}$
- $C_0 = \mathbb{Z} \times \mathbb{Z}$

with

$$C_4 = 0 \xrightarrow{\partial_4} C_3 = \mathbb{Z} \xrightarrow{\partial_3} C_2 = \mathbb{Z}^2 \xrightarrow{\partial_2} C_1 = \mathbb{Z}^3 \xrightarrow{\partial_1} C_0 = \mathbb{Z}^2 \xrightarrow{\partial_0} 0$$

Now, ∂_4 is the trivial map, so $Im(\partial_4) = 0$. Meanwhile, using the definition of the boundary map, $\partial_3[ABCD] = [BCD] - [ACD] + [ABD] - [ABC] = [BCD] - [BCD] + [ABC] - [ABC] = 0$ by our identifications. So $\partial_3[ABCD] = 0$ and $ker(\partial_3) = \mathbb{Z}$. Thus

$$H_3 = \frac{ker(\partial_3)}{Im(\partial_4)} = ker(\partial_3) = \mathbb{Z}$$

Next, we see that $\partial_2[ABC] = [BC] - [AC] + [AB]; \partial_2[ABD] = [BD] - [AD] + [AB]$. Thus, $Im(\partial_2(p, q, r)) = \langle -p[AC], q[BD] \rangle$ after our identifications.

Now $\partial_2(p,q,r) = 0 \implies p = 0, q = 0$, which means that $ker(\partial_2) = 0$. Thus

$$H_2 = \frac{ker(\partial_2)}{Im(\partial_3)} = ker(\partial_2) = 0$$

In this case we have 2 vertices, so $\partial_1[AB] = v_0 - v_1, \partial_1[BD] = \partial_1[AC] = 0$, and so $Im(\partial_1) = \langle a \rangle, ker(\partial_1) = \mathbb{Z}^2$. Thus

$$H_1 = \frac{ker(\partial_1)}{Im(\partial_2)} = \frac{\langle p, q \rangle}{\langle -p, q \rangle} = \frac{\langle p, q \rangle}{\langle p, q \rangle} = 0$$

Finally,

$$H_0 = \frac{ker(\partial_0)}{Im(\partial_1)} = \frac{\langle v_0, v_1 \rangle}{\langle v_0 - v_1 \rangle} = \frac{\langle v_0 - v_1, v_1 \rangle}{\langle v_0 - v_1 \rangle} = \langle v_1 \rangle = \mathbb{Z}$$

The homology groups obtained are

$$H_i = \begin{cases} \mathbb{Z} & i = 0, 3 \\ 0 & \text{else} \end{cases}$$

which are indeed the homology groups for S^3.

Our penultimate triangulation is that of the Lens space $L(4,1)$. The given triangulation has 1 vertex, 2 edges, 2 faces, 1 tetrahedron, and no higher dimensional cells. We use the identifications: ABC $\sim$ CAD, ABD $\sim$ BDC, $AB \sim CA \sim DC \sim BD$ and A $\sim C \sim B \sim$ D. Thus, we get the following chain complex:

- $C_4 = 0$
- $C_3 = \mathbb{Z}$
- $C_2 = \mathbb{Z} \times \mathbb{Z}$
- $C_1 = \mathbb{Z} \times \mathbb{Z}$
- $C_0 = \mathbb{Z}$

with

$$C_4 = 0 \xrightarrow{\partial_4} C_3 = \mathbb{Z} \xrightarrow{\partial_3} C_2 = \mathbb{Z}^2 \xrightarrow{\partial_2} C_1 = \mathbb{Z}^2 \xrightarrow{\partial_1} C_0 = \mathbb{Z} \xrightarrow{\partial_0} 0$$

Now, ∂_4 is the trivial map, so $Im(\partial_4) = 0$. Using the definition of the boundary map, $\partial_3[ABCD] = [BCD]-[ACD]+[ABD]-[ABC] = -[BDC]+[ABD]+[CAD]-[ABC] = 0$ by our identifications. So $\partial_3[ABCD] = 0$ and $ker(\partial_3) = \mathbb{Z}$. Thus

$$H_3 = \frac{ker(\partial_3)}{Im(\partial_4)} = ker(\partial_3) = \mathbb{Z}$$

Next, we see that $\partial_2[ABC] = [BC]-[AC]+[AB]; \partial_2[ABD] = [BD]-[AD]+[AB]$. Thus, $Im(\partial_2(p,q)) = \langle p-q, 2p+2q\rangle$ after our identifications. Now $\partial_2(p,q,r) = 0 \implies p = 0, q = 0$, which means that $ker(\partial_2) = 0$. Thus

$$H_2 = \frac{ker(\partial_2)}{Im(\partial_3)} = ker(\partial_2) = 0$$

Since we have only 1 vertex, $\partial_1[AB] = \partial_1[BC] = \partial_1[BD] = 0$, and so $Im(\partial_1) = 0, ker(\partial_1) = \mathbb{Z}^2$, and so $Im(\partial_1) = 0, ker(\partial_1) = \mathbb{Z}^2$. Thus

$$H_1 = \frac{ker(\partial_1)}{Im(\partial_2)} = \frac{\langle p,q\rangle}{\langle p-q, 2p+2q\rangle} = \frac{\langle p,q\rangle}{\langle p-q, 4p\rangle} = \frac{\langle p, p-q\rangle}{\langle p-q, 4p\rangle} = \frac{\langle p\rangle}{\langle 4p\rangle} = \mathbb{Z}_4$$

Finally,

$$H_0 = \frac{ker(\partial_0)}{Im(\partial_1)} = kernel(\partial_0) = \mathbb{Z}$$

The homology groups obtained are

$$H_i = \begin{cases} \mathbb{Z} & i = 0,3 \\ 0 & i = 2 \\ \mathbb{Z}_4 & i = 1 \end{cases}$$

which are indeed the homology groups for $L(4,1)$.

Finally, our last triangulation is that of the Lens space $L(5,2)$. The given triangulation has 1 vertex, 2 edges, 2 faces, 1 tetrahedron, and no higher dimensional cells. We use the identifications: ABC $\sim$ DCA, ABD $\sim$ BDC, $AB \sim DC \sim BD$, $AC \sim CB \sim DA$ and A $\sim C \sim B \sim$ D. Thus, we get the following chain complex:

- $C_4 = 0$
- $C_3 = \mathbb{Z}$

- $C_2 = \mathbb{Z} \times \mathbb{Z}$
- $C_1 = \mathbb{Z} \times \mathbb{Z}$
- $C_0 = \mathbb{Z}$

with

$$C_4 = 0 \xrightarrow{\partial_4} C_3 = \mathbb{Z} \xrightarrow{\partial_3} C_2 = \mathbb{Z}^2 \xrightarrow{\partial_2} C_1 = \mathbb{Z}^2 \xrightarrow{\partial_1} C_0 = \mathbb{Z} \xrightarrow{\partial_0} 0$$

Now, ∂_4 is the trivial map, so $Im(\partial_4) = 0$. Using the definition of the boundary map, $\partial_3[ABCD] = [BCD]-[ACD]+[ABD]-[ABC] = -[BDC]+[DCA]+[ABD]-[ABC] = 0$ by our identifications. So $\partial_3[ABCD] = 0$ and $ker(\partial_3) = \mathbb{Z}$. Thus

$$H_3 = \frac{ker(\partial_3)}{Im(\partial_4)} = ker(\partial_3) = \mathbb{Z}$$

Next, we see that $\partial_2[ABC] = [BC]-[AC]+[AB]; \partial_2[ABD] = [BD]-[AD]+[AB]$. Thus, $Im(\partial_2(p,q)) = \langle p+2q, -2p+q\rangle$ after our identifications. Now $\partial_2(p,q,r) = 0 \implies p = 0, q = 0$, which means that $ker(\partial_2) = 0$. Thus

$$H_2 = \frac{ker(\partial_2)}{Im(\partial_3)} = ker(\partial_2) = 0$$

Since we again have only 1 vertex, $Im(\partial_1) = 0, ker(\partial_1) = \mathbb{Z}^2$, and so $Im(\partial_1) = 0, ker(\partial_1) = \mathbb{Z}^2$. Thus

$$H_1 = \frac{ker(\partial_1)}{Im(\partial_2)} = \frac{\langle p,q\rangle}{\langle p+2q, -2p+q\rangle} = \frac{\langle p,q\rangle}{\langle p+2q, 5q\rangle} = \frac{\langle p, p+2qq\rangle}{\langle p+2q, 5q\rangle} = \frac{\langle q\rangle}{\langle 5q\rangle} = \mathbb{Z}_5$$

Finally,

$$H_0 = \frac{ker(\partial_0)}{Im(\partial_1)} = kernel(\partial_0) = \mathbb{Z}$$

The homology groups obtained are

$$H_i = \begin{cases} \mathbb{Z} & i = 0,3 \\ 0 & i = 2 \\ \mathbb{Z}_5 & i = 1 \end{cases}$$

which are indeed the homology groups for $L(5,2)$.

3.2 Normal surfaces

If $\tilde{\Delta}$ is a compact, convex, linear cell and $\tilde{\sigma}$ is a face of $\tilde{\Delta}$, we say a spanning arc in $\tilde{\sigma}$ is a *normal arc* if its endpoints are in distinct edges of $\tilde{\sigma}$. A *normal curve* in the boundary of a compact, convex, linear cell is a curve which meets each face in a collection of normal arcs. The elementary components of normal surface theory are the normal disks in the cells of the cell-decomposition. We call a properly embedded disk in a compact, convex, linear cell a *normal disk* if its boundary is a normal curve and it meets no edge more than once. If $\mathcal{T}$ is a cell-decomposition of the manifold M, then an isotopy of M is called a *normal isotopy* if it is invariant on the cells, faces, edges, and vertices of $\mathcal{T}$. Up to normal isotopy, there are only finitely many equivalence classes of normal disks in a compact, convex, linear cell; these are called *normal disk types.*

Suppose $\mathcal{T}$ is a cell-decomposition or ideal cell-decomposition of the 3-manifold M and S is a properly embedded surface transverse to the 2-skeleton of $\mathcal{T}$. Suppose c is a component of S in the cell Δ_i. Then c is the image of a properly embedded surface, $\tilde{c}$ in $\tilde{\Delta}_i$. We will call $\tilde{c}$ the *lift of c.*

Now, if $\mathcal{T}$ is a cell-decomposition of the 3-manifold M, we say a surface F is a *normal surface* in M (with respect to $\mathcal{T}$) if F meets each cell of $\mathcal{T}$ in the images of a collection of normal disks in the cells of $\mathbf{\Delta} = \{\tilde{\Delta}_1, \ldots, \tilde{\Delta}_n\}$. That is, the surface F is a normal surface if and only if the lift of every component of F in a cell of $\mathbf{\Delta}$ is a normal disk. The elementary components of normal surface theory for triangulations are the *normal triangles* and *normal quadrilaterals* (*normal quads*) in a tetrahedron. There are four types of normal triangles and three types of normal quads in each tetrahedron (no identification). The normal disk types give a normal surface F a cell-decomposition made up of normal quads and normal triangles; we call this the *cell-decomposition induced on F* (or the induced cell-decomposition). Finally, we note that an embedded normal surface must be properly embedded.

If S is a properly embedded surface in a 3–manifold M and $N(S)$ is a small regular neighborhood of S, the manifold $M' = M \backslash \mathring{N}(S)$ is said to be obtained from M by *splitting along S*. If S is one-sided in M, then S is nonorientable and there is a copy, say S', of the orientable double cover of S in $\partial M'$. If S is two-sided, then there are two homeomorphic copies, S' and S'', of S in $\partial M'$. They are in the same component of M' if and only if S does not

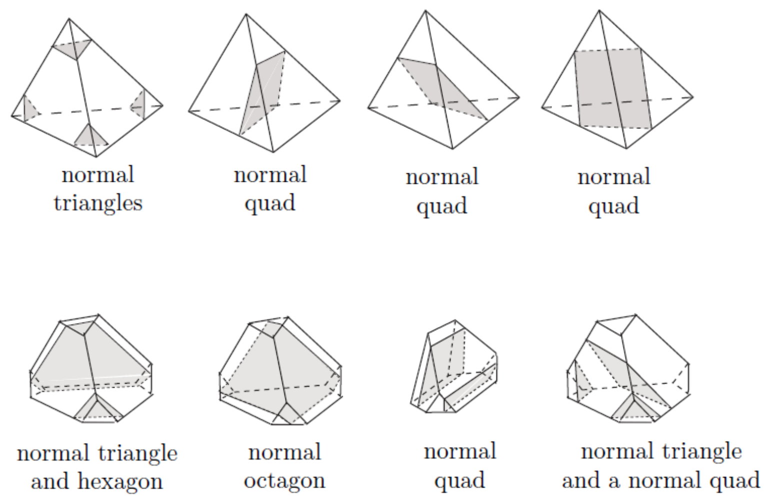


Figure 3.3: 7 types of normal disks

separate M. If S is a normal surface, then we choose $N(S)$ in such a way that the components of its frontier, S' and S'', are normally isotopic to S (or if S is one-sided, then $S' = 2S$). If S is a two-sided normal surface, M' is the manifold obtained by splitting M along S, and S' and S'' are the copies of S in $\partial M'$, then there is a natural identification of S' and S'' to recover the 3–manifold M with S' and S'' being identified to S in M. We will refer to this as *re-attaching* along S' and S''. In addition, if S^* is a normal surface in the cell-decomposition induced on M' (M split along S) and possibly now S^* meets $S' \cup S''$, then when we re-attach along S' and S'', we get a subcomplex, denoted $S \cup S^*$, which is the image of $S' \cup S'' \cup S^*$ in M. We will call this subcomplex the *piecewise linear normal surface* obtained from S and S^*. Finally, we say a 2–sphere S embedded in a 3–manifold M is *inessential* in M if S bounds a 3–cell in M; otherwise, S is *essential*. A properly embedded disk D in a 3–manifold M is *inessential* in M if ∂D bounds a disk D' in ∂M; otherwise, D is *essential* in M. We can now state some basic existence results:

Theorem 3.2.1. *Let M be a 3–manifold. If there is an essential, properly embedded disk in M, then for any cell-decomposition T of M, there is an essential, normal disk embedded in M.*

Theorem 3.2.2. *Let M be a 3–manifold. If there is an essential, embedded*

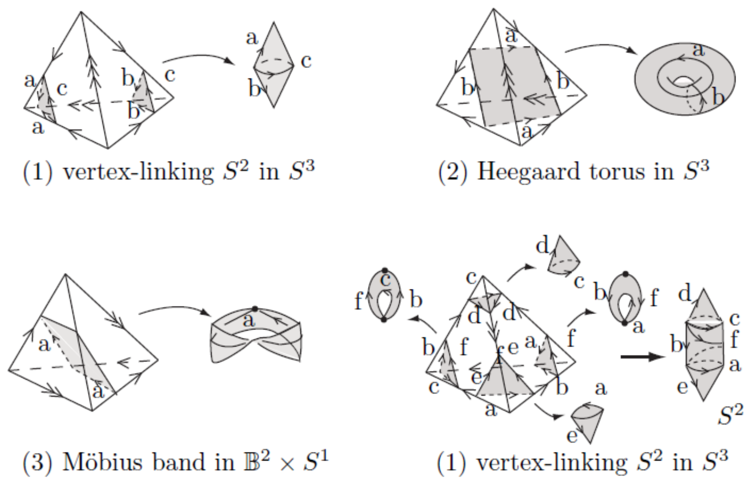


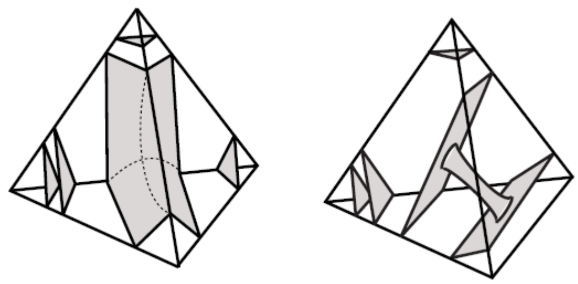
Figure 3.4: Examples of normal disk types

2–sphere in M, then for any cell-decomposition T of M, there is an essential, embedded, normal 2–sphere in M.

Theorem 3.2.3 (Kneser's Finiteness Theorem)**.** *Suppose T is a triangulation of the compact 3–manifold M. There is a nonnegative integer N_0 so that whenever $F_1, \ldots, F_n$ is a pairwise disjoint collection of normal surfaces in M and $n \geq N_0$, then for some $i \neq j$, $F_i = F_j$.*

Remark 6. The last equality, $F_i = F_j$ can mean that F_i is normally isotopic to F_j or, equivalently, they have the same parameterization. Furthermore, if M is closed, then one has, for example, $N_0 \leq 5t$, where t is the number of tetrahedra in T.

We also require the notion of almost normal surfaces in triangulations. A *normal octagon* is a properly embedded disk in a tetrahedron having boundary consisting of eight normal arcs in the boundary of the tetrahedron; whereas, a *normal tube* is a properly embedded annulus in a tetrahedron formed from two disjoint normal triangles, two disjoint normal quads, or a normal triangle and a disjoint normal quad by joining them via a tube parallel to an edge of the tetrahedron. A normal octagon and a normal tube are shown below. There are three types of normal octagons and twenty-five types of normal tubes in each tetrahedron.



(A) normal octagon (B) normal tube

Figure 3.5: A normal octagon and a normal tube between two normal quads

If T is a triangulation or ideal triangulation of the 3–manifold M, we say a surface F is *almost normal* (with respect to T) if and only if the lift of every component of F in a tetrahedron is either a normal triangle or a normal quadrilateral, except for at most one tetrahedron, where we allow the lift of F in the exceptional tetrahedron to be precisely one of:

- A collection of normal triangles and one normal octagon, or
- A collection of normal triangles and normal quads with one normal tube.

In the case of a normal tube, we do not allow that the normal tube is along an edge between two copies of the same normal surface. So, an almost normal surface never contains both a normal octagon and a normal tube but may contain one of them. An almost normal surface with a normal octagon is called an *octagonal almost normal surface*, and one with a normal tube is called a *tubed almost normal surface*. A compression of the tube gives a normal surface and does not give two copies of the same normal surface.

Finally, we have that if M has an essential 2-sphere, then for any triangulation of M, there is a vertex solution which is an essential, embedded, normal 2–sphere. Similarly, if M has an essential, properly embedded disk, then for any triangulation of M, there is a vertex solution which is an essential, normal disk. Under certain assumptions, one can also conclude that if there are almost normal 2–spheres, then there is one that is a vertex-solution.

3.3 Barrier surfaces

Suppose $\mathcal{T}$ is a cell-decomposition (or ideal cell-decomposition) of the 3-manifold M and S is a properly embedded surface in M, which is transverse to the 2-skeleton of $\mathcal{T}$, $\mathcal{T}^{(2)}$. The weight of S is the cardinality of $S \cap \mathcal{T}^{(1)}$, $\mathrm{wt}(S) = |S \cap \mathcal{T}^{(1)}|$. Now, the surface S is a normal surface if and only if the lift of every component of the intersection of S with a cell is a normal disk (i.e., a normal triangle or a normal quad). So, for normal surface theory, we would like the lifts of the components of S in the various cells to be disks, in general, and normal disks, in particular. If the lifts of all the components of S in the cells are not disks, we need a measure of this variance. We define the local Euler number of S, written $\lambda_\chi(S)$, to be the sum $\lambda_\chi(S) = \sum_{c \neq S^2}(1 - \chi(\bar{c}))$, where c runs over all non-2-sphere components of S in the cells of $\mathcal{T}$. Notice that $\lambda_\chi(S) = 0$ if and only if each non-spherical component of S in a cell of the decomposition $\mathcal{T}$ lifts to a disk. A 0-weight curve of intersection of S with $\mathcal{T}^{(2)}$ is a simple closed curve lying entirely in the interior of a face of $\mathcal{T}$. Let $\sigma(S)$ denote the number of 0-weight curves of the intersection of S with faces of $\mathcal{T}$, which are also in $\mathring{M}$, the interior of the 3-manifold M. We define the complexity of S to be $C(S) = (\mathrm{wt}(S), \sigma(S), \lambda_\chi(S))$, where we consider the set of triples under lexicographical order from the left.

Our goal here is to use four basic moves in shrinking a properly embedded surface S: a compression, an isotopy, a ∂-compression, and finally a "cleaning up" move. We begin with a properly embedded surface S meeting the 2-skeleton of the cell-decomposition transversely. To keep notation simple, we refer to the surface at each step of the shrinking as S, understanding that it may have changed considerably from the original surface S. The target of shrinking is to arrive at a surface (a stable surface) having components which are normal surfaces or are properly embedded, 0-weight 2-spheres and 0-weight disks, each of which is contained entirely in some cell of our cell-decomposition. Hence, the lifts of the components of S in a cell will be normal disks and properly embedded, 0-weight 2-spheres and 0-weight disks in the cell. Recall that normal disks are characterized by their boundary curves in the cells, which are made up of a finite number of normal spanning arcs in the faces of the cells and which do not meet an edge in the cell more than once.

The normal moves are:

1. A compression in the interior of a cell. This move reduces the local Euler number and does not change weight or the number of 0-weight curves of intersection of S with the faces of $\mathcal{T}$.

2. An isotopy reducing the number of times the boundary of a lift of a component meets an edge of a cell, where the edge is in the interior of M. This move reduces $wt(S)$.

3. A ∂-compression reducing the number of times a lift of a component meets an edge of a cell, where the edge is in the boundary of M. This move reduces $wt(S)$.

4. A compression eliminating 0-weight simple closed curve components from the intersection of S with the faces of $\mathcal{T}$ in the interior of M.

1. <u>A compression in the interior of a cell</u>
A compression reducing the local Euler number can be made in the interior of some cell whenever the local Euler number is not zero, $\lambda_\chi(S) \neq 0$. In this case, there is a component c of the intersection of S with some cell, say Δ_i, and for $\bar{c}$ the lift of c, we have $1 - \chi(\bar{c}) > 0$; hence, there is a compression of $\bar{c}$ along an essential curve in $\bar{c}$ in $\bar{\Delta}_i$. It follows, there is a disk $\bar{D}$ embedded in $\bar{\Delta}_i$ so that $\bar{D} \cap \bar{c} = \partial\bar{D}$ and $\partial\bar{D}$ is not trivial in $\bar{c}$. Of course, it is possible that $\bar{D}$ meets other lifts of the components of S in Δ_i. However, if this is the case, then we may assume the intersection of the lifts of the components of S in Δ_i meet $\bar{D}$ in simple closed curves in the interior of $\bar{D}$. Either we can change our choice of $\bar{D}$ to eliminate such intersections or there is a lift $\bar{c}'$ of a component c' of S meeting Δ_i and a disk $\bar{D}' \subset \bar{D}$ so that $\partial\bar{D}'$ is an essential curve in $\bar{c}'$ (in particular, $1 - \chi(\bar{c}') > 0$) and $\bar{D}'$ does not meet any other lifts of the components of S in Δ_i. Let's assume that D is such an innermost disk so we don't have to drag the prime notation along. We let D denote the image of $\bar{D}$ in Δ_i and compress c along D (which induces a compression of $\bar{c}$ along $\bar{D}$). Notice such a compression does not affect the weight but since $\partial\bar{D}$ is essential in $\bar{c}$, the compression decreases the local Euler number; furthermore, this move does not affect the intersection of S with the interior of the faces or the edges of $\mathcal{T}$. Hence, it reduces the complexity of the surface S. Note that this move may be an essential compression of the surface S and thereby a change of its topological type.

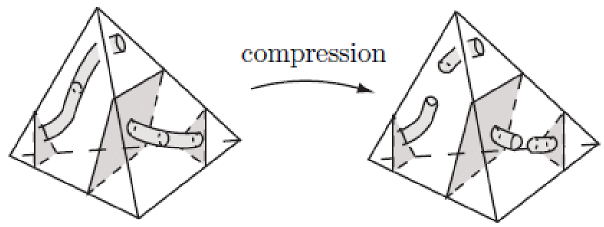


Figure 3.6: A compression in the interior

2. <u>An isotopy reducing the number of times the boundary of a lift of a component meets an edge of a cell, where the edge is in the interior of M.</u>
At this stage we may assume all the lifts of the components of S in a cell of $\mathcal{T}$ are either properly embedded disks or 2-spheres. Now, suppose $\bar{c}$ is a lift of a component of S in a cell Δ_i; $\bar{c}$ is a properly embedded disk; and $\bar{c}$ meets an edge e of $\bar{\Delta}_i$, the lift of Δ_i, more than once. Then if we consider the curve $\partial\bar{c}$, it divides the edge e into a number of subarcs and there is at least one, say $\bar{\beta}$, which has both its end points in $\partial\bar{c}$. Hence, there is a disk $\bar{D}$ embedded in $\bar{\Delta}_i$ so that $\bar{D}\cap\bar{c}$ is an arc $\bar{\alpha}\subset\partial\bar{D}$, $\bar{\alpha}\cup\bar{\beta}=\partial\bar{D}$, and $\bar{\alpha}\cap\bar{\beta}=\partial\bar{\alpha}=\partial\bar{\beta}$. See Figure 7. However, it is possible that $\bar{D}$ meets other lifts in $\bar{\Delta}_i$ of the components of S in Δ_i. If this is the case, then $\bar{D}$ meets lifts other than $\bar{c}$ in simple closed curves in the interior of $\bar{D}$ and spanning arcs in $\bar{D}$ having both their end points in $\bar{\beta}$. Standard techniques allow us to choose $\bar{D}$ so that there are no such simple closed curve components. Hence, if there are spanning arcs remaining, we can choose one $\bar{\alpha}'$, which is "outermost" in the sense that there is a disk $\bar{D}'\subset\bar{D}$ and a lift $\bar{c}'$ of a component c' of S in Δ_i so that $\bar{D}'\cap\bar{c}'=\bar{\alpha}'$, $\partial\bar{D}'=\bar{\alpha}'\cup\bar{\beta}'$, where $\bar{\beta}'\subset\bar{\beta}$ and $\bar{\alpha}'\cap\bar{\beta}'=\partial\bar{\alpha}'=\partial\bar{\beta}'$. Furthermore, $\bar{D}'$ does not meet any other lifts of components of S in Δ_i. As above, having demonstrated that we can find such an outermost disk, we assume the original disk $\bar{D}$ has this property so we do not need to drag along the prime notation. We consider the image D of $\bar{D}$ in Δ_i. Then D is an embedded disk in Δ_i; D only meets S in c; $D\cap c=\alpha\subset\partial D$ is a spanning arc of c; D meets the boundary of Δ_i in the arc β in the edge e of Δ_i (e is also used for the image of the edge e in $\bar{\Delta}_i$); and $\beta\subset\partial D$, where $\alpha\cap\beta=\partial\alpha=\partial\beta$ and $\alpha\cup\beta=\partial D$. of the 3-manifold, then there is an isotopy of S, splitting c into two disks and reducing $wt(S)$. Of course, this move may increase the

value $\sigma(S)$ and the local Euler number; however, it reduces the complexity of S.

3. A θ-compression reducing the number of times a lift of a component meets an edge of a cell, where the edge is in the boundary of M

Instead of, as above, where the edge e is in the interior of M, we now have the edge e in ∂M. In this case, the move must be accomplished by a θ-compression rather than an isotopy. This move reduces $wt(S)$; however, as above in the isotopy move, it may increase the value $\sigma(S)$ and the local Euler number. In any case, it reduces the complexity of S. Note that as above when we had a compression, a θ-compression can change the topological type of S.

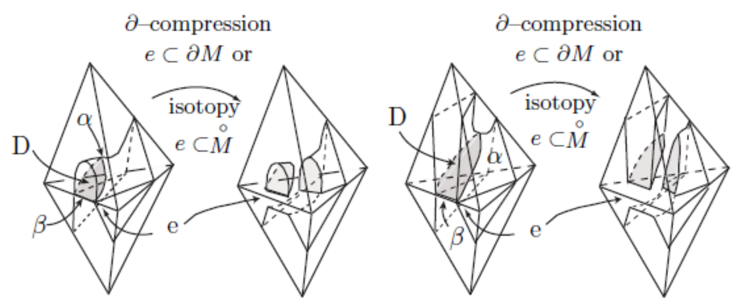


Figure 3.7: An isotopy or ∂-compression

4. A compression eliminating 0-weight simple closed curve components from the intersection of S with the faces of $\mathcal{T}$ in the interior of M

We may assume there are no essential compressions in the interior of any cell; actually, at this stage, we may assume the lift of any component of S in a cell is either a normal disk or a properly embedded, 0-weight 2-sphere or 0-weight disk in the cell. If there is a 0-weight simple closed curve common to S and a face σ of a cell, say Δ_i, then there is an innermost one. Such a 0-weight curve bounds a component c of S in Δ_i, which is a 0-weight disk, properly embedded in Δ_i, and having its boundary entirely in the interior of the face σ. Furthermore, its boundary bounds a disk D in the interior of σ. If σ is in the interior of the manifold, then there is a similarly embedded disk c' on the other side of this face, $\partial D = \partial c'$ and $c \cup c'$ is a small 2-sphere, which can be isotoped entirely into the interior of one of the cells or we can perform a compression along the disk D in σ and create two 0-weight 2-spheres, one in each cell sharing the face, σ. We choose to do the latter and therefore get two 2-spheres, each embedded entirely in the interior of a cell; this is one of our stable situations. Thus we eliminate all 0-weight simple closed curves in the interior of faces of $\mathcal{T}$, which are also in the interior of M. These moves reduce the value $\sigma(S)$ and do not affect the weight or the local Euler number. Note that we could just throw away all of the 0-weight pieces, which in practice is essentially what we do; but what we have done here reduces our work later when we need to analyze what we have after we shrink a surface.

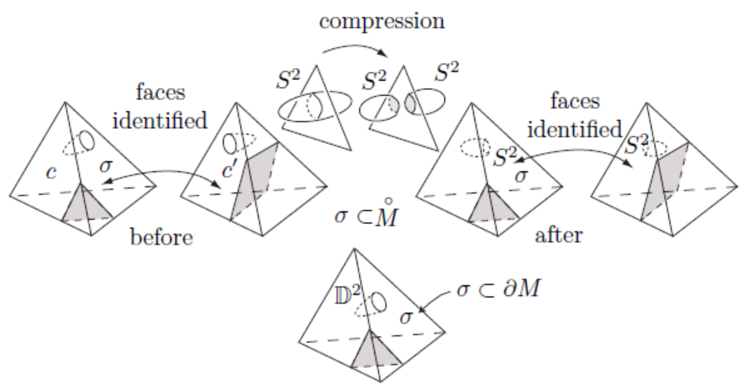


Figure 3.8: A compression removing a simple closed curve component

Remark 7. Compressions and θ-compressions can alter the properties of the surface S. We can place conditions on the surface and the 3-manifold so we can make these modifications and still maintain certain properties of the surface; for example, as indicated above, the surface is incompressible and θ-incompressible and the manifold is irreducible and θ-irreducible. However, these normal moves can be made on any surface. The moves never increase weight. We call a sequence of these normal moves a shrinking of the surface S and we allow the surface S to be compressed and θ-compressed, possibly resulting in a number of distinct components. After a finite number of steps, the components will either be normal or will be properly embedded, 0-weight 2-spheres and 0-weight disks contained entirely in various cells of our cell-decomposition. We say S has been shrunk to the resulting surfaces or sometimes we throw away the 0-weight components and say S has been shrunk to the resulting normal surfaces. Of course, it may be that S has shrunk until it disappears (has only 0-weight components).

Definition 20. *Suppose B is a properly embedded surface in a 3-manifold M, and let N be a component of the complement of B, $M \setminus B$. We say B is a barrier surface for N, or simply a barrier, if any properly embedded, compact surface F in N can be shrunk in N.*

We have the following criteria for a surface B, properly embedded in M, to be a barrier surface for a component N of its complement in M. Suppose Δ is a cell of $\mathcal{T}$, and C is the closure of a component of $\Delta \cap N$, $\bar{\Delta}$ the lift of Δ, and $\bar{C}$ the lift of C. Let $b = C \cap B$ and $\bar{b}$ denote the lift of b. A collection of pairwise disjoint disks in $\bar{C}$ is said to be a complete system of compressing disks for B in C if:

1. A disk, $\bar{D}$, in the collection meets $\bar{b}$ only in its boundary, which is an essential curve in $\bar{b}$; i.e., $\bar{D}$ is an essential compressing disk for $\bar{b}$ in $\bar{C}$.

2. A disk, $\bar{D}$, in the collection meets $\bar{b}$ in a properly embedded arc $\bar{\alpha}$ and meets the boundary of $\bar{\Delta}$ in an arc $\bar{\beta}$, which is entirely in the interior of an edge of $\bar{\Delta}$, $\bar{\alpha} \cup \bar{\beta} = \partial \bar{D}$, and $\bar{\alpha} \cap \bar{\beta} = \partial \bar{\alpha} = \partial \bar{\beta}$.

3. Each component remaining after $\bar{b}$ has been compressed and θ-compressed along the collection of disks is either a normal disk for $\mathcal{T}$ or a properly embedded 0-weight disk $\bar{E}$ having its boundary entirely in the interior of a face σ of $\bar{\Delta}$, $\partial \bar{E}$ bounds a disk $\bar{E}' \subset \sigma$, and the 2-sphere $\bar{E} \cup \bar{E}'$ bounds a 3-cell in $\bar{C}$.

Lemma 22. *Suppose $\mathcal{T}$ is a cell-decomposition (or ideal cell-decomposition) of the 3-manifold M, and B is a properly embedded surface in M. The surface B is a barrier surface for the component N of $M \setminus B$ if, for each cell Δ of $\mathcal{T}$ and the closure of each component C of $\Delta \cap N$, there is a complete system of compressing disks for B in C.*

Proof. Suppose F is a properly embedded surface in N. The surface F misses B (since N is a component of $M \setminus B$), and we may assume that F is transverse to $\mathcal{T}^{(2)}$, the 2-skeleton of $\mathcal{T}$. We choose a properly embedded surface S that has minimal complexity among all surfaces that can be obtained by shrinking F while missing B; i.e., sequences of compressions and θ-compressions,

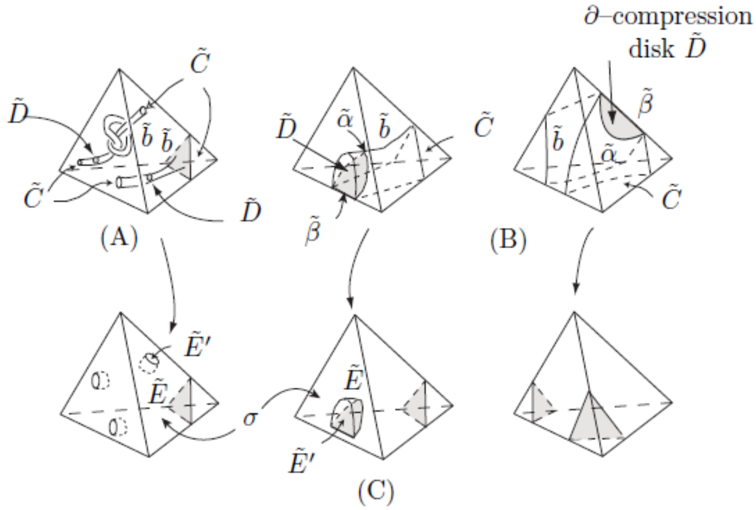


Figure 3.9: A system of compressing disks

missing B, and isotopies that are the identity on B.

We claim each component of S is either normal or is a 0-weight 2-sphere or disk properly embedded in a cell of $\mathcal{T}$. To see this, suppose Δ is a cell of $\mathcal{T}$. Let C be a component of $N \cap \Delta$. Let $\bar{C}$ be the lift of C; let $b = B \cap C$ and let $\bar{b}$ be the lift of b. By hypothesis, there is a complete system of disks for B in C. Suppose $c = S \cap C$ and $\bar{c}$ is the lift of c. If $\bar{c}$ meets the complete system of compressing disks for B in C, it does so in a very specific way. Specifically, if $\bar{D}$ is a compressing disk in $\bar{C}$ for $\bar{b}$, any component of $\bar{c} \cap \bar{D}$ is a simple closed curve in the interior of $\bar{D}$. If $\bar{D}$ is a θ-compressing disk in $\bar{C}$ for $\bar{b}$, any component of $\bar{c} \cap \bar{D}$ is either a simple closed curve in the interior of $\bar{D}$ or a spanning arc; furthermore, the spanning arcs have both endpoints in the same edge of $\bar{\Delta}$. It follows that none of these intersections of $\bar{c}$ with the complete system of compressing disks are essential in $\bar{c}$, or we could make a sequence of moves, as defined above, to reduce the complexity of S. Hence, we can assume that $\bar{c}$ misses a complete system of compressing disks for B in C. This is true in every cell of $\mathcal{T}$. Notice, a priori, it may seem we have to go back to a cell in which we have made these moves before. But this could only happen if we reduce the weight of S, which would contradict our choice of S.

Now, for any cell Δ and any component C of $N \cap \Delta$, if c is a component of S meeting C and $\bar{c}$ is the lift of c, we can make all compressions and θ-compressions along a complete system of disks for B in C, missing $\bar{c}$. Hence, $\bar{c}$ lies in a component of $\bar{\Delta}$ determined by normal disks and properly embedded 0-weight disks parallel into the interior of a face of $\bar{\Delta}$ through a 3-cell in $\bar{C}$. It follows that there is no obstruction to making further moves on $\bar{c}$ to reduce complexity unless each component of $\bar{c}$ is a normal disk or a properly embedded, 0-weight 2-sphere or 0-weight disk lying entirely in a cell of $\mathcal{T}$. $\square$

Some notation is in order here: if S is a two-sided, normal surface in M and M' is the manifold obtained from M by splitting along S, then we let S' and S'' denote the copies of S in $\partial M'$. Furthermore, if S^* is a normal surface in the induced cell structure on M', we let $S \cup S^*$ denote the piecewise normal surface obtained from S and S^*. Similarly, if $\mathcal{K}$ is a subcomplex of the cell-decomposition induced on M' by $\mathcal{T}$, then when we re-attach along S' and S'', we get a subpolyhedron, denoted $S \cup |\mathcal{K}|$, which is the image of $S' \cup S'' \cup |\mathcal{K}|$ in M. We now list some barrier surfaces that can be demonstrated easily:

1. If S is a normal surface or an almost normal surface in M and B is the boundary of a small regular neighborhood of S in M, then B is a barrier surface for each component of its complement not meeting S. Often in this case, we just say the normal surface or almost normal surface S is a barrier surface for each component of its complement.

2. If S is a two-sided, normal surface in M and S^* is a normal surface in the induced cell structure on M split along S and B is the boundary of a small regular neighborhood (in M) of the piecewise normal surface $S \cup S^*$, then B is a barrier surface for each component of its complement not meeting $S \cup S^*$.

3. If X is a finite union of normal surfaces in M, which meet transversely, and B is the boundary of a small regular neighborhood of X, then B is a barrier surface for each component of its complement not meeting X.

4. If $\mathcal{K}$ is a subcomplex of $\mathcal{T}$ and B is the boundary of a small regular neighborhood of the underlying point set of $\mathcal{K}$, $|\mathcal{K}|$, then B is a barrier surface for any component of its complement not meeting $|\mathcal{K}|$.

5. If S is a normal surface in M and $\mathcal{K}$ is a subcomplex of the cell structure induced by $\mathcal{T}$ on M split along S and B is the boundary of a small regular neighborhood of $S \cup |\mathcal{K}|$, then B is a barrier surface for any component of its complement not meeting $S \cup |\mathcal{K}|$.

6. If S is a two-sided, normal surface in M and $\mathcal{K}$ is a subcomplex of the cell structure induced by $\mathcal{T}$ on M', the manifold obtained by splitting M along S, and F is the frontier of a small regular neighborhood of $|\mathcal{K}|$, then F is a barrier surface in M' for any component of its complement not meeting $|\mathcal{K}|$.

We can now prove a couple of useful propositions regarding barrier surfaces that demonstrate their usefulness in shrinking. Suppose $B_1, \ldots, B_n$ is a pairwise disjoint collection of 3-cells in S^3. We call the 3-manifold $M = S^3 \setminus \bigcup_{i=1}^n \overset{\circ}{B}_i$ a punctured 3-sphere. In particular, the collection may be empty; so, we allow that the 3-sphere itself is a punctured 3-sphere, of course, without any punctures. In the case we know the boundary is not empty, we may also say we have a punctured 3-cell. In our definition of a barrier surface B, we have that whenever B is a barrier surface for a component of its complement, then in all cases B is two-sided. So, if B is a barrier surface for the component N of its complement, then by taking a small regular neighborhood of B, we have a copy of B in N. We can then shrink this copy, and call this procedure "shrinking B in N". Also, if the barrier surface B is a normal surface or has components which are normal surfaces, then we may say shrink B, understanding that each component of B which is normal is stable and there are no normalization moves on such components. Such components survive in the shrinking.

Proposition 4. *Suppose $\mathcal{T}$ is a triangulation of the closed, orientable, irreducible 3-manifold M, and suppose S is a normal 2-sphere in M which bounds a 3-cell E in M. Then there is a normal 2-sphere S' bounding a 3-cell E' in M, which contains E and all the vertices of $\mathcal{T}$ or $M = S^3$.*

Proof. If all the vertices of $\mathcal{T}$ are in E, then there is nothing to prove. Otherwise, split M along S and let M' denote the component not meeting E. We will continue to use S to denote the copy of S in the boundary of M'. There is a subcomplex Λ of the 1-skeleton of the induced cell structure on M' so that each component of Λ is a tree and meets S in precisely one point, and

Λ contains all vertices of $\mathcal{T}$ not in E. By the above theorem, the frontier B of a small regular neighborhood $N = N(S \cup \Lambda)$ of $S \cup \Lambda$ is a barrier surface in the component of its complement in M' not meeting $S \cup \Lambda$. Note that N is a punctured 3-cell in M' (actually, B is isotopic to S and in this case $N = S^2 \times I$), which contains all the vertices of $\mathcal{T}$ not in E.

If B is normal, then it is itself stable and does not shrink. In this case, we let $E' = E \cup N$ and let $S' = B$. So, we may assume B is not normal. We can shrink B in M'; furthermore, the point of B being a barrier is that this shrinking will not meet $S \cup \Lambda$. A shrinking (a finite sequence of normal moves) involves either a compression or an isotopy move (M is closed). So, assume we are at a stage in our shrinking where we have a finite number of pairwise disjoint 2-spheres $S_1, \ldots, S_n$, and a punctured 3-cell P_k with $S \cup \Lambda \subset P_k$ and ∂P_k includes S and the spheres $S_1, \ldots, S_n$. If the collection $S_1, \ldots, S_n$ is not stable, then there is either an isotopy normal move or a compression normal move on one of these 2-spheres.

An isotopy move is across an edge in N missing $S \cup \Lambda$; hence, we have an isotopy move of some 2-sphere, say S_i, in ∂P_k. We get a new collection of 2-spheres $S_1, \ldots, S_i', \ldots, S_n$ where S_i' replaces S_i and $wt(S_i') < wt(S_i)$. We let P_{k+1} denote the image of P_k under this isotopy. Then P_{k+1} is a punctured 3-cell containing $S \cup \Lambda$.

If there is a compression on one of the 2-spheres, say S_i, then let D denote the compressing disk. Not only does D not meet any 2-sphere in the collection except for S_i, which it meets in its boundary, D does not meet $S \cup \Lambda$. If $D \subset P_k$, then D splits P_k into two punctured 3-cells, one, say P_{k+1} containing $S \cup \Lambda$, and ∂D splits S_i into two 2-spheres, S_i' and S_i'', with, say $S_i' \subset P_{k+1}$. We have a new collection of 2-spheres, $S_{j_1}, \ldots, S_i', \ldots, S_{j_m}$, which along with S make up the boundary of our new punctured 3-cell P_{k+1}. If D is not in P_k, then again ∂D splits S_i into two 2-spheres, S_i' and S_i''; however, in this case, a compression is adding a 2-handle to S_i, and we get a new punctured 3-cell, P_{k+1}, containing P_k and having both S_i' and S_i'' in its boundary. Also, $S \cup \Lambda \subset P_{k+1}$.

It follows that in shrinking B and in the stable situation, we have a punctured 3-cell P, $S \cup \Lambda \subset P$, and the boundary of P consists of S along with possibly some other normal 2-spheres and possibly some 0-weight 2-spheres entirely in the interior of cells in the induced cell structure on M'. Each 0-weight 2-sphere bounds a 3-cell whose interior misses P. We fill in these 2-spheres with these 3-cells and continue to call our punctured 3-cell P. Now, since M is irreducible, each normal 2-sphere in the boundary of P bounds a 3-cell in M. If such a boundary component, other than S, bounds a 3-cell whose interior misses P, we add that 3-cell to P. We will continue to call the punctured 3-cell P. So, we now have that $S \cup \Lambda \subset P$, and any 2-sphere in the boundary of P other than S does not bound a 3-cell whose interior misses P. If S is the only boundary component of P, then M is S^3. If S' is a component of the boundary of P distinct from S, then by M irreducible, S' bounds a 3-cell, say E', in M. But then we have $E \cup P \subset E'$. So, such an S' and E' satisfy the conclusions of our proposition.

□

Remark 8. There is a useful variation to the previous proposition when we do not assume the 3-manifold M is irreducible; namely, we have either there is a collection $S_1, \ldots, S_n$ of normal 2-spheres bounding a punctured 3-cell P in M, where P contains E and all the vertices of $\mathcal{T}$, or M is S^3. We also can use the previous proposition by, say, choosing S to be a vertex-linking normal 2-sphere and E the 3-cell it bounds to conclude that for any triangulation $\mathcal{T}$ of a closed, orientable, irreducible 3-manifold M, there is a normal 2-sphere bounding a 3-cell containing all the vertices of $\mathcal{T}$ or it follows that M is S^3. There are triangulations of S^3 for which there is no normal 2-sphere bounding a 3-cell containing all the vertices of the triangulation.

Next, suppose F is a closed, orientable surface. Let $F \times [0,1]$ be the product of F with the unit interval, and let $\gamma_1, \ldots, \gamma_n$ be a finite, pairwise disjoint collection of simple closed curves in $F \times 0$; it is not necessary that the γ_i be essential. Choose small regular neighborhoods $N(\gamma_1), \ldots, N(\gamma_n)$ of the γ_i, $1 \leq i \leq n$, in $F \times 0$ so that $N(\gamma_i) \cap N(\gamma_j) = \emptyset$ for $i \neq j$. Let $D_1 \times [0,1], \ldots, D_n \times [0,1]$ be a collection of 2-handles, where D_i, $1 \leq i \leq n$, is a 2-cell. A 3-manifold is obtained by attaching the 2-handles, $D_i \times [0,1]$, along the γ_i; i.e., identifying the annulus $\partial D_i \times [0,1]$ with the annulus $N(\gamma_i)$ for $1 \leq i \leq n$. $F \times 1$ is a component of the boundary of this 3-manifold. There may be some number of 2-sphere components in the boundary as well. We may or may not fill in some of the 2-sphere boundary components with 3-cells (3-handles). We call the resulting 3-manifold, say H, a compression body and denote the boundary component $F \times 1$ by $\partial_+ H$ and denote the remaining boundary, which may not be connected, by $\partial_- H$. A component of $\partial_- H$, which is not a 2-sphere, is incompressible in H. If $\partial_- H = \emptyset$, then H is a handlebody, and if each component of $\partial_- H$ is a 2-sphere, then H is a punctured handlebody. Finally, $F \times [0,1]$ is itself a compression body, as is a punctured $F \times [0,1]$. We can then obtain a similar proposition.

Proposition 5. *Suppose $\mathcal{T}$ is a triangulation of the closed, orientable 3-manifold M, and suppose F is a normal, two-sided surface embedded in M. Then there are compression bodies H' and H'' embedded in M so that $H' \cap H'' = F = \partial_+ H' = \partial_+ H''$, each component of $\partial_- H'$ and $\partial_- H''$ is normal, and $H' \cup H''$ contains all vertices of $\mathcal{T}$.*

Proof. We proceed along similar lines. Let M' denote the manifold obtained by splitting M along F; let F' and F'' denote the copies of F in $\partial M'$. There are disjoint subcomplexes Λ' and Λ'' of the 1-skeleton of the induced cell structure on M' so that each component of Λ' and Λ'' is a tree and meets F' and F'', respectively, in precisely one point, $\Lambda' \cap F'' = \emptyset = \Lambda'' \cap F'$, and $\Lambda' \cup \Lambda''$ contains all vertices of $\mathcal{T}$. Let B' and B'' be the boundaries of small regular neighborhoods of $F' \cup \Lambda'$ and $F'' \cup \Lambda''$, respectively. Then $B' \cup B''$ is a barrier surface for the components of their complements not meeting $F' \cup \Lambda' \cup F'' \cup \Lambda''$; furthermore, B' and F' are the boundaries of a compression body, as are B'' and F''; actually, in these cases, the compression bodies are products. We shrink $B' \cup B''$. In shrinking $B' \cup B''$, we obtain two compression bodies G' and G'' so that $\partial_+ G' = F'$, $F' \cup \Lambda' \subset G'$, $\partial_+ G'' = F''$, $F'' \cup \Lambda'' \subset G''$, and each component of $\partial_- G'$ and of $\partial_- G''$ is either a normal surface or a 0-weight 2-sphere contained entirely in the interior of a cell in the induced cell structure on M'. Any such 0-weight 2-sphere bounds a 3-cell missing $F' \cup \Lambda'$ and $F'' \cup \Lambda''$. We fill in these 0-weight 2-spheres with such 3-cells. Finally, when we reattach F' and F'' to get M and set H' equal to the image of G' and set H'' equal to the image of G'', we get the desired compression bodies. $\square$

Proposition 6. *Suppose $\mathcal{T}$ is a triangulation of the closed, orientable, irreducible 3-manifold M, and suppose F is an embedded normal surface that bounds a handlebody H in M. If F is incompressible in $M \setminus \mathring{H}$ and is not contained in a 3-cell in M, then there is a normal surface F' embedded in M, F' is parallel to F, and bounds a handlebody H' in M so that $H \subset H'$ and H' contains all the vertices of $\mathcal{T}$.*

Proof. In this case, we split M along F and consider only the component that does not meet the handlebody H, call it M'. We denote the copy of F

in $\partial M'$ by F'. We have the subcomplex Λ' as above, and we let B' denote the boundary of a small regular neighborhood of $F' \cup \Lambda'$. B' is a barrier surface in the component of its complement not meeting $F' \cup \Lambda'$. Furthermore, F' and B' bound a compression body that is homeomorphic to $F' \times I$. We shrink B'. However, since F' is incompressible in M' (hence, B' is incompressible in M'), each normal move that is a compression is an inessential compression. It follows that in the stable situation, we have a surface that is a copy of F', and every other component is either a normal 2-sphere or a 0-weight 2-sphere contained entirely in the interior of a cell in the induced cell decomposition of M'. Since M is irreducible and F is not contained in a 3-cell, we can fill in each 2-sphere boundary component with a 3-cell whose interior does not meet the compression body. It follows that we have a compression body G' with $\partial_+ G' = F'$, $\partial_- G'$ a normal surface isotopic to F', and $F' \cup \Lambda' \subset G'$. When we reattach M' to H to get M, the image of G' along with H gives us the desired handlebody H'. □

Proposition 7 (Double barrier method)**.** *Suppose $\mathcal{T}$ is a triangulation of the compact, orientable 3-manifold M, and suppose K and L are disjoint subcomplexes in $\mathcal{T}$. Then there is a normal surface F in M separating K and L.*

Proof. Let B_K and B_L denote the frontiers of small regular neighborhoods of K and L, respectively, chosen so that $B_K \cap B_L = \emptyset$. Then B_K and B_L are barrier surfaces in the component of the complement of $B_K \cup B_L$ not meeting $K \cup L$. Furthermore, B_K separates K and L. We shrink B_K. In shrinking B_K, we have compressions, ∂-compressions, and isotopy moves. An isotopy move or θ-compression occurs through the interior of an edge that meets B_K and so is away from K or L. A compression is entirely in the interior of a cell or the face of a cell and so does not run through K or L. So, in our stable situation, we have components that are either normal surfaces or 0-weight 2-spheres and disks that are properly embedded in the cells of our induced cell decomposition; furthermore, the union of these components separates K from L. We wish to eliminate the 0-weight 2-spheres and disks. None of the 0-weight 2-spheres and disks separate any components of K from L; so, we can discard these components. Since we must have K separated from L, we have the desired normal surface. □

Remark 9. Notice that in shrinking a surface, we do not increase the genus of the surface, even in the bounded case. Hence, in the previous proposition, the separating normal surface may be found so that its genus is no more than the minimal genus of the surfaces B_K and B_L. In particular, if one of K or L is simply connected, then we can separate K and L by normal 2-spheres.

We also have the following proposition, which can be proved along the same lines as the previous proposition.

Proposition 8. *Suppose $\mathcal{T}$ is a triangulation of the compact, orientable 3-manifold M, and S is a closed, two-sided normal surface in M. Let M' be the manifold obtained by splitting M along S. Suppose $D_1, \ldots, D_n$ is a collection of pairwise disjoint, properly embedded disks in M', which are normal in the induced cell-decomposition on M'. Furthermore, suppose the D_i are all on the same side of S (only meet S', say, and so do not meet S'' in M'). Then there is a compression body H embedded in M, $\partial_+ H = S$, each component of $\partial_- H$ is a normal surface in M, and $D_1 \cup \ldots \cup D_n \subset H$.*