MORSE THEORY FOR CELL COMPLEXES

ROBIN FORMAN

Department of Mathematics
Rice University
Houston, TX 77251

§0 Introduction.

In this paper we will present a very simple discrete Morse theory for CW complexes. In addition to proving analogues of the main theorems of Morse theory, we also present discrete analogues of such (seemingly) intrinsically smooth notions as the gradient vector field and the gradient flow associated to a Morse function. Using this, we define a Morse complex, a differential complex built out of the critical points of our discrete Morse function which has the same homology as the underlying manifold.

This Morse theory takes on particular significance in the context of PL manifolds. To clarify this statement, we take a small historical digression. In 1961, Smale proved the h-cobordism theorem for smooth manifolds (and hence its corollary, the Poincaré conjecture in dimension ≥ 5) using a combination of handlebody theory and Morse theory [Sm2]. In [Mi2], Milnor presented a completely Morse theoretic proof of the h-cobordism theorem. In ([Ma], [Ba], [St]) Mazur, Barden and Stallings generalized Smale's theorem by replacing Smale's hypothesis that the cobordism be simply-connected by a weaker simple-homotopy condition. Along the way, this more general theorem (the s-cobordism theorem) was extended to other categories of manifolds. In particular, a PL s-cobordism theorem was established. In this case, it was necessary to work completely within the context of handlebody theory. The Morse theory presented in this paper can be used to give a Morse theoretic proof of the PL s-cobordism theorem, along the lines of the proof in [Mi2].

In the remainder of this introduction, we present an informal exposition of the contents of the paper. To avoid minor complications we will restrict attention, in this introduction, to simplicial complexes.

Let M be any finite simplicial complex, K the set of simplices of M, and K_p the simplices of dimension p. A discrete Morse function on M will actually be a function on K. That is, we assign a single real number to each simplex in M. Write $\sigma^{(p)}$ if σ has dimension p, and $\tau > \sigma$ if σ lies in the boundary of τ. We say a function

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

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

$$(1) \quad \#\{\tau^{(p+1)} > \sigma \mid f(\tau) \leq f(\sigma)\} \leq 1.$$

---

(2) $\#\{v^{(p-1)} < \sigma \mid f(v) \geq f(\sigma)\} \leq 1$.

For example, in figure 0.1, the function (i) is not a discrete Morse function as the edge $f^{-1}(0)$ violates rule (2) and the vertex $f^{-1}(3)$ violates rule (1). The function (ii) is a Morse function.
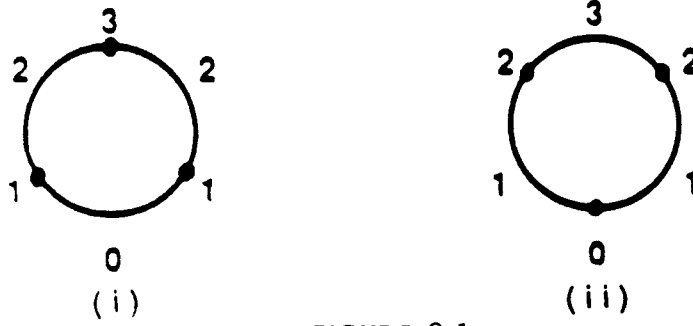

FIGURE 0.1

We say $\sigma^{(p)}$ is critical (with index p) if

(1) $\#\{\tau^{(p+1)} > \sigma \mid f(\tau) \leq f(\sigma)\} = 0$.
(2) $\#\{v^{(p-1)} < \sigma \mid f(v) \geq f(\sigma)\} = 0$.

For example, in figure 0.1(ii), $f^{-1}(0)$ is a critical point of index 0, $f^{-1}(3)$ is a critical simplex of index 1 and there are no other critical simplices. Note that if $\sigma^{(p)}$ is critical, it is necessarily critical of index p.

The above definition provides a discrete analogue of the smooth notion of a critical point of index p. For example, suppose x is a critical point of index 1 of a smooth Morse function F on a smooth manifold of dimension n. Then the Morse Lemma (see [Mil], Lemma 2.2) states that there are coordinates $(t_1, \ldots, t_n)$, with x corresponding to $(0, \ldots, 0)$, such that in these coordinates

$$F(t_1, \ldots, t_n) = F(x) - t_1^2 + \sum_{i=2}^{n} t_i^2.$$

That is, beginning at the point x, F decreases to both sides in the t_1 direction, and increases in the transverse directions. Now suppose σ is a critical edge of a discrete Morse function f. Then $f(\sigma)$ is greater than f at either boundary vertex, and less than f at any 2-simplex with σ in its boundary (see, for example, figure 0.2).
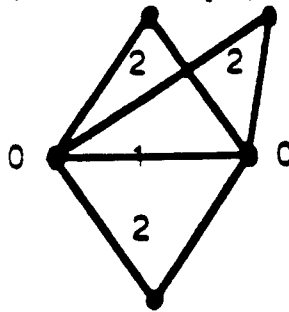

FIGURE 0.2

That is, f decreases as one moves from the edge to either boundary component, and increases in every transverse direction. We can see that this is, in fact, a discrete analogue of the smooth situation. Moreover, we see that, heuristically, if $\sigma^{(p)}$ is critical then the

2

simplex σ can be thought of as representing the p-dimensional "unstable" space at a smooth critical point of index p.

Before stating the main theorems, we need to present a definition. Suppose f is a discrete Morse function on a simplicial complex M. For any $c \in \mathbf{R}$ define the level subcomplex $M(c)$ by

$$M(c) = \bigcup_{f(\tau) \le c} \; \bigcup_{\sigma \le \tau} \sigma.$$

That is. $M(c)$ is the subcomplex of M consisting of all simplices τ with $f(\tau) \le c$, as well as all of their faces.

Theorem. *Suppose the interval $[a.b]$ contains no critical values of f. Then $M(a)$ is a deformation retract of $M(b)$. Moreover, $M(b)$ simplicially collapses onto $M(a)$.*

We will not define simplicial collapse here (the definition appears in section 1), but we will mention that the equivalence relation generated by simplicial collapse is called simple homotopy equivalence. This indicates that our Morse theory is particularly well suited to handling questions in this category.

Theorem. *Suppose $\sigma^{(p)}$ is a critical simplex with $f(\sigma) \in [a,b]$, and there are no other critical simplices with values in $[a.b]$. Then $M(b)$ is homotopy equivalent to*

$$M(a) \bigcup_{\dot{e}^{(p)}} e^{(p)}$$

where $e^{(p)}$ is a p-cell. and it is glued to $M(a)$ along its boundary $\dot{e}^{(p)}$.

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

This Corollary implies the standard Morse Inequalities (see Corollaries 3.6 and 3.7 and [Mil]). We present a simple example. See figure 0.3
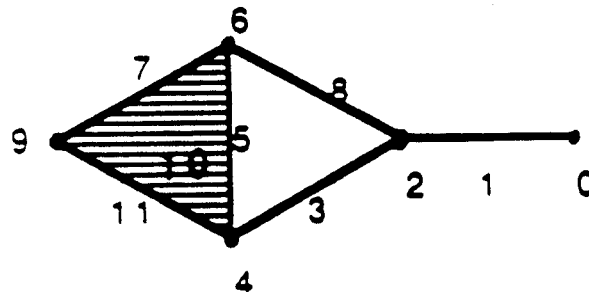

FIGURE 0.3

Here $f^{-1}(0)$ is a critical simplex of index 0, $f^{-1}(8)$ is a critical simplex of index 1, and there are no other critical simplices. Thus, it follows from the above corollary that M is homotopy equivalent to a circle, as is evident from the picture.

It is useful to indicate pictorially the simplicial collapse referred to in the theorem. Suppose $\sigma^{(p)}$ is a non-critical simplex with $\tau^{(p+1)} > \sigma$ satisfying $f(\tau) \leq f(\sigma)$. We then draw an arrow from σ to τ. The resulting diagram can be seen in figure 0.4. A simplex is critical if and only if it is neither the tail nor the head of an arrow. These arrows can be viewed as the discrete analogue of the gradient vector field of the Morse function.
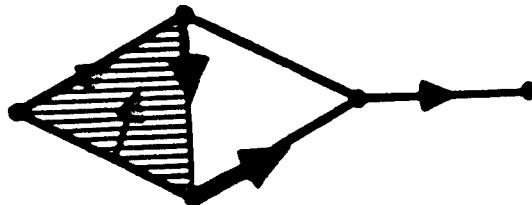

FIGURE 0.4

It is better to think of the gradient vector field, which we now call V, as a map of oriented simplices. That is, if v is a boundary vertex of an edge e with $f(e) \leq f(v)$, we want to think of $V(v)$ as a discrete tangent vector *leaving* v i.e., with e given the orientation indicated by the arrow in figure 0.5

FIGURE 0.5

More generally, if $\tau^{(p+1)} > \sigma^{(p)}$ satisfies $f(\tau) \leq f(\sigma)$ then we set $V(\sigma) = \pm\tau$ with the sign chosen so that

$$< \sigma, \partial V(\sigma) > \; = \; -1$$

where $<,>$ is the obvious inner product on oriented chains (with respect to which the oriented simplices are orthonormal). That is, $V(\sigma) = - < \sigma, \partial\tau > \tau$ (the integer $< \sigma, \partial\tau >$ is usually called the *incidence number* of τ and σ). Now V can be extended linearly to a map

$$V : C_p(M, \mathbf{Z}) \longrightarrow C_{p+1}(M, \mathbf{Z})$$

where, for each p, $C_p(M, \mathbf{Z})$ is the space of integer p-chains on M.

The next step is to define the discrete gradient flow. We define a map

$$\Phi : C_p(M, \mathbf{Z}) \longrightarrow C_p(M, \mathbf{Z}),$$

4

the discrete-time flow, by

$$\Phi = 1 + \partial V + V\partial.$$

We illustrate by an example. Consider the following complex, with the indicated gradient vector field V

FIGURE 0.6

Let e be the top edge oriented from left to right. We shall calculate

$$\Phi(e) = e + \partial V(e) + V\partial(e)$$
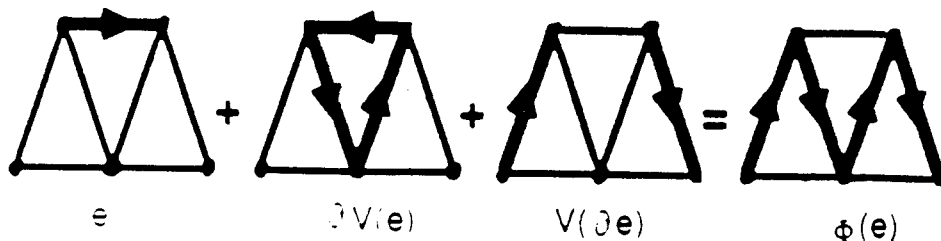
in figure 0.7.

$$e \qquad \partial V(e) \qquad V(\partial e) \qquad \Phi(e)$$

FIGURE 0.7

The properties of the map Φ are examined in section 6.

Let $C_p^\Phi \subseteq C_p(M, \mathbf{Z})$ denote the Φ-invariant p-chains, i.e., those p-chains c such that $\Phi(c) = c$. Since

$$\partial\Phi = \Phi\partial,$$

the Φ-invariant chains form a differential complex

$$C^\Phi : 0 \longrightarrow C_n^\Phi \xrightarrow{\partial} C_{n-1}^\Phi \xrightarrow{\partial} C_{n-2}^\Phi \xrightarrow{\partial} \cdots .$$

We prove in section 7 that

$$H_*(C^\Phi) \cong H_*(M, \mathbf{Z}).$$

That is, this complex, which we call the Morse complex, has the same homology as the underlying manifold. For discussions of the Morse complex in the smooth category see, for example, [Mi2] and [Kl].

The Morse complex can also be defined using critical simplices. For each p, let $M_p \subseteq C_p(M, \mathbf{Z})$ denote the span of the critical p-simplices. We prove in section 8 that for each p

$$M_p \cong C_p^\Phi.$$

Thus. the Morse complex C^Φ can be defined equivalently as the complex

$$\mathcal{M} : 0 \longrightarrow \mathcal{M}_n \xrightarrow{\;\partial\;} \mathcal{M}_{n-1} \xrightarrow{\;\partial\;} \mathcal{M}_{n-2} \xrightarrow{\;\partial\;} \cdots$$

where $\tilde{\partial}$ is the differential induced by the above isomorphism. We prove in section 8 that from this point of view, the differential can be defined by setting, for any oriented critical simplex $\tau^{(p-1)}$

$$\tilde{\partial}\tau = \sum_{\tilde{\sigma}^{(p)} < \tau} <\partial\tau, \tilde{\sigma}> \sum_{\substack{\text{oriented} \\ \text{critical } \sigma^{(p)}}} \left[\sum_{\gamma \in \Gamma(\tilde{\sigma}, \sigma)} m(\gamma) \right] \sigma$$

where $\Gamma(\tilde{\sigma}, \sigma)$ is the set of gradient paths from $\tilde{\sigma}$ to σ. Rather than define this term precisely, we again illustrate by an example. In figure 0.8 we show a single gradient path from the boundary of a critical 2-simplex τ to a critical edge σ, where the arrows indicate the gradient vector field.
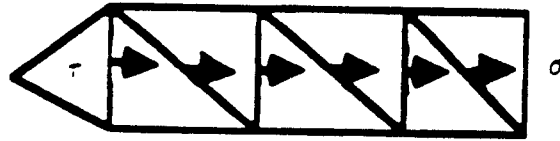

FIGURE 0.8

The coefficient $m(\gamma)$ is equal to ± 1 (in the case of a simplicial complex) and is determined by whether the orientation on τ induces (in a manner defined precisely in section 8) the chosen orientation on σ, or the opposite orientation.

It is frequently desirable to deform one Morse function into another. It is often convenient to work instead with the gradient vector field. When deforming the gradient vector field. it is necessary to know that the resulting vector field is, in fact, the gradient vector field of a Morse function. Thus, it is important to characterize those vector fields which are gradient vector fields of Morse functions. In the smooth case, Smale provided such a characterization in [Sm1]. In section 9, we provide an analogous characterization of discrete gradient vector fields. First, we define an object we call a discrete vector field. Such objects have been previously studied in [Du], [Sta] (under a different name) and the references therein (see [Fo1] for discrete vector fields on 1-dimensional complexes). We then prove, essentially, that any discrete vector field which has no closed loops is, in fact, a gradient vector field.

The results in section 9 are very powerful, and in the remainder of the paper we present some applications. For example, in section 11 we prove a "cancellation" theorem, a discrete analogue of Theorem 5.4 in [Mi2]. That is, if $\sigma^{(p)}$ and $\tau^{(p+1)}$ are 2 critical simplices, and if there is exactly 1 gradient path from $\partial\tau$ to σ, then σ and τ can be cancelled. More precisely, there is a Morse function with the same critical points except that σ and τ are

6

no longer critical. In the smooth case, the proof, either as presented originally by Morse in
[Mo] or as presented in [Mi2], is rather technical. In our discrete case the proof is simple.
If, in figure 0.8, the indicated gradient path is the only gradient path from $\partial \sigma$ to τ, then we
can vary the gradient vector field only along this path, replacing the figure by the vector
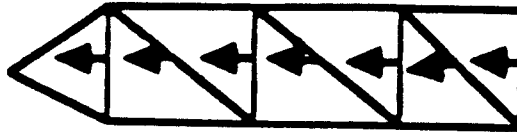field in figure 0.9.

FIGURE 0.9

Our characterization of gradient vector fields immediately implies that this new vector
field is the gradient vector field associated to some Morse function, and σ and τ are no
longer critical. This cancellation theorem is the fundamental tool in the Morse theoretic
proof of the s-cobordism theorem.

In analogy with [Sm1] and [Sm2], in sections 9 and 11, rather than working with a single
cell complex, we work instead with a cellular triad, which plays the role of a cobordism in
the cellular category.

In [Wi], Witten provided a Hodge-theoretic proof of the main theorems of smooth
Morse theory, as well as an analytic derivation of the Morse complex. In [Fo2], we provide
a discrete Hodge-theoretic derivation of the main theorems of discrete Hodge theory and
the Morse complex.

Before leaving this introduction, we note that there have been earlier attempts to de-
velop a Morse theory for PL manifolds, by restricting attention to piecewise-linear func-
tions (see, for example [K] and [B-K]). There are some drawbacks to such a theory. For
example, while it is easy to define a critical point for such functions, without further re-
strictions the notion of the index of a critical point is more difficult. More significantly,
the Morse complex does not appear in this setting. Hence there seems to be no direct way
of seeing the more subtle topological information contained in the Morse theory. The sim-
pler theory we present in this paper exhibits neither of these drawbacks, and, in addition,
applies to much more general cell complexes.

§1. Preliminaries.

Although the primary objects of study in this paper will be CW complexes, we will occasionally require extra structure. The four types of complexes we will study (in order of increasing structure) are: *finite CW complexes, regular CW complexes, polyhedra,* and *PL manifolds.* In this section, we define these different types of complexes, as well as some of their fundamental properties. The interested reader may skip this section and refer back as necessary.

Let M be a finite CW complex (see [L-W] for definitions and basic properties of CW complexes and regular CW complexes) and let K denote the set of open cells of M, with K_p the cells of dimension p. The notation $\sigma^{(p)}$ will indicate that σ is a cell of dimension p. To indicate relationships between cells, we write $\tau > \sigma$ (or $\sigma < \tau$) if $\sigma \neq \tau$ and $\sigma \subset \bar{\tau}$, where $\bar{\tau}$ is the closure of τ, and we say σ is a face of τ. We write $\tau \geq \sigma$ if either $\tau = \sigma$ or $\tau > \sigma$.

Suppose $\sigma^{(p)}$ is a face of $\tau^{(p+1)}$. Let B be a closed ball of dimension $p+1$, and

$$h : B \longrightarrow M$$

the characteristic map for τ. i.e., h is a continuous map that maps interior(B) homeomorphically onto τ.

Definition 1.1. Say $\sigma^{(p)}$ is a *regular face* of $\tau^{(p+1)}$ if

    (i) $h : h^{-1}(\sigma) \longrightarrow \sigma$ is a homeomorphism
    (ii) $\overline{h^{-1}(\sigma)}$ is a closed p-ball.

Otherwise we say σ is an *irregular face* of τ.

We note that if M is a regular CW complex (and hence if M is a simplicial complex or a polyhedron) then all faces are regular. Of crucial importance is the following property. Suppose $\sigma^{(p)}$ is a regular face of $\tau^{(p+1)}$. Choose an orientation for each cell in M and consider σ and τ as elements in the cellular chain groups $C_p(M, \mathbf{Z})$ and $C_{p+1}(M, \mathbf{Z})$, respectively. Then

$$(1.1) \qquad\qquad \langle \partial \tau, \sigma \rangle = \pm 1$$

where $\langle \partial \tau, \sigma \rangle$ is the incidence number of τ and σ (for a proof see Corollary V.3.6 of [L-W]).

We will require the following property of CW complexes

Theorem 1.2. *Suppose* $\tau^{(p+1)} > \sigma^{(p)} > v^{(p-1)}$, *then one of the following is true.*

    (i) σ *is an irregular face of* τ.
    (ii) v *is an irregular face of* σ.
    (iii) *There is a p-cell* $\bar{\sigma} \neq \sigma$ *satisfying*

$$\tau > \bar{\sigma} > v.$$

Proof. Suppose neither (i) nor (ii) is true. Choose an orientation for each cell of M. Since σ is a regular face of τ, (1.1) holds, so that

$$\partial\tau = \pm\sigma + \sum_{\substack{\bar{\sigma}^{(p)}\neq\sigma \\ \bar{\sigma}<\tau}} c_{\bar{\sigma}}\bar{\sigma}$$

for some integers $c_{\bar{\sigma}}$. Similarly, since υ is a regular face of σ,

$$\partial\sigma = \pm\upsilon + \sum_{\bar{\upsilon}\neq\upsilon} c_{\bar{\upsilon}}\bar{\upsilon}.$$

Therefore.

$$0 = \partial^2\tau = \pm\partial\sigma + \sum_{\tau>\bar{\sigma}\neq\sigma} c_{\bar{\sigma}}\partial\bar{\sigma}$$
$$= \pm\upsilon + \sum_{\tau>\bar{\sigma}\neq\sigma} c_{\bar{\sigma}}\partial\bar{\sigma} + \sum_{\bar{\upsilon}\neq\upsilon} c_{\bar{\upsilon}}\bar{\upsilon}.$$

For this equation to hold, there must be some $\bar{\sigma}$, with $\tau>\bar{\sigma}\neq\sigma$, satisfying

$$\partial\bar{\sigma} = c\upsilon + (\text{sum of } (p-1)\text{-cells other than } \upsilon)$$

for some $c\neq 0$. This implies $\bar{\sigma}>\upsilon$, so that $\tau>\bar{\sigma}>\upsilon$ as desired. $\square$

If M and N are CW complexes, say M and N are *isomorphic*, denoted by $M\cong N$, if there is a homeomorphism

$$h : N \longrightarrow M$$

which maps each cell of N homeomorphically onto a single cell of M. Say $\widetilde{M}$ is a *subdivision* of M if there is a homeomorphism

$$h : \widetilde{M} \longrightarrow M$$

which maps each cell of $\widetilde{M}$ into a single cell of M. Say M and N are *equivalent*, denoted by $M = N$. if there are finite subdivisions $\widetilde{M}$ of M and $\widetilde{N}$ of N with $\widetilde{M}\cong\widetilde{N}$.

Suppose M is a CW complex and $\sigma^{(p)} < \tau^{(p+1)}$ are 2 cells of M which satisfy

(i) σ is a regular face of τ.

(ii) σ is not a face of any other cell.

Let $N = M - (\sigma\cup\tau)$. We say M *collapses onto* N. More generally, we say M collapses onto N, and we write $M \searrow N$, if M can be transformed into N by a finite sequence of such operations. For example, figure 1.1 illustrates a 2-simplex collapsing onto a vertex.
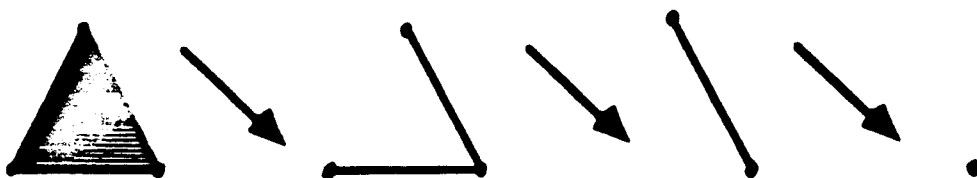

FIGURE 1.1

9

Note that if $M \searrow N$ then, in particular, N is a deformation retract of M.

Recall that a *regular cell complex* is a CW complex M in which for every p-cell σ of M there is a homeomorphism from the closed p-dimensional ball B into M which maps the interior of B homeomorphically onto σ. In particular, every simplicial complex, and more generally every polyhedra (with a fixed polyhedral decomposition), is a regular cell complex. We observe that if M is a regular cell complex, then if $\sigma^{(p)} < \tau^{(p+1)}$, σ is necessarily a regular face. From Theorem 1.2 we learn that if M is a regular cell complex, then $\tau^{(p+1)} > \sigma^{(p)} > v^{(p-1)}$ implies there is a $\bar{\sigma}^{(p)} \neq \sigma$ such that $\tau > \bar{\sigma} > v$. In fact, we will make use of the following generalization.

Theorem 1.3. *Suppose M is a regular cell complex, and for some p and $r \geq 1$ we have $\tau^{(p+r)} > v^{(p-1)}$. Then there are $p + r - 1$-cells $\sigma^{(p)}$ and $\bar{\sigma}^{(p)}$ such that $\sigma \neq \bar{\sigma}$ and*

$$\tau > \sigma > v, \quad \tau > \bar{\sigma} > v.$$

Proof. The proof is by induction on r. Suppose $r = 1$, that is, we have $\tau^{(p+1)} > v^{(p-1)}$. Since M is regular, the p-cells in $\dot{\tau}(= \bar{\tau} - \tau)$ are dense in $\dot{\tau}$, i.e.,

$$\overline{\bigcup_{\sigma^{(p)} < \tau} \sigma} = \bar{\tau} - \tau.$$

Thus, there is a p-cell $\sigma^{(p)}$ with $\tau > \sigma$ such that $\sigma > v$. Theorem 1.2 then guarantees the existence of $\bar{\sigma}^{(p)} \neq \sigma$ such that $\tau > \bar{\sigma} > v$.

For general r, we again have that the $(p + r - 1)$-cells in $\dot{\tau}$ are dense in $\dot{\tau}$. Thus we can find a $(p + r - 1)$-cell σ with $\tau > \sigma > v^{(p-1)}$. Continuing in this fashion, we can find a $(p + r - 2)$-cell $\bar{v}$ such that $\sigma > \bar{v} > v$. Applying Thoerem 1.2 to the triple $\tau > \sigma > \bar{v}$ we learn there is a $(p + r - 1)$-cell $\bar{\sigma} \neq \sigma$ such that $\tau > \bar{\sigma} > \bar{v}$. The cells σ and $\bar{\sigma}$ satisfy the desired properties. $\square$

We will occasionally require that M be a *polyhedron* (see [St] for the foundations of polyhedral topology). At such times, one can restrict attention to simplicial complexes without any significant loss of information.

A CW complex M is a polyhedron if M can be embedded in some Euclidean space such that every p-cell of M is convex and lies in a single p-dimensional affine subspace. Isomorphism, subdivision and equivalence are defined as for general CW complexes, with the proviso that when we speak of a map $h : N \to M$ between polyhedra we always mean a map which is linear on each cell of N, and we only consider subdivisions $\bar{M}$ of M which are themselves polyhedra (we will occasionally emphasize this point by refering to $\bar{M}$ as a *polyhedral subdivision*).

The main advantage of using the more general polyhedra, as opposed to simplicial complexes, is that one can restrict attention to a small set of subdivisions called *bisections*,

in which a single cell is divided into two (see [St] for a definition). In figure 1.2 we show a sequence of bisections of a simple polyhedron.
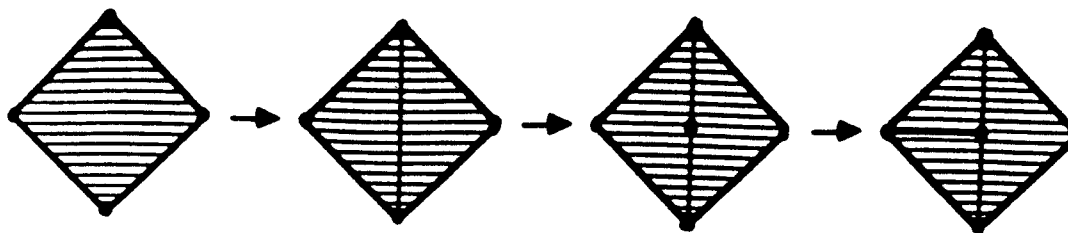

FIGURE 1.2

The following is Corollary 1.10.5 of [St].

Theorem 1.4. *If $\widetilde{M}$ is any finite polyhedral subdivision of a polyhedron M then there is a finite polyhedral subdivision $\widetilde{\widetilde{M}}$ of $\widetilde{M}$ (and hence of M) which can be produced from M by applying a finite sequence of bisections.*

Corollary 1.5. *If M and N are polyhedra, then M and N are equivalent if and only if there is a subdivision $\widetilde{M}$ of M resulting from a finite sequence of bisections, and a subdivision $\widetilde{N}$ of N resulting from a finite sequence of bisections, with $\widetilde{M} \cong \widetilde{N}$.*

At a few points, we will require that a polyhedron is a *PL n-manifold*, i.e. that each vertex of M has a PL neighborhood equivalent with the standard n-cell (see [St] for precise definitions). The reader need not be familiar with such objects. The main point is that we can make use of Whitehead's remarkable Theorem of Regular Neighborhoods (see [Wh] or [Gl] for simplicial complexes and [St] for PL manifolds). We quote a special case.

Theorem 1.6. *Let M be a PL n-manifold with boundary and v a vertex of M. If $M \searrow v$ then M is a PL n-cell (i.e., M is equivalent to an n-simplex with its standard triangulation).*

For applications of this theorem see Theorems 5.1 and 9.8.

11

§2 Combinatorial Morse Functions.

In this section we introduce the main definitions. Let M be a finite CW complex. All definitions and notation are as in §1.

Definition 2.1. A *discrete Morse function* on M is a function

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

satisfying for all $\sigma \in K_p$

(i) If σ is an irregular face of $\tau^{(p+1)}$ then $f(\tau) > f(\sigma)$. Moreover,

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

(ii) If $v^{(p-1)}$ is an irregular face of σ then $f(v) < f(\sigma)$. Moreover,

$$\#\{v^{(p-1)} < \sigma \mid f(v) \geq f(\sigma)\} \leq 1.$$

Definition 2.2. Given a combinatorial Morse function f on M we say $\sigma \in K_p$ is a *critical point of index p* if

(i) $\#\{\tau^{(p+1)} > \sigma \mid f(\tau) \leq f(\sigma)\} = 0$.
(ii) $\#\{v^{(p-1)} < \sigma \mid f(v) \geq f(\sigma)\} = 0$.

We note that a face σ of dimension p cannot be a critical point of any index other than p. Thus, without any loss of information, we may refer to σ simply as a critical point.

Example 2.3. Definitions 2.1 and 2.2 imply that if M is regular then the minimum of f must occur on a vertex, which must then be a critical point of index 0. This follows from the observation that if $p \geq 1$, then every p-cell has at least 2 $(p-1)$-dimensional faces.

Example 2.4. If M is a triangulated n-dimensional manifold without boundary, then the maximum of f must occur on a n-face, which must then be a critical point of index n. This follows from the observation that if $p \leq n-1$, then every p-cell is a face of at least 2 $(p+1)$-cells.

It follows from Definition 2.2 that a p-cell σ is not critical if and only if either of the following conditions holds

(i) $\exists \tau^{(p+1)} > \sigma$ such that $f(\tau) \leq f(\sigma)$.
(ii) $\exists v^{(p-1)} < \sigma$ such that $f(v) \geq f(\sigma)$.

Lemma 2.5. *Conditions (i) and (ii) cannot both be true.*

Proof. Condition (ii) requires $p \geq 1$ which we now assume. Suppose (i) is true. Then σ must be a regular face of τ. Moreover, from condition (i) from Definition 2.1, if $\bar{\sigma} \neq \sigma$ is any other p-face of τ, we must have

$$f(\bar{\sigma}) < f(\tau)$$

so that, in particular,

$$(2.1) \qquad\qquad f(\bar{\sigma}) < f(\sigma).$$

Now suppose (ii) is true. Then v must be a regular face of σ. By Theorem 1.2, there is a p-cell $\bar{\sigma} \neq \sigma$ satisfying

$$\tau > \bar{\sigma} > v.$$

From condition (ii) of Definition 2.1, $f(v)$ cannot be $\geq$ both $f(\sigma)$ and $f(\bar{\sigma})$. Thus $f(v) < f(\bar{\sigma})$ Combining this with (2.1) we learn

$$f(\sigma) \leq f(v) < f(\bar{\sigma}) < f(\tau) \leq f(\sigma)$$

which is a contradiction. $\square$

§3 The Morse Theorems for Regular Cell Complexes.

In this section we prove the main theorems of Morse theory in the case that M is a regular cell complex (the theorems are extended to general CW complexes in Corollary 8.3 and section 10). Let M be a regular cell complex and f a discrete Morse function on M.

Definition 3.1. For $c \in \mathbf{R}$, define

$$M(c) = \bigcup_{\substack{\sigma \in K \\ f(\sigma) \leq c}} \bigcup_{\tau \leq \sigma} \tau,$$

That is, $M(c)$ denotes all cells on which f is $\leq c$, as well as all of their faces. In particular, $M(c)$ is a subcomplex of M.

As it stands, to see if a cell σ with $f(\sigma) > c$ lies in $M(c)$ we must see if there is any τ with $\sigma < \tau$ and $f(\tau) \leq c$. In fact, it is enough to consider τ with

$$\dim \tau = \dim \sigma + 1.$$

This is the content of the following lemma.

Lemma 3.2. *Let σ be a p-cell of M and suppose $\tau > \sigma$. Then there is a $(p+1)$-cell $\tilde{\tau}$ with $\sigma < \tilde{\tau} < \tau$ and*

$$f(\tilde{\tau}) \leq f(\tau).$$

Proof. Since $\tau > \sigma$, it follows that $\dim \tau > \dim \sigma$. If $\dim \tau = p + 1$ we can let $\tilde{\tau} = \tau$. Assume

$$\dim \tau = p + r, \quad r > 1.$$

Then by Theorem 1.3 we can find two $(p + r - 1)$-faces v_1, v_2 satisfying

$$\tau > v_1 > \sigma$$
$$\tau > v_2 > \sigma.$$

From condition (ii) of Definition 2.1 either

$$f(v_1) < f(\tau)$$

or

$$f(v_2) < f(\tau).$$

In either case the result follows by induction. □

We now arrive at the main theorems of Morse theory.

Theorem 3.3. *If $a < b$ are real numbers such that $[a,b]$ contains no critical values of f, then*

$$M(b) \searrow M(a).$$

Proof. Note that if $\tau^{(p+1)} > \sigma^{(p)}$ satisfies $f(\tau) \le f(\sigma)$ then we may perturb f by replacing $f(\tau)$ by $f(\tau) - \epsilon$, or $f(\sigma)$ by $f(\sigma) + \epsilon$, for $\epsilon \ge 0$ small enough, without changing which cells are critical. If $\sigma^{(p)}$ satisfies $f(\tau^{(p+1)}) \ne f(\sigma) \ne f(v^{(p-1)})$ for each $\tau^{(p+1)} > \sigma > v^{(p-1)}$ then we may perturb f by changing $f(\sigma)$ to $f(\sigma) \pm \epsilon$, for ϵ small enough, without changing which cells are critical. Combining such operations, we may perturb f slightly without changing $M(b)$ or $M(a)$ so that

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

is 1-to-1.

If $f^{-1}([a,b]) = \emptyset$ then $M(a) = M(b)$ so there is nothing to prove. Otherwise, by partitioning $[a,b]$ into smaller intervals if necessary, we may assume there is a single non-critical cell σ with

$$f(\sigma) \in [a,b].$$

By Lemma 2.5 exactly 1 of the following holds:

   (i) $\exists \tau^{(p+1)} > \sigma$ with $f(\tau) \le f(\sigma)$.

   (ii) $\exists v^{(p-1)} < \sigma$ with $f(v) \ge f(\sigma)$.

In case (i), we must have $f(\tau) < a$. Thus $\tau \subseteq M(a)$. Since σ is a face of τ, we have $\sigma \subseteq M(a)$ so that

$$M(a) = M(b)$$

and again there is nothing to prove.

Suppose case (ii) is true. From Lemma 2.6, case (i) cannot be true, so for all $\tau^{(p+1)} > \sigma$ we have $f(\tau) > f(\sigma)$. In particular, $f(\tau) > b$. It follows from Lemma 3.2 that for any $\tau > \sigma$, $f(\tau) > b$. Therefore

$$\sigma \cap M(a) = \emptyset.$$

We have assumed there is a $v^{(p-1)} < \sigma$ with $f(v) > f(\sigma)$, so that, in particular, $f(v) > b$. If $\tilde{v}^{(p-1)} \ne v$ is any other $(p-1)$-face of σ we must have

$$f(\tilde{v}) < f(\sigma)$$

(from condition 2(ii) of Definition 2.1) so that

$$f(\tilde{v}) < a.$$

Thus, $\tilde{v}$ and all its faces are contained in $M(a)$.

Let $\tilde{\sigma}^{(p)} \ne \sigma$ be any other p-face of M with

$$\tilde{\sigma} > v.$$

Then condition 2(i) of Definition 2.1 implies

$$f(\tilde{\sigma}) > f(v) > b.$$

By Lemma 3.2, if $\bar{\sigma}$ is any face of any dimension such that $\bar{\sigma} > v$ then

$$f(\bar{\sigma}) > b$$

so that

$$v \cap M(a) = \emptyset.$$

It follows that $M(b)$ can be expressed as a disjoint union

$$M(b) = M(a) \cup \sigma \cup v$$

where v is a free face of σ. Therefore

$$M(b) \searrow M(a). \quad \square$$

Theorem 3.4. *Suppose $\sigma^{(p)}$ is a critical point of index p with*

$$f(\sigma) \in [a, b]$$

and $f^{-1}([a, b])$ contains no other critical points. Then $M(b)$ is homotopy equivalent to

$$M(a) \bigcup_{e^p} e^p$$

where e^p denotes a p-dimensional cell with boundary $\dot{e}^p$.

Proof. As in the proof of Theorem 3.3 we may assume f is $1 - 1$. Thus we can find a' and b' with

$$a < a' < b' < b$$

with

$$\sigma = f^{-1}([a', b']).$$

From Theorem 3.3 $M(b) \searrow M(b')$ and $M(a') \searrow M(a)$ so it is sufficient to prove that $M(b')$ is homotopy equivalent to

$$M(a') \bigcup_{\dot{e}^p} e^p.$$

Since σ is critical, if $\tau^{(p+1)} > \sigma$ we have

$$f(\tau) > f(\sigma)$$

so that

$$f(\tau) > b'.$$

From Lemma 3.2, if τ is any face of M with $\tau > \sigma$ then

$$f(\tau) > b'.$$

Thus

$$\sigma \cap M(a') = \emptyset.$$

It also follows from the criticality of σ that for every $v^{(p-1)} < \sigma$ we have

$$f(v) < f(\sigma)$$

so that

$$f(v) < a'$$

which implies

$$v \subseteq M(a').$$

Therefore,

$$\dot\sigma \subseteq M(a')$$

so that

$$M(b') = M(a') \bigcup_{\dot\sigma} \sigma.$$

Since σ is homeomorphic to e^p, the theorem follows. $\square$

Let $m_p(f)$ [or simply m_p if it will not cause confusion] denote the number of critical points of f of index p. The m_p's are called the *Morse numbers of f*. As an immediate corollary of Theorems 3.3 and 3.4 we learn

Corollary 3.5. *M is homotopy equivalent to a CW complex with exactly $m_p(f)$ cells of dimension p.*

Let **F** be a field and

$$b_i = \dim H_i(M, \mathbf{F})$$

the ith Betti number with coefficients in F.

Corollary 3.6. *(The Strong Morse Inequalities)* For any $N \geq 0$

$$m_N - m_{N-1} + \cdots \pm m_0 \geq b_N - b_{N-1} + \cdots \pm b_0.$$

Corollary 3.7. *(The Weak Morse Inequalities)*
   (i) *For every N*

$$m_N \geq b_N.$$

   (ii)

$$\chi(M) = b_0 - b_1 + b_2 - \cdots \pm b_{\dim M}$$
$$= m_0 - m_1 + m_2 - \cdots \pm m_{\dim M}.$$

For a discussion of how the corollaries follow from the preceding theorems, see [Mil].

§4 Examples of Discrete Morse Functions.

Every CW complex M has a discrete Morse function. For example, define a Morse function f by setting, for each $\sigma \in K$,

$$f(\sigma) = \dim \sigma.$$

Then every cell is critical. Corollary 3.5 (for regular cell complexes) and Theorem 10.3 (for general cell complexes) imply the tautological statement that a CW complex with m_p faces of dimension p is homotopy equivalent to a CW complex with m_p cells of dimension p.

We now examine ways in which a Morse function on one cell complex may induce a Morse function on another cell complex.

Lemma 4.1. *Let M be a CW complex and $N \subseteq M$ a subcomplex. Then any discrete Morse function on M restricts to a discrete Morse function on N. If $\sigma \subseteq N$ is a critical point for the original function, then it is critical for the restriction.*

Proof. The lemma follows directly from Definitions 2.1 and 2.2.

The following lemma is a converse to Lemma 4.1.

Lemma 4.2. *Let M be a cell complex and $N \subseteq M$ a subcomplex. Then any discrete Morse function on N can be extended to a discrete Morse function on M. That is, if f is a Morse function on N, then there is a Morse function g on M such that*

$$g(\sigma) = f(\sigma)$$

whenever $\sigma \subseteq N$.

Proof. Let $c = \max_{\sigma \subseteq N} f(\sigma)$. Define a combinatorial function g on M by setting, for each face σ on M

$$g(\sigma) = \begin{cases} f(\sigma) & \text{if } \sigma \subseteq N \\ c + \dim \sigma & \text{if } \sigma \not\subseteq N. \end{cases}$$

It can be easily seen that g is a Morse function on M that extends f. $\square$

The Morse function constructed in the above proof may be very inefficient. In particular, every face of $M - N$ is critical. There may exist extensions to M with many fewer critical points. This is partially rectified in the following lemma.

Lemma 4.3. *Let M be a CW complex and $N \subseteq M$ a subcomplex such that $M \searcloc N$. Let f be a Morse function on N and let $c = \max_{\sigma \subseteq N} f(\sigma)$. Then f can be extended to a Morse function on M with*

$$N = M(c)$$

and such that there are no critical points in $M - N$.

Proof. By induction on the number of elementary collapses required, it is sufficient to prove the lemma when M collapses onto N by a single elementary collapse. Suppose σ is a face of M with a free edge $\tau < \sigma$ such that M is a disjoint union

$$M = N \cup \sigma \cup \tau.$$

Define a Morse function g on M by setting

$$g(v) = f(v) \quad v \neq \sigma, \tau$$
$$g(\sigma) = c + 1$$
$$g(\tau) = c + 2.$$

Then it is easy to check that g satisfies the desired properties. $\square$

Corollary 4.4. *Let Δ^n denote the n-simplex with its standard triangulation, and $\dot{\Delta}^n$ its boundary. Then*

(i) *Δ^n has a Morse function with exactly 1 critical point.*

(ii) *$\dot{\Delta}^n$ has a Morse function with exactly 2 critical points.*

Proof. Part (i) follows from Lemma 4.3, since Δ^n collapses onto any vertex. Part (ii) follows from Lemmas 4.2 and 4.3 since, for any $(n-1)$-cell σ of $\dot{\Delta}^n$, $\dot{\Delta}^n - \sigma$ collapses onto any vertex.

For partial converses to Corollary 4.4 (ii) see §5.

We now present some general examples of how Lemma 4.3 may be used to construct Morse functions with desired properties.

Let M be an n-dimensional CW complex and f a discrete Morse function on M. Say f is a *polar Morse function* if

$$m_o(f) = m_n(f) = 1.$$

This definition was introduced by Morse in [Mo] where he proved that every smooth compact n-manifold without boundary has a smooth polar Morse function. We now prove a discrete analog. The proof proceeds in two steps.

Lemma 4.5. *Let M be a connected polyhedron. Then M has a Morse function f with $m_o(f) = 1$. Moreover, if $v \in K_0$ is a vertex of M, then f can be chosen so that v is the unique critical point of f of index 0.*

Proof. Let M_1 denote the 1-skeleton of M. Then M_1 is a connected graph. Let T be a maximal (= spanning) tree of M_1. That is, T is a connected, contractible subgraph of M_1 which contains every vertex of M_1 (and hence every vertex of M). It is easy to see if v is any vertex of T then

$$T \searrow v.$$

Thus, by Lemma 4.3, there is a Morse function on T such that v is the only critical point. By Lemma 4.2, this Morse function can be extended to a Morse function f on M. Since $M - T$ contains no vertices, v is the only critical point of f of index 0. $\square$

Theorem 4.6. *Let M be a connected polyhedron which is topologically a compact n-manifold without boundary. Then M has a polar Morse function. Moreover if v is a vertex of M and σ an n-face, then there is a polar Morse function f such that v is the unique critical point of index 0 and σ is the unique critical point of index n.*

Proof. If $\dim M = 0$ then M must be a single point and the theorem is trivial.

If $\dim M = 1$ then M must be a circle. Let $N = M - \sigma$. Then $N \searrow v$ so N has a Morse function f with v the only critical point. Let

$$c = \max_N f$$

and extend f to M by setting

$$f(\sigma) = c + 1.$$

Then f has the desired properties.

Suppose $\dim M \geq 2$. Let $N = M - \sigma$. Then N is a connected polyhedron which is topologically an n-manifold with boundary. Therefore $N \searrow L$, where L is a subcomplex of dimension $\leq n - 1$ which contains the $n - 2$ skeleton of N. (This can be seen by collapsing the n-faces of N one at a time along free $(n-1)$-faces. For a proof see [Gl] p. 52.) From Lemma 4.5, L has a Morse function f such that v is the only critical point of f of index 0. By Lemma 4.3 f can be extended to N without adding any critical points. Let $c = \max_N f$. Extend f to M by setting

$$f(\sigma) = c + 1.$$

Then f has the desired properties. $\square$

Our last goal of this section is to examine the relationship between Poincaré duality and discrete Morse theory. Recall that if M is a compact n-dimensional PL manifold without boundary then for any field $\mathbf{F}$ and any p, $0 \leq p \leq n$

$$H_p(M, \mathbf{F}) \cong H_{n-p}(M, \mathbf{F}).$$

If M is a smooth manifold this duality is reflected in Morse theory by the observation that if f is a smooth Morse function then $-f$ is also a smooth Morse function. Moreover, $p \in M$ is a critical point of f of index p if and only if p is a critical point of $-f$ of index $n - p$.

Now let us consider the discrete category. If M is a polyhedron and f is a discrete Morse function, then $-f$ is not a Morse function. Of course, a general polyhedron does not satisfy Poincaré duality. However, suppose M is a compact PL n-manifold without boundary. Then one can associate to M a dual polyhedron M^*. That is, M^* is a PL manifold homeomorphic to M. Moreover, for each p, $0 \leq p \leq n$, there is a 1-1 correspondence between $K_p(M)$ and $K_{n-p}(M^*)$. If $\sigma \in K_p(M)$ denote the corresponding $(n-p)$-face of M^* by σ^*. Then

$$\tau > \sigma \longleftrightarrow \tau^* < \sigma^*.$$

The following theorem follows directly from the definitions.

Theorem 4.7. *Let f be a Morse function on a compact PL n-manifold M without boundary. Then $-f$ is a Morse function on M^*. That is, we can define a Morse function g on M^* by setting*

$$g(\sigma^*) = -f(\sigma).$$

Moreover, $\sigma^{(p)} \in K_p(M)$ is critical if and only if $\sigma^{*(n-p)} \in K_{n-p}(M^*)$ is a critical point of $-f$ on M^*. In particular, for all p, $0 \leq p \leq n$

$$m_p(f) = m_{n-p}(-f).$$

§5. Sphere Theorems.

If M is a smooth compact manifold without boundary with a smooth Morse function with exactly 2 critical points then M is homeomorphic to a sphere ([Re], [Mil]). In this section we prove the analogous theorem in the cellular context under various assumptions on M.

Theorem 5.1. *1) Let M be a general CW complex with a Morse function f with exactly 2 critical points. Then M is homotopy equivalent to a sphere.*

2) If, in addition, M is a finite polyhedron which is topologically an n-manifold without boundary then M is homeomorphic to S^n.

3) If, in addition, M is a compact PL n-manifold without boundary then M is piecewise linear equivalent to a PL n-sphere (i.e., M is equivalent to $\dot\Delta^{n+1}$, the boundary of an $(n+1)$-simplex with its standard triangulation).

Remark. Our proof of (2) requires the Poincaré conjecture in dimensions $\neq 3$. However, our proof of (3) does not rely on the Poincaré conjecture, and instead follows from Whitehead's theorem on regular neighborhoods [Wh], as extended by Stallings to polyhedra [St].

Proof. (1) Since $H_0(M, \mathbf{R}) \neq 0$, the Morse inequalities imply that at least one critical point must have index 0. If the other critical point has index p, then Corollary 3.5 (if M is a regular cell complex) and Theorem 10.3 (if M is a general CW complex) imply M is homotopy equivalent to a p-cell with its boundary glued to a point, which is precisely a p-sphere.

(2) As in (1), at least one critical point must have index 0. Moreover, since $H_n(M, \mathbf{Z}/2\mathbf{Z}) \neq 0$ the Morse inequalities imply there must be at least one critical point of index n. Thus, from Corollary 3.5 and Theorem 10.3, M is a homotopy n-sphere. By the resolution of the generalized Poincaré conjecture in dimensions other than 3, if $n \neq 3$ then M is homeomorphic to S^n.

Suppose $n = 3$. Any finite polyhedron which is topologically a 3-manifold without boundary is, in fact, a PL 3-manifold ([Gl] Exercise II.8). Thus in dimension $n = 3$, the result follows from part (3).

(3) Let σ be the critical n-face of f and let $N = M - \sigma$. Then N has only 1 critical point occuring at a vertex v. Hence, by Theorem 3.3 $N \searrow v$. It follows from Whitehead's theorem (Theorem 1.6) that N is a PL n-cell. Clearly $\bar\sigma$ is a PL n-cell. Thus

$$M = N \cup_{\dot\sigma} \bar\sigma$$

is a PL n-sphere. □

The following theorem is a partial converse to Theorem 5.1 (compare with Corollary 4.4).

Theorem 5.2. *Suppose M is a PL n-sphere. Then, by performing a finite sequence of bisections, M can be subdivided to a polyhedron M' which has a Morse function with exactly 2 critical points.*

Proof. If $n = 0$ the theorem is trivially true so assume $n \geq 1$.

Since $M \approx \dot{\Delta}^{n+1}$, M and $\dot{\Delta}^{n+1}$ have isomorphic subdivisions. In fact, these subdivisions can be chosen to result from a finite sequence of bisections of M and $\dot{\Delta}^{n+1}$, respectively. Since $\dot{\Delta}^{n+1}$ has a Morse function with exactly 2 critical points (Corollary 4.4), the theorem follows from Theorem 11.1. $\square$

Theorems 5.1 and 5.2 yield an interesting reformulation of the PL Poincaré conjecture.

Corollary 5.3. *The following 2 statements are equivalent*

1) (PL Poincaré Conjecture) Let M be a PL n-manifold which is a homotopy n-sphere. Then M is a PL n-sphere.

2) Let M be a PL n-manifold which is a homotopy n-sphere. Then, by a series of bisections, M can be subdivided to a complex which has a Morse function with exactly 2 critical points.

§6. The Discrete Gradient Vector Field and the Associated Flow.

In smooth Morse theory, the gradient of the Morse function and the associated flow are essential tools in investigating the relationship between the critical points and the topology of the underlying manifold. In this section we define the gradient vector field and the corresponding flow in the cellular setting. In the following sections, the discrete gradient flow will be used to establish a more precise relationship between the critical points and the topology of the underlying complex than that provided by Theorems 3.3 and 3.4. We emphasize that in this section, and in the rest of the paper, M will denote a general CW complex, i.e., M need not be regular.

Our first goal is to define an object V_f which will represent a discrete analog of the gradient vector field $-\nabla f$, and the associated flow Φ_f. For now, we assume a Morse function f has been fixed, and we write V and Φ, for V_f and Φ_f, respectively. We begin our discussion with the vertices of M. Let $v \in K_0$. If v is a critical point, then any reasonable definition of gradient should vanish at v so we set

$$V(v) = 0.$$

If v is not critical. then there is a unique edge $e > v$ with

$$f(e) \leq f(v).$$

The edge e specifies the unique direction in which f is not increasing so $-\nabla f$ at v should be e. We must now be more precise. We wish to think of $V(v)$ as a discrete tangent vector pointing away from v. (See figure 0.6). That is, $V(v)$ is e with a chosen orientation.

We pause here to introduce the chain complex of M. Fix an orientation for each cell σ of M. Let $C_i(M, \mathbf{Z})$ denote the free abelian group generated over $\mathbf{Z}$ by these orientated cells of M. We now identify $-\sigma$ with σ given the opposite orientation. Let ∂ denote the usual boundary operator

$$\partial : C_p(M, \mathbf{Z}) \longrightarrow C_{p-1}(M, \mathbf{Z}).$$

Then

$$\partial \sigma = \sum_{v^{(p-1)} < \sigma} \epsilon(\sigma, v) v$$

where the ϵ's are integers called the incidence numbers.

It is convenient to introduce an inner product $<,>$ on C_* by declaring the cells of M to be an orthonormal basis. Using the inner product we can write

$$\partial \sigma = \sum_{v^{(p-1)} < \sigma} <\partial \sigma, v> v.$$

We are now ready to complete our definition of V. If $v \in K_0$ is not critical and the edge e satisfies

$$e > v, \quad f(e) \leq f(v)$$

we set

$$V(v) = \pm e$$

24

where the sign is determined so that

$$< \partial(V(v)), v > \; = \; -1.$$

It is now clear how to define the (discrete time) flow Φ on the vertices of $\mathcal{M}$. If v is critical, i.e.. $V(v) = 0$, then v is fixed under the gradient flow, so

$$\Phi(v) = v.$$

If v is not critical, and $V(v) = \pm e$, then v should flow to the "other end" of e. That is,

$$\Phi(v) = v + \partial(V(v)).$$

Note that this formula holds for all vertices, whether critical or not.

We can extend V and Φ linearly to maps (which we also call V and Φ) on chains

$$V : C_0(\mathcal{M}, \mathbf{Z}) \longrightarrow C_1(\mathcal{M}, \mathbf{Z})$$
$$\Phi : C_0(\mathcal{M}, \mathbf{Z}) \longrightarrow C_0(\mathcal{M}, \mathbf{Z}).$$

We now extend V to higher dimensional cells.

Definition 6.1. Let σ be a p-cell of $\mathcal{M}$ (with a fixed orientation). If there is a $\tau^{(p+1)} > \sigma$ with $f(\tau) \leq f(\sigma)$ we set

$$V(\sigma) \; = \; - < \partial \tau, \sigma > \tau.$$

note that σ must be a regular face of τ, so $\langle \partial \tau, \sigma \rangle = \pm 1$).If there is no such τ, then set

$$V(\sigma) = 0.$$

For each p. we extend V linearly to a map

$$V : C_p(\mathcal{M}, \mathbf{Z}) \longrightarrow C_{p+1}(\mathcal{M}, \mathbf{Z}).$$

Note that $V(\sigma) = 0$ does not imply σ is critical. Consider the following example (where the arrows indicate the chosen orientations on the edges).
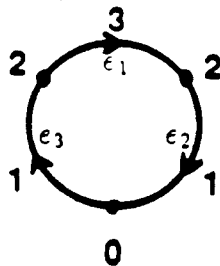

FIGURE 6.1

We observe that $V(e_1) = 0$. However, e_1 should not remain fixed under Φ. The boundary of e_1 moves downward, so e_1 should also. The main point is that for any face σ. $V(\sigma)$ can be thought of as representing the component of $-\nabla f$ which is transversal to σ. The component of $-\nabla f$ which is tangent to σ is determined by $V(\partial \sigma)$.

25

Definition 6.2. For any oriented face σ we define the (discrete time) gradient flow ϕ by

$$\Phi(\sigma) = \sigma + \partial V(\sigma) + V(\partial \sigma)$$

or, more succinctly,

$$\Phi = 1 + \partial V + V \partial$$

(see Figure 0.8).

The main properties of V and Φ are contained in the following theorems.

Theorem 6.3. *1)* $V \circ V = 0$.
2) If σ is an oriented p-cell, then

$$\#\{v^{(p-1)} \mid V(v) = \pm\sigma\} \leq 1.$$

3) If σ is an oriented p-cell of M then

$$\sigma \text{ is critical} \longleftrightarrow [\sigma \notin \text{Image}(V) \quad and \quad V(\sigma) = 0].$$

Proof. 1) If $V(v^{(p-1)}) = \pm\sigma^{(p)}$ then $v < \sigma$ and $f(\sigma) \leq f(v)$. By Lemma 2.5 there is no $\tau^{(p+1)} > \sigma$ with $f(\tau) \leq f(\sigma)$. Thus

$$V(\sigma) = 0.$$

2) If $V(v^{(p-1)}) = \pm\sigma^{(p)}$ then $v < \sigma$ and $f(\sigma) \leq f(v)$. By condition 2(i) of Definition 2.1, v is unique.

3) From Definition 2.2, σ is critical if and only if

    (i) There is no $v^{(p-1)} < \sigma$ with $f(v) \geq f(\sigma)$ and
    (ii) There is no $\tau^{(p+1)} > \sigma$ with $f(\tau) \leq f(\sigma)$.

These conditions are equivalent to

    (i) There is no $v^{(p-1)}$ with $V(v) = \pm\sigma$ and
    (ii) There is no $\tau^{(p+1)}$ with $V(\sigma) = \pm\tau$

i.e.,

    (i) $\sigma \notin \text{Image}(V)$ and
    (ii) $V(\sigma) = 0$. $\square$

Theorem 6.4. *1)* $\Phi\partial = \partial\Phi$.
Let $\sigma_1, \ldots, \sigma_r$ denote the p-cells of M each with a chosen orientation. Write

$$\Phi(\sigma_i) = \sum_j a_{ij}\sigma_j.$$

2) For every i, $a_{ii} = 0$ or 1, and $a_{ii} = 1$ if and only if σ_i is critical.
3) If $i \neq j$ then $a_{ij} \in \mathbf{Z}$. If $i \neq j$ and $a_{ij} \neq 0$ then $f(\sigma_j) < f(\sigma_i)$.

Proof. 1) Using $\Phi = 1 + \partial V + V\partial$ we find

$$\Phi\partial = (1 + V\partial + \partial V)\partial$$
$$= \partial + V\partial^2 + \partial V\partial$$
$$= \partial + \partial V\partial$$
$$\partial\Phi = \partial(1 + V\partial + \partial V)$$
$$= \partial + \partial V\partial + \partial^2 V$$
$$= \partial + \partial V\partial.$$

2) and 3). We prove these simultaneously. First, since ∂ and V both map integer chains to integer chains, each $a_{ij} \in \mathbf{Z}$.

By Theorem 6.3 each cell $\sigma^{(p)}$ satisfies exactly one of the following properties:

   (i) σ is critical.
   (ii) $\pm\sigma \in \text{Image}(V)$.
   (iii) $V(\sigma) \neq 0$.

We examine each case independently.

   (i) If σ is critical, then $V(\sigma) = 0$, so

$$\Phi(\sigma) = \sigma + V(\partial\sigma)$$
$$= \sigma + \sum_{v^{(p-1)} < \sigma} \langle \partial\sigma, v \rangle \, V(v).$$

Since σ is critical. for each $v^{(p-1)} < \sigma$

$$f(v) < f(\sigma).$$

For each such v, either $V(v) = 0$ or $V(v) = \tilde{\sigma}^{(p)}$ with

$$f(\tilde{\sigma}) \leq f(v) < f(\sigma).$$

Thus

$$\Phi(\sigma) = \sigma + \sum a_{\tilde{\sigma}}\tilde{\sigma}$$

where

$$a_{\tilde{\sigma}} \neq 0 \Rightarrow f(\tilde{\sigma}) < f(\sigma).$$

   (ii) Suppose $\pm\sigma \in \text{Image}(V) \subseteq \ker(V)$. Then

$$\Phi(\sigma) = \sigma + V(\partial\sigma)$$
$$= \sigma + \sum_{v^{(p-1)} < \sigma} \langle \partial\sigma, v \rangle \, V(v).$$

By Theorem 6.3 part (2) there is exactly one $(p-1)$-face $\tilde{v} < \sigma$ with

$$V(\tilde{v}) = \pm\sigma$$

27

and

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

Moreover. $\tilde{v} \neq v < \sigma$ implies $V(v) = 0$ or $V(v) = \tilde{\sigma}$ with $f(\tilde{\sigma}) \leq f(v) < f(\sigma)$. Thus

$$\Phi(\sigma) = \sum_{\tilde{\sigma}^{(p)}} a_{\tilde{\sigma}} \tilde{\sigma}$$

where $a_{\tilde{\sigma}} \neq 0$ implies $f(\tilde{\sigma}) < f(\sigma)$.

(iii) Suppose $V(\sigma) = - < \partial \tau, \sigma > \tau \neq 0$. Then

$$\Phi(\sigma) = \sigma + V(\partial \sigma) + \partial(V(\sigma)).$$

Since $V(\sigma) \neq 0. \pm \sigma \notin \text{Image}(V)$. Thus, for each $v^{(p-1)} < \sigma$, either $V(v) = 0$ or $V(v) = \pm \tilde{\sigma}$ where

$$f(\tilde{\sigma}) \leq f(v) < f(\sigma).$$

Moreover.

$$\begin{aligned}
\partial(V(\sigma)) &= - < \partial \tau, \sigma > \partial \tau = - < \partial \tau, \sigma >^2 \sigma + \sum b_{\tilde{\sigma}} \tilde{\sigma} \\
&= -\sigma + \sum b_{\tilde{\sigma}} \tilde{\sigma}
\end{aligned}$$

where $b_{\tilde{\sigma}} \neq 0$ implies $f(\tilde{\sigma}) \leq f(\tau) < f(\sigma)$.

This completes the proof. $\square$

Intuitively. Theorem 6.4 says that Φ decreases f and $\sigma \subseteq \Phi(\sigma)$ if and only if σ is critical. For example. consider again the circle in figure 6.1. In the smooth category, if Φ were the time 1 map corresponding to the flow along $-\nabla f$, where f is the height function, we would have

$$\begin{aligned}
\Phi(e_1) &\supseteq e_1 \\
\Phi(e_2) &\subseteq e_2 \\
\Phi(e_3) &\subseteq e_3.
\end{aligned}$$

In fact, with the combinatorial Morse function f indicated in the figure

$$\begin{aligned}
\Phi(e_1) &= e_1 + e_2 + e_3 \\
\Phi(e_2) &= \Phi(e_3) = 0.
\end{aligned}$$

§7. The Morse Complex and Invariant Chains.

Let $C_p^\Phi(M, \mathbf{Z})$ denote the Φ-invariant p-chains of M. That is,

$$C_p^\Phi(M, \mathbf{Z}) = \{c \in C_p(M, \mathbf{Z}) \mid \Phi(c) = c\}.$$

From Theorem 6.4 (i), the boundary operator ∂ maps C_p^Φ to C_{p-1}^Φ. Thus we have a differential complex

$$(7.1) \qquad \mathcal{C}^\Phi : 0 \longrightarrow C_n^\Phi(M, \mathbf{Z}) \overset{\partial}{\longrightarrow} C_{n-1}^\Phi(M, \mathbf{Z}) \overset{\partial}{\longrightarrow} \cdots \overset{\partial}{\longrightarrow} C_0^\Phi(M, \mathbf{Z}) \longrightarrow 0.$$

The complex $\mathcal{C}^\Phi$ is called the *Morse complex*. For discussions of the Morse complex in the smooth category, see, for example, [Mi2] and [Kl].

The goal of this section is to prove that the homology of $\mathcal{C}^\Phi$ is precisely the homology of M. The first step is to investigate the stabilization map $C_* \to C_*^\Phi$ given by

$$\Phi^\infty = \lim_{N \to \infty} \Phi^N.$$

In fact, the goal of the next 2 results is to show that for N large enough

$$\Phi^\infty = \Phi^N.$$

Lemma 7.1. *Let $c \in C_p^\Phi(M, \mathbf{Z})$ and write*

$$c = \sum_{\sigma \in K_p} a_\sigma \sigma.$$

Let

$$\sigma^* = \text{any maximizer of } \{f(\sigma) \mid a_\sigma \neq 0\}.$$

Then σ^* is a critical cell of f.

Proof. Since c is Φ-invariant

$$c = \Phi(c) = \sum_{\sigma \in K_p} a_\sigma \Phi(\sigma).$$

Therefore,

$$a_{\sigma^*} = <c, \sigma^*> = \sum_{\sigma \in K_p} a_\sigma < \Phi(\sigma), \sigma^* > .$$

From Theorem 6.4 (iii), if $\sigma \neq \sigma^*$ and $f(\sigma) \leq f(\sigma^*)$ then

$$< \Phi(\sigma), \sigma^* > = 0.$$

Thus,

$$0 \neq a_{\sigma^*} = a_{\sigma^*} < \Phi(\sigma^*), \sigma^* >$$

so

$$< \Phi(\sigma^*), \sigma^* > \neq 0.$$

Theorem 6.4 (ii) now implies σ^* is critical. $\square$

Theorem 7.2. *For N large enough $\Phi^N = \Phi^{N+1} = \cdots = \Phi^\infty$.*

Proof. Fix $\sigma \in K$. We will show that for N large enough

$$\Phi^N(\sigma) = \Phi^{N+1}(\sigma) = \cdots = \Phi^\infty(\sigma).$$

The proof is by induction on

$$r = \#\{\tilde\sigma \in K \mid f(\tilde\sigma) < f(\sigma)\}.$$

Suppose $r = 0$, then by Theorem 6.4, either $\Phi(\sigma) = \sigma$ or $\Phi(\sigma) = 0$. In either case $\Phi(\sigma) = \Phi^\infty(\sigma)$.

For general r, suppose first that σ is not critical. Then, by Theorem 6.4

$$\Phi(\sigma) = \sum_{f(\tilde\sigma) < f(\sigma)} a_{\tilde\sigma}\tilde\sigma.$$

By induction on r there is an $\tilde N$ so that $\Phi^{\tilde N}(\tilde\sigma)$ is Φ-invariant whenever $f(\tilde\sigma) < f(\sigma)$. Therefore, $\Phi^{\tilde N+1}(\sigma)$ is Φ-invariant.

Now suppose σ is critical, and let

$$c = V(\partial\sigma).$$

Then

$$\Phi^m(\sigma) = \sigma + c + \Phi(c) + \cdots + \Phi^{m-1}(c).$$

It follows that $\Phi^N(\sigma)$ is Φ-invariant if and only if $\Phi^N(c) = 0$ for some N. As seen in the proof of Theorem 6.4, c is the sum of p-cells $\tilde\sigma$ with $f(\tilde\sigma) < f(c)$. By induction, there is an $\tilde N$ so that $\Phi^{\tilde N}(c)$ is Φ-invariant.

We now observe that $c \in \text{Image}(V)$ and Image (V) is Φ-invariant, since

$$\Phi V = (1 + \partial V + V\partial)V = V(1 + \partial V)$$

(from Theorem 6.3 (i)). Thus, $\Phi^{\tilde N}(c) \in \text{Image}(V)$. From Theorem 6.3 (iii), the image of V is orthogonal to the critical faces. Hence $\Phi^{\tilde N}(c)$ is a Φ-invariant p-chain which is orthogonal to the critical faces and therefore, by Lemma 7.1, $\Phi^{\tilde N}(c) = 0$. That is, for N large enough $\Phi^N(c) = 0$ so $\Phi^N(\sigma)$ is Φ-invariant. $\square$

By Theorem 7.2, there is an N large enough so that for every chain c

$$\Phi^N(c) = \Phi^{N+1}(c) = \Phi^{N+2}(c) = \cdots.$$

Let $\Phi^\infty(c)$ denote this Φ-invariant chain. Then for each p we have maps

$$\Phi^\infty : C_p(M, \mathbf{Z}) \longrightarrow C_p^\Phi(M, \mathbf{Z})$$
$$i : C_p^\Phi(M, \mathbf{Z}) \hookrightarrow C_p(M, \mathbf{Z})$$

where i is the natural inclusion. Note that $\Phi^\infty \circ i$ is the identity on $C_p^\Phi(M, \mathbf{Z})$.

We are now ready to state and prove the main theorem of this section

Theorem 7.3. *Let C_*^Φ denote the Morse complex (7.1). Then for each p*

$$H_p(C_*^\Phi) \cong H_p(M, \mathbf{Z}).$$

Proof. Consider the following commutative diagram

$$0 \longrightarrow C_n(M, \mathbf{Z}) \xrightarrow{\partial} C_{n-1}(M, \mathbf{Z}) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C_0(M, \mathbf{Z}) \longrightarrow 0$$

$$\downarrow \Phi^\infty \quad \uparrow i \qquad \downarrow \Phi^\infty \quad \uparrow i \qquad \qquad \downarrow \Phi^\infty \quad \uparrow i$$

$$0 \longrightarrow C_n^\Phi(M, \mathbf{Z}) \xrightarrow{\partial} C_{n-1}(M, \mathbf{Z}) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C_0(M, \mathbf{Z}) \longrightarrow 0.$$

Let

$$\Phi_*^\infty : \quad H_*(M, \mathbf{Z}) \longrightarrow H_*(C_*^\Phi)$$
$$i_* : \quad H_*(C_*^\Phi) \longrightarrow H_*(M, \mathbf{Z})$$

denote the induced maps on homology. Our goal is to show that i_* and Φ_*^∞ are isomorphisms. In fact, we will see that they are inverses of each other. Since $\Phi^\infty \circ i = 1$ it follows that

$$1 = (\Phi^\infty \circ i)_* = \Phi_*^\infty \circ i_*.$$

To see that $i_* \circ \Phi_*^\infty = 1$ it is sufficient to find an operator

$$L : C_*(M, \mathbf{Z}) \longrightarrow C_{*-1}(M, \mathbf{Z})$$

such that

$$1 - i \circ \Phi^\infty = \partial L + L \partial$$

since then $1 - i \circ \Phi^\infty$ would map closed forms to exact forms and would thus be the zero map on homology. Since i is the identity map on chains and $\Phi^\infty = \Phi^N$ for some N large enough

$$1 - i \circ \Phi^\infty = 1 - \Phi^N = (1 - \Phi)(1 + \Phi + \cdots + \Phi^{N-1})$$
$$= (-\partial V - V\partial)(1 + \Phi + \cdots + \Phi^{N-1})$$
$$= \partial[-V(1 + \Phi + \cdots + \Phi^{N-1})] + [-V(1 + \Phi + \cdots + \Phi^{N-1})]\partial$$

(we have used Theorem 6.3 (i)). Let

$$L = -V(1 + \Phi + \cdots + \Phi^{N-1}). \quad \Box$$

§8. The Morse Complex and Critical Points.

In section 7 we defined the Morse complex, built out of the Φ-invariant chains of $\mathcal{M}$, which has the same integer homology as $\mathcal{M}$. In this section we will show that the Morse complex can be expressed directly in terms of the critical cells. The first goal of this section is to prove that the space of Φ-invariant chains is canonically isomorphic to the span of the critical cells. The second goal is to express the boundary operator of the Morse complex more explicitly in terms of the critical points.

For each p, let $\mathcal{M}_p$ denote the span of the critical p-cells, i.e.

$$\mathcal{M}_p = \left\{ \sum_{\sigma \in K_p} a_\sigma \sigma \mid a_\sigma \in \mathbf{Z} \text{ and } a_\sigma \neq 0 \Rightarrow \sigma \text{ critical} \right\}.$$

By restricting the map Φ^∞ defined in section 7, we get a map

(8.1) $$\Phi^\infty : \mathcal{M}_p \longrightarrow C_p^\Phi(M, \mathbf{Z}).$$

Fix an orientation for each p-cell σ and identify $-\sigma$ with σ given the opposite orientation.

Lemma 8.1. *Let σ be a critical p-cell. If $\tilde{\sigma} \neq \sigma$ is critical, then*

$$< \Phi^\infty(\sigma), \tilde{\sigma} > = 0.$$

Proof. As seen in the proof of Lemma 7.1

$$\Phi^\infty(\sigma) = \sigma + c$$

where $c \in \text{Image}(V) \subseteq \mathcal{M}_p^\perp$. This proves the lemma. $\square$

Theorem 8.2. *The map (8.1) is an isomorphism.*

Proof. (Onto) Suppose $c \in C_p^\Phi(M, \mathbf{Z})$, and let

$$\tilde{c} = \sum_{\sigma \text{ critical}} < c, \sigma > \sigma \in \mathcal{M}_p.$$

We shall see that

$$\Phi^\infty(\tilde{c}) = c.$$

It follows from Lemma 8.1 that for any critical σ

$$< \Phi^\infty(\tilde{c}), \sigma > = < c, \sigma > .$$

Therefore, $\Phi^\infty(\tilde{c}) - c$ is a Φ-invariant chain such that for any critical σ

$$< \Phi^\infty(\tilde{c}) - c, \sigma > = 0.$$

It now follows from Lemma 7.1 that

$$\Phi^\infty(\tilde{c}) - c = 0.$$

(One-to-one). Suppose $c \in \mathcal{M}_p$ satisfies $\Phi^\infty(c) = 0$. Then, for any σ critical

$$< \Phi^\infty(c), \sigma > = 0.$$

It follows from Lemma 8.1 that for any σ critical

$$< c, \sigma > = 0$$

which implies $c = 0$. $\square$

Theorem 8.2 implies that the Morse complex is isomorphic to

$$(8.2) \qquad \mathcal{M} : 0 \longrightarrow \mathcal{M}_n \xrightarrow{\partial} \mathcal{M}_{n-1} \xrightarrow{\partial} \cdots \xrightarrow{\partial} \mathcal{M}_0 \longrightarrow 0$$

where, for $c \in \mathcal{M}_p$, σ a critical $(p-1)$-face

$$(8.3) \qquad < \widetilde{\partial} c, \sigma > \; = \; < \partial \Phi^\infty c, \sigma > \; = \; < \Phi^\infty \partial c, \sigma >$$

(by Theorem 6.3 (i)). Since $H_*(\mathcal{M}) \cong H_*(M, \mathbf{Z})$, we learn from the Universal Coefficient Theorem that for any field $\mathbf{F}$,

$$H_*(\mathcal{M} \otimes \mathbf{F}) \cong H_*(\mathcal{M}) \otimes \mathbf{F} \cong H_*(M, \mathbf{Z}) \otimes \mathbf{F} \cong H_*(M, \mathbf{F}).$$

Thus, $\mathcal{M} \otimes \mathbf{F}$ is a differential complex of vector spaces over $\mathbf{F}$ with the same holmology as M. Since $\dim_{\mathbf{F}} \mathcal{M}_p \otimes \mathbf{F} = m_p(f)$, it follows from standard linear algebra that

Corollary 8.3. *If M is a finite CW complex, f is a discrete Morse function on M and $\mathbf{F}$ is any coefficient field, then the Strong Morse Inequalites (as stated in Corollary 3.6) and hence the Weak Morse Inequalities (as stated in Corollary 3.7) hold.*

Our goal now is to find a more convenient expression for $\widetilde{\partial}$. Namely, we will show that for $\tau^{(p+1)}$ and $\sigma^{(p)}$ critical, $< \widetilde{\partial} \tau, \sigma >$ can be expressed in terms of *gradient paths* from $\partial \tau$ to σ.

Definition 8.4. *A gradient path of dimension p is a sequence γ of p-cells of M*

$$\gamma \; = \; \sigma_0, \sigma_1, \sigma_2, \ldots, \sigma_r$$

such that for every $i = 0, \ldots, r-1$
  (i) if $V(\sigma_i) = 0$ then $\sigma_{i+1} = \sigma_i$.
  (ii) if $V(\sigma_i) \neq 0$ then $\sigma_{i+1} < V(\sigma_i)$ and $\sigma_{i+1} \neq \sigma_i$.
We say γ is a *gradient path from* σ_0 *to* σ_r, and the *length of* γ, denoted by $|\gamma|$, is equal to r.

We record in the following lemma two easily verified properties of gradient paths

Lemma 8.5. *(i) If $\gamma = \sigma_0, \sigma_1, \ldots, \sigma_s$ is a gradient path then for each $i = 0, 1, \ldots, s-1$ either $\sigma_i = \sigma_{i+1}$ or $f(\sigma_i) > f(\sigma_{i+1})$.*
  (ii) If $\gamma_1 = \sigma_0, \sigma_1, \ldots, \sigma_r$ and $\gamma_2 = \sigma_r, \sigma_{r+1}, \ldots, \sigma_{r+s}$ are two sequences of p-cells, then

$$\sigma_0, \sigma_1, \ldots, \sigma_r, \sigma_{r+1}, \ldots, \sigma_{r+s}$$

is a gradient path if and only if both γ_1 and γ_2 are gradient paths.

Suppose $\sigma \neq \bar{\sigma}$ are two p-cells of M and τ is a $(p+1)$-cell with $\sigma < \tau$ and $\bar{\sigma} < \tau$ and both are regular faces. Then an orientation on σ induces an orientation on $\bar{\sigma}$ in the following way. An orientation on σ induces an orientation on τ (so that $< \partial \tau, \sigma > = -1$). Given
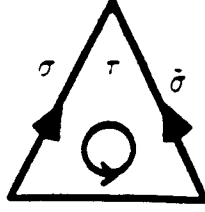
FIGURE 8.1

the orientation on τ, we choose the orientation on $\bar\sigma$ so that $< \partial\tau, \bar\sigma > = 1$. Equivalently, fixing an orientation on σ and τ, an orientation is induced on $\bar\sigma$ so that

$$< \partial\tau, \sigma >< \partial\tau, \bar\sigma > = -1.$$

Loosely speaking, we induce an orientation on $\bar\sigma$ by "sliding" σ across τ to $\bar\sigma$. In figure 8.1, an orientation is shown for σ and the induced orientations on τ and $\bar\sigma$ are indicated.

On the other hand, if $\sigma = \bar\sigma$, then an orientation on σ induces the same orientation on $\bar\sigma$. Thus, if $\gamma = \sigma_0, \sigma_1, \ldots, \sigma_r$ is a gradient path an orientation on σ_0 induces an orientation on each σ_i in turn, and, in particular, on σ_r. Recall that we have fixed an orientation for each face of M. Write $m(\gamma) = 1$ if the fixed orientation on σ_0 induces the fixed orientation on σ_r, and $m(\gamma) = -1$ otherwise. Equivalently,

$$(8.4) \qquad m(\gamma) = \prod_{\substack{i=0 \\ V(\sigma_i)\neq 0}}^{r-1} - < \partial V(\sigma_i), \sigma_i >< \partial V(\sigma_i), \sigma_{i+1} > .$$

We can use this formula to define the multiplicity of any gradient path.

Definition 8.6. Let

$$\gamma = \sigma_0, \sigma_1, \ldots, \sigma_r$$

denote a gradient path of dimension p. We define the *multiplicity* of γ, $m(\gamma)$, by the formula (8.4).

Note that if $\gamma_0 = \sigma_0, \ldots, \sigma_r$ and $\gamma_1 = \sigma_r, \ldots, \sigma_{r+s}$ are gradient paths, then

$$(8.5) \qquad m(\gamma_1)m(\gamma_0) = m(\gamma_1 \circ \gamma_0)$$

where

$$\gamma_1 \circ \gamma_0 = \sigma_0, \ldots, \sigma_r, \ldots, \sigma_{r+s}.$$

For p-cells σ and $\bar\sigma$, let $\Gamma_r(\sigma, \bar\sigma)$ denote the set of all gradient paths from σ to $\bar\sigma$ of length r. The remainder of this section is devoted to proving that if $\tau^{(p+1)}$ and $\sigma^{(p)}$ are critical, then

$$< \tilde\partial\tau, \sigma > = \sum_{\bar\sigma^{(p)} < \tau} < \partial\tau, \bar\sigma > \sum_{\gamma \in \Gamma_N(\bar\sigma, \sigma)} m(\gamma)$$

for any N large enough.

Definition 8.7. Define a *reduced gradient flow*

$$\widetilde{\Phi} : C_p(M, \mathbf{Z}) \longrightarrow C_p(M, \mathbf{Z})$$

by

$$\widetilde{\Phi} = 1 + \partial V.$$

Note that $\widetilde{\Phi}$ does not share most of the desirable properties of Φ listed in Theorem 6.4. However, $\widetilde{\Phi}$ is simpler to work with, and can be substituted into (8.3).

Lemma 8.8. *For any critical faces* $\tau^{(p+1)}$ *and* $\sigma^{(p)}$

$$< \widetilde{\partial}\tau, \sigma > \ = \ < \widetilde{\Phi}^{\infty}\partial\tau, \sigma > .$$

Proof. It is sufficient to prove that for every $r \geq 0$

$$< \widetilde{\Phi}^r \partial\tau, \sigma > \ = \ < \Phi^r \partial\tau, \sigma > .$$

This follows from the observation that for every chain c and every $r \geq 0$

$$\Phi^r(c) - \widetilde{\Phi}^r(c) \in \mathrm{Image}(V) \subsetneq \mathcal{M}_\bullet^\perp.$$

We prove this by induction on r. For $r = 0$ there is nothing to prove. For general r

$$\Phi^r(c) = \Phi(\Phi^{r-1}(c)) = \Phi(\widetilde{\Phi}^{r-1}(c) + V(\bar{c}))$$

for some chain $\bar{c}$ (by the inductive hypothesis)

$$
\begin{aligned}
&= (\widetilde{\Phi} + V\partial)(\widetilde{\Phi}^{r-1}(c) + V(\bar{c})) \\
&= \widetilde{\Phi}^r(c) + \widetilde{\Phi}(V(\bar{c})) + V\partial\widetilde{\Phi}^{r-1}(c) + V\partial V(\bar{c}) \\
&= \widetilde{\Phi}^r(c) + V(\bar{c} + \partial\widetilde{\Phi}^{r-1}(c) + V\partial V(\bar{c}))
\end{aligned}
$$

where the last equality follows from

$$\widetilde{\Phi}V = V + \partial V^2 = V. \quad \square$$

Lemma 8.9. *For any* $\sigma_1^{(p)}, \sigma_2^{(p)} \in K_p$

$$(8.6) \qquad < \widetilde{\Phi}\sigma_1, \sigma_2 > \ = \ \sum_{\gamma \in \Gamma_1(\sigma_1, \sigma_2)} m(\gamma).$$

Proof. First suppose $V(\sigma_1) = 0$. Then

$$< \tilde{\Phi}\sigma_1, \sigma_2 > \; = \; < \sigma_1, \sigma_2 > \; = \; \begin{cases} 1 & \sigma_1 = \sigma_2 \\ 0 & \sigma_1 \neq \sigma_2. \end{cases}$$

On the other hand, the only gradient path of length 1 beginning at σ, is the trivial gradient path $\gamma = \sigma_1, \sigma_1$ so that $m(\gamma) = 1$. Thus, if $\sigma_1 \neq \sigma_2$, $\Gamma_1(\sigma_1, \sigma_2) = \emptyset$ so

$$\sum_{\gamma \in \Gamma_1(\sigma_1, \sigma_2)} m(\gamma) = 0$$

and if $\sigma_1 = \sigma_2$, $\Gamma_1(\sigma_1, \sigma_2) = \gamma$ so

$$\sum_{\gamma \in \Gamma_1(\sigma_1, \sigma_2)} m(\gamma) = 1.$$

Now suppose $V(\sigma_1) \neq 0$. If $\sigma_1 = \sigma_2$ we calculate the left hand side of (8.6) to find

$$\begin{aligned} < \tilde{\Phi}\sigma_1, \sigma_1 > \; &= \; < \sigma_1, \sigma_1 > + < \partial V(\sigma_1), \sigma_1 > \\ &= \; 1 - 1 \; = \; 0. \end{aligned}$$

On the other hand, since $V(\sigma_1) \neq 0$ there is no gradient path of length 1 from σ_1 to σ_1 so $\Gamma_1(\sigma_1, \sigma_1) = \emptyset$ and

$$\sum_{\gamma \in \Gamma_1(\sigma_1, \sigma_1)} m(\gamma) = 0.$$

Now suppose $\sigma_1 \neq \sigma_2$. Then

$$\begin{aligned} < \tilde{\Phi}\sigma_1, \sigma_2 > \; &= \; < \sigma_1, \sigma_2 > + < \partial V(\sigma_1), \sigma_2 > \\ &= \; < \partial V(\sigma_1), \sigma_2 > . \end{aligned}$$

If σ_2 is not a face of $V(\sigma_1)$ then $< \tilde{\Phi}\sigma_1, \sigma_2 > = 0$. Moreover, in this case there are no gradient paths of length 1 from σ_1 to σ_2 so that

$$\sum_{\gamma \in \Gamma_1(\sigma_1, \sigma_2)} m(\gamma) = 0.$$

If σ_2 is a face of $V(\sigma_1)$ then there is exactly one gradient path from σ_1 to σ_2 of length 1, namely $\gamma = \sigma_1, \sigma_2$ and

$$\begin{aligned} m(\gamma) \; &= \; - < \sigma_1, \partial V(\sigma_1) > < \partial V(\sigma_1), \sigma_2 > \\ &= \; < \partial V(\sigma_1), \sigma_2 > \end{aligned}$$

as desired. $\quad\square$

Theorem 8.10. *For any critical cells $\tau^{(p+1)}$ and $\sigma^{(p)}$*

$$< \tilde{\partial}\tau, \sigma > = \sum_{\tilde{\sigma}^{(p)} < \tau} < \partial\tau, \tilde{\sigma} > \sum_{\gamma \in \Gamma_N(\tilde{\sigma}, \tilde{\sigma})} m(\gamma)$$

for N large enough.

Proof. From Lemma 8.8

$$< \tilde{\partial}\tau, \sigma > = < \tilde{\Phi}^N \partial\tau, \sigma >$$

for N large enough. Since

$$\partial\tau = \sum_{\tilde{\sigma}^{(p)} < \tau} < \partial\tau, \tilde{\sigma} > \tilde{\sigma}$$

we find

$$< \tilde{\partial}\tau, \sigma > = \sum_{\tilde{\sigma}^{(p)} < \tau} < \partial\tau, \tilde{\sigma} >< \tilde{\Phi}^N \tilde{\sigma}, \sigma >$$

for N large enough. We now prove inductively that for all $r \geq 0$

$$< \tilde{\Phi}^r \tilde{\sigma}, \sigma > = \sum_{\gamma \in \Gamma_r(\tilde{\sigma}, \sigma)} m(\gamma).$$

The case $r = 0$ is trivial, and $r = 1$ is Lemma 8.9. For general $r > 1$

$$
\begin{aligned}
< \tilde{\Phi}^r \tilde{\sigma}, \sigma > &= < \tilde{\Phi}(\tilde{\Phi}^{r-1}\tilde{\sigma}), \sigma > \\
&= \sum_{\sigma'^{(p)}} < (\tilde{\Phi}^{r-1}\tilde{\sigma}), \sigma' >< \tilde{\Phi}\sigma', \sigma > \\
&= \sum_{\sigma'} \sum_{\gamma \in \Gamma_{r-1}(\tilde{\sigma}, \sigma')} m(\gamma) < \tilde{\Phi}\sigma', \sigma > \\
&\quad \text{(by induction)} \\
&= \sum_{\sigma'} \sum_{\gamma \in \Gamma_{r-1}(\tilde{\sigma}, \sigma')} m(\gamma) \sum_{\gamma' \in \Gamma_1(\sigma', \sigma)} m(\gamma') \\
&\quad \text{(by Lemma 8.8)} \\
&= \sum_{\gamma \in \Gamma_r(\tilde{\sigma}, \sigma)} m(\gamma)
\end{aligned}
$$

(by Lemma 8.6 (ii) and (8.5)). $\square$

§9. Self-Indexing Morse Functions, Cellular Triads, and a Characterization of Gradient Vector Fields.

When manipulating Morse functions, one finds it is often convenient to manipulate instead the corresponding gradient vector field. It is important to know that after varying the vector field, one is left with the gradient vector field of some Morse function. With this in mind, our goal in this section is to characterize gradient vector fields of Morse functions. Along the way, we sha. .rove that every Morse function can be replaced with a particularly nice (i.e., self-indexing) Morse function with the same critical points. In this section we also introduce the notion of a cellular triad, which will play the role of a cellular cobordism in later sections. The results in this section are discrete analogues of those in [Sm1].

Definition 9.1. A *discrete vector field* is a map

$$W : K \longrightarrow K \cup \{0\}$$

satisfying

(1) For each p, $W(K_p) \subseteq K_{p+1} \cup \{0\}$
(2) For each $\sigma^{(p)} \in K_p$, either $W(\sigma) = 0$ or σ is a regular face of $W(\sigma)$.
(3) If $\sigma \in \text{Image}(W)$ then $W(\sigma) = 0$
(4) For each $\sigma^{(p)} \in K_p$

$$\#\{v^{(p-1)} \in K_{p-1} \mid W(v) = \sigma\} \leq 1.$$

(When M is a simplicial complex such objects have previously been considered under a different name, in [Du] and [Sta]. See also [Fo1] for discrete vector fields on 1-dimensional complexes).

This definition is not quite consistent with our definition of the discrete gradient vector field V_f in §6, in that V_f was defined to be a map of oriented cells. However, ignoring orientations, V_f gives rise to a discrete vector field as defined here. Conversely, if W is a discrete vector field, then by properties (2) and (4) if $W(\sigma^{(p)}) = \tau^{(p+1)}$ then σ is a regular face of τ. Thus, we can consider W to be defined on oriented cells by endowing σ and τ with orientations and setting $W(\sigma) = \pm\tau$, with the sign chosen so that

$$(9.1) \qquad\qquad < \sigma, \partial W(\sigma) > \ = \ -1.$$

With this in mind, we will frequently abuse notation and write expressions of the form $W = V_f$. This means either that if one ignores orientations then $W = V_f$ as maps of unoriented cells, or if one extends W to oriented cells via (9.1) then $W = V_f$ as maps of oriented cells. These two points of view are equivalent.

Definition 9.2. Let W be a combinatorial vector field. A *W-path of dimension p* is a sequence of p-cells

$$\gamma = \sigma_0, \sigma_1, \ldots, \sigma_r$$

such that

(i) if $W(\sigma_i) = 0$ then $\sigma_{i+1} = \sigma_i$
(ii) if $W(\sigma_i) \neq 0$ then $\sigma_{i+1} \neq \sigma_i$ and $\sigma_{i+1} < W(\sigma_i)$.

Say γ is a *closed path* if $\sigma_r = \sigma_0$ and γ is *non-stationary* if $\sigma_1 \neq \sigma_0$.

Theorem 9.3. *Let W be a discrete vector field. There is a discrete Morse function f with $W = V_f$ if and only if W has no non-stationary closed paths.*

Moreover, for every such W, f can be chosen to have the property that if $\sigma^{(p)}$ is critical, then

$$f(\sigma) = p.$$

Such a Morse function is said to be self-indexing.

Remark. Note that W determines the critical points of f. Namely, if $W = V_f$, then, by Lemma 6.3, $\sigma \in K_p$ is critical for f if and only if $W(\sigma) = 0$ and $\sigma \notin \text{Image}(W)$.

We will not prove Theorem 9.3, as it will follow directly from the more general Theorem 9.10. We pause here to note a corollary.

Corollary 9.4. *Let f be a discrete Morse function. Then there is a discrete self-indexing Morse function $\tilde{f}$ with the same discrete gradient vector field, and hence the same critical points.*

Proof. Given f, let $W = V_f$. By Theorem 9.3, $W = V_{\tilde{f}}$ for some self-indexing Morse function $\tilde{f}$.

Before going futher, we introduce the notion of a cellular triad.

Definition 9.5. A *cellular triad* (M, N^-, N^+) consists of a CW complex M and two disjoint subcomplexes N^- and N^+ with the property that for every p-cell $\sigma^{(p)}$ of $N^- \cup N^+$ there is exactly one $(p+1)$-cell $\tau^{(p+1)} \not\subseteq N^- \cup N^+$ satisfying $\tau > \sigma$. Moreover, we require that σ is a regular face of τ.

For example, if M is any CW complex, we can set $N^- = N^+ = \emptyset$. For a less trivial example, suppose M is a PL n-manifold such that $\dot{M}$ can be written as a disjoint union $\dot{M} = N^- \cup N^+$ where N^- and N^+ are subpolyhedra. Then (M, N^-, N^+) is a cellular triad. This is the most important class of examples, and such a triad will be called a *polyhedral triad*.

Definition 9.6. Let (M, N^-, N^+) be a cellular triad. A *discrete Morse function f on (M, N^-, N^+) with image $[a, b]$* is a function

$$f : K(M) \longrightarrow [a, b]$$

satisfying
  (1) $f^{-1}(a) = N^-$, $f^{-1}(b) = N^+$
  (2) For all $\sigma^{(p)} \in K_p(M)$, $\sigma \not\subseteq N^- \cup N^+$,

  (i) If σ is an irregular face of $\tau^{(p+1)}$ then $f(\tau) > f(\sigma)$. Moreover,

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

  (ii) If $v^{(p-1)}$ is an irregular face of σ then $f(v) < f(\sigma)$. Moreover,

$$\#\{v^{(p-1)} < \sigma \mid f(v) \geq f(\sigma)\} \leq 1.$$

Let f be a Morse function on (M, N^-, N^+). We call $\sigma^{(p)} \in K_p(M)$ a *critical point* of f if

(1) $\sigma^p \not\subseteq (N^- \cup N^+)$

(2) (i) $\#\{\tau^{(p+1)}\sigma \mid f(\tau) \leq f(\sigma)\} = 0$

    (ii) $\#\{v^{(p-1)} < \sigma \mid f(v) \geq f(\sigma)\} = 0$.

Definition 9.7. Let (M, N_-, N_+) be a cellular triad, and f a discrete Morse function on (M, N_-, N_+). We will now define the discrete gradient vector field V_f of f. Suppose that each cell of M has been endowed with an orientation. If $\sigma^{(p)} \not\subseteq N^- \cup N^+$ then we define $V_f(\sigma)$ as in Definition 6.1. If $\sigma \subseteq N^-$ then we set $V_f(\sigma) = 0$. If $\sigma \subseteq N^+$ then we set $V_f(\sigma) = \pm \tau^{(p+1)}$ where τ is the unique $(p+1)$ cell not in $N^- \cup N^+$ which has σ as a face (with the sign chosen as in Definition 6.1).

We observe that it is still true, in this more general setting, that if f has no critical points then $M \searrow N^-$. If (M, N^-, N^+) is a polyhedral triad then we can apply Whitehead's Theorem of Regular Neighborhoods ([Wh], see Theorem 1.6 for a special case) to learn

Theorem 9.8. *If (M, N^-, N^+) is a polyhedral triad, and f is a Morse function on (M, N^-, N^+) with no critical points, then*

$$M = N^- \times I$$

where I is a closed interval.

This is the crucial theorem which makes possible a Morse theoretic proof of the s-cobordism theorem.

Much of the previous chapters can be generalized to this setting. In particular, one can define the Morse complex $\mathcal{M}$ associated to a Morse function f on a cellular triad (M, N^-, N^+) exactly as in §8, with the differential $\widetilde{\partial}$ defined as in Theorem 8.10. Following the proofs of the preceeding sections we learn that in this case, the homology of $\mathcal{M}$ is not the homology of M, but rather the homology of M relative to N^-.

Theorem 9.9. *Let f be a Morse function on a cellular triad (M, N^-, N^+) and define the Morse complex as in (8.2) and Theorem 8.10. Then*

$$H_*(\mathcal{M}) \cong H_*(M, N^-, \mathbf{Z}).$$

We now present the main theorem of this section

Theorem 9.10. *Let (M, N^-, N^+) be a cellular triad, and*

$$W : K(M) \longrightarrow K(M) \cup \{0\}$$

a discrete vector field on M satisfying

(1) W *has no non-stationary closed paths*

(2) *For all $\sigma^{(p)} \subseteq N^-$, $W(\sigma) = 0$*

(3) *For all $\sigma^{(p)} \subseteq N^+$, $0 \neq W(\sigma) \not\subseteq N^- \cup N^+$.*

Then there is a Morse function f on the triad (M, N^-, N^+) such that $W = V_f$. Moreover f can be chosen to be self-indexing. That is, we can choose f to have image $[-1/2, n+1/2]$ so that

(1) $f^{-1}(-1/2) = N^-$

(2) $f^{-1}(n + 1/2) = N^+$

(3) *If $\sigma^{(p)}$ is a critical point of f then $f(\sigma) = p$.*

Proof. The proof is by induction on the skeleta of M. For each p, let M_p denote the p-skeleton of M. We restrict W to a combinatorial vector field W_p on M_p by setting

$$W_p(\sigma^{(l)}) = \begin{cases} W(\sigma^{(l)}) & \text{if } l < p \\ 0 & \text{if } l = p. \end{cases}$$

Define

$$N_p^- = N^- \cap M_{p-1}$$
$$N_p^+ = N^+ \cap M_{p-1}.$$

Then for each p, (M_p, N_p^-, N_p^+) is a cellular triad and W_p satisfies the hypotheses of the theorem.

We will prove inductively that there is a Morse function f_p on (M_p, N_p^-, N_p^+) with image $[-1/2, p - 1/2]$ satisfying

(1) $f^{-1}(-1/2) = N_p^-$

(2) $f^{-1}(p + 1/2) = N_p^+$

(3) If $\sigma^{(l)}$ is critical for f_p then $f_p(\sigma) = l$

(4) $W_p = V_{f_p}$.

$\mathbf{p = 0}$: In this case, M_0 is the 0-skeleton of M, W_0 maps every vertex to 0, and $N_0^- = N_0^+ = \emptyset$. The function f_0, where

$$f_0(v) = 0$$

for every vertex v, satisfies the desired properties.

General $\mathbf{p}$: Suppose we have a Morse function f_{p-1} defined on the triad $(M_{p-1}, N_{p-1}^-, N_{p-1}^+)$ with image $[-1/2, p - 1/2]$ satisfying

(1) $f_{p-1}^{-1}(1/2) = N_{p-1}^-$

(2) $f_{p-1}^{-1}(p - 1/2) = N_{p-1}^+$

(3) If $\sigma^{(l)} \subseteq M_{p-1}$ is critical for f_{p-1} then $f_{p-1}(\sigma) = l$

(4) $W_{p-1} = V_{f_{p-1}}$.

Our goal is to extend f_{p-1}, after minor modifications, to a suitable function on M_p.

Without loss of generality, we now assume $\dim M = p$, so that $(M_p, N_p^-, N_p^+) = (M, N^-, N^+)$ and $W_p = W$.

We now define f on M. For $l \leq p - 2$, set

$$f(\sigma^{(l)}) = \begin{cases} f_{p-1}(\sigma^{(l)}) & \text{if } \sigma^{(l)} \not\subseteq N^+ \\ p + \frac{1}{2} & \text{if } \sigma^{(l)} \subseteq N^+. \end{cases}$$

We can not define f on $(p-1)$-cells in such a straight forward manner. For suppose $W(\sigma^{(p-1)}) = \tau^{(p)}$. Then f must satisfy $f(\sigma) > f(\tau)$. Moreover, if $\bar{\sigma}^{(p-1)} \neq \sigma$ is any other $(p-1)$-face of τ we must have $f(\tau) > f(\bar{\sigma})$. In particular, f must satisfy $f(\sigma) > f(\bar{\sigma})$.

For each $\sigma^{(p-1)}$ define

$$d(\sigma) = \max\{r \mid \exists \text{ a } W \text{ path } \sigma, \sigma_1, \sigma_2, \ldots, \sigma_{r-1}, \sigma_r \text{ with } \sigma_{r-1} \neq \sigma_r \text{ and } W(\sigma_r) = 0\}$$
$$D = \max_{\sigma^{(p-1)}} d(\sigma).$$

Note that $d(\sigma)$ must be finite since W has no closed paths. Now set

$$f(\sigma^{(p-1)}) = \begin{cases} f_{p-1}(\sigma) + \frac{d(\sigma)}{2D+1} & \text{if } \sigma \not\subseteq (N^- \cup N^+) \\ p + \frac{1}{2} & \text{if } \sigma \subseteq N^+ \\ -\frac{1}{2} & \text{if } \sigma \subseteq N^-. \end{cases}$$

Note that $f_{p-1}(\sigma) \leq p - 1/2$ so if $\sigma \not\subseteq N^+$ then $f(\sigma) < p$.

We will now define f on the p-cells of M. Since for all $\sigma^{(p)}$ we have $W(\sigma) = 0$, $\sigma^{(p)}$ is critical if and only if $\sigma \notin \text{Image}(W)$. Thus, if $\sigma^{(p)} \notin \text{Image}(W)$ we set

$$f(\sigma) = p.$$

On the other hand, suppose $\sigma^{(p)} = W(v^{(p-1)})$. We will first see that if $\bar{v}^{(p-1)}$ is any other $(p-1)$-cell of σ then

$$(9.2) \qquad\qquad\qquad\qquad f(v) > f(\bar{v}).$$

Note that $\bar{v} \not\subseteq N^+$, since otherwise, by Definition 9.5, there is a unique p-cell $\bar{\sigma}$ such that $\bar{\sigma} \not\subseteq (N^- \cup N^+)$ and $\bar{\sigma} > \bar{v}$. Thus we must have $\bar{\sigma} = \sigma$. However, by hypothesis (3) of the theorem we must have $W(\bar{v}) = \sigma = W(v)$. This contradicts condition (iii) of Definition 9.1.

We observe that (9.2) is certainly true if $v \subseteq N^+$ since then $f(\bar{v}) < p + 1/2 = f(v)$. Suppose $v \not\subseteq N^+$. If $\bar{v}, v_1, \ldots, v_r$ is any W-path then $v, \bar{v}, v_1, \ldots, v_r$ is a W-path. Thus $d(v) > d(\bar{v})$, which implies (9.2). Now set $f(\sigma) = f(v)$.

We will now prove that f satisfies the desired properties. We must first check that f is, in fact, a Morse function on (M, N^-, N^+) with image $[-1/2, p + 1/2]$. By construction, $f^{-1}(N^-) = -1/2$, $f^{-1}(N^+) = p + 1/2$, and if $\sigma \not\subseteq N^- \cup N^+$, $f(\sigma) \in (-1/2, p + 1/2)$. This is condition (1) of Definition 9.6.

We now check condition (2) of Definition 9.6. Let $\sigma^{(l)} \in K_l(M)$. If $\sigma \subseteq (N^- \cup N^+)$ then 2(i) and 2(ii) are clearly true since $v < \sigma \Rightarrow v \subseteq N^- \cup N^+$ and there is exactly 1 $(l-1)$-face τ with $\tau > \sigma$ and $\tau \not\subseteq (N^- \cup N^+)$.

Suppose $\sigma \not\subseteq (N^- \cup N^+)$. If $l = p$ then 2(i) is vacuously true and 2(ii) is true by construction.

The following observation will simplify the remainder of the proof. Suppose $l < p$, $\sigma^{(l)} \not\subseteq N^+$, and $v^{(l-1)} < \sigma$. Then

$$(9.3) \qquad\qquad f(\sigma) < f(v) \longleftrightarrow f_{p-1}(\sigma) < f_{p-1}(v).$$

The proof is simple, but must be broken into a few cases.

If $v \subseteq N^+$ then $f(\sigma) < f(v) = p + 1/2$ and $f_{p-1}(\sigma) < f_{p-1}(v) = p - 1/2$, so (9.3) holds.

Suppose $v \not\subseteq N^+$. Then $f(v) = f_{p-1}(v)$. If $l < p - 1$ then $f(\sigma) = f_{p-1}(\sigma)$ so (9.3) holds. If $l = p - 1$ then $f(\sigma) \geq f_{p-1}(\sigma)$ so if $f_{p-1}(\sigma) > f_{p-1}(v)$ (9.3) holds. If $f_{p-1}(\sigma) < f_{p-1}(v)$ then $W(v) = \sigma$ which implies $W(\sigma) = 0$. Thus $d(\sigma) = 0$ and $f(\sigma) = f_{p-1}(\sigma)$ and again (9.3) holds. This proves (9.3) in all cases.

In particular, (9.3) implies, by induction, condition 2(i) if dimension $(\sigma) \leq p - 2$ and 2(ii) if dimension $(\sigma) \leq p - 1$. Suppose dimension $(\sigma) = p - 1$. By construction, if $\tau^{(p)} > \sigma$ then $f(\tau) \leq f(\sigma) \hookrightarrow W(\sigma) = \tau$. This implies condition 2(i) in this case. This completes the proof that f is a Morse function for the triad (M, N^-, N^+) with image $[-1/2, p + 1/2]$.

We will now check that $W = V_f$. That is, if $v^{(l-1)} < \sigma^{(l)}$ then

$$(9.4) \qquad\qquad f(v) \geq f(\sigma) \longleftrightarrow W(v) = \sigma.$$

By induction, if $l \leq p - 1$ then $f_{p-1}(\sigma) \leq f_{p-1}(v) \hookrightarrow \sigma = W_{p-1}(v) = W(v)$. Thus, in this case (9.4) follows from (9.3). If $l = p$ then (9.4) is true by construction.

Lastly, we must check that f is self-indexing. That is, if $\sigma^{(l)}$ is critical then $f(\sigma) = l$. If $l = p$ this is true by construction. If $l < p - 1$ then $f(\sigma) = f_{p-1}(\sigma)$ and σ is critical for f_{p-1} so it is true by induction. If $l = p - 1$ then if σ is critical we have $W(\sigma) = 0$ so $d(\sigma) = 0$. This implies $f(\sigma) = f_{p-1}(\sigma)$. Since σ is critical for f_{p-1} the result follows by induction.

This completes the proof. $\square$

We remark that it can be seen from the construction of the discrete Morse function f that

   (i) If $\sigma^{(p)} \in \mathrm{Image}(W)$ then $f(\sigma) \in (p - 1, p - \frac{1}{2})$.
   (ii) If $\sigma^{(p)}$ is critical then $f(\sigma) = p$.
   (iii) If $W(\sigma^{(p)}) \neq 0$ then $f(\sigma) \in (p, p + \frac{1}{2})$.

§10. The Morse Theorems for General CW Complexes.

In §3 we presented the Morse Theorems for a regular CW complex. In Corollary 8.3 we deduced the Strong Morse Inequalities for general CW complexes. Yet, for general CW complexes we still do not have the stronger topological statement of Corollary 3.5. In this section we fill this gap.

There are 2 difficulties in applying the proofs in Section 3 in this more general context. First, in Definition 3.1 we defined the level subcomplex $M(c)$ to be a union of some cells and all of their faces. However, unless M is normal [L-W p. 78] this need not be a subcomplex. Second, Lemma 3.2 may not hold. We need to restrict attention to those discrete Morse functions such that these difficulties do not arise. More precisely, for each cell σ, let Carrier(σ) denote the smallest subcomplex of M which contains σ. Given a CW complex M and a discrete Morse function f, say (M, f) satisfies the **Discrete Morse Hypotheses** if:

(DMH) (1) For every pair of cells σ and τ, if $\tau \subseteq$ Carrier(σ) and τ is not a face of σ, then $f(\tau) \leq f(\sigma)$. (This implies that for each c $M(c)$ is a subcomplex.)

(2) Whenever, for some p and $r \geq 0$, there is a $\tau^{(p+r+1)} > \sigma^{(p)}$ with $f(\tau) < f(\sigma)$ then there is a $\bar{\tau}^{(p+1)}$ with $\bar{\tau} > \sigma$ and $f(\bar{\tau}) \leq f(\tau)$.

Lemma 3.2 can now be restated as

Lemma 3.2. *If M is a regular CW complex and f is a discrete Morse function, then (M, f) satisfies the Discrete Morse Hypotheses.*

It is easy to check that if (M, f) satisfies the Discrete Morse Hypotheses then the proofs of Theorem 3.3 and Theorems 3.4, and hence Corollary 3.5, go through. That is, we have

Theorem 10.1. *If (M, f) satisfies the Discrete Morse Hypotheses then M is homotopy equivalent to a CW complex with exactly $m_p(f)$ cells of dimension p.*

We now apply the results of §9. Let M be a CW complex and f a discrete Morse function. Then, from Theorem 9.10 there is a Morse function $\bar{f}$ with the same critical cells such that

$$(10.1) \qquad \bar{f}(K_p \cap \text{Im } V) \subseteq (p-1, p-\tfrac{1}{2}) \quad \bar{f}(K_p \cap \text{Ker } V) \subseteq [p, p+\tfrac{1}{2})$$

(apply the theorem to the discrete vector field V_f and see the remark at the end of §9). Thus if $r \geq 0$ and

$$\bar{f}(\tau^{(p+r+1)}) \leq \bar{f}(\sigma^{(p)})$$

then necessarily $r = 0$. This is part (2) of **DMH**. Let $\sigma^{(p)}$ be any cell of M. If $\sigma \in \text{Ker } V$ if follows from (10.1) that for any cell $\tau \subseteq$ Carrier(σ) $\bar{f}(\tau) < \bar{f}(\sigma)$ (since dim $\tau < p$). If $\sigma \in \text{Im } V$, then σ must have a codimension 1 face. This implies that for any $\nu^{(p-1)} \subseteq$ Carrier(σ), ν is a face of σ. Thus, any $\tau \subseteq$ Carrier(σ) which is not a face of σ has dimension at most $p-2$, so that (10.1) implies $\bar{f}(\tau) < \bar{f}(\sigma)$. This is part (1) of **DMH**. Thus we can apply Theorem 10.1 to conclude

Theorem 10.2. *Let M be a CW complex and f a discrete Morse function. Then M is homotopy equivalent to a CW complex with exactly $m_p(f)$ cells of dimension p.*

§11. Cancelling Critical Points.

When using Morse theory to study the topology of a space, one frequently desires the Morse function to be as simple as possible, that is to have as few critical points as possible. In this section we use the results of §9 to investigate means of simplifying a given Morse function.

Theorem 11.1. *Let* (M, N^-, N^+) *be a cellular triad and* f *a Morse function on the triad. Suppose* $\tau^{(p+1)}$ *and* $\sigma^{(p)}$ *are critical points of* f *satisfying*

(i) *There is a regular face* $\bar{\sigma}^{(p)} < \tau$, *and a gradient path*

$$\bar{\sigma} = \sigma_0, \sigma_1, \ldots, \sigma_r = \sigma$$

from $\bar{\sigma}$ *to* σ *such that for each* $i = 0, \ldots, r-1$, σ_{i+1} *is a regular face of* $V(\sigma_i)$.

(ii) *There is no other gradient path from any p-face of* τ *to* σ.

(gradient paths are defined in Definition 8.4). Then there is a self-indexing Morse function $\bar{f}$ *on* M *such that*

$$(11.1) \qquad \{\text{critical points of } \bar{f}\} = \{\text{critical points of } f\} - \{\tau, \sigma\}.$$

Moreover, $V_f = V_{\bar{f}}$ *except along the unique gradient path from* $\partial\tau$ *to* σ.

Proof. Suppose $\bar{\sigma}^{(p)} < \tau$ and

$$\gamma : \bar{\sigma} = \sigma_0, \ \sigma_1, \ldots, \sigma_r = \sigma \qquad \sigma_{r-1} \neq \sigma_r$$

is the unique gradient path from a p-face of τ to σ. Define a combinatorial vector field W on M by setting

$$W(v) = V_f(v) \quad \text{if} \quad v \notin \{\bar{\sigma}, \sigma_1, \ldots, \sigma_{r-1}, \sigma\}$$
$$W(\sigma_i) = V_f(\sigma_{i-1}) \quad \text{for} \quad i = 1, 2, \ldots, r$$
$$W(\bar{\sigma}) = \tau$$

(note that none of the cells $\bar{\sigma}, \sigma_1, \ldots, \sigma_{r-1}, \sigma, V(\bar{\sigma}), V(\sigma_1), \ldots, V(\sigma_{r-1})$ lie in $N^- \cup N^+$). We have simply changed figure 0.8 to figure 0.9.

Now W satisfies the hypotheses of Theorem 9.10. Namely, it is clear that W is a combinatorial vector field (since V_f is). Moreover W has no non-stationary closed paths. To see this, suppose δ is a non-stationary closed W path. Since V_f has no such paths, δ must have dimension p and include at least one p-face from γ and at least one p-face which is not in γ. Therefore, δ must contain a segment $\bar{\delta}$ of the form

$$\bar{\delta} : \sigma_i, \ v_0, \ v_1, \ldots, v_r, \sigma_j$$

with $r \geq 0$, and $v_n \notin \{\sigma_0, \ldots, \sigma_r\}$ for all n (σ_i may equal σ_j). In particular, since $W(v_n) = V_f(v_n)$ for all n,

$$v_0, \ v_1, \ldots, v_r, \ \sigma_j$$

45

is a gradient path for f. In addition, if $i \neq 0$, then

$$v_0 \neq \sigma_{i-1}, \sigma_i \quad \text{and} \quad v_0 < W(\sigma_i) = V_f(\sigma_{i-1}).$$

Thus

$$\bar{\sigma} = \sigma_0, \ \sigma_1, \ldots, \sigma_{i-1}, \ v_0, \ v_1, \ldots, v_r, \ \sigma_j, \ \sigma_{j+1}, \ldots, \sigma_r = \sigma$$

is a second gradient path from $\partial \tau$ to σ. If $i = 0$ then

$$\bar{\sigma} \neq v_0 < W(\bar{\sigma}) = \tau$$

and

$$v_0, \ v_1, \ldots, v_r, \ \sigma_j, \ \sigma_{j+1}, \ldots, \sigma_r = \sigma$$

is a 2nd gradient path from $\partial \tau$ to σ. In either case we reach a contradiction.

Thus, by Theorem 9.10, there is a self-indexing Morse function $\tilde{f}$ with $W = V_{\tilde{f}}$. In particular, the critical points of $\tilde{f}$ are those faces v with $v \in \text{Kernel }(W)$, $v \notin \text{Image }(W)$. However,

$$\text{Kernel}(W) = \text{Kernel}(V_f) \setminus \{\sigma\}$$
$$\text{Image}(W) = \text{Image}(V_f) \cup \{\tau\}$$

which implies (11.1). $\square$

Corollary 11.2. *Let f be a Morse function on the triad (M, N^-, N^+). If $H_0(M, N^-) = 0$, then the critical points of f of index 0 can be cancelled against an equal number of critical points of index 1.*

Proof. Let v be a critical vertex of f. We will prove that there is a critical edge e such that there is exactly 1 gradient path from ∂e to v. From Theorem 10.1, the critical points v and e can then be cancelled. We can then repeat the process until all critical points of index 0 have been cancelled.

Applying Theorem 9.9 we learn that the map

$$\tilde{\partial} : C_1(M, N^-) \longrightarrow C_0(M, N^-)$$

is onto. If e is any critical edge then $\partial e = v_1 - v_2$ for some vertices v_1 and v_2 of M. There is a unique gradient path beginning at each of v_1 and v_2. Each such gradient path either ends at N^- or at a critical point of index 0. Thus, the image of $\tilde{\partial} e$ in $C_0(M, N^-)$ equals 0, $v_1, -v_1,$ or $v_1 - v_2$, where v_1 and v_2 are distinct critical points of index 0. In particular, if v is a critical vertex and $< \tilde{\partial} e, v > \neq 0$ then $< \tilde{\partial} e, v > = \pm 1$ and there is exactly 1 gradient path from ∂e to v. Since $\tilde{\partial}$ is onto, there is a linear combination

$$c = \sum_{e \text{ critical}} a_e e$$

such that

$$\tilde{\partial} \sum a_e e = \sum a_e \tilde{\partial} e = v.$$

Thus

$$\sum a_e < \tilde{\partial} e, v > = 1$$

so there must be at least one critical edge e with $< \tilde{\partial} e, v > \neq 0$. This edge has the desired property. $\square$

46

The process described in Theorem 11.1 can be reversed to create a pair of critical points.

Theorem 11.3. *Suppose*

$$\gamma : \sigma^{(p)} = \sigma_0, \ \sigma_1, \ldots, \sigma_r$$

is a gradient path of f of dimension p and $V_f(\sigma_r) = \tau \neq 0$. Suppose further that

$$\gamma \ \text{is the unique gradient path from} \ \sigma \ \text{to} \ \sigma_r.$$

Then there is a Morse function $\bar{f}$ such that

$$\{\text{critical points of } \bar{f}\} \ = \ \{\text{critical points of } .^r\} \cup \{\sigma, \tau\}.$$

Moreover, $V_f = V_{\bar{f}}$ except along the gradient path γ.

Proof. Define a combinatorial vector field W by

$$\begin{aligned} W(v) &= V_f(v) \qquad \text{if } v \notin \{\sigma_0, \sigma_1, \ldots, \sigma_r\} \\ W(\sigma_i) &= V_f(\sigma_{i-1}) \\ W(\sigma_0) &= 0. \end{aligned}$$

W has no closed paths because of (11.4). Hence by Theorem 9.10 $W = V_{\bar{f}}$ for a Morse function $\bar{f}$ with the desired properties. $\square$

The simplest example of the hypotheses of Theorem 11.3 is when $V_f(\sigma) = \tau$. Then we can simply take to be the trivial gradient path consisting only of σ. Then there is a Morse function $\bar{f}$ such that $V_{\bar{f}} = V_f$ on all cells except σ, and $V_{\bar{f}}(\sigma) = 0$. Changing f to $\bar{f}$ creates the pair of critical cells σ and τ.

§12. Invariance Under Subdivision.

As a final application of results in §9, we show that a discrete Morse function on a polyhedron has a natural extension to any subdivision resulting from a sequence of bisections. If one wishes to work in the category of simplicial complexes, one could show, by similar arguments, that a Morse function on a simplicial complex has a natural extension to any subdivision resulting from a sequence of stellar subdivisions.

Suppose M is a polyhedron with a Morse function f and $\widetilde{M}$ is a refinement of M resulting from a sequence of bisections (see §1). We will show that f induces an essentially equivalent Morse function $\tilde{f}$ on $\widetilde{M}$. In particular, the Morse complex associated to $\tilde{f}$ is canonically isomorphic to the Morse complex associated to f. Moreover, we can choose $\tilde{f}$ such that $V_{\tilde{f}}$ is equal to V_f on all cells of M which are also cells of $\widetilde{M}$. We state the theorem in terms of a single bisection, from which the general result clearly follows.

Theorem 12.1. *Suppose M is a polyhedron with a Morse function f, and $\sigma^{(p)}$ is a p-cell of M. Suppose $\widetilde{M}$ is the refinement of M resulting from a bisection*

$$\sigma^{(p)} = \sigma_1^{(p)} \cup v^{(p-1)} \cup \sigma_2^{(p)}.$$

Then there is a Morse function $\tilde{f}$ on $\widetilde{M}$ with the following properties

(i) $\tau \neq \sigma$ *is a critical point of f* $\longrightarrow$ τ *is a critical point of $\tilde{f}$.*

(ii) σ_2 *is not a critical point of $\tilde{f}$, and σ is critical for f* $\longrightarrow$ σ_1 *is critical for $\tilde{f}$.*

(iii) *Off of σ. $V_f = V_{\tilde{f}}$. That is, if $\tau_1 \neq \sigma \neq \tau_2$ then*

$$V_f(\tau_1) = \tau_2 \longleftrightarrow V_{\tilde{f}}(\tau_1) = \tau_2$$

and

$$V_f(\sigma) = \tau_1 \longleftrightarrow V_{\tilde{f}}(\sigma_1) = \tau_1 \quad or \quad V_{\tilde{f}}(\sigma_2) = \tau_1$$
$$V_f(\tau_1) = \sigma \longleftrightarrow V_{\tilde{f}}(\tau_1) = \sigma_1 \quad or \quad V_{\tilde{f}}(\tau_1) = \sigma_2.$$

(iv) *If $\tau_1^{(l)} \neq \sigma$ and $\tau_2^{(l+1)} \neq \sigma$ are critical points of f, then there is an algebraic-multiplicity preserving 1-to-1 correspondence between gradient paths of f from $\partial\tau_2$ to τ_1 in M and gradient paths of $\tilde{f}$ from $\partial\tau_2$ to τ_1 in $\widetilde{M}$. If σ is critical and either $\tau_1 = \sigma$ or $\tau_2 = \sigma$ then the same statement is true, replacing σ by σ_1 when considering $\widetilde{M}$.*

In particular, the Morse complex associated to f is canonically isomorphic to the Morse complex associated to $\tilde{f}$.

Proof. The proof is divided into 3 cases, depending on whether $\sigma \in \text{Image}(V_f)$, $V_f(\sigma) \neq 0$, or σ is critical.

Suppose first that $\sigma \in \text{Image}(V_f)$. Then $\sigma = V_f(\alpha^{(p-1)})$ for some $\alpha^{(p-1)} < \sigma$. Either $\alpha < \sigma_1$ or $\alpha < \sigma_2$. By relabeling if necessary, we assume $\alpha < \sigma_1$. Define a discrete vector field W on $\widetilde{M}$ by setting

$$W(\beta) = V_f(\beta)$$

48

if β is a cell of M and $\alpha \neq \beta \neq \sigma$

$$W(\alpha) = \sigma_1$$
$$W(\sigma_1) = 0$$
$$W(v) = \sigma_2$$
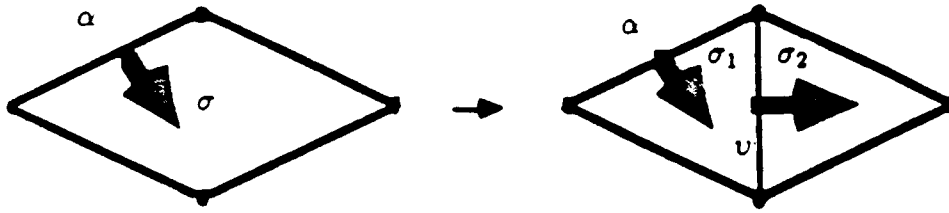$$W(\sigma_2) = 0$$

(see figure 12.1).

FIGURE 12.1

It is easy to see that since V_f has no closed paths, neither does W. Hence W is, in fact, a combinatorial vector field and, by Theorem 9.10, $W = V_{\tilde{f}}$ for some Morse function $\tilde{f}$ on $\widetilde{M}$. The critical points of f are identical to the critical points of $\tilde{f}$. The gradient paths of f which include neither α nor σ are precisely the gradient paths of $\tilde{f}$ which exclude α, σ_1, v and σ_2. Since $V_f(\sigma) = 0$, any gradient path of f of dimension p which includes σ must end at σ. Since σ is not critical, such paths do not play a role in the Morse complex associated to f. Similarly, gradient paths of $\tilde{f}$ which include σ_1 or σ_2 do not enter into the Morse complex associated to $\tilde{f}$.

If α is in a gradient path γ of f, then it must be followed by some $(p-1)$-cell $\tilde{\alpha} \neq \alpha$ of σ. If $\tilde{\alpha} < \sigma_1$ then γ is also a gradient path of $\tilde{f}$. If $\tilde{\alpha} < \sigma_2$ then replacing the segment α, $\tilde{\alpha}$ of γ by α, v, $\tilde{\alpha}$ we get a gradient path of $\tilde{f}$. This establishes a 1-1 correspondence between the gradient paths of f which contain α and the gradient paths of $\tilde{f}$ which contain α and v. The only gradient paths of $\tilde{f}$ which do not correspond to a gradient path of f are those that include v but not α. Such a path must begin at v. Since v is not the face of a critical p-cell, such paths play no role in the Morse complex. Thus, $\tilde{f}$ satisfies the conclusions of the theorem.

Suppose $V_f(\sigma) \neq 0$. Define a combinatorial vector field W on $\widetilde{M}$ by setting

$$W(\beta) = V_f(\beta)$$

if β is a cell of M and $\beta \neq \sigma$

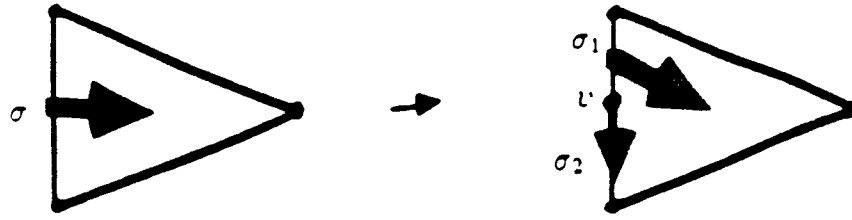$$W(\sigma_1) = V_f(\sigma)$$
$$W(v) = \sigma_2$$
$$W(\sigma_2) = 0$$

FIGURE 12.2

(see figure 12.2).

Then $W = V_{\bar{f}}$ for some Morse function $\bar{f}$ on $\widetilde{M}$. The gradient paths of f which do not include σ are precisely the same as the gradient paths of $\bar{f}$ which exclude σ_1, v and σ_2. Replacing σ by σ_1 establishes a 1-1 correspondence between the gradient paths of f which include σ and the gradient paths of $\bar{f}$ which include σ_1. The gradient paths of $\bar{f}$ which include v or σ_2 do not correspond to gradient paths of f. However, any gradient path which includes v must begin at v. Since v is not a face of a critical p-cell, such a path does not play a role in the Morse complex. Any gradient path which includes σ_2 must end at σ_2. Since σ_2 is not critical such a path does not play a role in the Morse complex. Thus $\bar{f}$ satisfies the conclusions of the theorem.

Lastly, suppose σ is critical. Define a discrete vector field W on $\widetilde{M}$ by setting

$$W(\beta) = V_f(\beta)$$

if β is a cell of M and $\beta \neq \sigma$

$$W(\sigma_1) = 0$$
$$W(v) = \sigma_2$$
$$W(\sigma_2) = 0$$

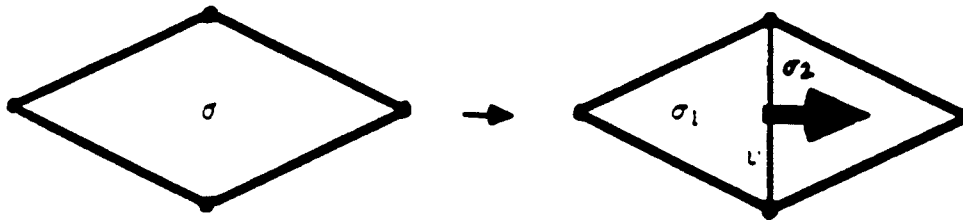so that σ_1 is critical for W (see figure 12.3)

FIGURE 12.3

Then $W = V_{\bar{f}}$ for some Morse function $\bar{f}$ on $\widetilde{M}$. Gradient paths of f which do not include σ are the same as gradient paths of $\bar{f}$ which exclude σ_1, v, and σ_2. Replacing σ by σ_1 gives a 1-1 correspondence between gradient paths of f which include σ and gradient paths of $\bar{f}$ which include σ_1. As in the other 2 cases, gradient paths which include σ_2 must end at σ_2 and hence play no role in the Morse complex. Gradient paths of $\bar{f}$ which include

50

v must begin at v. In this case, such paths are important in the Morse complex because v is a $(p-1)$-face of the critical p-cell σ_1. Suppose γ is a gradient path of f beginning at a $(p-1)$-face α of σ. Then either $\alpha < \sigma_1$ or $\alpha < \sigma_2$. If $\alpha < \sigma_1$ then γ is a gradient path of $\bar{f}$. If $\alpha < \sigma_2$ then v, γ is a gradient path of $\bar{f}$, and this establishes a 1-1 correspondence between gradient paths of f beginning at $\partial\sigma$ and gradient paths of $\bar{f}$ beginning at $\partial\sigma_1$. Thus, in this case too, $\bar{f}$ satisfies the conclusions of the theorem. $\square$

REFERENCES

[Ba] D. Barden, *The Structure of Manifolds*, Thesis, Cambridge University, 1963.

[B-K] U. Brehm and W. Kühnel, *Combinatorial Manifolds with Few Vertices*, Topology **26** (1987), 465–473.

[Du] A. Duval, *A Combinatorial Decomposition of Simplicial Complexes*, Israel J. of Math., to appear.

[Fo1] R. Forman, *Determinants of Laplacians on Graphs*, Topology **32** (1993), 35–46.

[Fo2] ———, *Witten-Morse Theory for Cell Complexes*, preprint.

[Gl] L.C. Glaser, *Geometrical Combinatorial Topology*, Van Nostrand Reinhold Company, 1970.

[Kl] W. Klingenberg, *The Morse Complex*, Symposia Mathematica, Vol. XXVI (Rome, 1980), Academic Press (1982), 117–122.

[K] A. Kosinski, *Singularities of Piecewise Linear Mappings. I Mappings into the Real Line*, Bull. AMS **68** (1962), 110–114.

[L-W] A. Lundell and S. Weingram, *The Topology of CW Complexes*, Van Nostrand Reinhold Company, 1969.

[Ma] B. Mazur. *Differential Topology from the Viewpoint of Simple Homotopy Theory*, Publ. Math. I.H.E.S. **15** (1963), 5–93.

[Mil] J. Milnor, *Morse Theory*, Annals of Mathematics Study No. 51, Princeton University Press, 1962.

[Mi2] ———, *Lectures on the h-Cobordism Theorem*, Princeton Mathematical Notes, Princeton University Press. 1965.

[Mo] M. Morse, *Bowls of a Non-Degenerate Function on a Compact Differentiable Manifold*, in *Differential and Combinatorial Topology*, S.S. Cairns ed., Princeton University Press (1965), 81–104.

[Sm1] S. Smale. *On Gradient Dynamical Systems*. Annals of Math. **74** (1961), 199–206.

[Sm2] ———. *The Generalized Poincaré Conjecture in Dimensions Greater than Four*, Annals of Math. **74** (1961). 391–406.

[St] J.R. Stallings, *Lectures on Polyhedral Topology*, Tata Institute of Fundamental Research, Bombay, 1968.

[Sta] R. Stanley. *A Combinatorial Decomposition of Acyclic Simplicial Complexes*, Discrete Math. **118** (1993).

[Wh] J.H.C. Whitehead. *Simplicial Spaces, Nuclei, and m-Groups*, Annals of Math.

[Wi] E. Witten. *Supersymmetry and Morse Theory*, J. Diff. Geom. **17** (1982), 661–692.