J. DIFFERENTIAL GEOMETRY
28 (1988) 59-92

# TOTAL ABSOLUTE CURVATURE AND EMBEDDED MORSE NUMBERS

R. W. SHARPE

**Abstract**

In this paper we use techniques of Morse theory to compute, under mild hypotheses, the infimum of the total absolute curvatures $\inf \tau(M^m \subset \mathbf{R}^w)$ for the smooth embeddings $M^m \subset \mathbf{R}^w$ in a given isotopy class.

## 1. Introduction

In 1929, W. Fenchel [8] showed that a circle immersed in $\mathbf{R}^3$ has (normalized) total absolute curvature (cf. §2 for the definition) at least 2 with equality only for the boundary of a convex planar disc. This was followed in 1949 by work of Fary [7] and Milnor [19] who showed that a knot in $\mathbf{R}^3$ has total absolute curvature more than 4. Since that time there has been considerable effort to obtain lower bounds for the total absolute curvature $\tau$ of a closed manifold immersed or embedded in Euclidean space in terms of the topological invariants of the situation and to study the consequences of small curvature (cf. e.g. Borsuk [1], Chern & Lashof [3], [4], Ferus [9], Fox [10], Kuiper & Meeks [15], Langevin & Rosenberg [17], Meeks [18], Pinkall [28], Sunday [32], and Wintgen [37]).

Recall that a Morse function on a smooth compact manifold $M$ is a smooth real valued function on $M$ whose critical points are all nondegenerate. The Morse number $\mu(M)$ is the minimum of the number of critical points of the Morse functions on $M$. For $m \neq 3$, 4 or 5 this is the same as the number of cells in the smallest CW complex with the simple homotopy type of $M$ (cf. Appendix 2.7). In 1958 Chern and Lashof [4] proved that for an immersion $\tau(i) \geq \mu(M)$, and raised the problem of determining the infimum of $\tau(i)$ as $i$ varies over some class of maps, such as all immersions, a regular homotopy class of immersions, all embeddings, or an isotopy class of embeddings. In particular they formulated:

Received May 19, 1986 and, in revised form, January 27, 1987. This research was supported in part by National Science and Engineering Research Council of Canada grant A8421.

**Chern-Lashof Conjecture.** $\inf \tau = \mu(M)$, *where the infimum is taken over all smooth immersions* $i\colon M^m \to \mathbf{R}^w$.

In this paper we study the case of embeddings. The first result is a simple consequence of the work of Kuiper [14], Wilson [36] and Perron [27].

**Theorem 1.1.** *Given a smooth embedding of a closed manifold* $i\colon M^m \to \mathbf{R}^w$, $5 < m < w - 2$, *then* $\inf \tau = \mu(M)$, *where the infimum is taken over all embeddings which are smoothly isotopic to the given one.*

When $2m < w$, every immersion is approximated by an embedding, so we get information about immersions too, and Theorem 1.1 implies:

**Theorem 1.2.** *The Chern-Lashof conjecture holds when* $5 < m$ *and* $2m < w$.

In a subsequent paper we shall give a proof of the Chern-Lashof conjecture for $m > 5$ without the codimension restriction.

The analogue of Theorem 1.1 for $w - m = 2$ or $1$ cannot hold in general. It is certainly false in codimension two for classical knots, since $\mu(S^1) = 2$, while $\tau > 4$ for such knots (Fary [7], Milnor [19]). In fact, Wintgen [37, p. 144] shows that every codimension two submanifold of Euclidean space whose complement has a noncyclic fundamental group provides a counterexample. Moreover, the normal circle bundles of these examples give counterexamples in codimension one for each dimension, such as the example in Figure 1.1 of the two-torus in $\mathbf{R}^3$ based on the trefoil knot.

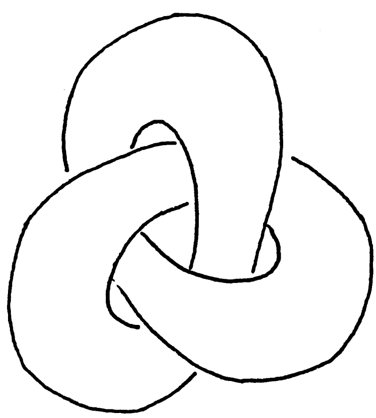


FIGURE 1.1

The next two theorems compute $\inf \tau$ when $m > 5$ in codimensions two and one in the cases when the fundamental groups are the simplest possible. It is convenient to fix an embedding $\mathbf{R}^w \subset S^w$, to set $X = S^w - D^{w-m}(M)$, where $D^{w-m}(M)$ is the normal disc bundle to $M$, and let $\tilde{X}$ denote the universal covering space.

**Theorem 1.3.** *Let $M^m \subset \mathbf{R}^{m+2}$ be a smooth embedding with $m > 5$, where $M$ is one-connected, and $X$ has infinite cyclic fundamental group. Then* $\inf \tau = \mu(X)$ *where the infimum is taken over all embeddings which are smoothly isotopic to the given one.*

**Theorem 1.4.** *Let $M^m \subset \mathbf{R}^{m+1}$ be a smooth embedding with $m > 5$, where $M$ is one-connected. Then* $\inf \tau(M) = \mu(X)$ *where the infimum is taken over all embeddings which are smoothly isotopic to the given one.*

When $M^m$ is simply connected with $m > 5$, a formula of Smale [31] gives

$$\mu(M) = \operatorname{rank}(H_*(M)/\text{torsion}) + 2 \cdot \operatorname{rank}(\text{torsion of } H_*(M)),$$

which can be used in Theorems 1.1, 1.2 and 1.4. When $X$ is nonsimply connected it does not seem possible to compute its Morse number except in special cases. We provide the following for use in Theorem 1.3:

**Theorem 1.5.** *Let $M^m \subset \mathbf{R}^{m+2}$ be a smooth embedding with $m > 5$, where $M$ is one-connected, and $X$ has the $[m/2]$-type of $S^1$. Then*

$$\mu(X) = \begin{cases} 2 + \operatorname{rank}_\Lambda H_k(\tilde{X}) & \text{if } m+2 = 2k, \\ 2 + 2\operatorname{rank}_\Lambda H_k(\tilde{X}) & \text{if } m+2 = 2k+1, \end{cases}$$

*where $\Lambda = \mathbf{Z}[t, t^{-1}]$, and multiplication by $t$ is the action in homology of the generator of the infinite cyclic group of covering transformations (i.e. the monodromy). By* $\operatorname{rank}_\Lambda$, *we mean the minimum number of generators as a $\Lambda$-module.*

(Note that, by duality, the hypotheses of 1.5 imply that $M$ is $[m/2] - 1$ connected.)

Now let us look at an explicit example in codimension two arising from a polynomial $f\colon \mathbf{C}^{n+1} \to \mathbf{C}$ with, at most, an isolated singularity at the origin. Let $V(f) \subset \mathbf{C}^{n+1}$ be the zero set of $f$, and let $M_\varepsilon = V(f) \cap S_\varepsilon^{2n+1} \subset S_\varepsilon^{2n+1}$, where $S_\varepsilon^{2n+1}$ is a small sphere of radius $\varepsilon$ centered at the origin. Then $X$ has the $(n-1)$-type of $S^1$ (Milnor [25]), so 1.3 applies when $n \geq 4$. Let $\rho = \operatorname{rank}_\Lambda H_k(\tilde{X})$. As is well known, $M_\varepsilon$ is a standard unknotted sphere if and only if $f$ is non-singular at 0 (Milnor [25]) and this in turn occurs if and only if $\rho = 0$ (Ferus [9]). Thus $f$ is singular at 0 if and only if $\rho \geq 1$. Exact computation of the ranks in 1.5 are not easy even for Brieskorn polynomials. However we offer:

**Proposition 1.6.** *If $n \geq 4$, then for the knotted pair $(S_\varepsilon^{2n+1}, M_\varepsilon)$ arising from the Brieskorn polynomial $f(z) = z_0^{d_0} + \cdots + z_n^{d_n}$, we have the estimate*

$$\frac{1}{d} \prod_{0 \leq r \leq n} (d_r - 1) \leq \operatorname{rank}_\Lambda H_n(\tilde{X}) \leq \prod_{1 \leq r \leq n} (d_r - 1),$$

*where $d = \text{l.c.m.}\{d_0, d_1, \cdots, d_n\}$.*

There is one special case where it is possible to give an exact answer:

**Proposition 1.7.** *If, in addition to the hypotheses of* 1.6, *all* $d_r$ *are relatively prime in pairs, then* $\operatorname{rank}_\Lambda H_n(\tilde{X}) = 1$.

It follows from this that, for any number of variables greater than four, and for any multiplicity greater than one, we can construct a Brieskorn polynomial in the given number of variables and with the given multiplicity for which $\inf \tau = 4$. This is quite different from the case of dimension 2 in which $\inf \tau$ is tied to the multiplicity by the formula $\inf \tau = 2\,\mathrm{mult}_0(z_0^{d_0} + z_1^{d_1}) = 2\min\{d_0, d_1\}$ (cf. Wintgen [37, p. 145]).

We give an outline of the paper. In §2 we present some background material concerning total absolute curvature. In particular we sketch the idea of Kuiper [14], improved by Wilson [36], relating $\inf \tau(M \subset \mathbf{R}^w)$ to the embedded Morse number. In §3 we recall the basic ideas of smooth embedded Morse theory due to B. Perron [27] which tells how the handle decomposition of the submanifold of a sphere induces a handle decomposition of its complement $X$. In §4 we refine the qualitative idea of §3 to give a quantitative description of the chain complex of the universal cover of the complement $X$. We also determine, on the chain level, the nature of the duality in this complement. It corresponds to the combination of Alexander and Poincaré duality given by the composite $H_r(X) \approx H^{(w-1)-r}(M) \approx H_{r-1}(M) \approx H^{(w-1)-(r-q)}(X)$. In §5 we study the effect on the chain complex of varying the Morse data, and in §6 the geometry is combined with the algebra of Appendix 2 to prove Theorems 1.3, 1.4 and 1.5. Finally, in §7, Propositions 1.6 and 1.7 are proved.

There are two appendices. In the first, we describe some of the conventions used in this paper. In the second, we study the algebra of chain complexes in a form suitable to our needs. The reader will find little that is really new here (cf., e.g., Cockcroft and Combes [5] and M. Cohen [6]).

We note that in the remainder of the paper we use $H_*(Y)$, $C_*(Y)$ etc. to denote the homology and chains of the universal cover of $Y$.

I would like to thank Bill Pardon for many helpful conversations during the preparation of this work, and Andrew Nicas who read portions of the manuscript. Thanks also go to Pierre Milman and Kunio Mursugi for their helpful interest. Finally I would like to thank the referees for their useful suggestions.

## 2. Total absolute curvature and embedded Morse numbers

Fix a smooth embedding (or immersion) $i: M \to \mathbf{R}^w$ of a smooth closed manifold into Euclidean space. Let $[i]$ denote the class of smooth embeddings isotopic to $i$ (or the class of smooth immersions regularly homotopic to $i$). Let

$\tau([i]) = \inf\{\tau(j) | j \in [i]\}$. In this section we recall the definition of $\tau(i)$ and describe the idea of Kuiper [14], and Wilson [36] which reduces the calculation of $\tau([i])$ to Morse theory.

Let $g\colon S^q(M) \to S^{w-1}$ be the Gauss map from the unit normal sphere bundle of $M$ to the unit sphere in $\mathbf{R}^w$ arising from the embedding (or immersion) $j\colon M \to \mathbf{R}^w$. The total absolute curvature of $j$ is defined to be

$$\tau(j) = \int_{S^q(M)} |g^*\omega|,$$

where $\omega$ is a rotation invariant volume form on $S^{w-1}$ of unit total volume.

A change of variables to the sphere $S^{w-1}$ yields the formula:

$$\tau(j) = 2\int_{S^{w-1}} \nu(e)\omega,$$

where $\nu(e) = \sharp\{x \in M | g(x) = e\}$. If $e$ is a regular value of $g$ (which holds for almost all $e \in S^{w-1}$ (cf. Chern-Lashof [4, p. 8])), then $h_e \equiv \langle e, \cdot\rangle|M$ is a Morse function and $\nu(e) + \nu(-e) = C(h_e) \equiv$ the number of critical points of $h_e$. Hence

$$\tau(j) = \int_{S^{w-1}} C(h_e)\omega.$$

Now define the embedded (or immersed) Morse number $\gamma \equiv \gamma([i])$ to be $\inf\{C(x_1 \circ j|M) | j \in [i]\}$. It follows immediately that $\tau \geq \gamma \geq \mu(M)$ (cf. [4, p. 8]).

Here is the observation due essentially to Kuiper.

**Theorem 2.1.** *Let $i\colon M^m \to \mathbf{R}^w$ be a smooth embedding (or immersion) of a smooth closed manifold in Euclidean space. Then $\tau([i]) = \gamma([i])$.*

*Sketch of Proof.* Choose $j\colon M^m \to \mathbf{R}^w$ isotopic (or regularly homotopic) to $i$ so that $C(x_1 \circ j|M) = \gamma$. Now "stretch out" the $x_1$ direction by means of the isotopy $D_t(x_1, x_2, \cdots, x_w) = (tx_1, x_2, \cdots, x_w)$, $1 \leq t < \infty$, so that the critical point structure of $h_{e_1}$ becomes predominate, i.e., $\forall\varepsilon > 0\ \exists t_0 > 1$ such that $\forall t \geq t_0$, and $\forall e \in S^{w-1}$ with $|e \cdot e_1| > \varepsilon$ we have $C(\langle e, \cdot\rangle | D_t(M')) = C(\langle e_1, \cdot\rangle | C_t(M'))$. It follows that

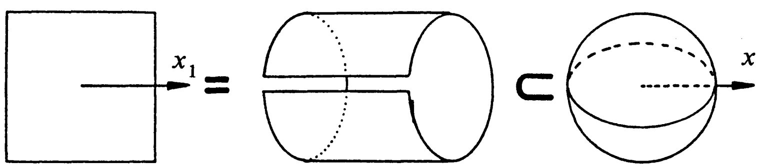
$$\lim_{t\to\infty} \int_{S^{w-1}} C(\langle e, \cdot\rangle | Dt(M'))\omega = C(\langle e_1, \cdot\rangle | M') = \gamma. \quad \text{q.e.d.}$$

Now let us consider an analogous circumstance in which we are given a smooth embedding (or immersion) $i\colon M \to S^w$. Fix a Morse function $f\colon S^w \to \mathbf{R}$ with just two critical points and define $\gamma \equiv \gamma([i])$ to be $\inf\{C(f \circ j|M) | j \in [i]\}$. When $w \geq 6$ this definition is independent of the particular choice of Morse function $f$, for given any two such functions, $f_1$ and $f_2$, they are related by an equation of the form $f_1 = h \circ f_2 \circ \Phi$ (cf. Cerf

[2, p. 189]), where $h$ is a diffeomorphism of $\mathbf{R}$, and $\Phi$ is a diffeomorphism of $S^w$ isotopic to the identity.

On the other hand, but only in the case of embeddings, we can define the embedded Morse number $\gamma = \gamma([i]) = \inf\{C(f \circ i) |$ where $f$ varies over all Morse function on $S^w$ with just two critical points$\}$. By the isotopy extension theorem (cf. Thom [33]) this depends only on the isotopy class $[i]$ of the embedding $i \colon M \to S^w$, and not on the particular representative given. It follows that the two definitions of the imbedded Morse number are equivalent.

Now let us compare the embedded Morse numbers of $M \subset \mathbf{R}^w$ and $M \subset S^w$, using the embedding $\varphi \colon \mathbf{R}^w \subset S^w$ as illustrated in Figure 2.1. Let $x_1 \colon S^w \subset \mathbf{R}^{w+1} \to \mathbf{R}$ be the standard first coordinate function, so that the composite $f = x_1 \circ \varphi$ is (left equivalent to) the standard first coordinate function of $\mathbf{R}^w$. Note that $S^w - \varphi(\mathbf{R}^w)$ is a great semicircle $J$ joining $\pm e_1$.




$R^w = R \times [S^{w-1} - p] \subset S^w$

FIGURE 2.1

**Proposition 2.2.** *Given a smooth embedding $i \colon M^m \to \mathbf{R}^w$ with $w > m+1$ the embedded Morse numbers $\gamma([i])$ and $\gamma([\varphi \circ i])$ are equal. In the case $w = m+1$, they are still equal provided we restrict the isotopies of $\varphi \circ i$ so that their images avoid the point $e_1 \in S^w$.*

*Proof.* Choose $i' \colon M^m \to \mathbf{R}^w$ isotopic to $i$ with $C(x_1 \circ i') = \gamma([i])$. Then $\gamma([i]) = C(x_1 \circ i') = C(x_1 \circ \varphi \circ i') \geq \gamma([\varphi \circ i])$. To obtain the inequality in the other direction choose $j \colon M^m \to S^w$ isotopic to $\varphi \circ i$ so that $C(x_1 \circ j) = \gamma([\varphi \circ i])$. By general position (or by assumption in the case $w = m+1$) we may assume that the image of the isotopy joining $\varphi \circ i$ and $j$ does not meet $e_1$. In fact we can deform the isotopy so that it does not meet $J$ either by pushing it off $J$ away from $e_1$. Since $j$ and the isotopy joining it to $\varphi \circ i$ now lie in $S^w - J = \varphi(\mathbf{R}^w)$ it follows that $C(x_1 \circ j) \geq \gamma([i])$. Hence $\gamma([\varphi \circ i]) = \gamma([i])$.

**Remark 2.3.** The proof of 1.4 can be regarded as showing that for $M$ simply connected, and $m > 5$, an imbedding $i \colon M^m \to S^{m+1}$, with both $\pm e_1$ in the same component of the complement $S^{m+1} - M^m$, is isotopic within $S^{m+1} - \{\pm e_1\}$ to an imbedding with $C(x_1 \circ i') = \mu(M)$. Since $\mu(M)$ is the minimum number of critical points possible, the restriction of the isotopies of $\varphi \circ i$ in 2.2 does not change the value of $\gamma([\varphi \circ i])$ in this case.

## 3. Relative Morse theory of a pair

In this section we give an informal review of how the handles of a submanifold give rise to handles in its complement in the ambient manifold. For the details we refer the reader to B. Perron [27].

To set the stage for the definitions we begin with the example of the standard torus $T^2 \subset \mathbf{R}^3$. In this example $T^2$ acquires its handle decomposition from the height function (cf. Milnor [21]). From Figure 3.1 it is clear that the portion of the complement of an open tubular neighborhood $N(T^2)$ of $T^2$ given by $X_t = \{x \in \mathbf{R}^3 - N(T^2) | x_3 \leq t\}$ changes topological type each time $t$ passes through a critical value of $x_3|T^2$. Figures 3.2 and 3.3 show how a 1-handle or a 2-handle is attached to $X_t$ as $t$ passes through a critical point of index one or two on $T^2$.

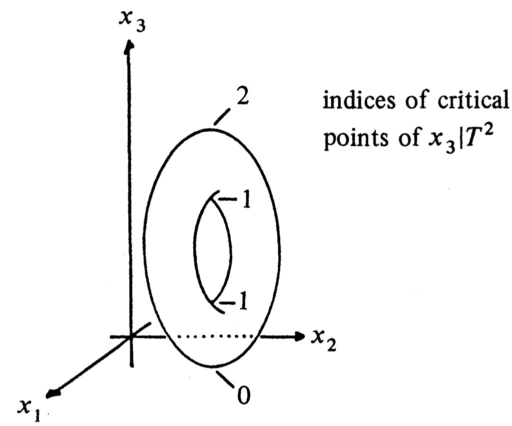



FIGURE 3.1

The idea that one can write down the cell structure of the complement in terms of the cell structure for an embedded circle has existed in the folklore of topology since the early days. It is the idea behind the Wirtinger presentation of the fundamental group of a classical knot. In higher dimensions, this basic idea has been studied in various forms by Rourke [29], Kearton [13] and Wintgen [37] as well as Perron [27]. We now pass to a general description of this phenomenon.

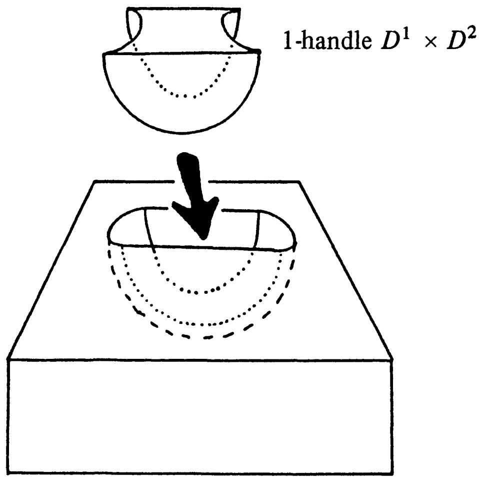



FIGURE 3.2

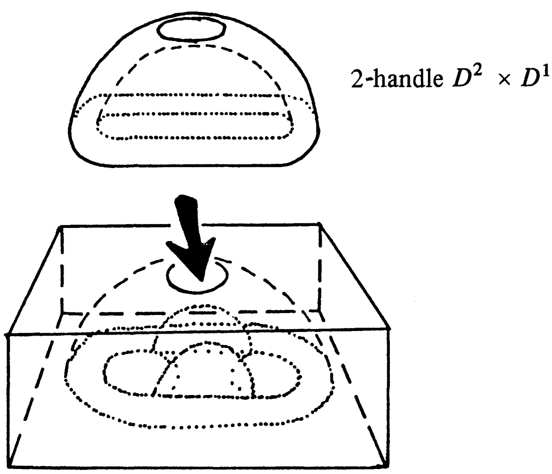



FIGURE 3.3

Let $(W, \partial_- W, \partial_+ W)$ be an oriented smooth manifold of dimension $w$, and $(M, \partial_- M, \partial_+ M) \subset (W, \partial_- W, \partial_+ W)$ an oriented submanifold of dimension $m$. Let

$$f\colon (W, \partial_- W, \partial_+ W) \to ([a, b], a, b)$$

be a Morse function on $W$ with no critical points on $M$, and assume that

$$f|M\colon (M, \partial_- M, \partial_+ M) \to ([a, b], a, b)$$

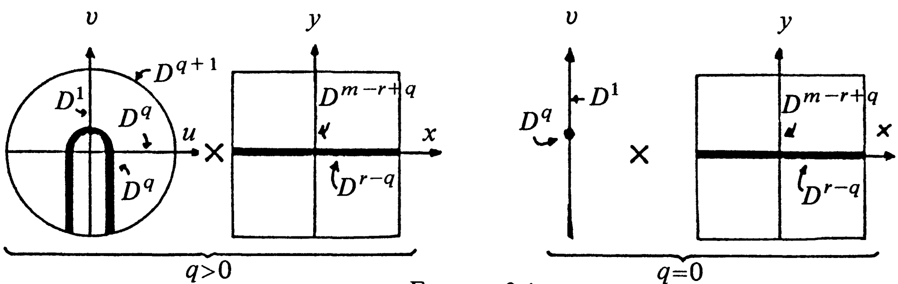



FIGURE 3.4

is also a Morse function. Then we shall say that $f$ is a Morse function for the pair $(W, M)$.

A vector field $\xi$ on $W$ is said to be gradient-like for $f$ if it satisfies:

**(3.1)** $\xi$ is gradient-like for $f$ except at the critical points of $f|M$.

**(3.2)** $\xi$ is tangent to $M$, and $\xi|M$ is gradient-like for $f|M$ in the usual sense (cf. Milnor [21])

**(3.3)** For each critical point $c$ of $f|M$, there is a neighborhood $U$ of $c$ in $W$ with a coordinate system $(u, v, x, y) \in \mathbf{R}^q \times \mathbf{R} \times \mathbf{R}^s \times \mathbf{R}^{m-2}$ such that:

(i) $M \cap U$ is given by $u = 0, v = 0$.

(ii) $f = f(c) + v - x^2 + y^2$ in $U$ (where $x^2 = x \cdot x$ etc.).

(iii) $\xi = (0, u^2 + v^2, -x, y)$ in $U$.

Such a pair $(f, \xi)$ we shall call Morse data for the pair $(W, M)$.

Given a pair $(W, M)$, the usual general position arguments apply to show that Morse data exist for it. Moreover, the usual Morse theory (cf. e.g., Milnor [22]) applies to show that $(f|M, \xi|M)$ gives rise to a handle decomposition of $M$ based on $\partial_- M$, with each critical point of index $r$ giving rise to a handle $(h^r, \partial_- h^r) = (D^r, S^{r-1}) \times D^{m-r}$. We denote by $\boldsymbol{B}_r(M)$ the set of $h$-handles on $M$.

As mentioned above in the case of $T^2 \subset \mathbf{R}^3$, we also obtain a handle decomposition of $X = W - M$. In this decomposition, the $r$-handles are of two types. A type I $r$-handle (denoted $h^r$) is just the usual one corresponding to a critical point of $f$ of index $r$. A type II $r$-handle (denoted $\mathfrak{h}^r$) corresponds to a critical point of $f|M$ of index $r - q$. We denote by $\boldsymbol{B}_r(X)$ the set of $r$-handles on $X$. Note the canonical inclusion $^\wedge$: $\boldsymbol{B}_{r-q}(M) \to \boldsymbol{B}_r(X)$.

The general picture of the core (Perron's *cupola*) of a type II handle $\mathfrak{h}^r$ is shown in Figure 3.4. In each of the cases in the figure the second factor (the square) with coordinates $(x, y)$ is the standard Morse theory picture for a neighborhood of the critical point $c$ in $M$ (with $D^{r-q}$ and $D^{m-r+q}$ corresponding to the negative and positive eigenspaces of the Hessian of $f$ at $c$ on

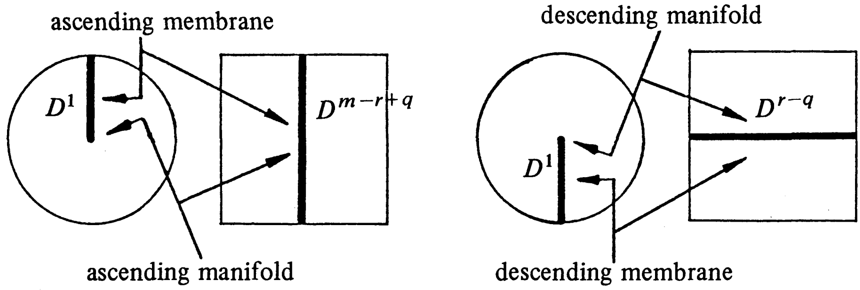
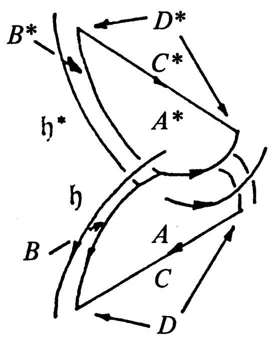



FIGURE 3.5




$D^*$ = coattaching sphere
$C^*$ = coattaching disc
$B^*$ = ascending manifold = core of $h^*$
$A^*$ = ascending membrane = cocore of $\mathfrak{h}$
$A$ = descending membrane = cocore of $\mathfrak{h}^*$
$B$ = descending manifold = core of $h$
$C$ = attaching disc
$D$ = attaching sphere

FIGURE 3.6

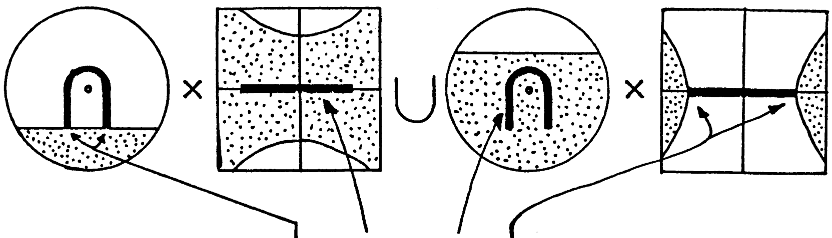
$M$ (cf. Milnor [21])). The first (disc) factor corresponds to the remaining coordinates $(u, v)$, with $v$ the vertical coordinate, and gives the fiber coordinates for the normal disc bundle to $M$ on a neighborhood of $c$. The dual Morse data $(-f, -\xi)$ gives rise to a handle $(h^r)^*$ of dimension $m - r$ dual to $h^r$, and a handle $(\mathfrak{h}^r)^*$ of dimension $w - (r + 1 - q)$ dual to $\mathfrak{h}^r$.

In addition to the usual ascending manifold, descending manifold, attaching sphere and co-attaching sphere of $c$ on $M$ we also have the ascending membrane, descending membrane, and their attaching discs (Perron's *projections*). These features are indicated in Figure 3.5. We can and shall always assume that these features are in general position with each other.

Figure 3.6 is a highly schematic diagram of these features. It also indicates our orientation conventions:

$$\partial A = B - C, \quad \partial B = \partial C = D; \qquad \partial A^* = B^* - C^*, \quad \partial B^* = \partial C^* = D^*,$$

where we follow the standard orientation convention for a manifold with boundary (Milnor [24, p. 27]): (Outward normal vector)$\times \partial Y = Y$. We shall also assume that $B \cdot B^* = 1$, $\mathfrak{h}^r \cdot A^* = (-1)^{r+1}$ and $A \cdot (\mathfrak{h}^r)^* = (-1)^{q(r-1)+1}$ so that once an orientation is chosen for $B$, all the others are determined.



$S^{q-1} \times D^{r-q} \cup D^q \times S^{r-q+1}$

= attaching sphere

FIGURE 3.7

In particular the orientation of $\mathfrak{h}$ is given by the orientation of $S^q \times h$ (with which it coincides along an open set), and similarly the orientation of $\mathfrak{h}^*$ is given by $S^q \times h^*$.

Now let us consider the handle $\mathfrak{h}^r$ in greater detail. Referring to Figure 3.7, the dark portion $D^q \times D^{r-q}$ is the core of $\mathfrak{h}^r$ (where the $D^q$ here refers to the upper hemisphere of an $\varepsilon$-sphere in $D^{q+1}$ surmounting the cylinder $S^{q-1} \times I$).

The attaching sphere of $\mathfrak{h}^r$ can be described as follows. Assume $f(c) = 0$, and set

$$X_s = \{x \in X | f(x) \leq s\},$$
$$V_s = \{(x,y) \in D^q \times D^{r-q} | -x^2 + y^2 \leq s\},$$
$$D_s^{q+1} = \{(u,v) \in D^q \times D^1 | v \leq s\}.$$

Then $X_{s-\varepsilon} \supset Y_\varepsilon = D_\varepsilon^{q+1} \times V_{s-2\varepsilon} \cup D_{-2\varepsilon}^{q+1} \times V_{s+\varepsilon}$, which is the shaded region in Figure 3.7. In the same figure the $r$-handle $\mathfrak{h}^r$ is the union of the products of the black regions and the attaching sphere $\partial\mathfrak{h}^r = \partial(D^q \times D^{r-q})$ in the level $f(c) - \varepsilon$ is indicated.

It is not difficult to show that $X_{s+\varepsilon}$ has $X_{s-\varepsilon} \cup \mathfrak{h}^r$ as a deformation retract. In fact, after rounding the corners, they are diffeomorphic.

Since the number of handles on $X$ = number of type I handles+number of type II handles=number of handles on $W$+number of handles on $M$, we recover Wintgen's inequality $\gamma(M \subset W) \geq \mu(X) - \mu(W)$ (cf. [37]).

Let us return to the oriented attaching sphere $S^{q-1} \times D^{r-q} \cup (-1)^q D^q \times S^{r-q-1}$, in the case when $q > 0$. The first portion forms part of the boundary of a certain $D_1^q \times D^{r-q}$ which lies in the level of the attaching sphere as in Figure 3.8. In fact, this $D_1^q \times D^{r-q}$ provides a homology between the attaching sphere of $\mathfrak{h}^r$ and the cycle $z = (-1)^q S^q \times S^{r-1-q}$ where $S^q \times S^{r-1-q}$ is the normal sphere bundle over the attaching sphere of the handle $h^{r-q}$ on $M$.

When $q = 1$ we will also need another description of the attaching sphere itself. Let $F_0$ be an oriented normal framing in $M$ for the core of the associated

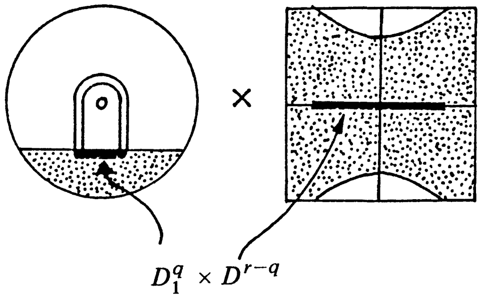



FIGURE 3.8

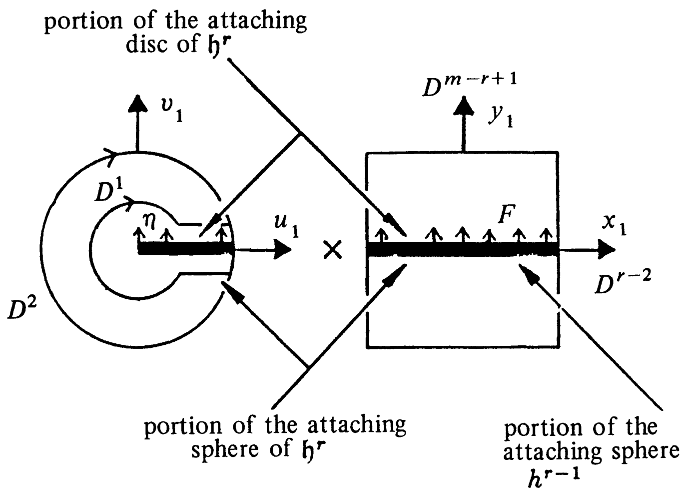



FIGURE 3.9

$r-1$ handle $h^{r-1}$. Augment $F_0$ to an oriented normal (in $W$) framing $(\eta_0, F_0)$ of the ambient descending manifold along its boundary. This framing extends uniquely (up to homotopy) over the ambient descending manifold to yield a framing $(\eta, F)$. Figure 3.9 is a picture of $\eta$ near $M$ in the level of the attaching sphere. (Note the new coordinates here, since we are no longer at the critical point.)

If we let $D^1$ be the normal circle with a small interval $(-\varepsilon, \varepsilon)$ of angles removed, and let $D_A^{r-1}$ denote the attaching disc of $\mathfrak{h}^r$, then we see that the oriented attaching sphere of $\mathfrak{h}^r$ is $-D^1 \times \partial(D_A^{r-1}) + \eta(D_A^{r-1}) - (-\eta)(D_A^{r-1})$.

When $q = 0$ the situation simplifies considerably. $\mathfrak{h}^r = \pm\sigma(h^r)$ where $\sigma$ is a normal vector field to $M$ along the core of $h^r$ which points up at $c$, and the sign compares $\sigma$ to the orientation of the normal bundle to $M$ in $W$. Thus the attaching sphere $\partial\mathfrak{h}^r$ is just $\pm\sigma(\partial h^r)$.

We now recall two fundamental results of Perron. Let us assume that $c_a, c_b \in M$ are critical points of index $a$ and $b$ for $f|M$ with $f(c_a) < v(c_b)$, and that $J = [f(c_a), f(c_b)]$ contains no critical values of either $f$ or $f|M$, other than its endpoints. Then:

**Ambient Rearrangement Theorem 3.4** (*Perron* [27]). *If $a \geq b$ and the ambient ascending and descending manifolds of $c_a$ and $c_b$ are disjoint, then there is a deformation of $f$ with support on a neighborhood of the portion of the ambient descending manifold of $c_a$ in $f^{-1}(J)$ to a Morse function of the pair $f_1 \colon W, M \to \mathbf{R}$ such that $f_1$ has the same critical points as $f$ except that $f_1(c_a) > f_1(c_b)$.*

**Ambient Cancellation Theorem 3.5** (*Perron* [27]). *Assume that $b = a + 1$, and that at some interior level of $J$ the descending manifold of $c_b$ and the ascending manifold of $c_a$ meet transversely in a single point. Assume in addition that the corresponding ambient ascending and ambient descending manifolds are disjoint. Then there is a deformation of $f$ with support as in* 3.4 *to a Morse function of the pair $f_1 \colon W, M \to \mathbf{R}$ such that $f_1$ has the same critical points as $f$, except for $c_a$ and $c_b$, which are no longer critical.*

**Remark 3.6.** The condition that the ambient descending and ambient ascending manifolds be disjoint follows by general position in 3.4 if $q > 0$, and in 3.5 if $q > 1$.

**Remark 3.7.** In the remainder of the paper we shall have $W^w = S^w$, and $f$ will be a Morse function on $S^w$ with just two critical points, a maximum and a minimum. By 3.4 and 3.5 we can, and do, assume that $f|M$ has all its index $i$ critical values in the interval $(i - 1/2, i + 1/2)$.

## 4. A chain complex for $X$

In this section we assume that we have a smooth embedding $M^m \subset S^w$ of codimension $q + 1 = w - m \geq 1$, that $M^m$ is connected, and, if $q = 1$, $M$ is orientable.

Let $C_*(Y)$ denote the chain complex of the universal cover $\tilde{Y}$ of $Y$ arising from the handle decomposition of $Y$. We refer the reader to the first appendix for our conventions concerning notation, orientations, base point paths, intersection numbers, etc.

By the results of Perron described in §3, we have a bijection between handles of $M$ and handles of type II of $X$. The aim of the present section is to analyze this bijection in order to realize Alexander duality and Poincaré duality at the level of chain complexes. That is, we show how the chain complex for $M$ (together with some extra structure if $q + 1 \leq 2$) determines the chain complex, and the dual complex, of $X$. The case of codimension $> 2$ appears essentially in Perron [27]. It is included here not only for completeness but because it forms the basis for the discussion of the cases of codimension $\leq 2$.

**4.A. The case of codimension greater than two** ($q > 1$).

**Theorem 4.A.1.** *If $q > 1$, then the cores of the 0-handles and $m$-handles of $M$ can be oriented so that the chain complex $(C_*(X), \delta_X)$ is isomorphic to the complex:*

$$\underset{\text{dim}:\ w}{0 \to \mathbf{Z}} \xrightarrow{\delta} C_m(M) \underset{w-1}{\otimes_\Xi} \mathbf{Z} \xrightarrow{\varepsilon_q \partial_m \otimes 1} C_{m-1}(M) \underset{w-2}{\otimes_\Xi} \mathbf{Z} \xrightarrow{\varepsilon_q \partial_{m-1} \otimes 1}$$
$$\cdots C_0(M) \underset{q}{\otimes_\Xi} \mathbf{Z} \to 0 \cdots 0 \to \underset{0}{\mathbf{Z}} \to 0.$$

*Here $\delta_w(1) = \sum [\hat{e}] \otimes 1$ (summation over $e \in \boldsymbol{B}_m(M)$; $[\hat{e}]$ denotes the chain corresponding to $\hat{e}$, and $\varepsilon_r = (-1)^r$.) Moreover, the dual complex, obtained from $(-f, -\xi)$ using the same orientations is:*

$$\underset{\text{dim}:\ w}{0 \to \mathbf{Z}} \xrightarrow{\delta^*} \underset{w-1}{\mathrm{Hom}_\Xi}(C_0(M), \mathbf{Z}) \xrightarrow{\varepsilon_1(\varepsilon_q \partial_1)^*} \underset{w-2}{\mathrm{Hom}_\Xi}(C_1(M), \mathbf{Z}) \xrightarrow{\varepsilon_2(\varepsilon_q \partial_2)^*}$$
$$\cdots \underset{q}{\mathrm{Hom}_\Xi}(C_m(M)\mathbf{Z}) \to 0 \cdots 0$$

*where $\delta^*(1) = \sum [\hat{e}]^*$ (summation over $e \in \boldsymbol{B}_0(M)$, and $[\hat{e}]^*$ is the dual basis of $[\hat{e}]$).*

*Proof.* Since $q > 1$, $X$ is simply connected. All the type II handles have dimension $\geq q \geq 2$. By our standing assumption, $X$ has only two type I handles, $h^w$ and $h^0$, corresponding to the maximum and minimum of $f$ on $S^w$. These remarks account for the copies of $\mathbf{Z}$ in dimensions 0 and $w$, and for the vanishing of the groups with dimensions in the range $0 < r < q$.

The co-core of a type II $(w-1)$-handle $\mathfrak{h}^{w-1}$ is a vertical interval which must meet the attaching sphere of $h^w$ transversely in a single point. We choose the orientations for the cores of the associated $m$-handles on $M$ so that these intersections all yield intersection numbers $+1$.

Now let $\mathfrak{h}^{r+1}$ and $\mathfrak{h}^r$ be type II handles on $X$ associated with the handles $h^{r+1-q}$ and $h^{r-q}$ on $M$. We want to compute the incidence number $[\mathfrak{h}^r, \mathfrak{h}^{r+1}]$, which we can obtain as the intersection number between $z$ and the co-core of $\mathfrak{h}^r$ (cf. §3).

Since the cycle $z$ in $X$ is localized near $M$, only the behavior of the co-core of $\mathfrak{h}^r$ near $M$ contributes to the incidence number, where it is a collar neighborhood of the co-core of $h^{r-q}$. In Figure 4.1 we picture a neighborhood in $S^w$ of an intersection point $(0,0)$ on $M$ between the attaching sphere of $h^{r+1-q}$ and the co-core of $h^{r-q}$ and illustrate how this intersection on $M$ gives rise to a corresponding intersection point $(P, 0)$ in $X$ between $z$ and the co-core of $\mathfrak{h}^r$. Call such an intersection *of type* A. Clearly every intersection between $z$ and the co-core of $\mathfrak{h}^r$ is of type A.

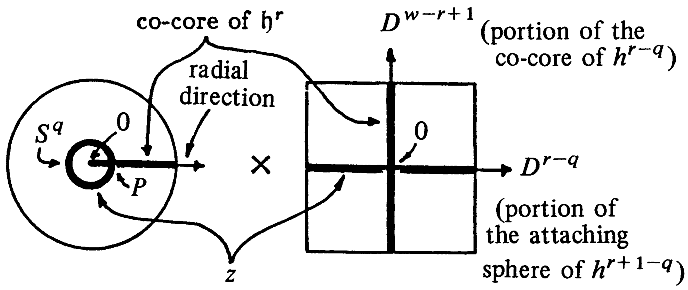



FIGURE 4.1

It follows from our orientation conventions that $\varepsilon_{(P,0)} = \varepsilon_q \varepsilon_{(0,0)}$, and from this we get:

$$[\mathfrak{h}^r, \mathfrak{h}^{r+1}] = \varepsilon_q [h^{r-q}, h^{r+1-q}] \otimes 1 \in \mathbf{Z} = \Xi \otimes_{\mathbf{Z}} \mathbf{Z},$$

which finishes the proof of the first part of the theorem.

The dual complex is obtained similarly.

**4.B. The case of codimension two** ($q = 1$). In this case $X$ is no longer simply connected so the choice of base point paths becomes important.

Choose a base point $*_X$ for $X$ in the normal circle bundle $S^1(M)$ of $M$, and choose the base point of $M$ to be its image $*_M = p(*_X)$, where $p\colon S^1(M) \to M$ is the bundle projection. For each critical point $c$ of $f|M$, choose $c' \in S^1(M)$ lying on the descending membrane below it, and select a path from it to $*_X$ in $S^1(M)$. This yields base point paths for the descending membranes. By composing with short paths in the fibers of $S^1(M)$, oriented positively, running from the ascending membrane to the descending membrane, we also obtain base point paths for the ascending membrane as well as the type II handle associated with $c$. Their images under projection yield base point paths for the handles of $M$. Finally we let $t \in \pi_1(S^1(M))$ be the class of the oriented fiber of $S^1(M)$ over $*_X$. By abuse of notation we also denote the image of $t$ in $\pi_1(X)$ by $t$. Since $M$ is orientable then $t$ is central in $\pi_1(S^1(M))$. We note however, that $t$ is not in general central in $\pi_1(X)$.

Set $\Lambda' = \mathbf{Z}\pi_1(S^1(M))$, and let $C_*(M)'$ be the (abstract) free $\Lambda'$-module with the handles of $M$ as basis. The canonical map $\Lambda' \to \Xi$ induces a $\Xi$-module isomorphism $C_*(M)' \otimes_{\Lambda'} \Xi = C_*(M)$.

Let $\partial_{\mathrm{I}}$ denote the boundary operator $\partial$ of $M$, but computed using the base point paths and intersections of type A in $S^1(M)$. That is, at a point of intersection $P$ between the attaching sphere $\partial h^{r+1}$ and the co-core of $h^r$ we join the ascending membrane of $h^r$ to the descending membrane of $h^{r+1}$ by the positively oriented path in $S^1$, and use it, together with the base point paths for the membranes described above, to yield an element $g'_P \in \pi_1(S^1(M))$ lifting $g_P \in \pi_1(M)$. Thus $\partial_{\mathrm{I}}$ operates on $C_*(M)'$, and

$$(C_*(M)', \partial_{\mathrm{I}}) \otimes_{\Lambda'} \Xi = (C_*(M), \partial).$$

(Note that we do not claim that $(C_*(M)', \partial_{\mathrm{I}})$ is a chain complex.)

Let $(\partial_{\mathrm{II}})_{r+1} \colon C_{r+1}(X) \to C_r(X)$ denote the intersection matrix measuring the intersections between the descending membrane and the ascending membrane of critical points of $f|M$ of adjacent indices $r$ and $r-1$. Transversality considerations show that there are no such intersections when $q > 1$, so this represents a new feature of the $q = 1$ case.

**Theorem 4.B.1.** *If $q = 1$, we can choose the orientations of the 0-handles and $m$-handles of $M$ so that $(C_*(X), \partial_X)$ is isomorphic to the chain complex:*

$$\underset{\dim:\ w}{0 \to \Lambda} \xrightarrow{\delta_w} \underset{w-1}{C_m(M)' \otimes \Lambda} \xrightarrow{\delta_{w-1}} \underset{w-2}{C_{m-1}(M)' \otimes \Lambda} \to \cdots \to \underset{1}{C_m(M)' \otimes \Lambda} \xrightarrow{\delta_1} \underset{0}{\Lambda} \to 0$$

*where the boundary operators are given by:*

$$\delta_{r+1} = -\{(\partial_{\mathrm{I}})_r \otimes 1 + (\partial_{\mathrm{II}})_{r+1}(t-1)\} \quad \textit{for } 1 < r+1 < w,$$
$$\delta_1[\hat{e}] = -g_e(t-1) \quad \textit{for } e \in \boldsymbol{B}_0(M), \textit{ and } g_e \in \pi_1(X),$$
$$\delta_w(1) = \sum [\hat{e}] g_e \quad (\textit{summation over } e \in \boldsymbol{B}_m(M), \textit{ and } g_e \in \pi_1(X)).$$

*Moreover, the dual complex arising from Morse data $(-f, -\xi)$ is*

$$0 \to \Lambda \xrightarrow{\delta^1} \mathrm{Hom}_\Xi(C_0(M), \Lambda) \xrightarrow{\delta^2} \mathrm{Hom}_\Xi(C_1(M), \Lambda) \xrightarrow{\delta^3} \cdots \mathrm{Hom}_\Xi(C_m(M), \Lambda) \xrightarrow{\delta^w} \Lambda \to 0.$$

*We are making the identification* $\mathrm{Hom}_\Xi(C_*(M), \Lambda) \approx \mathrm{Hom}_{\Lambda'}(C_*(M)', \Lambda)$, *and the boundary operators are given by:*

$$\delta^{r+1} = -\varepsilon_r\{(\partial_{\mathrm{I}})^*_r \otimes 1 + (\partial_{\mathrm{II}})^*_{r+1}(t^{-1} - 1)\} \quad \textit{for } 1 < r+1 < w,$$
$$\delta^1[\hat{e}]^* = -g_e^{-1}(t^{-1} - 1) \quad \textit{for } e \in \boldsymbol{B}_0(M),$$
$$\delta^w(1) = \sum [\hat{e}]^* g_e^{-1} \quad (\textit{summation over } e \in \boldsymbol{B}_m(M)).$$

(In particular, in the case when $t$ is central in $\pi_1(X)$ this complex and the original one are, after removing the extreme modules, and up to the signs of the boundary operators, *algebraic* duals.)

*Proof.* According to §3, the attaching sphere of $\mathfrak{h}^{r+1}$ is:

$$-D^1 \times \partial(D^r_A) + \eta(D^r_A) - (-\eta)(D^r_A).$$

We divide the intersections between the attaching sphere of $\mathfrak{h}^{r+1}$ and the co-core of $\mathfrak{h}^r$ into two types: those that lie on $D^1 \times \partial(D^r_A)$ (of type A), and those that lie on $\eta(D^r_A) - (-\eta)(D^r_A)$ (of type B).

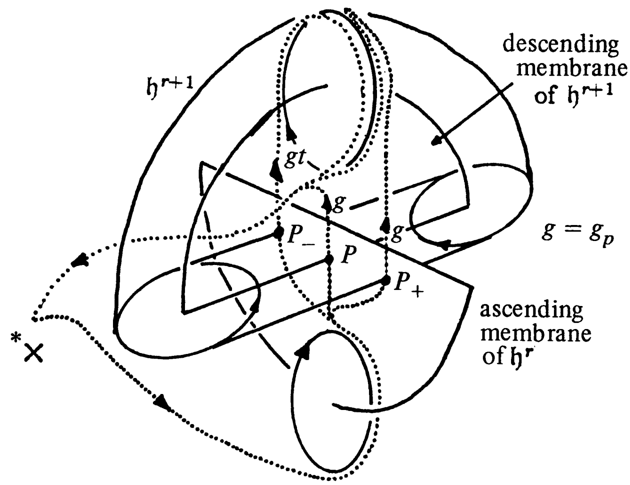



FIGURE 4.2

Let us first consider intersections of type A. We can compute the intersection of $D^1 \times \partial(D^r_A)$ with the co-core of $\mathfrak{h}^r$ by using the cycle $z$ instead, just as in the proof of 4.A.1, except that we must also pay attention to the base point paths. Our base point path conventions make $g'_P$ the correct choice. Since the computation of the signs is just as in 4.A.1, we have accounted for the term $-(\partial_{\mathrm{I}})_r \otimes 1$ in the boundary operator.

Now we consider the intersections of type B. These intersections arise in pairs $P_+$, $P_-$ of opposite sign, one pair for each transverse intersection point $P$ between the attaching disc $D_A^{r+1-q}$ of $\mathfrak{h}^{r+1}$ and the co-core of $\mathfrak{h}^r$. Let us assume that $P$ has intersection number $\varepsilon_P g_P \in \pm\pi_1(X)$.

From Figure 4.2 it is clear that $P_+$ and $P_-$ have intersection numbers $\varepsilon_P g_P$ and $-\varepsilon_P g_P t$, giving a total intersection number for the pair of $-\varepsilon_P g_P(t-1) \in \Lambda$. This accounts for the term $-(\partial_{\mathrm{II}})_r(t-1)$ in the boundary operator. We leave it to the reader to check the assertions about $\delta_w$ and $\delta_1$.

Now we pass to the duality statement. As in 4.A the type A intersections give rise to the term $-(\partial_{\mathrm{I}})^*_r$ in the dual boundary operator, so we need only understand the contribution of the type B intersections. Each such intersection point $P$, with intersection number $\varepsilon_P g_P \in \pm\pi_1(X)$ as above, when viewed dually gives the intersection number $(-\varepsilon_r)\varepsilon_P g_P^{-1}$. Moreover, $P$ gives rise to two intersections $P'_+$ and $P'_-$, one on each of the two sheets of the co-attaching sphere of $\mathfrak{h}^r$. As illustrated in Figure 4.3, these give a total intersection number of $(-\varepsilon_r)\varepsilon_P g_P^{-1}(t^{-1} - 1)$, which accounts for the term $-\varepsilon_r\{(\partial_{\mathrm{II}})^*_{r+1}(t^{-1} - 1)\}$ in the dual boundary operator. Again we leave the assertions about $\delta^w$ and $\delta^1$ to the reader.

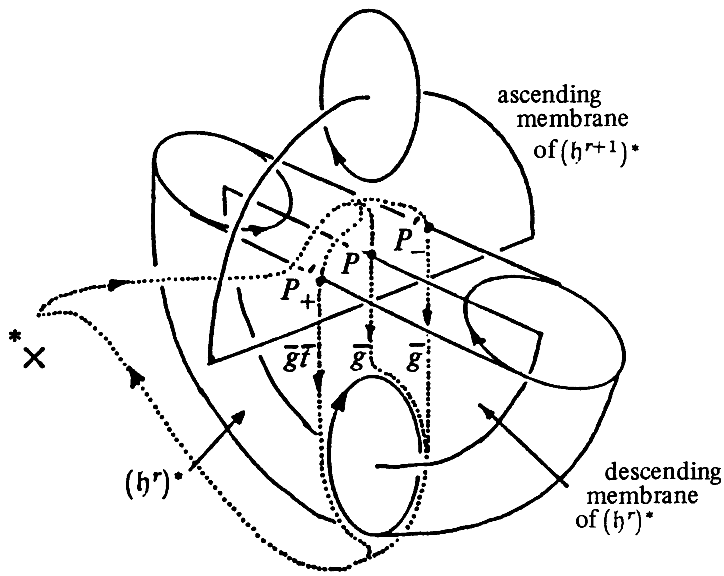



FIGURE 4.3

**4.C. The case of codimension one** ($q = 0$). Since a nonorientable manifold will not embed with codimension one in a sphere, $X = S^w - M$ splits into two components, $X = X_+ \cup X_-$. Let us assume for definiteness that $X_-$ contains the minimum and the maximum of $f$ and that $X_+$ is on the positive side of $M$. This corresponds to the assumption that the embedding $M^m \subset S^{m+1}$ arises from an embedding $M^m \subset \mathbf{R}^{m+1}$ (cf. §2). Set $\Lambda_\pm = \mathbf{Z}[\pi_1(X_\pm)]$, and let $\Xi \to \Lambda_\pm$ be the homomorphisms arising from the sections $\sigma_\pm$ of the normal bundle of $M$ in $S^w$. Let us partition $\boldsymbol{B}_*(M) = \boldsymbol{B}_*^+(M) \cup \boldsymbol{B}_*^-(M)$ so that the handles which are "locally below" $X_\pm$ lie in $\boldsymbol{B}_*^\pm(M)$. Similarly $C_*(M) = C_*^+ \oplus C_*^-$, where $C_*^\pm$ is the $\Xi$-submodule of $C_*(M)$ spanned by $\boldsymbol{B}_*^\pm(M)$. This yields a partition of the boundary map:

$$\delta = \begin{bmatrix} \delta_{00} & \delta_{10} \\ \delta_{01} & \delta_{11} \end{bmatrix}.$$

Choose a base point $*_M$ on $M$, and set $*_{X_\pm} = \sigma_\pm(*_M)$. Similarly, take the base point paths for the handles of $X_\pm$ to be the images of the base point paths on $M$ under $\sigma_\pm$.

**Theorem 4.C.1.** *If $q = 0$, we can choose the orientations and base point paths for the handles of $X$ so that the chain complexes are:*

*for* $X_+$:

$$\underset{\text{dim}:}{0} \to \underset{w-1}{C_m^+ \otimes_\Xi \Lambda_1} \xrightarrow{(\delta_{00})_m \otimes 1} \underset{w-2}{C_{m-1}^+ \otimes_\Xi \Lambda_1} \to \cdots \xrightarrow{(\delta_{00})_1 \otimes 1} \underset{0}{C_0^+ \otimes_\Xi \Lambda_1} \to 0,$$

*for* $X_-$:

$$\underset{\text{dim}:\ w}{0 \to \Lambda_2} \xrightarrow{\delta'} \underset{w-1}{C_m^- \otimes_\Xi \Lambda_2} \xrightarrow{-(\delta_{11})_m \otimes 1} \underset{w-2}{C_{m-1}^- \otimes_\Xi \Lambda_2} \to$$
$$\cdots \xrightarrow{\{-(\delta_{11})_1 + \varphi \circ \delta_{10}\} \otimes 1} \underset{0}{(C_0^- \otimes_\Xi \Lambda_2) \oplus \Lambda_2} \to 0$$

*where* $\delta'(1) = \sum [e] \otimes g_e$ *(summation over* $e \in \boldsymbol{B}_m^-(M)$*, and* $g_e \in \pi_1(X)$*) and* $\varphi([e]) = g_e$, $e \in \boldsymbol{B}_0^+(M)$.

*The complexes obtained from* $(-f, -\xi)$ *(but with the same orientations) are:*

*for* $X_+$:

$$\underset{\text{dim}:}{0} \to \underset{m}{\mathrm{Hom}_\Xi(C_0^-, \Lambda_1)} \xrightarrow{\varepsilon_1 (\delta_{11})_1^*} \underset{m-1}{\mathrm{Hom}_\Xi(C_1^-, \Lambda_1)} \to$$
$$\cdots \xrightarrow{\varepsilon_m (\delta_{11})_m^*} \underset{0}{\mathrm{Hom}_\Xi(C_m^-, \Lambda_1)} \to 0,$$

*for* $X_-$:

$$\underset{\text{dim}:\ m+1}{0 \to \Lambda_2} \xrightarrow{\delta''^*} \underset{m}{\mathrm{Hom}_\Xi(C_0^+, \Lambda_2)} \xrightarrow{-\varepsilon (\delta_{00})_1^*} \underset{m-1}{\mathrm{Hom}_\Xi(C_1^+, \Lambda_2)} \to$$
$$\cdots \xrightarrow{-\varepsilon_m (\delta_{00})_m^* + \varphi \circ \delta_{01}^*} \underset{0}{\mathrm{Hom}_\Xi(C_m^+, \Lambda_2) \oplus \Lambda_2} \to 0$$

*where* $\delta''^*(1) = \sum [e]^* g_e^{-1}$, *summation over* $e \in \boldsymbol{B}_m^+(M)$.

*Proof.* If $h^r$ is the handle of $M$ corresponding to the critical point $c$ with $X_\pm$ locally above $c$, then $\sigma_\pm(h^r) = \pm \mathfrak{h}^r$ (cf. §3). Let $\mathfrak{h}^r$ and $\mathfrak{h}^{r+1}$ be two type II handles of $X_\pm$ corresponding to the handles $h^r$ and $h^{r+1}$ of $M$. Note that the only intersections are of type A, since the ascending membrane of $\mathfrak{h}^r$ and the descending membrane of $\mathfrak{h}^{r+1}$ lie in distinct components of $X$. Thus $[\mathfrak{h}^r, \mathfrak{h}^{r+1}] = \pm \sigma_{j*}[h^r, h^{r+1}]$.

Now we pass to the duality statement. The critical points of $M$ which are locally below $X_+$ when we use $f$ as height become locally above when we use $-f$ as height. Thus if an $r$-handle $h^r$ on $M$ gives rise to a type II $r$-handle $\mathfrak{h}^r$ on $X_+$, then the dual of $\mathfrak{h}^r$ is an $(m-r)$-handle $(\mathfrak{h}^r)^*$ on $X_-$, arising from the dual of $h^r$. Given a pair of handles $h^{r+1}$ and $h^r$ on $M$ which are locally below $X_+$ with a local intersection number $\varepsilon_p g_p$ at $P \in M$, the associated type II handles $\mathfrak{h}^{r+1}$ and $\mathfrak{h}^r$ on $X_+$ have a corresponding local intersection number $\varepsilon_p \sigma_{+*}(g_p)$, and the dual handles $(\mathfrak{h}^{r+1})^*$ and $(\mathfrak{h}^r)^*$ have corresponding local intersection number $\varepsilon_{r+1}\varepsilon_p\sigma_{2*}(g_p^{-1})$ (cf. 4.A for the sign). The case of $X_-$ is similar.

## 5. Variations of $f$ and the membranes

In the last section we saw how a Morse function $f$ together with its membranes gives rise to a chain complex $(C_*(X), \partial_X)$. As is usual in Morse theory, the fact that $f$ and its membranes are not unique allows us to vary the chain complex by the operations of stabilization, cancellation and $\mathrm{Wh}_1$-simple change of basis (cf. Appendix 2 for descriptions of these operations of type A.2.1($r$), A.2.2($r$) and their inverses). The details of this investigation fall into three cases according as $q > 1$, $q = 1$ or $q = 0$ ($q = w - m - 1$). We shall study this question under the following:

**Hypothesis 5.1.** *Let $M^m \subset S^w$ be a smooth embedding of a closed connected manifold with $m > 5$. If $q = 0$ or 1 we assume that $M$ is simply connected. If $q = 1$ we assume in addition that $\pi_1(X) = \mathbf{Z}$.*

**5.A. The case of codimension greater than two** ($q > 1$). We know by 4.A.1 that $(C_*(M), \partial_M)$ determines $(C_*(X), \partial_X)$, which suggests that $(C_*(M), \partial_M)$ might be a complete invariant for our needs. The following result is essentially proved (but not stated) in Perron [27].

**Proposition 5.A.1.** *Let $M^m \subset S^w$ satisfy Hypothesis* 5.1. *Then every operation on $(C_*(M), \partial_M)$ of type* A.2.1($r$) *(where $0 \leq r < m$),* A.2.2($r$) *(where $0 < r < m$), and their inverses is realizable by variations in $f$ and its membranes.*

*Proof.* Consider first the inverse of an operation of type A.2.1($r$), where $0 \leq r < m$. If $\partial h^{r+1} = h^r \cdot \varepsilon g$ we can use 3.4 to put the corresponding critical points into adjacent position, and then use the Whitney trick [34] to perform an isotopy on $\xi|M$ so that the attaching sphere of $h^{r+1}$ meets the co-core of $h^r$ transversely in a single point, while leaving the other local intersections untouched. By the smooth isotopy extension theorem (Thom [33]), this can be accomplished by an isotopy of $\xi$ itself. Now apply 3.5 to obtain ambient cancellation of the critical points corresponding to $h^{r+1}$ and $h^r$. We leave to the reader the case of the operation A.2.1($r$) itself.

Now consider an operation of type A.2.2($r$), where $0 < r < m$. As usual, we can perform a handle addition on $(C_*(M), \partial_M)$ by an isotopy of $\xi|M$ and, as above, realize this by an isotopy of $\xi$ itself.

**5.B. The case of codimension two** ($q = 1$). This case is somewhat the reverse of 5.A in that, under Hypothesis 5.1, $(C_*(X), \partial_X)$ determines $(C_*(M), \partial_M)$, suggesting that $(C_*(X), \partial_X)$ is our complete invariant. Now it is not clear at first sight that the operations A.2.1($r$) and A.2.2($r$) are the appropriate ones for $(C_*(X), \partial_X)$, since the boundary operator is composed of two distinct parts. To remedy this situation we provide:

**Proposition 5.B.1.** *If* $q = 1$, *and assuming Hypothesis* 5.1, *then every algebraic decomposition* $(\partial_X)_{r+1} = -\{(\partial_{\mathrm{I}})_r \otimes 1 + (\partial_{\mathrm{II}})_{r+1}(t-1)\}$ *of the boundary operator in the range of dimensions* $0 < r \leq m$ *can be realized geometrically by a variation in the gradient-like vector field* $\xi$.

*Proof.* Any algebraic decomposition of

$$-(\partial_X)_{r+1} = (\partial_{\mathrm{I}})_r \otimes 1 + (\partial_{\mathrm{II}})_{r+1}(t-1)$$

can be obtained from any other by a sequence of operations of the following type. Choose basis elements $\mathfrak{h}^{r+1}$ and $\mathfrak{h}^r$. Given a local intersection between them of type A of intersection number $\varepsilon g t \in \pi_1(S^1(M))$, replace it by two local intersections, one of type A, of intersection number $\varepsilon g$, and another of type B, of intersection number $\varepsilon i_*(g)$. This corresponds to the algebraic identity $i_*(\varepsilon g t) = \varepsilon i_*(g) + \varepsilon i_*(g)(t-1)$, and can be accomplished by altering the gradient-like field $\xi$ by an isotopy with support in the normal disc bundle of $M$ in a band below the intermediate level (Milnor [23, Lemma 4.7, p. 43]). Hypothesis 5.1 insures that we reach every local intersection number of a type B intersection by means of this device. Figure 5.1 shows the effect of the isotopy in an intermediate level.

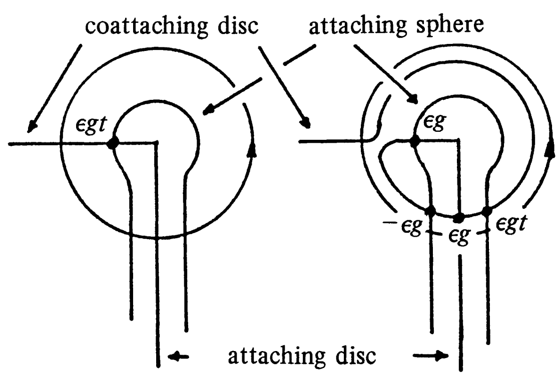



FIGURE 5.1

Now we can verify our hunch about $(C_*(X), \partial_X)$.

**Proposition 5.B.2.** *Let* $M^m \subset S^{m+2}$ *be a smooth embedding of a closed one-connected manifold with* $m \geq 6$. *Then every operation on* $(C_*(X), \partial_X)$ *of type* A.2.1($r$) (*where* $0 < r \leq m$), A.2.2($r$) (*where* $1 < r \leq m$) *and their inverses is realizable by variations in* $f$ *and its membranes.*

*Proof.* Consider first the inverse of an operation of type A.2.1($r$). Let $\mathfrak{h}^{r+1}$ and $\mathfrak{h}^r$ be two handles of $X$ whose total mutual incidence number is

$\varepsilon g \in \pm\pi_1(X)$, and let $h^r$ and $h^{r-1}$ be the handles of $M$ corresponding to them. Using Hypothesis 5.1, 5.B.1 and the Whitney trick, we may arrange that the ascending membrane associated to the $r$-handle $\mathfrak{h}^r$ has a single transverse intersection with the attaching sphere of the $r+1$ handle, and that this intersection is of type A. It therefore arises from a single transverse intersection of the ascending and descending manifolds on $M$, and moreover the ascending and descending membranes of $\mathfrak{h}^{r+1}$ and $\mathfrak{h}^r$ are disjoint. It follows from 3.5 that cancellation is possible.

Now let us consider operations of type A.2.2($r$). Let us fix an intermediate level between two adjacent critical points $p$ and $q$ on $M$, both of index $r-1$, where $0 < r-1 < m$, with $p$ above $q$. Let $h_p^{r-1}$, $h_q^{r-1}$ and $\mathfrak{h}_p^r$, $\mathfrak{h}_q^r$ be the corresponding handles on $M$ and $X$ respectively. In the intermediate level we have an *attaching pair* $(D_\delta^r(p), S_\delta^{r-1}(p)) \subset (S^w, M)$ for $\mathfrak{h}_p^r$ consisting of the attaching disc in $S^{m+2}$ and its bounding sphere in $M$, which is the attaching sphere of $h_p^{r-1}$ in $M$. Similarly we have a *co-attaching pair* $(D_\alpha^{m-r}(q), S_\alpha^{m-r-1}(q)) \subset (S^w, M)$ for $\mathfrak{h}_q^r$.

Now let us choose an isotopy of our intermediate level (respecting $M$, and with support off the attaching data of the remaining ambient $r$-handles) so that the pair $(D_\delta^r(p), S_\delta^{r-1}(p))$ is replaced by its boundary connected sum, "along a path in the normal sphere bundle of $M \subset S^w$," with the pair $(D^r, S^{r-1})$, where $D^r$ is a small normal disc of $S_\alpha^{m-r-1}(q)$ in $M$ whose interior has been pushed into $X$. This "path" is really an embedding of $I \times D^{r-1}$ into this level, where $I \times S^{r-1}$ is a narrow tube in $M$, and $I \times \operatorname{int}(D^{r-1})$ has been pushed into $X$. We fit this isotopy into a narrow band of levels below the given intermediate level, and use it to alter the gradient-like vector field $\xi$.

If we follow the small $(D^r, S^{r-1})$ down along trajectories to a level below $q$, it becomes a parallel copy of $(D_\delta^r(q), S_\delta^{r-1}(q))$. Thus the effect on the attaching pair of the critical point $p$ is to replace it by its boundary connected sum with the attaching pair for $q$ along a path in the normal disc bundle of $M$ (cf. Perron's *accident du type* (B) [27, p. 306]). Since a similar analysis holds also for the dual handles, we see that the effect on the chain complex $C_*(X)$ is an elementary change of basis in dimension $r$ where $1 < r \leq m$.

Hypothesis 5.1 insures that this handle addition can be used to replace $\mathfrak{h}_p^r$ by $\mathfrak{h}_p^r \pm \mathfrak{h}_q^r \cdot g$, where $g$ is an *arbitrary* element of $\pi_1(X) = \mathbf{Z}$.

**Remark 5.B.3.** Perhaps the restrictive conditions on $\pi_1(M)$ and $\pi_1(X)$ can be relaxed somewhat if the chain complex of $S^1(M)$ is used as the invariant (this complex has twice as many cells as $M$). However the allowable operations on it are of course more complicated than merely A.2.1($r$) and A.2.2($r$).

**5.C. The case of codimensions one** ($q = 0$). This case is analogous to the case $q > 1$ in that $(C_*(M), \partial_M)$, together with its partition of handles, is our complete invariant. We have:

**Proposition 5.C.1.** *Let $M^m \subset S^{m+1}$ satisfy Hypothesis* 5.1. *Then all the operations on $(C_*(M), \partial_M)$ of type* A.2.1($r$) (*where* $0 \leq r < m$, *and both handles belong to the same side of* $M$), A.2.2($r$) (*where* $0 < r < m$, *and both handles belong to the same side of* $M$), *and their inverses, are realizable by variations in $f$ and its membranes.*

*Proof.* The proof is exactly the same as the proof of 5.A.1, except that for realizing the operations A.2.1($r$) the disjointness required by 3.5 is automatic. This is because the two handles belong to the same side of $M$, and hence their ascending and descending membranes lie on opposite sides of $M$, and so lie in different components of $X$.

## 6. The Proof of Theorems 1.1, 1.3, 1.4 and 1.5

We begin by proving Theorem 1.1. Given $f$ and its membranes for the pair $M^m \subset S^w$, the restriction of this data to $M$ is the usual Morse data associated to $M$, and so there is a sequence of handle rearrangements, handle additions, stabilizations and cancellations replacing $f$ by a Morse function with $\mu(M)$ critical points. Now all of these operations can be realized ambiently: handle rearrangements by means of 3.4, handle additions by means of the smooth isotopy extension theorem, stabilizations obviously, and cancellations by means of 3.5 (this requires codimension greater than two). Thus $\gamma(M^m \subset S^w) \leq \mu(M)$, and since the reverse inequality is clear, we obtain $\gamma(M^m \subset S^w) = \mu(M)$. q.e.d.

We now prove Theorem 1.4. As before let $X = X_+ \cup X_- \subset S^{m+1}$, where $X_+$ corresponds to the "inside" of $M$ in $\mathbf{R}^{m+1}$. Note that $\pi_1(X_\pm) = 0$ since under Hypothesis 5.1 $M$ is simply connected, and hence by Van Kampen's theorem so are $X_+$ and $X_-$.

The correspondence of handles described in 4.C.1 shows that

$$2 + \sharp\{\text{handles of } M\} = \sharp\{\text{handles of } X_+\} + \sharp\{\text{handles of } X_-\}.$$

But in fact the explicit form of the boundary operator on the top dimensional handle of $X_-$ shows that two of its handles can be cancelled (nonambiently). This implies that $\gamma(M) \geq \mu(X_+) + \mu(X_-)$. Now we prove the reverse inequality. Let us consider the sequences of operations of types A.2.1 and A.2.2 on $C_*(X_+)$ and $C_*(X_-)$ reducing the numbers of cells to $\mu(X_+)$ and $\mu(X_-)$. Since our treatment of these is similar we work with $C_*(X_-)$ only.

Let us consider the class $\mathcal{D}$ of all finitely generated **Z**-free chain complexes which are chain homotopy equivalent to a chain complex arising from a CW structure on $X_-$. For example, the chain complex $C_*(X_-)$ of 4.C.1 lies in $\mathcal{D}$. Because of the explicit form of the boundary oeprator in 4.C.1 we have:

$$\underset{\dim:}{C_*(X_-)} \approx \{0 \to \underset{m+1}{\mathbf{Z}} \xrightarrow{\mathrm{id}} \underset{m}{\mathbf{Z}} \to 0\} \oplus C'_* \oplus \{0 \to \underset{1}{\mathbf{Z}} \xrightarrow{\mathrm{id}} \underset{0}{\mathbf{Z}} \to 0\}.$$

Now $C'_*$ can be nonzero in dimensions 0 through $m$ only. Choose $D_* \in \mathcal{D}$ with minimal total **Z**-rank $\mu$. In particular $\mu \leq \mu(X_-)$. Set

$$\underset{\dim:}{E_*} = \{0 \to \underset{m+1}{\mathbf{Z}} \xrightarrow{\mathrm{id}} \underset{m}{\mathbf{Z}} \to 0\} \oplus D_* \oplus \{0 \to \underset{1}{\mathbf{Z}} \xrightarrow{\mathrm{id}} \underset{0}{\mathbf{Z}} \to 0\}.$$

Clearly there is a (simple) chain equivalence $\theta\colon C_*(X_-) \to E_*$ which is an isomorphism outside the range $0 \leq r \leq m$. By A.2.6 we can change $C_*(X_-)$ to (an isomorphic copy of) $E_*$ by a sequence of operations of type A.2.1($r$), $0 \leq r < m$, and A.2.2($r$), $0 < r < m$. Now all these operations arise from the corresponding algebraic operations on $C_*(M)$ involving handles of $M$ on the $X_-$ side only. According to 5.C.1, all of these operations arise from variations in $f$ and its membranes for the pair $(S^{m+1}, M^m)$. Since similar remarks apply to $X_+$, it follows that $\gamma(M \subset S^{m+1}) \leq \mu(X_+) + \mu(X_-)$, and yields the equality of Theorem 1.4. q.e.d.

Now we prove Theorem 1.3 in an analogous way. Let us consider the class $\mathcal{D}$ of all finitely generated $\Lambda$ $(= \mathbf{Z}[t, t^{-1}])$-free chain complexes which are chain homotopy equivalent to a chain complex arising from the CW structure on $X$. (Note that $\mathrm{Wh}_1(\Lambda) = 0$, so that every chain equivalence is simple.) For example, the chain complex $C_*$ of Theorem 4.B.1 lies in $\mathcal{D}$. Since $X$ is connected, of dimension $m+2$, and not closed, from A.2.8 and A.2.9 it follows that

$$C_* \underset{\dim:}{\approx} \{0 \to \underset{m+2}{\Lambda} \xrightarrow{\mathrm{id}} \underset{m+1}{\Lambda} \to 0\} \oplus C'_* \oplus \{0 \to \underset{1}{\Lambda} \xrightarrow{t-1} \underset{0}{\Lambda} \to 0\}$$

where $C'_r \neq 0$ only in the range $1 \leq r \leq m+1$. It follows that the top two of the handles of $X$ can be cancelled (nonambiently), so that Wintgen's inequality (cf. §3) $\mu(X) - 2 \leq \gamma(M^m \subset S^{m+2})$ is improved to $\mu(X) \leq \gamma(M^m \subset S^{m+2})$.

Now choose $D^* \in \mathcal{D}$ with minimal total $\Lambda$-rank $\mu$. In particular $\mu \leq \mu(X)$. Again by A.2.8 and A.2.9,

$$\underset{\dim:}{D_*} = D'_* \oplus \{0 \to \underset{1}{\Lambda} \xrightarrow{t-1} \underset{0}{\Lambda} \to 0\}$$

where $D'_r \neq 0$ only in the range $1 \leq r \leq m+1$. Note that $C'_*$ and $D'_*$ are chain equivalent. Let

$$\underset{\dim:}{E_*} = \{0 \to \underset{m+2}{\Lambda} \xrightarrow{\mathrm{id}} \underset{m+1}{\Lambda} \to 0\} \oplus D_*$$

so that $\operatorname{rank}_\Lambda E_* = \mu + 2$. Clearly we can find a chain equivalence $\varphi_*\colon C_* \to E_*$ which is an isomorphism outside the range $1 \leq r \leq m+1$. By applying A.2.6 with $a = 1$ and $b = m+1$ we can alter $C_*$ by algebraic stabilization (in dimensions $1 \leq r \leq m$), cancellation (in dimensions $1 \leq r < m$), and handle additions (in dimensions $1 < r \leq m$) to obtain a chain equivalence $\varphi_*\colon C_* \to E_*$ which is a split epimorphism, and an isomorphism except in dimensions $m$ and $m+1$. By 5.B.2 each of these operations may be realized geometrically by changing $f$ and its membranes. Turning the situation upside down so that $C_*(M)$ is replaced by its dual $C^*(M)$ (cf. 4.B.1) we get a chain equivalence $\varphi^*\colon E^* \to C^*$ which is a split monomorphism, and an isomorphism except in dimensions 1 and 2. The left inverse for $\varphi^*$ gives a chain equivalence $\theta^*\colon C^*(M) \to E^*$ which is a split epimorphism, and an isomorphism except in dimensions 1 and 2. Now apply A.2.6 as above, but with $a = 1$ and $b = 3$ to make $\theta^*$ an isomorphism, and realize this (5.B.2) by geometric changes in $f$ and its membranes so that we have $\gamma(M \subset S^w) \leq \mu \leq \mu(X)$. Combining this with the reverse inequality given above shows all these inequalities are equalities, and proves Theorem 1.3. q.e.d.

Now consider Theorem 1.5. Since its hypotheses include those of Theorem 1.3, we can retain the notation and conclusions given above. Thus we have membranes for $f$ giving rise to a chain complex:

$$C_*(X) \underset{\dim:}{\approx} \{0 \to \underset{m+2}{\Lambda} \xrightarrow{\text{id}} \underset{m+1}{\Lambda} \to 0\} \oplus D'_* \oplus \{0 \to \underset{1}{\Lambda} \xrightarrow{t-1} \underset{0}{\Lambda} \to 0\}.$$

Now since $X$ has the $[m/2]$-type of $S^1$, it follows that $H_r(D'_*) = 0$ for $r \leq m/2$. We claim that in fact $D'_r = 0$ for $r \leq m/2$. For otherwise, choose $r \leq m/2$ minimal such that $D'_r \neq 0$. Then $D'_{r+1} \xrightarrow{\partial} D'_r$ is epic, and hence split. Thus $D'_*$ splits off a direct summand of the form $0 \to D'_r \to D'_r \to 0$, contradicting the minimality of $D'_*$.

Now by 4.B.1, turning the situation upside down replaces $C_*$ by the complex

$$\{0 \to \Lambda \xrightarrow{\text{id}} \Lambda \to 0\} \oplus D'^* \oplus \{0 \to \Lambda \xrightarrow{t-1} \Lambda \to 0\},$$

so $D'^r$ also vanishes for $r \leq m/2$.

If $m + 2 = 2k$, then $[m/2] = k - 1$ so that $D'_k = H_k(X)$ is the only nonvanishing module in $D'_*$, and hence $\mu = 2 + \operatorname{rank}_\Lambda H_k(X)$.

If $m + 2 = 2k + 1$, then $[m/2] = k - 1$, so that $D'_{k+1}$ and $D'_k$ are the only nonvanishing modules of $D'_*$. Now by (Milnor [25]) $H_{k+1}(X) = 0$, so we have a short exact sequence:

$$0 \to D'_{k+1} \to D'_k \to H_k(X) \to 0.$$

By duality we also have

$$0 \to D'^*_k \to D'^*_{k+1} \to H_k(X) \to 0,$$

and hence $x_\Lambda(H_k(X)) = 0$. Moreover, the fact that $D_*$ is minimal, together with A.2.10, shows that $\dim_\Lambda D'_{k+1} = \dim_\Lambda D'_k = \operatorname{rank}_\Lambda H_k(X)$, and hence $\mu = 2 + 2\operatorname{rank}_\Lambda H_k(X)$. These calculations finish the proof of Theorem 1.5.

## 7. Brieskorn varieties

Let $N_e = \mathbf{Z}[x]/(1 + x + \cdots + x^{e-1})$, and set $N_{\mathbf{d}} = N_{d_0} \otimes N_{d_1} \otimes \cdots \otimes N_{d_n}$ with $\Lambda$-module structure given by $t \cdot a_0 \otimes a_1 \otimes \cdots \otimes a_n = xa_0 \otimes xa_1 \otimes \cdots \otimes xa_n$.

It is known (Milnor [25]) that the "knot" arising from the Brieskorn polynomial $f(z) = z_0^{d_0} + z_1^{d_1} \cdots + z_n^{d_n}$ has complement $X$ with homology $H_n(X) = N_{\mathbf{d}}$ as a $\Lambda$-module. Thus Propositions 1.6 and 1.7 follow from:

**Proposition 7.1.** (a) *Let* $d = \mathrm{l.c.m.}(d_0, d_1, \cdots, d_n)$. *Then:*

$$\frac{1}{d} \prod_{0 \le r \le n} (d_r - 1) \le \operatorname{rank}_\Lambda N_{\mathbf{d}} \le \prod_{1 \le r \le n} (d_r - 1).$$

(b) *If* $d_0, d_1, \cdots, d_n$ *are relatively prime in pairs, then* $\operatorname{rank}_\Lambda N_{\mathbf{d}} = 1$.

*Proof.* First we obtain the upper bound of (a) by noting that $N_e$ is generated by a single element, and if $g_i$, $1 \le i \le s$, generate $N_{\mathbf{d}}$ then $g_i \otimes x^j$, $1 \le i \le s$, $0 \le j \le e - 2$, generate $N_{\mathbf{d}} \otimes N_e$. The upper bound

$$\operatorname{rank}_\Lambda N_{\mathbf{d}} \le \prod_{1 \le r \le n} (d_r - 1)$$

follows by induction. In case all the $d_r$ are relatively prime in pairs, the surjection $\mathbf{Z}[x]/(x^d - 1) \approx \bigotimes_r \mathbf{Z}[x]/(x^{d_r} - 1) \to N_{\mathbf{d}}$ shows that one generator is enough, proving (b).

To obtain the lower bound, we study the complex representation $N_{\mathbf{d}} \otimes \mathbf{C}$ of $\mathbf{C}[t, t^{-1}]$. Since $t^d$ acts by the identity, we have, in fact, a complex representation:

$$f(\tau) = \prod_{0 \le r \le n} (\tau^{e_r} + \tau^{2e_r} + \cdots + \tau^{(d_r - 1)e_r}) \in R(\mathbf{Z}/d) \approx \mathbf{Z}[\tau]/(\tau^d - 1)$$

where $d = e_r d_r$. Since each irreducible representation of $\mathbf{Z}/d$ occurs exactly once in a rank one free $\mathbf{Z}/d$ module, the rank of $N_{\mathbf{d}}$ is at least as large as the maximum coefficient of $f(\tau)$. Moreover:

$$\text{max. coeff. of } f(\tau) \ge \text{average coeff. of } f(\tau) = \frac{1}{d} f(1) = \frac{1}{d} \prod_{0 \le r \le n} (d_r - 1)$$

which gives the lower bound.

## Appendix 1. Various conventions

In this section we review the various conventions used in this paper. These are somewhat complicated in view of the fact that we need to compare many orientations and paths. If $(Y, *)$ is a (nice) pointed space, then its universal covering space is $\tilde{Y} = \{$homotopy classes of maps $\alpha\colon (I, 1) \to (Y, *)$, where the homotopies fix $\alpha(0)\}$. Using the standard multiplication of paths:

$$(\alpha * \beta)(t) = \begin{cases} \alpha(2t), & 0 \le t \le 1/2, \\ \beta(2t-1), & 1/2 \le t \le 1, \end{cases}$$

we obtain the group structure on $\pi_1(Y)$, as well as the *right* action of $\pi_1(Y)$ as the group of covering transformations on $\tilde{Y}$. Letting $C_*(Y)$ and $H_*(Y)$ denote the chains and homology of the *universal cover* $\tilde{Y}$ of $Y$, they become *right* $\Lambda = \mathbf{Z}\pi_1(Y)$-modules. For right modules, linear maps are written on the left, composed from right to left, and composition corresponds to the usual composition of matrices. That is, if we are given homomorphisms

$$\Lambda^c \xrightarrow{C} \Lambda^b \xrightarrow{B} \Lambda^a$$

and if $\{e_i\}$, $\{f_j\}$ and $\{g_k\}$ are the standard bases for $\Lambda^c$, $\Lambda^b$ and $\Lambda^a$ respectively, so that $B(f_j) = \sum g_k b_{kj}$ and $C(e_i) = \sum f_j b_{ji}$, then

$$(B \circ C)(e_i) = B(\sum f_j c_{ji}) = \sum B(f_j) c_{ji} = \sum g_k b_{kj} c_{ji} = \sum g_k (\sum b_{kj} c_{ji}).$$

Now let $Y$ be a handlebody [30]. We have in mind the handle decomposition arising from Morse data $(f, \xi)$ on a smooth manifold $M$. However most of our remarks apply equally well to any handle decomposition, such as the one for $X$ described in §3. Now recall the description of the chain complex $C_*(Y)$ associated to the handle decomposition. One first chooses orientations for the cores and paths to the base point for each of the handles (the resulting structures are called "based handles"). The co-core is then oriented so that (core) $\cdot$(co-core)$= 1$, with respect to the ambient orientation. (If $Y$ is not orientable we use the orientation at the base point of $M$ carried along the base point path to the intersection point.) Given these choices, the based $r$-handles lift uniquely to give a set $\boldsymbol{B}_r(Y)$ of based handles of $\tilde{Y}$. Regarding $\boldsymbol{B}_r(Y)$ as a set of chains in the free right $\Lambda$-module $C_r(Y)$ one sees that they yield a basis for it. Moreover the boundary map is given by $\partial_{r+1} h^{r+1} = \sum h^r [h^r, h^{r+1}]$, where the sum takes place over all $r$-handles $h^r$ and the incidence number $[h^r, h^{r+1}]$ is itself a sum $\sum \varepsilon_P g_P$ over the *local* intersection numbers corresponding to all the intersection points $P$ (assumed transverse) between the attaching sphere of $h^{r+1}$ and the co-core of $h^r$. The sign $\varepsilon_P = \pm 1$ is computed according to the formula:

$$(\text{Attaching sphere of } h^{r+1} \text{ at } P) \cdot (\text{co-core of } h^r \text{ at } P) = \varepsilon_P.$$

The fundamental group element $g_P$ is the class of the loop which passes from the base point along the inverse of the base point path to $h^r$, up the co-core of $h^r$ to $P$, change to the core of $h^{r+1}$, and then return along the base point path of $h^{r+1}$ to the base point, as indicated in the schematic Figure A.1.1.

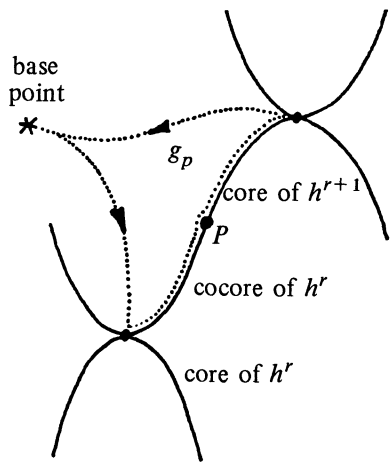



FIGURE A.1.1

Now let us consider "handle addition," which "adds $\varepsilon h_i^r$ to $h_j^r$" along a path $\lambda$ joining $h_i^r$ to $h_j^r$. The path $\lambda$ together with the base point paths of the handles provides an element $g \in \pi_1(M)$, and the handle addition has the effect of replacing the boundary operator $\partial_r$ by $E^{-1}\partial_r E$, where $E$ is the elementary matrix $e_{ij}(\varepsilon g)$. Thus such a handle addition can equally well be thought of as a simple change of basis among the $r$-handles. In the particular case when the handle decomposition arises from Morse data $(f, \xi)$ on $M$, such a "handle addition" can always be arranged by an isotopy of $\xi$ (cf. Milnor [22]).

We recall [24] that if

$$0 \to C_m(M) \xrightarrow{\partial_m} C_{m-1}(M) \xrightarrow{\partial_{m-1}} \cdots \xrightarrow{\partial_2} C_1(M) \xrightarrow{\partial_1} C_0(M) \to 0$$

is the chain complex arising from the Morse data $(f, \xi)$, then the dual complex (i.e. the one arising from $(-f, -\xi)$) is given by

$$0 \to \mathrm{Hom}_\Xi(C_0(M), \Xi) \xrightarrow{\varepsilon_1\partial_1^*} \mathrm{Hom}_\Xi(C_1(M), \Xi) \xrightarrow{\varepsilon_2\partial_2^*} \cdots \xrightarrow{\varepsilon_m\partial_m^*} \mathrm{Hom}_\Xi(C_m(M), \Xi) \to 0$$

where $\varepsilon_r = (-1)^r$.

## Appendix 2. Simplifying chain complexes

Let $k$ be a commutative ring and $G$ be a group. We set $\Lambda = kG$, the group algebra, and consider finitely generated, free, based right chain complexes $(C_*, \partial)$ over $\Lambda$. We study the following two operations on $(C_*, \partial)$:

**A.2.1($r$). Algebraic stabilization in dimension $r$.** This operation replaces $(C_*, \partial)$ by its direct sum with the canonically based chain complex:

$$\underset{\text{dim}:}{0} \to \underset{r+1}{\Lambda} \xrightarrow{\text{id}} \underset{r}{\Lambda} \to 0.$$

The inverse operation is called *algebraic cancellation in dimension $r$.*

**A.2.2($r$). Algebraic handle additions in dimension $r$.** This operation replaces the basis of $C_r$ by another which is $\mathrm{Wh}_1(G)$-simply equivalent to it.

We note in passing that neither of these operations alter the $\mathrm{Wh}_1(G)$-simple chain homotopy type of the complex.

Our first remarks about these operations corresponds to the basic simplification ideas of the $s$-cobordism theorem. In A.2.4, A.2.5, and A.2.6, we make the following standing assumption:

**Hypothesis A.2.3.** *$C_*$ and $D_*$ are finitely generated, free, based chain complexes over $\Lambda$, and $\varphi_*\colon C_* \to D_*$ is a $\mathrm{Wh}_1$-simple chain equivalence.*

**Lemma A.2.4.** *Assume $\varphi_r$ is an isomorphism for $r < n$ and an epimorphism for $r = n$.*

(i) *If $n + 1 = b$, then $\varphi_{n+1}$ is an epimorphism.*

(ii) *If $n+1 < b$, then we can stabilize $C_*$ by operations of type* A.2.1($n+1$) *so that $\varphi_*$ extends over the stabilization to yield an epimorphism in dimension $n + 1$.*

**Lemma A.2.5.** *Assume $\varphi_r$ is an isomorphism for $r < n$ and an epimorphism for $r = n$ and $r = n + 1$.*

(i) *If $n = b$, then $\varphi_n$ is an isomorphism.*

(ii) *If $n = b - 1$, then there is a combination of operations on $C_*$ consisting of stabilizations and their inverses of type* A.2.1($n$) *together with simple changes of base of type* A.2.2($n$), *so that $\varphi_*$ extends and restricts to give an isomorphism in dimension $n$, and is still an epimorphism in dimension $n+1$.*

(iii) *If $n < b-1$, then there is a combination of stabilizations and their inverses of types* A.2.1($n$) *and* A.2.1($n+1$), *together with simple changes of base of type* A.2.2($n + 1$), *so that $\varphi_*$ extends and restricts to give an isomorphism in dimension $n$, and is still an epimorphism in dimension $n + 1$.*

*Proof of* A.2.4. Consider the diagram:

$$\begin{tikzcd}
 & & & 0 \arrow[d] \\
C_{n+2} \arrow[r] \arrow[d, "\varphi_{n+2}"] & C_{n+1} \arrow[r] \arrow[d, "\varphi_{n+1}"] & C_n \arrow[r] \arrow[d, "\varphi_n"] & C_{n-1} \cdots \arrow[d, "\varphi_{n-1}"] \\
D_{n+2} \arrow[r] & D_{n+1} \arrow[r] & D_n \arrow[r] \arrow[d] & D_{n-1} \cdots \arrow[d] \\
 & & 0 & 0
\end{tikzcd}$$

Then (i) follows from a diagram chase. We obtain (ii) by induction on the number of basis elements of $D_{n+1}$ in the image of $\varphi_{n+1}$. If a basis element $\hat{e} \in D_{n+1}$ is not in this image we can include it by a single stabilization as follows: $\partial\hat{e} = \hat{u} = \varphi_n(u)$ and $\varphi_{n-1}(\partial u) = \partial\hat{u} = 0$. Thus $\partial u = 0$ so $u$ is a cycle. Since $\varphi_*\colon H_n(C_*) \approx H_n(D_*)$, and $\varphi_n(u) = \partial\hat{e}$, it follows that $u$ is a boundary, say $u = \partial e$. Hence $\partial\{\varphi_{n+1}(e) - \hat{e}\} = 0$. Moreover, since $\varphi_*\colon H_{n+1}(C_*) \approx H_{n+1}(D_*)$, there is a cycle $z \in C_{n+1}$ such that $\varphi_{n+1}(z) = \varphi_{n+1}(e) - \hat{e} + \partial\hat{a}$, $\hat{a} \in D_{n+2}$.

Now extend $\varphi_*$ over the stabilized version of $C_*$ by:

$$\begin{tikzcd}
0 \arrow[r] & \Lambda \arrow[r] & \Lambda \arrow[r] & 0 \\
 & 1 \arrow[r] \arrow[d] & 1 \arrow[d] & \\
 & \hat{a} \arrow[r] & \hat{e} - \varphi_{n+1}(e - z) & \\
\cdots \arrow[r] & D_{n+2} \arrow[r] & D_{n+1} \arrow[r] & \cdots
\end{tikzcd}$$

Clearly the extended map is still a simple chain equivalence, but now its enlarged image contains $\hat{e}$, completing the induction argument for A.2.4.

*Proof of* A.2.5. Consider the diagram:

$$
\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & & \\
\cdots \arrow[r] & K_{n+1} \arrow[r] \arrow[d] & K_n \arrow[r] \arrow[d] & 0 \arrow[d] & \\
\cdots \arrow[r] C_{n+2} \arrow[r] \arrow[d, "\varphi_{n+2}"] & C_{n+1} \arrow[r] \arrow[d, "\varphi_{n+1}"] & C_n \arrow[r] \arrow[d, "\varphi_n"] & C_{n-1} \arrow[r] \arrow[d, "\varphi_{n-1}"] & \cdots \\
\cdots \arrow[r] D_{n+2} \arrow[r] & D_{n+1} \arrow[r] \arrow[d] & D_n \arrow[r] \arrow[d] & D_{n-1} \arrow[r] \arrow[d] & \cdots \\
 & 0 & 0 & 0 &
\end{tikzcd}
$$

Again (i) comes from a diagram chase. Similar diagram chases show that in case (ii), $\partial\colon K_{n+1} \approx K_n$, while in case (iii) $\partial\colon K_{n+1} \to K_n$ is split by a map $s\colon K_n \to K_{n+1}$, so that $K_{n+1} = P \oplus s(K_n)$, where $P = \operatorname{Ker}\partial$. Clearly $K_n$, $K_{n+1}$, and hence $P$, are stably free. Stabilizations of type A.2.1($n$) will make $K_n$ and $K_{n+1}$ free, and in case (iii) additional stabilizations of type A.2.1($n+1$) will make $P$ free.

Choose a basis for $K_n$ (and hence for $s(K_n)$), and for $P$ so that $K_{n+1} = P \oplus s(K_n)$ is based, and the sequences $0 \to K_r \to C_r \to D_r \to 0$ (for $r = n$ and $r = n+1$) have no torsion. (If $n < b-1$ we may assume by stabilization that $P \neq 0$, and is indeed large enough to contain a basis to do the job. On the other hand if $n = b-1$, then $P = 0$, and we have an exact sequence of based chain complexes $0 \to K_* \to C_* \to D_* \to 0$. Since $\varphi_*$ is simple it follows that $K_*$ has no torsion.) Thus we can make a simple base change for $C_n$ and $C_{n+1}$ so that subsets of these bases give bases for $K_n$ and $s(K_n)$ which correspond under $\partial$. Now we can destabilize to remove $K_n$. This makes $\varphi_n$ an isomorphism, and completes the proof of A.2.5.

**Corollary A.2.6.** *Assume, in addition to Hypothesis* A.2.3, *that* $\varphi_*\colon C_* \to D_*$ *is an isomorphism outside the range of dimensions* $a \leq r \leq b$. *Then there is a sequence of operations on* $C_*$, *consisting of stabilizations (in dimensions* $a \leq r \leq b-1$*), cancellations (in dimensions* $a \leq r < b-1$*), and handle additions (in dimensions* $a < r \leq b-1$*), and a corresponding sequence of extensions and restrictions of the chain equivalence which yields a new chain equivalence* $\varphi_*\colon C_* \to D_*$ *which is a split epimorphism such that* $K_* = \operatorname{Ker}\varphi_*$ *is a free, based, acyclic complex with vanishing* $\mathrm{Wh}_1$ *torsion which is nonzero only in dimensions* $b-1$ *and* $b$.

*Proof.* Apply A.2.4 and A.2.5 inductively to obtain an epimorphism $\varphi_*: C_* \to D_*$ such that $K_* = \operatorname{Ker}\varphi_*$ is nonvanishing only in dimensions $b-1$ and $b$. Since $\varphi^*$ is a $\mathrm{Wh}_1$-simple chain equivalence, it follows that $K_*$ is a stably free, acyclic, torsion free complex. Make it free by stabilizing $C_*$ in dimension $b-1$. The splitting comes from a diagram chase.

**Corollary A.2.7.** *If $M^m$ is a smooth manifold of dimension $m \geq 6$, then every CW complex of the same simple homotopy type as $M$ has at least $\mu(M)$ cells. (And of course this minimum is attained by the CW complex corresponding to the minimal handlebody structure.)*

*Proof.* Fix a handlebody structure on $M$. Let $X$ be a finite CW complex simple homotopy equivalent to $M$ and choose a simple homotopy equivalence $f: M \to X$ inducing a simple chain equivalence $f_*: C_*(M) \to C_*(X)$ on the corresponding chain complexes. By A.2.6 (using $a = 0$ and $b = m$) there is a sequence of operations on $C_*(M)$, all realizeable by geometric operations on handles, by means of which $f_*$ can be converted to a split epimorphism whose kernel is a free based acyclic complex in dimensions $m-1$ and $m$ with vanishing Whitehead torsion. The dual map $f^*: C^*(X) \to C^*(M)$ is a split monomorphism, and the splitting yields a simple chain equivalence $g^*: C^*(M) \to C^*(X)$ which is again a split epimorphism. Its kernel is a free based acyclic complex in dimensions 0 and 1 with vanishing Whitehead torsion. Now $C^*(M)$ is the chain complex corresponding to the dual handle decomposition, so we can again invoke A.2.6 (This time using $a = 0$ and $b = 2$) to alter the handle decomposition so as to push the free acyclic kernel into dimensions 2 and 3. Finally in these dimensions we can use the fact that the kernel has vanishing Whitehead torsion to kill it by a geometric change in the handlebody. This yields an isomorphism $C^*(M) \to C^*(X)$, and hence $\mu(M) \leq$ number of cells of $X$.

**Lemma A.2.8** (*Wall*). *Let $C_*$ be a $\Lambda$ projective chain complex with $H^{r+1}(C_*, B) = 0$ for all coefficients $B$. Then $B_r$ is a direct summand of $C_r$.*

*Proof.* The hypothesis yields the exact sequence:

$$\begin{array}{ccccc} \operatorname{Hom}(C_{r+2}, B_r) & \longleftarrow & \operatorname{Hom}(C_{r+1}, B_r) & \longleftarrow & \operatorname{Hom}(C_r, B_r) \\ 0 & \longleftarrow & \partial & & \end{array}$$

Hence there exists $\varphi: C_r \to B_r$ such that $\varphi\partial = \partial$, i.e., $\varphi|B_r = \mathrm{id}$. This gives the required splitting.

Now we specialize to the case $\Lambda = \mathbf{Z}[t, t^{-1}]$. In particular, this insures that all projective modules are free, and all bases are stably equivalent (cf., e.g., [16]).

**Lemma A.2.9.** *Let $C_*$ be a $\Lambda$ free chain complex with $C_0 = \Lambda$, $H_1(C_*) = 0$, and $H_0(C_*) = \mathbf{Z}$ (with the trivial $\Lambda$ module structure). Then*

$$C_* \approx C'_* \oplus \{0 \to \Lambda \xrightarrow{t-1} \Lambda \to 0\}.$$

*Proof.* The short exact sequence $0 \to B_0 \to \Lambda \to \mathbf{Z} \to 0$ identifies $B_0$ with the augmentation ideal in $\Lambda$, which is a free rank one $\Lambda$ module. Hence $C_1 \to B_0$ splits, and we obtain our result.

**Lemma A.2.10.** *If $C_*$ and $D_*$ are finitely generated, $\Lambda$ free, chain equivalent chain complexes, with $n = \dim C_* > \dim D_*$, then*

$$C_* \approx C'_* \oplus \{0 \to C_n \xrightarrow{\approx} \partial C_n \to 0\}.$$

*Proof.* $H^n(C_*, B) = H^n(D_*, B) = 0$ for all coefficients $B$. By A.2.8, $\partial C_n$ is a projective, and hence free, summand of $C_{n-1}$. Since $H_n(C_*, B) = H_n(D_*, B) = 0$, we get $C_n \approx \partial C_{n+1}$, and hence the lemma.

**Lemma A.2.11.** *Let $N$ be a $\Lambda$ module with* $\operatorname{rank}_\Lambda N = g$, *Euler characteristic $\chi_\Lambda(N) = g - h$, and $h.d._\Lambda N = 1$. Then any free resolution $0 \to \Lambda^a \to \Lambda^b \to N \to 0$ can be altered by a sequence of stabilizations, cancellations and handle additions to a resolution of the form $0 \to \Lambda^h \to \Lambda^g \to N \to 0$.*

*Proof.* Since $\operatorname{rank}_\Lambda N = g$, there is an epimorphism $\Lambda^g \to N$, and since $\mathrm{h.d.}_\Lambda N = 1$, the kernel is projective, and hence free. Since the Euler characteristic of $N$ is $g - h$, the rank of the kernel must be $h$, so we obtain a resolution $0 \to \Lambda^h \to \Lambda^g \to N \to 0$ for $N$. The lemma now follows from A.2.6.

## References


[1] K. Borsuk, *Sur la courbure totale de courbes fermée*, Ann. Soc. Math. Polon. **20** (1947) 251–265.

[2] J. Cerf, *La stratification naturelle des espaces de fonctions différentielle réelles et la theoréme de la pseudo-isotopie*, Inst. Hautes Études Sci. Publ. Math. **39** (1970) 185–353.

[3] S. Chern & R. K. Lashof, *On the total curvature of immersed manifolds*. I, Amer. J. Math. **79** (1957) 306–318.

[4] ——, *On the total curvature of immersed manifolds*. II, Michigan J. Math. **5** (1958) 5–12.

[5] W. H. Cockcroft & J. A. Combes, *A note on the algebraic theory of simple equivalences*, Quart. J. Math. **22** (1971) 13–22.

[6] M. Cohen, *A course in simple-homotopy theory*, Springer, New York, 1973.

[7] I. Fary, *Sur la courbure total d'une courbe gauche faisant un noeud*, Bull. Soc. Math. France **77** (1949) 128–138.

[8] W. Fenchel, *Über Krümmung und Windung gesch lossener Raumkurven*, Math. Ann. **101** (1929) 238–252.

[9] D. Ferus, *Über die absolute Totalkrummung höher-dimensionaler Knoten*, Math. Ann. **171** (1967) 81–86.

[10] R. H. Fox, *On the total curvature of some tame knots*, Ann. of Math. (2) **52** (1950) 258–260.

[11] H. Hopf, *Über die Curvatura integra geschlossener Hyperfäschen*, Math. Ann. **96** (1926/27) 209–224.

[12] ——, *Vectorfelder in n-dimensionalen Mannigfaltkeiten*, Math. Ann. **96** (1926/27) 225–250.

[13] C. Kearton, *Presentation of n-knots*, Trans. Amer. Math. Soc. **202** (1975) 123–140.

[14] N. H. Kuiper, *Immersions with minimal total absolute curvature*, Colloque de Géométrie Différentalle CBRM, 1958, 75–88.

[15] N. H. Kuiper & W. Meeks, *Total curvature for knotted surfaces*, Invent. Math. **77** (1984) 25–69.

[16] T. Y. Lam, *Serre's conjecture*, Lecture Notes in Math., No. 635, Springer, Berlin, 1978.

[17] R. Langevin & H. Rosenberg, *On curvature integrals and knots*, Topology **15** (1976) 405–416.

[18] W. Meeks, *The topological uniqueness of minimal surfaces in three dimensional Euclidean space*, Topology **20** (1981) 389–410.

[19] J. W. Milnor, *On the total curvature of knots*, Ann. of Math. (2) **52** (1950) 248–257.

[20] ——, *On the immersions of n-manifolds in $(n+1)$-space*, Comment. Math. Helv. **30** (1955) 275–284.

[21] ——, *Morse theory*, Annals of Math. Studies, No. 51, Princeton University Press, Princeton, NJ, 1963.

[22] ——, *Lectures on the h-cobordism theorem* (Notes by L. Seibenmann and J. Sondow), Math. Notes, Princeton University Press, Princeton, NJ, 1965.

[23] ——, *Topology from the differentiable viewpoint*, University Press of Virginia, Charlottesville, VA, 1965.

[24] ——, *Whitehead torsion*, Bull. Amer. Math. Soc. **72** (1966), 358–426.

[25] ——, *Singular points of complex hypersurfaces*, Annals of Math. Studies, No. 91, Princeton University Press, Princeton, NJ, 1968.

[26] ——, *Infinite cyclic coverings*, Conference on the Topology of Manifolds (Michigan State University, East Lansing, MI, 1967), Prindle, Webber and Schmidt, Boston, MA, 1968, 115–133.

[27] B. Perron, *Pseudo-isotopies de Plongement en codimension* 2, Bull. Soc. Math. France **103** (1975) 289–339.

[28] U. Pinkall, *Tight surfaces and regular homotopy*, Topology **25** (1986) 475–482.

[29] C. Rourke, *Embedded handle theory*, Topology of Manifolds (C. Cantrell and C. Edwards, eds.), University of Georgia, 1969.

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

[31] ——, *On the structure of manifolds*, Amer. J. Math. **84** (1962) 387–399.

[32] D. Sunday, *The total curvature of knotted spheres*, Bull. Amer. Math. Soc. **82** (1976) 140–142.

[33] R. Thom, *La classification des immersions*, Séminaire Bourbaki, **157**, 1957.

[34] H. Whitney, *The self-intersections of a smooth n-manifold in 2n space*, Ann. of Math. (2) **45** (1944) 220–246.

[35] T. J. Willmore & B. A. Saleemi, *The total absolute curvature of immersed manifolds*, London Math. Soc. **41** (1956) 153–160.

[36] J. P. Wilson, *The total absolute curvature of immersed manifolds*, J. London Math. Soc. **40** (1965) 362–366.

[37] P. Wintgen, *Über die total Absolutkrümmung verknoteten Sphären*, Beiträge Algebra Geom. **9** (1980) 131–147.

UNIVERSITY OF TORONTO