SURGERY FOR AMATEURS

Andrew Ranicki
University of Edinburgh

John Roe
Penn State University

Author address:

Contents

CHAPTER 1

Questions About the Topology of Manifolds

This is a book about the topology of manifolds. One of the most important discoveries in topology — one that was the work of many mathematicians in the third quarter of the twentieth century — is that there is a systematic procedure for answering many natural questions about manifold topology, provided that the manifolds in question are sufficiently *high-dimensional*[1]. Insert quotation from Alexandroff here. A key geometric construction involved in this procedure is known as *surgery*, and the entire subject has taken on this name and is therefore often called 'surgery theory'. To do

Let's begin by reminding ourselves of the definitions of the objects that we want to study.

1.1. DEFINITION. A *topological n-manifold M* is a metrizable topological space that is locally homeomorphic to Euclidean space $\mathbb{R}^n$ — there is a cover of M by open sets U_α and there are homeomorphisms $\varphi_\alpha\colon U_\alpha \to \mathbb{R}^n$. (Such a cover $\{(U_\alpha, \varphi_\alpha)\}$ is called an *atlas*.)

The *transition functions* of an atlas are the functions $\varphi_{\alpha\beta} = \varphi_\alpha\varphi_\beta^{-1}$, which are homeomorphisms between open subsets of $\mathbb{R}^n$. An atlas is *smooth* if its transition functions are smooth (infinitely differentiable).

1.2. DEFINITION. A *smooth structure* on a topological manifold is a maximal smooth atlas. A *smooth manifold* is a manifold with a smooth structure.

Already some natural questions arise: Does every topological manifold admit a smooth structure? Is such a structure unique? As we shall see, the answers to both these questions are both negative in general.

A natural way to focus attention is to think about the *classification problem* — give a complete set of invariants which allows one to determine whether two manifolds are diffeomorphic, and give a list of representatives for the diffeomorphism classes. Of course, there is much more to differential topology than this, just as there is much more to group theory than trying to give a list of finite groups up to isomorphism; one wants to use the theory to say interesting things about non-trivial and natural examples. But classification is a good point at which to start our thinking. To solve a classification problem one needs to produce a list of *invariants* of the structure under consideration. What kind of invariants, then, are given to us by the statement that M is a smooth manifold?

Given any finite group presentation, one can construct a compact n-manifold, $n \geqslant 4$, whose fundamental group is given by the presentation. An effective classification of manifolds up to diffeomorphism (or even up to homotopy equivalence) would thus in particular include a classification of the groups given by finite presentations. It is known that there is no algorithm to accomplish such a classification. To avoid these logical issues, one usually formulates the classification problem in terms of classification of manifolds *within a given a homotopy type*: for some specified space X, how many 'essentially different' smooth manifolds are there homotopy equivalent to X?

---

[1] Usually, this means that the dimension needs to be at least 5 or 6.

In this chapter we want to review some of the invariants that can be used to approach this problem. We will also describe some key examples from the fifties and early sixties. These examples illustrate a number of mechanisms whereby the homotopy, homeomorphism and diffeomorphism of manifolds can be distinguished. Surgery theory proper tells us, in essence, that these mechanisms account for all the differences that there are between these various classifications.

1. Algebraic Topology

To begin with, we of course have the usual invariants of algebraic topology: homology, cohomology and homotopy groups. As a reference for these objects we recommend the text by Hatcher [**5**].

When the homology groups of a space X (or rather the associated numerical invariants — Betti numbers and torsion coefficients) were first defined by Poincaré and others, the definitions made use of a *triangulation* of X (that is, a representation of X as a simplicial complex). This led to the question whether homeomorphic polyhedra (or manifolds) are *combinatorially* equivalent (piecewise-linearly homeomorphic). The hypothesis that this is the case was known as the 'Main Conjecture' or *Hauptvermutung*. In fact the Hauptvermutung turned out to be false, even for manifolds — that is part of the story we have to tell in this book. However, long before these examples topologically invariant definitions of homology and cohomology had appeared (singular and Čech theories, for example). Thus the Hauptvermutung was no longer needed to prove the topological invariance of (co)homology.

When we deal with a *smooth* manifold M, it is also relevant to consider the *de Rham cohomology* groups. These are the cohomology groups of the complex

$$\Omega^0(M) \to \Omega^1(M) \to \Omega^2(M) \to \cdots$$

of differential forms on M. The *de Rham theorem* says that the de Rham cohomology of M is isomorphic to the usual cohomology with real coefficients. The usual proof of this establishes an isomorphism between de Rham and Čech cohomology; for this, and other matters relating to de Rham theory, our reference will be the book of Bott and Tu [**2**]. Cohomology has a ring structure (the cup product, given in de Rham theory by exterior product of forms) and this feature of cohomology will be crucially important in the discussion that follows.

One of the most notable features of the homology and cohomology of manifolds is *Poincaré duality.* Already in his 1895 memoir *Analysis Situs* [**?**], which founded the subject of topology, Poincaré had drawn attention to the fact that the Betti numbers of a compact oriented manifold exhibit a certain symmetry: $b_p = b_{n-p}$, if n is the dimension. Poincaré's 'proof' of this fact was severely criticized by Heegard, and in response he offered a second proof in [**?**]. This proof made use of dual cell decompositions in a manner that is still recognizable today. Poincaré also drew attention to the special rôle of the *middle dimension* in terms of duality. If $n = 2k$, then the k-dimensional (co)homology of M carries a nondegenerate bilinear form, the *intersection form*, which is defined in terms of the cup-product and is symmetric if k is even, skew-symmetric if k is odd. In particular, Poincaré pointed out, the middle Betti number of a (compact oriented) $4l+2$-dimensional manifold must be *even.* This is because the intersection form is nondegenerate and skew-symmetric, and such a form on a real vector space is a direct sum of copies of the form $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and in particular can exist only on an even-dimensional space.

In the case $n = 4l$ the intersection form is nondegenerate and symmetric. It is an elementary fact of linear algebra ("Sylvester's Law of Inertia") that any symmetric bilinear form over a finite-dimensional real vector space can be reduced, by a change of basis, to the form

$$B(\mathbf{x},\mathbf{y}) = x_1y_1 + \cdots + x_py_p - x_{p+1}y_{p+1} - \cdots - x_{p+q}y_{p+q},$$

and the number p of positive signs and q of negative signs appearing here are *invariants* of the form (in fact, they are the maximal dimensions of subspaces restricted to which the form is positive or negative definite). The difference $p-q$ is called the *signature* of the form, or of the manifold from which it arises.

1.3. EXERCISE. What is the signature of the complex projective space $\mathbb{CP}^{2k}$? Show that this space does not possess any orientation-reversing diffeomorphism.

Notice that in defining the signature we have neglected any finer arithmetic structure which arises from the fact that the intersection form is defined over $\mathbb{Z}$, not simply over $\mathbb{R}$. The classification of symmetric bilinear forms over $\mathbb{Z}$ is a much more subtle matter. For instance, a symmetric bilinear form over $\mathbb{Z}$ is called *even* if the diagonal entries in a matrix representation are even integers; equivalently, $B(\mathbf{x},\mathbf{x})$ is even for every integer vector $\mathbf{x}$. This notion is invariant under change of (integer) basis.

The *homotopy groups* of a space X are the groups $\pi_n(X) := [S^n, X]$ of homotopy classes of maps from the n-sphere to X. They are abelian when $n > 1$. In general, homotopy groups are much more mysterious that homology groups. The following example was known in the 1930s.

1.4. EXERCISE (The Hopf fibration). Regard S^3 as the group of unit quaternions and obtain a group homomorphism $S^3 \to SO(3)$ by sending a quaternion q to the transformation $x \mapsto qx\bar{q}$ of the purely imaginary quaternions. Since $SO(3)$ acts on S^2 by rotations, we obtain a map $S^3 \to S^2$. This map is called the *Hopf fibration*. Show that it represents a nonzero element in $\pi_3(S^2)$. (In fact, $\pi_3(S^2) = \mathbb{Z}$ and the Hopf map is the generator.)

hopf-fibration

1.5. EXERCISE. Following on from the above exercise, show that the Hopf fibration is a principal S^1-bundle over S^2. Give a complete classification of such bundles. (Any such bundle is trivial over the upper and lower hemispheres, so that it is determined by its *clutching function*, which is the map $S^1 \to S^1$ which shows how these two trivial bundles are joined together over the equator. Thus these bundles are classified by an integer $k \in \pi_1(S^1) = \mathbb{Z}$. This is an example of a *characteristic class*, see the next section. The Hopf fibration corresponds to $k=1$.)

1.6. EXERCISE. From a principal S^1-bundle over S^2 one can build an S^2-bundle over S^2 by fiberwise suspension. Show that the resulting S^2-bundles are classified by the residue class mod 2 of the integer k introduced in the previous exercise. (This is a matter of the homomorphism $\pi_1(SO(2)) = \mathbb{Z} \to \pi_1(SO(3)) = \mathbb{Z}/2$.)

1.7. EXERCISE. Show that the total space of the S^2-bundle over S^2 obtained in the previous section with $k=1$ is diffeomorphic to the connected sum $\mathbb{CP}^2\#(-\mathbb{CP})^2$, where $-\mathbb{CP}^2$ is the complex projective plane with the opposite of the standard orientation. (First show that the complement of a small 4-disk in $\mathbb{CP}^2$ is diffeomorphic to the total space of the complex line bundle associated to the Hopf bundle.)

projsum-exercise

We will need a number of key facts about the relationship between homotopy and homology. First notice the obvious map $h_n\colon \pi_n(X) \to H_n(X;\mathbb{Z})$, given by sending a map $f\colon S^n \to X$ to $f_*(x)$, where $x \in H_n(S^n;\mathbb{Z}) = \mathbb{Z}$ is a canonical generator.

1.8. THEOREM (Hurewicz Theorem). *Suppose that $\pi_n(X) = 0$ for $n < N$. Then $H_n(X;\mathbb{Z}) = 0$ for $n < N$ also, and moreover the Hurewicz map in dimension N, $\pi_N(X) \to H_N(X;\mathbb{Z})$, is an isomorphism.*

1.9. THEOREM (Whitehead Theorem). *Let $f\colon X \to Y$ be a map of connected CW-complexes inducing an isomorphism on all homotopy groups, or equivalently[2] inducing an isomorphism on π_1 and on all homology groups. Then f is a homotopy equivalence.*

[2] The equivalence follows from the Hurewicz theorem.

The reader will find the proofs of these results in [**5**, Chapter 4].

1.10. EXERCISE. Let M be a manifold of dimension $2k$ or $2k+1$. Show that if M is k-connected, then it is a *homotopy sphere* (i.e., homotopy equivalent to a sphere). (Use Poincaré duality and the Hurewicz and Whitehead theorems.)

homotopy-sphere

1.11. EXERCISE. Show that the smooth 4-manifolds $S^2 \times S^2$ and $\mathbb{CP}^2\#(-\mathbb{CP}^2)$ have isomorphic homotopy groups (in all dimensions), but are *not* homotopy equivalent. This shows that the condition of Whitehead's theorem cannot be weakened to *abstract* isomorphism of homotopy groups; it is necessary that the isomorphisms be induced by a map of spaces.

One way to show that these manifolds are not homotopy equivalent is to show that one has even intersection form but the other does not. On the other hand, one can represent $\mathbb{CP}^2\#(-\mathbb{CP}^2)$ as the total space of an S^2-bundle over S^2 which admits a cross section (see Exercise 1.7). Then its homotopy groups can be computed using the long exact homotopy sequence of a fibration.

2. Characteristic Classes

If M is a smooth manifold then its smooth structure provides a canonical (real) vector bundle, the *tangent bundle* TM over M. One can think of this as follows: let $\{U_\alpha\}$ be a coordinate cover of M; then the differentials of the transition functions of this cover provide maps $\varphi_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(n, \mathbb{R})$ that satisfy the *cocycle condition*

$$\varphi_{\alpha\beta}\varphi_{\beta\gamma}\varphi_{\gamma\alpha} = 1$$

where defined. Such a cocycle with values in $GL(n, \mathbb{R})$ can be used to construct a vector bundle by using the isomorphisms $\varphi_{\alpha\beta}$ to patch together trivial $\mathbb{R}^n$-bundles over U_α and U_β. Any cocycle with values in $GL(n, \mathbb{R})$ is cohomologous to one with values in the maximal compact subgroup $O(n)$ (one then speaks of a *reduction of structure group* to $O(n)$); this corresponds to the fact that every manifold can be given a Riemannian metric.

Diffeomorphic smooth manifolds have isomorphic tangent bundles. Therefore, invariants of smooth structure will be found from the *characteristic classes* of the tangent bundle.

Recall that a *characteristic class* for a certain category of bundles (the categories of real vector bundles and of complex vector bundles are the immediate examples) is just a natural map which associates, to each such bundle E over a base space B, a cohomology class $c(E) \in H^*(B)$, in such a way that isomorphic bundles E and E' have equal characteristic classes $c(E) = c(E')$. References for the theory include the classic notes by Milnor and Stasheff [**13**] and the more modern[3] book by Hatcher [**?**].

The most important characteristic classes for real vector bundles are the *Pontrjagin classes*. For a real vector bundle E over base B, these classes $p_k(E) \in H^{4k}(B; \mathbb{Z})$, $k = 1, 2, \ldots$ vanish for $k > \frac{1}{2}\dim E$, all vanish for a trivial bundle, and satisfy the *Whitney sum formula*: if we denote by $p(E)$ the 'total Pontrjagin class'

$$p(E) = 1 + p_1(E) + p_2(E) + \in H^*(B; \mathbb{Z})$$

then

$$p(E_1 \oplus E_2) = p(E_1) \cdot p(E_2) \quad \text{modulo 2-torsion.}$$

(The dot of course denotes the cup-product in the cohomology ring.)

Here is a very abbreviated account of the construction of the Pontrjagin classes. In classical differential geometry one encounters the *Gauss map* of an embedded k-submanifold $M \subseteq \mathbb{R}^n$. This is the map which to each point $m \in M$ associates the tangent

[3]But at this time unfortunately unfinished

plane to M at m, translated so as to pass through the origin in $\mathbb{R}^n$. It is a map from M to the *Grassmannian* $G_{k,n}(\mathbb{R})$ of k-dimensional subspaces of $\mathbb{R}^n$. The Grassmannian carries a 'tautological' k-dimensional vector bundle, and (by construction) the tangent bundle of M is the pull-back of this tautological bundle via the Gauss map. More generally, it is possible to show that *any* real vector bundle (at least over a compact base) is pulled back by some map from the universal bundle over some Grassmannian, and moreover the map is uniquely determined up to homotopy by the isomorphism class of the original bundle. This argument (sometimes called the *Yoneda lemma*) reduces the problem of finding characteristic classes to that of computing the cohomology of Grassmannians.

We denote the limit $\lim_{n\to\infty} G_{k,n}(\mathbb{R})$ by $BO(k)$ and call it the *classifying space* for bundles with structure group $O(k)$, that is k-dimensional real bundles. This construction is in fact a homotopy-theoretic one: for any group G, a space BG is defined uniquely up to homotopy equivalence by the requirement that it carry a *universal* principal G-bundle, one from which any other G-bundle is pulled back. (It turns out to be equivalent to require that the total space, denoted EG, of the universal bundle is contractible.) For similar reasons we denote by $BU(k)$ the limit $\lim_{n\to\infty} G_{k,n}(\mathbb{C})$, using the Grassmannian of k-dimensional *complex* subspaces of $\mathbb{C}^n$.

1.12. EXAMPLE. The spaces $BO(1)$ and $BU(1)$ are the infinite-dimensional real and complex projective spaces $\mathbb{RP}^\infty$ and $\mathbb{CP}^\infty$.

Although our interest is ultimately in real vector bundles, it turns out to be important to focus first on the classifying space $BU(1) = \mathbb{CP}^\infty$ for *complex* line bundles. This has a cell structure with cells only in even dimensions, and so its cohomology is $\mathbb{Z}$ in even dimensions and 0 in odd dimensions. Moreover, the cup-product of the generators in dimensions $2m$ and $2n$ is the generator in dimension $2(m+n)$ (geometric interpretation: in projective geometry the intersection of a codimension-m linear subspace and a codimension-n linear subspace is always a codimension-$(m+n)$ linear subspace). Thus

1.13. PROPOSITION. *The integral cohomology ring $H^*(BU(1);\mathbb{Z})$ is a polynomial ring $\mathbb{Z}[c]$ on one 2-dimensional generator.*

What this means for characteristic classes is that every complex line bundle L over a space X has a *first Chern class* $c_1(L) \in H^2(X;\mathbb{Z})$, and every other characteristic class for complex line bundles is just a polynomial in the first Chern class.

There are many other ways to define $c_1(L)$. For instance, the exponential map gives a short exact sequence of sheaves

$$0 \to \mathbb{Z} \to \mathcal{O}(\mathbb{R}) \to \mathcal{O}(S^1) \to 0;$$

and the associated Bockstein homomorphism $H^1(X;\mathcal{O}(S^1)) \to H^2(X;\mathbb{Z})$ maps a line bundle to its first Chern class.

What can be said about k-dimensional complex vector bundles? A simple example of such a bundle is a direct sum of k line bundles. It is a surprising fact that, for the purpose of characteristic class theory, one need only consider maps that *split* in this way. Here is the reason: Consider the product $BU(1) \times \cdots \times BU(1)$ (k copies). The cohomology of this space is a polynomial ring $\mathbb{Z}[x_1, \ldots, x_k]$, where $x_1, \ldots, x_k$ are the first Chern classes of the canonical line bundles over the various factors. The direct sum of all these line bundles is a k-dimensional vector bundle and this gives us a map

$$BU(1) \times \cdots \times BU(1) \to BU(k)$$

which classifies it. Now one has

1.14. PROPOSITION (Splitting Principle). *The map displayed above induces an injection on cohomology, whose image is the ring of* symmetric *polynomials in $x_1, \ldots, x_k$.*

It is a theorem of algebra [**8**, reference] that the ring of symmetric polynomials in $x_1, \dots, x_k$ is itself a polynomial ring, generated by the *elementary symmetric polynomials*

$$\begin{aligned} c_1 &= x_1 + \cdots x_k \\ c_2 &= x_1 x_2 + \cdots + x_{k-1} x_k \\ &\cdots \\ c_k &= x_1 \cdots x_k \end{aligned}$$

which are defined in general by

$$1 + c_1 t + c_2 t^2 + \cdots + c_k t^k = \prod_{i=1}^{k} (1 + t x_i).$$

Thus $H^*(BU(k); \mathbb{Z}) = \mathbb{Z}[c_1, \dots, c_k]$ where the generators c_i (of degree $2i$) are called the *i'th Chern classes*. These are the fundamental characteristic classes for k-dimensional complex vector bundles. Notice that the construction immediately gives us the Whitney sum formula for Chern classes,

$$c(V_1 \oplus V_2) = c(V_1) \cdot c(V_2),$$

where the total Chern class is defined by $c(V) = 1 + c_1(V) + c_2(V) + \cdots$.

1.15. EXERCISE. Show that $c_1(L_1 \otimes L_2) = c_1(L_1) + c_1(L_2)$ for complex line bundles L_1 and L_2.

1.16. EXERCISE. The *Chern character* is the characteristic class defined by the sum $e^{x_1} + \cdots + e^{x_k}$ (this is a symmetric formal power series rather than a symmetric polynomial, but things work in the same way). Using the previous exercise, show that the Chern character is a 'homomorphism' in the sense that

$$\mathrm{ch}(E_1 \oplus E_2) = ch(E_1) + \mathrm{ch}(E_2), \quad \mathrm{ch}(E_1 \otimes E_2) = \mathrm{ch}(E_1) \cdot \mathrm{ch}(E_2).$$

Now let us think about *real* rather than complex vector bundles. The process of complexifying (tensoring with $\mathbb{C}$) turns real vector bundles into complex ones and therefore provides a map $BO(k) \to BU(k)$. This pulls back the Chern classes to certain characteristic classes in $H^*(BO(k); \mathbb{Z})$. It turns out that the pullbacks of the *odd* Chern classes are 2-torsion elements (this is because the complexification of a real vector bundle is isomorphic to its complex conjugate bundle) but the pullbacks of the *even* Chern classes are significant and up to sign give the *Pontrjagin classes*

$$p_i(V) = (-1)^i c_{2i}(V \otimes \mathbb{C})$$

which generate a polynomial subring $\mathbb{Z}[p_1, p_2, \dots]$ of $H^*(BO(k); \mathbb{Z})$. Note that p_i has degree $4i$.

1.17. REMARK. When M is a smooth manifold, we refer to the 'Pontrjagin classes of M' instead of the Pontrjagin classes of the tangent bundle of M. By construction, these are diffeomorphism invariants of M.

mplex-projective-pont

1.18. EXAMPLE. Let us calculate the Pontrjagin classes of $M = \mathbb{CP}^n$, considered as a real $2n$-manifold. We recall that the cohomology of $\mathbb{CP}^n$ is a truncated polynomial ring $\mathbb{Z}[x]/(x^{n+1})$, where $x \in H^2(M; \mathbb{Z})$ is the first Chern class of the tautological line bundle L over M.

First we need

1.19. EXERCISE. Let T be the complex tangent bundle to M. Then one has an isomorphism of bundles $T \oplus \mathbb{C} = (n+1)\bar{L} = \bar{L} \oplus \cdots \oplus \bar{L}$. (Hint: Identify sections of $\bar{L}$ with homogeneous functions on $\mathbb{C}^{n+1}$, that is functions $f\colon \mathbb{C}^{n+1} \to \mathbb{C}$ such that $f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ for all $\lambda \in \mathbb{C}$. Identify sections of the bundle $T \oplus \mathbb{C}$ with homogeneous vector fields on $\mathbb{C}^n$. Choose a basis of $\mathbb{C}^n$ to get the desired isomorphism.)

It follows from the Whitney sum formula that $c(T) = (1+x)^{n+1}$. Now the complexification of the *real* tangent bundle to M (which is just the *real* vector-bundle underlying T) is isomorphic (as a *complex* vector-bundle) to $T \oplus \overline{T}$, and thus has total Chern class

$$c(T \oplus \overline{T}) = (1 - x^2)^{n+1}.$$

By definition, then, the k'th Pontrjagin class $p_k(M)$ is equal to $(-1)^k$ times the degree $2k$ term in the above polynomial, so it is equal to $\binom{n+1}{k} x^{2k}$. For instance, $p_1(\mathbb{CP}^2) = 3x^2$, $p_1(\mathbb{CP}^4) = 5x^2$, $p_2(\mathbb{CP}^4) = 10x^4$.

qpp-pont

1.20. EXERCISE. Calculate the Pontrjagin classes of quaternion projective space by a similar method. You should find that the total Pontrjagin class $p(\mathbb{HP}^n)$ equals $(1+x)^{2k+2}(1+4x)^{-1}$, where x is the generator of $H^4(\mathbb{HP}^n;\mathbb{Z})$; in particular, $p_1(\mathbb{HP}^n) = (2n-2)x$. See [**1**, page 519] or [**13**, Problem 20A]. Deduce that if $n > 1$, $\mathbb{HP}^n$ does not admit any orientation-reversing diffeomorphism.

If M is a *compact, oriented* manifold we define the *Pontrjagin numbers* of M to be the integers obtained by evaluating polynomials in the Pontrjagin classes on the fundamental homology class[4] $[M]$. Thus there is one Pontrjagin number for each polynomial in $\mathbb{Z}[p_1, p_2, \dots]$ of total degree equal to $\dim M$.

pont-cobord

1.21. LEMMA. *If the compact, oriented manifold M is the boundary of a compact manifold W, then all its Pontrjagin numbers vanish.*

PROOF. The restriction of TW to M is TM plus the normal bundle of M in W, which is a real line bundle and is trivial because of the orientation condition. Thus (by naturality and the Whitney sum formula) the Pontrjagin classes for M are the restrictions to M of cohomology classes on W. But any such restriction pairs trivially with $[M]$ (for those brought up on de Rham cohomology this is Stokes' theorem, but it can also be seen by calculating with exact sequences). □

This simple result shows the connection between Pontrjagin numbers and cobordism.

1.22. DEFINITION. Two compact oriented manifolds M and M' are *cobordant* if $M \sqcup (-M')$ is the boundary of a compact oriented manifold. The *oriented cobordism ring* Ω_* is the graded ring of cobordism classes of compact oriented manifolds: addition is by disjoint union, and multiplication is by Cartesian product.

From lemma 1.21 we see that each Pontrjagin number gives a group homomorphism $\Omega_* \to \mathbb{Z}$. Thom's computations of cobordism [**?**] (which we will review in Chapter **??**) showed that the Pontrjagin numbers are sufficiently rich to separate points on $\Omega_* \otimes \mathbb{Q}$. To put this another way, every group homomorphism $\Omega_* \to \mathbb{Z}$ is a Pontrjagin number with rational coefficients (an element of $\mathbb{Q}[p_1, p_2, \dots]$).

Now there is a completely different way to obtain a homomorphism from Ω_* to $\mathbb{Z}$: make use of Poincaré duality. We've seen above that every compact oriented manifold has a *signature*, defined using the intersection form on middle-dimensional cohomology, and it is not hard to check[5] that this quantity is cobordism invariant, so it defines a functional $\Omega_* \to \mathbb{Z}$. According to Thom's results, then, the signature is a Pontrjagin number. What number is it?

In low dimensions we can do some computations by hand. For instance, in dimension 4, the only Pontrjagin numbers are multiples of p_1. But for $M = \mathbb{CP}^2$ the signature is 1, whereas the Pontrjagin number $p_1(M)$ is 3, by the calculations of Example 1.18. Consequently we obtain

$$\mathrm{Sign}(M) = \frac{1}{3} p_1(M) \tag{1.23}$$

for any compact oriented 4-manifold M.

---

[4]Since the fundamental homology class depends on the choice of orientation, the Pontrjagin numbers depend on the choice of orientation, even though the Pontrjagin classes do not.

[5]See Proposition 2.48.

In dimension 8 there most general Pontrjagin number is $ap_1^2 + bp_2$, for some coefficients $a, b \in \mathbb{Q}$. Using the calculations of Example 1.18 again we obtain the equations

$$25a + 10b = 1, \qquad 18a + 9b = 1$$

by considering the 8-manifolds $M = \mathbb{CP}^4$ and $M = \mathbb{CP}^2 \times \mathbb{CP}^2$ respectively. These equations can be solved to yield $a = -1/45$, $b = 7/45$ and thus the formula

sig-8 (1.24)
$$\mathrm{Sign}(M) = \frac{1}{45}(7p_2 - p_1^2)[M]$$

for any compact oriented 8-manifold.

The general result was found by Hirzebruch. The *Hirzebruch Signature Theorem* gives an explicit procedure, in terms of certain power series, to build a characteristic class $L(M) = L(p_1, p_2, \dots)$, which in each degree is a polynomial in the Pontrjagin classes, such that

$$\mathrm{Sign}(M) = \langle L(M), [M] \rangle$$

for any compact oriented manifold M. The signature theorem expresses a deep and unexpected link between the algebra of intersection forms and the geometry of the tangent bundle. As we will see in a moment, it has very strong geometrical consequences.

1.25. REMARK. The L-class has components in degrees $0, 4, 8, \dots$. Moreover, by examining its explicit form one sees that, in rational cohomology $H^*(M; \mathbb{Q})$, the Pontrjagin classes can be recovered from the L-class. On the other hand, by the signature theorem the L-class determines not only the signature of M but also the signature of any submanifold N of M that has trivial normal bundle. (For then the Pontrjagin classes of M restrict to those of N.) Using arguments from homotopy theory (specifically Serre's theorem about the finiteness of the higher homotopy groups of spheres) it can be shown that there is also a converse here: to know the signatures of submanifolds with trivial normal bundle (in M and in certain 'stabilizations' of M) recovers the rational L-class. The conclusion is that the rational Pontrjagin classes determine and are determined by a list of signatures of submanifolds. Many of the deeper properties of Pontrjagin classes in differential topology depend on this fact.

There are other characteristic classes for real vector bundles, such as the Stiefel-Whitney classes in mod 2 cohomology, and the Euler class for oriented bundles, but we will not go into details here.

3. The Poincaré Conjecture

In the middle 1950s, shortly after the publication of the Hirzebruch signature theorem, Milnor was trying to understand the structure of $(n-1)$-connected manifolds of dimension $2n$. (His paper [**9**] gives some of the history.) Classical examples would be the complex projective plane $\mathbb{CP}^2$ of dimension 4, the quaternionic projective plane $\mathbb{HP}^2$ of dimension 8, and the Cayley projective plane of dimension 16. Each of these has $\pi_n(M) = \mathbb{Z}$, and $\pi_n(M)$ is generated by a single embedded n-sphere S^n in M. In an effort to generalize this construction, Milnor considered n-dimensional real vector bundles V over S^n. Taking the disk bundle of such a V gives a compact $2n$-manifold with boundary, say W; and if the closed $(2n-1)$-manifold ∂W happens to be a sphere, then we can attach a $2n$-disk to it and thus obtain a possibly exotic closed $2n$-manifold M.

The bundles V are classified by their 'clutching functions', which are maps $S^{n-1} \to SO(n)$ (or equivalently maps $S^n \to BSO(n)$, using the theory of classifying spaces discussed in the previous section). To begin his study Milnor asked for what choices of clutching function would the manifold ∂W constructed above have the *homotopy type* of a sphere.

Consider first the case $n = 2$. In this case the 2-plane bundles V over S^2 are completely determined by a single integer k in $\pi_1 SO(2) = \mathbb{Z}$ (that is the Euler class). The manifold ∂W is the total space of an S^1-bundle over S^2, and part of the homotopy exact sequence associated to this is

$$\pi_2(S^2) \xrightarrow{\times k} \pi_1(S^1) \longrightarrow \pi_1(\partial W) \longrightarrow \pi_1(S^2) = 0.$$

We see that ∂W is simply-connected (and thus a homotopy sphere) if and only if $k = \pm 1$. In this case the resulting 4-manifold M is simply $\pm\mathbb{CP}^2$, so the construction yields nothing new.

Look now at the case $n = 4$. The bundles V are 4-plane bundles over S^3, classified up to isomorphism by the homotopy class of the clutching map $S^3 \to SO(4)$, that is an element of the homotopy group $\pi_3(SO(4))$. One knows that the simply connected double cover of $SO(4)$ is $S^3 \times S^3$ (to see how an element of $S^3 \times S^3$ gives rise to a rotation, think of the points of S^3 as unit quaternions and associate to $(u, v) \in S^3 \times S^3$ the rotation $x \mapsto uxv$ of $\mathbb{H} = \mathbb{R}^4$). This gives us the calculation

$$\pi_3(SO(4)) = \pi_3(S^3 \times S^3) = \mathbb{Z} \oplus \mathbb{Z}.$$

So the possible bundles V are classified by *pairs* of integers i, j.

Now investigate what is the condition on i, j for the manifold ∂W constructed as above to be a homotopy sphere. W is the total space of an S^3-bundle over S^4 and part of the homotopy exact sequence associated to this is

$$\pi_4(S^4) \xrightarrow{\times(i+j)} \pi_3(S^3) \longrightarrow \pi_3(\partial W) \longrightarrow \pi_3(S^4) = 0.$$

Thus we conclude that W will be 3-connected (and therefore a homotopy sphere, see Exercise 1.10) if and only if $i + j = \pm 1$. In contrast to the case $n = 2$, this gives infinitely many possibilities. Let us fix $i + j = 1$ and consider the corresponding 8-manifolds W_i and their boundaries, the homotopy 7-spheres ∂W_i.

If $i = 1$, then $\partial W_i = S^7$; in fact, the 8-manifold M obtained by attaching a disk to W_1 is simply quaternion projective space. If $i = 2$, though, something strange happens. To see this, suppose for a moment that ∂W_i is also (diffeomorphic to) the 7-sphere, and let M_i be the closed 8-manifold obtained by attaching a disk. We ask: What are the Pontrjagin classes of M_i? Since the generator of $H_4(M; \mathbb{Z})$ is just the sphere S^4 that we started with, the Pontrjagin class $p_1(M_i)$ can be computed in a neighborhood of S^4, and thus from the data $i, j = 1 - i$ alone.

1.26. EXERCISE. Show that in the above situation we have $p_1(M_i) = 2(i - j) = 2(2i - 1)$ times the generator of $H^4(M; \mathbb{Z}) = \mathbb{Z}$. Check that this fits with the calculation of Pontrjagin classes for the quaternionic projective plane (Exercise 1.20).

The signature of M must be 1 (if we choose the orientation suitably) so the signature theorem for 8-manifolds, equation 1.24, yields

$$p_2[M] = \frac{p_1^2[M] + 45}{7} = \frac{4(2i-1)^2 + 45}{7}.$$

If i=1 this gives $p_2[M] = 7$, consistent with the calculations earlier (Exercise 1.20) for the quaternionic projective plane. But if $i = 2$ then we get $p_2[M] = 81/7$, which is ridiculous; Pontrjagin numbers are integers! The same integrality problem arises for any i not congruent to 0 or 1 modulo 7.

What can be the problem? The supposed smooth 8-manifold M cannot exist, and this means that the homotopy 7-sphere $\Sigma = \partial W$ cannot, after all, be the standard 7-sphere S^7. At this point two possibilities present themselves:

(a) Perhaps Σ is a homotopy 7-sphere which is not homeomorphic to the standard 7-sphere S^7 (and thus a counterexample to the Poincaré conjecture in dimension 7, see below)?

(b) Or, perhaps Σ is a smooth manifold homeomorphic but not diffeomorphic to S^7 — an 'exotic sphere'?

Milnor has recorded that he at first inclined to the view that (a) was true, but in fact the solution turned out to be (b), a conclusion that he announced in the revolutionary paper [**10**].

This is an appropriate point to state the Poincaré conjecture.

1.27. CONJECTURE (Generalized Poincaré Conjecture). *Every smooth homotopy n-sphere (that is, every smooth manifold homotopy equivalent to S^n) is homeomorphic to S^n.*

To do

Write a historical section about the PC. See Dieudonné, etc

In order to prove the Poincaré Conjecture one needs some mechanism for recognizing smooth manifolds homeomorphic to S^n. Such a mechanism is provided by the following theorem of Reeb.

1.28. THEOREM. *Let M be a compact smooth manifold. Suppose that $f\colon M \to \mathbb{R}$ is a smooth function having no critical points except for a single non-degenerate maximum and a single non-degenerate minimum. Then M is homeomorphic to a sphere.*

A *critical point* of f is a point where its gradient vanishes, and such a critical point is *non-degenerate* if the matrix of *second* derivatives of f has full rank there.

SKETCH PROOF. It is known that around a non-degenerate minimum point one can choose local coordinates so that

$$f(x_1, \ldots, x_n) = c + x_1^2 + \cdots + x_n^2$$

where $c = f(0, \ldots, 0)$ is the minimum value of f. (This is part of the *Morse Lemma* **??**.) Consequently, for sufficiently small $\varepsilon > 0$ the region $\{x : f(x) \leqslant c + \varepsilon\}$ is a closed n-disk in M. If we remove from M the interior of this disk, and of the corresponding disk around the maximum point, the part of M that remains can be given the structure of a cylinder $S^{n-1} \times I$ by making use of the gradient flow of f (see Figure 1). Thus M can be obtained by attaching two disks, by diffeomorphisms, to the ends of a cylinder. Since every homeomorphism of the boundary of a closed disk extends, by 'coning', to a homeomorphism of the whole disk, the resulting manifold is homeomorphic to the n-sphere. $\square$

1.29. REMARK. The process of extending a homeomorphism of a sphere to a homeomorphism of the disk that it bounds is called the *Alexander trick.* Note carefully that even if we start with a *diffeomorphism* of the sphere, the homeomorphism produced by the Alexander trick need not be smooth at the cone point (though of course it will be smooth everywhere else).

1.30. EXERCISE. Consider the Milnor 7-manifold M described above. Show that it can be obtained by identifying two copies of $\mathbb{R}^4 \times S^3$ in the following explicit way: the point (u, v) in the first copy of $(\mathbb{R}^4 \setminus \{0\}) \times S^3$ is identified with (u', v') in the second copy, where

$$u' = u/\|u\|^2, \quad v' = u^i v u^j / \|u\|$$

Size: 60 mm

FIGURE 1. The gradient flow provides a diffeomorphism to a cylinder grad-flow-fig

(using quaternion multiplication). Check that, if $i + j = 1$, the function

$$f(u, v) = \Re v/(1 + \|u\|^2)$$

extends smoothly to the whole of M and has precisely two critical points, both non-degenerate. Deduce that M is homeomorphic to S^7.

A few years after Milnor's work, Smale proved the Poincaré conjecture in high dimensions. (His famous remark that the proof occurred to him 'on the beaches of Rio' caused some upset back in the USA, see [**16**].) The idea of the proof is to study manifolds-with-boundary which have the homotopy-theoretic properties of the middle, cylindrical, region in the proof of Reeb's theorem above.

1.31. DEFINITION. Let W be a cobordism, that is, a compact manifold with boundary, whose boundary has two components $\partial_- W$ and $\partial_+ W$. It is said to be an *h-cobordism* if the inclusions $\partial_- W \to W$ and $\partial_+ W \to W$ are both homotopy equivalences.

A simple example of an h-cobordism is $M \times [0, 1]$, where M is compact without boundary. This is called a *product* cobordism. Smale proved

1.32. THEOREM (h-cobordism theorem). *Any simply-connected h-cobordism W of dimension $\geqslant 6$ is diffeomorphic to a product. In particular, $\partial_- W$ and $\partial_+ W$ are diffeomorphic to one another.*

We shall devote a large part of Chapter **??** to discussion of this theorem. The proof works with a smooth real-valued function $f\colon W \to \mathbb{R}$, constant on the two boundary components and having only non-degenerate critical points — a *Morse function*. If f has *no* critical points at all then we can use the gradient flow as in Reeb's proof to show that W is a product; so the idea is to modify f by 'canceling' its critical points until none are left. To give a simple example of how this might work, the cubic function on $\mathbb{R}$ given by $x \mapsto x^3 + 3x^2$ has critical points at 0 and -2; as one varies the function in the family $x^3 + 3x^2 + 3\lambda x$, $\lambda \in [0, 2]$, the two critical points coalesce (at $\lambda = 1$) and then both disappear. In order to carry out this cancellation in general there are some topological necessary conditions that must be satisfied (the h-cobordism condition) and the main part of the proof is to show geometrically that when these necessary conditions are satisfied, cancellation can always be carried out.

When we discuss the h-cobordism theorem we will, for technical reasons, use the language of *handlebody decompositions* rather than that of Morse functions. For a very careful account of the proof in Morse-theoretic language see [**12**].

Granted the h-cobordism theorem, the proof of the Poincaré conjecture, at least in dimensions 6 and above, is easy. We just follow the outline of the proof of Reeb's theorem, above. Let Σ be a homotopy sphere. Remove two small, disjoint disks. The resulting manifold-with-boundary is a simply-connected h-cobordism, hence a product. Gluing the disks back in gives a homeomorphism to the standard sphere, via the Alexander trick.

1.33. EXERCISE. Poincaré at first asked whether every *homology sphere* (a manifold having the same homology groups as S^n) is a standard sphere. However, he soon produced an example to show that the answer is 'no' in general [**?**]. Let us look at manifolds of the form $M = S^3/\Gamma$, where Γ is a discrete subgroup of $SO(4)$ acting freely on S^3. Show that if the group Γ is equal to its commutator subgroup $[\Gamma, \Gamma]$ (this is what is called a 'perfect' group), then M is a homology 3-sphere.

To get an explicit example, regard $S^3 = Sp(1)$ as the group of unit quaternions, which is the double cover of $SO(3)$. The inverse image of the symmetry group of the icosahedron, under this double cover, is a subgroup of $Sp(1)$ of order 120, called the *binary icosahedral group*. Show that the binary icosahedral group is perfect (use the fact that the symmetry group of the icosahedron is nonabelian and simple). Thus we obtain a homology sphere by dividing S^3 by Γ acting by group multiplication.

To do

refer to paper by Kirby and Schnarlemann, 8 faces of the Poincaré homology sphere. Connection with plumbing? (later)

theta-group

1.34. EXERCISE. Let Θ^n denote the collection of diffeomorphism classes of homology n-spheres. Show that the operation of connected sum gives a commutative, associative 'addition' on Θ^n for which the standard n-sphere is the identity element.

Let G be the group of diffeomorphisms of S^{n-1}, and let G_0 be the connected component of the identity in G. Show that the process of gluing two n-disks by a diffeomorphism gives a homomorphism (of monoids)

$$G/G_0 \to \Theta^n.$$

Using the h-cobordism theorem, show that this is actually an *isomorphism*. In particular, Θ^n is a group.

Θ^n is an example of a *structure set*, listing the smooth manifold structures within a given homotopy type. Surgery theory gives a systematic means to compute these structure sets.

pont-vary-sect

4. Variation of the Pontrjagin Classes

The Poincaré conjecture shows that for spheres, homotopy type determines homeomorphism type. This is not always true for more complicated manifolds. In this section we shall construct an example of a homotopy equivalence $f\colon M \to M'$ of smooth manifolds which does not preserve the Pontrjagin classes:

$$f^*(p_1(M')) \neq p_1(M) \in H^4(M;\mathbb{Q}).$$

It follows immediately that f cannot be homotopic to a diffeomorphism.

Once again the construction uses bundle theory. Let us consider 5-dimensional oriented vector bundles V over S^4. These are classified up to isomorphism by the homotopy classes of their clutching maps, which are elements of $\pi_3(SO(5))$. It is known that this group is the integers, $\mathbb{Z}$. Moreover, the integer $k \in \pi_3(SO(5))$ that classifies the bundle is just the Pontrjagin class $p_1(V) \in H^4(S^4)$.

One way to see this is to start with $\pi_3(SO(4)) = \mathbb{Z} \oplus \mathbb{Z}$ (see previous section) and calculate homotopy groups using the long exact sequence of the fibration $SO(4) \to SO(5) \to S^4$. Alternatively, the result is a special case of the Bott periodicity theorem for the orthogonal group. The statement about the Pontrjagin class follows from the rational Hurewicz isomorphism $\pi_4(BSO) \otimes \mathbb{Q} = H_4(BSO; \mathbb{Q})$.

Taking the boundary of the disk bundle associated to $V \oplus \mathbb{R}$, where $\mathbb{R}$ denotes a 1-dimensional trivial line bundle, we obtain a family M_k of closed 9-manifolds parameterized by integers $k \in \pi_3(SO(5)) = \mathbb{Z}$. The classification of these manifolds M_k up to homotopy type depends on the homotopy class of the clutching map, now considered as a map from $S^3 \times S^5 \to S^5$. Basepoints are preserved (if we take the basepoint in S^5 to be the 'north pole' associated to the added trivial line bundle) so that the clutching map is actually a map from $S^8 = S^3 \wedge S^5$ to S^5, and its homotopy class is an element of $\pi_8(S^5)$. Serre's results show that this group is $\mathbb{Z}/24$, so that if k is divisible by 24 the manifold M_k is homotopy equivalent to $M_0 = S^4 \times S^5$.

Lurking just beneath the surface of this discussion is a famous and important construction of homotopy theory, the *J-homomorphism*, which is the map $\pi_k(SO(m)) \to \pi_{m+k}(S^m)$ obtained by making $SO(m)$ act on S^m by rotations about the polar axis.

On the other hand, the trivial bundle factor gives a cross section to the fibration $S^5 \to M \to S^4$. This cross section is a copy N of S^4 which generates $H_4(M)$, and its normal bundle ν_N in M is just the original vector bundle V. Thus, evaluating on $[N]$ and using the Whitney sum formula

$$p_1(M) = p_1(\nu_N) + p_1(TN) = p_1(V) + 0 = k$$

since the tangent bundle to N (as to any sphere) is stably trivial. We conclude that $M_0 = S^4 \times S^5$ and M_{24} are homotopy equivalent, but their first Pontrjagin classes are different. The homotopy equivalence between them therefore cannot be homotopic to a diffeomorphism.

1.35. REMARK. We have chosen to work with a particular example here, but it is clear that similar constructions could be based on any element of the kernel $\operatorname{Ker} J$.

As in the previous section, two possibilities now present themselves.

(a) Perhaps M_0 and M_{24} are homotopy equivalent but not homeomorphic?
(b) Or, perhaps M_{24} is a smooth manifold homeomorphic but not diffeomorphic to $S^4 \times S^5$ — an 'exotic product of spheres'?

This time however it is (a) that is the true statement; M_{24} is not even homeomorphic to $S^4 \times S^5$. This follows from a deep theorem of Novikov:

novikovstheorem

1.36. THEOREM ([**14, 15**]). *If $f\colon M \to M'$ is a homeomorphism between smooth manifolds, then $f^*(p_i(M')) = p_i(M)$ as elements of the* rational *cohomology groups $H^*(M; \mathbb{Q})$.*

This result, proved in the middle 1960s, lies much deeper than anything else we have mentioned in this introduction. To prove it, Novikov devised an elaborate inductive technique for applying the methods of surgery theory, on non-simply-connected smooth manifolds, to problems about homeomorphisms. We will return to the study of Novikov's theorem in Chapter **??**.

In fact anything that is not 'nailed down' by the signature theorem can be modified by a homotopy equivalence — we need to discuss this explicitly somewhere.

To do

5. The manifold structure set

We now begin our introduction to surgery theory proper. The most basic objects of interest in surgery theory are *manifold structure sets*, which classify the manifold structures within a given homotopy type. The group Θ^n of exotic spheres, introduced in Exercise 1.34, is an example of a structure set.

Surgery theory can be carried out in many different categories of manifolds (smooth, piecewise linear, topological and so on). The basic structure of the theory is the same, although the geometric tools need to be forged anew for each category. It's convenient to have some terminology which covers all categories. We'll therefore make the convention that CAT denotes a category of manifolds: we write $CAT = TOP$ when we refer to topological manifolds, $CAT = PL$ for piecewise linear, $CAT = DIFF$ for smooth.

1.37. DEFINITION. Let X be a finite complex. A *CAT structure* on X is a homotopy equivalence $f\colon M \to X$, where X is a CAT manifold. Two such structures $f\colon M \to X$ and $f'\colon M' \to X$ are *CAT equivalent* if there is a homotopy commuting diagram

$$\begin{tikzcd} M \arrow[rr] \arrow[dr, "f"'] & & M' \arrow[dl, "f'"] \\ & X & \end{tikzcd}$$

in which the horizontal arrow is a CAT homeomorphism. The CAT structure set of X, $\mathcal{S}^{CAT}(X)$, is the collection of equivalence classes of CAT structures on X.

1.38. EXAMPLE. The classical *Poincaré conjecture* is that $\mathcal{S}^{TOP}(S^3)$ has just one element, the standard structure on S^3.

The Poincaré conjecture is usually stated in the following form: any manifold homotopy equivalent to S^3 is homeomorphic to S^3. To see that this statement is the same as the one made above one needs to know that every self-homotopy-equivalence of S^3 (or of any sphere) is homotopic to a homeomorphism. This is true by the Hurewicz theorem: homotopy classes of maps $S^n \to S^n$ are classified by their effect on homology, and in particular any homotopy equivalence $S^3 \to S^3$ is homotopic either to the identity map or to a reflection.

1.39. EXAMPLE. The structure set $\mathcal{S}^{DIFF}(S^n)$ is just Θ^n. It can be shown that this is a finite group for every n (see Chapter **??**).

Θ^n is a group, but in general there is no obvious group structure on $\mathcal{S}(X)$. It is a deep theorem in topological surgery that $\mathcal{S}^{TOP}(M)$ has a natural group structure — the structure does not have an straightforward 'geometrical' definition.

1.40. REMARK. Smale's proof of the high-dimensional Poincaré conjecture says that the image of $\mathcal{S}^{DIFF}(S^n)$ in $\mathcal{S}^{TOP}(S^n)$ is trivial. A stronger statement is in fact true, namely that $\mathcal{S}^{TOP}(S^n)$ is itself trivial. This is due to Newman [**?**]. The corresponding statement with TOP replaced by PL is also true, and is due to Stallings [**?**].

1.41. EXAMPLE. The calculations of the previous section show that $\mathcal{S}^{DIFF}(S^4 \times S^5)$ is infinite. Granted Novikov's Theorem 1.36, so is $\mathcal{S}^{TOP}(S^4 \times S^5)$.

1.42. EXERCISE. Let X be the 2-torus with one meridian collapsed to a point, see Figure 2. Then X has no manifold structures, i.e. $\mathcal{S}^{TOP}(X)$ is empty.

The point of the previous exercise is that, if there is to be any chance that X has a manifold structure at all, then it surely must be a *Poincaré duality space* in the sense that it has a fundamental class which induces duality isomorphisms between homology and cohomology[6].

[6]Later, we will need to consider more sophisticated notions of what is meant by a Poincaré duality space, but this will do for now.

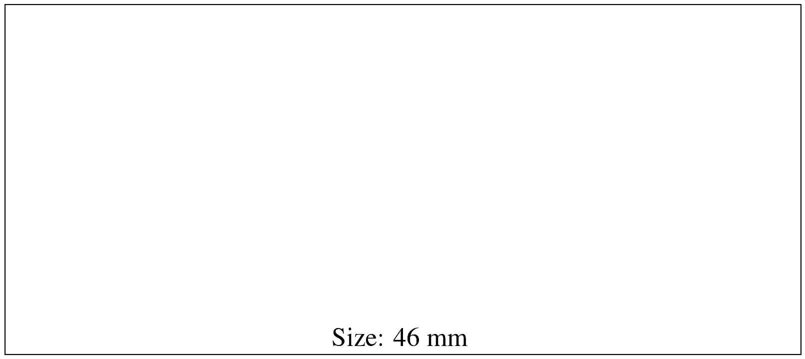

FIGURE 2. A homotopy type which supports no manifold structures squashed-torus

In 1962, Milnor and Kervaire [**6**] began the calculation of the groups of exotic spheres, introducing a new technique which they called ‘spherical modification’ but was soon renamed ‘surgery’. In a 1963 manuscript, Browder [**4**] showed how one can use these methods to study the more general problem of computing the structure set $\mathcal{S}(X)$ for a Poincaré duality space X. The starting point is to ask what additional conditions on a simply connected Poincaré duality space X are necessary to ensure that it admits *at least one* manifold structure.

6. Normal maps

If X is to have the structure of a DIFF manifold then there must be a vector bundle that will serve as its tangent bundle. In fact, it is often geometrically more convenient to think about the *stable normal bundle*. If X were a manifold and embedded in $\mathbb{R}^{n+k}$, k large, then it would have a tubular neighborhood diffeomorphic to the total space of a vector bundle ν_X, the *normal bundle* of the embedding. Once k is big enough, making it bigger just changes ν_X by the addition of a trivial bundle, so it makes sense to speak of a well-defined stable normal bundle. By construction, the tangent bundle plus the stable normal bundle is (stably) trivial.

The stable normal bundle cannot be prescribed arbitrarily. Its twofold interpretation (as a vector bundle over X and as a neighborhood of X in Euclidean space) gives a compatibility condition between Poincaré duality and the *Thom isomorphism* of the bundle ν_X. This condition is necessary (but not sufficient) for ν_X to be the stable normal bundle of a manifold structure on X.

Here is some background on the *Thom isomorphism*. Let V be a k-dimensional oriented vector bundle over a compact base space X (not necessarily a manifold). The *Thom space* $\operatorname{Th} V$ is the one-point compactification of the total space of V. The Thom isomorphism in cohomology is a natural isomorphism

$$\Phi^* \colon H^m(X) \to H^{m+k}(\operatorname{Th} V, \infty).$$

Dually, the Thom isomorphism in homology is a natural isomorphism

$$\Phi_* \colon H_m(\operatorname{Th} V, \infty) \to H_{m-k}(X).$$

One can think of these isomorphisms as given by cap or cup product with a generator for the cohomology of the fibers of M (the Thom class). The construction will be discussed in greater detail in Chapter **??**.

Suppose now that X is indeed a compact connected oriented n-manifold and that V is the normal bundle of some embedding of X in $\mathbb{R}^{n+k}$. Let us consider the group $H_{n+k}(\operatorname{Th} V, \infty)$. On the one hand, the total space of V is an open subset of Euclidean space $\mathbb{R}^{n+k}$. For any such open $U \subseteq \mathbb{R}^{n+k}$, there is an induced map one one-point compactifications

$$S^{n+k} = (\mathbb{R}^{n+k})^+ \to U^+$$

(note the 'wrong way' functoriality!), and the fundamental homology class of S^{n+k} thus gives rise to a homology class in $H_{n+k}(U^+, \infty)$, called a 'spherical' class. In fact, if U is connected, $H_{n+k}(U^+, \infty)$ is isomorphic to $\mathbb{Z}$ and is generated by a spherical class; this is a consequence of Alexander duality.

In particular then, $H_{n+k}(\mathrm{Th}\, V, \infty)$ is generated by a spherical class. On the other hand, the Thom isomorphism identifies this group with $H_n(X)$, a group that is generated by the fundamental homology class of X. Thus we obtain

1.43. CLAIM. In order that the Poincaré duality space X admit a manifold structure, it is necessary that there exist a k-vector bundle V over X, having the following property: if $\Phi\colon H_{r+k}(\mathrm{Th}\, V, \infty) \to H_r(X)$ denotes the Thom isomorphism, then $\Phi^{-1}([X])$ is a *spherical* class.

1.44. DEFINITION. A *normal invariant* for the Poincaré duality space X is a stable isomorphism class of vector bundles V as described in the Claim above. The pair (X, V) will be referred to as *normal data*.

1.45. REMARK. We shall see later that this notion can be expressed in homotopy-theoretic terms. In fact, any Poincaré duality space possesses a *Spivak normal bundle*, a spherical fibration canonically determined by the Poincaré duality structure. There is a forgetful functor from (stable) vector bundles to (stable) spherical fibrations, corresponding to a map of classifying spaces $BO \to BG$; a normal invariant is just a *reduction of structure* of the Spivak normal bundle to a vector bundle, or equivalently, a factorization of its classifying map $X \to BG$ through the space BO.

1.46. DEFINITION. A *normal map* associated to given normal data (X, V) is a degree-one map $f\colon M \to X$, where M is an oriented smooth manifold, such that $f^*(V)$ is (stably) isomorphic to the (stable) normal bundle of M.

We say that $f\colon M \to X$ is of *degree one* if $f_*[M] = [X]$, where the square brackets denote the fundamental homology classes; in particular the dimension of M equals the 'formal dimension' of X (the degree in which its fundamental class appears).

Surgery theory regards a normal map as a 'first approximation' to a homotopy equivalence from a manifold to X. The following theorem is therefore fundamental to the subject.

1.47. THEOREM. *Given normal data (X, V), there exist normal maps $f\colon M \to V$ associated to it.*

SKETCH OF PROOF. How shall we obtain a manifold M from normal data? The answer is to apply *transversality theory*. This theory — one of Thom's beautiful ideas — is about the 'generic' behavior of smooth maps. In its simplest form it concerns a smooth map between manifolds, $g\colon M^{n+k} \to N^n$. It is easy to see that any closed subset of M can appear as such an inverse image. Nevertheless, 'generically' the inverse image $g^{-1}\{p\}$ of a point $p \in N$ is a smooth k-dimensional submanifold. Notice that in linear algebra, 'generically' a linear map from $\mathbb{R}^{n+k}$ to $\mathbb{R}^n$ is surjective with k-dimensional kernel. The basic idea of transversality theory is that the generic behavior of *smooth* maps is modeled by the generic behavior of *linear* maps (which of course appear as the tangent maps to the smooth maps in question).

We will leave until later the question of what exactly is meant by 'generic'. Let us suppose that normal data (X, V) are given, and apply transversality to the map $g\colon S^{n+k} \to \mathrm{Th}\, V$ which is given to us by the assumption that the Thom class of V is spherical. What we mean by this is the following. Locally V is a product, so (away from the point at infinity) g looks like $(g_1, g_2)\colon U \to Y \times \mathbb{R}^k$, $Y \subseteq X$, $U \subseteq S^{n+k}$ and we may assume without loss of generality that g_2 is smooth. Transversality theory tells us that the 'generic' behavior of g_2 is as described above: the inverse image of a point (say the origin) is an n-dimensional submanifold. But the inverse image of a point under g_2 is just the inverse image of the *zero-section* of the bundle V under f. We therefore expect, and Thom's transversality theorem verifies, that 'generically' the inverse image of the zero section of $\mathrm{Th}\,V$ will be an n-dimensional submanifold of S^{n+k}, equipped with a map $f\colon M \to X$ (just the restriction of g) which pulls back V to the normal bundle of M. □

It should be underlined that this is a very *non-constructive* way to obtain the manifold M. The 'generic' perturbation of g cannot be precisely specified in advance.

1.48. REMARK. The proof is an example of the *Pontrjagin-Thom construction*, which relates homotopy theory to *bordism* theory. In fact, a refinement of the argument shows that each normal invariant corresponds exactly to a *normal bordism* class of normal maps. See **??** for more on this.

To do surgery theory in a category CAT, it is essential to have an appropriate CAT transversality theorem, which should tell us that in the situation of the above proof, the inverse image of the zero-section of the appropriate sort of CAT bundle (and of course this too needs to be defined) is a CAT submanifold. Thom's theorem gives DIFF transversality, and PL transversality (due to ??) incorporates still more directly the idea that transversality is a 'nonlinear version of general position'. By contrast, the absence of a satisfactory TOP transversality theory was for many years an obstruction to the development of the theory of TOP manifolds. The resolution of the problem and the construction of an appropriate transversality theory (due to Kirby-Siebenmann [**7**]) depends heavily on surgery theory for PL or DIFF manifolds.

7. Surgery on normal maps

We have seen that from normal data (X, V) it is always possible to construct normal maps $M \to V$ — in fact, the normal data determine an entire normal bordism class of normal maps. Browder now asks[7]: Does this normal bordism class contain at least one manifold structure, that is, a normal map which is actually a homotopy equivalence? This is where surgery proper enters the picture. Surgery gives a means of constructing normal bordisms, and conversely any normal bordism can be analyzed into a sequence of surgeries. The question is, therefore, whether starting with a 'generic' normal map produced by transversality, one can improve it by a sequence of surgeries until at last one obtains a homotopy equivalence.

Before giving the formal definition of a surgery let us think about its counterpart in pure homotopy theory (that is, without involving questions of manifold structure). This is the process of *attaching a cell*.

Let D^n denote the closed unit disk in $\mathbb{R}^n$, whose boundary S^{n-1} is the $(n-1)$-dimensional sphere.

1.49. DEFINITION. Let X be a topological space and $f\colon S^{n-1} \to X$ a map. The identification space

$$D^n \cup_f X$$

is said to be obtained from X be *attaching a cell*, with *attaching map f*.

[7] In the fundamental 1963 paper [] that we are following

For example, a finite CW-complex is just a space obtained by successively attaching cells to the empty set.

The construction in the following proposition is known as *Serre's procedure for killing homotopy groups*. Recall the definition of the homotopy groups of a map.

1.50. DEFINITION. Let $f\colon M \to X$ be a map. The homotopy group $\pi_r(f)$, $r \geqslant 1$, consists of homotopy classes of diagrams of continuous maps of the form

$$\begin{tikzcd} S^{r-1} \arrow[r] \arrow[d] & M \arrow[d, "f"] \\ D^r \arrow[r] & X \end{tikzcd}$$

map-homotopy-def

1.51. PROPOSITION. *Let $f\colon Y \to X$ be a map. Then there exist a space Y' obtained from Y by attaching cells*[8] *and a map $f'\colon Y' \to X$ extending f such that the groups $\pi_r(f')$ all vanish.*

PROOF. Each element of $\pi_r(f)$ is defined by a diagram as in Definition 1.50 above. For each such element, use the top row of the diagram as the attaching map of a cell to Y, and the bottom row to show how to extend f over this cell as a map to X. Details are left to the reader. $\square$

It is an easy consequence of Whitehead's Theorem that a map (between CW complexes) whose homotopy groups vanish is in fact a homotopy equivalence. Thus Serre's procedure shows us how to modify a map by successive cell attachments to get a homotopy equivalence. That is an easier homotopy-theoretic counterpart to our proposal to modify a normal map to get a homotopy equivalence. The procedure of attaching cells will not suffice for this as it stands, because attaching a cell to a manifold usually does not produce another manifold. We want to devise another procedure which is as similar as possible to the procedure of attaching cells but which also respects the manifold structure.

1.52. DEFINITION. Let M be a smooth manifold, and suppose that M contains an embedded sphere S^q whose normal bundle is trivial. Then S^q has a tubular neighborhood N which is a product $D^p \times S^q$. We note that

$$\partial N = S^{p-1} \times S^q = \partial N',$$

where $N' = S^{p-1} \times D^{q+1}$. The manifold M' *obtained from M by doing surgery on S^q* is defined by

$$M' = (M \setminus N) \cup_{\partial N'} N';$$

in words, M' is obtained from M by scooping out N and replacing it by N' instead.

1.53. EXAMPLE. The operation of connected sum is just surgery on an S^0.

We will skip for now the verification that this operation is well-defined, produces a manifold, and so on. Notice that the homotopy class of the embedded sphere S^q does vanish in M', in analogy with Serre's procedure. Unfortunately surgery messes up other homotopy groups as well, so it is not always clear that it effects any overall improvement. This problem will (eventually) lead us to the surgery obstruction groups.

---

[8] There may be infinitely many cells, but we don't stress this aspect of the construction as it is not relevant to the manifold case.

Size: 53 mm

FIGURE 3. A surgery

Our project is to kill the homotopy groups of a normal map $f\colon M \to X$ by means of a succession of surgeries. We will proceed inductively upwards over $r = 0, 1, \ldots$; once we have made f 1-connected, then we can take advantage of the Hurewicz theorem to say that we want to kill the *homology kernel*, that is the kernel of $f_*\colon H_*(M) \to H_*(X)$. It is convenient to maintain both points of view.

Two problems have to be addressed in carrying out the above procedure.

(a) The first problem is that to do surgery at all, the homotopy class we want to kill must be very geometrically represented. A mere map $S^r \to M$ will not suffice; we need an embedding, and moreover one whose normal bundle is trivial. We will therefore have to address the question of under what circumstances a homotopy class of maps from one manifold to another contains an embedding. The trivialization of the normal bundle turns out to be less of a difficulty; essentially, the facts that the tangent bundle to a sphere is stably trivial, the normal bundle to M is pulled back from X, and the embedding of S^r in M becomes contractible in X, combine to show that the normal bundle to S^r is *stably* trivial — and then one just has to figure out what is the stable range.

(b) The second problem is that the effect of surgery on the kernel (here it is best to think of the homology kernel) is not as clean as it might be. Sure, surgery will kill the class we are cutting out, in $H^r(M)$; but it will also inflict some collateral damage around $H^{n-r}(M)$ (it must, if Poincaré duality is to be preserved.) We may expect, then, that round about $r = n/2$ the method is going to get into trouble. By Poincaré duality , we need only kill the kernel up to dimension $n/2$ and we have killed the whole thing; but at the last stage there will in general be an obstruction arising from the algebra of Poincaré duality on the middle dimensional homology. Duality defines a quadratic form there (the intersection form) and what we will need is some 'stable algebra of quadratic forms'. This is called L-theory.

1.54. EXERCISE. Show that the homotopy effect of a surgery is a combination of attaching a cell and 'detaching' a cell, in the following sense: if M' is obtained from M by doing surgery on an S^q, then there are a space W obtained from M by attaching a $(q+1)$-cell, and a space W' obtained from M' by attaching a p-cell, and W is homotopy equivalent to W'.

homotopy-surgery

We can see a part of the L-theory obstruction (or surgery obstruction) quite directly; it is obtained from the signature. The signature of an oriented manifold (or Poincaré duality space) is clearly an invariant of homotopy type. Moreover we have

1.55. LEMMA. *If M' is obtained from M by a surgery, then M and M' have the same signature.*

PROOF. This can be proved by directly calculating the effect of surgery on homology (compare exercise 1.54). More geometrically, we will prove that M and M' are *cobordant*, that is, there is an oriented manifold W such that $\partial W = M \sqcup (-M')$; and it is known that cobordant manifolds have the same signature. □

Thus, if $f\colon M \to X$ is a normal map, and $\mathrm{Sign}(M) \neq \mathrm{Sign}(X)$, then f cannot possibly be surgered to a homotopy equivalence, and we have found an obstruction to the successful completion of our project.

Using the theory of quadratic forms over the integers, it can in fact be shown that $\mathrm{Sign}(M) - \mathrm{Sign}(X)$ must be a multiple of 8 whenever there exists a normal map $M \to X$. The integer $(\mathrm{Sign}(M) - \mathrm{Sign}(X))/8$ is referred to as the *simply-connected surgery obstruction* in this case.

It is important to see that this surgery obstruction can be computed directly from the data (X, ν_X), since the construction of M (by transversality) is not effective. The key to this is Hirzebruch's signature theorem, which tells us that the signature of M can be computed from the Pontrjagin classes of its tangent bundle. Equivalently, the signature can be computed from the Pontrjagin classes of the *normal* bundle of M; the tangent and normal bundles determine one another's Pontrjagin classes since their sum is trivial. Recall now that the normal bundle of M is pulled back from the bundle V which is prescribed as part of the bundle data; thus we obtain

1.56. PROPOSITION. *If $f\colon M \to X$ is a normal map, as above, then*

$$\mathrm{Sign}(X) - \mathrm{Sign}(M) = \mathrm{Sign}(X) - \langle L(\bar{p}_1(V), \bar{p}_2(V), \dots), [X]\rangle.$$

Thus, if the right-hand expression is non-zero, then no normal map associated to the data (X, V) can be modified by surgery so as to produce a homotopy equivalence.

Here the $\bar{p}_i$ are the 'dual Pontrjagin classes' defined in terms of the total class by $p(V)\bar{p}(V) = 1$. The point is that the dual Pontrjagin classes of the normal bundle to M are the ordinary Pontrjagin classes of its tangent bundle.

The right-hand side of the equation should be thought of as measuring the failure of the Poincaré duality space X (with the given normal data) to satisfy the Hirzebruch signature theorem.

After encountering all these obstructions, Browder **[4]** proved a positive result.

1.57. THEOREM. *Let X be a simply connected finite complex. Suppose that*

(i) *X is a Poincaré duality space of dimension $4j$, $j \geqslant 2$;*
(ii) *There is a vector bundle ν_X over X whose top class is spherical in the sense explained above;*
(iii) $\mathrm{Sign}(X) = \langle L_j(\bar{p}_1, \dots, \bar{p}_j), [X]\rangle$, *where the $\bar{p}_i$ denote the dual Pontrjagin classes of ν_X.*

Then X has the homotopy type of a smooth manifold, i.e., $\mathcal{S}^{DIFF}(X)$ is nonempty.

ROUGH OUTLINE OF THE PROOF. We start by getting a normal map $f\colon M \to X$. As was hinted above, there is no obstruction to doing surgery below the middle dimension so as to make f $2j$-connected. By Poincaré duality , all that we have left to do is to kill the kernel

$$K_{2j}(f) = \operatorname{Ker} f_* : H_{2j}(M) \to H_{2j}(X);$$

one can show that this kernel is a direct summand in $H_{2j}(M)$. The elements of the kernel are represented by $2j$-spheres which one can show are *embedded* in M.

Here is where it is crucial that we are in dimension $\geqslant 5$. General position arguments show that any map $S^k \to M$ can be deformed to an embedding if $k < 2j$, and to an immersion with isolated double points if $k = 2j$. To ensure that our $2j$-spheres are embedded we need a geometric technique for deforming an immersion so as to 'cancel' pairs of isolated double points. Such a technique, called the *Whitney trick*, is available in manifolds of dimension at least 5. It involves constructing an isotopy in a tubular neighborhood of an auxiliary embedded 2-disk. We need the dimension to be at least 5 so that any map of a 2-disk can be deformed to an embedding; and we need simple connectivity because we want to construct our 2-disks by filling in certain loops in M. The details will be clarified later.

Moreover the kernel has its own intersection form, a quadratic form over $\mathbb{Z}$; and because of our assumption that the surgery obstruction vanishes, the intersection form of the kernel has signature zero. Now we apply the theory of integral quadratic[9] forms. This tells us that any nonsingular quadratic form over $\mathbb{Z}$ with signature zero is *hyperbolic*, that is a direct sum of copies of the form with matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Translated into geometry, this tells us that the kernel K_{2j} is spanned by two sets of embedded $2j$-spheres, say $e_1, \dots, e_m, f_1, \dots, f_m$, all disjoint except that each e-sphere intersects the corresponding f-sphere in a single point.

How might this have come about? An obvious example of a manifold whose intersection form is hyperbolic is the product $S^{2j} \times S^{2j}$, in which the e and f spheres are just the two factors. More generally, if $M = M_0 \# S^{2j} \times S^{2j}$, then the intersection form of M is the intersection form of M_0 plus a hyperbolic piece. Moreover it is possible to show that, in our $2j$-connected situation, this is the *only* way that a hyperbolic piece of an intersection form can arise. Thus our manifold M is actually the connected sum of another manifold M' (still equipped with a $2j$-connected normal map to X) and a number of copies of $S^{2j} \times S^{2j}$, and these latter generate the whole kernel K_{2j}. It follows that the normal map $M' \to X$ is in fact $(2j+1)$-connected, hence (by Poincaré duality and the Hurewicz theorem) it is a homotopy equivalence. The construction is complete. □

Our computation of the structure set $\mathcal{S}(X)$ is not completed once we have shown that it is non-empty. It might seem as though the uniqueness problem — does X admit more than one manifold structure, and if so how many? — should require techniques rather different from the existence problem — does X admit any manifold structure at all?. In fact, however, surgery allows us to approach the uniqueness problem as well. The crucial tool is the h-cobordism theorem.

Suppose that $f, f'\colon M, M' \to X$ are two manifold structures for X, and we want to know whether or not they agree. We could ask first the weaker question: do they agree when considered simply as *normal maps*? If not, then certainly they do not agree as manifold structures. Let's suppose then that they do agree as normal maps — and, as we already remarked, it turns out that the correct notion of 'agreement' for normal maps is the existence of a *normal cobordism*, a cobordism W whose boundary components are M and M', and which is equipped with a normal map to $g\colon X \times [0,1]$ restricting on the

[9] The word 'quadratic' is somewhat nuanced here; over a ring where one can't divide by 2, a *quadratic form* is by definition a more refined object than a symmetric bilinear form. See later.

boundary components to f, f'. We can now apply the surgery technique to the normal map g, relative to the boundary of the cobordism W. If surgery succeeds, it replaces W by another normal cobordism W' with the same boundary components and such that $g' \colon W' \to X \times [0,1]$ is a homotopy equivalence. Thus, W' is an h-cobordism. If we are in the simply-connected case so that Smale's theorem applies, we may deduce that W' is a product, M is diffeomorphic to M' and our manifold structures are in fact the same.

Implicit in the above discussion is a sort of 'exact sequence', where the structure set for the n-dimensional space X is sandwiched between the normal cobordism classes of normal maps to X and the obstruction group to surgery in dimension $(n+1)$. Making this idea precise and giving an exposition of its proof is the main purpose of this book.

8. The non-simply connected case

Our discussion above has focused attention on the case of *simply connected* manifolds. There are good reasons to look first at the simply-connected case (as does one of the classic books on surgery, [**3**]). However, the geometric 'tool theorems' needed to carry out surgery work just as well on general manifolds. The difficulty lies rather in developing the appropriate algebra to keep track of the geometry.

An example (once again due to Milnor) shows some of the problems that can arise.

1.58. EXAMPLE. The *lens spaces* $L_{p,q}$ are certain 3-manifolds with finite cyclic fundamental groups, defined as follows. One begins with the representation of the 3-sphere as a subset of $\mathbb{C}^2$,

$$S^3 = \{(z_1, z_2) : |z_1|^2 + |z_2|^2 = 1\}.$$

Given coprime integers p, q there is a diffeomorphism $S^3 \to S^3$ defined by

$$(z_1, z_2) \mapsto (e^{2\pi i/p} z_1, e^{2\pi i q/p} z_2).$$

The quotient of S^3 by the action of this diffeomorphism is the *lens space* $L(p,q)$. (The reader should show that $L(p,q)$ is indeed a 3-manifold, which can also be obtained by identifying the boundaries of two solid tori in a suitable manner.)

It was shown by Reidemeister and Whitehead that the lens spaces $L(7,1)$ and $L(7,2)$ are homotopy equivalent but not diffeomorphic. Moreover, this is a stable phenomenon: $L(7,1) \times S^4$ is not diffeomorphic to $L(7,2) \times S^4$. But Milnor [**11**] showed that these two latter spaces are h-cobordant. Thus, *the h-cobordism theorem fails for non-simply connected manifolds.*

What is going on here? The answer lies in the proof of the h-cobordism theorem. This involves giving the h-cobordism a geometric structure — a handle decomposition or a Morse function — and then using the vanishing of the relative homology groups $H_*(W, \partial_\pm W)$ to simplify this geometric structure, step by step, until it corresponds to a product. The simplification depends on counting the points of intersection of certain submanifolds, and then applying geometric transformations (isotopies) to remove them. Speaking very roughly, one can say that in the simply-connected case the original intersection data correspond to an invertible matrix of integers (an element of $GL(N, \mathbb{Z})$). The geometric transformations correspond to integer row and column operations on this matrix. Success in proving the h-cobordism theorem depends on reducing the original matrix, via integer row and column operations, to the identity. The *Smith normal form theorem* on integer matrices assures us that this can always be done.

However, in the non-simply connected case the matrix of intersection data should be thought of as having entries not in $\mathbb{Z}$ but in the group ring $R = \mathbb{Z}[\pi]$ of the fundamental group. (This is because the geometric cancellation procedures all ultimately depend on spanning a loop by a disk; we need to make sure that the loops along which we try to cancel are nullhomotopic.) Now there is no counterpart of the Smith normal form theorem for matrices over an arbitrary ring R. Instead one defines the *algebraic K-theory group* $K_1(R)$ to be the quotient of the general linear group by the subgroup generated by elementary row and column operations. Any h-cobordism with fundamental group π has a *Whitehead torsion* invariant in (a certain quotient of) $K_1(\mathbb{Z}[\pi])$; and the h-cobordism is a product if and only if the Whitehead torsion is zero.

Similar remarks apply to the surgery obstruction. On a non-simply connected manifold M, the 'intersection number' of two submanifolds of complementary dimension is no longer a number, but an element of $\mathbb{Z}[\pi]$; and the 'intersection form' is a quadratic form over $\mathbb{Z}[\pi]$. The surgery obstruction will no longer be a simple signature, but a complicated algebraic object that keeps track of the stable classification of quadratic forms over $\mathbb{Z}[\pi]$. In fact, the surgery obstruction is an element of an L-*theory* group $L_n(\mathbb{Z}[\pi])$.

Of course we will need not just to *define* but also to *compute* the K-theory and L-theory groups if the theory is to be of any practical use. Because of the very algebraic way in which these groups are defined, the task is difficult. A conjectural link between the groups $L_*(\mathbb{Z}[\pi])$ and the topology of certain model spaces is provided by a raft of famous conjectures attributed to Borel, Novikov, Farrell–Jones, and others. The fundamental idea here is that once one completely understands the structure set for *just one* manifold with fundamental group π, one will be able to calculate the structure set for *any other* manifold with the same fundamental group. It requires the full power of surgery theory to prove this result.

CHAPTER 2

Poincaré Duality

Poincaré duality is the most basic homological property of manifolds. Geometrically, the duality between the homology and cohomology of manifolds corresponds to the *intersection pairing* on homology: two cycles of complementary dimension generically intersect in a finite set of points, and their *intersection number* — the signed count of those points — is a homological invariant. For applications in geometric topology we need to be in a situation where the algebraic intersection numbers provided by Poincaré duality reflect as closely as possible the actual geometric situation; in particular, if two cycles have zero algebraic intersection we would like to be able to make them geometrically *disjoint* by a suitable deformation. The need for this strong connection between algebra and geometry requires us to delve into two topics which are not part of the standard treatment of duality theory: the Whitney trick (a geometric device for canceling intersection points) and equivariant duality (needed because the Whitney trick explicitly involves the fundamental group). We begin however with the more classical form of duality and the Thom isomorphism.

1. Intersections and the Thom Isomorphism

In this section we will work with de Rham cohomology for smooth manifolds M. For On such a manifold we have the familiar de Rham complex of differential forms

$$\Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \cdots \xrightarrow{d} \Omega^n(M)$$

which computes the cohomology $H^*(M;\mathbb{R})$. If M is not compact, there is also the important subcomplex of *compactly supported* forms

$$\Omega^0_c(M) \xrightarrow{d} \Omega^1_c(M) \xrightarrow{d} \cdots \xrightarrow{d} \Omega^n_c(M);$$

its cohomology is the *compactly supported cohomology* $H^*_c(M;\mathbb{R})$. We will need to work with both of these cohomology theories.

compact-support

2.1. REMARK. Any definition of cohomology (singular, cellular, or even a generalized cohomology theory) has a 'compactly supported' variant defined for locally compact spaces X: it is the direct limit $\varinjlim H^*(X, X \setminus K)$ taken over the direct system of compact subsets of K of X ordered by inclusion. It is functorial for *proper* maps (a map is proper if the inverse image of any compact set is compact). The most familiar example of a compactly supported generalized cohomology theory is Atiyah-Hirzebruch K-theory.

2.2. EXERCISE. Show that our definition of compactly supported de Rham cohomology is consistent with the general definition in terms of relative groups given in the remark above. (Relative de Rham theory can be defined in terms of mapping cone complexes; see [**2**, page 78].)

Any calculation of de Rham cohomology begins with the *Poincaré lemma*, which gives the cohomology of Euclidean space. The result is:

poincare-lemma

2.3. LEMMA. *Let M be diffeomorphic to Euclidean space $\mathbb{R}^n$. Then*

(a) *$H^k(M;\mathbb{R})$ is isomorphic to $\mathbb{R}$ when $k = 0$, to 0 otherwise; the generator is the cohomology class of the constant function 1;*
(b) *$H^k_c(M;\mathbb{R})$ is isomorphic to $\mathbb{R}$ when $k = n$, to 0 otherwise; the generator is the cohomology class of a 'bump n-form' ω, compactly supported and with $\int_M \omega = 1$.*

Notice in (b) that the operation $\int$, and therefore the normalization of the generator, depend on the choice of an orientation of M.

Having understood the de Rham cohomology of a single Euclidean space, the next thing to understand is a smoothly varying collection of such spaces — that is, a vector bundle.

2.4. DEFINITION. Let V be an oriented ℓ-dimensional vector bundle over a compact manifold M. A *Thom form* for V is a closed ℓ-form α on the total space of V (considered as a noncompact manifold in its own right) such that

(a) α is compactly supported,
(b) α is closed, that is, $d\alpha = 0$, and
(c) for each fiber F of V (oriented by the orientation of V) we have $\int_F \alpha = 1$.

The cohomology class (in $H^\ell_c(V;\mathbb{R})$) of a Thom form is called a *Thom class* for V.

2.5. EXERCISE. Show that Thom forms exist. (Use a partition of unity to glue together local Thom forms.)

2.6. THEOREM. *(Thom Isomorphism Theorem) Let V be an oriented ℓ-dimensional vector bundle over a compact manifold M. All Thom forms for V define the same Thom class in $H^\ell_c(V;\mathbb{R})$. If π denotes the projection of the vector bundle V, then the map $\alpha \mapsto \pi^*\alpha \wedge \varphi$ gives an isomorphism (the* Thom isomorphism*) $\Phi\colon H^*(X;\mathbb{R}) \to H^{*+\ell}_c(V;\mathbb{R})$.*

SKETCH PROOF. Leave to one side for now the question of the uniqueness of the Thom classes; just choose a particular Thom form. Cap-product with it defines Thom maps Φ not just for M itself but for any open subset U: we have

$$\Phi_U : \Omega^*_c(U) \to \Omega^{*+\ell}_c(\pi^{-1}(U))$$

on the level of differential forms, and

$$\Phi_U\colon H^*_c(U;\mathbb{R}) \to H^{*+\ell}_c(\pi^{-1}(U);\mathbb{R})$$

on the level of cohomology. If U is a small open ball in a coordinate chart (so that the restriction of V to U is a trivial bundle) then both U and $\pi^{-1}(U)$ are diffeomorphic to Euclidean spaces and Φ_U is an isomorphism on cohomology by Lemma 2.3. We now piece these isomorphisms together using a Mayer-Vietoris argument. Suppose that U_1 and U_2 are open subsets of M and that it is known that Φ_{U_1}, Φ_{U_2}, and $\Phi_{U_1\cap U_2}$ are all isomorphisms. There is a commutative diagram of complexes and chain maps with exact columns

$$
\begin{tikzcd}
0 \arrow[d] & 0 \arrow[d] \\
\Omega^*_c(U_1 \cap U_2) \arrow[r, "\Phi"] \arrow[d] & \Omega^{\ell+*}_c(\pi^{-1}(U_1) \cap \pi^{-1}(U_2)) \arrow[d] \\
\Omega^*_c(U_1) \oplus \Omega^*_c(U_2) \arrow[r, "\Phi"] \arrow[d] & \Omega^{\ell+*}_c(\pi^{-1}(U_1)) \oplus \Omega^{\ell+*}_c(\pi^{-1}(U_2)) \arrow[d] \\
\Omega^*_c(U_1 \cup U_2) \arrow[r, "\Phi"] \arrow[d] & \Omega^{\ell+*}_c(\pi^{-1}(U_1) \cup \pi^{-1}(U_2)) \arrow[d] \\
0 & 0
\end{tikzcd}
$$

which gives rise to a commutative diagram of Mayer-Vietoris sequences in compactly supported cohomology. (See [**2**] for details of the construction of these Mayer-Vietoris sequences.) By our supposition, two out of the three vertical maps give rise to cohomology isomorphisms; so the five lemma tells us that the third one, $\Phi_{U_1 \cup U_2}$, will do so also. The proof that the Thom map for the compact manifold M is an isomorphism is now completed by an induction on the number of sets in a good open covering.

Finally, let us return to the question of the uniqueness of Thom classes. We may assume that M is connected. Then, by the isomorphism result that we have just proved, $H^\ell_c(V;\mathbb{R}) \cong H^0(M;\mathbb{R})$ is one-dimensional. Thus, any two Thom classes are scalar multiples of one another. The normalization condition (c) in the definition of a Thom class ensures that the multiple is 1. $\square$

2.7. REMARK. In particular note that $H^q_c(V;\mathbb{R}) = 0$ for $q < \ell$. The Mayer-Vietoris type argument used in this proof will recur frequently. We refer to it as as an *assembly* construction; it 'assembles' the local Thom isomorphisms given by the Poincaré lemma into a global isomorphism.

assembly-remark-1

fiber-integration

2.8. REMARK. The inverse of the Thom isomorphism can be described in de Rham theory as the operation of *integration over the fiber*. This process defines a 'wrong way' map on the complexes of differential forms, $\pi_* \colon \Omega^*_c(V) \to \Omega^{*-\ell}(M)$; one uses the local product structure to express a form as a product of terms coming from the root and from the fiber, and then integrates out the top-dimensional fiber component using the orientation. See [**2**, pages 61–63] for details. We will need to make use of the 'Fubini principle for integration along the fiber'

$$\int_V \pi^*\theta \wedge \varphi = \int_M \theta \wedge \pi_*\varphi,$$

where $\theta \in \Omega^*(M)$, $\varphi \in \Omega^*_c(V)$. This is proved by using a partition of unity to work in product neighborhoods.

Imagine now that the closed manifold M^m is embedded as a submanifold of the closed manifold W^n, and that both W and M are oriented. Then the normal bundle V of M in W is oriented also, and so it possesses a Thom class $[\alpha] \in H^{n-m}_c(V;\mathbb{R})$. Now, by the tubular neighborhood theorem (see Appendix ?), the total space of V may be identified with an open subset of W, and so there is a map on cohomology $H^*_c(V;\mathbb{R}) \to H^*(W;\mathbb{R})$

— in terms of differential forms, this is the operation of 'extension by zero' of a compactly supported form. Applying this map to $[\alpha]$ we see that each oriented submanifold M of W gives rise to a cohomology class

$$[\alpha_M] \in H^{n-m}(W;\mathbb{R}).$$

dual-class

2.9. DEFINITION. The cohomology class α_M defined above is called the *dual cohomology class* to M.

Thpd

2.10. PROPOSITION. *With M and W as above, the dual cohomology class $[\alpha_M]$ has the following property: for every closed form $\beta \in \Omega^m(W)$, we have*

$$\int_M \beta = \int_W \beta \wedge \alpha_M = \langle [\beta] \smile [\alpha_M], [W] \rangle.$$

PROOF. Denote by $i\colon M \to V$ the inclusion of the zero-section. Then

$$\int_W \beta \wedge \alpha_M = \int_V \beta \wedge \alpha_M = \int_V \pi^* i^* \beta \wedge \alpha_M,$$

the first equality by restriction and the second because i and π are mutually inverse homotopy equivalences between M and V. By the Fubini principle for integration along the fiber (Remark 2.8),

$$\int_V \pi^* i^* \beta \wedge \alpha_M = \int_M i^* \beta \wedge \pi_*(\alpha_M) = \int_M \beta,$$

since $\pi_*(\alpha_M) = 1$ by definition of the Thom form. □

Suppose now that M_1 and M_2 are two submanifolds of W^n, of dimensions m_1 and m_2 respectively, $m_1 + m_2 \geqslant n$. One says that M_1 and M_2 *intersect transversely* if, near any point of their intersection, there is a coordinate chart in which M_1 is represented by the subspace spanned by the first m_1 basis vectors of $\mathbb{R}^n$ and M_2 is represented by the subspace spanned by the last m_2 basis vectors.

This is a special case of the general notions of transversality that we will investigate in the next chapter. There we will see that given any two submanifolds it is possible to deform one of them slightly so as to make their intersection transverse.

In particular, the intersection $M_1 \cap M_2$ is a submanifold of dimension $m_1 + m_2 - n$. Moreover, the normal bundles are related by

$$\nu_{M_1 \cap M_2} = \nu_{M_1} \oplus \nu_{M_2}.$$

Since the Thom class of a direct sum of vector bundles is easily seen to be the product of the Thom classes of the summands, we have

intersect-dual

2.11. PROPOSITION. *If M_1 and M_2 intersect transversely, then the dual cohomology classes are related by*

$$\alpha_{M_1 \cap M_2} = \alpha_{M_1} \wedge \alpha_{M_2}.$$

In particular suppose that M_1 and M_2 intersect transversely and have complementary dimensions, $m_1 + m_2 = n$. The intersection $M_1 \cap M_2$ is then just a finite set of points p, each of which acquires a sign $\varepsilon(p) \in \{\pm 1\}$ according to whether or not the orientations of M_1 and M_2 at that point combine to yield the orientation of W. The signed count $\sum_{p \in M_1 \cap M_2} \varepsilon(p)$ of these points is called the *intersection number* $\lambda(M_1, M_2)$ of the two submanifolds. Plainly, this is just the integral over W of the dual class to the oriented 0-dimensional manifold $M_1 \cap M_2$. Thus, from Proposition 2.11 we obtain

intersect-equation

$$\lambda(M_1, M_2) = \int_M \alpha_{M_1} \wedge \alpha_{M_2} \tag{2.12}$$

Notice the important symmetry property

intersect-symmetry

$$\lambda(M_2, M_1) = (-1)^{m_1 m_2} \lambda(M_1, M_2) \tag{2.13}$$

which may be derived either by considering the orientation of the intersection points, or from the graded commutativity of the wedge product.

2.14. REMARK. When M_1 and M_2 are *not* transverse, we may use the homological formula as the *definition* of the intersection number; this will then be equal to the 'geometric' intersection number of M_1' and M_2, where M_1' is a small deformation of M_1 in the same homology class, and is transverse to M_2. An important example concerns the *self-intersection* $\lambda(M, M)$ of a middle-dimensional submanifold M. If the normal bundle of M admits a nowhere-vanishing section, then this section can be thought of as giving a small deformation of M to a submanifold M' which doesn't intersect M at all. The self-intersection will therefore be zero. In general this argument shows that the self-intersection of M is equal to the number of zeroes of a generic section of its normal bundle — that is (by Hopf's theorem) the *Euler number* of this bundle. See Appendix [**?**] for the obstruction theory that is involved here.

2. Duality in de Rham theory

The *de Rham homology groups* of a compact manifold W are the duals of the cohomology groups:

$$H_p(W; \mathbb{R}) = \operatorname{Hom}(H^p(W; \mathbb{R}); \mathbb{R}).$$

A closed oriented submanifold M^m gives rise to an m-dimensional homology class $[M]$ by

$$\langle [\theta], [M] \rangle = \int_M \theta;$$

Stokes' theorem shows that this operation is well-defined on cohomology.

2.15. REMARK. One can alternatively define de Rham homology as the homology of the complex of *currents* of M. The space of k-*currents* is the dual space of the locally convex topological vector space of k-forms, and a closed oriented submanifold defines a closed current by integration. This approach is equivalent to the simpler one above for a *closed* manifold W. For a manifold that is not closed, the dual of the LCTVS of *all* k-forms is the space of *compactly supported* currents, and the dual of the space of *compactly supported* k-forms is the space of *all* currents (without restriction on support). Thus we obtain two kinds of homology: ordinary homology (with currents of compact support) dual to ordinary cohomology (without support restriction), and *locally finite* or *closed* homology (using currents of unrestricted support) dual to compactly supported cohomology (compare Remark 2.1).

2.16. EXERCISE. Show that a bilinear *cap-product* pairing

$$H^p(W; \mathbb{R}) \otimes H_q(W; \mathbb{R}) \to H_{q-p}(W; \mathbb{R})$$

can be defined by dualizing the *cup-product*

$$H^p(W; \mathbb{R}) \otimes H^q(W; \mathbb{R}) \to H^{p+q}(W; \mathbb{R})$$

which is the map on cohomology induced by the exterior product of forms. Express this in terms of currents, and work out the various possible kinds of cap-product pairing on an *open* manifold.

We see that Proposition 2.10 can be expressed as follows: the homology class associated to a closed submanifold M of W can be represented by a dual cohomology class α_M, in such a way that

$$[M] = [\alpha_M] \frown [W]$$

where $\frown$ denotes the cap-product pairing between homology and cohomology. The *Poincaré duality theorem* states that every homology class arises in this way, and in fact that cap-product with the *fundamental class* $[W]$ (defined by integration over W) induces an isomorphism from cohomology to homology.

2.17. PROPOSITION. *Let W^n be a closed oriented manifold and let $[W] \in H_n(W;\mathbb{R})$ be the fundamental class. Then the* duality maps

$$D\colon H^q(W;\mathbb{R}) \to H_{n-q}(W;\mathbb{R}),$$

derham-duality *induced by cap-product with $[W]$, are isomorphisms.*

SKETCH PROOF. First, let us observe that the duality map D can be defined whether or not W is compact. If W is not compact then duality defines a map $H^q_c(W;\mathbb{R}) \to H_{n-q}(W;\mathbb{R})$ from *compactly supported* cohomology to homology. Moreover, direct calculation with the Poincaré lemma 2.3 shows that this map is an isomorphism when W is Euclidean space. Now cover a compact manifold W by finitely many open sets each of which, together with all their possible intersections, is either empty or diffeomorphic to Euclidean space. A Mayer-Vietoris 'assembly' argument (Remark 2.7) completes the proof. □

Notice that the 'dual class' to an oriented submanifold M, as we defined it in 2.9, is indeed the cohomology class Poincaré dual to $[M]$. We see therefore that Poincaré duality gives information about intersections of submanifolds; in fact, Equation 2.12 can be rephrased to say that for oriented submanifolds M_1, M_2 of complementary dimensions, the algebraic intersection number $\lambda(M_1, M_2)$ is equal to $\langle D^{-1}[M_1], [M_2]\rangle$.

This geometry explains the terminology 'intersection form' for the bilinear form $(\alpha,\beta) \mapsto \int \alpha\wedge\beta$ defined on cohomology, or its dual $(x,y) \mapsto D^{-1}(x)(y)$ defined on homology. Notice that since Poincaré duality is an isomorphism, the intersection form is nondegenerate (a bilinear form on a finite-dimensional vector space is called *nondegenerate* if it sets up an isomorphism between that vector space and its dual.) If M^{2m} is an even-dimensional manifold, then the intersection form restricts to the middle-dimensional cohomology $H^m(M;\mathbb{R})$; it is a symmetric bilinear form if m is even, a skew-symmetric form if m is odd.

2.18. DEFINITION. Let M be a compact oriented manifold of dimension $4j$. The *signature* of M is the signature of the intersection form in the middle dimension, that is the symmetric bilinear form

$$(\alpha,\beta) \mapsto \int \alpha\wedge\beta$$

on $H^{2j}(M;\mathbb{R})$.

By convention we say that the signature of M is zero if the dimension of M is not a multiple of 4.

2.19. REMARK. Here is an abstract framework in which one can understand the Mayer-Vietoris proof of Poincaré duality given in Proposition 2.17. Given a finite open cover $\mathcal{U}$ of the closed manifold W, we can build a simplicial complex called the *nerve* $N(\mathcal{U})$ of $\mathcal{U}$ as follows: the vertices of the nerve are the members of $\mathcal{U}$, and $U_1, \ldots, U_k \in \mathcal{U}$ span a simplex if and only if their intersection $U_1 \cap \cdots \cap U_k$ is a non-empty subset of W. Let F be a functor which attaches to each open subset U of X a chain complex (of real vector spaces) and which is covariant for inclusions; the examples we have in mind are $F_1(U) = \Omega_c^{n-*}(U)$ the compactly supported forms on U (with a shift of grading), and

$F_2(U) = \Omega^c_*(U)$ the compactly supported currents on U. Then to each simplex of $N(\mathcal{U})$ is associated a chain complex (via the functor F) and to each face map is associated a morphism of chain complexes. These data allow us to define a double complex (as in [**2**]) combining the given differentials on the functor F and the simplicial differential on the nerve $N(\mathcal{U})$. The duality map D defines a natural transformation of functors $F_1 \to F_2$ and the key point is that if such a natural transformation is an isomorphism 'locally' — over every simplex of N — then it is an isomorphism 'globally' — on the total complex of the double complex. We will develop these ideas in the next section.

functor-remark

3. Geometric Modules and Complexes

Remark 2.19 at the end of the last section suggests that, in order to understand the structure of Poincaré duality , it will be helpful to develop some 'geometric algebra' — algebra carried out on objects (such as modules) which are 'located' at some point of a 'control space'. In this section we shall develop one version of this idea, which is of central importance in modern topology.

Let K be a finite simplicial complex and let R be a commutative ring (usually it will be $\mathbb{Z}$).

2.20. DEFINITION. A *geometric R-module M over K* (or *(R, K)-module* for short) is a list $\{M_\sigma\}$ of R-modules parameterized by the simplices of σ. The *total module* of M is the direct sum $\bigoplus_\sigma M_\sigma$ (over all simplices of K). Usually we'll use the same notation M for the total module as we do for the geometric module itself. We will call M_σ the part of M *anchored* at σ.

2.21. DEFINITION. A *morphism* $\varphi\colon M \to N$ of (R, K)-modules is a list $\{\varphi_{\sigma,\tau}\}$ of R-module morphisms $M_\sigma \to N_\tau$, such that $\varphi_{\sigma,\tau}$ is zero unless $\sigma \leqslant \tau$ (that is, unless σ is a face of τ). We also use the notation φ for the *total morphism* induced by φ, that is the direct sum $\bigoplus_{\sigma,\tau} \varphi_{\sigma,\tau}$ considered as a morphism on the total modules.

Geometric R-modules and morphisms form an (additive) category.

2.22. EXAMPLE. Here is a key example. Let $C^q(K)$ be the geometric module whose component over a simplex σ is R if σ is a q-simplex, and 0 otherwise. The total module of this geometric module may be identified with the space of simplicial q-cochains of K (with coefficients in R). Moreover, the simplicial cochain complex of K,

$$C^0(K; R) \longrightarrow C^1(K; R) \longrightarrow C^2(K; R) \longrightarrow \cdots$$

now becomes a complex in the category of geometric modules. (This is because the coboundary of a simplex σ is a sum of simplices of which σ is a face.)

controlled-cochain

2.23. EXAMPLE. Let X be a topological space, $\mathcal{U}$ a finite open cover, $K = N(\mathcal{U})$ the nerve of $\mathcal{U}$ (as in Remark 2.19). Suppose that Γ is a sheaf of R-modules over X. Let $C^q(\mathcal{U}; \Gamma)$ be the geometric (R, K)-module which sends each q-simplex $\sigma = (U_1, \ldots, U_q)$ to the R-module $\Gamma(U_1 \cap \cdots \cap U_q)$, and is zero on simplices of other dimensions. The total module of this geometric module may be identified with the space of Čech q-cochains of the cover $\mathfrak{U}$ with coefficients in Γ. Moreover, the Čech cochain complex of the cover

$$C^0(\mathcal{U}; \Gamma) \to C^1(\mathcal{U}; \Gamma) \to C^2(\mathcal{U}; \Gamma) \to \cdots$$

now becomes a complex in the category of geometric modules.

Size: 50 mm

FIGURE 1. Barycentric subdivision, and a dual cell

2.24. EXERCISE. Let M and N be geometric (R, K)-modules. Show that the space $\mathrm{Hom}_{(R,K)}(M, N)$ of geometric morphisms from M to N is itself a geometric module, where we consider the component $\varphi_{\sigma,\tau}$ to be anchored at τ. Show that composition on the left with a geometric morphism $M' \to M$, or on the right with a geometric morphism $N \to N'$, themselves define geometric morphisms

$$\mathrm{Hom}_{(R,K)}(M, N) \to \mathrm{Hom}_{(R,K)}(M', N), \quad \mathrm{Hom}_{(R,K)}(M, N) \to \mathrm{Hom}_{(R,K)}(M, N'),$$

respectively.

Our definition of geometric morphism has a certain asymmetry, which is why it is easier to build cohomological examples than homological ones. However, homology can also be incorporated into the picture by the device of *dual cell decomposition*, which goes right back to Poincaré's proof of Poincaré duality.

Let K be a simplicial complex, as before. Remember that the *barycentric subdivision* K' of K may be defined (abstractly) as the simplicial complex whose vertices correspond to the simplices of K, with a simplex of K' being a *flag* of simplices of K. That is to say, the simplices of K' are spanned by vertices $\hat{\sigma}_0, \widehat{\sigma_1}, \ldots, \hat{\sigma}_q$ corresponding to simplices $\sigma_0, \sigma_1, \ldots, \sigma_q$ of K having $\sigma_0 < \sigma_1 < \cdots < \sigma_q$. The figure shows the geometric picture of a barycentric subdivision.

As a matter of terminology, if $[\hat{\sigma}_0, \ldots, \hat{\sigma}_q]$ is a simplex of K', we shall refer to the simplex σ_0 of K as its *root* and the simplex σ_q as its *tip*. If σ is a simplex of K, its *dual cell* $D(\sigma, K)$ is the subcomplex of K' comprising those simplices whose root σ_0 satisfies $\sigma \leqslant \sigma_0$; the condition of strict inequality $\sigma < \sigma_0$ defines a subcomplex of the dual cell which is called its *boundary* $\partial D(\sigma, K)$. The dual cell is contractible; there is an obvious 'linear' contraction to the vertex represented by σ.

2.25. EXAMPLE. Let K be a finite simplicial complex. Let $C_q(K', R)$ be the geometric (K, R)-module which assigns to a simplex $\sigma \in K$ the free R-module generated by those q-simplices of K' whose root is σ. As an R-module, this is canonically isomorphic to the q'th relative simplicial chain module of the pair $(D(\sigma, K), \partial D(\sigma, K))$. The total module of the geometric module $C_q(K', R)$ is just the module of simplicial chains on X'. Moreover, the simplicial chain complex of K'

$$C_0(K'; R) \longleftarrow C_1(K'; R) \longleftarrow C_2(K'; R) \longleftarrow \cdots$$

is now a complex in the category of geometric (R, K)-modules. This is because a face of a simplex of K' must have as vertices only simplices of K which have the root of the original simplex among their faces.

controlled-chain

We have constructed various chain complexes in the category of geometric modules. We will need a 'local-global principle' for deciding when two such complexes are chain equivalent.

2.26. DEFINITION. Let φ be a morphism of geometric (R, K)-modules. It is said to be *diagonal* if $\varphi_{\sigma,\tau} = 0$ unless $\sigma = \tau$. For a general morphism φ, its *diagonal part* is the diagonal morphism $\hat{\varphi}$ defined by

$$\hat{\varphi}_{\sigma,\tau} = \begin{cases} \varphi_{\sigma,\tau} & \text{if } \sigma = \tau \\ 0 & \text{otherwise} \end{cases}$$

2.27. EXERCISE. Check that this process of 'taking the diagonal part' is functorial (it preserves composition of morphisms). The reason is essentially that the map from upper triangular matrices to their diagonal part preserves matrix multiplication.

2.28. EXERCISE. Show that an (R, K)-module morphism is an isomorphism if and only if its diagonal part is an isomorphism. (Hint: If the diagonal part of φ is invertible, show that its inverse defines an (R, K)-module morphism ψ such that $\varphi\psi - 1$ and $\varphi\psi - 1$ are nilpotent. Remember that K is a *finite* complex.)

invert-ex

We can define a category (call it the category of 'diagonal modules') whose objects are geometric modules and whose morphisms are diagonal morphisms. The exercise shows that taking the diagonal part defines a functor from the category of geometric modules to the category of diagonal modules. In particular we can take the diagonal part of a chain complex of geometric modules, obtaining an associated chain complex of diagonal modules.

Notice that the total complex of a chain complex of diagonal modules splits into a direct sum of subcomplexes, one for each simplex σ. This means that properties of complexes of diagonal modules are 'local' — they can be verified one simplex at a time.

homology-ex1

2.29. EXAMPLE. The diagonal part of the cochain complex $C^\bullet(K; R)$, considered as a complex of geometric modules as in Example 2.22, assigns to each q-simplex σ the complex which has one free generator in dimension q and zero boundary maps. The diagonal part of the chain complex $C_\bullet(K'; R)$, considered as a complex of geometric modules as in Example 2.25, assigns to each q-simplex σ the relative chain complex of the pair $(D(\sigma, K), \partial D(\sigma, K))$ (the nontrivial statement here is that the diagonal part of the boundary map is exactly the relative boundary map of the pair).

We can now state the local-global principle

loctoglob

2.30. PROPOSITION. *A finite chain complex of geometric (R, K)-modules is chain contractible (in the category of geometric modules) if and only if its diagonal part is chain contractible (in the category of diagonal modules). Similarly, a chain map between such complexes is a chain equivalence if and only if the induced map on the diagonal parts is a chain equivalence.*

PROOF. Let (C, d) be a finite chain complex of geometric modules. It is clear that if C is chain contractible then so is $\hat{C}$. Conversely, suppose that $\hat{C}$ is chain contractible and let $\hat{\Gamma}\colon \hat{C} \to \hat{C}$ be a chain contraction, defined by diagonal (R,K)-module morphisms $\hat{\Gamma} : C_r \to C_{r+1}$ such that

$$\hat{d}\hat{\Gamma} + \hat{\Gamma}\hat{d} \;=\; 1 \;:\; C_r \to C_r \;.$$

The (R,K)-module morphisms defined by

$$\alpha \;=\; d\hat{\Gamma} + \hat{\Gamma} d \;:\; C_r \to C_r$$

have diagonal parts $\hat{\alpha} = 1$, so that they are automorphisms by Exercise 2.28. Moreover

$$d\alpha \;=\; d\hat{\Gamma} d \;=\; \alpha d \;:\; C_r \to C_{r-1} \;.$$

The (R,K)-module morphisms

$$\Gamma \;=\; \hat{\Gamma}\alpha^{-1} \;:\; C_r \to C_{r+1}$$

are such that

$$d\Gamma + \Gamma d \;=\; 1 \;:\; C_r \to C_r \;,$$

and so they define a chain contraction of C.

The second part of the proposition follows from the first by considering mapping cylinders. □

subdivide-cochain

2.31. EXERCISE. Show that the cochain complex $C^\bullet(K';R)$ of the barycentric subdivision of K becomes a chain complex of (R,K)-modules if we take each simplex $[\hat{\sigma}_0, \ldots, \hat{\sigma}_q]$ of K' to be anchored at its tip σ_q.

Show that the barycentric subdivision chain map [**?**] defines a chain equivalence of complexes of (R,K)-modules between $C^\bullet(K;R)$ and $C^\bullet(K';R)$.

4. Duality and Geometric Modules

As we have already seen, Poincaré duality is induced by the cap product. To use the category of geometric modules to discuss duality, we must therefore first show that the cap (and cup) products fit into that category.

Cup products in de Rham theory are represented simply by the exterior product of differential forms. In fact, one can think of the wedge product in the following way: if we identify the (suitably completed) tensor product $\Omega^*(M) \otimes \Omega^*(M)$ with the differential forms on the product manifold $M \times M$, then the wedge product is simply the map on forms

$$\Omega^*(M \times M) \to \Omega^*(M)$$

induced by the diagonal inclusion $M \to M \times M$.

When we use other homology and cohomology theories (such as simplicial theory), there is no longer such a canonical choice of *diagonal approximation*

$$C^\bullet(K) \otimes C^\bullet(K) \to C^\bullet(K).$$

Instead, there are theorems which show that diagonal approximations exist and are unique up to an appropriate notion of chain homotopy. (See Appendix **??**, 'Diagonal Approximations'). In our present discussion we shall make use of a particular diagonal approximation which is appropriate to simplicial homology, and which was first defined by Alexander and Whitney. To define it, we must first order (arbitrarily) the vertices of the complex K, and decide to represent each simplex by a symbol $[v_0 \cdots v_q]$ where the vertices appear in increasing order. Then the *Alexander-Whitney diagonal approximation* is the chain map

$$C_\bullet(K) \to C_\bullet(K) \otimes C_\bullet(K)$$

(or its dual on cochains) defined by

$$[v_0 \cdots v_q] \mapsto \sum_{i=0}^{q} [v_0 \cdots v_i] \otimes [v_i \cdots v_q].$$

We are going to apply the Alexander-Whitney diagonal approximation not to the complex K itself but to its barycentric subdivision K'. In order to do this we must order the vertices of K'. Remembering that each vertex of K' corresponds to a simplex of K, we order these by increasing dimension:

$$\text{0-simplices of } K < \text{1-simplices of } K < \cdots ;$$

and within each fixed dimension we order the simplices lexicographically. This choice of ordering gives us a chain level cap product map

capprod

$$C_\bullet(K';R) \to \mathrm{Hom}_R(C^\bullet(K';R), C_\bullet(K';R)) \tag{2.32}$$

which is defined by

$$[\hat\sigma_0 \cdots \hat\sigma_q] \mapsto \varphi([\hat\sigma_0 \cdots \hat\sigma_p])[\hat\sigma_p \cdots \hat\sigma_q]$$

if φ is a p-cochain.

controlled-capprod

2.33. PROPOSITION. *The pairing of Equation 2.32 in fact defines a (R,K)-module chain map*

$$C_\bullet(K';R) \to \mathrm{Hom}_{(R,K)}(C^\bullet(K';R), C_\bullet(K';R))$$

where the chain and cochain complexes are made into geometric modules as in Examples 2.25 and 2.31.

PROOF. There are two statements to verify here,

(i) that for a fixed simplex $[\hat\sigma_0 \cdots \hat\sigma_q]$ of K' the map $C^\bullet(K';R) \to C_\bullet(K';R)$ defined by $\varphi \mapsto \varphi([\hat\sigma_0 \cdots \hat\sigma_p])[\hat\sigma_p \cdots \hat\sigma_q]$ is an (R,K)-module homomorphism,

(ii) and that the map assigning to $[\hat\sigma_0 \cdots \hat\sigma_q]$ the (R,K)-module homomorphism defined in item (i) is itself an (R,K)-module homomorphism from $C_\bullet(K';R)$ to $\mathrm{Hom}_{(R,K)}(C^\bullet(K';R), C_\bullet(K';R))$.

It is easy to check these facts: remember that a simplex of the chain complex of K' is anchored at its root, whereas a simplex of the cochain complex is anchored at its tip. $\square$

2.34. EXERCISE. Show that the map

$$C_\bullet(K';R) \to \mathrm{Hom}_{(R,K)}(C^\bullet(K';R), C_\bullet(K';R))$$

defined in the proposition is in fact an (R,K)-module chain equivalence (use Proposition 2.30).

5. Geometric Poincaré Duality

Let K be a finite complex. For a vertex v of K, let $K \ominus v$ denote the subcomplex of K comprising all those simplices which do not have v as a vertex (this is the complement of the 'open star' of v in K).

2.35. DEFINITION. Let R be a commutative ring (usually $\mathbb{Z}$). The complex K is a (combinatorial) *homology n-manifold* (with coefficients R) if

$$H_k(K', K' \ominus \hat\sigma; R) = \begin{cases} R & \text{when } k = n \\ 0 & \text{otherwise} \end{cases}$$

homology-mfd-def

for every vertex $\hat\sigma$ of the barycentric subdivision K'.

2.36. EXERCISE. A compact Hausdorff space X is called a homology n-manifold if, for each point $x \in X$, one has

$$H_k(X, X \setminus \{x\}; R) = \begin{cases} R & \text{when } k = n \\ 0 & \text{otherwise} \end{cases}$$

using singular homology. Show that the complex K is a homology manifold by our definition in 2.35 if and only if its geometric realization $|K|$ is a homology manifold in the topological sense above.

2.37. EXAMPLE. Every compact smooth manifold can be triangulated (that is, it is homeomorphic to the geometric realization of a finite simplicial complex). This result is due to Cairns and Whitheead [**17**], and it can also be deduced from the handlebody decomposition of smooth manifolds which will be developed in Chapter **??**. Using excision and local coordinate charts, it is easy to check that a smooth manifold is a homology manifold, in the topological sense of the previous exercise. Therefore, by that exercise, any triangulation of a smooth manifold is a combinatorial homology manifold.

2.38. DEFINITION. Let K be a homology n-manifold (with coefficents R). An *orientation* for K is a homology class $[K] \in H_n(K'; R)$ (called a *fundamental class* for the orientation) which restricts to a generator of $H_n(K', K' \ominus \hat\sigma; R) \cong R$ for each vertex $\hat\sigma$ of K'.

Suppose that K is an oriented homology n-manifold, and pick a specific cycle representing the fundamental class $[K]$. By Proposition 2.33, cap-product with $[K]$ defines an (R, K)-module chain map from $C^{n-\bullet}(K'; R)$ to $C_\bullet(K'; R)$.

2.39. THEOREM (Geometric Poincaré Duality). *For an oriented homology n-manifold K as above, the (R, K)-module chain map defined by cap product with the fundamental class*

$$C^{n-\bullet}(K'; R) \to C_\bullet(K'; R)$$

is a chain equivalence (in the category of (R, K)-modules).

2.40. REMARK. In particular, cap-product with $[K]$ defines a chain equivalence in the category of R-modules, and therefore an isomorphism of homology and cohomology groups $H^{n-*}(K; R) \to H_*(K; R)$, which is the classical statement of Poincaré duality. But the local form of duality given by this theorem is more precise.

PROOF. According to Proposition 2.30 above, it will be enough to show that cap-product with $[K]$ gives a chain equivalence on the level of the diagonal parts of the (R, K)-module chain complexes $C^{n-\bullet}(K'; R)$ and $C_\bullet(K'; R)$.

The diagonal part of $C_\bullet(K'; R)$ anchored over a k-simplex σ is the simplicial chain complex of the dual cell $D(\sigma, K)$ relative to its boundary (see Example 2.29). Let us note that the k-fold suspension of the pair $(D(\sigma, K), \partial D(\sigma, K))$ is the pair consisting of the closed star of $\hat\sigma$ relative to its boundary, or equivalently (by excision) the pair $(K', K' \ominus \hat\sigma)$. See Figure 2. In particular, $H_\bullet(D(\sigma, K), \partial D(\sigma, K))$ is R in dimension $n-k$, 0 elsewhere.

Similarly, the diagonal part of $C^\bullet(K'; R)$ anchored over σ is spanned by all those simplices of K' which have σ as their tip. Let $(R)^k$ denote the cochain complex that has a single copy of R in dimension k, and zero elsewhere. There is a chain map $(R)^k \to C^\bullet(K'; R)_\sigma$ given by sending the generator to the sum of all the k-simplices of K' whose tip is σ; by Exercise 2.31, this chain map is a chain equivalence. Thus the cohomology of the diagonal part of $C^\bullet(K'; R)_\sigma$ is R in dimension k and 0 elsewhere.

One sees geometrically that the cap-product with the cohomology generator described above is just the suspension isomorphism

$$H_r(M, M \ominus \hat\sigma; R) \to H_{r-k}(D(\sigma, K), \partial D(\sigma, K)).$$

Size: 52 mm

FIGURE 2. Suspension of the dual cell gives a star suspstar

Taking $r = n$ this tells us that cap-product with the fundamental homology class maps the cohomology of the diagonal cochain complex anchored at σ isomorphically to the homology of the diagonal chain complex anchored at σ. Finally, we recall that a chain map between free chain complexes which induces a homology isomorphism is necessarily a chain equivalence. □

2.41. REMARK. The Poincaré duality map gives us an *intersection pairing*

$$H^r(M) \otimes H^{n-r}(M) \to R$$

by pairing homology with cohomology (alternatively, we can regard this as the cup-product on cohomology followed by evaluation against the fundamental class). It is well known that the cup-product is graded commutative: in particular, the intersection pairing on middle-dimensional cohomology ($n = 2r$) is skew-symmetric if r is odd and symmetric if r is even. This is of course consistent with the properties of intersection numbers of submanifolds (Equation 2.13).

It is worth asking where this symmetry comes from in our construction of Poincaré duality . It arises because the Alexander-Whitney diagonal approximation $\Delta\colon C(K) \to C(K) \otimes C(K)$ is chain homotopic to the map $T\Delta\colon C(K) \to C(K) \otimes C(K)$ obtained by composing it with the 'transposition' T which reverses the two factors of the tensor product. Thus the cup product is not symmetric on the nose, but only up to chain homotopy. Question: Is the *chain homotopy* symmetric on the nose? Answer: No, but it is symmetric up to a higher order chain homotopy, and so on (see Appendix **??**). The reader may be familiar with these higher order chain homotopies, as they are responsible for the construction of the *Steenrod squares* in classical algebraic topology. They will also play an important role when we come to construct an algebraic model for surgery theory.

2.42. REMARK. By elaborating these techniques slightly we can also prove the *Alexander duality theorem*: Let K be an oriented combinatorial homology n-manifold, and let L be a subcomplex of K'. Then cap-product with the fundamental class induces an isomorphism of (R, K)-module chain complexes

$$C^{n-\bullet}(L; R) \to C_\bullet(K', K' \ominus L; R).$$

Notice that when L consists of a single vertex, this is just the definition of orientation.

Although we have followed the classical approach to duality using triangulations and dual cells, Poincaré duality does not depend on the existence of such a combinatorial structure. Using Mayer-Vietoris arguments similar to those we employed for de Rham cohomology, one can for instance prove an Alexander duality theorem for topological homology manifolds:

2.43. THEOREM. *Let M be an oriented topological homology n-manifold (compact or not), and let $C \subseteq M$ be a compact subset. Then the cap-product with the orientation class defines duality isomorphisms*

$$D\colon \check{H}^r(C;R) \to H_{n-r}(M, M \setminus C; R)$$

where $\check{H}$ denotes Čech cohomology.

SKETCH OF PROOF. One verifies the theorem first when C is either empty (obvious) or is a *small cell* in M, that is a closed ball in some coordinate chart. In the latter case K is homotopy equivalent to a point and $M \setminus C$ is homotopy equivalent to the complement of that point, so the result follows from the definition of orientation. Now by the usual Mayer-Vietoris argument we can handle the case where C is a finite 'good' union of small cells. Given any closed set C and any open neighborhood U one can find C', $C \subseteq C' \subseteq U$, which is such a union of small cells; using the continuity property of Čech cohomology we can therefore complete the proof. For more details see Dold [?]. □

Some standard consequences are the separation theorems of Brouwer, generalizing the Jordan curve theorem.

sep

2.44. EXERCISE. Let C be a closed subset of a compact connected n-manifold. Show that the number of connected components of $M \setminus C$ is equal to $1 + \dim \operatorname{Coker}(H^{n-1}(M;\mathbb{Z}/2) \to \check{H}^{n-1}(C;\mathbb{Z}/2))$. (Use duality and exact sequences.)

2.45. EXERCISE. Prove the *Jordan-Brouwer separation theorem:* Any homeomorphic image K of a compact connected $(n-1)$-manifold (in particular, of S^{n-1}) in S^n separates S^n into two connected components, of which it is the common boundary. (Use the previous exercise.)

2.46. EXERCISE. Prove the theorem of *invariance of domain*: Let $U \subseteq \mathbb{R}^n$ be open, $f\colon U \to \mathbb{R}^n$ be continuous and injective; then $f(U)$ is open in $\mathbb{R}^n$. (Let $p \in U$ and surround p by a small sphere S^{n-1} in U; argue that $f(p)$ must belong to the unique bounded component of the complement of $f(S^{n-1})$, which must be the image under f of the interior disc to S^{n-1} in U; hence $f(p)$ belongs to the interior of the image.)

Suppose now that $(W, \partial W)$ is a compact smooth $(n+1)$-manifold with boundary. An *orientation* in this case is by definition a class $[W] \in H_n(W, \partial W; R)$ that restricts to a generator of $H_n(W, W \setminus \{x\}; R)$ for each $x \in W^\circ$, the interior of W. It is easy to check that $\partial[W] \in H_{n-1}(\partial W; R)$ is then an orientation for ∂W. Cap-product with the relevant orientation classes gives a diagram of duality maps

$$\begin{tikzcd}
{} \arrow[r] & H^{n-r+1}(W) \arrow[r] \arrow[d] & H^{n-r+1}(\partial W) \arrow[r] \arrow[d] & H^{n-r}(W, \partial W) \arrow[r] \arrow[d] & {} \\
{} \arrow[r] & H_r(W, \partial W) \arrow[r] & H_{r+1}(\partial W) \arrow[r] & H_{r+1}(W) \arrow[r] & {}
\end{tikzcd}$$

which commutes up to sign.

2.47. PROPOSITION (Lefschetz duality). *All the duality maps in the diagram above are isomorphisms.*

PROOF. We already know that the absolute duality map for ∂W is an isomorphism, so by the five lemma it suffices to prove that one of the relative duality maps is an isomorphism too, say the map $H^{n-r}(W) \to H_r(W, \partial W)$. One can regard this as an Alexander duality map for W considered as a closed subset of its 'double', obtained by joining two copies of W, with opposite orientations, along their common boundary. □

A corollary whose importance that we have already seen is

CIOS

2.48. PROPOSITION (Cobordism invariance of the signature). *Let W^{4j+1} be an oriented manifold with boundary ∂W. Then the signature* $\operatorname{Sign}(\partial W) = 0$.

PROOF. Let $M = \partial W$, let $i\colon M \to W$, and consider the subspace V which is the image of $i^*\colon H^{2j}(W) \to H^{2j}(M)$ in the middle-dimensional cohomology of M (we take coefficients in $\mathbb{R}$ throughout this proof). Then I claim that V is exactly equal to its own annihilator with respect to the intersection form $(x, y) \mapsto \langle x, D(y)\rangle$. For the proof, consider the diagram of duality maps, and write

$$x \in V \Leftrightarrow i_* D(x) = 0 \Leftrightarrow \langle H^{2j}(W), i_* D(x)\rangle = \{0\} \Leftrightarrow \langle V, D(x)\rangle = \{0\}.$$

But elementary linear algebra shows that if a symmetric bilinear form over $\mathbb{R}$ admits a subspace which is equal to its own annihilator (such a subspace is called *Lagrangian*) then it has signature zero. □

6. The Whitney Trick

Let N_1 and N_2 be smooth submanifolds of M having complementary dimensions and intersecting transversely. Clearly the *absolute* number of intersection points of N_1 and N_2 may be greater than the modulus of the algebraic intersection number $\lambda(N_1, N_2)$, because of the possibility of cancellation of intersection points of opposite sign. We would like to know that it is possible to deform the situation so that the superfluous intersection points cancel. For example, the graph of $y = x^3 - x$ intersects the x-axis in three points -1,0,1; the signs alternate, so the algebraic intersection number is 1. By continuous deformation one can move the x-axis up to the line $y = 2$, which now intersects $y = x^3 - x$ only in one point — the number of intersections required by the algebra.

Our objective in this section is to prove the *Whitney Lemma*, which is

2.49. LEMMA. *Let M be an n-dimensional manifold. Suppose that*

(a) *$N_1^{k_1}$ and $N_2^{k_2}$ are transversely intersecting oriented submanifolds of M, $n = k_1 + k_2$, $k_1, k_2 \geqslant 3$,*
(b) *P and P' are intersection points of N_1 and N_2, having opposite signs, and*
(c) *there exist paths γ_1 and γ_2 from P to P', lying in N_1 and N_2 respectively, such that the loop $\gamma_1^{-1}\gamma_2$ is nullhomotopic in M.*

Then there is an ambient isotopy of N_1 to a submanifold N_1' transverse to M and such that $N_1' \cap N_2 = N_1 \cap N_2 \setminus \{P, P'\}$ (including signs).

In other words, if two intersection points cancel algebraically (in an appropriate sense), then they can be canceled geometrically.

Ambient isotopy means that there exists a smooth family of diffeomorphisms $h_t\colon M \to M$ such that h_0 is the identity and $h_1(N_1) = N_1'$. With some more careful hypotheses one can relax the dimension requirements somewhat[1]; this is important for the proof of the h-cobordism theorem. The proof of the Whitney lemma depends on the easy part of Whitney's embedding theory for submanifolds, and the lemma itself is the key to the hard part of that theory.

If M is oriented we can give each of the intersection points P and P' a sign in the usual way (see the discussion preceding Equation 2.12). In fact, though, we do not need to assume that M is oriented; to define the signs of the intersection points we can choose an arbitrary orientation of M at P and then transport the orientation to P' along either of the (homotopic) paths γ_1 and γ_2. While the actual signs associated to the two intersection points will of course depend on the choice of orientation of M at P, the notion 'P and P' have opposite sign' will not.

PROOF. The idea of the proof is illustrated in the figure below. Suppose that we want to cancel the two intersection points P and P' of N_1 and N_2 shown in the figure. Join p and q by paths γ_1 and γ_2 in N_1 and N_2 respectively, so that the loop γ they form is

[1] To be precise, we can allow $n \geqslant 5$, $k_1 \geqslant 3$, provided that if $k_2 < 3$ we also assume that the induced map $\pi_1(M \setminus N_2) \to \pi_1(M)$ is an injection.

FIGURE 3. The Whitney Trick

nullhomotopic. We may assume without loss of generality that they do not meet any other intersection points. Now there is a homotopy class of maps $D^2 \to M$ with boundary γ realizing the nullhomotopy, and by the easy part of Whitney's embedding theorem, this homotopy class contains an embedding of a disc (because $n \geqslant 5 = 2 \cdot 2 + 1$); this disc may be assumed to be disjoint from N_1 and N_2 (this is an easy special case of transversality theory; remember that the codimension of both N_1 and N_2 is at least three). Such an embedded disc is called the *Whitney disc*. Now we use this as a guide for an isotopic 'push' of M parallel to the Whitney disc in a small neighborhood thereof; such a 'push' can be constructed to shove N_1 right through N_2, thereby getting rid of the intersection points, and to leave everything fixed outside a small neighborhood of the Whitney disc.

To be a bit more precise, let us define a 'standard Whitney model' to be the following configuration of two submanifolds intersecting transversely in $\mathbb{R}^n$. Write $\mathbb{R}^n = \mathbb{R}^{k_1-1} \times \mathbb{R}^2 \times \mathbb{R}^{k_2-1}$ and let γ_1 and γ_2 be the two transversely intersecting curves in the plane $\mathbb{R}^2$ given by the axis $y = 0$ and the parabola $y = x^2 - 1$. Let N_1 be the k_1-dimensional submanifold $\mathbb{R}^{k_1-1} \times \gamma_1 \times \{0\} \subseteq \mathbb{R}^n$, and let N_2 be the k_2-dimensional submanifold $\{0\} \times \gamma_2 \times \mathbb{R}^{k_2-1} \subseteq \mathbb{R}^n$. The proof of Whitney's lemma now has two parts:

(i) There is an ambient isotopy of the standard Whitney model which is equal to the identity off a compact set and which moves N_1 to a new submanifold N_1' of the model which does not intersect N_2;
(ii) In the situation specified by the Whitney lemma, there is a neighborhood (in M) of the Whitney disc which is diffeomorphic to the standard model.

To do

Rewriting done to this point

It is plain that rather more needs to be said, since in our outline above we never made use of the hypothesis that the intersection points had opposite sign! To see where this comes up, note that our construction of the 'push' isotopy assumed that the Whitney disc had a product neighborhood in a rather strong sense. Specifically, one needed to assume that there was a neighborhood of the disc in which the whole set-up was diffeomorphic to the 'standard model': here the neighborhood is $\mathbb{R}^2 \times \mathbb{R}^{n-2}$, γ_1 and γ_2 are two curves in the plane $\mathbb{R}^2$, say for definiteness[2] the axis $y = 0$ and the parabola $y = x^2 - 1$; D is the intervening disc $\{(x, y) : -1 \leqslant x \leqslant 1, x^2 - 1 \leqslant y \leqslant 0\}$; and $N_1 = \gamma_1 \times \mathbb{R}^{k_1 - 1} \times \{0\}$, $N_2 = \gamma_2 \times \{0\} \times \mathbb{R}^{k_2 - 1}$. In this model it is plain that we can construct the isotopy we need.

To embed the model we use normal bundles and the tubular neighborhood theorem. We will find an obstruction. So, in our original set-up, let D^+ be an open disc slightly extending the closed disc D (and similarly for γ_i^+); let ν be the normal bundle to D^+ in M, and let ν_i, $i = 1, 2$, be the normal bundles to γ_i^+ in N_i. These bundles are all over contractible spaces, so they are all trivial. The bundles ν_i are sub-bundles of the restriction of ν to γ_i. By the tubular neighborhood theorem, there are tubular neighborhoods of D^+ and γ_i^+ which are diffeomorphic to the (trivial) total spaces of the bundles ν and ν_i. Moreover, with care we can arrange[3] that the inclusions of the tubular neighborhoods correspond to the inclusions of the sub-bundles ν_i in ν.

To embed our standard model, what we now need to do is to choose an orthonormal frame $\{v_1^1, \dots, v_1^{k_1-1}, v_2^1, \dots, v_2^{k_2-1}\}$ in the normal bundle ν in such a way that the vectors v_1 form an orthonormal frame for ν_1 along γ_1 and the vectors v_2 form an orthonormal frame for ν_2 along γ_2. The only question is whether ν_1 and ν_2 match up correctly at the two points of intersection. Notice that $\nu_1 = \nu_2^\perp$ at the intersection points, so we can define a vector-bundle ξ over the circle γ whose fiber is ν_1 over γ_1 and $\nu_2^\perp$ over γ_2. A simple bundle theory argument shows that we can find the frames we require if and only if ξ is trivial. In general the bundle ξ defines a loop in $G(k_1 - 1, n - 2)$, the Grassmannian of $(k_1 - 1)$-planes in $(n - 2)$-space, and we need to know that this loop is null-homotopic.

Now our assumptions put us in the stable range for calculating the fundamental group of the Grassmannian, so $\pi_1 G(k_1 - 1, n - 1) = \mathbb{Z}/2$. There is just a single element of $\{\pm 1\}$ to calculate, which can be detected by considering orientations, so that ξ is trivial if and only if it is orientable. Since γ_1 and γ_2 intersect with opposite orientations at P and P', ξ will be orientable if and only if the intersection indices $\varepsilon(P)$ and $\varepsilon(P')$ are opposite. Since this was our assumption, the bundle ξ is trivial, and we can embed the standard model and complete the proof. $\square$

EXERCISE: Verify the assertion made above about the fundamental group of the Grassmannian (you will need to use its description as $G(k, n) = O(n)/O(k) \times O(n - k)$, together with the homotopy exact sequence). Also, compute $\pi_1 G(1, 2)$ by the same method; hence find another reason why the Whitney trick fails in dimension four.

WSC 2.50. COROLLARY. *If, in the situation of the Whitney lemma, M is simply-connected, then one can find an ambient isotopy of N_1 to a submanifold N_1'' which intersects N_2 in precisely $|\lambda(N_1, N_2)|$ points. In particular, if $\lambda(N_1, N_2) = 0$, then one can make N_1 and N_2 disjoint by an ambient isotopy.*

For a very careful account of the Whitney lemma and its consequences one should consult chapter 6 of Milnor's book *Lectures on the h-cobordism theorem.*

The hypothesis that M is simply-connected can be removed by developing a theory of equivariant intersection numbers. We will need the theory of group rings. Let π be a group, most likely the fundamental group of something. You will recall that the *group ring* $\mathbb{Z}[\pi]$ is by definition the collection of all finite formal linear combinations $\sum n_g g$, where g runs over the group π. These can be added and multiplied in the obvious way. Notice that

[2]Notice that we need to extend a little past the points of intersection, in order to have room to construct the Whitney isotopy; we might as well take the whole plane as our model, then.

[3]We use a Riemannian metric in which N_1 and N_2 are totally geodesic, and which is Euclidean near the intersection points.

$\mathbb{Z}[\pi]$ is an abelian group equipped with an action of π, and any such group is naturally a module over $\mathbb{Z}[\pi]$.

2.51. EXAMPLE. The group ring $\mathbb{Z}[\mathbb{Z}]$ is the ring $\mathbb{Z}[t, t^{-1}]$ of finite formal Laurent series over $\mathbb{Z}$.

2.52. DEFINITION. Let M be a compact connected manifold, with a preferred basepoint, and let $\pi_1 M = \pi$. A *π-trivial* submanifold N consists of a connected submanifold $N \subseteq M$ such that the image of $\pi_1 N \to \pi$ is the trivial group, together with a preferred homotopy class of paths from the basepoint of M to some fixed point of N.

Suppose that two oriented π-trivial submanifolds N_1 and N_2 of complementary dimensions intersect transversely. Then to each intersection point $p \in N_1 \cap N_2$ we may associate an element $g_p \in \pi$, namely the homotopy class of the path that runs from the basepoint in M, via the preferred route to N_1, then by a path in N_1 to p, then back from p by a path in N_2 to the preferred point in N_2, and back by the preferred route to the basepoint in M.

Suppose now that we choose an orientation for M at the basepoint (we can always do this, even though M need not be globally orientable). We can define a sign $\varepsilon(p) \in \{\pm 1\}$ for the intersection point p by comparing the orientation at p induced from the orientations of N_1 and N_2 with the orientation transported from the basepoint along the path for N_1.

equiv-intersect

2.53. DEFINITION. In the situation of the previous paragraphs, define the *equivariant intersection number* of N_1 and N_2 by $\lambda(N_1, N_2)_\pi \in \mathbb{Z}[\pi]$ by

$$\lambda(N_1, N_2)_\pi = \sum_{p \in N_1 \cap N_2} \varepsilon(p) g_p.$$

An alternative version of this definition can be given by considering the universal cover $\tilde{M}$ of M. It is easy to see that a π-trivial submanifold $N \subseteq M$ can equivalently be defined as one for which there is given a submanifold $\tilde{N}$ of $\tilde{M}$ that is mapped homeomorphically onto N by the covering map $\tilde{M} \to M$ of the universal cover. Now suppose that we have two transversely intersecting π-trivial submanifolds N_1 and N_2 as above. Notice that $\tilde{M}$ is oriented by the choice of orientation at the basepoint of M. Then, for each $g \in G$, the submanifolds $\tilde{N}_1$ and $g^{-1}\tilde{N}_2$ have an ordinary intersection number $[\tilde{N}_1 : g^{-1}\tilde{N}_2]$, and it is not hard to verify the identity

$$\lambda(N_1, N_2)_\pi = \sum_{g \in G} [\tilde{N}_1 : g^{-1}\tilde{N}_2] g.$$

2.54. DEFINITION. If $x = \sum n_g g$ belongs to $\mathbb{Z}[\pi]$, we define $|x| = \sum |n_g|$.

We can now generalize Corollary 2.50 as follows. The statement is the same, except that simple-connectedness of M has been weakened to π-triviality of the submanifolds, and we use equivariant intersection numbers.

WNSC

2.55. COROLLARY. *If, in the situation of the Whitney lemma, N_1 and N_2 are π-trivial in M, then one can find an ambient isotopy of N_1 to a submanifold N_1'' which intersects N_2 in precisely $|\lambda(N_1, N_2)_\pi|$ points. Hence, in particular, if $\lambda(N_1, N_2)_\pi = 0$, then one can make N_1 and N_2 disjoint by an ambient isotopy.*

For the proof, we merely note that superfluous intersections belonging to the same $g \in \pi$ can indeed be canceled by the Whitney trick, since the definition of equivariant intersection numbers provides the desired paths γ_1 and γ_2.

7. Algebra over group rings

We began this chapter with the algebraic theory of Poincaré duality. Then we developed the Whitney trick, which shows that for *simply-connected* manifolds (and in high dimensions) the *algebraic* intersection theory provided by Poincaré duality precisely matches the *geometric* intersection theory provided by counting intersection points. Studying the mechanism of the Whitney trick led us to the equivariant intersection numbers, which we need in order to carry out the trick when the ambient manifold M is not simply connected. In this and the following section we shall complete the chiasmus by relating equivariant intersection numbers back to algebra. This time, however, the algebra will be that of *equivariant* Poincaré duality. To begin with we study modules and homological algebra over group rings.

The basic objects of homological algebra are modules, tensor-products, and Hom-sets. When we work with modules over a group ring $R = \mathbb{Z}[\pi]$ we must face the fact that R is unlikely to be commutative. This means that there is a basic distinction to be drawn between

(a) *left modules* V over R (equipped with a multiplication $R \times V \to V$, satisfying the associativity law $(rs)v = r(sv)$),
(b) *right modules*, equipped with a multiplication $V \times R \to V$ satisfying the associativity law $v(rs) = (vr)s$, and
(c) *bimodules*, equipped with both a left and a right module structure and satisfying the compatibility law $(rv)s = r(vs)$.

The distinction corresponds to that between left, right, and two-sided ideals in a noncommutative ring.

One needs to exercise care in forming tensor products and Hom-sets. For instance, if V is a right R-module and W a left R-module, then $V \otimes_R W$ may be defined: it is the quotient of the tensor product in the category of additive groups by the subgroup generated by expressions

$$vr \otimes w - v \otimes rw, \quad v \in V, r \in R, w \in W.$$

Notice that this tensor product has no module structure — it is simply an abelian group. However, if V is a bimodule then the tensor product inherits a left R-module structure from V; if W is a bimodule it inherits a right R-module structure from W; and of course if both V and W are bimodules, then $V \otimes_R W$ is a bimodule as well. Similar remarks apply to Hom-sets $\mathrm{Hom}_R(V, W)$ (now V and W need to be modules of the same handedness, both left or both right, in order that $\mathrm{Hom}(V, W)$ be defined.)

Let X be a CW-complex[4]. The notation $\mathcal{C}(X)$ denotes the *cellular chain complex* of X, which has one generator in dimension q for each q-simplex of X. There is a natural map of complexes $\mathcal{C}(X) \to \mathcal{S}(X)$ including the cellular complex in the singular complex; this map is a chain equivalence.

Suppose now that X is a finite complex, with fundamental group π. The universal cover $\tilde{X}$ is then also a CW-complex, on which π acts freely[5] with one π-orbit of cells in $\tilde{X}$ corresponding to each individual cell of X. It follows then that $\mathcal{C}(\tilde{X})$ may be thought of as a complex of finitely generated, free right $\mathbb{Z}[\pi]$-modules.

It remains true that $\mathcal{C}(\tilde{X}) \to \mathcal{S}(\tilde{X})$ is a chain equivalence (in the category of complexes of right $\mathbb{Z}[\pi]$-modules). This proves the topological invariance of our constructions below.

2.56. DEFINITION. Let V be a left $\mathbb{Z}[\pi]$-module. The *homology of X with coefficients V*, written $H^\pi_*(X; V)$, is the homology of the complex

$$\mathcal{C}(\tilde{X}) \otimes_{\mathbb{Z}[\pi]} V$$

of abelian groups.

In general $H^\pi_*(X; V)$ is an abelian group; if V is a $\mathbb{Z}[\pi]$-bimodule, then the homology is naturally a right $\mathbb{Z}[\pi]$-module.

[4]See Appendix C for more about CW-complexes and cellular homology.
[5]Our convention is that π acts on the *right*.

2.57. DEFINITION. Let W be a right $\mathbb{Z}[\pi]$-module. The *cohomology of X with coefficients W*, written $H^*_\pi(X; W)$, is the cohomology of the complex

$$\mathrm{Hom}_{\mathbb{Z}[\pi]}(\mathcal{C}(\tilde{X}, W))$$

of abelian groups.

In general $H^*_\pi(X; W)$ is an abelian group; if W is a $\mathbb{Z}[\pi]$-bimodule, then the cohomology is naturally a left $\mathbb{Z}[\pi]$-module.

2.58. REMARK. There is a pairing

$$H^k_\pi(X; W) \otimes_{\mathbb{Z}} H^\pi_k(X; V) \to V \otimes_{\mathbb{Z}[\pi]} W$$

between homology and cohomology.

2.59. EXAMPLE. The group $\mathbb{Z}$ has a natural $\mathbb{Z}[\pi]$-bimodule structure in which every group element acts as the identity. The homology and cohomology groups $H^\pi_*(X; \mathbb{Z})$ and $H^*_\pi(X; \mathbb{Z})$ are canonically isomorphic to the usual homology and cohomology groups of X with integer coefficients. Here is why: there is an obvious map of complexes of abelian groups

$$\mathcal{C}(\tilde{X}) \to \mathcal{C}(X)$$

which sends each cell of $\tilde{X}$ to its image in X. Since the image of a cell of $\tilde{X}$ under this map is the same as the image of each of its π-translates, the map passes to the quotient

$$\mathbb{Z} \otimes_{\mathbb{Z}[\pi]} \mathcal{C}(\tilde{X}) \to \mathcal{C}(X).$$

This latter map is easily seen to be an isomorphism of chain complexes. A similar argument applies to cohomology.

2.60. EXAMPLE. Now consider homology and cohomology with coefficients in the bimodule $\mathbb{Z}[\pi]$. Since $\mathbb{Z}[\pi] \otimes_{\mathbb{Z}[\pi]} \mathcal{C}(X) = \mathcal{C}(X)$, the homology groups $H^\pi_*(X; \mathbb{Z}[\pi])$ with coefficients $\mathbb{Z}[\pi]$ are just the ordinary homology groups of $\tilde{X}$ with the natural right action of π.

The cohomology groups $H^*_\pi(X; \mathbb{Z}[\pi])$ with coefficients $\mathbb{Z}[\pi]$ are the *compactly supported* cohomology groups of $\tilde{X}$, with the natural left action of π. To see this, we note that a $\mathbb{Z}[\pi]$-module homomorphism from $\mathcal{C}(\tilde{X})$ to $\mathbb{Z}[\pi]$ is the same thing as a $\mathbb{Z}$-module homomorphism φ from $\mathcal{C}(\tilde{X})$ to $\mathbb{Z}$ with the additional constraint that for each cell σ, $\varphi(g \cdot \sigma)$ is nonzero for only finitely many $g \in \pi$. But since X has only finitely many cells, this is exactly the same as a compactly supported cochain for $\tilde{X}$.

This is a simple example of translating between *equivariance* and *geometric control* (in this case, compact support). The idea will become much more important later.

2.61. EXAMPLE. As an explicit example, let us consider $X = S^1$ with its usual cell structure with one 0-cell and one 1-cell. We take $\pi = \pi_1(S^1) = \mathbb{Z}$, so $\mathbb{Z}[\pi] = \mathbb{Z}[t, t^{-1}]$. The complex $\mathcal{C}(\tilde{X})$ is then

$$0 \longleftarrow \mathbb{Z}[t, t^{-1}] \xleftarrow{1-t} \mathbb{Z}[t, t^{-1}] \longleftarrow 0$$

One readily computes that its homology is $\mathbb{Z}$ in dimension 0 and trivial in dimension 1, in agreement with the ordinary homology of the universal cover $\tilde{X} = \mathbb{R}$.

Similarly the dual complex $\mathrm{Hom}(\mathcal{C}(\tilde{X}, \mathbb{Z}[\pi])$ is

$$0 \longrightarrow \mathbb{Z}[t, t^{-1}] \xrightarrow{1-t} \mathbb{Z}[t, t^{-1}] \longrightarrow 0$$

whose cohomology is trivial in dimension 0 and $\mathbb{Z}$ in dimension 1, in agreement with the compactly supported cohomology of $\mathbb{R}$. Notice that Poincaré duality apparently still holds! We will investigate this in general in the next section.

With reference to this example, you might be puzzled by the following question: homology (with $\mathbb{Z}[\pi]$ coefficients) is naturally a right $\mathbb{Z}[\pi]$-module, cohomology is naturally a left module. How then can there be a natural Poincaré duality isomorphism between them? The answer is that $\mathbb{Z}[\pi]$ is provided with some extra structure — an involution — which relates left and right actions. The involution allows us to understand the symmetry properties of Poincaré duality , which are critical in setting up the surgery obstruction groups.

2.62. DEFINITION. An *involution* on a ring R is a map $R \to R$, denoted $x \mapsto x^*$, which is a homomorphism of abelian groups, preserves the unit, and has $(xy)^* = y^*x^*$ and $x^{**} = x$ for all $x, y \in R$.

A ring with involution will be called a $*$-*ring*.

2.63. EXAMPLE. Conjugation on $\mathbb{C}$ or on $\mathbb{H}$ is an involution. The conjugate transpose on a ring of matrices over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ is an involution. The adjoint on the ring of bounded operators on a Hilbert space (or on any C^*-subalgebra, such as the ring of compact operators) is an involution.

More relevant to our purposes is the following.

2.64. PROPOSITION. *The map $g \mapsto g^{-1}$ extends (by linearity) to an involution on the group ring $\mathbb{Z}[\pi]$. More generally, the same is true of the map $g \mapsto w(g)g^{-1}$ where $w\colon \pi \to \{\pm 1\}$ is any group homomorphism.* □

These are called the *standard*, respectively the w-*twisted*, involutions on the group ring.

We may want to consider only the oriented case, and put the unoriented case into exercises. To do

In our situation π is usually the fundamental group $\pi_1(M)$ of some manifold. based loops γ in M can be classified as *orientation-preserving* or *orientation-reversing* according to whether a local orientation at the basepoint is preserved or reversed under smooth transport around γ. The homomorphism $\pi_1(M) \to \{\pm 1\}$ which sends orientation-preserving loops to $+1$ and orientation-reversing loops to -1 gives a canonical choice for the *orientation character* w on π. Note that M is orientable if and only if this w is identically 1.

2.65. EXERCISE. By abuse of notation, we will denote the 2-element group $\{\pm 1\}$ by $\mathbb{Z}/2$. By the Hurewicz theorem we have in fact $\mathrm{Hom}(\pi_1(M), \mathbb{Z}/2) = H^1(M; \mathbb{Z}/2)$. Show that under this isomorphism the orientation character w corresponds to the *first Stiefel-Whitney class* of M. (See [**13**] for the Stiefel-Whitney classes.)

Let R denote a ring with involution. Given a right R-module V, the *opposite* left R-module, V^o, is defined to be V with the left action of R given by

$$(r, v) \mapsto vr^*.$$

Similarly we can define the opposite of a left R-module, and even the opposite of an R-bimodule (take the opposite of each structure).

2.66. EXERCISE. Let R be a ring with involution. Verify that R^o is isomorphic to R
bimod-ex as an R-bimodule.

2.67. EXERCISE. Suppose that the group ring $\mathbb{Z}[\pi]$ is provided with the involution associated to $w\colon \pi \to \{\pm 1\}$. Give $\mathbb{Z}$ the trivial right $\mathbb{Z}[\pi]$-module structure in which each group element acts as 1. What is the structure of the left $\mathbb{Z}[\pi]$-module $\mathbb{Z}^o$? Show that the left and right actions commute so that $\mathbb{Z}$ becomes a $\mathbb{Z}[\pi]$-bimodule. (We denote it by $\mathbb{Z}^w$ when it is considered as a bimodule in this way.)

We isolate here a useful algebraic calculation.

diag-tensor

2.68. PROPOSITION. *Let U and V be right $\mathbb{Z}[\pi]$-modules. Equip $\mathbb{Z}[\pi]$ with the w-twisted involution, and let $\mathbb{Z}^w$ denote the integers considered as a left $\mathbb{Z}[\pi]$-module as in the previous exercise. The tensor product $U \otimes_{\mathbb{Z}} V$ in the category of abelian groups is made into a right $\mathbb{Z}[\pi]$-module by the diagonal action $(u \otimes v)g = ug \otimes vg$. Then there is a natural isomorphism*

$$(U \otimes_{\mathbb{Z}} V) \otimes_{\mathbb{Z}[\pi]} \mathbb{Z}^w \cong U \otimes_{\mathbb{Z}[\pi]} V^o$$

in the category of abelian groups.

PROOF. Send an element $x = u \otimes v \in U \otimes_{\mathbb{Z}[\pi]} V^o$ to the element $u \otimes v \otimes 1 \in (U \otimes_{\mathbb{Z}} V) \otimes_{\mathbb{Z}[\pi]} \mathbb{Z}^w$. The map is well-defined because if we represent x also by $(ug) \otimes (g^{-1}v)$, the image is

$$(ug) \otimes (g^{-1}v) \otimes 1 = w(g)(ug) \otimes (vg) \otimes 1 = u \otimes v \otimes 1 \in (U \otimes_{\mathbb{Z}} V) \otimes_{\mathbb{Z}[\pi]} \mathbb{Z}^w$$

using the definitions of the opposite module and of the involution in $\mathbb{Z}[\pi]$. The reader may verify similarly that the map is, in fact, an isomorphism. $\square$

The following example shows why involutions on $\mathbb{Z}[\pi]$ naturally arise in studying equivariant intersection theory.

2.69. EXAMPLE. Let N_1 and N_2 be transversely intersecting oriented π-trivial submanifolds of M, having complementary dimensions k_1 and k_2. Then their equivariant intersection numbers (Definition 2.53) are related by

$$[N_2 : N_1]_\pi = (-1)^{k_1 k_2} \lambda(N_1, N_2)^*_\pi,$$

where $*$ is the involution on $\mathbb{Z}[\pi]$ associated to the first Stiefel-Whitney class. Indeed, reversing the order of N_1 and N_2 replaces each group element g appearing in the sum defining the equivariant intersection number by its inverse, and also multiplies the orientation-dependent coefficient $\varepsilon(g)$ by the factor $w(g)$.

8. Orientations and equivariant duality

Our new homology and cohomology theories H^π_* and H^*_π enjoy suitable versions of the usual kinds of functorial properties, including excision, Mayer-Vietoris sequences, homotopy invariance and so on. In particular, it is true just as for ordinary homology that if M is an n-manifold, then $H^\pi_i(M, M \setminus \{x\}; \mathbb{Z}^w)$ is isomorphic to 0 for $i \neq n$, $\mathbb{Z}$ for $i = n$.

2.70. DEFINITION. An *orientation* of M for the orientation character w is a class $[M] \in H^\pi_n(M; \mathbb{Z}^w)$ which restricts to a generator of $H^\pi_n(M, M \setminus \{x\}; \mathbb{Z}^w)$ for all $x \in M$.

This is exactly the same definition as we previously gave in the non-equivariant case. But now we have

2.71. PROPOSITION. *Any (connected, compact) manifold M is orientable for the orientation character defined by the first Stiefel-Whitney class.*

PROOF. The manifold M has an *orientation cover* $\bar{M}$, a $\mathbb{Z}/2$ cover whose fiber over $p \in M$ consists of the possible orientations of M at p. The orientation cover $\bar{M}$ is the $\mathbb{Z}/2$-cover associated to the Stiefel-Whitney class $w = w_1 \colon \pi_1 M \to \mathbb{Z}/2$, and it is (tautologically) orientable. Now take a fundamental cycle for $\bar{M}$ and lift it (on the chain level) to a cycle on the universal cover; one sees directly that this lifted cycle belongs to $\mathcal{C}(\tilde{M}) \otimes \mathbb{Z}^w$, so it defines a w-twisted orientation for M. $\square$

2.72. REMARK. The usual notion of orientation is a special case, with $w_1 = 1$; you may prefer to focus on this case to start with. But our machinery handles the general case with almost no extra effort.

Our intention is to set up Poincaré duality for manifolds in the context of twisted homology. Having obtained the notion of orientation, the next task is to define a suitable cap-product. So, let X be a finite complex (or just a compact Hausdorff space, if we use singular theory), with fundamental group π, and let w be an orientation character for π. Let V be a right $\mathbb{Z}[\pi]$-module. For every $a \in H_r^\pi(X; \mathbb{Z}^w)$ we want to define a *cap-product*

$$\frown a : H^s_\pi(X; V^o) \to H^\pi_{r-s}(X; V)$$

which is a homomorphism of abelian groups. To do this, we begin with an Eilenberg-Zilber diagonal approximation

$$\mathcal{C}(\tilde{X}) \to \mathcal{C}(\tilde{X}) \otimes_{\mathbb{Z}} \mathcal{C}(\tilde{X}).$$

One can manufacture such a diagonal approximation (see Appendix D) which is *equivariant* with respect to the π-action on the tensor product by $(x \otimes y)g = (xg) \otimes (yg)$. Now tensor on the right by the module $\mathbb{Z}^w$. This gives a chain map

$$\mathcal{C}(\tilde{X}) \otimes_{\mathbb{Z}[\pi]} \mathbb{Z}^w \to (\mathcal{C}(\tilde{X}) \otimes_{\mathbb{Z}} \mathcal{C}(\tilde{X})) \otimes_{\mathbb{Z}[\pi]} \mathbb{Z}^w.$$

According to Proposition 2.68, the complex on the right of this display is naturally isomorphic to $\mathcal{C}(\tilde{X}) \otimes_{\mathbb{Z}[\pi]} \mathcal{C}(\tilde{X})^o$. Tensoring over $\mathbb{Z}$ with the complex $\mathrm{Hom}_{\mathbb{Z}[\pi]}(\mathcal{C}(X), V^o)$ which computes the cohomology, this gives us a diagram

$$\begin{tikzcd}
\mathrm{Hom}_{\mathbb{Z}[\pi]}(\mathcal{C}(X), V^o) \otimes_{\mathbb{Z}} (\mathcal{C}(\tilde{X}) \otimes_{\mathbb{Z}[\pi]} \mathbb{Z}^w) \arrow[d] \\
\mathrm{Hom}_{\mathbb{Z}[\pi]}(\mathcal{C}(X), V^o) \otimes_{\mathbb{Z}} (\mathcal{C}(X) \otimes_{\mathbb{Z}[\pi]} \mathcal{C}(X)^o) \arrow[d, "\text{evaluation}"] \\
V^o \otimes_{\mathbb{Z}[\pi]} \mathcal{C}(X)^o
\end{tikzcd}$$

in which the arrows are maps of complexes. Passing to homology this gives the desired
cap-def product

$$H_r^\pi(X; \mathbb{Z}^w) \otimes H^s_\pi(X; V^o) \to H^\pi_{r-s}(X; V) \tag{2.73}$$

2.74. DEFINITION. The product defined in Equation 2.73 above is called the *cap product* between H^*_π and H_*^π.

Assuming for simplicity that $w = +1$ we can obtain a more geometric picture of the cap product as follows. There is defined an *infinite transfer* map T from the ordinary homology $H_r(X; \mathbb{Z})$ to the locally finite homology $H_n^{lf}(\tilde{X}; \mathbb{Z})$ of the universal cover. The usual (locally finite) cap product with $T(a)$ defines a map

$$H^s_c(X; \mathbb{Z}) \to H_{r-s}(X; \mathbb{Z}).$$

The cap-product above is just this map.

2.75. REMARK. In the situation of the cap-product, above, suppose that V is not merely a right module but a *bimodule*. Then $H^s_\pi(X; V^o)$ is a right $\mathbb{Z}[\pi]$-module and $H^\pi_{r-s}(X; V)$ is a left $\mathbb{Z}[\pi]$-module. The cap-product with a is now a module map from the opposite of cohomology to homology.

When $V = \mathbb{Z}[\pi]$ itself, which will be the most important case, we can apply the result of Exercise 2.66 that $V \cong V^o$ as bimodules and express the cap-product with a as a map of left $\mathbb{Z}[\pi]$-modules

$$H^s_\pi(X; \mathbb{Z}[\pi])^o \to H^\pi_{r-s}(X; \mathbb{Z}[\pi]).$$

When calculating with this expression of the cup product it is important not to forget the extra involution that has been introduced by the isomorphism $\mathbb{Z}[\pi] \cong \mathbb{Z}[\pi]^o$.

Let M^n be a compact manifold, oriented with orientation character w. Then the cap-product with the fundamental class defines $\mathbb{Z}[\pi]$-module morphisms

duality-eq

$$D\colon H^r(M; \mathbb{Z}[\pi])^o \to H^w_{n-r}(M; \mathbb{Z}[\pi]). \tag{2.76}$$

Following word-for-word the proofs in the non-equivariant case, we find

To do

This 'following the proofs' needs to be expanded into a discussion of the assembly process from $(\mathbb{Z}, K)$-modules to $\mathbb{Z}[\pi_1 K]$ modules.

2.77. THEOREM (Universal Poincaré duality). *The equivariant duality maps D for a compact oriented manifold, defined in Equation 2.76 above, above are isomorphisms.* □

Now we make the connection to intersection theory. Let N_1 and N_2 be transversely intersecting oriented π-trivial submanifolds of M, of complementary dimensions. Recall that a π-trivial submanifold N^k has a preferred lift $\tilde{N}$ to a submanifold of the universal cover of M. The fundamental class of $\tilde{N}$ then maps to a class $[N] \in H^w_k(M; \mathbb{Z}[\pi])$. We use the orientation character coming from the first Stiefel-Whitney class.

2.78. THEOREM. *The equivariant intersection number of transversely intersecting submanifolds as above is related to equivariant Poincaré duality by*

$$\lambda(N_1, N_2)_\pi = D^{-1}[N_1][N_2].$$

PROOF. The equivariant intersection number of N_1 and N_2 is a sum, over $g \in \pi$, of the ordinary intersection numbers of $\tilde{N}_1$ and $g^{-1}\tilde{N}_2$. Now apply ordinary intersection theory in the universal cover. □

Together with corollary 2.55, this theorem provides the essential link between quadratic algebra over $\mathbb{Z}[\pi]$ and geometric intersections. Our application will be to discover when homology classes can be represented by disjoint embedded spheres, so that we can do surgery on them.

CHAPTER 3

General Position

In this chapter we will develop the 'general position' techniques, such as transversality, that will allow us to construct manifolds and embeddings. We have already seen that transversality is used at a key point in the argument of Browder (chapter 1).

The key to all these general position results is found in Sard's theorem. Let $f\colon M^m \to N^n$ be a smooth map of manifolds. The *critical set* $C_f \subseteq M$ of M is the set of $m \in M$ such that the tangent map $df_m\colon T_mM \to T_{f(m)}N$ fails to be surjective.

3.1. THEOREM. *(Sard) If f is a smooth map as above, then $f(C_f)$ has measure zero in N.*

The notion 'has measure zero' makes sense on any smooth manifold, even though there is no canonical choice of smooth measure. There is a disconcerting example which shows that high differentiability is necessary in Sard's theorem; this is an example of a C^1 map $f\colon \mathbb{R}^2 \to \mathbb{R}$, such that there is a (topological) arc γ in $\mathbb{R}^2$ for which $df = 0$ at all points of the arc, but nevertheless f is not constant along γ. The image of the critical set thus contains an open subset of $\mathbb{R}$.

Sard's theorem is certainly a basic, foundational result, and everyone should see the proof at least once. Nevertheless, we will not give it here.

AE 3.2. COROLLARY. *If $f\colon M^m \to N^n$ is smooth, and $m < n$, then the image of f has measure zero.*

For this corollary, only C^1 differentiability is in fact necessary.

1. Embeddings and Immersions

3.3. DEFINITION. Let $f\colon M \to N$ be a smooth map between manifolds. Then f is called an *immersion* if the tangent map $df_x\colon T_xM \to T_{f(x)}N$ is injective for all $x \in M$. It is an *embedding* if it is an immersion and, in addition, is a homeomorphism of M onto the image $f(M)$ (equipped with the topology it inherits as a subspace of N)

We will mostly be interested in compact manifolds; in this case any injective immersion is an embedding. For it is a standard result of elementary topology that a continuous bijection from a compact space to a Hausdorff space is in fact a homeomorphism.

The smooth map $\mathbb{R} \to T^2$ which wraps $\mathbb{R}$ densely around the 2-torus, using an irrational slope, is an example of an injective immersion of a non-compact manifold which is not an embedding.

We are going to construct embeddings and immersions of compact manifolds into Euclidean space, and eventually into other manifolds. A first step is provided by

3.4. PROPOSITION. *Let M^n be a smooth manifold, $K \subseteq M$ a compact subset. Then there is a smooth map $f\colon M \to \mathbb{R}^k$, for some large k, which is an embedding on a neighbourhood of K. In particular, any compact manifold can be embedded in a Euclidean space.*

PROOF. Cover K by finitely many coordinate charts $U_1, \dots, U_m$, with embeddings $h_i\colon U_i \to \mathbb{R}^n$. Let φ_i be a system of bump functions subordinated to U_i — by this I mean that φ_i is supported within U_i and that for each $x \in K$ there is some index i such that $\varphi_i(x) = 1$. Then the map $M \to \mathbb{R}^{n(m+1)}$ defined by

$$x \mapsto (\varphi_1(x)h_1(x), \dots, \varphi_m(x)h_m(x), \varphi_1(x), \dots, \varphi_m(x))$$

is easily seen to be an embedding. $\square$

The partition of unity construction yields a very large k. We now seek to reduce k. This we can do by general position arguments.

IIT 3.5. LEMMA. *Let $f\colon M \to \mathbb{R}^k$ be an embedding of a compact manifold (or of a compact piece of a manifold, as above), and suppose $k > 2n+1$. Then for almost all unit vectors $v \in \mathbb{R}^k$ the projection $P_v\colon \mathbb{R}^k \to \mathbb{R}^{k-1}$ orthogonal to v has the property that $P_v f$ is an embedding. If $k = 2n+1$, then for almost all unit vectors v, $P_v f$ is an immersion.*

PROOF. Think of M as a submanifold of $\mathbb{R}^k$. What can go wrong? $P_v f$ may fail to be one-to-one, or it may fail to be immersive. To say that $P_v f$ is not one-to-one is to say that v belongs to the image of the map

$$M \times M \setminus \Delta \to S^{k-1}, \qquad (x, y) \mapsto \frac{x-y}{\|x-y\|}.$$

However, by 3.2, the image of this map has measure zero, since $2n < k-1$. To say that $P_v f$ is not immersive is to say that v belongs to the image of the unit tangent bundle $T_1 M$ under the map $T_1 M \to S^{k-1}$ that sends each unit tangent vector to 'itself'. But again, by 3.2, the image of this map has measure zero; this is true even when $k = 2n+1$. This proves both results. $\square$

3.6. THEOREM. *(Whitney) Any map $M^n \to \mathbb{R}^{2n+1}$ may be arbitrarily well approximated by an embedding, and any map $M^n \to \mathbb{R}^{2n}$ may be arbitrarily well approximated by an immersion.*

PROOF. We do it for embeddings. Let $f\colon M \to \mathbb{R}^{2n+1}$ be the given map, and let $g\colon M \to \mathbb{R}^k$ be an embedding of M in some high-dimensional Euclidean space $\mathbb{R}^k$. Then $h = (f, g)\colon M \to \mathbb{R}^{2n+1+k}$ is an embedding, and $f = \Pi h$ where $\Pi\colon \mathbb{R}^{2n+1+k} \to \mathbb{R}^{2n+1}$ is the obvious projection. But by k applications of the previous lemma, we see that Π can be altered by an arbitrarily small amount so as to make Πh an embedding; and it is clear that this new Πh may be as close as we wish to f. $\square$

We would like to prove a similar result for maps of one manifold to another. It will be useful to note the obvious fact that the set of embeddings of M in N, and the set of immersions of M in N, are both open subsets of $C^\infty(M; N)$. (We assume M is compact here.)

WIT 3.7. THEOREM. *Let M, N be manifolds, $\dim N \geqslant 2\dim M + 1$, M compact. Then any smooth map $M \to N$ can be arbitrarily well approximated by an embedding. In fact, if $f\colon M \to N$ is a smooth map which is already an embedding on some closed subset C of M, then f can be arbitrarily well approximated by an embedding which agrees with f on C.*

PROOF. The relative version easily follows from the absolute version and the openness of the set of embeddings. Cover N by finitely many coordinate charts $V_1, \dots, V_\ell$, let $U_i = f^{-1}(V_i)$, and let $K_i \subseteq U_i$ be closed subsets such that $\bigcup K_i = M$. Assume that f is an embedding on $K_1 \cup \dots \cup K_{r-1}$. Using the previous theorem, we can make a small perturbation of f on U_r so that f becomes an embedding on K_r. Because the embeddings form an open set, if we choose the perturbation small enough it will not destroy the property that f is an embedding on $K_1 \cup \cdots \cup K_{r-1}$. An induction on r completes the proof. □

There is a corresponding result for immersions; the reader is left to state and prove it.

2. Transversality

Let $f\colon M \to N$ be a smooth map, and let X be a submanifold of N. Then for each $p \in f^{-1}(X)$, $d_p f$ induces a map from the tangent bundle of M to the normal bundle of X,

$$T_p M \xrightarrow{df} T_{f(p)} N \longrightarrow T_{f(p)} N / T_{f(p)} X$$

We say that f is *transverse to X at p* if this composite map is surjective, and that f is *transverse to* X if it is transverse at all $p \in f^{-1}(V)$. Thus, if $X = \{x\}$ consists of a single point, f is transverse to X if and only if x is a regular value, that is, $x \notin f(C_f)$. By Sard's theorem, this is a generic condition.

A generalization will be of value. In the definition of transversality, it is clear that only the part of f that maps M to some tubular neighbourhood of X is significant. Therefore we may consider the situation where M maps into the total space of some vector-bundle over X. In this situation, the condition that f be transverse does not involve the smooth structure on X at all. There is thus no reason to suppose X to be a manifold. We end up with the following definition:

3.8. DEFINITION. Let $\pi\colon V \to X$ be a vector-bundle over a space X and let $f\colon M \to \operatorname{Th} V$ be a vertically smooth map from M to the *Thom space* of V (that is, the space obtained by collapsing the complement of the unit disc bundle of V to a point). We say that f is *transverse at the zero section of* $\operatorname{Th} V$ if for all $p \in f^{-1}(X)$, the vertical tangent map

$$d_{p,v} f\colon T_p M \to (\pi^* V)_{f(p)}$$

is surjective.

EXERCISE: Check that the notions of 'vertically smooth' and the 'vertical tangent map' make sense. Since we are interested in transversality only at the zero-section, it makes no difference whether we consider maps to V itself or to its Thom space; the latter has the advantage of being compact (provided that X itself is compact).

If f is a transverse map to the Thom space of the bundle $V \to X$, then by the inverse function theorem $f^{-1}(X)$ is a smooth manifold of dimension equal to the dimension of M minus the fibre dimension of V, and its normal bundle in M is identified with the pull-back $f^* V$. In other categories, and with an appropriate notion of 'bundle', this conclusion of the inverse function theorem may be taken as the *definition* of transversality.

We say that two *submanifolds* of M are transverse if the inclusion of one is transverse to the other. (This is a symmetric condition, equivalent to the statement that the tangent spaces of the submanifolds together span the tangent space of M at each point of intersection.) Using the inverse function theorem, one sees that this is equivalent to the definition we gave on page **??** in the last chapter.

3.9. THEOREM. *(Thom) Any map from a compact manifold to the Thom space of a vector bundle can be arbitrarily well approximated by a map which is transverse at the zero section.*

PROOF. This is like the proof of 3.7. Let $f\colon M \to V$ be the given map. We start the proof by choosing an open cover $\{U_i\}$ of M, such that each $f(U_i)$ lies in a trivial part $\mathbb{R}^p \times V_i$ of the bundle, and such that there is a compact cover K_i contained in U_i. On

FIGURE 1. Neat and non-neat submanifolds

each U_i the map f can be represented as (g, h), where $g\colon U_i \to \mathbb{R}^p$ and $h\colon U_i \to V_i$, and transversality just says that zero is a regular value of g_i. Thus, by Sard's theorem, it is possible to make an arbitrarily small perturbation of f to make it transverse on K_i. (Just perturb by a small constant.)

Now we remark that transversality is an open condition (in a suitable topology). Thus we can carry out inductively a sequence of smaller and smaller perturbations over the sets K_r, $r = 1, 2, \ldots$, in order to make f transverse as required. □

It is convenient to make explicit a few points about transversality in the context of manifolds with boundary:

3.10. DEFINITION. Let $(M, \partial M)$ be a manifold with boundary. A submanifold $N \subseteq M$ is called *neat* if $\partial N = N \cap \partial M$ and N meets ∂M transversely.

In the figure, N_1 is a neat submanifold of M, but N_2 is not.

A map from $(M, \partial M)$ to the Thom space of a vector-bundle V is called *transverse at the zero-section* if it is transverse on the interior and its restriction to the boundary is transverse as a map from ∂M. If this is so, then the inverse image of the zero-section is a neat submanifold. The transversality theorem still applies: any map can be perturbed by an arbitrarily small amount so as to make it transverse.

As an illustration of the power of transversality, here is Hirsch's proof of the Brouwer fixed point theorem. You will observe that no algebraic topology is required.

As is well known, to prove Brouwer's theorem it is enough to show that there is no smooth retraction of D^n onto its boundary S^{n-1}. Suppose r is such a retraction. There is some point $p \in S^{n-1}$ such that r is transverse at p. Then $r^{-1}(p)$ is a 1-dimensional neat submanifold of D^n, so it is a finite union of circles and arcs with endpoints on the boundary. One of these arcs must run from p to some other point $q \in S^{n-1}$. Therefore, $r(q) = p$. But since r is a retraction, $q = p$, a contradiction.

3. More about Immersions and Embeddings

3.11. DEFINITION. Let $f\colon M^n \to N^{2n}$ be an immersion. We say that f has *clean double points* if, given any pair $m, m' \in M$ with $f(m) = p = f(m')$, there are neighbourhoods U, U' of m, m' with $f_{|U}$ and $f_{|U'}$ transverse embeddings and with $f(x) \neq p$ for all $x \notin U \cup U'$.

Any immersion $M^n \to N^{2n}$ (and hence, any map) can be approximated by one with clean double points. This uses a local-to-global argument like the ones we have already carried out: by dimension counting as in the proof of 3.5 we may assume there are no triple points, and then, since f is an embedding locally, we can make it transverse to itself locally by the transversality theorem, hence we can do so globally.

In some circumstances it is possible to cancel the double points of an immersion by the Whitney trick. Here is a classical example, due to Whitney (1946?).

hardwhitney

3.12. THEOREM. *Any smooth map $f\colon M^n \to N^{2n}$, with N simply-connected, and $n \geqslant 3$, is homotopic to an embedding.*

We do not assert that f can be approximated by embeddings; this is not true in general.

PROOF. We may assume, without loss of generality, that M is connected and that f is an immersion with clean double points. Consider now a pair of double points p and p'. At each such point two branches of M meet. Choose paths γ_1 and γ_2 in M between the double points, such that near each double point γ_1 and γ_2 lie in transverse branches of M. Choose arbitrarily an orientation of each branch of N at p (thus defining a self-intersection number $\varepsilon(p) \in \{\pm 1\}$) and transport these orientations by γ_1 and γ_2 to define a self-intersection number at p'. If $\varepsilon(p) = -\varepsilon(p')$, then we can apply the Whitney trick: there will be an isotopy of N (hence certainly a homotopy of f) which removes the two double points.

What if not all double points can be removed by this procedure? For any that are left over we can change f in its homotopy class to introduce extra double points to cancel them, by taking the connected sum of M with a sphere S^n immersed to have a single double point (in the case $n = 1$ this is the 'figure 8 immersion' of S^1 in $\mathbb{R}^2$). These extra double points can be manufactured with any desired orientation, and so we can use them to cancel all the double points of f. □

In this argument we defined the signs $\varepsilon(p)$ by making local choices of orientation. Beware of assuming that one might define a global self-intersection number of an immersion by choosing a global orientation of M! This cannot be done in general — the point is that where two branches of M cross we have no canonical way to decide which one is the 'first'. If M is even-dimensional this will not matter,as either choice will give the same sign; but if M is odd-dimensional the choice affects the sign, so the self-intersection will only be definable as an integer modulo 2. Of course we will need to consider the equivariant version of all this! That will lead us fairly directly to Wall's definition of the L-groups. Also beware that the 'self-intersections' we are talking about here are *not* the same as the intersection numbers $[M : M]$; for reasons that will become apparent, $[M : M]$ is referred to as the *symmetric* self-intersection, and the theory that we are sketching defines the *quadratic* self-intersection. What is the relationship between them? All will be revealed in Chapter 100000.

The argument above seems a little cheeky, especially the sudden introduction of the 'figure eight' immersions. Thinking about it, the reader may conclude that the following is a more appropriate equivalence relation than homotopy to consider on the set of immersions.

rh

3.13. DEFINITION. A *regular homotopy* of immersions is a smooth homotopy through immersions (in other words, a smooth map $h\colon M \times [0,1] \to N$ such that each map f_t, $f_t(x) = f(x,t)$, is an immersion.

The Whitney trick produces regular homotopies. However, the figure eight immersion is not regular homotopic to an embedding. To see this, and for other reasons, it will be necessary to understand the possible singularities of a regular homotopy.

cerf

3.14. PROPOSITION. *(Cerf?) Let h_t be a regular homotopy of immersions $M^n \to N^{2n}$. Then h can be arbitrarily well approximated by a regular homotopy in which the immersions h_t all have clean double points except for a finite number of values of t,*

FIGURE 2. A regular homotopy

at which there occur birth or death singularities, where two double points appear or disappear together. The track of each double point from birth to death is a continuous function of t.

The figure below shows a regular homotopy of an embedding of S^1 in $\mathbb{R}^2$ to an immersion with two double points. The birth-death singularity and the tracks of the double points can be easily seen.

PROOF. (Sketch) The regular homotopy can be thought of as an immersion $M \times I \to N \times I$ (defined by $(x, t) \mapsto (h(x, t), t)$). If we make this transverse to itself, a dimension count as above shows that the self-intersections are 1-dimensional submanifolds of $M \times I$, that is, circles and intervals with endpoints on the boundary. These define paths of double points except where their tangent vectors have no component in the direction of M, at which points they define birth-death singularities. $\square$

As a corollary, the number of double points modulo 2 of an immersion with clean double points is an invariant of regular homotopy. So, in particular, the figure eight immersion is not regular homotopic to an embedding.

To do

Insert more about immersions, and the Q-groups, here.

4. The Pontrjagin-Thom Construction

Transversality was applied by Thom to the calculation of the *cobordism* groups of smooth manifolds.

3.15. DEFINITION. Two compact oriented n-manifolds M and M' are *(oriented) cobordant* if there is an oriented compact $(n+1)$-manifold with boundary W such that $\partial W = M \sqcup (-M')$.

Here ∂W is assumed to have the orientation induced from the given orientation of W; and the notation $M \sqcup (-M')$ means the disjoint union of M and M' where M has its given orientation and M' the opposite of its given orientation. One can similarly define *unoriented* cobordism, *spin* cobordism, and other fancy kinds of cobordism.

A simple gluing argument shows that cobordism is an equivalence relation on manifolds. The set Ω_n of cobordism classes of n-dimensional oriented manifolds is an abelian group, using the empty set as the identity element, disjoint union as addition, and $-M$ as the inverse of M. More is true: The operation of cartesian product of manifolds passes to cobordism classes, giving $\Omega_* = \bigoplus_n \Omega_n$ the structure of a graded ring, the *oriented cobordism ring*.

BASIC PROBLEM: Compute the cobordism ring.

Thom approached this problem as follows. Recall that oriented k-dimensional vector bundles are *classified* by maps to a space $BSO(k)$, which one can think of as the Grassmannian of oriented k-planes in 'infinite dimensional Euclidean space'. (More formally, $BSO(k)$ is defined as a direct limit of finite-dimensional Grassmannians.) This means that given such a vector bundle over a space X, there is a unique homotopy class of maps $X \to BSO(k)$ that pulls back the universal bundle over $BSO(k)$ to the given vector bundle over X.

Suppose now we are given a map f from a sphere S^{n+k} to the Thom space $MSO(k)$ of the universal bundle over $BSO(k)$. (Notice that $BSO(k)$ is not compact. The Thom space in question is obtained by collapsing the outside of the unit disc subbundle of $ESO(k)$ to a point. Because the sphere is compact, we may as well think of a map from S^{n+k} to the Thom space of the universal bundle over some Grassmannian $G(k, N)$, where N is large.) By a small perturbation, we can make this map transverse at the zero-section; and $f^{-1}(BSO(k))$ then becomes a manifold M of dimension n.

Suppose now we change f by a homotopy to a new map g. The homotopy can itself be regarded as a map $h\colon S^{n+k} \times [0,1] \to MSO(k)$, and this map can itself be made transverse at the zero-section; so that $h^{-1}(BSO(k))$ becomes a manifold with boundary, which gives a cobordism between $f^{-1}(BSO(k)$ and $g^{-1}(BSO(k)$. In other words we have defined a map (easily seen to be a homomorphism)

$$\pi_{n+k}(MSO(k)) \to \Omega_n.$$

Moreover, by Whitney's embedding theorem, this map is surjective as soon as $k > n$. For any compact manifold M^n can be embedded in S^{2n+1}, and there M has a tubular neighbourhood which is diffeomorphic to the total space of an $(n+1)$-dimensional vector bundle ν. By collapsing all of S^{2n+1} outside the tubular neighbourhood we get a map $S^{2n+1} \to T(\nu)$, and then by composing with the classifying map for ν we get a map $S^{2n+1} \to MSO(n+1)$, which is already transverse at the zero-section and such that the inverse image of the zero-section is M.

A refinement of this argument (embedding a cobordism rel boundary) proves injectivity if k is a little larger. Thus we obtain

3.16. THEOREM. *(Thom) There is a canonical isomorphism*

$$\lim_{k\to\infty} \pi_{n+k}(MSO(k)) \to \Omega_n.$$

The limit makes sense because there is a canonical stabilisation map from the suspension $S(MSO(k))$ to $MSO(k+1)$, defined as follows. Consider the bundle V over $BSO(k)$ which is the direct sum of the universal k-plane bundle and a trivial line bundle. Note that the Thom space of V is $S(MSO(k))$. But V must be classified by a map $BSO(k) \to BSO(k+1)$, and passing to the Thom space we get the desired stabilisation map.

We can use this to compute the cobordism ring modulo torsion. This needs the theory of Pontrjagin numbers. Let $k = (k_1, \ldots, k_r)$ be a *partition* of n (that is, a list of nonnegative integers adding up to N). For an oriented $4n$-manifold M, the *Pontrjagin number* $p_k[M]$ corresponding to the partition **k** is the number

$$\langle p_{k_1}(TM)\dots p_{k_r}(TM),[M]\rangle$$

where on the left we have the Pontrjagin classes of the tangent bundle of M.

3.17. LEMMA. *Pontrjagin numbers are cobordism invariants.*

PROOF. Suppose $M = \partial W$. The Pontrjagin classes of TM are the same as the Pontrjagin classes of TW restricted to M; indeed, these two bundles differ only by a trivial 1-dimensional bundle. Let $i\colon M \to W$ be the inclusion. Then we have (denoting by p the relevant product of Pontrjagin classes)

$$\langle p(TM),[M]\rangle = \langle i^*p(TW),[M]\rangle = \langle p(TW), i_*[M]\rangle.$$

But $i_*[M] = 0$, since $[M]$ is the boundary of the orientation class $[W] \in H_{4n+1}(W, \partial W)$, so the Pontrjagin numbers are zero. □

The Pontrjagin numbers therefore give homomorphisms $\Omega_{4n} \to \mathbb{Z}$. In particular they give $\mathbb{Q}$-linear maps $\Omega_{4n}\otimes\mathbb{Q} \to \mathbb{Q}$.

Given a partition **k** of n, let $\mathbb{P}^k$ denote the product $\mathbb{CP}^{2k_1}\times\cdots\times\mathbb{CP}^{2k_r}$ of complex projective spaces, which is a $4n$-dimensional manifold. Let $\varphi(n)$ denote the number of partitions of n. Then we have

3.18. LEMMA. *For any n, the $\varphi(n)\times\varphi(n)$ matrix whose entries are the Pontrjagin numbers $p_j[\mathbb{P}^k]$ has nonzero determinant.*

PROOF. A computation with symmetric functions. See Chapter 16 of Milnor and Stasheff. □

As a corollary, the manifolds $\mathbb{P}^k$ are linearly independent elements of $\Omega_{4n}\otimes\mathbb{Q}$, and the dimension of this vector space is therefore at least $\varphi(n)$. In fact we have

TCT 3.19. THEOREM. *(Thom) The rational cobordism algebra $\Omega_*\otimes\mathbb{Q}$ is a polynomial algebra on the complex projective spaces $\mathbb{CP}^2, \mathbb{CP}^4, \dots$. In particular, the dimension of $\Omega_{4n}\otimes\mathbb{Q}$ is exactly $\varphi(n)$.*

EXERCISE: It is a consequence of the theorem that if M^r is a manifold and r is not a multiple of 4, then some finite disjoint union of copies of M is a boundary. Try to see this directly in some examples. For instance, what happened in the case of $\mathbb{CP}^m$, m odd? Hint: You should be able to represent $\mathbb{CP}^m$ in this case as the total space of a circle bundle over a quaternionic projective space.

PROOF. Given the linear independence lemma above, it is plain that all we need to do is to find an upper bound for the dimension $\dim_{\mathbb{Q}}\Omega_n\otimes\mathbb{Q}$; the upper bound should be $\varphi(m)$ if $n = 4m$ is a multiple of 4, and 0 otherwise.

We start by noting that by the Pontrjagin-Thom construction we can identify

$$\Omega_n\otimes\mathbb{Q} \simeq \pi_{n+k}(MSO(k))\otimes\mathbb{Q}$$

for k large. The space $MSO(k)$ is highly connected (in fact it is $(k-1)$-connected), and so we can apply the Hurewicz theorem tensored with[1] $\mathbb{Q}$; this theorem says that if

[1]This is a theorem of Serre. The main ingredient in the proof is the computation of the homotopy groups of spheres modulo torsion, which may be found for instance in Spanier, Chapter 9 Section 7. The computation is that $\pi_r(S^n)$ is a finite group for $r \neq n, 2n-1$, and this verifies that the theorem is true for a sphere. One then extends to prove the theorem for a bouquet of spheres, and then for an arbitrary finite complex X by considering a map $S^{r_1}\vee\cdots\vee S^{r_p}\to X$ obtained by combining the generators of the torsion-free parts of all the homotopy groups up to dimension $2k-1$. See Milnor and Stasheff, theorem 18.3.

X is a $(k-1)$-connected space then the Hurewicz map $\pi_r(X)\otimes\mathbb{Q}\to H_r(X;\mathbb{Q})$ is an isomorphism for $r<2k-1$. Now by the Thom isomorphism, $\tilde{H}_r(MSO(k);\mathbb{Q})\cong H_{r-k}(BSO(k);\mathbb{Q})$. Thus we find an isomorphism

$$\Omega_n\otimes\mathbb{Q}\to H_n(BSO(k);\mathbb{Q})$$

for k large.

Now the rational cohomology of the classifying space $BSO=\lim_k BSO(k)$ is well known; it is a polynomial algebra generated by the Pontrjagin classes. It follows that $\dim_{\mathbb{Q}} H_n(BSO(k);\mathbb{Q})$ is equal to $\varphi(n/4)$ if n is a multiple of 4, and zero otherwise. The proof is completed. □

DIGRESSION: The computation of a limit such as $lim_{k\to\infty}\pi_{n+k}MSO(k)$, as appears in Thom's calculation of the cobordism groups, is a problem of *stable homotopy theory*. "We say that some phenomenon is *stable*, if it can occur in any dimension, or in any sufficiently large dimension, and if it occurs in essentially the same way independent of dimension, provided perhaps that the dimension is sufficiently large." (J.F. Adams) The key notion in stable homotopy theory is that of a *spectrum*. A spectrum $\mathbb{X}$ is a sequence X_k of spaces with basepoint, equipped with structure maps $SX_k\to X_{k+1}$, where S denotes the suspension. As an elementary example one has the *suspension spectrum* $\mathbb{S}X$ of a space X; this is the spectrum whose k'th term is just the k'th suspension S^kX. More relevantly, we have the *Thom spectrum* $\mathbb{M}SO$ whose k'th term is $MSO(k)$, and whose structure maps are just the stabilisation maps we defined above.

A spectrum has (stable) homotopy groups, defined by $\pi_n\mathbb{X}=\lim_{r\to\infty}\pi_{n+r}X_r$, so that Thom's theorem says $\Omega_*=\pi_*\mathbb{M}SO$ in this language. There is an extensive theory of spectra starting from these basic ideas, and at certain points we may require to make use of it. A good reference is Adams' book.

5. The Hirzebruch Signature Theorem

We stated the signature theorem somewhat loosely in Chapter 1. Now we will give a more precise statement, and an outline of the proof.

We need the notion of a *multiplicative sequence*, due to Hirzebruch. This is a sequence of polynomials $K_0=1, K_1(p_1), K_2(p_1,p_2),\ldots$ and so on in the universal Pontrjagin classes, with $K_n\in H^{4n}(\cdot;\mathbb{Q})$, such that the *total K-genus* $K(V)=1+K_1(V)+K_2(V)+\cdots$ of a vector bundle V is multiplicative: $K(V_1\oplus V_2)=K(V_1)K(V_2)$. For example, the sequence of polynomials $K_n=p_n$ is multiplicative (this is just the Whitney sum formula.)

Recall the *splitting principle* from the theory of characteristic classes. The splitting principle tells us that given any reasonable space X (say a finite complex) and (real) vector-bundle V over X, we can find a map $f\colon Y\to X$ such that the induced map f^* on cohomology is injective and the pulled-back bundle f^*V splits as a direct sum of 2-plane bundles, together (possibly) with a line bundle. It follows that any multiplicative sequence K_n is determined uniquely by its value on 2-plane bundles, which is a formal power series $f(t)$ in the first Pontrjagin class. (The coefficients of $f(t)$ are just the coefficients of p_1^n in $K_n(p_1,\ldots,p_n)$.) Conversely, Hirzebruch showed that every formal power series with leading coefficient 1 determines uniquely a multiplicative sequence of polynomials, called the multiplicative sequence *belonging* to the given formal power series. The proof is a computation with symmetric functions: write formally

$$1+p_1t+p_2t^2+\cdots=(1+u_1t)(1+u_2t)\cdots,$$

where the formal variables u_i may be identified with the first Pontrjagin classes of the splitting 2-plane bundles. Then we must have

$$1+K_1t+K_2t^2+\cdots=f(1+u_1t)f(1+u_2t)\cdots,$$

so to find K_1, K_2 and so on we just expand the right-hand side as a power series in t whose coefficients are symmetric functions in the u_i, and then write these coefficients in terms of the elementary symmetric functions p_i.

3.20. EXAMPLE. The multiplicative sequence of polynomials belonging to the formal power series

$$f(t) = (1+t)^{-1} = 1 - t + t^2 - \cdots$$

expresses the dual Pontrjagin classes $\bar{p}_n$ (which we met in Chapter 1; they are the Pontrjagin classes of a stable inverse) in terms of the ordinary Pontrjagin classes.

3.21. THEOREM. *(Hirzebruch Signature Theorem) Let L_n be the multiplicative sequence of polynomials in the Pontrjagin classes belonging to the formal power series*

$$\frac{\sqrt{t}}{\tanh\sqrt{t}} = 1 + \frac{1}{3}t - \frac{1}{45}t^2 + \cdots .$$

Then for any compact oriented $4n$-manifold M we have

$$\operatorname{Sign} M = \langle L_n(p_1, \ldots, p_n), [M] \rangle$$

where the p_i are the Pontrjagin classes of the tangent bundle of M.

PROOF. (Sketch) Both sides of the equation define ring homomorphisms $\Omega_* \otimes \mathbb{Q} \to \mathbb{Q}$. For the left side this is obvious from the definition of a multiplicative sequence and the cobordism invariance of the Pontrjagin numbers. For the right side, we proved the cobordism invariance of the signature in 2.48 as a consequence of Poincaré duality for manifolds with boundary; the multiplicative property can similarly be proved using the Künneth theorem.

Since both sides of the Hirzebruch signature formula define ring homomorphisms from $\Omega_* \otimes \mathbb{Q}$, it suffices to check the theorem on a set of generators for this ring. By Thom's theorem 3.19, such a set of generators is provided by the even-dimensional complex projective spaces $\mathbb{CP}^{2k}$. These all have signature $+1$. On the other hand, the total Pontrjagin class of $\mathbb{CP}^{2k}$ is equal to $(1+a^2)^{2k+1}$, where $a \in H^2(\mathbb{CP}^{2k}; \mathbb{Q})$ is the canonical generator (the hyperplane class). Thus

$$L(p) = (a/\tanh a)^{2k+1}$$

and a direct calculation shows that the coefficient of a^{2k} in this power series is equal to 1. Thus the theorem is verified for the generators, hence it is true. □

This is Hirzebruch's proof, as given in his *New Topological Methods in Algebraic Geometry*. See also Milnor and Stasheff.

CHAPTER 4

Morse Theory and Handle Theory

In this chapter we will show how a Morse function allows us to build up a smooth manifold from a small number of elementary pieces. In particular, we will prove that any manifold is homotopy equivalent to a cell complex. The presentation we use will be very convenient for calculation. We will prove Smale's h-cobordism theorem and draw the standard consequences. Then we will go on to consider the s-cobordism theorem. We will meet an obstruction defined in terms of algebraic K-theory associated to the fundamental group. Besides its own significance, this obstruction serves as a model for the more complicated issues that arise when dealing with the surgery obstruction groups.

From the point of view of surgery theory itself, the importance of these results lies in questions about *how many* manifold structures there can be on a given Poincaré duality space X. Suppose that $M \to X$ and $M' \to X$ are two structures on X, and we want to know whether they are equivalent in $\mathcal{S}(X)$, that is, whether there is a diffeomorphism $M \to M'$ making the obvious diagram commute. We can first ask whether the structures are normally cobordant: that is, whether there is a normal map of a manifold W with boundary to $X \times [0,1]$, restricting on the boundary to the given structures. If the structures are equivalent, then certainly they are normally cobordant: the mapping cylinder of a diffeomorphism $M \to M'$ provides the necessary cobordism. If the structures are normally cobordant, we can apply the surgery technique to the cobordism W itself (leaving the boundary fixed). If surgery is successful it will produce a new normal cobordism between M and M' which is an h-*cobordism* in the sense of the following definition.

4.1. DEFINITION. An h-*cobordism* is a cobordism[1] $(W; \partial_- W, \partial_+ W)$ such that the inclusions $\partial_\pm W \to W$ are homotopy equivalences.

The notion of h-cobordism was introduced by Thom (circa 1957) into the study of exotic spheres, as a substitute for the relation of diffeomorphism, which seemed at that time to be inaccessible to algebraic study. If that were still the case, surgery theory would still work, but it would produce a classification of manifolds up to h-cobordism only[2]. However, around 1961 Smale proved the h-*cobordism theorem*, which states that any simply-connected h-cobordism has a product structure, and consequently that h-cobordant simply-connected manifolds are diffeomorphic. Thus the h-cobordism and the diffeomorphism classifications of simply-connected manifolds are the same. For general manifolds there is a difference between the classifications, but it is measured by an algebraic K-theory invariant (Whitehead torsion), and this invariant vanishes in several interesting cases (for example when the fundamental group is free abelian.)

---

[1]We will frequently use this notation for a cobordism. It means that W has a boundary which is a disjoint union of two pieces $\partial_- W$ and $\partial_+ W$, which are oriented respectively with the induced orientation from W and the opposite of that orientation.

[2]Something like this still *is* the case for the exotic homology manifolds of Bryant, Ferry, Mio and Weinberger.

Let us begin by recalling the notion of surgery. Let N^n be a manifold and suppose that N contains an embedded sphere S^q with trivial normal bundle. Then we get an embedding $D^p \times S^q \subset N$ as a tubular neighbourhood, and *surgery* consists of removing this $D^p \times S^q$ and gluing in $S^{p-1} \times D^{q+1}$ (which has the same boundary) instead.

EXERCISE: Worry about whether this operation is well-defined. A careful treatment is in Kosinski, *Differential Manifolds*.

In chapter 1 we remarked that, if N' is obtained from N by a surgery, then the manifolds N and N' will be cobordant. The construction which demonstrates this also has a number of other applications: it is called *handle attachment*. In order to define this in the smooth category, however, we need to digress briefly and discuss the notion of a *manifold with corners*.

A *manifold with corners* is a space locally modelled on an open subset of $(\mathbb{R}^+)^n$; the *corner set* is the set of points where two or more coordinates are zero in the local model. To be more precise, M has 'boundaries' $\partial_S M$, possibly empty, for each subset S of $\{1, \ldots, n\}$, where $\partial_S M$ is modelled locally by $\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_i = 0 \Leftarrow i \in S\}$. The transition functions are of course required to be smooth. Apart from their use in defining handle attachment, manifolds with corners will also inevitably have to be considered when we deal, as we eventually will have to, with cobordisms of manifolds with boundary.

Corners are a nuisance of the smooth category: in TOP or PL, $(\mathbb{R}^+)^n$ is homeomorphic to $\mathbb{R}^{n-1} \times \mathbb{R}^+$, so that manifolds with corners are the same thing as manifolds with boundary. We won't have occasion to consider manifolds with worse than second-order corners (I hope), so that $\partial_S M = \emptyset$ whenever $|S| \geqslant 3$. The corner set of such a manifold is then itself a closed manifold, and it has a tubular neighbourhood which is fibered (trivially) by quarter-spaces. We may turn such a manifold into an ordinary manifold with boundary by excising the tubular neighbourhood of the corners, doubling all the angles (thus turning the quarter-space into a half-space), and reattaching the resulting bundle of half-spaces. This process is called *unbending the corners*. Conversely, suppose that we are given an ordinary manifold M with boundary, and a codimension zero submanifold X (with boundary) of the closed manifold ∂M. Then we may *bend* the manifold M along M, obtaining a manifold M_c with corners such that $\partial_1 M_c = X$ and $\partial_2 M_c = \partial M \setminus X^\circ$.

Given two manifolds M and N with corners, such that there is an identification $\partial_1 M \cong \partial_1 N$, we may glue M to N by this identification. The resulting space is a manifold with boundary. This is just the 'cornered' version of the operation of joining two manifolds with boundary along their boundary to obtain a closed manifold. Indeed, this joining is exactly what happens to $\partial_2 M$ and $\partial_2 N$ in the example above.

Suppose now that W is a manifold with boundary, and that ∂W contains an embedded sphere S^q with trivial normal bundle. Let $X = D^p \times S^q$ be a tubular neighbourhood of S^q in ∂W; it is a codimension-zero submanifold with boundary. Bend along ∂X, as above, to obtain a manifold W with corners, having $\partial_1 W = D^p \times S^q$. Now consider the product $H = D^p \times D^{q+1}$; it is in a natural way a manifold with corners, having $\partial_1 H = D^p \times S^q$ and $\partial_2 H = S^{p-1} \times D^{q+1}$. Glue H to W_c along ∂_1, as above, to obtain a new manifold with boundary $W' = W \cup_X H$.

4.2. DEFINITION. The manifold W' so obtained is said to be gotten from W by *attaching a $q+1$-handle to $S^q \subseteq \partial W$.*

As a matter of jargon, we say that W is obtained by *attaching a $(q+1)$-handle* to $N \times [0,1]$ along the *attaching sphere* S^q in the boundary. The disc D^p is called the *belt disc* of the handle. Dually, S^{p-1} is the *belt sphere* and D^{q+1} the *attaching disc* (aka the *core disc* in some books).

The following observation is immediate:

4.3. LEMMA. *If W' is obtained from W by attaching a handle to the sphere $S^q \subseteq \partial W$, then $\partial W'$ is obtained from ∂W by performing surgery on this same sphere.*

Consequently, we have

4.4. PROPOSITION. *If N' is obtained from N by performing surgeries, then N' is cobordant to N.*

FIGURE 1. Attaching a handle to the boundary of a manifold

PROOF. It suffices to consider a single surgery; and then the manifold obtained by attaching a handle to the right-hand boundary of $N \times [0,1]$ provides a cobordism between N and N', by the previous proposition. $\square$

A cobordism obtained by attaching a single handle to a product, as in this proof, is called an *elementary cobordism*; it is also sometimes called the *trace* of the surgery that was done on N.

The advantage of the handle attachment procedure is that it makes new *manifolds* out of old ones. If however we are interested only in the homotopy type for some reason, and not in the manifold structure, then we can think on a much cruder level.

4.5. PROPOSITION. *Let W be a compact manifold with boundary, and let $S^q \subset \partial W$ be a q-sphere in the boundary with trivial normal bundle. Then attaching a $(q+1)$-handle to S^q has the same effect on homotopy type as attaching a $(q+1)$-cell D^{q+1} via the attaching map.*

PROOF. Up to homotopy equivalence we may contract the belt disc D^p to a point. $\square$

Morse theory provides a means of generating elementary cobordisms.

4.6. DEFINITION. Let $f\colon M \to \mathbb{R}$ be a smooth function on a manifold M. It is called a *Morse function* if the differential $df\colon M \to T^*M$ is transverse to the zero-section of T^*M.

4.7. PROPOSITION. *Morse functions are dense. More precisely, suppose that M is embedded in some $\mathbb{R}^k$. Then any smooth function f on M can be perturbed by an (arbitrarily) small linear function on $\mathbb{R}^k$, so as to make it a Morse function.*

PROOF. Look first on a coordinate patch, where T^*M is trivial. By Sard's theorem, we may perturb df by an arbitrarily small constant in order to make it transverse to the zero-section on this patch. Since a constant is the differential of a linear map, this means that we can perturb f by a small linear map to make df transverse on the given patch. Now, noting that the Morse condition is an open one, we may apply our usual local-to-global argument to get the result. $\square$

An equivalent definition of a Morse function is this: at each critical point of f (that is, point with $df = 0$), the *Hessian* — the symmetric matrix (in local coordinates) of second

FIGURE 2. The region between the level sets of a standard Morse function is a handle.

derivatives of f — should be nonsingular. By definition, the *index* of the critical point is the number of negative eigenvalues of the Hessian there. One can easily check that this does not depend on the choice of local coordinates.

The basic result of Morse theory is contained in the next proposition.

4.8. PROPOSITION. *Let W be a cobordism, with boundary $\partial W = \partial_- W \sqcup \partial_+ W$. Suppose that W admits a Morse function f with no critical values on the boundary. Then*

(i) *If f has no critical values on the interior of W, then W is a product;*
(ii) *If f has exactly one critical value, of index r say, then W is an elementary cobordism, obtained by attaching an r-handle to $\partial_- W \times [0,1]$.*

PROOF. (Sketch) First consider the case in which there are no critical points. Equip W with a Riemannian metric, which allows us to define the *gradient vector field* ∇f of f as the dual to df. The flow lines of this vector field foliate W, and they always run in the direction of decreasing f; so they give W a product structure.

Now consider the case of just one critical point. The first result that is needed is the *Morse lemma*: this says that in a neighbourhood of a critical point one can choose local coordinates so that the Morse function f is just a quadratic form

$$f(x_1, \dots, x_n) = \pm x_1^2 \pm x_2^2 \pm \dots \pm x_n^2$$

where the first r signs (r being the index) are negative and the last $n - r$ are positive. Using the first result we can localize matters to a neighbourhood of the critical point; we then just need to observe that if f is the quadratic form given above, the region $\{x \in \mathbb{R}^n : -1 \leqslant f(x) \leqslant 1\}$ is naturally diffeomorphic to $D^r \times D^{n-r}$ minus the corner set. (See the figure) $\square$

If M is a manifold, $f\colon M \to \mathbb{R}$ a Morse function, then it is quite easy to adjust f so that all the different critical points of f have different critical values. Consequently, we may pick a sequence $a_0 < a_1 < \dots$ of regular values for f such that there is exactly one critical value of f between a_i and a_{i+1} for each i. By the above result, $f^{-1}([a_i, a_{i+1}])$ is an elementary cobordism. Consequently, we obtain.

4.9. PROPOSITION. *Any (compact) manifold can be built up by successively attaching handles to the empty set. Any cobordism between manifolds can similarly be built up by* *successive attachment of handles. The number of r-handles equals the number of critical points of index r of the Morse function used to construct the handle decomposition.*

4.10. COROLLARY. *A compact manifold has the homotopy type of a finite CW-complex.*

For we saw that a handle attachment is the same thing, up to homotopy, as the attachment of a cell; and a CW-complex just is a space built up by attaching cells to the empty set.

Let us recall now the notion of *cellular homology* for CW-complexes. Suppose that X is such a complex, and recall the notion X^k for its k-skeleton. Choose (arbitrarily) an orientation for each cell. Using this orientation, we may identify the relative homology group $H_k(X^k, X^{k-1}; \mathbb{Z})$ with the free abelian group $C_k(X)$ generated by the k-cells of X.

We can define a boundary map $\partial\colon C_k(X) \to C_{k-1}(X)$ by composing pieces of the exact sequences for the pairs (X^k, X^{k-1}) and (X^{k-1}, X^{k-2}): it is the composite

$$H_k(X^k, X^{k-1}) \to H_{k-1}(X^{k-1}) \to H_{k-1}(X^{k-1}, X^{k-2}).$$

Notice that $\partial^2 = 0$ (because we compose two successive maps in the same exact sequence). Thus we have a complex.

4.11. PROPOSITION. *The homology of the complex $(C_*(X), \partial)$ is canonically isomorphic to the ordinary homology of X. There is a similar result for cohomology using the dual complex $C^*(X) = \mathrm{Hom}(C_*(X), \mathbb{Z})$.*

PROOF. Note that $H_r(X^k, X^{k-1}) = 0$ if $r \neq k$, and consider the spectral sequence associated to the filtration of X by skeleta. □

The advantage of cellular homology is the very geometric nature of the boundary map. Indeed, the map $\partial\colon C_k \to C_{k-1}$ is an integer matrix whose (ij) matrix entry is the degree of the map $S^{k-1} \to X^{k-1} \to S^{k-1}$, where the first map is the attaching map for the i'th k-cell and the second is the map which collapses the complement of the j'th $(k-1)$-cell to a point. More geometrically still, this matrix entry is the generic number of inverse images, under the attaching map f_i for the boundary sphere of the i'th k-cell, of a point in the interior of the j'th $(k-1)$-cell.

4.12. EXAMPLE. Suppose X is a finite ordered simplicial complex. Then it has a CW structure with simplices for cells. The complex $C_*(C)$ associated to this CW structure is the usual simplicial chain complex of X.

Now let us apply this to a manifold M^n with the CW structure determined by a handle decomposition. We choose (arbitrarily) an orientation for each handle, by which we mean an orientation for its attaching disc. Moreover, we will assume that the handle decomposition is *ordered* in the following sense: there is a sequence $\emptyset = M_{-1}, M_0, \dots, M_n = M$ of codimension zero submanifolds with boundary such that M_k is obtained by attaching k-handles to the boundary of M_{k-1} (so that M_k corresponds to the k-skeleton in the CW structure). It can be shown that it is always possible to attach handles in increasing order in this way[3]. Now the attaching sphere α of a k-handle is a submanifold of ∂M_{k-1}, and the belt sphere β of a $(k-1)$-handle is also a submanifold of ∂M_{k-1}; and their dimensions

rearrange

[3]This is a simple application of transversality. Suppose, for example, that we attach a 2-handle after a 3-handle. We can make the attaching sphere of the 2-handle transverse to the belt sphere of the 3-handle, and, counting dimensions, this means that they don't meet at all. So we can slide this attaching sphere right off the 3-handle, which means that we might as well have attached the 2-handle first.

add up to $\dim \partial M_{k-1} = n-1$, so that generically they intersect in a finite set of points. Furthermore, α is oriented (by the boundary of the orientation chosen for the attaching disc of the k-handle), and the *normal bundle* to β is oriented (because its fibre is just the attaching disc of the $(k-1)$-handle). In these circumstances we can attach a sign to each intersection point, so we can define an intersection number $[\alpha : \beta] \in \mathbb{Z}$.

HSC

4.13. PROPOSITION. *Let M be a manifold with an ordered handle decomposition, as above. Then the homology of M is computed by the following complex: the chain group C_k is the free abelian group on the k-handles, and the boundary map $C_k \to C_{k-1}$ is the matrix $\mathfrak{M}_k$ whose (ij)'th entry is equal to the intersection number $[\alpha_i : \beta_j]$ of the attaching sphere of the i'th k-handle with the belt sphere of the j'th $(k-1)$-handle.*

PROOF. Let $H = D^{n-k+1} \times D^{k-1}$ be the j'th $(k-1)$-handle, with $\partial_1 H = \beta_j \times D^{k-1}$ contained in ∂M_{k-1} and the remainder of the boundary contained in ∂M_{k-2}. According to our description above of the matrix entries of the boundary map, the matrix entry we want is just the generic number of preimages of some point p in the interior of D^{k-1} under the map $\alpha_i \to D^{k-1}$ which shrinks D^{n-k+1} to a point. But plainly this is just the intersection number of α_i with the belt sphere $S^{n-k+2} \times \{p\}$. $\square$

This result is due to Smale. The intersection numbers can also be described in terms of Morse theory as the numbers of flow lines of the gradient flow between a critical point of index k and one of index $k-1$. In this form, the result was rediscovered by Witten in the early eighties, with a highly original proof based on ideas of quantum-mechanical tunnelling. Witten's work set off an explosion of interest in the relationship between analysis and Morse theory.

A handle has a certain symmetry which we have not exploited so far. In fact, suppose given an ordered handle decomposition and corresponding filtration of M, as above. Let $\bar{M}_k = M \setminus M_{n-k+1}$. Then $\bar{M}_{k+1}$ is obtained from $\bar{M}_k$ by attaching k-handles; these are "the same" handles as the $(n-k)$-handles in the original decomposition, except that the rôles of the attaching and belt discs have been interchanged. We call this the dual handle decomposition of M (thinking of the Morse function as the "height" above some plane, what we have done is to turn M upside-down). Suppose now that M itself is oriented. Then an orientation of the belt disc of a handle determines an orientation of its attaching disc, by the requirement that the unique intersection point should have positive sign. Hence, an orientation for a handle decomposition determines an orientation for the dual handle decomposition. Plainly the matrices of intersection numbers for the dual handle decomposition are (up to some signs depending only on the dimension) just the transposes of the corresponding matrices for the original handle decomposition; or, to put it more canonically, there is an isomorphism of chain complexes

$$\bar{C}_*(M) \cong \operatorname{Hom}(C_{n-*}(M), \mathbb{Z}) = C^{n-*}(M).$$

Since both handle decompositions compute the (co)homology of M, we find another proof of Poincaré duality from this: the homology of the oriented manifold M is isomorphic to the cohomology in the complementary dimension.

REMARK: The *Morse inequalities* are another well-known consequence of 4.13, whose proof does not in fact require the explicit description of the boundary map. Let M^n be a manifold equipped with a Morse function f. Let b_i be the i'th Betti number of M, and let c_i be the number of critical points of f having index i. Then the Morse inequalities are

$$\begin{array}{rcl} b_0 & \leqslant & c_0 \\ b_1 - b_0 & \leqslant & c_1 - c_0 \\ b_2 - b_1 + b_0 & \leqslant & c_2 - c_1 + c_0 \\ & \cdots & \\ b_n - b_{n-1} + \cdots \pm b_0 & = & c_n - c_{n-1} + \cdots \pm c_0 \end{array}$$

To see this one simply notes that by 4.13 the c_i are the dimensions of the chain spaces in a complex that computes the rational homology. Elementary linear algebra now completes the proof.

1. Handle Calculus

The results above show that complete homological information about the structure of a manifold or a cobordism is contained in the matrices $\mathfrak{M}_k$ of intersection numbers between the attaching spheres of k-handles and the belt spheres of $(k-1)$-handles. *Handle calculus* shows that certain algebraic operations on these matrices can be 'performed geometrically', by changing the handle decomposition of the given manifold. In particular, if the matrices can be made trivial by suitable algebraic operations, the handle structure can be made trivial, and so the original cobordism will have a simple structure. One should not here the analogy with the Whitney lemma, which also shows that under certain circumstances algebra (intersection numbers, in this case) faithfully reflects geometry.

The following result, called the Cancellation Lemma, is due to Smale.

4.14. PROPOSITION. *Suppose that the manifold W' is obtained from W^n by adding successively a q-handle and a $(q+1)$-handle to the boundary, and suppose that the attaching sphere of the $(q+1)$-handle intersects the belt sphere of the q-handle transversely in a single point. Then W' is diffeomorphic to W.*

The application to intersection matrices is

4.15. COROLLARY. *Suppose that the simply-connected manifold W^n has a handle presentation in which one of the intersection matrices $\mathfrak{M}_k$, $4 \leqslant k \leqslant n-3$, has i'th row and j'th column intersecting in an entry ± 1 and having zeroes elsewhere. Then one can remove the corresponding k-handle and $(k-1)$-handle from the presentation without affecting the intersection numbers between other handles.*

To prove the corollary one uses the Whitney trick to ensure that all the algebraic intersection numbers are actually equal (in absolute value) to the number of geometric intersection points, and then uses Smale's lemma to cancel the two handles; because all the other intersection numbers are assumed to be zero, this operation can be carried out away from all the other handles so it doesn't affect them.

PROOF. (OF SMALE'S LEMMA) Let's begin by thinking what the region in ∂W must look like along which the two handles are attached. First, the q-handle is attached along $S^{q-1} \times D^{n-q}$. Then the q-handle is attached. Because of the transversality assumption, we may assume that its attaching sphere S^q is made up by joining two discs D^q, one of which is a disc $D^q \times \{p\}$, $p \in S^{n-q-1}$, in the other boundary of the q-handle, and the other is a disc D^q contained in ∂W and spanning $S^{q-1} \times \{p\}$. The whole attaching region in ∂W therefore looks like $S^{q-1} \times D^{n-q} \cup_{S^{q-1} \times D^{n-q-1}} D^q \times D^{n-q-1}$, and this is in fact naturally identified (as a manifold with corners) with $D^q \times D^{n-q-1}$. See the first figure.

Now we see after a little thought that the two attached handles together must make up the cell $D^1 \times D^q \times D^{n-q-1}$, attached along its face $\{0\} \times D^q \times D^{n-q-1}$. See the second figure. In this figure, the front face of the cube is the attaching region in ∂W. The light grey is the q-handle; it runs 'round the back' (invisibly in the figure) to form a neighbourhood of the two sides and the back of the cube. The dark grey is the $(q+1)$-handle.

Thus the effect of attaching two handles to W is simply to attach a disc along one face in its boundary. But plainly this produces a manifold diffeomorphic to W. (We leave corner-smoothing questions to the reader.) $\square$

FIGURE 3. The attaching region in the boundary

FIGURE 4. Two handles attached successively

We will now use Smale's lemma to start getting rid of handles. Our first victims are the 0-handles.

4.16. LEMMA. *Let W be a connected cobordism with $\partial_- W$ nonempty. Then W has a presentation without 0-handles. If $\partial_- W$ is empty, it has a presentation with just one 0-handle.*

PROOF. The second statement follows from the first (just punch a hole). To prove the first statement, consider a presentation with the minimal number of 0-handles. Since W is connected, W_1 is connected; so, if h^0 is a 0-handle, there is a 1-handle k^1 connecting it to somewhere else. But then its attaching sphere is an S^0 with one point on h^0 and one point somewhere else; so by Smale h^0 and k^1 can be cancelled. $\square$

The 1-handles are more recalcitrant. We have to pay a price in 3-handles for their elimination.

4.17. LEMMA. *Let W^n be a connected cobordism, with $\partial_- W \to W$ a 1-connected map and $n \geqslant 5$. Suppose that W is equipped with a presentation without 0-handles. Then there is a new presentation of W without 0-handles or 1-handles, and having the same handles as before in dimensions $\geqslant 4$.*

PROOF. (Outline) First we eliminate the 0-handles using the previous lemma. Then W_0 is connected. Consider now a presentation with the minimal number of 1-handles; for simplicity suppose that there is just one 1-handle h^1 attached to $\partial_+ W_0$. Let Γ_1 be a simple closed curve in $\partial_+ W_1$ which intersects the belt sphere of h^1 transversely in one point and returns through $\partial_+ W_0$. We may assume (by transversality) that Γ_1 is disjoint from the attaching spheres of all 2-handles, so it in fact lies in $\partial_+ W_2$. Now Γ_1 must be nullhomotopic in W (since it clearly doesn't come from $\partial_- W$); since W is obtained by adding handles of indices 3 and up to $\partial_+ W_2$, this means that Γ_1 is already nullhomotopic in $\partial_+ W_2$.

Now we attach to W_1 a cancelling 2-handle and 3-handle, $k^2 \cup l^3$, attached somewhere out of the way of everything else. This doesn't change W_1. The attaching sphere of k^2 is another null-homotopic simple closed curve Γ_2 in $\partial_+ W_2$, by the same argument as before.

Now we need an 'unknotting theorem': Two null-homotopic simple closed curves in a k-manifold, $k \geqslant 4$, are in fact isotopic there. At least for $k \geqslant 5$ this can easily be proved by using Whitney's embedding theorem to embed a cylindrical homotopy between the two curves, and then manufacturing an isotopy by 'pushing' along the cylinder. Thus Γ_1 and Γ_2 are isotopic, and we can use the isotopy to move our trivial handle pair so that k^2 actually is attached along Γ_1.

But now $h^1 \cup k^2$ is trivial, by the cancellation lemma. Thus

$$W_2 = W_0 \cup h^1 \cup k^2 \cup l^3 \cup \text{2-handles} = W_0 \cup l^3 \cup \text{2-handles}$$

and we have eliminated the 1-handle and acquired instead an extra 3-handle. □

This process is called *handle trading*. The two lemmas above are the first of a series of handle-trading lemmas, which culminate in the following

4.18. PROPOSITION. *Let W^n be a cobordism with $(W, \partial_- W)$ k-connected, $k \leqslant n-4$. Given any presentation of W, it can be modified to produce a new presentation without handles of dimension $\leqslant k$, and with the number of handles of dimensions $\geqslant k+3$ unchanged.*

Now remember that any cobordism can be given a dual presentation, in which the k-handles are the $(n-k)$-handles of the original presentation. This means that we can trade down as well as up (if things are connected enough). After sufficient trading we arrive at the following situation.

4.19. PROPOSITION. *Let W^n be an h-cobordism, $n \geqslant 5$. Then W has a presentation with only k-handles and $(k+1)$-handles; here k can be any integer between 2 and $n-3$.*

We now proceed towards the proof of the h-cobordism theorem. Let W^n be an h-cobordism where $n \geqslant 7$. (One can get the dimension down to 6 using the improved versions of the Whitney trick. We won't worry about this.) Then it has a presentation with 3-handles and 4-handles only. The homology is determined by a single matrix $\mathfrak{M}_4$ of intersection numbers; by the Whitney trick, we can assume that these intersection numbers are equal (in absolute value) to the actual geometric number of intersection points between the belt spheres of the 3-handles and the attaching spheres of the 4-handles.

Now the homology computed by the cellular chain complex in this case is the *relative* homology $H_*(W, \partial_- W)$, and consequently it is zero (because of the assumption that W is an h-cobordism). Thus $\mathfrak{M}_4$ is an *invertible* matrix of integers; in particular, it is square.

Now recall that the *elementary row operations* on a matrix of integers are the following:

(i) To interchange two rows;
(ii) To multiply a row by ± 1;
(iii) To add n times one row to another row, $n \in \mathbb{Z}$.

The *elementary column operations* are defined similarly.

4.20. PROPOSITION. (SMITH NORMAL FORM THEOREM) *Any integer matrix can be reduced by elementary row and column operations to a diagonal matrix in which each diagonal entry divides the next one. In particular, any invertible integer matrix can be reduced by elementary row and column operations to the identity matrix.*

This theorem is usually proved as part of the classification of finitely generated abelian groups. Its relevance here comes from the last of our geometric tool theorems:

4.21. PROPOSITION. *Let W be a cobordism equipped with a presentation having only k-handles and $(k+1)$-handles. Then any elementary row or column operation on $\mathfrak{M}_{k+1}$ can be implemented geometrically by a suitable change of the presentation.*

We won't prove this here.

4.22. THEOREM. (h-COBORDISM THEOREM) *A simply-connected h-cobordism in dimension 7 or above is a product.*

As we mentioned, the dimension 7 can be improved to 6; Donaldson gave counterexamples in dimension 5.

PROOF. By the results above, we can assume that there is a presentation having only 3-handles and 4-handles, and such that $\mathfrak{M}_4$ is the identity matrix. By the Whitney trick, we may adjust by isotopies so that the attaching sphere of each 4-handle meets the belt sphere of just one 3-handle transversely in a single point, and meets no other belt spheres. Now by the cancellation lemma we may get rid of all the handles. Thus W has a presentation with no handles at all, which makes it a product. $\square$

4.23. REMARK. Suppose that the inclusions $\partial_\pm W \to W$ are 1-connected and that $H_*(W, \partial_- W; \mathbb{Z}\pi) = 0$. Then W is an h-cobordism. This follows from the Hurewicz theorem and Poincaré duality.

2. Some consequences of the h-cobordism theorem

Here we follow Milnor's book in listing some results that can be easily deduced from the h-cobordism theorem.

DCC 4.24. THEOREM. (DISC CHARACTERIZATION THEOREM) *Let W^n be a compact simply-connected smooth manifold, $n \geqslant 6$, with simply-connected boundary and having the integral homology of a point. Then W^n is diffeomorphic to the n-disc.*

PROOF. Let D_0 be an n-disc in the interior of W. Then, by the remark above, $D \setminus D_0^\circ$ is a simply-connected h-cobordism; hence it is a product. Thus D is a disc with a cylinder attached to the boundary, i.e. a disc. $\square$

4.25. THEOREM. (GENERALIZED POINCARÉ CONJECTURE IN DIMENSION $\geqslant 6$) *If a smooth manifold M^n, $n \geqslant 6$, has the homotopy type of an n-sphere, then M is homeomorphic to S^n (by a homeomorphism which is smooth except at one point).*

PROOF. Remove the interior of a disc D_0 from M. The complement then satisfies the conditions of 4.24, so it is another disc. Thus M is a *twisted sphere* — it consists of two n-discs glued together by some diffeomorphism of their boundary $(n-1)$-spheres.

To prove that any twisted sphere is homeomorphic to an ordinary sphere, it clearly suffices to prove that any homeomorphism $S^{n-1} \to S^{n-1}$ can be extended to a homeomorphism $D^n \to D^n$. This is done by regarding D^n as the cone on S^{n-1} and extending the homeomorphism radially[4]. If the original homeomorphism is a diffeomorphism, the extended homeomorphism is smooth except at the cone point. $\square$

The argument does *not* show that M is smoothly a standard sphere, and in fact this is false, as we shall see. However, it does allow one to identify the following three objects:

(a) Θ^n, the set of smooth structures on the standard n-sphere (i.e. $\Theta^n = \mathcal{S}^{DIFF}(S^n)$;
(b) Γ^n, the group of diffeomorphisms of S^{n-1} modulo those that extend to D^n;
(c) A^n, the group of units in the monoid of all smooth n-manifolds under connected sum.

4.26. THEOREM. (DIFFERENTIABLE SCHOENFLIES THEOREM IN DIMENSIONS $\geqslant 6$) *Let Σ be a smoothly embedded $(n-1)$)-sphere in S^n, $n \geqslant 6$. Then there is an ambient isotopy of S^n that carries Σ onto the standard $S^{n-1} \subset S^n$.*

PROOF. By the Jordan-Brouwer separation theorem the complement of Σ has two components having Σ as their common boundary. Consider the closure C of one component. It is a simply-connected manifold with boundary Σ and having the homology of a point; hence, by 4.24, it is a disc. Now we appeal to the *disc theorem*[5] of Palais and Cerf: Any two orientation-preserving embeddings of a closed n-disc in a connected n-manifold are ambient isotopic. From this we obtain an isotopy that carries C to the lower hemisphere and hence Σ to the equator. $\square$

In the last two results the dimension 6 can be reduced to 5 by additional arguments.

3. The Whitehead group

I have just included the torsion chapter in with the previous. The material subsequent to this needs to be significantly shortened. We will express the book in terms of surgery up to h-cobordism, with some brief indications about how to go from h to s.

To do

4.27. DEFINITION. Let R be a ring with unit. The group $GL(R)$ is the inductive limit of the matrix groups $GL_n(R)$; in other words, $GL(R)$ is the group of infinite invertible matrices over R which differ in only finitely many places from the infinite identity matrix.

A matrix in $GL(R)$ is called *elementary* if it differs from the identity only in one place, not on the diagonal.

[4] This is called the *Alexander trick*.

[5] We have in fact already made extensive use of this theorem and its generalized version, the uniqueness of tubular neighbourhoods. They are needed to show that the operations of connected sum, surgery, and so on, are well-defined.

4.28. LEMMA. (WHITEHEAD LEMMA) *The subgroup $E(R)$ of $GL(R)$ generated by the elementary matrices is precisely the commutator subgroup of $GL(R)$.*

PROOF. The proof of the Whitehead lemma begins with the computation that any elementary matrix is the commutator of two other elementary matrices. Therefore

$$E(R) = [E(R), E(R)] \subseteq [GL(R), GL(R)]$$

and we need only prove the reverse inclusion. Now it is elementary (pardon the pun) to see that any upper or lower triangular matrix is the product of elementary matrices. The identity

$$\begin{pmatrix} 0 & x \\ -x^{-1} & 0 \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -x^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$$

shows that the left hand side belongs to $E(R)$ for any $x \in GL_n(R)$, and multiplying the left hand sides for $x = a$ and for $x = -1$ shows that

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$$

also belongs to $E(R)$. Now finally the identity

$$\begin{pmatrix} a^{-1}b^{-1}ab & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a^{-1}b^{-1} & 0 \\ 0 & ba \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} b & 0 \\ 0 & b^{-1} \end{pmatrix}$$

shows that $a^{-1}b^{-1}ab \in E(R)$, as required. □

Consequently, $GL(R)/E(R)$ is an abelian group.

4.29. DEFINITION. $K_1(R) = GL(R)/E(R)$.

It follows from the calculations above that the addition in $K_1(R)$ can be defined either by group product or by direct sum; these are equivalent.

4.30. REMARK. A pleasant description of algebraic K-theory is that it studies 'the deep structure of linear algebra over R'. To make some sense of this statement, suppose that we have an invertible matrix over R. Then we may ask: Why is it invertible? It might be invertible for the trivial reason that it is a product of elementary matrices: this is trivial in that it does not depend at all on the structure of the ring R. Or it might be invertible for some other reason, dependent on the structure of R. Now K_1R is just the 'group of non-trivial reasons for invertibility' in this sense.

For any commutative ring R one can define the determinant $\det\colon M_n(R) \to R$; it enjoys the usual multiplicative properties. Since the determinant of an elementary matrix is 1, we get an induced homomorphism

$$\det\colon K_1R \to R^\times, \tag{$*$}$$

where $R^\times$ denotes the group of units in R.

If R is a field, we know from undergraduate linear algebra that any invertible matrix of determinant 1 is a product of elementary matrices. Thus the determinant homomorphism $(*)$ is an isomorphism in this case. In view of the discussion above this can perhaps be phrased as follows.

4.31. PROPOSITION. *Linear algebra over a field is shallow.*

In the last chapter we discussed the Smith normal form theorem for matrices over $\mathbb{Z}$. This theorem works (with the same proof) for matrices over any Euclidean domain, so that an invertible matrix over such a domain is a product of elementary matrices and a diagonal matrix with units down the diagonal. This proves that linear algebra over a Euclidean domain is shallow too, or, more formally

4.32. PROPOSITION. *For any Euclidean domain R, $\det\colon K_1R \to R^\times$ is an isomorphism.*

In particular, $K_1\mathbb{Z} = \{\pm 1\}$.

Beware that this theorem does not remain true for a principal ideal domain R — despite the fact that the structure theory for finitely generated modules, usually proved as a consequence of the Smith theorem, still holds in this case! As an exercise, think about the proof of the structure theory in matrix language, and try to see why it does not prove $K_1 R = 0$.

Our interest is mainly in the functor K_1, but for completeness we will also define K_0. Let R be a ring, and consider the collection $P(R)$ of isomorphism classes of finitely generated projective modules over R. It is a monoid under direct sum.

4.33. DEFINITION. The group $K_0(R)$ is the Grothendieck group of the monoid $P(R)$.

To be precise, this means that $K_0(R)$ is an abelian group characterized by the following universal property: there is a homomorphism (of monoids) $P(R) \to K_0(R)$, and any other homomorphism from $P(R)$ to an abelian group factors through this one.

4.34. PROPOSITION. *If R is a principal ideal domain, then $K_0(R) = \mathbb{Z}$.*

PROOF. Notice first that since R is an integral domain (no zero-divisors) it has no non-trivial projections. Thus if $R \to P$ is a surjection, with P projective, we must have $P = R$ or $P = 0$. The structure theory for modules over a PID gives that any finitely-generated R-module M is the direct sum of finitely many modules M_i, each of which is a homomorphic image of R; if M is projective so is each M_i, so M is isomorphic to R^k for some k. We have proved that any finitely generated projective module over R is free, and it is clearly determined up to isomorphism by its rank k (tensor with the field of fractions to see this is well-defined); the map $[M] \mapsto \operatorname{rank} M$ gives the desired isomorphism. $\square$

Now we specialize to the case of group rings. Suppose that $R = \mathbb{Z}\pi$. Notice that for $g \in \pi$ the elements $\pm g$ are units of R, that is members of $GL_1(R)$, hence they define elements of $K_1(R)$.

4.35. DEFINITION. The *Whitehead group* $\mathrm{Wh}(\pi)$ is the quotient of $K_1(\mathbb{Z}\pi)$ by the subgroup $\{\pm g : g \in \pi\}$.

Naturally one would like to know about the Whitehead groups $\mathrm{Wh}(\pi)$. Perhaps they are all trivial? Unfortunately this is rather far from being the case in general.

4.36. EXAMPLE. Let $\pi = \mathbb{Z}/5$, generated say by t. The element $u = 1 - t - t^4$ is then a unit in $\mathbb{Z}\pi$, by direct computation; its inverse is $1 - t^2 - t^3$. Hence u defines a class $[u] \in K_1(\mathbb{Z}\pi)$. Now let $\theta = e^{2\pi i/5} \in \mathbb{C}$; then $t \mapsto \theta$ defines a homomorphism $\mathbb{Z}\pi \to \mathbb{C}$, so we get a homomorphism

$$K_1(\mathbb{Z}\pi) \to K_1(\mathbb{C}) \to \mathbb{R}^+$$

where the second arrow is induced by the determinant. Under the above homomorphism, the subgroup $\pm\pi \in K_1(\mathbb{Z}\pi)$ clearly maps to 1, so we obtain a homomorphism $\varphi\colon \mathrm{Wh}(\pi) \to \mathbb{R}^+$. But plainly $\varphi(u) = 1 - 2\cos\frac{2\pi}{5} \neq 1$, so that $[u]$ generates an infinite cyclic summand in $\mathrm{Wh}(\pi)$.

For torsion-free groups we are in better shape. In fact there is no counterexample to the following conjecture.

4.37. CONJECTURE. *Let π be a finitely generated torsion-free group; then* $\mathrm{Wh}(\pi) = 0$.

A key positive result here is

4.38. PROPOSITION. (BASS-HELLER-SWAN THEOREM) *Let π be a finitely generated free abelian group. Then* $\mathrm{Wh}(\pi) = 0$.

This result was proved in the sixties; the computation $\mathrm{Wh}(\mathbb{Z}) = 0$ (a special case) was done by Higman in the forties.

We won't go through the proof, but here are some of the ideas. Inductively it is sufficient to give a computation of the K-theory of the group ring $R[\mathbb{Z}]$ in terms of the K-theory of R. Now $R[\mathbb{Z}]$ can profitably be written as the ring of 'Laurent polynomials' $R[t, t^{-1}]$, and we see that it is some kind of algebraic analogue of the "R-valued functions on a circle." (In terms of C^*-algebras, this corresponds to the identification of the group C^*-algebra $C_r^*\mathbb{Z}$ with $C(S^1)$, via Fourier analysis.) In classical (topological) K-theory the effect of taking a product with a circle is

$$K^1(X \times S^1) = K^1(X) \oplus K^0(X)$$

and this suggests that something analogous should be true here. In fact Bass, Heller and Swan proved a formula

$$K^1 R[t, t^{-1}] = K_1 R \oplus K_0 R \oplus \mathrm{Nil}\, R \oplus \mathrm{Nil}\, R$$

where the groups $\mathrm{Nil}\, R$ vanish if R is a left regular ring (in particular, if R itself is a Laurent polynomial ring over $\mathbb{Z}$ in finitely many variables). Let's consider for simplicity the case $R = \mathbb{Z}$, so the Bass-Heller-Swan formula is computing $K_1(\mathbb{Z}[\mathbb{Z}])$. We know $K_1\mathbb{Z} = 0$, $K_0\mathbb{Z} = \mathbb{Z}$. The BHS map from $K_0 R$ to $K_1 R[t, t^{-1}]$ proceeds by associating to a projective module P over R the endomorphism t of the projective $R[t, t^{-1}]$-module $P \otimes_R R[t, t^{-1}]$. Here, then, we find that $K_1\mathbb{Z}[t, t^{-1}]$ is $\mathbb{Z}$ generated by t. But t is a group element, so $\mathrm{Wh}(\mathbb{Z}) = 0$, as asserted.

The Bass-Heller-Swan theorem leads to an inductive definition of 'lower K-theory' groups K_i, $i < 0$, such that the formula

$$K_i R[t, t^{-1}] = K_i R \oplus K_{i-1} R \oplus \mathrm{Nils}$$

remains true.

4. Whitehead torsion

Imagine that you are teaching a first course in algebraic topology, and you want to motivate the concept of homotopy equivalence. So, you explain, the disc is homotopy equivalent to a point; an annulus is homotopy equivalent to a sphere and so on.

Each of these homotopy equivalences is of a comparatively elementary sort. To be precise:

4.39. DEFINITION. Let X be a space, $f\colon D^{n-1} \to X$ a map. Let $Y = X \cup_f D^n$, where the identification is made along one hemisphere of $\partial D^n = S^{n-1}$. Then we say that Y is obtained from X by an *elementary expansion*, or that X is obtained from Y by an *elementary collapse*.

Plainly X and Y are homotopy equivalent. In terms of CW-theory, Y may be thought of as obtained from X by attaching first an $(n-1)$-cell (along $f_{|S^{n-2}}$) and then an n-cell (along a map half of which is f and the other half of which is the identity map to the previously attached $(n-1)$-cell). The two cells are analogous to the two cancelling handles in Smale's cancellation lemma.

4.40. DEFINITION. A homotopy equivalence is *simple* if it can be obtained by a succession of elementary expansions and collapses.

One reason for interest in this notion from the point of view of manifold theory is the following. Let P be a polyhedron and let K be a complex. K is said to be *full* if every simplex of P which has all its vertices in K is a simplex of K. In this case the *characteristic function* $\lambda\colon P \to [0, 1]$, defined to be the unique simplexwise affine function which is 1 on the vertices of K and 0 on the other vertices, has $\lambda^{-1}\{1\} = K$, and we may make the

FIGURE 5. Elementary expansion

FIGURE 6. The regular neighbourhood of a polyhedron and of an elementary expansion of it are homeomorphic.

4.41. DEFINITION. A *regular neighbourhood* of K in P is $\lambda^{-1}([0,t])$ for some $0 < t < 1$.

This is the PL substitute for the notion of tubular neighbourhood. Up to isotopy, the choice of t does not matter.

4.42. THEOREM. *Let X and Y be finite complexes in a PL manifold (Euclidean space, for example). Suppose that X and Y are related by an elementary expansion (or collapse). Then X and Y have PL homeomorphic regular neighbourhoods.*

I won't prove this, but hope the figure makes it plausible. For a proof see Rourke and Sanderson thm 3.26.

Two PL manifolds M and N are said to be *stably PL homeomorphic* if $M \times \mathbb{R}^n$ is PL homeomorphic to $N \times \mathbb{R}^n$ for some large n.

4.43. COROLLARY. *Let M and N be two stably parallelizable compact PL manifolds. Then M and N are stably PL homeomorphic if and only if they are simple homotopy equivalent.*

In the 1940's, Whitehead asked: Are there homotopy equivalences which are not simple? To answer this question he introduced the idea of the *torsion* of a homotopy equivalence. Let $i\colon X \to Y$ be a homotopy equivalence of CW complexes. By a classical construction (mapping cylinder) we may and shall assume that X is a subcomplex of Y and that i is the inclusion. Then (Y, X) is a relative CW-complex, and Y is obtained from X by attaching successively 0-cells, 1-cells, 2-cells and so on. We may then consider the relative cellular chain complex $C_*(Y, X)$; this is defined as in the last chapter, except that I now want to consider coefficients in $\mathbb{Z}\pi$. We need to choose a π-trivialization and an orientation of each cell; when we have done this, the cellular chain complex is a complex of finitely generated, based, free $\mathbb{Z}\pi$-modules, and it has zero homology (i.e. it is acyclic) because i is a homotopy equivalence.

We will define an invariant of such complexes, which one could think of as a 'deep Euler characteristic'. Let R be any ring and consider finite acyclic chain complexes C of fg based free R-modules.

4.44. DEFINITION. If such a chain complex contains only two non-vanishing groups and one non-trivial differential $d_n \colon C_n \to C_{n-1}$, let $\mathfrak{M}_n$ be the matrix of d_n relative to the given bases, and define $\tau(C) \in K_1(R)$ to be $(-1)^n$ times the class of $\mathfrak{M}_n$.

We would like to reduce the general case to the case of complexes of length two. In analysis, dealing say with elliptic complexes, one does this by taking adjoints and forming $d + d^* \colon C_{\text{even}} \to C_{\text{odd}}$. This technique is not available for complexes of general R-modules. However, the basic function of the inner product is to provide 'complements' to submodules, and this can be done by arguments using projectivity.

4.45. LEMMA. *Let (C, d) be a finite*[6] *acyclic chain complex of projective R-modules; then it is chain contractible, and each boundary submodule $B_k = Z_k = \ker d_k$ is a direct summand in C_k.*

PROOF. By induction, using the definition of 'projective'. □

Thus our complex has all the B_k finitely generated projective R-modules. By adding on trivial complexes of length 2 (geometrically this corresponds to introducing cancelling pairs of cells or handles) we can assume that all the B_k are actually *free* modules. We choose (arbitrarily) bases for them and complements X_k for them in C_k. Let $\mathfrak{N}_k$ be the matrix of

$$C_k = X_k \oplus B_k \to B_{k-1} \oplus B_k.$$

This is an invertible matrix and it is well defined up to the choice of projection onto B_k, an upper triangular term and hence a product of elementary matrices. Thus its class in $K_1 R$ is well defined. We now define the *torsion* in general to be the alternating sum

$$\tau(C) = \sum (-1)^k [\mathfrak{N}_k] \in K_1 R.$$

It is easy to see that this agrees with the previous definition for complexes of length two.

For a homotopy equivalence $f \colon X \to Y$, want to define the torsion as the torsion of the mapping cylinder. We have to choose bases for the free modules appearing in the cellular chain complex. These bases are not canonical[7], but the ambiguities appearing therein only change the torsion by the action of the group $\pm\pi$ on K_1. Thus the torsion

[6] Actually, we only require it to be bounded below.

[7] The *ordering* of the bases is not canonical either, but this does not matter. The reason for this is that it is possible to write the matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

as a product of elementary matrices. Since we have explicitly factored out the action of -1, it follows that any transposition of rows or columns acts trivially on the Whitehead group.

$\tau(f)$ is well-defined in $\mathrm{Wh}(\pi)$. Using cellular approximation, one can define $\tau(f)$ even for a non-cellular homotopy equivalence.

4.46. REMARK. There are various useful formulae for the torsions of certain kinds of composite maps. For instance, say that $f\colon X \to Y$ and $g\colon Y \to Z$ are homotopy equivalences. Then one has
$$\tau(g \circ f) = \tau(f) + \tau(g).$$
Suppose that $f\colon X \to Y$ and consider $f \times 1\colon X \times Z \to Y \times Z$. Then
$$\tau(f \times 1) = \tau(f)\chi(Z)$$
where χ denotes the Euler characteristic.

4.47. THEOREM. (WHITEHEAD) *A homotopy equivalence f between finite complexes is simple if and only if $\tau(f) = 0$.*

PROOF. (SKETCH) This is like the proof of the h-cobordism theorem. The easy way around, we verify by hand that the torsion of an elementary expansion or collapse is zero. Then the additivity of torsion tells us that any simple homotopy equivalence has torsion zero. The other way round, suppose we have a homotopy equivalence with zero torsion. Then we can do ‘cell-trading’ to arrange that the relative cellular complex only has cells in two consecutive dimensions. Since the torsion is zero, the boundary matrix of the relative cellular complex is a product of elementray matrices. One shows that these elementary matrices represent moves that can be effected geometrically, by expansions and collapses. □

Clearly the torsion of a cellular homeomorphism is zero. For a long time it was unknown whether the Whitehead torsion of a (non-cellular) homeomorphism must be zero. This was resolved by Chapman (1972) using infinite-dimensional topology.

4.48. DEFINITION. The *Hilbert cube Q* is the product $\prod_{i=1}^{\infty}[0,1]$ of countably many copies of the unit interval.

4.49. THEOREM. (CHAPMAN) *Two finite complexes X and Y are simple homotopy equivalent if and only if $X \times Q$ and $Y \times Q$ are homeomorphic. In fact, $f\colon X \times Y$ is a simple homotopy equivalence if and only if $f \times 1\colon X \times Q \to Y \times Q$ is homotopic to a homeomorphism.*

4.50. COROLLARY. (TOPOLOGICAL INVARIANCE OF WHITEHEAD TORSION) *If f is a homeomorphism, then it is a simple homotopy equivalence, i.e., $\tau(f) = 0$.*

There are now other proofs of this using controlled topology. See later, perhaps.

5. The s-cobordism theorem

Let us now consider the problem of determining whether a *non*-simply-connected h-cobordism is a product. The argument given in the last chapter for the h-cobordism theorem breaks down only at the very last stage, where we needed to apply the Whitney trick in order to see that the algebraic intersection numbers of the attaching and belt spheres were actually attained geometrically. We know that in the non-simply-connected case this need not be true if we use ordinary intersection numbers, but it becomes true once again if we use the $\mathbb{Z}\pi$ intersection numbers. What changes will we have to make to do this?

Clearly the attaching discs of handles are π-trivial submanifolds, since in fact they are contractible. Choosing (arbitrarily) π-trivializations for the attaching discs, the $\mathbb{Z}\pi$ intersection numbers can be organized into matrices giving us a complex of $\mathbb{Z}\pi$-modules.

Just as before, one can show that this complex computes the homology with coefficients in $\mathbb{Z}\pi$.

The handle trading lemmas still apply, so that a high-dimensional h-cobordism can be reduced to one with 3-handles and 4-handles only. The intersection matrix $\mathfrak{M}_4$ is now a matrix over $\mathbb{Z}\pi$, and the row operations that we can perform geometrically are:

(i) To interchange two rows;
(ii) To multiply a row by $\pm g$, $g \in \pi$ (this corresponds to a change of choice of π-trivialization and/or orientation);
(iii) To add a times one row to another row, $a \in \mathbb{Z}\pi$.

We need to know, then, whether an invertible matrix over $\mathbb{Z}\pi$ can always be reduced to the identity by employing these row operations (and the corresponding column operations). The answer is given by Whitehead torsion.

Let W be an h-cobordism, and let it be given a presentation with 3-handles and 4-handles only. The matrix $\mathfrak{M}_4$ is then an invertible matrix over the group ring $\mathbb{Z}\pi$, so it defines an element of $K_1(\mathbb{Z}\pi)$.

4.51. DEFINITION. The image of this class $[\mathfrak{M}_4]$ in $\mathrm{Wh}(\pi)$ is called the *torsion* $\tau(W)$ of the h-cobordism.

Plainly, this is just the Whitehead torsion of the inclusion $\partial_- W \to W$ (which is a homotopy equivalence).

4.52. THEOREM. (s-COBORDISM THEOREM) *A high-dimensional s-cobordism is a product if and only if its torsion vanishes.*

PROOF. In one direction, $M \times I$ has a handle-decomposition without handles, hence with torsion zero. This means that the torsion is an obstruction to the existence of a product structure. In the other direction, suppose that $\tau(W)$ vanishes. Then by means of elementary row and column operations $\mathfrak{M}_4$ can be brought to diagonal form where the diagonal entries are $\pm g$, $g \in \pi$. By the non simply connected Whitney trick, we can make further isotopies to arrange that the handles intersect in pairs with geometric intersection number 1; then we can cancel them by Smale's lemma as before. □

There is something a bit asymmetrical here: an h-cobordism is trivial if the torsion of the inclusion $\partial_- W \to W$ vanishes. What about the inclusion $\partial_+ W \to W$? The answer is that there is a duality formula, just like Poincaré duality for homology.

4.53. PROPOSITION. (MILNOR DUALITY) *If W^n is an h-cobordism, then $\tau(W) = (-1)^{n-1}\tau(\bar{W})^*$, where $\bar{W}$ denotes the dual h-cobordism, and $*$ is the involution on the Whitehead group induced by the involution on $\mathbb{Z}\pi$ coming from the first Stiefel-Whitney class; in other words, $\tau(\partial_- W \to W) = (-1)^{n-1}\tau(\partial_+ W \to W)$.*

This is pretty clear once we have rolled up to have handles in two consecutive dimensions only. One has a similar formula for a homotopy equivalence f of *closed* manifolds, namely $\tau(f) = (-1)^{n-1}\tau(f)^*$; that is, $\tau(f)$ lives in a certain involution-invariant subgroup of the Whitehead group.

Here is a geometric corollary (a 'rolling-up' argument). For it we need the following observation. Let W be an h-cobordism and Z a closed manifold. Then $W \times Z$ is also an h-cobordism. It can be shown that

$$\tau(W \times Z) = \tau(W)\chi(Z)$$

where χ denotes the Euler characteristic. This is just a corollary of the product formula for Whitehead torsion.

4.54. PROPOSITION. *Let M and N be compact high-dimensional manifolds. Suppose that $M \times \mathbb{R}$ and $N \times \mathbb{R}$ are diffeomorphic; then they are diffeomorphic by a diffeomorphism which is equivariant for the natural $\mathbb{Z}$-actions by translation.*

PROOF. By gluing M onto one end of $M \times \mathbb{R} \cong N \times \mathbb{R}$ and N onto the other (in the natural way) we obtain an h-cobordism between M and N. The torsion of this h-cobordism vanishes on crossing with S^1, so by the s-cobordism theorem $M \times S^1$ and $N \times S^1$ are diffeomorphic, by a diffeomorphism which on the homotopy level is just the product of a homotopy equivalence $M \to N$ with the identity on S^1. Passing to infinite cyclic covers we get the asserted periodic diffeomorphism between $M \times \mathbb{R}$ and $N \times \mathbb{R}$. $\square$

CHAPTER 5

Exotic spheres

The first application of surgery theory, by Milnor and Kervaire, was to the classification of 'exotic spheres', that is, non-standard differentiable structures on S^n. In this chapter we will outline some of the Milnor-Kervaire theory.

Need to included specific discussion of the Brieskorn examples

To do

5.1. DEFINITION. Θ^n denotes the smooth structure set of S^n, that is the collection of (oriented) manifolds having the (oriented) homotopy type of S^n, up to h-cobordism or equivalently up to diffeomorphism.

The project is to compute Θ^n.

5.2. PROPOSITION. *The operation of connected sum makes Θ^n into an abelian group in which the standard sphere S^n is the identity.*

PROOF. Only the existence of inverses is not apparent, and I assert that, in fact, the inverse of a homotopy sphere M is the sphere $-M$ with the opposite orientation. To prove this we appeal to the h-cobordism theorem, which implies that M is the union of two n-discs glued along their boundary by some $g \in \mathrm{Diff}(S^{n-1})$. Then $-M$ is the union of two discs glued by g^{-1}, and $M\#(-M) = D^n \cup_g S^{n-1} \times I \cup_{g^{-1}} D^n$ is plainly diffeomorphic to the standard sphere. □

5.3. DEFINITION. A manifold M is *parallelizable* if its tangent bundle is trivial.

Milnor and Kervaire's analysis begins by singling out a certain subgroup of Θ^n.

5.4. DEFINITION. The subgroup $bP_{n+1} \subseteq \Theta^n$ consists of those homotopy spheres which are the boundaries of parallelizable manifolds.

Implicit in this definition is the assertion that bP_{n+1} actually *is* a subgroup. Suppose M_1 and M_2 are homotopy spheres, bounding parallelizable manifolds W_1 and W_2 respectively. Then $W = W_1 \sqcup W_2$ is a manifold with two boundary components: attach a 0-handle to this manifold to join the boundary components. You obtain a manifold W whose boundary is $M_1\#M_2$ and which can be seen to be parallelizable. Thus bP_{n+1} is closed under the group operation.

There are two reasons why a sphere M might be exotic. First, it might not bound any parallelizable manifold at all. This phenomenon is related to the group Θ^n/bP_{n+1}, which we will shortly see is a subquotient of the stable homotopy group $\pi_n(\mathbb{S}) = \lim_k \pi_{n+k}(S^k)$. Secondly, however, M might bound a parallelizable manifold but not bound a contractible one[1]. The obstructions here have to do with signatures, and can be explored by means of surgery.

[1] If it bounds a contractible manifold it is certainly a standard sphere, by the Disc Characterization Theorem.

1. Framings

5.5. DEFINITION. Let V be an n-dimensional vector bundle over a space X. A *framing* of V is a set of n continuous sections that form a basis for the fibre at every point. A *stable framing* is a framing for $V \oplus \varepsilon^m$ for some trivial bundle ε^m.

Usually we will assume that our bundles are oriented and given euclidean metrics, and then we need only consider orthonormal framings. Clearly V has a framing iff it is a trivial bundle, and the framings are then in $1:1$ correspondence with maps $X \to SO(n)$.

obstruct 5.6. PROPOSITION. *Suppose that V is an n-dimensional vector bundle over a k-dimensional CW-complex X, and that $V \oplus \varepsilon^1$ admits a framing. Then, if $k < n$, then V itself admits a framing whose sum with the given framing of ε^1 is homotopic to the given framing of $V \oplus \varepsilon^1$. If $k = n$, then V admits such a framing if and only if a single obstruction in $H^n(X;\mathbb{Z})$ vanishes.*

PROOF. We try to construct the desired framing of V by induction over the simplices. The inductive step is then this: given a cell $(D^i, \partial D^i)$, and a framing of $V \oplus \varepsilon^1$ over D^i which arises on ∂D^i from a framing of V, deform rel boundary to obtain a framing of V on D^i. If we trivialize V over D^i the problem becomes one of filling in the dotted arrow in the diagram

$$\begin{tikzcd}
\partial D^i \arrow[r] \arrow[d] & SO(n) \arrow[d] \\
D^i \arrow[r, dotted] & SO(n+1)
\end{tikzcd}$$

to make the upper triangle commute. The possibility of doing this is controlled by the homotopy group $\pi_i(f)$ of the map $SO(n) \to SO(n+1)$; however, the homotopy sequence of the fibration $SO(n) \to SO(n+1) \to S^n$ shows that this group is $\mathbb{Z}$ if $i = n$ and 0 if $i < n$. Thus there is no obstruction to extending as far as the $(n-1)$-skeleton of X; if we want to go one step further and extend over the n-skeleton, then we have a bunch of $\mathbb{Z}$'s to sort out, one for each cell; these give a cellular cohomology class which is the obstruction. $\square$

A manifold is called *stably parallelizable* if its tangent bundle has a stable framing. Applying the above result inductively, we see that a manifold M is stably parallelizable if and only if $TM \oplus \varepsilon^1$ is a trivial bundle.

We now introduce a somewhat weaker notion.

5.7. DEFINITION. A manifold M is called *almost parallelizable* if $M \setminus \{p\}$ is parallelizable for $p \in M$.

Clearly, a homotopy sphere is almost parallelizable, since the result of removing a point from it is contractible.

The result of removing a point from a connected, compact manifold of dimension n has the homotopy type of a CW-complex of dimension $(n-1)$. (This is easy to see from a handle decomposition.) Therefore, by 5.6, a compact connected stably parallelizable manifold is almost parallelizable. We ask: When is the converse true?

APSP 5.8. THEOREM. *Let M be a compact, connected, oriented, almost parallelizable n-manifold. Then*

(a) *If n is not a multiple of 4, M is always stably parallelizable;*
(b) *If n is a multiple of 4, M is stably parallelizable if and only if its signature is zero.*

In the proof we will see a number of pieces of fancy homotopy theory put to good use. We start by figuring out the obstruction. Take a disc D^n in M. Then TM is trivial over D^n, trivial over the complement, so it is completely described by the map $S^{n-1} \to SO(n)$ relating the two trivializations. (Notice that this argument shows that TM is the pull-back of some bundle over S^n. In particular, all its lower Pontrjagin classes vanish.) The bundle $TM \oplus \varepsilon^1$ will be trivial if and only if the composite

$$\gamma\colon S^{n-1} \to SO(n) \to SO(n+1)$$

is nullhomotopic, i.e. if and only if a certain element of the group $\pi_{n-1}(SO(n+1))$ vanishes.

What is this group? The homotopy exact sequence shows that $\pi_{n-1}(SO(n+1)) = \pi_{n-1}(SO(n+k))$, $k \geqslant 1$; this common group is called $\pi_{n-1}(SO)$.

5.9. THEOREM. (BOTT) *The stable homotopy groups $\pi_*(SO)$ depend only on the congruence class of r modulo 8, according to the following table.*

r modulo 8	0	1	2	3	4	5	6	7
$\pi_{r-1}(SO)$	$\mathbb{Z}$	$\mathbb{Z}/2$	$\mathbb{Z}/2$	0	$\mathbb{Z}$	0	0	0

This deals with the case of dimensions 3,5,6, and 7 mod 8 straightaway: there is no obstruction. In the case of dimensions 0 and 4 mod 8, we need to know that the obstruction is a multiple of the signature. This is taken care of by the following computation.

p1 5.10. PROPOSITION. *The homomorphism $\pi_{4k-1}(SO) \to \mathbb{Z}$, which sends the map $u\colon S^{4k-1} \to SO$ to the k'th Pontrjagin class of the stable vector bundle over S^k obtained by clutching with u, is a monomorphism. In fact, a generator of $\pi_{4k-1}(SO)$ maps to the integer $(2k-1)!(3-(-1)^k)/2$.*

Nowadays we would think of this as a computation with the Chern (or Pontrjagin) character: the Bott generator for the real K-theory of a sphere maps to a non-trivial cohomology class.

Since TM is pulled back from a sphere, this shows that the obstruction is a certain (nonzero) multiple of the top Pontrjagin class of M. But now this Pontrjagin class is a nonzero multiple of the signature, by the HIrzebruch signature formula, since all the lower Pontrjagin classes of M are zero.

This leaves the cases of dimensions 1 and 2 modulo 8. Here we need to introduce another device from homotopy theory. Suppose that $f\colon S^k \to SO(m)$ is a map. Since $SO(m)$ acts on S^m (preserving basepoint) we get from f a map

$$S^k \to \mathrm{Maps}_\bullet(S^m, S^m)$$

where $\mathrm{Maps}_\bullet$ denotes base-point-preserving maps. By adjunction this gives us a map

$$g\colon S^{k+m} = S^k \wedge S^m \to S^m.$$

Clearly the homotopy class of g is determined by the homotopy class of f.

5.11. DEFINITION. The homomorphism $J\colon \pi_k(SO(m)) \to \pi_{m+k}(S^m)$ so obtained is called the *J-homomorphism.* If $m > k+1$ the groups stabilize and we get the *stable J-homomorphism* $J_k\colon \pi_k(SO) \to \pi_k(\mathbb{S})$.

The theorem in the cases of dimension congruent to 1 or 2 mod 8 now follows from two results.

5.12. PROPOSITION. (ROHLIN) *The obstruction γ to stable parallelizability lies in the kernel of J.*

5.13. THEOREM. (ADAMS) *The stable J-homomorphism J_k is injective if k is congruent to 1 or 2 modulo 8.*

I can't say more about the proof of Adams' result. The theorem of Rohlin depends on some results on framed cobordism, which have other important applications.

5.14. DEFINITION. Let N^k be a submanifold of M^{m+k}. A *normal framing* for N in M is a framing of its normal bundle.

A submanifold with a normal framing is usually (if confusingly) spoken of as a *framed submanifold*.

If N is a framed submanifold (codimension m) of M, its *Pontrjagin invariant* is a map $M \to S^m$ defined as follows: map M to the Thom space of the normal bundle of N by crushing the complement of a tubular neighbourhood of N to a point, and then map this Thom space to S^m using the given framing. There is a natural notion of *framed cobordism* between framed submanifolds of M — a framed cobordism is a neat framed submanifold of $M \times I$ — and one has

5.15. THEOREM. (PONTRJAGIN) *For a closed manifold M, the Pontrjagin construction gives a bijection*

$$\Omega^k_{fr}(M^{m+k}) \to [M, S^k]$$

where the left hand side denotes framed cobordism classes of k-submanifolds of M.

The proof is like that of Thom's theorem, but easier.

5.16. EXAMPLE. Consider the standard equatorial subsphere $S^k \subseteq S^{m+k}$. It comes equipped with a standard framing (of its normal bundle), and any other framing is related to this one by a map $S^k \to SO(m)$. The Pontrjagin invariant of the framing is an element of $\pi_{m+k}(S^k)$. In this way we get a homomorphism $\pi_k(SO(m)) \to \pi_{m+k}(S^k)$, and it is easy to see that this is simply the J-homomorphism.

This gives us

R1 5.17. PROPOSITION. *Let $f\colon S^k \to SO(m)$ be given. If $S^k \subset S^{m+k}$ with the normal framing given by f bounds a framed submanifold of D^{m+k+1}, then f maps to zero under the J-homomorphism.*

Now Rohlin's result is a corollary. Fix a small disc D in M, as before, and embed M in S^{n+m}, m large, in such a way that the boundary $(n-1)$-sphere of D is mapped into the equatorial $S^{n-1+m} \subset S^{n+m}$, with D being mapped into the lower hemisphere, the remainder M' of M into the upper hemisphere, and everything transverse at the equator (see figure). The tangent bundles to D and M' are trivial, so according to 5.6 their normal bundles will be trivial too, provided that we chose m big enough; and the obstruction we are studying (to stably trivializing the tangent bundle) is the same as the obstruction to trivializing the normal bundle of M. This obstruction, however, is just given by the map $S^{k-1} \to SO(m)$ that compares a trivialization of ν'_M with that of ν_D on $M' \cap D = S^{k-1}$; since the trivialization arising from the disc is the standard one, proposition 5.17 shows that the obstruction belongs to $\operatorname{Ker} J$.

We have finally completed the proof of theorem 5.8. An important corollary is

5.18. PROPOSITION. (KERVAIRE-MILNOR) *Homotopy spheres are stably parallelizable.*

For the signature of a homotopy sphere is certainly zero.

2. Spheres that do not bound parallelizable manifolds

We will now investigate the quotient group Θ^n/bP_{n+1}. Let M^n be a homotopy sphere; since, as we have just proved, it is stably parallelizable, its normal bundle will be trivial when it is embedded in a sphere S^{n+k} for sufficiently large k. Choosing a framing for the normal bundle we obtain a Pontrjagin invariant in $\pi_{n+k}(S^k) = \pi_n(\mathbb{S})$. Let $p(M) \subseteq \pi_n(\mathbb{S})$ denote the collection of all Pontrjagin invariants obtained in this way, for different framings of the normal bundle.

5.19. EXAMPLE. If $M = S^n$, the standard sphere, then $p(M) = \operatorname{Image} J$.

Now if M and M' are two homotopy spheres, embedded disjointly in S^{n+k} with framed normal bundles, then their disjoint union is framed cobordant to the connected sum $M\#M'$. Hence

$$p(M) + p(M') \subseteq p(M\#M').$$

It follows that the map p defines a homomorphism from Θ^n to $\pi_n(\mathbb{S})/\operatorname{Image} J_n = \operatorname{Coker} J_n$.

5.20. LEMMA. *The kernel* $\operatorname{Ker} p\colon \Theta^n \to \operatorname{Coker} J_n$ *is equal to* bP_{n+1}.

PROOF. If a sphere M belongs to $\operatorname{Ker} p$, then, by framed cobordism theory, it can be embedded in S^{n+k} as a (normally) framed boundary of a framed submanifold W of D^{n+k+1}. Since W is a framed submanifold of a parallelizable manifold, it is stably parallelizable. But a stably parallelizable manifold W with non-empty boundary is parallelizable (because it has the homotopy type of a CW-complex of dimension $\dim W - 1$.) $\square$

Thus we have shown that Θ^n/bP_{n+1} is isomorphic to a subgroup of $\operatorname{Coker} J_n$. In particular, we have

5.21. PROPOSITION. Θ^n/bP_{n+1} *is a finite group.*

For it is a result of Serre that $\pi_n(\mathbb{S})$ is finite, $n \geqslant 1$.

3. Signature obstructions

We now turn to an investigation of bP_{n+1} itself. We will assume that $n = 4k-1$ is congruent to 3 modulo 4. Then there is a signature-type obstruction to an $M \in bP_{n+1}$ being standard. In order to define it we need to know about the signatures of almost parallelizable manifolds. Because almost all of the Pontrjagin classes of such manifolds vanish, the Hirzebruch signature theorem places strong number-theoretical restrictions on these signatures.

TK 5.22. PROPOSITION. *The signatures of* $4k$*-dimensional almost parallelizable manifolds form a cyclic subgroup of* $\mathbb{Z}$*, generated by*

$$t_k = 2^{2k-1}(2^{2k-1}-1)\frac{3-(-1)^k}{2k}B_k|\operatorname{Image} J_{4k-1}|$$

where B_k *denotes the* k*'the Bernouilli number.*

Here is a table of the numbers t_k for some small values of k.

k	2	3	4	5
t_k	224	7936	65024	1046528

Adams calculated the size of $\operatorname{Image} J$ in terms of Bernouilli numbers, so that all terms in the expression for t_k are known.

PROOF. The group structure comes from the fact that the connected sum of two almost parallelizable manifolds is almost parallelizable. Let M be an almost parallelizable manifold. Its tangent bundle is pulled back from a bundle over the sphere S^{4k}, which must be obtained by clutching from an element of $\pi_{4k-1}(SO) = \mathbb{Z}$. This element must be in $\operatorname{Ker} J$, so it must be a multiple of $|\operatorname{Image} J|$ times the generator. Now by 5.10 we find that the Pontrjagin class p_k is a multiple of $(2k-1)!(3-(-1)^l)/2|\operatorname{Image} J|$. The lower Pontrjagin classes are zero, so the Hirzebruch signature theorem gives

$$\operatorname{Sign}(M) = \frac{2^{2k}(2^{2k-1}-1)B_k}{(2k)!} p_k.$$

Thus the group of signatures is contained in $t_k\mathbb{Z}$, and to show equality we just have to construct an almost parallelizable manifold with signature t_k. To do this, frame the normal bundle of S^{k-1} in some high-dimensional sphere by the generator of $\operatorname{Ker} J$. Then this framed submanifold is a framed boundary, and capping off by a disc we obtain an almost-parallelizable manifold with the required signature. □

Suppose now that $M \in bP_{n+1}$ is the boundary of a parallelizable manifold W. We know that M is homeomorphic to a standard sphere, and therefore we can construct a closed topological manifold W^* by $W^* = W \cup_h D^{n+1}$, where $h\colon \partial D^{n+1} = S^n \to M$ is an orientation-preserving homeomorphism. Now note that if M is actually *diffeomorphic* to S^n, W^* is an almost parallelizable *smooth* manifold, and consequently its signature is a multiple of t_k (as above). If we can construct examples of W^* with other signatures, then, their boundaries will be exotic spheres.

Clearly the signature of W^* depends only on W. In fact, the Poincaré duality theorem for manifolds with boundary shows that if ∂W is a homotopy sphere, then the intersection pairing for W is still well-defined and nondegenerate in the middle dimension; its signature is the signature of W^*.

5.23. DEFINITION. A manifold W with boundary is an (m,k)-*handlebody* if it is obtained by attaching a number of k-handles to the boundary of an m-disc.

An (m,k)-handlebody is therefore specified by a number of framed spheres S^{k-1} in $S^{m-1} = \partial D^m$. We will be particularly interested in the case $m = 2k$; note that in this case the *linking number* of two S^{k-1} in S^{m-1} is well-defined.

If W is a $(2k,k)$-handlebody, $k > 2$, then it is $(k-1)$-connected and its k'th homology is generated by the cores of the attached handles. We may construct an explicit set of embedded spheres representing the generators of the homology as follows: let $S \in \partial D^{2k}$ be a $(k-1)$-sphere along which a handle is attached. By Whitney's theorems, the embedding of S in ∂D can be extended to an embedding of a k-disc in D itself. Construct an embedded k-sphere in W as the union of this k-disc with the core k-disc of the attached handle. Let Σ_j denote the embedded sphere corresponding to the j'th handle, and let $\mathfrak{I}$ (the *intersection matrix*) be the matrix of intersection numbers $[\Sigma_i : \Sigma_j]$.

5.24. PROPOSITION. *Suppose that W is a $(2k,k)$-handlebody, $k > 2$, with intersection matrix $\mathfrak{I}$. If $\mathfrak{I}$ is invertible (over $\mathbb{Z}$), then ∂W is a homotopy sphere.*

PROOF. Clearly ∂W is $(k-2)$-connected; also one can show that the matrix $\mathfrak{I}$ is the matrix of the map $H_k(W) \to H_k(W, \partial W)$ (relative to the natural bases). Thus by the long exact sequence, if $\mathfrak{I}$ is invertible then ∂W is a homology sphere, hence a homotopy sphere since it is simply connected. □

We need also some remarks about symmetric bilinear forms over $\mathbb{Z}$. Such a form can be represented by a symmetric matrix of integers; it is *unimodular* if the matrix is invertible over $\mathbb{Z}$ (i.e. has determinant ± 1), and it is *even* if all the diagonal entries are even.

5.25. PROPOSITION. (VAN DER BLIJ'S LEMMA) *The signature of an even unimodular form is a multiple of 8.*

We will discuss this later. In particular, the first 'interesting' unimodular form is the positive definite form E_8, with matrix

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

plumb

5.26. PROPOSITION. (MILNOR) *Given any even unimodular form Q over $\mathbb{Z}$ of rank $4k$, one can find a parallelizable $(4k, 2k)$-handlebody W whose intersection matrix is Q.*

This result will be proved in the next section.

In particular, for any dimension $n = 4k - 1 \geqslant 7$ one can find a homotopy n-sphere which is the boundary of a parallelizable manifold W with signature 8. (One takes W to be a handlebody whose intersection form is the sum of E_8 and an appropriate number of copies of the 'hyperbolic' form $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.) We call this sphere the *Milnor manifold M_n* in dimension n. From the reasoning above, the Milnor manifold is exotic. In fact, we can state

5.27. PROPOSITION. *The Milnor manifold M_n generates a cyclic subgroup of bP_{n+1} of order at least $t_k/8$, where t_k is defined in (5.22).*

In fact Milnor and Kervaire went on to prove a sharper theorem.

5.28. THEOREM. *(In the case $n = 4k-1$) The group bP_{n+1} is cyclic of order exactly $t_k/8$, generated by the Milnor manifold.*

Let us see roughly how the proof of this result is accomplished. Clearly it will be enough to prove that any homotopy sphere which bounds a parallelizable manifold of signature a multiple of t_k also bounds a contractible manifold. Now we know by explicit construction that the standard sphere can be written as the boundary of a parallelizable manifold of signature any multiple of t_k; so, by taking connected sums along the boundary (aka attaching 0-handles to a disjoint union) we see that it's enough to prove that if M bounds a parallelizable manifold of signature zero, then it bounds a contractible manifold.

This is accomplished by surgery. Suppose that W is parallelizable with ∂W a homotopy sphere and Sign W equal to zero. One makes a succession of surgeries on W, keeping the boundary fixed, and killing more and more of the homotopy groups. As we explained in chapter 1, but have not yet proved, the only obstruction to killing all the homology of W in this way is the signature. Here, however, the signature is zero, so surgery can be done successfully. That completes the proof of the theorem.

Finally I would like to sketch out how this computation, special to the sphere, is going to fit into the general framework of surgery theory (when we get around to developing it!).

FIGURE 1. Plumbing to construct an intersection of 1

What we have produced is an exact sequence (in the case $n = 4k - 1$)

$$\mathbb{Z} \to \mathcal{S}(S^n) = \Theta^n \to \mathcal{N}(S^n) \to 0.$$

The $\mathbb{Z}$ should be thought of as a group of 'surgery obstructions', as the argument above shows. The $\mathcal{N}$ in this case is a subgroup of $\operatorname{Coker} J$. It is called the set of 'normal invariants' — basically, it is the bundle-theoretic data, corresponding in our discussion (in Chapter 1) to the choice of an appropriate 'stable tangent bundle'. The main result of surgery theory in general will be a similar 'exact sequence', allowing the structure set to be computed from bundle data and surgery obstructions.

4. Plumbing

We still need to explain how proposition 5.26 is proved. This uses a technique called 'plumbing'. We start by asking when a handlebody is parallelizable.

5.29. PROPOSITION. *A $(2k, k)$-handlebody is parallelizable if and only if the normal bundles to the presentation spheres Σ_j are stably trivial.*

PROOF. Since the tangent bundle to a sphere is stably trivial, the necessity of the condition is obvious. Conversely suppose that the normal bundles to the presentation spheres are stably trivial. Then the restriction of TW to each presentation sphere is stably trivial, hence trivial. Consider now the attachment of a single handle to a disc. The tangent bundle to W is trivial over the disc, trivial over the presentation sphere Σ, and these intersect in a contractible space (a hemisphere of Σ). Therefore the tangent bundle is trivial over $D \cup \bigcup \Sigma_i$. But W deformation retracts onto this set. Hence TW is trivial. $\square$

Now suppose that a given even $r \times r$ integer matrix $\mathfrak{I}$ is prescribed. We propose to manufacture a parallelizable handlebody which has $\mathfrak{I}$ as its intersection matrix. To do this we start with a trivial handlebody obtained by attaching the r handles to r distinct framed $(k-1)$-spheres sitting in distinct k-discs in the boundary $S^{2k-1} = \partial D^{2k}$. Clearly the intersection matrix here is zero. It will be sufficient now to prove that we can move the

FIGURE 2. Geometric and algebraic self-intersections

attaching (framed) spheres by regular homotopies so as to have either of the following effects:

(a) Change an off-diagonal matrix entry by ± 1;
(b) Change a diagonal matrix entry by ± 2

while at the same time not changing the stable normal bundles of the presentation spheres.

The trick that makes this possible is that a small piece of a k-handle will look like $D^k \times D^k$. Therefore, two distinct k-handles can be made to overlap at some common $D^k \times D^k$, and this will introduce an intersection number of ± 1 (depending on orientations). Thus, for example, to introduce an intersection number of 1 between two presentation spheres Σ_1 and Σ_2, we make the regular homotopy of the attaching spheres illustrated in the figure.

In the figure, the new handlebody is obtained by attaching handles to the larger disc along spheres S_1 and S_2. The shaded region between the smaller and larger discs represents the 'track' of the homotopy; the dark region is the $S^l \times D^k$ where the plumbing takes place.

One can also plumb to construct *self*-intersections (diagonal matrix elements). In this case, however, one can only change matters by ± 2. The reason is that a single 'geometric' self-intersection point changes the algebraic self-intersection number (which is just the Euler class of the normal bundle) by 2. For the algebraic self-intersection of Σ is the signed count of the number of intersection points of Σ and a generic section of ν_Σ; and two such intersections correspond to each 'geometric' self-intersection of Σ. In the dimensions that we are considering ($2k$-spheres in a $4k$-manifold) the signs are additive.

The normal bundles are affected by this procedure. However, to see that the stable normal bundles are unaffected, take the product of the whole picture with $\mathbb{R}$. Then by a small perturbation all the intersections may be eliminated, so the homotopies can be embedded and give isomorphisms between the normal bundles of the initial and final attaching spheres.

CHAPTER 6

Normal invariants and spherical fibrations

We are interested in studying homotopy equivalences between manifolds. As we have seen in the examples in Section 4, such a homotopy equivalence can change the tangent bundle. Moreover, the examples suggest that this 'flexibility' of the tangent bundle is related to the J-homomorphism $\pi_k(SO(m)) \to \pi_{m+k}(S^m)$. In this chapter we will develop this idea systematically. The central result is that if two manifolds are homotopy equivalent, the *stable spherical fibrations* underlying their tangent (or normal) bundles must be equivalent. The forgetful map from vector bundles to spherical fibrations is exactly the J-homomorphism.

1. Spherical Fibrations

6.1. DEFINITION. A *spherical fibration* $E \to B$ is a Serre fibration whose fibre is homotopy equivalent to a sphere.

Recall that a *Serre fibration* is a map that has the unrestricted homotopy lifting property: that is, $p\colon E \to B$ is a Serre fibration if any commutative diagram of the form

$$\begin{tikzcd} X \times \{0\} \arrow[r] \arrow[d] & E \arrow[d, "p"] \\ X \times [0,1] \arrow[r] \arrow[ur, dotted] & B \end{tikzcd}$$

can be completed by filling in the dotted arrow. The 'fibre' of such a fibration is the inverse image of a point in B; its homotopy type is well defined (provided that B is path connected). There is a standard procedure ("Serre's construction") in homotopy theory whereby any map $f\colon X \to B$ can be 'made into' a Serre fibration, that is, X can be replaced by a space E with a canonical homotopy equivalence $X \to E$ and a Serre fibration $p\colon E \to B$ such that the diagram

$$\begin{tikzcd} X \arrow[r] \arrow[dr, "f"'] & E \arrow[d, "p"] \\ & B \end{tikzcd}$$

commutes. (We define $E = \{(x, \varphi) \in X \times \mathrm{Maps}([0,1], B) : f(x) = \varphi(0)\}$.) The fibre of the new Serre fibration p is referred to as the *homotopy fibre* of f, and we will allow ourselves to speak of f as a spherical fibration if its homotopy fibre is a sphere, i.e. if p is a spherical fibration.

6.2. EXAMPLE. Suppose $V \to B$ is a vector bundle of fibre dimension n. Then the unit sphere bundle of V is an $(n-1)$-spherical fibration.

Corresponding to the addition of a trivial bundle to a vector bundle, there is an operation of 'fiberwise suspension' of a spherical fibration (replace E by $B \cup_p [0,1] \times E \cup_p B$, with the obvious map to B) which replaces an n-spherical fibration by an $(n+1)$-spherical fibration. This means that it makes sense to speak of 'stable spherical fibrations', just as it does to talk of 'stable vector bundles'.

The natural notion of equivalence for spherical fibrations (corresponding to isomorphism for vector bundles) is *fibre homotopy equivalence*: two spherical fibrations $E \to B$ and $E' \to B$ are fibre homotopy equivalent if there is a commutative diagram

$$\begin{tikzcd} E \arrow[r, "\simeq"] \arrow[d] & E' \arrow[d] \\ B \arrow[r, equal] & B \end{tikzcd}$$

in which the top row is a homotopy equivalence.

6.3. DEFINITION. $G(n)$ denotes the topological monoid of homotopy equivalences $S^{n-1} \to S^{n-1}$. G denotes the direct limit $\lim G(n)$ under suspension.

6.4. THEOREM. (STASHEFF) *There is a classifying space BG, with $\Omega BG \simeq G$, such that the fiber homotopy equivalence classes of stable spherical fibrations over a finite complex X are in natural $1:1$ correspondence with homotopy classes of maps $X \to BG$.*

Thus one can think of spherical fibrations loosely as 'fibre bundles' with 'group' G. There are also classifying spaces $BG(n)$, as well as corresponding spaces BSG, etc, for *oriented* spherical fibrations ($SG(n)$ consists of *orientation-preserving* homotopy equivalences $S^{n-1} \to S^{n-1}$).

6.5. REMARK. Every oriented vector-bundle gives rise to an oriented spherical fibration, so there is a map of classifying spaces $BSO \to BSG$. This map is closely related to the J-homomorphism which we studied in the previous chapter. Let us see why this is so.

Pick a base point $* \in S^n$. Then the action of $SG(n+1)$ on $*$ gives a map $SG(n+1) \to S^n$ and we have

6.6. LEMMA. *This map is a Serre fibration, with fiber the monoid $SF(n)$ of orientation preserving homotopy equivalences $S^n \to S^n$ that preserve the basepoint.*

The inclusion $SG(n) \to SG(n+1)$ has image in $SF(n)$, so that the limit SG might equally be called SF; some authors use this notation, especially if they want the letter G for other purposes.

6.7. PROPOSITION. *For $i \geqslant 1$, $\pi_i(SF(n)) \equiv \pi_{i+n}(S^n)$. Hence, $\pi_i(SF) = \pi_i(\mathbb{S})$.*

PROOF. By standard adjunction formulae,

$$\mathrm{Maps}_\bullet(S^n, S^n) = \Omega^n S^n;$$

the base-point-preserving maps $S^n \to S^n$ are just the n-fold loop space of S^n. This space is divided into connected components parameterized by the degree. A map $S^n \to S^n$ is an orientation-preserving homotopy equivalence if and only if it has degree 1. Hence, $SF(n)$ is a connected component (that corresponding to degree 1) of $\Omega^n S^n$; the result follows. □

We now see that the map $SO \to SF$ which associates to a (stable) orthogonal transformation the corresponding (stable) self-homotopy-equivalence of a sphere induces $\pi_i(SO) \to \pi_i(SF) = \pi_i(\mathbb{S})$, and it is plain that this is another description of the J-homomorphism.

6.8. LEMMA. *Any spherical fibration can be embedded (up to homotopy) as a subfibration of a fibration with contractible fiber (analogous to the disc bundle in the case of the spherical fibration associated to a vector bundle).*

The proof is homotopy-theoretic abstract nonsense, compare the proof that every map is homotopy-equivalent to a fibration, above.

As we will need to make use of this space and other spaces associated to a spherical fibration it will be helpful to change our notation somewhat. Let us denote a spherical fibration by a Greek letter, such as ξ. Its total space will be denoted by $S(\xi)$ and the total space of the associated contractible fibration by $D(\xi)$.

We can now see that all the usual operations on vector bundles now have counterparts for spherical fibrations. Thus one can define pullbacks, external products, Whitney sum, and so on of spherical fibrations. For example, to define the Whitney sum of two spherical fibrations ξ_1 and xi_2 of fiber dimensions $k_1 - 1$ and $k_2 - 1$ over B, one first forms the 'disc' fibrations $D(xi_1)$ and $D(xi_2)$. Their product is a fibration over $B \times B$, which can be restricted to the diagonal (i.e. pulled back over the diagonal map $B \to B \times B$) to yield a fibration over B. Then we define the Whitney sum of ξ_1 and ξ_2 to be the $(k_1 + k_2 - 1)$-spherical fibration over B with total space

$$S(xi_1) \times D(\xi_2) \cup D(\xi_1) \times S(xi_2) \subseteq D(xi_1) \times D(xi_2).$$

EXERCISE: Verify that this operation corresponds to the Whitney sum of vector bundles.

Since spherical fibrations are homotopically similar to vector bundles it is not surprising that the Thom isomorphism theorem can be proved for them as for vector bundles. We need to define the Thom space.

6.9. DEFINITION. Let ξ be an $(n-1)$-spherical fibration over B. The *Thom space* $T(\xi)$ is the mapping cone of the projection $p\colon S(\xi) \to B$; in other words, it is $S(\xi) \times [0,1] \sqcup B$ modulo the equivalence relation which identifies all points $(x, 0)$, $x \in S(\xi)$, with each other and each point $(x, 1)$, $x \in S(\xi)$, with $p(x) \in B$.

It is easy to verify that the Thom space (in the sense of this definition) of the spherical fibration underlying a vector bundle is homotopy equivalent to the Thom space (in the old sense) of the vector bundle itself. Notice that the pair $(T(\xi), \infty)$ (where ∞ denotes the cone point) is excision equivalent to $(D(\xi), S(\xi))$.

6.10. PROPOSITION. *Let $p\colon S(\xi) \to B$ be an oriented $(n-1)$-spherical fibration (i.e. classified by a map to $BSG(n)$). Then there is defined a* Thom class $\varphi \in H^n(T(\xi), \infty; \mathbb{Z})$ *such that cup and cap products with φ define isomorphisms*

$$H_{n+k}(T(\xi), \infty) \to H_k(B), \qquad H^k(B) \to H^{n+k}(T(\xi), \infty).$$

What about the case of non-orientable fibrations? One can still define the *first Stiefel-Whitney class* w of a spherical fibration: it is just the homomorphism $\pi_1 B \to \mathbb{Z}/2$ which tells us how the fundamental group of B acts on the top-dimensional cohomology of the fiber. Then there is still a Thom isomorphism theorem if one uses cohomology with w-twisted coefficients. In fact one gets isomorphisms

$$H_{n+k}(T(\xi), \infty) \to H_k(B; \mathbb{Z}^w), \qquad H^k(B) \to H^{n+k}(T(\xi), \infty; \mathbb{Z}^w)$$

in this case.

PROOF. We will stick with the oriented case for simplicity, but the proof is essentially the same in general — it just uses more machinery.

Observe that the homotopy groups $\pi_{i+1}(p) = \pi_{i+1}(D(\xi), S(\xi))$ of the projection are equal to π_i of the fiber (by the homotopy exact sequence of the fibration) so that the pair $(D(\xi), S(\xi))$ is $(n-1)$-connected and there is a canonical isomorphism $\pi_n(D(\xi), S(\xi)) \to \mathbb{Z}$. By the (relative) Hurewicz theorem we get

$$H_r(D(\xi), S(\xi); \mathbb{Z}) = \begin{cases} 0 & (r < n) \\ \mathbb{Z} & (r = n) \end{cases}$$

and hence by the Universal Coefficient Theorem $H^n(D(\xi), S(\xi); \mathbb{Z}) = \mathbb{Z}$; let φ be the (positive) generator of this group. This is the Thom class.

Now we appeal to the Serre spectral sequence of a fibration. (Because of our assumption of orientability we can use untwisted coefficients.) Specifically, we observe that cup product with φ induces an isomorphism on the E_2 terms from the spectral sequence of the trivial fibration $B \to B$ to the spectral sequence of $(D(\xi), S(\xi)) \to B$. However, both spectral sequences collapse at E_2, the first one for trivial reasons, and the second one because the E_2^{pq} vanish for $q \neq n$. Thus cup product with φ also induces an isomorphism on the E_∞ terms, that is, an isomorphism from $H^k(B)$ to $H^{n+k}(D(\xi), S(\xi))$. □

2. Spanier-Whitehead Duality

3. A theorem of Atiyah

[?]

4. The Spivak normal fibration

5. Normal maps and normal invariants

6. Framing obstructions for immersed spheres

Let X be a finite simplicial complex, with fundamental group π.

6.11. DEFINITION. A finite complex X is a *Poincaré duality space* (for some orientation character $w\colon \pi_1 X \to \mathbb{Z}/2$) of *formal dimension* n if it admits a fundamental class $[X] \in H_n^w(X;\mathbb{Z})$ that induces duality maps

$$D\colon H^r(M;\mathbb{Z}\pi) \to H_{n-r}^w(M;\mathbb{Z}\pi),$$

as in Chapter 2.

This definition is cribbed directly from chapter two. In that chapter, we proved that a manifold is a Poincaré duality space.

Using the ideas of torsion from chapter 5 we can define a stronger notion. A space X is a *simple Poincaré duality space* if there is a chain level map that implements Poincaré duality and which is a *simple equivalence*, meaning that it has zero Whitehead torsion. (This makes sense because the algebraic mapping cone of the Poincaré duality map is an acyclic complex of $\mathbb{Z}\pi$-modules.) A manifold is a simple Poincaré duality space.

To do

This gets skipped

6.12. EXAMPLE. Examples of Poincaré duality spaces may be obtained by gluing together manifolds with boundary by homotopy equivalences along the boundary. For example, consider the lens spaces $L(7,1)$ and $L(7,2)$, which we showed above are homotopy equivalent but not simple homotopy equivalent. Milnor showed that these spaces are in fact cobordant. Gluing the two ends of the cobordism by a homotopy equivalence, one obtains a Poincaré duality space which is not a simple Poincaré duality space, so in particular is not (PL homeomorphic to) a manifold. In fact, one can show by a further argument that this space is not even *homotopy equivalent* to a manifold.

To do

Suggest we construct the Spivak normal bundle *only* for a 'patch space' like that defined above. This is much easier, gives the general idea, but maintains the perspective that PD spaces are a 'technical tool' only.

To do

Remark on H spaces

Recall now that a manifold M has a well-defined *stable normal bundle*. To define this notion, embed M in a high-dimensional Euclidean space $\mathbb{R}^N$, with normal bundle νM satisfying $\nu_M \oplus TM = \varepsilon^N$. If we increase the dimension of the embedding we merely add a 1-dimensional trivial bundle to νM, so that the 'stable normal bundle' $\nu_s M$ of M is well defined[1]. As far as stable information (e.g. Pontrjagin classes) are concerned, knowledge of $\nu_s M$ is the same as knowledge of the tangent bundle TM.

Stable vector bundles over M are classified by maps $M \to BO$. Moreover, it can be shown that there exist other classifying spaces BPL, $BTOP$, which classify the appropriate notion of normal bundle for piecewise linear and topological manifolds. At present we want to ask: What is the appropriate kind of 'normal bundle' for a Poincaré duality space?

7. Spivak's theorems

Suppose that X is a finite complex embedded in some high-dimensional Euclidean space. Then a small neighbourhood of X will always be a manifold with boundary. If X is a manifold, this is a tubular neighbourhood, so it fibres over X; but in general we do not expect a neighbourhood to fibre over X in any very strong way. We thus see that the existence of a fibration structure on a small neighbourhood reflects the manifold structure of X.

We now ask: If X is a Poincaré duality space, does a neighbourhood of X have a nice fibration structure?

6.13. THEOREM. (SPIVAK) *Let X be a connected Poincaré duality space of formal dimension n, and suppose that X is embedded linearly in $\mathbb{R}^{n+k}$ (k large) with a closed regular neighbourhood $N(X)$ having boundary $\partial N(X)$. Then the map*

$$\partial N(X) \longrightarrow N(X) \xrightarrow{\simeq} X$$

is equivalent to a $(k-1)$-spherical fibration[2].

PROOF. The idea is to reverse the previous argument. Make the pair $(N(X), \partial N(X))$ into a Serre fibration (D, S), D having contractible fibers, over X. Taking k large we can arrange that the fiber of $S \to X$ is simply connected. Because both X (by hypothesis) and (D, S) (as a manifold with boundary) satisfy Poincaré duality in the appropriate form, we can define Thom isomorphisms by composing the induced maps from the projection $D \to X$ with Poincaré duality on both sides. The conclusion is that there exists a Thom class φ such that cap and cup products with it induce Thom isomorphisms on homology and cohomology. In particular, by the Hurewicz theorem, the fiber F of $S \to X$ has $\pi_r(F) = 0$ for $r \leqslant k-2$ and $\pi_{k-1}(F) = \mathbb{Z}$. We will show that F has the homology of a sphere, and then (by a theorem of Whitehead) since it is a simply-connected homology sphere, it will be a homotopy sphere.

We now compare the spectral sequences belonging to the two fibrations $X \to X$ and $(D, S) \to X$ under cup product with φ, as before. The second of these has E_2 term

$$E_2^{pq}(D, S) = H^p(X; H^q(cF, F)) = \begin{cases} 0 & (q < k) \\ H^p(X) & (q = k) \\ ? & (q > k) \end{cases}$$

We observe that $\smile \varphi$ gives an isomorphism $E_2^{p0}(X) \to E_2^{pk}(D, S)$ for all p. Moreover $\smile \varphi$ gives a monomorphism $E_\infty^{p0}(X) \to E_\infty^{pk}(D, S)$; that is because $E_\infty^{pk}(D, S)$ is the lowest non-trivial term in the associated graded group to some filtration of $H^{p+k}(D, S)$, and we know that $\smile \varphi$ induces an isomorphism $H^p(X) \to H^{p+k}(D, S)$. Putting these

[1] By definition, a *stable vector bundle* over M is an equivalence class of vector bundles under the relation of *stable isomorphism*, that is, isomorphism after the addition of trivial bundles.

[2] In other words, when we make it into a fibration by the Serre construction, its fibre is a homotopy S^{k-1}.

facts together we find that all the differentials of the spectral sequence that map into any terms $E_r^{pk}(D, S)$ must be zero.

In particular suppose that $H^l(cF, F)$ is the first non-zero cohomology group for $l > k$. Then $E_2^{0l}(D, S) = H^l(cF, F)$ and all the differentials mapping out of any term E_r^{0l} in the spectral sequence must be zero. Hence $E_2^{0l} = E_\infty^{0l}$ and so we deduce that $H^l(D, S)$ must contain an element which restricts nontrivially to a fiber (i.e. to $H^l(cF, F)$). But this is ridiculous: according to the Thom isomorphism such an element would have to be the image, under Thom, of a $(l-k)$-dimensional cohomology class of a point, and there is no such class. We deduce that $H^l(cF, F) = 0$ for $l > k$, hence F is a homology sphere, as asserted. $\square$

6.14. DEFINITION. The stable equivalence class of the spherical fibration just constructed is called the *Spivak normal bundle* νX of X.

Of course one needs to see that the various choices (of embedding, etc) wash out when one passes to stable equivalence classes. In fact we should note that there is a stronger uniqueness statement. To see it we note that the Spivak fibration is provided with a piece of extra structure. Namely, there is a class $b \in \pi_{n+k}(T(\nu X))$ whose image in homology corresponds to the fundamental class of X under the Thom isomorphism. This just comes from crushing the embedding sphere to the Thom class in the usual way. We say that the pair $(\nu X, b)$ is a *spherical fibration with spherical top class*.

Let ν_X denote the stable spherical fibration constructed in this way; it is not hard to see that it is independent of the choices involved. In fact, there is a stronger uniqueness result. To state it, observe that a spherical fibration over a Poincaré duality space X has a well-defined 'top homology class' for its Thom space, namely the image of the top class of X under the Thom isomorphism. The Spivak bundle, constructed above clearly has the property that its top class is *spherical*, that is, it arises from an element of $\pi_{n+k}(\nu_X)$ under the Hurewicz map.

6.15. THEOREM. *Let X be a Poincaré duality space. Suppose that ξ is a spherical fibration over X whose top class is spherical. Then ξ is stable fiber homotopy equivalent to the Spivak normal bundle of X.*

I don't want to go into the proof of this in detail. It depends on the theory of Spanier-Whitehead duality which is a stable homotopy interpretation of Poincaré duality. The idea is this: let ν be the Spivak normal bundle, ξ the bundle in the theorem, with its map $S^{n+k} \to T(\xi)$. One shows that $T(\nu)$ is a 'dual' of X; and under this duality the map $S^{n+k} \to T(\xi)$ passes to a map $T(\nu - \xi) \to S^k$ which gives a stable fiber homotopy trivialization of $\nu - \xi$. For details see Browder's book.

8. Normal maps and normal invariants

Now suppose that X is a manifold in the category CAT. (We are usually interested in the smooth category, of course, but this makes sense in general.) Then plainly its Spivak normal fibration arises from a CAT bundle, the CAT normal bundle of X. In other words, we have a lift

$$\begin{array}{ccc} & & BCAT \\ & & \downarrow \\ X & \longrightarrow & BG \end{array}$$

of the classifying map for the Spivak normal structure.

To put this another way, given a Poincaré duality space X, in order that it should have any CAT manifold structure at all, a first necessary condition is that the Spivak normal bundle νX should *reduce to* CAT. Moreover, to count structures, it will be necessary to count lifts. If one lifting is given canonically (as it is, for example, if X itself happens to be a CAT manifold), then the liftings are classified by maps $X \to G/CAT$, where G/CAT is the homotopy fiber of $BCAT \to BG$.

Let us try to describe CAT reductions more geometrically.

6.16. DEFINITION. (BROWDER) Let X be a Poincaré duality space of formal dimension n. A CAT *normal map* to X is given by an n-dimensional CAT manifold M and a degree one map $f\colon M \to X$ such that the stable normal bundle of M is pulled back by f from some CAT bundle over X.

If X is a PD space with orientation character w, then we require that w should be the first Stiefel-Whitney class of M. This allows the degree to be defined sensibly. A normal map is also sometimes called a *surgery problem*, although I will avoid this terminology.

Two normal maps are said to be *normally cobordant* if there is a 'normal map with boundary' $W \to X \times [0,1]$ whose boundary is their 'difference'. Let $\mathcal{N}^{CAT}(X)$ denote the set of normal cobordism classes of CAT normal maps to X (if CAT is omitted we mean the smooth category).

The Pontrjagin-Thom construction can be applied to $\mathcal{N}(X)$:

6.17. THEOREM. (SULLIVAN) *There is a* $1:1$ *correspondence between* $\mathcal{N}^{CAT}(X)$ *and the set of homotopy classes of* CAT *reductions of the Spivak normal bundle of* X.

PROOF. Suppose first that we are given a reduction (say a vector bundle reduction) of the Spivak normal bundle. This means that X can be embedded in a high-dimensional sphere in such a way that a regular neighbourhood has the structure of the Thom space of this vector bundle V. By collapsing we get a map $\varphi\colon S^{n+k} \to T(V)$ and making this transverse at the zero section we get a manifold $M = \varphi^{-1}(X)$ equipped with a normal map to X. The usual argument as in Thom's theorem shows that homotopic φ give normally cobordant normal maps.

Conversely, suppose that $f\colon M \to X$ is a normal map. Then $\nu_s M$ is pulled back from a stable vector bundle V over X whose top class is spherical (because the top class of νM is spherical). By the uniqueness of the Spivak normal structure, this vector bundle is a reduction of the Spivak normal bundle of X. $\square$

CHAPTER 7

The algebra of surgery obstructions

Let us review our progress. We are interested in studying $\mathcal{S}(X)$, the set of manifold structures on the Poincaré duality space X. We have noticed that any such structure gives rise to a reduction of the Spivak normal bundle of X, and we have introduced the normal invariant set $\mathcal{N}(X)$ to classify these reductions. This set has a direct homotopy-theoretic description and there is an obvious forgetful map

$$\mathcal{S}(X) \to \mathcal{N}(X)$$

which sends a structure to its underlying bundle data. Our next step is to investigate the 'kernel' and 'cokernel' of this forgetful map.

Recall that $\mathcal{N}(X)$ is a cobordism group, and cobordisms can be constructed by sequences of surgeries. The 'cokernel' of the forgetful map therefore measures the obstruction to the following problem: given a normal map $M \to X$, perform a sequence of surgeries on M (and on the map) which convert it into another normal map $M' \to X$ which is a homotopy equivalence.

In order to carry out surgeries on M we need some embedded spheres in M with trivial normal bundles. The embedded spheres are provided by Whitney's embedding and immersion theorems. The trivialization of the normal bundle is provided by the additional bundle data incorporated into the definition of the Whitney map. Thus we have

fs

7.1. LEMMA. *Let $f\colon M^m \to X$ be a normal map. Let $\alpha \in \pi_r(f)$ be a relative homotopy class, with $r \leqslant m/2$: recall that this means that α is represented by a commutative diagram*

$$\begin{tikzcd} S^{r-1} \arrow[r] \arrow[d] & M \arrow[d, "f"] \\ D^r \arrow[r] & X \end{tikzcd}$$

Then the homotopy class of α contains a framed embedding (that is an embedding with trivial normal bundle) $S^{r-1} \to M$.

PROOF. By the Whitney embedding theorem, the homotopy class contains an embedding. We need to see that it is framed. By hypothesis, the stable normal bundle νM is pulled back from some bundle over X. Its restriction to $S^{r-1} \subset M$ is therefore trivial, because it is pulled back from a bundle over a disc. The tangent bundle to M, restricted to S^{r-1}, breaks up into the sum of the tangent bundle to S^{r-1} (which is stably trivial) and the normal bundle V to S^{r-1} in M. Since the tangent bundle plus the stable normal bundle of M is trivial, we find that V plus two stably trivial bundles yields a trivial bundle. Hence V is stably trivial. This however implies that V is trivial, using the obstruction theory of 5.6.

Something funny about dimensions here! □ To do

1. Surgery below the middle dimension

Now we want to prove that we can always do surgery to make a normal map highly connected. The classical way to do this is to investigate the effect of surgeries on the homology groups of the map (having made it 1-connected by some preliminary skirmishing) and then appeal to some form of the Hurewicz theorem. Here, however, we will use direct handle arguments instead. The approach is based on Ferry and Pedersen's article.

7.2. THEOREM. *Let X be a Poincaré duality space of formal dimension $n \geqslant 5$, and let $f\colon M \to X$ be a normal map. Then f is normally cobordant to a normal map $f'\colon M' \to X$ which is $\lfloor n/2 \rfloor$-connected (meaning that $\pi_r(f') = 0$ for $r \leqslant \lfloor n/2 \rfloor$).*

PROOF. By replacing X by the mapping cylinder of f, we may assume that f is an inclusion. Thus (X, M) is a relative CW-complex and we let X^i be the i-skeleton, made up of M together with the cells of X of dimensions $\leq i$.

The plan of campaign is to define inductively a sequence W^i of bordisms between M and certain manifolds M_i, such that W^i is homotopy equivalent to X^i with $W^i \to X^i \to X$ a normal map, and subcomplexes K^i of W^i obtained by attaching cells of dimension $\leqslant i$ to M, such that $W^i \setminus K^i$ deformation retracts onto M_i.

We begin the induction by letting W^0 be the disjoint union of $W^{-1} = M \times [0,1]$ and an $(n+1)$-ball corresponding to each 0-simplex of X^0. The homotopy equivalence to X^0 is the obvious one, and it is easy to see that it gives a normal map $W^0 \to X$. We let K^0 be the union of $M \times \{0\}$ and the centers of the added $(n+1)$-balls. Then M_0 is the disjoint union of M and a collection of n-spheres, and it is a deformation retract of $W^0 \setminus K^0$.

Now for the next stage of the induction. Consider a 1-cell of (X, M) which is attached by a map $S^0 \to X^0$. Using the homotopy equivalence $W^0 \to S^0$ we can find a map $S^0 \to W^0$ which represents the same homotopy class. By general position, the image of this map can be assumed to miss K^0, so we may deform it to a map $S^0 \to M_0$. The image of this map in X is clearly inessential, so by lemma 7.1 we can further deform it to a framed embedding. Now we can use it to attach a 1-handle to W^0. Performing this operation for all the 1-cells of X we construct W^1.

Notice that W^1 has a natural homotopy equivalence to X^1 and that this is a normal map to X. We define K^1 to be the cores of the attached 1-handles together with additional 1-simplices linking them at each end to K^0. This completes the second step of the induction.

We can continue inductively in this way until the middle dimension, when we construct W^m, $m = \lfloor (n+1)/2 \rfloor$. At this point we are stopped by the failure of the embedding theorem 7.1 and also by the failure of general position: we cannot make the image of S^m miss K^m by general position arguments. Let $W = W^m$ and let $M' = M_m$. Then M' is normally cobordant to M and what remains to be seen is that $M' \to X$ is $\lfloor n/2 \rfloor$-connected. By construction, W is homotopy equivalent to the m-skeleton of X and so $W \to X$ is obtained by attaching cells of dimension $> m$. It is therefore a $\lfloor n/2 \rfloor$-connected map. On the other hand, by turning the cobordism upside down (ie making use of the dual presentation) we see that W is obtained from M' by attaching handles of index $> m$. Therefore $M' \to W$ is also an $\lfloor n/2 \rfloor$-connected map. The result follows. □

For various reasons it is necessary for us to think about surgery on manifolds with boundary. Such surgeries will give rise to cobordisms of manifolds with boundary; these cobordisms will of necessity be manifolds with corners. However there are a number of different possible definitions of such cobordisms, differing in the degree of restrictiveness placed on the boundary components. Correspondingly there are different notions of surgery.

The simplest is what we may call *restricted* or 'rel boundary' surgery. Here we are given a manifold with boundary $(M, \partial M)$ and we perform surgeries only on framed spheres embedded in the interior of M. The effect of this is that the boundary stays unchanged throughout the surgery process, and the cobordism of manifolds with boundary we produce is a product cobordism on the boundary. This kind of surgery is appropriate to a situation where we are given a normal map of pairs $(M, \partial M) \to (X, \partial X)$ which is already a homotopy equivalence on the boundary, and we seek to surger it to a manifold structure while preserving the given structure on the boundary.

This version of surgery theory is not significantly different from surgery theory of closed manifolds: the same results with the same proofs continue to hold. So for example the theorem on surgery below the middle dimension becomes that a normal map of pairs $(M, \partial M) \to (X, \partial X)$ of formal dimension n, which is a homotopy equivalence on the boundary, can be surgered rel boundary to a $\lfloor n/2 \rfloor$-connected normal map.

Do we need to discuss other versions. I've cut that out.

To do

2. The homology kernel of a normal map

We will investigate the homological properties of degree one normal maps. Let $f\colon M \to X$ be such a map (for simplicity we consider only maps between closed spaces, but later on, in the proof of the π–π-theorem for example, we will also need similar results for maps of pairs).

7.3. PROPOSITION. *Let $f\colon M \to X$ be a degree one normal map. Then*

(a) *The homology map $f_*\colon H_*(M) \to H_*(X)$ is split surjective, with kernel $K_*(f)$ such that $H_*(M) = H_*(X) \oplus K_*(f)$;*
(b) *The cohomology map $f^*\colon H^*(X) \to H^*(M)$ is split injective, with cokernel $K^*(f)$ such that $H^*(M) = H^*(X) \oplus K^*(f)$;*
(c) *$K_*(f)$ and $K^*(f)$ obey Poincaré duality in the sense that the duality map on X induces an isomorphism $K^*(f) \to K_{n-*}(f)$.*

PROOF. We split f_* by the map $f_!\colon H_*(X) \to H_*(M)$, defined by

$$H_*(X) \xrightarrow{D} H^{n-*}(X) \xrightarrow{f^*} H^{n-*}(M) \xrightarrow{D} H_*(M)$$

where D stands for Poincaré duality . In other words, the characteristic property of $f_!$is that $f_!([X] \frown x) = [M] \frown f^*(x)$ for all $x \in H^*(X)$. To check that this is a splitting we compute

$$f_* f_!([X] \frown x) = f_*([M] \frown f^*(x)) = f_*[M] \frown x = [X] \frown x$$

using the naturality of the cap product and the fact that f is degree one. The rest of the proof is similar. □

We have been a bit vague about what coefficients we are using here. In the non-simply-connected situation, we take $H_*(X)$ with coefficients in some $\mathbb{Z}\pi_1(X)$-module Λ, and then $H_*(M)$ has coefficients in the induced $\mathbb{Z}\pi_1(M)$ module $f^*\Lambda$. In the nonorientable case we have to twisted the coefficients by w, as well.

We note explicitly

7.4. PROPOSITION. *Let f be a degree one normal map. Suppose that f is 1-connected and that $K_*(f;\mathbb{Z}\pi) = 0$. Then f is an equivalence.*

PROOF. This follows from the Hurewicz theorem. □

As we have seen we can always do surgery to make a normal map connected up to half the dimension. For brevity we will refer to such a normal map as *highly connected.*

SF

7.5. PROPOSITION. *Let $f\colon M \to X$ be a highly connected normal map in even dimension $n = 2k$. Then $K_i(f) = 0$ for $i \neq k$, and $K_k(f)$ is a finitely generated, stably free $\mathbb{Z}\pi$-module.*

Stably free, applied to a module M, means that there is a free module N such that $M \oplus N$ is also free.

PROOF. $K_i(f)$ vanishes for $i < k$ by connectedness, and for $i > k$ by Poincaré duality ; the same vanishing applies to the cohomology kernels. By replacing f by the mapping cylinder we may and shall assume that f is an inclusion, hence that the kernels can be calculated as the (co)homology groups of a certain chain complex, the relative chain complex $C_* = C_*(X, M)$. By an induction argument, the modules Z_i of cycles are projective for $i \leqslant k$, and the whole chain complex is chain equivalent to the truncated complex

$$\cdots \to C_{k+2} \to C_{k+1} \to Z_k \to 0.$$

Now we play the following trick. Consider the boundary map $d\colon C_{k+1} \to B_k$. It can be thought of as a cochain in the cochain complex $\mathrm{Hom}(C_*, B)$, where $B = B_k$, and it is in fact a cocycle (tautology!). But $K^{k+1}(f;B) = 0$ so that this cocycle is a coboundary. Translated, this means that there is a map $p\colon Z_k \to B_k$ such that $d \circ p = d\colon C_{k+1} \to B_k$. But d is surjective by definition, so that p splits the inclusion $B_k \to Z_k$. It follows that $H_k = Z_k/B_k$ is projective, and isomorphic to a direct summand in C_k, so it's finitely generated.

Now that we know that the homology is projective we can return to the original chain complex and easily prove by induction

$$\bigoplus_i C_{k+2i} \equiv H_k \oplus \bigoplus_i C_{k+2i-1}$$

showing that H_k is stably free. □

3. Symmetric and quadratic forms

In the previous chapter we described, geometrically, certain groups $L_n(\Gamma)$ which classified the obstructions to surgery on manifolds with fundamental group Γ. (Strictly speaking our notation should have included, as well as the group Γ, a choice of orientation character $w\colon \Gamma \to \mathbb{Z}/2$, but for simplicity we considered only the case of trivial orientation character.) In this chapter we will describe functors $L_n(R)$ for any ring R with involution, purely in algebraic terms. It will turn out, however, that these functors coincide with the previous ones when the ring with involution R is taken to be the integral group ring $\mathbb{Z}\Gamma$. The algebraic description makes available many different computational techniques.

To do

This paragraph now does not quite fit

The idea is to set up some algebra which, when applied to the group ring $\mathbb{Z}\Gamma$, precisely reproduces the available geometric information about intersections and self-intersections of immersed middle-dimensional spheres in a manifold with fundamental group Γ. Using the results of chapter 2 we see that equivariant intersection numbers certainly define a symmetric or skew-symmetric bilinear form on the middle-dimensional homology (with $\mathbb{Z}\Gamma$ coefficients). Thus we begin by systematically studying such bilinear forms.

Recall some terminology from chapter 2. Let V be a (left) module over a ring R with involution. Then $\mathrm{Hom}(V, R)$ is a right R-module, and we define V' to be the *opposite* module to $\mathrm{Hom}(V, R)$, that is the left R-module which is $\mathrm{Hom}(V, R)$ as an additive group and with the left action of $a \in R$ defined to be the right action of a^*.

7.6. DEFINITION. Let V be an R-module, where R is a ring with involution. A *sesquilinear form* on V is a R-module homomorphism $\lambda\colon V \to V'$. It is *nondegenerate* if it is an isomorphism of R-modules. The set of all sesquilinear forms[1] on V is denoted by $\mathrm{Ses}(V)$.

Naturally, we identify λ with the function $V \times V \to R$ given by $(x, y) \mapsto (\lambda(x))(y)$. By a slight abuse of notation we denote this function by λ also. The sesquilinearity condition then states that

$$\lambda(ax, by) = b\lambda(x, y)a^*$$

for $a, b \in R$ and $x, y \in V$. We now study symmetry conditions on these forms.

7.7. DEFINITION. Let $\varepsilon = \pm 1$. The *ε-symmetrization map* $T_\varepsilon\colon \mathrm{Ses}(V) \to \mathrm{Ses}(V)$ is defined by

$$T_\varepsilon\lambda(x, y) = \varepsilon\lambda(y, x)^*.$$

We define $Q^\varepsilon(V) = \ker(1 - T_\varepsilon) = \{\lambda \in \mathrm{Ses}(V) : \lambda(y, x) = \varepsilon\lambda(x, y)^*\}$. This subspace is called the space of *ε-symmetric* forms (we can also say 'symmetric' or 'skew-symmetric' according as $\varepsilon = 1$ or $\varepsilon = -1$, but the more uniform terminology saves some writing).

7.8. EXAMPLE. Let M be a compact manifold with fundamental group Γ and dimension $2n$. Then the equivariant intersection form is a $(-1)^n$-symmetric form on the $\mathbb{Z}\Gamma$-module $H_n(M; \mathbb{Z}\Gamma)$ (see chapter 2).

7.9. EXAMPLE. Let V be any R-module at all. Then $W = V \oplus V'$ is also an R-module. We can define an ε-symmetric form on W by making use of the natural pairing between V and V':

$$\lambda\big((x_1, \varphi_1), (x_2, \varphi_2)\big) = \varphi_1(x_2) + \varepsilon\varphi_2(x_1)^*.$$

One checks easily that this is indeed a sesquilinear form and is ε-symmetric. It is called the *hyperbolic* ε-symmetric form associated to V.

Hyperbolic forms arise as the intersection forms of pairs of n-spheres in a $2n$-manifold which are embedded and meet transversely in a single point. This observation, made suitably precise, is at the core of surgery theory.

We have several times remarked (see, for example, the end of chapter 6) that the information given in the intersection form does not fully account for all the geometry of *self*-intersections of middle-dimensional spheres in a manifold. The simplest example of this is the observation that the intersection matrix of a parallelizable $(4k, 2k)$-handlebody has to be even, i.e. the self-intersection numbers (diagonal entries) must be multiples of 2. We now seek a more refined algebra which includes this extra information.

Some of this self-intersection and framing stuff gets moved to chapter 3

To do

7.10. DEFINITION. We define $Q_\varepsilon(V) = \mathrm{Coker}(1 - T_\varepsilon)$.

[1] This is just an abelian group, not any kind of R-module.

This is significant even when $V = R$. In this case one easily checks that $\mathrm{Ses}(R)$ is identified with (the additive group of) R itself: the element $a \in R$ corresponds to the sesquilinear form $\lambda(x, y) = yax^*$ on R. Thus if, for example, R is a commutative ring with the trivial involution $x^* = x$, $Q_{+1}(R) = R$ and $Q_{-1}(R) = R/\langle 2\rangle$. ($\langle 2\rangle$ denotes the principal ideal generated by $2 = 1 + 1$.)

Since $T_\varepsilon^2 = 1$, $\mathrm{Image}(1 - T_\varepsilon) \subseteq \mathrm{Ker}(1 + T_\varepsilon)$, so that $1 + T_\varepsilon$ gives a well-defined map $Q_\varepsilon(V) \to Q^\varepsilon(V)$.

7.11. PROPOSITION. *If $\frac{1}{2} \in R$, then the map $1 + T_\varepsilon$ gives an isomorphism between $Q_\varepsilon(V)$ and $Q^\varepsilon(V)$.*

PROOF. Let $x \in \mathrm{Ker}(1 + T)$. Then $(1 - T)x = 2x$, so $x = (1 - T)(x/2)$ belongs to the image of $1 - T$. $\square$

To make this argument work one needs, in fact, only the existence of a central element $s \in R$ such that $s + s^* = 1$. But such elements are rare in group rings. The main point is that there is a distinction between Q_ε and Q^ε, and that this distinction has to do with 2-primary torsion.

The Q_*-groups are geometrically significant for the following reason. Suppose that we are given an immersion (with clean double points) of a sphere S^n into a $2n$-dimensional manifold M with fundamental group Γ. We assume that this immersion is Γ-trivialized, i.e. there is given a homotopy class of paths from the base point in M to a base point in the immersed sphere, or equivalently that a choice of lifting of the immersion to the universal cover $\tilde{M}$ has been made. We want to 'count' the double points algebraically. Following the definition of the $\mathbb{Z}\Gamma$ intersection numbers, we proceed as follows. At a double point P, two 'branches' of S^n meet transversely. Pick (arbitrarily) one of these branches as the 'first' and the other as the 'second'. Choose a loop which runs from the base point of S^n through S^n (avoiding other double points) until it arrives at P along the 'first' branch, then leaves along the 'second' branch and returns (again, through S^n and avoiding other double points) to the base-point. This loop (conjugated if need be by the chosen path that runs from the basepoint of M to the basepoint of S^n) defines an element $\gamma \in \Gamma$. Now 'define' an element μ of $\mathbb{Z}\Gamma$, representing the self-intersection, to be $k\gamma$, where $k \in \{\pm 1\}$ is the local intersection number (depending only on orientations) of the two transversely intersecting branches.

The trouble with this as a definition is the ambiguity over which branch is 'first'. In fact, suppose we make the opposite choice. Then the loop γ is replaced by γ^{-1}, and k is replaced by $(-1)^n k$ in the case of trivial orientation character, or more generally by $(-1)^n w(\gamma)k$ if the orientation character is w. Thus μ is replaced by $(-1)^n \mu^*$, where $*$ is the w-twisted involution on $\mathbb{Z}\Gamma$. We conclude that μ is well-defined, not as an element of $\mathbb{Z}\Gamma$ itself, but as an element of $Q_\varepsilon(\mathbb{Z}\Gamma)$, where $\varepsilon = (-1)^n$. This is simply the natural place in which to count self-intersections.

7.12. DEFINITION. Let V be an R-module. An ε-*quadratic* form on V consists of an ε-symmetric form $\lambda\colon V \times V \to R$, together with a function $\mu\colon V \to Q_\varepsilon(R)$, which are related as follows:

(i) The identity $\mu(x + y) - \mu(x) - \mu(y) = [\lambda(x, y)]$ holds in $Q_\varepsilon(R)$ (here $[\lambda(x, y)]$ denotes the equivalence class of $\lambda(x, y) \in R$ under the quotient map $R \to Q_\varepsilon(R)$;

(ii) The identity $\mu(x) + \varepsilon\mu(x)^* = \lambda(x, x)$ holds in R. (Notice, here, that $\mu(x) + \varepsilon\mu(x)^* = (1 + T_\varepsilon)\mu(x)$ is a well-defined element of $Q^\varepsilon(R) \subseteq R$, since $1 + T_\varepsilon$ maps Q_ε to Q^ε.)

(iii) The identity $\mu(ax) = a\mu(x)a^*$ holds in $Q_\varepsilon(R)$. (Here we need to remark that even though $Q_\varepsilon(R)$ is *not* an R-module, the operation $\mu \mapsto a\mu a^*$ is well-defined on it.)

In the next chapter we will check in detail that the intersections and self-intersections of middle-dimensional spheres in an even-dimensional manifold define a quadratic form in this sense over the group ring of the fundamental group.

Once again, this more refined notion of *quadratic* form differs from that of *symmetric* form only so far as the prime 2 is concerned. In fact, suppose that 2 is invertible in R. Then $Q_\varepsilon(R) = Q^\varepsilon(R) \subseteq R$. Given λ, the second identity in the definition of a quadratic form shows that we must put $\mu(x) = ((1-T_\varepsilon)/2)\lambda(x,x)$. The first identity is then the familiar polarization identity for symmetric forms, and the third is apparent. Thus every symmetric form has a unique 'quadratic structure'.

We can go a little further with this idea. An ε-symmetric form λ is said to be *even* if $\lambda(x,x) \in \mathrm{Image}(1+T_\varepsilon)$ for all x. (In the case of symmetric forms over $\mathbb{Z}$ this coincides with the notion of 'even' form that we previously introduced.) Clearly any symmetric form which has a quadratic structure must be even. Conversely if $1+T_\varepsilon$ is *injective* (as it is for $\mathbb{Z}$) then there is a $1:1$ correspondence between quadratic forms and even symmetric forms.

7.13. EXAMPLE. Consider the hyperbolic ε-symmetric form on $W = V \oplus V'$, defined above. An underlying ε-quadratic structure is provided by the function $\mu\colon W \to Q_\varepsilon R$ with $\mu(x,\varphi) = [\varphi(x)]$. As a matter of notation, the space W equipped with this ε-quadratic form will be denoted by $\mathcal{H}_\varepsilon(V)$.

Let us classify the nonsingular[2] ε-quadratic forms over certain fields. We begin with the case $R = \mathbb{R}$. Since $\frac{1}{2} \in \mathbb{R}$, there is no difference between quadratic and symmetric forms. By Sylvester's law of inertia, nonsingular symmetric bilinear forms over $\mathbb{R}$ are classified completely by their rank and signature. On the other hand, *skew*-symmetric bilinear forms over $\mathbb{R}$ are necessarily hyperbolic, with no invariant other than their rank (which must be even).

The case $R = \mathbb{C}$ (with complex conjugation as the involution) is similar as regards symmetric (now usually known as *hermitian*) forms, which are classified by their rank and signature. Now, however, since the ring contains an element i such that $i^2 = -1$, there is no difference between symmetric and skew-symmetric forms; so skew-symmetric forms are also classified by their rank and signature.

By way of contrast we now consider the fundamental 2-torsion example, $R = \mathbb{Z}/2$. Thus we have a finite-dimensional vector space V over $F = \mathbb{Z}/2$, and on this we have a bilinear $\lambda\colon V \times V \to F$ and a function $\mu\colon V \to F$ such that

$$\mu(x+y) = \mu(x) + \mu(y) + \lambda(x,y). \qquad (*)$$

Notice that since $+1 = -1$ the distinction between symmetric and skew-symmetric forms has disappeared. By $(*)$ applied to $x+x$ we have that $\lambda(x,x) = 0$ for all x. I claim that the symmetric part λ of the given quadratic form is in fact a hyperbolic symmetric form, in fact it is a direct sum of elementary hyperbolic forms each of which has matrix

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

on a two-dimensional subspace. We prove this inductively, so let $x \in V$ be any element, and let $y \in V$ be an element such that $\lambda(x,y) = 1$, $\lambda(v,y) = 0$ for all $v \in V \setminus \{x\}$. (There is such a y since the map $x \mapsto 1$, $V \setminus \{x\} \mapsto 0$ is linear over F.) Then the subspace H spanned by x and y has $H \cap H^\perp = 0$, so that $V = H \oplus H^\perp$, and we have split off an elementary hyperbolic subspace. The assertion follows by induction.

Even though the *symmetric* structure of our form is now revealed to be hyperbolic, its *quadratic* structure need not be so. In fact, suppose now that x and y span an elementary hyperbolic subspace for λ. The identity $(*)$ easily shows that, of the three numbers $\mu(x)$, $\mu(y)$, $\mu(x+y)$, either all three are 1 (call this case $\mathcal{H}_1$) or two are 0 and the third is 1 (call this case $\mathcal{H}_0$). We have proved therefore that our given quadratic form is isomorphic to a direct sum of copies of $\mathcal{H}_0$ and $\mathcal{H}_1$.

There is however a relation to be taken into account: in fact, $\mathcal{H}_0 \oplus \mathcal{H}_0 \equiv \mathcal{H}_1 \oplus \mathcal{H}_1$. This can be proved by writing down an explicit isomorphism. In fact, if $\{x_1, y_1, x_2, y_2\}$ (with the obvious notation) is a basis for $\mathcal{H}_0 \oplus \mathcal{H}_0$, then the basis $\{x_1+y_1+x_2, x_1+y_1+y_2, x_1+x_2+y_2, y_1+x_2+y_2\}$ has $\mu = 1$ on each

---

[2] A quadratic form is said to be nonsingular if its symmetric part is nonsingular.

element, so exhibits an isomorphism with $\mathcal{H}_1 \oplus \mathcal{H}_1$. Thus we conclude that our form is in fact isomorphic to the direct sum of a number of copies of $\mathcal{H}_0$ together with at most one copy of $\mathcal{H}_1$.

This is as far as we can go: $\bigoplus^n \mathcal{H}_0$ is not isomorphic to $\bigoplus^{n-1} \mathcal{H}_0 \oplus \mathcal{H}_1$, because one can count the number of elements of the vector space on which μ is nonzero, and this number is greater in the second case. Thus we have a complete classification of quadratic forms (on finite-dimensional vector spaces) over $\mathbb{Z}/2$. If an $\mathcal{H}_1$ factor appears we say that the form has *Arf invariant* 1; otherwise it has Arf invariant 0. Notice that the Arf invariant (considered as a member of $\mathbb{Z}/2$) is additive on direct sums.

Since a symmetric form on V is by definition an element of $Q^\varepsilon(V)$, the reader may have expected us to define a quadratic form as an element of $Q_\varepsilon(V)$. Such a definition is possible for reasonable spaces V. In fact, for any V an element $\psi \in Q_\varepsilon(V)$ gives rise to a quadratic form (λ, μ), defined by

$$\lambda = (1 + T_\varepsilon)\psi, \quad \mu(x) = \psi(x)(x).$$

The converse is true if the module V is finitely generated and free. For in this case we may choose a basis $\{x_1, \dots, x_m\}$ for V. Now let (λ, μ) be a quadratic form on V and define $\lambda_{ij} = \lambda(x_i, x_j)$ for $i < j$, and $\mu_i = \mu(x_i)$. Now define a sesquilinear form ψ on V by

$$\psi\left(\sum a_i x_i, \sum b_j x_j\right) = \sum_{i<j} b_j \lambda_{ij} a_i^* + \sum_i b_i \mu_i a_i^*.$$

Then this form (or rather, its equivalence class in $Q_\varepsilon(V)$) maps to (λ, μ) under the construction described above.

EXERCISE: By embedding in a free module, extend the above equivalence to quadratic forms over finitely generated projective R-modules.

4. Definition of the L-groups

Suppose that V is an R-module equipped with a nonsingular bilinear form λ. Then, for any submodule $U \leq V$ there is defined the *orthogonal* $U^\perp$ in the natural way, namely

$$U^\perp = \{y \in V : \lambda(x, y) = 0 \forall x \in U\}.$$

In fact, we have already made use of this notion for vector spaces in our discussion of the Arf invariant. Notice that (in contrast to the familiar situation of orthogonal complements relative to a positive-definite inner product) it is possible for U and $U^\perp$ to intersect non-trivially.

The submodule U is said to be *complemented* in V if there is another submodule U' such that $U \oplus U' = V$.

7.14. LEMMA. *If U is complemented so is $U^\perp$.*

PROOF. Exercise. □

7.15. DEFINITION. Let V be an R-module equipped with a nonsingular ε-quadratic form (λ, μ). Let U be a submodule of V.

(i) U is called a *sublagrangian* if $U \subseteq U^\perp$, U is complemented, and $\mu_{|U} = 0$.
(ii) U is called a *lagrangian* if it is a sublagrangian and, in addition, $U = U^\perp$.

If V is a finite-dimensional vector space over a field R of characteristic not 2, then quadratic and symmetric forms coincide, and moreover all submodules are complemented. A sublagrangian is then what is usually called an 'isotropic subspace', and a lagrangian is a 'maximal isotropic subspace'.

7.16. EXAMPLE. Let V be any R-module, and consider the hyperbolic ε-quadratic form on $W = V \oplus V'$. Then V and V' are complementary lagrangians.

In fact, this is the only example which can occur.

7.17. THEOREM. (WITT) *If a nonsingular ε-quadratic form (V, λ, μ) on a finitely generated free R-module V admits a lagrangian U, then it is isomorphic to the hyperbolic form $\mathcal{H}_\varepsilon(U)$.*

PROOF. We note first that if we can find another lagrangian U' such that $V = U \oplus U'$, then λ will identify U' with the dual of U (as the notation suggests) and μ will be determined by the fact that it is zero on U and on U', so that it will follow that V is hyperbolic. We therefore aim to find a complementary lagrangian to U. The idea is to choose any complementary subspace and then modify it, à la Gram-Schmidt, so that it becomes lagrangian.

Let $i\colon U \to V$ be the inclusion. Dualizing and composing with the isomorphism $\lambda\colon V \to V'$ we get a map $j\colon V \to U'$. By definition, $\operatorname{Ker} j = U^\perp$, which equals U since U is lagrangian. Since a lagrangian is assumed to be complemented, j is a split surjection, and there is a $1:1$ correspondence between splittings $\theta\colon U' \to V$ of j and complementary submodules to U in V. Fix such a splitting θ; then any other splitting is of the form $\theta + \varphi$, where φ is a homomorphism from U' to $\operatorname{Ker} j = U$.

Now (using the fact that V is free) choose a sesquilinear form ψ on V such that $[\psi] \in Q_\varepsilon(V)$ represents the quadratic form (λ, μ). We would like to choose our splitting $\theta + \varphi$ so that it corresponds to a complementary lagrangian to U, which is to say that $(\theta + \varphi)'\psi(\theta + \varphi) = 0$. However we may compute

$$(\theta + \varphi)'\psi(\theta + \varphi) = \theta'\psi\theta + \varphi \in \operatorname{Hom}(U', U).$$

Thus we can achieve what we want by choosing $\varphi = -\theta'\psi\theta$. □

7.18. REMARK. By an extension of this argument we may also prove that if U is a *sub*lagrangian in V, then there is an isomorphism of quadratic forms $V \cong \mathcal{H}_\varepsilon(U) \oplus U^\perp/U$.

7.19. COROLLARY. *Let (V, λ, μ) be a quadratic form on a finitely generated free R-module; then $(V, \lambda, \mu) \oplus (V, -\lambda, -\mu)$ is isomorphic to a hyperbolic form.*

PROOF. The diagonally embedded copy of V is lagrangian. □

We may now define the even-dimensional L-groups. Let $\varepsilon = (-1)^n$. We define a group $L_{2n}(R)$ as follows: form a semigroup with one generator for each isomorphism class of ε-quadratic forms on f.g. free R-modules, with addition by direct sum, and with the imposed relation that every hyperbolic form represents zero (in other words, we take the free semigroup as described above, and divide by the subsemigroup generated by hyperbolic forms). By the corollary above, the quotient semigroup so defined is in fact a group; the inverse of (V, λ, μ) being $(V, -\lambda, -\mu)$.

7.20. DEFINITION. This quotient group is denoted $L_{2n}(R)$.

We observe that $L_{2n}(R)$ is a covariant functor of R (by "change of rings").

7.21. EXAMPLE. Here are the L-groups for the three fields $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}/2$ that we previously considered.

R	$L_0(R)$	$L_2(R)$
$\mathbb{R}$	$\mathbb{Z}$	0
$\mathbb{C}$	$\mathbb{Z}$	$\mathbb{Z}$
$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}/2$

These follow from the classification of symmetric and skew-symmetric forms over these fields, which we previously discussed. The isomorphisms $L_0(\mathbb{R}) \to \mathbb{Z}$, $L_0(\mathbb{C}) \to \mathbb{Z}$, and $L_2(\mathbb{C}) \to \mathbb{Z}$ are given by the signature. The isomorphisms $L_0(\mathbb{Z}/2) \to \mathbb{Z}/2$ and $L_2(\mathbb{Z}/2) \to \mathbb{Z}/2$ are given by the Arf invariant

7.22. EXAMPLE. An example of some interest to analysts occurs when the ring R is a complex C^*-algebra A. In this case there is an identification between $L_0(A) = L_2(A)$ and the topological K-theory of the algebra A. We will assume that the reader is familiar with K-theory for C^*-algebras.

We consider, then, symmetric forms on free A-modules; in fact, without essential loss of generality, it is enough to consider symmetric forms on A itself. Such a form is given by

$$\lambda(x, y) = yTx^*$$

for some self-adjoint $T \in A$. If the form is nondegenerate, then T is invertible, that is, the spectrum $\sigma(T)$ of T does not contain zero.

Choose a continuous and bounded function $f\colon \mathbb{R} \to \mathbb{R}$ which is equal to zero on $\mathbb{R}^-$ and equal to one on $\mathbb{R}^+ \cap \sigma(T)$. The operator $e_+(T) = f(T)$ defined by the functional calculus does not depend on the choice of function f; it is called the *positive spectral projection* of T. Similarly we may define the *negative spectral projection* $e_-(T)$ and we note that $e_-(T) + e_+(T) = 1$. Now define a new symmetric form λ' by

$$\lambda'(x, y) = y(-e_-(T) + e_+(T))x^*.$$

I claim that λ' and λ are isomorphic as forms; indeed, the spectral theorem provides an invertible operator $S = |T|^{1/2}$ such that $\lambda'(Sx, Sy) = \lambda(x, y)$. We conclude that, over a C^*-algebra A, any nondegenerate form is isomorphic to one arising as the 'difference' of two projections.

The *analytic signature* of λ is the class $[e_+] - [e_-]$ in $K_0(A)$. The analytic signature of a hyperbolic form is zero, so we get a map $L_0(A) \to K_0(A)$. The discussion above shows that this map is almost an isomorphism from L-theory to K-theory. The reason it isn't exactly an isomorphism is that the two projections are related by the requirement that their sum represent a free module. If we use the variant definition L^p of L-theory, made out of quadratic forms on f.g. *projective* modules, then the analytic signature gives an isomorphism $L_0^p(A) \to K_0(A)$ for any unital C^*-algebra A.

We can descripe our original L-theory $L_0(A) = L_0^h(A)$ in terms of K-theory as well. In fact, the map which sends the form to the pair (e_-, e_+) gives an isomorphism between $L_0(A)$ and the group G which consists of pairs $(x, y) \in K_0(A) \times K_0(A)$ such that $x + y$ vanishes in reduced K-theory $\tilde{K}_0(A)$, modulo the subgroup which is the image of the diagonal embedding $\mathbb{C} \to A \oplus A$. This group only differs from $K_0(A)$ up to 2-torsion; this is an example of a Ranicki-Rothenberg sequence (see later).

5. Calculation of $L_{2n}(\mathbb{Z})$.

The main point of L-theory is that it should be applicable to group rings, and the alert reader will have noticed that we have not computed the L-theory groups of a single group ring so far. In this section we will make some amends by calculating the simplest example, the L-theory of $\mathbb{Z}$ (which is of course the group ring of the trivial group). Even in this case, some substantial input from number theory is required.

Let $\alpha\colon \mathbb{Z} \to \mathbb{R}$ and $\beta\colon \mathbb{Z} \to \mathbb{Z}/2$ denote the obvious homomorphisms. We aim to prove the following two results.

L0Z 7.23. PROPOSITION. $\alpha_*\colon L_0(\mathbb{Z}) \to L_0(\mathbb{R}) = \mathbb{Z}$ *is injective, and its image is* $8\mathbb{Z}$.

L2Z 7.24. PROPOSITION. $\beta_*\colon L_2(\mathbb{Z}) \to L_2(\mathbb{Z}/2) = \mathbb{Z}/2$ *is an isomorphism.*

We begin with the symmetric case ($n = 0$). Here our task is to classify *even* nonsingular symmetric forms (= nonsingular quadratic forms) on finitely generated free abelian groups. It turns out to be helpful to begin with a 'stable' classification of *all* nonsingular symmetric forms.

7.25. DEFINITION. We define $\mathfrak{K}$ to be the Grothendieck group constructed from the semigroup of isomorphism classes of nonsingular symmetric forms on f.g. free abelian groups.

Let I_+ and I_- denote the rank 1 forms $\lambda(x, x') = xx'$ and $\lambda(x, x') = -xx'$.

7.26. PROPOSITION. *The group $\mathfrak{K}$ is free abelian generated by $[I_+]$ and $[I_-]$.*

To begin the proof we need a fact from number theory. A symmetric form λ on a $\mathbb{Z}$-module V is said to *represent zero* if there is some nonzero $v \in V$ such that $\lambda(v) = 0$.

7.27. LEMMA. *A nondegenerate symmetric form over $\mathbb{Z}$ represents zero over $\mathbb{Z}$ if and only if it represents zero over $\mathbb{R}$, that is, if and only if it is indefinite.*

This follows from the Hasse-Minkowski theorem according to which a rational quadratic form represents zero over $\mathbb{Q}$ if and only if it represents zero over $\mathbb{R}$ and over every p-adic completion $\mathbb{Q}_p$ of $\mathbb{Q}$. See Serre, *A Course in Arithmetic*.

PROOF. It will suffice to prove that any odd indefinite form is isomorphic to a direct sum of copies of I_+ and I_-. For certainly any form is can be made odd and indefinite by adding $I_+ \oplus I_-$, so this will prove that $[I_+]$ and $[I_-]$ generate $\mathfrak{K}$; on the other hand, the pair (rank,signature) gives a homomorphism $\mathfrak{K} \to \mathbb{Z} \times \mathbb{Z}$ under which the images of $[I_+]$ and $[I_-]$ are linearly independent, so $\mathfrak{K}$ must in fact be free on these classes.

We work by induction on the rank. Let λ be an odd indefinite form on a $\mathbb{Z}$-module V of rank n. By the lemma, above, there exists $x \in V$ such that $\lambda(x, x) = 0$. We may assume that x is *indivisible* (i.e. that it cannot be written kx' for any integer $k > 1$) and from this and the unimodularity of the form it follows that there exists $y \in V$ such that $\lambda(x, y) = 1$. Because λ is odd, a simple argument shows that we may choose such a y with $\lambda(y, y) = 2m + 1$ an odd number. Now let $x' = y - mx$, $y' = y - (m + 1)x$; then we have the following table of values for λ:

$$\begin{array}{ccc} & x' & y' \\ x' & 1 & 0 \\ y' & 0 & -1 \end{array}$$

and hence $V = I_+ \oplus I_- \oplus W$, where W is a module of rank $n - 2$. Now one of $I_+ \oplus W$, $I_- \oplus W$ is odd indefinite and of rank $n - 1$, so the induction may proceed. □

Now we can prove something we have already asserted and used, that the signature maps $L_0(\mathbb{Z})$ to $8\mathbb{Z}$.

7.28. PROPOSITION. *The signature of an even symmetric form (that is, a quadratic form) over $\mathbb{Z}$ is a multiple of 8.*

PROOF. We have observed that the signature gives a homomorphism $\mathfrak{K} \to \mathbb{Z}$. We will now define a related map $\sigma\colon \mathfrak{K} \to \mathbb{Z}/8$. Given a symmetric form λ on a free $\mathbb{Z}$-module V, let $\bar{\lambda}$ be the associated form on the vector space $\bar{V} = V \otimes \mathbb{Z}/2 = V/2V$ over $\mathbb{Z}/2$. On $\bar{V}$ the functional $\xi \mapsto \bar{\lambda}(\xi, \xi)$ is *linear* (the cross-terms vanish because $2 = 0$) and hence, by duality, there is a canonical element $\zeta \in \bar{V}$ such that

$$\lambda(\zeta, \xi) = \lambda(\xi, \xi)$$

for all $\xi \in \bar{V}$. Let $z \in V$ be a lift of ζ; it is unique modulo $2V$. Then $\lambda(z,z) \in \mathbb{Z}$ is well-defined modulo 8, since

$$\lambda(z+2x, z+2x) = \lambda(z,z) + 4(\lambda(x,z) + \lambda(x,x))$$

and $\lambda(x,x)$ agrees modulo 2 with $\lambda(z,x)$. This residue class modulo 8 is the invariant σ of the form (V,λ).

I now claim that σ is exactly the reduction of the signature modulo 8. Since σ and the signature both give homomorphisms on $\mathfrak{K}$, it suffices to check this assertion on the generators $[I_+]$ and $[I_-]$ of $\mathfrak{K}$, and there it is easy. But now, if λ is an *even* form, then $\bar{\lambda}(\xi,\xi)$ (in the notation above) vanishes identically on $\bar{V}$, so $\zeta = 0$ and we may take $z = 0$, whence $\sigma = 0$. The conclusion follows. □

Now we can do the computation of $L_0(\mathbb{Z})$. Let us first prove

7.29. LEMMA. *Every (nonzero) class in $L_0(\mathbb{Z})$ can be represented by a definite form.*

PROOF. It suffices to show that a hyperbolic summand can be split off from any even indefinite form. Let (V,λ) be such a form, let $x \in V$ be indivisible with $\lambda(x,x) = 0$ and let $y \in V$ have $\lambda(x,y) = 1$. Then $\lambda(y,y) = 2m$ for some m, and $x' = x$ and $y' = y - mx$ span a hyperbolic summand in V. □

This proves that the signature homomorphism $L_0(\mathbb{Z}) \to \mathbb{Z}$ is injective (since the signature of a definite form is equal to plus or minus its rank). On the other hand, we have seen that the image of this homomorphism is contained in $8\mathbb{Z}$; and the existence of the even definite form E_8 of rank 8 shows that the image is actually equal to $8\mathbb{Z}$. This completes the proof of 7.23.

The proof of 7.24 is easier. In fact, given a skew-quadratic form (λ,μ) over $\mathbb{Z}$, its symmetric part λ is easily seen simply to be a direct sum of hyperbolics. The quadratic part μ is a function to $Q_-(\mathbb{Z}) = \mathbb{Z}/2$ and the same discussion applies as over $\mathbb{Z}/2$ to show that there are essentially only two distinct possibilities for μ on each hyperbolic summand. Following through as in the previous discussion we find that the L-theory class of the form is entirely determined by that of its reduction modulo 2, and this is 7.24.

CHAPTER 8

The algebraic obstruction to surgery

We will show that the algebraically defined groups $L_{2n}(\mathbb{Z}\pi)$ which were introduced in the previous chapter exactly measure the obstruction to surgering a normal map to a homotopy equivalence. This is the main result of Wall's work on surgery. We begin by discussing the relevance of the μ-invariant to embeddings.

1. Deforming immersions to embeddings

Move some of this to chapter 3 — To do

Recall that in the previous chapter we have defined an invariant $\mu(f) \in Q_\varepsilon(\mathbb{Z}\Gamma)$ associated to any Γ-trivialized immersion of a manifold V^n into M^{2n}, where $\pi_1 M = \Gamma$, $\varepsilon = (-1)^n$. This invariant is the 'count' of the double points of the immersion, taking into account the ambiguities of sign that arise. To show that this is interesting it is necessary to show that $\mu(f)$ is invariant under the natural equivalence relation on such immersions, that is regular homotopy (3.13).

In order to clarify the proofs let us introduce some terminology. Let $f\colon V \to M$ be an immersion with clean double points. By an *ordered double point* $\mathfrak{p}$ of f we mean an ordered pair (v, v') of points of v such that $f(v) = f(v')$; in other words, an ordered double point is a double point together with a particular choice of which branch of $f(V)$ is 'first' near that double point. To an ordered double point $\mathfrak{p}$ is then associated an intersection number $k(\mathfrak{p}) \in \{\pm 1\}$ and a loop $\ell(\mathfrak{p}) \in \Gamma$; $\ell(\mathfrak{p})$ is defined as the loop that runs from the basepoint of M to the basepoint of $f(V)$ as required by then Γ-trivialization, then in V to v avoiding other double points, then from v' back to the basepoint. The element $\mu(f) \in Q_\varepsilon(\mathbb{Z}\Gamma)$ is then defined by making an arbitrary choice of ordering at each double point and setting

$$\mu(f) = \sum_{\mathfrak{p}} k(\mathfrak{p})\,\ell(\mathfrak{p}).$$

8.1. PROPOSITION. *If $f_0, f_1\colon V \to M$ are regularly homotopic immersions, then $\mu(f_0) = \mu(f_1)$.*

PROOF. Recall (3.14) that we may assume without loss of generality that the regular homotopy f_t is of generic form. This means that the immersions f_t all have clean double points except for a finite number of values of t, at which there occur birth or death singularities, where two double points appear or disappear together. Moreover, the track of each double point from birth to death is a continuous function of t. It is clear then that $\mu(f_t)$ will be continuous (hence constant) in t everywhere except possibly at the birth/death singularities. At such a singularity, two double points appear or disappear together. When the two double points are very close, it is clear that they can be ordered as $\mathfrak{p}_1$, $\mathfrak{p}_2$ so that $\ell(\mathfrak{p}_1) = \ell(\mathfrak{p}_2)$ and $k(\mathfrak{p}_1) = -k(\mathfrak{p}_2)$. Therefore the contributions from the two double points cancel out and so $\mu(f_t)$ is continuous even at the singularities. $\square$

As a consequence of this, if the immersion f is regularly homotopic to an embedding, then $\mu(f) = 0$. Generalizing the proof of the hard Whitney theorem, Wall showed that the converse is also true:

8.2. THEOREM. (WALL EMBEDDING THEOREM) *Suppose that* $f\colon V^n \to M^{2n}$ *is an immersion as above,* $2n \geqslant 6$*, and* $\mu(f) = 0$*. Then* f *is regularly homotopic to an embedding.*

PROOF. This uses the same device as in the proof of 3.12. Let us introduce some terminology. Suppose that $\mu(f) = 0$. Then the double points of f can be ordered and arranged in pairs $(\mathfrak{p}_i, \mathfrak{q}_i)$ having $\ell(\mathfrak{p}_i) = \ell(\mathfrak{q}_i)$ and $k(\mathfrak{p}_i) = -k(\mathfrak{q}_i)$. Hence for each such pair we can find a loop γ_i in V running from $\mathfrak{p}_i$ to $\mathfrak{q}_i$ along one branch of the immersion, and back along the other branch, with the signs of the intersections opposite at the two double points. Because of the dimension condition, we can use the Whitney trick to construct a regular homotopy which cancels the two double points. □

The original Whitney theorem, that any $f\colon V^n \to M^{2n}$, with M simply-connected and high-dimensional, is homotopic to an embedding, may be derived as a corollary of Wall's theorem. For in this case f is certainly Γ-trivialized, and the obstruction to regular homotopy to an embedding is an element of the group $Q_\varepsilon(\mathbb{Z})$, which is $\mathbb{Z}$ or $\mathbb{Z}/2$ according to the sign of ε. However, this obstruction may be varied at will by taking the connected sum with a "figure eight" immersion, which does not change the homotopy class (even though it does change the *regular* homotopy class).

Recall now that, in order to do surgery on a sphere in a manifold M equipped with a normal map $f\colon M \to X$, that sphere must be not merely embedded, but *framed*. That is, its normal bundle must be trivialized; and moreover, this trivialization of the normal bundle must be compatible with the stable trivialization determined by the diagram

$$\begin{tikzcd} S^{m-1} \arrow[r] \arrow[d] & M \arrow[d, "f"] \\ D^m \arrow[r] & X \end{tikzcd}$$

Recall that the stable normal bundle of M is pulled back from X (by definition of a normal map); since D^m is contractible, this implies that the stable normal bundle of M is trivialized over S^m, hence that the tangent bundle of M is stably trivialized there. Since the tangent bundle of S^{m-1} is (canonically) stably trivial, this gives a stable trivialization of the normal bundle of S^{m-1} in M. In order for it to be possible to perform surgery it is necessary that this stable trivialization of the normal bundle should be reducible to an actual trivialization.

In the case that we previously discussed, of surgery below the middle dimension, this is always possible for dimensional reasons. Now, however, we need to discuss surgery at the middle dimension (i.e. on n-spheres in a $2n$-manifold) and here there may *a priori* be an obstruction to destabilizing the framing of an embedding. We need to understand how this obstruction is related to the μ-invariant.

Consider immersions i of S^n into a manifold M^{2n} equipped with a normal map to X. In these circumstances the obstruction to destabilizing the framing of i is an element $\beta(i)$ of the relative homotopy group $\pi_{n+1}(BO, BO(n))$. What is this group?

[The identification of this group in Prop 6.5, as circulated, is wrong. Sorry.]

qhom 8.3. LEMMA. *There is a canonical identification of* $\pi_{n+1}(BO, BO(n))$ *with* $Q_\varepsilon(\mathbb{Z})$*,* $\varepsilon = (-1)^n$.

PROOF. (Stiefel manifolds etc) □

The boundary map $\pi_{n+1}(BO, BO(n)) \to \pi_n(BO(n))$ maps $\beta(i)$ to the absolute framing obstruction of the normal bundle (which is in fact the Euler class). This is the obstruction to framing the normal bundle *at all* (irrespective of whether this framing can be accomplished in a compatible manner).

Consider now the group ring $\mathbb{Z}\pi$. We note that there is a direct sum decomposition $\mathbb{Z}\pi = \mathbb{Z} \oplus \widetilde{\mathbb{Z}\pi}$, where $\widetilde{\mathbb{Z}\pi}$ is the set of elements in $\mathbb{Z}\pi$ such that the coefficient of the identity element of π is zero. (Note: This is a direct sum of additive groups with involution, not of rings!) There is a corresponding direct sum decomposition of Q-groups. For an immersion i we write

$$\mu(i) = (\mu_{\mathbb{Z}}(i), \tilde{\mu}(i)) \in \mathbb{Q}_\varepsilon(\mathbb{Z}) \oplus Q_\varepsilon(\widetilde{\mathbb{Z}\pi}),$$

and call these the *integral* and *reduced* components of the μ-invariant of i.

8.4. PROPOSITION. *Let $i_0, i_1 \colon S^n \to M^{2n}$ be homotopic immersions of a sphere into a manifold equipped with a normal map to X. Then*

(i) *The reduced components of their μ-invariants agree: $\tilde{\mu}(i_0) = \tilde{\mu}(i_1)$.*

(ii) *The difference of the integral components of their μ-invariants is equal to the difference of the framing destabilization obstructions:*

$$\mu_{\mathbb{Z}}(i_0) - \mu_{\mathbb{Z}}(i_1) = \beta(i_0) - \beta(i_1)$$

under the identification of lemma 8.3.

(iii) *For a fixed i_0, all possible values of $\mu_{\mathbb{Z}}$ can be realized by homotopic immersions i_1.*

OMITTED. □

This proposition allows us to define a refined μ-invariant for relative homotopy classes $x \in \pi_{n+1}(f)$. Namely, given such a homotopy class we deform it using general position so that it is represented by an immersion i with clean double points; we then define $\mu(x) = \mu(i) - \beta(i)$. The proposition above shows that this new μ-invariant depends only on the homotopy class of x, so it is well-defined. Moreover we have:

8.5. COROLLARY. *Let $x \in \pi_{n+1}(f)$ as above, where $f \colon M^{2n} \to X$ is a normal map. If $\mu(x) = 0$, then x can be represented by a unique regular homotopy class of embeddings $S^n \to X$ which are framed compatibly with the bundle data. Hence we can do surgery on x.*

PROOF. This is immediate from Wall's embedding theorem together with the above. □

2. The kernel form

More details are needed here

To do

We now consider the following situation. Suppose that $f \colon M^{2n} \to X$ is a normal map which is highly connected, meaning that $\pi_i(f) = 0$ for $i \leqslant n$. As we have seen, any normal map can be surgered to a highly connected one. The homology kernel $K_n(f)$ is then isomorphic to $\pi_{n+1}(f)$, by the relative Hurewicz theorem. We have seen 7.5 that $K_n(f)$ is a finitely generated stably free $\mathbb{Z}\pi$-module. Let us assume that $K_n(f)$ is in fact free; as we shall shortly see, we can always bring this about by doing some additional trivial surgeries which will have the effect of adding a free module to the homology kernel. An element $x \in K_n(f)$ is then represented by a homotopy class of π-trivialized maps (we may assume that they are immersions) $S^n \to M$ which become inessential in X.

On $K_n(f)$ two functions λ and μ are defined as follows:

(a) For $x, y \in K_n(f)$, $\lambda(x,y) \in \mathbb{Z}\pi$ is the $\mathbb{Z}\pi$ intersection number of the homology classes x and y, or equivalently the geometric count of the intersections of representatives of x and y by π-trivialized immersed spheres in general position;
(b) For $x \in K_n(f)$, $\mu(x) \in Q_\varepsilon(\mathbb{Z}\pi)$ is the refined μ-invariant of x defined in the previous section, that is, for an immersion i with clean double points representing x it is $\mu(i) - \beta(i)$, where β is the framing destabilization obstruction.

8.6. PROPOSITION. *The functions λ and μ defined above give a quadratic form on the free $\mathbb{Z}\pi$-module $K_n(f)$.*

PROOF. □

8.7. DEFINITION. The equivalence class of $(K_n(f), \lambda, \mu)$ in the group $L_{2n}(\mathbb{Z}\pi)$ is the *surgery obstruction* $\sigma(f)$ of the highly connected normal map f.

We propose to show that $\sigma(f)$ is invariant under surgery (so that it depends only on the normal bordism class of f) and that $\sigma(f) = 0$ if and only if f can be surgered to a homotopy equivalence. As a first step we investigate the effect of some simple surgeries.

8.8. DEFINITION. The hyperbolic form $H_\varepsilon(\mathbb{Z}\pi)$ will be called a *standard plane*.

Explicitly a standard plane is a free $\mathbb{Z}\pi$-module of rank[1] two, with basis $\{e_1, e_2\}$ say, and the quadratic form (λ, μ) defined by $\mu(e_i) = 0$, $\lambda(e_i, e_i) = 0$, $\lambda(e_1, e_2) = 1$. Any hyperbolic form on a free module is a direct sum of standard planes.

SN1 8.9. LEMMA. *Let $f\colon M^{2n} \to X$ be a highly connected normal map, and suppose that we perform surgery on some (nullhomotopic) embedded S^{n-1}. Then the result of the surgery is $f'\colon M' \to X$, such that f' is still highly connected, M' is diffeomorphic to the connected sum $M\#(S^n \times S^n)$, and $K_n(f')$ is the direct sum of $K_n(f)$ and a standard plane.*

PROOF. □

Notice in particular that this surgery therefore does not change the class $\sigma(f) \in L_{2n}(\mathbb{Z}\pi)$. Moreover, these trivial surgeries add free modules to the kernel, so they may be used to replace the original 'stably free' kernel by a free kernel, if required.

A more significant result is that this operation is reversible: *every* hyperbolic summand in $K_n(f)$ arises from trivial $(n-1)$-surgeries in the way just described.

SN2 8.10. PROPOSITION. *Let $f\colon M^{2n} \to X$ be a highly connected normal map, and suppose that $K_n(f) = U \oplus P$ (as quadratic forms), where P is a standard plane spanned by $\{e_1, e_2\}$. Then*
(i) *We can perform surgery on e_1;*
(ii) *The result of the surgery is a highly connected normal map $f'\colon M' \to X$ with $K_n(f') = U$;*
(iii) *$M = M'\#(S^n \times S^n)$;*

PROOF. Because $\mu(e_1) = 0$, the class e_1 is represented by an embedding $S^n \to M$ which is framed compatibly with the bundle data (by the results of the preceding section). We can use this embedding to perform surgery on e_1.

The effect of this surgery is to produce an elementary normal cobordism W from M to a new manifold M'; W is obtained by attaching a $(n+1)$-handle to $M \times [0,1]$ along the framed embedding $e_1\colon S^n \to M \times \{1\}$. The key assertion to prove is that the new

[1] The notion of *rank* is always well-defined for free modules over integral group rings.

normal map $f'\colon M' \to X$ obtained by surgery is still highly connected. For this purpose we consider the exact homology sequence of the pair (W, M'), as follows:

$$\to H_n(W) \to H_n(W, M') \to H_{n-1}(M') \to H_{n-1}(W) \to H_{n-1}(W, M') \to .$$

The cobordism W is obtained from M by adding an $(n+1)$-handle, or from M' by adding an n-handle. It follows that for $i \geqslant 1$ $\pi_{n-i}(f) = \pi_{n-i}(f')$, so that f' is certainly $(n-1)$-connected. In order that it should be n-connected it need only induce an isomorphism on H_{n-1}, by the Hurewicz theorem. But the map $W \to X \times [0,1]$ given by the normal cobordism is certainly n-connected, since W is n-equivalent to M. We see therefore that what we must prove is that the map $H_{n-1}(M') \to H_{n-1}(W)$ appearing in the exact sequence above is an isomorphism.

We have

$$H_i(W, M') = \begin{cases} \mathbb{Z}\pi & (i = n) \\ 0 & (i \neq n) \end{cases}$$

since W is obtained from M' by adding a single n-handle. To get the isomorphism it will therefore suffice to prove that the map $H_n(W) \to H_n(W, M') = \mathbb{Z}\pi$, appearing earlier in the exact sequence, is surjective. In fact we shall prove more: the composite

$$\alpha\colon H_n(M) \to H_n(W) \to H_n(W, M'') = \mathbb{Z}\pi$$

is surjective. Indeed, α has a direct geometrical description: it sends a class $x \in H_n(M)$ to the intersection number $\lambda(x, e_1) \in \mathbb{Z}\pi$. But $H_n(M)$ contains the class e_2, whose intersection number with e_1 is ± 1; hence α is surjective.

Once we have established that f' is highly connected the rest of the proof is easy. For (turning the cobordism W upside down) M is obtained from M' by doing surgery on an embedded S^{n-1}, which must be trivial. Thus lemma 8.9 applies: M is the connected sum of M' and $S^n \times S^n$. Hence the kernel form of M is the kernel form of M' plus a standard plane. Some simple identifications yield the result as stated. □

We will now prove that the surgery obstruction is well-defined.

8.11. PROPOSITION. *Let $f\colon M \to X$ and $f'\colon M' \to X$ be two highly connected normal maps which are normally cobordant. Then $\sigma(f) = \sigma(f') \in L_{2n}(\mathbb{Z}\pi)$.*

PROOF. Suppose that W is a normal bordism. We can consider W as a rel ∂ surgery problem whose ends, if not exactly homotopy equivalences, are at least n-connected; using surgery rel ∂ below the middle dimension, we may replace W by another normal bordism from M to M' which is itself n-connected. Let us continue to use the letter W for this new, highly connected, normal bordism[2]. Thus M, M', and W are all n-equivalent, and in terms of handle decompositions this means that W is obtained from M by adding some n-handles and some $(n+1)$-handles. By handle rearrangement (page 67 [November 19, 2004]) we can assume that all the n-handles are attached first (thereby obtaining a cobordism W_0 from M to some new manifold M'') and then all the $(n+1)$-handles are attached (obtaining a cobordism W_1 from M'' to M'). But now M'' is obtained from M by doing surgery on (trivial) S^{n-1}'s, so (by lemma 8.9) its kernel form is obtained from the kernel form of M by adding hyperbolics. By turning W_1 upside down we see also that M'' is obtained from M' by doing surgery on trivial S^{n-1}'s, so we deduce that the kernel form of M' and the kernel form of M are equivalent modulo the addition of hyperbolics. This is the required result. □

[2]To rephrase this argument: we have proved that if two highly connected normal maps are bordant at all, then they are 'highly connected bordant'.

A corollary is that we can define the surgery obstruction of *any* normal map f by first performing surgery to make f highly connected, and then taking the surgery obstruction of the resulting highly connected map. For any two ways of surgering f to be highly connected will themselves be related by a normal cobordism.

Notice that our definition of the surgery obstruction of a normal map involves a two-stage process: the first stage being geometric (to turn our normal map into a highly connected normal map) and the second algebraic (the L-theory classification of quadratic forms). It would be quite convenient to have a direct algebraic definition of the surgery obstruction which did not require preliminary geometric surgery below the middle dimension. Such a definition is provided by Ranicki's *algebraic surgery theory*. Crudely speaking, Ranicki regards the L-groups as being generated not by quadratic forms but by 'chain complexes equipped with a quadratic Poincaré duality'; a quadratic form just is such a chain complex which happens to be concentrated in the middle dimension. A normal map defines such a chain complex (without any preliminary surgery) and the L-theory class of this chain complex is the surgery obstruction.

8.12. THEOREM. (WALL) *A normal map $f\colon M^{2n} \to X$ can be surgered to a homotopy equivalence if and only if $\sigma(f) = 0 \in L_{2n}(\mathbb{Z}\pi)$.*

PROOF. The surgery obstruction is invariant under surgery (as we now know), and the obstruction of an equivalence is zero, so one direction of this proposition is now obvious. Consider the opposite direction: suppose that the surgery obstruction vanishes. This means that we can add standard planes to $K_n(f)$ so that the resulting sum $K_n(f)\oplus$standard planes is itself hyperbolic, i.e. a sum of standard planes. By 8.9 we can accomplish the first stage, that is the addition of standard planes to the kernel, geometrically, by performing surgery on trivial $(n-1)$-spheres. By repeatedly applying 8.10 we see that, once we have surgered so that $K_n(f)$ is hyperbolic, we can make a succession of further surgeries on n-spheres so that $K_n(f)$ becomes trivial. But then the resulting manifold admits an $(n+1)$-connected normal map to X, and by Poincaré duality and the Hurewicz theorem this map must be a homotopy equivalence. □

3. Realization and the surgery exact sequence

In the preceding section we have seen that the even-dimensional L-groups $L_{2n}(\mathbb{Z}\pi)$ measure exactly the obstruction to doing surgery on normal maps of even-dimensional manifolds. A major achievement of Wall's book on surgery was to also give an algebraic definition of the odd-dimensional L-groups and to show that in the same way they measure the obstruction to odd-dimensional surgery. *I will not attempt to describe this theory here*, however I ask the reader to take on faith that there exist such things as the groups $L_m(\mathbb{Z}\pi)$, m odd, and that they satisfy analogues of the theorems of the previous section. To save a little space, we will from now on use the abbreviated notation $L_m(\pi)$ for $L_m(\mathbb{Z}\pi)$.

Wall's results proved in the previous section show that the groups $L_m(\mathbb{Z}\pi)$ are in a certain sense the right place to measure surgery obstructions. However one might still wonder whether they are too big; does every element of an L-group arise as the surgery obstruction of some surgery problem? The following theorem provides a strong positive answer.

realz

8.13. THEOREM. (REALIZATION THEOREM) *Let X^{m-1} be a compact Poincaré duality space with fundamental group π, and let $M \to X$ be a manifold structure on X. Let $\alpha \in L_m(\pi)$ be given. Then there exists a cobordism W equipped with a normal map*

$f\colon W \to X \times [0,1]$ *such that there is a diagram*

$$\begin{tikzcd} M = \partial_- W \arrow[r] \arrow[d, "f_-"'] & W \arrow[d, "f"'] & \partial_+ W \arrow[l] \arrow[d, "f_+"'] \\ X \times \{0\} \arrow[r] & X \times [0,1] & X \times \{1\} \arrow[l] \end{tikzcd}$$

with $f_-\colon M \to X$ *the given structure,* f_+ *a homotopy equivalence, and the surgery obstruction for* f *rel* ∂ *equal to* α.

PROOF. This is done by plumbing, exactly like 5.26.... □

More information about realization is needed, and examples

To do

We need a result about gluing cobordisms of this kind. Specifically, suppose that W_0 is a normal cobordism between M and M', and W_1 is a normal cobordism between M' and M'', where M, M', and M'' are manifold structures on X. Then $W = W_0 \cup_{M'} W_1$ is a normal cobordism between M and M''. How are the rel ∂ surgery obstructions related?

8.14. LEMMA. *In the above situation the rel* ∂ *surgery obstructions of* W_0, W_1, *and* W *are related by* $\sigma(W_0) + \sigma(W_1) = \sigma(W)$.

PROOF. The homology kernel of W decomposes as a direct sum. □

4. Defining L-theory by algebraic bordism

Andrew to write this section

To do

CHAPTER 9

Surgery and manifold structures

1. The geometric surgery exact sequence

We can now describe the fundamental computational tool, the surgery exact sequence. Notice that our results so far provide us with a way to decide, in principle, whether a Poincaré duality space X admits a CAT manifold[1] structure: we decide first whether its Spivak normal bundle can be reduced to CAT, and, if so, for each such CAT reduction we compute the surgery obstruction. If one of these surgery obstructions is zero, the space admits a manifold structure, and otherwise it doesn't. For further study and especially for investigation of the uniqueness of manifold structures, it is convenient to assume that X is already equipped[2] with one 'reference' structure $M \to X$. As we explained, the existence of this reference structure allows us to identify the normal invariant set $\mathcal{N}(X)$ with the set $[X, G/CAT]$ of homotopy classes of maps from X into G/CAT.

9.1. THEOREM. (SURGERY EXACT SEQUENCE, FIRST FORM) *Let X^m be as above, $\pi = \pi_1 X$, $m \geqslant 5$. Suppose also that* $\mathrm{Wh}(\pi) = 0$. *Then there is an "exact sequence"*

$$\ldots \to [\Sigma X, G/CAT] \to L_{m+1}(\pi) \to \mathcal{S}_{CAT}(X) \to [X, G/CAT] \to L_m(\pi).$$

The exact meaning of 'exactness' will be explained in the proof. The Σ denotes suspension.

PROOF. What are the maps? The map $[X, G/CAT] \to L_m(\pi)$ is given by the surgery obstruction. So is the map $[\Sigma X, G/CAT] \to L_{m+1}(\pi)$; the left-hand side can be thought of as made up of bordism classes of degree one normal maps $W \to X \times [0,1]$ which restrict on each boundary component to the given structure of X, and such maps have a rel ∂ surgery obstruction. The map from the structure set to $[X, G/CAT]$ is just the natural one which associates to a structure the corresponding reduction of its normal bundle.

This leaves the 'map' from $L_{m+1}(\pi)$ to the structure set. Here, $L_{m+1}(\pi)$ is a group, whereas the structure set is just a set; the arrow denotes an *action* of the group on the set, so that given a structure $M_0 \to X$ and an $\alpha \in L_{m+1}(\pi)$ there is defined a new structure $\alpha \cdot M_0 = M_1$. (Of course, this action leads to a map by acting on the reference structure, but this does not capture all the information.) This action is defined by the Wall realization theorem. Specifically, according to 8.13 there exists a cobordism, with surgery obstruction

[1]Recall that 'CAT' means DIFF, PL, or TOP, according as whether we are considering smooth, piecewise-linear, or topological manifolds. We have worked throughout these notes in the smooth category. The same results are true in the piecewise-linear category, and the general outline of the proofs is the same, although of course some of the technology used in the smooth category (e.g. Morse functions) has to be replaced by appropriate PL technology (e.g. regular neighbourhood theory). The results are also true in the topological category; the proof of this, however, is a great deal harder, and involves a 'bootstrapping' argument which makes use of sophisticated PL or DIFF surgery in order to establish the foundational results (such as the existence of handle decompositions) in TOP. We will say something about this in the next (?) chapter; however, since the results are true in the end, allow me to state them for all categories at once.

[2]In the absence of existence, there is not much point talking about uniqueness.

α, from M_0 to a new structure M_1. We need to see that M_1 is uniquely determined by M_0 and α. Suppose that M_1 and M_1' are two structures both of which are obtained in this way; then (gluing the cobordisms, and using the lemma above) we see that M_1 and M_1' are (normally) cobordant by a cobordism W which has surgery obstruction 0. Since W has surgery obstruction 0, we can do surgery (rel ∂) on it to obtain a new cobordism W' such that in the diagram

$$\begin{tikzcd} M_1 = \partial_- W' \arrow[r] \arrow[d] & W' \arrow[d] & \partial_+ W' = M_1' \arrow[l] \arrow[d] \\ X \times \{0\} \arrow[r] & X \times [0,1] & X \times \{1\} \arrow[l] \end{tikzcd}$$

all the vertical maps are homotopy equivalences. But then W is an h-cobordism, hence (by our assumption about the Whitehead group) an s-cobordism. By the s-cobordism theorem, then, W is a product, and M_1 and M_1' represent the same element of the structure set.

Beware that we do not assert (and it is not in general true) that if $\alpha \neq 0$ then M_0 and $\alpha \cdot M_0$ represent *different* structures. The cobordism might represent a non-trivial 'automorphism' of a fixed structure.

We now proceed to interpret the 'exactness' of the surgery exact sequence. Exactness at $[X, G/CAT]$ just means that a normal invariant comes from a manifold structure if and only if its surgery obstruction vanishes; this is the main result of the previous section. Exactness at the structure set means that two elements of the structure set belong to the same orbit of the action of $L_{m+1}(\pi)$ if and only if they map to the same normal invariant, i.e., are normally cobordant. This is immediate from our definitions: M_0 and $\alpha \cdot M_0$ (above) are normally cobordant by construction, and, conversely, if M_0 and M_1 are normally cobordant, then the rel ∂ surgery obstruction of a normal cobordism between them gives an $\alpha \in L_{m+1}(\pi)$ such that $\alpha \cdot M_0 = M_1$. Finally, exactness at $L_{m+1}(\pi)$ means that an element $\alpha \in L_{m+1}(\pi)$ comes from $[\Sigma X, G/CAT]$ if and only if it maps the reference structure M in $calS(X)$ to itself. Since $[\Sigma X, G/CAT]$ can be represented by equivalence classes of normal bordisms from the reference structure to itself, this exactness is immediate from the definitions. □

The surgery exact sequence can be extended indefinitely to the left. However, the correct context in which to discuss this is that of 'spacified' surgery. We will not go into this here.

AWFUL WARNING: The space G/CAT is a homotopy commutative H-space, by Whitney sum of bundles. Hence $[X, G/CAT]$ is an abelian group. The surgery obstruction $\sigma\colon [X, G/CAT] \to L_m(\pi)$ is a map of abelian groups, which plainly takes zero to zero. Nevertheless, it is *not* in general a homomorphism[3]. However, the 'suspended' surgery obstruction map $[\Sigma X, G/CAT] \to L_{m+1}(\pi)$ is a homomorphism. This is because the group operation on the left hand side can be defined using either the H-space structure of G/CAT or the co-H-space structure of ΣX, and if we use the latter then addition corresponds to gluing normal cobordisms.

The Whitehead group assumption above is invidious. To get rid of it requires some changes to the definitions, and there is more than one place that these changes can be made. The most straightforward approach is to relax the definition of the structure set. Thus, let $\mathcal{S}^h(X)$ be the same thing as the structure set that we defined before, except that two structures are considered to be equivalent if they are merely *h-cobordant*, rather than diffeomorphic. Then plainly the arguments above yield an exact sequence

$$\ldots \to [\Sigma X, G/CAT] \to L^h_{m+1}(\pi) \to \mathcal{S}^h_{CAT}(X) \to [X, G/CAT] \to L^h_m(\pi),$$

[3] Wall adds: " as one readily sees by computing with Pontrjagin classes (the simplest example is the quaternion projective plane)."

where we have given our L-groups the "decoration" h to match the decoration h on the structure sets. A more sophisticated approach involves a modification of the definition of L-theory. Thus suppose that we manufacture a new L-theory out of *simple* isomorphism classes of quadratic forms on free and *(stably) based* $\mathbb{Z}\pi$-modules. (Simple isomorphism means isomorphism having zero Whitehead torsion.) This L-theory, denoted L^s, will measure the obstruction to surgering to get a *simple* homotopy equivalence. In the argument above we will therefore always get an s-cobordism where previously we got an h-cobordism. We will get another surgery exact sequence

$$\ldots \to [\Sigma X, G/CAT] \to L^s_{m+1}(\pi) \to \mathcal{S}^s_{CAT}(X) \to [X, G/CAT] \to L^s_m(\pi)$$

where now $\mathcal{S}^s$ refers to our original definition of the structure set, and L^s is our new "simple L-theory".

2. A chain complex model

Andrew will write

To do

3. The Naked Homeomorphism

In the last section we have constructed a chain complex model for the surgery exact sequence. This chain complex model does not preserve all the information about smooth manifolds; for instance, the groups of exotic spheres are all mapped to 0 in the algebraic model. It is a mysterious fact that although the algebraic model does not correspond exactly to the geometric and intuitive theory of *smooth* manifolds, it does correspond exactly to the apparently much more *outré*[4] theory of *topological* manifolds.

(preliminary sketch of how this will work and what it will imply; Rochlin's theorem.)

4. Notions of Controlled Topology

[4] Sherlock Holmes

CHAPTER 10

Applications of Surgery

CHAPTER 11

Appendices

Appendix C: CW Complexes

Appendix D: Diagonal Approximations

Appendix O: Obstruction Theory

Appendix Q: Quaternions and Octonions

Appendix S: Sard's Theorem

Appendix T: Tubular Neighborhoods

Bibliography

BoHi1 [1] A. Borel and F. Hirzebruch. Characteristic classes and homogeneous spaces I. *American Journal of Mathematics*, 80:458–538, 1958.

BoTu [2] R. Bott and L.W. Tu. *Differential Forms in Algebraic Topology*, volume 82 of *Graduate Texts in Mathematics*. Springer-Verlag, New York–Heidelberg–Berlin, 1982.

Browder [3] W. Browder. *Surgery on simply-connected manifolds*. Springer-Verlag, New York–Heidelberg–Berlin, 1972.

Browder2 [4] W. Browder. Homotopy type of differentiable manifolds. In S. Ferry, A. Ranicki, and J. Rosenberg, editors, *Proceedings of the 1993 Oberwolfach Conference on the Novikov Conjecture*, volume 227 of *LMS Lecture Notes*. Cambridge University Press, Cambridge, 1995.

Hatcher [5] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

KerMil [6] Michel A. Kervaire and John W. Milnor. Groups of homotopy spheres. I. *Annals of Mathematics*, 77:504–537, 1963.

KiSi [7] Robion C. Kirby and Laurence C. Siebenmann. *Foundational essays on topological manifolds, smoothings, and triangulations*. Princeton University Press, Princeton, N.J., 1977. With notes by John Milnor and Michael Atiyah, Annals of Mathematics Studies, No. 88.

Lang [8] S. Lang. *Algebra*. Addison-Wesley, 1995. Third edition.

Mil7 [9] J.W. Milnor. Classification of $(n-1)$-connected $2n$-dimensional manifolds and the discovery of exotic spheres.

Mil0 [10] J.W. Milnor. On manifolds homeomorphic to the seven-sphere. *Annals of Mathematics*, 64:399–405, 1956.

Mil8 [11] J.W. Milnor. Two complexes which are homeomorphic but combinatorially distinct. *Annals of Mathematics*, 74:575–590, 1961.

Mil6 [12] J.W. Milnor. *Lectures on the h-cobordism theorem*. Princeton University Press, Princeton, N.J., 1965.

MiS [13] J.W. Milnor and J.D. Stasheff. *Characteristic Classes*, volume 76 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, N.J., 1974.

Nov0 [14] S.P. Novikov. Homotopically equivalent smooth manifolds. *Mathematics of the USSR — Izvestija*, 28:365–474, 1964.

Nov1 [15] S.P. Novikov. Algebraic construction and properties of Hermitian analogs of K-theory over rings with involution from the viewpoint of hamiltonian formalism. applications to differential topology and the theory of characteristic classes I. *Mathematics of the USSR — Izvestija*, 4:257–292, 1970.

Sma0 [16] S. Smale. The story of the higher dimensional Poincaré conjecture: what actually happenned on the beaches of Rio. *Mathematical Intelligencer*, 12:44–51, 1990.

JHCW1 [17] J.H.C. Whitehead. On C^1 complexes. *Annals of Mathematics*, 41:809=824, 1940.