Handlebody Techniques in Geometric Topology

Remy Bohm

2020

Contents

Introduction

The goal of this paper is to provide a relatively self-contained introduction to handle decompositions of manifolds and their uses in various disciplines in geometric topology. In particular, we will prove the theorem that a handle decomposition exists for every compact smooth manifold using techniques from Morse theory. Sections 1 and 2 are devoted to building up the necessary machinery to discuss the proof of this fact, and the proof itself is in Section 3. In Section 4, we discuss an application of handle decompositions to algebraic topology, namely Poincaré duality. From there, we move on to more techniques using handle decompositions in Section 5, an application in 3-manifold theory in Section 6, and discuss the higher dimensional case, along with proving the h-cobordism theorem, in Section 7. Finally, we take a tour through 4-manifold theory with an emphasis on Kirby calculus in Sections 8 and 9.

We assume familiarity with some real analysis, linear algebra, and multivariable calculus. Basic algebraic topology is also used, and we will refer to homology and cohomology throughout. Several theorems in this paper rely heavily on commonplace results in these other areas of mathematics, and so in many cases, references are provided in lieu of a proof. This choice was made in order to avoid getting bogged down in difficult proofs that are not directly related to geometric and differential topology, as well as to make this paper as accessible as possible.

Before we begin, we introduce a motivating example to consider through the first few sections of this paper. Imagine a torus, standing up on its end, behind a curtain, and what the torus would look like as the curtain is slowly lifted. The pictures in Figure 1 show the portions of the torus that are visible at different moments as the curtain is lifted. A closer look will reveal that during this unveiling process, the topology of the revealed portion changes; at first it is simply a disk, then a tube, then a torus with one boundary component, and finally, a whole torus. This paper aims to provide an explanation for how the topology of the torus changes as it is unveiled, as well as how that informs other studies within topology.

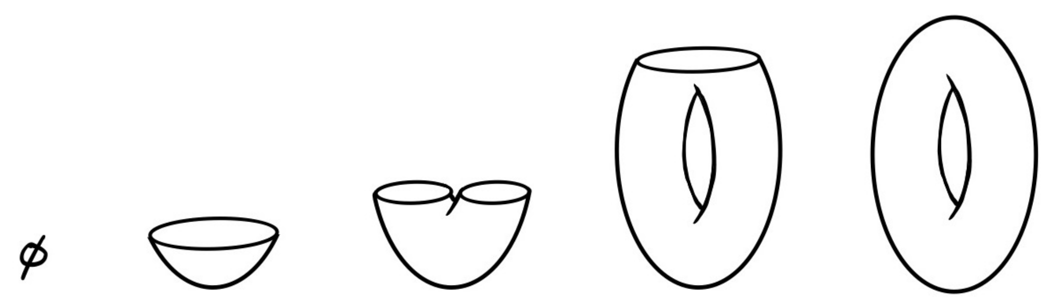

Figure 1: Unveiling a torus.

1 Smooth Manifolds and Handles

We begin by defining topological manifolds.

Definition 1.1. A *topological manifold* M is a second countable, Hausdorff topological space such that for all points p in M, there exists an open neighborhood N_p of p such that N_p is homeomorphic to the Euclidean open n-ball, $B^n := \{x \in \mathbb{R}^n \mid |x| < 1\}$.

It will be standard notation throughout this paper to use M^n to denote an n-dimensional manifold when the dimension of the manifold is relevant, after which the manifold may be simply referred to as M.

Definition 1.2. A *manifold with boundary* M is a second countable, Hausdorff topological space such that for all points p in M, there exists an open neighborhood N_p of p such that N_p is homeomorphic to either the Euclidean open n-ball $\{x \in \mathbb{R}^n : |x| < 1\}$ or the Euclidean open half-n-ball $\{x \in \mathbb{R}^n_+ : |x| < 1\}$.

Two points p and q in a manifold M may have neighborhoods that overlap, but are both homeomorphic to Euclidean balls B^n. We therefore introduce the idea of a transition map on the intersection.

Definition 1.3. Let M be a manifold, and let U, V be open subsets of M^n with homeomorphisms $P_U : U \to B^n$ and $P_V : V \to B^n$ such that $U \cap V \neq \emptyset$. The map $\phi \colon \mathbb{R}^n \to \mathbb{R}^n$ sending $P_U(U \cap V)$ to $U \cap V$ and then to $P_V(U \cap V)$ is called the *transition map* on $U \cap V$.

Transition maps are important in the study of manifolds, since they allow one to patch together local coordinate systems on manifolds to form globally defined structures.

We now proceed to definitions pertaining to smooth manifolds.

Definition 1.4. A function $f \colon \mathbb{R}^n \to \mathbb{R}^n$ is said to be C^∞ or *smooth* if it is infinitely differentiable.

Definition 1.5. A *smooth manifold* is a manifold M^n such that all of its transition maps are C^∞.

Definition 1.6. If M and N are smooth manifolds, then $f : U \to V$ is a *diffeomorphism* if it is a homeomorphism and if f and f^{-1} are smooth.

For our purposes, as is common in the literature, we will take diffeomorphisms to be infinitely differentiable, to match our infinitely differentiable manifolds.

Definition 1.7. A *manifold with corners* M is a second countable, Hausdorff topological space such that for all points p in M, there exists an open neighborhood N_p of p such that N_p is homeomorphic to one of the following:
(i) the Euclidean open n-ball $\{x \in \mathbb{R}^n : |x| < 1\}$
(ii) the Euclidean open half-n-ball $\{x \in \mathbb{R}^{n-1} \times \mathbb{R}_+ : |x| < 1\}$
(iii) other subsets of the Euclidean n-ball where more than one coordinate is restricted positive $\{x \in \mathbb{R}^{n-m} \times \mathbb{R}^m_{+^m} : |x| < 1\}$

Note that manifolds with corners are homeomorphic to manifolds with boundaries, but not necessarily diffeomorphic to them.

Definition 1.8. For each point p in a smooth manifold M, let N_p be a coordinate neighborhood with a local homeomorphism $\phi\colon N_p \to U \subset \mathbb{R}^n$. Consider the equivalence classes $[\gamma]$ of curves $\gamma\colon [-1,1] \to N_p$ passing through p such that $\gamma(0) = p$, under the equivalence relation $\gamma_1 \equiv \gamma_2$ if $\frac{\partial(\phi\circ\gamma_1)}{\partial t} = \frac{\partial(\phi\circ\gamma_2)}{\partial t}$ as maps $\phi \circ \gamma\colon [-1,1] \to U \subset \mathbb{R}^n$. We say that an equivalence class of such local paths v is a *tangent vector* to M at p, and that the vector space spanned by all such v is the *tangent space* of M at p, denoted T_pM.

The union of all the tangent spaces over M is called the *tangent bundle* on M, denoted TM.

Note that all topological manifolds have tangent spaces, but they do not necessarily patch together in a way that will be useful for our purposes without a smooth structure already in place. Nevertheless, tangent bundle on smooth manifolds is a major object of study in geometric and differential topology, and it comes with a lot of interesting structure. For a more detailed introduction, see [13].

There are two pieces of notation we will also introduce here, as they are standard across geometric topology. One is the *connect sum* of manifolds, which we will not formally define, but note that it will be donated with $\#$. The other is less widely known.

Definition 1.9. The *boundary connect sum* of two manifolds M and N with boundary, denoted $M \natural N$, is the manifold obtained by picking connected boundary components in M and N and taking the connect sum of those two submanifolds.

The main theme of this paper is to understand smooth manifolds by breaking them up into smaller, topologically trivial chunks called *handles*.

Definition 1.10. An *n-dimensional k-handle* is a copy of the contractible smooth manifold $D^k \times D^{n-k}$.

We specify the construction of k-handles as $D^k \times D^{n-k}$ so that we can denote with k the region of h^k along which we "glue" it to another topological space of the same dimension n. We make the notion of "gluing" precise below.

Definition 1.11. Let X, Y be topological spaces, and let $K \subset X$ and $L \subset Y$ be subspaces such that there exists a diffeomorphism $\phi\colon K \to L$. We obtain a new space, which we call X *glued to* Y *along* ϕ by taking ${}^{X \sqcup Y}/_{x \sim \phi(x)}$. We call ϕ the *attaching map*.

In general, we will use the notation $X \cup_\phi Y$ to denote "X glued to Y along ϕ", even if no attaching map ϕ is specified.

With the goal of gluing handles to other topological spaces in mind, we now define some useful parts of a k-handle.

Definition 1.12. There are five subsets of a k-handle which will be of interest to us. They are:
(i) the *attaching region*, defined to be $\partial D^k \times D^{n-k}$. In this paper, it is shown in bold in figures. Note that the attaching map of a k-handle is a homeomorphism of the attaching region into a subset of the space being glued to.
(ii) the *attaching sphere*, denoted A^k: $\partial D^k \times \{0\}$,
(iii) the *core*, denoted C^k: $D^k \times \{0\}$,
(iv) the *belt sphere*, denoted B^k: $\{0\} \times \partial D^{n-k}$,
(v) the *co-core*, denoted K^k: $\{0\} \times D^{n-k}$.

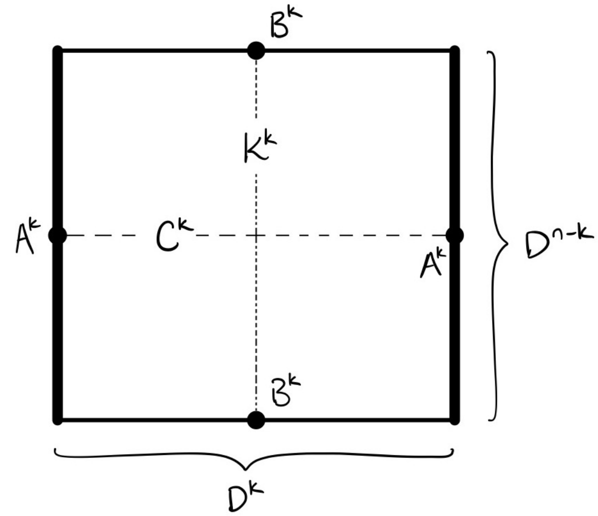

Figure 2: Anatomy of a k-handle, with attaching region shown in bold.

Envisioning D^k and D^{n-k} both as products of the unit interval, we draw the following diagram of a handle in Figure 2.

A k-handle in dimensions higher than 2 is difficult to draw, but to give a sense of how to interpret Figure 2, we show the attaching of several kinds of handles in Figure 3.

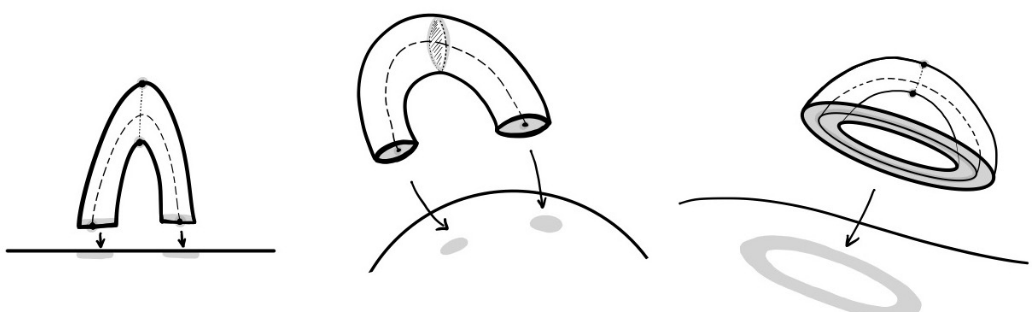

Figure 3: Left to right: Attaching a 2-dimensional 1-handle, attaching a 3-dimensional 1-handle, attaching a 3-dimensional 2-handle.

Definition 1.13. A *handle decomposition* of a compact manifold M is a finite sequence of manifolds $W_0, \ldots, W_l$ such that:
(i) $W_0 = \emptyset$,
(ii) W_l is diffeomorphic to M,
(iii) W_i is obtained from W_{i-1} by attaching a handle.
A *handlebody* is a compact manifold expressed as the union of handles.

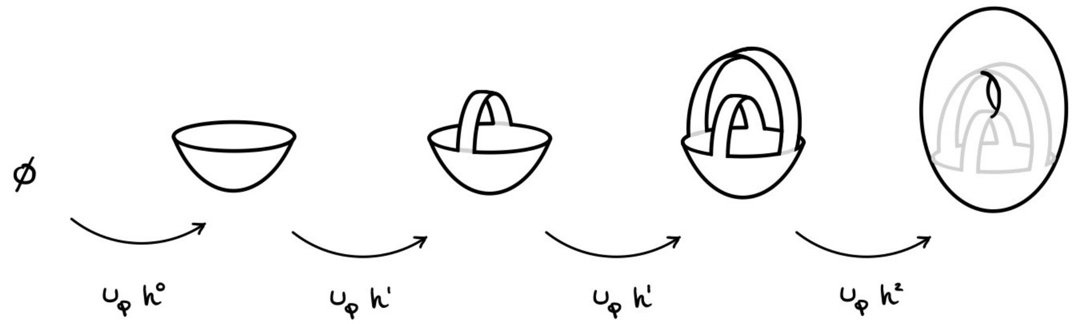

Figure 4: A handle decomposition of a torus.

Handle decompositions allow one to construct a manifold piece by piece, attaching one k-handle at a time. An example of a handle decomposition of a torus is shown below in Figure 4. The reader should take a moment to convince themselves that the final attachment of a 2-handle in the figure really does produce a torus. It is also important to note that a handle decomposition of a given manifold is not unique. For instance, below are two decompositions of the unit sphere S^2.

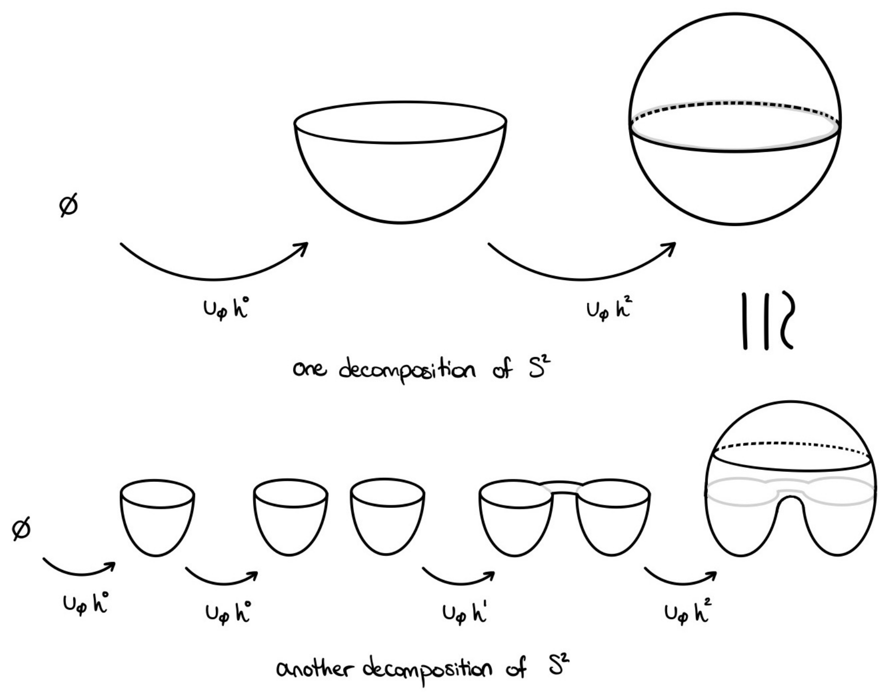

Figure 5: Two decompositions of S^2.

Even though any given manifold has many different handle decompositions, handle decompositions are nevertheless very useful tools for understanding the topology of manifolds, as they provide a "manual" of sorts for building a manifold piece by piece. In particular, all closed smooth manifolds admit handle decompositions, allowing many problems in topology to be studied purely in the context of handlebodies. The proof of this fact requires an understanding of some basic Morse theory, which we will now discuss.

2 Morse Functions and Flows

2.1 Definitions and Existence

The basic idea of Morse theory is to understand manifolds by studying certain real-valued maps, called Morse functions, on them. We begin by introducing some properties of smooth functions on manifolds.

Definition 2.1. The *gradient vector field* of a function f is the vector field on the domain of f that takes the value $(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})$ at each point. We denote this vector field ∇f and its value at a point p as $\nabla f|_p$.

Definition 2.2. A *critical point* of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is a point $p \in \mathbb{R}^n$ such that $\nabla f|_p = 0$. Similarly, a *critical value* of f is a value $c \in \mathbb{R}$ such that $f(p) = c$ for p a critical point of f.

Definition 2.3. The *Hessian* of a function $f\colon M \to \mathbb{R}$, denoted $\mathcal{H}_f$, is the matrix of mixed second order partial derivatives of f:

$$\mathcal{H}_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

The Hessian evaluated at a point p is written $\mathcal{H}_f(p)$.

Definition 2.4. A critical point p of a continuous function f is called *degenerate* if $\det(\mathcal{H}_f(p)) = 0$.

We can now define a Morse function.

Definition 2.5. Given a smooth manifold M and a smooth function $f\colon M \to \mathbb{R}$, we say that f is *Morse* if f has no degenerate critical points on M.

The prototypical example of a Morse function on a manifold is a height function on a surface. That is, imagine your favorite closed surface floating in space above a plane. Then let your Morse function simply measure the height of level sets of the surface above the plane. A visual of this example is shown in Figure 6.

The point of studying Morse functions is that if a function has only nondegenerate critical points, the function's local behavior in the neighborhood of its critical points can be further studied and classified, as is shown in the following definition.

Definition 2.6. Let M be a smooth manifold, $f\colon M \to \mathbb{R}$ be smooth, and p be a nondegenerate critical point of f. Then the *index* of f and p is defined to be the number of negative eigenvalues of the Hessian $\mathcal{H}_f$ evaluated at p.

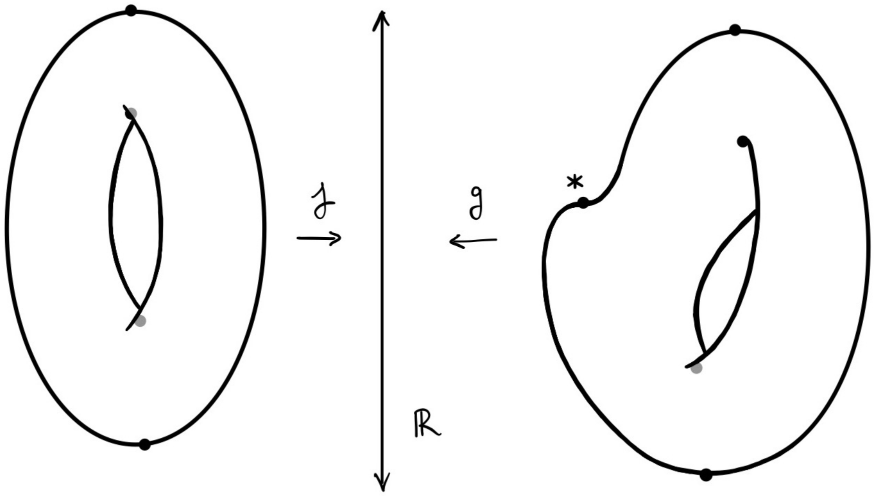

Figure 6: Two height functions on a torus, with critical points shown in bold. f is Morse; g is not. The starred critical point is degenerate.

Heuristically, the index of the Hessian tells us how many directions f is decreasing on. It will be the key to understanding how Morse functions relate to the actual attachment of handles to a manifold.

Proposition 2.7. *The nondegeneracy and the index of a function f at a critical point p do not depend on choice of local coordinates.*

Proof. We appeal to Sylvester's Law, which states that the number of negative eigenvalues of the Hessian is independent of the way it is diagonalized. Since diagonalization of a matrix corresponds to changing the basis of the source vector space so that the basis vectors are the eigenvectors of the matrix, this means that the number of negative eigenvalues of the Hessian is invariant under coordinate transformation. □

To make Morse functions effective tools in general, we must prove that they exist on all compact smooth manifolds. The proof of this fact is usually stated in the literature as the theorem that the set of Morse functions on a smooth, closed manifold M is dense in $C^\infty(M)$. In this treatment, we prove that one can always find a "very similar" function, or a (C^2, ε)-approximation, of any function such that the approximation function is Morse. Even with this modification, this is a rather involved proof requiring two fundamental lemmas dealing in real analysis. We therefore provide intuitive outlines for the proofs below, rather than fully rigorous ones. A more thorough treatment can be found in [14], from which these proofs are adapted.

We begin with the definition of a (C^2, ε)-approximation:

Definition 2.8. A function $f\colon K \subset \mathbb{R}^n \to \mathbb{R}$ is said to be a (C^2, ε)-*approximation* of a function

$g\colon K \to \mathbb{R}$ if the following inequalities hold for all points $p \in K$:

$$
\begin{aligned}
|f(p) - g(p)| &< \varepsilon & & \\
\left|\frac{\partial f}{\partial x_i}(p) - \frac{\partial g}{\partial x_i}(p)\right| &< \varepsilon & & i = 1, \ldots, n \\
\left|\frac{\partial^2 f}{\partial x_i \partial x_j}(p) - \frac{\partial^2 g}{\partial x_i \partial x_j}(p)\right| &< \varepsilon & & i, j = 1, \ldots, n
\end{aligned}
$$

We can now move on to the requisite lemmas from analysis.

Lemma 2.9. *Let U be an open subset of $\mathbb{R}^n$ with coordinates $\{x_1, \ldots, x_n\}$ and let $f\colon U \to \mathbb{R}$ be a smooth function. Then there exist real numbers $\{a_i\}$ such that $f(x_1, \ldots, x_n) - (a_1x_1 + \cdots + a_nx_n)$ is Morse on U. Moreover, for all $\varepsilon > 0$, each a_i can be chosen such that $|a_i| < \varepsilon$.*

Proof. The proof of this lemma is dependent on Sard's theorem, which states that the set of critical values of a continuous function $g\colon U \subset \mathbb{R}^n \to \mathbb{R}$ has measure 0 in $\mathbb{R}$. This result has very powerful applications in differential topology, but its proof is analytical, and so we refer the reader to Appendix C of [3] for a proof.

We begin with a function $f\colon U \to \mathbb{R}$ that may or may not have degenerate critical points in U. Let $h\colon U \to \mathbb{R}^n$ send $p \in U$ to $\nabla f(p)$. Then the matrix of partial derivatives of h is precisely the Hessian $\mathcal{H}_f$ at each point $p \in U$. Thus, critical points of h are precisely the degenerate critical points of f (points p where $\det(\mathcal{H}_f(p)) = 0$).

By Sard's Theorem, we can choose a point $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ such that a is arbitrarily close to 0 but a is not a critical value of h.

We now claim that $\bar{f} := f - (a_1x_1 + \cdots + a_nx_n)$ is Morse on U. To see this, let p be a critical point of $\bar{f}$. Then $h(p) = a$ since $\frac{\partial \bar{f}}{\partial x_i}(p) = \frac{\partial f}{\partial x_i}(p) - a_i = 0$. But since a was chosen to not be a critical value of h, p must not be a critical point of h, and hence $\det(\mathcal{H}_f(p)) \neq 0$. Furthermore, $\mathcal{H}_f = \mathcal{H}_{\bar{f}}$ since f and $\bar{f}$ differ only by linear terms which vanish under second derivatives. Conclude $\det(\mathcal{H}_{\bar{f}}(p)) \neq 0$, and so p is nondegenerate. ☐

The upshot of this lemma is that we only ever need to modify a smooth function on an open subset of a manifold by some arbitrarily small linear term to make it Morse.

Lemma 2.10. *Let K be a compact subset of a manifold M, and suppose that $g\colon M \to \mathbb{R}$ has no degenerate critical points in K. Then for sufficiently small $\varepsilon > 0$, any (C^2, ε)-approximation f of g has no degenerate critical point in K.*

Proof. Let $\{U_i\}$ be a finite cover by open coordinate neighborhoods of K. For any function f to have no degenerate critical points in a given U_i, it must have no points where all of its partial derivatives and the determinant of its Hessian matrix with respect to the coordinates $\{x_1, \ldots, x_n\}$ on U_i are all 0. Equivalently, it must satisfy the following inequality:

$$
\left|\frac{\partial f}{\partial x_1}\right| + \cdots + \left|\frac{\partial f}{\partial x_n}\right| + |\det(\mathcal{H}_f)| > 0
$$

But if f is a (C^2, ε)-approximation of g, which we know satisfies the above inequality, then we have that:

$$\left|\frac{\partial f}{\partial x_i} - \frac{\partial g}{\partial x_i}\right| < \varepsilon \qquad\qquad i = 1, \dots, n$$

$$\left|\frac{\partial^2 f}{\partial x_i \partial x_j} - \frac{\partial^2 g}{\partial x_i \partial x_j}\right| < \varepsilon \qquad\qquad i, j = 1, \dots, n$$

Therefore, for sufficiently small $\varepsilon > 0$, we have that f satisfies the desired inequality, and therefore has no degenerate critical points on U_i.

If we repeat this process on all the U_i, we get that f has no degenerate critical points on all of K as desired. □

Together, these lemmas allow us to perturb continuous functions on open subsets of a manifold to make them Morse, as well as ensure that these perturbations have minimal effect outside of the subsets on which they are defined. The content of the existence theorem, then, is stitching these *local* perturbations together to form a function that is *globally* Morse.

Theorem 2.11 (Existence of Morse functions)**.** *Let M be a compact manifold and $f_0 \colon M \to \mathbb{R}$ be smooth. Then there exists a Morse function f on M that is an arbitrarily close approximation of f_0.*

Proof. Let $\{U_l\}_{1 \le l \le k}$ be a finite open cover of M such that for each U_l, there exists a compact subset K_l of U_l with $\{K_l\}$ a cover of M by compact sets. We begin with some smooth function f_0 on M that may have degenerate critical points. The idea of this proof is to inductively define functions f_l on M such that f_l is Morse on $\bigcup_{j=1}^{l} K_j$, denoted C_l for brevity. When $l = k$, we will have f_k Morse on $C_k = M$.

Our base case for induction will be to let $K_0 := \{\emptyset\}$ with f_0 our base function.

For our inductive hypothesis, suppose that we already have $f_{l-1} \colon M \to \mathbb{R}$ such that f_{l-1} is Morse on C_{l-1}. We want now to show that there exists a function f_l that is Morse on $C_{l-1} \cup K_l = C_l$.

To do this, let $\{x_1, \dots, x_n\}$ be local coordinates on U_l. Lemma 2.9 then tells us that there exist real numbers $\{a_i\}$ such that $f_{l-1}(x_1, \dots, x_n) - (a_1 x_1 + \cdots + a_n x_n)$ is Morse on U_l.

We cannot simply set f_l to be this modified version of f_{l+1}, however, since the coordinates $\{x_1, \dots, x_n\}$ are local to U_l. To fix this, we introduce a smooth bump function on $h_l \colon U_l \to [0, 1]$ such that $h_l = 1$ on an open neighborhood V_l of K_l contained in U_l, but $h_l = 0$ outside of a compact neighborhood $\bar{V}_l$. This is a lot of sets to keep track of, so a picture is shown below in Figure 7.

We can now define f_l on all of M as follows:

$$f_l(p) = \begin{cases} f_{l-1}(p) - h_l(p) \cdot (a_1 x_1 + \cdots + a_n x_n) & p \in \bar{V}_l \\ f_{l-1}(p) & p \notin \bar{V}_l \end{cases}$$

All that remains is to check that f_l is a (C^2, ε)-approximation of f_{l-1}. Inside K_l, we can simply

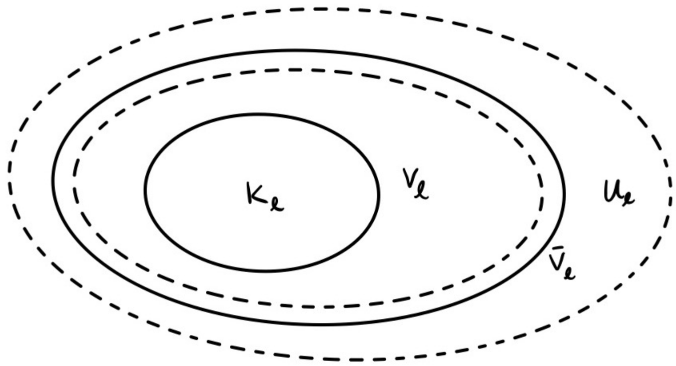

Figure 7: Relevant sets.

calculate the following inequalities:

$$|f_l - f_{l-1}| = |(a_1x_1 + \cdots + a_nx_n)|(h_l)$$

$$|\frac{\partial f_l}{\partial x_i} - \frac{\partial f_{l-1}}{\partial x_i}| = |a_ih_l + (a_1x_1 + \cdots + a_nx_n)\frac{\partial h_l}{\partial x_i}|$$

$$i = 1, \dots, n$$

$$|\frac{\partial^2 f_l}{\partial x_i \partial x_j} - \frac{\partial^2 f_{l-1}}{\partial x_i \partial x_j}| = |a_i\frac{\partial h_l}{\partial x_j} + a_j\frac{\partial h_l}{\partial x_i} + (a_1x_1 + \cdots + a_nx_n)\frac{\partial^2 h_l}{\partial x_i \partial x_j}|$$

$$i, j = 1, \dots, n$$

We know that h_l is bounded on $\bar{V}_l$, which is compact, and 0 elsewhere, and so $|\frac{\partial h_l}{\partial x_i}|$ and $|\frac{\partial^2 h_l}{\partial x_i \partial x_j}|$ must also be bounded on $\bar{V}_l$. Therefore for all $\varepsilon > 0$, by choosing each a_i small enough, we can make $|f_l - f_{l-1}|$, $|\frac{\partial f_l}{\partial x_i} - \frac{\partial f_{l-1}}{\partial x_i}|$, and $|\frac{\partial^2 f_l}{\partial x_i \partial x_j} - \frac{\partial^2 f_{l-1}}{\partial x_i \partial x_j}|$ all less than ε. Therefore, on K_l, f_l is a (C^2, ε)-approximation of f_{l-1}.

Outside of $\bar{V}_l$, $f_l = f_{l-1}$, but we know that some compact sets K_j must intersect K_l since all the K_j together cover M. We therefore must check that f_l is a (C^2, ε)-approximation on the overlaps $K_l \cap K_j$, $j \neq l$, that is to say, with respect to the coordinates on U_j for those U_j which intersect U_l. Fortunately, because M is smooth, all of its transition maps between overlapping open sets are C^∞, and so the composition of any transition map from U_l to U_j with f_l differs from f_{l-1} on U_j by a bounded term. Therefore, for all $\varepsilon > 0$, we can adjust the a_i to be even smaller such that f_l is a (C^2, ε)-approximation of f_{l-1} on the overlaps $K_l \cap K_j$. Outside of $\bar{V}_l$, $f_l = f_{l-1}$, and so we conclude that f_l is a (C^2, ε)-approximation of f_{l-1} on all of M.

By Lemma 2.10, we can now say that if f_{l-1} had no degenerate critical points in any K_j for $j \neq l$, then f_l must also have no degenerate critical points in any of those sets. Thus, after inducting on l until $l = k$, we have f_k Morse on M. $\square$

2.2 The Morse Lemma

Now that we have familiarized ourselves with the idea and existence of Morse functions, we can proceed to the first result of Morse theory: the use of the index of a critical point to define new coordinate systems on neighborhoods of critical points. This observation is made rigorous in the Morse lemma below. Proving the Morse lemma is a key step in the construction of handlebody decompositions, as it allows us to reduce the behavior of a Morse function near a critical point to simply telling us how many coordinates f is increasing on, and how many it is decreasing on. If f is a height function, as in our examples, this is equivalent to walking along inside our manifold at p and noting how many directions one could walk in to go "down" and how many one could walk in to go "up".

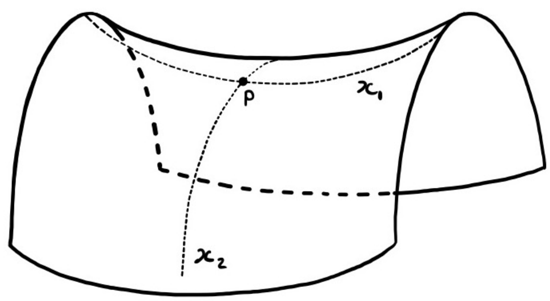

Figure 8: An index 1 critical point p on a surface, with local coordinates $\{x_1, x_2\}$.

This is shown in Figure 8. Note that in this figure, if one were walking along the surface at p, one could walk along the x_1 axis to walk "up" or along the x_2 axis to walk "down". Figure 8 therefore corresponds to an index 1 critical point.

Before we can get to the actual proof of the Morse lemma, we will need a lemma from multivariable calculus.

Lemma 2.12. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be C^∞ on a convex neighborhood $U \subset \mathbb{R}^n$ containing the origin, and suppose that $f(0, \ldots, 0) = 0$. Then there exist C^∞ functions $\{g_i\}_{1 \le i \le n}$ defined on U such that:*

$$f = \sum_{i=1}^{n} x_i g_i$$

with:

$$g_i(0, \ldots, 0) = \frac{\partial f}{\partial x_i}(0, \ldots, 0)$$

We will omit the proof of this lemma for brevity, as it is an application of the multivariable chain rule. For a concise proof, see Part I, Chapter 2 of [15].

We are finally ready to prove the Morse lemma. There are many proofs of the Morse lemma out there, all in varying levels of detail. Here we provide a general idea of the proof that is palatable for those not in the mood to do lots of coordinate transformations, with a particular emphasis on aspects of the proof that are enlightening for its later use in handlebody decompositions. Milnor in [15] has a similar proof.

Theorem 2.13 (Morse lemma)**.** *Let f be a Morse function on a manifold M and p be a nondegenerate critical point of f. Then there exist local coordinates $\{x_1, \ldots, x_n\}$ on a neighborhood N_p such that on N_p, $f(x_1, \ldots, x_n)$ has the form:*

$$f(x_1, \ldots, x_n) = -x_1^2 - \cdots - x_k^2 + x_{k+1}^2 + \cdots + x_n^2$$

Where k is the index of f at p, and p corresponds to the origin of this coordinate system.

Proof. We begin by letting $\{y_1, \ldots, y_n\}$ be local coordinates on N_p and considering $f(y_1, \ldots, y_n)$. To rearrange the coordinates such that f takes a quadratic form on a neighborhood of p, we would like to rewrite f in such a way that allows us to see the behavior of its second partial derivatives at $p = (0, \ldots, 0)$.

To do this, we use Lemma 2.12 twice; one iteration of the lemma applied to f defines functions $g_i \colon \mathbb{R}^n \to \mathbb{R}$ such that $f(y_1, \ldots, y_n) = \sum_{i=1}^n y_i g_i(y_1, \ldots, y_n)$ and $g_i(0) = \frac{\partial f}{\partial y_i}(0) = 0$, and the second iteration applies the lemma to each g_i to define functions $h_{i,j} \colon \mathbb{R}^n \to \mathbb{R}$ such that $h_{i,j}(0) = \frac{\partial g_i}{\partial x_j}(0)$. The details of this calculation are not enlightening, but it is necessary to obtain the form below for f:

$$f(y_1, \ldots, y_n) = \sum_{i=1}^n \sum_{j=1}^n y_i y_j h_{i,j}(y_1, \ldots, y_n)$$

Because all partial derivatives of f are assumed to exist on N_p, $h_{i,j} = h_{j,i}$. Furthermore, if we compute the 2^{nd} partial derivatives of f at $p = (0, \ldots, 0)$ in terms of $\{h_{i,j}\}$, we see that:

$$\frac{\partial^2 f}{\partial y_i \partial y_j}(p) = \begin{cases} 2h_{i,j}(p) & i = j \\ h_{i,j}(p) & i \neq j \end{cases}$$

Crucially, this observation means that $2h_{i,i}(p)$ is equal to the i^{th} diagonal entry of $\mathcal{H}_f(p)$ in terms of coordinates $\{y_1, \ldots, y_n\}$, and that $h_{i,j}$ is equal to the $(i,j)^{\text{th}}$ entry for $i \neq j$. Our goal, therefore, is to diagonalize $\mathcal{H}_f(p)$ in order to get a quadratic form for f.

Now because second partial derivatives commute, the Hessian is a symmetric matrix with real entries and is therefore diagonalizable. There therefore exists a coordinate basis $\{\bar{x}_1, \ldots, \bar{x}_n\}$ on N_p such that in this basis, $\mathcal{H}_f(p)$ is diagonal. If we let λ_i be the i^{th} diagonal entry of $\mathcal{H}_f(p)$, then we know that in this basis, f takes the following form:

$$f(\bar{x}_1, \ldots, \bar{x}_n) = \sum_{i=0}^n \frac{\lambda_i}{2} \bar{x}_i^2$$

We perform one more coordinate transformation J on N_p to get rid of the $\frac{\lambda_i}{2}$ coefficients, being careful to keep their sign:

$$x_i = J(\bar{x}_i) := \text{sign}(\lambda_i) \cdot \sqrt{\frac{|\lambda_i|}{2}} \bar{x}_i$$

Thus we obtain a quadratic form for f in the coordinates $\{x_1, \ldots, x_n\}$:

$$f(x_1, \ldots, x_n) = \text{sign}(\lambda_1) x_1 + \cdots + \text{sign}(\lambda_n) x_n$$

If we wish, we can now permute the coordinates to group them by the signs of their coefficients to achieve the desired form for f.

$$f(x_1, \cdots, x_n) = -x_1^2 - \cdots - x_k^2 + x_{k+1}^2 + \cdots + x_n^2$$

□

If f can be locally modeled by quadratics in each coordinate, then because quadratics have no critical points that cannot be isolated from their extrema by open neighborhoods (in fact, they have no other critical points at all), f must not have any critical points too near to any other. We therefore obtain the following result:

Corollary 2.14. *Nondegenerate critical points on any manifold can be isolated by open neighborhoods.*

This is an important corollary, as it allows us to consider nondegenerate critical points one by one.

2.3 Flows on Manifolds

We now take a brief detour into another topic in differential topology and geometry: flows. Morally, a flow is a group of diffeomorphisms associated to a smooth vector field X on a manifold M that sends a point on M in the direction of the vector associated to that point by X. We make these two definitions precise below:

Definition 2.15. A *smooth vector field* X on a manifold M is a smooth map from M into its tangent bundle TM that assigns to each p in M a vector v_p in T_pM.

Definition 2.16. A *flow* on a manifold M is a map $\Phi\colon \mathbb{R} \times M \to M$ with the following properties:
(i) $\varphi_t(p) := \Phi(t, p)$ is a diffeomorphism of M,
(ii) $\varphi_0(p)$ is the identity diffeomorphism of M,
(iii) For all $s, t \in \mathbb{R}$, $\varphi_{s+t} = \varphi_s \circ \varphi_t$.

The *trajectory* of a point p in M of a flow is a map $\psi_p\colon \mathbb{R} \to M$ that sends $t \in \mathbb{R}$ to $\varphi_t(p)$. In particular note that, $\psi_p(0) = \varphi_0(p) = p$.

Note that the definition of a flow is equivalent to that of a smooth group action of $\mathbb{R}$ on M by diffeomorphisms.

One can think of a flow as being generated by a smooth vector field X on M by defining X such that $\frac{\partial \Phi(t,p)}{\partial t} = X \circ \varphi_t(p)$. In this framework, the trajectory of a point is equivalent to an integral curve of the vector field.

For our purposes, we will want to understand when this generation of flows by smooth vector fields is unique. It turns out that the restriction we put on smooth vector fields on M to have them generate unique flows on M is the property of being *compactly supported*, or taking on a value of 0 outside of some compact subset of M.

The following proof of this fact is adapted from [15], but it can be found in a variety of different texts in differential topology in varying forms. A particularly thorough treatment can be found in Chapter 12 of [13].

Theorem 2.17. *Let X be a smooth vector field on a manifold M and suppose that X is compactly supported on $K \subset M$. Then X generates a unique flow Φ on M.*

Proof. Given X, consider the set of differential equations on t parametrized by points p in K given by

$$\frac{\partial \Phi(t,p)}{\partial t} = X \circ \varphi_t(p)$$

with the initial condition $\varphi_0(p) = p$ at all points p in K.

Existence and uniqueness of Φ both, then, are a consequence of the common result in ordinary differential equations that guarantees that an ordinary differential equation with an initial condition has a unique, smooth solution that depends smoothly on the initial condition. So for any given point $p \in M$, there exists an open neighborhood N_p of p with a unique $\Phi(t,p)$ defined on it that satisfies the above differential equation for $t \in (-\varepsilon, \varepsilon)$. Furthermore, because X was smooth, for a given set of solutions $\{\varphi_{t,\alpha}\}$ such that each $\varphi_{t,\alpha}$ is defined on the same open set $(-\varepsilon', \varepsilon') \subset \mathbb{R}$, we can patch these local solutions on M together to define φ_t globally on M within $(-\varepsilon', \varepsilon')$.

It now remains to show that we can find such an open set in $\mathbb{R}$ that all φ_t are defined on. To do this, note that because K is compact, we can restrict a cover of K by neighborhoods N_p of individual points p_i indexed by $i \in I \subset \mathbb{N}$ to finitely many open neighborhoods $\{N_{p_i}\}$. Let ε_0 be the smallest of the ε_i corresponding to these neighborhoods N_{p_i}. Note that ε_0 is well-defined and nonzero because there are only finitely many neighborhoods N_{p_i}.

If $\varphi_t(q) = q$ for all $q \notin K$ and for all $t \in \mathbb{R}$, then we know that φ_t is defined for all of M for $t \in (-\varepsilon_0, \varepsilon_0)$. Furthermore, as each φ_t is generated by X, $\varphi_t \circ \varphi_s = \varphi_{t+s}$ for all $t, s \leq \varepsilon_0$ and so we can simply iterate φ_t on itself to generate $\varphi_{t'}$ for all t' such that $|t'| > \varepsilon_0$. Thus we obtain $\Phi(t,p)$ defined globally on M and for all $t \in \mathbb{R}$. □

This theorem will be essential in showing the existence of a diffeomorphism between submanifolds of M whose boundaries' image under f do not include a critical point.

3 From Morse Functions to Handle Decompositions

We have now constructed enough machinery to obtain a handlebody decomposition of any smooth, compact manifold. This theorem will follow from the following two intermediate results about the local behavior of a Morse function on a manifold, Theorem 3.2 and Theorem 3.4.

Definition 3.1. The *sublevel set* of a Morse function on M at a point $a \in \mathbb{R}$ is $\{p \in M \mid f(p) \leq a\}$. It is denoted M_a.

Theorem 3.2. *Let M be a compact manifold and $f : M \to \mathbb{R}$ be a Morse function. Suppose that $a, b \in \mathbb{R}$ are regular values of f such that $f^{-1}[a,b]$ is nonempty. If $f^{-1}[a,b]$ does not contain a critical point of f, then M_a is diffeomorphic to M_b.*

Proof. The idea of this proof is to find a vector field that we can associate to f and use that vector field to generate a flow that will give us a diffeomorphism from M_a to M_b.

Choose a Riemannian metric on M with inner product $\langle\ ,\ \rangle$ and let $||\ ||$ denote the induced norm on M. Note that we can choose such a metric because M is assumed to be smooth. For more information, see Chapter 8 of [13].

To construct a satisfactory flow, we need our vector field to have only unit vectors on $f^{-1}[a,b]$. To do this, define a new function $g : M \to \mathbb{R}$ such that $g = \frac{1}{||\nabla f||^2}$ on $f^{-1}[a,b]$ and g vanishes outside of a compact neighborhood of $f^{-1}[a,b]$. Note that the existence of such a g is a consequence of the existence of bump functions on manifolds; see [13].

We then define a vector field X on M by:

$$X(p) := g(p) \cdot \nabla f(p)$$

So on $f^{-1}[a, b]$, X takes the form:

$$X(p) = \frac{\nabla f}{||\nabla f||^2}(p)$$

However, X is compactly supported, since g was defined to vanish outside of a compact neighborhood of $f^{-1}[a, b]$. This allows us to apply Theorem 2.17 to generate a flow Φ on M.

We want to show that Φ contains a diffeomorphism that sends M_a to M_b. To do this, let $F\colon \mathbb{R} \to \mathbb{R}$ be defined by $F(t) = f \circ \varphi_t(p)$. We now calculate the derivative of F with respect to t.

$$\frac{\partial F}{\partial t} = \left\langle \frac{\partial \Phi}{\partial t}, \nabla f \right\rangle = \langle X, \nabla f \rangle = 1$$

This tells us that F, as a function from $\mathbb{R}$ to $\mathbb{R}$, is linear with slope 1. Therefore, φ_0 is the identity diffeomorphism on M_a, and $\varphi_{b-a}(M_a) = M_b$.

We have found a diffeomorphism from M_a to M_b, thus completing the proof. □

We now deal with the case where $f^{-1}[a, b]$ contains a critical point. We will need to use the following lemma.

Lemma 3.3. *Let M be a smooth manifold with corners. Then there exists a smooth manifold M' without corners that is homeomorphic to M and diffeomorphic to M outside of a neighborhood of the corner points. Furthermore, M' is unique up to diffeomorphism.*

We refer the reader to [21] for a proof.

Theorem 3.4. *Let M, f, and a, b be as in Theorem 3.2. If $f^{-1}[a, b]$ contains one critical point of f with index k, then M_b is diffeomorphic to the union of M_a with a k-handle.*

Proof. Let p be the critical point in $f^{-1}(a, b)$, denote its image under f as c, and let k be the index of p. Because of Corollary 2.14, we can assume, up to adjusting a and b to decrease $|a - c|$ and $|b - c|$, that there exists a coordinate neighborhood N_p that intersects both preimages $f^{-1}(a)$ and $f^{-1}(b)$ and that contains no other critical points of f. By the Morse lemma, we can alter the coordinates on N_p such that f takes the form $f(x_1, \dots, x_n) = -x_1^2 - \cdots - x_k^2 + x_{k+1}^2 + \cdots + x_n^2$. With respect to these coordinates, we are able to draw a contour map of f on N_p as in Figure 9.

The Morse lemma allows us to split the n dimensions of our manifold into two subspaces, one of dimension k on which f takes values less than c, and the other of dimension $n - k$ on which f takes values greater than c. Note that the level set $f^{-1}(a)$ in N_p intersects the coordinate axes of $\{x_1, \dots, x_k\}$, and that it does not intersect the coordinate axes of $\{x_{k+1}, \dots, x_n\}$. To see this, imagine standing at p and noting that walking along any of the axes x_i for $i \leq k$ will lead you down towards $f^{-1}(a)$, while walking along any of the axes x_i for $i > k$ will lead you "up". The same logic allows us to represent $f^{-1}(b)$ as intersecting the axes $\{x_{k+1}, \dots, x_n\}$ and avoiding the others.

We will be interested in the intersections of M_a and M_b with N_p, since the local behavior of f at p occurs within N_p. The intersections $N_p \cap M_b$ and $(N_p \cap M_b) - M_a$ are shown in Figure 10.

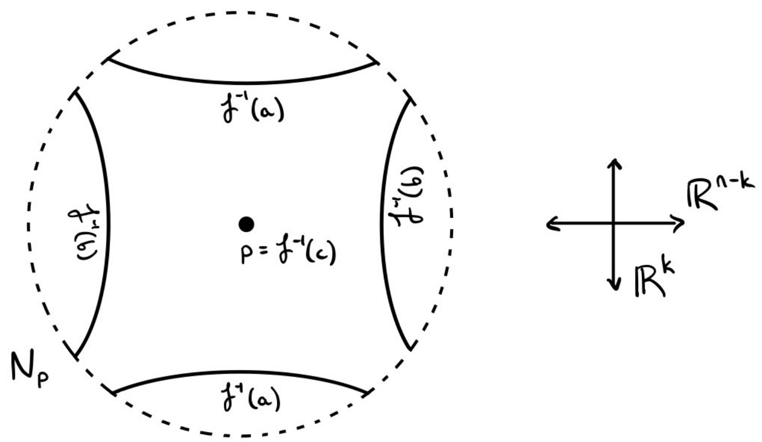

Figure 9: A neighborhood of a nondegenerate critical point p, with local coordinate axes shown separately.

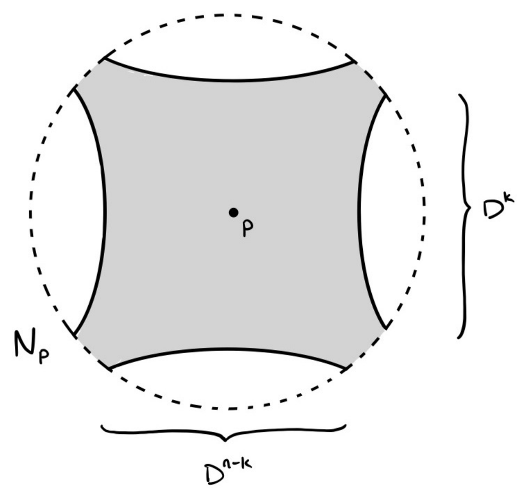

Figure 10: $(N_p \cap M_b) - M_a$.

Let H be the subset $(N_p \cap M_b) - M_a$. In fact, H is, as a topological space, a $k-$handle, just with an unfamiliar shape. Recalling that our goal is to show a diffeomorphism between $M_a \cup H$ and M_b, we apply Lemma 3.3 to round out the corners of H where it does not meet M_a. This is shown in Figure 12. H then is diffeomorphic, as a manifold with corners, to $D^k \times D^{n-k}$.

It now remains to show that $M_a \cup H$ is diffeomorphic to M_b. Rather than construct such a map explicitly, we appeal to Theorem 3.2. The existence of Morse functions (Theorem 2.11) guarantees the existence of another Morse function g such that for some $a' \in \mathbb{R}$, $M_{g=a'} = M_a \cup H$

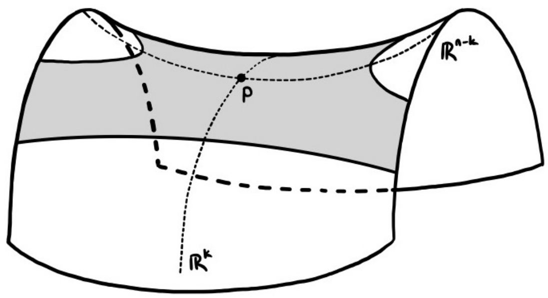

Figure 11: H on a torus.

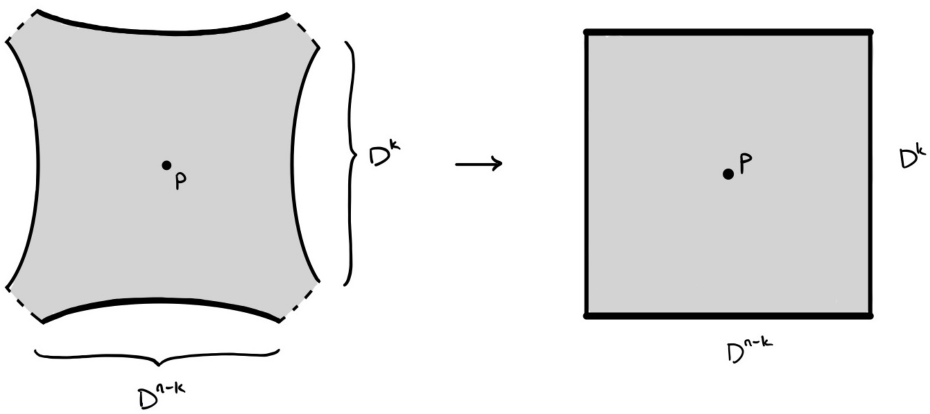

Figure 12: Smoothing corners of H, with attaching region of H to M_a shown in bold.

and $M_{g=b} = M_b$. In particular, g can be chosen such that M does not contain a critical point of g in $g^{-1}[a', b]$ (since f does not by assumption). So by Theorem 3.2, we have that $M_{g=a'} = M_a \cup H$ is diffeomorphic to $M_{g=b} = M_b$. $\square$

We need one more lemma in order to prove the existence of handle decompositions for all compact smooth manifolds.

Lemma 3.5. *Let M be a smooth manifold with f a Morse function on M. Then if p and q are both critical points of f such that $f(p) = f(q)$, then there exists a smooth manifold M' that is diffeomorphic to M such that $f(p) \neq f(q)$.*

We refer the reader to [15] for a proof.

Theorem 3.6 (Existence of handle decompositions)**.** *There exists a handle decomposition for every compact smooth manifold.*

Proof. Let M be a compact smooth manifold. By Theorem 2.11, there exists a Morse function f on M. Because M is compact, we know that there exist $A, B \in \mathbb{R}$ such that $M_A = \{\emptyset\}$ and $M_B = M$. Compactness also guarantees us that there are only finitely many critical points p_i of f, and Lemma 3.5 guarantees that we can adjust M by diffeomorphism such that $f(p_i) \neq f(p_j)$ for all $i \neq j$. Our goal now is to use these critical points to build a handle decomposition of M.

Let L be the total number of critical points on M. Index each critical point p_i such that if $i < j$, then $f(p_i) < f(p_j)$. This way, p_1 is the lowest critical point, and p_L is the highest. For each pair p_i, p_{i+1} for $i = 1, \ldots, L-1$, let $a_i = \frac{f(p_i)+f(p_{i+1})}{2}$. Note that a_i is defined such that the only critical points that the sublevel set M_{a_i} contains are $p_1, \ldots, p_i$. For notation, set $M_0 = \{\emptyset\}$, and set $M_L = M$.

If we compare M_{a_i} with $M_{a_{i+1}}$, we see that $f^{-1}[a_i, a_{i+1}]$ contains exactly one critical point, p_{i+1}. By Theorem 3.4, we have that $M_{a_{i+1}}$ is diffeomorphic to M_{a_i} with the attachment of a k-handle, where k is the index of a_{i+1}. Furthermore, by Theorem 3.2, we have that the topology of two sublevel sets M_{a_i} and M_b for $b > a_i$ only differs when $b > f(p_{i+1})$. Therefore, the sequence $\{M_0, M_1, \ldots, M_{L-1}, M_L\}$ is a handle decomposition for M. □

4 Interlude: Poincaré Duality

Our goal for this section is to illustrate an application of handlebodies to algebraic topology, specifically, Poincaré duality. Note that Poincaré duality deals in the homology and cohomology of manifolds, and therefore is not sensitive to differential structure, only the homotopy type of a manifold. To accommodate this, our first step in proving Poincaré duality is to come up with a way to view a handlebody as a CW complex.

Proposition 4.1. *A handlebody decomposition of a manifold M defines a CW complex X that is homotopy equivalent to M.*

Proof. The idea here is to recognize that up to homotopy equivalence, the structure of a k-handle can be reduced to that of a k-cell where the attaching sphere of the k-handle becomes the boundary of the k-cell glued on. This is made precise by noting that the key information contained in the gluing map of a k-handle can be reduced to the dimension of the attaching sphere A^k and its placement on the manifold.

Recall from Chapter 1 that the attaching sphere of a k-handle was defined to be $\partial D^k \times \{0\}$, which is homeomorphic to S^{k-1}. It is the boundary of the core $D^k \times \{0\}$. Since h^k can be viewed as the direct product of its core with D^{n-k}, which is contractible, h^k deformation retracts onto its core.

This deformation retraction shrinks the attaching region to the attaching sphere A^k. However, this does not change the topology of the attaching map, since the attaching region is just $A^k \times D^{n-k}$. Thus, deformation retracting h^k to its core induces a deformation retraction on the space h^k is attached to, but does not change the topology of either. We therefore say that the k-cell e^k associated to h^k is the core of h^k.

Note that the attaching map of the handle h^k restricted to A^k is now precisely the attaching map of the cell e^k. □

Armed with a cellular description of M, we can now proceed to construct its cellular homology and cohomology. Below, we provide the reader with a brief overview of these theories. For more information, we recommend [9]. It is important to note that this is not intended to be a sufficiently thorough introduction without prior familiarity with homology and cohomology.

Definition 4.2. The *cellular chain complex* of a CW complex X is the chain complex associated to a space X where the $k-$dimensional chain groups $C_k(X)$ are defined to be the free abelian groups generated by the set of k-cells in X.

In a cellular chain complex, the boundary map sends a k-cell e^k to the formal sum of the (k-1)-cells in the image of the attaching map of e^k. This is formalized in the *cellular boundary formula*:

$$\partial_k(e_i^k) = \sum_{j \in J} d_{ij} e_j^{k-1}$$

where d_{ij} is the degree of the composition of the following three maps: the attaching map of e_i^k sending $\partial e_i^k \cong S_i^{k-1}$ into X^{k-1}, the quotient map sending X^{k-1} to X^{k-1}/X^{k-2}, and the collapsing map sending all the copies of S^{k-1} in X^{k-1}/X^{k-2} to a single S_j^{k-1}. The composition of all of these maps defines a single map from S_i^{k-1} to S_j^{k-1}, the degree of which is well-defined.

Again, more information on cellular homology, as well as a proof of the cellular boundary formula from the typical construction of cellular homology, can be found in [9].

To each chain complex we can associate the dual complex, called the cochain complex. For the cellular case, it is defined as follows:

Definition 4.3. The *cochain complex* associated to a cellular chain complex is the chain complex in which the $k-$dimensional chains $C^k(X)$ are defined to be $C^k(X) := \mathrm{Hom}(C_k(X), \mathbb{Z})$.

For a given $f \in C^k(X)$, the composition $f \circ \partial_{k+1}$ defines a homomorphism from $C_{k+1}(X)$ to $\mathbb{Z}$, which is precisely an element of $\mathrm{Hom}(C_{k+1}(X), \mathbb{Z})$, or $C^{k+1}(X)$. We therefore define the *cellular coboundary map* $\delta^k \colon C^k(X) \to C^{k+1}(X)$ as follows:

$$\delta^k(f) = f \circ \partial_{k+1}$$

The chain and cochain complexes of a CW decomposition X of a compact, orientable smooth manifold M can be related using Morse functions on M! The following proposition constructs the foundation for this relationship.

Proposition 4.4. *If $f \colon M^n \to \mathbb{R}$ is a Morse function, then $-f \colon M^n \to \mathbb{R}$ is also a Morse function with the same critical points as f.*

Furthermore, if p is an index k critical point of f, then p is an index $n-k$ critical point of $-f$.

Proof. If f is smooth, then $-f$ is smooth, as it is the composition of f with the map $\mathbb{R} \to \mathbb{R}$ sending x to $-x$.

Let p be a critical point of f. Then on a neighborhood of p with local coordinates $\{x_i\}$, $\frac{\partial f}{\partial x_i}(p) = 0$. Hence $\frac{\partial(-f)}{\partial x_i}(p) = -\frac{\partial f}{\partial x_i}(p) = 0$, and so p is a critical point of $-f$. Furthermore, by our assumption that f is Morse, $\det(\mathcal{H}_f(p)) \neq 0$. But $\mathcal{H}_{-f}(p) = -\mathcal{H}_f(p)$, and so $\det(\mathcal{H}_{-f}(p)) = -\det(\mathcal{H}_f(p)) \neq 0$. So p is a nondegenerate critical point of $-f$. This completes the proof that $-f$ is Morse.

As for the index of a critical point p of $-f$, note that multiplication by -1 of a matrix $\mathcal{H}_f(p)$ flips the sign of all of its eigenvalues. Therefore, since $\mathcal{H}_{-f}(p) = -\mathcal{H}_f(p)$, the number of negative eigenvalues of the $n \times n$ matrix $\mathcal{H}_{-f}(p)$ is $n-$ the number of negative eigenvalues of $\mathcal{H}_f(p)$. $\square$

This lemma leads us to the following key corollary, which forms the foundation for Poincaré duality.

Corollary 4.5. *Up to diffeomorphism, the k-handles of the decomposition associated to f are equal as submanifolds to the $(n-k)$-handles of the decomposition associated to $-f$.*

We are now ready to prove Poincaré duality.

Theorem 4.6 (Poincaré duality)**.** *Let M^n be a closed, orientable manifold. Then for all $0 \leq k \leq n$, $H^{n-k}(M) \cong H_k(M)$.*

Proof. To begin, let f be a Morse function on M. Then by Proposition 4.4, $-f$ is a Morse function on M.

Let $\{M_r\}$ be the handle decomposition of M obtained from f, and let X be the CW complex obtained from $\{M_r\}$. Let $\{W_s\}$ be the handle decomposition of M obtained from $-f$, and let Y be the CW complex obtained from $\{W_s\}$. Note that M is homotopy equivalent to both X and Y.

We will refer to $n-k$-handles of $\{M_r\}$ as x_i^{n-k}, and k-handles of $\{W_s\}$ as y_i^k. Note that we may index these both with the same variable because Corollary 4.5 guarantees us a map from $n-k$-handles of $\{M_r\}$ to k-handles of $\{W_s\}$. In fact, they are equal as submanifolds. Therefore, for a given handle H_i (a $n-k$-handle when viewed in $\{M_r\}$ and a k-handle when viewed in $\{W_s\}$), we define new variables for its core and co-core, since those terms are no longer well defined when switching between decompositions. Define α_i to be the core of H_i seen as y_i^k, or equivalently, the co-core of H_i seen as x_i^{n-k}. Similarly, define β to be the co-core of H_i seen as y_i^k, or the core of H_i seen as x_i^{n-k}. This can be seen in Figure 13.

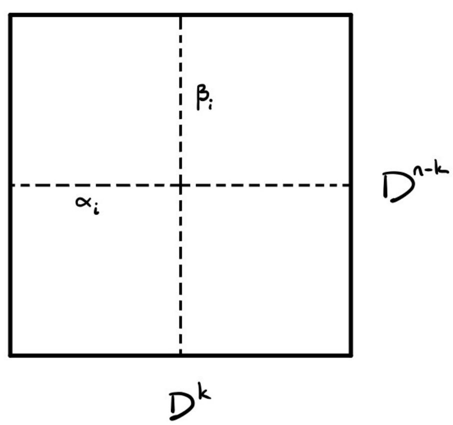

Figure 13: Anatomy of H_i. Note that when H_i is viewed as a k-handle y^k, its core is α_i, but when it is viewed as a $(n-k)$-handle, its core is β_i.

In Proposition 4.1, we saw that to consider a handlebody as a CW complex, each handle was shrunk to its core. We can therefore say that for every handle H_i in $\{M_r\}$ and $\{W_s\}$, β_i, its core as a $n-k$-handle in $\{M_r\}$ is a generator of $C_{n-k}(x)$. Similarly, α_i, its core as a k-handle in $\{W_s\}$, is a generator of $C_k(Y)$. We will use this duality to show that the cellular cochain complex of X is isomorphic to the cellular chain complex of Y.

We begin by defining a homomorphism $\psi_k \colon C_k(Y) \to C^{n-k}(X)$ given by:

$$\psi_k(\alpha_i) = c_i^{n-k}$$

where c_i^{n-k} denotes the element of $C^{n-k}(X)$ that maps β_i to $1 \in \mathbb{Z}$ and all other β_j to 0 for $j \neq i$. Note that because $\{\beta_i\}$ generate $C_{n-k}(X)$ and that X was assumed to be compact, the maps c_i which send β_i to 1 and all other β_j to 0 generate $C^{n-k}(X)$. Furthermore, $\mathrm{rank}(C_{n-k}(X)) = \mathrm{rank}(C^{n-k}(X))$ as free abelian groups.

The equivalence of ranks of these chain and co-chain groups guarantees that ψ_k is bijective for all k. Every handle H_i has exactly one core/co-core α_i and one co-core/core β_i. Furthermore, for every β_i in $C_{n-k}(X)$, there is exactly one c_i in $C^{n-k}(X)$.

To extend the maps ψ_k to an isomorphism of chain complexes, we must show that they commute with the boundary maps of each complex. Specifically, we want to show that the following diagram commutes:

$$\begin{tikzcd}
C_k(Y) \arrow[r, "{}_Y\partial_k"] \arrow[d, "\psi_k"'] & C_{k-1}(Y) \arrow[d, "\psi_{k-1}"] \\
C^{n-k}(X) \arrow[r, "{}_X\delta^{n-k}"'] & C^{n-k+1}(X)
\end{tikzcd}$$

To avoid confusion, we take a moment to give names to the elements of these chain groups:

Elements of $C_k(Y)$ are denoted α_i.
Elements of $C_{k-1}(Y)$ are denoted $\bar{\alpha}_j$.
Elements of $C_{n-k}(X)$ are denoted β_i.
Elements of $C_{n-k+1}(X)$ are denoted $\bar{\beta}_j$.
Elements of $C^{n-k}(X)$ are denoted c_i.
Elements of $C^{n-k+1}(X)$ are denoted $\bar{c}_j$.

Consider first $\psi_{k-1} \circ {}_Y\partial_k$.

The map ${}_Y\partial_k$ sends the core of a k-handle y_i^k to a formal sum of cores of $(k-1)$-handles. The image of α_i, the core of y_i^k, under ${}_Y\partial_k$ is then:

$$\sum_{j=1}^{\mathrm{rank}(C_{k-1}(Y))} A_{i,j}\bar{\alpha}_j$$

Where $\bar{\alpha}_j$ is the core of y_j^{k-1}, and thus a generator of $C_{k-1}(Y)$.

Formally, $A_{i,j}$ is the degree of the attaching map of y_i^k. Geometrically, the coefficients $A_{i,j}$ represent the number of times the attaching region of y_i^k "wraps around" the core of each y_j^{k-1}, with sign determined by the orientation of the cores. However, because the core and co-core of any handle intersect transversely exactly once, the signed number of times y_i^k "wraps around" each y_j^{k-1} is precisely the signed transverse intersection number of α_i with the co-cores of each y_j^{k-1}. Note that these co-cores, which we denote $\bar{\beta}_j$, are the generators of $C_{n-k+1}(X)$.

If we now apply ψ_{k-1} to $\sum_{j=1}^{\mathrm{rank}(C_{k-1}(Y))} A_{i,j}\bar{\alpha}_j$, we obtain the following:

$$\psi_{k-1} \circ {}_Y\partial_k(\alpha_i) = \sum_{j=1}^{\mathrm{rank}(C^{n-k+1}(X))} A_{i,j}\bar{c}_j$$

We now examine ${}_X\,\delta^{n-k} \circ \psi_k$.

Recall that ${}_X\,\delta^{n-k}$ was defined so that ${}_X\,\delta^{n-k}(c_i) = c_i \circ {}_X\partial_{n-k+1}$. The map ${}_X\partial_{n-k+1}$ sends the core of a $(n-k+1)$-handle x^{n-k+1} to a formal sum of cores of $(n-k)$-handles. We can therefore denote the image of $\bar{\beta}_j$ under ${}_X\partial_{n-k+1}$ as follows:

$$\sum_{i=1}^{\mathrm{rank}(C_{n-k}(X))} B_{i,j}\beta_i$$

As before, note that geometrically the coefficients $B_{i,j}$ represent the number of times the attaching region of x_j^{n-k+1} "wraps around" the core of each x_i^{n-k}, with signs determined by orientation. And again, the core and co-core of each x_i^{n-k} intersect transversely exactly once, and so $B_{i,j}$ is equivalent to the signed transverse intersection number of $\bar{\beta}_j$ with the co-cores of each x_i^{n-k}. But the co-core of x_i^{n-k} is precisely α_i! So $B_{i,j}$ = the signed transverse intersection number of α_i with $\bar{\beta}_j = A_{i,j}$. Note that we implicitly used orientability of our manifold here to ensure that the orientations chosen for α_i and β_i are consistent with their boundary components under both ${}_Y\partial_k$ and ${}_X\partial_{n-k+1}$, thus giving a well-defined signed transverse intersection number.

With this in mind, we can rewrite the image of $\bar{\beta}_j$ under ${}_X\partial_{n-k+1}$ as:

$$\sum_{i=1}^{\mathrm{rank}(C_{n-k}(X))} A_{i,j}\beta_i$$

Applying c_i to this sum, we obtain the following form for ${}_X\,\delta^{n-k}$:

$${}_X\,\delta^{n-k}(c_i) = \sum_{j=1}^{\mathrm{rank}(C^{n-k+1}(X))} A_{i,j}\bar{c}_j$$

If we precompose this map with ψ_k, since ψ_k is an isomorphism sending α_i to c_i, we obtain:

$${}_X\,\delta^{n-k} \circ \psi_k(\alpha_i) = \sum_{j=1}^{\mathrm{rank}(C^{n-k+1}(X))} A_{i,j}\bar{c}_j$$

Thus we have:

$$\psi_{k-1} \circ {}_Y\partial_k = {}_X\,\delta^{n-k} \circ \psi_k$$

Therefore, the chain complex of Y is isomorphic to the co-chain complex of X.

Isomorphic chain complexes have isomorphic homology groups, and so we can conclude that $H_k(Y) \cong H^{n-k}(X)$. But recall that X and Y were merely two CW structures on the same manifold, and since the homology of a manifold is independent of its CW structure, we know that $H_k(Y) \cong H_k(X)$.

We can therefore conclude that $H_k(X) \cong H^{n-k}(X)$. $\square$

This proof is particularly neat because it uses only the very rudimentary facts about handle decompositions. As we will see in the next several sections, the theory surrounding handle decompositions is far wider and more complicated than their existence. Moving forward, we will examine some of the techniques used in handlebody theory, and then their applications in various dimensions.

5 Manipulations of Handlebodies

This section is devoted to building the tools for analyzing handlebodies that are foundational for the further study of manifolds in this context. We begin with a relatively simple observation:

Proposition 5.1. *Let M be a smooth manifold. Then there exists a handle decomposition of M such that all k-handles are attached before any higher index handles are attached for $1 \leq k < n$.*

Proof. We will show this by inductively constructing diffeomorphic manifolds M_i such that M_{i+1} has the same handles as M_i, but two of its subsequently attached handles are attached in the opposite order.

Suppose M_i is given, with $M_0 = M$, and with it, a handle decomposition. If there are no handles that are not attached in increasing order, then we're done. If there is one, then we attach handles of M one by one until we reach a handle h^k that is attached immediately after another handle h^l where $k \leq l$.

If the attaching sphere A^k of h^k is disjoint from h^l, then we can shrink the D^{n-k} factor of the attaching region until the two handles are attached completely disjointly, at which point we may as well have attached h^k first. Therefore, we can swap the attaching order of h^k and h^l without changing M at all. In this case, let M_{i+1} be the manifold obtained via this reordered handle decomposition. We can now go back and begin checking its handle ordering from the starting 0-handle again to obtain a decomposition for M_{i+2}. Note that we can not assume that M_{i+1} now has all of its handles in order up to the flipped pair h^k and h^l, and so we must begin checking the order of handles of M_{i+1} from the first h^0.

If the attaching region of h^k is not disjoint from h^l, then we show that $\bar{A}^k$ can be isotoped away from h^l to obtain the case above. First of all, if $\bar{A}^k$ misses the belt sphere B^l of h^l, then as the remaining boundary of h^l to which A^k is attached is diffeomorphic to $B^l \times D^{l-k}$, we can isotope A_k along ∂h^l radially away from B^l to the attaching region of h^l and eventually off of h^l entirely. Furthermore, $\bar{A}^k$ can in fact always be isotoped to not intersect B^l. We see this via a general position argument: in $n-1$-dimensional boundary of an n-dimensional manifold, $A^k \cong S^{k-1}$ has codimension $n-k$, and $B^l \cong S^{n-l-1}$ has codimension l. Their generic intersection, then, has codimension $n-k+l$, and thus dimension $k-l-1$. But $l > k$, so this shows that A^k and B^l do not generically intersect, and so A^k can be separated from B^l in ∂M_i so that the above cases apply. □

While the above proof shows that handles of different indices can be reordered within a given manifold, the fact that the dimension of the generic intersection of the attaching and belt spheres is $k-l-1$ also implies that k-handles can be perturbed to be disjoint from each other and moved around each other. These sorts of isotopies are called handle slides.

Definition 5.2. Let M^n be a handlebody and h_1^k and h_2^k be two handles attached to ∂M. A *handle slide* of h_1^k over h_2^k is an isotopy of the attaching region of h_1^k through $\partial(M \cup h_2^k)$ to another map of the attaching region into ∂M.

Handle slides vastly complicate the theory of handlebodies, as they allow for the manipulation of handles after attaching them. Additionally, since isotopies of handles don't change diffeomorphism type of the manifold they define, handle slides define another way in which a different handle decomposition for the same manifold can be obtained without the need for defining a new Morse function. Some examples of handle slides in various dimensions can be found below in Figure 14.

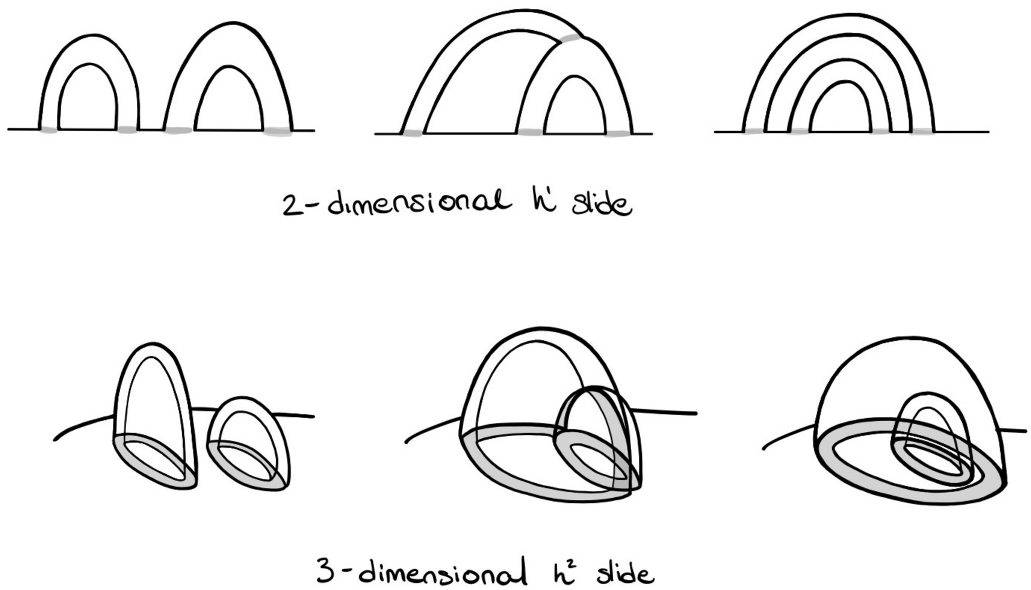

Figure 14: Some handle slides, in various dimensions.

It is possible to obtain, from a Morse function or otherwise, a handle decomposition of a manifold that carries with it more handles than are necessary to describe it. In this case, we introduce the notion of handle cancellation.

Definition 5.3. We say that two handles h^k and h^{k+1} *cancel* if the attaching sphere of h^{k+1} intersects the belt sphere of h^k geometrically exactly once.

Intuition behind handle cancellation is best found in Figure 15. On the level of Morse functions, a handle decomposition with a pair of cancelling handles can be thought of as a sort of unnecessary ripple in the function that conveys no extra information about the manifold. To make this precise, we prove the following.

Proposition 5.4. *Let M be a manifold and suppose we attach two cancelling handles h^k and h^{k+1} to obtain N. Then M is diffeomorphic to N.*

Proof. If the attaching sphere of h^{k+1} intersects the belt sphere of h^k geometrically exactly once, then up to isotopy, we can split the attaching region of h^{k+1} into two parts: the part attached to ∂M, and the part attached to h^k. Note that both parts are diffeomorphic to $D^k \times D^{n-k-1}$, as they glue together along $\partial D^k \times D^{n-k-1}$ to form the $S^k \times D^{n-k-1}$ shape we expect of A^{k+1}. With this in mind,

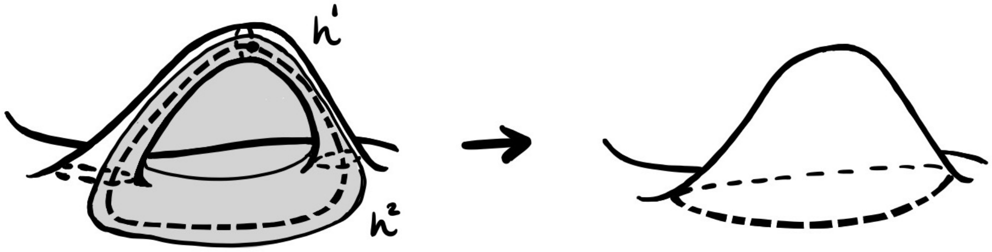

Figure 15: A 3-dimensional 1-2 cancelling handle pair.

we can consider $h^k \cup_\phi h^{k+1}$ by themselves, before being attached to M. Attaching h^{k+1} to h^k pushes a $D^k \times D^{n-k-1}$-shaped segment of $\partial h^k \setminus A^k$ into the interior of $h^k \cup_\phi h^{k+1}$. The remaining part of A^{k+1} is then another copy of $D^k \times D^{n-k-1}$, but when we attach $h^k \cup_\phi h^{k+1}$ to M, it will be attached as well. By attaching h^{k+1} to h^k, then, we have effectively decreed that a portion of ∂h^k that is not in A^k but that A^k bounds will also be attached to M. Hence, the subset of $h^k \cup_\phi h^{k+1}$ that attaches to M is diffeomorphic to $D^k \times D^{n-k-1}$, and so $M \cup_\phi (h^k \cup_\phi h^{k+1}) \cong M \natural (h^k \cup_\phi h^{k+1})$. Finally, note that as $h^k \cup_\phi h^{k+1}$ is simply two copies of D^n glued along a copy of D^{n-1}, $M \natural (h^k \cup_\phi h^{k+1}) \cong M$. □

We can use this notion to strengthen Proposition 5.1 to streamline handle decompositions even further.

Proposition 5.5. *Let M^n be a connected, compact, smooth manifold. Then there exists a handle decomposition of M with a unique 0-handle, a unique n-handle, and all intermediate handles attached in order of increasing index.*

Proof. We start by reordering the handles of M as in Proposition 5.1. Pick some 0-handle h_0^0 to be the base 0-handle in this new decomposition. For each extra 0-handle h_i^0 in the decomposition, note that because 1-handles are the only handles whose attaching regions are disconnected, either h_i^0 is connected to h_0^0 by a 1-handle, or it is connected to some other 1-handle-linked chain of 0-handles that terminates at h_0^0 by a 1-handle. In either case, the belt sphere of h_i^0 is simply ∂h_i^0, and so any 1-handle attached to it on one end necessarily cancels with it. We can therefore use the assumed connectedness of M to cancel all 0-handles with 1-handles attached on one end to them except our chosen h_0^0, after which all 1-handles will be attached at both ends to h_0^0. This gives us a unique 0-handle. To obtain a unique n-handle, simply turn the corresponding Morse function to this modified handle decomposition upside down as in Proposition 4.4. This allows us to use connectedness to cancel all extra n-handles with $n-1$-handles until a unique n-handle is obtained. □

In addition to this useful fact, we can also use handle cancellation to invent yet another way of producing new handle decompositions: handle trading.

Definition 5.6. Given a manifold M with a handle decomposition and a handle h^k in it, if there exists an attaching map for a $k+1$-handle into M such that h^k and h^{k+1} cancel, then we can *trade* h^k for a $k+2$-handle to obtain a new manifold N which is diffeomorphic to M by attaching a $k+1$, $k+2$-cancelling handle pair to M such that h^{k+1} cancels with h^k.

Note that such a map existing is by no means a given. Knowing that one exists is actually stronger than knowing the homology of M.

The machinery worked through in this section turns out to be enough to classify manifolds up to diffeomorphism.

Theorem 5.7 (Cerf 1970)**.** *Two smooth manifolds, possibly with boundary, are diffeomorphic if and only if they are related by a series of handle slides, addition or removal of cancelling handle pairs, and isotopies to their handle attaching regions.*

This is a fundamental result in geometric topology that will inform our further study of manifolds from here on out. A proof can be found in [5]. In fact, all of these manipulations of handlebodies will make some sort of an appearance later on when we examine the h-cobordism theorem. Handle trading, in particular, plays an important role.

6 The H-Cobordism Theorem and the Generalized Poincaré Conjecture

6.1 The H-Cobordism Theorem

High dimensional manifolds are unwieldy creatures. Homeomorphic manifolds of dimension at least 4 can carry different smooth structures. Without visual intuition, proving anything about them is all the more difficult. Because of this, techniques for analyzing higher dimensional manifolds vary drastically from those used in low-dimensional topology. That being said, one of the great successes made by mathematicians in high-dimensional topology, the h-cobordism theorem, is built on handle decompositions, and armed with the tools presented in the previous section, we are ready to tackle it. We begin with the notion of a cobordism.

Definition 6.1. A *cobordism* between two n-manifolds M and N is a $n+1$-manifold W and a splitting of ∂W such that $\partial W = M \sqcup N$.

We say that W is an *h-cobordism* between M and N when the inclusion maps $M \hookrightarrow W$ and $N \hookrightarrow W$ are both homotopy equivalences.

A cobordism really is just a manifold, but we refer to them as cobordisms when we want to emphasize them as relating their boundary components in some way. Nevertheless, they have some familiar structure.

Definition 6.2. A *relative handle decomposition* of a cobordism W^{n+1} between M^n and N^n is a sequence of $n+1$-manifolds $W_0 \dots W_m$ such that $W_0 = M \times [0,1]$, $W_m = W$, and W_{k+1} is obtained from W_k by attaching an $n+1$-dimensional handle.

Such a decomposition exists by the work we have already done on normal handle decompositions. In the proof of Theorem 3.6, nowhere did we use the assumption that our manifold had no boundary. Typically, then, handle decompositions of cobordisms are obtained by setting one boundary component to be the "low" component, where the Morse function takes its lowest value, and the other to be the "high" component.

Proposition 6.3. *Proposition 5.1, handle rearrangement, holds for relative handle decompositions.*

Proof. Suppose W^{n+1} is a cobordism between M^n and N^n, with a relative handle decomposition of W. Then with our Morse function taking its minimum value on M, we can treat M as the boundary of a 0-handle, and follow the same procedure as in Proposition 5.1 to slide 1-handles off of higher index handles until they are attached to M, and do the same for each increasing index. If N is nonempty, then there will be no $n+1$-handle in the decomposition; rather, N will be the boundary of the handles attached to M of index n and lower. $\square$

Theorem 6.4 (h-cobordism theorem, Smale 1962)**.** *Let M, N be simply connected manifolds of dimension $n \geq 5$ and let W be an h-cobordism between them. Then W is diffeomorphic to $M \times [0,1]$.*

Proof. The basic plan of attack for this proof is the following: if W is an h-cobordism between simply connected manifolds, then $\pi_1(W) \cong H_i(W, M; \mathbb{Z}) \cong 0$ for all i. Therefore, the homology coming from the handles of W relative to M should totally vanish. Now $M \times [0,1]$ has the same fundamental group and relative homology groups, and that manifold has a handle decomposition with no handles at all. Therefore, our goal is to use the algebraic structure of W, combined with techniques from Section 5, to slide and cancel handles in the decomposition of W until we are left with a handle-less decomposition, at which point we will have shown that W is diffeomorphic to $M \times [0,1]$.

Step 1 of this procedure is easy, as we have already proven it: apply Proposition 6.3 to rearrange the handles in W so that there are no 0- or $n+1$-handles, and the rest of the handles are attached in increasing index order.

Step 2 is also relatively simple. In lieu of cancelling them directly, we trade the 1- and n-handles for 3- and $n-2$-handles. To do this, pick a 1-handle in M. Since M is simply connected and of dimension greater than or equal to 5, the handle bounds an embedded disk that we can thicken into a cancelling 2-3 handle pair. This allows us to cancel our 1-handle with the 2-handle, but leaves a 3-handle for us to deal with. We can then perform the same procedure on M with its Morse function flipped, which turns k-handles into $(n+1)-k$-handles and so allows us to trade n-handles for $n-2$-handles.

Step 3 of the proof is to slide the remaining ordered handles until they appear algebraically cancelling. Specifically, recall that we can obtain the homology of M from its handle decomposition by taking cellular k-chain groups as generated by the cores of the k-handles and the boundary map ∂ being determined by the attaching maps. If we take a matrix presentation for ∂_k for each k with respect to the basis generated by the handle decomposition obtained from Step 1, we can therefore realize handle manipulations as actions on this matrix. In particular, given two k-handles h_i^k and h_k^j, a slide of h_i^k over h_j^k in the agreeing orientation adds the $i^{\text{ th}}$ row to the $j^{\text{ th}}$ row of ∂_k and the $i^{\text{ th}}$ column to the $j^{\text{ th}}$ column of ∂_{k+1}. Reversing the orientation of a handle multiplies its corresponding row and column by -1. Thus, because we can simultaneously perform Gaussian elimination on each boundary operator over $\mathbb{Z}$ (for a detailed explanation of how to do this, see page 60 of [19]), we can perform the corresponding handle modifications until the matrix forms of the boundary maps corresponding to the handle decomposition are all populated with 0's except for 1's on the diagonal. This tells us that each remaining handle in the new decomposition is paired up with another handle either one above or one below it in index such that their algebraic intersection is 1. Note that we can talk about their algebraic intersection as a number because the two submanifolds we're considering are really the attaching sphere of a k-handle (dimension $k-1$) and the belt sphere of a $k-1$-handle (dimension $n-k+1$). These submanifolds are of complementary dimension in the n-dimensional boundary of our $n+1$-dimensional cobordism, and so they intersect generically in points. To say that they have algebraic intersection 1, then, is to say that the signs of the points in which they intersect sum to 1.

Algebraic intersection 1, however, is not enough to cancel pairs of handles. We need geometric intersection 1, and because algebraic intersection is sensitive to the sign of intersection points, two handles that algebraically intersect once might intersect geometrically some odd number of times with $+1, -1$ pairs of intersection points. Therefore, before we can cancel each pair, we need to make sure that we can actually eliminate these extraneous intersection points geometrically.

Step 4 is this cancellation. Consider a pair of handles with index k and $k-1$ (recall that as we took care of our 1- and n-handles, $3 \leq k \leq n-1$) with algebraic intersection 1, and suppose that they intersect geometrically more than once. Then the extra geometric intersections must algebraically cancel, and so we can pick two intersection points of opposite sign. Call them p and q. We can now pick two paths from p to q, denoted α and β, such that α lies entirely in the attaching sphere A^k of our k-handle and β lies in the belt sphere B^{k-1} of its paired $k-1$-handle. In particular, by pushing α slightly away from A^k along its normal bundle and β away from B^{k-1} along its normal bundle, we can see that if the complement of $A^k \cup B^{k-1}$ in the boundary of the handlebody is simply connected, then $\alpha \cup \beta$ bounds a disk in the complement.

Now the union $A^k \cup B^{k-1}$ is $\max\{k-1, n-k+1\}$-dimensional, and so its codimension in the handlebody boundary is $\min\{k-1, n-k-1\}$-dimensional. Now if the codimension is at least 3, then the complement of $A^k \cup B^{k-1}$ is automatically simply connected. To see this, pick a loop in the complement, and note that it bounds a disk in the entire boundary. However, if $A^k \cup B^{k-1}$ have codimension at least 3, then we have enough room to isotope the bounded disk away from A^k and B^k until it lies entirely in their complement.

Now by this reasoning, in the case of 2-3 and $(n-1)-(n-2)$ cancelling pairs, it is not clear that this complement is actually simply connected. It turns out that it is, but to see this, we refer the reader to [19], as the argument is more involved. We will therefore continue assuming that the complement of $A^k \cup B^{k-1}$ is simply connected, and that as a result, $\alpha \cup \beta$ bounds a disk, which we call the Whitney disk and denote D. Now because the boundary has dimension at least 5, we can take D to be smoothly embedded, since smooth embeddings of m-dimensional manifolds are dense in the space of maps to $\mathbb{R}^{2m+1}$. Finally, we can use D as a guide to isotope A^k away from B^{k-1} until they do not intersect. Specifically, if we thicken the Whitney disk to n dimensions by ε in each dimension, then this isotopy is given by smoothly extending to A^k an isotopy of the path α across D to a path parallel to β pushed out of B^{k-1} by ε. This is shown in Figure 16.

Note that because the interior of D was assumed to be disjoint from A^k and B^{k-1}, this isotopy introduces no more intersections of A^k with B^{k-1}, and in fact eliminates the intersections at p and q. Therefore, using this technique, we can geometrically eliminate algebraically cancelling pairs of points, until each intersecting pair of handles geometrically intersects only once, allowing us to cancel them. After this procedure, we are left with a handle-less decomposition of W relative to M, meaning that it is diffeomorphic to $M \times [0, 1]$. $\square$

One might hope that the h-cobordism theorem might be true in dimensions less that 5. There are several obstructions to this approach, however, notably the failure of the smooth Whitney trick in dimension 4. It has been shown that the h-cobordism theorem as stated above is false in dimension 4. Amazingly, the *topological* version of the h-cobordism theorem is true in dimension 4, which is a consequence of Freedman's classification, which we'll discuss later. The proof of this fact, as well as Freedman's classification, is beyond the scope of this paper. We will look more closely at his results, however, later on in Section 8.1. For now, we turn to one of the oldest problems in

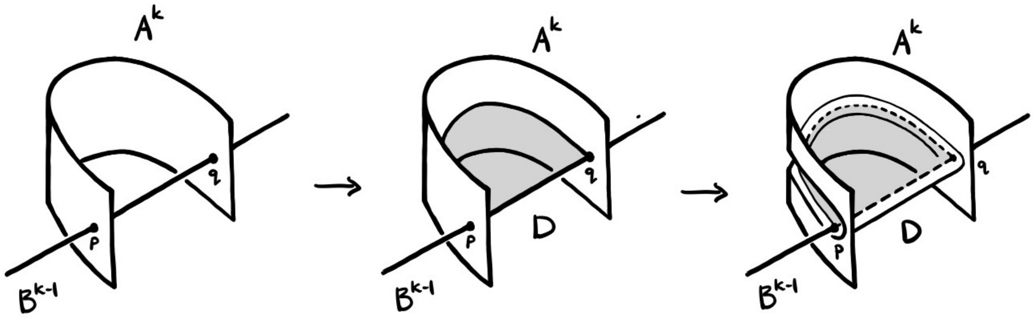

Figure 16: Isotopy along the Whitney disk.

geometric topology, and its relationship to h-cobordisms.

6.2 The Generalized Poincaré Conjecture

One of the ways in which the h-cobordism theorem was used almost immediately after Smale proved it is in the proof of the topological generalized Poincaré conjecture. For the unfamiliar reader, we recall Poincaré's original conjecture:

Theorem 6.5 (Poincaré conjecture 1904, proven by Perelman 2003)**.** *Every 3-manifold that is homotopy equivalent to the 3-sphere is homeomorphic to the 3-sphere.*

We'll briefly take a detour back to the world of 3-manifolds to prove that it can be equivalently stated in a different way:

Proposition 6.6. *The Poincaré conjecture is equivalent to the statement that every simply connected, closed 3-manifold is homeomorphic to S^3.*

Proof. Though it is stated in such a way as to make its relationship to the Poincaré conjecture more apparent, the point of this proposition is that for closed 3-manifolds, the properties of being simply connected and being homotopy equivalent to S^3 mutually imply each other. It is a standard result that the lower homotopy groups of spheres vanish, and so $\pi_1(S^3)$ is trivial. For the reverse direction of implication, let M be a simply connected, closed 3-manifold. Then since $H_1(M;\mathbb{Z}) \cong \mathrm{Ab}(\pi_1(M))$, $H_1(M;\mathbb{Z}) = 0$. By Poincaré duality, we deduce that $H^2(M;\mathbb{Z}) = 0$, and by the Universal Coefficient Theorem for cohomology, $H^1(M;\mathbb{Z}) \cong 0$ as well. We can then apply Poincaré duality again to get that $H_2(M;\mathbb{Z}) = 0$, and so we have deduced:

$$H_k(M;\mathbb{Z}) \cong \begin{cases} \mathbb{Z} & k = 0,3 \\ 0 & k = 1,2 \end{cases}$$

Now by the Hurewicz theorem, since H_1 and H_2 both vanish, $H_3(M;\mathbb{Z}) \cong \pi_3(M)$, and so we can represent a generator for $H_3(M;\mathbb{Z})$ by a map $f\colon S^3 \to M$. In particular, f induces isomorphisms on H_k for all k, and so, noting that both M and S^3 are both simply connected and can be given CW structures, we can apply a corollary of Whitehead's theorem (see [9], Ch. 4) to get that f is actually a homotopy equivalence. □

We will not discuss the proof of the Poincaré conjecture here; rather we will investigate its generalizations to higher dimensions and other categories.

Conjecture 6.7 (generalized Poincaré conjecture)**.** *Every n-manifold that is homotopic to the standard n-sphere is $\mathscr{C}$-isomorphic to the standard n-sphere, where $\mathscr{C}$ is one of the following categories: $\mathscr{Top}, \mathscr{PL}, \mathscr{Diff}$.*

The veracity of the various forms of the generalized Poincaré conjecture are mostly known; a table of the conjectures can be found in Table 6.2.

Proposition 6.8. *The h-cobordism theorem implies the generalized Poincaré conjecture in dimensions 6 and higher in $\mathscr{Top}$, but not necessarily in $\mathscr{PL}$ or $\mathscr{Diff}$.*

The proof of this result depends on the following lemma, commonly known as Alexander's trick, named for J. W. Alexander.

Lemma 6.9 (Alexander)**.** *Any homeomorphism of S^{n-1} can be extended to a homeomorphism of D^n.*

We can now proceed to the proof of the topological generalized Poincaré conjecture.

Proof of Prop. 6.8. Let M^n be a smooth manifold that is homotopy equivalent to S^n for $n \geq 6$. Select points p and q in M with disjoint open neighborhoods N_p and N_q that are homeomorphic to the open disk D^n, and let $N = M \setminus N_p, N_q$. Then as the boundaries of N_p and N_q in M are each copies of S^{n-1}, we have that N is a cobordism between two copies of S^{n-1}.

As M is homotopy equivalent to the n-sphere, we can deduce that the images of ∂N_p and ∂N_q in S^n under this homotopy equivalence are homotopic to copies of S^{n-1}, and so the homotopy equivalence restricts on N to an h-cobordism between copies of S^{n-1} inside S^n. Pulling back by the homotopy equivalence, we therefore have that N is also an h-cobordism between the two copies of S^{n-1}.

Now recalling that $n \geq 6$, and so $n-1 \geq 5$, we can apply the h-cobordism theorem to see that N is diffeomorphic to $S^{n-1} \times [0,1]$. Let f denote this diffeomorphism. We can now extend f to some map $\bar{f}\colon M \to S^n$ via the following: because N was obtained from M by removing two open disks, we can realize M as the quotient of $N \cup D^n \cup D^n$ by some gluing diffeomorphisms of S^{k-1}. The composition of these diffeomorphisms with f therefore yields two diffeomorphism g_p and g_q from S^{n-1} to $S^{n-1} \times [0,1]$, one for each boundary component, $S^{n-1} \times \{0\}$ and $S^{n-1} \times \{1\}$. Relaxing our consideration of g_p and g_q to homeomorphisms instead of diffeomorphisms of S^{n-1}, we apply Alexander's trick to get extensions $\bar{g}_p, \bar{g}_q$ from D^n to the manifold W obtained from $S^{n-1} \times [0,1]$ by gluing in copies of D^n to $S^{n-1} \times \{0\}$ and $S^{n-1} \times \{1\}$ via the identity map on ∂D^n. Finally, we can piece together a homeomorphism $\bar{f}$ from M to W given by:

$$\bar{f}(x) = \begin{cases} \bar{g}_p(x) & x \in N_p \\ \bar{g}_q(x) & x \in N_q \\ id & x \in N \subset M \end{cases}$$

We also see that W, being a capped-off S^{n-1}-cylinder, is homeomorphic to S^n. Thus, we have that M is homeomorphic to S^n. $\square$

In literature and historically, the h-cobordism theorem refers to the version stated in Theorem 6.4, in which the dimension of the cobordant manifolds is greater than or equal to 5, and equivalence is shown in $\mathscr{Diff}$. However, like the generalized Poincaré conjectures, one can also consider the question of whether h-cobordisms are trivial in other dimensions and various categories. We therefore state the following conjecture, many cases of which are proved or disproved, and one of which, notably, is open:

Conjecture 6.10. *Let M, N be simply connected manifolds of dimension n and let W be an h-cobordism between them. Then W is $\mathscr{C}$-isomorphic to $M \times [0,1]$, where $\mathscr{C}$ is one of the following categories: $\mathscr{Top}, \mathscr{PL}, \mathscr{Diff}$.*

To avoid confusion, we will refer to the theorem proved by Smale as the h-cobordism theorem, and the lower dimensional statements in various categories with the statement "n-dimensional h-cobordisms are $\mathscr{C}$-trivial".

In lower dimensions, it becomes possible to engineer a full equivalence between the $n-1$-dimensional h-cobordism conjectures and the n-dimensional Poincaré conjectures in various categories. We will take a moment to examine two such equivalences: $n = 3$, in which $\mathscr{Top}$, $\mathscr{PL}$, and $\mathscr{Diff}$ are equivalent, and $n = 4$, in which $\mathscr{Top}$ differs from $\mathscr{Diff}$ and $\mathscr{PL}$. Although their proofs are very similar, the $n = 4$ case requires slightly more work and an extra lemma, and so we state them separately.

Proposition 6.11. *The 3-dimensional $\mathscr{Top}$ Poincaré conjecture is equivalent to 2-dimensional h-cobordisms being $\mathscr{Top}$-trivial.*

Note again that in dimensions 2 and 3, the $\mathscr{Top}$, $\mathscr{PL}$, and $\mathscr{Diff}$ categories are all equivalent, and so we might as well have replaced $\mathscr{Top}$ with either $\mathscr{PL}$ or $\mathscr{Diff}$.

Proof of Proposition 6.11. The proof that 2-dimensional h-cobordisms being $\mathscr{Top}$-trivial implies the 3-dimensional Poincaré conjecture follows exactly the same structure as the proof of Proposition 6.8. In particular, note that while Smale's h-cobordism theorem gives the diffeomorphism type of h-cobordisms, we ultimately only used their homeomorphism types in obtaining the result of the generalized Poincaré conjecture. Here too we are only proving the topological Poincaré conjecture, and so we only need $\mathscr{Top}$-triviality of h-cobordisms.

For the other direction of implication: suppose that homotopy 3-spheres are standard and let M and N be simply connected 2-manifolds with W an h-cobordism between them. By the classification of surfaces, M and N are both homeomorphic to S^2, and so we can consider W to be an h-cobordism between copies of S^2. Consider the manifold W' obtained by gluing in copies of D^3 to the two copies of S^2 via the identity homeomorphism on ∂D^3. Applying the Seifert-van Kampen theorem from algebraic topology, we see that $\pi_1(W')$ is trivial, and so W' is a simply connected, compact 3-manifold, and by Proposition 6.6, W' is therefore homotopy equivalent to S^3. But by assumption, W' is now homeomorphic to S^3, and so by restricting this homeomorphism to W, we have that W is homeomorphic to S^3 with two disjoint copies of D^3 removed, which is simply an S^2-cylinder. Hence, we have that W is homeomorphic to $S^2 \times [0,1]$. □

Before moving on to the $n = 4$ case, we state a theorem of Cerf that is effectively a strengthening of Alexander's trick.

Theorem 6.12 (Cerf 1968)**.** *Any diffeomorphism of S^3 extends to a diffeomorphism of D^4.*

Cerf's original proof ([4]) of this theorem is in French and not immediately accessible. Another proof can be found in [7], although it uses contact structures and other techniques outside the scope of this paper. We therefore omit a discussion of the proof, and move on to the $n = 4$ case. Here, we will actually use the 3-dimensional Poincaré conjecture.

Proposition 6.13. *The 4-dimensional $\mathscr{C}$ Poincaré conjecture is equivalent to 3-dimensional h-cobordisms being $\mathscr{C}$-trivial, where $\mathscr{C}$ is either $\mathscr{Top}$ or $\mathscr{PL}/\mathscr{Diff}$.*

Proof. Again, the proof that 3-dimensional h-cobordisms being $\mathscr{C}$-trivial implies the 4-dimensional $\mathscr{C}$ Poincaré conjecture follows the same structure as Proposition 6.8 and Proposition 6.11. However, in this case, we can leverage Theorem 6.12 as a stronger version of Alexander's trick to extend the map from our cobordism N to an S^3-cylinder to a diffeomorphism from the full manifold M to S^4 to take care of the smooth case. In the topological case, exactly the same argument as before holds.

For the reverse direction: suppose that homotopy 4-spheres are $\mathscr{C}$-standard and let M and N be simply connected 3-manifolds and W an h-cobordism between them. By Proposition 6.6 and the work of Perelman, we can take M and N to both be S^3, leaving W an h-cobordism between two copies of S^3. As in the $n = 3$ case, we glue in copies of D^4 to the two copies of S^3 along the identity on ∂D^4, extending smoothly with Theorem 6.12 if desired, to obtain W', a simply connected 4-manifold. Now there do exist simply connected 4-manifolds that are not homotopy equivalent to S^4, so we must do a little more work to show that W' is. Because W is an h-cobordism between simply connected boundary components, $H_k(W, M; \mathbb{Z}) = 0$ for all $k \geq 0$. From the long exact sequence of a pair, we deduce that $H_k(W; \mathbb{Z}) \cong H_k(M; \mathbb{Z})$ for all k, which is, of course, simply the k^{th} homology group of S^3. From here, we can apply the Mayer-Vietoris sequence twice, once for each copy of D^4 glued to W, to obtain $H_k(W'; \mathbb{Z})$. Omitting the actual calculation, the resulting homology groups are:

$$H_k(W'; \mathbb{Z}) \cong \begin{cases} \mathbb{Z} & k = 0, 4 \\ 0 & k = 1, 2, 3 \end{cases}$$

We can now follow the same logic used in Proposition 6.6: by the Hurewicz theorem, elements of $H_4(W'; \mathbb{Z})$ are represented by maps of S^4 into W', and so a generator for $H_4(W'; \mathbb{Z})$ induces isomorphisms on homology. Such a map is therefore, by Whitehead's theorem, a homotopy equivalence, and we have that W' is homotopy equivalent to S^4. By assumption, then, W' is $\mathscr{C}$-isomorphic to S^4, and so we can restrict this isomorphism to $W \subset W'$ to obtain a $\mathscr{C}$-isomorphism between W and $M \times [0, 1] \cong S^3 \times [0, 1]$. $\square$

Again, the topological case of the 4-dimensional Poincaré conjecture and, thus, the topological triviality of 3-dimensional h-cobordisms follows from the work of Freedman. The smooth and piecewise-linear cases, being equivalent, are currently open.

For the first example of an exotic smooth structure on a sphere, see Milnor in [16]. Further discussions of exotic smooth structures in high dimensions are both fantastically interesting and far beyond the scope of this paper.

7 Knots and 3-Manifolds

There are a lot of reasons to care about 3-manifolds. Not only are they easier to picture than higher dimensional ones, they also carry with them extra intriguing structure that has been a focus

	$\mathscr{Top}$	$\mathscr{PL}$	$\mathscr{Diff}$
$n = 1$	True	True	True
2	True	True	True
3	True (Perelman 2003)	True	True
4	True (consequence of Freedman 1982)	Open	Open
5	True (Smale 1961, [20])	True, equivalent to $\mathscr{Diff}$	True (Kervaire-Milnor 1963, [10])
6+	True (Smale 1962, consequence of h-cobordism theorem. Prop. 6.8)	True	True in dimension 6, but in general, false. Finitely many unique smooth structures (Kervaire-Milnor 1963, [10])

Table 1: The state of the generalized Poincaré conjectures, in various dimensions and categories. All manifolds are assumed to be smoothable.

of study for low-dimensional topologists since Poincaré. Sadly, in this paper we explore only a very small portion of 3-manifold theory, in particular its relationship to knots and links. The reason for this is not rooted in spite at the glory of 3-manifolds, but rather that their connections to knot theory become their most important aspect to us later on, when we consider 3-manifolds arising as boundaries of 4-manifolds.

7.1 A Brief Introduction to Knot Theory

Definition 7.1. A *knot* in S^3 is an embedding $K\colon S^1 \to S^3$.

While a knot refers to an embedding, usually up to ambient isotopy, they are represented by *knot diagrams*, or projections of the knot onto $\mathbb{R}^2$, where the information about the shape of the knot is encoded in the crossings. Below are some knot diagrams.

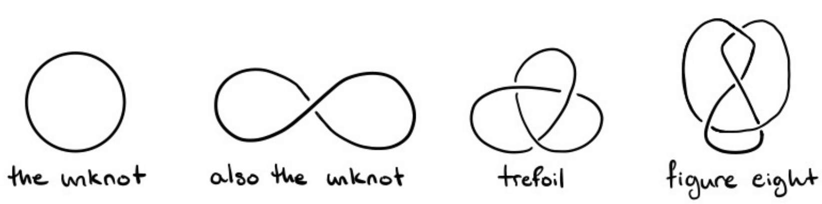

Figure 17: Some knot diagrams.

It is important to note that many knot diagrams may define the same knot up to isotopy. For instance, some equivalent unknots are shown in Figure 18.

Classifying knots is a mathematical endeavor with a long and storied history that we do not have time to go into. A good reference for the interested reader, particularly one who is new to knot theory and enjoys interesting, open-ended exercises, is Colin Adams' aptly named book, The

Figure 18: Some surprising unknots.

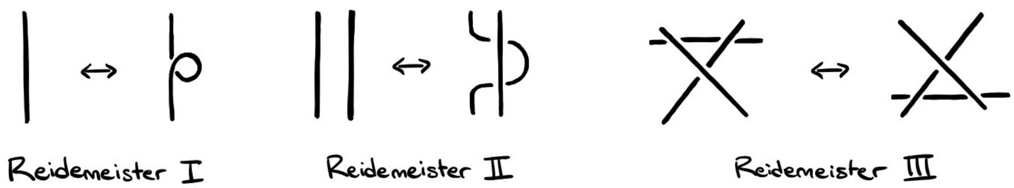

Figure 19: The Reidemeister moves.

Knot Book [1]. Others may prefer Rolfsen ([17]). We'll now review some basic definitions from knot theory that will be of use later. Central to a mathematical description of knots is the notion of a *crossing*, which is exactly what it sounds like. Knots are by no means determined by the number of crossings they have in a given diagram (see Figure 18), and so we need to introduce a way to modify knot diagrams in such a way that they represent isotopic embeddings of S^1 in S^3.

Definition 7.2. The *Reidemeister moves* are a collection of three modifications one can make to a knot diagram that leave the isotopy type of the knot invariant. They are shown in Figure 19.

It is common in the literature to call the three moves R1, R2, and R3.

It may be trivial to see that the Reidemeister moves leave a knot invariant, but they are still helpful to knot theorists, as they allow us to test possible knot invariants against isotopy in a controlled way. Specifically, a knot property is only a isotopy invariant if it remains the same under modifications made to the knot by all three Reidemeister moves.

Another important bit of knot theory notation is the parity of a crossing, which is meant to distinguish between "right-handed" and "left-handed" crossings. To define the sign of a crossing in a knot, we will need to give our knot diagram an orientation by picking a point on it and a direction along the knot from that point.

Definition 7.3. The *sign* or *parity* of a crossing in an oriented knot diagram is said to be either +1 or -1, depending on which of the following diagrams it looks like:
For those who have studied a bit of physics, this choice is consistent with the right-hand rule for

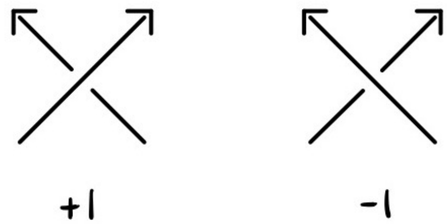

Figure 20: Parity of crossings.

determining torque or magnetic field direction. As such, it is also consistent with the standard cross product of vectors in $\mathbb{R}^3$.

It is important to note that reversing the orientation of a knot diagram does not affect the sign of its crossings. Taking a mirror image, however, will. We can now define the writhe:

Definition 7.4. The *writhe* of an oriented knot diagram is the sum of the signs of its crossings.

By our above comment, the writhe of a knot is invariant under choice of orientation. As if knots themselves were not complicated enough, we will also deal with links in this paper.

Definition 7.5. A *link* in S^3 is a collection of maps $K_1, \ldots, K_n \colon S^1 \to S^3$ such that the images of these maps are all disjoint.

As in the case with knots, we draw *link diagrams*, which can be altered by the Reidemeister moves without changing the isotopy class of the link they represent.

A very rough invariant we can give to links is the linking numbers of their individual components:

Definition 7.6. The *linking number* $\mathrm{lk}(K_1, K_2)$ of two link components K_1 and K_2 is $k/2$, where k is the sum of signed crossings of K_1 with K_2.

We can aggregate all the linking numbers of the various components of a link into symmetric a *linking matrix*. The (i, j) entry of the linking matrix is the linking number of K_i with K_j, with the diagonal entries being 0.

Once again, reversing the orientation of a link component will negate its linking number with each other component. We have one final definition to consider this section.

Definition 7.7. A *Seifert surface* for a knot $K \colon S^1 \to S^3$ is an embedding $S \colon D^2 \to S^3$ such that S restricts to K on ∂D^2.

Theorem 7.8 (Frankl-Pontryagin 1930, algorithm Seifert 1934)**.** *Embedded Seifert surfaces exist for all knots.*

An explanation of Seifert's algorithm can be found in [8].

Seifert surfaces are useful tools for knot classification, although we will not use them this way. Instead, we will use the following property.

Proposition 7.9. *The linking number of two knots K_1 and K_2 is equal to the signed intersection of K_1 with a Seifert surface for K_2.*

A proof of this can be found in [17].

As we move on to looking at 3-manifolds, only some of the above definitions and theorems are used. The remainder will be used later on, when we look at Kirby calculus and 4-manifolds.

7.2 Heegaard Splittings and Dehn Surgery

This section gives an introduction to some of the basic techniques used in 3-manifold theory. In particular, we will prove the Lickorish-Wallace theorem, which shows the remarkable relationship between links and closed, orientable 3-manifolds. We begin by defining Heegaard splittings, which were not originally conceived using handle decompositions, but which handles allow for a very clean definition of.

Definition 7.10. A *Heegaard splitting* of a 3-manifold M is a triple (H_1, H_2, f) where H_1 and H_2 are homeomorphic genus g handlebodies, $f\colon \partial H_1 \to \partial H_2$ is a homeomorphism on their boundaries, and M is homeomorphic to $H_1 \cup_f H_2$.

The following theorem is a consequence of the work of Moise and Bing on the triangulation of 3-manifolds, but the proof we present is not theirs; rather, it falls quite neatly out of the work of Smale and handlebody theory.

Theorem 7.11. *Every closed, orientable 3-manifold admits a Heegaard splitting.*

Proof. Let M be a closed, orientable 3-manifold. By Proposition 5.5, we can find a handle decomposition of M consisting of a unique 0-handle, a unique 3-handle, and some other 1- and 2-handles. Consider the corresponding Morse function $f\colon M \to \mathbb{R}$.

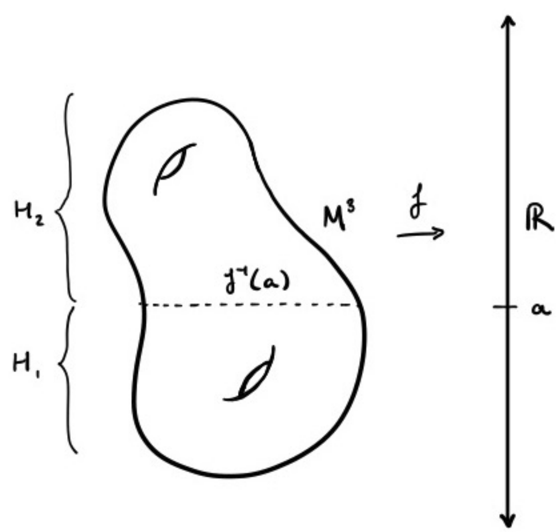

Figure 21: Obtaining a Heegaard splitting from a Morse function.

Because of our reordering of the handles, there exists some $a \in \mathbb{R}$ such that the sublevel set M_a contains only the 0-handle and 1-handles, with all 2-handles and the 3-handle sitting above

$f^{-1}(a)$. (Note that by our earlier argument about flipping our Morse function upside down, $M \setminus M_a$ is also homeomorphic to a 0-handle with some 1-handles attached.) We can therefore cut along $f^{-1}(a)$ in M to obtain two handlebodies $H_1 = M_a$ and $H_2 = M \setminus M_a$ with common boundary $f^{-1}(a)$ homeomorphic to some closed orientable genus g surface Σ_g. H_1 and H_2 therefore are also genus g handlebodies. Furthermore, since M was obtained by cutting along Σ_g, there exists a homeomorphism of surfaces $h\colon \partial H_1 \to \partial H_2$ such that $M \cong H_1 \cup_h H_2$. Thus, (H_1, H_2, h) is a Heegaard splitting of M. □

Another way to obtain closed 3-manifolds is via Dehn surgery on S^3.

Definition 7.12. Let K be a knot in S^3. The process of cutting out a tubular neighborhood of K homeomorphic to a solid torus T and gluing it back in via some homeomorphism on its boundary torus is called *Dehn surgery*. The specific homeomorphism ϕ of ∂T can be specified with two coprime integers p and q as follows: let γ be a meridinal curve (a simple closed curve that bounds a disk in T). Then $\phi\colon \partial T \to \partial T$ sends γ to a curve going around ∂T p times meridinally and q times longitudinally.

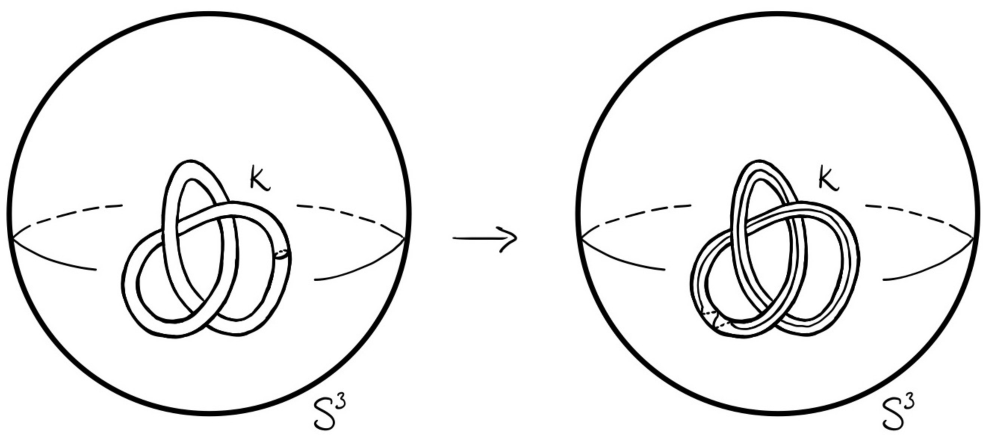

Figure 22: Dehn surgery on S^3.

Remarkably, every closed 3-manifold can be obtained via Dehn surgery on some link in S^3, and we spend the remainder of this section proving that theorem. Before we can prove the ubiquity of Dehn surgery as a tool for obtaining 3-manifolds, we will need the following lemma. Be warned: it requires the Lickorish twist theorem, which is well beyond the scope of this paper.

Lemma 7.13. *Let H be a 3-dimensional handlebody and let ϕ be a homeomorphism of ∂H. Then there exists a collection of disjoint solid tori $V_1, \ldots, V_r$ in H such that ϕ extends to a homeomorphism $\bar{\phi}$ on $H \setminus V_1, \ldots, V_r$.*

Proof. By the Lickorish twist theorem, ϕ is isotopic to a product of Dehn twists about a collection of simple closed curves $\{\alpha_1, \ldots, \alpha_r\}$ on ∂H.

Let T_1 be a Dehn twist about α_1 such that T_1 is the identity on ∂H except for on an open neighborhood of α_1 in ∂H, denoted $N(\alpha_1)$. We'll use $N(\alpha_1)$ to find our first solid torus V_1. To do this, consider the obvious embedding of $\partial H \times [0, \varepsilon]$ into H such that $\partial H \times \{0\}$ maps to ∂H. Then $\partial H \times \{\varepsilon\}$ is a copy of ∂H sitting inside H. We can now project $N(\alpha_1)$ down from $\partial \times \{0\}$ to $\partial H \times \{\frac{\varepsilon}{2}\}$ and to $\partial H \times \{\varepsilon\}$ to get two copies of $N(\alpha_1)$ sitting inside H, one on top of the other. Then, take V_1 to be the solid torus (with corners, but this distinction disappears up to homeomorphism) bounded above by $N(\alpha_1) \times \{\frac{\varepsilon}{2}\}$, below by $N(\alpha_1) \times \{\varepsilon\}$, and on the right and left $\partial N(\alpha_1) \times [\frac{\varepsilon}{2}, \varepsilon]$.

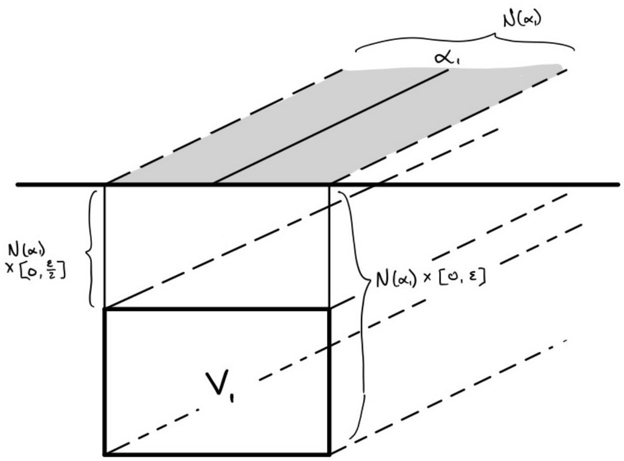

Figure 23: V_1 sitting inside H.

We can now extend T_1 to all of $H \setminus V_1$ as follows: on $N(\alpha_1) \times [0, \frac{\varepsilon}{2}]$, define $\bar{T}_1(x,t) = (T_1(x), t)$, and outside of that region, take $\bar{T}_1$ to be the identity. Note that $\bar{T}_1$ is still continuous: since T_1 is a continuous map and is the identity outside of $N(\alpha_1)$, $\bar{T}_1$ is continuous inside $N(\alpha_1) \times [0, \frac{\varepsilon}{2}]$, is the identity map on $\partial(N(\alpha_1)) \times [0, \frac{\varepsilon}{2}]$, and so agrees with the identity map outside of $N(\alpha_1) \times [0, \frac{\varepsilon}{2}]$.

We can now proceed to extend T_i to $\bar{T}_i$ on $H \setminus V_i$, where V_i is obtained by the same process as above, with the following caveat that we need to account for the case where α_i and α_j intersect for some $i > j$. In this case, simply define V_i to be below V_j inside H; that is, increase ε for V_i until the two solid tori do not intersect.

Define $\bar{\phi}$ to be the composition $\bar{T}_r \circ \cdots \circ \bar{T}_1$ restricted to $H \setminus V_1, \ldots, V_r$. Each $\bar{T}_i$ is the identity outside of $N(\alpha_i) \times [0, \varepsilon]$ and on the boundary and continuous inside $N(\alpha_i) \times [0, \varepsilon]$, and so we only need to convince ourselves that continuity is preserved on those subsets of H that lie above or between two solid tori corresponding to intersecting α_i and α_j. The relevant regions for this case are shown in Figure 25.

On region $A \subset \partial H$, the composition $\bar{T}_j \circ \bar{T}_i$ restricts to the composition of two Dehn twists $T_j \circ T_i$, which we know is continuous. Furthermore, as T_i agrees with the identity map on $\partial N(\alpha_i)$ and the same for T_j, $T_j \circ T_i$ agrees with T_i on the rest of $N(\alpha_i)$ and T_j agrees with T_j on the rest of $N(\alpha_j)$. By the definition of $\bar{T}_j \circ \bar{T}_i$, we also see that continuity and agreement on boundary is preserved on $B \subset H$. Note that on the interface between B and $C \subset V_i$, $\bar{T}_j \circ \bar{T}_i$ defines a nontrivial homeomorphism on ∂V_i, but that as C is contained in V_i, $\bar{\phi}$ is not defined on C even though $\bar{T}_j$ is. On D, $\bar{T}_i$ is simply the identity, since D is beneath V_i, and so $\bar{T}_j \circ \bar{T}_i = \bar{T}_j$ on this subset. Thus, continuity on the interior and agreement with the surrounding maps are preserved on D. Finally, region E is in V_j, and so will not be in the domain of $\bar{\phi}$.

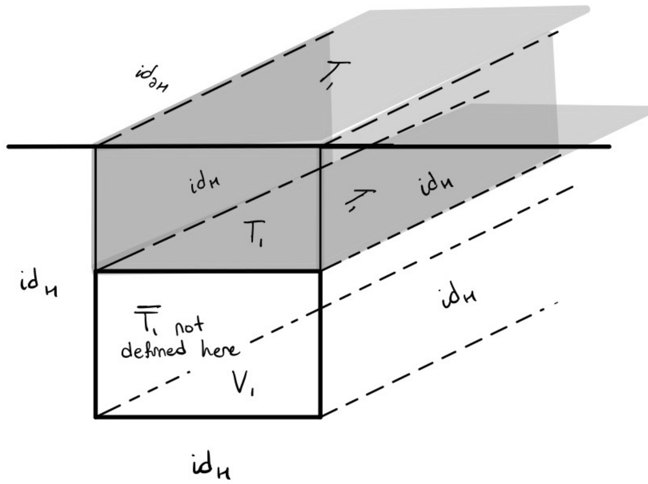

Figure 24: Extending T_1 to $\bar{T}_1$.

As all homeomorphisms of ∂H can be taken to be twists about transversely intersecting simple closed curves, the two-twist case is the only extra case we need to consider. $\square$

After convincing ourselves of the lemma above, the following result follows surprisingly quickly:

Theorem 7.14 (Lickorish-Wallace)**.** *Every compact, orientable 3-manifold M can be obtained via Dehn surgery on a link in S^3.*

Proof. Let (H_1, H_2, f) be a Heegaard splitting of genus g for M, and let (J_1, J_2, h) be the standard genus g Heegaard splitting of S^3. Then H_1, H_2, J_1, and J_2 are all homeomorphic, but in particular, we want to consider the homeomorphism $\phi\colon H_1 \to J_1 \subset S^3$. Note in particular that ϕ restricted to ∂H_1 is a (probably nontrivial) homeomorphism sending ∂H_2 to ∂J_2.

If we could extend ϕ to a homeomorphism $\bar{\phi}$ on the interior of H_2 in addition to H_1 and their common boundary, then we would have $M \cong S^3$, which is probably not the case. Instead, we will have to settle for extending it to a subset of H_2 using Lemma 7.13.

Let $V_1, \ldots, V_r$ be the solid tori in H_2 from Lemma 7.13 that $\bar{\phi}$ does not extend to. Then for each V_i, consider $\bar{\phi}$ restricted to ∂V_i. Each of these restricted maps is a homeomorphism of ∂V_i, which is therefore a homeomorphism of a torus. If we then remove the solid tori bound by $\bar{\phi}(\partial V_i)$ for each i from J_2 and glue them back in according to the restriction $\bar{\phi}|_{V_i}$, we can extend $\bar{\phi}$ to all of H_2. In removing and regluing the solid tori in J_2, however, we have performed Dehn surgery on $J_1 \cup_h J_2 \cong S^3$, and therefore have extended $\bar{\phi}$ to a homeomorphism from $H_1 \cup_f H_2 = M^3$ to a Dehn-surgered S^3. $\square$

It should be noted that this procedure can also be modified so as to produce only integral surgery coefficients. A proof can be found in [18].d

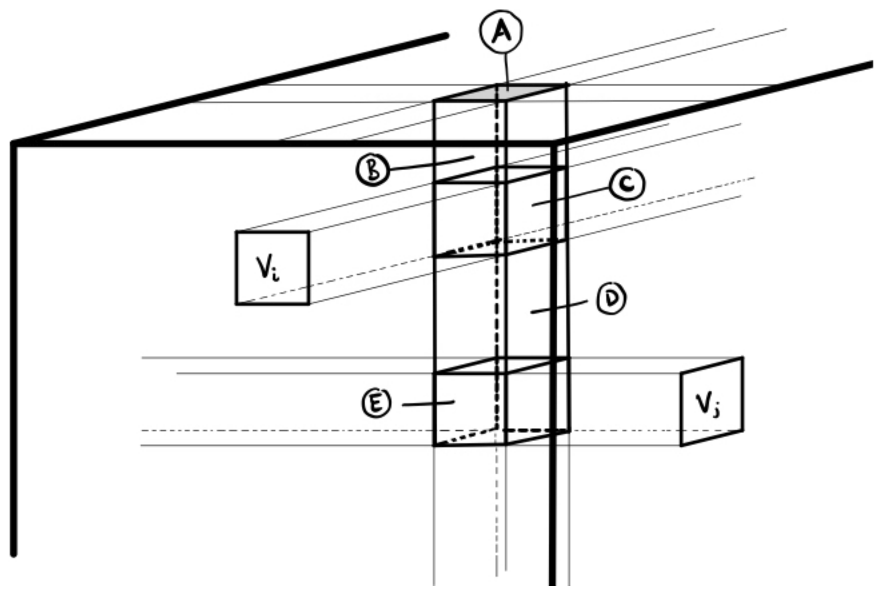

Figure 25: Map of regions between two solid tori V_i and V_j.

8 4-Manifolds

After having examined manifolds of both higher and lower dimensions, we arrive at the turning point: dimension 4. 4-manifolds are particularly special. In many contexts, they resist both analytical tools from 3-manifolds and high-dimensional techniques. Ultimately, we will be working towards an understanding of Kirby calculus, which is a handle-oriented approach to understanding 4-manifolds. However, for now, we'll take a look at some of the earlier techniques, namely, intersection forms. This section should provide some context for the study of 4-manifolds.

8.1 Intersection Forms, 4-Manifold Cobordisms, and Related Results

One of the fundamental tools in classical 4-manifold theory is the intersection form, which is used to understand the interactions of surfaces embedded in 4-manifolds. We take a moment to formally define this form:

Definition 8.1. The *intersection form* of a closed 4-manifold M is a particular bilinear function $Q_M\colon H_2(M;\mathbb{Z}) \times H_2(M;\mathbb{Z}) \to \mathbb{Z}$. We will denote the intersection form of two 2-homology classes α and β as $\alpha \cdot \beta$.

There are several different, but equivalent, ways to construct Q_M. Those who are familiar with the cup product on cohomology may treat the following proposition as a definition of the intersection form.

Proposition 8.2. *Let α, β be classes in $H_2(M;\mathbb{Z})$ and α^*, β^* denote their Poincaré dual classes in $H^2(M;\mathbb{Z})$. Then $\alpha \cdot \beta = (\alpha^* \cup \beta^*)[M]$.*

The more geometric definition requires a little more work to construct, but is more intuitive from the topological perspective. In short, the intersection form $\alpha \cdot \beta$ is the signed intersection number of two embedded surfaces in M representing α and β. To state this as a definition, however, requires us to justify representing all 2-homology classes with embedded surfaces.

Proposition 8.3. *Every homology class α in $H_2(M;\mathbb{Z})$ for M a closed, orientable, smooth 4-manifold can be represented by an embedded orientable surface in M.*

Proof. We begin by noting that for any smooth manifold N^l, regardless of dimension, every 2-homology class is represented by a map of a surface into N. To see this, note that so long as N is smooth, it admits a piecewise linear structure and is hence triangulable. We can therefore consider the 2-homology classes as generated by simplicial maps of cycles into the triangulated N. From here, it is simple to check that the cycle condition requires that open neighborhoods in cycles are homeomorphic to open sets in $\mathbb{R}^2$, even neighborhoods of vertices.

It remains to be shown that in a 4-manifold M, 2-homology classes can be represented by embedded surfaces, not simply immersed ones. Consider a 2-homology class α and an immersed surface Σ representing it, as above. If Σ cannot be perturbed to eliminate self-intersections, then as Σ is 2-dimensional and M is 4-dimensional, Σ generically intersects itself in 0-dimensional manifolds, and so it can be perturbed to reduce self-intersections to transverse intersections at points. We can now eliminate these self-intersection points by replacing Σ with some Σ' of higher genus.

To see this, model a neighborhood N_p of a self-intersection point p as a copy of the open disk in $\mathbb{R}^4$. The two subsets of Σ that intersect in N, then, can be modeled as two copies of $\mathbb{R}^2$ meeting in a point at the origin in $\mathbb{R}^4$. Their traces on the boundary of this open disk - that is, in the boundary of N_p - are copies of S^1 sitting inside S^3 in a Hopf link. We can therefore eliminate the self-intersection point by replacing Σ with some surface Σ' where, inside N_p, these two copies of S^1 are joined by an embedding of the Seifert surface of the Hopf link (a twice-twisted annulus), rather than by the disks they bound in Σ. As the annulus can be embedded in S^3, Σ' misses the origin in our model of N_p entirely, and so we have elmininated the intersection point p. In doing this, we have actually replaced Σ with Σ' with one higher genus, as we have linked two copies of S^1 that previously bound disjoint disks in Σ. ☐

Many classical matrix properties correspond to properties of intersection forms in matrix form. For instance, the rank and definiteness of the intersection form matrix are often referenced in 4-manifold literature. One such property that holds particular importance in the theory is the signature of the intersection form.

Definition 8.4. Consider the intersection form Q_M for a 4-manifold in matrix form. The *signature* σ of Q_M is the number of positive eigenvalues of this matrix minus the number of negative eigenvalues.

While geometrically unintuitive, the signature allows us to detect the bounding and co-bounding properties of 4-manifolds. For instance, the following is not too difficult to prove, and a proof can be found in [19]:

Proposition 8.5. *If M^4 bounds some oriented 5-manifold W, then the signature of Q_M is 0.*

Amazingly, the converse to this proposition is also true.

Theorem 8.6 (Rokhlin)**.** *If the signature of Q_M is 0, then M bounds a 5-manifold.*

This statement is significantly more involved to prove so we omit a discussion of it. A proof can be found in [11] for interested parties.

The intersection form itself was shown over the years to be a stronger and stronger invariant. This can be seen in the following theorems:

Theorem 8.7 (Milnor 1958, from work of Whitehead)**.** *Two simply connected 4-manifolds are homotopy equivalent if and only if their intersection forms are isomorphic.*

A proof outline for this theorem can be found in [19].

Theorem 8.8 (Wall 1964)**.** *Two simply connected 4-manifolds with isomorphic intersection forms are h-cobordant.*

Finally, with the work of Freedman on 4-dimensional h-cobordisms, we obtain the holy grail of 4-manifold classification:

Theorem 8.9 (Freedman 1982)**.** *Let M and N be simply connected 4-manifolds, and let W be an h-cobordism between them. Then W is homeomorphic to $M \times [0,1]$.*

In addition to solving the 4-dimensional $\mathfrak{Top}$ Poincaré conjecture, this result actually gave a complete topological classification for simply connected smooth manifolds.

Corollary 8.10. *Suppose M and N are simply connected, smooth 4-manifolds with isomorphic intersection forms. Then M and N are homeomorphic.*

This set of results, combined with Wall's theorem, opened up the possibility for the existence of h-cobordant 4-manifolds that are not diffeomorphic. Examples of such manifolds were found by Simon Donaldson ([6]). In fact, Freedman's theorem extends beyond the statement given above, and can tell us, given a simply-connected 4-manifold homeomorphism type, how many smooth structures that type supports. Explaining this, however, requires further understanding of intersection forms and other invariants called Kirby-Siebenmann invariants, which we will not go into.

8.2 The Handlebody Perspective

We now move into looking at the handle decomposition view of 4-manifolds. As we will be working with 4-dimensional k-handles substantially, it is worth reminding ourselves what they look like. Table 8.2 is provided for quick reference.

Index	Structure	Attaching sphere	Core
0	$\{0\} \times D^4$	$\emptyset$	$\{0\}$
1	$D^1 \times D^3$	S^0	D^1
2	$D^2 \times D^2$	S^1	D^2
3	$D^3 \times D^1$	S^2	D^3
4	$D^4 \times \{0\}$	S^3	D^4

Table 2: 4-dimensional handle lookup table.

Recall that when attaching a handle h^k to a manifold M via some attaching map, the diffeomorphism class of the resulting manifold is determined by the isotopy class of the attaching map on the entire attaching region $S^{k-1} \times D^{n-k}$. To aid in our later discussions of handlebodies, we break this data into two pieces of information: the isotopy class of the attaching map *restricted to the attaching sphere* $S^{k-1} \times \{0\}$, and the way in which the D^{n-k} factor "twists" around the sphere, called the *framing*.

Definition 8.11. Let M^n be a handlebody and suppose we wish to attach an n-dimensional k-handle $h^k = D^k \times D^{n-k}$ with attaching sphere $A^k = S^{k-1}$ to it. Then a *framing* of the handle is a choice identification of the normal bundle to A^k with the normal bundle to its image under the attaching map.

It behooves us, being interested in 4-manifolds, to classify the possible framings for 4-dimensional k-handles.

Lemma 8.12. *There is a bijection between isotopy classes of framings of* S^{k-1} *in* ∂M^n *and elements of* $\pi_{k-1}O(n-k)$.

Proof. It is a fact from vector bundle theory that in general, given a fiber bundle E with base space B^n and fibers $\mathbb{R}^m$, then there are $|\{f\colon B\times\mathbb{R}^m \to B\times\mathbb{R}^m\}| = |\{g\colon B \to \mathrm{GL}(m) \text{ fixing B pointwise}\}|$ unique isotopy classes of $\mathbb{R}^m$-bundles over B. Treating the D^{n-k} factor of the attaching region as a topological subset of the $n-k$-dimensional normal bundle over S^{k-1}, we can therefore say that framings of S^{k-1} are classified by maps $g\colon S^{k-1} \to \mathrm{GL}(n-k)$ based at the identity $\mathbb{1} \in \mathrm{GL}(n-k)$ up to homotopy. Additionally, since $\mathrm{GL}(n-k)$ deformation retracts onto $\mathrm{O}(n-k)$, we have a bijection between framings of S^{k-1} and $\pi_{k-1}(\mathrm{O}(n-k), \mathbb{1})$. ☐

With this in mind, we can now calculate the number of isotopy classes of framings of 4-dimensional handles by computing $\pi_{k-1}(O(4-k), \mathbb{1})$ for $0 \leq k \leq 4$. A table of these homotopy groups is shown below.

Index	Structure	Homotopy group to compute	Number of isotopy classes of framings
0	$\{0\} \times D^4$	N/A	1 (disjoint union)
1	$D^1 \times D^3$	$\pi_0(\mathrm{O}(3), \mathbb{1})$	2 (oriented and unoriented)
2	$D^2 \times D^2$	$\pi_1(\mathrm{O}(2), \mathbb{1})$	In bijection with $\mathbb{Z}$ ($\mathrm{SO}(2) \cong S^1$)
3	$D^3 \times D^1$	$\pi_2(\mathrm{O}(1), \mathbb{1})$	1 ($\mathrm{SO}(1) \cong \{0\}$)
4	$D^4 \times \{0\}$	N/A	1 (Att. map $\phi\colon S^3 \to S^3$ is isotopic either to id. or reflection, and both extend over D^4)

Table 3: Framings of 4-dimensional handles.

We will revisit this table in Section 9. For now, however, note that the possible framings for 4-dimensional handles are all finite, with the exception of 2-handles. Later on, we will see that this is indicative of the fact that for orientable 4-manifolds, their complexity is mainly due to their 2-handles.

We need one more theorem about 4-dimensional handlebodies before moving on.

Theorem 8.13 (Laudenbach-Poénaru)**.** *Let* $N = \sharp^m S^1 \times S^2$ *and* $M = \natural^m S^1 \times D^3$ *with* $\partial M = N$. *Then any self-diffeomorphism of* N *can be realized by composition of handle slides and ambient isotopies of handles in* M.

A proof of this fact can be found in [12]. It is not incredibly long, but quite technical, and thus will be omitted. However, it leads us to the following important result:

Corollary 8.14. *Let H be a 4-dimensional handlebody comprising only 0-, 1-, and 2-handles, and let M and N be closed, oriented 4-manifolds obtained from H by attaching 3- and 4-handles. Then M and N are diffeomorphic.*

Proof. By Proposition 5.5, we can arrange the handles of M and N and rescale corresponding Morse functions on them such that there exists some $a \in \mathbb{R}$ such that the sublevel set of the Morse functions of M and N below a are both diffeomorphic to H. We can then cut M along the preimage of a to obtain a copy of H glued via some diffeomorphism ϕ of ∂H to the union of some 3- and 4-handles, denoted B_M. Similarly, cutting N along the preimage of a yields another diffeomorphism ψ gluing another union B_N of 3- and 4-handles to ∂H. By turning the handle decompositions for M and N upside down, we see that B_M and B_N can both be viewed as the union of some 1-handles to a 0-handle, and that therefore, their boundaries are both diffeomorphic to the connect sum of some number of copies of $S^1 \times S^2$. As ϕ and ψ are both diffeomorphisms from ∂B_M and ∂B_N to ∂H, we can conclude that not only are B_M and B_N diffeomorphic, but that $\phi \circ \psi^{-1}$ defines a diffeomorphism of $\sharp^m S^1 \times S^2$. We can now apply Theorem 8.13 to realize this diffeomorphism by handle slides in B_M, meaning that M and N themselves are related by handle slides and are therefore diffeomorphic. $\square$

This corollary is foundational to developing Kirby calculus, as it allows us to represent closed, oriented 4-manifold with only information about their 1- and 2-handles.

9 Kirby Calculus

9.1 Handles and Diagrams

The goal of Kirby diagrams is to represent 4-manifolds on a page in an intuitive way using handle decompositions. We know from Proposition 5.5 that any smooth n-manifold has a handle decomposition with a unique 0- and n-handle, and Corollary 8.14 tells us that to represent a closed, orientable 4-manifold, we need only specify the 0-, 1-, and 2-handles. From this, we deduce that a closed, orientable 4-manifold is defined by its 1- and 2-handles, as we can always take there to be a unique 0-handle. We therefore introduce the following notation for 1- and 2-handles.

Figure 26: On the left, a 1-handle, with reflecting axis shown. On the right, a 2-handle, with attaching region knotted in a trefoil.

1-handles are represented by pairs of spheres, with the convention that they are identified via a reflection between a (usually omitted) axis between them. The intuition is that these spheres bound 3-balls, which form the attaching region of a 1-handle. One can imagine the actual 1-handle as a 4-dimensional "bridge" coming out of the page and connecting the two balls. 2-handles are represented by their attaching spheres, as maps of S^1 into S^3, or knots.

For the most part, we will be dealing with oriented 4-manifolds, and even 4-manifolds which have no 1-handles at all. Therefore, we omit a framing on the 1-handles, and assume that they are all oriented. For the 2-handles, recall that there are a $\mathbb{Z}$'s worth of possible framing coefficients for a chosen orientation of the attaching sphere S^1. We can therefore either write an integer next to every 2-handle to denote its framing, or we can use the *double strand notation*, which shows, with a parallel knot, a section of the normal bundle to the attaching sphere as it twists around.

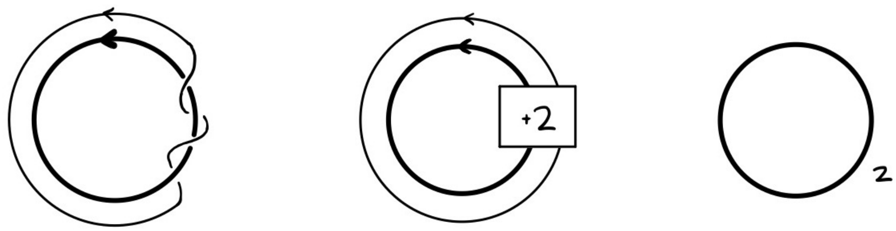

Figure 27: Equivalent ways of representing framings on an unknotted 2-handle.

Now as diffeomorphism type is invariant under isotopy of attaching regions of handles, these diagrams, called Kirby diagrams, represent the same manifold if they are related by Reidemeister moves on the 2-handle attaching regions or isotopies of the 1-handle attaching region in S^3.

There are several notes to make on this notation. First of all, to take the connect sum of two closed 4-manifolds, we need only draw their Kirby diagrams next to each other. This makes seeing when 4-manifolds split as connect sums quite easy. However, there are some subtleties to consider as well. It is tempting to say that an attached 2-handle with a knotted attaching sphere K is 0-framed if, using the double strand notation, the two strands do not link. This is misguided.

Definition 9.1. The framing on a knotted 2-handle attaching region represented by a knot diagram K where the double strand notation appears trivial is called the *blackboard framing* of K.

In fact, the blackboard framing of a 2-handle represented by a knot diagram is actually the writhe of that knot diagram, meaning that not only is it in general nonzero, but it is not even invariant under the Reidemeister moves. Therefore, when attaching 2-handles, their framings can be specified as the blackboard framing only if a knot diagram representing the 2-handle is specified as well.

Another point to mention is that we can see cancelling handles in Kirby diagrams and modify the diagrams as such without changing the diffeomorphism type of the represented manifold. The diagrams shown in Figure 28 represent 1-2 and 2-3 cancelling handle pairs.

There is an unfortunate ambiguity to Kirby diagrams the way we have defined them. It is quite intuitive to draw 1-handles as pairs of D^3s, however, this representation does not actually respect isotopy invariance of 2-handle framings. For instance, consider the following procedure, shown in

Figure 28: Cancelling Kirby diagrams. Left, a cancelling 1-2 pair, for any knot K and any framing coefficient k. Right, a 2-3 cancelling pair. Note that the 3-handle is not pictured.

Figure 29, by which we change the framing of a 2-handle passing through a 1-handle with isotopies alone.

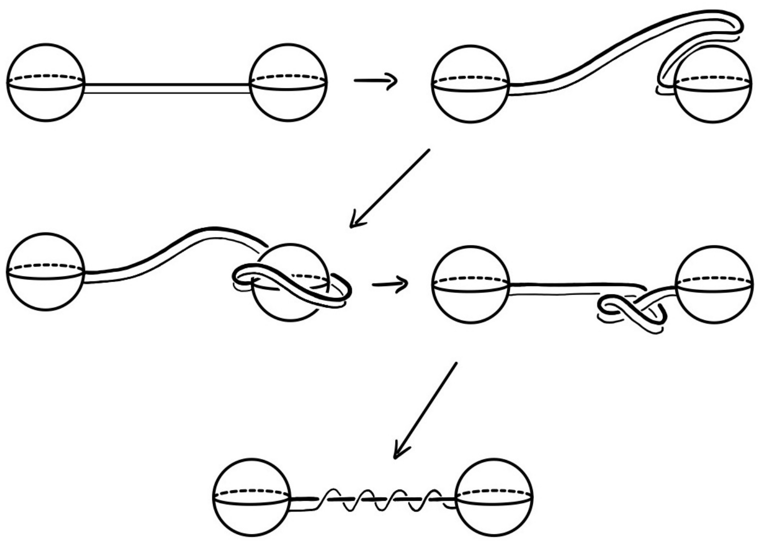

Figure 29: An isotopy changing the framing of a 2-handle passing through a 1-handle.

We therefore make the following observation regarding 1-handles. Attaching a 1-handle to a manifold M is equivalent to carving out a 2-handle-shaped hole in M "below" the attaching region of the 1-handle. This is illustrated in Figure 30.

We therefore change our notation for 1-handles from a pair of 3-balls representing the attaching region to a circle with a dot (referred to as a "dotted circle", despite the fact that the line type of the circle is not actually dotted), which represents the boundary of the disk which, if pushed from the boundary into the 4-manifold and thickened to 4 dimensions, would be carved out. 2-handles passing through 1-handles are now drawn as linked with dotted circles. This is shown in Figure 31.

With this new notation, we can represent 4-manifolds uniquely up to diffeomorphism just with link diagrams containing dotted and undotted components.

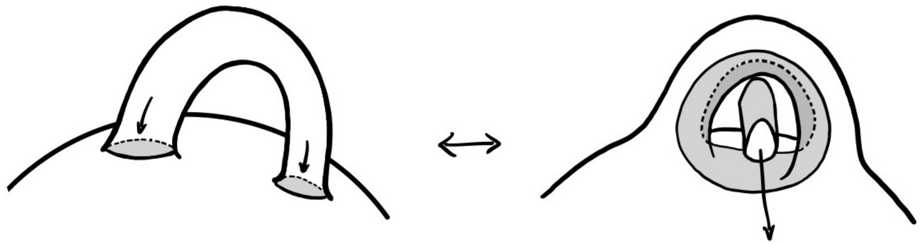

Figure 30: Equivalence of attaching a 1-handle and carving out a 2-handle.

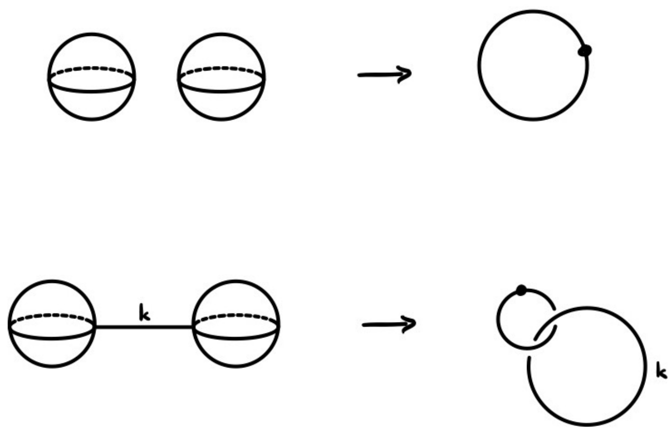

Figure 31: Dotted circle 1-handle notation.

Before proceeding to introduce more moves on Kirby diagrams, we pause for a moment to introduce some examples and intuition to keep in mind while reading further.

9.2 Basic Identification Methods

We take a moment to prove the following theorem, which relates the Kirby diagram and intersection form for a closed 4-manifold with only 2-handles.

Theorem 9.2. *Let M be a 4-manifold with no 1- or 3-handles, and let $\mathcal{L}$ be a Kirby diagram for M with link components $K_1, \ldots, K_m$ representing the attaching spheres of 2-handles $H_1, \ldots, H_m$. Then the matrix representation of the intersection form Q_M as the linking matrix N for $\mathcal{L}$, modified such that the diagonal elements are the framing coefficients of each link component, up to elementary row and column operations over $\mathbb{Z}$.*

Proof. For each link component in the Kirby diagram, we want to associate a surface in M. To do this, recall that each link component represents the attaching sphere of some 2-handle, and so bounds the core of the 2-handle, which is a copy of D^2 embedded in M. However, being knots, each link component also bounds an embedded Seifert surface in S^3. The union of the Seifert surface and the core of each handle is then a surface in M that K_i lies in. Furthermore, we can extend the orientation on each link component to its corresponding surface. We can use these surfaces, then, as generators for $H_2(M;\mathbb{Z})$, and use them to compute the intersection form for M.

To do this, and to see that the intersection form is isomorphic to the linking matrix, we need to rearrange our surfaces a bit. Consider two such surfaces, S_i and S_j, obtained from link components K_i and K_j, for $i \neq j$. Leaving K_i in S^3, we push S_i into the interior of the 4-manifold so that $S_i \cap \partial M = K_i$. Leaving the Seifert surface component of S_j in the boundary as well, and noting that the core component is in the interior of the 4-manifold as well and is disjoint from S_i, we have now realized every intersection of S_i with S_j as an intersection of K_i with a Seifert surface for K_j in S^3. We can now leverage Proposition 7.9 to obtain our desired equivalence between the signed intersection of K_i and S_j (and, by our earlier argument, the intersection of S_i with S_j) and the linking number of K_i with K_j.

To see that the diagonal terms of the linking matrix and the intersection form agree, note that a second representative for a 2-homology class can be constructed from S_i by taking a parallel copy S_i' inside a thickened neighborhood of S_i. Because this thickened neighborhood must restrict on a tubular neighborhood of K_i to a framed D^2-bundle over K_i (the attaching region of the 2-handle), the parallel copy S_i' intersects S_i generically in points coming from the crossing of a parallel knot K_i' over K_i. This is precisely the double strand notation for framing, and so the intersection form takes the homology class represented by S_i to the framing on K_i.

The reason this matrix is well-defined up to elementary row and column operations over $\mathbb{Z}$ comes from two equivalent ambiguities. In the intersection form case, it is easy to see that choosing a different basis for $H_2(M;\mathbb{Z})$ would yield a different matrix form for Q_M, but one that could be related by elementary operations. In the link diagram, we can see this ambiguity in our choice of orientation for each link component, but also in the fact that the linking matrix, which is invariant under isotopy, can actually be changed for a given Kirby diagram via handle slides. This will become clearer as we examine handle slides in Kirby diagrams in the next section. □

This observation will allow us to recognize some of our favorite 4-manifolds by their Kirby diagrams. In particular, we compute the following specific example.

Example 9.3. We will now show that the Kirby diagram for $S^2 \times S^2$ is a 0-framed Hopf link. Our explanation of this example, as well as the notation used, is based on the explanation given in [8].

Take f to be the standard Morse function on S^2, yielding a handle decomposition with one 0-handle and one 2-handle, denoted D_- and D_+. The function $g\colon S^2 \times S^2 \to \mathbb{R}$ given by $g(p,q) = f(p)f(q)$ is Morse on $S^2 \times S^2$, and yields a handle decomposition with four handles: one 0-handle, two 2-handles, and one 4-handle, pictured in Figure 32.

Now if we think about attaching just one of the 2-handles to the 0-handle, we cap off one factor of S^2, and so we get either $S^2 \times D_-$ or $D_- \times S^2$, depending on which 2-handle we attach first. Either way, $S^2 \times D^2$ has the 0-framed unknot for a Kirby diagram, and so the diagram for $S^2 \times S^2$ must therefore have two 0-framed unknots in it. Furthermore, the signed intersection of two representative copies of S^2, one in each factor, intersect in a point in $S^2 \times S^2$, and so by Theorem 9.2, the two unknots in the diagram for $S^2 \times S^2$ must have linking number 1. Finally, we see that they must be actually arranged in a Hopf link, because the attaching region of each 2-handle is precisely the boundary of a copy of D_- in our 0-handle $D_- \times D_-$. The full boundary of $D^2 \times D^2$ is, of course, S^3, but it decomposes as $S^1 \times D^2 \cup_\phi D^2 \times S^1$, and the boundary of each D_- is, as subsets of those component solid tori, $\{0\} \times S^1$ and $S^1 \times \{0\}$, arranged in a Hopf link in

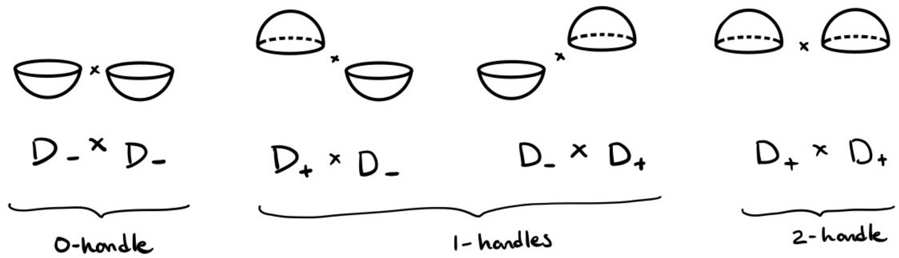

Figure 32: Handles of $S^2 \times S^2$.

$S^1 \times D^2 \cup_\phi D^2 \times S^1 \cong S^3$. We conclude that the Kirby diagram for $S^2 \times S^2$ is a 0-framed Hopf link.

Figure 33 contains some other basic Kirby diagrams that are identifiable as commonplace 4-manifolds. The reader should use this table to check their understanding of Kirby diagrams thusfar.

9.3 Handle Slides

Where Kirby calculus shines as a calculation tool is in its treatment of handle slides. We will concentrate on the notation for 2-handle slides over other 2-handles, as they are the most pertinent for our later questions. Throughout the discussion, we will denote the handle being slid as H_1, and the handle being slid over as H_2.

An oriented k_1-framed 2-handle slide over another oriented k_2-framed 2-handle, where both are unlinked and their orientations agree, is shown, with framings, in Figure 34.

Note that the framing on H_1 after the slide is the sum of its original framing with the framing coefficient of H_2. The fact that they sum comes from our choice to slide H_1 in such a way that its orientation agrees with the orientation of H_2. Such a slide is called *adding* H_1 to H_2. Similarly, we can *subtract* H_1 from H_2 by sliding it so that the orientations disagree.

It is worth pausing and drawing, with the double strand notation, the slide of H_1 over H_2 and combing the resulting knots to check that the framing of H_1 really does change as we have claimed.

Of course, we can also slide H_1 over H_2 even when their attaching regions are linked. However, when such a slide is performed, H_1 traverses the framed D^2 bundle over H_2, and so ends up linked with H_2 with linking number equal to the framing of H_2. Furthermore, the following formula holds for the new framing coefficient for H_1, where the + operation occurs when adding handles, and the − when subtracting them:

$$k_1' = k_1 \pm k_2 \pm 2\ \mathrm{lk}(K_1, K_2).$$

An example of such a handle slide is shown below in Figure 36. As it illustrates, most of the time, sliding 2-handles over other 2-handles only complicates the Kirby diagram. In rare cases, however, the right handle slide can untangle linked components, as shown in Figure 37. This example, incidentally, shows that the 1-twist S^2-bundle over S^2 is diffeomorphic to $\mathbb{CP}^2 \# \overline{\mathbb{CP}^2}$.

There is one more operation we need to define, and it comes from the ambiguity we eliminated earlier in the previous section regarding 1-handles. Although we've changed our notation so that

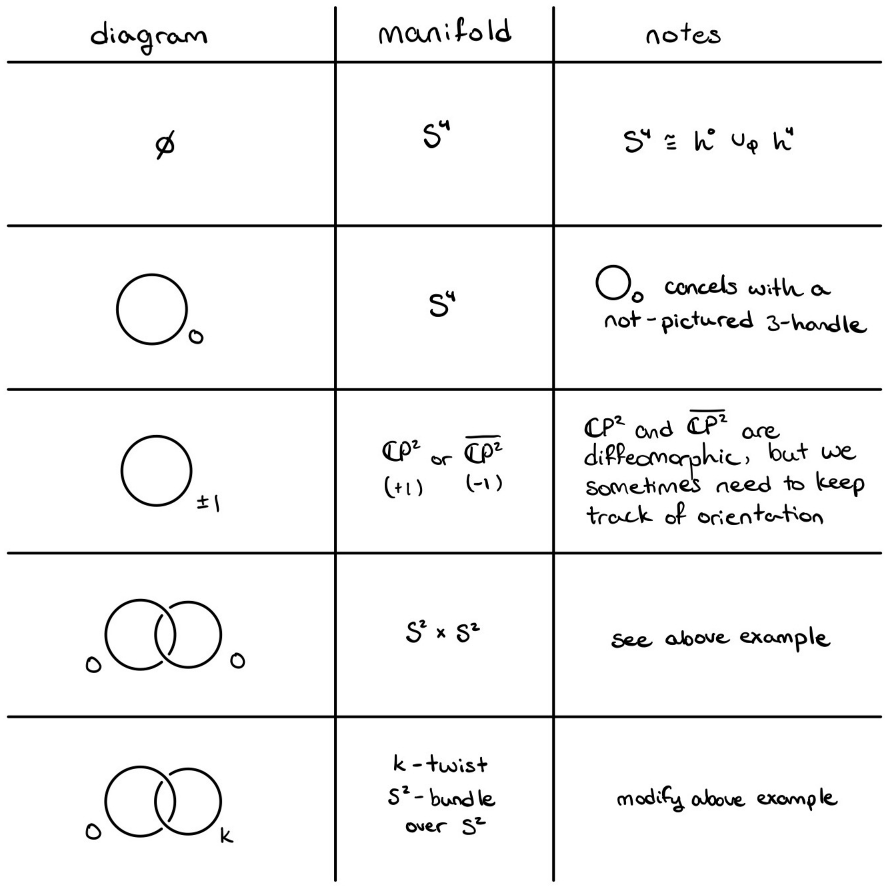

Figure 33: Some basic Kirby diagrams.

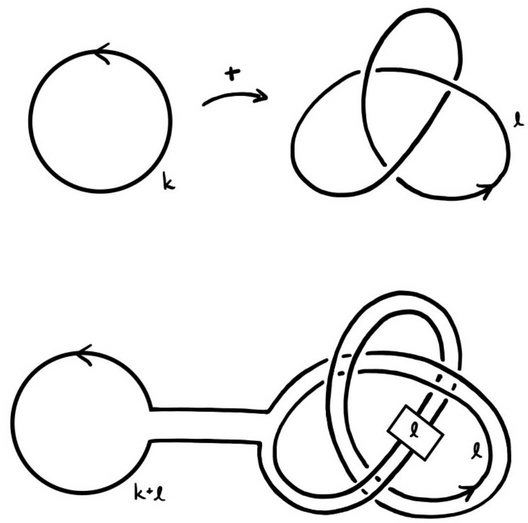

Figure 34: 2-handle additive slide over another 2-handle, where the two are unlinked.

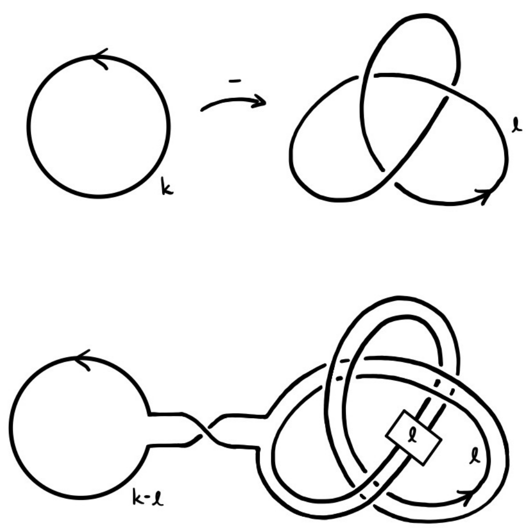

Figure 35: 2-handle subtraction slide over another 2-handle, where the two are unlinked.

Kirby diagrams now are invariant up to ambient isotopy, the process of looping 2-handles around 1-handles that they pass through is still a valid self-diffeomorphism of a 4-manifold, and so we need to account for it in our new notation. Rather than sacrificing isotopy invariance, however, we

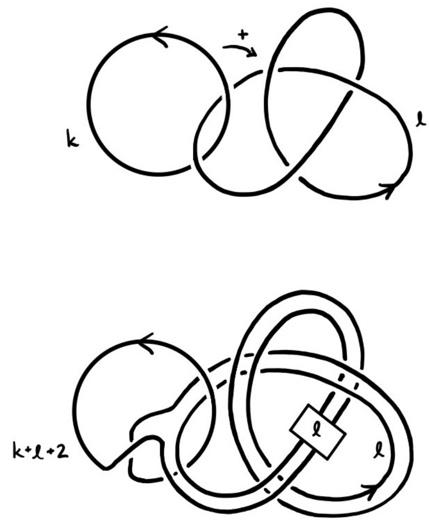

Figure 36: An additive slide of a 2-handle over another 2-handle with linking number 1.

simply introduce a “2-handle slide over a 1-handle” to represent this isotopy. If we treat the dotted circle representing a 1-handle as a 0-framed 2-handle, then in fact, the rules for changing framings that we specified above describe perfectly what happens to a 2-handle’s framing if it is “slid over” a 1-handle it passes through. Therefore, we allow the move shown in Figure 38 on a Kirby diagram.

From these basic moves, an entire theory of Kirby calculus techniques has emerged over the past several decades, including such combination moves as the “blowup” of a diagram or the “slam-dunk”. The classic text on Kirby diagrams is [8], but we recommend supplementation with [2] for an alternate perspective.

9.4 Boundaries of Kirby Diagrams

Recall that from Laudenbach and Poénaru’s result (Theorem 8.13), we derived Corollary 8.14, which tells us that if M and N are two closed 4-manifolds obtained by attaching 3- and 4-handles to the same union of 0-, 1-, and 2-handles, then M and N are diffeomorphic. This result is foundational to the techniques of Kirby calculus; however, it does leave open the following question:

Question 9.4. *Given a framed link diagram, is this diagram a Kirby diagram for some closed 4-manifold?*

The beauty of this question lies in the fact that if the answer, for some link diagram, is yes, then in fact Corollary 8.14 tells us that the manifold it defines is unique up to diffeomorphism. In general, this is a very difficult question to answer, as really it is asking whether or not the 3-manifold boundary of the union of 0-, 1-, and 2-handles defined by a framed link diagram is diffeomorphic to $\sharp^m S^1 \times S^2$ for some m. However, we do introduce one tool that can help us begin to understand the boundaries of Kirby diagrams.

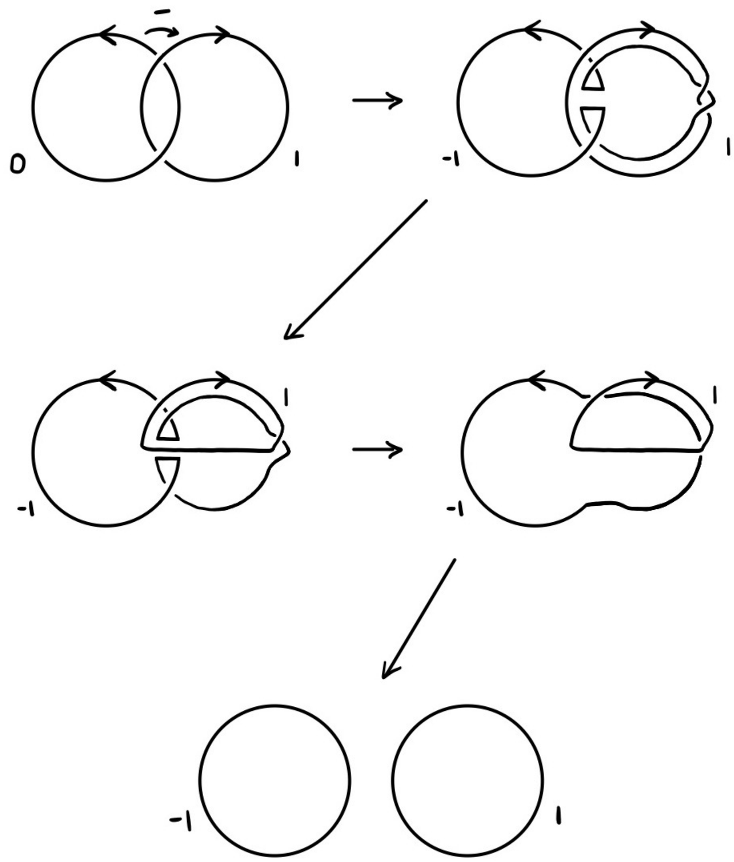

Figure 37: Unlinking handles via sliding.

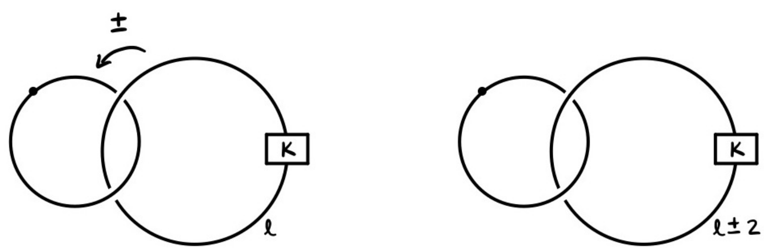

Figure 38: 2-handle "slide" over a 1-handle, changing its framing coefficient by 2.

Proposition 9.5. *Let M be a collection of 0-, 1-, and 2-handles represented by a Kirby diagram $\mathcal{L}$. Then attaching a k-framed 2-handle H along a knot K by a gluing map ψ changes ∂M by doing* *k-framed Dehn surgery on the knot K in ∂M.*

Proof. Recall that a 2-handle is $D^2 \times D^2$, and so we can decompose its boundary S^3 into two copies of $S^1 \times D^2$, one playing the role of the attaching region, and the other the remaining boundary. This is nothing special - we have simply identified that S^3 can be written as the union of two solid tori, glued to each other along a diffeomorphism of their boundary tori that sends a meridinal curve on one to a longitudinal curve on the other. We can identify the attaching sphere of the 2-handle with this identified curve, and in particular, with the longitudinal curve on the attaching region solid torus. This is shown in Figure 39.

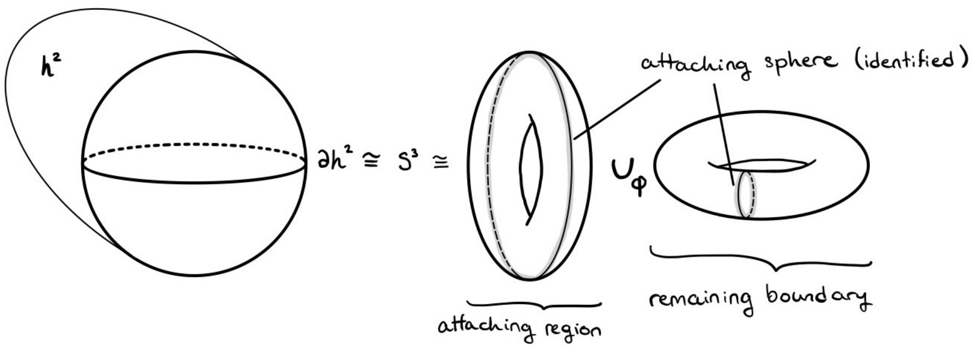

Figure 39: Decomposing ∂h^2 into two solid tori.

When we attach H to a 4-manifold along some knot in its boundary, we effectively push a tubular neighborhood of that knot corresponding to the attaching region copy of $S^1 \times D^2$ in ∂H into the interior of the 4-manifold. However, its boundary torus is still identified with the boundary of the other copy of $D^2 \times S^1$ in ∂H. We can therefore think of the attaching process as removing from ∂M a solid torus (the image of the attaching map of H) and gluing in another one ($\partial H\setminus$ its attaching region) along a diffeomorphism of the boundary torus. Now recall that framing of H is really an identification of the attaching region $S^1 \times D^2$ with the normal D^2-bundle over K in S^3. Specifically, a framing coefficient k corresponds to a section γ of the normal bundle over K that wraps around K k times. But γ is identified with a meridian of $D^2 \times S^1$, and so the diffeomorphism of the boundary torus that we glue along sends γ to the $(k, 1)$ torus knot on the torus boundary of the attaching region. This is precisely our definition of k-framed Dehn surgery on ∂M. A diagram of the whole procedure is shown in Figure 40. $\square$

This connection between Kirby diagrams and Dehn surgery actually gives us a way to prove the following result.

Theorem 9.6. *Every compact, orientable 3-manifold bounds a compact, orientable 4-manifold. Furthermore, this 4-manifold can be taken to contain only 2-handles attached to a 0-handle, meaning that in particular, it is simply connected.*

Proof. Let M be a compact, orientable 3-manifold. Then by Theorem 7.14, M is given by Dehn surgery on some framed link in S^3. Considering a framed link diagram $\mathcal{L}$ for that link as a Kirby diagram, Proposition 9.5 tells us that the boundary of the 4-manifold N described by $\mathcal{L}$ is the

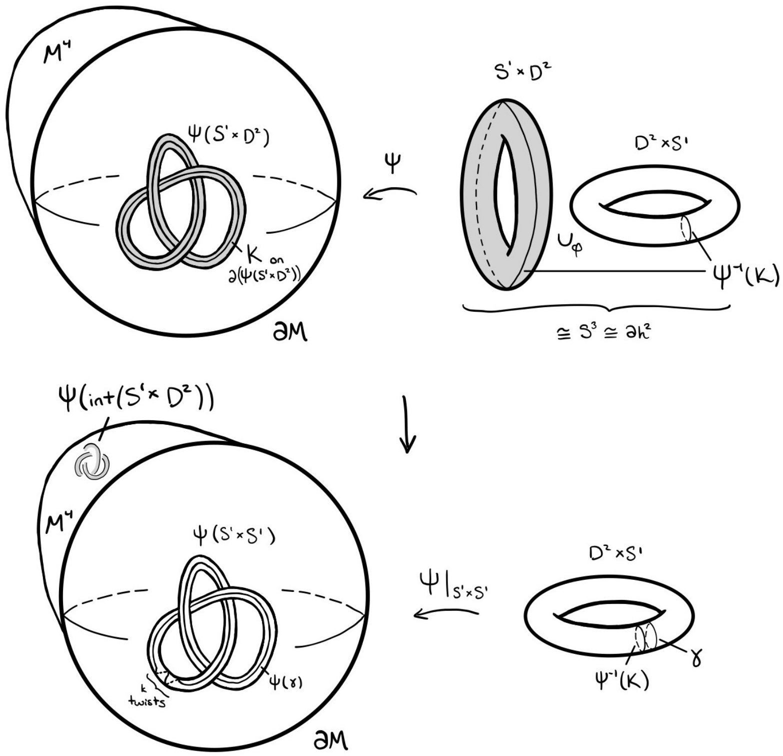

Figure 40: On the top, attaching a 2-handle with attaching map ψ sending the attaching sphere to K. Below, Dehn surgery as given by the attaching of the remainder of ∂h^2 to ∂M along $\psi_{S^1\times S^1}$.

3-manifold obtained by performing exactly that Dehn surgery on the boundary of a 0-handle, S^3. Note that the Kirby diagram for N contains only 2-handles, and so N is simply connected. $\square$

Acknowledgements

The first four sections of this paper were researched and written during the second half of University of Chicago Mathematics Research Experience for Undergraduates program, in the late summer of 2019. The remainder was researched and written in the following 2019-2020 academic year.

I would first and foremost like to thank my thesis advisors, Genevieve Walsh and Robert Kropholler, for their instruction, advice, and support. For his supervision during the University of Chicago REU, I thank Peter May, and for the interesting and engaging lectures and conversations during the Park City Mathematics Institute program in 2019 which inspired this project, I thank John Etnyre and Paul Melvin. Finally, I thank my housemates, Ella, Haley, Hannah, and Petra, for their patience with my ramblings this past year.

References

[1] Colin C. Adams. *The knot book.* American Mathematical Society, Providence, RI, 2004. An elementary introduction to the mathematical theory of knots, Revised reprint of the 1994 original.

[2] Selman Akbulut. *4-manifolds*, volume 25 of *Oxford Graduate Texts in Mathematics.* Oxford University Press, Oxford, 2016.

[3] Glen E. Bredon. *Topology and geometry*, volume 139 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1993.

[4] Jean Cerf. *Sur les difféomorphismes de la sphère de dimension trois* ($\Gamma_4 = 0$). Lecture Notes in Mathematics, No. 53. Springer-Verlag, Berlin-New York, 1968.

[5] Jean Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. *Inst. Hautes Études Sci. Publ. Math.*, (39):5–173, 1970.

[6] S. K. Donaldson. Irrationality and the h-cobordism conjecture. *J. Differential Geom.*, 26(1):141–168, 1987.

[7] Hansjörg Geiges and Kai Zehmisch. Eliashberg's proof of Cerf's theorem. *J. Topol. Anal.*, 2(4):543–579, 2010.

[8] Robert E. Gompf and András I. Stipsicz. *4-manifolds and Kirby calculus*, volume 20 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1999.

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

[10] Michel A. Kervaire and John W. Milnor. Groups of homotopy spheres. I. *Ann. of Math. (2)*, 77:504–537, 1963.

[11] Robion C. Kirby. *The topology of 4-manifolds*, volume 1374 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1989.

[12] François Laudenbach and Valentin Poénaru. A note on 4-dimensional handlebodies. *Bull. Soc. Math. France*, 100:337–344, 1972.

[13] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics.* Springer, New York, second edition, 2013.

[14] Yukio Matsumoto. *An introduction to Morse theory*, volume 208 of *Translations of Mathematical Monographs.* American Mathematical Society, Providence, RI, 2002. Translated from the 1997 Japanese original by Kiki Hudson and Masahico Saito, Iwanami Series in Modern Mathematics.

[15] J. Milnor. *Morse theory.* Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.

[16] John Milnor. On manifolds homeomorphic to the 7-sphere. *Ann. of Math. (2)*, 64:399–405, 1956.

[17] Dale Rolfsen. *Knots and links.* Publish or Perish, Inc., Berkeley, Calif., 1976. Mathematics Lecture Series, No. 7.

[18] Dale Rolfsen. Rational surgery calculus: extension of kirby's theorem. *Pacific J. Math.*, 110(2):377–386, 1984.

[19] Alexandru Scorpan. *The wild world of 4-manifolds.* American Mathematical Society, Providence, RI, 2005.

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

[21] C. T. C. Wall. *Differential topology*, volume 156 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2016.