2022

Decomposing Manifolds in Low-dimensions: from Heegaard Splittings to Trisections

Suixin "Cindy" Zhang
Colby College

Decomposing Manifolds in Low-dimensions: from Heegaard Splittings to Trisections

Suixin (Cindy) Zhang

A thesis submitted in partial fulfillment of the requirements for the degree of

Bachelor of Arts with Honors

Examined and approved on

____

by the following examiners:

____

Dr. Scott A. Taylor (advisor)

____

Dr. Fernando Q. Gouvêa (reader)

Department of Mathematics

Colby College

Acknowledgments

I would like to first thank Professor Scott Taylor for advising me on this project. He has not only sparked my interest in knot theory and low-dimensional topology, but also offered me an invaluable opportunity to do original research in mathematics. His patience and trust have empowered me to believe in myself, to keep trying and exploring.

Next, I would like to thank my mentors Román Aranda and Puttipong Pongtanapaisan for being my collaborators and putting faith in my ability. They have explained mathematics to me on many weekdays and weekends and given me lots of kind encouragement, without which this thesis would not have been possible.

I would also like to thank Professors Tamar Friedmann, Jan Holly, Leo Livshits, Benjamin Mathes, Evan Randles, and Nora Youngs for enriching my mathematical education in amazing ways. Thank you to Thomas Hameister and Professor Akhil Mathew from the University of Chicago for being the most passionate and inspiring mentors one can have for my REU. Thank you to Sarah Sculley for being both my Davis Connects advisor and host mum at Colby and helping with many aspects of my college life.

Many thanks also go to my family and friends. To my family: Thank you for loving and supporting me unconditionally and letting me find my own way. To my dearest friends Christine Chang and Rachel Covitz: Thank you for always being there for me and brightening my world with your courage, kindness and love.

Finally, I'd like to thank my academic advisor Professor Fernando Gouvêa, for sparking my interest in mathematics with his enthusiasm, encouraging me to major in mathematics, and inspiring me to see the beauty of many different mathematical subjects as well as their interconnections. Without his constant loving support, I would not have been able to get through the darkest moments in the past four years and become who I am today.

Contents

Chapter 1

Introduction

The decomposition of a topological space into smaller and simpler pieces is useful for understanding the space. In 1898, Poul Heegaard introduced the concept of a Heegaard splitting which is a bisection of a 3-manifold in his thesis [10]. Heegaard diagrams, which describe Heegaard splittings combinatorially, have been recognized as a powerful tool for classifying 3-manifolds and producing important invariants of 3-manifolds. Handle decomposition, invented by Stephen Smale in [11] in 1962, describes how an n-manifold can be constructed by successively adding handles. In 2012, Gay and Kirby introduced trisections of 4-manifolda in [4], which are a four-dimensional analogues of Heegaard splittings in dimension three. Trisection diagrams give a simple way of understanding and studying 4-dimenisonal spaces.

This thesis is intended to give a friendly introduction to the analogies between the theory of Heegaard splittings of 3-manifolds and trisections of 4-manifolds. The way we introduce these two decompositions is based on Gay's "From Heegaard splittings to trisections; porting 3-dimensional ideas to dimension 4" [5], which is a summary and expansion on a mini-course given at CIRM in 2018. To develop the slogan "Trisections are to 4-manifolds as Heegaard splittings are to 3-manifolds", the author maintains a 2–column format for most definitions and theorems, with parallel bulleted items for the 4–dimensional setting on the left and the 3–dimensional setting on the right. Instead of following this format, we will first talk about Heegaard splittings and then discuss trisections. For each kind of decomposition, we will follow the layout of three ways of thinking about it suggested by [5]: the basic definitions as decompositions of manifolds, the Morse theoretic perspective and the diagrammatic descriptions. Additionally, we discuss how the algebraic structures are also encoded in the splittings. We will set up the terminologies carefully, give definitions and statements of basic results (mostly without proofs), work out some explicit examples suggested as exercises in [5]. In particular, we will restrict ourselves to decomposing only the smooth, closed, orientable manifolds in this paper. We hope that the limited material we present to our readers can serve as a launching point for further exploration of the topics.

Chapter 2

Preliminaries

Among all the topological spaces, we will be interested in a special kind—those that resemble Euclidean space locally. In this chapter, we present selected definitions and results from topology and algebra that will be applied in the later chapters. The definitions and results in Section 2.1, unless otherwise specified, are taken from [1] and [7], and those in Section 2.2 are taken from [6] and [7].

2.1 Homotopy

Definitions and results in this section are based on [1].

Definition 2.1.1. A **homeomorphism** is a continuous function between topological spaces that has a continuous inverse. A homeomorphism onto its image is called an **embedding**.

Definition 2.1.2. Let X, Y be spaces. A **homotopy** is a family of maps $f_t : X \to Y$, $t \in I = [0,1]$, such that the associated map $F : X \times I \to Y$ given by $F(x,t) = f_t(x)$ is continuous. Two maps $f_0, f_1 : X \to Y$ are **homotopic** if there exists a homotopy f_t connecting them.

If, additionally, each f_t is an embedding, i.e. the restriction $f_t : X \to f_t(X)$ is a homeomorphism, then F is called an **isotopy** and we say that f_0 is **isotopic** to f_1.

A special case of the notion of homotopy is the following.

Definition 2.1.3. A **deformation retraction** of a space X onto a subspace A is a homotopy $f_t : X \to X, t \in I = [0,1]$, such that $f_0 = \mathbb{1}$ (the identity map), $f_1(X) = A$, and $f_t|_A = \mathbb{1}$ for all t.

Definition 2.1.4. A **path** is a continuous map $f : I = [0,1] \to X$.

The space X is called **path-connected** if there is a path joining any two points in X.

A **homotopy of paths** in X is a family of paths $f_t : I \to X, 0 \leq t \leq 1$, such that

1. The endpoints $f_t(0) = x_0$ and $f_t(1) = x_1$ are independent of t;
2. The associated map $F : I \times I \to X$ defined by $F(s,t) = f_t(s)$ is continuous.

Two paths f_0 and f_1 connected in such way by a homotopy f_t are said to be **homotopic**. The notation for this is $f_0 \simeq f_1$.

Definition 2.1.5. A map $f : X \to Y$ is called a **homotopy equivalence** if there is a map $g : Y \to X$ such that $fg \simeq \mathbb{1}_Y$ and $gf \simeq \mathbb{1}_X$. The spaces X and Y are said to be **homotopy equivalent** or to have the same **homotopy type**. The notation is $X \simeq Y$.

A space having the homotopy type of a point is called **contractible**.

The relation of homotopy on paths with fixed endpoints in any space turns out to be an equivalence relation. We will denote the equivalence class of a path f under the equivalence relation of homotopy by $[f]$, which will be called the **path homotopy class of** f**.**

We will now define an operation on the path homotopy equivalence class.

Definition 2.1.6. Given two paths $f, g : I \to X$ such that $f(1) = g(0)$, the **concatenation or product path** $f \cdot g$ traverses first f and then g. It is defined by the formula

$$f \cdot g(s) = \begin{cases} f(2s), & 0 \leq s \leq \frac{1}{2} \\ g(2s-1), & \frac{1}{2} \leq s \leq 1 \end{cases}$$

This product operation respects homotopy classes, i.e.

$$[f] \cdot [g] = [f \cdot g].$$

Definition 2.1.7. Paths $f : I \to X$ with the same starting and ending point $f(0) = f(1) = x_0 \in X$ are called **loops**. The common starting and ending point x_0 is called the **basepoint**. The set of all homotopy classes $[f]$ of loops $f : I \to X$ at the basepoint x_0 is denoted $\pi_1(X, x_0)$.

Proposition 2.1.8. *$\pi_1(X, x_0)$ is a group with respect to the product $[f] \cdot [g] = [f \cdot g]$, which is called the **fundamental group of X at the basepoint** x_0.*

It turns out that, if X is path-connected, the group $\pi_1(X, x_0)$ is, up to isomorphism, independent of the choice of basepoint x_0. This is recorded in the following proposition.

Proposition 2.1.9. *Let X be a path-connected space. Then, for any point $x_0, x_1 \in X$, $\pi_1(X, x_1) \cong \pi_1(X, x_0)$.*

Proposition 2.1.10. *A homotopy equivalence $\varphi : X \to Y$ induces an isomorphism $\pi_1(X, x_0) \cong \pi_1(Y, \varphi(x_0))$ for all $x_0 \in X$.*

A very nice tool used to compute fundamental groups is the Seifert-van Kampen theorem.

Theorem 2.1.11 (The Seifert-van Kampen theorem)**.** *Let X be the union of path-connected open sets A_α each containing the basepoint $x_0 \in X$ and $i_{\alpha\beta} : \pi_1(A_\alpha \cap A_\beta) \to \pi_1(A_\alpha)$ is the homomorphism induced by the inclusion $A_\alpha \cap A_\beta \hookrightarrow A_\alpha$. If each intersection $A_\alpha \cap A_\beta$ is path-connected, then the homomorphism $\Phi : *_\alpha \pi_1(A_\alpha) \to \pi_1(X)$ induced by the inclusions $A_\alpha \to X$ is surjective. If in addition each intersection $A_\alpha \cap A_\beta \cap A_\gamma$ is path-connected, then the kernel of Ψ is the normal subgroup N generated by all elements of the form $i_{\alpha\beta}(\omega) i_{\beta\alpha}(\omega)^{-1}$ for $\omega \in \pi_1(A_\alpha \cap A_\beta)$, and hence Φ induces an isomorphism $\pi_1(X) \cong *_\alpha \pi_1(A_\alpha)/N$.*

When there are only two sets A_α and A_β in the cover of X, the condition on the triple intersection in Theorem 2.1.11 is superfluous. In this case, given that the double intersection $A_\alpha \cap A_\beta$ is path-connected, there is an isomorphism $\pi_1(X) \cong (\pi_1(A_\alpha) * \pi_1(A_\beta))/N$.

2.2 Topology

In this section, we will cover the basic concepts of manifolds and their classifications.

Definition 2.2.1 ([2])**.** A topological space X is a **Hausdorff space** if for each pair of distinct points $x, y \in X$ there exist open neighborhoods U of x and V of y that are disjoint.

Definition 2.2.2 ([2])**.** A topological space X is **second countable** if its topology has a countable basis.

An example of a topological space which is both Hausdorff and second countable is $\mathbb{R}$. For each pair of distinct points $x, y \in \mathbb{R}$, there exist disjoint open intervals I, J such that $x \in I$ and $y \in J$; the collection of open intervals with rational end-points forms a basis. More generally, $\mathbb{R}^n$ is second countable and Hausdorff.

Definition 2.2.3 (Topological n-manifold)**.** A **topological n-manifold**, (or a **manifold**), is a second countable Hausdorff space M for which there exists a family of pairs $\{(M_\alpha, \phi_\alpha)\}$ with the following properties:

1. for all α, M_α is an open subset of M and $M = \bigcup_\alpha M_\alpha$;
2. ϕ_α is a homeomorphism from M_α to an open subset of $\mathbb{R}^n$.

A pair (M_α, ϕ_α) is called a **chart** of M. The family of pairs $\{(M_\alpha, \phi_\alpha)\}$ is called an **atlas** for M.

The **dimension** of an n-manifold is n. Two n-manifolds are called **equivalent** if and only if they are homeomorphic.

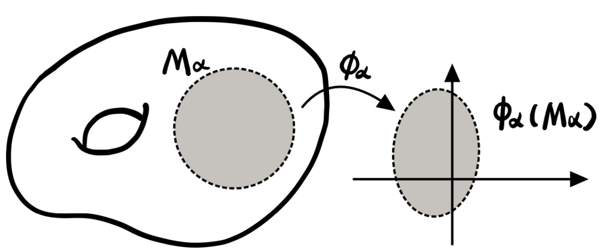

Figure 2.1: A chart of a 2-manifold.[3]

Definition 2.2.4 (Manifolds with boundary)**.** Let $H^n = \{(x_1, ..., x_n) \in \mathbb{R}^n : x_1 \geq 0\}$. An **$n$-manifold with boundary** is a second countable Hausdorff space M with an atlas such that, for all α, ϕ_α, there is a homeomorphism from M_α to an open subset of $\mathbb{R}^n$ or H^n.

The **boundary** of M, denoted by ∂M, is the set of all points in M that have a neighborhood homeomorphic to H^n but no neighbourhood homeomorphic to $\mathbb{R}^n$. Points not on the boundary are called **interior points**.

Definition 2.2.5. A topological space X is said to be **compact** if every open cover of X has a finite subcover.

Definition 2.2.6. An n-manifold M is **closed** if and only if M is compact and $\partial M = \emptyset$.

Finally, we will define what we mean by gluing two simpler manifolds together to obtain a more complicated manifold.

Definition 2.2.7 (Gluing)**.** Let X, Y be closed subsets of the same dimension in the boundaries of the manifolds A and B respectively. Form the disjoint union $A \sqcup B$. Let $\phi : A \to B$ be a homeomorphism and define an equivalent relation $\sim$ by $x \sim y$ for $x \in X$ and $y \in Y$ iff $y = \phi(x)$. Then, the result of **gluing of** A **along** B is the quotient space $A \sqcup B/\sim$.

2.2.1 Differentiable Manifolds

Definition 2.2.8. A map $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be C^q if it has continuous partial derivatives of order q. A map is **smooth** (or C^∞) if it has partial derivatives of all orders.

Definition 2.2.9. A C^q**-manifold**, for $q \in [0, \infty]$, is a topological manifold M with an atlas such that, for any pair of charts $(M_\alpha, \phi_\alpha), (M_\beta, \phi_\beta)$ in this atlas, the map $\phi_\beta \circ \phi_\alpha^{-1}$ is C^q where defined.

A C^∞**-manifold** is also called a **differentiable manifold**, or a **smooth manifold**.

Given an atlas for a manifold, a map of the form $\phi_\beta \circ \phi_\alpha^{-1}$, denoted by $\phi_{\alpha\beta}$, is called a **transition map**.

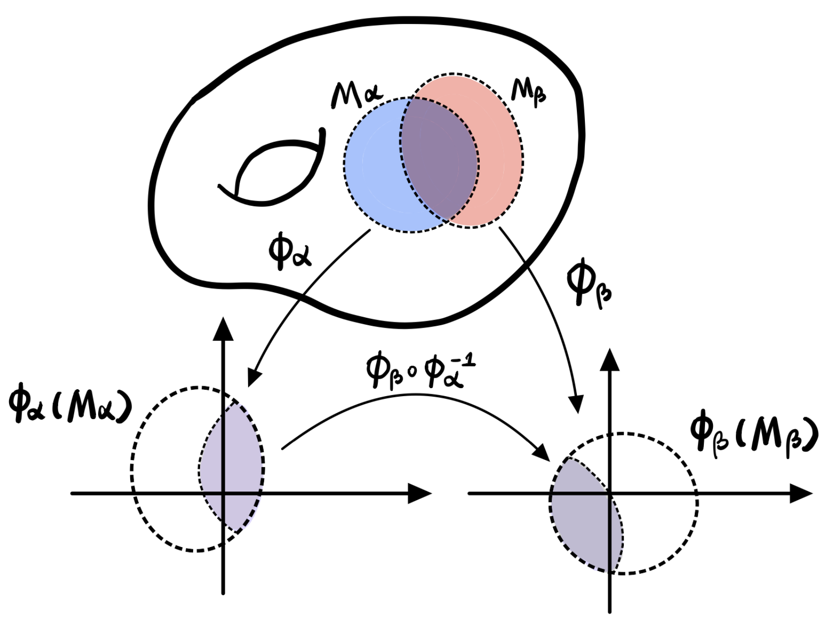

Figure 2.2: A transition map.[3]

Note that the transition maps (wherever they are defined) are maps going from $\mathbb{R}^n$ to $\mathbb{R}^n$. Just as we have differentiable maps from $\mathbb{R}^n$ to $\mathbb{R}^m$, we can extend some of the concepts to manifolds too.

Definition 2.2.10. Let M be a manifold with atlas $\{(M_\alpha, \phi_\alpha)\}$ and let N be a manifold with atlas (N_β, ψ_β). A map $f : M \to N$ is C^q if for all α, β, the map $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ (where it is defined) is C^q. In particular, a map $f : M \to N$ is **smooth** if for all α, β, the map $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ (where it is defined) is **smooth**.

Here are a few useful examples of smooth maps.

Example 2.2.11 (Smoothness of a map to a circle)**.** [12] The map $F : \mathbb{R} \to S^1$ defined by $F(t) = (\cos(t), \sin(t))$ for all $t \in \mathbb{R}$ is smooth.

Example 2.2.12 (Smoothness of a projection map)**.** [12] Let M and N be manifolds. The projection map $\pi : M \times N \to M$ defined by $\pi((p, q)) = p$ for all $p \in M$ and $q \in N$ is a smooth map.

Definition 2.2.13. (Diffeomorphisms) A C^q-map between C^q-manifolds with a C^q-inverse is called a C^q**-diffeomorphism.** A C^∞-diffeomorphism is simply called a **diffeomorphism**.

Example 2.2.14 (Coordinate map)**.** [12] If (M_α, ϕ_α) is a chart on an n-manifold M, then the coordinate map $\phi_\alpha : M_\alpha \to \phi_\alpha(M_\alpha) \subset \mathbb{R}^n$ is a smooth map with smooth inverse and therefore a diffeomorphism.

Definition 2.2.15 (Equivalence of C^q-manifolds)**.** Two C^q-manifolds are considered to be equivalent if there is a C^q-diffeomorphism between them.

One can extend these notions to the settings of submanifolds, manifolds with boundary, and submanifold of a manifold with boundary (see [6]).

2.2.2 Orientable Manifolds

Definition 2.2.16. A C^∞-manifold M with boundary is **orientable** if it has an atlas such that the Jacobians of all transition maps have positive determinants. If no such atlas exists, then M is **non-orientable**. Such an atlas is called an **orientation** of M, and we often write (M, ϕ_α) to denote an oriented manifold.

Let M be a differentiable manifold with two orientations $(M, \{\phi_\alpha\})$ and $(M, \{\psi_\beta\})$. The subset of M on which $\phi_\alpha \circ \psi_\beta^{-1}$ is defined and has a Jacobian with a positive determinant is the subset where the orientations of $(M, \{\phi_\alpha\})$ and $(M, \{\psi_\beta\})$ are said to **coincide**. The subset of M on which $\phi_\alpha \circ \psi_\beta^{-1}$ is defined and has a Jacobian with a negative determinant is the subset where the orientations of $(M, \{\phi_\alpha\})$ and $(M, \{\psi_\beta\})$ are said to **differ**. Both subsets above are open. Thus, for a connected manifold M, two orientations either coincide on all of M or differ on all of M.

Given an oriented manifold $(M, \{\phi_\alpha\})$, we can create an oriented manifold $(M, \{\psi_\alpha\})$ for which the orientations differ on the whole M. Here is how we do it: For each $\phi_\alpha : U_\alpha \to \mathbb{R}^n$, defined by $\phi_\alpha(x) = (x_1, ..., x_n)$, we substitute $\psi_\alpha : U_\alpha \to \mathbb{R}^n$ defined by $\psi_\alpha(x) = (-x_1, ..., x_n)$. In particular, the resulting orientation is called the **opposite orientation** of the original one, and the resulting oriented manifold is denoted by $-M$.

Example 2.2.17 (Some orientable and non-orientable manifolds.)**.** The n-sphere S^n and n-ball B^n for any $n \in \mathbb{Z}_{\geq 0}$ are orientable. In dimension-2, the torus $T^2 = S^1 \times S^1$ is orientable, while the Mobius strip and the Klein bottle are non-orientable. In dimension-3, the solid torus $S^1 \times B^2$ is orientable, while the solid Klein bottle is non-orientable.

Definition 2.2.18. For oriented C^∞-manifolds $(M, \{\phi_\alpha\})$ and $(N, \{\psi_\beta\})$ of the same dimension, a smooth map $h : M \to N$ is **orientation-preserving** if the Jacobians of the maps $\psi_\beta \circ \phi_\alpha^{-1}$ (where they are defined) all have positive determinant. If they all have negative determinant, it is **orientation-reversing**.

Definition 2.2.19 (Connected sum, [8])**.** Let M_1 and M_2 be two connected manifolds. Choose embeddings $i_1 : B^n \to M_1$ and $i_2 : B^n \to M_2$, so that i_1 preserves orientation and i_2 reverses orientation. Then, the **connected sum** $M_1 \# M_2$ is the gluing of M_1 and M_2 via identifying $i_1(tu)$ with $i_2((1-t)u)$ for each unit vector $u \in S^{n-1}$ and each $0 < t < 1$. Choose the orientation for $M_1 \# M_2$ which is compatible with that of M_1 and that of M_2. This operation is well-defined up to orientation preserving diffeomorphism.

2.2.3 Morse Functions

Taking $N = \mathbb{R}$ in Definition 2.2.10, we have the definition of smoothness of a function going from a manifold to the real line.

Definition 2.2.20. Given a smooth manifold M, a function $f : M \to \mathbb{R}$ is **smooth** if the function $\hat{f} : f \circ \phi_\alpha^{-1}$ is smooth for every chart (M_α, ϕ_α) in the atlas.

Definition 2.2.21. Given a differentiable map $\hat{f} : \mathbb{R}^n \to \mathbb{R}$ and a point $p \in \mathbb{R}^n$, the **gradient** of $\hat{f}$ at p is the vector $\nabla_p \hat{f} = (\frac{\partial \hat{f}}{\partial x_1}|_p, \cdots, \frac{\partial \hat{f}}{\partial x_n}|_p)$.

Definition 2.2.22. Let M be a smooth manifold. Given a function $f : M \to \mathbb{R}$ and a point $p \in M$, we will write $\nabla_p f = 0$ iff $\nabla_p \hat{f} = \nabla_p(f \circ \phi_\alpha) = 0$ in some chart (M_α, ϕ_α) with $p \in M_\alpha$. Otherwise, we will write $\nabla \hat{f} \neq 0$.

The following lemma says that Definition 2.2.22 does make sense.

Lemma 2.2.23. *Let (M_α, ϕ_α) and (M_β, ϕ_β) be charts of some atlas of M, $f : M \to \mathbb{R}$ a smooth function, $\hat{f}_\alpha = f \circ \phi_\alpha^{-1}$ and $\hat{f}_\beta = f \circ \phi_\beta^{-1}$. For any $p \in M_\alpha \cap M_\beta$, let $x = \phi_\alpha(p)$ and $x' = \phi_\beta(p)$. Then, the gradient $\nabla_x \hat{f}_\alpha = \mathbf{0}$ if and only if $\nabla_{x'} \hat{f}_\beta = \mathbf{0}$.*

Therefore, the gradient $\nabla_p \hat{f}_\alpha = 0$ for some chart (M_α, ϕ_α) if and only if $\nabla_p \hat{f} = 0$ for all charts.

So $\nabla f = 0$ if and only if $\nabla_p \hat{f} = 0$ for all charts.

Definition 2.2.24. The **Hessian matrix** of a function $\hat{f} : \mathbb{R}^n \to \mathbb{R}$, denoted by $H_x(\hat{f})$, is the matrix of second partial derivatives of $\hat{f}$:

$$H_{\hat{f}} = \begin{bmatrix} \frac{\partial^2 \hat{f}}{\partial x_1^2} & \cdots & \frac{\partial^2 \hat{f}}{\partial x_1 \partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial^2 \hat{f}}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 \hat{f}}{\partial x_n^2} \end{bmatrix}.$$

The Hessian evaluated at a point $x \in \mathbb{R}^n$ is denoted by $H_{\hat{f}}(x)$.

Definition 2.2.25. For $f : M \to \mathbb{R}$ and $p \in M$, we will say that the Hessian of f at p, denoted by $H_f(p)$, is **singular** if in some chart (M_α, ϕ_α), $\det(H_{\hat{f}}(\phi_\alpha(p))) = 0$.

Similarly, the following lemma assures us that this is a useful definition.

Lemma 2.2.26. *For $p \in M$, if $\nabla_p f = 0$ and $\det(H_{\hat{f}}(\phi_0(p))) = 0$ for some chart (M_0, ϕ_0), then $\det(H_{\hat{f}}(\phi_\alpha(p))) = 0$ for all charts (M_α, ϕ_α).*

Having a way of talking about derivatives on smooth manifolds allows us to define the notion of critical points of smooth functions on such manifolds.

Definition 2.2.27. Let $f : M \to \mathbb{R}$ be smooth. A point $p \in M$ is called a **critical point** of f iff $\nabla_p(f) = 0$. The value $f(p) \in \mathbb{R}$ is called a **critical value**.

If the preimage $f^{-1}(a)$ does not contain any critical points, then $a \in \mathbb{R}$ is called a **regular value**.

At a critical point $p \in M$, Lemma 2.2.26 ensures that $\det(H_f(p)) = 0$ is well defined.

Definition 2.2.28. A critical point $p \in M$ of a function $f : M \to \mathbb{R}$ is called **non-degenerate** iff $\det(H_f(p)) \neq 0$. Otherwise, we call the critical point **degenerate**.

Let $M = \mathbb{R}$ or $\mathbb{R}^2$. Here are some example of degenerate and non-degenerate critical points of functions on $\mathbb{R}$ and $\mathbb{R}^2$.

Example 2.2.29 (Monkey saddle, [9])**.** Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$(x, y) \mapsto x^3 - 3xy^2.$$

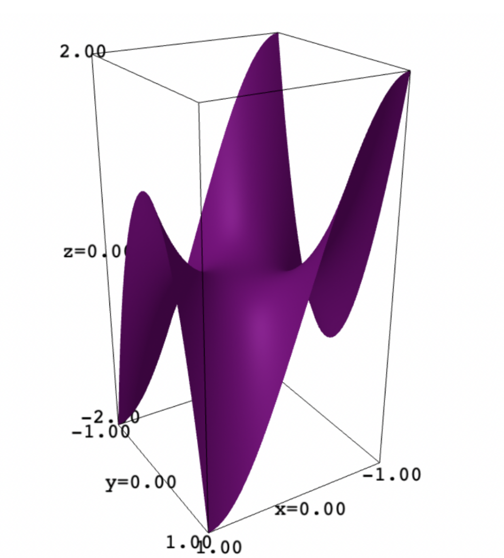

Figure 2.3: Monkey saddle.

Then, we have

$$\frac{\partial f(x,y)}{\partial x} = 3x^2 - 3y^2,$$
$$\frac{\partial f(x,y)}{\partial y} = -6xy,$$

so that the only critical point of f is $(0,0)$. The Hessian matrix is given by

$$H_f = \begin{bmatrix} 6x & -6y \\ -6y & -6x \end{bmatrix},$$

and thus

$$\det(H_f((0,0))) = (-36x^2 - 36y^2)|_{(x,y)=(0,0)} = 0.$$

So $(0,0)$ is a degenerate critical point of f. f is called the Monkey saddle, because it is a saddle a monkey (with a tail) would use. (See Figure 2.3.)

Example 2.2.30 (A non-degenerate critical point)**.** Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$(x,y) \mapsto x^2 - 3y^2.$$

Then, we have

$$\frac{\partial f(x,y)}{\partial x} = 2x,$$
$$\frac{\partial f(x,y)}{\partial y} = -6y,$$

so that the only critical point of f is $(0,0)$. The Hessian matrix is given by

$$H_f = \begin{bmatrix} 2 & 0 \\ 0 & -6 \end{bmatrix},$$

and thus

$$\det(H_f((0,0))) = (-6 \cdot 2 - 0 \cdot 0)|_{(x,y)=(0,0)} = -12 \neq 0.$$

So $(0,0)$ is a non-degenerate critical point of f.

Definition 2.2.31 (Morse functions)**.** A smooth function $f : M \to \mathbb{R}$ is a **Morse function** if f has no degenerate critical points on M.

Lemma 2.2.32 (Morse lemma)**.** *Let f be a Morse function on an n-manifold M and $p \in M$ be a non-degenerate critical point of f. Then there exists a neighborhood $N \subset M$ and a map $\phi : N \to \mathbb{R}^d$ sending p to the origin, such that if $\hat{f} = f \circ \phi^{-1}$ then*

$$\hat{f}(x_1, \cdots, x_d) = \pm x_1^2 \pm \cdots \pm x_d^2 + f(p).$$

If p is not a critical point of f then there is a chart (N, ϕ) such that

$$\hat{f}(x_1, ..., x_d) = x_1 + f(p).$$

*The number of minus signs in the quadratic polynomial is independent of the choice of chart and is called the **index** of the critical point.*

Remark 2.2.33. The index of a critical point determines the behavior of f at that point, and Morse's Lemma classifies non-degenerate critical points into exactly $n+1$ types.

Note that the origin $\mathbf{0} = (0, \cdots, 0)$ is the only critical point of the function $\hat{f}(x_1, \cdots, x_d) = \pm x_1^2 \pm \cdots \pm x_d^2 + f(p)$. On the other hand, for the function $\hat{f}(x_1, ..., x_d) = x_1 + f(p)$, there are no critical points. Therefore, every regular point and every nondegenerate critical point has a neighborhood containing at most one critical point.

The remark above leads to the following corollary.

Corollary 2.2.34. *Suppose f is a function on a compact, smooth manifold M such that every critical point of f is non-degenerate. Then, f has a finite number of critical points.*

It follows from Corollary 2.2.34 that a Morse function contains finitely many critical points, all of which are non-degenerate. Furthermore, any two critical points correspond to different critical values.

It's a fact that every smooth function is pretty close to a Morse function, and Morse functions are "stable", meaning that small perturbation will not change the behavior of the function significantly (See [7]).

Remark 2.2.35. The index of a Morse function $f : M \to \mathbb{R}$ at a critical point p defined in Definition 2.2.32 is equal to the number of negative eigenvalues of the Hessian H_f evaluated at p, which does not depend on choice of local coordinates. (See [9].)

Example 2.2.36. (The height function on S^2, [13]) Let $M = S^2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\} \subset \mathbb{R}^3$ and consider the height function $f : M \to \mathbb{R}$ defined by $f(x_1, x_2, x_3) = x_1$. Let's show that f is a Morse function.

Let $M_+ = S^2 \setminus \{(-1, 0, 0)\}$ and $M_- = S^2 \setminus \{(1, 0, 0)\}$. Let

$$\phi_+(x_1, x_2, x_3) = (\frac{x_2}{1-x_1}, \frac{x_3}{1-x_1}),$$

$$\phi_-(x_1, x_2, x_3) = (\frac{x_2}{1+x_1}, \frac{x_3}{1+x_1}).$$

Then, $\{(M_+, \phi_+), (M_-, \phi_-)\}$ forms an atlas of S^2, where (M_+, ϕ_+) is the coordinate chart around $(-1, 0, 0)$ and (M_-, ϕ_-) is the coordinate chart around $(1, 0, 0)$. The inverses of ϕ_+ and ϕ_- are given by:

$$\phi_+^{-1}(y_1, y_2) = (\frac{y_1^2 + y_2^2 - 1}{y_1^2 + y_2^2 + 1}, \frac{2y_1}{y_1^2 + y_2^2 + 1}, \frac{2y_2}{y_1^2 + y_2^2 + 1}),$$

$$\phi_-^{-1}(y_1, y_2) = (\frac{1 - y_1^2 - y_2^2}{y_1^2 + y_2^2 + 1}, \frac{2y_1}{y_1^2 + y_2^2 + 1}, \frac{2y_2}{y_1^2 + y_2^2 + 1}).$$

To find the critical points of f, we need to look at the maps $\hat{f}_+ = f \circ \phi_+^{-1}$ and $\hat{f}_- = f \circ \phi_-^{-1}$ from $\mathbb{R}^2$ to $\mathbb{R}$. Note that

$$\hat{f}_-(y_1, y_2) = \frac{1 - y_1^2 - y_2^2}{y_1^2 + y_2^2 + 1},$$

and the gradient at (y_1, y_2) is given by

$$\nabla_{(y_1,y_2)}\hat{f}_- = (\frac{-4y_1}{(y_1^2+y_2^2+1)^2}, \frac{-4y_2}{(y_1^2+y_2^2+1)^2}).$$

Thus, we have $\nabla_{(y_1,y_2)}\hat{f}_- = 0$ if and only if $y_1 = y_2 = 0$, which implies that the point $p_- = (1,0,0) \in M$ is the only critical point of f in $M_- \subset M$.

The Hessian of $\hat{f}$ is given by

$$H_{\hat{f}_-} = \begin{bmatrix} \frac{-4(1-3y_1^2+y_2^2)}{(y_1^2+y_2^2+1)^3} & \frac{16y_1y_2}{(y_1^2+y_2^2+1)^3} \\ \frac{16y_1y_2}{(y_1^2+y_2^2+1)^3} & \frac{-4(1-3y_2^2+y_1^2)}{(y_1^2+y_2^2+1)^3} \end{bmatrix}.$$

So we have

$$H_{\hat{f}_-}(\phi_-(p_-)) = H_{\hat{f}_-}((0,0)) = \begin{bmatrix} -4 & 0 \\ 0 & -4 \end{bmatrix},$$

$$\det(H_{\hat{f}_-}(\phi_-(p_-))) = \det(H_{\hat{f}_-}(0,0)) = 16 \neq 0.$$

Hence, the point $p_- = (1,0,0) \in M$ is a non-degenerate critical point. In view of Remark 2.2.35, this is a critical point of index 2.

One can go through the same calculation for $\hat{f}_+$ and show that the point $p_+ = (-1,0,0)$ is the only critical point in $M_+ \subset M$, which is non-degenerate and has index 0. This shows that f is indeed a Morse function on S^2.

Theorem 2.2.37. *Let M be a smooth manifold. Then, there exists Morse function on M. In fact, the set of Morse functions is dense in $C^\infty(M)$.*

Chapter 3

Handles, Handlebodies and Morse functions

To set us up for the discussion about Heegaard splittings in dimension-3 and trisections in dimension-4, in this chapter we will introduce handles and handlebodies as well as how they are related to Morse functions on manifolds. Definitions and results in this chapter are from [6] and [7] unless otherwise specified.

3.1 Handle Decompositions

Handle decompositions exist for smooth manifolds of any dimension. It is a particular way to build an n-manifold from n-balls.

Definition 3.1.1 (k-handle)**.** An n-dimensional **k-handle** is an n-ball, thought of as $[0,1]^k \times [0,1]^{n-k}$, that is attached to some pre-existing submanifold along $\partial[0,1]^k \times [0,1]^{n-k}$.

In Figure 4.2, we show all the k-handles in 3-dimension.

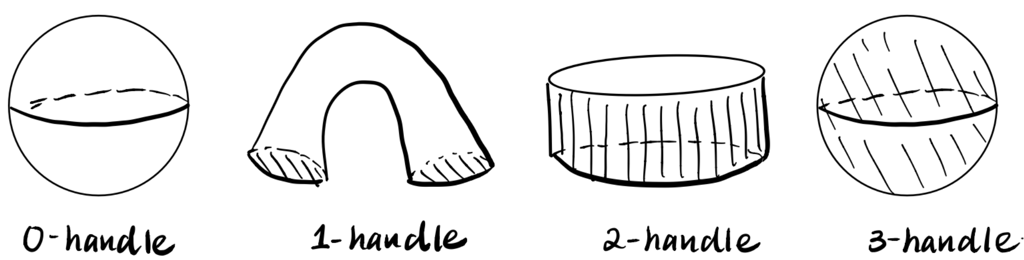

Figure 3.1: k-handles in 3-dimension with the attaching regions shaded.

Let N be an n-dimensional manifold with boundary. Let h be a k-handle, $\varphi : [0,1]^k \times \partial[0,1]^{n-k} \to \partial N$ be a one-to-one continuous map, and N' be the result of gluing h to N by the map φ. Then, we call N' the result of **attaching a k-handle to** N. When $k = 0$, N' is the disjoint union of N and the closed ball h.

Definition 3.1.2 (Core and cocore)**.** [6] The **core** of a 3-dimensional k-handle is $\{\frac{1}{2}\}^{3-k} \times [0,1]^k$. The **cocore** is $[0,1]^{3-k} \times \{\frac{1}{2}\}^k$.

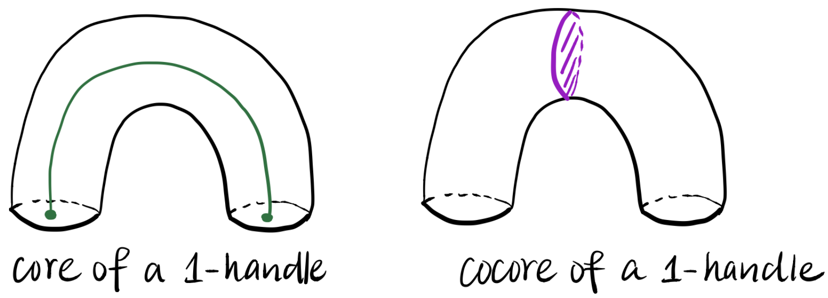

Figure 3.2: Core and cocore of a 3-dimensional 1-handle.

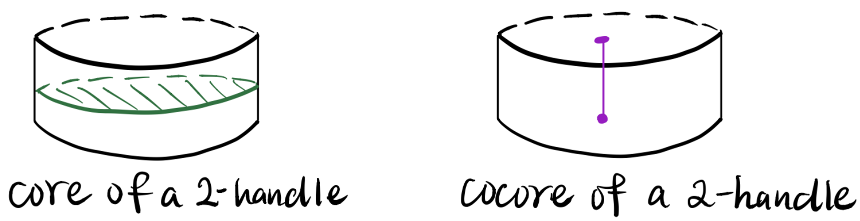

Figure 3.3: Core and cocore of a 3-dimensional 2-handle.

Definition 3.1.3 (Handlebody)**.** In dimension-3, a **genus-g handlebody** is a compact, connected, and orientable 3-manifold that admits a handle decomposition consisting of 0-handles and 1-handles and has genus-g boundary.

More generally, a manifold in n-dimension is called a k**-handlebody** if it is the union of r-handles, for r at most k.

The following lemma justifies our definition of the genus of a handlebody.

Lemma 3.1.4. *Two 3-dimensional handlebodies are homeomorphic if and only if their boundaries have the same genus.*

The handle decomposition of a 3-manifold is not unique. In fact, given a handlebody in 3-dimensions, there are infinitely many ways of decomposing it into 0-handles and 1-handles. Luckily, Proposition 3.1.5 says that there is a "universal" construction.

Proposition 3.1.5. *A genus-g 3D handlebody can be constructed from one 0-handle and g 1-handles.*

Let H be a handlebody. We will call a properly embedded disk $D \subseteq H$ **essential** if its boundary does not bound a disk in ∂H. Starting with a handle decomposition of H consisting of 0-handles and 1-handles, we can obtain a collection of properly embedded, essential disks which cut H into balls by taking the set of cocores of the 1-handles in this decomposition. In

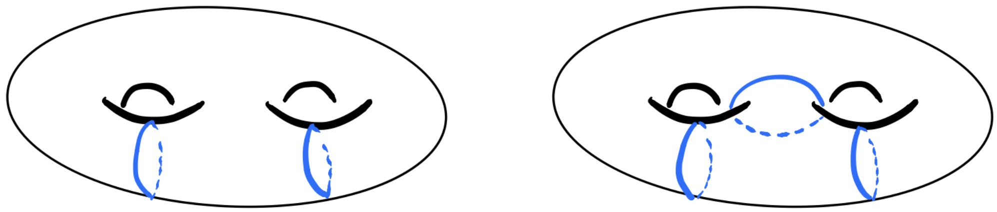

Figure 3.4: Diagram of a non-minimal collection (RHS) and a minimal collection of disks (LHS).

fact, we could have equivalently defined a 3D handlebody as a connected manifold obtained by gluing a collection of closed 3-balls along a collection of pairwise disjoint disks in the union of their boundaries. [7]

Definition 3.1.6 (System of disks)**.** A **system of disks** for a 3D handlebody is a collection of properly embedded, essential disks that cut the handlebody into a collection of 3-balls.

Definition 3.1.7 (Cut system)**.** A **cut system** of a genus g surface Σ_g is a collection of g disjoint simple closed curves such that surgery along these curves produces S^2.

Since the boundary of a genus-g handlebody is $\partial H = \Sigma_g$, a cut system $\{\alpha_1, ...\alpha_g\}$ of ∂H corresponds to a system of disks $\{D_1, ...D_g\}$ for H, where $\partial D_i = \alpha_i$. In particular, such system of disks is called **minimal**. It's not hard to see that a system of disks for a handlebody is minimal if and only if its complement is connected. [7]

While Definition 3.1.3 suggests a way of building a handlebody from 3-balls, one can also construct a handlebody by "thickening" a graph.

Definition 3.1.8 (Spine of a handlebody)**.** A **spine** of a handlebody H is a piecewise linear graph K embedded in H so that $H \setminus K$ is homeomorphic to $\partial H \times (0, 1]$.

In particular, if H is a regular neighborhood of a graph K embedded in it, then K is a spine of H.

Lemma 3.1.9. *If H is a handlebody, then there is a spine for H.*

Proof. By definition, every handlebody H consists of 0-handles and 1-handles. We can construct a spine for H as follows. Put a vertex w_j in the interior in each of the 0-handles. If a 1-handle is attached to a 0-handle, so that one endpoint v_i of its core lies on the boundary of the 0-handle, connect this endpoint to the chosen vertex with an unknotted edge e. The graph consists of all the w_j's, v_i's, e's and the cocores is a spine for H. □

Lemma 3.1.10. *If H is a 3D handlebody, then there is a 1-vertex spine for H.*

Proof. We know from Proposition 3.1.5 that H can be constructed from a single 0-handle and some number of 1-handles as needed. Now the construction provided by the proof of Lemma 3.1.9 gives a 1-vertex spine for H. □

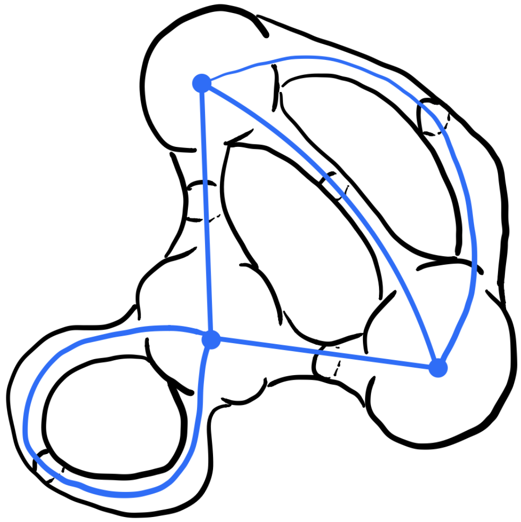

Figure 3.5: A spine of a handlebody.

3.2 Sliding

In section 3.1, we saw two ways of understanding the structure of a handlebody: via systems of disks and via embedded graphs. In this section, we will explore the connection between these two perspectives on handlebodies and introduce an important concept—sliding.

Before we do any sliding, we shall first build connections between systems of disks and spines for a handlebody.

Definition 3.2.1. A spine K is **dual** to a collection of meridian disks $\mathcal{D}$ for the handlebody H if each edge of K intersects a single disk of $\mathcal{D}$ exactly once, each disk intersects exactly one edge and each ball of $H \setminus \mathcal{D}$ contains exactly one vertex of K.

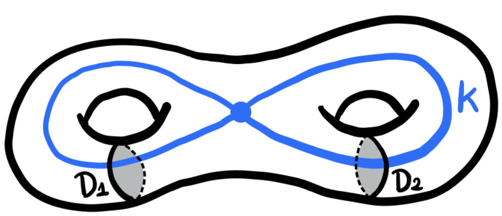

Figure 3.6: A system of disk dual to a spine K.

The following propositions explain how a collection of disks and a spine for a handlebody encode the same information. In particular, a spine K will induce a system of disks for H uniquely up to isotopy. Conversely, given a collection of disk, we can always find a spine K dual to it, and such a spine is also unique up to isotopy.

Proposition 3.2.2. *Given a spine K of a handlebody H, there is a system of disks dual to K. If $\mathcal{D}$ and $\mathcal{D}'$ are both minimal systems of disks dual to K then the systems of disks are isotopic.*

Proposition 3.2.3. *Given a minimal system of disks $\mathcal{D}$ for a handlebody H, there exists a spine dual to $\mathcal{D}$. If spines K and K' are dual to $\mathcal{D}$, then there is an isotopy of H taking K onto K'.*

We are now ready to discuss the move of sliding—a move that allows us to turn a system of disks into another system of disks, or to turn a spine into another spine for a handlebody.

3.2.1 Disk Slides

Definition 3.2.4 (Sliding of disks)**.** Let H be a 3-dimensional handlebody and $\mathcal{D} = \{D_1, ..., D_n\}$ a system of disks for H. Let a be an arc in ∂H with endpoints on ∂D_i and ∂D_j for some $i \neq j \in \{1, ..., n\}$. Assume the interior of a is disjoint from ∂D_i for each i. Let D_a be the frontier of a closed regular neighborhood of $D_i \cup a \cup D_j$. Then, D_a is called a **sliding** of D_i along a over D_j. (See Figure 3.7.)

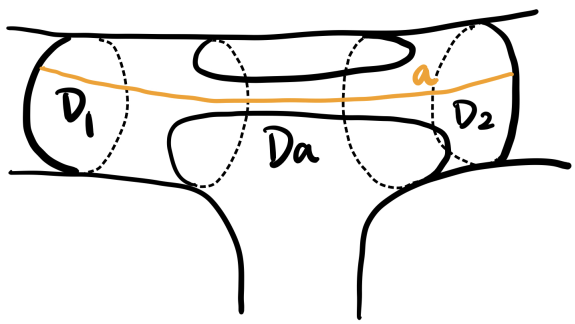

Figure 3.7: A sliding of D_1 along a over D_2.

It turns out that if we replace D_j (or D_i) with D_a in $\mathcal{D}$ in Definition 3.2.4, $\mathcal{D}$ remains a system of disks for H. This motivates the following definition.

Definition 3.2.5 (Disk slide and slide equivalence)**.** [7] Two systems of disks for a 3D handlebody H are called **isotopic** of there is an isotopy of H which takes one system of disks to the other.

A system of disks that is isotopic to $\mathcal{D} \cup \{D_\alpha\} \setminus \{D_j\}$ for some α and D_j as defined in Definition 3.2.4 is called a **disk slide** of $\mathcal{D}$.

Two system $\mathcal{D}$ and $\mathcal{D}'$ of disks are called **slide equivalent** if there is a sequence $\mathcal{D} = \mathcal{D}_0, \mathcal{D}_1, ..., \mathcal{D}_k = \mathcal{D}'$, where $\mathcal{D}_{i+1}$ is a disk slide of $\mathcal{D}_i$ for each $i \in \{0, 1, ..., k-1\}$.

For any handlebody with genus $g > 1$, there are infinitely many non-isotopic systems of disks. However, it turns out that we can avoid such ambiguities if we insist our system of disks to be minimal, thanks to the theorem below.

Theorem 3.2.6. *Any two minimal systems of disks for a handlebody H are slide equivalent.*

Taking the boundary a system of disks, we obtain from Definition 3.2.4 the notion of the sliding of simple closed curves on an oriented surface.

Definition 3.2.7. (Sliding of curves) Given a finite collection $\alpha = \{\alpha_1, ..., \alpha_n\}$ of disjoint simple closed curves on an oriented surface Σ and two of these curves α_i and α_j and an arc a joining α_i and α_j and otherwise disjoint from the α curves, one can produce a new collection of curves $\alpha' = \{\alpha'_1, ..., \alpha'_n\}$ by **sliding** α_i **over** α_j **along** a as follows: Let $\alpha'_k = \alpha_k$ for all $k \neq i$, while α'_i is the unique boundary component of a regular neighborhood of $\alpha_i \cup a \cup \alpha_j$ which is not isotopic to either α_i or α_j.

Two collections of disjoint simple closed curves on the same surface are called **slide equivalent** if one can be transformed to the other by a sequence of sliding and isotopies.

Two pairs of collections of simple closed curves (α, β) and (α', β') on the same surface are **slide equivalent** if α is slide equivalent to α' and β is slide equivalent to β'. Two triples (Σ, α, β) and $(\Sigma', \alpha', \beta')$ are **slide diffeomorphic** if (α, β) is slide equivalent to some (α'', β'') such that $(\Sigma, \alpha'', \beta'')$ is orientation preserving diffeomorphic to $(\Sigma', \alpha', \beta')$. [5]

3.2.2 Edge Slides

The duality between disk systems and spines allows us to translate sliding into a move on a graph.

Let K be a spine of a handlebody H and e_1, e_2 be edges in K sharing a vertex v. Let D be a disk with $\text{int}(D) \cap K = \emptyset$, whose boundary ∂D is divided into 3 segments $\alpha_1, \alpha_2, \alpha_3$, where α_1 is closed half of e_1 containing v, α_2 is e_2, and α_3 is an arc in $\text{int}(H)$. Let K' be the result of replacing α_1 with α_3 in K.

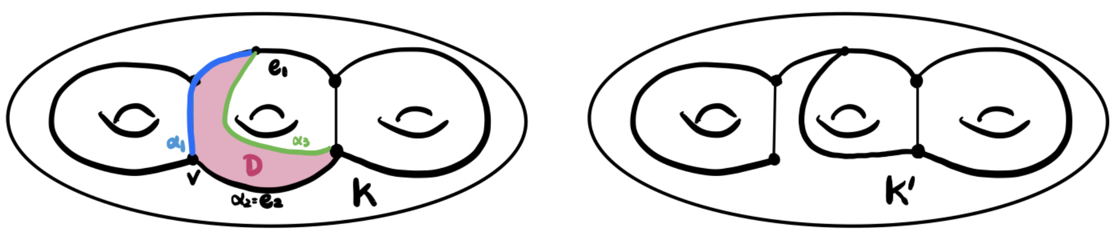

Figure 3.8: Edge Slides

Definition 3.2.8 (Edge slides)**.** Any graph isotopic to the graph K' described above is called an **edge slide** of the spine K.

Two spines K and K' of a handlebody H are called **slide equivalent** if there is a sequence of edge slides and isotopies starting at K and ending at K'.

Just as sliding disks produces a large family of systems of disks for a handlebody, sliding edges enables us to form a large family of embedded graphs which are spines of the handlebody.

Proposition 3.2.9. *Let K be a spine for the handlebody H. Suppose K' is an edge slide of K. Then, K' is also a spine for H.*

The following proposition says that the duality between disks and spine extends over the move of sliding.

Proposition 3.2.10. *Let $\mathcal{D}$ be a minimal collection of disks for a 3D handlebody H. Let $\mathcal{D}'$ be a disk slide of $\mathcal{D}$. If K is a spine dual to $\mathcal{D}$, then there is an edge slide K' of K that is dual to $\mathcal{D}'$.*

In an isotopy of a manifold, the 2-handles may follow paths which take them over other 2-handles, while the 1-handles may slide their ends onto other 1-handles. Both of these can be understood through disk slides and edge slides.

Definition 3.2.11 (Handle slides)**.** A **1-handle slide** is the result of thickening an edge slide of the handlebody. A **2-handle slide** is the result of thickening a disk slide of the handlebody.

3.3 Morse Functions and Handle Decompositions

Definition 3.3.1 (Handle decomposition)**.** A **handle decomposition** of a compact n-manifold M is a finite sequence of manifolds $M_0, ..., M_l$ such that:

1. $M_0 = \emptyset$;
2. M_l is diffeomorphic to M;
3. M_i is obtained from M_{i-1} by attaching a handle.

It turns out that Morse functions on smooth manifolds, whose existence is guaranteed by Theorem 2.2.37 naturally induce handle decompositions of the manifolds, which we shall explain now.

Definition 3.3.2. Let $f : M \to \mathbb{R}$ be a Morse function. Then, for each $x \in \mathbb{R}$, the **level set** of f is $f_x = f^{-1}(x) \subseteq M$.

Note that f_x is always closed, because it is the pre-image of a closed set under a continuous function. Additionally, since M is compact, f_x must also be compact. In particular, if $[a, b] \subseteq \mathbb{R}$ is an interval with no critical values, then the level sets must fit together in a rather simple way.

Lemma 3.3.3. *Let $f : M \to \mathbb{R}$ be a Morse function and $a, b \in \mathbb{R}$ are such that $a < b$ and a, b are both regular values of f. Let $f_a = f^{-1}(a)$. Suppose there is no critical values in the interval $[a, b]$. Then, $f^{-1}([a, b])$ is homeomorphic to $f_a \times [a, b]$.*

Definition 3.3.4. A **sublevel set** $f_x^- = f^{-1}((-\infty, x])$ is everything below a level set.

If $x \in \mathbb{R}$ is a regular value, then f_x^- is an n-dimenisional submanifold.

Lemma 3.3.5. *Let $f : M \to \mathbb{R}$ be a Morse function and $a, b \in \mathbb{R}$ are such that $a < b$ and a, b are both regular values of f. Suppose there is no critical value in the interval $[a, b]$. Then, there is an isotopy of M which maps f_a^- onto f_b^-.*

Lemma 3.3.6. *If x is a regular value, then f_x is a closed submanifold of M of dimension $d - 1$, where $d = \dim(M)$.*

A Morse function naturally decomposes a manifold into regular and singular level sets. Let $a, b \in \mathbb{R}$ be two regular values of f. Suppose there is no critical value in $[a, b]$. Then, the level sets $f^{-1}(a)$ and $f^{-1}(b)$ are diffeomorphic. If, additionally, we have $c \in (a, b)$, then the level sets $f^{-1}((a, b))$ is diffeomorphic to $f^{-1}(c) \times (a, b)$.

Therefore, to understand the manifold, we only need to understand how the level set changes as we cross through a critical value. In particular, we want to observe how the submanifold $f^{-1}((-\infty, y])$ changes as we pass through a critical value y_0 via local coordinates.

We shall consider a Morse function on a closed 3-dimensional manifold. In this case, the regular level sets are surfaces. Thus, for any $a, b \in \mathbb{R}$ which lie strictly between two critical values, we have

$$f^{-1}((a, b)) = \text{surface} \times I.$$

A critical point of index 0 is a point $x_0 \in M$ for which there exist local coordinates $\{x_1, ..., x_n\}$ such that on a neighborhood of x_0, f has the form $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$. We say that x_0 is modeled on $x_1^2 + x_2^2 + x_3^2$. In particular, as we move from the level sets below the critical point to the level sets above the critical point, we see a point appear and expand into a 2-sphere.

A critical point of index 1 is modeled on $-x_1^2 + x_2^2 + x_3^2$. As we move from the level sets below the critical point to the level sets above the critical point, we see a hyperboloid of 2 sheets meet at 1 point and then merge into 1 sheet.

A critical point of index 2 is modeled on $-x_1^2 - x_2^2 + x_3^2$. As we move from the level sets below the critical point to the level sets above the critical point, we see a hyperboloid of 1 sheet collapse by pinching along a circle and break into a hyperboloid of 2 sheets.

A critical point of index 3 is modeled on $-x_1^2 - x_2^2 - x_3^2$. In particular, as we move from the level sets below the critical point to the level sets above the critical point, we see a 2-sphere collapse into a point and disappear.

More generally, let M be a closed n-dimensional manifold and a Morse function $f : M \to \mathbb{R}$. Let $x_0 \in M$ be a critical point of f. Then, there is a closed neighborhood N_{x_0} of X_0 homeomorphic to $[0, 1]^n$. By Morse lemma, there exist local coordinates about x_0 in which

$$f(x_0) = f(x_1, \cdots, x_n) = -\sum_{i=1}^{k} x_i^2 + \sum_{i=k+1}^{n} x_i^2.$$

Let $y_0 \in \mathbb{R}$ be the critical value corresponding to x_0. The homeomorphism type of $f^{-1}((-\infty, y])$ changes as we pass through y_0. In particular, if $a, b \in \mathbb{R}$ are such that y_0 is the only critical value in (a, b), $f^{-1}((-\infty, b])$ is homeomorphic (diffeomorphic, in fact) to the result of attaching a copy of $[-1, 1]^n$ to $\partial f^{-1}((-\infty, a])$ along $\partial[-1, 1]^k \times [-1, 1]^{n-k}$, i.e. a k-handle!

This suggests that every closed manifold can be built from handles starting from the 0-handle, a copy of $[-1, 1]^n$ appearing out of thin air. To build a given n-manifold M, we can choose a Morse function $h : M \to \mathbb{R}$ and mimic the growth of $h^{-1}((-\infty, y])$ by attaching a k-handle when y passes through a critical value of index k.

A k-handle is **dual** to an $n - k$-handle in the following sense. Given any Morse function $h : M \to \mathbb{R}$, $-h : M \to \mathbb{R}$ is also a Morse function. Due to the sign change, critical points of index k for h are in 1-1 correspondence with critical points of index $n - k$ for $-h$. Thus,

for each k-handle attached along $\partial[-1,1]^k \times [-1,1]^{n-k}$ to $h^{-1}((-\infty, y_0 - \epsilon])$ of M, we can instead attach an $(n-k)$-handle to $h^{-1}([y+\epsilon, +\infty))$ along $\partial[-1,1]^{n-k} \times [-1,1]^k$.

Example 3.3.7. (The height function on S^2 again) We showed in Example 2.2.36 that the height function h defined by $(x_1, x_2, x_3) \mapsto x_1$ on the unit 2-sphere in $\mathbb{R}^3$ is a Morse function. In particular, the only two critical points $(-1,0,0)$ and $(1,0,0)$ correspond to critical values -1 and 1 in $\mathbb{R}$. One can check that $(-1,0,0)$ is of index 0 and $(1,0,0)$ is of index 2. Thus, we see a disk, a 2-dimensional 0-handle, appears as we pass through -1 and then cap it off with another disk, a 2-dimensional 2-handle, by gluing its boundary to that of the pre-existing one.

The function $-h$ defined by $(x_1, x_2, x_3) \mapsto -x_1$ is also Morse, and it still has two critical points $(-1,0,0)$ and $(1,0,0)$. However, as a result of the sign change, $(-1,0,0)$ now has index 2, and $(1,0,0)$ has index 0.

Chapter 4

Heegaard Splittings of 3-Manifolds

Complicated and interesting manifolds can be built from handlebodies. In this chapter, we will introduce Heegaard splittings for 3-manifolds, which is a systematic way of decomposing "nice" 3-manifolds into handlebodies. Definitions and results in this chapter come from [5], [6] and [7] unless otherwise specified. We follow closely the way how Heegaard splittings are introduced in [5].

4.1 The Basic Definitions: Decompositions

Definition 4.1.1 (Heegaard splittings)**.** A genus-g **Heegaard splitting** of a closed, connected, oriented 3–manifold M is a decomposition $M = H_1 \cup_\Sigma H_2$, usually denoted by a triple $\mathcal{S} = (\Sigma, H_1, H_2)$, such that:

1. For each i, H_i is diffeomorphic to H_g, a genus-g 3-dimensional handlebody.
2. $H_1 \cap H_2 = \Sigma$ is diffeomorphic to Σ_g, a genus-g surface. We orient $H_1 \cap H_2$ as $\partial H_1 = -\partial H_2$.

Σ is called the **Heegaard surface** or **splitting surface** of $M = H_1 \cup_\Sigma H_2$. The **Heegaard genus** of M is the smallest possible genus of a Heegaard splittinig of M.

We will define two different notions of equivalence for Heegaard splittings.

Definition 4.1.2 (Equivalence of Heegaard splittings)**.** Two Heegaard splittings (Σ, H_1, H_2) and (Σ', H_1', H_2') of M are **isotopic** if and only if their splitting surfaces Σ and Σ' are isotopic.

Two Heegaard splittings (Σ, H_1, H_2) and (Σ', H_1', H_2') are **homeomorphic** if and only if there is a homeomorphism $h : M \to M$ that takes Σ to Σ', H_1 to H_1', and H_2 to H_2'.

Example 4.1.3 (A genus-0 Heegaard splitting of S^3)**.** The decomposition $S^3 = B^3 \cup_\Sigma B^3$, where $\Sigma = S^2$, is a genus-0 Heegaard splitting of S^3, regardless of homeomorphism.

Example 4.1.4 (A genus-1 Heegaard splitting of S^3)**.** The decomposition $S^3 = H_1 \cup_\Sigma H_2$, where $\Sigma = S^1 \times S^2$ and $H_1 \cong S^1 \times B^2$ for $i = 1, 2$, is a genus-1 Heegaard splitting of S^3. In particular, the meridian of H_2 is glued along the longitude of H_1, and the longitude of H_2 is glued along the meridian of H_1.

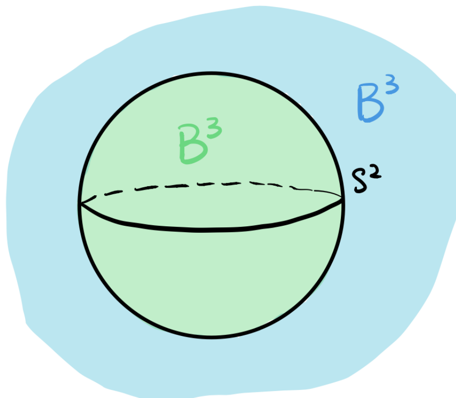

Figure 4.1: The Genus-0 Heegaard Splitting of S^3.

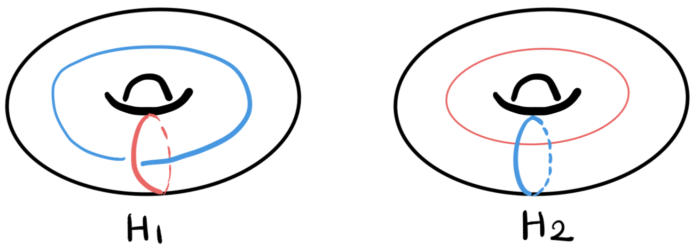

Figure 4.2: A genus-1 Heegaard splitting of S^3.

Provided that we have a Heegaard splitting $M = H_1 \cup_\Sigma H_2$, the spine of each of the two handlebodies is a graph embedded in the ambient 3-manifold M. The following lemma says that the isotopy class of the Heegaard surface and hence the Heegaard splitting is completely determined by the spine of H_1.

Lemma 4.1.5. *Let (Σ, H_1, H_2) and (Σ', H_1', H_2') be Heegaard splittings of a manifold M. Let K and K' be spines of H_1 and H_1', respectively. If K is isotopic to K', then Σ and Σ' are isotopic. As a consequent of Definition 4.1.2, if K is isotopic to K', then (Σ, H_1, H_2) and (Σ', H_1', H_2') are isotopic splittings.*

It is known that such decomposition always exists for a large family of 3-manifolds.

Theorem 4.1.6 (Existence of Heegaard splittings, Moise)**.** *Every compact, closed, connected, orientable 3-manifold allows a Heegaard splitting.*

Theorem 4.1.6 follows from the existence of a special type of Morse functions, which we'll discuss next.

4.2 The Morse Theoretic Perspective

Definition 4.2.1. A Morse function $f : M \to \mathbb{R}$ is **self-indexing** if for all k, the value of f at any critical point of index k is less than the value of f at any critical point of index $k+1$.

Let M be a 3-manifold and $h : M \to \mathbb{R}$ a self-indexing Morse function. Then, there is some regular point $r \in \mathbb{R}$ such that all critical points of index 0 or 1 occur below r and all critical points of index 2 or 3 occur above r, so that $h^{-1}((-\infty, r])$ is a handlebody. On the other hand, we've seen in Section 3.3 that $h^{-1}([r, +\infty))$ is a handlebody too, since all critical points contained in $[r, +\infty)$ of $-h$ are of index 0 and 1. A self-indexing Morse function thus defines a Heegaard splitting of the 3-manifold M, which justifies the following definition.

Definition 4.2.2. A self-indexing Morse function on a 3-manifold is called a **Heegaard splitting Morse function**.

In [5], Gay equivalently defines a Heegaard splitting as follows. Let $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ and $R_\theta \subset \mathbb{R}^2$ be the ray making angle $-\theta$ with the positive x–axis and parametrize R_θ using $[0, \infty) \to (t\cos(-\theta), t\sin(-\theta))$. Thus, we can think of $\mathbb{R}$ as the x-axis by identifying $[0, \infty)$ with R_0 and $(-\infty, 0]$ with R_π.

Definition 4.2.3 (Heegard splitting Morse function)**.** A genus-g **Heegard splitting Morse function** f on a 3–manifold M is a Morse function $f : M \to \mathbb{R}$ such that:

1. 0 is a regular value of f, and thus $f^{-1}(0) = \Sigma$ is a closed surface, which we require to be connected of genus g.

2. On each of the two rays R_0 and R_π, f has exactly g index 2 and one index 3 critical points, all of which have distinct critical values.

Note that the fact that f has critical points of index 2 and index 3 is because that we oriented $(-\infty, 0]$ of the x-axis away from the origin.

Conversely, a Heegaard splitting can be used to define a Morse function as well, though it's usually not unique, due to the non-uniqueness of handles structures on handlebodies.

Example 4.2.4 (Standard genus-0 Heegaard splitting of S^3)**.** Consider $S^3 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1\} \subset \mathbb{R}^4$. The projection $(x_1, x_2, x_3, x_4) \mapsto x_1$ on S^3 is a Heegaard splitting Morse function, giving the standard genus-0 Heegaard splitting of S^3.

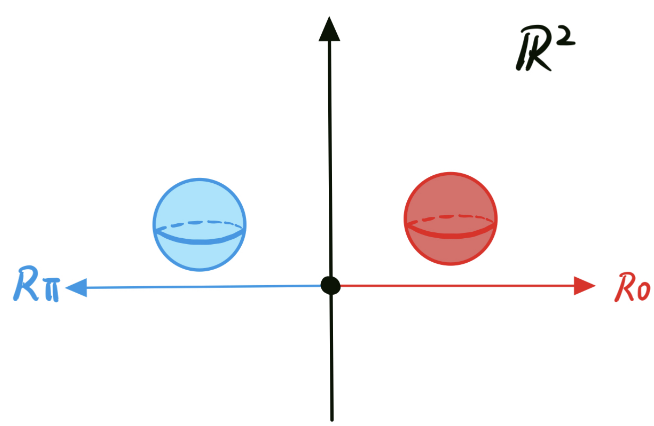

Figure 4.3: A genus-0 Heegard splitting Morse function

In particular, as illustrated in Figure 4.3, we have

$$f^{-1}(R_0) = f^{-1}([0,1]) = \{(x_1, x_2, x_3, \sqrt{1-(x_1^2+x_2^2+x_3^2)}) : x_1^2+x_2^2+x_3^2 \leq 1\} \cong B^3,$$

$$f^{-1}(R_\pi) = f^{-1}([-1,0]) = \{(x_1, x_2, x_3, -\sqrt{1-(x_1^2+x_2^2+x_3^2)}) : x_1^2+x_2^2+x_3^2 \leq 1\} \cong B^3,$$

$$f^{-1}(0) = \{(x_1, x_2, x_3, 0) : x_1^2+x_2^2+x_3^2 \leq 1\} \cong S^2.$$

In addition to the existence statement in Theorem 4.1.6, there is a uniqueness statement. To discuss that, we need to introduce an important operation on Heegaard splittings called stabilization.

4.3 Stabilization

Given a 3-manifold M and a genus g Heegaard splitting (Σ, H_1, H_2) of M, we can construct a genus $g+1$ Heegaard splitting of it as follows:

1. Attach an unknotted 1-handle h to H_1. The core of $h \subset H_2$ together with an arc in Σ bounds a disk D in M.
2. Let $C = D \times I$ be the thickened D. Then, $H_1 \cup h \cup C = H_1 \cup B^3$ is isotopic to H_2, and the closure $\overline{H_2 \setminus (h \cup C)}$ is isotopic to H_2.
3. Let $H_1' = H_1 \cup h$ and $H_2' = \overline{H_2' \setminus (h \cup C)} \cup C$. So $M = H_1' \cup H_2'$, and $\Sigma' := H_1' \cap H_2' = \partial H_1 \cup \partial h$.

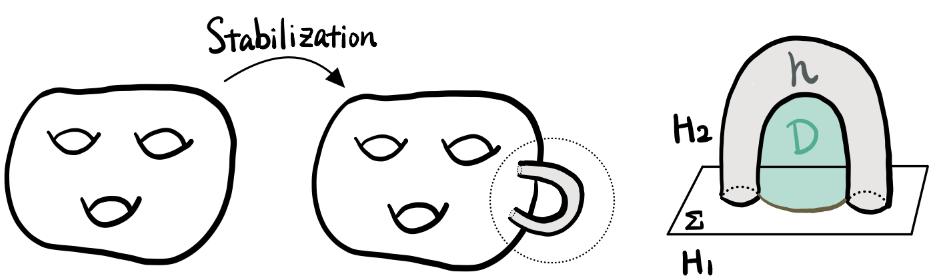

Figure 4.4: A stabilization of a Heegaard splitting.

Definition 4.3.1 (Stabilization)**.** Any Heegaard splitting (Σ', H_1', H_2') constructed by iterating the above process one or more times, possibly reversing the roles of H_1 and H_2, is called a **stabilization** of the Heegaard splitting (Σ, H_1, H_2).

Since we would mostly like to consider Heegaard splittings which are not stabilizations, criterion for deciding whether a splitting is a stabilization of another splitting would be useful. This motivates the following definition, which is justified by Lemma 4.3.3 that follows.

Definition 4.3.2. A Heegaard splitting (Σ', H_1', H_2') is called **stabilized** if there is a pair of essential, properly embedded disks D_1', D_2' in H_1', H_2', respectively, such that $\partial D_1' \cap \partial D_2'$ consists of a single point in Σ'.

Lemma 4.3.3. *A Heegaard splitting* (Σ', H_1', H_2') *is stabilized if and only if it is a stabilization of a Heegaard splitting* (Σ, H_1, H_2)*.*

As stabilization allows us to look at Heegaard splittings with different genera of the same 3-manifold, we may ask in what ways these Heegaard splittings are related. Quite remarkably, it turns out that for any two splittings of the same manifold, there exists a third splitting which is isotopic to a stabilization of each. This is discovered by Reidemeister and Singer independently in 1935.

Theorem 4.3.4 (Reidemeister-Singer)**.** *Let* (Σ, H_1, H_2) *and* (Σ', H_1', H_2') *be Heegaard splittings (not necessarily of the same genus) of a compact 3-manifold* M*. Then there is a Heegaard splitting surface* Σ'' *which is isotopic to both a stabilization of* Σ *and a stabilization of* Σ'*.*

In 1968, Waldhausen proved the following theorem about Heegaard splittings of S^3.

Theorem 4.3.5 (Waldhausen)**.** *Every genus* $g \geq 1$ *Heegaard splitting of* S^3 *is stabilized.*

We will refer our readers to [6] for proofs of Theorem 4.3.4 and Theorem 4.3.5. As a corollary, we have the following.

Corollary 4.3.6. *Every Heegaard splitting of* S^3 *is a stabilization of a genus-zero Heegaard splitting.*

We can now add a uniqueness statement to Theorem 4.1.6.

Theorem 4.3.7. *Every closed, connected, oriented 3–manifold has a Heegaard splitting, and any two Heegaard splittings of the same 3–manifold become isotopic after some number of stabilizations.*

4.4 The Diagrammatic Perspective: Heegaard Diagrams

Since all handlebodies of the same genus are homeomorphic, given a Heegaard splitting $M = H_1 \cup_\Sigma H_2$ of genus g, it is not too hard to visualize the genus-g handlebodies H_1 and H_2 in $\mathbb{R}^3$ individually. Since $\partial H_1 = \Sigma$, to reconstruct M together with its Heegaard splitting from H_1, it suffices to understand how a system of disks for H_2 is attached to ∂H_1, since there is only one way to attach the 3-handles along meridian disks by Alexander's theorem, as stated below.

Theorem 4.4.1 (Alexander's theorem, [2])**.** *Every embedded* S^2 *in* $\mathbb{R}^3$ *bounds an embedded 3-ball.*

Since a system of disks is defined by a collection of curves on ∂H_1, the Heegaard splitting can be encoded by a closed surface and a set of curves on it diagrammatically, which determines the homeomorphism glueing the surfaces ∂H_1 and ∂H_2 together, This motivates the definition of Heegaard diagrams.

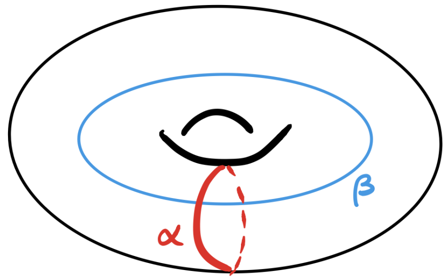

Figure 4.5: A genus-1 Heegaard Diagram of S^3

Definition 4.4.2 (Heegaard diagrams)**.** A **Heegaard diagram** $(\Sigma_g, \alpha, \beta)$ is a closed, orientable surface Σ of genus g that is equipped with two collections of disjoint, essential, simple closed curves $\alpha = \{\alpha_1, ..., \alpha_g\}$ and $\beta = \{\beta_1, ..., \beta_g\}$, each of which forms a cut system of the surface.

The following theorem connects Heegaard splittings to Heegaard diagrams.

Theorem 4.4.3. *Given a Heegaard diagram $\mathcal{D} = (\Sigma, \alpha, \beta)$, there is a 3–manifold $M = M(D)$ with Heegaard splitting $S(\mathcal{D}) = (M, M_1, M_2)$ such that $\Sigma = M_1 \cap M_2$, oriented according to the conventions in Definition 4.1.1, and such that the α curves bound embedded disks in M_1 and the β curves bound embedded disks in M_2. Then:*

1. *Any other Heegaard split 3–manifold satisfying these same properties with respect to the given diagram $\mathcal{D}$ is in fact orientation preserving diffeomorphic to $\mathcal{S}(\mathcal{D})$.*
2. *For every Heegaard splitting $\mathcal{S} = (M, M_1, M_2)$ of a 3–manifold M, there is a Heegaard diagram $\mathcal{D}$ such that $\mathcal{S} \cong \mathcal{S}(\mathcal{D})$.*

In other words, every Heegaard diagram corresponds to a Heegaard splitting up to diffeomorphism. Conversely, for every Heegaard splitting, we can draw a Heegaard diagram describing it.

When do two Heegaard diagrams describe equivalent Heegaard splittings? While the concept of disk slides and edge slides in Section 3.2 themselves seem to have no depth, they allow us to study Heegaard splittings diagrammatically without ambiguities.

Proposition 4.4.4. *Two Heegaard diagrams $(\Sigma_g, \alpha, \beta)$ and $(\Sigma'_g, \alpha', \beta')$ correspond to equivalent Heegaard splittings if and only if there exists a diffeomorphism $P : \Sigma_g \to \Sigma'_g$ and a sequence of handle slides which take $P(\alpha)$ to α' and $P(\beta)$ to β'.*

Uniqueness of Heegaard splittings translates into diagrams as well.

Definition 4.4.5. The **standard genus-**1 **Heegaard diagram** for S^3 is the diagram $\mathcal{D}^* = (T^2, \alpha, \beta)$ shown in Figure 5.5.

We can relate stabilization of Heegaard splittings and Heegaard diagrams as follows.

Theorem 4.4.6. *Given a Heegaard diagram $\mathcal{D}$ with associated Heegaard split 3–manifold $\mathcal{S} = \mathcal{S}(\mathcal{D})$, let $\mathcal{S}'$ be the result of stabilizing $\mathcal{S}$. Then $\mathcal{S}' \cong \mathcal{S}(\mathcal{D}\#\mathcal{D}^*)$.*

Given two Heegaard diagrams D and D', with

$$\mathcal{S}(\mathcal{D}) = (M, M_1, M_2)$$

and

$$\mathcal{S}(\mathcal{D}') = (M', M_1', M_2'),$$

we have that $M \cong M'$ if and only if, for some k and k', the two Heegaard diagrams

$$\mathcal{D}\#(\#^k\mathcal{D}^*)$$

and

$$\mathcal{D}\#(\#^{k'}\mathcal{D}^*)$$

are slide diffeomorphic.

4.5 The Fundamental Group

Finally, we will take some time to explore the algebraic data one can obtained from a Heegaard splitting of a 3-manifold. In particular, given a Heegaard splitting, we can get a presentation for the fundamental group of the 3-manifold. Let's start from the simplest kind of all—the handlebodies. The proofs in this section follow the spirit of [7].

Lemma 4.5.1. *Let H be a handlebody and let K be a spine for H with a single vertex, x, and g oriented edges $e_1, ..., e_g$. The fundamental group of H_g with base point x is the free group F^g generated by the elements $f_1, ..., f_g$ determined by the path $l_1, ..., l_g$, where $l_i : [0, 1] \to H$ is a loop whose image is the edge e_i.*

Proof. Let K be the one vertex spine of a genus g handlebody H_g, which is the wedge sum of g circles. Denote the vertex by x and the edges by e_i for $i \in \{1 \pmod g), ..., g \pmod g)\}$.

Consider the open sets $A_i = e_i \cup U$ for each i, where U is an open neighborhood of x such that $U \cap e_i \neq \emptyset$ for all i and U deformation retracts to x. $K = \bigcup_i A_i$, and each intersection $A_i \cap A_j = U$, which is path-connected. Additionally, each triple intersection $A_i \cap A_j \cap A_k = U$, which is also path-connected. Since U deformation retracts to x, $\pi_1(A_i \cap A_j) = \pi_1(U) = \{\mathbb{1}\}$. Hence, by Theorem 2.1.11, $\pi_1(K) \cong *_{1\leq i \leq g}\pi_1(A_i) = *_g\mathbb{Z} = F^g$.

Since the handlebody H_g is homotopy equivalent to its spine K, by Propsition 2.1.10, $\pi_1(H_g) \cong \pi_1(K) = F^g$. $\square$

Now we are ready to glue the simple pieces together and find a presentation for the fundamental group of more complicated 3-manifolds from their Heegaard splittings.

Lemma 4.5.2. *Let (Σ, H_1, H_2) be a Heegaard splitting. Let $D_1, D_2, ..., D_g$ be a minimal disk system for H_2 so that ∂D_i is properly embedded in $\partial H_i = \Sigma$.*

For each $i \in \{1, ..., g\}$, choose a point $x_i \in \partial D_i$ and let $k_i : [0, 1] \to \partial D_i$ be a map such that $k_i(0) = k_i(1) = x_i$.

Let $j_i : [0,1] \to H_1$ be a path such that $j_i(0) = x$ and $j_i(1) = x_i$, so that $j_i^{-1} k_i j_i$ is a closed path based at x.

The fundamental group of M is isomorphic to $\langle f_1, f_2, ..., f_g : r_1, ..., r_g \rangle$, where each $r_i \in \pi_1(M)$ is determined by the loop $j_i^{-1} k_i j_i$ and the f_i are defined as in Lemma 4.5.1 with $H = H_1$.

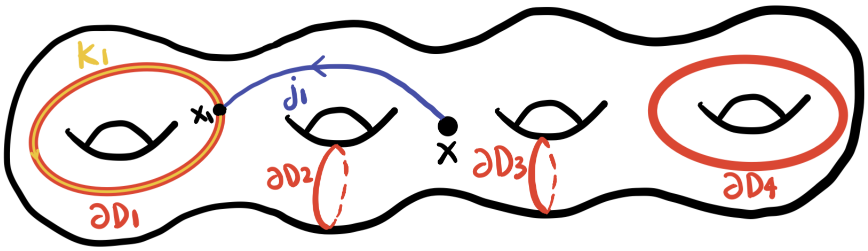

Figure 4.6: A diagram illustrating the setup of Lemma 4.5.2.

Proof. Let M_1 be the result of gluing D_1 to H_1 along the loop l_1. Let A_1 be a regular neighborhood of H_1 in M_1 and A_2 be the interior of D_1 denoted by $\text{int}D_1$. We know from Lemma 4.5.1 that $\pi_1(A_1) \cong \pi_1(H_1) \cong F^g$. On the other hand, $\pi_1(A_2) \cong \{\mathbb{1}\}$. The intersection $A_1 \cap A_2$ is an open annulus, and so $\pi_1(A_1 \cap A_2) \cong \pi_1(S^1 \times (0,1)) \cong \mathbb{Z}$. Van Kampen's theorem says that there is a homomorphism $\phi_1 : F^g * \{\mathbb{1}\} \to \pi_1(M_1)$ with $\ker(\phi_1) = \langle \{ i_{12}(w) i_{21}(w)^{-1} : w \in \pi_1(A_1 \cap A_2)\} \rangle$, where $i_{12} : \pi_1(A_1 \cap A_2) \to \pi_1(A_1)$ and $i_{21} : \pi_1(A_1 \cap A_2) \to \pi_1(A_2)$ are the natural inclusions. Note that any loop in $A_1 \cap A_2$ included into $A_1 = H_1$ must be a power of r_1. Hence, it follows that

$$\pi_1(A_1 \cap A_2) = F^g * \{\mathbb{1}\} / \ker(\phi_1) = F^g / \langle\langle r_1 \rangle\rangle,$$

where $\langle\langle r_1 \rangle\rangle$ denotes the normal closure of r_1.

Let M_2 be the result of gluing D_2 to ∂M_1 along the loop l_1. Then, by Theorem 2.1.11, there is a homomorphism $\phi_2 : \pi_1(M_1) * \{1\} \to \pi_1(M_2)$ with $\ker(\phi_2) = \langle\langle r_2 \rangle\rangle$. So the composition $\phi_2 \circ \phi_1 : \pi_1(H_1) \cong F^g * \{\mathbb{1}\} \to \pi_1(M_2)$ is a homomorphism with $\ker(\phi_2 \circ \phi_1) = \langle\langle r_1, r_2 \rangle\rangle$.

Inductively, we will obtain $\pi_1(M_g) \cong F^g / \langle\langle r_1, ..., r_g \rangle\rangle$. where M_g is the result of gluing D_i along the loop l_i for all $i \in \{1, ..., g\}$. Since the collection of disks $\{D_i : 1 \leq i \leq g\}$ form a minimal disk system, there exists an embedding $h : M_g \to M$ such that $(h(M_g))^c \subseteq M$ is an open ball B^3.

Let $A := B^3$ and $\tilde{A}$ be a regular neighborhood of $h(M_g)$, so that $M = A \cup \tilde{A}$ and $A \cap \tilde{A} \cong S^2$. Appealing to Theorem 2.1.11, again it follows that

$$\pi_1(M) \cong \{\mathbb{1}\} * \pi_1(M_g) / \pi_1(S^2) \cong \pi_1(M_g).$$

□

Example 4.5.3. ($S^1 \times S^2$) Consider the Heegaard splitting specified by the Heegaard diagram in Figure 4.7. The embedded graph with a single vertex x and a single edge e_1 is a spine for H_1, which define the generator for the free group $\pi_1(H_1)$. The simple closed curve

k_1 bounds a properly embedded disk D_1 which forms a minimal disk system for H_2, and banding k_1 with an arc j_1 form a loop based at x which is contractible in H_1. By Lemma 4.5.2, the fundamental group $S^1 \times S^2$ has presentation

$$\pi_1(S^1 \times S^2) = \langle f_1 \mid \mathbb{1} \rangle = \langle f_1 \mid \ \rangle \cong \mathbb{Z}.$$

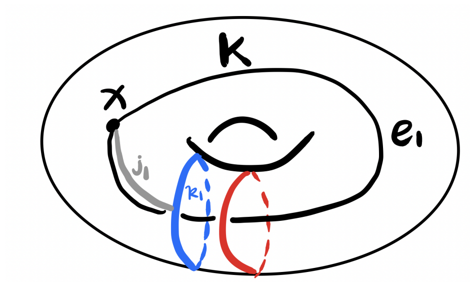

Figure 4.7: A genus-1 Heegaard splitting for $S^1 \times S^2$.

4.5.1 Stabilization and Fundamental Group

Consider a presentation of the fundamental group $\pi_1(M) = \langle f_1, ..., f_g : r_1, ..., r_g \rangle$ obtained from a Heegaard splitting (Σ, H_1, H_2), where the generators f_i come from H_1 and the relations r_j come from H_2. (This is an arbitrary choice.)

Now, when perform a stabilization, we attach to H_1 an unknotted 1-handle B, whose core union with an arc α in H_1 bounds a disk in M. Moreover, $B \cup N(\alpha)$ (a regular neighborhood of α) is homotopic to a solid torus.

Let $A_\alpha := H_1$ and $A_\beta := B \cup N(\alpha)$, so that $H_1' = A_\alpha \cup A_\beta$. The double intersection is $A_\alpha \cap A_\beta \cong B^2 \times I$, which is path connected. Moreover, $\pi_1(A_\alpha \cap A_\beta) = \pi_1(B^2 \times I) = \{\mathbb{1}\}$. Thus, by Theorem 2.1.11,

$$\begin{aligned}\pi_1(H_1') &= \pi_1(A_\alpha) * \pi_1(A_\beta)/N = \pi_1(H_1) * \pi_1(S^1)/\{\mathbb{1}\} \\ &= \langle f_1, ..., f_g, f_{g+1} \rangle.\end{aligned}$$

On the other hand, drilling out a $B^2 \times I$ from H_2 is equivalent to adding a one handle to it and thus adding a disk to the minimal disk system, whose boundary is x_{g+1} on Σ'. Thus, the relations coming from H_2' are $r_1, ..., r_g, x_{g+1}$. So this presentation is actually identical to the original one up to triviality. Let's record this as the following lemma.

Lemma 4.5.4. *Let $\pi_1(M) = \langle f_1, ..., f_g : r_1, ..., r_g \rangle$ obtained from a Heegaard splitting (Σ, H_1, H_2). Then, performing a stabilization yields the new presentation of the fundamental group of M*

$$\pi_1(M) \cong \langle f_1, ..., f_g, f_{g+1} : r_1, ..., r_g, f_{g+1} \rangle.$$

Chapter 5

Trisecting 4-Manifolds

In this chapter, we will venture into dimension four and generalize many ideas from the previous section to 4-manifolds. Trisections in dimension-4 are analogous to Heegaard splittings in dimension-3. The whole chapter follows [5] closely. It is also where the definitions and lemmas, unless otherwise specified, come from.

5.1 The Basic Definitions: Decompositions

We will denote a 4-dimensional 1-handlebody with a handle structure of one 0-handle and k 1-handles by $\natural^k S^1 \times B^3$.

Definition 5.1.1. A $(g; k_1, k_2, k_3)$ trisection of a closed, connected, oriented 4–manifold X is a decomposition $X = X_1 \cup X_2 \cup X_3$ such that:

1. For each $i \in \{1, 2, 3\}$, X_i is diffeomorphic to $\natural^{k_i} S^1 \times B^3$, i.e. a "genus-$k_i$" 4-dimensional 1-handlebody.

2. Taking indices mod 3, $X_i \cap X_{i+1} = H_{i(i+1)}$ is diffeomorphic to $\natural^g S^1 \times B^2$ for each $i \in \{1, 2, 3\}$, i.e. a 3D 1-handlebody. In particular, we orient $X_i \cap X_{i+1}$ as a submanifold of ∂X_{i+1}, and $\partial X_i = H_{i(i-1)} \cup H_{i(i+1)}$ is a genus-g Heegaard splitting for $\partial X_i = \natural^{k_i} S^1 \times S^2$.

3. $\Sigma_g := X_1 \cap X_2 \cap X_3 = \partial H_{i(i+1)}$ for $i \in \{1, 2, 3\}$ is a genus-g surface. We orient $X_1 \cap X_2 \cap X_3$ as $\partial(X_1 \cap X_2) = \partial(X_2 \cap X_3) = \partial(X_3 \cap X_1)$.

The surface Σ_g is called the **trisection surface**, and we will denote the trisection by the quadruple $\mathcal{T} = (X, X_1, X_2, X_3)$. (See Figure 5.1.)

A trisection is called **balanced** if $k_1 = k_2 = k_3 = k$, in which case we call it a (g, k)-trisection.

We again define two notions of equivalence for trisections. We will say that two trisections of a 4-manifold $\mathcal{T} = (X, X_1, X_2, X_3)$ and $\mathcal{T}' = (X, X_1', X_2', X_3')$ are **diffeomorphic** if there is a diffeomorphism of X such that $f(X_i) = X_i'$ for $i \in \{1, 2, 3\}$. Two trisections are **isotopic** if there is an isotopy f_t of X such that $f_0 = id$ and $f_1(X_i) = X_i'$ for $i \in \{1, 2, 3\}$.

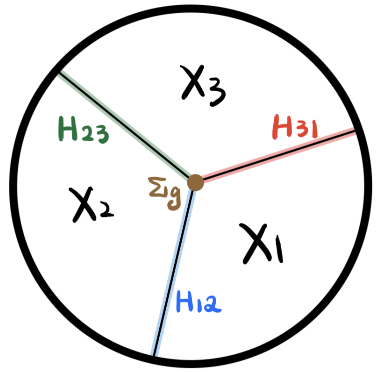

Figure 5.1: Schematic of a trisection

Remark 5.1.2. Given a $(g,k)-$trisection of X, $\chi(X) = 2 + g - 3k$, which means that k can be determined by g and X. Thus, we usually refer to a $(g,k)-$trisection simply as a genus-g trisection of X. [4]

We can reconstruct from $H_{12} \cup H_{23} \cup H_{31}$ uniquely the 4-manifold X and the trisection structure on X as follows:

1. Thicken $H_{12} \cup H_{23} \cup H_{31}$ by taking the product $\Sigma_g \times B^2$ and $H_{ij} \times I$ for $i \neq j \in \{1,2,3\}$. This leaves a manifold with 3 boundary component, each diffeomorphic to $\natural^k S^1 \times S^2$;

2. One can show that any diffeomorphism of $\natural^{k_i} S^1 \times S^2$ extends across H^4_k. Consequently, each of the boundary components can be capped off with H^4_k uniquely.

Since we made no choices throughout the process, we have uniquely constructed a trisected 4-manifold.

Trisections exist for a large family of nice 4-manifolds.

Theorem 5.1.3 ([4])**.** *Every closed, connected, oriented 4–manifold has a trisection.*

Just as Heegaard splittings in dimension-3 arises naturally from Morse function on the 3-manifold, trisections arise naturally from Morse 2-functions on 4-manifolds, which we will define and discuss in the next section.

5.2 The Morse Theoretic Perspective

Definition 5.2.1. A **Morse 2-function** on a 4-manifold X is a smooth function $f : X \to \mathbb{R}^2$ which, at every point $p \in X$, has one of the following three forms with respect to appropriate local coordinates (t,x,y,z) near p and u,v near $f(p)$:

1. $(t,x,y,z) \mapsto (u = t, v = x)$; here p is called a **regular point**.

2. $(t,x,y,z) \mapsto (u = t, v = \pm x^2 \pm y^2 \pm z^2)$; here p is called a **fold point**, and p is called **definite** or **indefinite** depending on whether the quadratic form $\pm x^2 \pm y^2 \pm z^2$ is definite or indefinite.

3. $(t, x, y, z) \mapsto (u = t, v = x^3 - tx \pm y^2 \pm z^2)$; here p is called a **cusp point**.

Both fold and cusp points are **critical points**. A point $q \in \mathbb{R}^2$ is called a **regular value** if all points $p \in f^{-1}(q)$ are regular points; otherwise, q is a **critical value**.

The inverse image of a regular value $(u, v) \in \mathbb{R}^2$ is a closed surface, as it is given by $f^{-1}(u, v) = \{(u, v, y, z) : y, z \in \mathbb{R}\}$ locally. The singular locus, i.e. the set of all critical points, is a finite collection of embedded circles in dimension four. The cusp points form a finite collection of points on the singular locus.

If $f : X \to \mathbb{R}^2$ is a Morse 2-function and A is an arc in $\mathbb{R}^2$ avoiding the cusps and transverse to the image of the singular locus, then the inverse image $M = f^{-1}(A)$ is a 3-manifold in X, with $\partial M = f^{-1}(\partial A)$. Moreover, if we embed A in $\mathbb{R}$, then, $f|_M : M \to A$ is a Morse function with critical points of index 0 and 3 where A crosses definite folds and index 1 and 2 where A crosses indefinite folds. If we reverse the orientation of A, every critical point of index j is changed to one of index $3 - j$.

Crossing a definite fold in the index 0 direction adds a new S^2 component to the preimage of regular value, while crossing a definite fold in the index 3 direction cap off such a component.

Crossing an indefinite fold in the index 1 either increases the genus of a fiber component by 1, or connects two disconnected components. Crossing an indefinite fold in the index 2 cuts the fibre along a compressing circle, either decreasing the genus by 1 or splitting a connected component into two pieces.

Passing immediately adjacent to a cusp and thus crossing two successive folds, the corresponding critical points are two canceling critical points with successive indices.

In [5], David Gay includes the following lovely diagram illustrating the features of a Morse 2-function mentioned above.

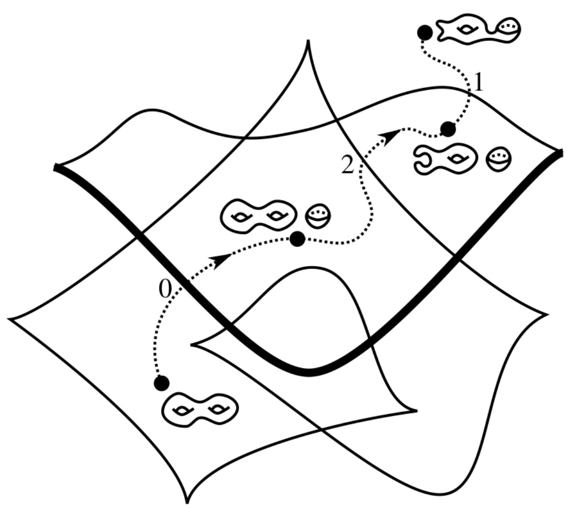

Figure 5.2: The darker arc is a definite fold and the remaining solid arcs are indefinite folds. The dotted arc is an oriented arc A transverse to the folds, with indices of the critical points of the associated Morse function on $f^{-1}(A)$ labelled at the crossings. The surfaces are representative inverse images of regular values along the arc. [5]

To emphasize the parallel, we will set up the same notations as that for Heegaard splitting Morse function to describe what Morse 2-function will produce a trisection. [5]

Let $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ and denote $R_\theta \subset \mathbb{R}^2$ be the ray starting at the origin and making angle $-\theta$ with the positive x-axis. Parametrize R_θ with $R_\theta(t) = (t\cos(-\theta), t\sin(-\theta))$ for $t \in [0,\infty)$, so that we always see R_θ oriented away from the origin. For $i \in \mathbb{Z}/3\mathbb{Z}$, let A_i be the sector bounded by the rays $R_{2\pi(i-1)/3}$ and $R_{2\pi(i)/3}$, which produces a decomposition $\mathbb{R}^2 = A_1 \cup A_2 \cup A_3$.

Definition 5.2.2 (Trisecting Morse 2–function)**.** A $(g;k_1,k_2,k_3)$ **trisecting Morse 2–function** f on a 4–manifold X is a Morse 2–function $f : X \to \mathbb{R}^2$ such that:

1. $(0,0)$ is a regular value of f, so that $f^{-1}(0) = \Sigma$ is a closed surface. In particular, we require Σ to be connected of genus g.

2. On each of the three rays $R_0, R_{2\pi/3}$ and $R_{4\pi/3}$, f has exactly g index 2 and one index 3 critical points, all of which have distinct critical values.

3. Over each sector A_i, the singular locus of f has exactly $g+1$ components, all of which are arcs from $R_{2\pi i/3}$ to $R_{2\pi(i+1)/3}$. Among the arcs, $g-k_i$ of them are indefinite folds each with exactly one indefinite cusp, g of them are indefinite folds with no cusps, and one of them (which is the outermost) is a definite fold. Furthermore, each of the arcs is transverse to the ray R_θ for all $\theta \in [2\pi i/3, 2\pi(i+1)/3]$ except at cusps, where they are tangent to the rays. f restricted to the singular locus is an immersion with cusps and double points avoiding the cusps.

Again, we include an illustrating diagram for a trisection Morse 2-function is Figure 5.3 given in [5]. On top of the original figure, we highlight the definite folds in purple, the indefinite folds without cusps in yellow, and the indefinite folds with cusps in pink.

A $(g;k_1,k_2,k_3)$ trisecting Morse 2–function induces a trisection as desired.

Lemma 5.2.3 ([5])**.** *Given a $(g;k_1,k_2,k_3)$ trisecting Morse 2–function $f : X \to \mathbb{R}^2$, let $X_i = f^{-1}(A_i)$. Then, $X = X_1 \cup X_2 \cup X_3$ is a $(g;k_1,k_2,k_3)$ trisection of X.*

Example 5.2.4 ((0,0)-trisection of S^4 produced by a Morse 2-function, [5])**.** Write $S^4 = \mathbb{R}^2 \times \mathbb{R}^3 = \{(r,\theta,x_3,x_4,x_5) : r^2 + x_3^2 + x_4^2 + x_5^2 = 1\}$. Define the projection $p : B^4 \to \mathbb{R}^2$ as

$$p(r,\theta,x_3,x_4,x_5) = (r,\theta)$$

and trisect the plane into

$$A_1 = \{(r,\theta) : 0 \le r \le 1, 0 \le \theta \le \frac{2\pi}{3}\},$$

$$A_2 = \{(r,\theta) : 0 \le r \le 1, \frac{2\pi}{3} \le \theta \le \frac{4\pi}{3}\},$$

$$A_3 = \{(r,\theta) : 0 \le r \le 1, \frac{4\pi}{3} \le \theta \le 2\pi\}.$$

Then, letting $X_i := p^{-1}(D_i)$ for each $i \in \{1,2,3\}$, we'll obtain $X_i \cong B^4$ and $S^4 = X_1 \cup X_2 \cup X_3$. Moreover, $X_i \cap X_j \cong B^3$ and

$$\Sigma = X_1 \cap X_2 \cap X_3 = \{(0,0,x_3,x_4,x_5) : x_3^2 + x_4^2 + x_5^2 = 1\} \cong S^2.$$

So (X, X_1, X_2, X_3) is a $(0,0)$-trisection of S^4.

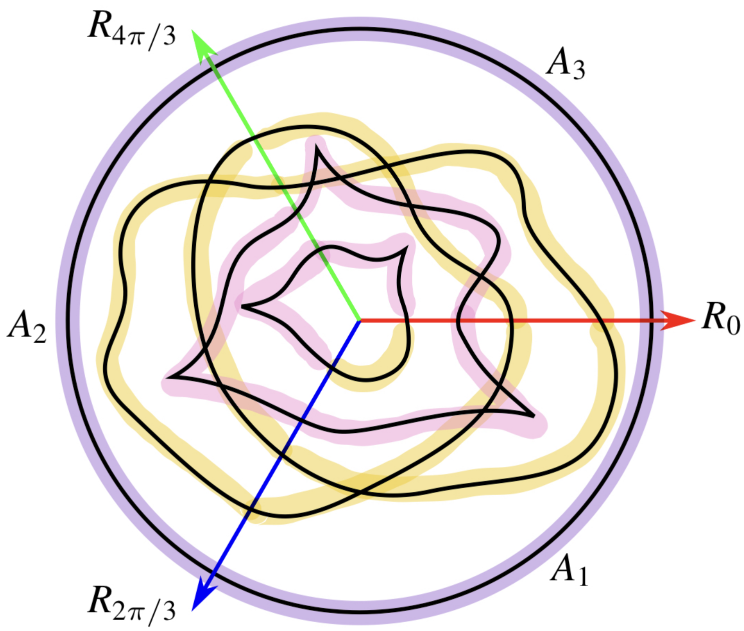

Figure 5.3: The singular value set of a $(4; 3, 2, 2)$ trisecting Morse 2-function. [5]

As mentioned in [5], one can write down a trisecting Morse 2-function on $\mathbb{CP}^2$ by perturbing the moment map $\mu : \mathbb{CP}^2 \to \mathbb{R}^2$ defined by

$$[z_1 : z_2 : z_3] \mapsto \left(\frac{|z_1|^2}{|z_1|^2 + |z_2|^2 + |z_3|^2}, \frac{|z_2|^2}{|z_1|^2 + |z_2|^2 + |z_3|^2}\right).$$

However, we can actually extract a trisection from this map already without perturbing it into a Morse-2 function.

Example 5.2.5 ((1,0)-trisection of $\mathbb{CP}^2$ produced by the moment map)**.** Write $\mathbb{CP}^2 = \{[z_1, z_2, z_3] : z_i \in \mathbb{C}, [z_1 : z_2 : z_3] = [\lambda z_1 : \lambda z_2 : \lambda z_3], \lambda \in \mathbb{C} \setminus \{0\}\}$. Consider the map $\mu : \mathbb{CP}^2 \to \mathbb{R}^2$ defined by

$$[z_1 : z_2 : z_3] \mapsto \left(\frac{|z_1|^2}{|z_1|^2 + |z_2|^2 + |z_3|^2}, \frac{|z_2|^2}{|z_1|^2 + |z_2|^2 + |z_3|^2}\right).$$

One can verify that the image of the map is the triangle T with vertices $(0,0), (0,1), (1,0)$ on the plane. In particular:

1. For $p \in \{(0,0), (0,1), (1,0)\}$, $\mu^{-1}(p)$ are points;

2. For $p \in \text{int}(T)$, $\mu^{-1}(p)$ are tori;

3. For $p \in \partial T \setminus \{(0,0), (0,1), (1,0)\}$, $\mu^{-1}(p)$ are circles.

To obtain a trisection of $\mathbb{CP}^2$, we will "trisect" this triangle T in $\mathbb{R}^2$ and lift this decomposition to $\mathbb{CP}^2$. Mark $(\frac{1}{3}, \frac{1}{3}) \in T$ and connect it to the midpoints of the three edges respectively. This decomposes T into three closed disks D_i for $i \in \{1, 2, 3\}$ with boundary $\partial D_i = l_i$. Lifting this decomposition to $\mathbb{CP}^2$, we will see that:

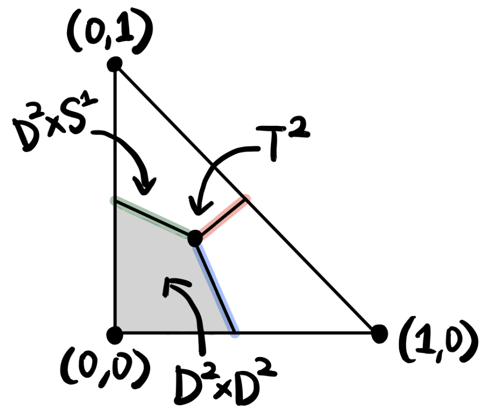

Figure 5.4: The trisected triangle $T \subset \mathbb{R}^2$.

1. $\Sigma = \mu^{-1}((1,1)) \cong T^2$;
2. $\mu^{-1}(l_i) \cong D^2 \times S^1$;
3. $\mu^{-1}(D_i) \cong B^4$.

5.3 Stabilization of Trisection

Definition 5.3.1 (Stabilization)**.** Given a trisection $\mathcal{T} = (X, X_1, X_2, X_3)$ of a 4–manifold X and an index $i \in \mathbb{Z}/3\mathbb{Z}$, an i–**stabilization** of this trisection is a trisection $\mathcal{T}' = (X, X_1', X_2', X_3')$ obtained as follows:

1. Choose some properly embedded and boundary parallel arc a in $H_{(i-1)(i+1)} = X_{i-1} \cap X_{i+1}$, with a regular neighborhood $\nu \cong B^3 \times a$, so that $\nu \cap X_i \cong B^3 \times \partial a$ and and $\nu \cap X_1 \cap X_2 \cap X_3 \cong B^2 \times \partial a$.
2. Let $X_i' = X_i \cup \nu$;
3. Let $X_{i\pm1}' = X_{i\pm1} \setminus \mathring{\nu}$.

Any two i–stabilizations of the same trisection are isotopic. However, 1–stabilizations, 2 –stabilizations and 3–stabilizations are necessarily different. In particular, an i–stabilization turns a $(g; k_1, k_2, k_3)$ trisection into a $(g+1; k_1', k_2', k_3')$ trisection where $k_i' = k_i + 1$ and, for $j \neq i$, $k_j' = k_j$. For a balanced (g,k)-trisection, a "stabilization" means performing one 1–stabilization, one 2–stabilization and one 3–stabilization, which is defined more precisely in Definition 5.3.2 below. This contrasts with stabilizations of Heegaard splittings in dimension-3. While a stabilization of a Heegaard splitting also increases the genus by 1, there is a unique stabilization up to ambient isotopy.[5]

Definition 5.3.2 (Balanced stabilization of a trisection)**.** Given a (g,k)-trisection of a 4-manifold (X, X_1, X_2, X_3), we can construct a new trisection (X, X_1', X_2', X_3') as follows: For every $i \neq j \in \{1,2,3\}$, let H_{ij} be the handlebody $X_i \cap X_j$, with boundary $\Sigma_g = X_1 \cap X_2 \cap X_3$.

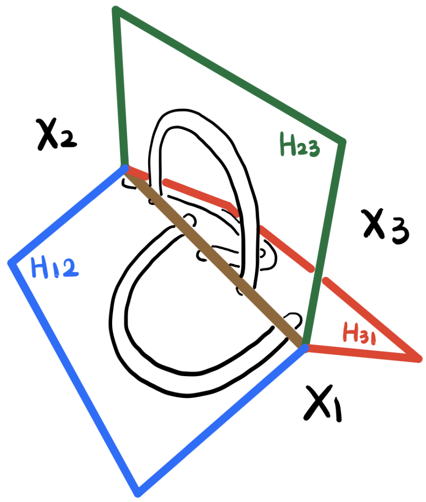

Figure 5.5: Balanced stabilization in dimension 3. [4]

Let a_{ij} be a properly embedded boundary parallel arc in each H_{ij}, such that the endpoints of a_{12}, a_{23} and a_{31} are disjoint in Σ_g. Let ν_{ij} be a closed 4-dimensional regular neighborhood of a_{ij} in X (and thus homeomorphic to B^4), with ν_{12}, ν_{23} and ν_{31} disjoint and $\mathring{\nu}_{ij}$ an open 4-dimensional regular neighborhood.

Define

1. $X_1' = (X_1 \cup \nu_{23}) \setminus (\mathring{\nu}_{31} \cup \mathring{\nu}_{12})$;
2. $X_2' = (X_2 \cup \nu_{31}) \setminus (\mathring{\nu}_{23} \cup \mathring{\nu}_{12})$;
3. $X_1' = (X_3 \cup \nu_{12}) \setminus (\mathring{\nu}_{23} \cup \mathring{\nu}_{31})$.

This operation of replacing (X, X_1, X_2, X_3) with (X, X_1', X_2', X_3') is called the **balanced stabilization** of a (g, k)-trisection.

Since any two boundary parallel arcs in a handlebody are isotopic, this operation does not depend on the choice of arcs or neighborhoods. [4]

Just as a stabilization of a Heegaard splitting is again a Heegaard splitting, the stabilization (X, X_1', X_2', X_3') of a trisection of X is again a trisection of X.

For each X_i, adding the 4-dimensional neighborhood ν_{jk} of an arc, which meets X_i in two B^3, is equivalent to adding a 4-dimensional 1-handle to X_i. On the other hand, since $\nu_{ij} \subset H_{ij} \subset \partial X_i$ and $\nu_{ki} \subset H_{ki} \subset \partial X_i$, removing the two open neighborhood $\mathring{\nu}_{ij} \cup \mathring{\nu}_{ki}$ from the boundary of X_i can be realized by an isotopy of X_i and thus does not change its diffeomorphism type. Hence, after a balanced stabilization, $X_i = \natural^k S^1 \times B^3$ becomes $X_i' = \natural^{k+1} S^1 \times B^3$ for each $i \in \{1, 2, 3\}$ and the genus becomes $g' = g + 3$.

We can now add a uniqueness statement to our existence statement in Theorem 5.1.3.

Theorem 5.3.3. *Every closed, connected, oriented 4–manifold has a trisection, and any two trisections of the same 4–manifold become isotopic after some number of stabilizations.*

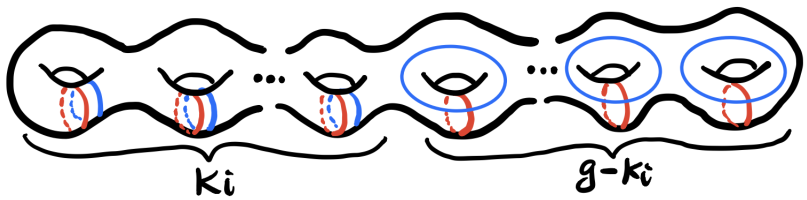

Figure 5.6: Standard diagrams needed for the definition of trisection diagrams.[5]

More precisely, let (X, X_1, X_2, X_3) and (X, X_1', X_2', X_3') be two trisections of X. Then, after stabilizing each trisection some number of times, there is a diffeomorphism $h : X \to X$ isotopic to the identity with the property that $h(X_i) = X_i'$ for each i. In particular, $h(X_i \cap X_j) = X_i \cap X_j'$, for $i \neq j \in \{1, 2, 3\}$, and $h(X_1 \cap X_2 \cap X_3) = h(X_1' \cap X_2' \cap X_3')$. [4]

5.4 The Diagrammatic Perspective

Recall that a Heegaard diagram encodes the splitting of a 3-manifold with an oriented surface and two sets of simple closed curves. In this section, we will see how one can understand trisections of 4-manifolds diagrammatically in an analogous way.

Definition 5.4.1 (Trisection diagrams)**.** A $(g; k_1, k_2, k_3)$ **trisection diagram** is a tuple $(\Sigma_g, \alpha, \beta, \gamma)$, where Σ_g is a closed oriented surface of genus g and the triples $(\Sigma, \alpha, \beta), (\Sigma, \beta, \gamma)$ and (Σ, γ, α) are each slide diffeomorphic to the standard Heegaard diagram $(\Sigma_g, \alpha^{g,k_i}, \beta^{g,k_i})$ shown in Figure 5.6. (Here $i = 1$ for (α, β), $i = 2$ for (β, γ) and $i = 3$ for (γ, α).) [5]

Or equivalently, α, β, γ each form a cut system of Σ, and the triples $(\Sigma, \alpha, \beta), (\Sigma, \beta, \gamma)$ and (Σ, γ, α) each forms a Heegaard diagram for $\#^{k_i} S^1 \times S^2$. [4]

Now let's relate trisection diagrams to trisections.

Lemma 5.4.2 ([5])**.** *Given a trisection diagram $\mathcal{D} = (\Sigma, \alpha, \beta, \gamma)$, there is a 4–manifold $X = X(\mathcal{D})$ with trisection $\mathcal{T}(\mathcal{D}) = (X, X_1, X_2, X_3)$ such that $\Sigma = X_1 \cap X_2 \cap X_3$, oriented according to the conventions in Definition 5.1.1, and such that the α curves bound embedded disks in $X_3 \cap X_1$, the β curves in $X_1 \cap X_2$, and the γ curves in $X_2 \cap X_3$.*

Moreover, any other trisected 4–manifold satisfying these same properties with respect to the given diagram $\mathcal{D}$ is in fact orientation preserving diffeomorphic to $\mathcal{T}(\mathcal{D})$.

For every trisection $\mathcal{T} = (X, X_1, X_2, X_3)$ of a 4–manifold X there is a trisection diagram $\mathcal{D}$ such that $\mathcal{T} \cong \mathcal{T}(\mathcal{D})$.

Figure 5.7 exhibits a selection of trisection diagrams. The easiest way to see this is probably the first characterization in Definition 5.4.1. The triples (Σ, α, β), (Σ, β, γ), and (Σ, γ, α) of all the trisection diagrams but the genus-3 one already forms standard Heegaard diagrams for $\#^{k_i} S_1 \times S^2$.

Example 5.4.3. Let's show that Figure 5.8 is a $(1; 1, 0, 0)-$trisection of S^4.

Let $B^5 = B^3 \times B^2 = \{(x, y) : x \in B^3, y \in B^2\}$ and $S^4 = \partial B^5$. Then, we can write

$$\begin{aligned} S^4 &= \partial(B^3 \times B^2) = (\partial B^3 \times B^2) \cup (B^3 \times \partial B^2) \\ &= (S^2 \times B^2) \cup (B^3 \times S^1) \\ &= (B_+^2 \times B^2) \cup (B_-^2 \times B^2) \cup (B^3 \times S^1), \end{aligned} \tag{5.1}$$

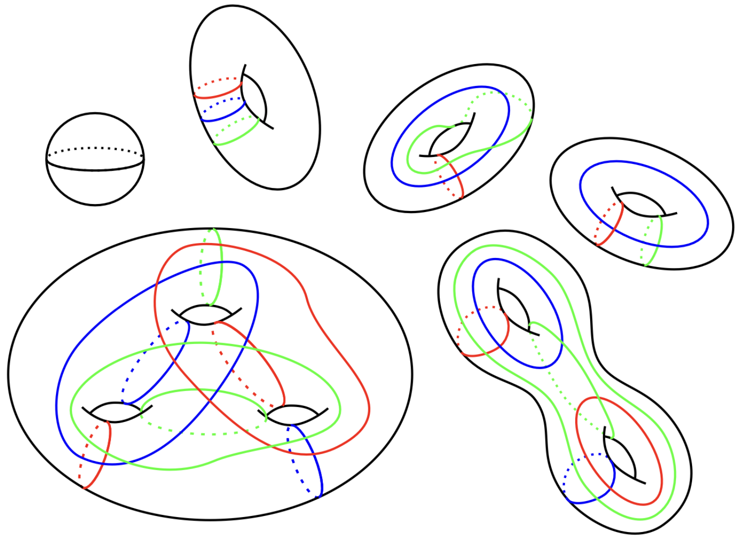

Figure 5.7: A selection of trisection diagrams. Red is α, blue is β and green is γ. [5]

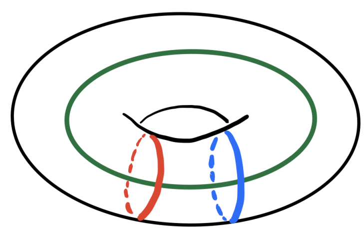

Figure 5.8: The standard $(1; 1, 0, 0)$-trisection of S^4.

where the last equality was obtained by writing $S^2 = B^2_+ \cup B^2_-$. Now, let $X_2 = B^2_+ \times B^2$, $X_3 = B^2_- \times B^2$, and $X_1 = B^3 \times S^1$. Then:

1. $X_2 \cong X_3 \cong B^4$ and $X_1 = B^3 \times S^1$;

2. $X_i \cap X_{i+1} \cong S_1 \times B^2$. Indeed,

$$X_2 \cap X_3 = (B^2_+ \times B^2) \cap (B^2_- \times B^2) = (B^2_+ \cap B^2_-) \times B^2 = S^1 \times B^2,$$

$$X_3 \cap X_1 = (B^2_- \times B^2) \cap (B^3 \times S^1) = (B^2_- \cap B^3) \times (B^2 \cap S^1) = B^2_- \times S^1,$$
$$X_1 \cap X_2 = (B^3 \times S^1) \cap (B^2_+ \times B^2) = (B^3 \cap B^2_+) \times (S^1 \cap B^2) = S^1 \times B^2_+;$$

3. The triple intersection is

$$\begin{aligned} X_1 \cap X_2 \cap X_3 &= (X_3 \cap X_1) \cap (X_2 \cap X_3) \\ &= (S^1 \times B^2_+) \cap (B^2_- \times S^1) \\ &= (B^2_+ \cap B^2_-) \times S^1 \\ &= S^1_x \times S^1_y, \end{aligned}$$

where $S^1_x \subset B^3$ and $S^1_y = \partial B^2$.

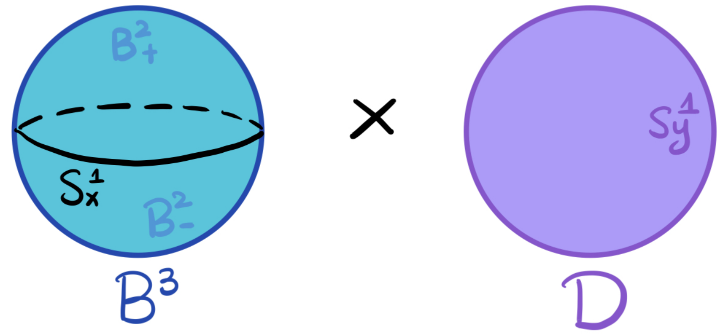

Figure 5.9: Diagram illustrating Example .

In particular, for every $x \in S^1_x$, $\{x\} \times S^1_y$ bounds a disk in $X_2 \cap X_3$; for every $y \in S^1_y$, $S^1_x \times \{y\}$ bounds a disk in both $X_3 \cap X_1$ and $X_1 \cap X_2$. In view of Definition 5.4.1, the tuple $(S^4; X_1, X_2, X_3)$ thus forms a genus-1 trisection corresponding to the diagram in Figure 5.8.

Recall that stabilization of Heegaard splittings can be related to Heegaard diagrams. Similarly, stabilization of trisections also translates into operation on the trisection diagrams. As before, we will define diagrams which will be called "standard", which we've just shown to be trisection diagrams of S^4.

Definition 5.4.4. The standard genus 1 $(1; 1, 0, 0)$ trisection diagram for S^4 is the diagram $\mathcal{D}^*_1 = (T^2, \mu, \mu, \lambda)$ shown in Figure 5.8. Cyclically permuting the curve systems gives the standard $(1; 0, 1, 0)$ diagram $\mathcal{D}^*_2$ and the standard $(1; 0, 0, 1)$ diagram $\mathcal{D}^*_3$.

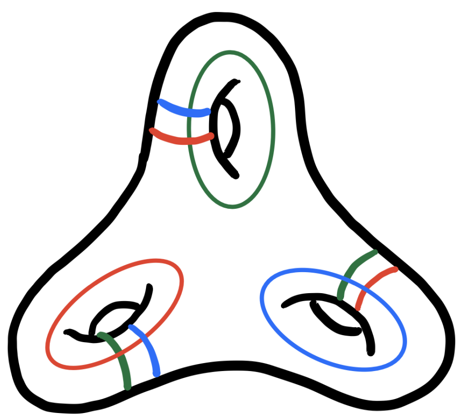

Figure 5.10: The connected sum of three unbalanced trisections of genus 1 gives a balanced $(3, 1)$-trisection for S^4.

Lemma 5.4.5 ([5])**.** *Given a trisection diagram $\mathcal{D}$ with associated trisected 4–manifold $\mathcal{T} = \mathcal{T}(\mathcal{D})$, let $\mathcal{T}'$ be the result of an i–stabilization of $\mathcal{T}$. Then $\mathcal{T}' \cong \mathcal{T}(\mathcal{D}\#\mathcal{D}^*_i)$.*

Given two trisection diagrams $\mathcal{D}$ and $\mathcal{D}'$, with $\mathcal{T}(\mathcal{D}) = (X, X_1, X_2, X_3)$ and $\mathcal{T}(\mathcal{D}') = (X', X'_1, X'_2, X'_3)$, we have that $X \cong X'$ if and only if, for some k_1, k_2, k_3 and k'_1, k'_2, k'_3, the following two trisection diagrams are slide diffeomorphic:

$$\mathcal{D}\#(\#^{k_1}\mathcal{D}^*_1)\#(\#^{k_2}\mathcal{D}^*_2)\#(\#^{k_3}\mathcal{D}^*_3)$$

and

$$\mathcal{D}\#(\#^{k_1'}\mathcal{D}_1^*)\#(\#^{k_2'}\mathcal{D}_2^*)\#(\#^{k_3'}\mathcal{D}_3^*).$$

As a consequence of Lemma 5.4.4, a balanced stabilization for a balanced trisection translates to trisection diagram as follows. Given a trisection diagram $\mathcal{D} = (\Sigma_g, \alpha, \beta, \gamma)$ for a balanced (g, k)-trisection, $\mathcal{T} \cong \mathcal{T}(\mathcal{D}\#\mathcal{D}^*)$, where $\mathcal{D}^* = \mathcal{D}_1^*\#\mathcal{D}_2^*\#\mathcal{D}_3^*$ is a balanced $(3,1)$-trisection of S^4, as shown in Figure 5.10. In particular, $\mathcal{T}(\mathcal{D}\#\mathcal{D}^*)$ is a $(g+3, k+1)$-trisection.

5.5 The Fundamental Group

Finally, let's just briefly remark that one can extract information about the fundamental group of a 4-manifold from a trisection just as we did for Heegaard splitting of a 3-manifold.

Provided that we have a $(g; k_1, k_2, k_3)$-trisection $\mathcal{T} = (X, X_1, X_2, X_3)$, we have $\pi_1(H_{i(i+1)}) \cong \pi_1(\natural^g S^1 \times B^2) \cong F^g$, the free group on g generators. Moreover, each of the X_i has fundamental group $\pi_1(X_i) \cong \pi_1(\natural^{k_i} S^1 \times B^3) \cong F^{k_i}$, the free group on k_i generators.

Since $X_1 \cap X_2 = H_{12}$ is path-connected, we can apply Theorem 2.1.11 to first find a presentation of the fundamental group $\pi_1(X_1 \cup X_2)$ by taking the free product of $\pi_1(X_1)$ and $\pi_1(X_2)$ amalgamated along $\pi_1(X_1 \cap X_2)$. Then, take $\tilde{X} = X_1 \cup X_2$, so that $\tilde{X} \cup X_3 = X$ and $\tilde{X} \cap X_3 = H_{23} \cup H_{31}$. Applying Theorem 2.1.11 again, we can find a presentation for $\pi_1(X)$ by taking the amalgamated free product of $\pi_1(\tilde{X})$ and $\pi_1(X_3)$ along $\pi_1(H_{23} \cup H_{31})$.

Bibliography

[1] Hatcher, Allen. *Algebraic Topology.* Cambridge University Press, 2019.

[2] Hatcher, Allen. *Notes on introductory point-set topology.* Cornell university. Retrieved May 6, 2022, from https://pi.math.cornell.edu/ hatcher/Top/TopNotes.pdf

[3] Lee, John Marshall. *Introduction to Smooth Manifolds.* Springer, 2013.

[4] David Gay and Robion Kirby. *Trisecting 4-manifolds.* Geom. Topol., 20(6):3097–3132, 2016.

[5] David T Gay. "From Heegaard splittings to trisections; porting 3-dimensional ideas to dimension 4". Winter Braids Lecture Notes, Tome 5 (2018), Exposé no. 4, 19 p. doi : 10.5802/wbln.24.

[6] Schultens, Jennifer. *Introduction to 3-manifolds.* American Mathematical Society, 2014.

[7] Johnson, Jesse. *Notes on Heegaard Splittings.*

[8] Kervaire, M. A., and John W. Milnor. *Groups of Homotopy Spheres, I.* New York University, Courant Institute of Mathematical Sciences, 1961.

[9] Milnor, John W. Morse Theory. Based on Lecture Notes by M. Spivak and R. Wells. 1963.

[10] Heegaard, Poul (1898), "Forstudier til en topologisk Teori for de algebraiske Fladers Sammenhang", Thesis (in Danish).

[11] Smale, Stephen (1898), "On the structure of manifolds" Amer. J. Math. , 84 (1962) pp. 387–399.

[12] Tu, Loring W. *An Introduction to Manifolds.* Springer New York, 2011.

[13] *Introduction to Differential Geometry-Math.toronto.edu.* https://www.math.toronto.edu/~mein/teaching/MAT367/DiffGeomNotes.pdf.