2017

Normal Surfaces and 3-Manifold Algorithms

Josh D. Hews
Colby College

Normal Surfaces and 3-Manifold Algorithms

Joshua D. Hews

May 15, 2017

Abstract

This survey will develop the theory of normal surfaces as they apply to the S^3 recognition algorithm. Sections 2 and 3 provide necessary background on manifold theory. Section 4 presents the theory of normal surfaces in triangulations of 3-manifolds. Section 6 discusses issues related to implementing algorithms based on normal surfaces, as well as an overview of the Regina, a program that implements many 3-manifold algorithms. Finally section 7 presents the proof of the S^3 recognition algorithm and discusses how Regina implements the algorithm.

Acknowledgements

I would first like to thank Professor Scott Taylor for adivising me on this project. I would not have been able to complete this survey without his incredible support. I would also like to thank Professor David Krumm for being my second reader. His comments were extremely helpful. Finally I would like to acknowledge my reliance on *Introduction to 3-Manifolds* by Jennifer Schultens and *Algorithmic Topology and Classification of 3-Manifolds* by Sergei Matveev. These texts were a signficant resource in developing the background necessary to understand the S^3 recognition algorithm and the program Regina.

Contents

1 Introduction

This paper is meant to be a survey of algorithmic 3-manifold theory, from the perspective of normal surfaces. We begin with an overview of manifolds before exploring 3-manifolds in more depth. The goal of section 2 and section 3 is to provide the language necessary to properly discuss the theory of normal surfaces and 3-manifold algorithms. The section 4 fully develops normal surface theory. This is the largest section and represents the meat of this paper. After this, we present the general structure of 3-manifold algorithms that utilize normal surfaces. We forgo an example that follows this structure, as these examples tend to quickly reduce to technical arguments that are not informative from the viewpoint of a survey. Instead we discuss questions of implementation, focusing on the program Regina [3] that implements many 3-manifold algorithms. Finally we conclude with an examination of the algorithm to recognize the 3-sphere. The 3-sphere recognition problem is interesting both from the perspective of the topological arguments necessary to prove the existence of the algorithm, as well as the tools needed for implementing the algorithm in software.

2 Manifolds

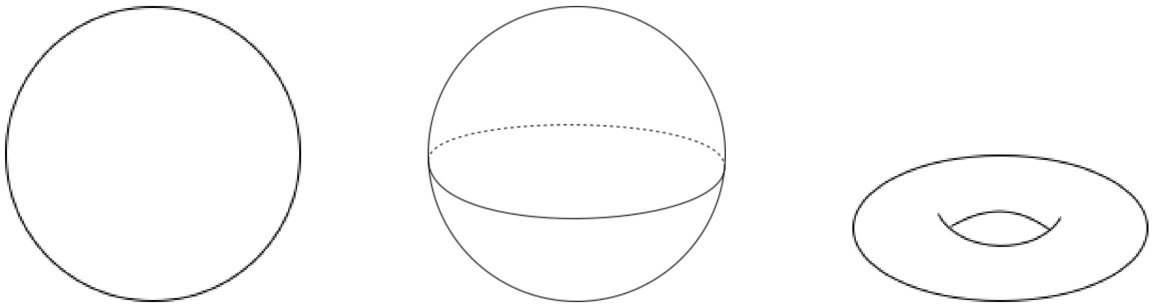
We will be focusing on 3-manifolds, and in particular triangulations of 3-manifolds. However it will be useful to introduce manifolds in full generality. There are multiple ways of viewing and defining manifolds depending on the analysis one wants to do. Basically these can be split into **Topological Manifolds**, **Differentiable Manifolds**, and **Triangulated Manifolds**. Topological manifolds are the broadest form of manifolds, whereas the other three are topological manifolds with additional structure. We will begin our discussion with Topological Manifolds before focusing most of our attention on Triangulated Manifolds.

2.1 Topological Manifold

Definition 2.1.1. A **Topological Manifold** M is a topological space[1] with a family of open sets and functions $\{(M_\alpha, \phi_\alpha)\}$ such that,

- $M = \cup_\alpha M_\alpha$
- $\forall \alpha,\ \phi_\alpha : M_\alpha \to U_\alpha$ is a homeomorphism onto an open subset $U_\alpha \subset \mathbb{R}^n$.

[1] For completeness we actually need to require also that M is a second countable Hausdorff topological space



(a) S^1 (b) S^2 (c) T^2

Figure 1: Some examples of manifolds.

Each pair (M_α, ϕ_α) is known as a *chart*, and the family $\{(M_\alpha, \phi_\alpha)\}$ is called an *atlas*. Unsurprisingly, the *dimension* of the manifold is n.

Example 2.1.2. $\mathbb{R}^n$ and every open subset of $\mathbb{R}^n$ are n-manifolds.

Example 2.1.3. The set $S^n = \left\{ x \in \mathbb{R}^{n+1} \;\middle|\; ||x|| = 1 \right\}$ is an important example of an n-manifold, known as the n-sphere. See Figure 1.

Example 2.1.4. Another class of manifolds are the tori $\mathbb{T}^n$. $\mathbb{T}^n$ if $\underbrace{S^1 \times \cdots \times S^1}_{n}$ with the product topology, but you can also construct $\mathbb{T}^n$ by taking the unit cube in $\mathbb{R}^n$ and identifying opposite faces. Figure 1(c) shows the result of gluing opposite edges of the unit square.

Definition 2.1.5. Let M be an n-manifold. A *p-dimensional submanifold* of M is a closed subset L of M where there exists an atlas $\{(M_\alpha, \phi_\alpha)\}$ of M such that for all $x \in L$ there is a chart in the atlas with $x \in M_\alpha$ and

$$\phi_\alpha(L \cap M_\alpha) = \{0\} \times \mathbb{R}^p \subset \mathbb{R}^n$$

Remark 2.1.6. A submanifold is itself a manifold. Note that the atlas described in the definition induces an atlas for the submanifold. However not every subset of M that is a manifold is a submanifold.[2]

Example 2.1.7. The equator of the 2-sphere in Figure 1(b) is a one dimensional submanifold.

Definition 2.1.8. Let L and M be manifolds. A continuous function $f : L \to M$ is an *embedding* if it is a homeomorphism onto its image $f(L)$ and the image $f(L)$ is a submanifold of M.

Example 2.1.9. A knot is an embedding of S^1 into $\mathbb{R}^3$ (see Figure 2).

[2]For an example I refer the reader to the *Alexander Horned Sphere*, which is an embedding of S^2 in $\mathbb{R}^3$ that is not a submanifold

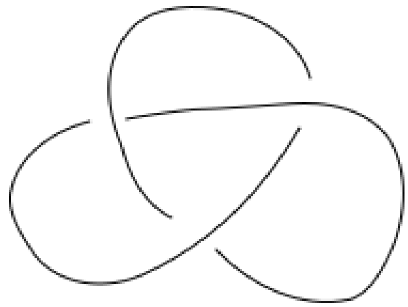
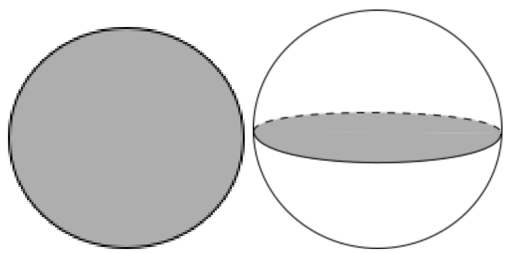


Figure 2: An embedding of S^1 into $\mathbb{R}^3$.



(a) B^2 (b) B^3

Figure 3: The 2 and 3 dimensional balls.

Definition 2.1.10. An *n-manifold with boundary* is again a topological space M with atlas $\{(M_\alpha, \phi_\alpha)\}$ such that for each chart (M_α, ϕ_α), ϕ_α is a homeomorphism onto an open subset of $\mathbb{R}^n$ or $\mathbb{H}^n = \{\ x \in \mathbb{R}^n \mid x_1 \geq 0\ \}$.

The boundary of M, denoted by ∂M, consists of the points of M with a neighborhood homeomorphic to an open subset of $\mathbb{H}^n$ but no neighborhood homeomorphic to an open subset of $\mathbb{R}^n$. Similarly the interior of M, denoted by M^o, consists of the points with a neighborhood homeomorphic to a ball in $\mathbb{R}^n$. Notice that if M is an n-manifold with boundary, then ∂M is an $(n-1)$-manifold without boundary.

Example 2.1.11. The closed n-ball $\mathbb{B}^n = \{\ x \in \mathbb{R}^n \mid ||x|| \leq 1\ \}$ is an n-manifold with boundary. Notice that $\partial \mathbb{B}^n = S^{n-1}$; see Figure 3.

Definition 2.1.12. Let M be an n-manifold with boundary. A p-dimensional submanifold of M is a closed subset L of M where there exists an atlas $\{(M_\alpha, \phi_\alpha)\}$ of M such that the following hold:

1. $\forall x \in L$ in the interior of M, there is a chart (M_α, ϕ_α) such that $x \in M_\alpha$ and

$$\phi_\alpha(L \cap M_\alpha) = \{0\} \times \mathbb{R}^p \subset \mathbb{R}^n.$$

2. $\forall x \in L$ in the boundary of M, there is a chart (M_α, ϕ_α) such that $x \in M_\alpha$ and

$$\phi_\alpha(L \cap M_\alpha) = \mathbb{H}^p \subset \mathbb{H}^n \text{ and } \phi_\alpha(x) \in \{0\} \times \partial\mathbb{H}^p \subset \partial\mathbb{H}^n.$$

Example 2.1.13. A chord of a disc is a submanifold with boundary.
Remark 2.1.14. While the boundary ∂M is a subset of M that is a manifold, it is *not* a submanifold of M.
Definition 2.1.15. An n-manifold M is *closed* if M is compact and $\partial M = \emptyset$.
Example 2.1.16. Both S^n and $\mathbb{T}^n$ are closed manifolds.

2.2 Triangulated Manifolds

Triangulations are the primary method to make algorithmic investigations of 3-manifolds possible. The most important aspect is that they allow us to discretize the underlying space of the manifold. In order to see this we must first discuss the building blocks of the triangulations.
Definition 2.2.1. Let v_i denote the standard basis vectors of $\mathbb{R}^{k+1}$. The *standard (closed) k simplex*, denoted by $[v_0, \ldots, v_k]$, or $[s]$ when the dimension of the simplex is clear, is the set

$$\left\{ a_0v_0 + \cdots + a_kv_k \;\middle|\; a_i \geq 0 \text{ and } \sum_{i=0}^{k} a_i = 1 \right\}$$

In other words $[s]$ is the convex hull of the $k+1$ standard basis vectors of $\mathbb{R}^{k+1}$. Similarly the *standard open k simplex*, denoted by $(v_0, \ldots, v_k)$ or (s), is the set

$$\left\{ a_0v_0 + \cdots + a_kv_k \;\middle|\; a_i > 0 \text{ and } \sum_{i=0}^{k} a_i = 1 \right\}$$

The k-simplices form the basic building blocks of triangulations. The following definition shows how to connect the standard simplex to a general simplex in a topological space X.
Definition 2.2.2. A k-simplex in a topological space X is a continuous function $f : [s] \to X$ such that the restriction $f|_{(s)}$ is a homeomorphism onto its image.

Strictly speaking the k-simplex is the entire function $f : [s] \to X$, but it is usually more convenient to think of the image of f as a k-simplex. Figure 4 shows a 3-simplex.
Definition 2.2.3. A *face* of the standard k-simplex $[s]$ is a subset of the form

$$\left\{ a_0v_0 + \cdots + a_kv_k \;\middle|\; a_{i_1} = \cdots = a_{i_j} = 0 \right\} \text{ for } j = 0, \ldots, k.$$

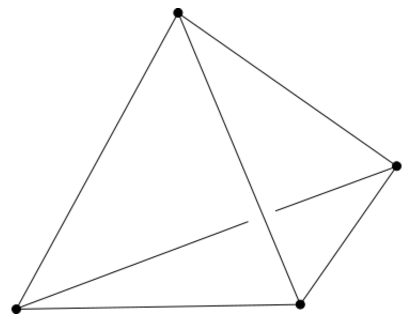


Figure 4: A tetrahedron. Each triangular face is a 2-simplex, each edge a 1-simplex, and each vertex a 0-simplex.

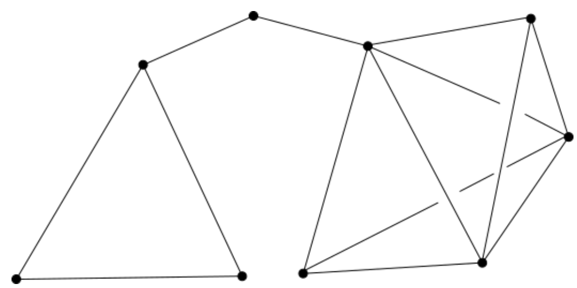


Figure 5: A 3-dimensional simplicial complex.

A face of a general k-simplex f is then the restriction $f|_{[t]}$ where $[t]$ is a face of the standard k-simplex. The dimension of a face is just $k - j$.

Examples 2.2.4.

1. The only face of the 0-simplex $[v_0]$ is itself.
2. The 1-simplex $[v_0, v_1]$ has a single 1 dimensional face and two 0 dimensional faces.
3. The 2-simplex $[v_0, v_1, v_2]$ has a single 2 dimensional face, three 1 dimensional faces, and three 0 dimensional faces.
4. The 3-simplex $[v_0, v_1, v_2, v_3]$ has a single 3 dimensional face, four 2 dimensional faces, six 1 dimensional faces, and four 0 dimensional faces.

Definition 2.2.5. A *simplicial complex* based on the space X is a collection K of simplices in X such that

- For each $f \in K$, every face of f is in K.
- For each pair of simplices $f_1, f_2 \in K$, if the $\mathrm{Im}(f_1|_{(s_1)})$ and $\mathrm{Im}(f_2|_{(s_2)})$ have non-empty intersection, then $\mathrm{Im}(f_1|_{(s_1)}) = \mathrm{Im}(f_2|_{(s_2)})$.

The union of the images of the simplices in K is denoted by $|K|$ and is called the *underlying space* of K. Given an integer i we define K^i to be the collection of i-simplices in K.

An important note is that we do not require the closed simplices of the simplicial complex to be embeddings. Also two distinct simplices can intersect at more than one face.

We finally have the tools necessary to define a triangulated manifold.
Definition 2.2.6. A *triangulated n-manifold* is a pair (M, K), where M is a topological manifold of dimesnion n and K a simplicial complex such that

- $|K| = M$
- For every compact subset $C \subset M$, the set $\{ f \in K \mid \operatorname{Im}(f) \cap C \neq \emptyset \}$ is finite.

We call K a *triangulation* of M. Triangulations are a powerful tool for studying manifolds algorithmically because of the second condition in the definition. This is known as being *locally finite*. In particular we are often interested in studying compact 3-manifolds, so the locally finite condition guarantees that any triangulation we work with must be finite.

While the above definition is sufficient to examine triangulations of manifolds, it is still relatively cumbersome to work with. However the following proposition will provide a constructive way to view triangulations that will be easier to think about.

Consider a triangulated k-manifold (M, K). Let C be the disjoint union of copies of standard k-simplices, one for each k-simplex in K. We now create a quotient space M' from C via the following quotient: Identify two points in C if an only if the corresponding points $x \in [s], y \in [t]$ for $f : [s] \to M$ and $g : [t] \to M$ in K satisfy $f(x) = g(y)$.
Proposition 2.2.7. *Endow M' with the quotient topology. Then M and M' are homeomorphic.*

The above proposition allows us to think of the triangulated k-manifold M either as the pair (M, K) for a simplicial complex K based on M, or as being built out of k-simplices by identifying faces.

Finally, the following theorem will allows us to restrict our attention to triangulations of the manifolds for the remainder of the survey.
Theorem 2.2.8. *Every compact (1, 2, or 3)-manifold admits a triangulation.*

The 1 dimensional case is relatively straightforward, while the 2 and 3 dimensional cases are more substantial. In [7, Chapter 1.4], Schultens cites Tibor Radó and Béla Kerékjártó for proving the 2 dimensional case, and R. H. Bing and Edwin Moise for proving the 3 dimensional case.

The obvious question to ask is whether higher dimensional manifolds can be triangularized. An important result from algebraic topology shows that for all $n \geq 4$, there exist

(a) Annulus

(b) Möbius Strip

Figure 6: Orientable and non-orientable 2-manifolds.

n-manifolds that do not admit any triangulations.

2.3 Orientability

Manifolds can be split into two categories, orientable and non-orientable. The canonical examples of these types are the annulus and the Möbius strip (Figure 6). There are multiple ways to define orientation of a manifold, but we will limit ourselves to defining orientation for triangulated manifolds. In general this is not sufficient to define a concept of orientation to all manifolds, but 2.2.8 assures us that we can define a concept of orientation for all (1, 2, or 3)-manifolds.

We begin by orienting simplices:

Definition 2.3.1. An *oriented k-simplex* is a k-simplex together with an ordering of the vertices $(v_0, v_1, \dots, v_k)$. We say that two orientations are the same if the ordering of the vertices differs by an even permutation. In particular this means that there are exactly two orientations for a k-simplex.

For 1 and 2 simplices, the orientation can be easily interpreted. An orientation of a 1-simplex is either moving to the left or right along the edges, while an orientation for a 2-simplex is moving clockwise or counterclockwise around the edges. See Figure 7.

Remark 2.3.2. Notice that orienting a k-simplex induces orientations for each face by taking the ordering of the vertices of the face from the top level orientation.

Definition 2.3.3. A manifold M is *orientable* if given a triangulation K of M, we can assign a consistent orientation to each simplex in the triangulation. Otherwise M is non-orientable. By consistent orientation we mean orienting each k-simplex of the triangulation K so that, given any two faces that are glued together, the induced orientations are opposite each other (Figure 8).

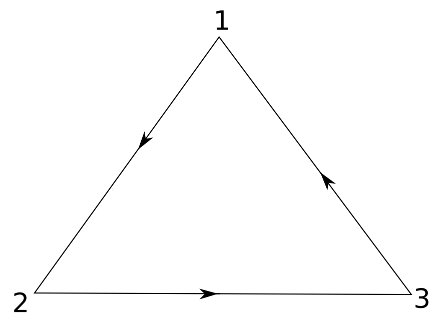


Figure 7: An oriented 2-simplex.

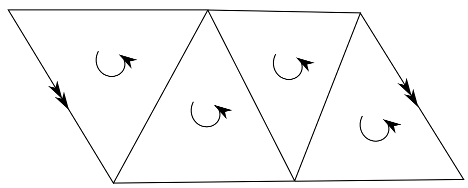
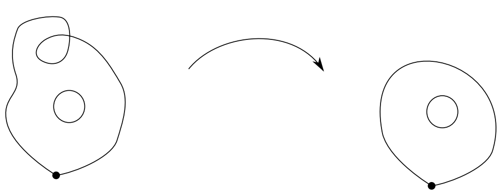


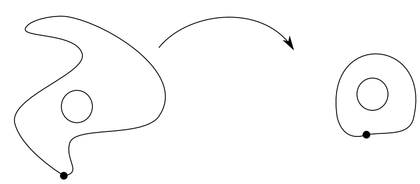
Figure 8: An oriented Annulus.

2.4 Homotopies

Definition 2.4.1. Let M be an n-manifold and S a k-manifold with $k < n$. Suppose f and g are continuous functions from S to M. f and g are *homotopic* if there exists a continuous function $H : [0,1] \times S \to M$ such that $H(0,x) = f(s)$ and $H(1,x) = g(x)$.
Definition 2.4.2. Suppose further that f and g are embeddings of S into M. Then f and g are *isotopic* if for every $i \in [0,1]$ the restriction of H to $\{i\} \times S$ is also an embedding.



(a) Homotopy of a loop



(b) Isotopy of a loop

Figure 9: Examples of a homotopy and isotopy.

2.5 Connected Sum

Suppose we have two disjoint n-manifolds M and N. An obvious question to ask is whether we can somehow combine these two manifolds into a new manifold L. This brings us to the concept of connected sum.

Definition 2.5.1. Let B_M, B_N be n-balls in M and N respectively. Delete the interior of B_M and B_N. Pick a homeomorphism f from B_M to B_N and let L be the result of gluing B_M and B_N via f. We call L the connected sum of M and N, and denote the operation by $M\#N$.

There is a point of concern by calling L *the* connected sum when we consider the fact that there were many possible choices of balls B_M and B_N. However the following theorem should put the reader at ease:

Theorem 2.5.2 ([7, Theorem 1.6.4])**.** *If B_1 and B_2 are n-balls in the interior of a connected n-manifold M, then there is an isotopy $f : [0,1] \times M \to M$ such that $f(0,)|_{B_1}$ is the identity and $f(1,)|_{B_1}$ is a homeomorphism onto B_2.*

Definition 2.5.3. An n-manifold M is *prime* if $M = M_1\#M_2$ implies that either M_1 or M_2 is the n-sphere.

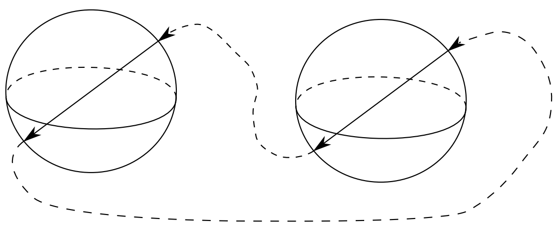
3 3-Manifolds

We now restrict our attention to 3-manifolds for the rest of this survey. We begin with some examples, after which we discuss embedded surfaces in 3-manifolds which is a powerful way to study the structure of 3-manifolds.

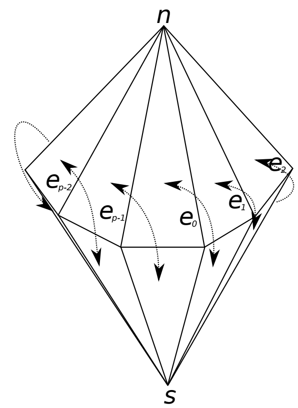
3.1 Examples

Example 3.1.1 (The 3-Sphere)**.** As mentioned in 2.1.3, as a set the 3-sphere S^3 consists of the unit vectors of $\mathbb{R}^4$. However for our purposes this is not the most useful description to understand the structure of S^3. Another way to view S^3 is by gluing two copies of B^3 together along their 2-sphere boundaries via an orientation reversing homeomorphism. As an analogy to the 2-sphere, we can think of the centers of the two balls as the "north" and "south" poles of S^3. Then the straightline from either pole will form a closed loop that passes through both poles, see Figure 10(a).

Example 3.1.2 (Lens Spaces)**.** Lens Spaces form a class of examples of 3-manifold $L_{p,q}$ parameterized by two integers. Consider a regular planar p sided polygon P, together with two points n and s above and below. We connect each vertex of P to n and s forming a bipyramid that we then fill in. Now label each edge of P, $e_0, \dots e_{p-1}$, and label the corresponding triangular faces above and below as n_i and s_i. Form a quotient space



(a) S^3



(b) The lens space $L_{p,2}$

Figure 10: Examples of 3 Manifolds

by identifying n to s and the triangular faces n_i with $s_{(i+q) \mod p}$. This forms a closed 3-manifold. See Figure 10(b).

3.2 Embedded Surfaces

An important method used to study 3-manifolds is to look at embedded surfaces inside the 3-manifold. This is analogous to studying surfaces by examining the simple closed curves. Just as some simple closed curves in surfaces are more interesting than others, some embedded surfaces in a 3-manifold are more interesting than others. The important surfaces we consider here are analogous to essential simple closed curves in surfaces.

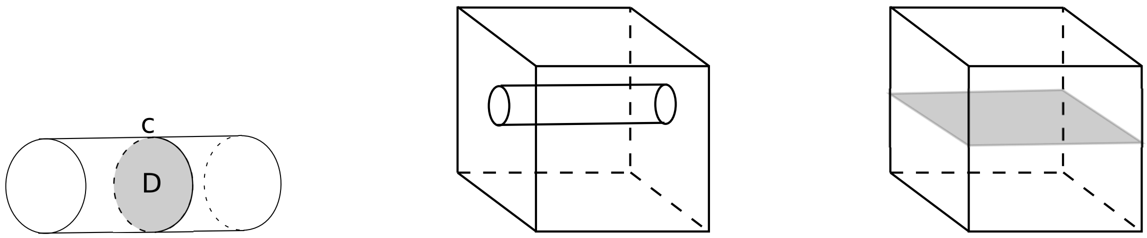
Definition 3.2.1. A submanifold S in a compact manifold M is *proper* if $\partial S = S \cap \partial M$.

Example 3.2.2. The equitorial disc in Figure 3 is a proper surface.

Definition 3.2.3. Let M be a 3-manifold. A surface $S \subset M$ is *compressible* (Figure 11(b)) if either

1. S is a 2-sphere that bounds a 3-ball in M.
2. There exists a simple closed curve c in S that bounds a disc D with interior in $M \backslash S$ but does not bound any disc with interior a component of $S \backslash c$. The disc D is known as a *compressing disc* for S; see Figure 11(a).

A proper surface S is *incompressible* if it is not compressible, see Figure 11(c).



(a) A compressing disc D for a surface.

(b) A compressible torus in T^3.

(c) An incompressible torus surface in T^3.

Figure 11:

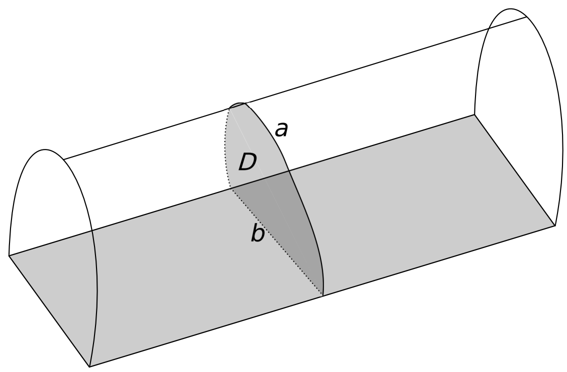


Figure 12: A boundary compressing disc D.

Definition 3.2.4. Let M be a 3-manifold with boundary. A surface $S \subset M$ is *boundary compressible*, or *∂-compressible*, if there is an essential simple arc α in S and an essential simple arc β in ∂M such that $\alpha \cup \beta$ is the boundary of a disc D in M with interior disjoint from S (Figure 12) A surface S is *boundary incompressible*, or *∂-incompressible*, if it is not compressible.

Notice that for closed 3-manifolds, incompressible surfaces always contain relevant information about the 3-manifold. However this need not be the case for manifolds with boundary. Suppose we have a connected 3-manifold $M \neq \mathbb{B}^3$ such that $\partial M \neq \emptyset$. We can construct an incompressible surface $F \subset M$ as follows. First make a copy of one of the boundary components C of M. Then push this copy into the manifold, creating a proper surface F homeomorphic to C. If we cut M along F one of the components will be homeomorphic to $F \times [0,1]$ and the other to M. Therefore F did not capture any of the relevant information about M. This leads us to the following definition.

Definition 3.2.5. A surface $F \subset M$ is *boundary parallel* if it is separating and a component

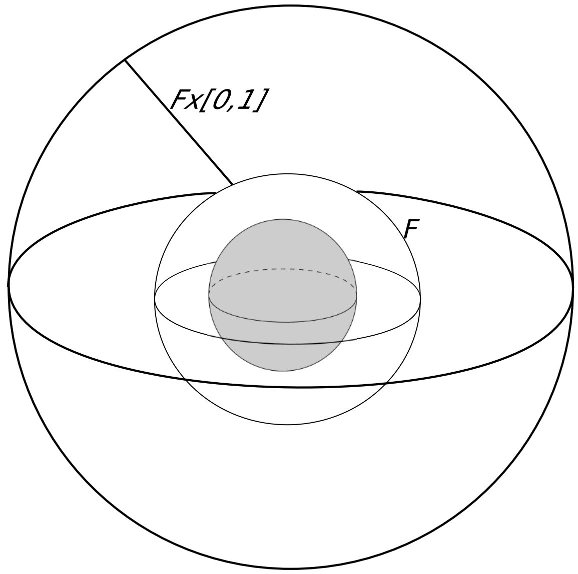


Figure 13: A surface that is boundary parallel.

of $M \backslash F$ is homeomorphic to $F \times [0,1]$; see Figure 13.

We are now ready to present the surfaces of a 3-manifold that are analogous to essential simple curves in surfaces.

Definition 3.2.6. Let M be a connected 3-manifold. A surface F is *essential* if either it is a 2-sphere that does not bound a 3-ball or it is incompressible, ∂-incompressible, and not boundary parallel.

Example 3.2.7. The incompressible surface in Figure 11 is essential.

Essential surfaces play an important role in decomposing 3-manifolds into component parts. Remember from subsection 2.5 we defined prime manifolds as having only trivial connected sum decompositions. Essential spheres play a key role in identifying prime manifolds, which the following proposition develops.

Proposition 3.2.8. *Suppose M is a 3-manifold. Then M is prime if and only if M does not contain any separating essential 2-spheres.*

Proof. Suppose M is prime. Let S be a separating 2-sphere. Then S decomposes M into two components M_1 and M_2. We can then set $M = M_1' \# M_2'$ where M_i' is constructed from M_i by gluing on a 3-ball along the 2-sphere boundary component from the cutting operation. Since M is prime one of M_i' must be S^3, and therefore M_i is a 3-ball. Therefore S is non-essential.

Now suppose that M does not contain any separating essential 2-spheres. Suppose

$M = M_1 \# M_2$. Then M contains a separating 2-sphere S, namely the one used to glue $M_1 - B^3$ and $M_2 - B^3$. Since M does not contain separating essential 2-spheres, S must bound a 3-ball. Therefore one of $M_i - B^3$ must be a 3-ball, and so M_i is S^3. Therefore M is prime. ■

Notice that we had to make the requirement that we only consider separating 2-spheres. What if we want to drop the separating condition? If we do that we get the following definition.

Definition 3.2.9. A 3-manifold M is *irreducible* if every 2-sphere in M is non-essential, i.e. it bounds a 3-ball. M is *reducible* if it contains an essential 2-sphere.

Notice that if M is irreducible then in particular it does not contain separating essential 2-spheres, and so by Proposition 3.2.8 M is prime. So irreducibility implies primeness. Unfortunately the reverse is not true. However all is not lost. It turns out there are exactly two reducible prime manifolds, $S^2 \times S^1$ and $S^2 \tilde{\times} S^1$. $S^2 \tilde{\times} S^1$ is constructed as follows: start with $S^2 \times [0,1]$ and then identify $S^2 \times \{0\}$ with $S^2 \times \{1\}$ via the antipodal map. $S^2 \tilde{\times} S^1$ can be considered a non-orientable version of $S^2 \times S^1$.

Proposition 3.2.10. *$S^2 \times S^1$ and $S^2 \tilde{\times} S^1$ are reducible.*

For either case we can take a copy of $S^2 \times \{i\}$ as our proper 2-sphere. After cutting along this 2-sphere we get $S^2 \times [0,1]$ and so in particular $S^2 \times \{i\}$ is not separating, and so must be essential.

Theorem 3.2.11 ([7, Theorem 3.3.4])**.** *An irreducible closed connected 3-manifold is prime. A closed connected prime 3-manifold is irreducible, $S^2 \times S^1$, or $S^2 \tilde{\times} S^1$.*

4 Normal Surfaces

As explained in [6, Chapter 3], Wolfgang Haken was one of the first topologists to recognize that the best way to explore 3-manifolds is to examine embedded surfaces within the 3-manifold. However an uncountable set of surfaces does not lend itself well to algorithmic methods. This is where normal surfaces become relevant. Basically normal surfaces are embedded surfaces with a restricted structure. They form a class of surfaces large enough to capture important information about the 3-manifold, while having an algorithmically enumerable finite basis.

4.1 Haken's Scheme

While we will focus our attention exclusively on triangulations of 3-manifolds to develop our algorithms, Haken's scheme for normal surfaces need not be limited to triangulations. In fact the general scheme provides a way to develop normal surfaces for any decomposition of a 3-manifold M we choose. The basic steps are as follows [6, Chapter 3.1]:

1. Find a way to decompose M into a collection of objects that have a finite number of different types. Denote the decomposition as ξ.

2. A proper surface F will be decomposed by ξ into a set of elementary pieces, i.e. the connected components of F intersected with each element of ξ. We then define a set of allowable elementary pieces. A *normal surface* is a surface that is decomposed by ξ into allowable pieces. We specify normal surfaces up to isotopies that are normal at each step, which we call a *normal isotopy*.[3] Of course we also want to make sure that the collection of normal surfaces has surfaces that actually contain information about M.

3. Identify each normal surface with an element in $\mathbb{Z}^k$ based on the count of each elementary piece in each element of ξ.[4]

4. Determine the conditions on a vector in $\mathbb{Z}^k$ that allow it to be identified with a normal surface. These conditions determine what is known as a matching system of linear equations and subset of admissible solutions.

5. Form a system of integer linear equations E based on the above conditions. Determine a finite set of fundamental solutions for E. These fundamental solutions generate the *fundamental normal surfaces*. All that remains is to determine a geometric interpretation of vector addition, after which each normal surface can be constructed as a linear combination of fundamental normal surfaces.[5]

In our case the decomposition of M will always be a triangulation, where the basic type is only the 3-simplex. Together with theorem 2.2.8, we have our decomposition of M. In our case the first part of Step 2 is straightforward, while making sure we still have informative normal surfaces becomes the hard part. Step 3 is basically automatic. Step 4 is just a careful examination of how to build a normal surface out of elementary pieces. Finally the important aspect of Step 5 is the geometric interpretation of adding two normal surfaces together.

[3] We must take care when specifying the allowable elementary pieces. We require that if a collection of allowable pieces in an element of ξ can be constructed from disjoint pieces, this construction is unique up to normal isotopy.

[4] We must also ensure that two normal surfaces with the same integer vector are normally isotopic.

[5] Of course the coefficients in the linear combination will be non-negative integers.

(a) A normal curve

(b) A non-normal curve

Figure 14: Examples of curves in a tetrahedron.

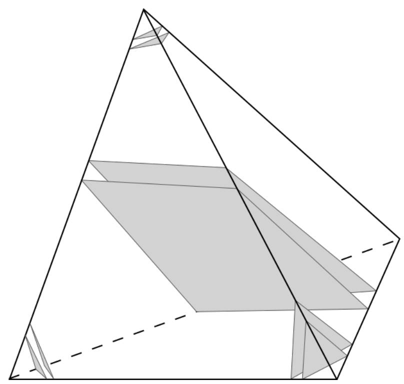


Figure 15: Normal discs in a tetrahedron.

4.2 Normal Surfaces in Triangulations

In order to define our normal surfaces we need only describe the allowable elementary pieces of a surface in a tetrahedron.

Definition 4.2.1. A simple arc on a 2 dimensional face $[f]$ of a tetrahedron $[s]$ is a *normal arc* if its endpoints lie on distinct edges of $[f]$. A simple closed curve c on the boundary of $[s]$ is a *normal curve* if $c \cap [f]$ consists of normal arcs for every 2 dimensional face of $[s]$, see Figure 14.

Definition 4.2.2. The *length* of a normal arc c, $l(c)$, is the number of intersections with c and the edges of $[s]$.

We can now define the elementary pieces, which fall into two types, *normal triangles* and *normal quadrilaterals*. Each elementary pieces is a disc in $[s]$ whose boundary is a normal curve of length 3 or 4, respectively. For a tetrahedron, there are 4 normal triangles, one around each vertex, and 3 normal quadrilaterals, see Figure 15.

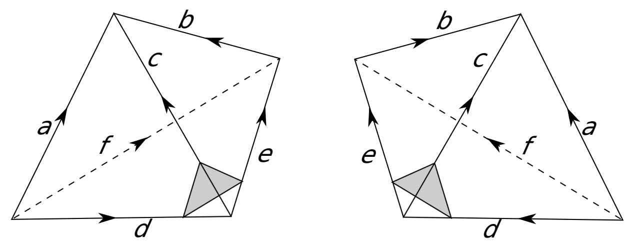


Figure 16: A normal sphere in S^3.

Now let (M, K) be a triangulated 3-manifold.

Definition 4.2.3. A proper surface F in M is a *normal surface* relative to the triangulation K, if $F \cap [s]$ is a collection of disjoint normal triangles and quadrilaterals for every tetrahedron in K, see Figure 16.

The following definitions are helpful to prove things about normal surfaces.

Definitions 4.2.4. The *weight*, $w(F)$, of a normal surface F in (M, K) is the number of components of $F \cap |K^1|$. The *measure* of F, $m(F)$, is the number of components of $F \cap (|K^2| \backslash |K^1|)$.

Finally the following theorem from [7, Chapter 5.2] tells us that the normal surfaces with respect to triangulations form a rich and informative class of surfaces.

Theorem 4.2.5 ([7, Theorem 5.2.14])**.** *Let M be a closed irreducible 3-manifold containing an incompressible surface F. For any triangulation K of M, there exists an isotopy that takes F to a normal surface in K.*

There is also an analogous result where we only require M be compact and irreducible, but we must also require that S is incompressible and ∂-incompressible.

Proof. Let (M, K) be a triangulation of M and S an incompressible surface in M. Isotope S to that $(w(S), m(S))$ is minimal (ordered lexicographically). Let Δ^3 be a tetrahedron of K and let f be a face of Δ^3.

Suppose $S \cap f$ contains simple closed curves. Let c be an innermost disc[6] in f and let $D \subset f$ be the disc that c bounds. Since S is incompressible, c also bounds a disc $D' \subset S$. Therefore $D \cup D'$ forms a 2-sphere and so must bound a 3-ball by irreducibility of M. This 3-ball defines an isotopy that pushes S through f (see Figure 17), and thus reducing $(w(S), m(S))$ by $(0, 1)$. So by minimality of $(w(S), m(S))$, $S \cap f$ does not contain any closed components.

[6] An innermost disc is one that bounds a disc that does not contain any other simple closed curves.

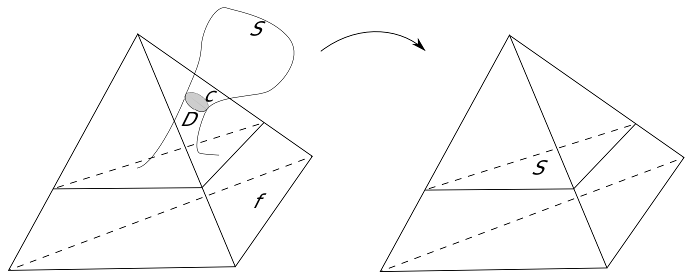


Figure 17: Removing simple arcs from $S \cap f$.

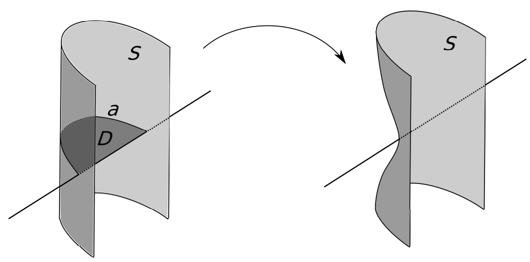


Figure 18: Removing arcs with endpoints on the same edge.

Next suppose that $S \cap f$ contains arcs whose endpoints lie on the same edge. Let a be an outermost arc in f, and D be the disc that a forms with the component of the edge joining the endpoints of a. This disc defines an isotopy that removes a from the intersection (see Figure 18), reducing $(w(S), m(S))$ by $(2, 1)$. So by minimality of $S \cap f$ consists of normal arcs. Since this holds for all faces of Δ^3, $S \cap \partial\Delta^3$ consists of normal curves.

Let $\tilde{c}$ be a normal curve in $S \cap \partial\Delta^3$. First notice that $\tilde{c}$ bounds a disc E in Δ^3 and since S is incompressible $\tilde{c}$ must also bound a disc $\tilde{S}$ in S. Using a similar innermost disc argument as above we can assume that $E \cap \tilde{S} = \tilde{c}$, and therefore $E \cup \tilde{S}$ is a 2-sphere and so bounds a 3-ball. If $\tilde{S}$ is not contained in Δ^3, the 3-ball defines an isotopy that reduces $w(S)$ or $m(S)$, and so by minimality of $(w(S), m(S))$ we must have $\tilde{S} = S \cap \Delta^3$.

To complete the proof we must show that $\tilde{S}$ is a normal disc in Δ^3, i.e. $\tilde{c} = \partial\tilde{S}$ has length 3 or 4. It turns out that showing $\tilde{c}$ has length 3 or 4 is equivalent to showing that $\tilde{c}$ does not meet any edges e of Δ^3 more than once. Suppose that $\tilde{c}$ intersects an edge e of Δ^3 more than once. First notice that $\tilde{c}$ partitions $\partial\Delta^3$ into two discs D, D'. One of these discs, say D, meets the interior of e along a subarc α that connects two adjacent intersection points of $\tilde{c} \cap e$. Furthermore, $\tilde{S}$ is isotopic to D. Therefore there must be another arc β in the interior of $\tilde{S}$ and a disc E with $\partial E = \alpha \cup \beta$ and interior of E in Δ^3. This disc describes an isotopy that moves β outside of Δ^3 and thus reduces $w(S)$ by 2. Therefore by minimality of $(w(S), m(S))$, $\partial\tilde{S}$ does not meet an edge of Δ^3 more than once, and so $\tilde{S}$ must be normal.

Applying the above argument to all tetrahedron in K shows that after the isotopy S is normal. ■

4.3 Matching System and Admissible Solutions

Steps 3 and 4 of Haken's Scheme are perhaps the most important aspects of the scheme that allows us to algorithmically explore manifolds. In particularly they provide us with a method to express every possible normal surface in a triangulation from a finite basis of vectors in $\mathbb{R}^n$.

4.3.1 Correspondence With $\mathbb{R}^n$

We first aim to develop the correspondence between normal surfaces in a triangulated manifold (M, K) to integer vectors in $\mathbb{R}^n$. It is easy to assign a vector to every normal surface S; we simply count the number of normal quadrilaterals and triangles the surface creates in each tetrahedron of the triangulation. More specifically, suppose $\{T_1, \ldots, T_k\}$ are the tetrahedrons of K. For each T_i we have 4 possible types of normal triangles and 3 possible types of normal quadrilaterals. We now let $(p_{i1}, p_{i2}, p_{i3}, p_{i4}, q_{i1}, q_{i2}, q_{i3}) \in \mathbb{R}^7$ be the count of each type of normal triangle and quadrilateral formed by the intersection of S with T_i. Now by stacking the k vectors together, we get a vector $v(S) \in \mathbb{R}^{7k}$ that represents the surface S.

Given a vector $v \in \mathbb{R}^{7k}$, how do we build a surface S in M with $v(S) = v$? The basic method is to simply put p_i normal discs for each type in v, however for an arbitrary v this will just result in a disjoint collection of discs and not a closed surface. So let's look at the conditions v must satisfy for the above method to result in a valid surface. We must consider both the local conditions on a single tetrahedron and the global conditions for the entire triangulation.

First we restrict our attention to a single tetrahedron in K. In this case we need to make sure that the surface will not have any self intersections. It's not hard to see that by placing the normal triangles arbitrarily close to their respective vertices, we can place any number of triangles in a single tetrahedron without having self intersections. But what about normal quadrilaterals? It turns out that we can have exactly one type of normal quadrilateral, but we can have arbitrarily many of that type. So for a vector $(p_{i1}, p_{i2}, p_{i3}, p_{i4}, q_{i1}, q_{i2}, q_{i3}) \in \mathbb{R}^7$ to be valid for a tetrahedron, we must have that exactly one of the q_{ij} be non-zero.

We call the above condition the square condition. So for a vector $v \in \mathbb{R}^{7k}$ to be valid it must satisfy the square condition for each tetrahedron. We call such a v *admissible*.

Now we look at the global conditions. We need to guarantee that after we've placed the normal discs in each tetrahedron, we can glue the faces together so that the edges of the disc end up being glued together. First consider a single face f of a tetrahedron T_i. For each pair of edges in f, there is exactly one normal triangle and one normal quadrilateral whose boundary in f will connect the two edges. So when we glue two faces together, we need to make sure that the number of arcs connecting each pair of edges is the same in each face. Consider two tetrahedron T_i, T_j that share a face. For each pair of edges T_i will have $p_{ik} + q_{il}$ arcs and T_j will have $p_{jk} + q_{jl}$ arcs. So we need to have that,

$$p_{ik} + q_{il} = p_{jk} + q_{jl}$$

Now each tetrahedron has 4 faces and each face has 3 possible arcs, so assuming each face is glued to another face[7], the above procedure gives us $6k$ equations that v must satisfy to form a valid surface. We call this system of linear equations the *matching system* for K.

4.3.2 Integer Linear Equations

From the correspondence, we can now express each normal surface uniquely as vector in $\mathbb{R}^{7k}$ and we have a system of equations E whose admissible solutions will represent every normal surface in the manifold. So to explore the normal surfaces in a triangulated manifold (M, K), we only need to know how to find admissible solutions to the matching system. Therefore we need to develop some tools for solving a system of integer linear equations.

Suppose we have a system of integer linear homogenous equations E,

$$a_{i1}x_1 + \cdots + a_{in}x_n = 0, 1 \leq i \leq m.$$

Definition 4.3.1. A non-negative integral solution x to E is *fundamental* if it cannot be represented as the sum of two non-trivial non-negative integral solutions y, z.

Theorem 4.3.2 ([6, Theorem 3.2.8])**.** *The set of fundamental solutions to E is finite and can be constructed algorithmically. Also any solution to E can be expressed as a linear combination of fundamental solutions with integer coefficients.*

[7] This happens when M is a closed manifold. When M has boundary, there will be faces not glued to other faces. This will only change the number of equations we have but will not change overall structure.

Proof. Let σ^{n-1} be the standard simplex in $\mathbb{R}^n$. Let S be the set of solutions to E over $\mathbb{R}^n$, L the support plane for σ^{n-1}, and $P = L \cap S$.

Remarks 4.3.3.

1. P is the intersection of L with the hyperplanes defined by the equations in E and the half spaces defined by $x_i \geq 0$.
2. P is contained in σ^{n-1} and so is necessarily bounded.

Therefore P must be a convex polyhedron of dimension $m < n$. S can be thought of as the union of straight rays from the origin passing through points in P. Notice that the vertices of P must be rational. If we multiply each vertex $\bar{v}$ by the smallest integer $k > 0$ such that the coordinates of $k\bar{v}$ are integers, we get the set $\mathcal{V}$ of *vertex* solutions. The vertex solutions must be fundamental.

Since P is a convex polyhedron of dimension m, we can decompose P into m-simplices without introducing extra vertices. Therefore we can present S as the union of cones over m-simplices with vertices in $\mathcal{V}$. We must now show that each of these cones contains only finite number of fundamental solutions.

Let δ be an m-simplex with vertices $\bar{V}_0, \bar{V}_1, \dots, \bar{V}_m \in \mathcal{V}$, and $S_\delta \subset S$ be the cone over δ. Any point inside an m-simplex can be expressed as a non-negative linear combination of its vertices. We have that for any integer point $\bar{x} \in S_\delta = \sum_{i=0}^{m} \alpha_i \bar{V}_i$, where all $\alpha_i \geq 0$. Suppose one of $\alpha_i > 1$. Then $\bar{x} - \bar{V}_i$ is an integer point with non-negative coordinates, and so $\bar{x}$ can be presented as $\bar{x} = (\bar{x} - \bar{V}_i) + \bar{V}_i$. Therefore $\bar{x}$ is not fundamental.

Therefore we must have that all fundamental solutions fall in the compact set

$$U_\delta = \left\{ \sum_{i=0}^{m} \alpha_i \bar{V}_i \;\middle|\; 0 \leq \alpha_i \leq 1 \right\}.$$

Since all the solutions are integer vectors, the fundamental solutions fall in $\mathbb{Z}^n \cap U_\delta$, which must be finite.

The algorithm to find the fundamental solutions reduces to taking solutions $\bar{x}$ from $\mathbb{Z}^n \cap U_\delta$, and reducing them via known fundamental solutions until all fundamental solutions have been found.[8]

■

[8]Notice that this proves the *existence* of an algorithm, but does not provide a *practical* algorithm for finding fundamental solutions. We discuss this problem in subsection 6.1.

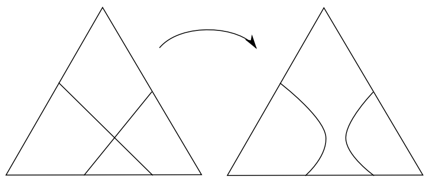
4.4 Haken Sum

At this point we have discussed what our normal surfaces are within the triangulation and how to construct a correspondence between normal surfaces and solutions to a homogeneous system of integer linear equations. Based on this correspondence we can construct a set of fundamental normal surfaces from the admissible fundamental solutions to the matching system. However we are still missing a key connection between the normal surfaces and the solutions to the matching system. Namely how we interpret the addition of two admissible solutions to the matching system, whose sum is also admissible.

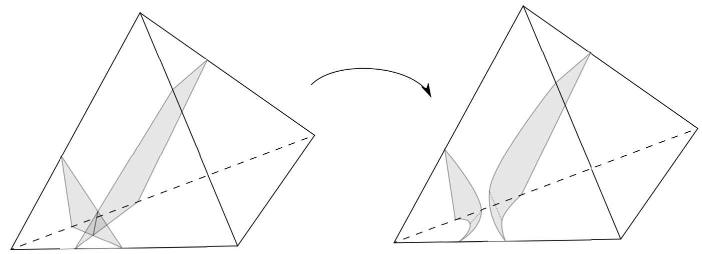
Suppose we have two normal surface S_1, S_2 in M such that $v(S_1) + v(S_2)$ is an admissible solution. In particular this means that for each tetrahedron Δ^3 of M, both S_1 and S_2 contribute exactly one type of quadrilateral disc. First we use a normal isotopy to put S_1 and S_2 into positions such that on the interior of any tetrahedron Δ^3 of the triangulation, S_1 and S_2 meet along double arcs that have endpoints on the interior of faces. In fact we can take this further and require that any two elementary discs of S_1 and S_2 intersect at no more that two arcs.

Now consider a double line c of S_1 and S_2. Decompose it into arc pieces according to the intersection of c with each tetrahedron that it meets. Consider one of there arcs l in a tetrahedron Δ^3. Note that l is the intersection of two normal discs in Δ^3. We can perform the following operation on the discs. First cut each disc along l and then glue adjacent pieces together. This is known as a *switch*. Obviously we have a choice as to which pair of adjacent pieces get glued back together. Since l has endpoints on the interior of faces, the endpoint is the intersecting point of the two normal arcs for the normal discs. Now we have two ways to split these arcs and glue them back together, one which creates two normal arcs of the same type, and another that produces two arcs that are no longer normal. We call the first choice a *regular switch* and the second an *irregular switch*, see Figure 19. Notice that a regular switch at one tetrahedron matches with the regular switch at neighboring tetrahedra. Therefore we can determine a global regular switch along c that is well defined.

Perform the regular switch for each double line c of S_1 and S_2. By the previous argument the resulting surface will be normal, and by construction we have that $v(S_1 + S_2) = v(S_1) + v(S_2)$.



(a) A regular switch of intersecting arcs.



(b) A regular switch of intersecting discs.

Figure 19: Regular switches.

5 Algorithms

Armed with our theory of normal surfaces, we can begin developing algorithms. All of our algorithms end up following the same basic structure.

5.1 General Scheme

Suppose we have some 3-manifold M and we want to determine if it has some property P. The following procedure gives us a method to generate an algorithm that will recognize if M has property P.

1. First translate P into a property that can be determined from M containing a proper surface F with a property P'.
2. Triangulate M and prove that if M contains a surface with property P', then there exists a normal surface with property P'. Often this is done by showing that the property P' is preserved by the normalization procedure.
3. Show that if there exists a normal surface with property P', then there exists a fundamental surface with property P'.
4. Finally generate an algorithm to determine if a surface has property P'.

Once the above steps are completed, the algorithm to detect property P in a 3-manifold M is as follows.

1. Choose a triangulation K of M.
2. Determine the matching system of linear equations.
3. Solve the matching system and determine the finite set of fundamental surfaces.

4. Check each fundamental surface for property P'.

6 Implementation Considerations & Regina

The strategy we have developed thus far for creating recognition algorithms for 3 manifolds works well enough in the abstract, but in general they are all well outside the computationally tractable bounds. In fact the complexity of all the algorithms we can create will grow exponentially in the number of triangles in our triangulation, and this is directly related to the complexity of manifolds that we can look at. It would be nice to somehow make these algorithms tractable, but to do that we must first determine the bottlenecks in these algorithms. After determining the bottlenecks, we can begin to find ways to work around them.

6.1 Computational Complexity and Bottlenecks

Considering the general set up of every 3-manifold algorithm from item 5.1, there are two places where the main bottleneck in the algorithm could occur. Notice that specifying a triangulation and generating the matching system are implicit in an implementation to represent the manifold of consideration and so do not contribute to the complexity of any algorithm. Therefore the bottleneck either occurs from enumerating the fundamental solutions to the matching system, or from the algorithm concerning surfaces. Most algorithms concerning surfaces are able to run in polynomial time, and so the main bottleneck comes down to enumerating the fundamental surfaces.

It turns out that for many algorithms we do not need to enumerate all normal surfaces, and instead we need only enumerate the so called *vertex* solutions to the matching system. These solutions correspond to the vertices of the polytope defined by the matching system. Since we must at least enumerate all of these solutions, there are two obvious ways to get at lower bounds for the complexity of the enumeration given a triangulation K:

- *Combinatorial Complexity* $\sigma(K)$, which measures the total number of vertex fundamental normal surfaces in K.
- *Algebraic Complexity* $\kappa(K)$, which measures the magnitude of the largest coordinate of the vertex fundamental normal surfaces in K.

The number $\sigma(K)$ provides an immediate lower bound on the complexity of the enumeration since it defines the output size of any enumeration algorithm. The number $\kappa(K)$

provides another important measure when considering actually implementing these algorithms. Indeed if $\kappa(K)$ is small enough[9] the values can be natively manipulated in constant time, whereas if $\kappa(K)$ is large we must use arbitrary precision integers which cannot be manipulated in constant time.

In [4], Burton et. al. produce a class of triangulations with $\sigma(K) = \Theta(2.317^n)$ where n is the number of tetrahedron. They also produce class of triangulations such that $\kappa(K)$ follows a fibonacci growth rate in the number of tetrahedrons. These results show that the worst case for a general triangulation is necessarily very bad.

6.2 Regina

Regina [3] is a software program that implements many of the 3-manifold algorithms that are based on normal surface theory. It manages to overcome the above time complexity issues by using heuristics to lower the practical run time of the algorithms it implements.

6.3 Simplifying Triangulations

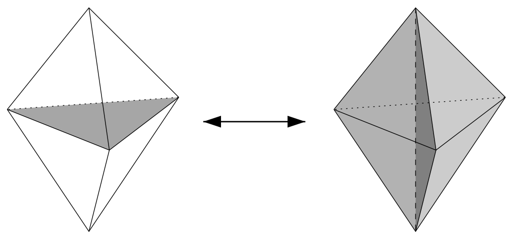
One of the most important algorithms when trying to implement 3-manifold algorithms is being able to simplify a given triangulation K to an equivalent triangulation K' with fewer tetrahedra. There are two major reasons for this:

- Since 3-manifold algorithms run exponentially in the number of tetrahedrons, being able to quickly simplify a triangulation provides a massive speed up in the total running time of any 3-manifold algorithm.
- Even better, by having a powerful simplification algorithm you can sometimes avoid having to run the expensive 3-manifold algorithm by simply comparing the simplified triangulation to a known triangulation that has the property you are testing for.

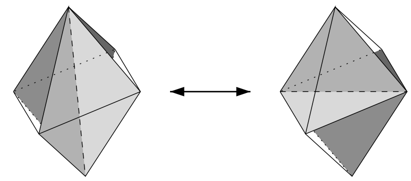
Regina implements a powerful simplification algorithm based on local simplification moves. These moves are split into 4 types:

1. *Pachner Type Moves*, which never change the underlying manifold.
2. *Moves around low degree edges and vertices*, which can potentially change the underlying manifold.

[9] Current CPU's have 64-bit registers and so can represent unsigned numbers in the range $[0, 2^{65} - 1]$.



(a) The 2-3 and 3-2 Pachner moves.



(b) A 4-4 Pachner move.

Figure 20: Examples of Pachner moves (taken from [1])

3. *Boundary Moves* which only apply to triangulations with boundary and can also potentially change the underlying manifold.
4. *Edge Collapse*, the most powerful of the move, but also the most complex to guarantee that it does not change the underlying manifold.

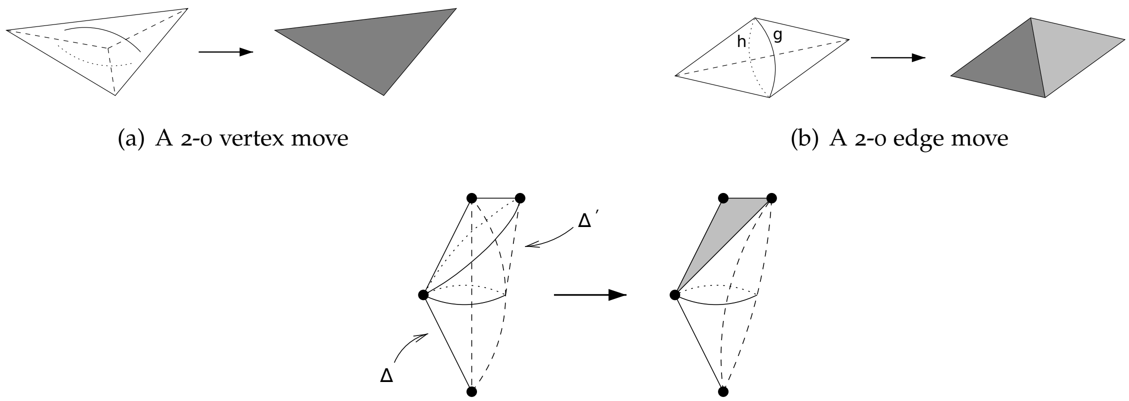
6.3.1 Pachner Type Moves

- Consider a triangular bipyramid, which can be triangulated either with (i) 2 tetrahedron that share an internal face, or (ii) 3 tetrahedron that share an internal edge. A *2-3 Pachner move* replaces (i) with (ii), and a *3-2 Pachner move* replaces (ii) with (i), see Figure 20(a).
- Consider an octahedron, which can be triangulated with four tetrahedron that share an internal edge. There are 3 ways of triangulating it in this way, where the shared edge is any of the main diagonals. A *4-4 move* replaces one of these triangulations with another, see Figure 20(b).

6.3.2 Moves around low degree edges and vertices

There are three moves that operate on vertices and edges with low degree.

- Consider a triangular pillow that is formed by two distinct tetrahedron surrounding a degree 2 vertex. A *2-0 vertex move* flattens this pillow and replaces the two tetrahedron with a single face, see Figure 21(a).
- Consider a bigon pyramid that is formed from two district tetrahedron surrounding a degree 2 edge. A *2-0 edge move* flattens the pyramid to a pair of faces that share



(a) A 2-0 vertex move

(b) A 2-0 edge move

(c) A 2-1 edge move

Figure 21: The local moves around vertices and edges with low degrees (taken from [1])

an edge, see Figure 21(b).

- Consider a tetrahedron Δ^3 that has two faces identified around one of it's edges. Let $\Delta^{3\prime}$ be an adjacent tetrahedron. A *2-1 edge move* flattens two faces of $\Delta^{3\prime}$ together and re-triangulates the region with a single tetrahedron, see Figure 21(c).

These moves have the potential to change the underlying manifold, but the below conditions guarantee when the moves are safe.

- A 2-0 vertex move is safe if the faces bounding the pillow are not identified and not both on the boundary.
- A 2-0 edge move is safe if the two edges being flattened together are not identified in the triangulation. Also the two faces on either sides of the flattening edges cannot be identified together or both be on the boundary (similar to the above condition). Finally both sets of faces on either side of the flattening edge cannot be identified, or two cannot be identified if the other two are both on the boundary.
- A 2-1 edge move is safe if the two faces of $\Delta^{3\prime}$ that are identified are not identified and both are not on the boundary.

6.3.3 Moves on the boundary

The following moves relate to modifying the boundary of the triangulation.

(a) Book opening and closing moves.

(b) A boundary shelling move.

Figure 22: Moves on the boundary (taken from [1]).

- A *book opening move* takes a face with exactly two edges on the boundary and unfolds the tetrahedrons that are identified at the face adding two triangles to the boundary. A *book closing move* is the inverse and identifies two triangles on the boundary that share an edge. See Figure 22(a).
- A *boundary shelling move* removes a tetrahedron Δ^3 that has one, two, or three faces on the boundary of the triangulation. See Figure 22(b).

As with the low degree moves, the boundary moves have the potential to change the underlying manifold.

- A book opening move is always safe, but a book closing move is only safe if the two faces do not make up the entire boundary component and the vertices opposite the shared edge are not already identified.
- A boundary shelling move's safety depends on the number of faces the tetrahedron Δ^3 being removed has on the boundary.
  - If Δ^3 has three faces on the boundary, removing it is always safe.
  - If Δ^3 has two faces on the boundary, removing it is only safe if the two faces not on the boundary are not identified and the edge not on these faces is internal to the triangulation.
  - If Δ^3 has one face on the boundary, removing it is only safe if the none of the edges not on the face are identified and the vertex not on the face is internal to the triangulation.

6.3.4 Edge Collapse

Finally the edge collapse move is the most powerful move available. Given an edge e that joins two distinct vertices, the *edge collapse move* crushes e to a point and flattens every tetrahedron that contains e to a face, see Figure 23. The conditions to make sure an edge

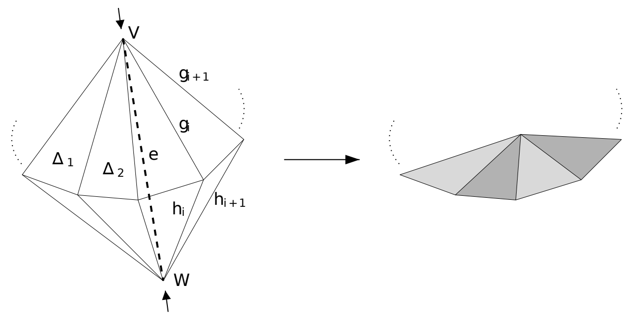


Figure 23: Edge collapse move (taken from [1])

collapse move is safe are much more complex than the previous moves. Let $\Delta_1, \ldots, \Delta_d$ be the d tetrahedra that share e as an edge and V, W as the distinct vertices being collapsed. If the following four conditions are met[10], then the edge collapse move is safe.

1. $\Delta_1, \ldots, \Delta_d$ are all distinct.
2. V and W do not both lie on the boundary.
3. Let $g_1, \ldots, g_d$ be the edges that touch V and $h_1, \ldots, h_d$ be the corresponding edges that touch W. We construct a graph Γ as follows:
   - The nodes of Γ are the edges of the triangulation.
   - We add an edge connecting each pair of nodes corresponding to edges g_i, h_i.
   - We add a "boundary" node ∂ with edges connecting it to each edge on the boundary of the triangulation.

   Γ must not contain any cycles.
4. We construct a similar graph for the faces touching V and W and require that the face graph also contain no cycles.

6.3.5 Simplification Algorithm

With the above moves we can describe in full the simplification algorithm employed in Regina. The input to the algorithm is a triangulation K and the output is a simpler triangulation (fewer tetrahedra) representing the same underlying manifold.

[10] These conditions only consider internal edges. The conditions for boundary edges are similar but slightly more complex

Algorithm 6.3.1.

1. *Greedily reduce the number of tetrahedra by repeatedly applying the following moves in the given priority until no moves are possible:*
   (a) *Edge collapse*
   (b) *3-2 Pachner move, 2-0 edge move, or 2-1 edge move*
   (c) *2-0 vertex move*
   (d) *boundary shelling move*
2. *Make up to $5R$ random 4-4 moves, with R being the maximum number of available 4-4 moves for the current triangulation. If after any 4-4 move a move from step 1 becomes available, return to step 1.*
3. *If K has boundary then make book opening moves until none are available. If at any point we can collapse an edge, return to step 1. If after we've performed all available book opening moves we cannot collapse an edge, then undo each move and continue.*
4. *If there is an available book closing move, then perform it and return to step 1. Otherwise terminate and return the current triangulation.*

As shown in [1], the above simplification algorithm runs in $O(n^4 \log n)$ time where n is the number of tetrahedra in the input triangulation. This may not seem very good, but compared to the exponential time of the high level 3-manifold algorithms, this simplification algorithm provides a powerful preprocessing step that makes the high level algorithms computationally tractable.

7 S^3 Recognition

We end this survey of normal surface theory with an exploration of an important problem in 3-manifold algorithms, S^3 recognition. The S^3 recognition problem is a particularly interesting problem to explore because it requires a sophisticated application of normal surface theory to first prove the existence of an algorithm, and requires state of the art implementation techniques to make the algorithm computational tractable. Unfortunately the normal surfaces we have examined thus far are not sufficient to develop the algorithm and we must introduce the concept of *almost normal surfaces*.

Definition 7.0.1. Given a triangulated manifold (M, K), a proper surface S is *almost normal* if for exactly one tetrahedron Δ^3 in K, $S \cap \Delta^3$ consists of exactly one disc whose boundary is normal curve c with length 8 and possibly other normal triangles, and for every other tetrahedron Δ^3, $S \cap \Delta^3$ is normal discs.

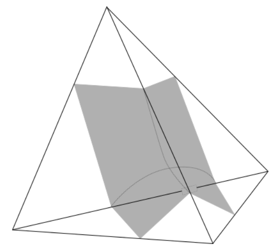


Figure 24: Almost normal disc

(a)

(b)

Figure 25: Types of components that can result from cutting along Σ.

7.1 Algorithm

Theorem 7.1.1 (Thompson, 1994; Rubinstein, 1992)**.** *There exists an algorithm to determine if a compact 3-manifold is S^3.*

Outline of the proof of 7.1.1. Suppose we have a compact triangulated manifold (M, K). Let Σ be a maximal collection of disjoint non-parallel normal 2-spheres in M. First notice that Σ is not empty because around every vertex of K, there is a normal 2-sphere made up of triangular pieces. Σ cuts M into three types of components M_0:

1. M_0 is a 3-ball containing a single vertex of K, as in Figure 25(a).
2. M_0 contains more than one boundary component as in Figure 25(b).
3. M_0 has exactly one boundary component but is not of type 1.

Using simplicial homology we can algorithmically guarantee that all 2-spheres in M are separating. Therefore determining if M is a 3-sphere is equivalent to determining if each component M_0 is a punctured 3-ball. The meat of the proof is proving the following two lemmas:

Lemma 7.1.2 ([8, Lemma 2]). *A component of type 2 is a punctured 3-ball.*

Lemma 7.1.3 ([8, Lemma 4]). *A component of type 3 is a 3-ball if and only if it contains an almost normal 2-sphere.*

Assuming 7.1.2 and 7.1.3 we can present the algorithm to determine if M is a 3-sphere:

1. Search Σ, a maximal collection of disjoin non-parallel normal 2-spheres. This is done by searching through the fundamental surfaces for the matching system for disjoint non-parallel normal 2-spheres until no more can be added to the collection.
2. For each component of $M \backslash \Sigma$ with a single boundary component, search it for an almost normal 2-sphere.
3. M is a 3-sphere if and only if each component from step 2 contains an almost normal 2-sphere.

■

Presented in this form, the proof of the 3-sphere recognition algorithm is actually quite simple. There are three points we must address to fully complete the proof:

1. Searching for almost normal 2-spheres, which requires a modification to our method for finding normal surfaces.
2. The proof of 7.1.2 and 7.1.3 can be found in [8]. The argument for 7.1.3 becomes very technical and requires tools from knot theory.

7.2 Searching for almost normal 2-spheres

The first step is to create a matching system E for almost normal surfaces. This requires only slight modification to our setup from before. Fix a tetrahedron Δ^3 of the triangulation K, and fix a normal curve c of length 8 on the boundary of Δ^3. Now create a system of matching equations E such that for the admissible solutions of E will contain normal triangles and quads in every tetrahedron different from Δ^3, and will contains normal triangles and possible an octagonal component in Δ^3. All of our results from before concerning fundamental surfaces still apply in this case, and we can still interpret the admissible sum of solutions as performing regular switches along intersection curves. Finally to complete the argument we need to show that if M contains an almost normal 2-sphere it contains a fundamental almost normal 2-sphere. Unfortunately we can't quite get there; instead we restrict our attention to a component M_0.

Lemma 7.2.1. *If there exists an almost normal 2-sphere in M_0 then there exists one which is a fundamental solution to E.*

Proof. Let S be an almost normal 2-sphere in M_0 and suppose $S = G_1 + \cdots + G_k$, where G_i is a fundamental solution to E. Notice that since the addition of two almost normal solutions is not admissible, exactly one of G_i is almost normal and the rest are normal surfaces. Notice that since S lies inside M_0, each G_i must also lie in M_0. Therefore, since M does not contain a closed non-orientable surface (it has trivial homology), and the Euler characteristic is additive under Haken Sum, at least one of the G_i, say G_1, is a 2-sphere. If G_1 is an almost normal surface then we are done. Supposing otherwise, G_1 must be parallel to the boundary of M_0, since otherwise Σ would not have been a maximal collection.

We can construct a normal isotopy i of M_0 such that $i(G_1)$ is disjoint from $i(G_2 \cup \cdots \cup G_k)$. i can be described as follows: for each tetrahedron of the triangulation, simply push the components of $G_2 \cup \cdots \cup G_k$ into M_0 so that they are disjoint from the $S^2 \times [0, 1]$ component formed by G_1 and the boundary.

Since i does not change the structure of G_i or $G_2 \cup \cdots \cup G_k$, we must have that,

$$G_1 + G_2 + \cdots + G_k = i(G_1) + i(G_2) + \cdots + i(G_k).$$

However $i(G_1) + i(G_2) + \cdots + i(G_k)$ is not connected since $i(G_1)$ is disjoint from $i(G_2 \cup \cdots \cup G_k)$. Therefore S cannot be written as a non-trivial sum of fundamental surfaces and thus must itself be a fundamental surface. ■

7.3 Implementation

We now have an algorithm to detect S^3, but as mentioned before this algorithm is not sufficient if we want to implement in software. The main difficulty in implementing S^3 recognition is the cutting procedure when examining the components of $M - \Sigma$. Cutting a manifold is a non-trivial operation. After cutting a triangulated manifold, the resulting components are no longer triangulated. Therefore we must re-triangulate, and in general re-triangulating the components will result in adding more tetrahedra than before. So the exponential behavior we see when enumerating normal surfaces becomes even worse.

Given these difficulties, how does Regina implement 3-sphere recognition? The key machinery comes from Jaco and Rubinstein. In [5] they develop an efficient procedure to cut a 3-manifold along a normal surface, known as *crushing*. The main idea is instead of cutting the manifold along the surface and re-triangulating the components with boundary, you identify the new boundary components to a point. After the crushing, the resulting triangulation is guaranteed to have fewer tetrahedra than the original triangulation. In [2], Burton explores the crushing procedure in the context of implementing 3-manifold algorithms and summarizes the results of [5]. We first begin with an outline of the crushing procedure.

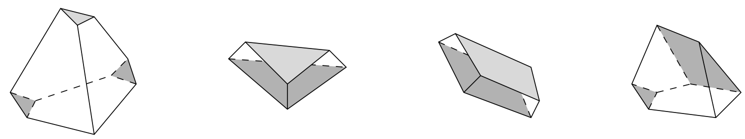
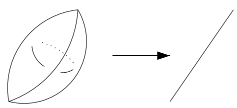
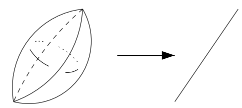


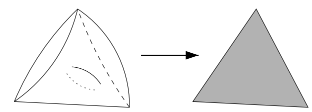
Figure 26: Examples of cells obtained after cutting along a surface.



(a) A 3-sided football



(b) A 4-sided football



(c) A triangular purse

Figure 27: Non tetrahedron components obtained after crushing.

Definition 7.3.1 ([2, 5])**.** The *Jaco-Rubinstein Crushing Procedure* operates on a normal surface S in a triangulated manifold (M, K) as follows:

1. We first cut K along S, producing what is known as a cell decomposition. Essentially a cell decomposition is a generalization of a triangulation where the number of allowable building blocks is much larger. In this case, the cell types take a variety of forms resulting from cutting tetrahedra along intersecting planes, see Figure 26. If S is two-sided in K, then after the cutting operation we get two new boundary components in the resulting cell decomposition that are homeomorphic to S. If S is one-sided we get one new boundary component that is homeomorphic to S.

2. Next we collapse each copy of S on the boundary to a point. This then converts the cell decomposition from before to a new cell decomposition C with four types of cells: tetrahedron, 3-sided football (Figure 27(a)), 4-sided football (Figure 27(b)), and triangular purses (Figure 27(c)).

3. Finally we eliminate any non-tetrahedron cells, as shown in Figure 27. To get a valid triangulation K', we keep only the surviving tetrahedra and the gluings along 2-faces. We remove degenerate components as outlined in Figure 28.

An important note is that the crushing procedure 7.3.1 is "destructive" in the sense that the resulting triangulation K' will in general change the topology of the underlying manifold more than just cutting along the surface. However the following theorem from [2] that summarizes the results of [5] describes exactly the possible changes that could

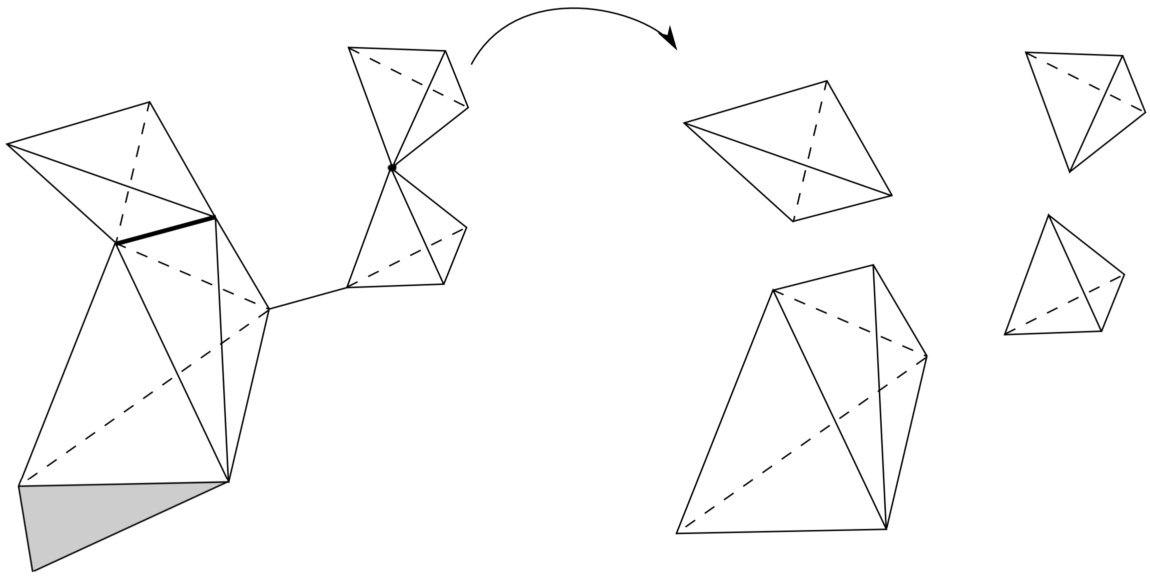


Figure 28: Removing degenerate components after crushing.

be introduced when crushing along a sphere or disc.

Theorem 7.3.2 (Jaco and Rubinstein [5]). *Let K be a triangulation of a compact orientable 3-manifold M, and let S be a normal sphere or disc in K. Then, if we crush S using the Jaco-Rubinstein crushing procedure 7.3.1, the underlying manifold M' of the resulting triangulation K' can be obtained from M by zero or more of the following operations:*

- *Undoing a connected sum, i.e. replacing some component $M'' = M_1''\#M_2''$ with the disjoint union $M_1'' \cup M_2''$;*
- *Cutting open along proper discs;*
- *Filling boundary spheres with 3-balls;*
- *Deleting a 3-ball, 3-sphere, $\mathbb{R}P^3$, $L_{3,2}$, or $S^2 \times S^1$ components.*

The above theorem is of particular interest in the context of the 3-sphere recognition algorithm because it allows us to efficiently cut along 2-spheres, which is the most expensive aspect of Algorithm 7.1.1.

The following is the complete 3-sphere recognition algorithm as implemented in Regina. The input is a triangulation K and the algorithm outputs if K is a triangulation of the 3-sphere.

Algorithm 7.3.3.

1. *Test if K is closed, connected, and orientable. If K fails any of these tests return false.*

2. *Simplify K using 6.3.1.*
3. *Test if K has trivial homology and return false if it does not.*
4. *Create a list of L of triangulations to process, initialized with K.*
5. *While L is not empty:*
    - *Remove the first triangulation N to process from L.*
    - *Search N for a quadrilateral normal 2-sphere F.*
        - *If F exists, use the Jaco-Rubinstein Crushing Procedure 7.3.1 on F. Simplify each new component N' and add it back to L.*
        - *If no F can be found and N contains a single vertex, search N for an almost normal quadrilateral 2-sphere. If none can be found return false.*
6. *Once L has been fully processes, terminate and return true.*

Notice that the overall structure of Algorithm 7.3.3 is the same as the original algorithm 7.1.1. There are two differences worth noting. The first is that instead of searching for Σ, the maximal collection of disjoint non-parallel normal 2-spheres, explicitly and then examining the resulting components, we recursively reduce M via normal 2-spheres until we have examined all of the components. This is largely a product of implementation, as it provides a way to throw out components as we process them. The second difference is more substantial, and it is the use of the Jaco-Rubinstein crushing procedure as a proxy for cutting. As previously discussed, this is a hugely important tool that makes Algorithm 7.3.3 possible to implement. However it does not represent a fundamental change in the original algorithm, and instead is a more sophisticated tool to accomplish the subproblem of getting the components.

8 Conclusion

In this survey we have examined how to approach questions involving 3-manifolds from an algorithmic viewpoint. In the process we developed significant machinery to make 3-manifolds computationally approachable, and looked at a high level algorithm to recognize the 3-sphere.

References

[1] Benjamin A. Burton. Computational topology with Regina: Algorithms, Heuristics and Implementations. *Contemporary Mathematics*, 596:195–224, 2013.

[2] Benjamin A. Burton. A new approach to crushing 3-manifold triangulations. *Discrete and Computational Geometry*, 52(1):116–139, 2014.

[3] Benjamin A. Burton, Ryan Budney, William Pettersson, et al. Regina: Software for low-dimensional topology, 1999-2017.

[4] Benjamin A. Burton, João Paix, and Jonathan Spreer. Computational topology and normal surfaces: Theoretical and experimental complexity bounds. In *Proceedings of the Meetings on Algorithms Engineering and Experiments*, pages 78–87. SIAM, 2013.

[5] William Jaco and J. Hyam Rubinstein. 0-efficient triangulations of 3-manifolds. *Journal of Differential Geometry*, 65(1):61–168, 2003.

[6] Sergei Matveev. *Algorithmic Topology and Classification of 3-Manifolds*. Springer, Berlin, 2007.

[7] Jennifer Schultens. *Introduction to 3-Manifolds*. American Mathematics Society, Providence, 2014.

[8] Abigail Thompson. Thin position and the recognition problem for s^3. *Mathematical Research Letters*, 1:613–630, 1994.