The number of Reidemeister Moves Needed for Unknotting

Joel Hass and Jeffrey C. Lagarias *

May 7, 1998

Abstract

There is a positive constant c_1 such that for any diagram $\mathcal{D}$ representing the unknot, there is a sequence of at most $2^{c_1 n}$ Reidemeister moves that will convert it to a trivial knot diagram, where n is the number of crossings in $\mathcal{D}$. A similar result holds for elementary moves on a polygonal knot K embedded in the 1-skeleton of the interior of a compact triangulated orientable PL 3-manifold M. There is a positive constant c_2 such that for each $t \geq 1$, if M consists of t tetrahedra, and K is unknotted, then there is a sequence of at most $2^{c_2 t}$ elementary moves in M which transforms K to a triangle contained inside one tetrahedron of M. We obtain explicit values for c_1 and c_2.

Keywords: knot theory, knot diagram, Reidemeister move, normal surfaces, computational complexity

*This paper grew out of work begun while the authors were visiting the Mathematical Sciences Research Institute in Berkeley in 1996/7. Research at MSRI is supported in part by NSF grant DMS-9022140. The first author is partially supported by NSF grant DMS-9704286.

The Number of Reidemeister Moves Needed for Unknotting

Joel Hass
University of California, Davis
Davis, CA 95616

Jeffrey C. Lagarias
AT&T Labs – Research
Florham Park, NJ 07932

(May 7, 1998)

1. Introduction

A knot is an embedding of a circle S^1 in a 3-manifold M, usually taken to be $\mathbb{R}^3$ or S^3. In the 1920's Alexander and Briggs [2, §4] and Reidemeister [22] observed that questions about ambient isotopy of polygonal knots in $\mathbb{R}^3$ can be reduced to combinatorial questions about knot diagrams. These are labeled planar graphs with overcrossings and undercrossings marked, representing a projection of the knot onto a plane. They showed that any ambient isotopy of a polygonal knot can be achieved by a finite sequence of piecewise linear moves which slide the knot across a single triangle, which are called *elementary moves* (or Δ-moves). They also showed that two knots were ambient isotopic if and only if their knot diagrams were equivalent under a finite sequence of local combinatorial changes, now called *Reidemeister moves*, see §7.

A triangle in M defines a *trivial knot*, and a loop in the plane with no crossings is said to be a *trivial knot diagram.* A knot diagram $\mathcal{D}$ is *unknotted* if it is equivalent to a trivial knot diagram under Reidemeister moves.

We measure the complexity of a knot diagram $\mathcal{D}$ by using its *crossing number*, the number of vertices in the planar graph $\mathcal{D}$, see §7. A problem of long standing is to determine an upper bound for the number of Reidemeister moves needed to transform an unknotted knot diagram $\mathcal{D}$ to the trivial knot diagram, as an explicit function of the crossing number n, see Welsh [29, p. 95]. This paper obtains such a bound.

Theorem 1.1 *There is a positive constant c_1, such that for each $n \geq 1$, any unknotted knot diagram $\mathcal{D}$ with n crossings can be transformed to the trivial knot diagram using at most $2^{c_1 n}$ Reidemeister moves.*

We obtain the explicit value of 10^{11} for c_1; this value can clearly be improved. In this paper we have not attempted to get an optimal value for c_1, and several of our estimates are accordingly rather rough.

In 1934 Goeritz [8] showed that there exist diagrams $\mathcal{D}$ of the unknot such that any sequence of Reidemeister moves converting $\mathcal{D}$ to the trivial knot must pass through some intermediate knot diagram $\mathcal{D}'$ that has more crossings than $\mathcal{D}$.

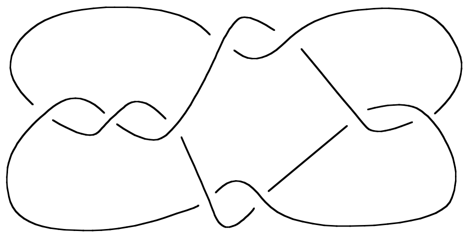

Figure 1: Goeritz's unknot.

It is not known how much larger the number of crossings needs to be for a general n-crossing knot, but Theorem 1.1 implies the following upper bound.

Corollary 1.1 *Any unknotted knot diagram $\mathcal{D}$ with n crossings can be transformed by Reidemeister moves to a trivial knot diagram through a sequence of knot diagrams each of which has at most $2^{c_1 n}$ crossings.*

This holds because each Reidemeister move creates or destroys at most two crossings, and all crossings must be eliminated by the final Reidemeister move. It seems possible that for each knot diagram representing the unknot there exists some sequence of Reidemeister moves taking it to the trivial knot diagram which keeps the crossing number of all intermediate knot diagrams bounded by a polynomial in n.

In a general triangulated 3-manifold, a trivial knot is a triangle lying inside one tetrahedron and an elementary move takes place entirely within a single tetrahedron of the triangulation. The proof of Theorem 1.1 is based on a general result that bounds the number of elementary moves needed to transform a closed polygonal curve K embedded in the 1-skeleton of a triangulated PL 3-manifold M to a trivial knot.

Theorem 1.2 *There is a constant c_2 such that for each $t \geq 1$, any compact PL 3-manifold with boundary $(M, \partial M)$ triangulated by t tetrahedra has the following property: If K is an unknot embedded in the 1-skeleton of interior (M), then K can be isotoped to the trivial knot using at most $2^{c_2 t}$ elementary moves.*

We obtain the explicit value $c_2 = 10^7$. The hypothesis that K lies in the interior of M can be weakened to requiring only that K lies in M, as we indicate in §6.

The methods we use are based on the normal surface theory of Kneser [19] and Haken [9]. Haken applied this theory to obtain an algorithm to decide if a knot is trivial. See [10] for a survey of algorithms to recognize unknotting, anD [5] and [4] for possible alternate approaches based on braids.

The paper is organized as follows. In §2 we define elementary moves and outline the proof of Theorem 1.2. In §3, §4 and §5 we bound the number of elementary moves needed to isotop a curve in three distinct settings. First we consider an isotopy of a curve to a point across a compressing disk in a 3-manifold, secondly we consider an isotopy through a solid torus from a core to a longitude, and finally we consider an isotopy across a surface between two isotopic curves. In §6 we tie these results together to complete the proof of Theorem 1.2. We then consider the more special case of knots in $\mathbb{R}^3$. In §7 we outline the proof of Theorem 1.1 and begin it by relating elementary moves to Reidemeister moves. In §8 we start with a knot diagram and show

how to construct a triangulation of a convex submanifold of $\mathbb{R}^3$ containing the corresponding knot on its 1-skeleton. This allows us to apply Theorem 1.2. In §9 we complete the proof of Theorem 1.1.

2. Elementary Moves in a 3-manifold

In this section we define elementary moves for knots and links. Elementary moves are well-defined in arbitrary triangulated PL 3-manifolds. In section 7 we will discuss Reidemeister moves, which are associated to knots and links in $\mathbb{R}^3$, and relate these two concepts for manifolds PL-embedded in $\mathbb{R}^3$.

A (piecewise linear, unoriented) *knot* K in a triangulated PL manifold M is a closed embedded polygonal curve. An (unoriented) *link* L is a finite union of nonintersecting knots in M. Let $|L|$ denote the number of vertices of L.

An *elementary move* (or Δ-*move*) on a knot or link consists of one of the following operations, see Burde and Zieschang [6, p. 4], Murasugi [20, p. 7]. Each of these operations is required to take place in a single tetrahedron of M.

(1) Split an edge into two edges by adding a new vertex to its interior.

(1′) [Reverse of (1)] Combine two edges whose union is a line segment, erasing their common vertex.

(2) Given a line segment AB of L that contains a single interior vertex D and a point C not in L such that the triangle ABC intersects L only in edge AB, erase edge AB and add two new edges AC and BC and new vertex C.

(2′) [Reverse of (2)] If for two adjacent edges AC and BC of K the triangle ABC intersects L only in AC and BC then erase AC, BC and vertex C and add the edge AB, with a new interior vertex D.

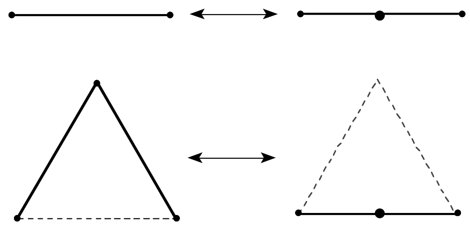

Figure 2: Two types of elementary moves.

Corresponding to elementary moves of type (2), (2′) is the piecewise-linear operation of linearly pulling the edges of a link across a triangle ABC. To define this operation, an extra vertex D on the segment AB must be specified; this is the role of elementary moves (1) and (1′).

Two knots K_1 and K_2 are *equivalent* if one can be reached from the other by a finite sequence of elementary moves. This gives the same equivalence classes as requiring K_1 and K_2 to be *ambient isotopic* [24] in M. The *unknot* is the class of knots equivalent to a triangle contained in a tetrahedron, and a knot K is *unknotted* if it is equivalent to the unknot. Equivalently, an unknot is the boundary of a PL embedded disk in M [24]. We will work with knots embedded in the interior of M. This is not a significant restriction, since it is possible to perturb any knot which meets ∂M into the interior of M after some simplicial subdivision.

We conclude this section by outlining the proof of Theorem 1.2. We first construct a compact triangulated PL submanifold M_K by removing a solid torus neighborhood of K. By removing the interior of a solid torus neighborhood R_K of K we construct a new triangulated manifold $M_K = M - R_K$. The boundary of M_K consists of the original boundary ∂M plus a 2-torus ∂R_K, called the *peripheral torus.*

If K is unknotted then M_K contains an *essential normal disk* S. Essential means that $\partial S \subseteq \partial R_K$ is not the boundary of any disk in the torus ∂R_K. Drawing on Haken's normal surface theory, a result of Hass, Lagarias and Pippenger [11] implies that S can be chosen to contain at most 2^{8t+6} triangles. We present this argument for completeness in §2. Its boundary ∂S is a curve K_2. K_2 is isotopic to K within $\bar{R}_K$. There is a sequence of at most 2^{8t+7} elementary moves which deform K_2 to a triangle across triangles in S. This is carried out in §3. It remains to relate K_2 to K by elementary moves.

In §3 we construct a curve K_1 isotopic to K in R_K which is a *longitude*, a curve on the peripheral torus, ∂R_K, which bounds an embedded disk in M_K but not inside the peripheral torus. Such curves exist if K is unknotted in M. In §4 we construct a longitude K_2 that lies in the 1-skeleton of the peripheral

2-torus ∂R_K, consists of at most $O(2^{c_3 t})$ line segments and can be isotoped to K using at most $O(2^{c_3 t})$ elementary moves. In §5 we prove results which are used to show that there is a sequence of at most $O(2^{c_4 t})$ elementary moves which deform K_1 to K_2 while remaining entirely in the peripheral 2-torus ∂R_K. In §6 we combine these results to complete the proof. All the constants and bounds above are explicitly computed.

The proof of Theorem 1.2 in effect deforms the original knot K to a trivial knot with elementary moves made inside three surfaces: an annulus is used to go from K to the longitude K_1, a 2-torus is used to go from K_1 to the boundary of a normal disk K_2, and a disk is used to go from K_2 to the trivial knot.

3. Normal surfaces and elementary moves across an essential disk

In this section we assume given a compact PL 3-manifold with (possibly empty) boundary $(M, \partial M)$, which is triangulated using t tetrahedra. We will work with a knot K embedded in the 1-skeleton of Interior(M). By barycentrically subdividing M twice, we obtain a manifold $(M'', \partial M'')$ in which K has a regular neighborhood R_K. The regular neighborhood, obtained by taking the closed star of K in M'', is a solid torus with ∂R_K disjoint from ∂M. We let $M_K = (M'' \backslash R_K) \cup \partial R_K$ and note that its boundary $\partial M_K = \partial M'' \cup \partial R_K$. We call the manifold $(M_K, \partial M_K)$ with ∂R_K marked a *truncated knot complement* for K.

A *meridian* on the peripheral torus ∂R_K is an oriented simple closed curve μ whose homology class $[\mu]$ in $H_1(\partial R, \mathbb{Z})$ generates the kernel of the map

$$j_* : H_1(\partial R_K, \mathbb{Z}) \to H_1(R_K, \mathbb{Z}),$$

where $j : \partial R_K \hookrightarrow R_K$ is the inclusion map. Equivalently, it is a simple closed curve which bounds a disk in R_K but not in ∂R_K.

A *longitude* on the peripheral torus ∂R_K is an oriented simple closed curve λ which intersects a meridian transversely at a single point and whose homology class $[\lambda]$ in $H_1(\partial R, \mathbb{Z})$ generates the kernel of the map

$$i_* : H_1(\partial R, \mathbb{Z}) \to H_1(M_K, \mathbb{Z}),$$

where $i : \partial R_K \hookrightarrow M_K$ is the inclusion map, provided that this kernel is non-trivial. Note that the kernel is always infinite cyclic or trivial [6].

If the core K of the solid torus R_K is unknotted in M, then a longitude exists, and it bounds a disk in M_K. For general K, a longitude does not exist when K represents an element of infinite order in $H_1(M, \mathbb{Z})$. If $M \cong S^3$, or more generally a homology sphere, then a longitude exists for any K. This curve is sometimes called a *preferred longitude.*

When a longitude λ of ∂R_K exists, its intersection number on ∂R_K with a meridian μ is ± 1, and the homology classes $[\mu]$ and $[\lambda]$ generate $H_1(\partial R, \mathbb{Z}) \cong \mathbb{Z} \oplus \mathbb{Z}$.

A *properly embedded disk* in a 3-manifold with boundary $(M, \partial M)$ is a disk S PL-embedded in M which satisfies $S \cap \partial M = \partial S$. An *essential disk* in M is a properly embedded disk such that ∂S does not bound a disk in ∂M, i.e. the homotopy class of ∂S in $\pi_1(\partial M)$ is non-trivial.

Lemma 3.1 *Let M_K be a triangulated PL 3-manifold which is a truncated knot complement for K. Then K is unknotted if and only if M_K contains an essential disk S with ∂S a longitude in ∂R_K.*

Proof. If K is unknotted then it bounds a disk which is embedded in M. This disk can be intersected with M_K to give an essential PL disk S in M_K. Conversely if M_K contains an essential PL disk S with $\partial S \subseteq \partial R_K$, then the boundary ∂S of S represents a longitude on ∂R_K, which is isotopic to K in R_K [6, p. 29, Thorem 3.1], and therefore is equivalent to K. This disk can be used to contract ∂S until it becomes a single triangle in M_K. The first homology group of a torus, $Z \oplus Z$, is canonically isomorphic to its fundamental group, and the statement about the homology class of $[\partial S]$ is a consequence. $\square$

We will use normal surface theory in the form described in Jaco and Rubinstein [17, Section 1]. A *normal surface* S in a triangulated compact 3-manifold

M is a PL-surface whose intersection with each tetrahedron in M consists of a disjoint set of *elementary disks.* These are either triangles and quadrilaterals. A quadrilateral consists of two triangles glued together along an edge lying interior to the tetrahedron. All other edges of each triangle or quadrilateral are contained in the 2-skeleton of the tetrahedron, and do not intersect any vertices. (See Figure 3).

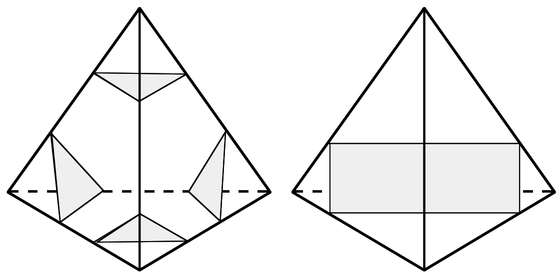

Figure 3: Elementary disks in a normal surface.

Within each tetrahedron of M there are only 7 kinds of possible triangle or PL quadrilateral, up to PL-isotopies that map the tetrahedron to itself. These are labeled by the 7 possible ways of partitioning the vertices of the tetrahedron into the two nonempty sets determined by cutting the tetrahedron along the triangle or quadrilateral. W. Haken observed* that the isotopy-type of a normal surface is completely determined by the number of pieces of each of the 7 kinds that occur in each tetrahedron. This can be described by a nonnegative integer vector $\mathbf{v} = \mathbf{v}(S) \in \mathbb{Z}^{7t}$, which gives the *normal coordinates* of S. He also showed that the set of allowable values $\mathbf{v}(S)$ for a normal surface S lie in a certain homogeneous rational cone $\mathcal{C}_M$ in $\mathbb{R}^{7t}$ which we call the *Haken normal cone.* If $\mathbf{v} = (v_1, v_2, \ldots, v_{7t}) \in \mathbb{R}^{7t}$, then Haken's normal cone is specified by a set of linear equations and inequalities of the form

$$v_i \geq 0 \ , \quad \text{for} \ \ 1 \leq i \leq 7t \ , \tag{3.1}$$

$$v_{i_1} + v_{i_2} = v_{i_3} + v_{i_4} \quad \text{(up to } 6t \text{ equations)} \ . \tag{3.2}$$

*Haken's [9] treatment of normal surfaces was based on an equivalent formulation using cells. We follow an alternate development based on tetrahedra, as in Jaco and Rubinstein [17, Section 1] and Jaco and Tollefson [18]

The second set of equations expresses *matching conditions* which say that the number of edges on a common triangular face of two adjacent tetrahedra coming from regions in each of the tetrahedra must match. For each triangular face there are three types of edges (specified by a pair of edges on the triangle), which yields 3 matching conditions per face. Triangular faces in the boundary ∂M give no matching equations. The cone $\mathcal{C}_M$ is *rational* because all equations 3.1, 3.2 have integer coefficients. We let

$$\mathcal{C}_M(\mathbb{Z}) = \mathcal{C}_M \cap \mathbb{Z}^{7t} \ , \tag{3.3}$$

which is the set of integral vectors in the Haken normal cone $\mathcal{C}_M$. Haken defined a *fundamental surface* to be a normal surface S such that

$$\mathbf{v}(S) \neq \mathbf{v}_1 + \mathbf{v}_2 \ , \quad \text{with } \mathbf{v}_1, \mathbf{v}_2 \in \mathcal{C}_M(\mathbb{Z}) \backslash \{\mathbf{0}\} \ .$$

Such a vector $\mathbf{v}(S)$ is called an element of the *minimal Hilbert basis* $\mathcal{H}(\mathcal{C}_M)$ of $\mathcal{C}_M$, using the terminology of integer programming, see Schrijver [26, Theorem 16.4].

Theorem 3.1 *(Haken) If a triangulated irreducible compact 3-manifold with boundary $(M, \partial M)$ contains an essential disk, with boundary in a component ∂R of ∂M, then it contains such an essential disk S with boundary in ∂R which is a fundamental normal surface.*

Proof. See Jaco and Rubinstein [17, Theorem 2.3] for a treatment of the existence of an essential normal disk in M with boundary on ∂R. For a fundamental disk, we refer to Jaco and Tollefson [18, Corollary 6.4], which establishes the stronger result that there exists such a disk that is a fundamental surface and is also a *vertex* surface, which means that $\mathbf{v}(S)$ lies on an extreme ray of Haken's normal cone $\mathcal{C}_M$. $\square$

Remark. This result applies in arbitrary 3-manifolds. The proof in [18] assumes compactness and irreducibility of M, but these hypotheses can be removed. A classical theorem of Alexander implies that knot complements in the 3-sphere are always irreducible.

Hass, Lagarias and Pippenger [11] give a simple bound for the complexity of any fundamental surface, whose proof we include for completeness.

Lemma 3.2 *Let M be a triangulated compact 3-manifold, possibly with boundary, that contains t tetrahedra.*

(1) *Any vertex minimal solution $\mathbf{v} \in \mathbb{Z}^{7t}$ of the Haken normal cone C_M in $\mathbb{R}^{7t}$ has*

$$\max_{1 \leq i \leq 7t} (v_i) \leq 2^{7t-1} . \tag{3.4}$$

(2) *Any minimal Hilbert basis element $\mathbf{v} \in \mathbb{Z}^{7t}$ of the Haken fundamental cone C_M has*

$$\max_{1 \leq i \leq 7t} (v_i) < t \cdot 2^{7t+2} . \tag{3.5}$$

Proof. (1) Choose a maximal linearly independent subset of matching condition equations 3.2. There will be $7t - d$ of them, where $d = \dim_{\mathbb{R}}(\mathrm{C}_M)$. Any vertex ray is determined by adjoining to these equations $d-1$ other binding inequality constraints

$$\{v_{i_k} = 0; \quad 1 \leq k \leq d-1\} ,$$

with the proviso that the resulting system have rank $7t-1$. These conditions yield a $(7t-1) \times 7t$ integer matrix M of rank $7t-1$, and the vertex ray elements $\mathbf{z} = (z_1, \ldots, z_{7t})$ satisfy

$$\mathrm{M}\ \mathrm{z} = \mathbf{0} .$$

In order to get a feasible vertex ray in C_M all nonzero coordinates z_i must have the same sign.

At least one of the unit coordinate vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_{7t}$ must be linearly independent of the row space of M. Adjoin it as a first row to M and we obtain a full rank $7t \times 7t$ integer matrix

$$\tilde{\mathrm{M}} = \begin{bmatrix} \mathbf{e}_k \\ \mathrm{M} \end{bmatrix}, \quad \det(\tilde{\mathrm{M}}) \neq 0 .$$

Consider the adjoint matrix $adj(\tilde{\mathrm{M}}) = \det(\tilde{\mathrm{M}})\, \tilde{\mathrm{M}}^{-1}$, which has integer entries

$$w_{ij} = adj(\tilde{\mathrm{M}})_{ij} = (-1)^{i+j} \det\ \tilde{\mathrm{M}}[j|i] , \tag{3.6}$$

in which $\tilde{\mathrm{M}}[j|i]$ is the minor obtained by crossing out the j^{th} row and i^{th} column of $\tilde{\mathrm{M}}$. Let

$$\mathbf{w} = [w_{11}, w_{21}, \ldots, w_{7t,1}]^t$$

be the first column of $adj(\tilde{\mathrm{M}})$. Since $\tilde{\mathrm{M}}adj(\tilde{\mathrm{M}}) = \det(\tilde{\mathrm{M}})\,\mathrm{I}$ this yields

$$\mathrm{M}\mathbf{w} = \mathbf{0} \ ,$$

and $\mathbf{w} \neq \mathbf{0}$ because $\langle \mathbf{e}_k, \mathbf{w} \rangle = \det(\tilde{\mathrm{M}}) \neq 0$. We bound the entries of $\mathbf{w}$ using *Hadamard's inequality*, which states for an $m \times m$ real matrix N that

$$\det(\mathrm{N})^2 \leq \prod_{i=1}^{m} \|\mathbf{n}_i\|^2 \ , \tag{3.7}$$

in which $\|\mathbf{n}_i\|^2$ is the Euclidean length of the i^{th} row $\mathbf{n}_i$ of N. We apply this to equation 3.6, and observe that each row of the $(7t-1) \times (7t-1)$ matrix $\tilde{\mathrm{M}}[j|i]$ has squared Euclidean length at most 4, because this is true for all row vectors in the system of equations 3.2 and for $\mathbf{e}_k$. Applied to equation 3.6 this gives

$$|w_{ij}|^2 \leq 4^{7t-1} \ .$$

However equation 3.6 shows that $\mathbf{w} \in \mathbb{Z}^{7t}$, and a vertex minimal solution $\mathbf{v}$ in the extreme ray is obtained by dividing $\mathbf{w}$ by the greatest common divisor of its elements, hence

$$\max_{1 \leq i \leq 7t} |v_i| \leq \max_{1 \leq i \leq 7t} |w_{i1}| \leq 2^{7t-1} \ .$$

(2) A *simplicial cone* C in $\mathbb{R}^{7t}$ is a d-dimensional pointed cone which has exactly d extreme rays. Let $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)} \in \mathbb{Z}^{7t}$ be the vertex minimal solutions for the extreme rays. Each point in C can be expressed as a nonnegative linear combination of the $\mathbf{v}^{(j)}$, as

$$v = \sum_{j=1}^{d} c_j \mathbf{v}^{(j)}, \quad \text{each} \quad c_j \geq 0 \ .$$

If $\mathbf{v}$ is in the minimal Hilbert basis $\mathcal{H}(\mathrm{C})$ of the cone C, then $0 \leq c_j \leq 1$, for otherwise one has

$$\mathbf{v} = (\mathbf{v} - \mathbf{v}^{(j)}) + \mathbf{v}^{(j)} \ ,$$

and both $\mathbf{v}-\mathbf{v}^{(j)}$ and $\mathbf{v}^{(j)}$ are nonzero integer vectors in C, which is a contradiction. Thus, any minimal Hilbert basis element $\mathbf{v} = (v_1, \ldots, v_{7t})$ of a simplicial cone satisfies

$$|v_i| \le \sum_{j=1}^{d} |v^{(j)}| \quad \text{for} \quad 1 \le i \le 7t \ . \tag{3.8}$$

The cone C_M may not be simplicial, but we can partition it into a set of simplicial cones $\{\ \mathrm{C}_m^{(k)}\}$ each of whose extreme rays are extreme rays of C_M itself. We have

$$\mathcal{H}(\ \mathrm{C}_M) \subseteq \bigcup_k \mathcal{H}(\ \mathrm{C}_M^{(k)}) \ .$$

Thus all Hilbert basis elements of $\mathcal{H}(\ \mathrm{C}_M)$ satisfy the bound 3.8 for Hilbert basis elements of $\mathcal{H}(\ \mathrm{C}_M^{(k)})$. Using equation 3.4 to bound $|v^{(j)}|$ we obtain

$$|v_i| \le d2^{7t-1} < 7t2^{7t-1} < t2^{7t+2} \ ,$$

as claimed. $\square$

A PL-embedded normal disk S can be used as a template to transform its boundary ∂S to a single triangle by a sequence of elementary moves, in such a way that all intermediate curves lie on the surface S.

Lemma 3.3 *Let M be a piecewise-linear triangulated compact 3-manifold with boundary, and let S be a normal disk in M that consists of w triangles. Then ∂S can be isotoped to a triangle by a sequence of at most $2w$ elementary moves on M, each of which takes place in a triangle or edge contained within S.*

Proof. We will construct a sequence of w simple closed curves

$$\partial S = C_0, C_1, C_2, \ldots, C_{w-1} = \partial T$$

where T is a triangle in S, and such that for $0 \le i \le w-1$:

(i) C_{i+1} is obtained from C_i by at most two elementary moves; either a move of type $2'$ across a single triangle in S followed by a move of type $1'$ which removes the extra vertex created by the first move, or a move of type 1 followed by a move of type 2 across a single triangle in S.

(ii) There is a triangulated disk S_i contained in S which consists of $w - i$ triangles, such that $C_i = \partial S_i$.

The construction proceeds by eliminating triangles in S one at a time by a sequence of steps, each consisting of two elementary moves, in such a way as to preserve the property that each S_j is a topological disk. We proceed by induction on j, with the base case $j = 0$ holding by hypothesis. For the induction step, we note that the only way that removing from S_j a triangle that contains at least one edge of ∂S_j can yield a surface S_{j+1} that is not a topological disk is by producing a vertex $\mathbf{v}$ that is a cut-point, i.e. $\mathbf{v}$ is visited twice in the curve ∂S_{j+1}. In that case $\mathbf{v}$ is a vertex of ∂S_j, which implies that the triangle that is removed must have exactly one edge on ∂S_j, and the pair of elementary moves pulls this edge across to the two edges of this triangle that are adjacent to $\mathbf{v}$. We therefore proceed by first checking if S_j has any triangle having two edges on ∂S_j. If so, we remove this triangle to obtain S_{j+1}. If this cannot be done, then S_j must contain an interior vertex, because a triangulated polygon that has no interior vertices and contains more than one triangle contains at least 2 triangles having two edges on the boundary. In this case there exists a triangle having one edge on the boundary which also has an interior vertex, and we can pull across this triangle to obtain S_{j+1}. This completes the induction step. $\square$

Putting all these results together yields:

Lemma 3.4 *Let M_K be a triangulated PL 3-manifold consisting of t tetrahedra which is a truncated knot complement for K. If K is unknotted, then there is an essential normal disk S in M_K that contains at most 2^{8t+6} triangles, and S has a polygonal boundary $K_1 = \partial S$ which lies on the peripheral 2-torus ∂R_K and consists of at most 2^{8t+7} line segments. The boundary K_1 of S can be transformed to a single triangle by at most 2^{8t+7} elementary moves, with all intermediate curves being the boundary of some triangulated subdisk of S.*

Proof. By Lemma 3.1 and Theorem 3.1 there exists an essential normal disk which is a fundamental surface for M_K. The bound of Lemma 3.2 implies that this surface contains at most $7t^2 \cdot 2^{7t+2} < t^2 \cdot 2^{7t+5} < 2^{8t+5}$ triangles and quadrilaterals. Since each quadrilateral consists of two triangles in the triangulation, we obtain the upper bound 2^{8t+6} for the total number of triangles. The boundary ∂S contains at most two edges of each triangle, for a total of at most 2^{8t+7} edges. Lemma 3.3 gives the stated bound on elementary moves. $\square$

Remark. We could improve the above estimate by using a vertex minimal solution which is a disk.

4. Longitudes and cores of solid tori

In this section we bound the number of elementary moves needed to isotop a knot K in the 1-skeleton of a triangulated compact 3-manifold $(M, \partial M)$ across an annulus to a curve K_2 which is a longitude on the peripheral torus ∂R_K of K. The main result of this section is the following:

Theorem 4.1 *Let $(M, \partial M)$ be a triangulated compact PL 3-manifold with t tetrahedra, let K be an unknotted curve embedded in Interior(M), and let ∂R_K be the peripheral torus of K in the second barycentric subdivision of M. Then there exists a PL closed curve K_2 in ∂R_K which is a longitude for ∂R_K in M_K such that:*

> *(i). There is an embedded PL annulus S in the solid torus R_K whose two boundary components are K_2 and K, and which consists of at most 2^{1858t} triangles.*
>
> *(ii). There is an isotopy from K_2 to K across S that consists of at most $2^{1858t+1}$ elementary moves.*

Remarks. The closed curve K_2 is embedded in ∂R_K but generally does not lie in its 1-skeleton. In the special case that $(M, \partial M)$ is PL-homeomorphic to a triangulated convex polyhedron in $\mathbb{R}^3$, as occurs with a standard knot, the exponential bounds in (i), (ii) can be improved to bounds polynomial in t, as we indicate after the proof.

The proof of Theorem 4.1 proceeds in three steps. We first construct a closed curve α embedded in the 1-skeleton of R_K which is *parallel* to the core K, i.e. its homology class $[\alpha]$ in $H_1(\partial R_K, \mathbb{Z})$ has intersection number ± 1 with a meridian $[\mu]$. We next observe that the homology class in $H_1(\partial R_K, \mathbb{Z})$ of a longitude K_2 is necessarily

$$[K_2] = \pm([\alpha] + k[\mu]) \tag{4.1}$$

for some integer k. The second step is to bound $|k|$ by an exponential function of t. The third step is to obtain K_2 from α by adding k twists to α near a fixed meridian μ. Given K_2, we construct a triangulated annular surface S embedded in R_K which lies in the 2-skeleton of R_K except in a "collar" of one meridian disk, into which the k twists are inserted. All vertices of the triangulated annulus S lie on its boundary $\partial S = K_2 \cup K$. Finally we obtain an isotopy giving the elementary move bound by pulling K across the triangles in S to K_2.

We begin with some preliminary lemmas involving PL topology, which pin down some of the combinatorial structure of the regular neighborhood of K. Background for PL topology can be found in [16] or [25].

Let $v_0, \ldots, v_r$ be the vertices of the knot K in M. Since R_K lies in the second barycentric subdivision of M, each edge $[v_i, v_{i+1}]$ of K is subdivided into four edges by the addition of three new vertices . We will continue to denote the subdivided curve by K. Denote the original vertices v_i by w_{4i} and let $w_{4i+1}, w_{4i+2}, w_{4i+3}$, $0 \le i \le r$ denote the new vertices added in the two barycentric subdivisions. We also use the convention that $w_s = w_0$, where $s = 4r + 4$ is the total number of vertices, so that the vertices w_i of K are cyclically ordered.

We will break up the solid torus R_K into a union of what we call *wheels* and *wedges*. For each edge $[w_{i-1}, w_i]$ let the *wheel* around that edge be the closed star $St([w_{i-1}, w_i], R_K)$ of the edge $[w_{i-1}, w_i]$ in the solid torus R_K. This is the union of all the tetrahedra in R_K having $[w_{i-1}, w_i]$ as an edge, and is homeomorphic to a closed ball. For each vertex w_i on K let the *wedge around* w_i be the union of all the tetrahedra in R_K meeting K only at the point w_i.

Let

$$\mu_i = lk([w_{i-1}, w_i], R_K) \ , \quad 1 \leq k \leq s \ , \tag{4.2}$$

be the link of the edge $[w_{i-1}, w_i]$ in the solid torus R_K.

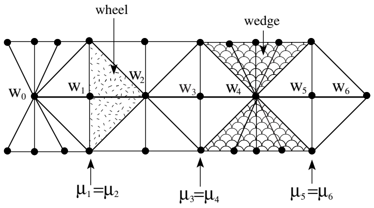

Figure 4: A 2-dimensional section of R_K.

Lemma 4.1 *(1) Each μ_i lies in ∂R_K and is a meridian.*
(2) μ_i and μ_j coincide if i is odd and $j = i + i$ or if i is even and $j = i - 1$. Otherwise μ_i and μ_j are disjoint.
(3) An edge on ∂R_K either lies on some μ_i or else lies on a unique triangle in R_K whose opposite vertex w lies on K.

Proof. (1) The closed star $St([w_{i-1}, w_i], R_K)$ is supported on the set of tetrahedra in R_K having $[w_{i-1}, w_i]$ as an edge. Each such tetrahedron has one edge which is disjoint from $[w_{i-1}, w_i]$, and the link μ_i is the union of these edges. The surface D_i which is obtained by coning μ_i to the vertex w_i forms a topological disk in R_K whose boundary is μ_i, and which intersect K transversely in a single point. Thus μ_i is a meridian; see Figure 5.

(2) This follows since R_K is a barycentric subdivision. The vertices w_i with i odd were added during the second barycentric subdivision. Prior to the second subdivision, these vertices lay along interior points of edges of K. The two meridians associated to the edges $[w_{i-1}, w_i]$ and $[w_i, w_{i+1}]$ coincide; see Figure 4 for a cross sectional sketch of R_K. For i even, the distance between w_{i-1} and w_{i+1} in $link(w_i)$ is at least four, since the link has been barycentrically subdivided. If $\mu_i \cap \mu_{i+1} \neq \emptyset$ then this distance equals two.

(3) Each edge in ∂R_K is an edge of a triangle with a vertex w_i on K, since R_K is obtained by taking the closed star of K in M''. If the edge is in a meridian

μ_i, then it lies in triangles with vertices at both w_{i-1} and w_i or w_i and w_{i+1}. If it doesn't lie on a meridian then it is contained in the wedge around w_i. Any triangle containing the edge and a point w' on K cannot cross the two wheels meeting w_i, and the point w' must coincide with w. $\square$

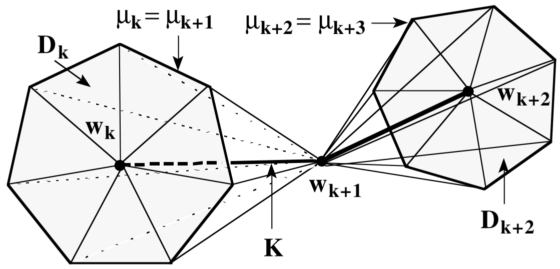

Figure 5: Adjacent meridians. The index k is odd.

Lemma 4.2 *There exists an embedded closed curve α in the 1-skeleton of ∂R_K which is parallel to K in R_K. There is an annulus S_0 embedded in the 2-skeleton of the solid torus R_K whose boundary components are α and K, and all vertices of S_0 lie in its boundary $\alpha \cup K$.*

Proof. We construct α as the union of a sequence of arcs. For each odd i, $1 \leq i \leq s$, we pick a path α_i in the 1-skeleton of ∂R_K that connects a vertex x_i of $\mu_i = \mu_{i+1}$ to a vertex x_{i+1} on $\mu_{i+2} = \mu_{i+3}$, does not cross μ_{i-1}, and is shortest among all such paths. (We don't allow α_i to cross μ_{i-1} to prevent it from running between the curves in the wrong direction.) We set $x_s = x_0$ as before. Note that α_i is embedded and that it meets the collection of meridians only at its endpoints, since otherwise it can be shortened.

We then add for each odd i a connecting path β_i on μ_i from the endpoint x_{i-1} of α_{i-2} to the initial point x_i of α_i, and we set

$$\alpha = \beta_1 \alpha_1 \beta_3 \alpha_3 \beta_5 \ldots \beta_{s-1} \alpha_{s-1} \ . \tag{4.3}$$

The path α is a closed curve, and is embedded in ∂R_K.

We construct the surface S_0 as follows. Each path β_i, $1 \leq i \leq s$, i odd, lies on a meridian $\mu_i = \mu_{i+1}$ which is the boundary of a meridian disk D_{i+1}

centered at the vertex w_{i+1} on K, and we cone β_i to w_{i+1} along D_{i+1}, obtaining a subdisk of D_{i+1}. For each edge α_i that lies on the boundary of the wedge around the vertex w_{i+1}, there is a unique triangle containing that edge and w_{i+1}. For fixed i, the union of these triangles forms a disk in R_K meeting K at the point w_{i+1}, and ∂R_K along $\beta_i \cup \alpha_i$, and whose boundary consists of $\beta_i \cup \alpha_i$ together with two interior edges of R_K, one running from the initial point x_{i-1} of β_i to w_{i+1} and one running from the endpoint x_{i+1} of α_i to w_{i+1} .

Finally for each meridian μ_i with i odd, there is a pair of triangles

$$[w_{i-1}, w_i, x_{i-1}]$$

and

$$[w_i, w_{i+1}, x_{i-1}],$$

whose union is a disk contained in the two wheels around $[w_{i-1}, w_i]$ and $[w_i, w_{i+1}]$, and connecting the free edge running from the endpoint x_{i-1} of α_{i-2} to w_{i-1} to the free edge running from the initial point (also x_{i-1}) of β_i to w_{i+1}. The union of all these disks forms an annulus which is embedded in the 2-skeleton of R_K, whose boundary components are α and K, and all vertices of which lie in $\alpha \cup K$. See Figure 6. $\square$

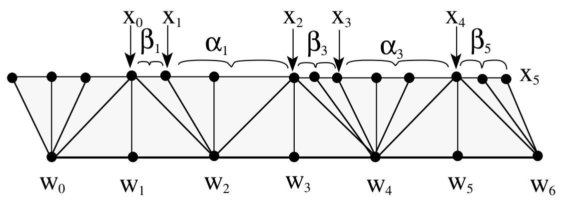

Figure 6: The annulus S_0.

Proof of Theorem 4.1. We wish to convert the closed curve α given in Lemma 4.2 to a longitude K_2 by adding "twists". The homology class $[K_2]$ of a longitude K_2 satisfies

$$[K_2] = \pm([\alpha] + k[\mu])$$

where μ is a fixed meridian. According to Lemma 4.1 we may choose μ to be μ_1 taken with a fixed orientation.

The number k represents the number of "twists" that we must insert in α to convert it to the knot K_2. We proceed to bound k in terms of the number of tetrahedra t in the original manifold M. We may assume $k \neq 0$ for otherwise we are done. The value of k is uniquely determined by the requirement that the homology class $[\alpha] + k[\mu]$ vanish in $H_1(M_K, \mathbb{Z})$. (Here we use the fact that $[\mu]$ is nonvanishing in $H_1(M_K, \mathbb{Z})$.) This is equivalent to the curve $\alpha + k\mu$ being a boundary in the cycle group. The manifold M_K contains at most $576t$ tetrahedra, since it is contained in the second barycentric subdivision, hence the number e of edges in M_K is at most $3456t$. The group of boundaries in the PL chain complex of M_K is generated by the paths around all 2-simplices in M_K, and the number f of these is at most $2304t$. We choose fixed orientations for the edges and faces, and we can then describe a basis of generators for the group of boundaries using an $e \times f$ matrix $\mathrm{M} = [\,\mathrm{M}_{ab}]$ with entries

$$\mathrm{M}_{ab} = \begin{cases} \pm 1 & \text{if } a \text{ is an edge in the oriented 2-simplex } b\ , \\ 0 & \text{otherwise}\ . \end{cases}$$

in which each column of M has exactly three nonzero entries. Let $\mathbf{c}_\alpha$ be the $e \times 1$-vector describing α as an oriented cycle in M_K, and $\mathbf{c}_\mu$ be the $e \times 1$ vector describing μ as an oriented cycle in M_K. All entries in $\mathbf{c}_\alpha$ and $\mathbf{c}_\mu$ are 0,1 or -1. The condition for $\mathbf{c}_\alpha + k\mathbf{c}_\mu$ to be a boundary is that the linear system

$$\mathrm{M}\mathbf{x} = \mathbf{c}_\alpha + k\mathbf{c}_\mu \tag{4.4}$$

have a solution in integer unknowns $\mathbf{x} = (x_1, \ldots, x_f)^T$. The hypotheses imply that this linear system is solvable over $\mathbb{Z}$ for a unique value of k. By taking a generating set of boundary faces, we have that M is of full column rank. The meridian class $\mathbf{c}_\mu$ is not in the column space of M since otherwise μ would bound a surface in M_K and the unknotted curve K would represent an element of infinite order in $H_1(M, \mathbb{Z})$. It follows that there is at most one *real* value of k for which the system (4.4) is solvable, for if there were two then $\mathbf{c}_\mu$ would be in the column space of M. Thus $\mathbf{c}_\mu$ is linearly independent of the column space of M and the linear system

$$[\,\mathrm{M}, \mathbf{c}_\mu] \begin{bmatrix} \mathbf{x} \\ \hline -k \end{bmatrix} = \mathbf{c}_\alpha\ ,$$

has full column rank. Choosing a suitable $f+1$ rows of $[\ \mathrm{M}\mathbf{c}_\mu]$ yields an invertible matrix $\tilde{\mathrm{M}}$ and

$$[\frac{x}{-k}] = \tilde{\mathrm{M}}^{-1}\tilde{\mathbf{c}}_\alpha \ ,$$

in which $\tilde{\mathbf{c}}_\alpha$ is the subset of $\mathbf{c}_\alpha$ determined by the same $f+1$ rows. By Cramer's rule we have

$$-k = \frac{1}{\det(\tilde{\mathrm{M}})} \sum_{i=1}^{f+1} \tilde{\mathrm{M}}[i, f+1](\tilde{\mathbf{c}}_\alpha)_i \ ,$$

where $\tilde{\mathrm{M}}[i, f+1]$ denotes a minor of $\tilde{\mathrm{M}}$ on the last column. Since $|(\tilde{\mathbf{c}}_\alpha)_i| \leq 1$ we have

$$|k| \leq \sum_{i=1}^{f+1} |\ \tilde{\mathrm{M}}[i, f+1]| \ . \tag{4.5}$$

We bound the terms in this expression using Hadamard's determinant bound (3.7). Each column of $\tilde{\mathrm{M}}$ has at most three nonzero entries, each equal to ± 1, except the last column, which does not appear in the $f \times f$ minor $\tilde{\mathrm{M}}[i, f+1]$. This yields the bound

$$|\ \tilde{\mathrm{M}}[i, f+1]|^2 \leq 3^f \ , \quad 1 \leq i \leq f+1 \ ,$$

which with (4.5) yields

$$|k| \leq (f+1)3^{\frac{f}{2}} \leq (2304t+1)3^{1152t} \ . \tag{4.6}$$

We now describe how the "twists" are added to α and the surface S_0 of Lemma 4.2. to yield an annulus S with boundary a longitude. The disk D_2 cuts the core K at the point w_2 and stretches to the boundary. We add the twists to S_0 by modifying it inside the wedge around w_2. Let A denote the surface consisting of the set of closed triangles in the wedge around w_2 which intersect the meridian μ_2 and which lie in ∂R_K. Then A is an annulus, and we suppose it contains ℓ triangles. The first edge e of α_1 has one endpoint on $\mu_1 = \mu_2$ and the other endpoint disjoint from μ_2 and on the other boundary component of A.

We replace the edge $[x_1, v]$ of α with a path α' of $k\ell$ edges which form a spiral going k times around A, starting at x_1 and ending at v; see Figure 8. To convert the surface S_0 of Lemma 4.2 to the desired surface S, we first subdivide the interval $[x_1, v]$ into $k\ell$ segments by adding vertices $y_1, y_2, \ldots, y_{k\ell-1}$. We remove the triangle $[x_1, v, w_2]$ from S_0 and add $k\ell$ new triangles, obtained by coning the edges of α' to w_2. This construction produces the surface S, which is also an annulus and which also has its vertices either on K or ∂R_K.

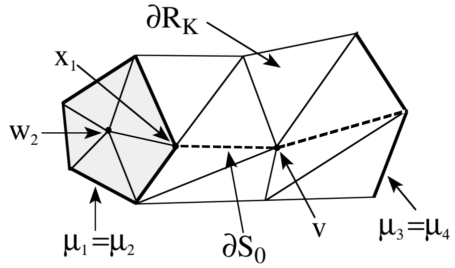

Figure 7: S_0.

It is now straightforward to complete the proof. To prove (i) we observe that the number of triangles in S is at most the number of triangles in R_K, which is at most $2304t$, plus the number $k\ell$ of extra triangles added in the "twisting" step in constructing S from S_0. The trivial bound for ℓ is the number of triangles in M'' which is at most $2304t$. Combining this with (4.6) yields at most

$$2304t + k\ell \leq 2304t + (2304t+1)3^{1152t}(2304t) \leq t^2 \cdot 3^{1152t+20} < 2^{1858t}$$

triangles in S.

To prove (ii) we pull the curve K across S to K_2 one triangle at a time. Using the fact that all vertices of S are in the boundary ∂S, a suitable legal order of triangles to pull across is easily found. Counting two elementary moves

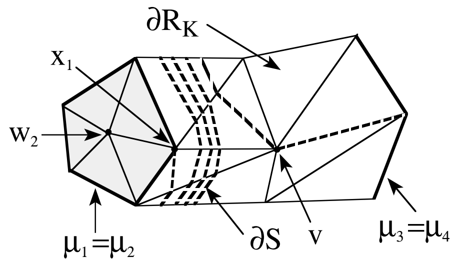

Figure 8: Adding twists replaces S_0 by S.

per triangle, this yields at most $2^{1858t+1}$ elementary moves in total. $\square$

Remark. In the special case that the manifold M is embedded in $\mathbb{R}^3$, as when we start with a standard knot complement, we can improve the bounds (i), (ii) of Theorem 4.1 to be polynomial in n. We sketch the idea. The knot K has at most $2304t$ edges in M. Using the embedding of the manifold M in $\mathbb{R}^3$ we can find a regular projection of the knot to a knot diagram $\mathcal{D}$, and this has $O(n^2)$-crossings, as we show in Lemma 7.1 of §7. We can specify a curve α in the 1-skeleton of ∂R_K which is parallel to the core K whose projection runs just to the right of K in the projection plane. Choose an orientation of K. A longitude of a knot K in $\mathbb{R}^3$ is characterized by having a linking number with K equal to zero. The linking number is easily calculated from the projection. To each crossing of K is associated a sign, and the sum of these signs, called the *writhe* of the knot, gives the linking number of a curve with K. To ensure that the linking number is zero, we alter the construction of α. Near every overcrossing of K, we add a full twist of α around K. This twist is positive if the sign of the crossing is negative, and vice versa. The result is a knot K_2 in $\mathbb{R}^3$ that lies on the peripheral torus whose linking number with K is zero. We can construct the associated annulus S in R_K with boundary components K and K_2 similar to the proof given for Theorem 4.1, and it has $O(n^2)$ triangles. In this argument the key step is that the writhe is bounded in terms of the number of crossings of $\mathcal{D}$, which is $O(n^2)$, and this gives a bound $k = O(n^2)$ in place of (4.6).

5. Isotopies of curves on a surface

In this section we examine isotopies between two PL curves α and β on an oriented triangulated surface F, and bound the number of elementary moves needed to move α to β. In the application, the surface will be the peripheral torus ∂R_K of a knot K, α will be K_1, the boundary of a normal disk, isotopic on ∂R_K to β, and β will be a fixed longitudinal curve K_2, parallel to K in its regular neighborhood R_K. We derive the results of this section for a general surface, since no extra work is needed.

Our main result is the following.

Theorem 5.1 *Let F be a triangulated orientable surface with u triangles, and let α and β be embedded isotopic PL curves on F, each having at most s line segments. Then there is an isotopy carrying β to α via no more than $2^{11}(s+u^2)^4u^3$ elementary moves.*

Corollary 5.1 *Let F be a triangulated orientable surface with at most $2^{10}t$ triangles, and let α and β be embedded isotopic PL curves on F, each having at most $2^{1070209t}$ line segments. Then there is an isotopy carrying β to α via no more than 2^{10^7t-1} elementary moves.*

We will work with specially positioned curves and isotopies, called *basic*, which we now define. Let E be a unit edge length equilateral triangle. Let E' be an equilateral triangle of half the size; see Figure 9. A *basic arc* in E is one of two types:

(a) *Type 1.* An arc running between the interiors of two distinct edges of E. This type of arc contains three straight line segments. The first segment runs from the initial point towards the center of E until it meets E'. The third segment runs from the final point of the arc towards the center of E until it meets E'. The middle line segment connects the two in E'.

(b) *Type 2.* An arc with both of its endpoints on the interior of a single edge of E. This type of arc contains two straight line segments, each of which makes an angle of 1^o with the edge.

Arcs of type 2 are disjoint from any arcs of type 1 with non-linking endpoints, since an arc of type one makes an angle larger than 30^0 with an edge of E, and arcs of type 2 are disjoint from the smaller equilateral triangle.

Now let E' be any triangle and fix a linear map from E to E'. A *basic arc in* E' is defined to be the image of a basic arc in E.

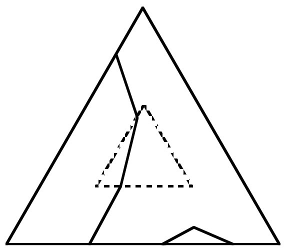

Figure 9: Type 1 and type 2 basic arcs.

A *basic curve* in a triangulated surface is a curve which meets each triangle in a union of basic arcs.

Lemma 5.1 *Any two basic arcs in a triangle with distinct endpoints which lie on the interiors of the edges of E intersect in either zero or one point. Any collection of disjoint embedded curves in a triangle can be isotoped, without moving their endpoints, to a collection of embedded basic arcs.*

Proof. Two basic arcs with distinct endpoints intersect if and only if their endpoints link on the boundary of the triangle. Any family of disjoint embedded arcs in a triangle is isotopic, rel boundary, to any other such family. $\square$

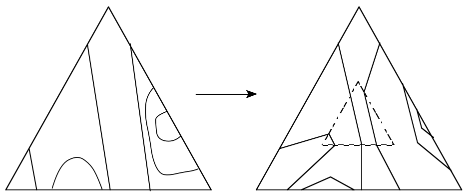

Figure 10: Any set of arcs can be isotoped to a set of basic arcs.

A *basic move* of a basic curve in a triangulated surface is defined to be one of the following curve isotopies, each of which produces a new basic curve:

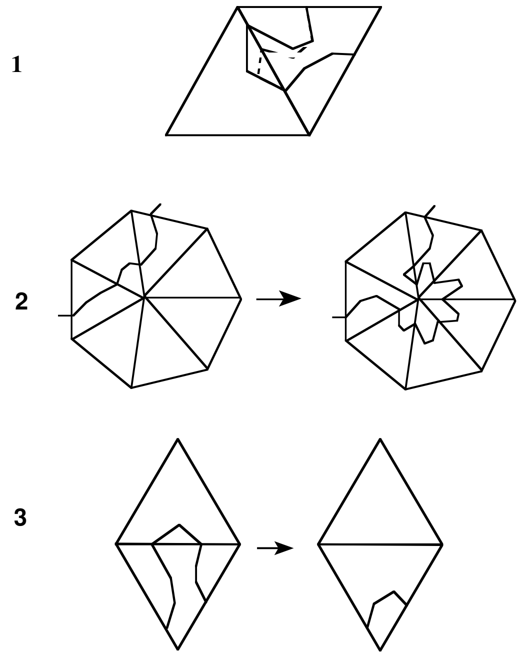

Figure 11: Basic move of types 1, 2 and 3.

(a) *Type 1.* An isotopy that slides an arc of the curve that crosses an edge along that edge.

(b) *Type 2.* An isotopy that isotops an arc of the curve that crosses an edge of the triangulation across an adjacent vertex, and then isotops each of the arcs of intersection with a triangle created or altered in this process to basic arcs.

(c) *Type 3.* An isotopy that slides an arc of the curve with both endpoints on one edge of a triangle E across the 2-gon it forms with that edge. This results in the joining of two arcs in an adjacent triangle E'. If the two arcs give a closed curve in E', the closed curve is removed. Otherwise, this new arc is isotoped in E' to a basic arc.

We count the number of elementary moves required by each basic move. We first look in a single triangle.

Lemma 5.2 *Let $\alpha = \alpha_1 \cup \alpha_2 \cup ... \cup \alpha_n$ and $\beta = \beta_1 \cup \beta_2 \cup ... \cup \beta_n$ each be a union of n connected polygonal arcs in a triangle with the endpoints of α_i and β_i equal, and having a total of a and b edges respectively. Then α can be isotoped to β in the triangle using at most $a + b$ type 2 elementary moves, and $2(a + b)$ elementary moves in total.*

Proof. We will count the number of required elementary moves in two cases. In each case we will isotop both α and β to an intermediate curve γ, which stays close to the boundary of the triangle. To construct γ, we first pick θ to be a positive angle which is small enough so that for any vertex v in $\alpha \cup \beta$ which lies on an edge of the triangle and any vertex w in $\alpha \cup \beta$ which does not lie on that edge, the angle between that edge and the line segment vw is greater than θ.

Given an arc α_i of α, we construct an arc γ_i with the same endpoints, but with interior disjoint from α and containing just two segments. If α_i runs between distinct edges of the triangle then it cuts off a subdisk containing a single vertex of the triangle. Construct γ_i to consist of two segments, each of which makes an angle of θ with the edges going towards that vertex. By definition of θ, γ_i will be disjoint from the interiors of α_i and β_i. If α_i has both endpoints on a single edge of the triangle, then it cuts off a subdisk containing no vertices and a segment of one edge. Construct γ_i to consist of two segments contained in this disk, each of which makes an angle of θ with its edge. Again,

γ_i will be disjoint from the interiors of α_i and β_i.

(1) α and β are connected.

If α does not meet both sides of β, then the region between them can be triangulated with $a+b-2$ triangles, which can be used to isotop α to β using $a+b-2$ type 2 elementary moves. If they intersect, then applying the previous argument shows that each can be isotoped to the disjoint curve γ by using a and b type 2 elementary moves respectively. It follows that we can always isotop α to β using at most $a+b$ type 2 elementary moves, and $2(a+b)$ elementary moves of all types.

(2) α and β are not connected.

We will construct an isotopy from α to β which keeps the curves embedded at each stage. Let a_i be the number of segments of α_i, so that $\sum_{i=1}^{n} a_i = a$. Each arc α_i of α cuts off a disk meeting exactly one vertex of the triangle in which it lies, if it runs between distinct edges of the triangle, or no vertices if it runs from an edge back to that edge. An arc of α is said to be *outermost* if it cuts off a disk which contains no other arcs of α. Let α_i be outermost. Then γ_i is disjoint from all the other arcs in $\alpha \cup \beta$ and meets $\alpha_i \cup \beta_i$ only in their endpoints. The number of segments of γ_i is two, and $\gamma = \bigcup \gamma_i$ is embedded. We can isotop the outermost arc α_i of α to the arc γ_i with the same endpoints without ever introducing any new intersections in α, since the other arcs in α lie on the other side of α_i from γ_i. This isotopy requires at most a_i type 2 elementary moves, as in the previous case. We then repeat for an outermost arc α_j among the collection of arcs of $\alpha - \{\alpha_i\}$. This isotopy takes place between α_j and γ_j, and in particular never meets γ_i. Repeating for all arcs in α, we can isotope the entire collection α to γ using at most $\sum_{i=1}^{n} a_i = a$ type 2 elementary moves. Similarly, we can isotop β to γ using at most b type 2 elementary moves. It follows that we can isotop α to β using at most $a+b$ type 2 elementary moves.

The total number of elementary moves required in these isotopies is at most twice the number of type 2 elementary moves.

□

Define the *length* of a PL curve transverse to the 1-skeleton of a triangulated surface to be the number of edges of the 1-skeleton that it crosses, counted with multiplicity.

Lemma 5.3 *A basic move in a surface with a triangulation of valence at most V can be realized by at most 28 elementary moves for a type 1 basic move, $6V + 24$ elementary moves for a type 2 basic move, and 22 elementary moves for a type 3 basic move. In all cases $17V$ elementary moves suffice.*

Proof. We count the number of elementary moves required for each type of basic move.

Lemma 5.2 implies that a type 1 basic move requires at most 14 elementary moves in each of the two triangles meeting the edge along which the curve is isotoped.

The type 2 basic move can be done in the following sequence of steps.

(1) Two type 2 elementary moves slide the arc in a neighborhood of an edge to pass through the adjacent vertex.

(2) Two type 2 elementary moves slide the curve to run along edges in a neighborhood of the adjacent vertex.

(3) $V - 2$ type 2 elementary moves slide the arc over a neighborhood of the vertex in each adjacent triangle.

(4) At most 7 type 2 elementary moves slide the arc to basic curves in each of the original two triangles, by Lemma 5.2.

(5) Two type 2 elementary moves slide the arc to basic curves in each of the $V - 2$ triangles where new intersections were introduced.

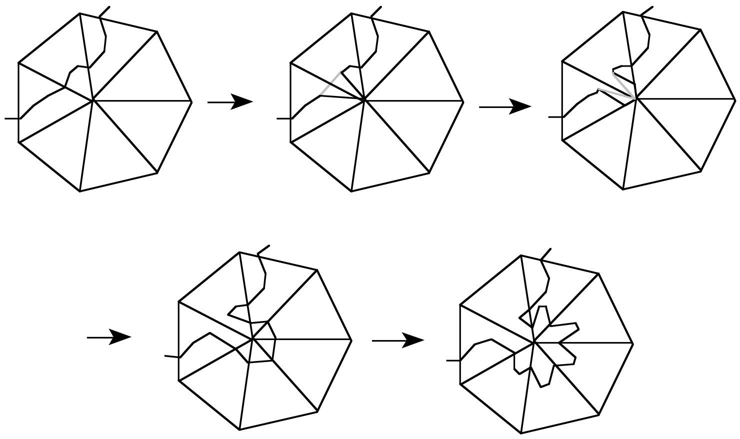

Figure 12: Elementary moves for a type 2 basic move.

Summing, this basic move requires less than $3V + 12$ type 2 elementary moves, and $6V + 24$ in total.

The type three basic move requires

(1) One type 2 elementary move to slide the arc to coincide with a subarc of its adjacent edge.

(2) At most 10 type 2 elementary moves slide the resulting arc to a basic curve with the same endpoints in the adjacent triangle, by an application of Lemma 5.2. This sums to at most 11 elementary moves of type 2 and 22 elementary moves of all types.

Since $V \geq 3$, in all cases $17V$ elementary moves suffice. $\square$

If two basic curves are isotopic using only basic moves of type 1, we say that they are *parallel* basic curves.

Lemma 5.4 *Let α and β be two parallel basic curves with total length ℓ. Then α can be isotoped to coincide with β by $\ell/2$ basic moves of type 1. Similarly, α can be isotoped across β by $\ell/2$ basic moves of type 1.*

Proof. Each of α and β has length $\ell/2$ and they intersect each triangle of the surface in pairs of parallel basic arcs. Using $\ell/2$ basic moves of type 1, we can either move the intersections of α along each edge to coincide with the intersection of β with that edge, or move the intersections past to the other side of β. $\square$

Let α and β be intersecting simple closed curves in general position on a surface F. A *2-gon* is a subdisk of F whose boundary consists of one subarc of each of α and β. A 2-gon is *innermost* if no arc of α or β meets its interior. In some cases, it can be shown that curves with excess intersection create innermost 2-gons. This is analyzed in some generality in [13]. The following is Lemma 3.1 of [13].

Lemma 5.5 *Let F be a compact orientable surface and let α and β be isotopic simple closed curves on F, intersecting transversely. Then either α and β are disjoint or there is an innermost 2-gon on F bounded by one arc from each of α and β.*

Proof. The lift of α to the universal cover of F is a family of disjoint lines $\tilde{\alpha}$, and the lift of β is a similar family $\tilde{\beta}$. Assume α and β intersect. Then a lift α_1 of α and a lift β_1 of β also intersect. The curves α_1 and β_1 are isotopic to disjoint lines by a lift of an isotopy making α and β disjoint. The fundamental group of F acts on the universal cover, and α_1 and β_1 have a common cyclic stabilizer. It follows that they intersect infinitely often. An arc of each between adjacent crossing points cuts off a 2-gon in the universal cover. A finite number of arcs of $\tilde{\alpha} \cup \tilde{\beta}$ meet this 2-gon. Each such arc is disjoint from either $\tilde{\alpha}$ or $\tilde{\beta}$, so that it cuts off a smaller 2-gon inside the previous one. We can pass to an innermost 2-gon whose interior meets no arc of $\tilde{\alpha} \cup \tilde{\beta}$. This 2-gon is disjoint from any of its translates, and must project 1-1 down to F. $\square$

Lemma 5.6 *Let F be an orientable triangulated surface with u triangles, and let α and β be isotopic simple basic curves on F intersecting transversely. Let l be the length of $\alpha \cup \beta$, and let V be the maximal valence of a vertex in F. There is an isotopy carrying β to α via no more than l^4uV basic moves.*

Proof. We treat cases. Note that $V \geq 3, l \geq 4$ and $u \geq 4$.

Case 1: α and β are disjoint.

Then α and β bound a subsurface of F homeomorphic to an annulus. We will perform a series of basic moves which isotop the curves to coincide.

The edges of F induces a cellulation of the annulus. The cells are subdisks of triangles, bounded by arcs in $\alpha \cup \beta$. If there is a vertex in this cellulation, we can slide α or β over a vertex by a basic move of type 2. Repeating at most u times, we eliminate all vertices in the annulus. The length of $\alpha \cup \beta$ is increased by at most $V - 2$ each time we do a basic move, so the final curves have total length at most $l + u(V - 2) = l + uV - 2u$.

The annulus between the curves is cut by the edges of F into a collection of cells. Since α and β are disjoint, and each cell boundary contains an edge lying on the 1-skeleton between each edge lying on $\alpha \cup \beta$, all cells are even sided. An Euler characteristic calculation shows that either all cells are 4-gons, or there is a 2-gon cell in the annulus, bounded by a basic arc of type 2 and an arc on an edge of the triangulation.

If there is a cell which is a 2-gon, we can perform a basic move of type 3 which slides an arc of α or β across that 2-gon. This decreases the length of $\alpha \cup \beta$ by two. After repeating at most $(l + uV - 2u)/2$ times, we eliminate all 2-gon cells from the annulus, and arrive at a cellulation in which all cells are 4-gons. When all cells are 4-gons, the two boundary components of the annulus are parallel basic curves, each of length at most $(l + uV - 2u)/2$. Using at most $(l + uV - 2u)/2$ basic moves of type 1, we can make the two curves coincide.

The total number of basic moves required in the three steps above is bounded by

$$u + (l + uV - 2u)/2 + (l + uV - 2u)/2 = l + uv.$$

Case 2: α and β intersect.

Two basic arcs in a triangle with distinct endpoints intersect at most once, so the total number of intersections of α and β is at most $l^2/4$. Lemma 5.5 gives the existence of an innermost disk D_1 on the surface, bounded by subarcs

α_1 of α and β_1 of β. The 1-skeleton of F induces a cellulation of D_1. The cells are bounded by arcs in $\alpha_1 \cup \beta_1$ and arcs in the 1-skeleton of F. If D_1 contains a vertex of F, we can isotop α_1 or β_1 over a vertex (possibly a different one) by a basic move of type 2, keeping $\alpha \cup \beta$ embedded. The number of vertices is bounded above by u, so repeating at most u times, we eliminate all vertices in D_1, using at most u basic moves of type 2. The length of $\alpha_1 \cup \beta_1$ is increased by at most $V-2$ by each basic move, so α and β are isotoped to curves with total length less than $l + u(V-2) = l + uV - 2u$.

Since α_1 and β_1 meet at exactly two points in the boundary of D_1, there are at most two cells in D_1 with an odd number of sides. An Euler characteristic calculation shows that one of three cases occurs:

Case 1. All of D_1 is contained in a single triangle, or

Case 2. There is a 2-gon cell in D_1, or

Case 3. All cells of D_1 are 4-gons, other then two 3-gons containing points of $\alpha_1 \cap \beta_1$.

The first case is impossible because basic arcs don't intersect in more than one point in a single triangle. A 2-gon cell in D_1 gives a basic arc of type 2. We can use this to perform a basic move of type 3, which slides an arc of α_1 or β_1 across a 2-gon. This decreases the length of $\alpha_1 \cup \beta_1$ by two, and thus can be repeated at most $(l + uV - 2u)/2$ times. This leaves us in Case 3. We can then isotop β_1 across α_1, using at most $(l + uV - 2u)/2$ basic moves of type 1. We have then eliminated the innermost 2-gon D_1, reducing the number of intersection points of α and β by two, while using at most $l + uV$ basic moves.

The length of $\alpha \cup \beta$ is now bounded by the larger value $l + uV$. If α and β still intersect, we repeat with another innermost disk. Eliminating this disk requires at most $(l + uV) + uV = l + 2uV$ basic moves. Repeating at most $l^2/4$ times we eliminate all intersections and arrive at disjoint curves. Summing the resulting series, which has at most $l^2/4$ terms, gives a bound of

$$l + uv + (l + 2uv) \ldots + (l + (l^2/4)uv) \leq l^3/4 + \frac{l^2}{4}(\frac{l^2}{4} - 1)uv/2 \leq l^3/4 + l^4uV/32$$

basic moves. With the curves now disjoint, and having length less than $l + \frac{l^2}{4}uV$, case 1 isotops them to coincide using at most $l+\frac{l^2}{4}uV+uV$ basic moves. The total number of basic moves used is less than

$$l^3/4+l^4uV/32+l+\frac{l^2}{4}uV+uV.$$

Since $l\geq 4$ and $u\geq 4$,

$$l^4uV \geq l^4uV/32{+}2l^3uV \geq l^4uV/32{+}l^3/4{+}4l^2uv \geq l^4uV/32{+}l^3/4{+}l{+}l^2uV/4{+}uv$$

so the sum is less than l^4uV as claimed. $\square$

Lemma 5.7 *Let F be an orientable triangulated surface with u triangles, and let α and β be isotopic basic curves on F in general position. Let l be the total length of $\alpha\cup\beta$. Then there is an isotopy carrying β to α via no more than $17l^4u^3$ elementary moves.*

Proof. The required number of basic moves is at most l^4uV, and each basic move requires at most $17V$ elementary moves, so the total number of elementary moves required is less than $17l^4uV^2$. We know $V\leq u$ so that $17l^4u^3$ gives a bound on the number of elementary moves. $\square$

We now give the proof of the main result of this section.

Proof of Theorem 5.1: We begin by isotoping α to a basic curve. First we perturb α to miss the zero-skeleton of the triangulation. There are at most u vertices, and each has at most V edges coming into it, where V is the maximal valence of the triangulation of the surface. We can perturb α near any vertex which it passes through so that it crosses at most $V/2$ edges near that vertex. This requires a total of at most $u(V/2+1)$ type 2 elementary moves, and $uV+2$ elementary moves of both types. The number of segments on the curve has now been increased to at most $s+uV/2+u<s+uV$.

Next we isotop the intersection of α with each triangle to a union of basic arcs. This can increase the number of segments by at most a factor of three, to $3(s+uV)$ and requires less than $4(s+UV)$ elementary moves by Lemma 5.2. We can do similar moves on β, resulting in new curves α and β which are transversely intersecting basic curves with a total of at most $6(s+uV)$ segments. Note that the length of a basic curve is between one third and one half of the number of segments it contains, therefore $\alpha \cup \beta$ has length at most $3(s+uV)$. We now apply Lemma 5.7, isotoping the curves together using at most $17[3(s+uV)]^4u^3$ elementary moves. We can bound V by u, obtaining a bound of at most

$$u^2 + 2 + 4(s+u^2) + 1377[(s+u^2)]^4u^3 < 2^{11}(s+u^2)^4u^3$$

elementary moves. $\square$

6. Proof of Theorem 1.2

We combine the results of §2–§5 to prove Theorem 1.2.

Proof of Theorem 1.2: To prove Theorem 1.2, we first construct a compact triangulated PL submanifold M_K of M obtained by removing a solid torus neighborhood of K. We barycentrically subdivide M twice and construct a new triangulated manifold $M_K = M - R_K$, by removing the interior of a regular neighborhood R_K of K. The boundary of M_K consists of a subdivision of the original boundary ∂M plus a 2-torus ∂R_K, the peripheral torus. The manifolds M_K and R_K contain between them $576t$ tetrahedra. As outlined in §3, we deform the original knot K across an annulus by elementary moves to a longitude K_1 on the peripheral 2-torus ∂R_K. Then we deform K_1 on the peripheral torus by elementary moves to the boundary K_2 of the embedded normal disk given by Haken's Theorem 3.1. Finally we deform the curve K_2 across this disk using elementary moves to a triangle.

We count elementary moves. By Lemma 3.4 the normal disk with K_2 as boundary contains at most $2^{8\cdot 576t+6} = 2^{4608t+6}$ triangles, and the number of edges $|K_2|$ of K_2 is at most

$$|K_2| \leq 2^{4608t+7} < 2^{4609t} = s_2 \ . \tag{6.1}$$

Also by Lemma 3.4 it takes at most $2^{4608t+7}$ elementary moves to transform K_2 to a triangle.

By Theorem 4.1 the longitude K_1 consists of at most

$$|K_1| \leq 2^{1858(576t)+1} < 2^{1070209t} = s_1 \tag{6.2}$$

edges, and the isotopy from K to K_1 requires at most the same number of elementary moves.

It remains to bound the number of elementary moves needed to carry K_1 to K_2, using Theorem 5.1. We note that K_1 is isotopic to K_2 if properly oriented. Theorem 5.1 applies with $s = \max(s_1, s_2) = s_1$ to yield a bound of at most $2^{11}(s + u^2)^4 u^3$ elementary moves taking K_1 to K_2, where u is the number of triangles on the peripheral torus. Since $u \leq 576t < 2^{10}t$, Corollary 5.1 yields a bound of $2^{10^7 t - 1}$ for the number of elementary moves taking K_1 to K_2. For the three steps together we obtain a bound of $2^{4609t} + 2^{1070209t} + 2^{10^7 t - 1} < 2^{10^7 t}$. $\square$

Remark. The assumption $K \subseteq Interior(M)$ was used to allow the peripheral torus R_K to be constructible by two barycentric subdivisions and to have ∂R_K disjoint from ∂M. For the general case that $K \subseteq M$, we could proceed by barycentrically subdividing M twice, then isotoping K off the boundary ∂M to a "parallel" knot K'' contained strictly in the interior of M'', using $O(t)$ elementary moves. Now Theorem 1.2 applies to K'' in $(M'', \partial M'')$.

7. Reidemeister moves and elementary moves in $\mathbb{R}^3$

The remainder of the paper is concerned with 3-manifolds embedded in $\mathbb{R}^3$, and is devoted to proving the Reidemeister move bound of Theorem 1.1.

We outline the proof of Theorem 1.1. We have given a knot diagram $\mathcal{D}$ with n vertices. We use it to construct a triangulated PL manifold M embedded in $\mathbb{R}^3$ which contains a knot K in its 1-skeleton whose projection on the xy-plane in $\mathbb{R}^3$ is the knot diagram $\mathcal{D}$. In section 8 we show that such a manifold M can be constructed using $O(n)$ tetrahedra. Now Theorem 1.2 shows that we can isotop K to a trivial knot using $2^{c_3 n}$ elementary moves. We project these elementary moves down onto the xy-plane and obtain a sequence of Reidemeister moves.

The proof is completed by bounding the number of Reidemeister moves in terms of elementary moves, which we do below.

In the rest of this section we describe Reidemeister moves and relate them to elementary moves. We recall basic facts about knot and link diagrams.

A projection of a link in $\mathbb{R}^3$ into a plane is *regular* if all crossings are in general position, and at most 2 edges intersect the pre-image of any point. Corresponding to a regular projection of a link is a *link diagram*. This is a labeled planar graph having the following properties:

(i) Any connected component of $\mathcal{D}$ with no vertices is an (isolated) loop.

(ii) Each vertex has exactly four adjacent edges, with two edges labeled "undercrossing" and two labeled "overcrossing" at the vertex.

(iii) The labeling prescribes a cyclic ordering of edges at each vertex, and "overcrossing edges" are adjacent to "undercrossing edges" in this ordering.

Conversely, for any labeled planar graph $\mathcal{D}$ satisfying (i)–(iii) there exists an unoriented link L having projection $\mathcal{D}$. We note that the cyclic ordering of vertices in (iii) actually specifies a planar embedding of the graph $\mathcal{D}$, since it determines the faces of the planar embedding.

A *link component* in a link diagram $\mathcal{D}$ is obtained by tracing a closed path of edges such that each pair of consecutive edges match at their common vertex, either both overcrossing or both undercrossing. The *order* of $\mathcal{D}$ is the number of components it contains. A *link diagram component* of $\mathcal{D}$ is a connected component of the underlying graph; it can be the union of several link components. The *crossing measure* $|\mathcal{D}|$ of a link diagram is

$$|\mathcal{D}| = \#(\text{vertices}) + \#(\text{link diagram components}) - 1 \ . \tag{7.1}$$

A *knot diagram* is a link diagram that consists of one link component. A *trivial knot diagram* is a knot diagram that has no vertices. If $\mathcal{D}$ is a knot diagram, then $|\mathcal{D}|$ is the *crossing number* of the knot projection.

Lemma 7.1 *If a polygonal link L has regular projection $\mathcal{D}$, then*

$$|\mathcal{D}| \leq |L|^2 \ . \tag{7.2}$$

Proof. The number of vertices $|L|$ is at least as large as the number of distinct line segments it contains. Since line segments project to line segments, each pair of line segments produces at most one crossing. There are less than $|L|^2$ pairs. Furthermore line segments that project into different connected components of $|\mathcal{D}|$ cannot produce any crossings. Thus the number of nonintersecting pairs of line segments is at least $\#(\text{components of } \mathcal{D}) - 1$, so 7.2 follows. □

The bound 7.2 is the correct order of magnitude, because one can easily construct a sequence of knots K with $|K| \to \infty$, which have the property that $|\mathcal{D}| > \frac{1}{2}|K|^2$. There is no upper bound for $|L|$ in terms of $|\mathcal{D}|$, but there does exist a link L with a projection having diagram $\mathcal{D}$, with $|L| \leq c|\mathcal{D}|$, see [11, Lemma 5.1] and §3.

Reidemeister moves are local moves on link diagrams that are analogous to elementary moves. They convert a link diagram $\mathcal{D}$ to another link diagram $\mathcal{D}'$ by a local change in its structure. For unoriented links there are four kinds of Reidemeister moves, examples of which are pictured in Figure 13, see one of [20, p. 48], [1],[6],[],[24]. (For oriented links there are more Reidemeister moves, see Wu [32].)

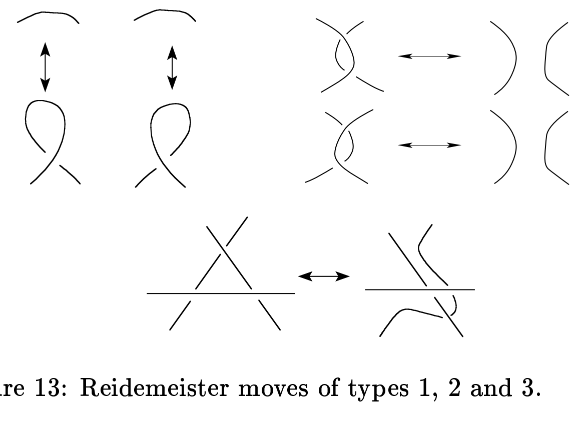

Figure 13: Reidemeister moves of types 1, 2 and 3.

We call two knot diagrams $\mathcal{D}_1$ and $\mathcal{D}_2$ *equivalent* if and only if one can be converted into the other by a finite sequence of Reidemeister moves. This relation corresponds to knot equivalence: Knots K_1 and K_2 have knot diagrams $\mathcal{D}_1$ and $\mathcal{D}_2$ that are equivalent if and only if $\mathcal{D}_1$ and $\mathcal{D}_2$ are equivalent.

Reidemeister moves on a link diagram $\mathcal{D}$ are related to elementary moves on a link L whose projection is $\mathcal{D}$, but the correspondence of moves is not one-to-one.

Lemma 7.2 *Let L be a polygonal link and suppose that it has diagram $\mathcal{D}$ as regular projection on a plane. Let L' be a link derived from L by an elementary move of type (2) or $(2')$, and suppose that it has regular projection $\mathcal{D}'$ on the same plane. Then there is a sequence of at most $2|L|+2|\mathcal{D}|$ Reidemeister moves that converts $\mathcal{D}$ to $\mathcal{D}'$.*

Proof. As a line segment is moved affinely in $\mathbb{R}^3$ with one end fixed, its projection intersects the projection of any vertex of the polygonal link L in at most one point. Therefore it can generate at most one type I or type II Reidemeister move involving the projection of that vertex. The number of vertices is bounded by $|L|$. Similarly, each pair of crossing segments in $\mathcal{D}$ can produce at most one type III Reidemeister move. Since two line segments are moved in a type (2) or $(2')$ elementary move, this yields at most $2|L| + 2|\mathcal{D}|$ Reidemeister moves. $\square$

The bound of Lemma 7.2 is nearly best possible.

We use these lemmas to prove a bound for Reidemeister moves in terms of elementary moves.

Theorem 7.1 *Let L and L' be polygonal links each consisting of at most n line segments and suppose that $\mathcal{D}$ and $\mathcal{D}'$ are link diagrams obtained by regular projection of L and L' on a fixed plane. If L can be transformed to L' using k elementary moves, then $\mathcal{D}$ can be transformed to $\mathcal{D}'$ using at most $2k(n+\frac{1}{2}k+1)^2$ Reidemeister moves.*

Proof. Let $L = L_0, L_1, L_2, \ldots, L_k = L'$ be the set of intermediate links occurring in the sequence of elementary moves, and let $\mathcal{D}_i$ denote the projection of L_i on the fixed plane. We may assume that all projections are regular, by applying a a transversality argument and infinitesimally moving the plane of projection [6, Prop. 1.12]. Since an elementary move changes the number of vertices by at most one, and $|L_0|, |L_k| \leq n$, we have, for $1 \leq i \leq k$, that

$$|L_i| \leq \min(n+i, n+k-i) \leq n + \frac{1}{2}k \ . \tag{7.3}$$

Lemma 7.1 gives

$$|\mathcal{D}_i| \leq |L_i|^2 \leq (n + \frac{1}{2}k)^2 = n^2 + nk + \frac{k^2}{4} \ . \tag{7.4}$$

Thus we obtain

$$2|L_i| + 2|\mathcal{D}_i| \leq 2(n + \frac{1}{2}k + 1)^2 \ ,$$

and applying Lemma 7.2 to each L_i separately yields the upper bound $2k(n + \frac{1}{2}k + 1)^2$. $\square$

8. Triangulation of knot complements in $\mathbb{R}^3$

In this section we construct a triangulated 3-manifold M_K which is a knot complement in a ball embedded in $\mathbb{R}^3$. That is, $M_K = S^3 - (R_K \cup B)$, in which R_K is an open regular neighborhood of the knot K, B is an open ball containing the point at infinity in its interior, and $\bar{R}_K \cap \bar{B} = \emptyset$. The manifold M_K has boundary $\partial M_K = \partial R_K \cup \partial B$ which consists of two connected components, which are topologically a 2-sphere and a torus. As in §2, we call such a manifold M_K a *truncated knot complement.*

We carry out this construction more generally for link diagrams, as it requires no extra work. Given a link diagram $\mathcal{L}$ with crossing measure n, we construct a triangulated manifold M in $\mathbb{R}^3$ which contains $O(n)$ tetrahedra and has a link L in its 1-skeleton which has a regular projection that gives the link diagram $\mathcal{L}$.

Theorem 8.1 *Given a link diagram $\mathcal{L}$ of crossing measure n, one can construct a triangulated convex polyhedron P in $\mathbb{R}^3$ such that:*

(i) The triangulation contains at most $840n$ tetrahedra.

(ii) Every vertex in the triangulation is a lattice point $(x, y, z) \in \mathbb{Z}^3$*, with* $0 \le x \le 30n$*,* $0 \le y < 30n$ *and* $-6 \le z \le 6$*.*

(iii) There is a link $\mathcal{L}$ *embedded in the 1-skeleton of the triangulation which lies entirely in the interior of* P*, and whose orthogonal projection on the plane* $z = 0$ *is regular and is a link diagram isomorphic to* $\mathcal{L}$*.*

Proof. The link diagram $\mathcal{L}$ comes with a topological planar embedding which determines the faces. We begin by adding extra vertices to $\mathcal{L}$. To each edge or non-isolated loop we add two new vertices of degree 2, which splits it into three edges. To each isolated loop we add three new vertices of degree 2, making it an isolated triangle. The resulting labeled graph $\mathcal{G}$ is still planar, has no loops or multiple edges, and has at most $5n$ vertices. (The worst case consists of several disjoint single crossing projections.) Let m denote the number of vertices of $\mathcal{G}$, and call the $m - n$ vertices added *special vertices.* Note that each vertex of the original graph has associated to it four distinct special vertices.

Using the topological planar embedding of $\mathcal{G}$, we next add extra edges to triangulate each face of $\mathcal{G}$, obtaining a topologically embedded triangulated planar graph $\mathcal{G}'$. The graph $\mathcal{G}'$ has m vertices and $2m - 4$ bounded triangular faces, and the unbounded face is also a triangle. We next encase the graph $\mathcal{G}'$ in a triangulated planar graph $\mathcal{H}'$ with $m + 3$ vertices and $2m + 2$ bounded triangular faces, by framing the outside triangular face with a larger triangle and then adding 6 edges to subdivide the resulting concentric region into 6 triangles, obtaining $\mathcal{G}''$. See Figure 14.

Our next goal is to replace the topological embedding of $\mathcal{G}''$ in the plane with a straight-line embedding. In 1952 Fary showed that straight-line embeddings always exist, and a recent result of de Frajsseix, Pach and Pollack [7] states that the vertices can be chosen to lie in a small rectangular grid of integer lattice points. Their results imply that there exists a planar embedding of $\mathcal{H}'$ whose vertices $V(\mathcal{H}')$ lie in the plane $z = -1$, and are contained in the grid

$$\{(x, y, -1) : 0 \le x, y \le 10n - 1 \ ; \quad x, y \in \mathbb{Z}\} \ ,$$

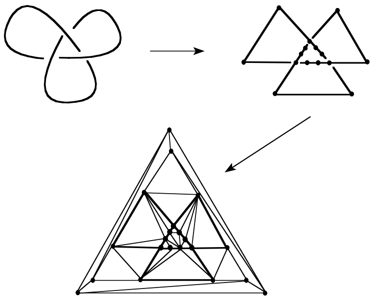

Figure 14: Triangulating a knot projection.

and all edges of the graph $\mathcal{H}'$ are straight line segments.

Next we make an identical copy $\mathcal{H}''$ of $\mathcal{H}'$ on the plane $z = 1$ with $\mathcal{H}'' = \mathcal{H}' + (0, 0, 2)$, and vertex set $V(\mathcal{H}'') = V(\mathcal{H}') + (0, 0, 2)$. We now consider the polyhedron P' which is the convex hull of $V(\mathcal{H}')$ and $V(\mathcal{H}'')$. It is a triangular prism, because the outside face of $\mathcal{H}'$ is a triangle. We add m_1 vertical edges connecting each vertex $\mathbf{v} \in V(\mathcal{H}')$ to its copy $\mathbf{v}+(0, 0, 1) \in V(\mathcal{H}'')$. Let $\mathcal{E}$ denote these edges, together with all the edges in $\mathcal{H}'$ and $\mathcal{H}''$. Using these edges, the polyhedron P' decomposes into $2m+2$ triangular prisms $\{Q_j : 1 \leq j \leq 2m-4\}$, with the top and bottom faces of each being congruent triangular faces of $\mathcal{H}'$ and $\mathcal{H}''$.

We next triangulate P' by dissecting each of the $2m + 2$ triangular prisms into 14 tetrahedra, as follows. We subdivide each vertical rectangular face into four triangles using its diagonals. Then we cone each rectangular face to the centroid of P', and note that the centroid lies in the plane $z = 0$. We add 4 new vertices to each prism, one on each rectangular face and one in the center. The point of this subdivision is that the triangulations of adjacent prisms are compatible. Note also that all new vertices added lie in the lattice $\frac{1}{3}\mathbb{Z}^2$.

Let $\mathcal{E}'$ denote all the edges in the union of these triangulated prisms. We identify the link diagram $\mathcal{G}$ with the graph $\mathcal{G}$ embedded in $\mathcal{H}'$. We next observe that there is a polygonal link L embedded in the edge set $\mathcal{E}'$ whose projection is the link diagram $\mathcal{L}$. We insist that any edge in $\mathcal{D}$ that runs to an undercrossing have its vertex which is an undercrossing be in the plane $z = -1$, while any edge that runs to an overcrossing have the overcrossing vertex in the plane $z = 1$. Each such edge goes from one of the n original vertices of $\mathcal{L}$ to a special vertex adjacent to that vertex. Edges that do not meet vertices labeled overcrossing or undercrossing are assigned to the $z = -1$ plane. The edge corresponding to an edge running from the $z = -1$ to the $z = 1$ plane is contained in one of the diagonals added to a prism. The resulting embedding L in P' has a regular orthogonal projection onto the (x, y)-plane, since no edges are vertical and only transverse double points occur as singularities in the projection. Note that no edge of L touches the outer triangle of $\mathcal{H}'$ because all edges project to the graph $\mathcal{G}'$ contained strictly inside $\mathcal{H}'$. This will be important for property (iii).

The knot edges lie in the upper and lower boundary of P', but do not touch its sides. To have them not touch the boundary we take two additional copies of P' and glue one to its top, along the plane $z = 1$, and one to its bottom, along the plane $z = -1$. The knot now lies entirely in the interior of the resulting polyhedron P''. All the vertices added in this construction lie in $\frac{1}{3}\mathbb{Z}^2$. The total number of tetrahedra used in triangulating P'' is $84(m + 1)$, which is at most $420(n + 1)$, or at most $840n$ for $n \geq 1$.

Now triple all coordinates $(x, y, z) \to (3x, 3y, 3z)$ to obtain a triangulated polytope $P = 3P''$ with integer vertices, with a triangulation that satisfies (i)–(iii). $\square$

Remark. (1). The construction in Theorem 8.1 can be effectively computed in time $O(n \log n)$ using the Hopcroft-Tarjan planarity algorithms [15] and the embedding algorithm of de Frajsseix, Pach and Pollack [7, Section 4].

(2). One can ask whether a polygonal knot K in $\mathbb{R}^3$ that consists of n line segments can be embedded in the 1-skeleton of some triangulated polyhedron that uses $O(n)$ tetrahedra. We do not know if this can always be done. There exist polygonal knots whose planar projections in any direction have at least $\Omega(n^2)$ crossings. Avis and El Gindy [3] showed that a set of n points in general position in $\mathbb{R}^3$ can be embedded as vertices of a triangulation of their convex hull using $O(n)$ tetrahedra, but that in the general case $\Omega(n^2)$ tetrahedra are sometimes needed.

9. Proof of Theorem 1.1

We combine the earlier results to prove Theorem 1.1 with an explicit bound on the number of Reidemeister moves required to isotop an unknot to the trivial projection.

Proof of Theorem 1.1. Applying Theorem 8.1 to the knot diagram $\mathcal{D}$ yields a triangulated convex polyhedron M embedded in $\mathbb{R}^3$ with t tetrahedra, where,

$$t \leq 840n \ ,$$

and with a knot K embedded in the 1-skeleton of the interior of M that projects to $\mathcal{D}$. The hypotheses of Theorem 1.2 are satisfied for M, hence there is a sequence of at most $2^{c_2 t}$ elementary moves transforming K to the trivial knot in M, with $c_2 = 10^7$. Now Theorem 7.1 gives a bound of at most

$$2^{c_2 t}\left(6 \cdot 840n + \frac{1}{2} \cdot 2^{c_2 t} + 1\right)^2 \ . \tag{9.1}$$

Reidemeister moves, since K contains at most $6 \cdot 840n$ line segments.

The bound (9.1) yields the desired bound of the form $2^{c_1 n}$ for an explicit constant c_1. Using the value $c_2 = 10^7$ from the proof of Theorem 1.2, one finds that $c_1 = 10^{11}$ suffices. $\square$

References

[1] C. Adams, *The Knot Book. An elementary introduction to the mathematical theory of knots.* W. H. Freeman Co., New York 1994.

[2] J. W. Alexander and G. B. Briggs, On types of knotted curves, Ann. Math. **28** (1926/27) 562–586.

[3] D. Avis and H. El Gindy, Triangulating Point Sets in Space, Disc. & Comp. Geom. **2** (1987), 99–111.

[4] J. Birman and M.D. Hirsch, Recognizing the unknot, preprint, 1997.

[5] J. Birman and W. Menasco, Studying links via closed braids V: The unlink, Trans. AMS **329**, No. 2, Feb. 1992.

[6] G. Burde and H. Zieschang, *Knots*, de Gruyter 1985.

[7] H. de Frajsseix, J. Pach and R. Pollack, How to draw a planar graph on a grid, Combinatorica **10** (1990), 41–51.

[8] L. Goeritz, Bemerkungen zur knotentheorie, Abh. Math. Sem. Univ. Hamburg **10** (1934), 201–210.

[9] W. Haken, Theorie der Normalflachen, ein Isotopie Kriterium für ein Kreis, Acta Math. **105** (1961), 245–375.

[10] J. Hass, Algorithms for recognizing knots and 3-manifolds, to appear in Chaos, Fractals and Solitons.

[11] J. Hass, J. C. Lagarias and N. Pippenger, The computational complexity of knot and link problems, preprint, 1997.

[12] J. Hass, J. C. Lagarias and N. Pippenger, The computational complexity of knot and link problems, preliminary report, Proc. 38th Annual Symposium on Foundations of Computer Science, (1997) 172–181.

[13] J. Hass and G. P. Scott, Intersections of curves on surfaces, Israel Math. J. 51 (1985), 90–120.

[14] G. Hempel, *3-Manifolds,* Princeton University Press 1976.

[15] J. E. Hopcroft and R. E. Tarjan, Efficient planarity testing, J. Assoc. Comp. Mach. **21** (1974), 549–568.

[16] J.F.P. Hudson, *Piecewise-Linear Topology*, W.A. Benjamin Inc. (1969).

[17] W. Jaco and H. Rubinstein, *PL* Equivariant Surgery and Invariant Decompositions of 3-Manifolds, Advances in Math. **73** (1989), 149–191.

[18] W. Jaco and J. L. Tollefson, Algorithms for the complete decomposition of a closed 3-manifold, Ill. J. Math. **39** (1995), 358–406.

[19] H. Kneser, Geschlossene Flachen in dreidimensionalen Mannigfaltigkeiten, Jahresbericht Math. Verein. **28** (1929), 248–260.

[20] K. Murasugi, *Knot Theory and Its Applications*, Birkhauser: Boston 1996.

[21] H. Reidemeister, Knoten und Gruppen, Abh. Math. Sem. Univ. Hamburg **5** (1926), 7–23.

[22] H. Reidemeister, Elementare Begründang der Knotentheorie, Abh. Math. Sem. Univ. Hamburg **5** (1926), 24–32.

[23] H. Reidemeister, *Knotentheorie*, J. Springer: Berlin 1932; Chelsea; New York, 1948. [English translation: *Knot Theory*, BCS Associates: Moscow, Idaho 1984.]

[24] D. Rolfsen, *Knot and Links*, Publish or Perish Inc.: Berkeley 1976.

[25] C.P. Rourke and B.J. Sanderson, *Introduction to Piecewise-Linear Topology*; Springer-Verlag: Berlin Heidelberg, New York 1982.

[26] A. Schrijver, *Theory of Linear and Integer Programming*; John Wiley and Sons: New York 1986.

[27] H. Schubert, Bestimmung der Primfactor zerlegung von Verkettungen, Math. Z. **76** (1961), 116–148.

[28] A. Sebö, Hilbert bases, Caraetheodory's theorem and combinatorial optimization, in: *Integer Programming and Combinatorial Optimization* (R. Kannan and W. R. Pulleybank, Eds.), U. Waterloo Press 1990, pp. 431–455.

[29] D. J. A. Welsh, *Complexity: Knots, Colourings and Counting*, Cambridge University Press: Cambridge 1993.

[30] D. J. A. Welsh, The complexity of knots, in: *Quo Vadis, Graph Theory?* (Ed. J. Gimbel, J. Kennedy, L. V. Quintoo), North-Holland: Amsterdam 1993, pp. 159–173. (Also: Annals Disc. Math. **55** (1993), 159–173.)

[31] D. J. A. Welsh, Knots and braids: some algorithmic questions, in: *Graph structure Theory* (Seattle, WA 1991), Contemporary Math. Vol. 147, AMS: Providence 1993, pp. 109–123.

[32] F. Y. Wu, Knot Theory and Statistical Mechanics, Rev. Mod. Phys. **64** (1992), 1099–1131.

email: `hass@math.ucdavis.edu`
`jcl@research.att.com`

address: Prof. J. Hass
Dept. of Mathematics
Univ. of California, Davis
Davis, CA 95616

Dr. J. C. Lagarias
Room C235, AT&T Labs–Research
180 Park Avenue
Florham Park, NJ 07932-0971