The Computational Complexity of Knot and Link Problems
(Preliminary Version)

Joel Hass
Deapartment of Mathematics
University of California
Davis, CA 95616
hass@math.ucdavis.edu

Jeffrey C. Lagarias
Information Sciences Research
AT&T Labs - Research
Florham Park, New Jersey 07932
jcl@research.att.com

Nicholas Pippenger
Depaartment of Computer Science
University of British Columbia
Vancouver, BC V6T 1Z4 Canada
nicholas@cs.ubc.ca

Abstract

We consider the problem of deciding whether a polygonal knot in 3-dimensional Euclidean space is unknotted (that is, whether it is capable of being continuously deformed without self-intersection so that it lies in a plane). We show that this problem, *UNKNOTTING PROBLEM*, is in **NP**. We also consider the problem, *SPLITTING PROBLEM*, of determining whether two or more such polygons can be split (that is, whether they are capable of being continuously deformed without self-intersection so that they occupy both sides of a plane without intersecting it), and show that it also is in **NP**. Finally, we show that the problem of determining the genus of a polygonal knot (a generalization of the problem of determining whether it is unknotted) is in **PSPACE**.

1. Introduction

The problems dealt with in this paper might reasonably be called "computational topology"; that is, we study classical problems of topology (specifically, the topology of 1-dimensional curves in 3-dimensional space) with the objective of determining their computational complexity. One of the oldest and most fundamental of such problems is that of determining whether a closed curve embedded in space is unknotted (that is, whether it is capable of being continuously deformed without self-intersection so that it lies in a plane). Topologists study this problem at several levels, with varying meanings given to the terms "embedding" and "deformed". The level that seems most appropriate for studying computational questions is that which topologists call "piecewise-linear". At this level, a closed curve is embedded in space as a simple (non-self-intersecting) polygon with finitely many edges. Such an embedding is called a "knot". (Operating at the piecewise-linear level excludes "wild" knots such as those given by polygons with infinitely many edges, but finite total length.) More generally, one may study "links." A link is a finite collection of simple polygons disjointly embedded in 3-dimensional space. The individual polygons are called components of the link and a knot is a link with one component.

A continuous deformation is required to be piecewise-linear; that is, it consists of a finite number of stages, during each of which every vertex of the polygon moves linearly with time. From stage to stage the number of edges in the polygon may increase (by subdivision of edges at the beginning of a stage) or decrease (when cyclically consecutive edges become collinear at the end of a stage). If the polygon remains simple throughout this process, the deformation is called an "isotopy" between the initial and final knots. Knot isotopy defines an equivalence relation, called "equivalence" of knots. It is easy to see that all knots that lie in a single plane are equivalent; knots in this equivalence class are said to be "unknotted" or "trivial" knots.

While it is "intuitively obvious" that there are non-trivial knots, it is not at all obvious how to prove this. Stillwell [31] traces the mathematical notion of knot back to a paper of A.T. Vandermonde in 1771; the first convincing proof of the non-triviality of a knot seems to be due to Max Dehn [5] in 1910.

There are a great many alternative formulations of the notion of knot equivalence. Here are some.

1. One can consider sequences of "elementary moves", which are very simple isotopies that move a single edge across a triangle to the opposite two sides, or vice versa.

2. One can consider "ambient isotopies" that move not only the knot, but also the space in which it is embedded, in a piecewise-linear way.

3. One can consider "homeomorphisms" (continuous bijections that have continuous inverses) that map the space to itself in a piecewise-linear way, are orientation preserving, and send one knot to the other.

One can also study knots or links by looking at their "projections" onto a generic plane. In this way, a knot or link may be represented by a planar graph, called a "knot diagram" or "link diagram", in which all vertices (representing the "crossings" of edges of the polygon) have degree four, and for which an indication is given at each crossing of which edge goes "over" and which edge goes "under". This gives an additional formulation of equivalence:

4. One may consider sequences of "Reidemeister moves", which are simple transformations on the diagram of a knot that leave the equivalence class of the knot unchanged.

For more details on piecewise-linear topology, the various formulations of knot and link equivalence, and many other aspects of knot theory, we recommend the books of Adams [1] and Burde and Zieschang [4].

In order to study the computational complexity of knot and link problems, we must agree on a finite computational representation of a knot or link. There are two natural representations: a polygonal representation in 3-dimensional space, or a link diagram representing a 2-dimensional projection.

A polygonal representation of a link L consists of a set of simple polygons in 3-dimensional space described by listing the vertices of each polygon in order; we assume that these vertices have rational coordinates. We can reduce to the case of integer lattice point vertices by replacing L by a scaled multiple mL for a suitable integer m. This does not change the equivalence class of L. A particularly simple kind of polygonal representation uses only integer lattice points as vertices and edges of unit length, so that the polygon is a closed self-avoiding walk on the integer lattice; a sequence of moves (up, down, north, south, east, west) that traverse the polygon, returning to the starting point without visiting any other point twice. (This formulation was used by Pippenger [22] and Sumners and Whittington [32] to show that "almost all" long self-avoiding polygons are non-trivially knotted.) The size of a polygonal representation L is the number of edges in L; its input length is the number of bits needed to describe its vertices, in binary.

A link diagram $\mathcal{D}$ is a planar graph with some extra labeling for crossings that specifies a (general position) two-dimensional projection of a link. A precise definition is given in section 3. The size of a link diagram is the number of vertices in $\mathcal{D}$ plus the number of isolated loops.

These two representations are polynomial-time equivalent in the following sense. Given a polygonal representation L one can find in polynomial time in its input length a planar projection yielding a link diagram $\mathcal{D}$; if L has n edges then the graph $\mathcal{D}$ has at most $O(n^2)$ vertices. Conversely given a link diagram $\mathcal{D}$ with n vertices and l components, one can compute in time polynomial in $n+l$ a polygonal link L with $O(n+l)$ edges that has integer vertices and input length $O(n+l)$ and which projects in the z-direction onto the link diagram $\mathcal{D}$; see Section 7.

In this paper we consider knots and links as represented by link diagrams and take the crossing number as the measure of input size. We can now formulate the computational problem of recognizing unknotted polygons as follows:

Problem: UNKNOTTING PROBLEM
Instance: A link diagram $\mathcal{D}$.
Question: Is $\mathcal{D}$ a knot diagram that represents the trivial knot?

See Welsh [34]—[36] for more information on this problem. The main result of this paper is the following.

Theorem 1.1. *The* UNKNOTTING PROBLEM *is in* **NP**.

The UNKNOTTING PROBLEM was shown to be decidable by Haken [7]; the result was announced in 1954, and the proof published in 1961. From then until now, we know of no strengthening of Haken's decision procedure to give an explicit complexity bound.

We also study the splittability of links. A link is said to be "splittable" if it can be can be continuously deformed (by a piecewise-linear isotopy) so that one or more curves of the link can be separated from one or more other curves by a plane that does not itself intersect any of the curves. We note that this notion remains unchanged if we replace "plane" by "sphere" in the definition. We formulate the computational problem of recognizing splittable links as follows.

Problem: SPLITTING PROBLEM
Instance: A link diagram $\mathcal{D}$.
Question: Is the link represented by $\mathcal{D}$ splittable?

The SPLITTING PROBLEM was shown to be decidable by Schubert [27] in 1961. We establish the following result.

Theorem 1.2. *The* SPLITTING PROBLEM *is in* **NP**.

Another generalization of the unknotting problem concerns an isotopy invariant of a knot K called the "genus" $g(K)$ of K. This was defined by Seifert [29] in 1935; an informal account of the definition follows. Given a knot K, consider the class $\mathcal{S}(K)$ of all orientable spanning surfaces for K; that is, embedded orientable surfaces that have K as their boundary. Seifert showed that this class is non-empty for any knot K. (We shall assume in this discussion that all surfaces are triangulated and embedded in a piecewise-linear way.) Up to piecewise-linear homeomorphism, an orientable surface is characterized by the number of boundary curves and the number of "handles", which is called the "genus" of the surface. The genus $g(K)$ of the knot K is defined to be the minimum genus of any surface in $\mathcal{S}(K)$. Seifert showed that a trivial knot K is characterized by the condition $g(K) = 0$. This means that a knot is trivial if and only if it has a spanning disk.

The notion of genus gives us a natural generalization of the problem of recognizing unknotted polygons; we formulate the problem of computing the genus as a language-recognition problem in the usual way.

Problem: GENUS PROBLEM
Instance: A link diagram $\mathcal{D}$ and a natural number k.
Question: Does the link diagram $\mathcal{D}$ represent a knot K with $g(K) \leq k$?

Haken [7] observed that his methods also suffice to show the decidability of the GENUS PROBLEM. We establish the following result.

Theorem 1.3. *The* GENUS PROBLEM *is in* **PSPACE**.

2. Historical background

The problem of recognizing whether two knots are equivalent has been one of the motivating problems of knot theory. A great deal of effort has been devoted to a quest for algorithms for recognizing the unknot, beginning with the work of Dehn [5] in 1910. Dehn's idea was to look at the fundamental group of the complement of the knot, for which a finite presentation in terms of generators and relations can easily be obtained from a standard presentation of the knot. Dehn claimed that a knot is trivial if and only if the corresponding group is infinite cyclic. The proof of what is still known as "Dehn's Lemma" had a gap, which remained until filled by Papakyriakopoulos [21] in 1957. A consequence is the criterion that a curve is knotted if and only if the fundamental group of its complement is nonabelian. Dehn also posed the question of deciding whether a finitely presented group is isomorphic to the infinite cyclic group. During the 1950s it was shown that many such decision problems for finitely presented groups (not necessarily arising from knots) are undecidable (see Rabin [23], for example), thus blocking this avenue of progress. (The avenue has been traversed in the reverse direction, however: there are decision procedures for restricted classes of finitely presented groups arising from topology. In particular, computational results for properties of knots that are characterized by properties of the corresponding groups can be interpreted as computational results for knot groups.)

Abstracting somewhat from Dehn's program, we might try to recognize knot triviality by finding an invariant of the knot that (1) can be computed easily and (2) assumes some particular value only for the trivial knot. (Here "invariant" means invariant under isotopy.) Thus Alexander [2] defined in 1928 an invariant $A_K(x)$ (a polynomial in the indeterminate x) of the knot K that can be computed in polynomial time. Unfortunately, it turns out that many non-trivial knots have Alexander polynomial $A_K(x) = 1$, the same as the Alexander polynomial of a trivial knot.

Another invariant that has been investigated with the same hope is the Jones polynomial $J_K(x)$ of a knot K, discovered by Jones [18] in 1985. In this case the complexity bound is less attractive: the Jones polynomial for links (a generalization of the Jones polynomial for knots) is $\#\mathbf{P}$-hard and in $\mathbf{FP}^{\#\mathbf{P}}$ (see Jaeger, Vertigan and Welsh [17]). It is an open question whether trivial knots are characterized by their Jones polynomial. Even this prospect, however, has led Welsh [34] to observe that an affirmative answer to the last open question would yield an algorithm in $\mathbf{P}^{\#\mathbf{P}}$ for recognizing trivial knots, and to add: "By the standards of the existing algorithms, this would be a major advance."

The revolution started by the Jones polynomial has led to the discovery of a great number of new knot and link invariants, including Vassiliev invariants and invariants associated to topological quantum field theories, see Birman [3] and Sawin [25]. The exact ability of these invariants to distinguish knot types has not been determined.

A different approach to the problems of recognizing unknottedness and deciding knot equivalence eventually culminated in decision procedures. This is based on the study of normal surfaces in 3-manifolds (defined in section 3), which was initiated by Kneser [19] in 1929. In the 1950's Haken elaborated the theory of normal surfaces, and in 1961 published his decision procedure for unknottedness. Schubert [27] extended Haken's procedure to decide the link splitting problem and related problems. Haken also outlined an approach via normal surfaces to decide the knot equivalence problem [33]. The final step in this program was completed by Hemion [10] in 1979. This approach actually solves a more general decision problem, concerning a large class of 3-manifolds, now called Haken manifolds, which can be cut into "simpler" pieces along certain surfaces (incompressible surfaces), eventually resulting in a collection of 3-balls. Knot complements are Haken manifolds. It gave a procedure to decide if two Haken manifolds are homeomorphic [14]. Recently Jaco and Tollefson [15] further simplified some of these algorithms.

Apart from these decidability results, there appear to be no explicit complexity bounds, either upper or lower, for any of the three problems that we study. The work of Haken [7] and Schubert [27] predates the currently used framework of complexity classes and hierarchies. Their algorithms were originally presented in a framework (handlebody decompositions) that makes complexity analysis appear difficult, but it was recognized at the time that implementation of their algorithms would require at least exponential time in the best case. More recently Jaco and others reformulated normal surface theory using piecewise linear topology, but did not determine complexity bounds. Other approaches to 3-manifold algorithms include methods related to Thurston's geometrization program for 3-manifolds [8]; these currently have unknown complexity bounds.

Our results are obtained using normal surface theory. Among other things we show that Haken's original approach yields an algorithm which determines if a knot diagram with n crossings is unknotted in time $O(2^{cn^2})$, and that the improved algorithm of Jaco and Tollefson runs in time $O(2^{cn})$, see Theorem 8.1. The complexity class inclusions that we prove require some additional observations.

3. Knots and links

A *knot* is an embedding $f : \mathbf{S}^1 \to \mathbf{R}^3$, although it is usually identified with its image $K = f(\mathbf{S}^1)$. (Thus we are considering *unoriented* knots.) A *link* with k components is a collection of k knots with disjoint images. An equivalent formulation regards a knot as an embedding in the one-point compactification $\mathbf{S}^3$ of $\mathbf{R^3}$, and we will sometimes use this setting.

Two knots K and K' are *ambient isotopic* if there exists a homotopy $h_t : \mathbf{R}^3 \to \mathbf{R}^3$ for $0 \le t \le 1$ such that h_0 is the identity, each h_t is a homeomorphism, and $h_1(K) = K'$. We shall also say in this case that K and K' are *equivalent* knots. A knot or link is *tame* if it is ambient isotopic to a piecewise-linear knot or link, also called a polygonal knot or link. We restrict our attention to tame knots and links. Given this restriction, we can without further loss of generality restrict our attention to piecewise-linear manifolds and maps (see Moise [20]).

A *regular projection* of a knot or link is an orthogonal projection into a plane (say $z = 0$) that contains only finitely many multiple points, each of which is a double point with transverse crossing. Any regular projection of a link gives a *link diagram*, which is an undirected labeled planar graph such that:

1. Connected components with no vertices are loops.
2. Each non-loop edge meets a vertex at each of its two ends, and has a label at each end indicating an overcrossing or undercrossing at that end.
3. Each vertex has exactly four incident edges, two labeled as overcrossings and two labeled as undercrossings, and has a cyclic ordering of the incident edges that alternates overcrossings and undercrossings.

Conversely, every labeled planar graph satisfying these conditions is a link diagram for some link.

Given a link diagram, if we connect the edges across vertices according to the labeling, then the diagram separates into k edge-connected components, where k is the number of components in the link. A *knot diagram* is a link diagram having one component. A *trivial* knot diagram is a single loop with no vertices.

We define the *crossing measure* to be the number of vertices in the diagram, plus the number of connected components in the diagram, minus one. For knot diagrams, the crossing measure is equal to the *crossing number*, which is the number of vertices in the diagram. A trivial knot diagram is the only link diagram with crossing measure zero. All other link diagrams have strictly positive crossing measure. A knot diagram is the *unknot* (or is *unknotted*) if there is a knot K having this diagram that is ambient isotopic to a knot K' having a trivial knot diagram.

4. An unknottedness criterion

Our algorithm to solve the UNKNOTTING PROBLEM, like that of Haken, relies on the following criterion for unknottedness. A knot K is unknotted if and only if there exists a piecewise-linear disk D embedded in $\mathbf{R}^3$ whose boundary ∂D is the knot K. We call such a disk a *spanning disk*. We shall actually use a slightly weaker unknottedness criterion, given in Lemma 4.1 below. It does not deal with with a spanning disk of K, but rather with a spanning disk of another knot K' that is ambient isotopic to K.

Given a knot K, let $\mathcal{T}$ be a finite triangulation of $\mathbf{S}^3$ containing K in its 1-skeleton, where the 3-sphere $\mathbf{S}^3$ is the one-point compactification of $\mathbf{R}^3$, and the "point at infinity" is a vertex of the triangulation. Barycentrically subdivide $\mathcal{T}$ twice to obtain a triangulation $\mathcal{T}''$, and let $M_K = \mathbf{S}^3 - R_K$ denote the compact triangulated 3-manifold with boundary obtained by deleting the open regular neighborhood R_K of K. Here R_K consists of all 0-simplices and open

1-, 2- and 3-simplices whose closure intersects K. The closure $\overline{R}_K$ is a tubular neighborhood of K, that is, a solid torus containing K as its core, and its boundary $\partial R_K = \partial M_K$ is topologically a 2-torus. Each of R_K, M_K and $\partial R_K = \partial M_K$ are triangulated by simplices in $\mathcal{T}''$. We call such a manifold M_K a "knot complement manifold".

We call a triangulation of $M_K = \mathbf{S}^3 - R_K$ as above a *good triangulation* of M_K. Similarly we define a good triangulation of a link complement manifold. For any good triangulation of M_K, the homology group $H_1(\partial M_K, \mathbf{Z}) \approx \mathbf{Z} \oplus \mathbf{Z}$, since ∂M_K is a 2-torus. We take as generator $(1, 0)$ the homology class of a fixed closed oriented boundary ∂B of an essential disk B in $\overline{R}_K$ (a "meridian"), and as generator $(0, 1)$ the homology class of a fixed closed oriented circle in ∂M_K that has algebraic intersection 1 with the meridian and algebraic linking number 0 with K (a "longitude"). (A simple closed curve in ∂R_K whose homology class is the identity in the 3-manifold $\overline{R}_K$ but not in the surface ∂R_K is a meridian. A simple closed curve in ∂R_K whose homology class is the identity in the 3-manifold M_K but not in the surface ∂R_K is a longitude. The homology classes of a meridian and longitude are well-defined up to orientation.)

A surface S with boundary ∂S contained in a 3-manifold M with boundary ∂M is said to be *properly embedded* if it does not intersect itself and if $S \cap \partial M = \partial S$. A surface S is *essential* for M if it is properly embedded in M, cannot be homotoped into ∂M while holding ∂S fixed, and has fundamental group which injects into the fundamental group of M (what topologists call an "incompressible" surface, see Hempel [11]). In particular, a surface of smallest genus with boundary a longitude of a knot is an example of an essential surface.

Lemma 4.1. *Let K be a polygonal knot, and take any good triangulation of M_K.*

1. *If K is knotted, then there exists no essential disk in M_K.*

2. *If K is unknotted, then there exists an essential disk in M_K, and any such essential disk S has (oriented) boundary ∂S in a homology class $[\partial S] = (0, \pm 1)$ in $H_1(\partial M_K, \mathbf{Z})$.*

Roughly speaking, Lemma 4.1 replaces the problem of finding a spanning disk for K with that of finding a spanning disk for a longitude. The condition on $[\partial S]$ has the convenient property that it can be detected by homology with coefficients in $\mathbf{Z}/2\mathbf{Z}$. This will play a crucial role in reducing the complexity of our algorithm from **PSPACE** to **NP**.

Lemma 4.2. *If S is a connected triangulated surface embedded in $\mathbf{R}^3$ with Euler characteristic $\chi(S) = 1$, then S is a topological disk.*

Indeed, the only compact connected surfaces with Euler characteristic 1 are the disk and the projective plane, and the latter cannot be embedded in $\mathbf{R}^3$. This lemma will also play a crucial role in reducing the complexity of our algorithm.

5. Normal surfaces

Let M be a triangulated compact 3-manifold with boundary ∂M. Let t denote the number of tetrahedra (that is, 3-simplices) in the triangulation of M.

A *normal surface* of M (with respect to the given triangulation) is a surface $S \subseteq M$ such that:

1. S is properly embedded in M.

2. The intersection of S with any tetrahedron in the triangulation is a finite disjoint union of triangles and quadrilaterals whose vertices are contained on different edges of the tetrahedron.

(There are some differences in the literature in the definitions concerning normal surfaces. Our usage follows Kneser [19], Jaco and Rubinstein [16], Hemion [10] and Jaco and Tollefson [15]. The definitions used by Haken [7] and Schubert [27] are based on a handlebody decomposition of a 3-manifold, rather than a triangulation. We allow a normal surface to have more than one component, and the individual connected components may be orientable or non-orientable. Some authors require a normal surface to be connected, and refer to what we have defined as a *system of normal surfaces*.)

A normal surface has associated to it combinatorial data that specify the number and type of regions (triangles and quadrilaterals) that appear in the intersection of S with each tetrahedron in the triangulation of M. For a given tetrahedron, each of these triangles or quadrilaterals separates the four vertices into two non-empty sets; there are thus seven possibilities: four types of triangles, which separate one vertex from the other three, and three types of quadrilaterals, which separate two vertices from the other two. If there are t tetrahedra in the triangulation of M, then there are $7t$ pieces of combinatorial data (each a non-negative integer), which specify the number of regions of each of the seven types in each of the t tetrahedra. We represent this combinatorial data as a vector $v = v(S) \in \mathbf{Z}^{7t}$ by choosing a fixed ordering of the region types and tetrahedra. We call $v(S)$ the *normal coordinates* of S.

When is a vector $v \in \mathbf{Z}^{7t}$ the normal coordinates for some normal surface? We shall call such a vector an *admissible* vector. Admissible vectors satisfy the following conditions:

1. *Non-negativity conditions:* Each component v_i of v (for $1 \leq i \leq 7t$) satisfies $v_i \geq 0$.

2. *Matching conditions:* Suppose two tetrahedra T and T' in the triangulation have a common face F. Each region type in T and T' produces either zero or one edge in F, which intersects a given two of the three sides of F. For each pair of sides of F, the number of edges coming from regions in T must equal that coming from regions in T'. These conditions each have the form $v_a + v_b = v_c + v_d$.

3. *Quadrilateral conditions:* In each tetrahedron of the triangulation at most one of the three types of quadrilaterals can occur. (If two quadrilaterals of different types occurred in some tetrahedron, they would intersect, contradicting the condition that a normal surface must be properly embedded.)

Haken [7] (Hauptsatz 2) proved that these necessary conditions for a vector to be admissible are also sufficient:

Theorem 5.1. *Let M be a triangulated compact 3-manifold with boundary, comprising t tetrahedra. Any integer vector $v \in \mathbf{Z}^{7t}$ that satisfies the non-negativity conditions, matching conditions and quadrilateral conditions gives the normal coordinates $v(S)$ of some normal surface S in M, which is unique up to ambient isotopy.*

This result characterizes the set $\mathcal{W}_M$ of all admissible vectors of normal surfaces as a certain set of integer points in a rational polyhedral cone in $\mathbf{R}^{7t}$. We define the *Haken normal cone* $\mathcal{C}_M$ to be the polyhedral cone in $\mathbf{R}^{7t}$ defined by the non-negativity conditions and matching conditions. The points in $\mathcal{W}_M$ are then just the points in the Haken normal cone $\mathcal{C}_M$ that satisfy the quadrilateral conditions.

The usefulness of normal surfaces is that any surface on a 3-manifold M can be simplified by ambient isotopies and "compressions" (removing a handle - a kind of surgery on the the surface) to an incompressible normal surface. In particular this applies to essential surfaces.

The "simplest" normal surfaces are surfaces S such that $v(S)$ cannot be written as $v(S') + v(S'')$ for any non-empty normal surfaces S' and S''. Haken calls these surfaces *fundamental surfaces*, and the corresponding vectors $v(S)$ *fundamental solutions*. Fundamental surfaces are always connected, since otherwise their vectors would be a sum of the vectors of their corresponding components. Such vectors are in the *minimal Hilbert basis* of the cone $\mathcal{C}_M$, which is a finite set (see Schrijver [26] or Sebö [28]). Haken [7] (Chapter 5) proved the following result.

Theorem 5.2. *Let M be a triangulated compact 3-manifold M with non-empty boundary ∂M.*

(1) If M is irreducible, then any essential surface S in M is ambient isotopic in $(M, \partial M)$ to an essential normal surface.

(2) If M contains an essential normal surface S then it contains an essential normal surface S' that is a fundamental surface, such that the entries of $v(S')$ are componentwise less than or equal to those of $v(S)$.

Any knot complement manifold M_K is irreducible (see [13]), and a surface of smallest genus in the class of surfaces whose boundary is a longitude of ∂M_K is an essential surface in M_K. Theorem 5.2 implies that there is a fundamental surface that is such a surface of minimal genus for the knot K; if K is unknotted this surface is an essential disk.

A normal surface S in M is a *vertex surface*, and the corresponding vector $v(S)$ is a *vertex solution*, if $v(S)$ lies on an extremal ray of the Haken normal cone $\mathcal{C}_M$. The notion of a vertex surface was introduced by Jaco and Oertel [14]. A vertex surface is *minimal* if it is a fundamental surface. Jaco and Tollefson [15] (Corollary 6.4) recently obtained the following strengthening of Theorem 5.2 in the case that there is an essential disk.

Theorem 5.3. *If a triangulated compact 3-manifold M with non-empty boundary ∂M contains an essential disk, then it contains such a disk that is a minimal vertex surface.*

The key advantage of this theorem over Theorem 5.2 is that it is possible to test in polynomial time whether a solution to the non-negativity conditions and matching conditions is a vertex solution, just by verifying that non-negativity conditions that are "tight" determine the solution.

6. Bounds for fundamental solutions and Hilbert bases

We bound the number and size of fundamental solutions in the Haken normal cone $\mathcal{C}_M$ of an arbitrary triangulated compact 3-manifold M with boundary ∂M that contains t tetrahedra. The system of linear inequalities and equations defining the Haken normal cone $\mathcal{C}_M$ has the form:

$$v_i \geq 0,$$

where i runs from 1 to $7t$, and

$$v_{a_j} + v_{b_j} = v_{c_j} + v_{d_j},$$

where j runs from 1 to some limit that is at most $6t$.

Lemma 6.1. *Let M be a triangulated compact 3-manifold, possibly with boundary, that contains t tetrahedra in the triangulation.*

(1) Any minimal vertex solution $v \in \mathbf{Z}^{7t}$ of the Haken normal cone $\mathcal{C}_M$ has

$$\max_{1 \le i \le 7t} v_i \le 2^{7t-1}.$$

(2) Any minimal Hilbert basis element $v \in \mathbf{Z}^{7t}$ of the Haken normal cone $\mathcal{C}_M$ has

$$\max_{1 \le i \le 7t} v_i \le t2^{7t+2} - 1.$$

Proof sketch: Assertion (1) uses Hadamard's inequality to bound the determinants in an application of Cramer's rule to the equations that determine an extreme ray. Assertion (2) follows easily from assertion (1),using a standard bound, see Sebö [28], (Theorem 1.1).

This lemma give a bound on the "complexity" of a spanning disk when one exists. Specifically, it shows that for a diagram of the unknot with n crossings, there exists a triangulated spanning disk with at most 2^{cn} triangles, for some constant c. Hass and Lagarias [9] have used this bound to show that such an unknot diagram can be transformed to the trivial knot diagram with at most $2^{c'n}$ Reidemeister moves, for some explicitly given constant c'. Snoeyink [30] has announced that there exist polygons with n sides for which any triangulated spanning disk must have at least $2^{c''n}$ triangles, for some constant c''; as of this writing, however, a gap remains in the proof of this claim.

Lemma 6.2. *(1) The Haken normal cone $\mathcal{C}_M$ has at most 2^{7t} vertex fundamental solutions.*

(2) The Haken normal cone $\mathcal{C}_M$ has at most $t^{7t}2^{49t^2+14t}$ elements in its minimal Hilbert basis.

Proof sketch: This is an easy counting argument using Lemma 6.1 for (2).

7. Triangulations

Given a link diagram $\mathcal{D}$, we show how to construct a triangulated 3-manifold $M_L \cong \mathbf{S}^3 - R_L$, where $\overline{R}_L$ is a regular neighborhood of a link L which has a regular projection that is the link diagram $\mathcal{D}$. The construction takes time that is polynomial in the crossing measure of $\mathcal{D}$, and the triangulations of M_L and $\overline{R}_L$ each contain $O(n)$ tetrahedra.

Lemma 7.1. *Given a link diagram $\mathcal{D}$ of crossing measure n, one can construct in time $O(n \log n)$ a link L in $\mathbf{R}^3$ having regular projection $\mathcal{D}$ in the z-direction and a triangulated 3-manifold $M_L \cong \mathbf{S}^3 - R_L$ which has a good triangulation containing at most $O(n)$ tetrahedra. Furthermore the triangulation of ∂M_L is supplied with marked sets of edges for a meridian on each 2-torus component of ∂M_L, and a marked set of edges for an arc joining each pair of 2-torus components of ∂M_L.*

Proof sketch: We first construct a link L in $\mathbf{R}^3$ which is embedded in the 1-skeleton of a triangulated convex polyhedron having $O(n)$ tetrahedra, with all vertices being integer lattice points using integers bounded by $O(n)$, and which projects in the z-direction to the link diagram $\mathcal{D}$. To do this we extend $\mathcal{D}$ to a maximal planar graph and then use de Frajsseix *et al.* [6] to construct a planar embedding of this graph with small integer lattice point vertices. We take two copies of the graph in the plane $z = -1$ and $z = 1$ and use them for overcrossings and undercrossing planes, respectively. We can embed this in a triangulated convex polyhedron with the lifted link L in its interior using $840n$ tetrahedra. We barycentrically subdivide twice to obtain a regular neighborhood of L, remove its interior, and construct the marked edges. Finally we extend to a triangulation of $\mathbf{S}^3$ by coning the triangular faces on the surface of the polyhedron to a point "at infinity".

8. Certifying unknottedness

To show that the UNKNOTTING PROBLEM is in **NP**, we must construct for any unknotted knot diagram $\mathcal{D}$ a polynomial length certificate, that can be verified in polynomial time, for the unknottedness of $\mathcal{D}$. The construction of the certificate, and its verification, take place in the following steps.

1. Given a link diagram $\mathcal{D}$, verify that it is a knot diagram. (This can be done in deterministic polynomial time.)

2. Construct a piecewise-linear knot K in $\mathbf{R}^3$ that has regular projection $\mathcal{D}$, together with a good triangulation. From it construct a good triangulation of $M_K \cong \mathbf{S}^3 - R_K$ which contains t tetrahedra, with $t = O(n)$, and with a meridian marked in ∂M_K. (Use Lemma 7.1.)

3. Guess a suitable fundamental vertex solution $v \in \mathbf{Z}^{7t}$ to the Haken normal equations for M_K. (This solution can be written in polynomial length by Lemma 6.1.) Verify the quadrilateral disjointness conditions. Let S denote the associated normal surface, so $v = v(S)$.

4. Verify that S is an essential disk for ∂M_K.

   (a) Verify that S is connected by verifying that v is a minimal vertex solution.

   (b) Verify that S is a disk by verifying that $\chi(S) = 1$. The Euler characteristic can be calculated as an appropriate linear combination of components of v. Since the connected surface S is embedded in $\mathbf{R}^3$, $\chi(S) = 1$ implies that S is orientable and that ∂S is homeomorphic to a circle.

   (c) Verify that S is essential by verifying that the homology class $[\partial S] = (0, \pm 1)$ in $H_1(\partial M_K, \mathbf{Z})$. The only possibilities for $[\partial S]$ in $H_1(\partial M_K, \mathbf{Z})$ are $(0,0)$ or $(0, \pm 1)$, so this can be done by verifying that the number of intersections of ∂S with the marked meridian of ∂M_K is odd.

This certificate specifies a normal surface which may contain exponentially many pieces, but labels it using the vector v which is of polynomial size, and is able to verify its properties using polynomial time computations on v, using Lemma 4.1 and Lemma 4.2.

The correctness of this certificate relies on the result of Jaco and Tollefson [15] given as Theorem 5.3. Without using this result we could still obtain the weaker result that the UNKNOTTING PROBLEM is in $\Sigma_2\mathbf{P}$. In step 3 we guess a suitable fundamental solution, not known to be a vertex solution. The only step that must be changed is Step 4a, which we change to verify that v is a fundamental solution.. This can be done by verifying that all decompositions of the form $v = v' + v''$, with v' and v'' being solutions of the Haken normal equations for M_K, have either $v' = 0$ or $v'' = 0$.

This approach also yields an algorithm to decide unknottedness of a link diagram, which proceeds by systematically searching for a certificate of the kind above. The algorithm generates all vertex solutions sequentially, and deterministically tests the steps of the certificate above on each one. This yields the following result.

Theorem 8.1. *There is a constant c and a Turing machine that can decide for any n-crossing knot diagram whether it represents the trivial knot in time $O(2^{cn})$ and space $O(n^2 \log n)$.*

For Haken's original approach we obtain the running time bound $O(2^{cn^2})$ by finding and testing all fundamental solutions sequentially, using the bound of Lemma 6.2.

9. Certifying splittability

We treat the SPLITTING PROBLEM with a modification of the method described above. We use the splittability criterion of Schubert [27] (Satz 4.1), according to which a link is splittable if and only if there is a normal sphere separating two components of the boundary of the link complement. We also use a result of Jaco and Tollefson [15] (Theorem 5.2), according to which, if there is such a normal sphere, then there is one associated to a vertex solution of the Haken normal equations. The construction of the certificate, and its verification, take place in the following steps.

1. Given a link diagram $\mathcal{D}$, construct a piecewise-linear link L in $\mathbf{R}^3$ that has regular projection $\mathcal{D}$. From it construct a good triangulation of $M_L \cong \mathbf{S}^3 - R_L$ which contains t tetrahedra, with $t = O(n)$, and with a meridian marked in each component of ∂M_L. (Use Lemma 7.1.)

2. Guess a suitable vertex solution $v \in \mathbf{Z}^{7t}$ to the Haken normal equations for M_L. (This solution can be written in polynomial length by Lemma 6.1.) Verify the quadrilateral disjointness conditions. Let S denote the associated normal surface, so $v = v(S)$.

3. Verify that S is a sphere that splits two components of ∂M_L.

   (a) Verify that S is connected by verifying that v is a minimal vertex solution.

   (b) Verify that S is a sphere by verifying that $\chi(S) = 2$.

   (c) Verify that S separates two components T and T' of ∂M_L by verifying that the number of intersections of S with the marked arc joining T and T' is odd.

10. Determining the genus

Finally, the algorithm of section 8 can easily be generalized to solve the GENUS PROBLEM in polynomial space.

1. Given a link diagram $\mathcal{D}$ and a genus k, verify as before that $\mathcal{D}$ is a knot diagram.

2. Construct a piecewise-linear knot K in $\mathbf{R}^3$ that has regular projection $\mathcal{D}$, together with a good triangulation. From it construct $M_K \cong \mathbf{S}^3 - R_K$, which contains t tetrahedra, with $t = O(n)$, with a meridian marked in ∂M_K.

3. Guess a suitable fundamental solution $v \in \mathbf{Z}^{7t}$ to the Haken normal equations for M_K. (Here we use Theorem 5.2.) Verify the quadrilateral disjointness conditions. Let S denote the associated normal surface, so $v = v(S)$.

4. Verify that S is a connected orientable surface ∂S a circle, with genus $g(S) \leq k$ and with ∂S a longitude in ∂M_K.
   (a) Verify that S is connected by verifying the connectedness of an undirected graph with nodes corresponding to triangles in the triangulation of S and edges joining matching triangles.
   (b) Verify that S is orientable by verifying the non-connectedness of an undirected graph with nodes representing each of the two sides of triangles in the triangulation and edges joining matching sides of matching triangles. (Since the surface S is embedded and connected in an orientable manifold, S is orientable if and only if it is two-sided.)
   (c) Verify that ∂S is a single circle by verifying that it is non-empty and connected (as an undirected graph).
   (d) Verify that S has genus at most k by verifying that $\chi(S) \geq 1 - 2g$.
   (e) Verify that ∂S is a longitude in ∂M_K by verifying that the homology class $[\partial S] = (0, 1)$ in $H_1(\partial M_K, \mathbf{Z}/2\mathbf{Z})$. This can be done by verifying that the number of intersections of ∂S with the marked meridian of ∂M_K is odd.

In Steps 4a, 4b and 4c, we use the fact that in an undirected graph in which nodes can be written down in polynomial length and in which adjacency of nodes can be tested in polynomial space, the connectedness of the graph can be determined in polynomial space (see Savitch [24]). Since all other steps can clearly be implemented in at most polynomial space, this yields an algorithm for the GENUS PROBLEM in polynomial space.

11. Conclusion

We know of no non-trivial lower bounds or hardness results for any of the problems we have discussed; in particular, we cannot even refute the implausible hypothesis that they can all be solved in logarithmic space. There are also a great many other knot properties and invariants apart from those considered here, and for many of them it is a challenging open problem to find complexity bounds.

One interesting question is whether the UNKNOTTING PROBLEM is in **co-NP**. Thurston's geometrization theorem for Haken manifolds implies that knot groups are residually finite [12]. It follows that a non-trivial knot has a non-cyclic representation into a finite permutation group. Unfortunately no way is yet known to bound the size of this group; if the number of symbols in the smallest such permutation group were bounded by a polynomial in the number of crossings, then the UNKNOTTING PROBLEM would be in **co-NP**. In practice the order of such a group seems to be quite small.

Perhaps the most ambitious of the open problems is to determine the complexity of the KNOT EQUIVALENCE PROBLEM (see Waldhausen [33] and Hemion [10]).

References

[1] C. C. Adams, *The knot book. An elementary introduction to the mathematical theory of knots.* W. H. Freeman and Company, New York, 1994.

[2] J. W. Alexander, "Topological Invariants of Knots and Links", *Trans. AMS*, 30 (1928) 275–306.

[3] J. Birman, "New Points of View in Knot Theory", *Bull. AMS*, 28 (1993) 253–287.

[4] G. Burde and H. Zieschang, *Knots*, De Gruyter, 1985.

[5] M. Dehn, "Über die Topologie des dreidimensional Raumes", *Math. Annalen*, 69 (1910) 137–168.

[6] H. de Frajsseix, J. Pach and R. Pollack, "How to Draw a Planar Graph on a Grid", *Combinatorica*, 10 (1990) 41–51.

[7] W. Haken, "Theorie der Normalflächen: Ein Isotopiekriterium für den Kreisknoten", *Acta Math.*, 105 (1961) 245–375.

[8] J. Hass "Algorithms for recognizing knots and 3-manifolds", (to appear in *Chaos, Solitons and Fractals*).

[9] J. Hass and J. C. Lagarias, "The number of Reidemeister moves needed for unknotting", preprint.

[10] G. Hemion, *The Classification of Knots and 3-Dimensional Spaces*, Oxford University Press, 1992.

[11] J. Hempel, *3-Manifolds*, Princeton University Press, Princeton, NJ 1976.

[12] J. Hempel, "Residual finiteness for 3-manifolds", Combinatorial group theory and topology, 379–396, Ann. of Math. Stud., 111, Princeton Univ. Press, Princeton, NJ, 1987.

[13] W. Jaco *Lectures on three-manifold topology*, CBMS Regional conference series in mathematics, 43 AMS, Providence, RI, 1980.

[14] W. Jaco and U. Oertel, "An Algorithm to Decide If a 3-Manifold Is a Haken Manifold", *Topology*, 23 (1984) 195–209.

[15] W. Jaco and J. L. Tollefson, "Algorithms for the Complete Decomposition of a Closed 3-Manifold", *Illinois J. Math.*, 39 (1995) 358–406.

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

[17] F. Jaeger, D. L. Vertigan and D. J. A. Welsh, "On the Computational Complexity of the Jones and Tutte Polynomials", *Math. Proc. Cambridge Phil. Soc.*, 108 (1990) 35–53.

[18] V. F. R. Jones, "A Polynomial Invariant of Knots via von Neumann Algebras", *Bull. AMS*, 12 (1985) 103–111.

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

[20] E. E. Moise, "Affine Structures in 3-Manifolds, V: The Triangulation Theorem and Hauptvermutung", *Ann. Math.*, 56 (1952) 96–114.

[21] C. D. Papakyriakopoulos, "On Dehn's Lemma and the Asphericity of Knots", *Ann. Math.*, 66 (1957) 1–26.

[22] N. Pippenger, "Knots in Random Walks", *Discr. Appl. Math.*, 25 (1989) 273–278.

[23] M. O. Rabin, "Recursive Unsolvability of Group-Theoretic Problems", *Ann. Math.*, 67 (1958) 172–194.

[24] W. J. Savitch, "Relationship between Nondeterministic and Deterministic Tape Classes", *J. Computer and System Sciences*, 4 (1970) 177–192.

[25] S. Sawin, "Links, quantum groups and TQFTs", *Bull. AMS*, 33 (1996) 413-445.

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

[27] H. Schubert, "Bestimmung der Primfaktorzerlegung von Verkettungen", *Math. Zeitschr.*, 76 (1961) 116–148.

[28] A. Sebö, "Hilbert Bases, Caratheodory's Theorem and Combinatorial Optimization", in: R. Kannan and W. R. Pulleyblank (Eds.), *Integer Programming and Combinatorial Optimization*, University of Waterloo Press, 1990, pp. 431–455.

[29] H. Seifert, "Über das Geschlecht von Knoten", *Math. Annalen*, 110 (1935) 571–592.

[30] J. Snoeyink, "A Trivial Knot Whose Spanning Disks Have Exponential Size", *Proc. ACM Conf. on Computational Geometry*, 6 (1990) 139–147.

[31] J. Stillwell, *Classical Topology and Combinatorial Group Theory*, Springer-Verlag, 1980.

[32] D. W. Sumners and S. G. Whittington, "Knots in Self-Avoiding Walks", *J. Phys. A*, 21 (1988) 1689–1694.

[33] F. Waldhausen, "Recent Results on Sufficiently Large 3-Manifolds", *Proc. Symp. Pure Math.*, 32, 2 (1978) 21–38.

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

[35] D. J. A. Welsh, "The Complexity of Knots", *Ann. Discr. Math.*, 55 (1993) 159–173.

[36] D. J. A. Welsh, "Knots and Braids", *Contemp. Math.*, 147 (1993) 109–123.