A characterisation of alternating knot exteriors

JOSHUA A HOWIE

We give a topological characterisation of alternating knot exteriors based on the presence of special spanning surfaces. This shows that being alternating is a topological property of the knot exterior and not just a property of diagrams, answering an old question of Fox. We also give a characterisation of alternating link exteriors which have marked meridians. We then describe a normal surface algorithm which can decide if a knot is alternating given a triangulation of its exterior as input.

1 Introduction

Let L be a link in S^3, and let $N(L)$ be a regular open neighbourhood. Then the link exterior $X \cong S^3 \setminus N(L)$ is a compact 3–manifold with torus boundary components. A planar link diagram $\pi(L)$ is a projection

$$\pi \colon S^2 \times I \to S^2,$$

where L has been isotoped to lie in some $S^2 \times I \subset S^3$, together with crossing information. A diagram $\pi(L)$ is alternating if the crossings alternate between over- and under-crossings as we traverse the projection of the link. A nontrivial link is alternating if it admits an alternating diagram. We take the convention that the unknot is not alternating.

A simple Euler characteristic argument shows that if $\pi(L)$ is a planar diagram with n crossings and checkerboard surfaces Σ and Σ', then

$$\chi(\Sigma) + \chi(\Sigma') + n = 2.$$

Furthermore, if $\pi(L)$ is reduced, nonsplit and alternating, then Σ and Σ' are both π_1–essential in X and $2n$ is the difference between the aggregate slopes of Σ and Σ'. If K is a knot, then $2n$ is the difference between their boundary slopes.

Our main result is to prove the converse, where we think of the difference in boundary slopes of Σ and Σ' as the minimal geometric intersection number of $\partial\Sigma$ and $\partial\Sigma'$ on ∂X, which we denote by $i(\partial\Sigma, \partial\Sigma')$.

Theorem 3.1 *Let K be a nontrivial knot in S^3 with exterior X. Then K is alternating if and only if there exist a pair of connected spanning surfaces Σ, Σ' in X which satisfy*

$(\star)$ $\qquad\qquad\qquad \chi(\Sigma) + \chi(\Sigma') + \frac{1}{2}i(\partial\Sigma, \partial\Sigma') = 2.$

Notice that all the conditions on X in this characterisation are topological in nature. Theorem 3.1 answers an old question, attributed to Ralph Fox, "What is an alternating knot?" This question has been interpreted as requesting a nondiagrammatic description of alternating knots; see Lickorish [16, page 32].

A similar characterisation of alternating knot exteriors has been independently obtained by Greene [7]. Alternating knot exteriors are characterised by a pair of spanning surfaces which are positive and negative definite with respect to the Gordon–Litherland bilinear form.

By a theorem of Gordon and Luecke [6], a knot exterior has a unique meridian so the concept of a spanning surface is well-defined in X. For link exteriors, this is not true (see Gordon [5]) and there are 3–manifolds which are homeomorphic to the exterior of both alternating and nonalternating links. However, X together with a marked meridian on each boundary component does uniquely determine a link in S^3. With the addition of an extra condition on intersection numbers, we are able to give a characterisation of alternating link exteriors with marked meridians in Theorem 3.2.

For the second half of this article we turn to normal surface theory, and show that given a 3–manifold with connected torus boundary X, it is possible to decide if X is the exterior of an alternating knot in S^3. We also show that given a nonalternating planar diagram of a knot K, we can decide if K is alternating, and if so, we can produce an alternating diagram of either K or its mirror image.

Theorem 5.2 *Let X be the exterior of a knot $K \subset S^3$. Given X, there is a normal surface algorithm to decide if K is alternating.*

In an appendix to [7], Juhász and Lackenby have also found a normal surface algorithm to decide if a knot is prime and alternating based on Greene's characterisation.

Organisation In Section 2, we describe how two spanning surfaces for a link intersect. In Section 3, we prove Theorem 3.1, and give the version for links. In Section 4, we give some background on normal surface theory and the boundary solution space. In Section 5, we detail the algorithm which can decide if a knot manifold is the exterior of an alternating knot.

2 Intersections of spanning surfaces

Let L be a link with m components denoted by L_j for $j = 1, \ldots, m$. Denote the boundary components of X by $C_j = \partial N(L_j)$ where each C_j is a torus.

A curve $\mu_j \subset C_j$ is a meridian of X if μ_j bounds an embedded disk in $N(L_j)$ which intersects L_j transversely exactly once. Given X, a set of marked meridians is a set of curves $\{\mu_j\}$ with one μ_j on each C_j such that Dehn filling along each μ_j produces the 3–sphere.

We define a preferred longitude λ_j of C_j to be the unique nontrivial curve on C_j which meets μ_j exactly once and bounds an orientable surface S_j in $S^3 \setminus N(L_j)$. Note that S_j is not necessarily embedded in X since it may intersect other components of L.

Let $\overline{\Sigma}$ be a compact surface embedded in S^3. Then $\overline{\Sigma}$ is a spanning surface for a link L if $\partial \overline{\Sigma} = L$. Let Σ be a surface with boundary, properly embedded in X. Then Σ is a spanning surface for X if each component of $\partial \Sigma$ has minimal geometric intersection number one with the meridian μ_j of C_j. These two notions of spanning surface are related since $\Sigma = X \cap \overline{\Sigma}$ whenever $\overline{\Sigma}$ is in general position with respect to ∂X, and Σ can be extended to $\overline{\Sigma}$ by attaching a small annulus in each component of $N(L)$.

Spanning surfaces Σ and Σ' are in general position in X if they intersect in a set of properly embedded arcs and embedded loops. In particular, there are no triple points or branch points, since each spanning surface is properly embedded.

For $j = 1, \ldots, m$, let $\{\sigma_j\}$ be the components of $\partial \Sigma$ and let $\{\sigma'_j\}$ be the components of $\partial \Sigma'$. Fix an orientation on each longitude λ_j and define the orientation of each meridian μ_j so that the (λ_j, μ_j) form a right-handed basis for each torus boundary component C_j. Then $[\sigma_j] = p_j[\mu_j] + [\lambda_j]$ and $[\sigma'_j] = p'_j[\mu_j] + [\lambda_j]$, where $p_j, p'_j \in \mathbb{Z}$. We define the algebraic intersection number of σ_j and σ'_j to be

$$i_a(\sigma_j, \sigma'_j) = p_j - p'_j = -i_a(\sigma'_j, \sigma_j),$$

while the geometric intersection number is

$$i(\sigma_j, \sigma'_j) = |i_a(\sigma_j, \sigma'_j)| = |p_j - p'_j|.$$

Define the geometric intersection number of $\partial \Sigma$ and $\partial \Sigma'$ to be

$$i(\partial \Sigma, \partial \Sigma') = \sum_{j=1}^{m} i(\sigma_j, \sigma'_j).$$

This geometric intersection number measures the difference in aggregate slopes of the two spanning surfaces, as defined in [1]. For a knot this is the difference in boundary slopes.

If we isotope Σ and Σ' so that they realise $i(\partial\Sigma, \partial\Sigma')$, then σ_j and σ_j' form a quadrangulation Q_j of C_j, where each quadrangular face has one pair of nonadjacent vertices on its boundary which are identified. There are $i_j = i(\sigma_j, \sigma_j')$ vertices, $2i_j$ edges, and i_j faces in Q_j. We refer to $\bigsqcup_{j=1}^{m} Q_j$ as a boundary quadrangulation of ∂X. When forming a quadrangulation on C_j, every intersection of σ_j and σ_j' has the same sign.
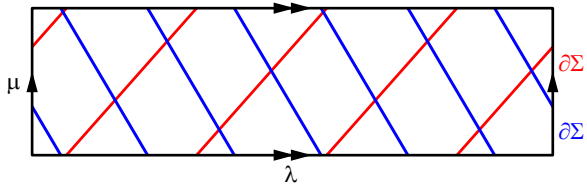

Figure 1: A boundary quadrangulation of ∂X formed by $\partial\Sigma$ and $\partial\Sigma'$ in the case of a knot exterior

An arc of intersection between two spanning surfaces Σ and Σ' in a link exterior X is called a double arc. Let α be a double arc and let $\bar{\alpha}$ be its extension to S^3 such that $\partial\bar{\alpha} \subset L$. There are two types of double arc.

Let W be a regular neighbourhood of $\bar{\alpha}$ in S^3. Let β and β' be the two components of $W \cap L$. We can choose W so that $V = W \cap X$ is a compact handlebody of genus two, and so that $\Sigma \cap V$ and $\Sigma' \cap V$ are both disks.

Fix an orientation on β. This induces orientations on the disks $\Sigma \cap V$ and $\Sigma' \cap V$. If these both induce the same orientation on β', then α is a *parallel* arc. If they induce opposite orientations on β', then α is a *standard* arc.
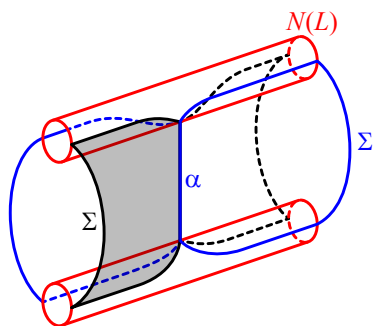

Figure 2: A parallel arc of intersection between two spanning surfaces

This is equivalent to saying $\partial(\Sigma \cap V)$ and $\partial(\Sigma' \cap V)$ have algebraic intersection number zero if α is a parallel arc, and algebraic intersection number two if α is a standard arc. The intersections on ∂X at the endpoints of a parallel arc α have the same sign if α is standard, but opposite signs when α is parallel.

If we collapse a standard arc to a point, then $(\overline{\Sigma} \cup \overline{\Sigma}') \cap W$ collapses to a disk. If we try to collapse a parallel arc to a point, then $(\overline{\Sigma} \cup \overline{\Sigma}') \cap W$ collapses to an object homeomorphic to a neighbourhood of the apex of a double cone.

Lemma 2.1 *Let Σ and Σ' be spanning surfaces for a link L, isotoped so that their boundaries realise the intersection number $i(\partial\Sigma, \partial\Sigma')$ on ∂X. If*

$$i(\partial\Sigma, \partial\Sigma') = \left| \sum_{j=1}^{m} i_a(\sigma_j, \sigma_j') \right|,$$

then every double arc of intersection between Σ and Σ' is standard.

Proof By definition $i(\partial\Sigma, \partial\Sigma') = \sum_{j=1}^{m} i(\sigma_j, \sigma_j')$. If

$$\left| \sum_{j=1}^{m} i_a(\sigma_j, \sigma_j') \right| = \sum_{j=1}^{m} i(\sigma_j, \sigma_j'),$$

then for every $j = 1, \ldots, m$, either every intersection between σ_j and σ_j' is positive, or every intersection between σ_j and σ_j' is negative. A parallel arc only occurs when a double arc connects a positive intersection to a negative intersection. $\square$

Note that Lemma 2.1 implies that parallel arcs of intersections can only occur when the double arc runs between different components of ∂X. Hence if K is a knot, then every arc of intersection between two spanning surfaces realising minimal intersection number must be standard.
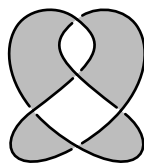

Figure 3: Black and white checkerboard surfaces for a knot

For any nonsplit planar link diagram $\pi(L)$, there is a standard position for the associated checkerboard surfaces. Away from a crossing the checkerboard surfaces are embedded in S^2, but in a small regular neighbourhood of a crossing, we think of the link lying on the surface of a ball U. The ball U intersects S^2 in an equatorial disk, and the over-strand runs over the upper hemisphere, while the under-strand runs under the lower hemisphere. Each checkerboard surface intersects U in a half-twisted band. The ball U is called a bubble and this viewpoint of checkerboard surfaces for planar alternating diagrams was introduced by Menasco [17].
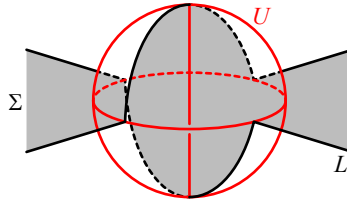
Figure 4: A checkerboard surface in standard position near a bubble U

In standard position, the checkerboard surfaces Σ and Σ' do not intersect in any loops, and intersect only in double arcs corresponding to the north-south axis of each bubble. These double arcs are all standard. The intersection of the corresponding spanning surfaces $\overline{\Sigma}$ and $\overline{\Sigma}'$ is a disjoint union of trivalent graphs and loops consisting of the link L and the collection of vertical axes of the bubbles. If we assume that $\pi(L)$ is nonsplit, then every component of L is involved in a crossing of $\pi(L)$, so that $\overline{\Sigma} \cap \overline{\Sigma}'$ forms a graph Γ', where each connected component is 3–regular.

For a nonsplit alternating planar projection $\pi(L)$ in standard position, as we traverse the image of L_j in $\pi(L)$, it can be seen that σ_j rotates in a positive manner say, with respect to S^2, while σ'_j rotates in a negative manner. Hence the checkerboard surfaces in standard position already realise the minimal geometric intersection number of their boundaries, and their boundaries form a boundary quadrangulation of ∂X.

Thus if $\pi(L)$ is a nonsplit planar alternating diagram of an m–component link L, which has n crossings, then

$$2n = i(\partial\Sigma, \partial\Sigma') = \sum_{j=1}^{m} i(\sigma_j, \sigma'_j).$$

A method for calculating the boundary slopes of the checkerboard surfaces associated to a reduced alternating knot diagram is detailed in [3], where it can be seen that $2n$ is the difference between the boundary slopes of the two checkerboard surfaces. If Σ is a spanning surface for a knot and $[\partial\Sigma] = p[\mu] + [\lambda]$, then the boundary slope of Σ is $p \in 2\mathbb{Z}$.

Note that if $\pi(L)$ is a nonalternating planar diagram with n crossings, then somewhere there are two consecutive over-crossings, which forces the boundaries of the associated checkerboard surfaces to create a bigon on some component of C_j. In this case $n > \frac{1}{2}i(\partial\Sigma, \partial\Sigma')$.

The Euler characteristics of the checkerboard surfaces arising from a planar projection are related to the Euler characteristic of the projection sphere by the equation

$$\chi(\Sigma) + \chi(\Sigma') + n = \chi(S^2) = 2.$$

A spanning surface Σ is π_1–essential in a knot exterior X if the induced homomorphism

$$\pi_1(\Sigma) \to \pi_1(X)$$

and the induced map

$$\pi_1(\Sigma, \partial\Sigma) \to \pi_1(X, \partial X)$$

are both injective. This implies that Σ is both incompressible and boundary-incompressible in X. Aumann [2] proved that both checkerboard surfaces associated to a planar reduced alternating projection of a knot K are π_1–essential in X.

3 Characterisation

We now give the proof of the nondiagrammatic characterisation of alternating knot exteriors.

Theorem 3.1 *Let K be a nontrivial knot in S^3 with exterior X. Then K has an alternating projection onto S^2 if and only if there exist a pair of connected spanning surfaces Σ, Σ' for X which satisfy*

$(\star)$ $$\chi(\Sigma) + \chi(\Sigma') + \tfrac{1}{2} i(\partial\Sigma, \partial\Sigma') = 2.$$

Proof One direction follows from the discussion in Section 2. For the converse, let Σ_0 and Σ'_0 be a pair of connected spanning surfaces for X which satisfy $(\star)$.

Since X is not a solid torus, X is boundary-irreducible, so K does not bound a disk in S^3. Hence $\chi(\Sigma_0) + \chi(\Sigma'_0) \leq 0$, so $i(\partial\Sigma_0, \partial\Sigma'_0) \neq 0$ by $(\star)$.

Isotope Σ_0 and Σ'_0 in X so that their boundaries realise the minimal geometric intersection number $i(\partial\Sigma_0, \partial\Sigma'_0)$. Hence $\partial\Sigma_0$ and $\partial\Sigma'_0$ form a boundary quadrangulation Q, and Q will remain fixed throughout the proof. We may assume that Σ_0 and Σ'_0 are in general position, so that they intersect in a set of proper arcs $\mathcal{A}$ and a set of embedded loops $\mathcal{L}_0$.

Recall that $\overline{\Sigma}_0$ is the extension of Σ_0 to S^3, so that $\partial\overline{\Sigma}_0 = K$. Assume that the interiors of $\overline{\Sigma}_0$ and $\overline{\Sigma}'_0$ are in general position, and no loops of intersection have been introduced by the extension process. Let $F'_0 = \overline{\Sigma}_0 \cup \overline{\Sigma}'_0$ and let $\Gamma' = (\overline{\Sigma}_0 \cap \overline{\Sigma}'_0) \setminus \mathcal{L}_0$, both of which are connected. Let $\overline{\mathcal{A}}$ be the extension of $\mathcal{A}$ to S^3, or in other words, let $\overline{\mathcal{A}}$ be the closure of $\Gamma' \setminus K$.

If we collapse each component of $\overline{\mathcal{A}}$ to a point, then F'_0 collapses to an immersed surface F_0 because X is a knot exterior and every arc of $\mathcal{A}$ is standard by Lemma 2.1.

Cutting F_0', $\overline{\Sigma}_0$ and $\overline{\Sigma}_0'$ along Γ' allows us to calculate that

$$\chi(F_0') = \chi(\overline{\Sigma}_0) + \chi(\overline{\Sigma}_0') + \tfrac{1}{2}i(\partial\Sigma_0, \partial\Sigma_0'),$$

from which the equation ($\star$) tells us that $\chi(F_0') = 2$. The 3–regular graph Γ' collapses to a 4–regular graph Γ. Let

$$f_0 \colon S_0 \looparrowright S^3$$

be the immersion of a surface S_0 such that $f_0(S_0) = F_0$. There are no triple points of self-intersection in F_0 since Σ_0 and Σ_0' are embedded, and the only double loops of self-intersection are precisely the elements of $\mathcal{L}_0$. It follows that

$$\chi(S_0) = \chi(F_0) = \chi(F_0') = 2,$$

which implies that S_0 is a 2–sphere.

Suppose $\mathcal{L}_0 \neq \varnothing$ and let $\mathcal{B}_0$ be the collection of loops $f_0^{-1}(\mathcal{L}_0)$ on S_0. Since $\mathcal{B}_0$ is the preimage of double loops, we know that $\mathcal{B}_0$ contains an even number of elements. Because S_0 is a 2–sphere, each loop $\beta \in \mathcal{B}_0$ is separating, and $\mathcal{B}_0$ cuts S_0 into a collection of planar surfaces with boundary. Let z_h be the number of planar surfaces in $S_0 \setminus \mathcal{B}_0$ which have h boundary components. Exactly one component of $S_0 \setminus \mathcal{B}_0$ contains the connected graph $f_0^{-1}(\Gamma)$.

Using an Euler characteristic argument, Nowik [19] points out that

$$\sum_{h \geq 1} (2 - h) z_h = 2,$$

which in particular implies that z_1, the number of disk regions in $S_0 \setminus \mathcal{B}_0$, is at least 2. Of course, for any collection of disjoint curves on a 2–sphere, there is an innermost one. Thus there is at least one loop in $\mathcal{L}_0$ which bounds a disk in F_0.

Let ℓ_0 be a loop in $\mathcal{L}_0$ which bounds a disk D_0 in either Σ_0 or Σ_0'. Without loss of generality assume $D_0 \subset \Sigma_0$. Let $\{\beta_0, \beta_0'\} = f_0^{-1}(\ell_0)$ where $f_0(N(\beta_0)) \subset \Sigma_0$ and $f_0(N(\beta_0')) \subset \Sigma_0'$. Notice that f_0^{-1} restricted to the interior of D_0 is a homeomorphism onto the interior of a disk in $S_0 \setminus \mathcal{B}_0$.

Let $A_0 = \ell_0 \times (-1, 1)$ be a regular neighbourhood of ℓ_0 in Σ_0'. We perform surgery on Σ_0' along D_0, by removing the annulus A_0 from Σ_0' and gluing in the two disks $D_0 \times \{-1\}$ and $D_0 \times \{1\}$.

Let (Σ_1, Σ_1') be the result of performing surgery along D_0 on (Σ_0, Σ_0'). Define S_1 to be the result of doing surgery along $f_0^{-1}(D_0)$ in S_0 to remove the loop β_0', and deleting the curve β_0.
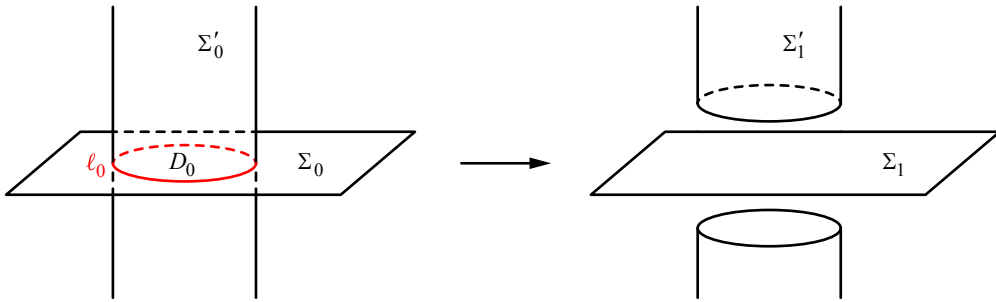
Figure 5: Surgery on Σ'_0 along the disk $D_0 \subset \Sigma_0$

Suppose $|\mathcal{L}_0| = k \geq 0$. For $0 \leq j \leq k-1$, inductively define $(\Sigma_{j+1}, \Sigma'_{j+1})$ to be the result of performing surgery along the disk D_j, where $\partial D_j = \ell_j \subset \mathcal{L}_j$ and D_j is a subdisk of either Σ_j or Σ'_j. Let $\mathcal{L}_{j+1} = \mathcal{L}_j \setminus \ell_j$ so that $\mathcal{L}_{j+1}$ is the set of loops of intersection of Σ_{j+1} and Σ'_{j+1}. Let F'_{j+1} be the result of the corresponding surgery on the 2–complex F'_j. Define S_{j+1} to be the result of doing surgery along $f_j^{-1}(D_j)$ in S_j and deleting the curve β'_j or β_j.

A similar calculation to Nowik's shows that

$$\sum_{h \geq 0}(2-h)z_h = \chi(S_j) = 2 + 2j,$$

where S_j is a collection of $j+1$ closed 2–spheres since β_j and β'_j are separating in S_j. Since Γ is connected and disjoint from $\mathcal{L}_0$, $f_j^{-1}(\Gamma)$ is contained in exactly one component of S_j. Hence $z_0 \leq j$, so that $z_1 \geq 2$, and therefore the disk D_j exists.

Continue this inductive surgery process until we have constructed Σ_k and Σ'_k. At this stage $\mathcal{L}_k$ is empty, and S_k consists of $k+1$ 2–spheres. Hence F'_k consists of k unmarked embedded 2–spheres and one 2–complex, denoted F', which contains Γ'.

Define $\overline{\Sigma}$ and $\overline{\Sigma}'$ to be the components of $\overline{\Sigma}_k$ and $\overline{\Sigma}'_k$, respectively, which constitute F'. Then Σ and Σ' are connected spanning surfaces for X which satisfy $(\star)$, and whose intersection is exactly $\mathcal{A}$. Collapsing the arcs of $\overline{\mathcal{A}}$ to points collapses F' and Γ' to F and Γ, respectively. Since $\chi(F') = 2$, it follows that $\chi(F) = 2$, so F is an embedded 2–sphere, which will be our desired projection surface.

Since $\partial\Sigma$ and $\partial\Sigma'$ realise $i(\partial\Sigma, \partial\Sigma') = i(\partial\Sigma_0, \partial\Sigma'_0)$, we can recover the crossing information of $\pi(K)$ from $\Gamma \subset F$. This is because every double arc is standard, so instead of collapsing every arc of $\overline{\mathcal{A}}$ to a point, we could collapse a regular neighbourhood of each $\overline{\alpha}$ to a bubble. The diagram $\pi(K)$ must be alternating since, otherwise, there would be a bigon between $\partial\Sigma$ and $\partial\Sigma'$ on ∂X, which contradicts that the boundary quadrangulation Q has remained fixed. Note that $\pi(K)$ is not necessarily reduced. $\square$

If $\pi(K)$ is a nonalternating planar diagram with checkerboard surfaces Σ and Σ', then there is a bigon on ∂X between $\partial\Sigma$ and $\partial\Sigma'$. Any attempt to isotope the checkerboard surfaces to remove all bigons and obtain a boundary quadrangulation causes Σ and Σ' to create an alternating diagram onto a nonplanar surface F. This surface F could be either embedded or immersed, and in the latter case may even be nonorientable [10, page 146].

Note that Theorem 3.1 is not concerned with primeness. However, a knot is prime if there are no essential annuli properly embedded in X at meridional slope. If a nontrivial knot is prime and alternating, then X is atoroidal, since no prime alternating knot is satellite [17].

In [10, Theorem 4.1], using a different method, the author also proved a variation of Theorem 3.1 which required both spanning surfaces to be π_1–essential in X. We also note that a more complicated characterisation can be obtained as a special case of a theorem proved in [10, Theorem 3.33] which gives a topological characterisation of a class of links which have certain alternating diagrams onto orientable surfaces of higher genus. This will be written up in a forthcoming article with Rubinstein [11].

We have not stated any of the theorems in this section for links. The issue is that, given two spanning surfaces, there could be parallel arcs between different components of ∂X. If the two spanning surfaces have been isotoped to create a boundary quadrangulation and there are parallel arcs of intersection, then the complex F_0' does not collapse to a surface. We note that there exists an example [10, page 139] of a pair of spanning surfaces for a nonsplit 2–component link which intersect in parallel arcs, yet still satisfy ($\star$).

However, if we assume that all arcs of intersection are standard, then we have the following theorem.

Theorem 3.2 *Let L be a nontrivial nonsplit link in S^3 with exterior X which has a marked meridian on each boundary component. Then L has an alternating projection onto S^2 if and only if there exist a pair of connected spanning surfaces Σ, Σ' for X which satisfy*

$$(\star) \qquad\qquad \chi(\Sigma) + \chi(\Sigma') + \tfrac{1}{2}i(\partial\Sigma, \partial\Sigma') = 2,$$

and

$$i(\partial\Sigma, \partial\Sigma') = \left| \sum_{j=1}^{m} i_a(\sigma_j, \sigma_j') \right|.$$

Proof Let $\pi(L)$ be a reduced nonsplit alternating projection of L onto S^2, and let Σ and Σ' be the associated checkerboard surfaces in standard position. Then $\pi(L)$, Σ and Σ' are connected, so every arc of intersection between Σ and Σ' is standard.

This means every algebraic intersection number $i_a(\sigma_j, \sigma'_j)$ is either positive or every $i_a(\sigma_j, \sigma'_j)$ is negative, which implies the second equation. The first equation follows from the two displayed equations on page 2358.

For the converse, the concept of a spanning surface for X is well-defined since a set of meridians for X are specified. Lemma 2.1 ensures that every arc of intersection is standard. Then the rest of the proof goes through as in Theorem 3.1. Since Σ and Σ' are connected, $\pi(L)$ must be nonsplit. $\qquad\square$

4 Normal surface theory

In this section we provide the background material necessary to construct our algorithm. Kneser introduced the concept of a normal surface, before Haken [8] developed normal surface theory into an important tool for algorithmic topology. We will give a brief outline of the theory; for full details the reader is referred to [15].

A knot manifold is a compact irreducible 3–manifold with connected torus boundary. A triangulation $\mathcal{T}$ of a knot manifold M is a collection of t tetrahedra and a set of equations which identify some pairs of faces of the tetrahedra, so that the link of every vertex is either a 2–sphere or a disk, and the unglued faces form the boundary torus ∂M.

A normal surface S is a properly embedded surface in M which is transverse to the 2–skeleton of $\mathcal{T}$, and such that $S \cap \triangle$ is a collection of triangular or quadrilateral disks, where $\triangle$ is any tetrahedron of $\mathcal{T}$, and each disk intersects each edge of $\triangle$ in at most one point. There are seven normal isotopy classes of normal disks, four are triangular and three are quadrilateral, and each of these is known as a disk type.

If we fix an ordering of the disk types $d_1, d_2, \ldots, d_{7t}$, then a normal surface S can be represented uniquely up to normal isotopy by a $7t$–tuple of nonnegative integers $\boldsymbol{n}(S) = (x_1, x_2, \ldots, x_{7t})$, where x_i is the number of disks of type d_i, and t is the number of tetrahedra in $\mathcal{T}$.

Conversely, given a $7t$–tuple of nonnegative integers $\boldsymbol{n}$, we can impose restrictions on the x_i so that $\boldsymbol{n}$ represents a properly embedded normal surface. We require that at least two of the three quadrilateral disk types are not present in each tetrahedra. This ensures that the surface is embedded. We also need to make sure that the disk types match up with the disk types in neighbouring tetrahedra.

An arc type is the normal isotopy class of the intersection of a normal surface with a face of a tetrahedron. There are three arc types in each face of each tetrahedron, and

each arc type is contributed to by two different disk types, one triangular, the other quadrilateral. We require that the number of each arc type in each face agrees with the number of arcs of the corresponding type in the face of the tetrahedron which is glued to it. This condition can be described by a linear equation for each arc type. Together they are called the matching equations for the normal surface S, and in a one-vertex triangulation of a knot manifold M there are $6t - 3$ matching equations.

The set of nonnegative integer solutions to the normal surface equations lie within an infinite linear cone $\mathcal{S}_{\mathcal{T}} \subset \mathbb{R}^{7t}$. The linear cone $\mathcal{S}_{\mathcal{T}}$ is called the solution space.

The additional condition that

$$\sum_{i=1}^{7t} x_i = 1$$

turns the solution space into a compact, convex, linear cell $\mathcal{P}_{\mathcal{T}} \subset \mathcal{S}_{\mathcal{T}}$. We call $\mathcal{P}_{\mathcal{T}}$ the projective solution space, and we let $\hat{\boldsymbol{n}}(S)$ represent the projective class of the normal surface S. The carrier of a normal surface S, denoted $\mathcal{C}_{\mathcal{T}}(S)$, is defined to be the unique minimal face of $\mathcal{P}_{\mathcal{T}}$ which contains $\hat{\boldsymbol{n}}(S)$.

Let S be a properly embedded surface in a 3-manifold M with triangulation $\mathcal{T}$. Haken [8] showed that after a series of isotopies, compressions, boundary-compressions and the removal of trivial 2–spheres and disks, S' can be represented as the union of properly embedded normal surfaces with respect to $\mathcal{T}$. In particular, if S is π_1–essential in M, then S can be isotoped to be normal with respect to $\mathcal{T}$.

Two normal surfaces S and S' are compatible if, for each tetrahedron $\triangle$ of $\mathcal{T}$, S and S' do not contain quadrilateral disks of different types. If S and S' are compatible, then we can form the Haken sum of S and S', which we denote $S \oplus S'$. The Haken sum is a geometric sum along each arc and loop of intersection between S and S', which is uniquely defined by the requirement that $S \oplus S'$ is a normal surface. Any other choice of geometric sums would produce a surface with folds. If $\boldsymbol{n}(S) = (x_1, x_2, \ldots, x_{7t})$ and $\boldsymbol{n}(S') = (x'_1, x'_2, \ldots, x'_{7t})$ are representatives of compatible normal surfaces S and S' in a triangulation $\mathcal{T}$ of a 3–manifold M, then $\boldsymbol{n}(S \oplus S') = \boldsymbol{n}(S) + \boldsymbol{n}(S') = (x_1 + x'_1, x_2 + x'_2, \ldots, x_{7t} + x'_{7t})$. Also, $\chi(S \oplus S') = \chi(S) + \chi(S')$.

A normal surface S is called a vertex surface if $\hat{\boldsymbol{n}}(S)$ lies at a vertex of the projective solution space. This means that whenever some multiple of S can be written as a Haken sum of two surfaces, both the summands are also multiples of S.

A normal surface S is called a fundamental surface if $\boldsymbol{n}(S)$ cannot be written as the sum of two solutions to the normal surface equations. Every vertex surface is a fundamental surface, but there exist fundamental surfaces which are not vertex surfaces.

All normal surfaces can be written as a finite sum of fundamental surfaces. There are a finite number of fundamental surfaces. They can be found algorithmically and Haken used this fact to construct his algorithms. Many of these algorithms have been subsequently improved so that they use vertex solutions rather than fundamental solutions, which makes the algorithms more efficient.

A triangulation is 0–efficient if the only normal disks or normal spheres are vertex-linking. Jaco and Rubinstein [12] showed that every compact orientable irreducible and boundary-irreducible 3–manifold with nonempty boundary admits a 0–efficient triangulation. Since the solid torus admits a one-vertex triangulation, it then follows that every knot exterior admits a one-vertex triangulation.

Let $\mathcal{T}$ be a one-vertex triangulation of a knot manifold M. Then there is an induced one-vertex triangulation $\mathcal{T}_\partial$ of ∂M. The boundary triangulation $\mathcal{T}_\partial$ consists of one vertex, three edges and two faces. There are six normal arc types; however, a normal curve is determined by just three of these arc types. Every curve on ∂M has a unique normal representative. This means that isotopy classes of curves on ∂M correspond to normal isotopy classes of curves on $\mathcal{T}_\partial$.
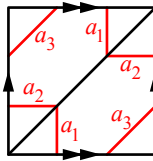

Figure 6: Normal arc types in the boundary triangulation

Fix an ordering of the disk types in $\mathcal{T}$ such that $d_1, \ldots, d_7$ represent the disk types in one of the tetrahedra which meets ∂M in a face ϕ. Furthermore, let $d_1, \ldots, d_4$ represent triangular disk types, and d_5, d_6, d_7 represent quadrilateral disks, such that d_i and d_{i+4} meet ϕ in the same arc type a_i for $i = 1, 2, 3$. Here d_4 is the triangular disk type which is disjoint from ϕ.

Let y_i be the number of arcs of type a_i in ϕ. It follows that

$$y_i = x_i + x_{i+4}$$

for $i = 1, 2, 3$. Jaco and Sedgwick [14, Theorem 3.6] showed that y_1, y_2, y_3 and the matching equations for normal curves determine the number of arcs of each type in the other 2–simplex of $\mathcal{T}_\partial$. We define the boundary solution space of ∂M to be

$$\mathcal{S}_{\mathcal{T}_\partial} = \{(y_1, y_2, y_3) \mid y_i \in \mathbb{N}_0\} \subset \mathbb{R}^3,$$

where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. If ∂S is the boundary of a properly embedded normal surface, then ∂S is represented by $\boldsymbol{n}(\partial S) = (y_1, y_2, y_3)$ in $\mathcal{S}_{\mathcal{T}_\partial}$.

If each coordinate of $\boldsymbol{n}(\partial S)$ is nonzero, then ∂S contains a trivial curve. Hence if S is an incompressible surface, then at least one of the coordinates of $\boldsymbol{n}(\partial S)$ is zero.

Jaco and Sedgwick [14, Section 3.3] proved that if S is a properly embedded π_1–essential normal surface with boundary in a compact 3–manifold M with triangulation $\mathcal{T}$, then every surface in $\mathcal{C}_{\mathcal{T}}(S)$ is either closed, or has the same slope as S. This means that if p/q is a boundary slope of X, then there is a vertex surface S which has slope p/q. Hence it is only necessary to check the vertices of $\mathcal{P}_{\mathcal{T}}$ in order to list all boundary slopes of X. In proving this theorem, Jaco and Sedgwick have given another proof of a theorem of Hatcher [9] that there are only a finite number of slopes bounding π_1–essential surfaces in any knot exterior. Recall that the set of slopes of a link exterior which bound π_1–essential surfaces is not necessarily finite, so our algorithm is only designed to work for knots.

Jaco and Sedgwick [14, Theorem 6.4] also gave an algorithm to decide if a knot manifold M is a knot exterior in S^3, which is also an algorithm to find the unique meridian μ of a knot exterior X. This algorithm makes use of the Rubinstein–Thompson algorithm [21; 22] which can decide if a 3–manifold is homeomorphic to the 3–sphere. There is also an algorithm which can decide if X is a solid torus [8], which is equivalent to deciding if K is the unknot, and it is now known that some spanning disk for K can be found as a vertex solution if K is the unknot.

The boundary triangulation $\mathcal{T}_\partial$ consists of two 2–simplices and three edges. We can modify the triangulation $\mathcal{T}$ by gluing two faces of a tetrahedron $\triangle$ to $\mathcal{T}_\partial$. The resulting triangulation $\mathcal{T}' = \mathcal{T} \cup \triangle$ is another one-vertex triangulation of M, and $\mathcal{T}'$ is called a layered triangulation. In effect, this is a $(2, 2)$–Pachner move on $\mathcal{T}_\partial$. The other two faces of $\triangle$ form the boundary triangulation $\mathcal{T}'_\partial$.

Layering a tetrahedron changes the slope of one of the edges in the boundary triangulation. It is always possible to layer a triangulation with a sequence of tetrahedra so that the edges of $\mathcal{T}'_\partial$ have slopes $\infty, k, k+1$, for some $k \in \mathbb{Z}$. We will choose to do this so that $(1, 0, 0) \in \mathcal{S}_{\mathcal{T}_\partial}$ represents the meridian μ.

5 Alternating algorithm

We now describe an algorithm to decide if a knot is alternating on S^2. The input is a triangulation $\mathcal{T}$ of a knot exterior X. If instead we are given a nonalternating planar diagram $\pi(K)$, then there is a method of Petronio [20] to construct a spine of the knot complement $S^3 \setminus K$ from the diagram $\pi(K)$. Dual to this spine is an ideal triangulation of $S^3 \setminus K$. We can then use an inflation of Jaco and Rubinstein [13] to construct a one-vertex triangulation of the knot exterior X from the ideal triangulation.

First we need one more result on spanning surfaces.

Lemma 5.1 *Suppose that Σ and Σ' are spanning surfaces for a knot $K \subset S^3$. If*

$$\chi(\Sigma) + \chi(\Sigma') + \tfrac{1}{2}i(\partial\Sigma, \partial\Sigma') > 2,$$

then at least one of Σ or Σ' is not connected.

Proof Isotope Σ and Σ' so that they are in general position and realise $i(\partial\Sigma, \partial\Sigma')$ on ∂X. As in the proof for Theorem 3.1, we form the pseudo-2–complex $F' = \overline{\Sigma} \cup \overline{\Sigma}'$ and collapse the arcs of intersection to points to obtain an immersed surface F, where $\chi(F) = \chi(\Sigma) + \chi(\Sigma') + \frac{1}{2}i(\partial\Sigma, \partial\Sigma') > 2$. Let $f\colon S \looparrowright S^3$ be an immersion of a closed surface S such that $f(S) = F$. The only possible self-intersections of $f(S)$ are loops, so $\chi(S) = \chi(F) > 2$, and therefore S is not connected. Hence either Σ or Σ' is not connected. $\square$

Theorem 5.2 *Let X be the exterior of a knot $K \subset S^3$. Given X, there is a normal surface algorithm to decide if K is alternating.*

Proof Let $\mathcal{T}'$ be a one-vertex triangulation of X. As shown in [14], any other triangulation of X can be modified to a one-vertex triangulation.

Use the Jaco–Sedgwick algorithm to find the unique meridian μ of X. Included in this process is a check whether X is a solid torus. If X is a solid torus, then K is the unknot, which by our convention is not alternating.

Layer the triangulation until one of the edges in the boundary is parallel to μ. Then the other edges in the boundary are parallel to $\lambda + k\mu$ and $\lambda + (k+1)\mu$ for some $k \in \mathbb{Z}$. Call this triangulation $\mathcal{T}$.

Let $\triangle$ be one of the two tetrahedra that meets the boundary and let ϕ be a face of $\triangle$ which lies in the boundary. Let $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ describe the normal coordinates of $S \cap \triangle$ where S is a properly embedded surface with boundary in X. Let (y_1, y_2, y_3) describe the normal coordinates of $\partial S \cap \phi$. As described in Section 4, we label the arc and disc types so that $y_i = x_i + x_{i+4}$ for each $i = 1, 2, 3$, and so that $(1, 0, 0)$ represents μ in ∂X. Let $(0, 1, 0)$ represent $\lambda + k\mu$ so that $(0, 0, 1)$ represents $\lambda + (k+1)\mu$ for some $k \in \mathbb{Z}$.

Any spanning surface for X meets μ exactly once. It follows that if (y_1, y_2, y_3) represents a spanning surface, then $y_2 + y_3 = 1$. So there are two types of coordinates in $\mathcal{S}_{\mathcal{T}_\partial}$ which can represent spanning surfaces: $(y, 1, 0)$ and $(y, 0, 1)$ for some $y \in \mathbb{N}_0$.

Let Σ and Σ' be normal spanning surfaces in X. We can read off the minimal geometric intersection number of their boundaries from their coordinates in $\mathcal{S}_{\mathcal{T}_\partial}$. Let $y, y' \in \mathbb{N}_0$. There are three cases:
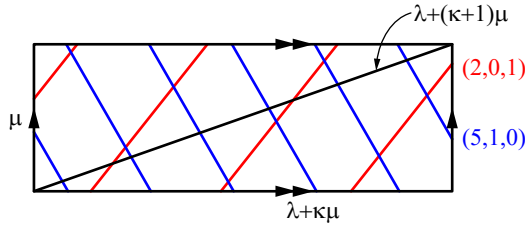
Figure 7: Normal curves in the boundary triangulation after layering

(1) If $\partial\Sigma$ and $\partial\Sigma'$ are represented by $(y, 1, 0)$ and $(y', 1, 0)$, respectively, then $i(\partial\Sigma, \partial\Sigma') = |y - y'|$.

(2) If $\partial\Sigma$ and $\partial\Sigma'$ are represented by $(y, 0, 1)$ and $(y', 0, 1)$, respectively, then $i(\partial\Sigma, \partial\Sigma') = |y - y'|$.

(3) If $\partial\Sigma$ and $\partial\Sigma'$ are represented by $(y, 1, 0)$ and $(y', 0, 1)$, respectively, then $i(\partial\Sigma, \partial\Sigma') = y + y' + 1$. See Figure 7 for an example of this case.

Note that we could continue layering the triangulation until $k = 0$, which would require detection of a Seifert surface, but this is not necessary since we are only interested in the differences of spanning slopes, and not the boundary slopes themselves.

Theorem 3.1 tells us that we need to find a pair of connected spanning surfaces at even boundary slope, which satisfy

$$(\star) \qquad\qquad \chi(\Sigma) + \chi(\Sigma') + \tfrac{1}{2} i(\partial\Sigma, \partial\Sigma') = 2.$$

The checkerboard surfaces Σ and Σ' associated to a reduced alternating diagram of K are one such pair of surfaces. Aumann [2] showed that they are both π_1–essential in X, so we know that both Σ and Σ' must have normal representatives in their isotopy classes.

Let Σ and Σ' be a pair of connected π_1–essential normal spanning surfaces which satisfy $(\star)$. Suppose that $n(\Sigma)$ is not a fundamental solution. Then Σ can be written as a Haken sum of fundamental surfaces,

$$\Sigma = \Sigma_1 \oplus \cdots \oplus \Sigma_a \oplus S_1 \oplus \cdots \oplus S_b,$$

where each Σ_i is a properly embedded compact surface with boundary, and each S_j is a properly embedded closed surface.

Since Σ is π_1–essential, it follows from Jaco and Sedgwick [14, Section 3.3] that that each Σ_i must have the same slope as Σ. Since Σ is a spanning surface, then a must equal 1, and Σ_1 is also a spanning surface at the same slope as Σ. In fact, Σ_1 must be fundamental.

Since X is irreducible and embedded in S^3, it follows that $\chi(S_j) \leq 0$ for each $j = 1, \ldots, b$. Suppose that for some k, we have $\chi(S_k) \leq -2$. But then $\chi(\Sigma) < \chi(\Sigma_1)$ which implies that

$$\chi(\Sigma_1) + \chi(\Sigma') + \tfrac{1}{2} i(\partial \Sigma_1, \partial \Sigma') > 2.$$

Hence, by Lemma 5.1, at least one of Σ_1 or Σ' must be disconnected. Every fundamental surface is connected, so Σ' must be disconnected, which gives a contradiction.

Hence $\chi(S_i) = 0$ for each i, and every S_i is an embedded torus. Moreover, $\chi(\Sigma) = \chi(\Sigma_1)$. Similarly,

$$\Sigma' = \Sigma_1' \oplus S_1' \oplus \cdots \oplus S_c',$$

where Σ_1' is fundamental spanning surface with the same slope and Euler characteristic as Σ', and S_j' is an embedded torus for each $j = 1, \ldots, c$. Therefore Σ_1 and Σ_1' are fundamental spanning surfaces which satisfy ($\star$).

Let $\mathcal{F}$ be the set of all fundamental spanning surfaces in X. For each pair of surfaces $\Sigma, \Sigma' \in \mathcal{F}$, calculate the intersection number $i(\partial \Sigma, \partial \Sigma')$, and calculate $\chi(\Sigma)$ and $\chi(\Sigma')$. There is an algorithm to compute the Euler characteristic of a properly embedded normal surface described in [15, Algorithm 9.1]. If Σ and Σ' satisfy ($\star$), then K is alternating by Theorem 3.1. If no pair of surfaces from $\mathcal{F}$ satisfy ($\star$), then K is not alternating. $\qquad\square$

Let $\pi(K)$ be an alternating diagram of the prime knot K with associated checkerboard surfaces Σ and Σ'. Let $\pi_*(K)$ be a different alternating diagram of K with associated checkerboard surfaces Σ_* and Σ_*'. If $\pi(K)$ and $\pi_*(K)$ are both reduced, then we know from a theorem of Menasco and Thistlethwaite [18] that $\pi(K)$ and $\pi_*(K)$ are related by a sequence of flypes. In that case, Σ and Σ_* are homeomorphic and have the same boundary slope, but Σ and Σ_* may not be isotopic in X. The same is true for Σ' and Σ_*'.

However, every checkerboard surface for a reduced alternating diagram is π_1–essential, and thus will appear amongst our collection of fundamental spanning surfaces $\mathcal{F}$. The collection $\mathcal{F}$ may also contain some pairs of surfaces which correspond to an alternating diagram which is not reduced. In this case, at least one of the checkerboard surfaces fails to be π_1–essential.

Let Σ and Σ' have minimal intersection number amongst all surfaces from $\mathcal{F}$ which satisfy ($\star$). Place an orientation on $\partial \Sigma$, and label the vertices of $\partial \Sigma \cap \partial \Sigma'$ in the order they are encountered as one traverses $\partial \Sigma$ by $1, \ldots, i$, where $i = i(\partial \Sigma, \partial \Sigma')$. Then each arc of intersection between Σ and Σ' is labelled by two numbers, one even and one odd. These pairs of numbers, listed as a sequence of even positive integers

in the order of their paired odd numbers, correspond to the Dowker–Thistlethwaite notation [4] of a planar alternating diagram of K or its mirror image.

Therefore, given a nonalternating planar diagram of a knot K, there is an algorithm to decide if K is alternating, and if so, there is an algorithm to produce an alternating diagram of K up to chirality.

References

[1] **C Adams**, **T Kindred**, *A classification of spanning surfaces for alternating links*, Algebr. Geom. Topol. 13 (2013) 2967–3007 MR

[2] **R J Aumann**, *Asphericity of alternating knots*, Ann. of Math. 64 (1956) 374–392 MR

[3] **C L Curtis**, **S J Taylor**, *The Jones polynomial and boundary slopes of alternating knots*, J. Knot Theory Ramifications 20 (2011) 1345–1354 MR

[4] **C H Dowker**, **M B Thistlethwaite**, *Classification of knot projections*, Topology Appl. 16 (1983) 19–31 MR

[5] **C M Gordon**, *Links and their complements*, from "Topology and geometry: commemorating SISTAG" (A J Berrick, M C Leung, X Xu, editors), Contemp. Math. 314, Amer. Math. Soc., Providence, RI (2002) 71–82 MR

[6] **C M Gordon**, **J Luecke**, *Knots are determined by their complements*, J. Amer. Math. Soc. 2 (1989) 371–415 MR

[7] **J E Greene**, *Alternating links and definite surfaces*, preprint (2015) arXiv To appear in Duke Math. J.

[8] **W Haken**, *Theorie der Normalflächen*, Acta Math. 105 (1961) 245–375 MR

[9] **A E Hatcher**, *On the boundary curves of incompressible surfaces*, Pacific J. Math. 99 (1982) 373–377 MR

[10] **J A Howie**, *Surface-alternating knots and links*, PhD thesis, University of Melbourne (2015) Available at `http://hdl.handle.net/11343/58991`

[11] **J A Howie**, **J H Rubinstein**, *Weakly generalised alternating links* In preparation

[12] **W Jaco**, **J H Rubinstein**, *0–efficient triangulations of 3–manifolds*, J. Differential Geom. 65 (2003) 61–168 MR

[13] **W Jaco**, **J H Rubinstein**, *Inflations of ideal triangulations*, Adv. Math. 267 (2014) 176–224 MR

[14] **W Jaco**, **E Sedgwick**, *Decision problems in the space of Dehn fillings*, Topology 42 (2003) 845–906 MR

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

[16] **W B R Lickorish**, *An introduction to knot theory*, Graduate Texts in Mathematics 175, Springer, New York (1997) MR

[17] **W Menasco**, *Closed incompressible surfaces in alternating knot and link complements*, Topology 23 (1984) 37–44 MR

[18] **W Menasco**, **M Thistlethwaite**, *The classification of alternating links*, Ann. of Math. 138 (1993) 113–171 MR

[19] **T Nowik**, *Dissecting the* 2–*sphere by immersions*, Geom. Dedicata 127 (2007) 37–41 MR

[20] **C Petronio**, *Ideal triangulations of link complements and hyperbolicity equations*, Geom. Dedicata 66 (1997) 27–50 MR

[21] **J H Rubinstein**, *An algorithm to recognize the* 3–*sphere*, from "Proceedings of the International Congress of Mathematicians" (S D Chatterji, editor), volume 1, Birkhäuser, Basel (1995) 601–611 MR

[22] **A Thompson**, *Thin position and the recognition problem for* S^3, Math. Res. Lett. 1 (1994) 613–630 MR

School of Mathematics and Statistics, University of Melbourne
Melbourne VIC 3010, Australia

`josh.howie@monash.edu`