*Math. Proc. Camb. Phil. Soc.* (2024), **177**, 67–102
doi:10.1017/S0305004124000148
First published online 19 September 2024



# Generalised knotoids

BY COLIN ADAMS AND ZACHARY ROMRELL
*Department of Mathematics,* 18 *Hoxsey St., Williams College, Williamstown, MA* 01267, *U.S.A.*
*e-mails*: cadams@williams.edu, zr3@williams.edu

ALEXANDRA BONAT
*Department of Mathematics,* 480 *Lincoln Drive, University of Wisconsin, Madison, WI* 537067, *U.S.A.*
*e-mail*: bonat@math.wisc.edu

MAYA CHANDE
*Department of Mathematics, Fine Hall, Washington Rd., Princeton University, Princeton, NJ* 08544, *U.S.A.*
*e-mail*: mchande@princeton.edu

JOYE CHEN, MAXWELL JIANG AND DANIEL SANTIAGO
*Department of Mathematics,* 77 *Massachusetts Ave., M.I.T., Cambridge, MA* 021397, *U.S.A.*
*e-mails*: joyec@mit.edu, mdjiang@mit.edu, dsantiag@mit.edu

BENJAMIN SHAPIRO
*Department of Mathematics,* 27 *N. Main St., Dartmouth College, Hanover, NH* 03755, *U.S.A.*
*e-mail*: benjamin.i.shapiro.gr@dartmouth.edu

AND DORA WOODRUFF
*Department of Mathematics,* 1 *Oxford St., Harvard University, Cambridge, MA* 021387, *U.S.A.*
*e-mail*: dorawoodruff@college.harvard.edu

(*Received* 30 *November* 2022*; revised* 15 *January* 2024*; accepted* 10 *January* 2024)

*Abstract*

In 2010, Turaev introduced knotoids as a variation on knots that replaces the embedding of a circle with the embedding of a closed interval with two endpoints which here we call poles. We define generalised knotoids to allow arbitrarily many poles, intervals and circles, each pole corresponding to any number of interval endpoints, including zero. This theory subsumes a variety of other related topological objects and introduces some particularly interesting new cases. We explore various analogs of knotoid invariants, including height, index polynomials, bracket polynomials and hyperbolicity. We further generalise to knotoidal graphs, which are a natural extension of spatial graphs that allow both poles and vertices.

© The Author(s), 2024. Published by Cambridge University Press on behalf of Cambridge Philosophical Society.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

2020 Mathematics Subject Classification: 57K12 (Primary); 57K14, 57K32 (Secondary)

## 1. *Introduction*

Knotoids were introduced by Turaev in [31] as an extension of classical knot theory. Knotoids are equivalence classes of generic immersions of [0,1] into $S^2$ considered up to isotopy and Reidemeister moves away from the endpoints. At the endpoints, here called poles, there is a forbidden fourth move, disallowing an endpoint from moving over/under a strand. Knotoids have been studied recently in various papers, including among many others [7, 17, 24, 26]. They have also received attention for their potential as protein models, for example in [10] and [14].

We introduce *generalised knotoids* and define analogs of knotoid invariants for generalised knotoids. The theory of generalised knotoids in $S^2$ subsumes knot theory and its various extensions, including both planar and spherical (multi-)knotoids as originally defined in [31], the multi-linkoids of [12], long knots as in [22], string links as in [30], the polar knots of [8] and so-called staked links coming from adding valency zero poles to link diagrams. Generalised knotoids are also interesting in their own right, and we include several conjectures concerning their properties.

We also introduce *knotoidal graphs* as a further extension of generalised knotoids. This definition is a natural extension of spatial graphs, bonded knotoids as in [15] and [16] and graphoids as in [16] and [19]. Knotoidal graphs may provide a useful model for proteins that consist of more than one protein molecule bonded together as occurs for hemoglobin and insulin for instance. Finally, we introduce hyperbolicity as an invariant for knotoidal graphs.

### 1·1. *Organisation*

In Section 2, we define generalised knotoid theory and describe its relation to various existing extensions of knot theory. The three sections that follow describe analogs of knotoid invariants for generalised knotoids. In Section 3, we generalise the notion of height and extend a theorem of Kauffman and Gügümcü [18] regarding height in minimal-crossing diagrams. In Section 4, we define a family of polynomial invariants for generalised knotoids that extend existing index polynomials for knotoids, and in Section 5, we define a bracket polynomial that generalises existing bracket polynomial invariants for knotoids.

In Section 6, we define knotoidal graphs. We also describe *rail diagrams* for knotoidal graphs, which provide a useful topological point of view. In Section 7, we use rail diagrams to extend a map defined in [2] for spherical knotoids to knotoidal graphs. This map takes knotoidal graphs to spatial graphs in manifolds which are either handlebodies or thickened surfaces. Then we define hyperbolicity and hyperbolic volumes of knotoidal graphs and discuss some applications. In particular, we prove that every knot in $S^3$ can be staked to be hyperbolic and thus has a well-defined minimum staked volume.

## 2. *Generalised knotoids*

Let $\Sigma$ denote a closed orientable surface and $G$ a finite graph. We do not require $G$ to be connected or simple, and $G$ may have valency-zero vertices. Let $\tilde{G}$ denote the disjoint union of $G$ with a finite collection of circles. The edges and circles of $\tilde{G}$ are called its *constituents*.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

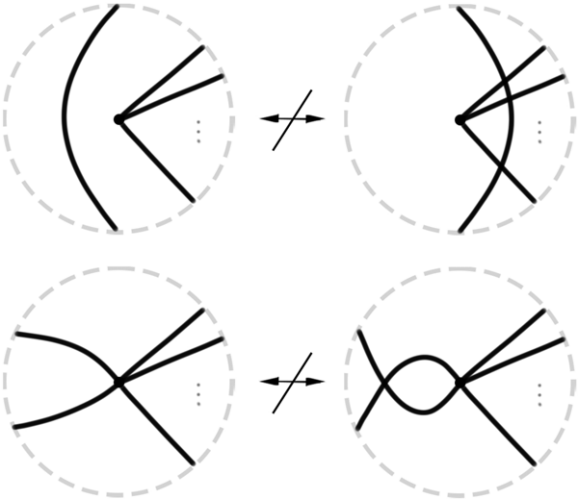


Fig. 1. The pole slide move (top) and the pole twist move (bottom) are forbidden, irrespective of the crossing data chosen for the diagrams on the right.

A *generalised knotoid diagram* $\mathcal{D}$ is a generic immersion of $\tilde{G}$ in $\Sigma$ whose only singularities are transverse double points, called *crossings*, which are labelled with over/undercrossing data. For brevity, we will also use $\mathcal{D}$ to refer to the image of the immersion. The graph $G$ (resp. $\tilde{G}$) is called the *underlying graph* (resp. *underlying looped graph*) of $\mathcal{D}$. Let $P(\mathcal{D})$ denote the set of images of the vertices of $\tilde{G}$, called the *poles* of $\mathcal{D}$. The *valency* of a pole $p \in P(\mathcal{D})$ is the valency of its corresponding vertex in the underlying graph $G$. A valency zero pole is also called an *isolated pole*.

Let $E(\mathcal{D})$ denote the set of images of the edges of $\tilde{G}$, called *segment constituents* of $\mathcal{D}$, and let $L(\mathcal{D})$ denote the images of the circles of $\tilde{G}$, called *loop constituents* of $\mathcal{D}$. Let $C(\mathcal{D}) := E(\mathcal{D}) \cup L(\mathcal{D})$ denote the set of *constituents* of $\mathcal{D}$. A labelling of the vertices of $G$ induces a labelling of the poles of $\mathcal{D}$. Similarly, a labelling (resp. an orientation) of the constituents of $\tilde{G}$ induces a labelling (resp. orientation) of the constituents of $\mathcal{D}$.

We introduce an equivalence relation on generalised knotoid diagrams in $\Sigma$ generated by ambient isotopy and the three standard Reidemeiester moves away from the poles. Observe that it is forbidden for a constituent to pass through a pole and for "twists" to be created or destroyed near a pole. See Figure 1. We remark that the forbidden twist move of Figure 1 parallels the forbidden move for vertices in a rigid-vertex spatial graph.

A *generalised knotoid* $\kappa$ is an equivalence class of generalised knotoid diagrams. It is clear that equivalency respects the number of poles and constituents as well as their associated data; in particular, we may speak of labelled and/or oriented generalised knotoids as equivalence classes of labelled and/or oriented generalised knotoid diagrams. Note that allowing twisting near poles yields a different theory of generalised knotoids, which we do not consider in this paper. However, some of the invariants we define later are able to distinguish between generalised knotoids which differ only by twists near poles and are otherwise identical.

An equivalent, topological viewpoint is to define a generalised knotoid diagram as a generic immersion of $\tilde{G}$, where all vertices of $G$ have valency at least one, in a compact

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

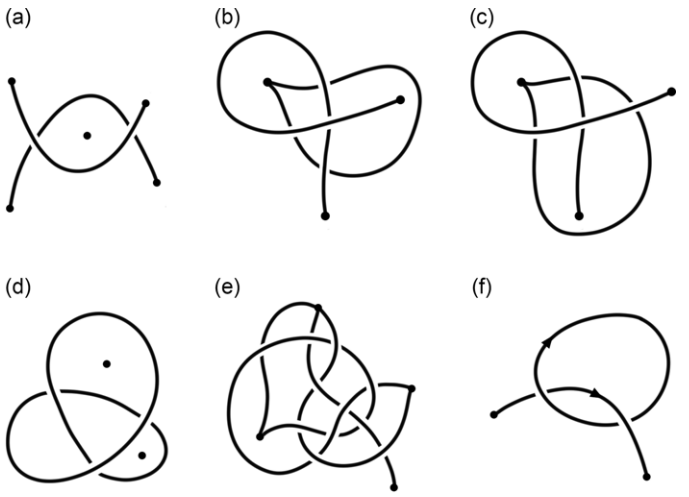



Fig. 2. Generalised knotoid diagrams.

orientable surface $\Sigma$ with or without boundary. Given such a diagram on a compact surface, we may recover a diagram on a closed surface by capping off each boundary component with a disk, then collapsing each disk to a point representing an isolated pole. Conversely, given a generalised knotoid diagram on a closed surface with isolated poles, we may remove an open disk neighbourhood of each isolated pole. That is, the theory of generalised knotoids on closed surfaces with isolated poles is equivalent to the theory of generalised knotoids on compact surfaces without isolated poles. We rely on both viewpoints, especially in Section 7 where the topological viewpoint proves useful.

We denote the case $\Sigma = S^2$ by *classical generalised knotoid theory*. In the remainder of this paper, a generalised knotoid is assumed to be classical unless otherwise stated.

*Example 2·1.* See Figure 2 for some diagrams of generalised knotoids. The reader may verify that Figures 2(b) and 2(c) represent equivalent generalised knotoids. Figure 2(f) represents an oriented generalised knotoid.

*Example 2·2.* A spherical knotoid is a generalised knotoid with two poles and a single segment constituent between them.

*Example 2·3.* A planar knotoid is a generalised knotoid with three poles and a single segment constituent connecting two of them. The equivalency of the two notions is obtained by identifying the isolated pole of the generalised knotoid with the point at infinity of the planar knotoid diagram. This example provides a useful viewpoint for a number of invariants defined for planar knotoids, and it motivates parts of the constructions of invariants for generalised knotoids in Sections 4 and 5.

*Example 2·4.* A classical knot or link is a generalised knotoid with no poles.

*Example 2·5.* The multi-linkoids of [**12**] are generalised knotoids whose poles all have valency one.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

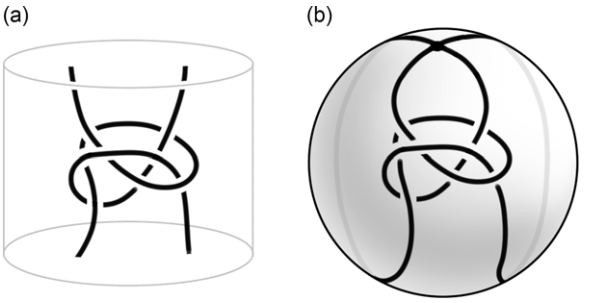



Fig. 3. The string link in Figure 3(a) corresponds to the generalised knotoid in Figure 3(b). The fact that Reidemeister moves must occur away from the pole captures the fact that the endpoints of the string link's segments are fixed.

*Example 2·6.* For a positive integer $n$, an *n-polar knot* [**8**] is a generalised knotoid with no loop constituents and whose underlying graph is the cycle graph $C_n$. An $n$-polar knot diagram looks like a classical knot diagram with $n$ poles placed on the knot, away from the crossings. Long knots [**22**] can be thought of as 1-polar knots, and the equivalency of the two notions is obtained by identifying the single pole of a polar knot diagram with the point at infinity in a long knot diagram.

*Example 2·7.* A *string link* with $n$ strings is defined in [**30**] as the embedding of a finite set of closed intervals $I_1, I_2, \ldots, I_n$ in $D \times I$ where $D$ is the unit disk in the $xy$-plane, such that the initial point of $I_i$ is sent to $(x_i, 0, 0)$, and the final point is sent to $(x_i, 0, 1)$, where $-1 < x_1 < x_2 < \cdots < x_n < 1$. This is defined up to ambient isotopy of the strings with fixed endpoints.

We represent a string link as a generalised knotoid with a single pole of valency $2n$. See Figure 3. If we thicken the sphere, we may consider the pole as the removal of a neighbourhood of a single vertical line in $S^2 \times I$ (bringing the ambient space to $D \times I$) and the constituents as embedded segments with endpoints on the circle boundary of $D \times \{1/2\}$. (See Section 6·2 for a discussion on *rail diagrams*.) The legs of the strings appear in clockwise order $\ell_1, \ell_2, \ldots, \ell_n$ around the circle boundary of $D \times \{1/2\}$ followed by the heads of the strings in clockwise order $h_n, h_{n-1}, \ldots, h_1$. Since the pole twist move is disallowed, these endpoints are fixed in their order around the circle, and we recover the notion of a string link.

*Example 2·8.* We define *staked links* on a surface $\Sigma$ to be the class of generalised knotoids on $\Sigma$ with no segment constituents and at least one pole. A diagram of a staked link looks like a link diagram on $\Sigma$ with isolated poles placed in the complementary regions of the link diagram on $\Sigma$. An example appears in Figure 2(d).

We have subsequently learned that in [**13**], Goldman and Kauffman defined tunnel links that are equivalent to staked links with two stakes and that were generated by considering conductance of link diagrams. In more recent as-of-yet unpublished work, independent of our work, Kaufmman and Gügümcü have been investigating properties of staked links under the name starred links.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

## 3. *Height for generalised knotoids*

We review the notion of height for knotoids, defined originally as *complexity* in [**31**]. Let $\mathcal{D}$ be a knotoid diagram on $\Sigma$, and let $\alpha$ be an embedded arc in $\Sigma$ connecting the endpoints that intersects $\mathcal{D}$ transversely and away from the crossings. We call $\alpha$ a *shortcut*. The *diagram height* $h(\mathcal{D})$ is the number of intersections of $\alpha$ with $\mathcal{D}$ not including the endpoints, minimised over all shortcuts $\alpha$. The *height* of a knotoid $k$ is the minimum of $h(D)$ over all diagrams $\mathcal{D}$ representing $k$, and a diagram $\mathcal{D}$ attaining this minimum *realises the height*.

This definition extends to generalised knotoids in a natural way. Let $\mathcal{D}$ be a pole-labelled generalised knotoid diagram on $\Sigma$ and let $A, B \in P(\mathcal{D})$ be distinct poles. A *shortcut* $\alpha$ between $A$ and $B$ is an embedded arc in $\Sigma$ with endpoints at $A$ and $B$ that intersects $\mathcal{D}$ transversely and away from crossings and poles. (If $A = B$, we define a shortcut in the same way with "embedded arc" replaced by "simple closed curve".) The *diagram height* between $A$ and $B$, denoted $h_{\mathcal{D}}(A, B)$, is the number of intersections of $\alpha$ with $\mathcal{D}$ not including $A$ and $B$, minimised over all shortcuts $\alpha$ between $A$ and $B$. An $\alpha$ attaining this minimum is a *minimal shortcut*. Given a generalised knotoid $\kappa$ and poles $A, B \in P(\kappa)$, the *height* between $A$ and $B$, denoted $h_{\kappa}(A, B)$, is the minimum of $h_{\mathcal{D}}(A, B)$ over all diagrams $\mathcal{D}$ representing $K$. A diagram $\mathcal{D}$ attaining this minimum *realises the height between A and B*. Given an ordering on the poles $P_1, \ldots, P_n$ of $\kappa$, the *height spectrum* of $\kappa$ is the symmetric matrix $(h(P_i, P_j))_{ij}$.

Theorem 3·1 was conjectured by Turaev in [**31**] and proved by Gügümcü and Kauffman in [**18**].

THEOREM 3·1 ([**18**, Theorem 13]). *Let $k$ be a spherical knotoid with $h(k) = 0$. Then any minimal-crossing diagram $\mathcal{D}$ of $k$ realises the height.*

The proof of theorem 3·1 given in [**15**] considers the *virtual closure* of a knotoid and applies results from virtual knot theory derived from the method of *parity projection* [**28**]. In theorem 3·2, we extend theorem 3·1 to generalised knotoids. The proof of theorem 3·2 adapts a topological approach to parity via colourings described in [**9**].

THEOREM 3·2. *Let $\kappa$ be a classical generalised knotoid and suppose poles $A, B \in P(\kappa)$ satisfy $h_{\kappa}(A, B) = 0$. Then any minimal-crossing diagram $\mathcal{D}$ of $\kappa$ realises the height between A and B.*

*Proof.* If $A = B$, the result is trivial, so assume $A \neq B$. Fix a diagram $\mathcal{D}_0$ such that $h_{\mathcal{D}_0}(A, B) = 0$, and let $\alpha$ be a minimal shortcut between $A$ and $B$. Suppose for contradiction's sake that there exists a minimal-crossing diagram $\mathcal{D}$ with $h_{\mathcal{D}}(A, B) > 0$. We may assume that there is a sequence of diagrams $\mathcal{D}_0 \to \mathcal{D}_1 \to \cdots \to \mathcal{D}_n = \mathcal{D}$ where for each $0 \leq i < n$, the diagrams $\mathcal{D}_i$ and $\mathcal{D}_{i+1}$ are related by a Reidemeister move. Recall that Reidemeister moves take place away from poles, so we may assume that each pole $Q$ of $\mathcal{D}_0$ has a neighbourhood $N_Q$ that is unchanged as the moves occur. We declare that the shortcut $\alpha$ is fixed across all the diagrams $\mathcal{D}_i$, and we may assume that $\alpha$ intersects each $\mathcal{D}_i$ transversely and away from crossings.

CLAIM 3·3. *The number of intersections of $\alpha$ with a constituent of $\mathcal{D}_i$, excluding potential intersections at poles, is even.*

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

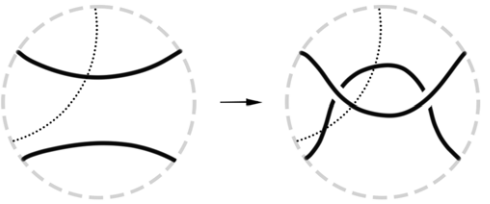


Fig. 4. A Reidemeister move occuring in the disk region $D$, whose boundary is indicated by dashed line. The shortcut $\alpha$ is indicated with dotted line. The parity of the number of intersections of $\alpha$ with each participating strand is preserved.

*Proof.* Let $e \in C(\kappa)$ be a constituent. For each $0 \leq j < n$, let $\#(\alpha \cap e)_j$ denote the number of intersections of $\alpha$ with $e$ in the diagram $\mathcal{D}_j$, excluding poles. We show that $\#(\alpha \cap e)_j$ and $\#(\alpha \cap e)_{j+1}$ have the same parity, and the conclusion follows from the assumption that $\#(\alpha \cap e)_0 = 0$.

Let $D$ be the disk region where the Reidemeister move $\mathcal{D}_j \to \mathcal{D}_{j+1}$ is applied. For each $t \in \{1, 2, 3\}$, a type $t$ Reidemeister move has $t$ *participating strands* $s_1, \ldots s_t$, some of which may belong to $e$. Each $s_\ell$ meets $\partial D$ at two *endpoints* that are fixed by the Reidemeister move.

If $\alpha$ does not intersect $D$, then the conclusion is clear. Otherwise, we may assume that $\alpha$ meets $\partial D$ at a finite number of points $P_1, P_2, \ldots, P_k$ and $Q_1, Q_2, \ldots, Q_k$, distinct from the endpoints of the participating strands, such that each pair $P_m$ and $Q_m$ is connected by a non-self-intersecting arc $\alpha_m \subset \alpha$ contained in the interior of $D$. Each $\alpha_m$ divides $D$ into two regions, and $\#(\alpha_m \cap s_\ell)_j$ is odd if the endpoints of $s_\ell$ are in opposite regions of $D$ and even otherwise. Since the Reidemeister move fixes the endpoints, it follows that $\#(\alpha_m \cap s_\ell)_j$ and $\#(\alpha_m \cap s_\ell)_{j+1}$ have the same parity for each $m$ and $\ell$, and the conclusion follows. See Figure 4.

Given a constituent $e$ in a diagram $\mathcal{D}_i$, an *$\alpha$-colouring* of $e$ is an assignment of black or gray to each point of $e$ such that the colour of $e$ changes precisely at the points where it intersects $\alpha$. It follows from Claim 3·3 that an $\alpha$-colouring exists for each $e$. We define an *$\alpha$-colouring* of the diagram $\mathcal{D}_i$ to be a choice of $\alpha$-colouring for each of its constituents such that each segment constituent is black in a neighbourhood of its endpoint poles. We inductively fix an $\alpha$-colouring $C_i$ for each $\mathcal{D}_i$ as follows. We declare $C_0$ to be the $\alpha$-colouring in which all constituents of $\mathcal{D}_0$ are coloured entirely black. Once $C_i$ is chosen, the Reidemeister move $\mathcal{D}_i \to \mathcal{D}_{i+1}$ uniquely determines $C_{i+1}$ if we require $C_i$ and $C_{i+1}$ to agree outside the disk region where the Reidemeister move is applied. See Figure 5. Under the $\alpha$-colouring $C_i$, a crossing $c$ in $\mathcal{D}_i$ is assigned a colour from each of the constituents meeting at $c$. We say $c$ is *even* if these colours agree and *odd* if they differ.

Our next aim is to show that the diagram $\mathcal{D} = \mathcal{D}_n$ has at least one odd crossing. To do this, we introduce some notation.

Define a *region* (sometimes also called a complementary region) to be a connected component of the complement $S^2 \setminus \mathcal{D}$. Observe that a region need not be an open disk, and its boundary need not be connected. For a given region $R$, define $\partial_0 R$ to be the subset of $\partial R \subset \mathcal{D}$ consisting of points $x$ such that every neighbourhood of $x$ intersects at least two distinct regions. In particular, the shortcut $\alpha$ intersects $\partial_0 R$ each time it enters or exits a region $R$. Conversely, each intersection of $\alpha$ with $\partial_0 R$ corresponds to an entering or exiting of $R$.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

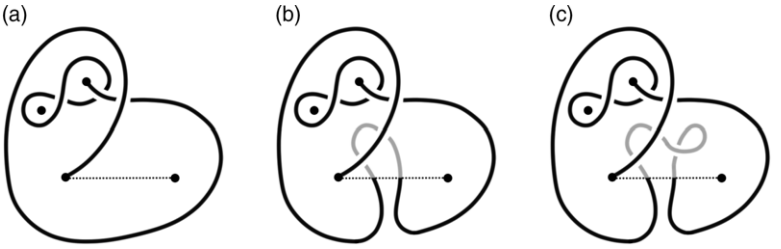



Fig. 5. Figures 5(a), 5(b), 5(c) show an example sequence of diagrams $\mathcal{D}_0$, $\mathcal{D}_1$, $\mathcal{D}_2$ equipped with their $\alpha$-colourings. The shortcut $\alpha$ is indicated with dotted line.

Define a *looped graph* to be the disjoint union of a finite graph with a collection of circles. A looped graph is *Eulerian* if the edges of its graph portion can be partitioned into cycles (i.e. closed paths). It is a classical result that a looped graph is Eulerian if and only if all of its vertices have even valency.

We use the term *singularity* to refer to a pole or crossing of $\mathcal{D}$. We use the term *border* to refer to a portion of a constituent $e$ that connects consecutive singularities along $e$, or all of $e$ if $e$ is a loop constituent with no singularities. Every singularity $s$ has a disk neighbourhood $N_s$ that contains no other singularities. In particular, $\mathcal{D}$ meets $N_s$ at $m$ radii of $N_s$, where $m$ is the number of borders meeting at $s$.

CLAIM 3·4. *The set $\partial_0 R$ forms an Eulerian looped graph whose vertices are singularities and whose edges and circles are borders.*

*Proof.* First note that if any point of a border is in $\partial_0 R$, then the entire border (including singularities at its endpoints) is contained in $\partial_0 R$. It now suffices to show that each singularity $s \in \partial_0 R$ has even valency. Consider $N_s$ and $m$ as defined above. The diagram $\mathcal{D}$ divides $N_s$ into $m$ connected components belonging to (not necessarily distinct) regions $R_1, \ldots, R_m$, some of which are equal to $R$. The radii belonging to $\partial_0 R$ are those that separate an $R$-region and a non-$R$-region. There are an even number of such radii, and the conclusion follows.

CLAIM 3·5. *The diagram $\mathcal{D} = \mathcal{D}_n$ has at least one odd crossing.*

*Proof.* For sufficiently small choices of neighbourhoods $N_A$ and $N_B$, there is a unique *starting region* $R_A$ satisfying $N_A \cap \alpha \setminus \{A\} \subset R_A$. See Figure 6(a). Similarly, there is a unique *ending region* $R_B$ satisfying $N_B \cap \alpha \setminus \{B\} \subset R_B$. By Claim 3·4, we may view $\partial_0 R_A$ as an Eulerian looped graph and partition its edges into cycles. By the assumption that $h_{\mathcal{D}}(A, B) \neq 0$, we have $R_A \neq R_B$. Thus, $\alpha$ intersects $\partial_0 R_A$ an odd number of times, and it follows that $\alpha$ intersects some cycle or circle $C \subset \partial_0 R_A$ an odd number times.

Consider the $\alpha$-colouring of $\mathcal{D}$ restricted to $C$. Recall that constituents must be black near poles, so $C$ does not change colour at poles. Thus, each point of $C$ at which its colour changes falls into one of two categories:

(i) an intersection of $C \setminus \{A, B\}$ with $\alpha$. (Every such intersection yields a colour change.)

(ii) a point of $C$ at which $\mathcal{D}$ has an odd crossing.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

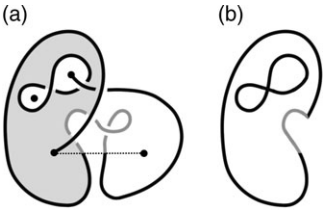



Fig. 6. The diagram from Figure 5(c). Assuming $A$ is the left endpoint pole of the shortcut $\alpha$, Figure 6(a) shows the starting region $R_A$, shaded. Figure 6(b) shows $\partial_0 R_A$. Note that one of the colour changes on the outer cycle of $\partial_0 R_A$ occurs at an odd crossing.

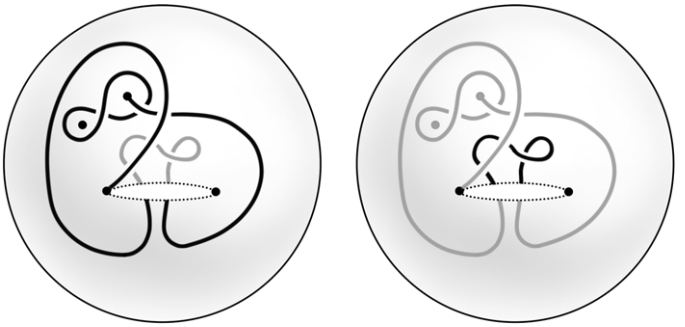


Fig. 7. One possible visualisation of the branched double cover. On the left is $H_1$, on the right is $H_2$, and they are glued together along their boundaries (the dotted circles obtained by cutting a sphere along $\alpha$) such that the top and bottom arcs on the left are identified with the bottom and top arcs on the right. The preimage $p^{-1}(\mathcal{D})$ and its induced $\alpha$-colouring are shown for the diagram $\mathcal{D}$ from Figure 5(c).

As we make a full traversal around $C$, the colour changes an even number of times. An odd number of colour changes fall under (i), so an odd number of colour changes fall under (ii). In particular, $\mathcal{D}$ has at least one odd crossing.

Fix a double branched cover $p\colon S^2 \to S^2$ branched over $\alpha$. Let $e \in C(\mathcal{D}_i)$ be a constituent. (If $e$ has an endpoint at $A$ or $B$, replace $e$ by $e \setminus \{A, B\}$.) The preimage $p^{-1}(e)$ has two connected components $e'$, $e''$, each homeomorphic to $e$. The preimage $p^{-1}(\mathcal{D}_i)$ defines a generalised knotoid diagram on the double cover sphere with poles at the points of $p^{-1}(P(\mathcal{D}_i))$ and constituents $e'$, $e$" for each $e \in C(\mathcal{D}_i)$. Its over/under crossing data is inherited from $\mathcal{D}_i$ in the natural way.

The preimage $p^{-1}(\alpha)$ is a circle, and $S^2 \setminus p^{-1}(\alpha)$ is the disjoint union of two open disks $H_1$ and $H_2$. Let $p_1 : H_1 \to S^2 \setminus \alpha$ and $p_2 : H_2 \to S^2 \setminus \alpha$ be the restrictions of $p$ to $H_1$ and $H_2$. Note that $p_1$ and $p_2$ induce bijections $p_1 : p^{-1}(e) \cap H_1 \to e$ and $p_2 : p^{-1}(e) \cap H_2 \to e$, and recall that $e$ has an $\alpha$-colouring dictated by $C_i$. We colour each point of $p^{-1}(e)$ such that the bijections induced by $p_1$ and $p_2$ are colour-preserving and colour-reversing, respectively. Doing so for each $e \in C(\mathcal{D}_i)$ yields the *induced $\alpha$-colouring* of $p^{-1}(\mathcal{D}_i)$. See Figure 7. Using the induced $\alpha$-colourings of the preimages $p^{-1}(e)$, we define *even* (resp. *odd*) crossings

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

of $p^{-1}(\mathcal{D}_i)$ as crossings where the colours agree (resp. differ). Observe that $p$ defines a parity-preserving, 2-to-1 map from the crossings of $p^{-1}(\mathcal{D}_i)$ to the crossings of $\mathcal{D}_i$.

CLAIM 3·6. *Under the induced $\alpha$-colouring of $p^{-1}(\mathcal{D}_i)$, one of e' and e" is entirely black and the other is entirely gray. Moreover, if e has an endpoint at a pole $Q \notin \{A, B\}$, then the black component of $p^{-1}(e)$ has a corresponding endpoint at $p_1^{-1}(Q) \in H_1$.*

*Proof.* It is clear from the definition of the induced $\alpha$-colouring that $p^{-1}(e)$ has both black and gray points. Each time $e$ crosses $\alpha$ and changes colour, the lift $e_1$ crosses from $H_1$ to $H_2$ or vice versa. Since $p_1$ is colour-preserving and $p_2$ is colour-reversing, $e_1$ does not change colour. It follows that $e_1$ is monochromatic. Similarly, $e_2$ is monochromatic. This proves the first statement.

Now suppose $e$ has an endpoint at a pole $Q \notin \{A, B\}$. By definition of $\alpha$-colouring, there is a neighbourhood $N_Q$ of $Q$ such that $C_i$ assigns black to $e \cap N_Q$. Since $p_1$ is colour-preserving, there is a neighbourhood of $p_1^{-1}(Q)$ in which $p^{-1}(e)$ is black. This proves the second statement.

In light of Claim 3·6, we let $e'$ denote the component of $p^{-1}(e)$ coloured black and define a generalised knotoid diagram $\mathcal{D}_i' \subset p^{-1}(\mathcal{D}_i)$ as follows. The diagram $\mathcal{D}_i'$ has poles at $p^{-1}(A)$, $p^{-1}(B)$, and $p_1^{-1}(Q) \in H_1$ for each $Q \in P(\mathcal{D}_i) \setminus \{A, B\}$, and it has the constituent $e'$ for each $e \in C(\mathcal{D}_i)$. Its over/under crossing data is inherited from $\mathcal{D}_i$ in the natural way. The black portions of Figure 7 show $\mathcal{D}'$, where $\mathcal{D} = \mathcal{D}_2$ is the diagram from Figure 5(c).

CLAIM 3·7. *The crossings of $\mathcal{D}_i'$ are in bijection with the even crossings of $\mathcal{D}_i$.*

*Proof.* Recall that, as a map from the crossings of $p^{-1}(\mathcal{D}_i)$ to the crossings of $\mathcal{D}_i$, the map p preserves parity. Each crossing $c'$ of $\mathcal{D}_i'$ is even (black-black), so $p(c')$ is an even crossing of $\mathcal{D}_i$. Thus $p$ defines map from the crossings of $\mathcal{D}_i'$ to the even crossings of $\mathcal{D}_i$. For any even crossing $c$ in $\mathcal{D}_i$, the preimage $p^{-1}(c) = \{p_1^{-1}(c), p_2^{-1}(c)\}$ consists of two crossings, one black-black and one gray-gray. Only the black-black crossing is a crossing of $\mathcal{D}_i'$. It follows that $p$ gives the desired bijection.

Let $\sim$ denote the equivalence generated by ambient isotopies and Reidemeister moves.

CLAIM 3·8. *For each $0 \leq i < n$, we have $\mathcal{D}_i' \sim \mathcal{D}_{i+1}'$.*

*Proof.* Let $D$ be the disk region where the type $t$ Reidemeister move $\mathcal{D}_i \to \mathcal{D}_{i+1}$ is applied and let $s_1, \ldots, s_t$ be the participating strands, with endpoints on $\partial D$, as in the proof of Claim 3·3. The preimage $p^{-1}(D)$ is the disjoint union of two disk regions $D_1$ and $D_2$, and the move $\mathcal{D}_i \to \mathcal{D}_{i+1}$ lifts to a move $p^{-1}(\mathcal{D}_i) \to p^{-1}(\mathcal{D}_{i+1})$ given by a type $t$ Reidemeister move in each of $D_1$ and $D_2$. By Claim 3·6, each participating strand $s_\ell$ lifts to a monochromatic strand $s_{\ell,1} \subset D_1$ and a monochromatic strand $s_{\ell,2} \subset D_2$ (of the opposite colour). Since $\mathcal{D}_i \to \mathcal{D}_{i+1}$ preserves the colours of the endpoints, the move $p^{-1}(\mathcal{D}_i) \to p^{-1}(\mathcal{D}_{i+1})$ also preserves the colours of the endpoints and thus preserves the colours of $s_{\ell,1}$ and $s_{\ell,2}$. It follows that restricting the move $p^{-1}(\mathcal{D}_i) \to p^{-1}(\mathcal{D}_{i+1})$ to only the black participating strands yields either a Reidemeister move or an ambient isotopy that takes $\mathcal{D}_i'$ to $\mathcal{D}_{i+1}'$. See Figure 8.

We finish the proof of Theorem 3·2. Since $D_0$ does not intersect $\alpha$ except at $A$ and $B$, it follows by construction that $\mathcal{D}_0'$ is contained in $H_1 \cup \{p^{-1}(A), p^{-1}(B)\}$, and that $\mathcal{D}_0$ and $\mathcal{D}_0'$

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

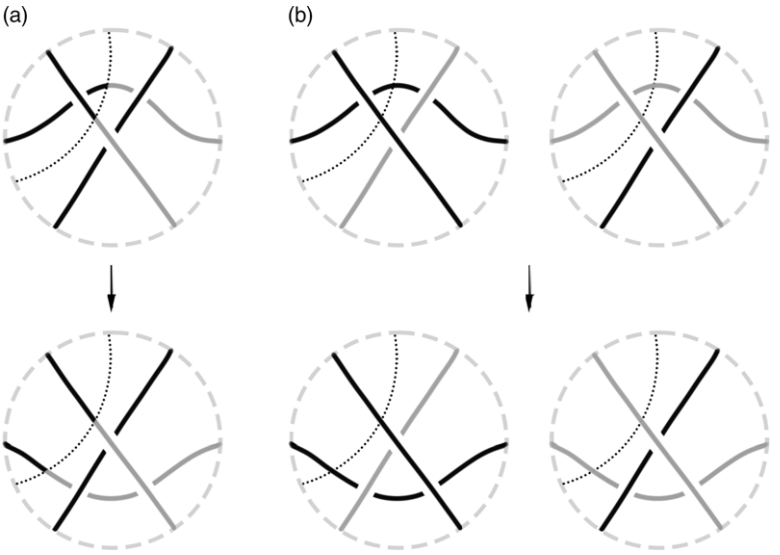



Fig. 8. A Reidemeister move $\mathcal{D}_i \to \mathcal{D}_{i+1}$ (Figure 8(a)) lifts to a move $p^{-1}(\mathcal{D}_i) \to p^{-1}(\mathcal{D}_{i+1})$ in two disk regions (Figure 8(b)). Restricting to only the black strands in Figure 8(b) yields a move $\mathcal{D}'_i \to \mathcal{D}'_{i+1}$. In this example it is an ambient isotopy.

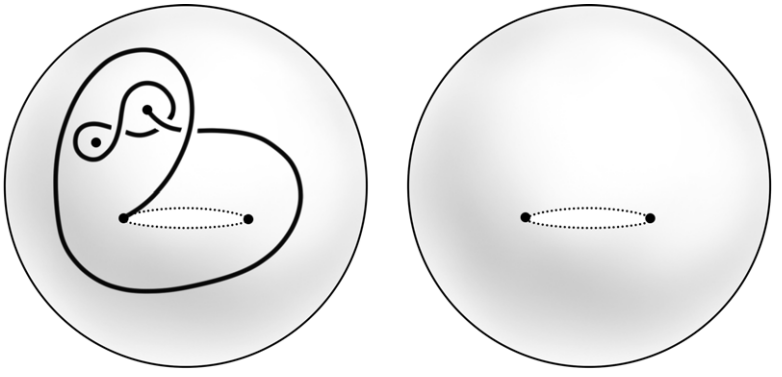


Fig. 9. The diagram $\mathcal{D}'_0$ for the diagram $\mathcal{D}_0$ from Figure 5(a), shown in the same schematic as Figure 7. The diagrams $\mathcal{D}_0$ and $\mathcal{D}'_0$ are ambiently isotopic.

are ambiently isotopic diagrams of $\kappa$. See Figure 9. We now have

$$\mathcal{D}_n \sim \mathcal{D}_0 \sim \mathcal{D}'_0 \sim \mathcal{D}'_n,$$

where the last equivalence follows from Claim 3·8. Hence $\mathcal{D}_n$ and $\mathcal{D}'_n$ are both diagrams of $\kappa$. However, Claim 3·5 and Claim 3·7 together imply that $\mathcal{D}'_n$ has fewer crossings than $\mathcal{D}_n$. This contradicts the assumption that $\mathcal{D}_n$ is a minimal-crossing diagram of $\kappa$.

We conclude this section with some open questions.

*Question 3·9.* Does every generalised knotoid $\kappa$ have a diagram $\mathcal{D}$ that simultaneously realises the height between every pair of poles of $\kappa$?

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

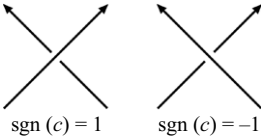



Fig. 10. Positive and negative crossing, respectively.

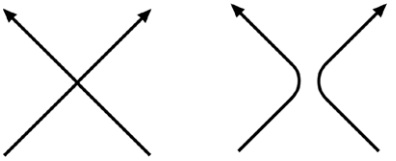


Fig. 11. A crossing and its oriented smoothing. The smoothing does not depend on over/under data.

As a corollary to Theorem 3·2, we obtain an affirmative answer to Question 3·9 in the case that the height spectrum of $\kappa$ is the zero matrix: we may take $\mathcal{D}$ to be a minimal-crossing diagram of $\kappa$.

*Question 3·10.* Let $\kappa$ be a generalised knotoid and let $A, B \in P(\kappa)$ be poles. Does every minimal-crossing diagram of $\kappa$ realise the height between $A$ and $B$?

Question 3·10 asks a generalisation of Theorem 3·2 from height 0 to arbitrary heights, and an affirmative answer to Question 3·10 implies an affirmative answer to Question 3·9.

*Question 3·11.* Does Theorem 3·2 generalise to non-classical generalised knotoids (for example, generalised knotoids on a genus $g$ surface)?

The proof we have given for Theorem 3·2 does not immediately generalise to higher genus surfaces because it uses the fact that there is a double branched cover of the sphere by itself. Generalising this method may require rephrasing the argument in terms of a "virtual theory" for generalised knotoids.

*Question 3·12.* Which matrices of integers are realised as the height spectrum of a generalised knotoid?

## 4. *Index polynomials for generalised knotoids*

### 4·1. *Notation*

Given a crossing $c$ in an oriented diagram, we let $\mathrm{sgn}(c) \in \{1, -1\}$ denote the *sign* of the crossing. See Figure 10. The *oriented smoothing* at $c$ is the smoothing that respects the orientation of the incoming and outgoing strands. See Figure 11. Given a diagram $\mathcal{D}$ and two constituents $\alpha, \beta \in C(\mathcal{D})$, we define their *linking number* by the half-integer

$$\mathrm{lk}(\alpha, \beta) = \frac{1}{2} \sum_{c} \mathrm{sgn}(c),$$

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

where the sum is taken over all crossings between $\alpha$ and $\beta$. Observe that the linking number between $\alpha$ and $\beta$ is an invariant of the generalised knotoid represented by $\mathcal{D}$.

Given two (possibly identical) oriented generalised knotoid diagrams $\mathcal{D}_1, \mathcal{D}_2$ considered on the same surface $\Sigma$ and constituents $\alpha \in C(\mathcal{D}_1)$ and $\beta \in C(\mathcal{D}_2)$, let $T(\alpha, \beta)$ denote the finite subset of $\alpha \cap \beta$ consisting of transverse intersections between $\alpha$ and $\beta$, excluding potential intersections at poles. At each $c \in T(\alpha, \beta)$, the tangent vectors $v_\alpha, v_\beta$ of $\alpha$ and $\beta$ form a basis for the tangent space of $\Sigma$ at $c$. We set $\varepsilon(c) = 1$ if $(v_\beta, v_\alpha)$ is a positively oriented basis, and $\varepsilon(c) = -1$ otherwise. (Equivalently, $\varepsilon(c)$ gives the sign of $c$ if $\beta$ is taken to be the over-strand.) The *algebraic intersection number* of $\alpha$ and $\beta$, denoted $\alpha \cdot \beta$, is defined as

$$\alpha \cdot \beta = \sum_{c \in T(\alpha,\beta)} \varepsilon(c).$$

The algebraic intersection number does not require over/under data. The sign convention we have adopted matches the convention of the algebraic intersection number defined in [**31**].

Proposition 4·1 and Corollary 4·2 are well known.

PROPOSITION 4·1. *If $\alpha$ and $\beta$ are loop constituents on $S^2$ such that all intersections in $\alpha \cap \beta$ are transverse and away from crossings, then $\alpha \cdot \beta = 0$.*

COROLLARY 4·2. *Let $\beta$ be a loop constituent on $S^2$ and fix two points $A$, $B$ in $S^2 \setminus \beta$. Let $\alpha$ be a segment constituent between $A$ and $B$ intersecting $\beta$ transversely and away from crossings. Then $\alpha \cdot \beta$ does not depend on the choice of $\alpha$.*

The *index polynomial* is an invariant for virtual knots introduced by Turaev in [**30**] and Henrich in [**21**]. A similar *affine index polynomial* for oriented virtual knots was defined by Kauffman in [**23**] and extended to knotoids in [**17**]. Kim, Im, and Lee defined an index polynomial for knotoids in [**24**] that is distinct from the affine index polynomial.

These polynomial invariants are based on the following idea: consider a classical crossing $c$ in a diagram of a knotoid or virtual knot and apply an oriented smoothing at $c$. The smoothing produces a diagram with two constituents $\alpha_c, \beta_c$. In the case of a virtual knot, the smoothed diagram is a two-component virtual link, and in the case of a knotoid, it is a loop constituent and a segment constituent. An *intersection index*, denoted $\mathrm{in}(c)$, is computed from the algebraic intersection number of $\alpha_c$ and $\beta_c$, and the polynomial invariant takes the form

$$F_{\mathcal{D}}(t) = \sum_c \mathrm{sgn}(c)(t^{\mathrm{in}(c)} - 1), \tag{4·1·1}$$

where the sum is taken over all crossings of the diagram $\mathcal{D}$. The values of $\mathrm{in}(c)$ computed for the various index polynomials differ only by sign:

(i) in Henrich's index polynomial for virtual knots, the intersection index is taken to be $\mathrm{in}(c) = |\alpha_c \cdot \beta_c|$;

(ii) in Kauffman's affine index polynomial $F^{\mathrm{aff}}(t)$ for oriented virtual knots and knotoids, the constituents $\alpha_c, \beta_c$ are labelled such that $\alpha_c$ contains the incoming portion of the over-strand and $\beta_c$ contains the incoming portion of the under-strand. The intersection

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

index is taken to be $\mathrm{in}(c) = w(c) := \alpha_c \cdot \beta_c$. (This is not Kauffman's original formulation of the affine index polynomial in [23], but it was shown to be equivalent in [11].)

(iii) in Kim, Im and Lee's index polynomial $F^{\mathrm{ind}}(t)$ for knotoids, the constituents $\alpha_c, \beta_c$ are labelled so that $\alpha_c$ is the loop constituent and $\beta_c$ is the segment constituent, and the intersection index is taken to be $\mathrm{in}(c) = \mathrm{ind}(c) := \alpha_c \cdot \beta_c$. See the discussion in [27, section 5].

In [17], it is shown that for a knotoid $k$, the affine index polynomial $F_k^{\mathrm{aff}}(t)$ satisfies $F_k^{\mathrm{aff}}(t) = F_k^{\mathrm{aff}}(t^{-1})$. It follows that the affine index polynomial for knotoids is recoverable from the index polynomial via

$$F_k^{\mathrm{aff}}(t) = \frac{1}{2}(F_k^{\mathrm{ind}}(t) + F_k^{\mathrm{ind}}(t^{-1})).$$

In [24], the index polynomial is strengthened to a two-variable polynomial

$$F_{\mathcal{D}}^{\mathrm{ind}}(s,t) = \sum_{c \in U(\mathcal{D})} \mathrm{sgn}(c)(t^{\mathrm{ind}(c)} - 1) + \sum_{c \in O(\mathcal{D})} \mathrm{sgn}(c)(s^{\mathrm{ind}(c)} - 1).$$

Here, $U(\mathcal{D})$ (resp. $O(\mathcal{D})$) denotes the set of *early undercrossings* (resp. *early overcrossings*) of the diagram $\mathcal{D}$, the crossings $c$ such that $c$ is first encountered as an undercrossing (resp. overcrossing).

4·2. *Generalised index polynomials*

We now define index polynomial invariants for generalised knotoids. Let $\mathcal{D}$ be an oriented, constituent-labelled generalised knotoid diagram. Let $c$ be a crossing, and let $o_c$ and $u_c$ denote the constituents of $\mathcal{D}$ containing the over-strand and under-strand at $c$, respectively. Suppose an oriented smoothing is applied to $c$. Let $\alpha_c$ denote the constituent of the smoothed diagram containing the incoming portion of $o_c$, and let $\beta_c$ denote the constituent of the smoothed diagram containing the incoming portion of $u_c$. Note that $o_c$ and $u_c$ are not necessarily distinct, and $\alpha_c$ and $\beta_c$ are not necessarily distinct.

We introduce variables $t_e$, and $s_e$ for constituents $e$ and $r_{e_1,e_2}$ for pairs of (not necessarily distinct) constituents $(e_1, e_2)$ and define

$$g(c) = r_{o_c,u_c}\left(\prod_{e \in C(\mathcal{D})} t_e^{\alpha_c \cdot e} - 1\right)\left(\prod_{e \in C(\mathcal{D})} s_e^{\beta_c \cdot e} - 1\right).$$

We define the *generalised index polynomial* by

$$G_{\mathcal{D}}(r,s,t) = \sum_c \mathrm{sgn}(c)g(c). \tag{4·2·1}$$

Here $r$ represents a vector whose components are the variables $r_{e_1,e_2}$ in some fixed order, and $s$ and $t$ are similar. Operations on $r$, $s$, and $t$ are applied component-wise. For example, the polynomial $G_{\mathcal{D}}(r, s^{-1}, t)$ is obtained by replacing $s_e$ with $s_e^{-1}$ for all $e \in C(\mathcal{D})$, and $G_{\mathcal{D}}(r, t, s)$ is obtained by swapping $t_e$ and $s_e$ for all $e \in C(\mathcal{D})$. For $r$, we let $\bar{r}$ indicate the operation that swaps $r_{e_1,e_2}$ and $r_{e_2,e_1}$ for all constituents $e_1, e_2$.

THEOREM 4·3. *The polynomial $G_{\mathcal{D}}$ is a generalised knotoid invariant.*

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

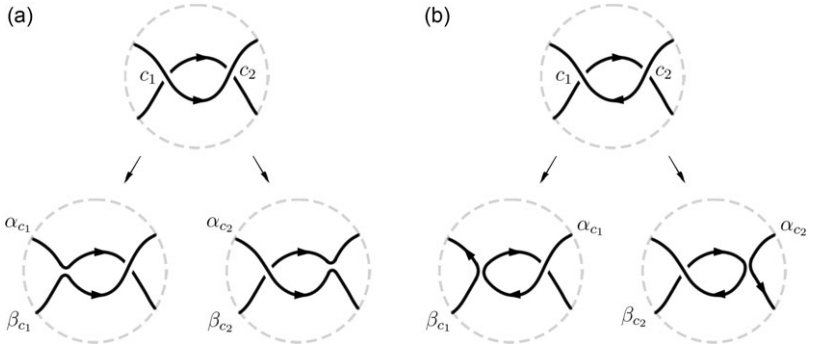



Fig. 12. Two possible configurations for a Type II Reidemeister move. In each of Figures 12(a) and 12(b), the two diagrams obtained by smoothing at $c_1$ and $c_2$ carry the same algebraic intersection data.

*Proof.* Suppose a generalised knotoid $\kappa$ has diagrams $\mathcal{D}, \mathcal{D}'$ related by a Reidemeister move. The crossings of $\mathcal{D}$ and $\mathcal{D}'$ are in natural correspondence except potentially for crossings created or destroyed by a Type I or II Reidemeister move. For corresponding crossings, the values of $\mathrm{sgn}(c)$, $o_c$, $u_c$, $\alpha_c \cdot e$, and $\beta_c \cdot e$ are identical for any $e \in C(\kappa)$. It remains to verify that the crossings created or destroyed by a Type I or Type II Reidemeister move do not contribute to $G_{\mathcal{D}}$.

(i) Consider a Type I Reidemeister move, and assume without loss of generality that a crossing $c$ is created. Observe that at least one of $\alpha_c, \beta_c$ is a loop with no crossings, so $g(c) = 0$.

(ii) Consider a Type II Reidemeister move, and assume without loss of generality that two crossings $c_1, c_2$ are created. Observe that $c_1$ and $c_2$ have opposite sign and satisfy $(o_{c_1}, u_{c_1}) = (o_{c_2}, u_{c_2})$. Outside of the neighbourhood where the Reidemeister move is performed, the constituents $\alpha_{c_1}$ and $\alpha_{c_2}$ are identical, and the constituents $\beta_{c_1}$ and $\beta_{c_2}$ are identical. Inside the neighbourhood, the oriented smoothings at $c_1$ and $c_2$ each leave one crossing, the two of which are indistinguishable by algebraic intersection numbers. See Figure 12. Thus $\alpha_{c_1} \cdot e = \alpha_{c_2} \cdot e$ and $\beta_{c_1} \cdot e = \beta_{c_2} \cdot e$ for any $e$, and it follows that $g(c_1) = g(c_2)$. Finally, $\mathrm{sgn}(c_1)g(c_1) + \mathrm{sgn}(c_2)g(c_2) = 0$, as desired.

The generalised index polynomial extends the affine index polynomial in the following sense.

PROPOSITION 4·4. *Two knotoids have identical affine index polynomials if and only if they have identical generalised index polynomials.*

*Proof.* If $\mathcal{D}$ is a spherical or planar knotoid diagram, the generalised index polynomial becomes a three-variable polynomial

$$G_{\mathcal{D}}(r, s, t) = \sum_c \mathrm{sgn}(c) r (t^{\alpha_c \cdot e} - 1)(s^{\beta_c \cdot e} - 1).$$

Here, $e$ is the single segment constituent of $\mathcal{D}$. Because self-crossings of $\alpha_c$ create two intersections of $\alpha_c$ with $e$ which cancel each other out, all contributions to $\alpha_c \cdot e$ come about

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

through intersections of $\alpha_c$ with $\beta_c$. Hence, $\alpha_c \cdot e = \alpha_c \cdot \beta_c = -\beta_c \cdot e = w(c)$. It follows that the affine index polynomial $F^{\text{aff}}_{\mathcal{D}}(t)$ is recovered from $G_{\mathcal{D}}(s,t)$ by

$$F^{\text{aff}}_{\mathcal{D}}(t) = -G_{\mathcal{D}}(1,0,t).$$

Conversely, $G_{\mathcal{D}}(r,s,t)$ is recovered from $F^{\text{aff}}_{\mathcal{D}}(t)$ by

$$G_{\mathcal{D}}(r,s,t) = r(F^{\text{aff}}_{\mathcal{D}}(ts^{-1}) - F^{\text{aff}}_{\mathcal{D}}(t) - F^{\text{aff}}_{\mathcal{D}}(s^{-1})).$$

By Proposition 4·1, the generalised index polynomial is trivial for any classical link.

If a generalised knotoid diagram $\mathcal{D}$ is the disjoint union of diagrams $\mathcal{D}_1$ and $\mathcal{D}_2$ with no crossings occuring between constituents of $\mathcal{D}_1$ and $\mathcal{D}_2$, then the generalised index polynomial $G_{\mathcal{D}}$ provides no information on their relative positioning. For example, a valency-zero pole can be moved to any region of a diagram without changing the generalised index polynomial.

We extend the generalised index polynomial for pole-labelled generalised knotoids with at least one pole. Given a diagram $\mathcal{D}$ and a fixed pole $P \in P(\mathcal{D})$, we choose oriented shortcuts $\{\gamma_Q \mid Q \in P(\mathcal{D})\}$ from $P$ to each pole in $\mathcal{D}$. For a constituent $e$, let $\mathbf{1}_L(e)$ denote the indicator function that is 1 if $e$ is a loop constituent and 0 otherwise. We introduce variables $a_Q, b_Q$ for each $Q \in P(\mathcal{D})$ and define, for each crossing $c$, the polynomial

$$h_P(c) = \left(\prod_{Q \in P(\mathcal{D})} a_Q^{\gamma_Q \cdot \alpha_c}\right)^{\mathbf{1}_L(\alpha_c)} \left(\prod_{Q \in P(\mathcal{D})} b_Q^{\gamma_Q \cdot \beta_c}\right)^{\mathbf{1}_L(\beta_c)}.$$

We define the *base-pointed index polynomial* by

$$G_{\mathcal{D},P}(r,s,t,a,b) = \sum_c \operatorname{sgn}(c) g(c) h_P(c). \tag{4·2·2}$$

The notational remarks following Equation (4·2·1) apply to the variables $a$ and $b$ as well.

By Corollary 4·2, the base-pointed index polynomial is independent of the choice of shortcuts $\gamma_Q$, and the proof that $G_{\mathcal{D},P}$ is a generalised knotoid invariant is identical to the proof of Theorem 4·3. The choice of the base-point pole $P$ is immaterial. If $P'$ is another pole, the polynomial $G_{\mathcal{D},P'}$ is obtained from $G_{\mathcal{D},P}$ in the following way. Since we can take as shortcut from $P'$ to any other pole $Q$ the path obtained by reversing the orientation of $\gamma_{P'}$ and then adding to it the shortcut from $P$ to $Q$, this modifies each monomial in $G_{\mathcal{D},P}$ by subtracting the exponent of $a_{P'}$ from the exponent of each $a_Q$ and subtracting the exponent of $b_{P'}$ from the exponent of each $b_Q$.

*Example 4·5.* The generalised knotoid $\kappa$ in Figure 13(a) has

$$G_{\kappa,P} = r_{e_1,e_1}((s_{e_2}^2 - 1)a_Q^{-1}a_R^{-1}b_Q b_R^{-1} - (t_{e_2} - 1)(s_{e_2} - 1)a_Q^{-1}a_R^{-2}b_Q)$$

and

$$G_{\kappa,R} = r_{e_1,e_1}((s_{e_2}^2 - 1)a_P b_P b_Q^2 - (t_{e_2} - 1)(s_{e_2} - 1)a_P^2 a_Q^1 b_Q).$$

*Example 4·6.* The generalised knotoid $\kappa$ in Figure 13(b) has $G_{\kappa,P} = 0$, so the base-pointed generalised index polynomial cannot distinguish $\kappa$ from a diagram with no crossings. However, every diagram of $\kappa$ has at least one crossing since $\operatorname{lk}(e_1, e_2) = 1/2$.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

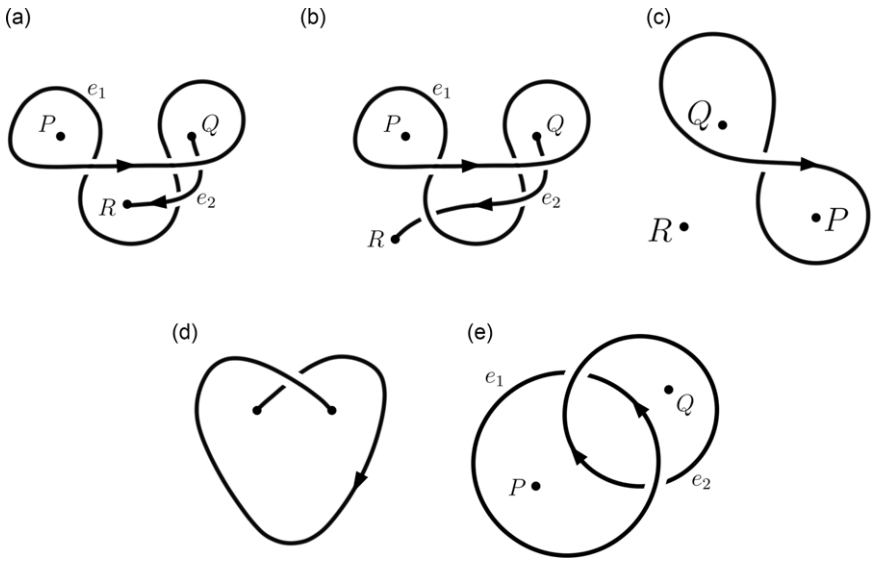



Fig. 13. Finding polynomials of certain generalised knotoids.

We now define a variant of the base-pointed index polynomial. With all notation as before, define

$$\tilde{g}(c) = r_{o_c,u_c} \prod_{e \in C(\mathcal{D})} t_e^{\alpha_c \cdot e} s_e^{\beta_c \cdot e}$$

and

$$\tilde{h}_P(c) = \left( \mathbf{1}_L(\alpha_c) \prod_{Q \in P(\mathcal{D})} a_Q^{\gamma_Q \cdot \alpha_c} - 1 \right) \left( \mathbf{1}_L(\beta_c) \prod_{Q \in P(\mathcal{D})} b_Q^{\gamma_Q \cdot \beta_c} - 1 \right).$$

We define the *pole-centric base-pointed index polynomial* by

$$\tilde{G}_{\mathcal{D},P}(r, s, t, a, b) = \sum_c \operatorname{sgn}(c) \tilde{g}(c) \tilde{h}_P(c).$$

The proof that $\tilde{G}_{\mathcal{D},P}$ is a generalised knotoid invariant is nearly identical to the proof of Theorem 4·3. The remarks following Equation (4·2·2) for $G_{\mathcal{D},P}$ hold for $\tilde{G}_{\mathcal{D},P}$ as well.

In analogy to Proposition 4·4, the pole-centric base-pointed index polynomial extends the strengthened index polynomial $F^{\text{ind}}(s, t)$ for knotoids. Suppose $k_1$ and $k_2$ are two spherical knotoids, each with poles labelled $L$ and $H$ and a constituent oriented from $L$ to $H$. In what follows, we take $\alpha_c$ and $\beta_c$ to be as in the definition of the generalised index polynomial as in Section 4·2, and we take ind(c) to be as defined in Kim, Im, and Lee's index polynomial in point (iii) under Equation 4·1·1.

PROPOSITION 4·7. *If* $\tilde{G}_{k_1,L} = \tilde{G}_{k_2,L}$, *then* $F^{\text{ind}}_{k_1}(s, t) = F^{\text{ind}}_{k_1}(s, t)$.

*Proof.* For a knotoid diagram $\mathcal{D}$ oriented from $L$ to $H$, the pole-centric base-pointed index polynomial based at $L$ is a five-variable polynomial $\tilde{G}_{\mathcal{D},L}(r, s, t, a_H, b_H)$. Set $r = s = t = 1$ to

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

obtain a two-variable polynomial $\tilde{G}_{\mathcal{D},L}(a_H, b_H)$. For each crossing $c$ in $\mathcal{D}$, exactly one of $\alpha_c, \beta_c$ is a loop constituent. The crossings for which $\alpha_c$ is the loop constituent is precisely the set of early undercrossings, and the crossings for which $\beta_c$ is the loop constituent is precisely the set of early overcrossings. Thus

$$\tilde{G}_{\mathcal{D},L}(a_H, b_H) = -\sum_{c\in U(k)} \operatorname{sgn}(c)(a_H^{\gamma_H\cdot\alpha_c} - 1) - \sum_{c\in O(\mathcal{D})} \operatorname{sgn}(c)(b_H^{\gamma_H\cdot e} - 1).$$

Suppose $\alpha_c$ is the loop constituent. Observe that $\beta_c$ defines an oriented shortcut from $L$ to $H$. It follows from Corollary 4·2 that

$$\gamma_H \cdot \alpha_c = \beta_c \cdot \alpha_c = -\text{ind(c)}.$$

Similarly, if $\beta_c$ is the loop constituent, then $\gamma_H \cdot \beta_c = \alpha_c \cdot \beta_c = -\text{ind}(c)$. Thus we obtain

$$\begin{aligned} -\tilde{G}_{\mathcal{D},L}(s^{-1}, t^{-1}) &= \sum_{c\in U(\mathcal{D})} \operatorname{sgn}(c)(s^{\text{ind}(c)} - 1) + \sum_{c\in O(\mathcal{D})} \operatorname{sgn}(c)(t^{\text{ind}(c)} - 1) \\ &= F^{\text{ind}}_{\mathcal{D}}(s, t). \end{aligned}$$

In particular, the strengthened index polynomial is recoverable from the pole-centric base-pointed index polynomial, and the conclusion follows.

*Example 4·8.* The staked knot $\kappa$ from Figure 13(c) has $\tilde{G}_{\kappa,P} = a_Q b_Q b_R$. On the other hand, the base-pointed index polynomial $G_{\kappa,P}$ is trivial for any staked link $\kappa$ by Proposition 4·1.

*Question 4·9.* Does there exist a pole-labelled, constituent-labelled, oriented generalised knotoid $\kappa$ with a pole $P$ such that $\tilde{G}_{\kappa,P} = 0$ but $G_{\kappa,P} \neq 0$?

The index polynomials of the form given by Equation (4·1·1) cannot distinguish between inequivalent planar knotoids with diagrams that are equivalent when considered on the sphere. By viewing a planar knotoid as a generalised knotoid with three poles $L$, $H$, and $\infty$ and a constituent oriented from $L$ to $H$, as in Example 2·3, the base-pointed index polynomial and its pole-centric variant yield new invariants for planar knotoids. For example, the planar knotoid $k$ in Figure 13(d) has trivial index polynomial but has $\tilde{G}_{k,\infty} = a_L a_H$.

Following the notation of [**31**] and [**27**], we extend the basic knotoid involutions to generalised knotoids and state their effects on the generalised index polynomials. Given a diagram $\mathcal{D}$, we let $\text{mir}(\mathcal{D})$ denote the diagram obtained by toggling all over/under crossing data, and we let $\text{sym}(\mathcal{D})$ denote the diagram obtained by reflecting across a great circle of $S^2$ (and preserving crossing data). We let rot denote the composition $\text{mir} \circ \text{sym} = \text{sym} \circ \text{mir}$. For an oriented diagram $\mathcal{D}$, we let $\text{rev}(\mathcal{D})$ denote the diagram obtained by reversing the orientation on all constituents. It is clear that mir, sym, rot, and rev define involutions on the set of generalised knotoids, possibly labelled or oriented. Note that mir, sym, and rot preserve the underlying directed graph of an oriented generalised knotoid, while rev may not.

PROPOSITION 4·10. *Let $\kappa$ be pole-labelled, constituent-labelled, oriented generalised knotoid and let $P \in P(\kappa)$. The base-pointed index polynomial satisfies*:

*(a)* $G_{\text{mir}(\kappa),\text{mir}(P)}(r, s, t, a, b) = -G_{\kappa,P}(\bar{r}, t, s, b, a)$;

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

(*b*) $G_{\mathrm{sym}(\kappa),\mathrm{sym}(P)}(r,s,t,a,b) = -G_{\kappa,P}(r,s^{-1},t^{-1},a^{-1},b^{-1})$;

(*c*) $G_{\mathrm{rot}(\kappa),\mathrm{rot}(P)}(r,s,t,a,b) = G_{\kappa,P}(\bar{r},t^{-1},s^{-1},b^{-1},a^{-1})$.

*Identical relationships hold for the pole-centric variant.*

*Proof.* Part (*a*) follows from the fact that, for each crossing $c$, the operation mir negates $\mathrm{sgn}(c)$ and swaps $\alpha_c$ and $\beta_c$. Part (*b*) follows from the fact that sym negates both the sign of each crossing and all algebraic intersection numbers. Part (*c*) follows from (*a*) and (*b*).

Thus the base-pointed index polynomials may be used to distinguish between $\kappa$ and its images under these three basic involutions. But no such relationship holds for rev.

The base-pointed index polynomials also yield lower bounds on the height between poles in a generalised knotoid. Given a Laurent polynomial $F \in \mathbb{Z}[x_1^{\pm 1},\dots,x_n^{\pm 1}]$, we let $[x_i]F \in \mathbb{Z}[x_1^{\pm 1},\dots,x_{i-1}^{\pm 1},x_{i+1}^{\pm 1},\dots,x_n^{\pm 1}]$ denote the coefficient of $x_i$ in $F$. For any subset $S = \{x_{i_1},\dots,x_{i_m}\}$ of the variables and a monomial $M = \prod_{j=1}^{n} x_j^{u_j}$, we define $\deg_S(M) = \sum_{k=1}^{m} |u_{i_k}|$, and we let $\deg_S(F)$ denote the maximum of $\deg_S(M)$ over all monomials $M$ with nonzero coefficient $F$. (We omit the set notation from $S$ for brevity.)

PROPOSITION 4·11. *Let $\kappa$ be a pole-labelled, constituent-labelled, oriented generalised knotoid and let $P, Q \in P(\kappa)$ be poles. Then*

$$h_\kappa(P,Q) \geq \sum_{e \in C(\kappa)} \deg_{a_Q,b_Q}([r_{e,e}]G_{\kappa,P}).$$

*An identical inequality holds with $G_{\kappa,P}$ replaced by $\tilde{G}_{\kappa,P}$.*

*Proof.* The following argument holds for both $G_{\kappa,P}$ and $\tilde{G}_{\kappa,P}$. Let $\mathcal{D}$ be any diagram representing $\kappa$ and let $\gamma$ be any shortcut oriented from $P$ to $Q$. It suffices to show that, for any constituent $e \in C(\mathcal{D})$, the number of intersections $\#(\gamma \cap e)$ of $\gamma$ with $e$ away from poles is at least $d := \deg_{a_Q,b_Q}([r_{e,e}]G_{\kappa,P})$.

If $d = 0$, this is trivial. Assuming $d > 0$, the diagram $\mathcal{D}$ must contain a crossing $c$ with $o_c = u_c = e$ and $\mathbf{1}_L(\alpha_c)|\gamma_Q \cdot \alpha_c| + \mathbf{1}_L(\beta_c)|\gamma_Q \cdot \beta_c| = d$. Note that $\alpha_c \neq \beta_c$, and

$$\begin{aligned}\#(\gamma \cap e) &= \#(\gamma \cap \alpha_c) + \#(\gamma \cap \beta_c)\\ &\geq |\gamma_Q \cdot \alpha_c| + |\gamma_Q \cdot \beta_c|\\ &\geq d.\end{aligned}$$

Proposition 4·11 implies that the generalised knotoid $\kappa_1$ from Example 4·5 has $h_{\kappa_1}(P,Q) = h_{\kappa_1}(P,R) = h_{\kappa_1}(R,Q) = 2$, and the generalised knotoid $\kappa_2$ from Example 13(c) has $h_{\kappa_2}(P,Q) = 2$ and $h_{\kappa_2}(P,R) = h_{\kappa_2}(Q,R) = 1$.

For certain generalised knotoids $\kappa$, the polynomials $G_{\kappa,P}$ or $\tilde{G}_{\kappa,P}$ give more information than the bound stated in Proposition 4·11, as the bound only considers intersections of a shortcut $\gamma$ with loops $\alpha_c, \beta_c$ that result from smoothing at a self-intersection of a constituent $e$. The generalised knotoid $\kappa$ in Figure 13(e) has $\tilde{G}_{\kappa,P} = 2r_{e_1,e_2}a_Q^{-2}$, and an argument similar to the proof of Proposition 4·11 shows that $h_\kappa(P,Q) = 2$ despite the fact that Proposition 4·11 only yields the trivial bound $h_\kappa(P,Q) \geq 0$.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

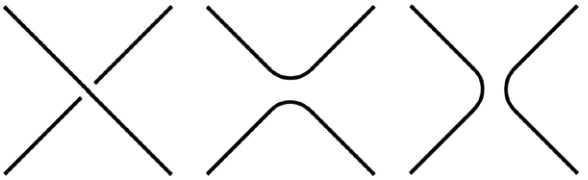


Fig. 14. A crossing, its $A$-type smoothing, and its $B$-type smoothing, respectively.

We remark on non-constituent-labelled generalised knotoids. If $\kappa$ is simply an oriented, generalised knotoid, we obtain a two-variable generalised index polynomial $G_\kappa(s, t)$ by computing the generalised index polynomial of $\kappa$ with constituents arbitrarily labelled, and then setting $r_{e_1,e_2} = 1$, $t_{e_1} = t$, and $s_{e_1} = s$ for all $e_1, e_2 \in C(\kappa)$. In the case that $\kappa$ is pole-labelled, we again label the constituents arbitrarily and compute the base-pointed index polynomial (or the pole-centric variant). We then re-index the variables $r_{e_1,e_2}, t_e, s_e$ as follows. Each subscript $e$ corresponding to a segment constituent $e$ oriented from pole $P$ to pole $Q$ is replaced by $(P, Q)$, and all subscripts corresponding to loop constituents are replaced by a single symbol $L$.

## 5. *A bracket polynomial for generalised knotoids*

### 5·1. *Generalised bracket polynomial*

We define an extension of the Kauffman bracket polynomial for oriented, pole-labelled, edge-labelled generalised knotoids.

Given a generalised knotoid diagram $\mathcal{D}$ and a crossing $c$ in $\mathcal{D}$, a *smoothing at c* is one of the two operations pictured in Figure 14, and is either $A$-type or $B$-type. Observe that a smoothing preserves the number of segment constituents. A *state s* of $\mathcal{D}$ is a diagram obtained by smoothing at each of the crossings of $\mathcal{D}$. Let $\sigma(s)$ denote the number of $A$-type smoothings of $s$ minus the number of $B$-type smoothings of $s$. Observe that $s$ has no crossings and has the same number of segment constituents as $D$, though in general, there is no natural correspondence between the segment constituents of $D$ and $s$.

For every state $s$, we define auxiliary polynomials $G_s$, $L_s$, and $E_s$ to record the data of the state's underlying graph, loop constituents, and segment constituents, respectively. They are as follows.

Introduce variables $\lambda_{\{P,Q\}}$ for every unordered pair $\{P, Q\}$ of not necessarily distinct poles $P, Q \in P(\mathcal{D})$ and let $s(P, Q)$ denote the number of edges between $P$ and $Q$ in the underlying graph of $s$. Define the polynomial

$$G_s = \prod_{\{P,Q\}} \lambda_{\{P,Q\}}^{s(P,Q)}.$$

Let $2^{P(\mathcal{D})}/\sim$ denote the set of unordered bipartitions of the poles of $\mathcal{D}$. (The equivalence relation $\sim$ identifies each subset of $P(\mathcal{D})$ with its complement.) Each loop constituent of $s$ separates the sphere into two regions and determines an element of $2^{P(\mathcal{D})}/\sim$. For each $U \in 2^{P(\mathcal{D})}/\sim$, let $n_U(s)$ denote the number of loop constituents of $s$ corresponding to $U$. Let $n_0(s) := n_{[\emptyset]}(s)$ denote the number of loops that create a region with no poles. We call such loops *nullhomotopic*. We define the polynomial $L_s$ in variables $\{x_U \mid U \in 2^{P(\mathcal{D})}/\sim, U \neq [\emptyset]\}$ by

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

$$L_s = \prod_{\substack{U \in 2^{P(\mathcal{D})}/\sim \\ U \neq [\emptyset]}} x_U^{n_U(s)}.$$

Fix a set of shortcuts $\{\alpha_{P,Q} \mid P, Q \in P(\mathcal{D})\}$ between every pair of poles of $\mathcal{D}$, where $\alpha_{P,Q}$ is oriented from $P$ to $Q$. Suppose $e, f$ are segment constituents, each belonging to either $\mathcal{D}$ or one of its states. Given orientations on $e$ and $f$, define

$$\mu_{e,f}(P, Q) = e \cdot \alpha_{P,Q} - f \cdot \alpha_{P,Q}.$$

We say $e$ and $f$ are *aligned* and write $e \parallel f$ if their endpoint poles coincide and they are identical (as unoriented constituents) in a neighbourhood of each of the endpoint poles.

Suppose $e \in E(\mathcal{D})$ is oriented from pole $e_0$ to pole $e_1$. For a given state $s$, there is at most one $e_s \in E(s)$ aligned with $e$. If such an $e_s$ exists, we give it the orientation that agrees with $e$ near $e_0$ and $e_1$. Define the polynomial $E_s$ in variables $\{y_{e,P} \mid e \in E(\mathcal{D}), P \in P(\mathcal{D})\}$ by

$$E_s = \prod_{\substack{e \parallel e_s \\ p \in P(\mathcal{D})}} y_{e,P}^{\mu_{e_s,e}(e_0,P)},$$

where the product is taken over all $e \in E(\mathcal{D})$, $P \in P(\mathcal{D})$ such that $s$ has a segment constituent $e_s \in E(s)$ with $e \parallel e_s$.

Finally, define the *generalised bracket polynomial* $\ll \mathcal{D} \gg$ as the Laurent polynomial in variables $A$, $\{\lambda_{\{P,Q\}}\}$, $\{x_U\}$, and $\{y_{e,P}\}$ given by

$$\ll \mathcal{D} \gg = \sum_s A^{\sigma(s)}(-A^2 - A^{-2})^{n_0(s)} \cdot G_s \cdot L_s \cdot E_s. \tag{5·1·1}$$

We show in Lemma 5·1 that $E_s$ is well-defined, independent of the choice of shortcuts $\{\alpha_{P,Q}\}$. The lemma should be thought of as an analog of Corollary 4·2 for closed (but not generic) curves. We argue similarly to the proof given for lemma 8·1 in [**31**]. (See also the *shortcut moves* of [**27**].)

LEMMA 5·1. *Suppose $e$, $f$ are segment constituents, each belonging to either $\mathcal{D}$ or one of its states, with $e \parallel f$ and orientations that agree near their endpoints. Let $P, Q \in P(\mathcal{D})$. Then $\mu_{e,f}(P, Q)$ does not depend on the choice of shortcut $\alpha_{P,Q}$.*

*Proof.* Observe that any shortcut between $P$ and $Q$ is obtainable from $\alpha_{P,Q}$ by a sequence of local transformations of the following types:

(i) pulling $\alpha_{P,Q}$ across a strand of $\mathcal{D}$ (c.f. Type II Reidemeister);
(ii) pulling $\alpha_{P,Q}$ across a crossing of $\mathcal{D}$ (c.f. Type III Reidemeister);
(iii) pulling $\alpha_{P,Q}$ across a pole other than $P$ or $Q$ (c.f. pole slide move);
(iv) pulling $\alpha_{P,Q}$ across a strand of $\mathcal{D}$, near $P$ or $Q$ (c.f. pole twist move).

Observe that (i) and (ii) each preserve $e \cdot \alpha_{P,Q}$ and $f \cdot \alpha_{P,Q}$, while (iii) and (iv) each preserve $e \cdot \alpha_{P,Q} - f \cdot \alpha_{P,Q}$ by the alignment condition.

COROLLARY 5·2. *With $e, f$ as in Lemma 5·1, we have $\mu_{e,f}(P, P) = 0$.*

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

The *normalised generalised bracket polynomial* $\ll \mathcal{D} \gg_*$ is defined by

$$\ll \mathcal{D} \gg_* = (-A^3)^{-\mathrm{wr}(\mathcal{D})} \ll \mathcal{D} \gg .$$

Here, the *writhe* $\mathrm{wr}(\mathcal{D})$ of the oriented generalised knotoid diagram $\mathcal{D}$ is defined by the sum of $\mathrm{sgn}(c)$ over all crossings $c$ in $\mathcal{D}$.

PROPOSITION 5·3. *The polynomial $\ll \mathcal{D} \gg_*$ is a generalised knotoid invariant.*

*Proof.* Observe that the generalised bracket polynomial satisfies the usual disjoint union and skein relations

$$\ll \mathcal{D} \cup \bigcirc \gg = (-A^2 - A^{-2}) \ll \mathcal{D} \gg \quad \text{and} \quad \ll \mathcal{D} \gg = A \ll \mathcal{D}_+ \gg + A^{-1} \ll \mathcal{D}_- \gg,$$

where $\mathcal{D}_+$ and $\mathcal{D}_-$ denote the diagrams obtained by performing an $A$-smoothing and a $B$-smoothing, respectively, at a particular crossing of $\mathcal{D}$. As with the traditional bracket polynomial for knots, this implies that the normalised polynomial $\ll \mathcal{D} \gg_*$ is preserved by Reidemeister moves.

In analogy to Proposition 4·11 for the base-pointed index polynomial, Proposition 5·5 gives lower bounds on heights between poles in terms of the generalised bracket polynomial. Before stating the bounds, we record a lemma.

LEMMA 5·4. *Let $e,f$ be as in Lemma 5·1 and let $P, Q, R \in P(\mathcal{D})$. Then*

$$\mu_{e,f}(P, Q) + \mu_{e,f}(Q, R) = \mu_{e,f}(P, R).$$

*Proof.* If $P = Q$ or $Q = R$, the result follows from Corollary 5·2, so assume this is not the case. Without loss of generality, we choose $\alpha_{P,Q}$ and $\alpha_{Q,R}$ with no intersections other than poles; by Lemma 5·1, this does not change any of the values in question. Then we obtain a shortcut $\tilde{\alpha}_{P,R}$ by concatenating $\alpha_{P,Q}$ with $\alpha_{Q,R}$ and perturbing the resulting curve in a neighbourhood of $Q$ so that, locally, it neither passes through $Q$ nor intersects $e$ or $f$. (This is possible even if $e$ and $f$ have an endpoint at $Q$ since $e \parallel f$.) Observe that $e \cdot \alpha_{P,Q} + e \cdot \alpha_{Q,R} = e \cdot \tilde{\alpha}_{P,R}$ and $f \cdot \alpha_{P,Q} + f \cdot \alpha_{Q,R} = f \cdot \tilde{\alpha}_{P,R}$, whence

$$\begin{aligned} \mu_{e,f}(P, Q) + \mu_{e,f}(Q, R) &= (e \cdot \alpha_{P,Q} - f \cdot \alpha_{P,Q}) + (e \cdot \alpha_{Q,R} - f \cdot \alpha_{Q,R}) \\ &= e \cdot \tilde{\alpha}_{P,Q} - f \cdot \tilde{\alpha}_{P,Q} \\ &= \mu_{e,f}(P, R), \end{aligned}$$

where the last equality follows from Lemma 5·1.

Following the notation of [**31**], we define the *$x_i$-span*, denoted $\mathrm{spn}_{x_i}(F)$, of a Laurent polynomial $F \in \mathbb{Z}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ as the non-negative difference between the largest and smallest exponents of $x_i$ among monomials with nonzero coefficients in $F$ (by convention, $\mathrm{spn}_{x_i}(0) = -\infty$). Given a generalised knotoid $\kappa$ and poles $P, Q \in P(\kappa)$, we let $\{P \mid Q\}$ denote the set of variables $x_U$ with $U \in 2^{P(\kappa)}/\sim$ such that $P$ and $Q$ are separated by $U$.

PROPOSITION 5·5. *Let $\kappa$ be a pole-labelled, constituent-labelled, oriented generalised knotoid and let $P, Q \in P(\kappa)$ be poles. Then*:

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

(*a*) $h_\kappa(P,Q) \geq \deg_{\{P|Q\}}(\ll \kappa \gg_*)$;

(*b*) $h_\kappa(P,Q) \geq \frac{1}{2}\max_{e\in E(\kappa)} \deg_{y_{e,P}}\left(\ll \kappa \gg_* \Big|_{y_{e,Q}=y_{e,P}^{-1}}\right)$;

(*c*) $h_\kappa(P,Q) \geq \frac{1}{2}\max_{e\in E(\kappa)} \mathrm{spn}_{y_{e,P}}\left(\ll \kappa \gg_* \Big|_{y_{e,Q}=y_{e,P}^{-1}}\right)$.

*Proof.* Let $\mathcal{D}$ be any diagram of $\kappa$ and let $\gamma$ be any shortcut oriented from $P$ to $Q$. Let $d_a$ denote the right side of the inequality in (*a*). Then there is a state $s$ of $\mathcal{D}$ with $d_a = \sum_{U\in\{P|Q\}} n_U(s)$. Since $\gamma$ has at least one intersection with each loop constituent of $s$ that separates $P$ and $Q$, part (*a*) follows.

We now prove (*b*) and (*c*). First, suppose $e, f$ are oriented segment constituents, each belonging to either $\mathcal{D}$ or a state of $\mathcal{D}$. Each point of intersection of $\gamma$ with $e \cup f$ away from poles is counted by $e\cdot\gamma - f\cdot\gamma$ with multiplicity in $\{-2,-1,0,1,2\}$. It follows that the number of such points, denoted $\#(\gamma \cap (e\cup f))$, satisfies

$$\#(\gamma \cap (e\cup f)) \geq \frac{1}{2}|e\cdot\gamma - f\cdot\gamma|. \tag{5·1·2}$$

Let $d_b$ denote the right-hand side of the inequality in (*b*). Suppose $d_b > 0$. Then there is a segment constituent $e \in E(\mathcal{D})$, a state $s$, and a segment constituent $e_s \in E(s)$ such that $e \parallel e_s$ and

$$\begin{aligned} d_b &= \frac{1}{2}|\mu_{e_s,e}(e_0,P) - \mu_{e_s,e}(e_0,Q)| \\ &= \frac{1}{2}|\mu_{e_s,e}(P,Q)| \\ &= \frac{1}{2}|e_s\cdot\gamma - e\cdot\gamma|, \end{aligned}$$

where the second equality uses Lemma 5·4 and the third uses Lemma 5·1. It follows from Inequality (5·1·2) that $\#(\gamma\cap(e_s\cup e)) \geq d_b$, and part (*b*) follows.

Finally, let $d_c$ denote the right-hand side of the inequality in (*c*). Assume $d_c > d_b$. Then there is a segment constituent $e \in E(\mathcal{D})$, two states $s_1, s_2$, and two segment constituents $f_1 \in E(s_1), f_2 \in E(s_2)$ such that $e \parallel f_1 \parallel f_2$ and

$$\begin{aligned} d_c &= \frac{1}{2}|(\mu_{f_1,e}(e_0,P) - \mu_{f_1,e}(e_0,Q)) - (\mu_{f_2,e}(e_0,P) - \mu_{f_2,e}(e_0,Q))| \\ &= \frac{1}{2}|-\mu_{f_1,e}(P,Q) + \mu_{f_2,e}(P,Q)| \\ &= \frac{1}{2}|\mu_{f_2,f_1}(P,Q)| \\ &= \frac{1}{2}|f_2\cdot\gamma - f_1\cdot\gamma|, \end{aligned}$$

where the second equality uses Lemma 5·4 and the fourth uses Lemma 5·1. It follows from Inequality (5·1·2) that $\#(\gamma\cap(f_1\cup f_2)) \geq d_c$, and part (*c*) follows.

*Example 5·6.* Let $\kappa$ denote the generalised knotoid represented by the diagram in Figure 15. Computations yield

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

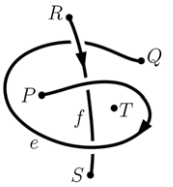



Fig. 15. An oriented generalised knotoid with five poles $P$, $Q$, $R$, $S$, $T$ and two constituents $e, f$.

$$\begin{aligned}\ll \kappa \gg_* = &-\lambda_{QR}\lambda_{PS}(1 + A^{-2} + A^{-2}x_{[\{T\}]} + A^{-4}) - \lambda_{PR}\lambda_{QS}A^{-4} \\ &- \lambda_{PQ}\lambda_{RS}y_{e,Q}y_{e,R}y_{e,S}y_{f,P}^{-1}(A^{-2}y_{f,T} + A^{-4}x_{[\{T\}]} + A^{-6}y_{e,T}^{-1}).\end{aligned}$$

For brevity, we have omitted the set notation in the subscripts of the $\lambda$ variables. Proposition 5·5($a$) yields the height bounds $h_\kappa(T, P)$, $h_\kappa(T, Q)$, $h_\kappa(T, R)$, $h_\kappa(T, S) \geq 1$. Meanwhile, Proposition 5·5($b$) and 5·5($c$) yield the height bounds $h_\kappa(P, Q)$, $h_\kappa(P, R)$, $h_\kappa(P, S) \geq 1/2$ and $h_\kappa(T, P)$, $h_\kappa(T, Q)$, $h_\kappa(T, R)$, $h_\kappa(T, S) \geq 1$. In this case, the bounds give enough information for a complete description of the height spectrum of $\kappa$.

We remark on the case of generalised knotoids that are not necessarily oriented or fully labelled. Consider an oriented, pole-labelled, constituent-labelled diagram $\mathcal{D}$. As we remove orientation or labelling data from $\mathcal{D}$, we alter the polynomial $\ll \mathcal{D} \gg_*$ accordingly, as follows.

(i) If constituent labels are removed, we re-index the variables $\{y_{e,P}\}$ by replacing each subscript component $e \in E(\mathcal{D})$ with the ordered pair $(e_0, e_1)$ corresponding to its endpoint poles.

(ii) If pole labels are removed, we remove $G_s$ and $E_s$ from the definition given in Equation (5·1·1), and we re-index the variables $\{x_U\}$ as follows. For each subscript $U$, write $U = [S]$ for a set $S \subset P(\mathcal{D})$, then replace $U$ with the integer $\min(|S|, |P(\mathcal{D})| - |S|)$.

(iii) If orientation is disregarded, we remove $E_s$ from the definition given in Equation (5·1·1) and redefine the normalised bracket polynomial by $\ll \mathcal{D} \gg_* = (-A^3)^{-\sum_{e \in C(\mathcal{D})} \mathrm{wr}(e)} \ll \mathcal{D} \gg$, where $\mathrm{wr}(e)$ denotes the sum of the signs of the self-crossings of an arbitrarily-oriented constituent $e$. (In particular, $\mathrm{wr}(e)$ is independent of orientation.)

Given an oriented, unlabelled diagram $\mathcal{D}$ (for example), we obtain the bracket polynomial $\ll \mathcal{D} \gg_*$ by labelling $\mathcal{D}$ arbitrarily and then applying (i) and (ii). (Note that (ii) and (iii) subsume (i).) It is straightforward to verify that, in general, this method yields a bracket polynomial invariant for each class of generalised knotoids with specified labelling and orientation data.

5·2. *Recovering existing bracket polynomials*

We survey some existing bracket polynomial constructions for classes of objects subsumed by generalised knotoids and show that they are recovered by the generalised bracket polynomial.

In the case of a classical link diagram $L$, there are no poles, so all loop constituents of any state $s$ are nullhomotopic and $n_0(s)$ is simply the number of constituents of $s$. Moreover, $G_s$,

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

$L_s$, and $E_s$ are constant, so $\ll L \gg_*$ recovers the traditional bracket polynomial $\langle L \rangle$ of the link diagram.

In [12], a bracket polynomial for oriented, ordered (i.e. pole-labelled) multi-linkoids $L$ is defined by

$$\langle L \rangle_\bullet(A, \{\lambda_{ij}\}) = (-A^3)^{-\mathrm{wr}(L)} \sum_s A^{\sigma(s)}(-A^2 - A^{-2})^{\|s\|} \prod_\Lambda \lambda_{ij}.$$

Here, the poles of the multi-linkoid are labelled $1, 2, \ldots, 2n$, the symbol $\Lambda$ indicates that the product is taken over all $i < j$ such that poles $i$ and $j$ are connected by a segment constituent in $s$, and $\|s\|$ denotes the number of loop constituents of $s$. It is clear that $\langle L \rangle_\bullet$ is recovered from $\ll L \gg_*$ by identifying $G_s$ with $\prod_\Lambda \lambda_{ij}$ and making the substitutions $x_U = -A^2 - A^{-2}$ for all $U \in (2^{P(\mathcal{D})}/\sim) \setminus [\emptyset]$ and $y_{e,P} = 1$ for all $e \in E(L), P \in P(L)$.

In [31], Turaev defines a two-variable bracket polynomial for oriented spherical knotoids. Let $K$ be a spherical knotoid oriented from $L$ to $H$, and let $\alpha = \alpha_{L,H}$ be a shortcut oriented from $L$ to $H$. (We also use $K$ to denote the segment constituent of the knotoid.) Turaev's polynomial is given by

$$\ll K \gg_\circ (A, u) = (-A^3)^{-\mathrm{wr}(K)} \sum_s A^{\sigma(s)}(-A^2 - A^{-2})^{\|s\|} u^{k_s \cdot \alpha - K \cdot \alpha},$$

where $k_s$ denotes the unique segment constituent of the state $s$, oriented from $L$ to $H$. Note that for any state $s$, the segment constituent $k_s$ aligns with $K$, and all loop constituents of $s$ are nullhomotopic. In terms of the generalised bracket polynomial, the polynomials $G_s = \lambda_{LH}$ and $L_s = 1$ are constant, and $n_0(s) = \|s\|$. It follows that $\ll K \gg_\circ$ is recovered from $\ll K \gg_*$ by the substitutions $\lambda_{LH} = 1$ and $y_{K,H} = u$. It also follows that the height bound in Proposition 5·5(c) is equivalent to the complexity bound given in [31, equation 8·3·1].

In [31], Turaev extends $\ll K \gg_\circ$ to a bracket polynomial for oriented planar knotoids given by

$$[K]_\circ(A, B, u) = (-A^3)^{-\mathrm{wr}(K)} \sum_s A^{\sigma(s)}(-A^2 - A^{-2})^{p(s)} B^{q(s)} u^{k_s \cdot \alpha - K \cdot \alpha},$$

where $p(s)$ denotes the number of loop constituents of $s$ that do not enclose the segment constituent $k_s$ and $q(s)$ denotes the number of loop constituents of $s$ that enclose $k_s$. Represent $K$ as a generalised knotoid with poles $L, H, \infty$ and constituent $K$ oriented from $L$ to $H$, equipped with shortcuts $\{\alpha_{P,Q} \mid P, Q \in \{L, H, \infty\}\}$, where $\alpha_{L,H} = \alpha$. Then $p(s) = n_0(s)$ is simply the number of nullhomotopic loop constituents of $s$, and $q(s) = n_{[\{\infty\}]}(s)$ is the number of loop constituents of $S$ that determine the bipartition $\{\{\infty\}, \{L, H\}\}$. Thus $[K]_\circ$ is recoverable from $\ll K \gg_*$ by the substitutions $\lambda_{LH} = 1$, $x_{[\{\infty\}]} = B$, $y_{K,\infty} = 1$, and $y_{K,H} = u$.

In [26], Kutluay refines $[K]_\circ$ by replacing the term $u^{k_s \cdot \alpha - K \cdot \alpha}$ with the term $\ell^{w_{\gamma_s}(L) - w_\gamma(L)} h^{w_{\gamma_s}(H) - w_\gamma(H)}$, obtaining a polynomial $[K]_\bullet(A, B, \ell, h)$. Here, $w_\beta : \mathbb{R}^2 \to \mathbb{Z}/2$ is the *winding potential function* of an oriented generic closed curve $\beta$ that maps each point $x \in \mathbb{R}^2$ to the winding number of $\beta$ around $x$. (For $x \in \beta$, the winding number is taken to be the average of the winding numbers of the regions adjacent to $x$.) The closed curve $\gamma$ (resp. $\gamma_s$) is defined by $K \cup \alpha^r$, (resp. $k_s \cup \alpha^r$), where $\alpha^r$ is $\alpha$ with reversed orientation. Kutluay observes that

$$k_s \cdot \alpha - K \cdot \alpha = [w_{\gamma_s}(L) - w_\gamma(L)] - [w_{\gamma_s}(H) - w_\gamma(H)],$$

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press



Fig. 16. A knotoidal graph with three poles and four spatial vertices.

so Turaev's polynomial $[K]_\circ$ is recoverable from Kutluay's polynomial $[K]_\bullet$ via the substitutions $\ell = u$ and $h = u^{-1}$.

It is well known that for a generic closed curve $\beta$ on the plane and a point $x \notin \beta$, the winding number $w_\beta(x)$ is given by $\beta \cdot \alpha_{x,\infty}$, where $\alpha_{x,\infty}$ is any path from $x$ to the unbounded region of the plane that intersects $\beta$ transversely and away from self-intersections of $\beta$. Using this fact, it is straightforward to verify that

$$\begin{aligned} w_{\gamma_s}(L) - w_\gamma(L) &= \gamma_s \cdot \alpha_{L,\infty} - \gamma \cdot \alpha_{L,\infty} \\ &= \mu_{k_s,K}(L, \infty), \end{aligned}$$

and similarly

$$\begin{aligned} w_{\gamma_s}(H) - w_\gamma(H) &= \mu_{k_s,K}(H, \infty) \\ &= -\mu_{k_s,K}(L, H) + \mu_{k_s,K}(L, \infty). \end{aligned}$$

It follows that $[K]_\bullet$ is recoverable from $\ll K \gg_*$ by the substituions $\lambda_{LH} = 1$, $x_{[\{\infty\}]} = B$, $y_{K,\infty} = \ell h$, and $y_{K,H} = h^{-1}$.

## 6. *Knotoidal graphs*

### 6·1. *Knotoidal graphs*

We now define an extension of generalised knotoids. Let $\Sigma$ denote a closed orientable surface, and let $G$ be a finite graph. As with generalised knotoids, we do not require $G$ to be connected or simple, and $G$ may have valency-zero vertices. Let $\tilde{G}$ denote the disjoint union of $G$ with a finite collection of circles. The edges and circles of $\tilde{G}$ are still called constituents.

A *knotoidal graph diagram* on $\Sigma$ is a pair $(\mathcal{D}, V(\mathcal{D}))$, where $\mathcal{D}$ is a generic immersion of $\tilde{G}$ in $\Sigma$ whose only singularities are transverse double points, called crossings, with over/undercrossing data, and where $V(\mathcal{D})$ is a chosen subset of the images of the vertices of $G$, each of which has valency at least one. We call the elements of $V(\mathcal{D})$ the *spatial vertices* of $\mathcal{D}$. The images of the remaining vertices of $\tilde{G}$ are called *poles* of $\mathcal{D}$ and the set of poles is denoted by $P(\mathcal{D})$. Figure 16 shows an example.

The rest of the terminology for generalised knotoids carries over. we call the graph $G$ (resp. $\tilde{G}$) the *underlying graph* (resp. *underlying looped graph*) of $\mathcal{D}$. The *valency* of a pole or spatial vertex is the valency of the corresponding vertex in the underlying graph G. Let $E(\mathcal{D})$ denote the set of images of the edges of $\tilde{G}$, called *segment constituents* of $\mathcal{D}$, and let $L(\mathcal{D})$ denote the images of the circles of $\tilde{G}$, called *loop constituents of* $\mathcal{D}$. Let $C(\mathcal{D}) := E(\mathcal{D}) \cup L(\mathcal{D})$ denote the set of *constituents of* $\mathcal{D}$.

We consider knotoidal graph diagrams in $\Sigma$ up to ambient isotopy and *generalised Reidemeister moves*: the three standard Reidemeister moves away from poles and spatial vertices, along with the *vertex slide move* and the *vertex twist move* near spatial vertices.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

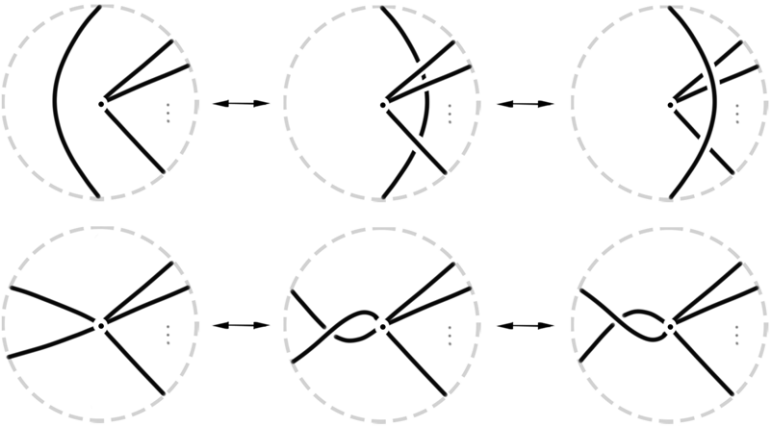


Fig. 17. The vertex slide move (top) and the vertex twist move (bottom).

See Figure 17. The forbidden pole moves of Figure 1 remain in effect. A *knotoidal graph* $KG$ is an equivalence class of knotoidal graph diagrams. We may speak of knotoidal graphs possibly with labels assigned to poles or spatial vertices, and labels or orientations assigned to constituents.

With the vertex twist move, the theory of knotoidal graphs is akin to the theory of spatial graphs with pliable vertices (as opposed to rigid vertices, where the twist move is not allowed). One could require knotoidal graphs to have rigid spatial vertices, which would yield a different theory that we do not consider in this paper.

As we remarked for generalised knotoids, an equivalent theory arises by considering knotoidal graphs in compact surfaces with boundary while disallowing valency-zero poles.

*Example 6·1* Generalised knotoids are knotoidal graphs with no spatial vertices.

*Example 6·2* Spatial graphs are knotoidal graphs with no poles.

*Example 6·3.* A *graphoid* (resp. *multi-graphoid*) as defined in [**16**] is a knotoidal graph with exactly two poles (resp. $2n$ poles for some $n \geq 1$), each of valency one. A *bonded knotoid*, used to model bonded proteins [**15**], is particular type of graphoid obtained by adding edges to a knotoid with endpoints at distinct points of the knotoid away from its endpoints. So all resulting spatial vertices have valency three.

### 6·2. *Rail diagrams*

A topological perspective on knotoids is introduced in [**25**] by means of *rail diagrams*. For planar knotoids, consider the thickened disk $D^2 \times I$ together with two distinct lines $\ell_1 = \{x\} \times I$, $\ell_2 = \{y\} \times I$, called *rails*. A *planar rail diagram* is a proper embedding of an arc into $(D^2 \times I) \setminus \mathring{N}(\ell_1 \cup \ell_2)$ such that one endpoint is embedded in $\partial N(\ell_1)$ and one is embedded in $\partial N(\ell_2)$, as shown in Figure 18. We consider planar rail diagrams up to isotopies of the embedded arc in $(D^2 \times I) \setminus \mathring{N}(\ell_1 \cup \ell_2)$. Intuitively, the removal of the rails has the effect of topologically enforcing the forbidden move for knotoid diagrams. It is shown in [**25**] that the theory of planar rail diagrams is equivalent to the theory of planar knotoids,

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

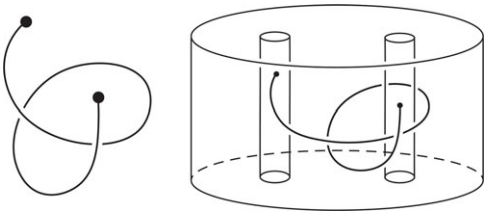


Fig. 18. A planar rail diagram for a knotoid.

and if the disk $D^2$ is replaced by an arbitrary surface $\Sigma$, the arguments in [25] generalise readily to show that the theory of $\Sigma$-rail diagrams is equivalent to the theory of knotoids on $\Sigma$.

We extend $\Sigma$-rail diagrams to knotoidal graphs. For a given knotoidal graph $KG$ in $\Sigma$ with diagram $\mathcal{D}$, take the thickened surface $\Sigma \times I$ and $n$ distinct rails $\ell_i = \{x_i\} \times I$, $i = 1, \ldots, n$, where $n = |P(\mathcal{D})|$, each rail corresponding to a pole $p_i$. Let $\tilde{G}'$ be the graph obtained from the underlying looped graph $\tilde{G}$ of $\mathcal{D}$ by removing a disk neighbourhood of each pole. Then for each pole $p_i$ there are $v(p_i)$ endpoints in $\tilde{G}'$ created by removing the disk neighbourhood. Now properly embed $\tilde{G}'$ in $(\Sigma \times I) \setminus \mathring{N}(\bigcup_{i=1}^{n} \ell_i)$ so that the projection of the embedding to $\Sigma \times \{0\}$ recovers $\mathcal{D}$ away from the poles, and so the $v(p_i)$ endpoints of $\tilde{G}'$ are embedded on the boundary $\partial N(\ell_i)$ of the corresponding rail neighbourhood.

Again, we consider these diagrams up to isotopies of the embedded looped graph in $(\Sigma \times I) \setminus \mathring{N}(\bigcup_{i=1}^{n} \ell_i)$, with one major difference. We may identify the cylindrical boundary portion of each $\partial N(\ell_i)$ with $S^1 \times I$. If an endpoint $v$ of $\tilde{G}'$ is embedded as the point $(x, t)$ in $\partial N(\ell_i)$, then we require the isotopies to keep $v$ in the vertical line segment $\{x\} \times I$. That is, we only allow endpoints of the graph $\tilde{G}'$ embedded in rail neighbourhoods to slide vertically up and down. This prevents arbitrary twisting near poles, which we disallow in the diagrammatic theory of knotoidal graphs.

The theory of $\Sigma$-rail diagrams of knotoidal graphs is equivalent to the theory of knotoidal graphs in $\Sigma$. Consequently, we may derive invariants of knotoidal graphs using topological methods, as we do in Section 7.

Although we do not address it here, one can also allow virtual crossings in a spherical or planar knotoidal graph. This would generate classical knotoidal graphs in surfaces as we have described, however, in this theory, different projections can generate surfaces of different genus and equivalence between diagrams must then allow adding and removing handles for the corresponding surfaces. See [19] for this theory as applied to graphoids.

## 7. *Hyperbolicity for knotoidal graphs*

We review the notion of hyperbolicity for knotoids, studied in [2]. Two maps $\phi^D_{S^2}$ and $\phi^G_{S^2}$ are defined in [2], each sending a spherical knotoid to a knot in the thickened torus $T \times (0, 1)$. Two maps $\phi^D_{\mathbb{R}^2}$ and $\phi^G_{\mathbb{R}^2}$ are also defined in [2], each sending a planar knotoid to a knot in the genus three handlebody $H_3$ or to a knot in the genus two handlebody $H_2$, respectively.

The map $\phi^D_{S^2}$ is termed the *spherical reflected doubling map* and is defined as follows.

*Definition 7·1.* Consider a spherical rail diagram for a knotoid $k$, namely a thickened cylinder $M = S^2 \times (0, 1) \setminus \mathring{N}(\ell_1 \cup \ell_2)$ with an embedded arc, where each $\ell_i = \{x_i\} \times (0, 1)$

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

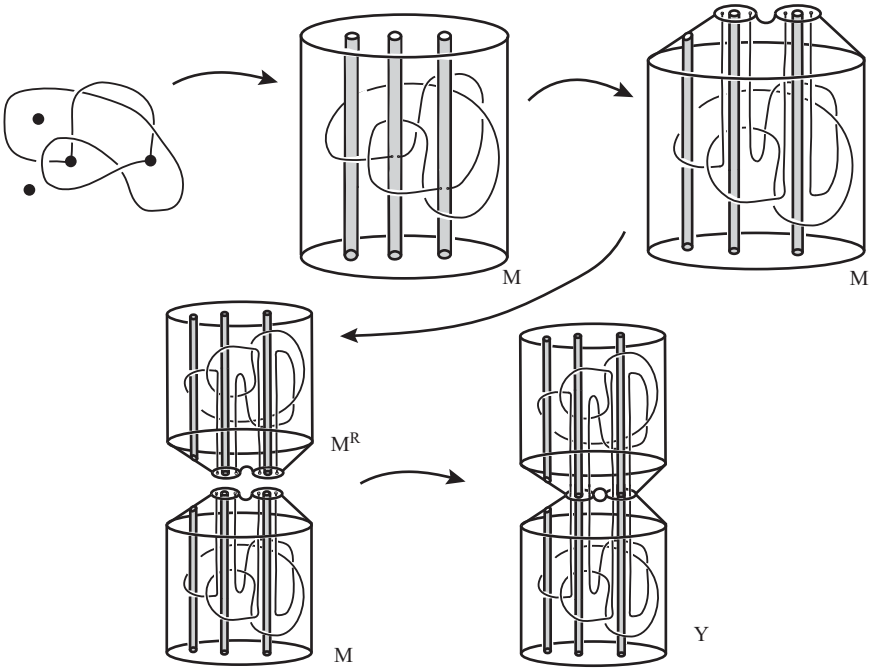



Fig. 19. The construction for a generalised knotoid or knotoidal graph.

is a rail. Let $C_i = \partial N(\ell_i)$ denote the boundaries of the regular neighbourhoods of the rails, which are homeomorphic to annuli and are each punctured once by the knotoid (we may take the neighbourhoods to be sufficiently small so that this is true). Label the punctures by $z_i$. Take a reflected copy $M^R$ of $M$ which contains reflected copies $C_i^R$ of the rail boundaries and copies $z_i'$ of the punctures. Then we glue together $M$ and $M^R$ by gluing $C_1$ to $C_1^R$ and $C_2$ to $C_2^R$ via a reflection as appears in Figure 19. This yields a well-defined knot $K$ in the thickened torus $T \times (0, 1)$. We then set $\phi^D_{S^2}$ to $K \in \mathcal{K}(T^2 \times (0, 1))$.

We say that $k$ is hyperbolic if $K$ is hyperbolic in $T \times (0, 1)$, that is, $(T \times (0, 1)) \setminus N(K)$ admits a complete hyperbolic metric. We define the volume of $k$, denoted $\mathrm{Vol}_{S^2}(k)$, to be half the volume of $(T \times (0, 1)) \setminus N(K)$.

7·1. *Hyperbolicity for knotoidal graphs*

We extend the domain of the map $\phi^D_{S^2}$ to knotoidal graphs.

*Definition 7·2.* Consider a spherical rail diagram for a knotoidal graph $KG$ in $S^2$. Denote by $\ell_1, \ldots, \ell_m$ the rails corresponding to nonzero-valency poles $p_1, \ldots, p_m$, and denote by $\ell_{m+1}, \ldots, \ell_n$ the rails corresponding to valency-zero poles. Let $C_i$ be the boundary of $N(\ell_i)$ for $i = 1, \ldots, m$; we may choose the regular neighbourhoods to be sufficiently small such that each $C_i$ is punctured exactly $v(p_i)$ times, where $v(p_i)$ is the valency of the pole $p_i$. Let $M$ be the manifold $(S^2 \times I) \setminus N(\bigcup_{i=1}^n \ell_i)$, and take a reflected copy $M^R$ which contains reflected copies $C_i^R$ of the rail boundaries. Then glue together $M$ and $M^R$ by gluing $C_i$ to $C_i^R$ via reflection for $i = 1, \ldots, m$, so that the punctures in the $C_i$ and $C_i^R$ line up. Note that we do not glue the boundaries corresponding to valency-zero poles, and if there are no poles of nonzero valency, then we do not take a reflected copy, leaving us with just $M$.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

This process yields a spatial graph $\mathcal{G}$ in a manifold $Y$. If there are no valency-zero poles, then $Y$ is homeomorphic to a thickened closed orientable surface of genus $m-1$. Otherwise, $Y$ is homeomorphic to a handlebody of genus $m+(m-1)+2(n-m-1)=2n-3$ when $n-m>1$: each valency-zero pole after the first one increases the genus by two. If there is exactly one valency-zero pole, then $Y$ is homeomorphic to a handlebody of genus $m+(m-1)=2m-1$. Hence, we obtain an extension of the *spherical reflected doubling map* $\phi^D_{S^2}$ to knotoidal graphs. It associates to every knotoidal graph in $S^2$ a spatial graph in a 3-manifold $Y$, and we write $\phi^D_{S^2}(KG)=(\mathcal{G},Y)$.

When restricted to spherical knotoids (knotoidal graphs with two poles and a single segment constituent between them), the map $\phi^D_{S^2}$ agrees with the map given in Definition 7·1, justifying the notation. Similarly, when restricted to planar knotoids (knotoidal graphs with three poles and a single segment constituent between two of them), the map $\phi^D_{S^2}$ is precisely the planar reflected doubling map $\phi^D_{\mathbb{R}^2}:\mathbb{K}(\mathbb{R}^2)\to\mathcal{K}(H_3)$ defined in [**2**].

Note that we can also extend this definition to knotoidal graphs on a general closed orientable projection surface $\Sigma$, by replacing every occurrence of $S^2$ in Definition 7·2 with $\Sigma$. Again, we obtain a spatial graph $\mathcal{G}$ in a manifold $Y$, and we denote the map associating $(\mathcal{G},Y)$ to $KG$ by $\phi^D_\Sigma$. If $\Sigma$ has genus $g$ and there are no valency-zero poles, then $Y$ is homeomorphic to a thickened closed orientable surface of genus $2g+m-1$. Otherwise, $Y$ is homeomorphic to a handlebody of genus $(2g+m-1)+2(n-m)=2g+2n-m-1$. We can now define hyperbolicity for knotoidal graphs on any closed orientable projection surface $\Sigma$.

*Definition 7·3.* Let $KG$ be a knotoidal graph in a closed orientable surface $\Sigma$. Then we say $KG$ is *hyperbolic* if its image $\phi^D_\Sigma(KG)=(\mathcal{G},Y)$ under the $\Sigma$-reflected doubling map is tg-hyperbolic. That is, the manifold $Y\setminus N(\mathcal{G})$ admits a complete hyperbolic metric such that its higher genus boundary components are totally geodesic in the metric.

We require the stronger condition of tg-hyperbolicity on $Y\setminus N(\mathcal{G})$ to guarantee that it has a well-defined, finite hyperbolic volume.

*Definition 7·4.* Let $KG$ be a hyperbolic knotoidal graph in $S^2$, and let $(\mathcal{G},Y)$ be its image under $\phi^D_{S^2}$ with a tg-hyperbolic metric. Then the *hyperbolic volume* $\mathrm{Vol}(KG)$ of $KG$ is defined as

$$\mathrm{Vol}(KG):=\frac{1}{2}\mathrm{Vol}(Y\setminus N(\mathcal{G})).$$

We remark that hyperbolic volume does not distinguish between a knotoidal graph $KG$ and the knotoidal graph obtained by applying a pole twist move at a pole of $KG$ since the reflected doubling operation "undoes" the twist. (However, the polynomial invariants for generalised knotoids defined in Sections 4 and 5 may detect twists.)

### 7·2. *Staked links*

We defined staked links in Example 2·8. Given any link $L$ with projection on a closed orientable surface $\Sigma$, we may *stake* it by adding any number of valency-zero poles to its diagram, and we refer to these poles as *stakes*. By means of their rail diagrams, staked links are equivalent to links in handlebodies.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

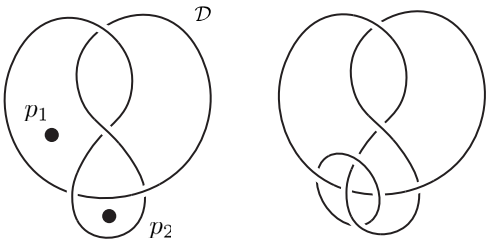



Fig. 20. Adding two stakes to a link diagram on $S^2$ is equivalent to adding an unknot component.

We would like to consider when staked links are hyperbolic. As a simple example, let $L$ be a link with diagram on $\Sigma = S^2$. Suppose $L'$ is a staked link obtained by adding a single stake to the diagram of $L$. If $(\mathcal{G}, Y)$ is the image of $L'$ under $\phi^D_{S^2}$, then $Y$ is just a 3-ball and $\mathcal{G}$ is a link so that $\mathcal{G}$ in $Y$ is equivalent to $L$ in $S^2 \times I$ under the correspondence between links in $S^3$ and links in $S^2 \times I$. In particular, a 1-staked link $L'$ obtained from $L \subset S^2 \times I$ is hyperbolic if and only if $L$ is hyperbolic. Contrast this with 1-polar links: they are never hyperbolic since there is always an essential annulus in $Y \setminus N(\mathcal{G})$.

As another example, let $L$ be a link with diagram $\mathcal{D}$ on $\Sigma = S^2$, and suppose $L'$ is a staked link obtained by adding two stakes to the diagram of $L$. If $(\mathcal{G}, Y)$ is the image of $L'$ under $\phi^D_{S^2}$, then the complement $Y \setminus N(\mathcal{G})$ is homeomorphic to the complement of a link $L''$ in $S^3$ obtained by adding an unknotted component to $L$ in the following way: let $p_1$ and $p_2$ denote points on $S^2$ where the stakes are added, and add two arcs to the diagram $\mathcal{D}$ with endpoints on $p_1$ and $p_2$ such that one arc crosses over every strand of $\mathcal{D}$ it meets and the other crosses under every strand it meets. This yields a new link diagram $\mathcal{D}''$. Then $L'$ is hyperbolic if and only if $L''$ is hyperbolic. See Figure 20 for an example.

Observe that when $p_1$ and $p_2$ are in distinct, non-adjacent regions, then $L''$ is obtained by augmenting $L$ as in [**1**]. When $L$ is a prime, non-split, alternating link which is not a 2-braid, Theorem 2·1 from [**1**] shows that such a 2-staked link $L''$ is hyperbolic.

More generally, if $L$ has a *cellular alternating* diagram on $\Sigma$, then we can determine precisely when a staked link $L'$ obtained by staking the diagram of $L$ is hyperbolic.

*Definition 7·5* Let $\Sigma$ be a closed, orientable surface, and let $L \subset \Sigma \times I$ be a staked link with diagram $\mathcal{D}$. We say $L$ is *weakly prime* if every disk $D \subset \Sigma$ containing at least one crossing of $\mathcal{D}$ such that the circle $\partial D$ intersects $\mathcal{D}$ transversely in two points also contains at least one stake in its interior in a complementary region that does not touch $\partial D$.

*Definition 7·6.* Let $\Sigma$ be a closed, orientable surface and let $\mathcal{D}$ be a projection diagram for a link $L$ in $\Sigma$. We say $\mathcal{D}$ is *cellular* if every complementary region in $\Sigma \setminus \mathcal{D}$ is a disk.

Then Theorem 1·6 from [**4**], stated below in the language of stakes, allows us to determine whether or not a staking yields a hyperbolic staked link.

PROPOSITION 7·7. *Let $\Sigma$ be a closed projection surface, and let $L \subset \Sigma \times I$ be a link with a connected, reduced, cellular alternating projection diagram $\mathcal{D} \subset \Sigma \times \{1/2\}$ with at least one crossing. Let L' be a staked link obtained by staking $\mathcal{D}$, and let $\mathcal{D}'$ be the reduced diagram for L' in $\Sigma$. Let $M = (\Sigma \times I) \setminus N(L)$. Then M is tg-hyperbolic if and only if the following four conditions are satisfied*:

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

*(i)* $\mathcal{D}'$ *is weakly prime on* $\Sigma \times \{1/2\}$;

*(ii) every complementary region of* $\Sigma \setminus \mathcal{D}'$ *contains at most one stake*;

*(iii) adjacent regions of* $\Sigma \setminus \mathcal{D}'$ *do not both contain stakes*;

*(iv) there is no simple closed curve* $\alpha$ *in* $\Sigma$ *that intersects* $\mathcal{D}'$ *exactly in a nonempty collection of crossings, such that for each such crossing,* $\alpha$ *bisects the crossing and the two opposite complementary regions meeting at that crossing that do not intersect* $\alpha$ *near that crossing are staked.*

These conditions are all relatively easy to check directly on the diagram $\mathcal{D}'$.

*Proof.* As in the remarks following the definitions of generalised knotoids and knotoidal graphs, adding a stake to a link diagram $\mathcal{D}$ on $\Sigma$ is equivalent to removing an open disk neighbourhood of $\Sigma$. Hence, we can view $\mathcal{D}'$ as a link diagram on a compact surface $\Sigma'$ with boundary. Then the proposition is a restatement of [**4**, theorem 1·6 and corollary 1·7].

This proposition gives us a large class of hyperbolic staked links. The following theorem uses it to show that an even larger class is hyperbolic. We say a projection of a link on a surface is checkerboard-colourable if we can shade a subset of the complementary regions so that no two adjacent regions are both shaded or both unshaded. This is equivalent to the fact that every simple closed curve that intersects the projection transversely away from the crossings does so an even number of times.

THEOREM 7·8. *Let* $\Sigma$ *be a closed, orientable surface and let* $L \subset \Sigma \times I$ *be a link with a projection diagram* $\mathcal{D}$ *that is checkerboard-colourable. Then there exists another projection* $\mathcal{D}'$ *of* $L$ *such that* $\mathcal{D}'$ *can be staked so that the resulting staked link is hyperbolic.*

*Proof.* Given our projection $\mathcal{D}$, we first apply Type II Reidemeister moves to make it connected. We then apply further Type II Reidemeister moves to make it cellular and weakly prime. None of these moves change the fact it is checkerboard-colourable. We can also ensure that there are multiple crossings.

For any such projection on a closed surface, we can always find a subset of the crossings to change in order to make it alternating. We see this by checkerboard shading the complementary regions and then changing crossings so each shaded region is oriented clockwise with respect to the overcrossings and each unshaded region is oriented counterclockwise with respect to the overcrossings.

Although we do not need this, we can always change at most half the crossings, since initially, if we need to change more than half, we can instead just change the complementary crossings.

As we will see, it is enough to stake a pair of opposite regions at a crossing to allow us to change it. However, at this stage, we cannot stake accordingly at each crossing we wish to change as this can either create adjacent staked regions from different crossings, or it can create a simple closed curve $\alpha$ that fails condition (iv) of Proposition 7·7. So instead, at each of the crossings we desire to change, we apply the move to the projection that appears in Figure 21. This move is obtained by applying a swirl move that appears twice in [**3**, figure 8]. This yields a new projection $\mathcal{D}'$.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

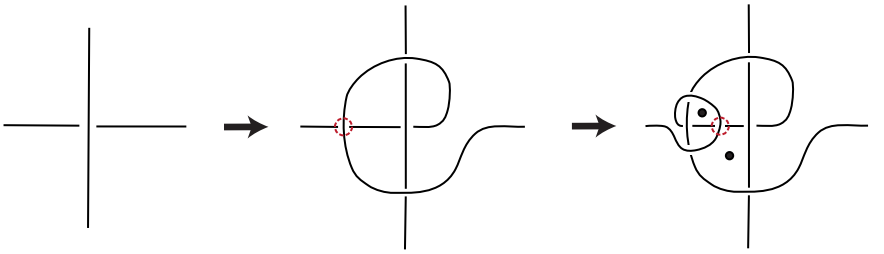


Fig. 21. For a crossing we want to change, we apply a swirl move twice and then stake with two stakes.

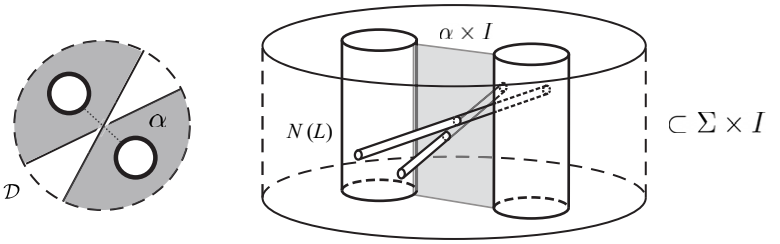



Fig. 22. Staking two opposite regions near a crossing; crossing the arc $\alpha$ by $I$ yields the twice-punctured disk $F = \alpha \times I$ in $\Sigma \times I$, shown on the right.

In order to make the corresponding projection alternating, the crossing highlighted in red becomes the one we need to change. To do so, we stake each of the regions in the diagram as shown that are opposite one another and adjacent to the new crossing.

Note that in the resulting diagram, no two staked regions are adjacent. Furthermore, there can be no simple closed curve $\alpha$ that bisects crossings in the projection and is staked in a pair opposite regions it does not pass through at each crossing. This is because when $\alpha$ passes through such an appropriately staked crossing, the other crossings that are in the region $\alpha$ just entered do not have the requisite stakes to either side for $\alpha$ to exit.

Now that we have staked two opposite regions at a crossing we want to change, we see a twice-punctured disk $F$ in $\Sigma \times I$ intersecting the link at the crossing as in Figure 22.

We can cut along $F$, yielding copies $F_1$ and $F_2$, rotate copy $F_2$ by $2\pi$, and reglue the two copies back together, changing the crossing to the opposite crossing. This is a homeomorphism of $M$ so the resulting link in the handlebody $\Sigma' \times I$ is hyperbolic if and only if the original link is hyperbolic. Upon applying this move to the crossings we need to change, we obtain a staked projection that is cellular alternating and weakly prime. We can now apply Proposition 7·7.

COROLLARY 7·9. *Every link $L$ in $S^2 \times I$ has a diagram that can be staked so that the resulting staked link is hyperbolic.*

*Proof.* Every link diagram in $S^2$ is checkerboard-colourable. So, this follows immediately from the theorem.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

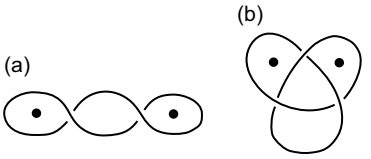



Fig. 23. Staked diagrams of the unknot and trefoil realising the minimum volume over all staked diagrams.

*Definition 7·10.* We define the *staked volume* of a link $L$ in $\Sigma \times I$, where $\Sigma$ is a closed surface, to be the minimum volume over all diagrams $\mathcal{D}$ of $L$ in $\Sigma \times \{1/2\}$ of staked links obtained by adding stakes to $\mathcal{D}$ so that the result is hyperbolic.

Note that this is well-defined because hyperbolic volumes are known to be well-ordered. By Corollary 7·9, we may associate a finite staked volume to every link in $S^3$, including links that are not hyperbolic in the classical sense. In this regard, staked volume is similar to Turaev volume as defined in [**5**].

*Example 7·11* The staked volume of the trivial knot in $S^3$ is 3.6638 . . . , and it is achieved by the staking configuration in Figure 23(a). To see this, observe that for any staked link on $S^2$ with exactly two stakes, its volume as a knotoidal graph is computed as the volume of a link complement in the solid torus. In particular, this is a 2-cusped manifold. In [**6**], Agol proved that the minimum volume of a 2-cusped manifold is 3.6638 . . . . Hence, this must be the minimum volume of any staked diagram of the trivial knot with two stakes. Any diagram of the trivial knot with one stake is isotopic to the trivial diagram with one stake, which yields a trivial knot in the 3-sphere, which is not hyperbolic. Any diagram of the trivial knot with three or more stakes yields a knot in a handlebody of genus at least two. However, for this to be tg-hyperbolic, the boundary must be totally geodesic and by Theorem 5·2 of [**29**], we know that any manifold with totally geodesic boundary must have volume at least 3.6638 . . . and if it has volume exactly 3.6638 . . . , it cannot be the complement of a knot or link in a handlebody.

*Example 7·12.* The staked volume of the trefoil knot in $S^3$ is also 3.6638 . . . , and it is achieved by the staking configuration in Figure 23(b). Similarly to Example 7·11, it is the minimum volume of any staked diagram of the trefoil with exactly two stakes. As in the previous example, one stake will not make it hyperbolic and three or more stakes will cause it volume to be larger.

*Example 7·13.* For any twist knot other than the trefoil knot, its staked volume is its hyperbolic volume. If we use one stake, then removing the corresponding rail from $S^2 \times I$ yields the link in a ball, and capping that ball with another ball yields the knot in $S^3$, with volume the original hyperbolic volume of the knot. If we utilise two stakes, the result will be the complement of a 2-component link in $S^3$, which by [**6**] has volume at least 3.6638 . . . . In fact there is a 2-pole staking for each twist knot that yields this volume. However, every twist knot comes from Dehn filling the Whitehead link, which has this volume, and Dehn filling always lowers volume, so the volume from one stake is strictly less than the volume from two stakes. And exactly as in the last two examples, the volume from three or more stakes is strictly greater than 3.6638 . . . .

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

With this, we end our discussion of generalised knotoids and knotoidal graphs. There are many interesting questions that remain to be explored.

*Acknowledgements.* The research was supported by Williams College and NSF Grant DMS-1947438 supporting the SMALL Undergraduate Research Project. Thanks to the referee for helpful comments.

## REFERENCES

[1] C. ADAMS. Generalised augmented alternating links and hyperbolic volumes. *Algebr. Geom. Topol.* **17**(6) (2017), 3375–3397.
[2] C. ADAMS, A. BONAT, M. CHANDE, J. CHEN, M. JIANG, Z. ROMRELL, D. SANTIAGO, B. SHAPIRO and D. WOODRUFF. Hyperbolic knotoids. ArXiv: 2209.04556 (2022).
[3] C. ADAMS, J. BROCK, J. BUGBEE, T. COMAR, K. FAIGIN, AMY HUSTON, A. JOSEPH and D. PESIKOFF. Almost alternating links. *Topology Appl.* **46** (1992), 151–165.
[4] C. ADAMS and J. CHEN. Hyperbolicity of alternating links in thickened surfaces with boundary. ArXiv: 2309.04999 (2023).
[5] C. ADAMS, O. EISENBERG, K. KAPOOR, J. GREENBERG, Z. LIANG, K. O'CONNOR, N. PACHECO-TALLAJ and Y. WANG. Turaev hyperbolicity of classical and virtual knots. *Algebr. Geom. Topol.* **21** (2021), 3459–3482.
[6] I. AGOL. The minimal volume orientable hyperbolic 2-cusped 3-manifolds. *Proc. Amer. Math. Soc.* **138**(10) (2010), 3723–3732.
[7] K. BHATAINEH. An invariant of planar knotoids and finite-type invariants (2020). Preprint available at https://www.researchgate.net/publication/344446254_AN_INVARIANT_OF_PLANAR_KNOTOIDS_AND_FINITE-TYPE_INVARIANTS.
[8] K. BHATAINEH. New polynomial invariants of knotoids and the theory of polar knots. *Mediterranean Journal of Mathematics* **19**(40) (2022), 33 pages.
[9] H. U. BODEN and W. RUSHWORTH. Minimal crossing number implies minimal supporting genus. *Bull. London Math. Soc.* **53** (2021), 1174–1184.
[10] J. DORIER, D. GOUNDAROULIS, F. BENEDETTI and A. STASIAK. Knotoid: a tool to study the entanglement of open protein chains using the concept of knotoids. *Bioinformatics* **34**(19) (2018), 3402–3404.
[11] L. FOLWACZNY and L. KAUFFMAN. A linking number definition of the affine index polynomial and applications. *J. Knot Theory Ramifications* **22** (1341004) (2013), 30 pages.
[12] B. GABROVŠEK and N. GÜGÜMCÜ. Invariants of multi-linkoids. *Mediterranean Journal of Mathematics* **20** (165) (2023), 22 pages.
[13] J. GOLDMAN and L. KAUFFMAN. Knots, tangles and electrical networks. *Adv. Math.* **14** (1993), 267–306.
[14] D. GOUNDAROULIS, J. DORIER, and A. STASIAK. Knotoids and protein structure. *Topol. Geom. Biopolym.* **74**, 2020.
[15] D. GOUNDAROULIS, N. GÜGÜMCÜ, S. LAMBROPOULOU, J. DORIER, A. STASIAK and L. KAUFFMAN. Topological models for open-knotted protein chains using the concepts of knotoids and bonded knotoids. *Polymers* **9** (2017), 444–460.
[16] N. GÜGÜMCÜ, B. GABROVSEK and L.H. KAUFFMAN. Invariants of bonded knotoids and applications to protein folding. *Symmetry* **14**(8) (2022), 1724–1737.
[17] N. GÜGÜMCÜ and L. KAUFFMAN. New invariants of knotoids. *European J. Combin.* **65** (2017), 186–229.
[18] N. GÜGÜMCÜ and L. KAUFFMAN. Parity, virtual closure and minimality of knotoids. *J. Knot Theory Ramifications* **30** (2150076) (2022), 28 pages.
[19] N. GÜGÜMCÜ, L. KAUFFMAN and P. PONGTANAPAISAN. Graphoids. ArXiv: 2209.09086 (2022).
[20] N. HABEGGER and X.-S. LIN. The classification of links up to link-homotopy. *J. Amer. Math. Soc.* **3** (1990), 389–419.
[21] A. HENRICH. A sequence of degree-one vassiliev invariants for virtual knots. *J. Knot Theory Ramifications* **19** (2010), 461–487.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press

[22] R. KASHAEV. Invariants of long knots. In *Representation Theory, Mathematical Physics and Integrable Systems*, (Birkhäuser, 2019), pp. 431–451.
[23] L. KAUFFMAN. An affine index polynomial invariant of virtual knots. *Journal of Knot Theory and its Ramifications*, **22** (1340007) (2013), 30 pages.
[24] S. KIM, Y. H. IM and S. LEE. A family of polynomial invariants for knotoids. *J. Knot Theory Ramifications* **27** (1843001) (2018), 15 pages.
[25] D. KODOKOSTAS and S. LAMBROPOULOU. Rail knotoids. *J. Knot Theory Ramifications* **28** (1940019) (2019), 19 pages.
[26] D. KUTLUAY. Winding homology of knotoids. PhD. thesis. ArXiv: 2002.07871 (2020).
[27] L. LINOV. Signed heights of knotoids. *J. Knot Theory Ramifications* **31** (2250037) (2022), 25 pages.
[28] V. O. MANTUROV. Parity and projection from virtual knots to classical knots. *J. Knot Theory Ramifications* **22** (1350044) (2013), 20 pages.
[29] Y. MIYAMOTO. Volumes of hyperbolic manifolds with totally geodesic boundary. *Topology* **33**(4) (1994), 613–629.
[30] V. TURAEV. Cobordism of knots on surfaces. *Journal of Topology* **1** (2008), 285–305.
[31] V. TURAEV. Knotoids. *Osaka J. Math.* **49**(02) (2010), 195–223.

https://doi.org/10.1017/S0305004124000148 Published online by Cambridge University Press