

GENERALIZATIONS OF KNOTOIDS AND SPATIAL GRAPHS

COLIN ADAMS, ALEXANDRA BONAT, MAYA CHANDE, JOYE CHEN, MAXWELL
JIANG, ZACHARY ROMRELL, DANIEL SANTIAGO, BENJAMIN SHAPIRO,
AND DORA WOODRUFF

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 generalized 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 generalize to knotoidal graphs, which are a natural extension of spatial graphs that allow both poles and vertices.

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, 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 [6], [16], [23], [25]. They have also received attention for their potential as protein models, for example in [9] and [13].

We introduce *generalized knotoids* and define analogs of knotoid invariants for generalized knotoids. The theory of generalized 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 [11], long knots as in [20], string links as in [29], and the polar knots of [7]. Generalized 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 generalized knotoids. This definition is a natural extension of spatial graphs, bonded knotoids as in [14] and [15] and graphoids as in [15] and [18]. 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. Organization. In Section 2, we define generalized knotoid theory and describe its relation to various existing extensions of knot theory. The three sections that follow describe analogs of knotoid invariants for generalized knotoids. In Section 3, we generalize the notion of height and extend a theorem of Kauffman and Gügümcü [17] regarding height in minimal-crossing diagrams. In Section 4, we define a family of polynomial invariants for generalized knotoids that extend existing index polynomials for knotoids, and in Section 5, we define a bracket polynomial that generalizes 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, particularly to staked knots.

Acknowledgements. The research was supported by Williams College and NSF Grant DMS-1947438 supporting the SMALL Undergraduate Research Project.

2. GENERALIZED KNOTOIDS

Let Σ 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*.

A *generalized knotoid diagram* \mathcal{D} is a generic immersion of \tilde{G} in Σ whose only singularities are transverse double points, called *crossings*, which are labeled 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 labeling of the vertices of G induces a labeling of the

poles of \mathcal{D} . Similarly, a labeling (resp. an orientation) of the constituents of \tilde{G} induces a labeling (resp. orientation) of the constituents of \mathcal{D} .

We introduce an equivalence relation on generalized knotoid diagrams in Σ generated by ambient isotopy and the three standard Reidemeister 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.

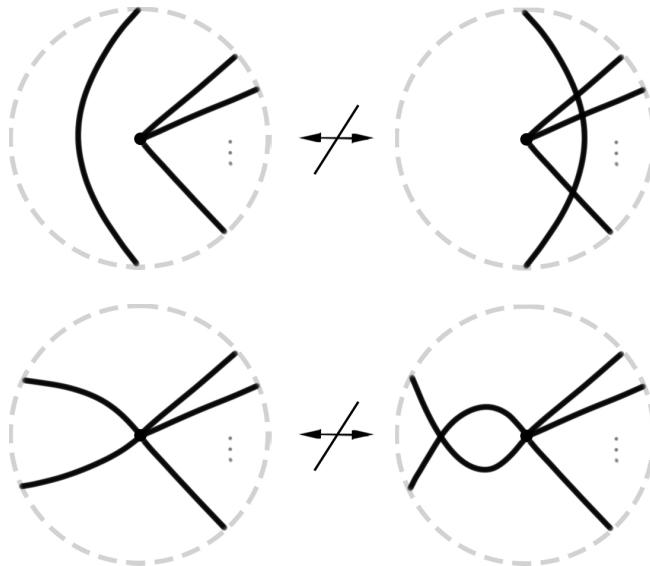


FIGURE 1. The pole slide move (top) and the pole twist move (bottom) are forbidden, irrespective of the crossing data chosen for the diagrams on the right.

A *generalized knotoid* κ is an equivalence class of generalized 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 labeled and/or oriented generalized knotoids as equivalence classes of labeled and/or oriented generalized knotoid diagrams. Note that allowing twisting near poles yields a different theory of generalized knotoids, which we do not consider in this paper. However, some of the invariants we define later are able to distinguish between generalized knotoids which differ only by twists near poles and are otherwise identical.

An equivalent, topological viewpoint is to define a generalized knotoid diagram as a generic immersion of \tilde{G} , where all vertices of G have valency at least one, in a compact orientable surface Σ with or without boundary.

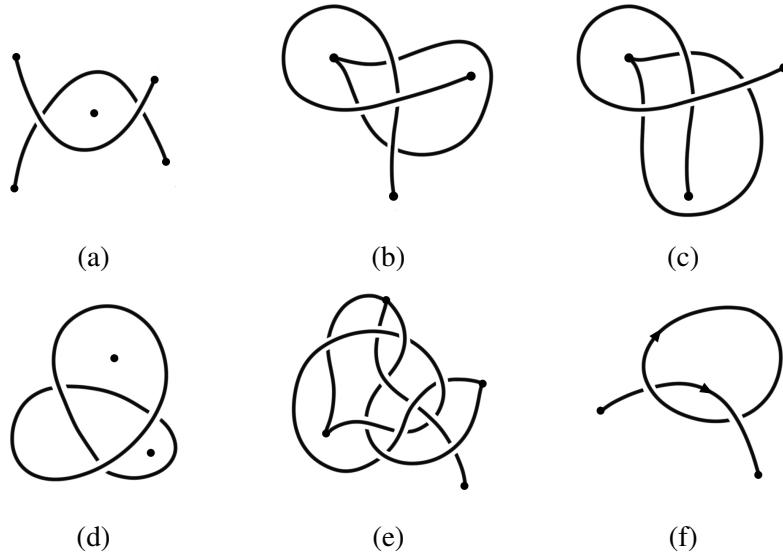


FIGURE 2. Generalized knotoid diagrams.

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 generalized knotoid diagram on a closed surface with isolated poles, we may remove an open disk neighborhood of each isolated pole. That is, the theory of generalized knotoids on closed surfaces with isolated poles is equivalent to the theory of generalized 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 generalized knotoid theory*. In the remainder of this paper, a generalized knotoid is assumed to be classical unless otherwise stated.

Example 2.1. See Figure 2 for some diagrams of generalized knotoids. The reader may verify that Figures 2(b) and 2(c) represent equivalent generalized knotoids. Figure 2(f) represents an oriented generalized knotoid.

Example 2.2. A spherical knotoid is a generalized knotoid with two poles and a single segment constituent between them.

Example 2.3. A planar knotoid is a generalized 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 generalized 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 generalized knotoids in Sections 4 and 5.

Example 2.4. A classical knot or link is a generalized knotoid with no poles.

Example 2.5. The multi-linkoids of [11] are generalized knotoids whose poles all have valency one.

Example 2.6. For a positive integer n , an n -polar knot [7] is a generalized 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 [20] 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 [29] as the embedding of a finite set of closed intervals I_1, I_2, \dots, 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 $-1 < x_1 < x_2 < \dots < x_n < 1$. This is defined up to ambient isotopy of the strings with fixed endpoints.

We represent a string link as a generalized 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 neighborhood 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 $\ell_1, \ell_2, \dots, \ell_n$ around the circle boundary of $D \times \{1/2\}$ followed by the heads of the strings in clockwise order h_n, h_{n-1}, \dots, 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 Σ to be the class of generalized knotoids on Σ with no segment constituents and at least one pole. A diagram of a staked link looks like a link diagram on Σ with isolated poles placed in the regions of the link's complement in Σ . An example appears in Figure 2(d).

We have subsequently learned that in [12], 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, Kauffman and Gügümcü have been investigating properties of staked links under the name starred links.

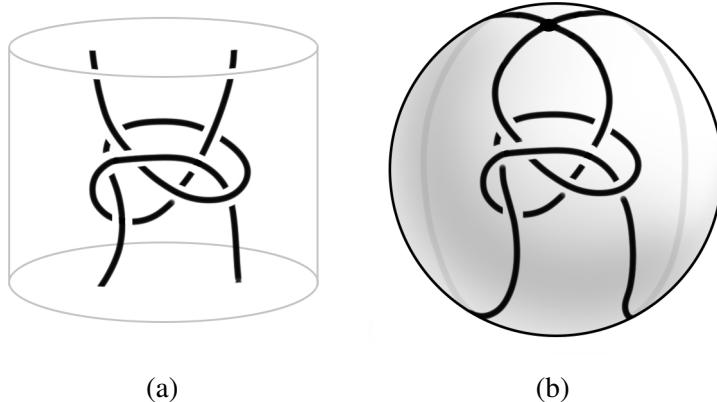


FIGURE 3. The string link in Figure 3(a) corresponds to the generalized knotoid in Figure 3(b). The fact that Reidemeister moves must occur away from the pole captures the fact that the endpoints of the string link's segments are fixed.

3. HEIGHT FOR GENERALIZED KNOTOIDS

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

This definition extends to generalized knotoids in a natural way. Let \mathcal{D} be a pole-labeled generalized knotoid diagram on Σ and let $A, B \in P(\mathcal{D})$ be distinct poles. A *shortcut* α between A and B is an embedded arc in Σ 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 α with \mathcal{D} not including A and B , minimized over all shortcuts α between A and B . An α attaining this minimum is a *minimal shortcut*. Given a generalized knotoid κ 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 *realizes the height between A and B* . Given an ordering on the poles P_1, \dots, P_n of κ , the *height spectrum* of κ 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 [17].

Theorem 3.1 ([17], Theorem 13). *Let k be a spherical knotoid with $h(k) = 0$. Then any minimal-crossing diagram \mathcal{D} of k realizes the height.*

The proof of Theorem 3.1 given in [17] considers the *virtual closure* of a knotoid and applies results from virtual knot theory derived from the method of *parity projection* [27]. In Theorem 3.2, we extend Theorem 3.1 to generalized knotoids. The proof of Theorem 3.2 adapts a topological approach to parity via colorings described in [8].

Theorem 3.2. *Let κ be a classical generalized knotoid and suppose poles $A, B \in P(\kappa)$ satisfy $h_\kappa(A, B) = 0$. Then any minimal-crossing diagram \mathcal{D} of κ realizes 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 α 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 $\mathcal{D}_0 \rightarrow \mathcal{D}_1 \rightarrow \dots \rightarrow \mathcal{D}_n = \mathcal{D}$ 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 neighborhood N_Q that is unchanged as the moves occur. We declare that the shortcut α is fixed across all the diagrams \mathcal{D}_i , and we may assume that α intersects each \mathcal{D}_i transversely and away from crossings.

Claim 3.3. *The number of intersections of α with a constituent of \mathcal{D}_i , excluding potential intersections at poles, is even.*

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 α 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 $\mathcal{D}_j \rightarrow \mathcal{D}_{j+1}$ is applied. For each $t \in \{1, 2, 3\}$, a type t Reidemeister move has t *participating strands* s_1, \dots, s_t , some of which may belong to e . Each s_ℓ meets ∂D at two *endpoints* that are fixed by the Reidemeister move.

If α does not intersect D , then the conclusion is clear. Otherwise, we may assume that α meets ∂D at a finite number of points P_1, P_2, \dots, P_k and Q_1, Q_2, \dots, 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 α_m divides D into two regions, and $\#(\alpha_m \cap s_\ell)_j$ is odd if the endpoints of s_ℓ 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 ℓ , and the conclusion follows. See Figure 4. \square

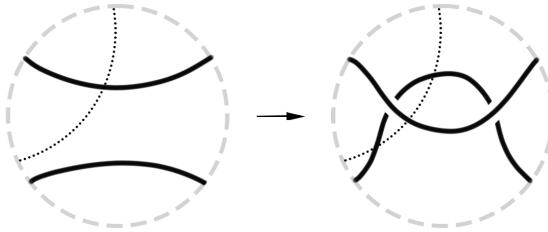


FIGURE 4. A Reidemeister move occurring in the disk region D , whose boundary is indicated by dashed line. The shortcut α is indicated with dotted line. The parity of the number of intersections of α with each participating strand is preserved.

Given a constituent e in a diagram \mathcal{D}_i , an α -coloring of e is an assignment of black or gray to each point of e such that the color of e changes precisely at the points where it intersects α . It follows from Claim 3.3 that an α -coloring exists for each e . We define an α -coloring of the diagram \mathcal{D}_i to be a choice of α -coloring for each of its constituents such that each segment constituent is black in a neighborhood of its endpoint poles. We inductively fix an α -coloring \mathcal{C}_i for each \mathcal{D}_i as follows. We declare \mathcal{C}_0 to be the α -coloring in which all constituents of \mathcal{D}_0 are colored entirely black. Once \mathcal{C}_i is chosen, the Reidemeister move $\mathcal{D}_i \rightarrow \mathcal{D}_{i+1}$ uniquely determines \mathcal{C}_{i+1} if we require \mathcal{C}_i and \mathcal{C}_{i+1} to agree outside the disk region where the Reidemeister move is applied. See Figure 5. Under the α -coloring \mathcal{C}_i , a crossing c in \mathcal{D}_i is assigned a color from each of the constituents meeting at c . We say c is *even* if these colors 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* 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 neighborhood of x intersects at least two distinct regions. In particular, the shortcut α intersects $\partial_0 R$ each time it enters or exits a region R . Conversely, each intersection of α with $\partial_0 R$ corresponds to an entering or exiting of R .

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

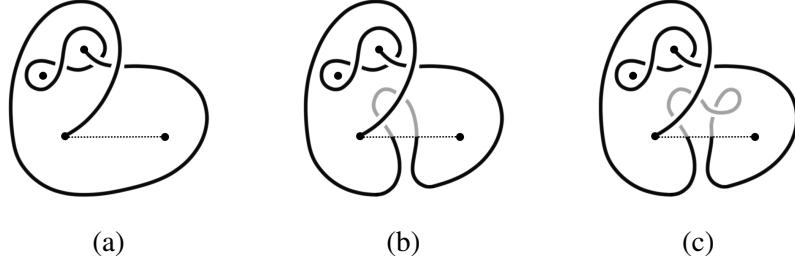


FIGURE 5. Figures 5(a), 5(b), 5(c) show an example sequence of diagrams \mathcal{D}_0 , \mathcal{D}_1 , \mathcal{D}_2 equipped with their α -colorings. The shortcut α is indicated with dotted line.

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 neighborhood 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, \dots, 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. \square

Claim 3.5. *The diagram $\mathcal{D} = \mathcal{D}_n$ has at least one odd crossing.*

Proof. For sufficiently small choices of neighborhoods 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 $\partial_0 R_A$ 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, α intersects $\partial_0 R_A$ an odd number of times, and it follows that α intersects some cycle or circle $C \subset \partial_0 R_A$ an odd number times.

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

- (i) An intersection of $C \setminus \{A, B\}$ with α . (Every such intersection yields a color change.)
- (ii) A point of C at which \mathcal{D} has an odd crossing.

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

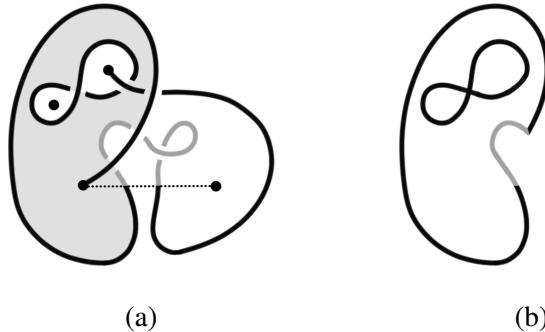


FIGURE 6. The diagram from Figure 5(c). Assuming A is the left endpoint pole of the shortcut α , Figure 6(a) shows the starting region R_A , shaded. Figure 6(b) shows $\partial_0 R_A$. Note that one of the color changes on the outer cycle of $\partial_0 R_A$ occurs at an odd crossing.

Fix a double branched cover $p: S^2 \rightarrow S^2$ branched over α . 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 generalized 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 $p_1: H_1 \rightarrow S^2 \setminus \alpha$ and $p_2: H_2 \rightarrow S^2 \setminus \alpha$ be the restrictions of p to H_1 and H_2 . Note that p_1 and p_2 induce bijections $p_1: p^{-1}(e) \cap H_1 \rightarrow e$ and $p_2: p^{-1}(e) \cap H_2 \rightarrow e$, and recall that e has an α -coloring dictated by \mathcal{C}_i . We color each point of $p^{-1}(e)$ such that the bijections induced by p_1 and p_2 are color-preserving and color-reversing, respectively. Doing so for each $e \in C(\mathcal{D}_i)$ yields the *induced α -coloring*

of $p^{-1}(\mathcal{D}_i)$. See Figure 7. Using the induced α -colorings of the preimages $p^{-1}(e)$, we define *even* (resp. *odd*) crossings of $p^{-1}(\mathcal{D}_i)$ as crossings where the colors 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 .

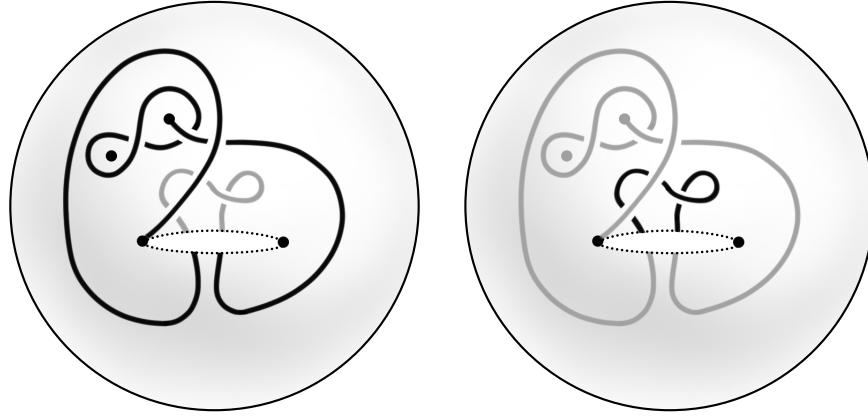


FIGURE 7. One possible visualization of the branched double cover. On the left is H_1 , on the right is H_2 , and they are glued together along their boundaries (the dotted circles obtained by cutting a sphere along α) such that the top and bottom arcs on the left are identified with the bottom and top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

Claim 3.6. *Under the induced α -coloring 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 α -coloring that $p^{-1}(e)$ has both black and gray points. Each time e crosses α and changes color, the lift e_1 crosses from H_1 to H_2 or vice versa. Since p_1 is color-preserving and p_2 is color-reversing, e_1 does not change color. 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 α -coloring, there is a neighborhood N_Q of Q such that \mathcal{C}_i assigns black to $e \cap N_Q$. Since p_1 is color-preserving, there is a neighborhood of $p_1^{-1}(Q)$ in which $p^{-1}(e)$ is black. This proves the second statement. \square

In light of Claim 3.6, we let e' denote the component of $p^{-1}(e)$ colored black and define a generalized 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. \square

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 $\mathcal{D}_i \rightarrow \mathcal{D}_{i+1}$ is applied and let s_1, \dots, s_t be the participating strands, with endpoints on ∂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 $\mathcal{D}_i \rightarrow \mathcal{D}_{i+1}$ lifts to a move $p^{-1}(\mathcal{D}_i) \rightarrow p^{-1}(\mathcal{D}_{i+1})$ given by a type t Reidemeister move in each of D_1 and D_2 . By Claim 3.6, each participating strand s_ℓ lifts to a monochromatic strand $s_{\ell,1} \subset D_1$ and a monochromatic strand $s_{\ell,2} \subset D_2$ (of the opposite color). Since $\mathcal{D}_i \rightarrow \mathcal{D}_{i+1}$ preserves the colors of the endpoints, the move $p^{-1}(\mathcal{D}_i) \rightarrow p^{-1}(\mathcal{D}_{i+1})$ also preserves the colors of the endpoints and thus preserves the colors of $s_{\ell,1}$ and $s_{\ell,2}$. It follows that restricting the move $p^{-1}(\mathcal{D}_i) \rightarrow p^{-1}(\mathcal{D}_{i+1})$ 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. \square

We finish the proof of Theorem 3.2. Since D_0 does not intersect α 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 are ambiently isotopic diagrams of κ . 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 κ . However, Claim 3.5 and Claim 3.7 together imply that

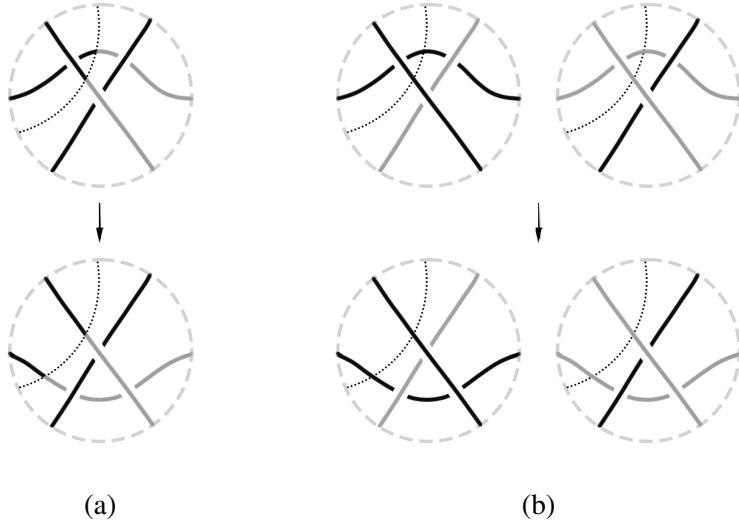


FIGURE 8. A Reidemeister move $\mathcal{D}_i \rightarrow \mathcal{D}_{i+1}$ (Figure 8(a)) lifts to a move $p^{-1}(\mathcal{D}_i) \rightarrow p^{-1}(\mathcal{D}_{i+1})$ in two disk regions (Figure 8(b)). Restricting to only the black strands in Figure 8(b) yields a move $\mathcal{D}'_i \rightarrow \mathcal{D}'_{i+1}$.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

\mathcal{D}'_n has fewer crossings than \mathcal{D}_n . This contradicts the assumption that \mathcal{D}_n is a minimal-crossing diagram of κ . \square

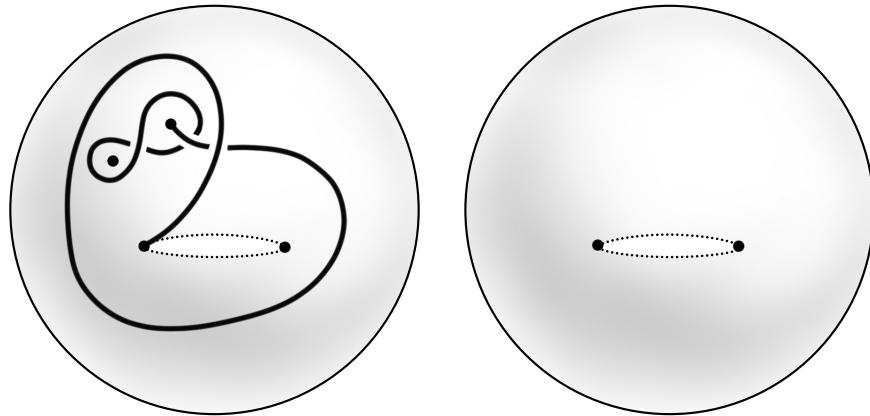


FIGURE 9. The diagram \mathcal{D}'_0 for the diagram \mathcal{D}_0 from Figure 5(a), shown in the same schematic as Figure 7. The diagrams \mathcal{D}_0 and \mathcal{D}'_0 are ambiently isotopic.

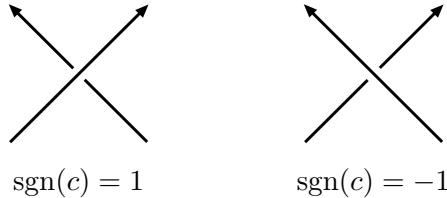


FIGURE 10. Positive and negative crossing, respectively

We conclude this section with some open questions.

Question 3.9. Does every generalized knotoid κ have a diagram \mathcal{D} that simultaneously realizes the height between every pair of poles of κ ?

As a corollary to Theorem 3.2, we obtain an affirmative answer to Question 3.9 in the case that the height spectrum of κ is the zero matrix: we may take \mathcal{D} to be a minimal-crossing diagram of κ .

Question 3.10. Let κ be a generalized knotoid and let $A, B \in P(\kappa)$ be poles. Does every minimal-crossing diagram of κ realize the height between A and B ?

Question 3.10 asks a generalization 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 generalize to non-classical generalized knotoids (for example, generalized knotoids on a genus g surface)?

The proof we have given for Theorem 3.2 does not immediately generalize to higher genus surfaces because it uses the fact that there is a double branched cover of the sphere by itself. Generalizing this method may require rephrasing the argument in terms of a “virtual theory” for generalized knotoids.

Question 3.12. Which matrices of integers are realized as the height spectrum of a generalized knotoid?

4. INDEX POLYNOMIALS FOR GENERALIZED KNOTOIDS

4.1. Notation. Given a crossing c in an oriented diagram, we let $\text{sgn}(c) \in \{1, -1\}$ 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

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

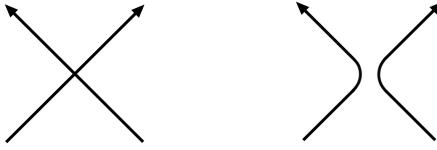


FIGURE 11. A crossing and its oriented smoothing. The smoothing does not depend on over/under data.

where the sum is taken over all crossings between α and β . Observe that the linking number between α and β is an invariant of the generalized knotoid represented by \mathcal{D} .

Given two (possibly identical) oriented generalized knotoid diagrams $\mathcal{D}_1, \mathcal{D}_2$ considered on the same surface Σ 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 α and β , excluding potential intersections at poles. At each $c \in T(\alpha, \beta)$, the tangent vectors v_α, v_β of α and β form a basis for the tangent space of Σ at c . We set $\varepsilon(c) = 1$ if (v_β, v_α) is a positively oriented basis, and $\varepsilon(c) = -1$ otherwise. (Equivalently, $\varepsilon(c)$ gives the sign of c if β is taken to be the over-strand.) The *algebraic intersection number* of α and β , 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 α and β 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 β be a loop constituent on S^2 and fix two points A, B in $S^2 \setminus \beta$. Let α be a segment constituent between A and B intersecting β transversely and away from crossings. Then $\alpha \cdot \beta$ does not depend on the choice of α .*

The *index polynomial* is an invariant for virtual knots introduced by Turaev in [30] and Henrich in [19]. A similar *affine index polynomial* for oriented virtual knots was defined by Kauffman in [21] and extended to knotoids in [16]. Kim, Im, and Lee defined an index polynomial for knotoids in [23] 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 α_c, β_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 $\text{in}(c)$, is computed from the algebraic intersection number of α_c and β_c , and the polynomial invariant takes the form

$$(4.1.1) \quad F_{\mathcal{D}}(t) = \sum_c \text{sgn}(c)(t^{\text{in}(c)} - 1),$$

where the sum is taken over all crossings of the diagram \mathcal{D} . The values of $\text{in}(c)$ 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 $\text{in}(c) = |\alpha_c \cdot \beta_c|$.
- (ii) In Kauffman's affine index polynomial $F^{\text{aff}}(t)$ for oriented virtual knots and knotoids, the constituents α_c, β_c are labeled such that α_c contains the incoming portion of the over-strand and β_c contains the incoming portion of the under-strand. The intersection index is taken to be $\text{in}(c) = w(c) := \alpha_c \cdot \beta_c$. (This is not Kauffman's original formulation of the affine index polynomial in [21], but it was shown to be equivalent in [10].)
- (iii) In Kim, Im, and Lee's index polynomial $F^{\text{ind}}(t)$ for knotoids, the constituents α_c, β_c are labeled so that α_c is the loop constituent and β_c is the segment constituent, and the intersection index is taken to be $\text{in}(c) = \text{ind}(c) := \alpha_c \cdot \beta_c$. See the discussion in Section 5 of [26].

In [16], it is shown that for a knotoid k , the affine index polynomial $F_k^{\text{aff}}(t)$ satisfies $F_k^{\text{aff}}(t) = F_k^{\text{aff}}(t^{-1})$. It follows that the affine index polynomial for knotoids is recoverable from the index polynomial via

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

In [23], the index polynomial is strengthened to a two-variable polynomial

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

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. Generalized index polynomials. We now define index polynomial invariants for generalized knotoids. Let \mathcal{D} be an oriented, constituent-labeled generalized 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 α_c denote the constituent of the smoothed diagram containing the incoming portion of o_c ,

and let β_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 α_c and β_c are not necessarily distinct.

We introduce variables r_{e_1, e_2} , t_{e_1} , and s_{e_1} for every pair of (not necessarily distinct) constituents $e_1, e_2 \in C(\mathcal{D})$ 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 *generalized index polynomial* by

$$(4.2.1) \quad G_{\mathcal{D}}(r, s, t) = \sum_c \text{sgn}(c) g(c).$$

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 generalized knotoid invariant.*

Proof. Suppose a generalized knotoid κ 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 $\text{sgn}(c)$, 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 α_c, β_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 neighborhood where the Reidemeister move is performed, the constituents α_{c_1} and α_{c_2} are identical, and the constituents β_{c_1} and β_{c_2} are identical. Inside the neighborhood, 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, $\text{sgn}(c_1)g(c_1) + \text{sgn}(c_2)g(c_2) = 0$, as desired.

□

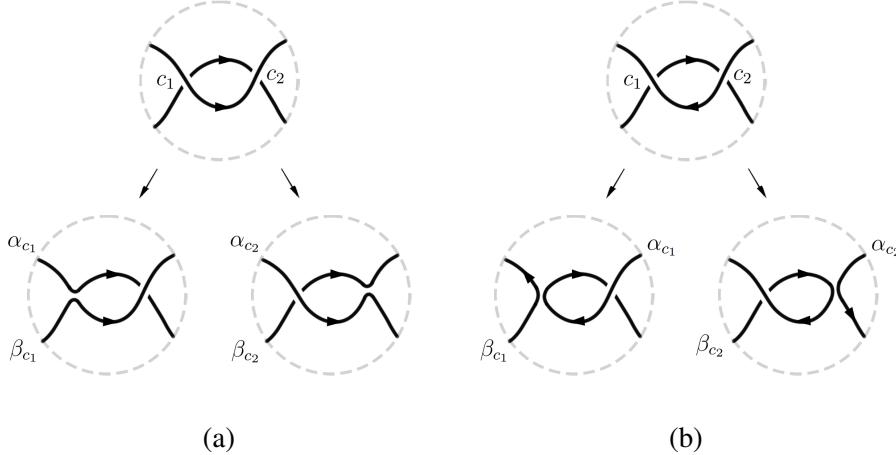


FIGURE 12. Two possible configurations for a Type II Reidemeister move. In each of Figures 12(a) and 12(b), the two diagrams obtained by smoothing at c_1 and c_2 carry the same algebraic intersection data.

The generalized 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 generalized index polynomials.*

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

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

Here, e is the single segment constituent of \mathcal{D} . Observe that $\alpha_c \cdot e = \alpha_c \cdot \beta_c = -\beta_c \cdot e = w(c)$. It follows that the affine index polynomial $F_{\mathcal{D}}^{\text{aff}}(t)$ is recovered from $G_{\mathcal{D}}(s, t)$ by

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

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

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

□

By Proposition 4.1, the generalized index polynomial is trivial for any classical link.

If a generalized knotoid diagram \mathcal{D} is the disjoint union of diagrams \mathcal{D}_1 and \mathcal{D}_2 with no crossings occurring between constituents of \mathcal{D}_1 and \mathcal{D}_2 ,

then the generalized 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 generalized index polynomial.

We extend the generalized index polynomial for pole-labeled generalized 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 $1_L(e)$ 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

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

We define the *base-pointed index polynomial* by

$$(4.2.2) \quad G_{\mathcal{D},P}(r, s, t, a, b) = \sum_c \operatorname{sgn}(c) g(c) h_P(c).$$

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 γ_Q , and the proof that $G_{\mathcal{D},P}$ is a generalized 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 $G_{\mathcal{D},P}$ by the following procedure: modify 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 generalized knotoid κ in Figure 13(a) has

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

and

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

Example 4.6. The generalized knotoid κ in Figure 13(b) has $G_{\kappa,P} = 0$, so the base-pointed generalized index polynomial cannot distinguish κ from a diagram with no crossings. However, every diagram of κ has at least one crossing since $\operatorname{lk}(e_1, e_2) = 1/2$.

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

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

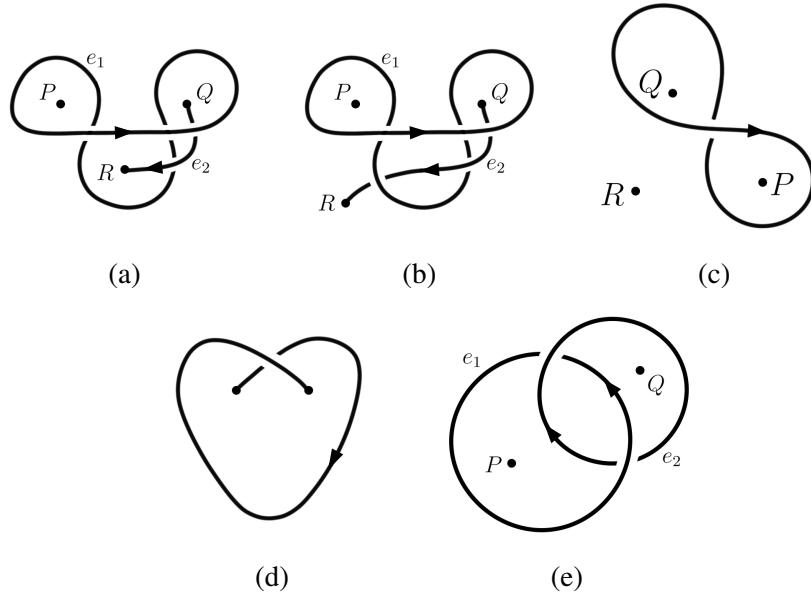


FIGURE 13

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

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

The proof that $\tilde{G}_{\mathcal{D},P}$ is a generalized 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 labeled L and H and a constituent oriented from L to H .

Proposition 4.7. If $\tilde{G}_{k_1,L} = \tilde{G}_{k_2,L}$, then $F_{k_1}^{\text{ind}}(s,t) = F_{k_2}^{\text{ind}}(s,t)$.

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 obtain a two-variable polynomial $\tilde{G}_{\mathcal{D},L}(a_H,b_H)$. For each crossing c in \mathcal{D} , exactly one of α_c, β_c is a loop constituent. The crossings for which α_c is the loop constituent is precisely

the set of early undercrossings, and the crossings for which β_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 α_c is the loop constituent. Observe that β_c defines an oriented shortcut from L to H . It follows from Corollary 4.2 that

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

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

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

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

Example 4.8. The staked knot κ 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 κ by Proposition 4.1.

Question 4.9. Does there exist a pole-labeled, constituent-labeled, oriented generalized knotoid κ 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 generalized knotoid with three poles L , H , and ∞ 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 [26], we extend the basic knotoid involutions to generalized knotoids and state their effects on the generalized index polynomials. Given a diagram \mathcal{D} , we let $\operatorname{mir}(\mathcal{D})$ denote the diagram obtained by toggling all over/under crossing data, and we let $\operatorname{sym}(\mathcal{D})$ denote the diagram obtained by reflecting across a great circle of S^2 (and preserving crossing data). We let rot denote the composition $\operatorname{mir} \circ \operatorname{sym} = \operatorname{sym} \circ \operatorname{mir}$. For an oriented diagram \mathcal{D} , we let $\operatorname{rev}(\mathcal{D})$ 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 generalized knotoids, possibly labeled or oriented. Note that mir , sym , and rot preserve the underlying directed graph of an oriented generalized knotoid, while rev may not.

Proposition 4.10. *Let κ be pole-labeled, constituent-labeled, oriented generalized knotoid and let $P \in P(\kappa)$. The base-pointed index polynomial satisfies:*

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

Identical relationships hold for the pole-centric variant.

Proof. Part (a) follows from the fact that, for each crossing c , the operation mir negates $\text{sgn}(c)$ and swaps α_c and β_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). \square

Thus the base-pointed index polynomials may be used to distinguish between κ and its images under the basic involutions.

The base-pointed index polynomials also yield lower bounds on the height between poles in a generalized 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 κ be a pole-labeled, constituent-labeled, oriented generalized 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 κ and let γ 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 γ 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

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

\square

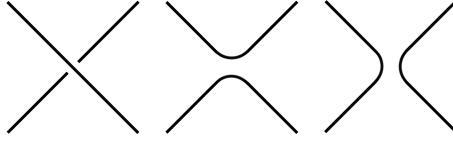


FIGURE 14. A crossing, its *A*-type smoothing, and its *B*-type smoothing, respectively.

Proposition 4.11 implies that the generalized knotoid κ_1 from Example 4.5 has $h_{\kappa_1}(P, Q) = h_{\kappa_1}(P, R) = h_{\kappa_1}(R, Q) = 2$, and the generalized knotoid κ_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 generalized knotoids κ , 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 γ with loops α_c, β_c that result from smoothing at a self-intersection of a constituent e . The generalized knotoid κ 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 $h_{\kappa}(P, Q) \geq 0$.

We remark on non-constituent-labeled generalized knotoids. If κ is simply an oriented, generalized knotoid, we obtain a two-variable generalized index polynomial $G_{\kappa}(s, t)$ by computing the generalized index polynomial of κ with constituents arbitrarily labeled, 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 κ is pole-labeled, 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 GENERALIZED KNOTOIDS

5.1. Generalized bracket polynomial. We define an extension of the Kauffman bracket polynomial for oriented, pole-labeled, edge-labeled generalized knotoids.

Given a generalized 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

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

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 neighborhood 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 *generalized bracket polynomial* $\langle\!\langle \mathcal{D} \rangle\!\rangle$ as the Laurent polynomial in variables A , $\{\lambda_{\{P,Q\}}\}$, $\{x_U\}$, and $\{y_{e,P}\}$ given by

$$(5.1.1) \quad \langle\!\langle \mathcal{D} \rangle\!\rangle = \sum_s A^{\sigma(s)} (-A^2 - A^{-2})^{n_0(s)} \cdot G_s \cdot L_s \cdot E_s.$$

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 [26].)

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. \square

Corollary 5.2. *With e, f as in Lemma 5.1, we have $\mu_{e,f}(P, P) = 0$.*

The *normalized generalized bracket polynomial* $\langle\!\langle \mathcal{D} \rangle\!\rangle_*$ is defined by

$$\langle\!\langle \mathcal{D} \rangle\!\rangle_* = (-A^3)^{-\text{wr}(\mathcal{D})} \langle\!\langle \mathcal{D} \rangle\!\rangle.$$

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

Proposition 5.3. *The polynomial $\langle\!\langle \mathcal{D} \rangle\!\rangle_*$ is a generalized knotoid invariant.*

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

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

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 normalized polynomial $\langle\!\langle \mathcal{D} \rangle\!\rangle_*$ is preserved by Reidemeister moves. \square

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 generalized 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 neighborhood 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. \square

Following the notation of [31], we define the x_i -span, denoted $\text{spn}_{x_i}(F)$, of a Laurent polynomial $F \in \mathbb{Z}[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$ as the non-negative difference between the largest and smallest exponents of x_i among monomials with nonzero coefficients in F (by convention, $\text{spn}_{x_i}(0) = -\infty$). Given a generalized knotoid κ 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 κ be a pole-labeled, constituent-labeled, oriented generalized knotoid and let $P, Q \in P(\kappa)$ be poles. Then*

- (a) $h_\kappa(P, Q) \geq \deg_{\{P \mid Q\}}(\langle\langle \kappa \rangle\rangle_*)$,
- (b) $h_\kappa(P, Q) \geq \frac{1}{2} \max_{e \in E(\kappa)} \deg_{y_{e,P}} \left(\langle\langle \kappa \rangle\rangle_* \Big|_{y_{e,Q}=y_{e,P}^{-1}} \right)$,
- (c) $h_\kappa(P, Q) \geq \frac{1}{2} \max_{e \in E(\kappa)} \text{spn}_{y_{e,P}} \left(\langle\langle \kappa \rangle\rangle_* \Big|_{y_{e,Q}=y_{e,P}^{-1}} \right)$.

Proof. Let \mathcal{D} be any diagram of κ and let γ 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 $d_a = \sum_{U \in \{P \mid Q\}} n_U(s)$. Since γ 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 γ 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

$$(5.1.2) \quad \#(\gamma \cap (e \cup f)) \geq \frac{1}{2} |e \cdot \gamma - f \cdot \gamma|.$$

Let d_b denote the right 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 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. \square

Example 5.6. Let κ denote the generalized knotoid represented by the diagram in Figure 15. Computations yield

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

For brevity, we have omitted the set notation in the subscripts of the λ 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 κ .

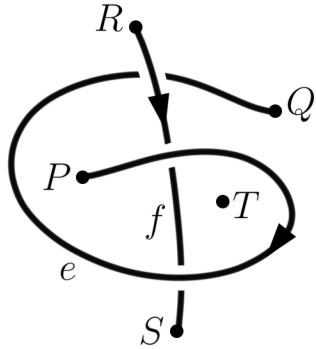


FIGURE 15. An oriented generalized knotoid with five poles P, Q, R, S, T and two constituents e, f .

We remark on the case of generalized knotoids that are not necessarily oriented or fully labeled. Consider an oriented, pole-labeled, constituent-labeled diagram \mathcal{D} . As we remove orientation or labeling data from \mathcal{D} , we alter the polynomial $\langle\!\langle \mathcal{D} \rangle\!\rangle_*$ 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 normalized bracket polynomial by $\langle\!\langle \mathcal{D} \rangle\!\rangle_* = (-A^3)^{-\sum_{e \in C(\mathcal{D})} \text{wr}(e)} \langle\!\langle \mathcal{D} \rangle\!\rangle$, where $\text{wr}(e)$ denotes the sum of the signs of the self-crossings of an arbitrarily-oriented constituent e . (In particular, $\text{wr}(e)$ is independent of orientation.)

Given an oriented, unlabeled diagram \mathcal{D} (for example), we obtain the bracket polynomial $\langle\!\langle \mathcal{D} \rangle\!\rangle_*$ by labeling \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 generalized knotoids with specified labeling and orientation data.

5.2. Recovering existing bracket polynomials. We survey some existing bracket polynomial constructions for classes of objects subsumed by generalized knotoids and show that they are recovered by the generalized 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 , L_s , and E_s are constant, so $\langle\langle L \rangle\rangle_*$ recovers the traditional bracket polynomial $\langle L \rangle$ of the link diagram.

In [11], a bracket polynomial for oriented, ordered (i.e. pole-labeled) multi-linkoids L is defined by

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

Here, the poles of the multi-linkoid are labeled $1, 2, \dots, 2n$, the symbol Λ 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 $\langle\langle L \rangle\rangle_*$ 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

$$\langle\langle K \rangle\rangle_{\circ}(A, u) = (-A^3)^{-\text{wr}(K)} \sum_s A^{\sigma(s)} (-A^2 - A^{-2})^{\|s\|} u^{k_s \cdot \alpha - K \cdot \alpha},$$

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 generalized bracket polynomial, the polynomials $G_s = \lambda_{LH}$ and $L_s = 1$ are constant, and $n_0(s) = \|s\|$. It follows that $\langle\langle K \rangle\rangle_{\circ}$ is recovered from $\langle\langle K \rangle\rangle_*$ 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 Equation 8.3.1 of [31].

In [31], Turaev extends $\langle\langle K \rangle\rangle_{\circ}$ to a bracket polynomial for oriented planar knotoids given by

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

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 generalized knotoid with poles L, H, ∞ 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 $\langle\!\langle K \rangle\!\rangle_*$ by the substitutions $\lambda_{LH} = 1$, $x_{[\{\infty\}]} = B$, $y_{K,\infty} = 1$, and $y_{K,H} = u$.

In [25], 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, $w_\beta: \mathbb{R}^2 \rightarrow \frac{1}{2}\mathbb{Z}$ is the *winding potential function* of an oriented generic closed curve β that maps each point $x \in \mathbb{R}^2$ to the winding number of β 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 γ (resp. γ_s) is defined by $K \cup \alpha^r$, (resp. $k_s \cup \alpha^r$). Kutluay observes that

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

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 β 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 β transversely and away from self-intersections of β . 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 $\langle\!\langle K \rangle\!\rangle_*$ by the substitutions $\lambda_{LH} = 1$, $x_{[\{\infty\}]} = B$, $y_{K,\infty} = \ell h$, and $y_{K,H} = h^{-1}$.

6. KNOTOIDAL GRAPHS

6.1. Knotoidal graphs. We now define an extension of generalized knotoids. Let Σ denote a closed orientable surface, and let G be a finite graph. As with generalized 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 Σ is a pair $(\mathcal{D}, V(\mathcal{D}))$, where \mathcal{D} is a generic immersion of \tilde{G} in Σ 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.

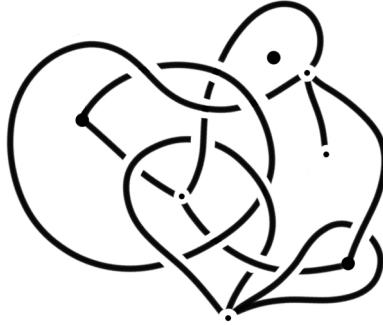


FIGURE 16. A knotoidal graph with three poles and four spatial vertices.

The rest of the terminology for generalized 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 Σ up to ambient isotopy and *generalized 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. 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 generalized knotoids, an equivalent theory arises by considering knotoidal graphs in compact surfaces with boundary while disallowing valency-zero poles.

Example 6.1. Generalized knotoids are knotoidal graphs with no spatial vertices.

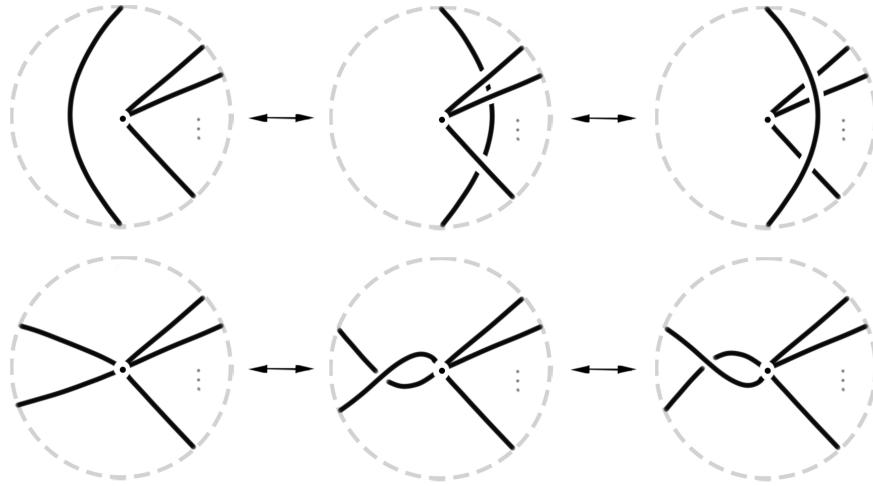


FIGURE 17. The vertex slide move (top) and the vertex twist move (bottom).

Example 6.2. Spatial graphs are knotoidal graphs with no poles.

Example 6.3. A *graphoid* (resp. *multi-graphoid*) as defined in [22] 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 [14], 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 [24] 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 $(D^2 \times I) \setminus \overset{\circ}{N}(\ell_1 \cup \ell_2)$ 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 $(D^2 \times I) \setminus \overset{\circ}{N}(\ell_1 \cup \ell_2)$. Intuitively, the removal of the rails has the effect of topologically enforcing the forbidden move for knotoid diagrams. It is shown in [24] that the theory of planar rail diagrams is equivalent to the theory of planar knotoids, and if the disk D^2 is replaced by an arbitrary surface Σ , the arguments in [24] generalize readily to show that the theory of Σ -rail diagrams is equivalent to the theory of knotoids on Σ .

We extend Σ -rail diagrams to knotoidal graphs. For a given knotoidal graph KG in Σ with diagram \mathcal{D} , take the thickened surface $\Sigma \times I$ and n distinct rails $\ell_i = \{x_i\} \times I, i = 1, \dots, n$, where $n = |P(\mathcal{D})|$, each rail

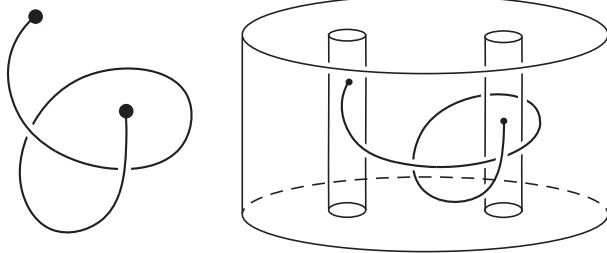


FIGURE 18. A planar rail diagram for a knotoid.

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 neighborhood of each pole. Then for each pole p_i there are $v(p_i)$ endpoints in \tilde{G}' created by removing the disk neighborhood. 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 neighborhood.

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 neighborhoods 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 Σ -rail diagrams of knotoidal graphs is equivalent to the theory of knotoidal graphs in Σ . 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 [18] 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 $\phi_{S^2}^D$ and $\phi_{S^2}^G$ are defined in [2], each sending a spherical knotoid to a knot in the thickened torus $T \times (0, 1)$. Two maps $\phi_{\mathbb{R}^2}^D$ and $\phi_{\mathbb{R}^2}^G$ 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 $\phi_{S^2}^D$ 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 $M = S^2 \times (0, 1) \setminus \dot{N}(\ell_1 \cup \ell_2)$ with an embedded arc, where each $\ell_i = \{x_i\} \times (0, 1)$ is a rail. Let $C_i = \partial N(\ell_i)$ denote the boundaries of the regular neighborhoods of the rails, which are homeomorphic to annuli and are each punctured once by the knotoid (we may take the neighborhoods 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. This yields a well-defined knot K in the thickened torus $T \times (0, 1)$. We then set $\phi_{S^2}^D$ 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 $\text{Vol}_{S^2}(k)$, 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 $\phi_{S^2}^D$ to knotoidal graphs.

Definition 7.2. Consider a spherical rail diagram for a knotoidal graph KG in S^2 . Denote by ℓ_1, \dots, ℓ_m the rails corresponding to nonzero-valency poles p_1, \dots, p_m , and denote by $\ell_{m+1}, \dots, \ell_n$ the rails corresponding to valency-zero poles. Let C_i be the boundary of $N(\ell_i)$ for $i = 1, \dots, m$; we may choose the regular neighborhoods 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 $i = 1, \dots, m$, 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 .

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 $\phi_{S^2}^D$ to knotoidal graphs. It associates to every knotoidal graph in S^2 a spatial graph in a 3-manifold Y , and we write $\phi_{S^2}^D(KG) = (\mathcal{G}, Y)$.

When restricted to spherical knotoids (knotoidal graphs with two poles and a single segment constituent between them), the map $\phi_{S^2}^D$ 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 $\phi_{S^2}^D$ is precisely the planar reflected doubling map $\phi_{\mathbb{R}^2}^D : \mathbb{K}(\mathbb{R}^2) \rightarrow \mathcal{K}(H_3)$ defined in [2].

Note that we can also extend this definition to knotoidal graphs on a general closed orientable projection surface Σ , by replacing every occurrence of S^2 in Definition 7.2 with Σ . Again, we obtain a spatial graph \mathcal{G} in a manifold Y , and we denote the map associating (\mathcal{G}, Y) to KG by ϕ_Σ^D . If Σ 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 Σ .

Definition 7.3. Let KG be a knotoidal graph in a closed orientable surface Σ . Then we say KG is *hyperbolic* if its image $\phi_\Sigma^D(KG) = (\mathcal{G}, Y)$ under the Σ -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 $\phi_{S^2}^D$ with a tg-hyperbolic metric. Then the *hyperbolic volume* $\text{Vol}(KG)$ of KG is defined as

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

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 generalized 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 on a closed orientable surface Σ , we may *stake* it by adding any number of valency-zero poles to its diagram, and we refer to these poles as

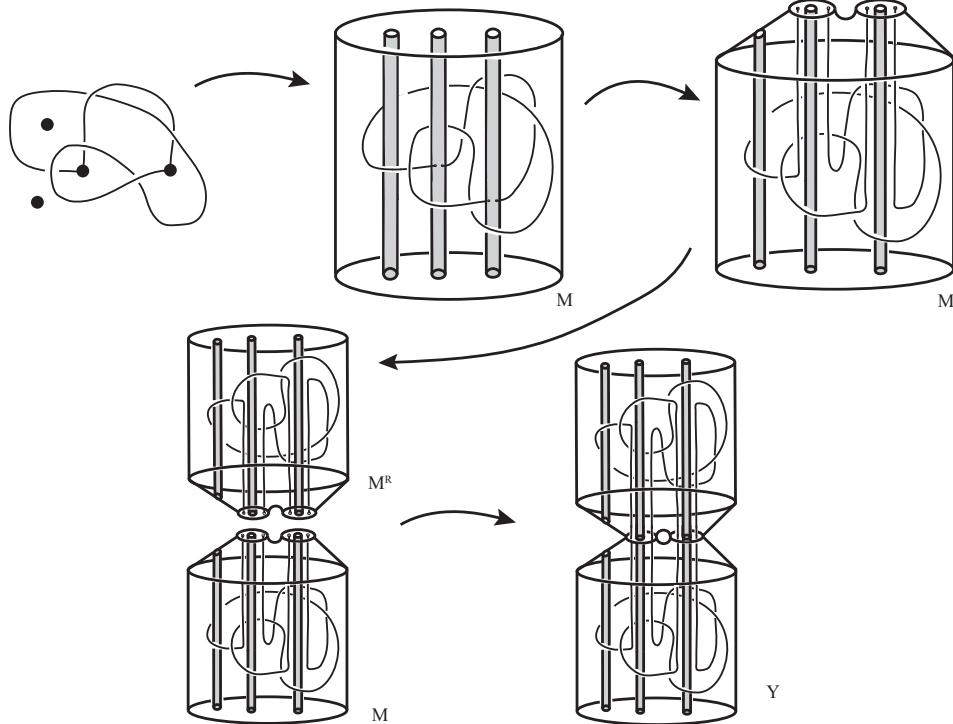


FIGURE 19. The construction for a generalized knotoid or knotoidal graph.

stakes. By means of their rail diagrams, staked links are equivalent to links in handlebodies.

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 $\phi_{S^2}^D$, 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 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 $\phi_{S^2}^D$, 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

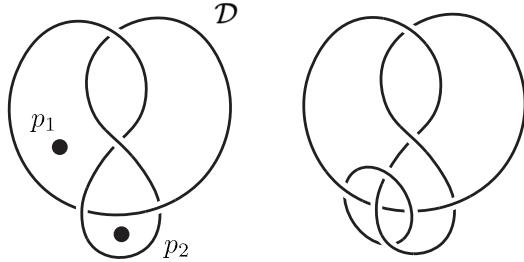


FIGURE 20. Adding two stakes to a link diagram on S^2 is equivalent to adding an unknot component.

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 \mathcal{D}'' is hyperbolic. See Figure 20 for an example.

Observe that when p_1 and p_2 are in distinct, non-adjacent regions, then \mathcal{D}'' 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 \mathcal{D}'' is hyperbolic.

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

Definition 7.5. Let Σ 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 ∂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 ∂D .

Definition 7.6. Let Σ be a closed, orientable surface and let \mathcal{D} be a projection diagram for a link L in Σ . We say \mathcal{D} is *cellular* if every complementary region $\Sigma \setminus \mathcal{D}$ is a disk.

Then Theorem 1.6 from [3], stated below, allows us to determine whether or not a staking yields a hyperbolic staked link.

Proposition 7.7. *Let L be a link on a closed orientable surface Σ and suppose it has a cellular alternating diagram \mathcal{D} . Let L' be a staked link obtained by staking L , and let \mathcal{D}' be a reduced diagram for L' in Σ . If L is not the unknot or a $(2, q)$ -torus link on $\Sigma = S^2$, then L' is hyperbolic if and only if*

- (1) \mathcal{D}' is weakly prime on Σ ;

- (2) every complementary region of $\Sigma \setminus \mathcal{D}'$ contains at most one stake;
- (3) adjacent regions of $\Sigma \setminus \mathcal{D}'$ do not both contain stakes.

If L is the unknot or a $(2, q)$ -torus link on $\Sigma = S^2$, then L' is hyperbolic if and only if conditions (1)-(3) are satisfied and

- (4) L' has at least two stakes.
- (5) the rail diagram for L' is not homeomorphic to the complement of a (p, q) -torus link in the solid torus, where this link may be isotoped to sit on the boundary without crossings;
- (6) the rail diagram for L' is not homeomorphic to the thickened torus.

These conditions are all relatively easy to check directly on the diagram \mathcal{D}' . Note that if L is the unknot or a $(2, q)$ -torus link on S^2 , then conditions (4) and (5) are automatically satisfied if L' has at least three stakes. Also note that this result subsumes Theorem 2.1 from [1].

Proof. As in the remarks following the definitions of generalized knotoids and knotoidal graphs, adding a stake to a link diagram \mathcal{D} on Σ is equivalent to removing an open disk neighborhood of Σ . Hence, we can view \mathcal{D}' as a link diagram on a compact surface Σ' with boundary. Then the proposition is a restatement of Theorem 1.6 and Corollary 1.7 from [3]. \square

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.

Theorem 7.8. *Let Σ be a closed, orientable surface and let $L \subset \Sigma \times I$ be a link with a cellular projection diagram \mathcal{D} with at least two crossings such that $\Sigma \setminus \mathcal{D}$ is checkerboard-colorable. If \mathcal{D} is not a diagram of the 2-component unlink or Hopf link on S^2 with exactly two crossings, then \mathcal{D} can be staked such that the resulting staked link is hyperbolic.*

Proof. Consider a checkerboard coloring of $\Sigma \setminus \mathcal{D}$ and in every shaded face, add a single stake. We call this a *checkerboard staking* of \mathcal{D} , and we claim that the resulting staked link is hyperbolic. We may view it as a link in the thickened surface Σ' with boundary obtained by removing an open disk neighborhood of each stake in Σ . By abuse of notation, we refer to the link viewed in the handlebody $\Sigma' \times I$ by L . We want to show that $M = (\Sigma' \times I) \setminus N(L)$ is tg-hyperbolic.

Each crossing of the resulting diagram \mathcal{D}' locally divides Σ' into four regions, and two of the regions opposite to each other contain a circle boundary where a neighborhood of the stake was removed. For some crossing c , denote these regions by R_1 and R_2 , and denote the respective circle boundaries by C_1 and C_2 . Let α be an arc in Σ' with endpoints on the C_i such that it intersects \mathcal{D} exactly once through the crossing c . Then $\alpha \times I$ is a properly

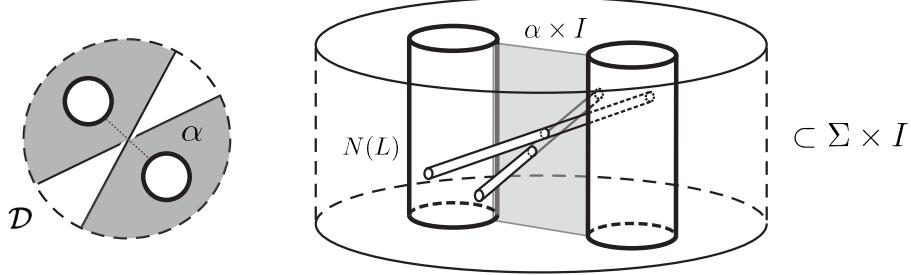


FIGURE 21. A checkerboard staking near a crossing; crossing the arc α by I yields the twice-punctured disk $F = \alpha \times I$ in $\Sigma \times I$, shown on the right.

embedded disk F in $\Sigma' \times I$ which is twice-punctured by L' at the crossing. See Figure 21.

We can cut along F , yielding copies F_1 and F_2 , rotate copy F_2 by 2π , 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. We may switch any of the crossings of L in the handlebody while preserving hyperbolicity. Therefore, we can perform this move several times to obtain an alternating link L' in a handlebody.

Now, we apply Proposition 7.7. Conditions (2) and (3) are automatically satisfied by construction. If $\Sigma = S^2$, then we need to consider conditions (4)-(6). Condition (4) is satisfied since \mathcal{D} has at least two crossings: hence, there are at least four regions of $\Sigma \setminus \mathcal{D}$. Note that conditions (5) and (6) are automatically satisfied if there are at least three stakes. We can choose such a checkerboard staking if there are at least five regions; hence, we may assume that there are exactly two crossings and four regions of $\Sigma \setminus \mathcal{D}$ such that exactly two regions are shaded and two are white in a checkerboard coloring. This only occurs if \mathcal{D} is the diagram of a 2-component unlink or Hopf link with exactly two crossings.

It remains to check condition (1). We may assume that \mathcal{D} is a reduced diagram on Σ' . By definition, it suffices to check that every circle $\gamma \subset \Sigma$ which bounds a disk D in Σ containing crossings of \mathcal{D} and is twice-punctured by \mathcal{D} contains a region of each color. Since D contains at least one crossing, it must contain at least one region. Without loss of generality, say this region is shaded. If it is staked, then we are done; otherwise, \mathcal{D} is staked in the unshaded regions. Since γ meets \mathcal{D} exactly twice, it intersects the shaded faces in a single arc and the unshaded faces in another arc. Let W be the unshaded face intersecting γ ; since γ contains only shaded faces, W is the

only unshaded face γ runs through. Let c be a crossing contained in D and let μ be a small arc in $W \cup D$ which intersects D exactly once through the double point of c . Since W is connected, there is an arc $\sigma \subset W$ connecting the endpoints of μ . Then $\mu \cup \gamma$ is a circle which bounds a disk in Σ' and intersects D precisely once through a double point. This contradicts D' being reduced on Σ' . Hence, all of the conditions of Proposition 7.7 are satisfied and the theorem follows. \square

Corollary 7.9. *Every link L in $S^2 \times I$ has a diagram which can be staked such that the resulting staked link is hyperbolic.*

Proof. Every link diagram in S^2 is checkerboard-colorable. Also, every link diagram in S^2 can be made connected with sufficiently many crossings (via Type I and II Reidemeister moves) so that the previous theorem applies. \square

The checkerboard staking is maximal in the sense that adding an additional stake to D anywhere in the diagram causes it to no longer be hyperbolic. There are examples of links in Σ which are not hyperbolic prior to staking but become hyperbolic after adding just a few stakes. For instance, if L becomes alternating after switching crossings c_1, \dots, c_n , then we may repeat the proof of Theorem 7.8 after staking once in each shaded region adjacent to one of the crossings c_i . In particular, if L is a link which is *almost alternating*, that is, becomes alternating after switching just one crossing, then it can be made hyperbolic by staking just twice (assuming the other conditions in the statement of Proposition 7.7 are satisfied).

However, not all hyperbolic staking configurations arise this way. For example, see the staking configurations of the unknot and trefoil in Figure 22. Hence, given a link L , it would be interesting to determine exactly the staking configurations that yield a hyperbolic staked link, that is, extend Proposition 7.7 to larger classes of links.

Definition 7.10. We define the *staked volume* of a link L in $\Sigma \times I$, where Σ is a closed surface, to be the minimum volume over all diagrams D of L in $\Sigma \times \{1/2\}$ of staked links obtained by adding stakes to D . If no such staking exists for any diagram D , then we set the staked volume of L to infinity.

By Corollary 7.9, we may associate a finite hyperbolic 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 [4].

Example 7.11. The staked volume of the trivial knot in S^3 is $3.6638\dots$, and it is achieved by the staking configuration in Figure 22(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

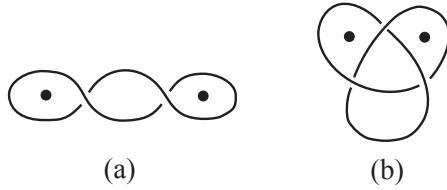


FIGURE 22. Staked diagrams of the unknot and trefoil realizing the minimum volume over all staked diagrams.

solid torus. In particular, this is a 2-cusped manifold. In [5], Agol proved that the minimum volume of a 2-cusped manifold is $3.6638\dots$. 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 [28], we know that any manifold with totally geodesic boundary must have volume at least $3.6638\dots$ and if it has volume exactly $3.6638\dots$, 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 $3.6638\dots$, and it is achieved by the staking configuration in Figure 22(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 utilize two stakes, the result will be the complement of a 2-component link in S^3 , which by [5] has volume at least $3.6638\dots$. 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 $3.6638\dots$.

REFERENCES

- [1] C. Adams. Generalized augmented alternating links and hyperbolic volumes. *Algebraic & Geometric Topology*, 17(6):3375–3397, Oct 2017.
- [2] C. Adams, A. Bonat, M. Chande, J. Chen, M. Jiang, Z. Romrell, D. Santiago, B. Shapiro, and D. Woodruff. Hyperbolic knotoids. *preprint*, 2022.
- [3] C. Adams and J. Chen. Hyperbolicity of alternating links in thickened surfaces with boundary. *in preparation*, 2022.
- [4] 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. *Algebraic and Geometric Topology*, 21:3459–3482, 2021.
- [5] I. Agol. The minimal volume orientable hyperbolic 2-cusped 3-manifolds. *Proceedings of the American Mathematical Society*, 138(10):3723–3732, May 2010.
- [6] 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_ANDFINITE-TYPE_INVARIANTS.
- [7] K. Bhataineh. New polynomial invariants of knotoids and the theory of polar knots. *Mediterranean Journal of Mathematics*, 19:Article 40, 2022.
- [8] H. U. Boden and W. Rushworth. Minimal crossing number implies minimal supporting genus. *Bulletin of the London Mathematical Society*, 53:1174–1184, 2021.
- [9] 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):3402–3404, 2018.
- [10] L. Folwaczny and L. Kauffman. A linking number definition of the affine index polynomial and applications. *Journal of Knot Theory and its Ramifications*, 22:1341004 (30 pages), 2013.
- [11] B. Gabrovšek and N. Gügümcü. Invariants of multi-linkoids. *arXiv:2204.11234*, 2022.
- [12] J. Goldman and L. Kauffman. Knots, tangles and electrical networks. *Advances in Mathematics*, 14:267–306, 1993.
- [13] D. Goundaroulis, J. Dorier, and A. Stasiak. Knotoids and protein structure. *Topol. Geom. Biopolym.*, 746, 2020.
- [14] 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*, 9:444, 2017.
- [15] N. Gügümcü, B. Gabrovsek, and L.H. Kauffman. Invariants of bonded knotoids and applications to protein folding. *Symmetry*, 14, 2022.
- [16] N. Gügümcü and L. Kauffman. New invariants of knotoids. *European Journal of Combinatorics*, 65:186–229, 2017.
- [17] N. Gügümcü and L. Kauffman. Parity, virtual closure and minimality of knotoids. *Journal of Knot Theory and its Ramifications*, 30:2150076 (28 pages), 2022.
- [18] N. Gügümcü, L. Kauffman, and P. Pongtanapaisan. Graphoids. *preprint to appear on the ArXiv*, 2022.
- [19] A. Henrich. A sequence of degree-one vassiliev invariants for virtual knots. *Journal of Knot Theory and its Ramifications*, 19:461–487, 2010.
- [20] R. Kashaev. Invariants of long knots. In *Representation Theory, Mathematical Physics, and Integrable Systems*, pages 431–451. Birkhäuser, 2019.
- [21] L. Kauffman. An affine index polynomial invariant of virtual knots. *Journal of Knot Theory and its Ramifications*, 22:1340007 (30 pages), 2013.

- [22] L. Kauffman, B. Gabrovšek, and N. Gügümcü. Topological invariants of bonded proteins, 2022. Preprint available at https://www.researchgate.net/publication/361633189_TOPOLOGICAL_INVARIANTS_OF_BONDED_PROTEINS.
- [23] S. Kim, Y. H. Im, and S. Lee. A family of polynomial invariants for knotoids. *Journal of Knot Theory and its Ramifications*, 27:1843001 (15 pages), 2018.
- [24] D. Kodokostas and S. Lambropoulou. Rail knotoids. *Journal of Knot Theory and its Ramifications*, 28:1940019 (19 pages), 2019.
- [25] D. Kutluay. Winding homology of knotoids, 2020. Ph.D. thesis; arXiv:2002.07871.
- [26] L. Linov. Signed heights of knotoids. *Journal of Knot Theory and its Ramifications*, 31:2250037 (25 pages), 2022.
- [27] V. O. Manturov. Parity and projection from virtual knots to classical knots. *Journal of Knot Theory and its Ramifications*, 22:1350044, 2013.
- [28] Yosuke Miyamoto. Volumes of hyperbolic manifolds with totally geodesic boundary. *Topology*, 33(4):613–629, 1994.
- [29] X.-S. Lin N. Habegger. The classification of links up to link-homotopy. *J. Amer. Math. Soc.*, 3:389–419, 1990.
- [30] V. Turaev. Cobordism of knots on surfaces. *Journal of Topology*, 1:285–305, 2008.
- [31] V. Turaev. Knotoids. *Osaka Journal of Mathematics*, 49:195–223, 02 2010.

DEPARTMENT OF MATHEMATICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267
Email address: cadams@williams.edu

DEPARTMENT OF MATHEMATICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267
Email address: ajb10@williams.edu

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NJ 08544
Email address: mchande@princeton.edu

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NJ 08544
Email address: joyec@princeton.edu

DEPARTMENT OF MATHEMATICS, M.I.T., CAMBRIDGE, MA 02139
Email address: mdjiang@mit.edu

DEPARTMENT OF MATHEMATICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267
Email address: zr3@williams.edu

DEPARTMENT OF MATHEMATICS, M.I.T., CAMBRIDGE, MA 02139
Email address: dsantiag@mit.edu

DEPARTMENT OF MATHEMATICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267
Email address: bis1@williams.edu

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138
Email address: dorawoodruff@college.harvard.edu