Open Knots

Julien Dorier*[1], Dimos Goundaroulis†[2,3], Eric J. Rawdon‡[4], and Andrzej Stasiak§[5]

[1]Bioinformatics Competence Center, University of Lausanne, Lausanne, Switzerland,
[2]The Center for Genome Architecture, Baylor College of Medicine, Houston, TX, USA
[3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA
[4]University of St. Thomas, Saint Paul, MN, USA
[5]University of Lausanne, Lausanne, Switzerland

Publishing information

This article appears as Chapter 86 of the book:

> Colin Adams, Erica Flapan, Allison Henrich, Louis H. Kauffman, Lewis D. Ludwig, and Sam Nelson, editors. *Encyclopedia of Knot Theory.* Chapman and Hall/CRC, first edition, ISBN 9781138297845, 2020.

1 Introduction

The goal of this chapter is to describe different techniques used to measure knotting in open curves. Note that there is no "agreed upon" definition for describing knotting in open curves. As a result, we describe the context motivating each definition and then describe some advantages and disadvantages of the different approaches.

Traditional knot theory focuses on closed curves. For closed curves, ambient isotopy separates the essence of the knotting in the curve (i.e. the topology or knot type) from the geometry of the configurations. Unfortunately, all non-self-intersecting open curves are ambient isotopic. Therefore, a different paradigm is needed to extract information about knotting in open curves.

Measuring knotting in open curves seems like a reasonable thing to do mathematically. Furthermore, there are scientific motivations. In the 1980s, knot theory started to be used as a tool to understand biological systems involving objects (like DNA and proteins) that could be modeled (coarsely) as polygons. On the other end of the length-scale spectrum, objects such as solar flares and solar wind could be entangled. So while there are closed knotted curves in nature (or curves

---

*julien.dorier@unil.ch
†dimos.gkountaroulis@bcm.edu
‡ejrawdon@stthomas.edu
§andrzej.stasiak@unil.ch

that behave like they are closed, e.g. when both ends are anchored), oftentimes one finds open curves. In fact, there are likely many more entangled open curves in nature than closed ones.

Most techniques to measure knotting in open curves require closing the curve and then rely on knot type classifications from traditional knot theory to describe the entanglement of the open curve. However, some new approaches utilize projections of an open curve (knotoids) or closing the curve using virtual crossings (See Chapter 36). It is important to note that the study of entanglement on open curves, to date, is more of a geometric exercise than a topological one. In particular, these techniques measure the entanglement of any given fixed open configuration, rather than studying classes of configurations with the same entanglement profile (i.e. knot type).

In the following sections, we explore different techniques for measuring knotting in an open curve configuration. We begin by motivating properties we might like to have in such a definition. Following that section, we present the different techniques grouped by class: single closure techniques, probabilistic closure techniques, projection techniques, and topological techniques.

2 Motivation

We might begin by asking what sorts of properties a measure of knotting in open curves should possess.

As a matter of everyday life, we see lots of entanglement: in headphone cables, garden hoses, climbing rigs, etc.. If a person is given a piece of rope and told to tie a knot, the person typically has no problem creating something that most people would classify as being a knot. Thus, we might like the definition of open knotting to be consistent with the opinion of reasonable people, such as the authors of this chapter (this was termed the "reasonable-person test" in [15]). In particular, we are motivated by the three examples in Figure 1. Figure 1(a) is an example of what most people think of as a knot. Figures 1(b) and 1(c) are closed knots with small portions of the curve removed. These three examples all seem to be uncontroversially knotted, and the knotting appears to be trefoil-esque.

So, first, we might like the definition to assess these curves as being knotted, and to associate them with the trefoil in some fashion. We refer to this property as matching our intuition.

Second, we might like open knotting to converge to what we know about closed knots in some sense. For example, if we decrease the length of the missing pieces in Figures 1(b) and 1(c), we might hope for the open knotting classification to converge to the trefoil. Similarly, in Figure 1(a), as the tails get longer and longer we might hope for the open knotting classification to converge to the trefoil. We refer to this property as convergence.[1]

Third, the definition should be well-defined mathematically and able to be measured by a computer (at least for polygons), without human interaction, in some reasonable amount of time. We refer to this property as computability.

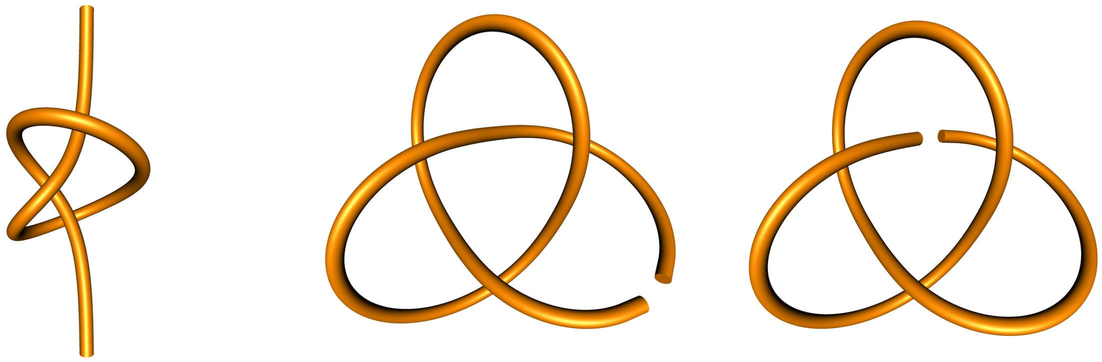
Fourth, the definition should be continuous as a function of the curve (in some sense) and stable relative to small perturbations of the curve.[2] In particular, the classification scheme should minimize the effect of small perturbations on the endpoints. We refer to these properties as continuity and stability.

There are some cases where the "correct" classification seems obvious, e.g. the examples in Figure 1. Another case occurs with full protein chains (see Chapter 90). Proteins have the pleasant

---

[1] Millett called this property "continuity" [17].

[2] Millett called this property "robustness" [17].



(a) (b) (c)

Figure 1: Three examples of open curves that many people would classify as being knotted. In (a) we show an overhand knot which, when closed in a simple fashion, becomes a trefoil knot, 3_1. In (b) and (c), small portions of a KnotPlot 3_1 configuration [21] have been removed. A reasonable classification scheme should call these knotted, although the positioning of the opening near the center in (c) may make us think that it is not as robustly knotted as (a) or (b).

property that the endpoints of protein chains typically lie on or very near the surface of the convex hull. In such a case (we also see this in Figure 1(a)), one can simply connect the endpoints via a circular arc lying outside of the convex hull to produce a closed knot, and that classification should satisfy any reasonable person.

King et al. [13] did the first search for protein slipknots, i.e. sub-arcs of proteins which are knotted, but for which the full protein chain is unknotted. When studying knotting in protein subarcs, we necessarily enter the uncomfortable situation where the endpoints are embedded within the entanglement, and the knotting is ambiguous. One could lament. However, perhaps this issue is more of an opportunity than a problem. Really, we should strive for a definition that is equally applicable to all open curves, not just the simplest cases.

3 Closure techniques

The techniques described in this section are split into two classes: various methods that perform a single closure (Section 3.1) and various probabilistic methods that apply multiple closures (Section 3.2) to describe the entanglement of an open curve. The closure techniques classify the entanglement of open curves in terms of traditional knot types. Note that for many of the definitions below, we include an example image (Figures 4 and 5).

3.1 Single closure techniques

In this section, we review five single closure techniques. The underlying philosophy with these approaches is to select one closed configuration for each open configuration, and then describe the knot type of the open configuration as the knot type of the corresponding closed configuration.

In general, single closure methods have trouble with continuity and stability: in particular, small changes in the position of the endpoints can change the measured knotting immensely. Also, these methods tend to be undefined for some set of configurations (an issue of computability) and will generate classifications that do not match our intuition. Mathematicians tend to be disturbed by these fundamental issues. Physicists are more forgiving in this sense since these methods are quick to compute, and in many cases the classifications agree with the more involved methods. For example, when gathering data on large ensembles of random configurations, the statistics may be similar between using single closure methods versus more involved methods, and the trade-off for speed can justify using the single closure methods.

3.1.1 Direct closure

The easiest way to create a closed configuration from an open configuration is to add a line segment between the endpoints of the open configuration. See Figure 4(a). Of course, there are some configurations for which the additional segment will intersect the curve, but perhaps that set is small enough that one could find a way around that problem. The bigger issue with this approach is that the line segment can pierce the region defining the entanglement, which can cause the resulting classification to be in conflict with our intuition. For example, in Figure 1(a), the direct closure yields an unknot. Similarly, one can create examples where the direct closure would be a more complicated knot than one might classify by eye (see [2]).

3.1.2 Simplification

Many of the early schemes to measure knotting in proteins involved simplifying the open curve to the point where the knot type of the curve becomes clearer. In the protein world, this is called the KMT algorithm [14, 23], although the idea has deep mathematical roots, tracing back to at least [20]. Proteins are modeled, coarsely, as polygons, and the simplification sequentially removes edges as demonstrated in Figure 2. However, Millett et al. [16] showed that the order in which the moves are performed can result in different knot types after simple post-simplification closure, even when the endpoints are on the surface of the convex hull.

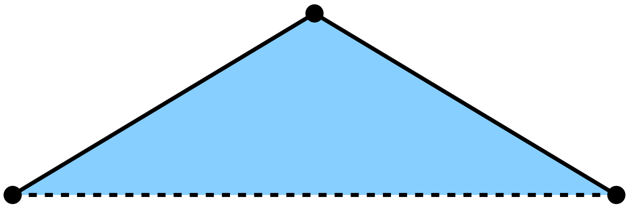


Figure 2: A configuration simplifying move from the KMT algorithm [14, 23]. To simplify an open polygon, one sequentially replaces the two adjacent edges with the dashed edge whenever the shaded triangle is not pierced by any other edge of the polygon. For closed polygons, such moves do not change the knot type. For open polygons, Millett et al. [16] showed that the order in which these moves are performed can affect the knot type classification.

To be fair, simplification is not a definition of knotting, per se, and the KMT algorithm has been employed primarily for proteins (whose endpoints generally are near the convex hull). The algorithm is effective in simplifying closed knots to make the computation of a knot type quicker (since knot type classifications can be exponential in the number of crossings).

3.1.3 Center of mass closure

In the direct closure method, the closure segment stays within the convex hull of a given open curve, thereby making it likely that the segment interacts in some fashion with the curve. The center of mass closure method is, at least in part, an attempt to minimize such interactions. Consider an open curve with endpoints v_1 and v_2 and center of mass c (note: for polygons, researchers typically use the center of mass of the vertex set instead of the edges). The closed polygon is then the union of the curve with two rays: one starting at v_1 with direction $v_1 - c$ and one starting at v_2 with direction $v_2 - c$. In other words, one extends rays from the endpoints in the direction away from the center of mass, and then the curve is closed at infinity. Of course, in practice, one only needs to add line segments along these rays of a sufficient length that one is guaranteed to be outside of the convex hull, and then find a way to connect the endpoints of the segments in a reasonable fashion. See Figure 4(b).

If the endpoints of the curve are on the surface of the convex hull, then this procedure is equivalent to connecting via a circular arc (mentioned earlier), and thus works roughly as desired. In particular, the method sees the trefoil for the images in Figure 1(a) and Figure 1(b), but not for Figure 1(c).

Beyond the continuity problem and the non-definability problem for some configurations, a main problem with this approach occurs when one or both endpoints are near the center of mass. In such a case, a small perturbation of an endpoint can result in huge differences in knot type (i.e. of arbitrarily large differences in crossing number).

3.1.4 Minimal interference

The minimal interference method [25] is an effort to wed the best of the direct closure method and something akin to the center of mass method. In particular, if the endpoints are closer together than the sum of the minimal distances from the endpoints to the surface of the curve's convex hull, then one uses direct closure. Otherwise, one adds segments from each of the endpoints to the nearest point (which may not be unique, but usually is) on the convex hull and then connects the two points on the convex hull with a circular arc lying outside of the convex hull. See Figure 4(c).

This method seems quite reasonable. We have the typical single closure method problems of sometimes not matching our intuition, continuity, stability, and having the quantity being undefined due to intersections between the added segments and the existing curve. Continuity and stability are the major issues. For example, one can construct examples where small perturbations lead to arbitrarily large changes in crossing number [2]. Still, intuitively, this method seems the best of the quantities presented to this point.

3.1.5 Nearest neighbors between open curves and closed curves

There is a standard way to identify almost all equilateral open polygons (or arc-length parametrized open curves) with corresponding closed equilateral polygons (or arc-length parametrized curves) [12, 19] by viewing the closed polygons as a quotient space of the open polygons. Cantarella and Shonkwiler [6] point out that this is a natural closure algorithm, and Cantarella et al. [3] use a variant of this method to construct the closest closed equilateral polygon to a given open equilateral polygon. See Figure 4(d). The advantage of this approach is that there is a firm mathematical underpinning, so theorems are within reach. In particular, when the endpoints are relatively close, the local knotting structure is likely to be preserved. However, when the endpoints are far apart, the motion obtained via the geodesic between an open configuration and its corresponding closed configuration is likely to pass many edges through each other, leading to classifications that do not match our intuition. For example, if we trim the tails a bit on the configuration in Figure 1(a), the corresponding closed configuration is unknotted.

This technique is well defined and continuous on all but a set of polygons of measure zero. It satisfies our convergence and computability properties. However, again, one can construct examples where small perturbations lead to arbitrarily large changes in crossing number, leading to issues with stability. Overall, this technique has many attractive properties.

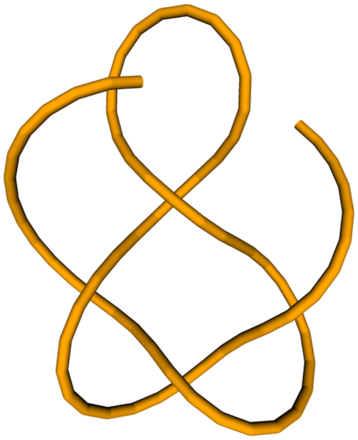


Figure 3: A polygonal arc which is a portion of the KnotPlot 6_2 knot configuration [21]. This is a base configuration used to demonstrate the different techniques in Figures 4, 5, 6, and 7.

3.2 Probabilistic closure techniques

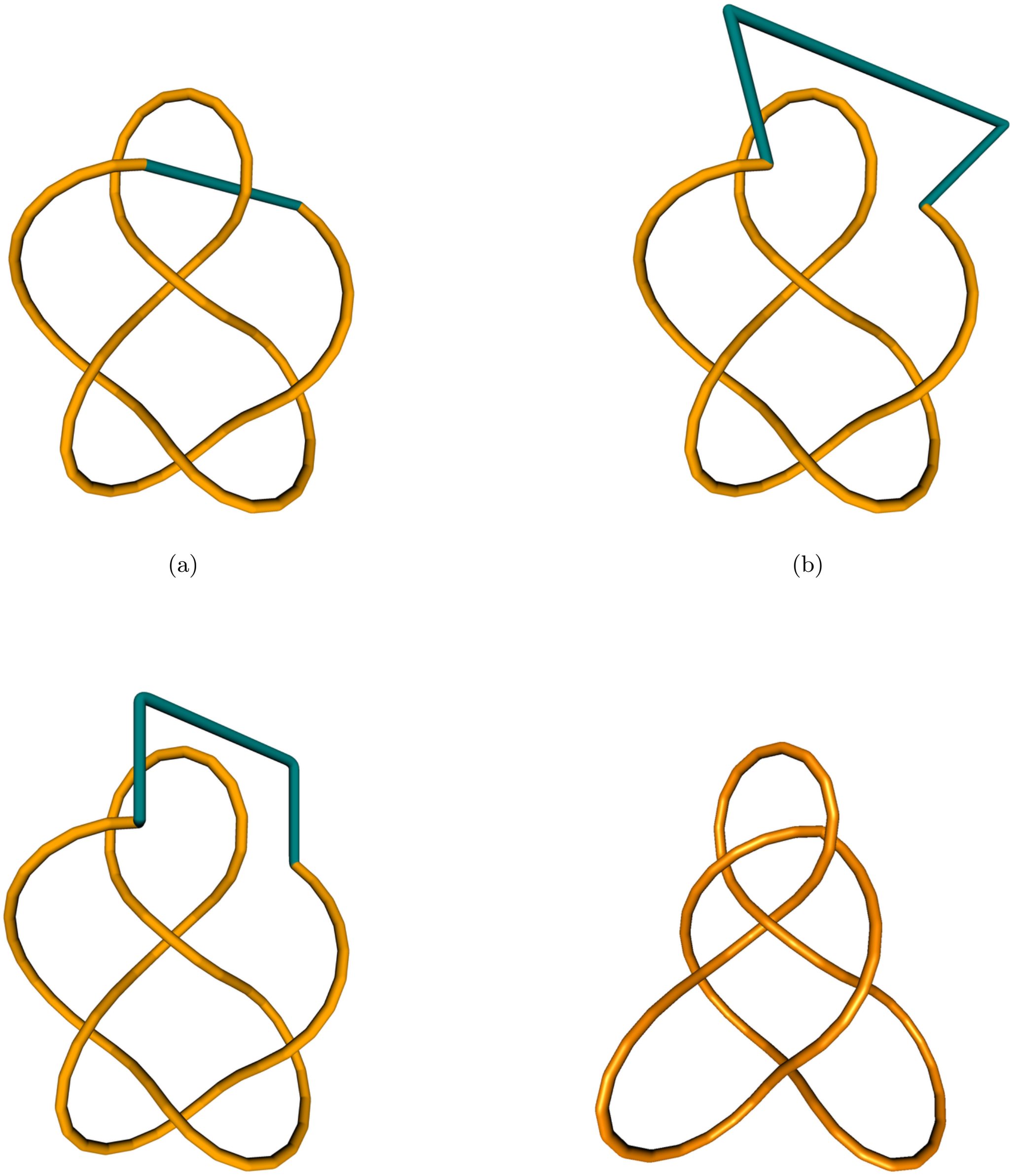
Another overall approach is to think of the open curve as being an incomplete form, and then consider the knotting as a distribution of knot types that the curve could become. The main mathematical problem with approaches of this sort is that it is unclear what the theorems should be, or whether there even are good theorems. The main practical problem is that these approaches can take orders of magnitude longer to compute than the single closure methods, and one can only approximate the distributions.

On the other hand, these methods do generally satisfy all of desired properties mentioned earlier.

Overall, the idea here is to assign a distribution of objects to a given open curve. The main question is: what is the right set of objects?

3.2.1 Double infinity closure

If one were handed an entangled shoelace string and told to pull it tight, the person would have to choose in which direction to pull each end. Humans would generally choose to pull the ends of



(a) (b) (c) (d)

Figure 4: Examples of single closure techniques applied to the configuration in Figure 3. The single closure techniques (a) direct closure, (b) center of mass closure, (c) minimal interference, and (d) nearest closed neighbor applied to this configuration. For (c), one of the endpoints is on the convex hull and the other is very close to the convex hull. The teal arc in (c) is simply one example of a simple arc lying outside of the convex hull which connects the two endpoints.

the string in fixed (but possibly different) directions as opposed to performing some complicated pattern. In the double infinity closure method, one extends rays from each endpoint in independently chosen directions and then closes at infinity [15]. There is an $S^2 \times S^2$ of such choices. Then the knotting of the open curve is the distribution of knot types obtained from this procedure. For any given open curve, there are pairs of directions which yield degenerate self-intersecting closed curves, but the set of such pairs is of measure zero. Thus, this method is well defined. See Figure 5(a).

In the following section, we discuss a more commonly used probabilistic method, a sibling of this method called the infinity closure. Since the two methods share many advantages and disadvantages, we reserve our discussion of their properties for the next section.

3.2.2 Infinity closure

In the previous section, the curve is closed at infinity by extending rays in independent directions from the endpoints. In the infinity closure method, one uses a common direction for the rays (so there is an S^2 worth of direction choices). See Figure 5(b). Also, see Figure 7 for an example of the map of knot types as a function of the ray directions. From a computational perspective, sampling from S^2 is significantly less strenuous than sampling from $S^2 \times S^2$. We typically have used 100 closures per open curve, so one would need $100^2 = 10000$ closures with the double infinity closure method to obtain the same density of information. Furthermore, there is precedence in knot theory for analyzing knotting over a single sphere of directions. In particular, the average crossing number and writhe are both computed by averaging over a sphere. The average crossing number is the average over S^2 of the number of crossings observed when the knot is projected in each of the S^2 different directions. The writhe is computed similarly, but using the sum of the values $+1$ for each right-handed crossing and -1 for each left-handed crossing in the projections.

This method was originally proposed by Millett et al. [18], although in lieu of parallel rays, they closed to a common point on a large sphere encompassing the open curve.

There are many things to like about this strategy: it yields a trefoil for a great majority of the directions in all three examples from Figure 1; the distributions change continuously as a function of the endpoint positions; no human interaction is needed in the computation; the method (while slower than the single closure methods) is straightforward to compute; the method sees a straight line segment as being 100% unknotted; the method provides a well-defined measurement of entanglement for open curves regardless of the position of the endpoints; and as the gap in Figure 1(b) and 1(c) approaches zero, the distribution approaches 100% trefoil knots.

On the negative side: one can only approximate the distribution of knot types (although one could, perhaps, compute the exact distributions for polygons with small numbers of edges); the computations are time-consuming; the definition seems difficult to use theoretically; and a distribution of knot types might not be the most desirable way to describe knotting.

Regarding the last point, researchers typically use the knot type seen with the highest percentage over the different closures to be the "knot type" of the open curve. Millett [17] observed that 996 of 1000 random equilateral open polygons with 300 edges had one knot type that appeared in at least 50% of the closure directions. This result suggests that assigning the predominate knot type in this fashion is generally quite reasonable. Of course, one can concoct open curves where two or more knot types occur in a near tie for the highest percentage, or where the direction sampling (with say 100 closure directions) is not sufficient to declare a clear predominate knot type. However, from the Section 3.1 we know that assigning a single knot type to an open curve can produce awkward moments. At the least, the percentage of the predominate knot type provides a measure of confidence in the assignment. One could use a cut-off (such as 50%) in place of using the predominate knot type, or devise some other scheme, if one is weary of commitment.

3.2.3 Random arc closure

Instead of closing via rays heading to infinity, we could connect the endpoints via random arcs. In particular, for a given open polygonal arc, there is some fixed distance between the endpoints. One need only randomly sample from open polygons (or, conceivably, curves) with that endpoint distance. One could then use the distributions of knot types, or possibly the difference between that distribution and some other base distribution of knot types, to describe the knotting of the arc.

Cantarella's lab's `plCurve` library is able to generate random equilateral polygons with a given distance between the endpoints and a given number of edges, and is derived from theory in [4]. See Figure 5(c).

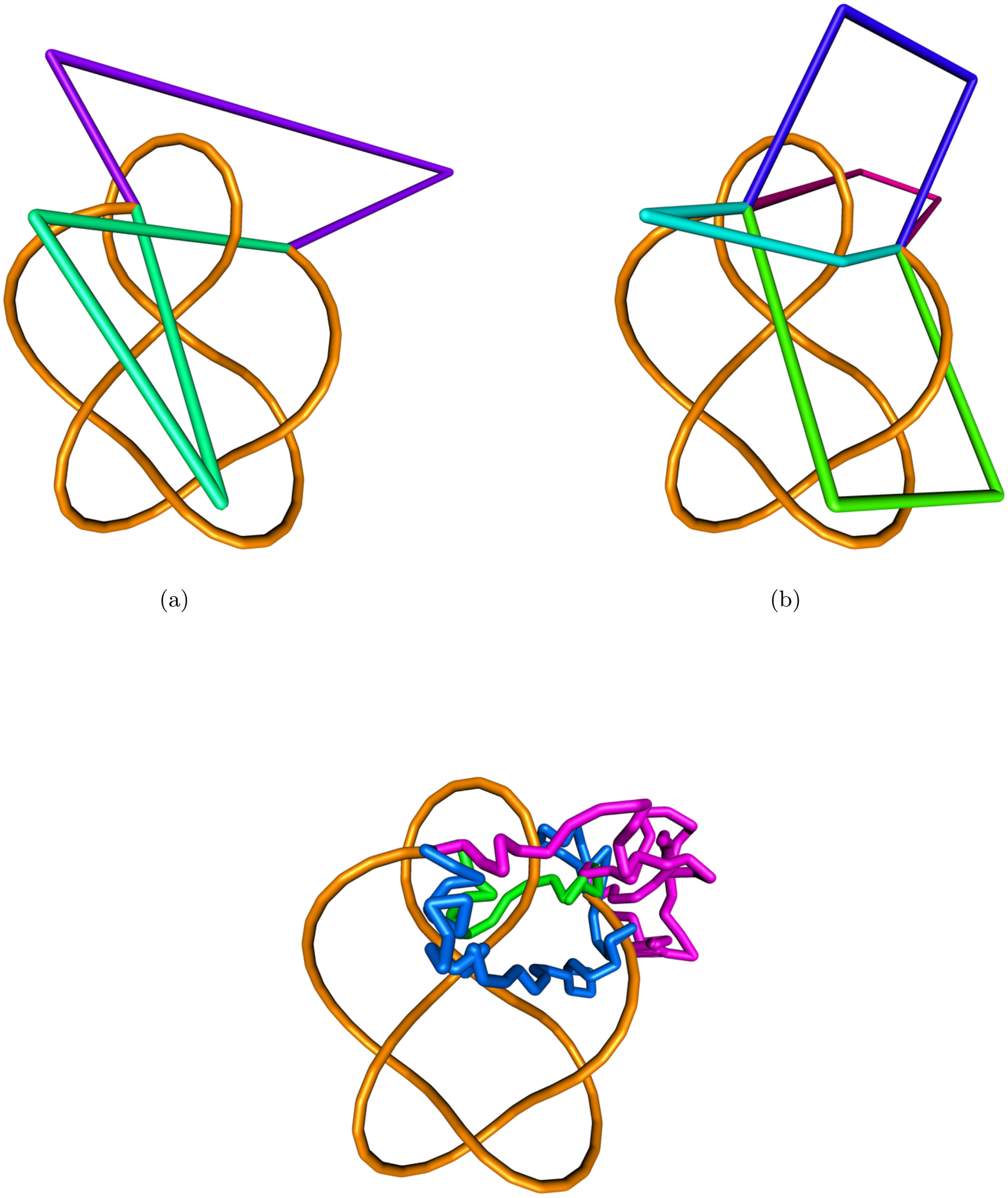
The main issue with this technique is how to choose the length, or lengths, of the connecting arcs. Perhaps there is a way around this problem, for example by assigning a likely length or a distribution of likely lengths. If one chooses a fixed number of equilateral edges with which to close a given equilateral open chain, then one is in an honest-to-goodness knot space, one even with a metric. This underlying measure makes this technique attractive mathematically.

Beyond the issue of choosing the length of the connecting arcs, the random closures technique shares the same practical problems as the single and double infinity closures: namely, one can only approximate the distribution of knot types; the computations are time-consuming; the definition seems difficult to use theoretically; and a distribution of knot types might not be the most desirable way to describe knotting.

4 Measuring entanglement without requiring a closure

The introduction of the concept of knotoids by V. Turaev [26] brought into play new ways of measuring the entanglement of an open curve. In particular, one may leave the curve open and rely on the classification of its projections by their knotoid type to describe the entanglement. Knotoids are equivalence classes of open-ended diagrams that comprise an extension of classical knot theory (see Chapter 36) for an introduction to knotoids). Furthermore, knotoids generalize the notion of a long knot or 1-1 tangle, allowing the definition of an open knot type, at least at the diagrammatic level. In analogy to the techniques that search for a closure for the curve, it is the projections this time that are functions of the geometry of the curve. Here the underlying philosophy is to assign a knotoid type (or a probability distribution of knotoid types) to the open configuration.

For a given open curve, each projection direction (there are S^2 of these) determines a knotoid. Different choices of projection planes can change the placement of the endpoints which, in turn, can change the measured entanglement (i.e. knotoid type) of the curve dramatically, see Figure 6(a). Notice that the three projections shown are not isotopic to each other as knotoid diagrams. More precisely, knotoid (A) is a knotoid with 6 crossings, knotoid (B) is isotopic to a knotoid diagram with just two crossings, and knotoid (C) reduces to the trivial knotoid diagram which is just a single unknotted long segment. Choosing a single projection (which would yield a single knotoid) suffers from the same continuity and stability issues as the single closure techniques. Thus, we deduce that the knotoids approach is best applied as a probabilistic method.

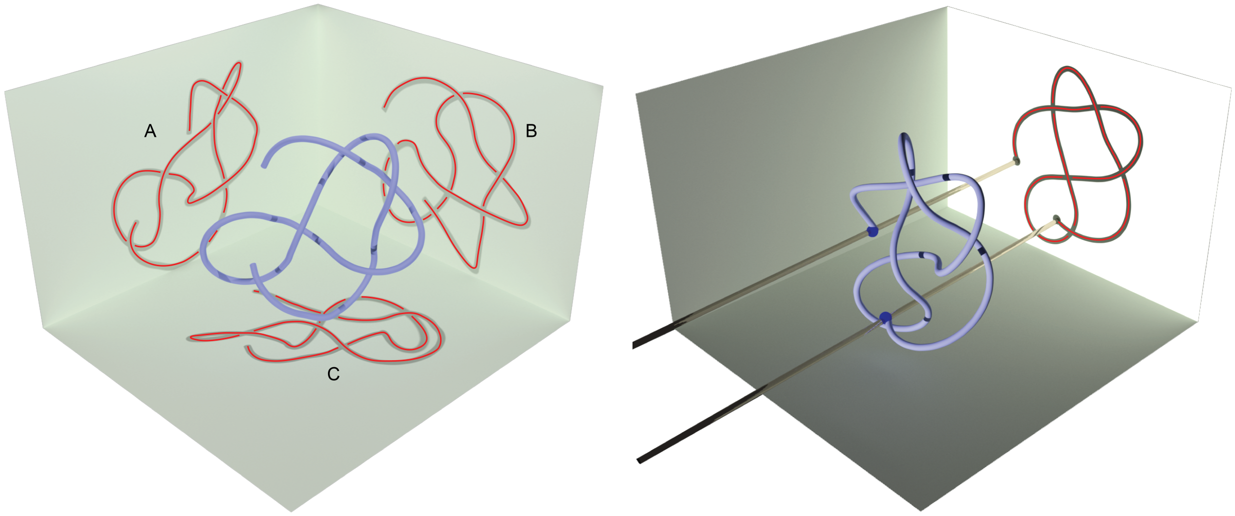


(a)

(b)

(c)

Figure 5: Examples of probabilistic closure techniques on the configuration from Figure 3. In (a), we show two examples of closures from the double infinity closure technique. While the polygons are officially closed at infinity, one can connect the endpoints of the line segments once both of the rays pierce the convex hull. In (b), we show four examples of parallel rays from the infinity closure technique. In (c), we show three connecting random arcs with different numbers of edges: 20 (green), 50 (blue), and 60 (purple).



(a) (b)

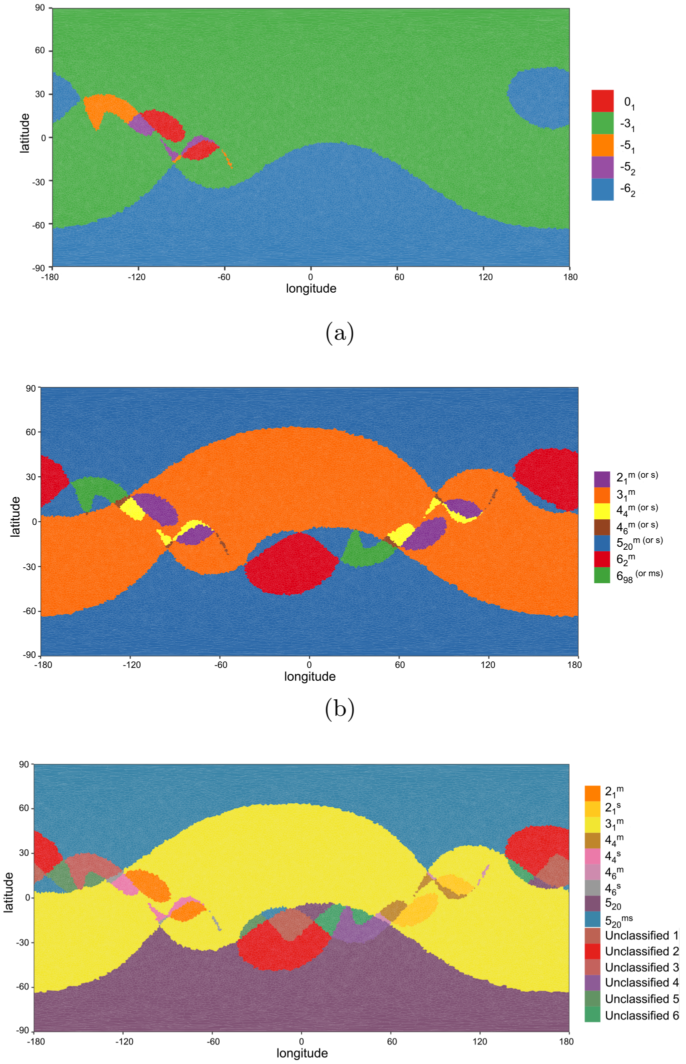
Figure 6: Knotoid projections from the open configuration of Figure 3. In (a), three knotoids are produced by projecting the configuration onto three different planes. In (b), a knotoid coming from a projection of the open configuration is shown together with the infinite lines that fix the endpoints region-wise.

Similar to the probabilistic closure method that produces a spectrum of knots, the knotoids approach assigns to a curve a spectrum of knotoids resulting from different directions of projections.

To compute knotoid types more efficiently, the knotoid projections can often be simplified using an algorithm similar to the KMT algorithm [14, 23]. To do so, one adds two parallel lines to the knotoid projection: these lines are perpendicular to the chosen projection plane and pass through one of the endpoints, see Figure 6(b). Then a triangle of the polygonal knotoid is simplified (as in Figure 2) only if it is not pierced by any other arc of the knotoid or by either of the two infinite lines. These KMT-like simplifications do not change the knotoid type and thus are able to give precise classifications even for sub-arcs where the endpoints lie inside the convex hull. On a side note, the introduction of the two infinite lines allows for the development of theorems, such as in [11] where the conditions are shown for when two such ensembles of curves and lines are isotopic with respect to the same plane of projection.

As in the case of the probabilistic closure methods, the knotoid approach classifies the configurations of Figure 1 as a trefoil for most of the projections, the distributions change continuously as a function of the endpoint positions, and, in general, it carries all the advantages and disadvantages of the infinity closure.

It is important to mention here that knotoids can be studied on any oriented surface like, for example, on S^2. In fact, there is a well-defined map between the set of all planar knotoids (knotoids in $\mathbb{R}^2$) and the set of all spherical knotoids (knotoids in S^2) which is induced by the inclusion of $\mathbb{R}^2$ in S^2 [11, 26]. Moreover, it can be shown that this map is surjective but not injective since one can find examples of knotoids that are nontrivial in $\mathbb{R}^2$ and trivial in S^2. It was shown recently that the entanglement analysis of subchains of an open curve using spherical knotoids provides a more detailed overview of the topology of a curve than using knot types via infinity closures [8].



(a)

(b)

(c)

Figure 7: The knot types, spherical knotoid types, and planar knotoid types obtained as a function of the ray direction for the configuration from Figure 3. Each latitude/longitude pair on the unit sphere specifies a unique closure or projection direction. In (a), a sampling of different spherical coordinates yields a map of knot types (here, flattened from the sphere). The predominate knot type is the one with the largest area on the sphere, and this knot type is typically used as a proxy to describe the knotting of the open arc. In this case, the green region, corresponding to the -3_1 knot, has the largest area, so this configuration normally would be referred to as forming an open -3_1 knot. The maps in (b) for spherical and (c) for planar knotoids are similar. For (b) and (c) the latitude/longitude pair define a projection direction, which in turn defines a spherical or planar knotoid. The coloring is based on the spherical or planar knotoid types. For all three analyses, a sample of 40,000 directions was used.

This analysis can be further refined if one chooses to work with planar knotoids [10].

In Figure 7 one can see a comparison between the probabilistic methods using (single) infinity closures, spherical knotoids, and planar knotoids. Each small region in these maps corresponds to a region on the sphere (shown in spherical coordinates). The different knot/knotoid types are indicated by the different colors in the legend of each image. Each point corresponds to either a closure or a projection direction for the studied open curve. The notation for knotoids follows the classification in [9]. The exponents m, s, and ms correspond to the following involutive forms of the respective knotoids shown in the table of [9]: mirror reflected, symmetric, and mirror symmetric. Note that the projection map for the spherical case shows antipodal symmetry, while the one for the planar case does not. The knotoid types are computed using invariants. The Jones-type invariants for spherical knotoids available in the literature (like the Jones polynomial and the arrow polynomial) cannot distinguish between a knotoid and its mirror symmetric involution (nor between the mirror reflection of a knotoid and its symmetric involution). However, this issue may resolve itself in the future with a new invariant for spherical knotoids. The loop arrow polynomial [9, 10] is an invariant for planar knotoids that, in principle, can distinguish a planar knotoid from its involutions.

Returning to the comparison of projection maps between knots and knotoids, we see that new regions appear in the projection map as we transition from infinity closures to spherical knotoids and to planar knotoids. This behavior is good news and bad news. The knotoid methods provide additional information about the knotting in an open curve. However, since there are more regions with the knotoids than with the knot types using the infinity closure, the percentages associated with each knotoid type are smaller, which makes the assignment of a predominate knotoid type precarious.

We note here that a computational tool that analyzes open curves, using both knotoid approaches as well as the infinity closure approach, has been implemented [7]. The projection maps in Figure 7 were produced using this software.

An equivalent approach to spherical knotoids is the virtual closure technique [1] where the knotoid diagram is closed to a virtual knot (see Chapter 36). More precisely, the endpoints are joined by an arc that crosses the rest of the diagram using virtual crossings. The resulting virtual knot is independent of the choice of the closing arc. Since all virtual crossings occur along the closure arc, one may consider this arc as living in a single handle between the endpoints of the open arc and, thus, they do not interfere with the rest of the diagram. Therefore, the virtual closure technique is a probabilistic method equivalent to the spherical knotoids approach with one extra step, the addition of the virtual closure on a knotoid diagram.

5 Topological techniques

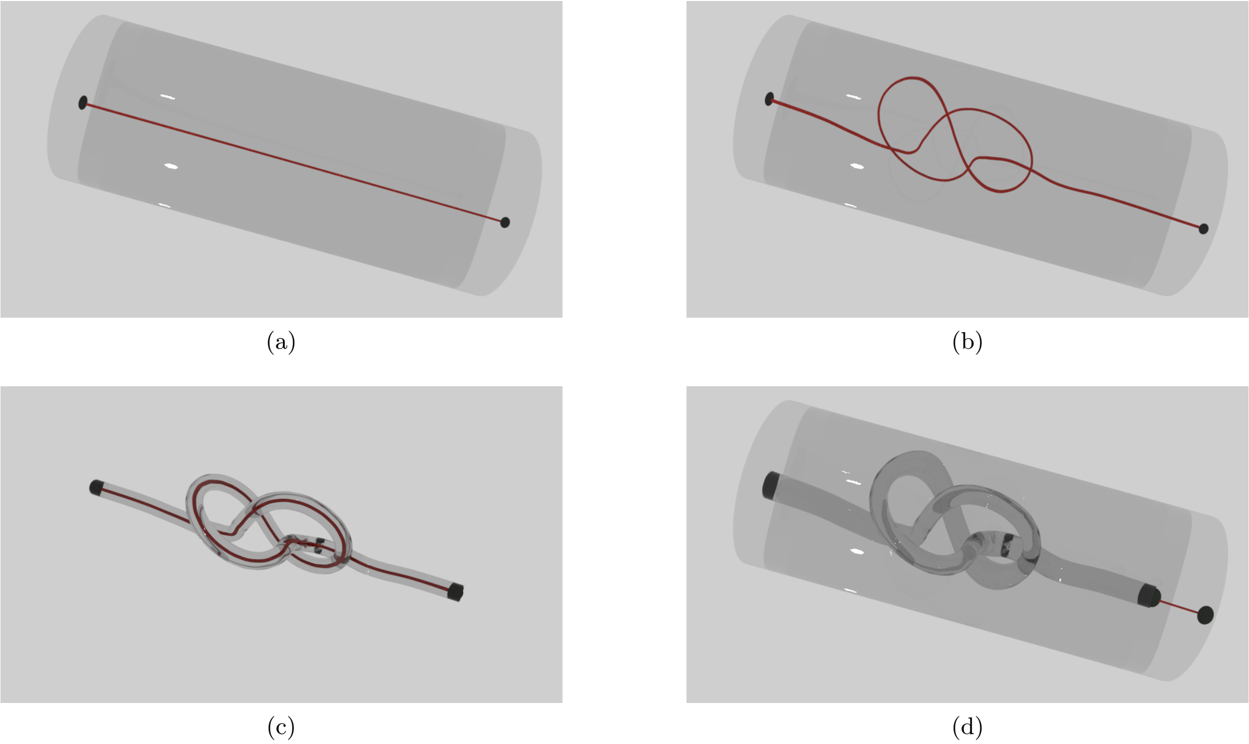
Below we review two topological approaches to measure knotting in open curves.

5.1 Ball-arc pair

One topological approach is to use a ball-arc pair. In particular, a *ball-arc pair* is an embedded open curve (the arc) and a topological sphere (i.e. an object homeomorphic to a round sphere) that does not intersect the curve except at the two endpoints of the curve. One can analyze ball-arc pairs as topological objects. However, as we show below, the prescription of the ball to the open curve is equivalent to assigning a closing curve [2]. So while the ball-arc pair does allow one to analyze the knotting of an open curve topologically, that is because the geometry of the arc acts in conjunction with the choice of the topological sphere.

To illustrate this issue, we show a classic example from [22] in Figure 8. The topological spheres are shown as transparent tubes, the arcs are shown in red, and the intersection between the balls and the arcs are shown as black round spheres. Note that in (d), the topological sphere is the clear outside and the dark gray part "inside" (that looks like an overhand knot) while the arc is the short straight segment on the right. By deforming the sphere in (c), one can obtain (a), so the ball-arc pairs of (a) and (c) are equivalent. Similarly, the ball-arc pairs of (b) and (d) are equivalent. While the arcs of (a) and (d) are both straight, the ball-arc pairs of (a) and (d) are not equivalent, due to what might be considered an unusual choice of a topological sphere in (d). In (d), the union of the straight segment and any non-self-intersecting curve lying on the topological sphere that connects the two segment endpoints is a trefoil. Thus, choosing a topological sphere is equivalent to choosing a closing curve.

Certainly, it seems like one would want a straight line segment to be considered mainly unknotted, so this behavior limits the applicability of this approach. In the end, the ball-arc pair approach does allow one to prove theorems, but cannot independently describe the knotting of an open curve in a way that matches our intuition. To be fair, the ball-arc pair was not defined as a technique to classify knotting in a fixed open configuration.



(a) (b)

(c) (d)

Figure 8: Four ball-arc pairs originally presented in [22]. The topological spheres are shown as transparent tubes, the arcs are shown in red, and the intersection between the balls and the arcs are shown as black round spheres. The ball-arc pairs from (a) and (c) are equivalent, as are the pairs in (b) and (d). However, the ball-arc pairs from (a) and (c) are not equivalent to the ball-arc pairs in (b) and (d).

5.2 Polygons with fixed edge lengths

If we concentrate on open polygons, it is possible to use a topological approach. Consider the class of open polygons where the edge lengths are fixed (although they need not be all the same) and define equivalence in terms of ambient isotopies that keep the edge lengths fixed. In this situation, Cantarella and Johnston [5] showed that there are “stuck unknots.” In other words, there exist open polygons that cannot be deformed into (non-self-intersecting) collinear segments via this restricted class of ambient isotopies. One could then define “knot types” as connected components of the given polygon space, i.e. polygons with a given array of edge lengths using ambient isotopies that fix the edge lengths. Again, this approach allows one to prove theorems. One might imagine that this technique could be useful for biological chains since chemical bonds are modeled as edges, and these lengths are usually fixed (or close to fixed). However, biological chains tend to have some electrostatic repulsion which keeps the edges from getting too close together. With this property, it is unlikely that the biological chains could be “stuck.” This approach can be used for applications such as the movement of robot arms and motion planning (see Chapter 9 of [24]).

6 Conclusion

We have presented techniques used by researchers to describe the knotting of open curves, focusing primarily on open polygons. There are many options, and this chapter provides researchers with insights into which approach might be most useful for a given project. Furthermore, we hope that the discussions inspire researchers to develop new techniques and create new mathematics to understand knotting in open arcs.

Acknowledgments

We thank Jason Cantarella for helpful comments. This material is based upon work supported by the National Science Foundation under Grant Numbers 1418869 and 1720342 to EJR, and by the Swiss National Science Foundation [31003A_166684] and the Leverhulme Trust [RP2013-K-017] to AS.

References

[1] Keith Alexander, Alexander J. Taylor, and Mark R. Dennis. Proteins analysed as virtual knots. *Sci. Rep.*, 7:42300, 2017.

[2] Erin Brine-Doyle, Madeline Shogren, Emily Vecchia, and Eric J. Rawdon. Open knotting. In Philipp Reiter, Simon Blatt, and Armin Schikorra, editors, *New Directions in Geometric and Applied Knot Theory.* De Gruyter, 2018.

[3] Jason Cantarella, Kyle Chapman, Philipp Reiter, and Clayton Shonkwiler. Open and closed random walks with fixed edgelengths in $\mathbb{R}^d$. *J. Phys. A*, 51(43):434002, 2018.

[4] Jason Cantarella, Bertrand Duplantier, Clayton Shonkwiler, and Erica Uehara. A fast direct sampling algorithm for equilateral closed polygons. *J. Phys. A*, 49(27):275202, 2016.

[5] Jason Cantarella and Heather Johnston. Nontrivial embeddings of polygonal intervals and unknots in 3-space. *J. Knot Theory Ramifications*, 7(8):1027–1039, 1998.

[6] Jason Cantarella and Clayton Shonkwiler. Personal communication, 2018.

[7] Julien Dorier, Dimos Goundaroulis, Fabrizio Benedetti, and Andrzej Stasiak. Knoto-ID: A tool to study the entanglement of open protein chains using the concept of knotoids. *Bioinformatics*, 34(19):3402–3404, 2018.

[8] Dimos Goundaroulis, Julien Dorier, Fabrizio Benedetti, and Andrzej Stasiak. Studies of global and local entanglements of individual protein chains using the concept of knotoids. *Sci. Rep.*, 7:6309, 2017.

[9] Dimos Goundaroulis, Julien Dorier, and Andrzej Stasiak. A systematic classification of knotoids on the plane and on the sphere. arXiv:1902.07277, 2019.

[10] Dimos Goundaroulis, Neslihan Gügümcü, Sofia Lambropoulou, Julien Dorier, Andrzej Stasiak, and Louis Kauffman. Topological models for open-knotted protein chains using the concepts of knotoids and bonded knotoids. *Polymers*, 9(9):444, 2017.

[11] Neslihan Gügümcü and Louis H. Kauffman. New invariants of knotoids. *European J. Combin.*, 65:186–229, 2017.

[12] Michael Kapovich and John J. Millson. The symplectic geometry of polygons in Euclidean space. *J. Differential Geom.*, 44(3):479–513, 1996.

[13] Neil P. King, Eric O. Yeates, and Todd O. Yeates. Identification of rare slipknots in proteins and their implications for stability and folding. *J. Mol. Biol.*, 373:153–166, 2007.

[14] Kleanthes Koniaris and Murugappan Muthukumar. Self-entanglement in ring polymers. *J. Chem. Phys.*, 95(4):2873–2881, 1991.

[15] Marc L. Mansfield. Are there knots in proteins? *Nat. Struct. Biol.*, 1:213–214, 1994.

[16] Kenneth Millett, Akos Dobay, and Andrzej Stasiak. Linear random knots and their scaling behavior. *Macromolecules*, 38(2):601–606, 2005.

[17] Kenneth C. Millett. The length scale of 3-space knots, ephemeral knots, and slipknots in random walks. *Prog. Theor. Phys. Suppl.*, 191:182–191, 2011.

[18] Kenneth C. Millett, Akos Dobay, and Andrzej Stasiak. Linear random knots and their scaling behaviour. *Macromolecules*, 38:601–606, 2005.

[19] John J. Millson and Brett Zombro. A Kähler structure on the moduli space of isometric maps of a circle into Euclidean space. *Invent. Math.*, 123(1):35–59, 1996.

[20] Kurt Reidemeister. Elementare begründung der knotentheorie. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 5(1):24–32, 1927.

[21] Robert G. Scharein. KnotPlot. http://www.knotplot.com, 1998. Program for drawing, visualizing, manipulating, and energy minimizing knots.

[22] De Witt Sumners and Stuart G. Whittington. Knots in self-avoiding walks. *J. Phys. A*, 21(7):1689–1694, 1988.

[23] William R. Taylor. A deeply knotted protein and how it might fold. *Nature*, 406:916–919, 2000.

[24] Csaba D. Toth, Joseph O'Rourke, and Jacob E. Goodman, editors. *Handbook of Discrete and Computational Geometry.* CRC Press, third edition, 2017.

[25] Luca Tubiana, Enzo Orlandini, and Cristian Micheletti. Probing the entanglement and locating knots in ring polymers: A comparative study of different arc closure schemes. *Prog. Theor. Phys. Suppl.*, 191:192–204, 2011.

[26] Vladimir Turaev. Knotoids. *Osaka J. Math.*, 49(1):195–223, 2012.