UNIVERSITÉ LIBRE DE BRUXELLES
Faculté des Sciences

Geometric Reconfigurations

Thèse présentée en vue de l'obtention
du grade de Docteur Agrégé de l'Enseignement Supérieur

Stefan LANGERMAN
Année académique 2006–2007

φ

Contents

Chapter 1

Introduction

Geometry represents real world objects that often change, move, are transformed or manipulated. The study of *geometric reconfiguration* analyzes the transformations of such geometric objects under simple rules and constraints. By composing elementary transformations, complex behaviors emerge, which, if understood, can lead to a deep comprehension of those objects, and consequently of the disciplines that study them.

We start with the following example: a *linkage* is a sequence of fixed-length disjoint straight bars (or *edges*) attached together at their endpoints using joints around which the bars can rotate freely. Now, suppose you are given such a linkage in a fairly tangled *configuration*. Can you move the edges continuously, without bending them or altering their lengths, without letting edges cross, and keeping them attached at their joints and reach a straight configuration, i.e., where all edges are on a same line?

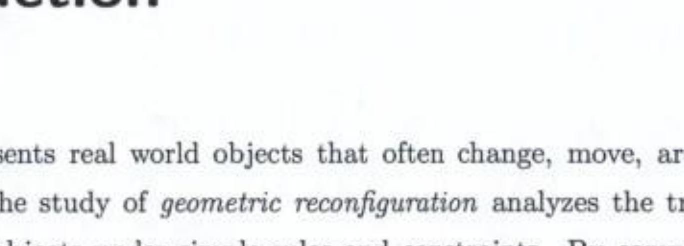


Figure 1.1: A locked universal chain with 5 bars. Figure from [DLOS03].

While for some 3D linkages with only 5 edges, this is known to be impossible (we then say the linkage is *locked*, see Figure 1.1), some restrictions on the starting configuration suffice to ensure straightening is always achievable (e.g. all edges in a plane). Still, elementary variants of this problem remain elusive, for instance, if we restrict our attention to linkages where all edges are of the same length, it is unknown whether there exists a linkage that is locked or if any configuration can be straightened.

Linkages are elementary, and can be used, for example, to study the range of motions of robotic arms. The reconfigurations of equilateral polygons are of intense interest to knot theorists as well. In biology, they are often used to represent the backbone of proteins. One of our hopes is that elementary questions such as the one mentioned above might help in understanding of how proteins fold, which is one of the most important challenges of bioinformatics, or at least to reduce the computational power required to simulate a folding motion.

This can be exemplified by considering the computational counterpart of the above question: given two specific configurations of a chain, can you reconfigure from one to the other? For unconstrained 3D linkages, this problem has recently been shown to be PSPACE-complete, while for some of the more restricted configurations mentioned above, the problem is trivial: the answer is always yes. Yet knowing that the answer is yes is not necessarily the end of the story, and if for some classes of configurations straightenability is guaranteed, it does not necessarily mean that the path to the straight configuration is known or easy to find. These are only a few of the problems we discuss in this thesis.

This work is not intended to be an exhaustive survey of geometric configuration. Several excellent surveys have been published on the subject [DO05, O'R07], including a very recent 450 pages book by Erik D. Demaine and Joseph O'Rourke [DO07]. Instead, we will focus on some key topics where we have made a contribution, or where we feel advances are within reach.

In the next chapter, we begin by discussing the reconfiguration of linkages in the plane and in $\mathbb{R}^3$ under continuous motions. After defining the problem (Section 2.2), we discuss which 2D linkages can lock or not and how the situation changes if we thicken the bars of a linkage or replace them by polygons [CDD$^+$] (Section 2.3). We then turn to linkages in 3D and perform a nearly exhaustive analysis of which combinations of small chains can lock and/or interlock [DLOS03, DLOS02, GLO$^+$04] (Section 2.4). We then focus on the motions of 3D linkages for which the angle between adjacent bars is fixed [ADD$^+$02] (Section 2.5) and devise a model connecting our theories to that of protein folding in molecular biology [DLO06] (Section 2.6)

In Chapter 3, we analyze combinatorial reconfigurations of linkages, i.e. when the number of possible operations that can be applied on a linkage is finite. After describing the background, the *pocket flip* problem by Erdős and the *flipturn* operation (Sections 3.1 and 3.2), we present a more thorough analysis of operations that permute the order of the edges of linkage [ABD+02, ABB+] (Section 3.3).

Finally, in Chapter 4 we discuss several problems related to the unfolding of polyhedra: the unfolding of polyhedral bands [ADL+] (Section 4.1) and of orthogonal polyhedra [DIL] (Section 4.2), and how to wrap smooth convex surfaces with a piece of paper [DDIL07] (Section 4.3).

Most of the results described in this thesis have been published in a series of articles [CDD+, DLOS03, DLOS02, GLO+04, ADD+02, DLO06, ABD+02, ABB+, ADL+, DIL, DDIL07]. A copy of these articles is provided in the appendix.

Chapter 2

Linkages

As mentioned in the introduction, a *linkage* (or *polygonal chain* or just *chain*) is a sequence of fixed-length *bars* (also called *links* or *edges*) connected at their endpoints (or *vertices*) using *joints*.

A *configuration* describes the position of every edge in space such that all edge lengths are respected. The configuration will be called *simple* if no two edges intersect except at their joints. Constraints can be imposed on the joints. For example, if the angle between any two consecutive edges is fixed, the chain is called a *fixed-angle* linkage.

In this chapter, we discuss reconfigurations of linkages under continuous motions.

Contents

2.1 Definitions

The polygonal chain P in $\mathbb{R}^d$ has $n+1$ vertices $V = \langle v_0, \ldots, v_n \rangle$, and is specified by the fixed edge lengths d_i between v_i and v_{i+1}, $i = 0, \ldots, n-1$. We write $P[i,j]$, $i \leq j$, for the polygonal subchain composed of vertices $v_i, \ldots, v_j$. Internal vertices $v_1, \ldots, v_{n-1}$ are called *joints*. A *closed* chain is a chain that connects its two endpoints, i.e., $v_0 = v_n$. Otherwise, the chain is *open*.

A *configuration* $Q = \langle q_0, \ldots, q_n \rangle$ of the chain P (see Fig. 2.1) is an embedding of P into $\mathbb{R}^d$, i.e., a mapping of each vertex v_i to a point $q_i \in \mathbb{R}^d$, satisfying the constraints that the distance between q_i and q_{i+1} is d_i. The points q_i and q_{i+1} are connected by a straight line segment e_i.

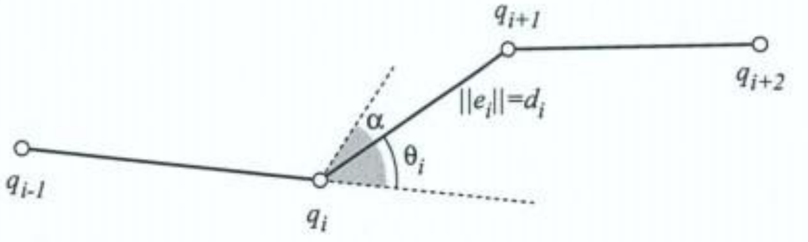


Figure 2.1: Notation for a configuration Q. Figure from [DLO06].

Thus, a configuration in $\mathbb{R}^2$ can be specified by the position of e_0 and clockwise turn angles θ_i at each vertex v_i, $i = 1, \ldots, n-1$. A configuration in $\mathbb{R}^3$ can be specified by the position of e_0 and e_1, turn angles θ_i, and dihedral angles δ_i, $i = 1, \ldots, n-2$, where δ_i is the angle between half-planes $e_i e_{i-1}$ and $e_i e_{i+1}$ whose boundaries contain e_i, and whose interiors contain e_{i-1} and e_{i+1} respectively. The configuration is *simple* if no two nonadjacent segments intersect.

In $\mathbb{R}^3$, we will discuss several restrictions to valid configurations. In a *rigid* chain, values of turn angles θ_i and dihedral angles δ_i are specified along with the edge lengths d_i in the definition of the chain, and a configuration of that chain must satisfy those constraints. The position of a rigid chain can thus be fully specified by the position of edges e_0 and e_1. In a *fixed-angle* chain (also sometimes called *dihedral* or *revolute*), only values of the turn angles θ_i are specified along with the edge lengths d_i in the definition of the chain. A configuration can thus be described by the position of e_0 and e_1 and the dihedral angles δ_i. A fixed-angle chain where all angles $\theta_i \leq \alpha$ for some $0 < \alpha < \pi$, P is called a $(\leq \alpha)$-*chain*.[1]

The most general type of chain, in which only edge lengths are specified, is called *flexible*.

A *motion* $M = \langle m_0, \dots, m_n \rangle$ of a chain P is a list of $n+1$ continuous functions $m_i : [0, \infty] \to \mathbb{R}^d$, $i = 0, \dots, n$, such that $M(t) = \langle m_0(t), \dots, m_n(t) \rangle$ is a configuration of P for all $t \in [0, \infty]$. The motion is said to be *simple* if all such configurations $M(t)$ are simple. We often assume that the motion is *finite* in the sense that, after some time T, M becomes independent of t.

The *configuration space* of a chain P is the set of simple configurations of that chain. Then a simple motion can be seen as the parametric equation of a connected curve in the configuration space.

2.2 Reachability

The main problem that will be studied throughout this chapter is that of *reachability*:

Problem 2.1 *Given two configurations Q_1 and Q_2 of a same chain P, is there a (simple) motion M that transforms Q_1 into Q_2, i.e. $M(0) = Q_1$ and $M(\infty) = Q_2$?*

Note that this is equivalent to asking if Q_1 and Q_2 are in the same component of the configuration space of P. This problem has many facets each of which will be considered in turn. Already, as we saw above, the definition of chains and configurations come in many flavors: dimension of the space, type of joints etc., and more refinements of those will be defined along the way. Each of these variants will motivate an equal number of reachability questions to solve.

Although reachability is stated as a decision problem, of which the complexity will have to be analyzed, it is natural to first ask whether the answer to Problem 2.1 is always "Yes", i.e., whether the configuration space is always connected. In such situations, we would talk about a *universality* result. The decision problem in that case is computationally trivial, but this does not necessarily imply that it is easy to find the transforming motion M.

The complexity of deciding reachability is only one of the computational questions that can be asked. Another important one is that of the complexity of the simplest motion for transforming a chain into another, as well as the complexity of computing

[1]Some work [ADD⁺02, ADM⁺02] focuses on the angle between adjacent edges, which for us is $\pi - \alpha$. Thus "nonacute chains" in that work corresponds to $(\leq \pi/2)$-chains here.

such a simple motion. This will require defining the complexity of a motion, and there several models have been considered.

One standard technique for proving a universality result is to define a *canonical configuration* for a chain. For example in the case of open universal chains, the canonical configuration will often be chosen to be the *straight configuration*, in which all turn angles θ_i have value 0. A configuration or a chain that can always reach the straight configuration will be qualified of *straightenable*. Then, if one can show that both Q_1 and Q_2 are straightenable, then they can reach each other by combining the straightening motion from Q_1 and the reverse of the straightening motion from Q_2. More generally, if both chains can reach some canonical configuration, then they can reach each other.

However, it is often the case that a single canonical configuration is difficult to define, and it is instead more natural to consider a class of canonical configurations. For example, for closed universal chains, we will define *convex* configurations, and a chain that can reach some convex configuration will be qualified of *convexifiable*. Likewise, for fixed-angle chains in $\mathbb{R}^3$, we define a *flat* configuration to be a simple configuration such that all edges and vertices lie in a common plane. A chain or a configuration that can reach some flat configuration will be called *flattenable*. Then, in order to follow the same reasoning as before, it will be necessary to show that any canonical configuration in the class can reach any other. Thus this motivates the study of reachability and the search for universality results in restricted classes of configurations (such as convex or flat configurations).

While many classes of chains and configurations exhibit universality, many others do not. In order to show this, it will suffice to describe two configurations Q_1 and Q_2 of a same chain P and prove that no motion exists that can transform one into the other. When Q_2 is a canonical configuration (e.g. Q_1 is not straightenable), then we say that Q_1 is *locked*. The existence of locked configurations is fundamental in many respects. They not only to demonstrate the absence of a universality result, but they have also helped in gadgets that were used in proving the intractability of some versions of the reachability problem. Furthermore, as we show in [DLO06], if small locked chains exist, then the class of configurations that can reach the canonical configuration is a vanishingly small portion of the entire configuration space as n, the number of links, grows.

Note that claiming the existence of a locked configuration is equivalent to stating that the configuration space is not connected. Thus, the validity of the claim is independent of the choice of the canonical configuration(s). Because of this, we will allow ourselves to claim the existence of locked configurations even when the canonical configuration has not been explicitly defined.

We also study the reconfiguration of several of chains. The same reachability questions can then be asked, and be enriched with the concept of separability. A configuration of two or more chains is said to be *separable* if there is a simple motion that moves them arbitrarily far apart. If no such motion exist, the configuration is said to be *interlocked.* Notice that the existence of an interlocked configuration implies that the configuration space is not connected, while the converse is not necessarily true.

While seemingly simple, the question[2] of whether a chain can be locked in $\mathbb{R}^2$ has challenged researchers for over 25 years until it was finally answered negatively in 2000 by Connelly, Demaine, and Rote [CDR03]. The proof relies on techniques borrowed from rigidity theory and convex programming, as well as the concept of *expansive motions* (motions where every pair of points on the chain moves apart). Since then, two further algorithms for straightening open chains and convexifying closed chains were presented, one using the theory of frameworks and pseudo-triangulations [Str00] and the other by minimizing some energy function [CDIO04]. A complete history of this important discovery can be found in [DO05, DO07, Sos01].

In $\mathbb{R}^3$, it is known that a single open chain having as few as 5 bars can be locked, see Figure 1.1. Two proofs of this exist, the first proof for this by Canterella and Johnston [CJ98] and uses a precise case analysis. A somewhat simpler proof using elements of knot theory was presented in [BDD$^+$99]. Moreover, it was recently shown that deciding reachability in $\mathbb{R}^3$ is PSPACE-hard [AKRW04].

Surprisingly, $\mathbb{R}^3$ is the only dimension where the configuration space of universal chains is not always connected, as it was shown that chains and cycles in $\mathbb{R}^d$ for $d \geq 4$ can always be straightened [CO01].

2.3 2D Chains of Polygons

Of course, the models described in the previous sections are purely theoretical, if only because the edges are assumed to be infinitely thin. In a recent paper [CDD$^+$], we thicken the edges of chains in $\mathbb{R}^2$ by gluing shapes (*adornments*) on them, which are required not to collide. In that case, we show that it might happen that the chains become locked. However, we manage to characterize quite precisely the adornments for

[2] Posed independently by Stephen Schanuel and George Bergman in the early 1970's, Ulf Grenander in 1987, William Lenhart and Sue Whitesides in 1991, and Joseph Mitchell in 1992 [CDR03]

which the configurations space is always connected. We show that if the adornments on an open chain are *slender* (roughly, if the inward normal from any boundary point intersects the edge), then expansive motions will be simple and thus will straighten the chain without collisions. The same, however, is not true for closed chains. We also show that our characterization is tight, in the sense that isosceles triangles with any desired apex angle $< 90°$ admit locked chains, which is precisely the threshold beyond which the slender property no longer holds.

Our work has already found applications in the analysis of hinged collections of polygons. For instance, a *hinged dissection* is a chain or a tree of polygons that can be reconfigured into two or more self-touching configurations with desired silhouettes. For example, Figure 3 in [CDD+] shows a classic hinged dissection from 1902 for transforming a triangle into a square. While general families of dissections have been extensively analyzed[AN98, DDE+05, DDLS05, Epp01, Fre02], nearly none of these studies consider the problem of actually moving from one configuration to the other without collisions. Our results provide potential tools for resolving these problems, and already solve an open problem from [DDE+05].

2.4 Small (inter)locked linkages

In this section, we come back to the reachability problem in 3D. In particular we try to determine the number of bars needed for one or more linkages to lock and/or interlock.

Problem 2.2 *What 3D chains can lock/interlock?*

As we saw in Section 2.2, Canterella and Johnston [CJ98] have proved that there is a universal open chain with 5 bars that can lock. They also showed that no chain with 4 bars or less can lock, that a closed hexagon can lock but no closed pentagon. It is also known that fixed-angle open chains with 4 bars can lock, but not 3. The configuration space of closed fixed angle chains with 4 bars or more can have more than one component, but the configuration space of a triangle is trivially connected. Finally, the configuration space of a rigid chain is always connected. These results are summarized in Table 2.1

We now turn to determining which sets of linkages can interlock. This work was motivated by an open problem by a question posed by Anna Lubiw [DO00]: Into how many pieces must a chain be cut so that the pieces can be separated and straightened? This problem is motivated by protein molecules, which can be modeled by polygonal

	Open	Closed
Universal	5	6
Fixed-angle	4	4
Rigid	∞	∞

Table 2.1: Minimum number of edges needed for a 3D chain to have a non-connected configuration space

chains, and, according to some theories, temporarily split apart in order to reach the minimum-energy folding.

First note that configuration spaces involving two closed chains can never been connected, as it always possible to construct a configuration topologically equivalent to the link 0_1^2 composed of two disjoint unknots, and another one for the link 2_1^2 composed of two unknots, one going through the other, see Figure 2.2[3]. Since those are not topologically equivalent, it is impossible to move from one to the other without causing the configuration to be non simple at some time.

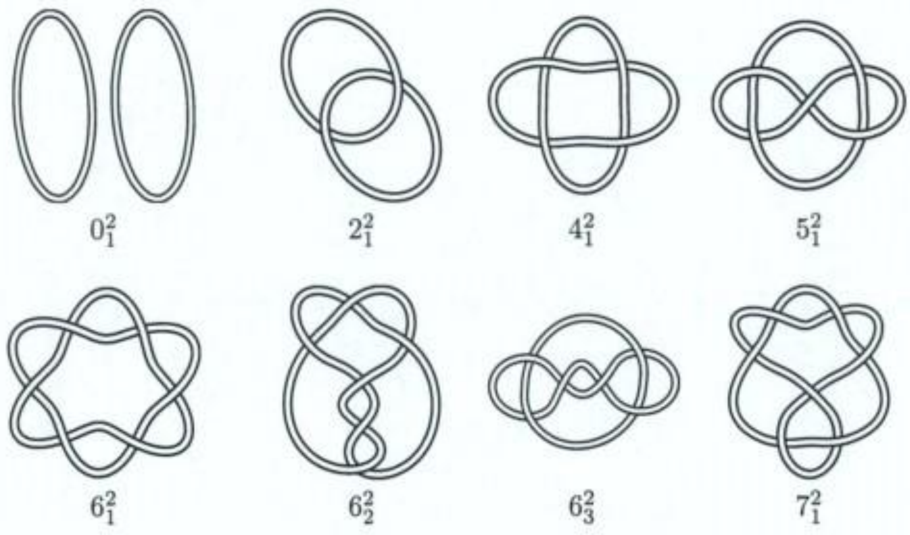


Figure 2.2: The first few two-component links.

In [DLOS03], we study the small interlocked configurations composed of one closed and one open chain. The results are summarized Table 2.2. The proofs are similar in spirit to the argument above, with a twist: if we construct an open chain that have long

		Closed		
		3	4	5
	2	-	-	-
Open	3	-	+	+
	4	+	+	+

Table 2.2: Open chains and closed chains that can interlock (+) or not (-), depending on the number of edges. A claim that a k-chain can interlock holds also for any l-chain with $l > k$, and a claim that a k-chain cannot interlock holds also for any l-chain with $l < k$.

end bars, then we can simulate a two component link by supposing that the endpoints of one chain are tied together with a rope.

To prove the above results, we (conceptually) close an open chain by adding a piece of rope, then argue that geometric properties keeps the rope from interfering with any motion, and that topological invariants demonstrated that the resulting closed links are interlocked. However, this approach does not extend to proving interlocking between two open chains: we cannot simply close two or more open chains with ropes because the ropes may interfere with one another. Instead we establish geometric invariants, typically about the convex hull of joints and the relations of the end bars, often by considering convenient projections of the linkage [DLOS02]. In [GLO$^+$04] it is shown that a universal 2-chain can interlock with a universal open 19-chain. The 19-chain was then subsequently reduced to an 11-chain [GLOZ06]. Finding the smallest chain that can interlock with a 2-chain remains an open problem. The results are shown Table 2.3

For sets of 3 chains, we show [DLOS02] that three open universal chains with 3 bars each can interlock, but two 3-chains along with any number of 2-chains cannot. Thus we almost completely characterize the size of combinations of chains that can interlock.

Important special cases remains open as, for example, we do not know whether a chain can lock if all its edges are of unit length. For fixed-angle chains, we do not know whether such a unit chain can lock as soon as the fixed angles are larger than 60°. This problem is of particular interest since proteins have all edges nearly equal, and bond angles around 110°.

[3]link images produced by Robert Scharein's `knotplot` program `http://www.cs.ubc.ca/nest/imager/contributions/scharein/KnotPlot.html`.

		2-chain		3-chain			4-chain			5-chain
		flexible	rigid	flexible	fixed-angle	rigid	flexible	fixed-angle	rigid	rigid
2-chain	flexible	$-$	$-$	$-$	$-$	$-^{5}$	$-^{2}$	$-^{5}$	$-^{5}$	$+^{15}$
	rigid	$-$	$-$	$-$	$-^{4}$	$+^{12}$	$+^{14}$	$+^{14}$	$+^{14}$	$+$
3-chain	flexible	$-$	$-$	$-^{1}$	$-^{6}$	$+^{18}$	$+^{11}$	$+$	$+$	$+$
	fixed-angle	$-$	$-^{4}$	$-^{6}$	$+^{21}$	$+$	$+$	$+$	$+$	$+$
	rigid	$-^{5}$	$+^{12}$	$+^{18}$	$+$	$+$	$+$	$+$	$+$	$+$
4-chain	flexible	$-^{2}$	$+^{14}$	$+^{11}$	$+$	$+$	$+$	$+$	$+$	$+$
	fixed-angle	$-^{5}$	$+^{14}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
	rigid	$-^{5}$	$+^{14}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
5-chain	rigid	$+^{15}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
11-chain	flexible	$+^{*}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$

Table 2.3: Results on interlocking pairs of open chains. $(+)$ = can, $(-)$ = cannot interlock. In superscript is the number of the theorem from [DLOS02] proving the result, or * for [GLOZ06], the other entries are implied.

2.5 Flat to flat

We highlighted earlier the importance of defining a natural set of canonical configurations in order to solve some instances of the reachability problem, and while straight configurations seem like an obvious option for universal chains, the choice is more difficult to make when it comes to fixed-angle chains. One might be tempted to choose *flat* configurations, i.e. configurations where all edges lie in a common plane. Unfortunately, a flat configuration for a given chain might not be unique, and even worse, it might not exist.

Nevertheless, many classes of chains (e.g. if all edges have equal length, or if all turning angles are $\leq 90°$) do have a flat state. In that case, is there a simple motion from any flat configuration to any other flat configuration? If yes, we say that the chain is *flat-state connected*.

Problem 2.3 *Is every open fixed-angle chain flat-state connected?*

Although the general problem is still unsolved, a series of papers partially answers this question for some important classes of chains [ADM$^+$02, ADD$^+$02, AM06]. The currently best known results are shown Table 2.4

Notice that linkages are not the only structures to be represented in the table. In an attempt to understand the nature of the problem, we have analyzed the flat-state

Constraints on Fixed-Angle Linkage				Flat-state
Connectivity	Angles	Lengths	Motions	connectivity
Open chain	—	—	—	?
	has a monotone state		—	Connected [AM06]
	nonacute	—	—	Connected [ADD+02]
	equal acute	—	—	Connected [ADM+02]
	each in $(\delta, 2\delta)$ for $\delta \leq \pi/3$		—	Connected [AM06]
	each in $(60°, 90°]$	unit	—	Connected [ADM+02]
	—	—	180° edge spins	Disconnected [ADD+02]
	orthogonal	—	180° edge spins	Connected [ADD+02]
Set of chains, each	orthogonal	—	—	Connected [ADD+02]
pinned at one end	orthogonal	—	partially rigid	Disconnected [ADD+02]
Tree	orthogonal	—	partially rigid	Disconnected [ADD+02]
Graph	orthogonal	—	—	Disconnected [ADD+02]

Table 2.4: Known results. The '—' means no restriction of the type indicated in the column heading. Entries marked '?' are open problems

connectivity of more complex structures and restrictions on the motions. In particular, we study sets of chains with one endpoint stuck on the xy plane, partially rigid trees, and graphs. A detailed description of these structures and of the results can be found in [ADD+02].

2.6 Protein machine

Fixed-angle chains are of particular interest, partly because they can be used to represent the backbone of protein chains. In this part of our work, we ask the following, intentionally vague, question:

Problem 2.4 *What classes of configurations correspond to folded proteins?*

As a first step towards this goal, we study the impact of the way a protein is produced on the structure of the folded backbone [DLO06]. Our inspiration derives from the ribosome, which is the "machine" that creates protein chains in biological cells. Figure 2.3 shows a schematic cross section of a ribosome and its exit tunnel, based on a model developed by Nissen et al. [NHB+00].

We consider a very simple geometric model that roughly captures the exit point x of the ribosome: the chain is produced inside a cone of some half-angle β, with the chain emerging through the cone's apex. All through the production process, we require that the chain being produced remains simple at all times and the portion already out of the

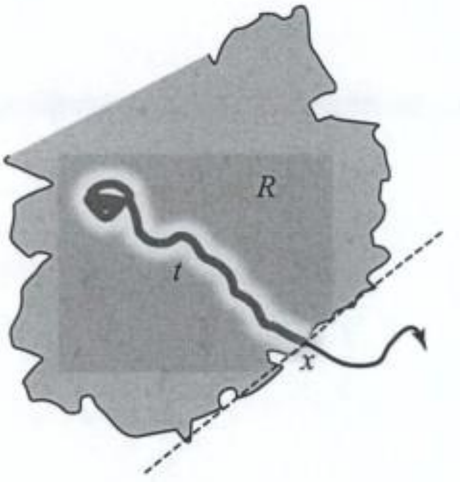


Figure 2.3: The ribosome R in cross-section. The protein is created in tunnel t and emerges at x. Figure from [DLO06].

cone never intersects the cone. This constraint immediately implies that the maximum turn angle α in the produced chain is at most 2β.

Our main result is that a chain can be produced in our model if and only if it can be flattened. Furthermore, since we use a unique canonical configuration to describe the motion, this provides an alternate proof of the flat-state connectivity of those chains.

We also show that, if some small locked chain exists, the the portion of the configuration space that is producible (and thus flattenable) is vanishingly small as the number of edges grows.

Interestingly, the canonical configuration we use has an helical form that is in many ways similar to the shape a protein takes as it naturally comes out of the ribosome.

Chapter 3

Combinatorial moves

As the full range of motions for linkages is often quite complex, many authors have considered restricting the reconfigurations to be from a finite set of *moves*. One of the earliest occurrences of such a problem was posed by Paul Erdős in 1935 [Erd35], and has spawned a large number of studies and variants.

Contents

3.1 Erdős flips

In 1935, Paul Erdős posed the following problem [Erd35]: given a simple closed polygon P with n edges, let $CH(P)$ be the convex hull of P. If we remove P from $CH(P)$, we are left with *pockets*, maximal connected regions that are in $CH(P)$ but outside of P. The boundary of the pocket is composed of one convex hull edge called the *lid*, and a contiguous subsequence of the edges of P. A *pocket flip* or just *flip*, transforms the polygon by reflecting the pocket's subsequence of edges across the lid. It can also be seen as a $180°$ rotation of the pocket's edges about the supporting line of the lid in 3D. Every non-convex polygon has at least one pocket. Can every polygon be convexified by a finite number of simultaneous pocket flips?

A few years later, Nagy [Nag39] noted that performing simultaneous flips might result in a self-intersecting polygon. He then proposed a proof that any polygon can become convex after a finite number of single pocket flips. Several further proofs were published by different authors since then. Unfortunately most of these proofs, including the first one by Nagy, contain flaws. Those were noticed and corrected in a recent paper by Demaine, Gassend, O'Rourke and Toussaint [DGOT06]. It was also proved by Joss and Shannon in 1973 that it is not possible to bound the number of flips needed to convexify a polygon as a function of n [Grü95], in fact it is possible to construct a quadrilateral requiring arbitrarily many flips to be convexified.

Wegner [Weg93] defined the *deflation* as the inverse operation from a flip, and conjectured that any polygon can be deflated at most a finite number of times (while remaining simple) until no further deflation is possible. This conjecture was disproved by Fevens et al. [FHM$^+$01] who presented a family of quadrilaterals that can be deflated infinitely many times.

Combining flips and deflations, one could again pose the reachability problem: given two configurations of a same chain P in $\mathbb{R}^2$, is there a sequence of flips and deflation operations to go from one to the other. We can rapidly obtain a negative answer by a counting argument: the number of configurations reachable from a starting configuration by flips and deflations is countable, while the set of possible configurations is not. Still it would be interesting to know when the combination of flips and deflations are as universal as combinatorial moves could get, i.e., are there polygons and configurations that can reach any neighborhood of the configuration space via a sequence of flips and deflations?

3.2 Flipturns

Flipturns, defined by Joss and Shannon, are a simple variation on pocket flips where instead of reflecting the pocket through the lid, the pocket is rotated 180° about the center of the lid. This has for effect of reversing the order of the pocket's edges while maintaining their lengths and orientations. Thus, because a flipturn only produces a permutation of the same edges, they show that at most $(n-1)!-1$ flipturns can be made before the polygon becomes convex. This bound was improved by Ahn et al to $n(n-3)/2$ or $n(s-1)/2-s$ when the polygon edges have only s different slopes [ABC$^+$00]. This result uses a modified definition of flipturns to handle degenerate cases more easily. Aichholzer et al. [ACD$^+$02]. showed that a bound of n^2-4n+1 holds for the original definition of flipturns, and showed that computing the longest sequence of flipturns is NP-hard. Whether finding the shortest convexifying sequence of flipturns can be done in polynomial time remains an open problem. They also prove that the convex polygon obtained after any maximal sequence of flipturns is always the same and can be computed in $O(n \log n)$ time by sorting the edges by slope, and provide data structures for maintain the polygon and its convex hull after flipturn operations, in $O(\log^4 n)$ amortized time per flipturn.

In [Bie06], Biedl presented a polygon such that the shortest convexifying flipturn sequence is of length $(n-1)^2/8$ and another one whose longest sequence has $(n-2)^2/4$ flipturns.

3.3 Permutations

In [ABD$^+$02], we analyze several further operations whose effect is to produce a different permutation of the same edges, preserving edge lengths and orientations. In general, we will view each edge as a rooted vector, and the polygon is a sequence of these vectors. When talking about a simple polygon, we often require the polygon to be clockwise (counterclockwise), i.e., the interior of the polygon is on the right (left) side of every vector.

A *reversal* or *generalized flipturn* [ACD$^+$02] reverses the order of some subsequence of edges, and is thus a generalization of the flipturn operation. A *transposition* swaps two subsequences of edges. A *single edge transposition* moves an edge to another position in the sequence. It can be seen as a transposition between a single edge and an empty subsequence at some arbitrary position. An *edge swap* of *popturn* reverses the order of two adjacent edge. It is thus a special case of a reversal and at the same time a single edge transposition. All those operations are called *crossing-free* if the resulting polygon is simple, and *simple* if, starting with a simple clockwise polygon, the resulting polygon is simple and the clockwise.

Interestingly, questions about those operators without requiring simplicity requires very little geometry. It is in fact equivalent to the problem of sorting a circular permutation of the ordered set of edge orientations using those operators. These problems have received considerable attention in the literature. Table 3.1 summarizes known results. The remainder of this discussion focuses on simple and crossing-free operations.

		Convexifiable?	min # moves to convex	finding min # moves
Non simple	popturn	always	$\Theta(n^2)$ [Bal03]	?
	single-edge transp.	always	$\Theta(n)$	?
	transposition	always	$\Theta(n)$	?
	flipturn	always	$O(n!)$ [GZ98]	?
	reversal	always	$\Theta(n)$	NP-hard [SSL03]
Crossing free	popturn	NP-hard [ABB+]	—	—
	single-edge transp.	not always [ABD+02]	—	—
	transposition	?	?	?
	flipturn	always[Grü95]	$\Theta(n^2)$ [ABC+00]	? [ACD+02]
	reversal	always	$O(n^2)$	NP-hard [ACD+02]
Simple	popturn	iff no *purse*, $O(n)$ [ABB+]	$\Theta(n^2)$ [ABD+02] [ABB+]	$O(n \log n)$ [ABB+]

Table 3.1: Known results for combinatorial moves

Because reversals are a generalization of flipturns, we can easily conclude that any polygon can be convexified after $O(n^2)$ reversals. Aichholzer et al. [ACD+02] proved that computing the shortest convexifying sequence of reversals for a given polygon is NP-hard. In [ABD+02], we show polygons that cannot be convexified using popturns or single-edge transpositions without introducing crossings. Recently [ABB+], we refined the analysis of popturns and obtained several new results. Interestingly, the distinction between crossing-free and simple popturns influences dramatically the complexity of determining if the polygon is convexifyable. In the same paper, we extend further those operations for the case of unit orthogonal polygons and show that permitting 180° *rotations* or *untwist* operations suffice to make any such polygon convexifiable. It would be interesting to see whether those extensions could be generalized to general polygons.

In [ABD$^+$02], we study several further computational problems on permutations of the edges of a polygon, such as finding the permutation of smallest or largest diameter.

Chapter 4

Polyhedra and surfaces

One of the most classical reconfiguration problems is that of polyhedron *edge unfolding*: Given a convex polyhedron, can we cut along some of its edges in such a way that its surface can be unfolded into a single planar polygon without overlap? This problem implicitly goes back to 1525 when Albrecht Dürer published edge unfoldings of convex polyhedra, and is to this day unsolved.

Problem 4.1 *Does every convex polyhedron have an edge unfolding?*

If we are allowed to cut the polyhedron through its faces, then we talk about a *general unfolding*. Two methods exist that will produce a general unfoldings for any convex polyhedron: source unfoldings and star unfoldings. On the other hand, it is unknown whether any (non-convex) polyhedron has a general unfolding.

Problem 4.2 *Does every polyhedron (not necessarily convex) have a general unfolding?*

Although those two important open problems have remained open for quite some time, progress has been made at steady pace over the past few years, every time increasing our understanding of the problems. We here illustrate two approaches that have lead to interesting results and end the chapter by discussing the wrapping of smooth surfaces.

Again, we refer the reader to the recent book of Demaine and O'Rourke [DO07] for an extensive survey of the area of polyhedra unfolding.

Contents

4.1 Bands

One approach towards understanding edge unfoldings is to study polyhedra that have a special structure. Another one is to loosen the requirements of the resulting unfolding. Here, rather than trying to unfold the entire polyhedron, we attempt to edge-unfold a large portion of a convex polyhedron.

In 1998, E. Demaine, M. Demaine, A. Lubiw, and J. O'Rourke posed the following problem: A *polyhedral band* is the surface of a convex polyhedron enclosed between parallel planes, and containing no polyhedron vertices. Can every polyhedral band be unfolded without overlap by cutting an appropriate single edge? This case is in some sense the simplest sort of edge unfolding task one could imagine since only one edge has to be cut. At the same time, this question can be seen as a way to exploit the knowledge we have of polygonal chains by thickening them into bands.

We first studied the *nested* case where, in some projection, one border of the band is in the interior of the other. For that case, we show [ADL$^+$] that it is always possible to cut an edge to unfold the band without overlap. Furthermore, we show two different continuous unfolding motions from the original band to its unfolded configuration. The result was further improved by Aloupis [Alo05] to include non-nested bands and bands whose border contains vertices.

4.2 Orthogonal polyhedra

Here again, in order to facilitate the study of Problems 4.1 and 4.2, we will restrict our attention to a special class of polyhedra: the class of *orthogonal polyhedra*, in which every face is orthogonal to the x, y or z axis.

One of the earliest results concerning orthogonal polyhedra came from Biedl et al. [BDD$^+$98] where two subclasses of polyhedra are analyzed: *orthostacks* and *orthotubes*. An orthostack is an orthogonal polyhedron of which every horizontal planar slice not including a horizontal face is a single simple (orthogonal) polygon. An orthotube is a (possibly cyclic) sequence of rectangular boxes glued together according to some rules.

They show the existence of orthostacks that cannot be edge-unfolded, but show that both orthostacks and orthotubes can be generally unfolded. In their unfoldings, they use cuts along planes orthogonal to the coordinate axes. In the case of orthotubes, all such planes go through at least one vertex of the polyhedron. Such an unfolding is called *grid-edge* unfolding. For orthostacks however, they need to use planes that do not go through any vertex and ask whether every orthostack can be grid-edge unfolded. In subsequent articles, the question was generalized to the following:

Problem 4.3 *Can every orthogonal polyhedron be grid-edge unfolded?*

A normal edge unfolding cuts the polyhedron along edges in such a way that the resulting cut surface is edge-connected and can be embedded onto a plane without overlap. In a *vertex unfolding*, the polyhedron is again cut along edges but the resulting surface only needs to be connected through vertices. Vertex unfoldings were introduced in [DEE$^+$02, DEE$^+$03].

In [DIL], we prove that every orthostack can be *grid-vertex* unfolded, i.e., cut along axis-parallel planes incident to vertices of the polyhedron such that the resulting surface is vertex connected and can be embedded onto a plane without overlap.

This result was subsequently generalized in several ways: Damian, Flatland and O'Rourke [DFO06] show that every orthogonal polyhedron can be vertex-unfolded, and in another paper [DFO07] that if we are allowed to cut along a refined grid subdividing each slab of the grid through every vertex into $2^{O(n)}$ pieces, then any orthogonal polyhedron after that refinement can be grid-edge unfolded, thus providing the first general unfolding of any orthogonal polyhedron. Nevertheless, Problem 4.3 remains open to this day.

A recent survey of unfolding results for orthogonal polyhedra was recently published by O'Rourke [O'R07].

4.3 Wrapping convex smooth surfaces

We end this chapter with a preliminary result, and an open-ended discussion, on how to wrap smooth surfaces.

All the unfolding works described above as well as in mathematical origami describe a mapping between a flat piece of paper and a surface. The mapping described is isometric in the sense that pairwise distances (defined as shortest path on the piece of paper) are preserved. A problem naturally arises if we consider convex smooth surfaces, e.g., that have positive curvature everywhere.

Thus, we must find a way to allow changing the curvature while preventing from stretching the piece of paper. For this, we define a concept of *wrapping* as a contractive mapping, i.e. where every pairwise distance on the piece of paper either stays the same or shrinks.

In [BCO04], Benbernou, Cahn, and O'Rourke defined a similar notion of unfolding, but restricted to certain piecewise ruled surfaces and they analyzed the *volcano unfolding* of smooth prizmatoids. In such unfoldings, the paper is wrapped from each point from the base along a straight geodesic line on the surface.

In [DDIL07], we formalize the concept of wrapping for smooth convex surfaces and focus on *stretched wrappings* that can be specified by a surface-covering tree of stretched paths, i.e. paths along which the wrapping is isometric. We then analyze the sizes and shapes of a piece of paper that can cover a sphere of unit radius. For instance, we show that a sphere can be covered by a disk of radius π (surface π^3), a square of diameter 2π (surface $2\pi^2$) and a rectangle $2\pi x \pi$ (surface $2\pi^2$). Surprisingly, we then show that it is possible to cover the sphere with an equilateral triangle of area $1.998626\pi^2$, a 0.1% improvement. We then discussed how to tile the plane with shapes that wrap the sphere, while keeping the shapes fat. This could be useful, e.g., if the wrapping paper is to be produced in large quantities.

Chapter 5

Conclusion

In this thesis, we have presented a variety of problems and results on the geometric reconfiguration of linkages and polyhedra. Although we could only touch on a few topics, we hope to have illustrated some of the richness and diversity of the subject.

The reader probably noticed the many doors that every new result opens, as is clearly illustrated by the several question marks left in most table of results. Beyond these obvious holes to fill in and the main questions discussed throughout the text, several problems deserve discussion and further work. I will highlight several of them and indicate some possible directions to attack them.

Problem 5.1 *Can a unit linkage lock?*

This problem has puzzled researchers for many years, still several variants are within reach: what locked linkage has the minimum ratio between its longest and shortest edge lengths? Does that ratio depend on n, the number of bars in the linkage? How about fixed angle chains? Is there a tradeoff between the fixed angle, the number of bars and the ratio?

As discussed earlier, the existence of small chains has strong implications on other results, such as the protein machine result, or the existence of hardness proofs.

Problem 5.2 *What chains correspond to proteins?*

Biologists constantly make observations about the geometric structure of proteins. For example, that the endpoints of a protein always remain close to its outer layer[1]. If we could formalize those observations, it could be possible to reduce the search space of protein folding simulations, and increase the efficiency of the simulation.

[1]R. Brasseur, personal communication

Problem 5.3 *Which results on combinatorial reconfigurations of chains can be generalized to 3D?*

Grünbaum and Zaks [GZ98] have asked whether it is always possible to convexify a polygon in 3D using 3D flips (for some definition of a 3D flip), however, it seems like some other operations, such as the flipturn, can generalize more easily and more naturally.

Problem 5.4 *Can the band unfoldings be combined to generate a new general unfolding algorithm?*

Or another interesting question is to characterize which bands can be thickened so as to contain the entire faces it touches, while still being edge-unfoldable.

Problem 5.5 *Can the unfolding algorithms for orthogonal polyhedra be generalized to c-oriented polyhedra?*

Maybe this way it would be possible to leverage the general unfolding algorithm for orthogonal polyhedra.

Bibliography

[ABB+] G. Aloupis, B. Ballinger, P. Bose, M. Damian, E.D. Demaine, M.L. Demaine, R. Flatland, F. Hurtado, S. Langerman, J. O'Rourke, P. Taslakian, and G. Toussaint. Vertex pops and popturns. In *Canadian Conference on Computational Geometry (CCCG 2007)*. Submitted.

[ABC+00] Hee-Kap Ahn, Prosenjit Bose, Jurek Czyzowicz, Nicolas Hanusse, Evangelos Kranakis, and Pat Morin. Flipping your lid. *Geombinatorics*, 10(2):57–63, 2000.

[ABD+02] G. Aloupis, P. Bose, E.D. Demaine, S. Langerman, H. Meijer, M. Overmars, and G.T. Toussaint. Computing signed permutations of polygons. In *Proceedings of the 14th Canadian Conference on Computational Geometry (CCCG 2002)*, 2002.

[ACD+02] Oswin Aichholzer, Carmen Cortés, Erik D. Demaine, Vida Dujmović, Jeff Erickson, Henk Meijer, Mark Overmars, Belén Palop, Suneeta Ramaswami, and Godfried T. Toussaint. Flipturning polygons. *Discrete & Computational Geometry*, 28(2):231–253, August 2002.

[ADD+02] G. Aloupis, E.D. Demaine, V. Dujmovic, J. Erickson, S. Langerman, H. Meijer, J. O'Rourke, M. Overmars, M. Soss, I. Streinu, and G.T. Toussaint. Flat-state connectivity of linkages under dihedral motions. In *Proceedings of the 13th Annual International Symposium on Algorithms and Computation (ISAAC 2002)*, volume 2518 of *LNCS*, pages 369–380. Springer-Verlag, 2002.

[ADL+] G. Aloupis, E.D. Demaine, S. Langerman, P. Morin, J. O'Rourke, I. Streinu, and G.T. Toussaint. Unfolding polyhedral bands. *Computational Geometry: Theory and Applications*. To appear. Special issue of selected papers from the 16th Canadian Conference on Computational Geometry (CCCG'04). Preliminary version: [ADL+04].

[ADL+04] G. Aloupis, E.D. Demaine, S. Langerman, P. Morin, J. O'Rourke, I. Streinu, and G.T. Toussaint. Unfolding polyhedral bands. In *Proceedings of the 16th Canadian Conference on Computational Geometry (CCCG 2004)*, 2004.

[ADM+02] Greg Aloupis, Erik D. Demaine, Henk Meijer, Joseph O'Rourke, Ileana Streinu, and Godfried Toussaint. Flat-state connectedness of fixed-angle chains: Special acute chains. In *Proceedings of the 14th Canadian Conference on Computational Geometry*, pages 27–30, Lethbridge, Canada, August 2002.

[AKRW04] Helmut Alt, Christian Knauer, Günter Rote, and Sue Whitesides. On the complexity of the linkage reconfiguration problem. In J. Pach, editor, *Towards a Theory of Geometric Graphs*, volume 342 of *Contemporary Mathematics*, pages 1–14. American Mathematical Society, 2004.

[Alo05] G. Aloupis. *Reconfigurations of Polygonal Structures*. PhD thesis, McGill University, 2005.

[AM06] G. Aloupis and H. Meijer. Reconfiguring planar dihedral chains. In *22nd European Workshop on Computational Geometry (EuroCG 2006)*, 2006.

[AN98] Jin Akiyama and Gisaku Nakamura. Dudeney dissection of polygons. In *Revised Papers from the Japan Conference on Discrete and Computational Geometry*, volume 1763 of *Lecture Notes in Computer Science*, pages 14–29, Tokyo, Japan, December 1998.

[Bal03] B. Ballinger. *Length-Preserving Transformations on Polygons*. PhD thesis, Univ. Calif. Davis, 2003.

[BCO04] Nadia Benbernou, Patricia Cahn, and Joseph O'Rourke. Unfolding smooth prismatoids. In *14th Annual Fall Workshop on Computational Geometry*, Cambridge, MA, 2004. arXiv:cs.CG/0407063.

[BDD+98] Therese Biedl, Erik Demaine, Martin Demaine, Anna Lubiw, Mark Overmars, Joseph O'Rourke, Steve Robbins, and Sue Whitesides. Unfolding some classes of orthogonal polyhedra. In *Proceedings of the 10th Canadian Conference on Computational Geometry*, Montréal, Canada, August 1998.

[BDD+99] T. Biedl, E. Demaine, M. Demaine, S. Lazard, A. Lubiw, J. O'Rourke, M. Overmars, S. Robbins, I. Streinu, G. Toussaint, and S. Whitesides. Locked and unlocked polygonal chains in 3D. In *Proc. 10th ACM-SIAM Sympos. Discrete Algorithms*, pages 866–867, January 1999. Full version LANL arXive cs.CG/9910009.

[Bie06] T. Biedl. Polygons needing many flipturns. *Discrete and Computational Geometry*, 35(1):131–141, 2006.

[CDD+] R. Connelly, E.D. Demaine, M.L. Demaine, S.P. Fekete, S. Langerman, J.S.B. Mitchell, A. Ribó, and G. Rote. Locked and unlocked chains of planar shapes. *Discrete & Computational Geometry*. Submitted. Preliminary version: [CDD+06].

[CDD+06] R. Connelly, E.D. Demaine, M.L. Demaine, S.P. Fekete, S. Langerman, J.S.B. Mitchell, A. Ribó, and G. Rote. Locked and unlocked chains of planar shapes. In *Proceedings of the 2006 ACM Symposium on Computational Geometry (SoCG 2006)*, pages 61–70, 2006.

[CDIO04] Jason H. Cantarella, Erik D. Demaine, Hayley N. Iben, and James F. O'Brien. An energy-driven approach to linkage unfolding. In *Proceedings of the 20th Annual ACM Symposium on Computational Geometry*, pages 134–143, Brooklyn, New York, June 2004.

[CDR03] Robert Connelly, Erik D. Demaine, and Günter Rote. Straightening polygonal arcs and convexifying polygonal cycles. *Discrete & Computational Geometry*, 30(2):205–239, September 2003.

[CJ98] J. Cantarella and H. Johnston. Nontrivial embeddings of polygonal intervals and unknots in 3-space. *J. Knot Theory Ramifications*, 7(8):1027–1039, 1998.

[CO01] Roxana Cocan and Joseph O'Rourke. Polygonal chains cannot lock in 4D. *Discrete & Computational Geometry*, 20(3):105–129, November 2001.

[DDE+05] Erik D. Demaine, Martin L. Demaine, David Eppstein, Greg N. Frederickson, and Erich Friedman. Hinged dissection of polyominoes and polyforms. *Computational Geometry: Theory and Applications*, 31(3):237–262, June 2005.

[DDIL07] E.D. Demaine, M.L. Demaine, J. Iacono, and S. Langerman. Wrapping the mozartkugel. In *Abstracts from the 23rd European Workshop on Computational Geometry (EuroCG 2007)*, pages 14–17, 2007.

[DDLS05] Erik D. Demaine, Martin L. Demaine, Jeffrey F. Lindy, and Diane L. Souvaine. Hinged dissection of polypolyhedra. In *Proceedings of the 9th Workshop on Algorithms and Data Structures*, volume 3608 of *Lecture Notes in Computer Science*, pages 205–217, Waterloo, Canada, August 2005.

[DEE+02] Erik D. Demaine, David Eppstein, Jeff Erickson, George W. Hart, and Joseph O'Rourke. Vertex-unfolding of simplicial manifolds. In *Proceedings of the 18th Annual ACM Symposium on Computational Geometry*, pages 237–243, Barcelona, Spain, June 2002.

[DEE+03] Erik D. Demaine, David Eppstein, Jeff Erickson, George W. Hart, and Joseph O'Rourke. Vertex-unfolding of simplicial manifolds. In *Discrete Geometry: In Honor of W. Kuperberg's 60th Birthday*, pages 215–228. Marcer Dekker Inc., 2003.

[DFO06] M. Damian, R. Flatland, and J. O'Rourke. Grid vertex-unfolding orthogonal polyhedra. In *Proceedings of the 23rd Symposium on Theoretical Aspects of Computer Science (STACS 2006)*, volume 3884 of *LNCS*, pages 264–276. Springer-Verlag, 2006.

[DFO07] M. Damian, R. Flatland, and J. O'Rourke. Epsilon-unfolding orthogonal polyhedra. *Graphs and Combinatorics*, 2007.

[DGOT06] Erik D. Demaine, Blaise Gassend, Joseph O'Rourke, and Godfried T. Toussaint. Polygons flip finitely: Flaws and a fix. In *Proceedings of the 18th Canadian Conference on Computational Geometry (CCCG 2006)*, pages 109–112, August 14–16 2006.

[DIL] E.D. Demaine, J. Iacono, and S. Langerman. Grid vertex-unfolding orthostacks. *International Journal of Computational Geometry and Applications.* To appear. Preliminary version: [DIL05].

[DIL05] E.D. Demaine, J. Iacono, and S. Langerman. Grid vertex-unfolding orthostacks. In *Proceedings of the Japan Conference on Discrete and Com-*

putational Geometry (JCDCG 2004), volume 3742 of *LNCS*, pages 76–82. Springer-Verlag, 2005.

[DLO01] E.D. Demaine, S. Langerman, and J. O'Rourke. Short interlocked linkages. In *Proceedings of the 13th Canadian Conference on Computational Geometry (CCCG 2001)*, 2001.

[DLO03] E.D. Demaine, S. Langerman, and J. O'Rourke. Geometric restrictions on polygonal protein chain production. In *Proceedings of the 14th Annual International Symposium on Algorithms and Computation (ISAAC 2003)*, volume 2906 of *LNCS*, pages 395–404. Springer-Verlag, 2003.

[DLO06] E.D. Demaine, S. Langerman, and J. O'Rourke. Geometric restrictions on polygonal protein chain production. *Algorithmica*, 44(2):167–181, February 2006. Special issue of selected papers from the 14th Annual International Symposium on Algorithms and Computation (ISAAC 2003). Preliminary version: [DLO03].

[DLOS02] E.D. Demaine, S. Langerman, J. O'Rourke, and J. Snoeyink. Interlocked open linkages with few joints. In *Proceedings of the 2002 ACM Symposium on Computational Geometry (SoCG 2002)*, 2002.

[DLOS03] E.D. Demaine, S. Langerman, J. O'Rourke, and J. Snoeyink. Interlocked open and closed linkages with few joints. *Computational Geometry: Theory and Applications*, 26(1):37–45, 2003. Special issue of selected papers from the *13th Canadian Conference on Computational Geometry (CCCG'01)*. Preliminary version: [DLO01].

[DO00] E. D. Demaine and J. O'Rourke. Open problems from CCCG'99. In *Proc. 12th Canad. Conf. Comput. Geom.*, pages 269–272, August 2000.

[DO05] Erik D. Demaine and Joseph O'Rourke. A survey of folding and unfolding in computational geometry. In Jacob E. Goodman, János Pach, and Emo Welzl, editors, *Combinatorial and Computational Geometry*, volume 52 of *Mathematical Sciences Research Institute Publications*, pages 167–211. Cambridge University Press, August 2005.

[DO07] Erik D. Demaine and Joseph O'Rourke. *Geometric Folding Algorithms: Linkages, Origami, Polyhedra.* Cambridge University Press, 2007.

[Epp01] David Eppstein. Hinged kite mirror dissection. arXiv:cs.CG/0106032, June 2001. `http://arXiv.org/abs/cs.CG/0106032`.

[Erd35] Paul Erdős. Problem 3763. *American Mathematical Monthly*, 42:627, 1935.

[FHM+01] Thomas Fevens, Antonio Hernandez, Antonio Mesa, Patrick Morin, Michael Soss, and Godfried Toussaint. Simple polygons with an infinite sequence of deflations. *Beiträge zur Algebra und Geometrie (Contributions to Algebra and Geometry)*, 42(2):307–311, 2001.

[Fre02] Greg N. Frederickson. *Hinged Dissections: Swinging & Twisting*. Cambridge University Press, August 2002.

[GLO+04] J. Glass, S. Langerman, J. O'Rourke, J. Snoeyink, and J.K. Zhong. A 2-chain can interlock with a k-chain. In *14th Annual Fall Workshop on Computational Geometry*, pages 18–19, 2004.

[GLOZ06] J. Glass, B. Lu, J. O'Rourke, and J.K. Zhong. A 2-chain can interlock with an open 11-chain. *Geombinatorics*, XV(4):166–176, April 2006.

[Grü95] Branko Grünbaum. How to convexify a polygon. *Geombinatorics*, 5:24–30, July 1995.

[GZ98] Branko Grünbaum and Joseph Zaks. Convexification of polygons by flips and by flipturns. Technical Report 6/4/98, Department of Mathematics, University of Washington, Seattle, 1998.

[Nag39] Béla de Sz. Nagy. Solution to problem 3763. *American Mathematical Monthly*, 46:176–177, March 1939.

[NHB+00] P. Nissen, J. Hansen, N. Ban, B. Moore, and T. A. Steitz. The structural basis of ribosome activity in peptide bond synthesis. *Science*, 289:920–930, August 2000.

[O'R07] J. O'Rourke. Unfolding orthogonal polyhedra, 2007.

[Sos01] Michael Soss. *Geometric and Computational Aspects of Molecular Reconfiguration*. PhD thesis, School of Computer Science, McGill University, 2001.

[SSL03] A. Solomon, P. Sutcliffe, and R. Lister. Sorting circular permutations by reversal. In *Proceedings of the 8th International Workshop on Algorithms* *and Data Structures (WADS 2003)*, volume 2748 of *LNCS*, pages 319–328. Springer, 2003.

[Str00] Ileana Streinu. A combinatorial approach to planar non-colliding robot arm motion planning. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, pages 443–453, Redondo Beach, California, November 2000.

[Weg93] Bernd Wegner. Partial inflation of closed polygons in the plane. *Beiträge zur Algebra und Geometrie*, 34(1):77–85, 1993.

Appendix A

Articles

[CDD$^+$] R. Connelly, E.D. Demaine, M.L. Demaine, S.P. Fekete, S. Langerman, J.S.B. Mitchell, A. Ribó, and G. Rote. Locked and unlocked chains of planar shapes. *Discrete & Computational Geometry*. Submitted. Preliminary version: [CDD$^+$06].

[DLOS03] E.D. Demaine, S. Langerman, J. O'Rourke, and J. Snoeyink. Interlocked open and closed linkages with few joints. *Computational Geometry: Theory and Applications*, 26(1):37–45, 2003. Special issue of selected papers from the *13th Canadian Conference on Computational Geometry (CCCG'01)*. Preliminary version: [DLO01].

[DLOS02] E.D. Demaine, S. Langerman, J. O'Rourke, and J. Snoeyink. Interlocked open linkages with few joints. In *Proceedings of the 2002 ACM Symposium on Computational Geometry (SoCG 2002)*, 2002.

[GLO$^+$04] J. Glass, S. Langerman, J. O'Rourke, J. Snoeyink, and J.K. Zhong. A 2-chain can interlock with a k-chain. In *14th Annual Fall Workshop on Computational Geometry*, pages 18–19, 2004.

[ADD$^+$02] G. Aloupis, E.D. Demaine, V. Dujmovic, J. Erickson, S. Langerman, H. Meijer, J. O'Rourke, M. Overmars, M. Soss, I. Streinu, and G.T. Toussaint. Flat-state connectivity of linkages under dihedral motions. In *Proceedings of the 13th Annual International Symposium on Algorithms and Computation (ISAAC 2002)*, volume 2518 of *LNCS*, pages 369–380. Springer-Verlag, 2002.

[DLO06] E.D. Demaine, S. Langerman, and J. O'Rourke. Geometric restrictions on polygonal protein chain production. *Algorithmica*, 44(2):167–181, February 2006.

Special issue of selected papers from the 14th Annual International Symposium on Algorithms and Computation (ISAAC 2003). Preliminary version: [DLO03].

[ABD+02] G. Aloupis, P. Bose, E.D. Demaine, S. Langerman, H. Meijer, M. Overmars, and G.T. Toussaint. Computing signed permutations of polygons. In *Proceedings of the 14th Canadian Conference on Computational Geometry (CCCG 2002)*, 2002.

[ABB+] G. Aloupis, B. Ballinger, P. Bose, M. Damian, E.D. Demaine, M.L. Demaine, R. Flatland, F. Hurtado, S. Langerman, J. O'Rourke, P. Taslakian, and G. Toussaint. Vertex pops and popturns. In *Canadian Conference on Computational Geometry (CCCG 2007)*. Submitted.

[ADL+] G. Aloupis, E.D. Demaine, S. Langerman, P. Morin, J. O'Rourke, I. Streinu, and G.T. Toussaint. Unfolding polyhedral bands. *Computational Geometry: Theory and Applications*. To appear. Special issue of selected papers from the 16th Canadian Conference on Computational Geometry (CCCG'04). Preliminary version: [ADL+04].

[DIL] E.D. Demaine, J. Iacono, and S. Langerman. Grid vertex-unfolding orthostacks. *International Journal of Computational Geometry and Applications*. To appear. Preliminary version: [DIL05].

[DDIL07] E.D. Demaine, M.L. Demaine, J. Iacono, and S. Langerman. Wrapping the mozartkugel. In *Abstracts from the 23rd European Workshop on Computational Geometry (EuroCG 2007)*, pages 14–17, 2007.

Locked and Unlocked Chains of Planar Shapes*

Robert Connelly[†] Erik D. Demaine[‡] Martin L. Demaine[†] Sándor P. Fekete[§]
Stefan Langerman[¶] Joseph S. B. Mitchell[||] Ares Ribó[**] Günter Rote[||]

Dedicated to Godfried Toussaint
on the occasion of his 60th birthday.

Abstract

We extend linkage unfolding results from the well-studied case of polygonal linkages to the more general case of linkages of polygons. More precisely, we consider chains of nonoverlapping rigid planar shapes (Jordan regions) that are hinged together sequentially at rotatable joints. Our goal is to characterize the familes of planar shapes that admit *locked chains*, where some configurations cannot be reached by continuous reconfiguration without self-intersection, and which families of planar shapes guarantee *universal foldability*, where every chain is guaranteed to have a connected configuration space. Previously, only obtuse triangles were known to admit locked shapes, and only line segments were known to guarantee universal foldability. We show that a surprisingly general family of planar shapes, called *slender adornments*, guarantees universal foldability: roughly, the distance from each edge along the path along the boundary of the slender adornment to each hinge should be monotone. In constrast, we show that isosceles triangles with any desired apex angle $< 90°$ admit locked chains, which is precisely the threshold beyond which the slender property no longer holds.

1 Introduction

In this paper, we explore the motion-planning problem of *reconfiguring* or *folding* a hinged collection of rigid objects from one state to another while avoiding self-intersection. This general problem has been studied since the beginnings of the motion-planning literature when Reif [Rei79] proved that deciding reconfigurability of a "tree" of polyhedra, amidst fixed polyhedral obstacles, is PSPACE-hard. This result has been strengthened in various directions over the years, although the cleanest

*A preliminary version of this paper appeared in *Proceedings of the 22nd Annual ACM Symposium on Computational Geometry*, June 2006, pages 61–70. R. Connelly supported in part by NSF grant DMS-0209595. E. Demaine supported in part by NSF grant CCF-0347776 and DOE grant DE-FG02-04ER25647. S. Langerman is Chercheur qualifié du FNRS. J. Mitchell supported in part by NSF (CCR-0098172, ACI-0328930, CCF-0431030, CCF-0528209), the U.S.-Israel Binational Science Foundation (2000160), Metron Aviation, and NASA Ames (NAG2-1620).

[†]Department of Mathematics, Cornell University, Ithaca, NY 14853, USA. connelly@math.cornell.edu

[‡]MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar St., Cambridge, MA 02139, USA, {edemaine,mdemaine}@mit.edu

[§]Abteilung für Optimierung, Technische Universität Braunschweig, Pockelsstrasse 14, D-38106 Braunschweig, Germany. s.fekete@tu-bs.de

[¶]Chercheur qualifié du FNRS, Université Libre de Bruxelles, Département d'informatique, ULB CP212, Belgium. Stefan.Langerman@ulb.ac.be

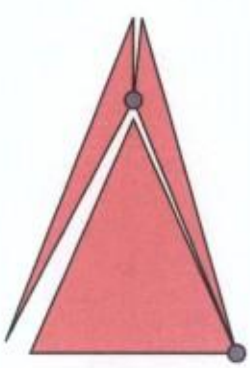
[||]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794-3600, USA. jsbm@ams.sunysb.edu

[**]Institut für Informatik, Freie Universität Berlin, Takustraße 9, D-14195 Berlin, Germany. {ribo,rote}@inf.fu-berlin.de

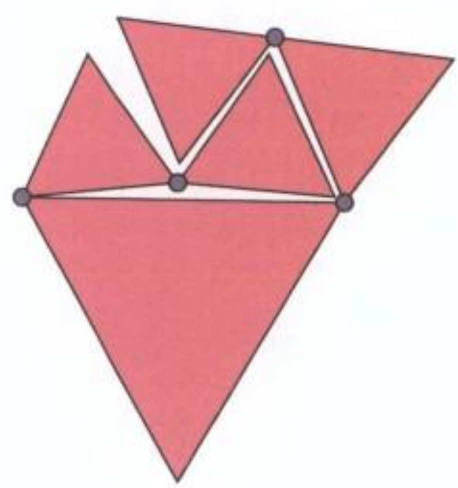
versions were obtained only very recently: deciding reconfigurability of a tree of line segments in the plane, and deciding reconfigurability of a chain of line segments in 3D, are both PSPACE-complete [AKRW04]. This result is tight in the sense that deciding reconfigurability of a chain of line segments in the plane is easy, in fact, trivial: the answer is always yes [CDR03].

These results illustrate a rather fine line in reconfiguration problems between computationally difficult and computationally trivial. The goal of our work is to characterize what families of planar shapes and hingings lead to the latter outcome, a *universality result* that reconfiguration is always possible. The only known example of such a result, however, is the family of chains of line segments, and that problem was unsolved for about 25 years [CDR03]. (A *chain* is a sequence of line segments joined end-to-end that are disjoint except except for consecutive endpoints, where they are hinged. It is called an *open chain* when the first line segment is not joined to the last, and *closed* when it is joined to the last segment in a closed cycle.) Even small perturbations to the problem, such as allowing a single point where three line segments join, leads to *locked* examples where reconfiguration is impossible [CDR02].

What about chains of shapes other than line segments? It is easy to see that a shape tucked into a "pocket" of a nonconvex shape immediately makes trivial locked chains with two pieces. Back in January 1998, the third author showed how to simulate this behavior with convex shapes, indeed, just three triangles; see Figure 1(a). This example has circulated throughout the years to many researchers (including the authors of this paper) who have asked about chains of 2D shapes. The only really unsatisfying feature of the example is that some of the angles are very obtuse. But with a little more work, one can find examples with acute angles, indeed, equilateral triangles, albeit of different size; see Figure 1(b). What could be better than equilateral triangles?



(a)



(b)

Figure 1: Simple examples of locked chains of triangles. (a) A locked chain of three triangles. (b) A locked chain of equilateral triangles of different sizes. The gaps should be tighter than drawn.

It is therefore reasonable to expect, as we did for many years, that there is no interesting class of shapes, other than line segments, with a universality result—essentially all other shapes admit locked chains. We show in this paper, however, that this guess is wrong.

We introduce a family of shapes, called *slender adornments*, and prove that all open chains, made up of arbitrarily many different shapes from this family, can be universally reconfigured between any two states. Indeed, we show that these chains have a natural canonical configuration, analogs of the straight configuration of an open chain. Our result is based on the existence of "expansive motions", proved in [CDR03]. Our techniques build on the theory of unfolding chains of line segments, substantially generalizing and extending the results from that theory. Indeed, the results in this paper essentially piggy-back onto the results of [CDR03] or any of the other results and algorithms such as [Str00] or [Str05] that provide a continuous expansive motion of the base chain. Our results go far beyond what *we* thought was possible (until recently). As part of the methods that we describe here, we also consider discrete expansive motions of the base chain, that do not necessarily come from a continuous expansive motion. (A *discrete expansion* of a chain C is simply another corresponding chain C' such that if x and y are two points in C the distance between corresponding points x' and y' is not smaller than the distance between x and y.) In that case if all the slender adornments are *symmetric* under the reflection about the line of the base chain, then any expansive discrete motion of the base chain will have the property that the attached adornments will not overlap. It turns out that the continuous case, when the adornments are not necessarily symmetric, follows from the discrete symmetric case.

The family of slender adornments has several equivalent definitions. The key idea is to distinguish the two hinge points on the boundary of the shape connecting to the adjacent shapes in the chain, and view the shape as an *adornment* to the line segment connecting those two hinges, called the *base*. This view is without loss of generality, but provides additional information relating the shapes and how they are attached to neighbors, which turns out to be crucial to obtaining a universality result. An adornment is *slender* if the distance from either endpoint of the base to a point moving along one side of the adornment changes monotonically. If the boundary curve, defining the adornment, is sufficiently smooth, it is slender if and only if every inward normal of the shape hits the base. Equivalently, an adornment is slender if it is the union, in each half-plane having the base as a boundary, of the intersection of pairs of disks centered at the two endpoints of the base.

Slender adornments are quite general. Figure 2 shows several examples of slender adornments. These examples are themselves slender adornments, but also any pair can be joined along their bases so that the union makes another slender adornment. Our results imply that one can take any of these slender adornments, link the bases together into an open chain in any way that the chain does not self-intersect, and the resulting chain can be unfolded without self-intersection to a straight configuration, and thus the chain can be folded without self-intersection into every configuration.

We also demonstrate the tightness of the family of slender adornments by giving examples of locked chains of shapes that are not slender. Specifically, we show that, for any desired angle $\theta < 90°$, there is a locked chain of isosceles triangles with apex angle θ. This is precisely the family of isosceles-triangle adornments that are not slender. Thus, for chains of triangles, obtuseness is really desirable, contrary to our intuition from Figure 1(a): the key is that the apex angle opposite the base (in the adornment view) be nonacute, not any other angle. The proof that our examples are locked uses the self-touching theory developed for trees of line segments in [CDR02].

Motivation. Hinged collections of rigid objects have been studied previously in many contexts, particularly robotics. One recent body of algorithmic work by Cheong et al. [CvdSG$^+$] considers how chains of polygonal objects can be *immobilized* or *grasped* by a robot with a limited number of actuators. Grasping is a natural first step toward robotic *manipulation*, but the latter challenge requires a better understanding of reconfigurability. This paper offers the first theoretical underpinnings for reconfiguration of chains of rigid objects (other than line segments).

Another potential application is to continuous folding of hinged dissections. Hinged dissections are chains or trees of polygons that can be reconfigured into two or more self-touching configurations with desired silhouettes. For example, Figure 3 shows a classic hinged dissection from 1902 of a

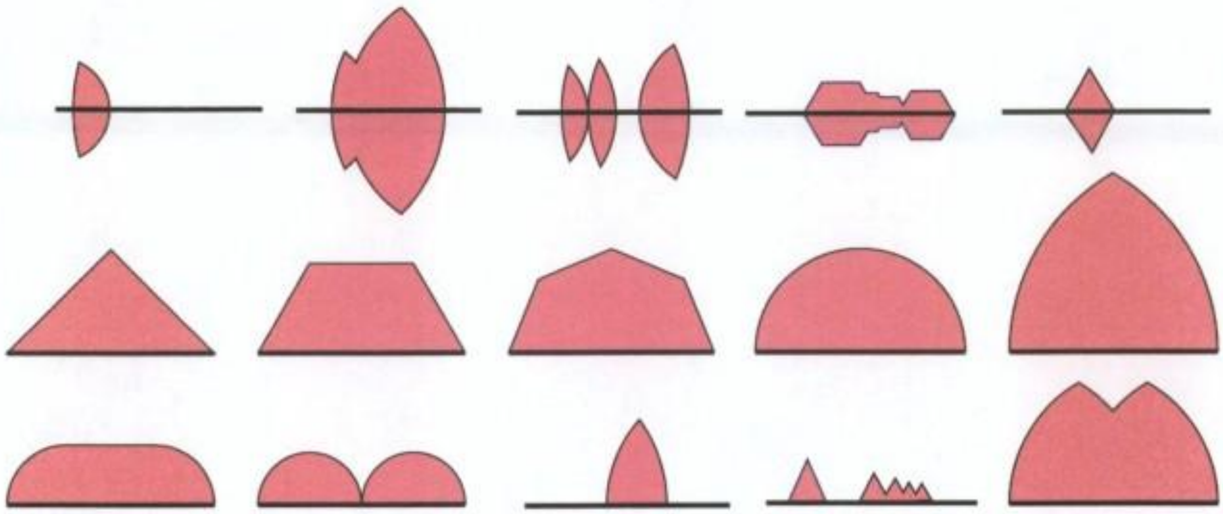


Figure 2: Examples of slender adornments. The base is drawn bold. The examples in the top row are symmetric. Any two of these examples can be glued together along a common base so that the union also becomes a slender adornment.

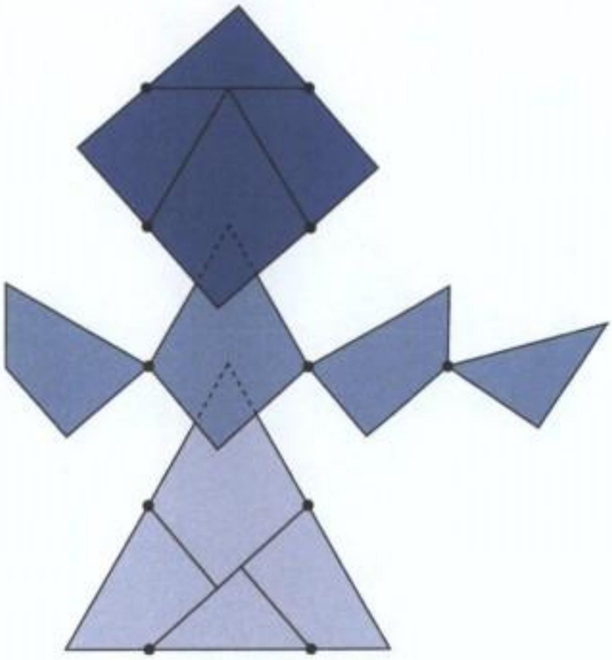


Figure 3: Hinged dissection of square to equilateral triangle, described by Dudeney [Dud02]. Different shades show different folded states (overlapping slightly).

square into an equilateral triangle of the same area. Many general families of hinged dissections have been established in the recent literature [AN98, DDE$^+$05, DDLS05, Epp01, Fre02]. One problem not addressed in this literature, however, is whether the reconfigurations can actually be executed without self-intersection, as in Figure 3. Our results provide potential tools, previously lacking, for addressing this problem. While hinged dissections have frequently been considered in recreational contexts, they have recently found applications in nanomanufacturing [MTW$^+$02] and reconfigurable robotics [DDLS05].

Outline. This paper is organized as follows. Section 2 defines the model and slender adornments more precisely, and proves several basic properties. Section 3 describes the case when each adornment is symmetric about its base and is important for proving, in Section 4, that simple chains of slender adornments can always be unfolded so that the base is convex or straight. In Section 5 we discuss the situation when the adornments are permitted to overlap. Section 6 describes our examples of locked chains of isosceles triangles, including the necessary background from self-touching trees. The Appendix describes an example of a closed chain, with slender adornments attached, that has infinitely many components in its configuration space.

2 Slender Adornments

This section provides a formal statement of the objects we consider—adorned chains consisting of slender adornments—and proves several basic results about them.

2.1 Adorned Chains

Our object of study is a chain of nonoverlapping rigid planar shapes (Jordan regions) that are hinged together sequentially at rotatable joints. Another way to view such a chain is to consider the *underlying polygonal chain*, the *core*, of line segments connecting successive joints. (For an open chain, there is some freedom in choosing the endpoints for the first and the last bar of the chain.) On the one hand, these line segments can be viewed as *bars* that move rigidly with the shapes to which they belong. On the other hand, the shapes can be viewed as "adornments" to the bars of an underlying polygonal chain. This view leads to the concept of an "adorned polygonal chain", which we now proceed to define more precisely.

An *adornment* is a simply connected compact region in the plane, called the *shape*, together with a line segment xy connecting two boundary points, called the *base*. There are two *boundary arcs* from x to y that enclose the shape, called *sides*. We require the base to be contained in the shape; i.e., the base must be a chord of the shape.

We say that two distinct adornments *overlap* when some point of one adornment lies in the interior of the other, and we insist that the relative interiors of the base chains be disjoint. Thus, the bases of two shapes are not allowed to touch except at common hinges of the polygonal chain. An *adorned polygonal chain* is a set of nonoverlapping adornments whose bases form a polygonal chain. We permit the shapes to touch on their boundary and to slide along each other.

For our main result, Theorem 3, where we assume that the motion of the base is expansive, it is not necessary to assume that the base chain is simple. It can be any finite embedded graph with straight edges whose relative interiors are pairwise disjoint; a vertex may touch an edge. When the base chain is simple the results of [CDR03] or [Str05] guarantee that there is such an expansive motion. On the the other hand, although an expansive motion of the base chain of a strictly simple closed polygon to a convex convex configuration can be guaranteed, it may happen that two realizations are not in the same configuration component, as shown in Figure 10, and in the Appendix there is a description of a case when there are infinitely many components in the configuration space.

The viewpoint of a chain of shapes as an adorned polygonal chain is useful for two reasons. First, we can more easily talk about the kinds of shapes, and their relation to the locations of the incident hinges, in a family of chains: this information is captured by the adornments. Second, the underlying polygonal chain provides a mechanism for folding the chain of shapes, as well as a natural *unfolding* goal: straighten the underlying open chain or convexify the underlying closed chain. Indeed, we show that, in some cases, unfolding motions of the polygonal chain induce valid unfolding motions of the chain of shapes.

2.2 Slender Adornments

An adornment is defined to be *slender* if, for a point moving on either side of the shape, the distance to each endpoint of the base changes monotonically (possibly not strictly monotonically). An adornment is called *symmetric* if it is symmetric about the line through the base. An adornment is called *one-sided* if it lies in just one of the closed half-planes whose boundary contains the base. Clearly, a general adornment is the union of two one-sided adornments, and a one-sided adornment is the intersection of a symmetric adornment with a closed half-plane whose boundary contains the base. For any base interval with endpoints x and y, and a point z in the plane, where $||z-x|| \leq ||y-x||$ and $||z-y|| \leq ||x-y||$, let $L(z)$ be the intersection of the disk with z on its boundary centered at x, and the intersection of the disk with z on its boundary centered at y. We call $L(z)$ the *lens determined by z associated to the base $[x,y]$*. A *half-lens*, denoted as $\hat{L}(z)$, is the intersection of $L(z)$ and the closed half-plane through the base containing z. See Figure 4 for a picture of a half-lens and lens. The following are some simple, but useful, properties of lenses.

Proposition 1 *For any point z in a (symmetric) slender adornment A, $\hat{L}(z) \subset A$ ($L(z) \subset A$).*

Proof: Let z be a point on the defining boundary of A. Since the distance to x along the boundary is monotone, no point along the path from z to y intersects the interior of the circle centered at x through z. Similarly, no point along the path from z to x intersects the interior of the circle centered at y through z. Thus, the intersection of the circular disks centered at x and y, with z on their boundary, and the closed half-plane containing z, $\hat{L}(z)$ is contained in the adornment. In the symmetric case, the intersection of the circular disks with z on their boundary $L(z)$ is contained in A. See Figure 4. □

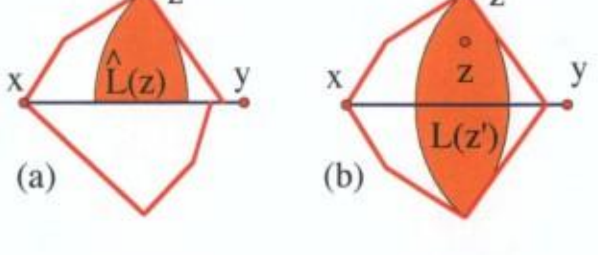


Figure 4: Figure (a) shows a half-lens in non-symmetric adornment, and Figure (b) show a symmetric lens with a point z in the interior of the lens and the adornment.

Proposition 2 *For any point z in the interior of a (symmetric) slender adornment A, there is a half-lens $\hat{L}(z') \subset A$ (lens $L(z') \subset A$) that has z in its interior.*

Proof: The half-lens (lens) through z is contained in A by Proposition 1. Since z is in the interior of A, there is another point z' in A on the line perpendicular to the base segment slightly further away from the base. Then z is in the interior of the half-lens (lens) defined by z'. □

Proposition 3 *A symmetric adornment is slender if and only if it is the union of the intersection of pairs of disks one centered at each endpoint x and y of its base.*

Proof: By Proposition 1, $L(z) \subset A$. Thus, the union of the lenses $L(z)$ for z on the boundary of the slender adornment is contained in the adornment.

To show the reverse containment, any point z in the interior of the adornment lies on a circle centered at x, and this circle must intersect the boundary of the adornment at (at least) one point z'. Then z is in $L(z')$. Thus, the union of the lenses $L(z')$ for z' on the boundary of the slender adornment contains the adornment.

To show the converse suppose a point z is on the boundary of A, any closed adornment defined by an continuous path from x to y on one side. Then z must lie in some adornment and so the lens $L(z)$ itself must be in A. Thus, the path along the boundary, away from x, must not enter the circle through z centered at x. Thus, the path is monotone at z. A similar argument applies to y. □

Proposition 4 *Finite unions and arbitrary intersections of slender adornments are slender adornments.*

Proof: This follows from Proposition 3 in the symmetric case, and the non-symmetric case follows from the symmetric case by intersecting with the closed half-plane containing the line segment. □

Proposition 5 *Every slender adornment is contained in the symmetric lens determined by either of the points equidistant from the endpoints of the base as in Figure 5.*

Proof: Any slender adornment must be contained in the disk through the other end of the base, and thus it is in the intersection of those two disks. □

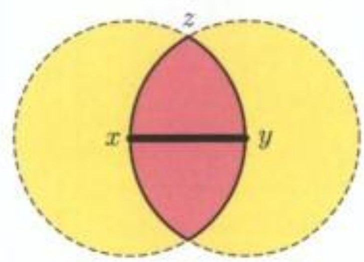


Figure 5: The largest slender adornment with a given base is a lens $L(z)$, where $||z - x|| = ||y - x||$ and $||z - y|| = ||y - x||$.

Figure 2 shows examples of slender adornments. If a single triangle is attached, where the base forms one side, then the angle at the other vertex must be obtuse or a right angle.

2.3 Remark

Suppose that an adornment is such that its boundary is a differentiable curve from one endpoint x of the base to the other endpoint y. Then the condition of being slender is equivalent to requiring

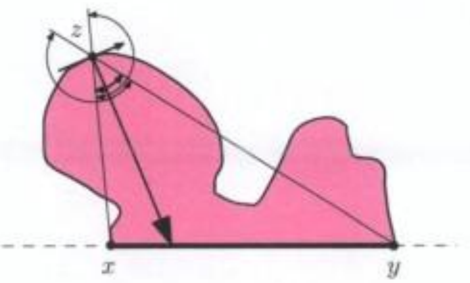


Figure 6: The normal property for slender shapes.

that every inward normal of the shape intersects the base, before exiting the shape. This is shown in Figure 6. This property was our original motivation for defining the property of being slender. It insures that slender adornments will not get closer together during an expansive motion of the base. But the monotone distance property is easier to handle, and it does not raise questions of differentiability.

2.4 Kirszbraun's Theorem

In what follows it is very handy to have the following theorem of Kirszbraun [Kir34].

Theorem 1 *Suppose a finite set of closed circular disks in Euclidean space are rearranged so that no pair of centers gets strictly closer together. If the original set has an empty intersection, so does the rearranged set.*

There is a discussion and proof of this in [Ale84] as well as references to other proofs. We only need this result for four disks in the Euclidean plane.

3 Expanded Slender Symmetric Adornments Never Overlap

We first prove the following for the case of symmetric slender adornments. Note that the following result is for discrete expansions of the base chain. Recall that two adornments *overlap* if a point in one adornment lies in the interior of the other. This allows their boundaries to touch, but not to penetrate each other. Note that the bases of a chain do not cross as well, by the expansive property of a discrete motion. We do not need the continuous expansive property for this result.

Theorem 2 *Consider two configurations X and Y of corresponding chains with symmetric slender adornments such that the base chain of Y is an expansion of the base chain of X. We assume that the adornments attached to the base chain of X do not overlap. Then, when the corresponding adornments are attached to the base chain of Y, they also do not overlap.*

Proof: Suppose A_X and B_X are two slender adornments attached to different links of the base chain of X, and A_X and B_X do not overlap. Let A_Y and B_Y be the corresponding adornments for Y. Suppose that z is a point in the intersection $A_Y \cap B_Y$, where z is in the interior of, say, A_Y. We wish to find a contradiction.

Let z_A and z_B be the corresponding distinct points in A_X and B_X, respectively, that map to z under the expanding map of their bases. Thus, the lenses $L_A(z_A)$ and $L_B(z_B)$ for A_X and B_X have disjoint interiors, since the adornments do not overlap. Since z is in the interior of A_Y, we can assume that $L_A(z_A)$ can be chosen so that the closed lenses $L_A(z_A)$ and $L_B(z_B)$ are disjoint also. Thus the four circular disks that correspond to the circular disks that define $L_A(z_A)$ and $L_B(z_B)$ have an empty intersection. By Kirszbraun's Theorem 1, and the expansion property of the endpoints of the bases of A_X and B_X, which are the centers of the four circular disks, the intersection of the corresponding lenses for A_Y and B_Y must also be empty, contradicting the assumption that A_Y and B_Y overlap. See Figure 7. □

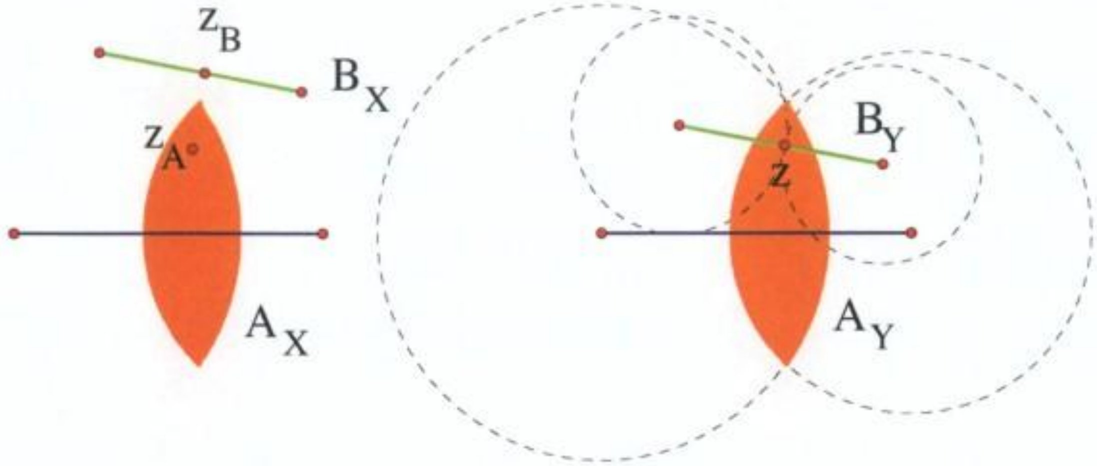


Figure 7: This is the situation when two adornments overlap. The four circles that used in the application of Kirszbraun's Theorem are indicated. Note that, in this figure, the motion from X to Y is not an expansion, since that would contradict Theorem 2.

For discrete expansions, it is not possible to deal with non-symmetric adornments. Figure 8 shows an example of two chains with corresponding slender adornments, one an expansion of the other. One starts with no overlap, and the other has such an overlap.

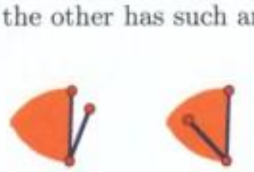


Figure 8: This shows two chains, with slender but not symmetric adornments, where one is an expansion of the other, while there is an overlap in the expanded configuration, but not the original.

4 Slender Adornments Cannot Lock

We now consider the general case, assuming a continuous expansive motion.

Theorem 3 *Suppose there is a continuous expansive motion of the base chain with slender non-overlapping, not necessarily symmetric, adornments attached. Then the adornments never overlap during the motion.*

Proof: Because of the expansive property, two segments of the base chain can only intersect at common endpoints of adjacent segments. Thus, suppose z_A, in the interior of adornment A, intersects z_B in adornment B at some time t_1 during the motion. We look for a contradiction. By Proposition 2, there is a closed half-lens L_A for A that contains z_A in its interior and there is a first time $t_0 < t_1$ when L_A intersects another half-lens L_B for B that contains z_B. Necessarily, that intersection must be on the common boundary of L_A and L_B. (Note that L_B could be a single point on a base segment.) Then there are three cases that can occur. In each case, we will show that when the motion is continued from t_0 to t_1, z_A and z_B cannot intersect.

Case 1: The bases of A and B intersect in the interior of at least one of the bases. This cannot happen because the bases are initially disjoint and the motion is expansive. See Figure 9(a).

Case 2: The base of A or B intersects the half-lens of the other. The half lens can be extended to a full symmetric lens without overlaping the base of the other. Applying Theorem 2 we see that z_A and z_B cannot intersect upon further expansion. See Figure 9(b).

Case 3: The half lenses of A and B intersect. In this case both half lenses can be extended to non-overlapping symmetric lenses. Again we apply Theorem 2 to see that z_A and z_B cannot intersect upon further expansion. See Figure 9(c).

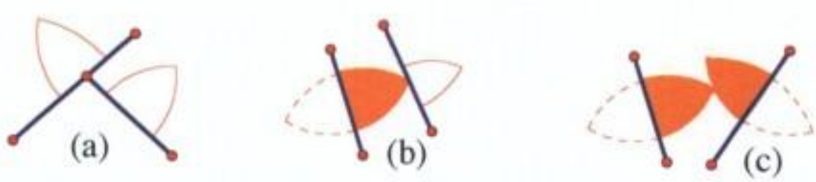


Figure 9: This presents the cases when one adornment with its base might start to overlap with the other. The dashed lines indicate where one or both of the lens of the adornment can be extended so that it does not intersect the relevant part of the other. The thick lines indicate the part of the adornment that is not to be penetrated by the other lens or base. The thin lines indicate where some of the rest of the adornment might lie, containing the point z_A, say, in the proof.

□

Corollary 4 *A strictly simple polygonal chain with slender adornments attached can always be straightened or convexified by a continuous motion.*

Proof: By [CDR03] There is a continuous expansive motion of the base chain, where the final configuration is convex in the case of a closed chain and straight in the case of an open chain. Then Theorem 3 implies that they can be carried along without overlap. □

Corollary 5 *Strictly simple open polygonal chains with slender adornments, attached on the same corresponding sides, can always be continuously reconfigured between any two states.*

Proof: By Corollary 4, both states can be continuously expansively reconfigured so that the base chains are straight. But there is only one way to do this, since the adornments are attached on the same sides. Thus, one state can be expanded to have a straight base configuration, and then contracted to the other configuration by running its expansion backwards. □

It is interesting to note that the conclusion of Corollary 5 does not hold for closed chains, even though any two convex chains with no adornments can be continuously reconfigured from one to the other. Figure 10 shows an example, where the configuration space has two components, where the base chain is a quadrilateral, and where each adornment is a triangle attached to its base.

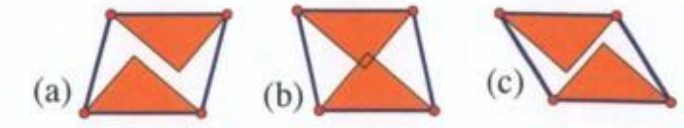


Figure 10: This shows two configurations (a) and (c) of a quadrilateral with two slender adornments attached such that it is not possible to continuously move from one to the other without colliding. Figure (b) shows how the two adornments collide as the quadrilateral is deformed from (a) to (c).

Indeed, in the appendix it is shown how to create a quadrilateral with two slender adornments such that the configuration space has infinitely many components.

It is also interesting to note that when the slender adornments are attached, and the base chain is expanded, often it happens that the motion on the adorned configuration is not expansive. Figure 11 shows an example.

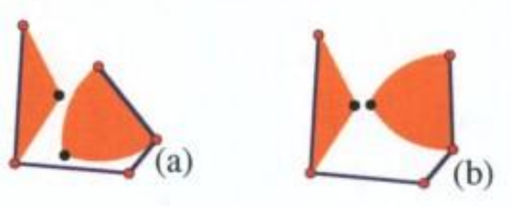


Figure 11: The base chain of Figure (a) expands to Figure (b). But the dark points on the corresponding slender adornments get closer together.

5 Generalizations: Overlapping Adornments and Generalized Slender Symmetric Adornments

In the discussion so far, we have assumed, when the adornments are attached to their chains, that they do not overlap. What happens when the slender adornments do overlap? It turns out that we can apply some of the results of [BC02] related to problems concerning areas of unions and intersections of circular disks in the plane to the case when the adornments are all symmetric.

Proposition 3 shows that any symmetric adornment is the infinite union of symmetric lenses $L(z)$ for all z on the boundary of the adornment. To apply the theory of [BC02] it is more convenient that there only be a finite number of sets involved in the union of lenses. But it is easy to see that each adornment can be approximated by a finite union of lenses.

We first define a *flower* as a set in the plane that can be described in terms of finite unions and intersections of circular disks, where each disk appears once and only once in the in the Boolean expression that describes the set. For the special case at hand we need only be concerned with flowers F of the following sort:

$$F = (B_1 \cap B_2) \cup (B_3 \cap B_4) \cup (B_5 \cap B_6) \cup \ldots (B_{N-1} \cap B_N), \tag{1}$$

where each $B_i, i = 1, \ldots, N$ is a circular disk in the plane. Flowers were defined by [GM92], and a special case of Corollary 8 in [BC02] shows the following. Let $B(x, r)$ denote a disk in the plane of radius r centered at x.

Theorem 6 *Let $B(p_i, r_i)$ and $B(q_i, r_i), i = 1, \ldots, N$ be two sets planar disks, where $||p_i - p_{i+1}|| \geq ||q_i - q_{i+1}||$ for i odd, and $||p_i - p_{i+1}|| \leq ||q_i - q_j||$ for all other pairs $i < j$. Then the area of the flower F in (1) defined for the configuration of p_i is less than or equal to the area of the flower F defined for the configuration of q_i.*

The crucial observation is that the union of the slender adornments can be approximated by flowers. Each lens is the intersection of two disks, one of the terms in (1), and Proposition 3 implies the following.

Theorem 7 *Suppose that one chain is a discrete expansion of the other and slender symmetric adornments are attached to each chain. Then the area of the union of the adornments does not decrease.*

Figure 12 shows an example of overlapping symmetric adornments. Figure 13 shows an example of

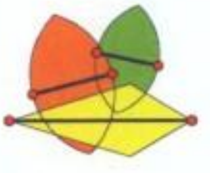


Figure 12: This Figure shows three intervals with overlapping symmetric slender adornments.

a chain with non-symmetric adornments that expands to another chain and the area of the slender adornments with an expanded core decreases.

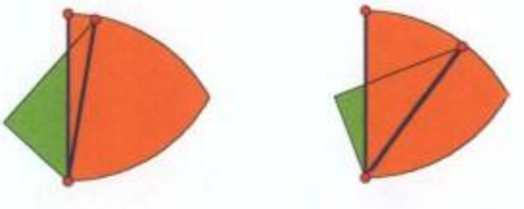
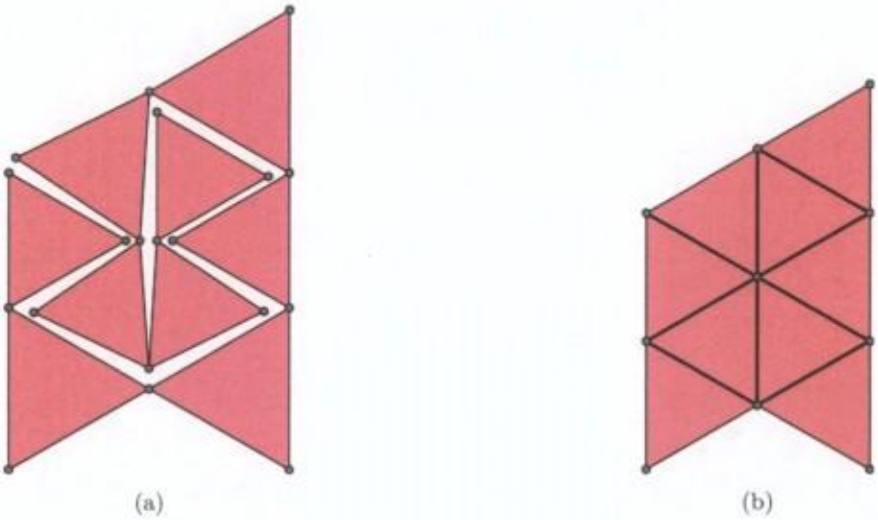


Figure 13: This is an example of two chains with non-symmetric slender adornments, where the expanded chain with adornments has smaller area.

Another possible generalization is to attach the analogue of slender adornments to simplicial complexes in higher dimensions. For example, a set A in three-space would be called *slender with* *respect to a triangle base* B if for any plane P perpendicular to the plane of A, $P \cap A$ is slender with respect to $B \cap A$. Then the analogue of Corollary 4 should also hold using the notion of symmetric slender adornments. Even the analogue of Theorem 7 for the volumes of symmetric slender adornments would still hold, but it could only be asserted for continuous expansions of the base chain. The higher dimensional version of Corollary 8 in [BC02] is not known for discrete expansions. However, in [Csi01], there is a continuous version that will suffice. In higher dimensions, the idea is to assume simply that the base chain, to which the adornments are attached, is expanded.

6 Locked Chains of Sharp Triangles

An isosceles triangle with an apex angle of $\geq 90^\circ$ and with the nonequal side as the base is a slender adornment. By Corollary 4, any chain of such triangles can be straightened. In this section we show that this result is tight: for any isosceles triangle with an apex angle of $< 90^\circ$ and with the nonequal side as a base, there is a chain of these triangles that cannot be straightened.



(a) (b)

Figure 14: A locked chain of nine equilateral triangles. (a) Drawn loosely. Separations should be smaller than they appear. (b) Drawn tightly, with no separation, as a self-touching configuration.

Figure 14(a) shows the construction for equilateral triangles (of slightly different sizes). This figure is drawn with the pieces loosely separated, but the actual construction has arbitrarily small separations and arbitrarily closely approximates the self-touching geometry shown in Figure 14(b). Stretching the triangles in this self-touching geometry, as shown in Figure 15, defines our construction for any isosceles triangles with an opposite angle of any value less than 90°. In this case, however, our construction uses two different scalings of the same triangle.

6.1 Theory of Self-Touching Configurations

This view of the construction as a slightly separated version of a self-touching configuration allows us to apply the program developed in [CDR02] for proving a configuration locked. This theory allows us to study the rigidity of self-touching configurations, which is easier because vertices cannot move even slightly, and obtain a strong form of lockedness of non-self-touching perturbations drawn with sufficiently small (but positive) separations.

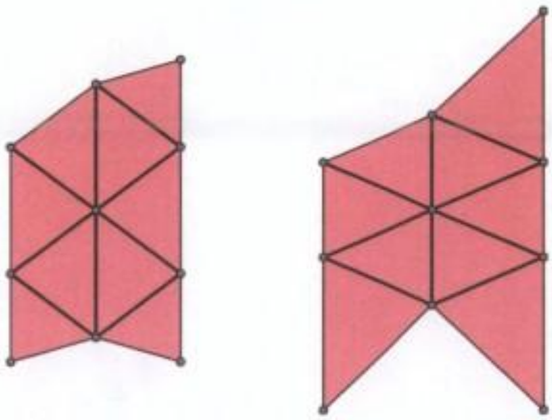


Figure 15: Variations on the self-touching configuration from Figure 14(b) to have any desired angle $< 90°$ opposite the base of each triangle.

To state this relation precisely, we need some terminology from [CDR02]. Call a linkage configuration *rigid* if it cannot move at all. Define a δ*-perturbation* of a linkage configuration to be a repositioning of each vertex within distance δ of its original position, without regard to preserving edge lengths (better than $\pm 2\delta$), but consistent with the combinatorial information of which vertices are on which side of which bar. Call a linkage *locked within* ε if no motion that leaves some bar pinned to the plane moves any point by more than ε. Call a self-touching linkage configuration *strongly locked* if, for any desired $\varepsilon > 0$, there is a $\delta > 0$ such that all δ-perturbations are locked within ε. Thus, if a self-touching configuration is strongly locked, then the smaller we draw the separations in a non-self-touching perturbation, the less the configuration can move. In particular, if we choose ε small enough, the linkage must be locked in the standard sense of having a disconnected configuration space locally.

Theorem 8 [CDR02, Theorem 8.1] *If a self-touching linkage configuration is rigid, then it is strongly locked.*

Therefore, if we can prove that the self-touching configuration in Figure 14(b) (and its variations in Figure 15) are rigid, then sufficiently small perturbations along the lines shown in Figure 14(a) are rigid.

The theory of [CDR02] also provides tools for proving rigidity of a self-touching configuration. Specifically, we can study *infinitesimal motions*, which just define the beginning of a motion to the first order. Call a configuration *infinitesimally rigid* if it has no infinitesimal motions.

Lemma 9 [CDR02, Lemma 6.1] *If a self-touching linkage configuration is infinitesimally rigid, then it is rigid.*

A final tool we need from [CDR02] is for proving infinitesimal rigidity. For each vertex u wedged into a convex angle between two bars $\{v, w_1\}$ and $\{v, w_2\}$, we say that there are two *zero-length connections* between u and v, one perpendicular to each of the two bars $\{v, w_i\}$.[1] See Figure 16. These connections must increase to the first order because u must not cross the two bars $\{v, w_i\}$. In proving infinitesimal rigidity, we can choose to discard any zero-length connections we wish, because ignoring some of the noncrossing constraints only makes the configuration more flexible. Together,

[1]The definition of such connections in [CDR02] is more general, but this definition suffices for our purposes.

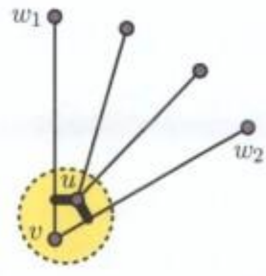


Figure 16: Two zero-length connections between vertices u and v.

the bars and the zero-length connections are the *edges* of the configuration. Define a *stress* to be an assignment of real numbers (*stresses*) to edges such that, for each vertex v, the vectors with directions defined by the edges incident to v, and with magnitudes equal to the corresponding stresses, sum to the zero vector. We denote the stress on a bar $\{v, w\}$ by ω_{vw}, and we denote the stress on a zero-length connection between vertex u and vertex v perpendicular to $\{v, w\}$ by $\omega_{u,vw}$.

Lemma 10 [CDR02, Lemma 7.2] *If a self-touching configuration has a stress that is negative on every zero-length connection, and if the configuration is infinitesimally rigid when every zero-length connection is treated as a bar pinning two vertices together, then the self-touching configuration is infinitesimally rigid.*

6.2 Locked Chains

We are now in the position to state the precise senses in which the chains of isosceles triangles in Figures 14 and 15 are locked:

Theorem 11 *The self-touching chains of nine isosceles triangles shown in Figures 14(b) and 15 are rigid provided that the apex angle is $< 90°$.*

Applying Theorem 8, we obtain the desired result:

Corollary 12 *The self-touching chains of nine isosceles triangles shown in Figures 14(b) and 15 are strongly locked provided that the apex angle is $< 90°$. Therefore, any sufficiently small non-self-touching perturbation, similar to the one shown in Figure 14(a), is locked.*

Sections 6.3–6.4 prove Theorem 11.

6.3 Simplifying Rules

We introduce two rules that significantly restrict the allowable motions of the self-touching configuration of isosceles triangles.

Rule 1 *If a bar b is collocated with another bar b' of equal length, and the bars incident to b' form angles less than $90°$ on the same side as b, then any motion must keep b collocated with b' for some positive time. See Figure 17.*

Proof: The noncrossing constraints at the endpoints of b and b' prevent b from moving relative to b' until the angles at the endpoints of b' open to $\geq 90°$, which can only happen after a positive amount of time. □

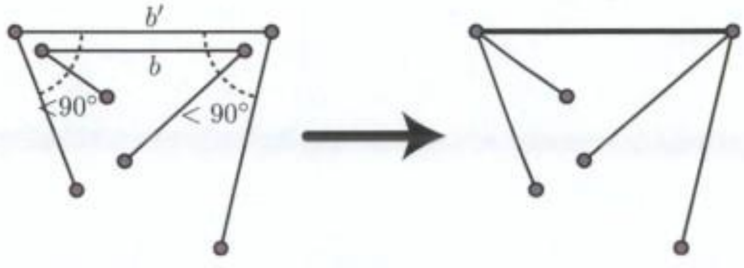


Figure 17: Rule 1 for simplifying self-touching configurations.

We can apply this rule to the region shown in Figure 18, resulting in a simpler linkage with the same infinitesimal behavior. Although the figure shows positive separations for visual clarity, we are in fact acting on the self-touching configuration of Figure 14(b).

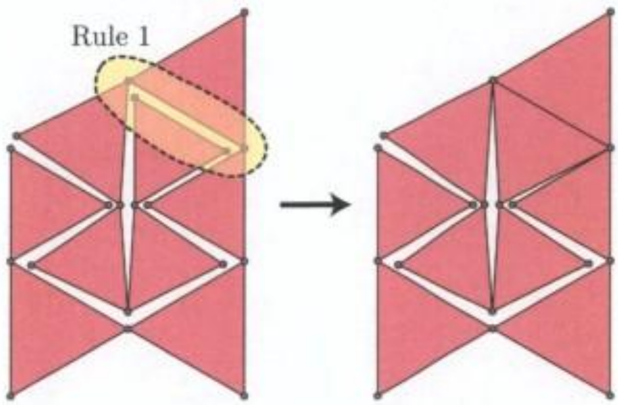


Figure 18: Applying Rule 1 to the chain of nine equilateral triangles from Figure 14.

Rule 2 *If a bar b is collocated with an incident bar b' of the same length whose other incident bar b'' forms a convex angle with b' surrounding b, then any motion must keep b collocated with b' for some positive time. See Figure 19.*

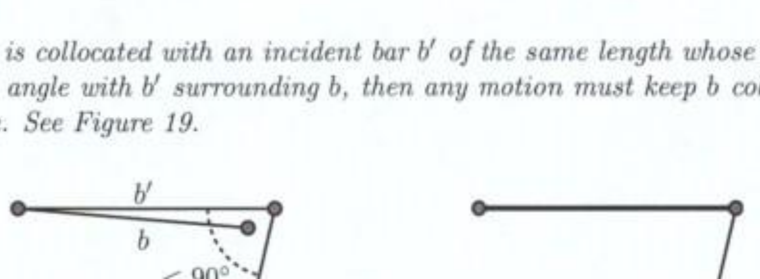


Figure 19: Rule 2 for simplifying self-touching configurations.

Proof: The noncrossing constraints at the endpoint of b surrounded by the convex angle formed by b' and b'' prevent b from moving relative to b' until the convex angle opens to $\geq 90°$, which can only happen after a positive amount of time. □

We can apply this rule twice, as shown in Figure 20, to further simplify the linkage.

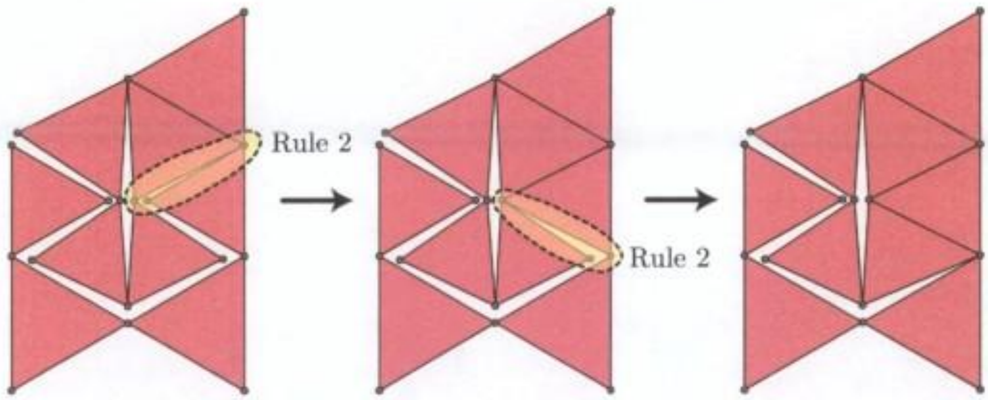


Figure 20: Applying Rule 2 twice to the configuration from Figure 18.

The final simplification comes from realizing that the central quadrangle gap between triangles is effectively a triangle because the right pair of edges are a rigid unit. Thus the gap forms a rigid linkage (though it is not infinitesimally rigid, because a horizontal movement of the central vertex would maintain distances to the first order), so we can treat it as part of a large rigid block. Figure 21 shows a simplified drawing of this self-touching configuration, which is rigid if and only if the original self-touching configuration is rigid.

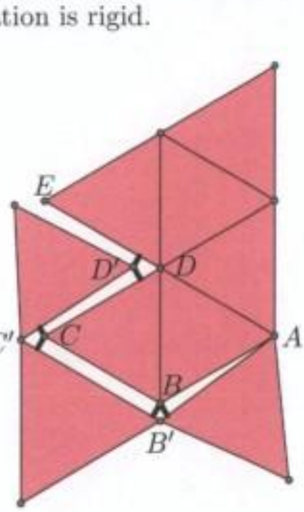


Figure 21: The simplified configuration from Figure 20.

6.4 Stress Argument

Finally we argue that the simplified configuration of Figure 21 is infinitesimally rigid using Lemma 10. The configuration is clearly infinitesimally rigid if B is pinned against B', C is pinned against C', and D is pinned against D'. It remains to construct a stress that is negative on all length-zero connections. The stress we construct is nonzero only on the edges connecting points with labels in Figure 21; we also set $\omega_{AD} = 0$.

We start by assigning the stresses incident to A. We choose $\omega_{AB} < 0$ arbitrarily, and set $\omega_{AB'} := -\omega_{AB} > 0$. A is now in equilibrium because these stress directions are parallel.

We symmetrically assign $\omega_{BC} := \omega_{AB} < 0$ and $\omega_{B'C'} = \omega_{A'B'} > 0$. The resulting forces on B and

B' are vertical. They can be balanced by an appropriate choice of the stresses $\omega_{B,B'A} = \omega_{B,B'C'} < 0$, which, taken together, also point in the vertical direction.

Vertex D' has exactly three incident stresses—$\omega_{C'D'}$, $\omega_{D',DC}$, and $\omega_{D',DE}$—which do not lie in a halfplane. Thus there is an equilibrium assignment to these stresses, unique up to scaling, and the stresses all have the same sign. Because zero-length connections must be negative, we are forced to make all three of these stresses negative. We also choose this scale factor to be substantially smaller than the stresses that have been assigned so far.

By assigning $\omega_{CD} = -\omega_{C'D'}$, we establish equilibrium at vertex D as well: the forces at D are the same as at D', only with reversed signs.

Vertex C feels two stresses assigned so far—$\omega_{CD} > 0$ and $\omega_{BC} < 0$. By the choice of scale factors, the latter force dominates, leaving us with a negative force in the direction close to CB, and two stresses $\omega_{C,C'B'}$ and $\omega_{C,C'D'}$ which can be used to balance this force. The three directions do not lie in a halfplane. Therefore $\omega_{C,C'B'}$ and $\omega_{C,C'D'}$ can be assigned negative stresses.

Finally, vertex C' is also in equilibrium because $\omega_{B'C'} = -\omega_{BC}$, $\omega_{C'D'} = -\omega_{CD}$, and the stress from the zero-length connections are the same as for C but in the opposite direction.

In summary, we have shown the existence of a stress that is positive on all zero-length connections. By Lemma 10, the self-touching configuration is infinitesimally rigid, so by Lemma 9, the configuration is rigid. By the simplification arguments above, the original self-touching configuration is also rigid. By Theorem 8, the original self-touching configuration is strongly locked, so sufficiently perturbations are locked.

We remark that an argument similar to the one above, using an assignment of stresses, can also be used for proving Rules 1 and 2, with an appropriate modification of Lemma 10; however, the direct argument that we have given is simpler.

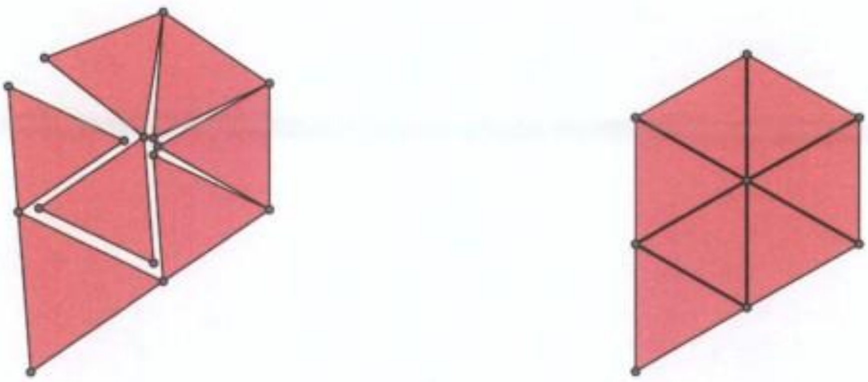
The argument relied on the isosceles triangles having an apex angle of $< 90°$ (but no more) in order to guarantee that particular triples of stress directions are or are not in a halfplane. It also relies on the symmetry of the configuration through a vertical line (excluding the triangle in the upper right). Thus the argument generalizes to all isosceles triangles sharper than $90°$.

6.5 Locked Equilateral Triangles

Figure 22 shows another, simpler example of a locked chain of equilateral triangles, using just seven triangles instead of nine. However, this example cannot be stretched into a locked chain of triangles with an arbitrary apex angle of $< 90°$, as in Figure 15.

To prove that this example is locked, we first apply Rule 1 and then Rule 2, as shown in Figure 23. Unlike the previous example, the resulting simplified configuration is not infinitesimally rigid (the middle vertex can move infinitesimally horizontally), so we cannot use a stress argument. In this case, however, we can use a more direct argument to prove rigidity of the simplified configuration (and thus of the original self-touching configuration).

Let ℓ denote the side length of the triangles in any of the self-touching configurations. Consider the two dashed chains connecting vertices A and B in the simplified configuration. The left chain of two bars forces the distance between A and B to be at most 2ℓ, with equality as in the original configuration only if the angle between the two bars remains straight. The right chain of three bars can only open its angles, because of the three triangles on the inside, so the right chain acts as a *Cauchy arm*. The Cauchy-Steinitz Arm Lemma (see, e.g., [Conn82] or [SZ67]) proves that the endpoints of such a chain can only get farther away from each other. Thus the distance between A and B is at least 2ℓ, with equality only if the angles in the right chain do not change. These upper and lower bounds of 2ℓ on the distance between A and B force the bounds to hold with equality,



(a) (b)

Figure 22: A locked chain of seven equilateral triangles. (a) Drawn loosely. Separations should be smaller than they appear. (b) Drawn tightly, with no separation, as a self-touching configuration.

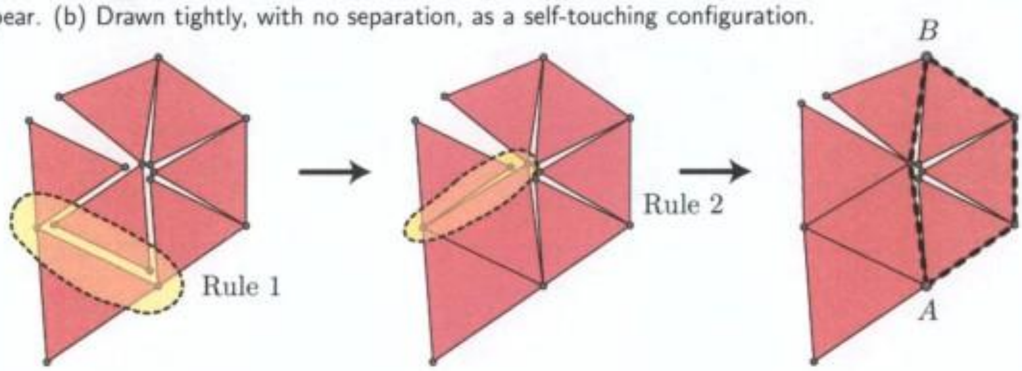


Figure 23: Applying Rules 1 and 2 to the chain of seven equilateral triangles from Figure 22.

which prevents any angles from changing except possibly for the angles at A and B. However, it is impossible to change fewer than four angles of a closed chain such as the one formed by the left and right dashed chains. (This simple fact was also proved by Cauchy [Cro97].) Therefore, the configuration is rigid.

Applying Theorem 8, we obtain that the self-touching configuration is strongly locked:

Theorem 13 *The self-touching chain of seven equilateral triangles shown in Figure 22(b) is rigid and thus strongly locked. Therefore, any sufficiently small non-self-touching perturbation, similar to the one shown in Figure 22(a), is locked.*

Appendix

Here we describe a construction of a closed chain, a convex parallelogram, with slender adornments attached, where each adornment together with its base is convex, such that the configuration space has infinitely many components.

We attach a single obtuse triangle as a slender adornment to the top base segment, as with Figure 10. If the bottom segment is fixed the path of the bottom vertex in the upper adornment traces out a circle, which is shown as a dashed circular arc C in Figure 24(a).

The second slender adornment is attached to the bottom segment and is the convex hull of infinitely many points, each slightly above C. The points form an infinite sequence $p_1, p_2, \ldots$

converging to a point on the right p_∞, and they are chosen so the straight line interval from p_i to p_{i+1} intersects the lower portion of C (the open circular disk determined by C). An exaggerated picture of this construction is in Figure 24(b). Thus, the upper slender adornment intersects the lower adornment and misses it alternately infinitely often.

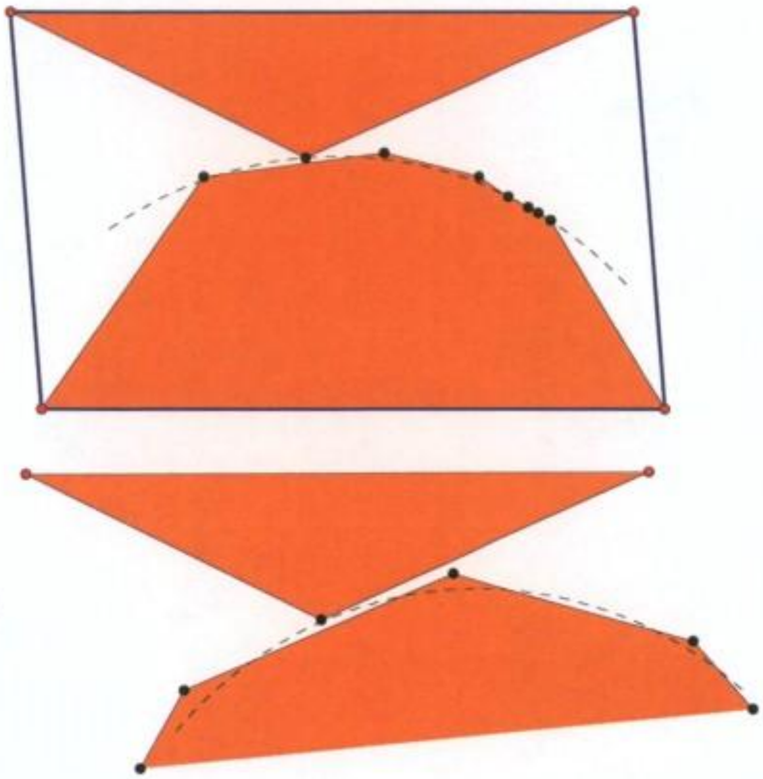


Figure 24: Figure (a) shows the overall set-up of a parallelogram with two convex slender adornments attached such that the configuration space has infinitely many components. Figure (b) is an exaggerated close-up of where the two adornments are close.

References

[AKRW04] Helmut Alt, Christian Knauer, Günter Rote, and Sue Whitesides. On the complexity of the linkage reconfiguration problem. In J. Pach, editor, *Towards a Theory of Geometric Graphs*, volume 342 of *Contemporary Mathematics*, pages 1–14. American Mathematical Society, 2004.

[AN98] Jin Akiyama and Gisaku Nakamura. Dudeney dissection of polygons. In *Revised Papers from the Japan Conference on Discrete and Computational Geometry*, volume 1763 of *Lecture Notes in Computer Science*, pages 14–29, Tokyo, Japan, December 1998.

[Ale84] Ralph Alexander. The circumdisk and its relation to a theorem of Kirszbraun and Valentine. *Math. Mag.*, 57 (1984), no. 3, pages 1165–169.

[CDIO04] Jason H. Cantarella, Erik D. Demaine, Hayley N. Iben, and James F. O'Brien. An energy-driven approach to linkage unfolding. In *Proceedings of the 20th Annual ACM Symposium on Computational Geometry*, pages 134–143, Brooklyn, New York, June 2004.

[BC02] Károly Bezdek and Robert Connelly. Pushing disks apart–the Kneser-Poulsen conjecture in the plane. In *J. Reine Angew. Math.* 553 (2002), pages 221–236.

[Conn82] Robert Connelly. Rigidity and energy. *Invent. Math.* 66 (1982), no. 1, 11–33..

[CDR02] Robert Connelly, Erik D. Demaine, and Günter Rote. Infinitesimally locked self-touching linkages with applications to locked trees. In J. Calvo, K. Millett, and E. Rawdon, editors, *Physical Knots: Knotting, Linking, and Folding of Geometric Objects in 3-space*, pages 287–311. American Mathematical Society, 2002.

[CDR03] Robert Connelly, Erik D. Demaine, and Günter Rote. Straightening polygonal arcs and convexifying polygonal cycles. *Discrete & Computational Geometry*, 30(2):205–239, September 2003.

[Cro97] Peter R. Cromwell. Equality, rigidity, and flexibility. In *Polyhedra*, chapter 6, pages 219–247. Cambridge University Press, 1997.

[Csi01] B. Csikós. On the volume of flowers in space forms. *Geom. Dedic.* 1-3 (2001), no. 86, pages 57–59.

[CvdSG+] Jae-Sook Cheong, A. Frank van der Stappen, Ken Goldberg, Mark H. Overmars, and Elon Rimon. Immobilizing hinged polygons. *International Journal on Computational Geometry and Applications*. To appear.

[DDE+05] Erik D. Demaine, Martin L. Demaine, David Eppstein, Greg N. Frederickson, and Erich Friedman. Hinged dissection of polyominoes and polyforms. *Computational Geometry: Theory and Applications*, 31(3):237–262, June 2005.

[DDLS05] Erik D. Demaine, Martin L. Demaine, Jeffrey F. Lindy, and Diane L. Souvaine. Hinged dissection of polypolyhedra. In *Proceedings of the 9th Workshop on Algorithms and Data Structures*, volume 3608 of *Lecture Notes in Computer Science*, pages 205–217, Waterloo, Canada, August 2005.

[Dud02] Henry E. Dudeney. Puzzles and prizes. *Weekly Dispatch*, 1902. The puzzle appeared in the April 6 issue of this column. An unusual discussion followed on April 20, and the solution appeared on May 4.

[Epp01] David Eppstein. Hinged kite mirror dissection. arXiv:cs.CG/0106032, June 2001. http://arXiv.org/abs/cs.CG/0106032.

[Fre02] Greg N. Frederickson. *Hinged Dissections: Swinging & Twisting*. Cambridge University Press, August 2002.

[GM92] Y. Gordon and M. Meyer. On the volume of unions and intersections of balls in Euclidean space. *Geometric aspects of functional analysis(Israel, 1992-1994)*. pages 91–101. *Oper. Theory Adv. Appl., 77, Birkhäuser, Basel, 1995.*

[MTW+02] Chengde Mao, Venkat R. Thallidi, Daniel B. Wolfe, Sue Whitesides, and George M. Whitesides. Dissections: Self-assembled aggregates that spontaneously reconfigure their structures when their environment changes. *Journal of the American Chemical Society*, 124:14508–14509, 2002.

[Kir34] M. D. Kirszbraun. Über die zusammenziehenden und Lipschitzchen Transformationen. In *Fund. Math.*, 22 (1934), pages 77–108.

[Rei79] John H. Reif. Complexity of the mover's problem and generalizations. In *Proceedings of the 20th Annual IEEE Symposium on Foundations of Computer Science*, pages 421–427, 1979.

[SZ67] I. J. Schoenberg and S. K. Zaremba On Cauchy's Lemma Concerning Convex Polygons. Canad. J. of Math. 19, 1062-1071 (1967)

[Str00] Ileana Streinu. A combinatorial approach to planar non-colliding robot arm motion planning. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, pages 443–453, Redondo Beach, California, November 2000.

[Str05] Ileana Streinu. Pseudo-triangulations, rigidity and motion planning. Discrete Comput. Geom. 34 (2005), no. 4, pages 587–635.

Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT•

ELSEVIER

Computational Geometry 26 (2003) 37–45

Computational Geometry
Theory and Applications

www.elsevier.com/locate/comgeo

Interlocked open and closed linkages with few joints

Erik D. Demaine [a], Stefan Langerman [b,*], Joseph O'Rourke [c,1], Jack Snoeyink [d,2]

[a] *MIT Laboratory for Computer Science, 200 Technology Square, Cambridge, MA 02139, USA*
[b] *School of Computer Science, McGill University, 3480 University Street, Suite 318, Montreal, QC, H3A 2A7, Canada*
[c] *Department of Computer Science, Smith College, Northampton, MA 01063, USA*
[d] *Department of Computer Science, UNC Chapel Hill, Chapel Hill, NC 27599-3175, USA*

Received 13 November 2001; received in revised form 13 September 2002; accepted 15 September 2002

Communicated by T. Biedl

Abstract

We study collections of linkages in 3-space that are *interlocked* in the sense that the linkages cannot be separated without one bar crossing through another. We explore pairs of linkages, one open chain and one closed chain, each with a small number of joints, and determine which can be interlocked. In particular, we show that a triangle and an open 4-chain can interlock, a quadrilateral and an open 3-chain can interlock, but a triangle and an open 3-chain cannot interlock.

Keywords: Linkages in S^3; Knots; Configurations

1. Introduction

Consider a simple polygonal chain, either an open *arc* or a closed *polygon*, that is embedded in 3-space. We view the vertices of the chain (except the endpoints of an open chain) as universal *joints*, and the edges of the chain as rigid *bars*. We call a chain with k bars a k-*chain*. A *motion* of the chain is a motion of the vertices that preserves the length of the bars, and never causes bars to cross. In particular, a *straightening of an open chain* is a motion that makes all joint angles become 180°. We say that a collection of disjoint, simple chains can be *separated* if, for any distance d, there is a motion whose

* Corresponding author.
E-mail addresses: edemaine@mit.edu (E.D. Demaine), sl@cgm.cs.mcgill.ca (S. Langerman), orourke@cs.smith.edu (J. O'Rourke), snoeyink@cs.unc.edu (J. Snoeyink).

[1] Supported by NSF Distinguished Teaching Scholars award DUE-0123154.

[2] Partially supported by NSF grants 9988742 and 0076984.

0925-7721/03/$ – see front matter
doi:10.1016/S0925-7721(02)00171-2

result is that every pair of points on different chains has distance at least d. If a collection cannot be separated, we say that its chains are *interlocked*. If a single chain cannot be straightened, we say that it is *locked*.

It is known that a single, open chain in 3-space, having as few as 5 bars, can be locked [1,2]. Other classes of chains are known to be unlocked, but the complexity of deciding whether a given chain can be unlocked is not known. One decision procedure applies the roadmap algorithm for general motion planning [3,4], which runs in exponential time.

Our work is inspired by a question posed by Anna Lubiw [5]: Into how many pieces must a chain be cut so that the pieces can be separated and straightened? This problem is motivated by protein molecules, which can be modeled by polygonal chains, and, according to some theories, temporarily split apart in order to reach the minimum-energy folding.

We can observe easy upper and lower bounds for Lubiw's problem: some n-chains require cutting at least $\lfloor (n-1)/4 \rfloor$ vertices for separation, and no chain requires cutting of more than $\lfloor (n-1)/2 \rfloor$ vertices. The lower bound is obtained by concatenating many copies of the 5-bar "knitting needles" example from [1,2], each sharing one bar with the next as in Fig. 1. Observe that each copy of the locked 5-bar chain must have one of its four interior vertices cut. The upper bound is obtained by cutting every second joint of a chain, and observing that the resulting 2-bar pieces ("hairpins") can be rigidly separated arbitrarily far by dilating from a point, because the pieces are starshaped sets. This separation motion dates back at least to de Bruijn in 1954 [6], where he used it to prove separability of convex objects; the same motion was shown to apply to the more general situation of starshaped objects by Dawson in 1984 [7], and the algorithmic side of this result is described by Toussaint in 1985 [8]. See also [9].

While Lubiw's problem motivated our original interest in interlocked open chains, we explore here interlocking for combinations of open and closed chains. In the next section, we resolve how many bars are needed by each chain in order to obtain an interlocked pair, as summarized in Table 1.

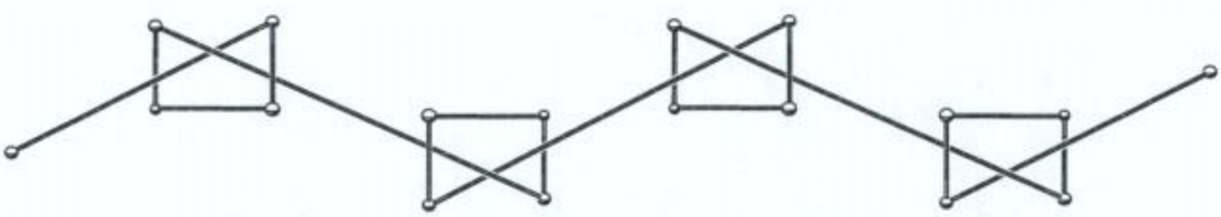


Fig. 1. An $n = 17$ bar chain that requires cutting at least $\lfloor (n-1)/4 \rfloor = 4$ vertices to separate.

Table 1
Our results on when an open chain and a closed chain can interlock. A claim that a k-chain can interlock holds also for any l-chain with $l > k$, and a claim that a k-chain cannot interlock holds also for any l-chain with $l < k$

Sec	Chain 1	Chain 2	Result
2	closed triangle	open 3-chain	Cannot Interlock
3.1	closed triangle	open 4-chain	Can Interlock
3.2	closed quadrilateral	open 3-chain	Can Interlock

2. Triangle and 3-chain cannot interlock

We begin by showing that a triangle and a 3-chain cannot interlock. As we will see later, this is in some sense a maximal non-interlocking configuration.

Theorem 1. *An open* 3*-chain cannot interlock with a triangle.*

Proof. We follow this notation: $\triangle abc$ lies in plane H, and the 3-chain C has vertices (p_0, p_1, p_2, p_3) and bars (l_0, l_1, l_2). First assume C is not planar; otherwise, make C nonplanar by a small motion. Let L_i be the support line of l_i and define points $q_i = L_i \cap H$.

(1) Bar l_1 intersects the closed $\triangle abc$. In this case, it is possible to move bar l_0 and bar l_2 within the plane that it forms with l_1 so that the angle at the joint shared with l_1 is arbitrarily close to either 0 or π, because one of the two wedges spanned by these two motions does not intersect any other edge. Once both end bars have been moved to that position, C is arbitrarily close to a single bar which can be translated in the direction $\overrightarrow{p_1 p_2}$.
(2) Bar l_1 does not intersect the closed $\triangle abc$. Because configuration C is non-self-intersecting, we can assume that the points $\{q_0, p_1, p_2, q_2\}$ do not lie on a common plane, or equivalently $\{q_0, q_1, q_2\}$ are not collinear. Denote the line containing q_0 and q_2 by $Q_{0,2}$, as in Fig. 2. In fact, for any position of l_1 such that $(L_1 \cap H) \notin Q_{0,2}$, the lines containing $q_0 p_1$ and $p_2 q_2$ do not intersect, and do not intersect the edges of $\triangle abc$. Thus the motion that translates l_1 in a direction orthogonal to $Q_{0,2}$ and parallel to H, away from $\triangle abc$, while maintaining L_0 and L_2 through the original points q_0 and q_2, will avoid self-intersection.[3] $\square$

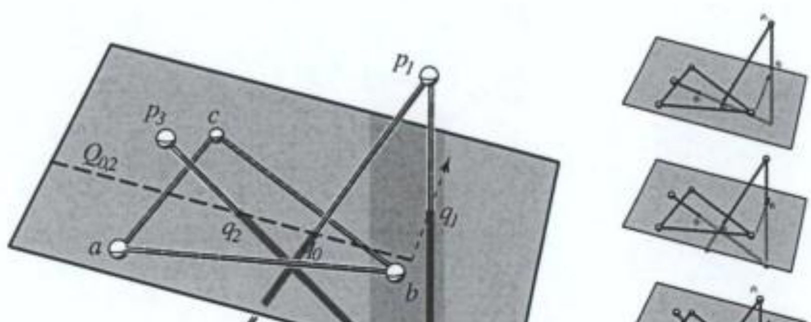


Fig. 2. Translate l_1 so that the point $q_1 = L_1 \cap H$ moves away from $Q_{0,2}$. Keep the points q_0 and q_2 fixed in H, so that the lines L_0 and L_2 pivot about $q_0 = L_0 \cap H$ and $q_2 = L_2 \cap H$ as l_1 moves. This separates the 3-chain from $\triangle abc$.

[3] See http://www.cs.smith.edu/~orourke/Interlocked/ for an animation of this motion.

3. Interlocked examples and the topological method

Our two proofs that chains are interlocked follow a similar structure in what we call the *topological method*. We imagine tying the two ends of the open chain with a long rope near infinity, which defines a topological *link* (multicomponent knot) [10, p. 17]. For the two chains to separate, they must form the trivial link (referred to as 0_1^2; see later). First we show that before this happens, the ends of the open chain must get close to the closed chain. Second we argue that this proximity is impossible before changing the topology of the link. Finally we prove that this circularity leads to a contradiction, so the chains are interlocked.

To make connections to known mathematics for links, we will refer to some links by their numbers from standard tables. See [10, p. 287] or [11, p. 1086]. Tables of links are often organized by (minimum) crossing number. The superscript in the link notation is the number of components, for us always 2. The subscript is an arbitrary table index. See Fig. 3.[4]

3.1. Triangle and 4-chain

We begin with the configuration illustrated in Fig. 4.

Theorem 2. *A triangle can interlock with a 4-chain.*

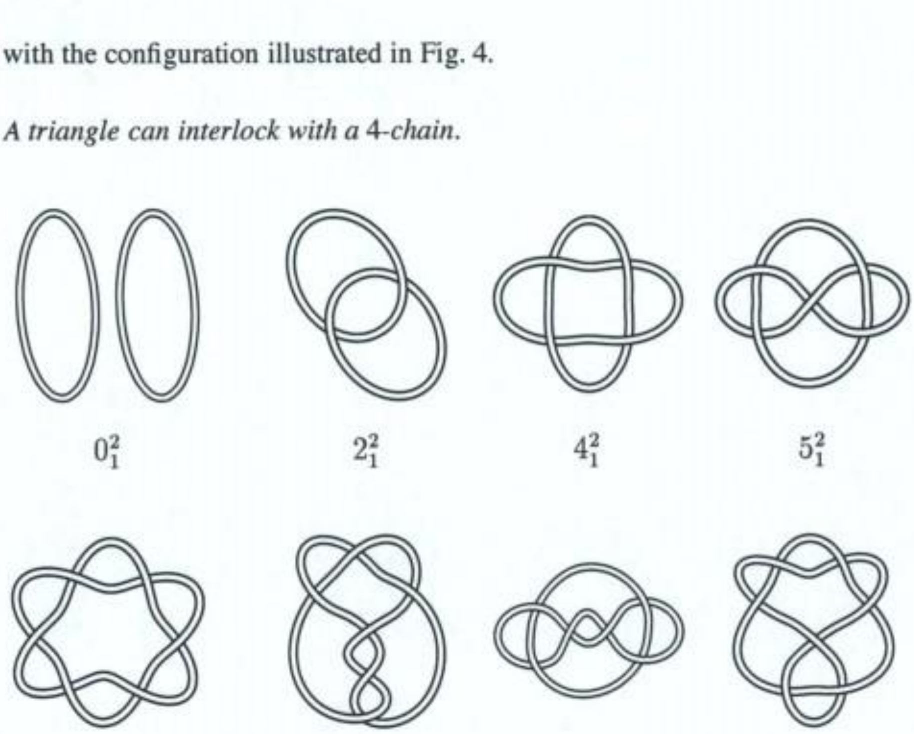


Fig. 3. The first few two-component links.

[4] Link images produced by Robert Scharein's `knotplot` program http://www.cs.ubc.ca/nest/imager/contributions/scharein/KnotPlot.html.

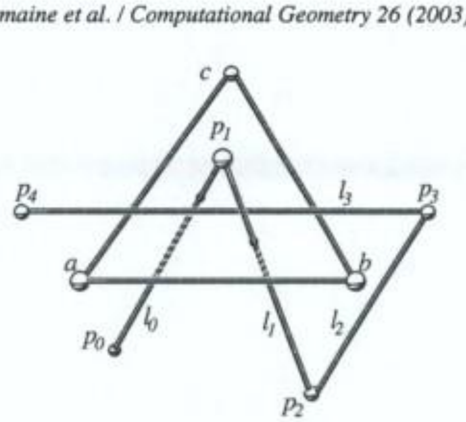


Fig. 4. A triangle and a 4-chain can lock.

Proof. We choose the following notation for the configuration of Fig. 4: A triangle abc lies in a plane H, with H^+ the halfspace above and H^- the halfspace below H. Let the circumcircle of $\triangle abc$ have center o, and radius r.

The 4-chain alternates points and bars $p_0, l_0, p_1, l_1, \dots, l_3, p_4$ with the following placements: p_0 is in H^-, bar l_0 crosses the interior of $\triangle abc$, and ends at a point p_1 above o. Bar l_1 crosses the interior of $\triangle abc$ again, so $p_2 \in H^-$. Bar l_2 crosses H outside of $\triangle abc$, and l_3 crosses the wedge formed by l_0 and l_1 above H. So $\{p_0, p_2\} \subset H^-$ and $\{p_1, p_3, p_4\} \subset H^+$.

Let R be the real number $r + |l_1| + |l_2|$, and set the length of l_0 and l_3 to $20R$. Consider the open ball B of radius $15R$, and the ball B' of radius $4R$, both centered at o. Initially, p_0 and p_4 lie outside of B, while a, b, c, p_1, p_2 and p_3 all lie inside $B' \subset B$. As long as p_0 and p_4 stay outside B and all other vertices stay inside B, we can attach a sufficiently long unknotted string between p_0 and p_4 that remains outside B, and thus is never crossed by any of the bars, and our configuration is equivalent to the link 5_1^2. The non-interlocked configuration corresponds to two separable unknots 0_1^2, so any motion separating this configuration would require p_0 or p_4 to enter the ball B or p_1, p_2 or p_3 to leave B.

Consider the first event when any p_i, $i = 0, \dots, 4$, touches the boundary of B. Then before or at that event, points p_1, p_2 and p_3 must be out of B' but still inside B: When p_0 touches B, point p_1 must be

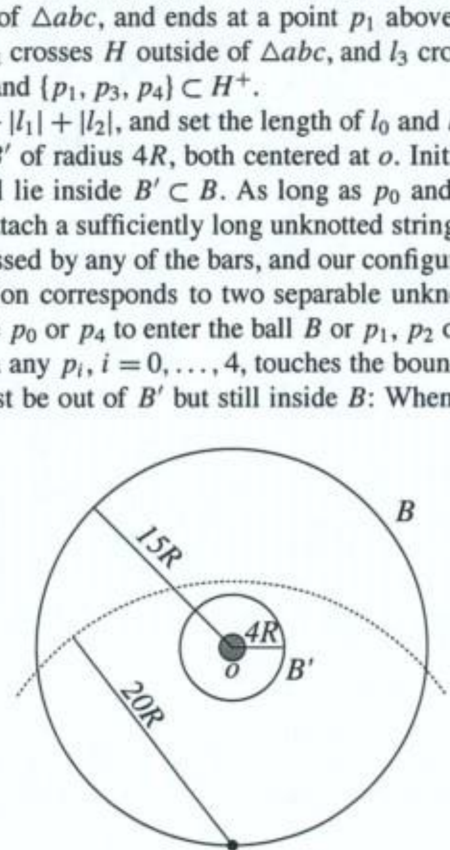


Fig. 5. When p_0 touches B, point p_1, p_2 and p_3 must be exterior to B'.

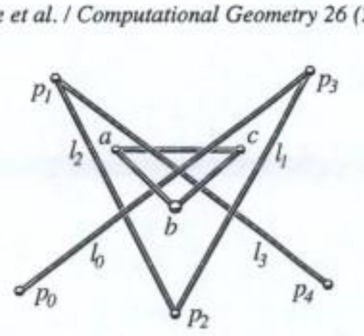


Fig. 6. This configuration is incompatible with the fact that p_0 or p_4 touches the boundary of B.

exterior to B' by at least R, and therefore p_2 and p_3 are also exterior to B'. See Fig. 5. The same applies for when p_4 touches the boundary of B. When any one of p_1, p_2 or p_3 touches the boundary, the other two are at least at a distance $14R$ from o and so are outside of B'. Since we consider the first such event, there must be an instant before that when all three points are outside B' but still inside B.

At this time, the only elements possibly inside B', besides $\triangle abc$, are the two bars l_0 and l_3. Then either one of l_0 and l_3 crosses the interior of $\triangle abc$, or both do, or neither do. The first case corresponds to a link 2^2_1 and the third case to two separable unknots 0^2_1; neither of these are equivalent to our starting configuration (in the knot-theoretical sense). Since the rope and the bars have not crossed, the topology of the configuration cannot have changed and so these cases lead to a contradiction.

The case in which both l_0 and l_3 cross $\triangle abc$ requires a careful analysis. Because end vertices p_0 and p_4 are still outside of the open ball B, we can replace the string joining them by a great arc γ on the boundary of B. Let T be the plane parallel to l_0 and l_3, and passing through o. Consider the orthogonal projection of the 4-bar linkage onto T. Note that in the projection, the lengths of bars l_0 and l_3 are preserved, and all other segment lengths are at most their original lengths. Let q_0 be the intersection of l_0 and plane H. The triangle $\triangle abc$ is contained in a ball of radius $2R$ centered at q_0, and joints p_1, p_2 and p_3 lie in a ball of radius R centered at p_1. Since p_1 is outside B' and q_0 is inside the circumcircle of $\triangle abc$, the distance between those two points is larger than $3R$, and that distance is preserved in the projection. Thus, the projections of the two balls are disjoint and we can separate the projections of p_1, p_2 and p_3 from the projections of p_0, p_4 and $\triangle abc$ by a line (this separation is necessary to exclude cases such as the one shown in Fig. 6), and the two bars l_1 and l_2 can be replaced by a single bar joining p_1 and p_3 without changing the topology of the link. By enumerating all possible above/below combinations for the crossings in that projection, we can infer that configuration is equivalent to 0^2_1, which is two separated, unknotted links, or to 4^2_1, which is shown in Fig. 7. But neither of these are topologically equivalent to our starting configuration, so this first event could never happen.

Note that a similar argument can be used to show that the chains in Fig. 6 are interlocked as well. □

3.2. *Quadrilateral and* 3*-chain*

In the following, we will use what is known as the *linking number* of a two component link. We first arbitrarily orient both components of the link. Then each crossing drawn in the projection of the link has one of two types, associated with a value $+1$ or -1. See Fig. 8.

The *linking number* of the link is half the sum of the values of all crossings between the different components; crossings of a component with itself are not counted. For example, the link 5^2_1 has 5 crossings, but only four of them involve both components. The sum of the values of the four crossings

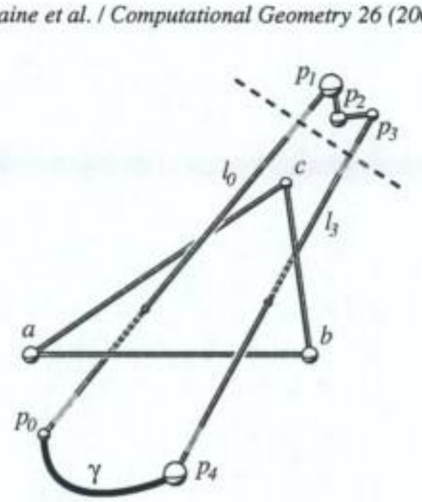


Fig. 7. The link 4_1^2, formed when bars l_0 and l_3 both pass through the interior of $\triangle abc$. (Not to scale; gray segments indicate omissions.) Joints $\{p_1, p_2, p_3\}$ can be separated from $\{a, b, c, p_0, p_4\}$.

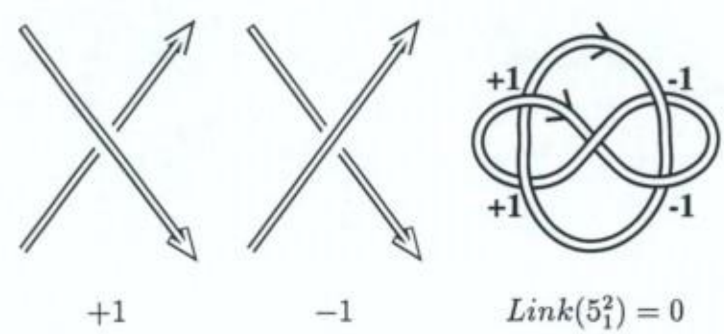


Fig. 8. Sign of a crossing.

is 0, which yields a linking number of 0. Note that if the orientation of one of the components is reversed, then the linking number is negated. It can be proved using some elementary knot theory that the linking number of an oriented link is an *invariant*, that is, it has the same value for all drawings of the oriented link [10, p. 21].

Theorem 3. *A* 4*-gon can interlock with a* 3*-chain.*

Proof. Let the 4-gon be $abcd$, and again use (l_0, l_1, l_2) and (p_0, p_1, p_2, p_3) to represent the bars and vertices of the 3-chain. Starting with the configuration of Fig. 9, let $R = |ab| + |bc| + |cd| + |l_1|$ and set the length of l_0 and l_2 to $20R$. Consider the open ball B of radius $15R$, the ball B' of radius $4R$, and the ball B'' of radius R, all three centered at a. As in the previous proof, we connect p_0 to p_3 by a string exterior to B. The resulting link is now 6_1^2. We again argue that in order to separate the 4-gon from the 3-chain, p_0 or p_3 has to enter the ball B or p_1 or p_2 have to leave B. Before that, there must be an instant when p_0 and p_3 are still outside B, p_1 and p_2 are still inside B but out of B', and the only elements possibly inside B', besides $abcd$, are the two edges l_0 and l_2.

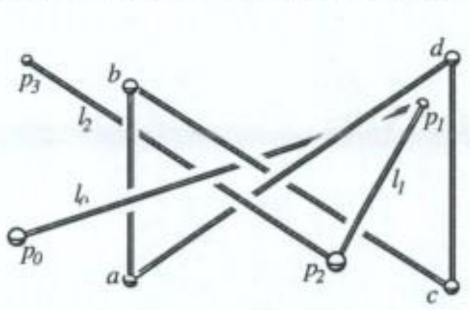


Fig. 9. A quadrilateral and a 3-chain can lock.

If neither l_0 nor l_2 intersects B'', then the configuration is the link 0_1^2, contradicting that the topology cannot have changed. If one of the two end bars, say l_0, intersects B'', let q_0 be a point of $l_0 \cap B''$. We project the configuration onto a plane parallel to l_0 and l_2, preserving the distances along those two bars. As in the previous proof, because the length of the segment $q_0 p_1$ is preserved in the projection, only the interiors of l_0 and l_2 can intersect the projection of B''. This implies that the linking number of the configuration will be the sum of the values induced by l_0 and $abcd$, and the values induced by l_2 and $abcd$, divided by 2. Notice that the total of the values induced by a straight edge and a 4-gon is at most 2, and so the linking number of the configuration is at most $(2+2)/2 = 2$. But the linking number of 6_1^2 is 3. Because the linking number is an invariant, the topology of the configuration must have changed, a contradiction. □

4. Open problems

Many open problems remain in the context of interlocking pairs of open chains, which have close connections to the motivating problem of Lubiw. For each value of i, what is the smallest j for which an i-chain can interlock with a j-chain?

The topological method of Theorems 2 and 3, where we used a "rope" to close one open chain to form a topological linkage, does not easily extend to pairs of open chains. Two ropes would be needed, and their potential interactions would need to be controlled. To extend this work, therefore, we will be investigating a geometric method that establishes a collection of geometric facts and shows that there can be no first violation. We believe that we can use such a method to establish three conjectures: that a 3-chain can interlock with a 4-chain, that three 3-chains can interlock, but that two 3-chains cannot interlock even in the presence of any finite number of 2-chains.

The proof of Theorem 3 depends upon a tetrahedron formed by the 4-gon, and does not show that a 3-chain and a k-gon can interlock for any $k > 4$. In fact, adding any small edge to the 4-gon would allow the 3-chain to escape. On the other hand, our conjecture that a 3-chain can interlock with a 4-chain, once established, would imply that a 3-chain can interlock with a k-gon for any $k \geqslant 5$ by connecting the endpoints of the 4-chain with one or more edges.

Chains that model physical objects, such as robot arms or protein backbones, often have restrictions placed on the motion of a joint. There are a number of interesting problems for open and closed chains under various restrictions on motions. For example, we conjecture that a rigid, open 3-chain can interlock with a flexible, open 3-chain.

Acknowledgements

This work was initiated in the Waterloo algorithmic open-problem session, and continued at CCCG 2000, with contributions by Therese Biedl, Hamish Carr, Eowyn Čenek, Timothy Chan, Beenish Chaudry, Martin Demaine, Rudolf Fleischer, John Iacono, Anna Lubiw, Dessislava Michaylova, Veronica Morales, Katherine Sinclair, Geetika Tewari and Ming-wei Wang. The authors wish to thank the anonymous referees for many useful comments and for suggesting a more elegant proof for Theorem 1.

References

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

[2] T. Biedl, E. Demaine, M. Demaine, S. Lazard, A. Lubiw, J. O'Rourke, M. Overmars, S. Robbins, I. Streinu, G. Toussaint, S. Whitesides, Locked and unlocked polygonal chains in 3D, Discrete Comput. Geom. 26 (2) (2001) 269–282.

[3] J. Canny, A new algebraic method for robot motion planning and real geometry, in: Proceedings of the 28th Annual Symposium on Foundations of Computer Science, Los Angeles, California, 1987, pp. 39–48.

[4] J. Canny, Some algebraic and geometric computations in PSPACE, in: Proceedings of the 20th Annual ACM Symposium on Theory of Computing, Chicago, Illinois, 1988, pp. 460–469, http://www.acm.org/pubs/citations/proceedings/stoc/6221%2/p460-canny/.

[5] E.D. Demaine, J. O'Rourke, Open problems from CCCG'99, in: Proc. 12th Canad. Conf. Comput. Geom., 2000, pp. 269–272.

[6] N.G. de Bruijn, Nieuw Archief voor Wiskunde 2 (1954) 67, problems 17 and 18. Answers in Wiskundige Opgaven met de Oplossingen 20 (1955) 19–20.

[7] R.J. Dawson, On removing a ball without disturbing the others, Math. Mag. 57 (1984) 27–30.

[8] G.T. Toussaint, Movable separability of sets, in: G.T. Toussaint (Ed.), Computational Geometry, North-Holland, Amsterdam, 1985, pp. 335–375.

[9] J. Snoeyink, J. Stolfi, Objects that cannot be taken apart with two hands, Discrete Comput. Geom. 12 (1994) 367–384.

[10] C.C. Adams, The Knot Book, W.H. Freeman, New York, 1994.

[11] E.W. Weisstein, CRC Concise Encyclopedia of Mathematics, CRC Press LLC, 1999.

Interlocked Open Linkages with Few Joints*

Erik D. Demaine
MIT LCS
Cambridge, MA
edemaine@mit.edu

Stefan Langerman
School of Comp. Sci.
McGill Univ.
Montreal, Quebec
sl@cgm.cs.mcgill.ca

Joseph O'Rourke
Dept. Comp. Sci.
Smith College
Northampton, MA
orourke@cs.smith.edu

Jack Snoeyink
Dept. Comp. Sci.
UNC Chapel Hill
Chapel Hill, NC
snoeyink@cs.unc.edu

ABSTRACT

We advance the study of collections of open linkages in 3-space that may be *interlocked* in the sense that the linkages cannot be separated without one bar crossing through another. We consider chains of bars connected with rigid joints, revolute joints, or universal joints and explore the smallest number of chains and bars needed to achieve interlock. Whereas previous work used topological invariants that applied to single or to closed chains, this work relies on geometric invariants and concentrates on open chains.

Categories and Subject Descriptors

F.2.2 [**Analysis of Algorithms**]: Nonnumerical Algorithms—*Geometrical problems and computations*

General Terms

Theory

Keywords

Linkages, Knots, Geometry, Configurations, Robotic Arms, Protein Models

1. INTRODUCTION

Consider a simple polygonal chain that is embedded in 3-space with disjoint, straight-line edges, which we think of as fixed-length *bars*. We call a chain with k bars a *k-chain*. The $k+1$ vertices of a k-chain are the two end points, adjacent to the end bars, and $k-1$ internal vertices, or *joints*. We can place restrictions that each joint be *rigid*, permitting no relative motion between its two incident bars, or be *revolute*, a term that we will consistently use for a rotational joint that preserves the angle between its two incident bars, or be *flexible*, serving as a universal joint that allows any rotation.

A *motion* of a chain is a motion of the vertices that preserves the length of the bars, respects the restrictions on joints, and never causes nonadjacent bars to touch. We say that a collection of disjoint, simple chains can be *separated* if, for any distance d, there is a motion whose result is that every pair of points on different chains has distance at least d. If a collection cannot be separated, we say that its chains are *interlocked*.

In this paper, we characterize collections of open chains with small numbers of bars that can interlock. Our results on pairs of chains, summarized in Table 1, explore when it is possible for an open k-chain to interlock with an open m-chain. A result that an open k-chain can interlock with an m-chain also implies that open or closed l-chain, with $l > k$, can interlock with an m-chain, and a result that no open k-chain can interlock with an m-chain also implies that no open l-chain with $l < k$ can interlock with an m-chain.

In addition, we show that

- Two flexible 3-chains with any finite number of flexible 2-chains cannot interlock, but three flexible 3-chains can interlock.
- A flexible 4-chain with any finite number of flexible 2-chains cannot interlock, but a flexible 3-chain and 4-chain can interlock.

We prove results on separability of chains in Section 2, and on interlocked chains in Section 3. Our proofs assume general position, namely that no nonincident bars are coplanar and no three joints collinear. Since we can enforce general position by a small perturbation, this assumption can be made without loss of generality. We list some remaining open problems in Section 4.

Previous work has considered motions of single chains and of closed chains. A *straightening* of a flexible chain is a motion that makes all joint angles become 180°. If a single chain cannot be straightened, we say that it is *locked*. It is known that a single, open chain in 3-space, having as few as 5 bars, can be locked [4, 1]. In a companion paper [7], we showed examples with open and closed chains that were interlocked, including an open 3-chain with a quadrilateral and an open 4-chain with a triangle. In these previous works it was possible to (conceptually) close an open chain by adding a piece of rope, then argue that geometric properties kept the rope from interfering with any motion, and that topological invariants demonstrated that the resulting closed links were interlocked. However, this approach does not ex-

*The second author was partially supported by grants from MITACS, FCAR and CRM, the third by NSF Distinguished Teaching Scholars award DUE-0123154, and the fourth by NSF grants 9988742 and 0076984.

SoCG'02, June 5–7, 2002, Barcelona, Spain.

		2-chain		3-chain			4-chain			5-chain
		flexible	rigid	flexible	revolute	rigid	flexible	revolute	rigid	rigid
2-chain	flexible	$-$	$-$	$-$	$-$	$-^{5}$	$-^{2}$	$-^{5}$	$-^{5}$	$+^{15}$
	rigid	$-$	$-$	$-$	$-^{4}$	$+^{12}$	$+^{14}$	$+^{14}$	$+^{14}$	$+$
3-chain	flexible	$-$	$-$	$-^{1}$	$-^{6}$	$+^{18}$	$+^{11}$	$+$	$+$	$+$
	revolute	$-$	$-^{4}$	$-^{6}$	$+^{21}$	$+$	$+$	$+$	$+$	$+$
	rigid	$-^{5}$	$+^{12}$	$+^{18}$	$+$	$+$	$+$	$+$	$+$	$+$
4-chain	flexible	$-^{2}$	$+^{14}$	$+^{11}$	$+$	$+$	$+$	$+$	$+$	$+$
	revolute	$-^{5}$	$+^{14}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
	rigid	$-^{5}$	$+^{14}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
5-chain	rigid	$+^{15}$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$

Table 1: Our results on interlocking pairs of open chains. $(+)$ = can, $(-)$ = cannot interlock. In superscript is the number of the theorem proving the result, the other entries are implied.

tend: we cannot simply close two or more open chains with ropes because the ropes may interfere with one another. Instead we establish geometric invariants, typically about the convex hull of joints and the relations of the end bars, often by considering convenient projections of the linkage. We emphasize the different proof techniques used within each section.

One of the inspirations for our work was a question posed by Anna Lubiw [6]: into how many pieces must a chain be cut so that the pieces can be separated and straightened? This question is motivated by proteins, which may, according to some theories, temporarily split apart in order to reach the minimum-energy folding. Our results on open flexible chains, along with the locked 5-chain of [4, 1], imply that a set of chains can always be separated and every chain straightened if the total number of middle bars is less than three. If the end bars are long enough, there are interlocked configurations whenever the number of middle bars is at least three. Soss [8] investigated revolute chains[1], also motivated by proteins, and created a "staple and hook" example of an interlocked revolute 3-chain and 4-chain. We have an interlocked example with two revolute 3-chains.

The complexity of deciding whether a given chain can be unlocked is not known. One decision procedure applies the roadmap algorithm for general motion planning [2, 3], which runs in polynomial space but exponential time. Because all of our results are for a few chains, each of a few joints, the roadmap algorithm could in principle establish interlock for our examples, but couldn't discover them and probably wouldn't give insight into their structure. On the other hand, the separability proofs apply to general classes of sets of chains, rather than the specific instances handled by the algorithm.

2. SEPARABLE CHAINS

In this section, we prove that certain configurations are separable by extending a scaling idea (whose earliest reference we know is de Bruijn [5]) and other arguments to find a separating motion. Except for a couple of cases involving a flexible 2-chain, the theorems in this section are tight in the sense that, for the chains considered, any additional bars or further restrictions on the motion can allow an interlocked configuration.

2.1 Two 3-flexible chain+many 2-flexible cannot interlock

We show that two 3-chains (even with added 2-chains) never form an interlocked configuration.

THEOREM 1. *Two open, flexible 3-chains and any finite number of flexible 2-chains can always be separated.*

PROOF. Consider two 3-chains C_1 and C_2, and especially their middle bars, k_1 and k_2. By our general position assumption, non-adjacent bars are not coplanar. Let K be a plane between, and parallel to, the middle links k_1 and k_2. We may choose the coordinate system such that K is the yz plane. If necessary, apply another small perturbation to ensure that no two vertices have the same x coordinates except for the vertices of k_1 and of k_2.

Now, consider the affine transformation $x \rightarrow \alpha x$ for any real $\alpha \geq 1$. Note that this is a non-uniform scaling that increases all distances between pairs of points with different x-coordinates. Thus, it preserves the lengths of k_1 and k_2, and increases the length of all the other edges.

Create a motion parameterized by time $t \geq 1$ by placing the chains according to the transform for $\alpha = t$, and truncating the edges at both ends of each chain to preserve the lengths. Because affine transformations preserve incidence relationships among lines, the motion cannot cause any bars to touch. As t becomes large, the chains separate arbitrarily far, so they are not interlocked. □

We can prove a similar theorem for an open 4-chain and 2-chains.

THEOREM 2. *An open, flexible 4-chain and any finite number of flexible 2-chains can always be separated.*

PROOF. As in the proof of Theorem 1, rotate the configuration so that the three joints of the 4-chain are parallel to

[1]Here, and throughout this paper, a revolute joint is one that preserves the angle between the adjacent bars, which is called an "edge spin" in [9] and a "dihedral motion" in [8]. In some areas "revolute" is used for the larger class of pin joints whose axes need not align with one of the bars; we use only the restricted definition.

the yz plane, and apply the affine transformation $x \rightarrow \alpha x$ for any real $\alpha \geq 1$ to increase the distance between all vertices except the joints of the 4-chain. Each end bar can be truncated to obtain a separating motion. □

A corollary improves the bound for a problem posed by Lubiw, and first addressed in [7].

COROLLARY 3. *Given a n-chain, it is always possible to cut $\lfloor (n-3)/2 \rfloor$ vertices so that the pieces obtained can be separated and straightened.*

PROOF. Cut the 4th joint, then cut every other joint to obtain one 4-chain and many 2-chains. □

The next three subsections establish theorems on pairs of chains with restricted motions.

2.2 2-rigid+3-revolute cannot interlock

THEOREM 4. *A rigid 2-chain and a revolute 3-chain cannot interlock.*

PROOF. Consider the rigid 2-chain $P = (p_0, p_1, p_2)$ and the revolute 3-chain $R = (r_0, r_1, r_2, r_3)$. The general position assumption ensures that no two non-adjacent edges are coplanar. Let H be the plane containing P. Then R intersects H in at most three points: let r'_i be the intersection between $r_i r_{i+1}$ and H, if it exists.

The two lines containing p_0p_1 and p_1p_2 divide H into 4 quadrants $Q_1, \ldots, Q_4$. If quadrant Q_1, Q_2, or Q_3 contains no intersection point r'_i, then P can be separated by a translation in H: if Q_1 is empty, we translate P in the direction $\overrightarrow{p_1p_2}$, if Q_2 is empty, we translate P in the direction $\overrightarrow{p_2p_1}$, and if Q_3 is empty, we translate P in the direction $\overrightarrow{p_0p_1}$. Otherwise, we may assume $r'_{i_1} \in Q_1$, $r'_{i_2} \in Q_2$ and $r'_{i_3} \in Q_3$.

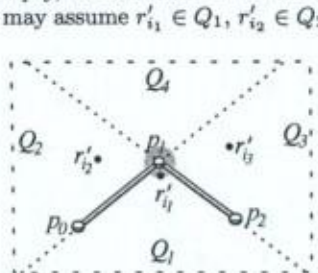


Figure 1: 2-chain P in its plane H.

Now, translate P in H so that joint p_1 is within a distance ε of r'_{i_1} as shown in Fig. 1, where $\varepsilon > 0$ is a small value to be chosen later. This can be done without intersections. If the segment $r'_{i_2}r'_{i_3}$ does not intersect P, then we can rotate P counterclockwise about r'_{i_1} until quadrant Q_2 (which is changing shape as P rotates) becomes empty—then translate P in the direction p_2p_1. There remains the case in which segment $r'_{i_2}r'_{i_3}$ intersects P. We analyze two subcases: either $i_1 = 1$ and r'_1 is in Q_1, or $i_1 \neq 1$.

If $r'_1 \in Q_1$, suppose that the middle bar of R is fixed. Then the end bar r_0r_1 can move in a cone with apex r_1 and axis r_1r_2 passing through r'_1. If ε was chosen small enough, this cone intersects H in a curve (a conic section) that connects point r'_0 to some point in quadrant Q_4 without intersecting Q_1. Bar r_0r_1 can rotate until it reaches the ray from r'_1 through $r'_{i_3} \in Q_3$ without intersecting bar r_2r_3, so we can rotate r'_0 into Q_4, then can separate P by a translation in H.

For the last case, we assume without loss of generality that $r'_1 \in Q_2$, $r'_0 \in Q_1$ and $r'_2 \in Q_3$. Then, for any $\delta > 0$, we can choose ε small enough so that P can be translated to be at distance at most δ from r_1 without crossings. Because the vertex angles at r_1 and r_2 are fixed, we can choose δ small enough in order to rotate r_2r_3 without crossings to bring it arbitrarily close to r_0r_1. Then, for some small values of δ and then ε, the cone describing the motions of p_1p_2 when p_0p_1 is fixed does not intersect r_1r_2, and we can move p_1p_2 until we fall into one of the previous cases. □

2.3 2-flexible+4-rigid cannot interlock

When the 2-chain is flexible, the extra degree of freedom allows it to escape in its plane from any chain that intersects the plane in at most four points.

THEOREM 5. *A flexible 2-chain and a rigid 3-chain, 4-chain, or closed 5-chain cannot interlock.*

PROOF. As in the previous theorem, let the 2-chain $P = (p_0, p_1, p_2)$ define a plane H and four quadrants, $Q_1, \ldots, Q_4$. Consider the at most four points where the other chain R intersects H. If one of the quadrants Q_j, for $j \in \{1, 2, 3\}$, does not contain at least one intersection point, then we can separate P from R by translation in H.

We could move point p_1 along $\overrightarrow{p_2p_1}$, allowing p_0p_1 to rotate if it reaches any point in Q_2, unless and until p_1 approaches a ray $\rho_{12} = \overrightarrow{r_1r_2}$, with $r_1 \in R \cap Q_1$ and $r_2 \in R \cap Q_2$. We could then move p_1 along ray ρ_{12}, until p_1 approaches a ray $\rho_{13} = \overrightarrow{r'_1r_3}$, with $r'_1 \in R \cap Q_1$ and $r_3 \in R \cap Q_3$. If these motions do not separate the chains, then we have found two rays that cross in Q_4. This implies that $r_1 \neq r'_1$ and we know the quadrants of all four points of $R \cap H$. We can now straighten the 2-chain P by a motion in H that preserves the ray/chain intersection points $r_{12} \cap P$ and $r_{13} \cap P$. Then we can separate P from R by translation. □

2.4 3-flexible+3-revolute cannot interlock

THEOREM 6. *A flexible 3-chain and a revolute 3-chain cannot interlock.*

PROOF. Let $P = (p_0, \ldots, p_3)$ denote the flexible 3-chain and $R = (r_0, \ldots, r_3)$ denote the revolute 3-chain. Consider the projection of the two chains from the viewpoint p_1 onto a sphere. All three bars of R and p_2p_3 project to segments of great arcs of angle $< \pi$, and p_0p_1 and p_1p_2 project to points. Thus p_0p_1 can be moved arbitrarily close to r_1r_2 unless its projection is enclosed in a triangle formed by r_0r_1, r_2r_3 and p_2p_3. But then, looking at the projection from viewpoint p_2 instead, p_2p_3 can be moved arbitrarily close to r_1r_2. Once one of the end bars of P is moved close to r_1r_2, the second end bar can be moved close to r_1r_2 as well, and they can then both be moved close to the midpoint of r_1r_2.

So we have reached a configuration where both p_0p_1 and p_2p_3 are at a distance at most ε from the midpoint r'_1 of r_1r_2 for some appropriate $\varepsilon > 0$. Let H be the plane containing p_1, p_2 and r'_1, and project P onto H in the direction r_1r_2. For any given $\delta > 0$, we can choose the value of ε so that for any bar ab intersecting H at a distance $> \delta$ from any point of the projection of P, that segment does not touch any bar of P.

Let r'_i be the intersection of r_ir_{i+1} with H. If we fix the position of r_1r_2, the possible positions of r_0r_1 and r_2r_3 intersect H in two curves (conic sections). Both these curves

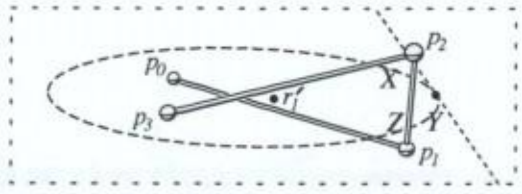


Figure 2: Possible positions for r_0' on components of the dotted ellipse in H.

are cut into pieces by the projection of P. Those pieces will be called *components* for r_0' or r_2'.

We will describe several motions of the chain P where p_1p_2 will remain in H and will be translated in some specified direction, while the support lines of p_0p_1 and p_2p_3 will slide around r_1' and remain within a distance ε of that point. We will call any such motion *feasible* if there exists a simultaneous motion of R, with r_1r_2 fixed, that introduces no crossings. This motion will not introduce crossings between P and itself, or between P and r_1r_2. Also, r_0r_1 and r_2r_3 only intersect if the rays $\overrightarrow{r_1'r_0'}$ and $\overrightarrow{r_1'r_2'}$ are equal, so we will have to preserve the radial ordering of r_0' and r_2' with respect to r_1' during the motion. The last kind of possible crossings would be between the end bars of R and the chain P. For those, we observe the possible movements of those end bars, which correspond to the *components* for r_0' or r_2'. If the component for r_0' or r_2' is unbounded (e.g. the end of a parabola), then the corresponding bar can be moved to stop intersecting H, which can only help moving P away. If the component for r_0' or r_2' is bounded but never dissapears during the entire motion, the corresponding bar can be continuously moved within that component to avoid crossings with P. So, if r_0' and r_2' are contained in unbounded components, or in bounded components that never disappear during the entire motion, then the motion is feasible. Conversely, the only way for a motion not to be feasible is when either r_0' or r_2' is contained in a bounded component that disappears. Because the curves are convex, and r_1' is inside their convex hull, the disappearance of a component during a motion of the kind described above must involve p_1p_2.

Fig. 2 denotes by X, Y and Z the three kinds of components that could disappear. Since we have only two points to place in those components, at least one of X, Y or Z contains neither r_0' nor r_2', and perhaps does not exist. If X is empty or non-existent, then we can translate p_1p_2 in the direction p_2p_1. This translation does not reduce the size of Z until p_1p_2 stops bounding Z, and Y remains unchanged by the motion, and so the motion is feasible. If Z is empty or non-existent, then translating p_1p_2 in the direction p_1p_2 produces a feasible motion for the same reasons. If Y is non-existent for at least one of the two curves, then X and Z are the same component for that curve and we fall into the previous case. Finally, suppose Y exists and is empty for both curves, and there is a non-empty X component and a non-empty Z component. Assume that X contains r_0' and Z contains r_2'. Then one of the two curves must be an ellipse; assume that it is the curve containing r_0'. We can translate p_1p_2 with r_0' along its component, away from r_1', until the Y component of r_0' disappears, connecting the X and Z components of r_0' and falling back into the previous case. □

3. INTERLOCKED CHAINS

To show that two or more chains are interlocked we establish geometric invariants, often regarding the convex hull of selected vertices or joints. We begin with some useful preliminaries. We use a bracket $[abcd]$ to denote the 4×4 orientation determinant of the homogeneous coordinates of four points a, b, c, and d. It will be positive if the ray $\overrightarrow{ab}$ is consistently oriented with ray $\overrightarrow{cd}$ according to a right-hand rule.

Since we are concerned with invariants under motion, the points in a bracket will move over time. We can make statements about the invariance of faces of convex hulls like the following two lemmas; see Figure 3 for an illustration.

LEMMA 7. *Under continuous motion of a, b, c, and d, determinant $[abcd]$ is positive iff the convex hull $CH(a,b,c,d)$ is a tetrahedron with edges to a, b, and c appearing in counterclockwise (ccw) order around d.*

PROOF. This is a consequence of properties of the orientation determinant. □

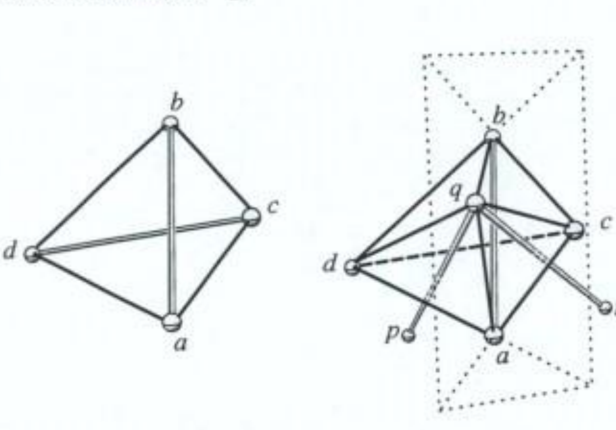


Figure 3: The configurations for Lemmas 7 and 8.

LEMMA 8. *Suppose, as depicted at the right of Figure 3, that the convex hull $CH(a,b,c,d,q)$ initially has six faces $\triangle qac$, $\triangle qcb$, $\triangle qbd$, $\triangle qda$, $\triangle adc$, and $\triangle bcd$. As long as three conditions hold under motion of a, b, c, and d—that $\triangle adc$ and $\triangle bcd$ are faces of convex hull $CH(q,a,b,c,d)$, that bar pq intersects $CH(a,b,c,d)$ with $[pqab] > 0$, and that qr intersects $CH(a,b,c,d)$ with $[qrab] > 0$—the convex hull $CH(a,b,c,d,q)$ retains its face structure. In particular, ab pierces $\triangle qcd$.*

PROOF. Since $\triangle adc$ and $\triangle bcd$ remain faces of the convex hull $CH(q,a,b,c,d)$, they remain faces of $CH(a,b,c,d)$, which must be a tetrahedron. By Lemma 7, $[abcd] > 0$.

We claim that q remains in the intersection of halfspaces bounded by planes through acd, bcd, abd, and acb. These planes are indicated by dotted lines at the right of Figure 3. If point q would exit this intersection by first reaching planes through acd or bcd, then a or b would no longer be a vertex of the convex hull $CH(q,a,b,c,d)$. If q first reached abd or acb, then pq or qr could no longer intersect the tetrahedron $abcd$ and maintain a positive orientation determinant with ab. (Note that reaching two or more planes simultaneously still violates the conditions.) Thus, q remains on the convex hull and keeps all its incident faces. □

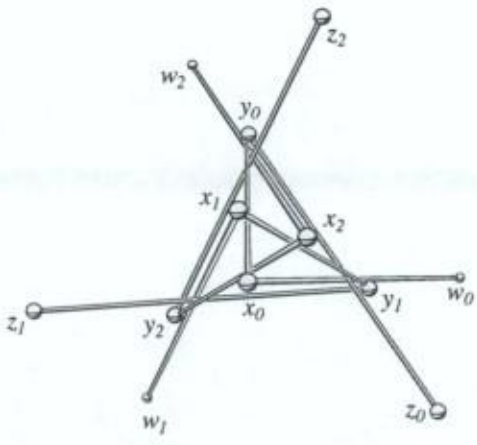


Figure 4: Three flexible 3-chains that interlock.

3.1 Three flexible 3-chains can interlock

In this section we show that the three open 3-chains of Fig. 4 interlock. We say that chain i, for $i \in \{0,1,2\}$ has vertices w_i, x_i, y_i, and z_i, as illustrated. We will use index arithmetic modulo 3.

To make this example, one could start with Borromean rings made of triangular chains with $w_i = z_i$, then extend the end bars of chain i above and below the surrounding chain $(i+1)$ until the end bars are at least three times longer than the middle bars. Let us assume that the middle bars have unit length.

THEOREM 9. *Three flexible 3-chains can interlock.*

PROOF. We can make a number of initial geometric observations, which we will show are geometric invariants of this linkage. When we say a segment pq *pierces a triangle* $\triangle abc$, it is a shorthand for saying that five brackets are positive: $[pabc]$, $[abcq]$, $[pqab]$, $[pqbc]$, and $[pqca]$—that is, points p and q are on opposite sides of the plane abc and $\triangle abc$ is oriented consistent with a right-hand rule around pq. We have the following for all $i \in \{0,1,2\}$:

(1) The convex hull of the joints $\mathcal{Q} = CH(\{x_j, y_j \mid 0 \le j \le 2\})$ is an octahedron with edges to x_{i+1}, y_{i-1}, y_{i+1} and x_{i-1} appearing counter-clockwise (ccw) around x_i and clockwise (cw) around y_i.

(2) Middle bar x_iy_i pierces $\triangle x_{i-1}y_{i-1}x_{i+1}$.

(3) End bar x_iw_i pierces $\triangle y_{i-1}x_{i-1}y_{i+1}$, forms positive determinants $[x_iw_ix_{i+1}y_{i+1}]$ and $[x_iw_iy_{i+1}z_{i+1}]$, and exits the hull $\mathcal{Q}$.

(4) End bar y_iz_i pierces $\triangle x_{i-1}y_{i-1}y_{i+1}$ (the same triangle with the opposite orientation), forms positive determinants $[y_iz_iz_{i+1}y_{i+1}]$ and $[y_iz_iy_{i+1}x_{i+1}]$, and exits the hull $\mathcal{Q}$.

As the points and vertices move, let us consider which of these conditions could fail first. We divide them into two classes: *hull conditions*, where a joint or end point goes inside the hull $\mathcal{Q}$ or a hull edge disappears as two adjacent faces become coplanar, and *piercing conditions*, where a bar fails to pierce its triangle or one of its orientation determinants becomes zero.

We begin by showing that the first change cannot be a joint disappearing inside the convex hull. Consider vertex x_i. Segment x_iw_i pierces $\triangle y_{i-1}x_{i-1}y_{i+1}$ and segment x_iy_i pierces $\triangle x_{i-1}y_{i-1}x_{i+1}$. Since both enter the tetrahedron formed by the middle bars $x_{i-1}y_{i-1}$ and $x_{i+1}y_{i+1}$, we can apply Lemma 8 to the 2-chain $w_ix_iy_i$ to see that joint x_i cannot be first joint to disappear inside the convex hull. Similarly, the two segments $x_iy_iz_i$ intersect the convex hull of the two middle bars such that we can apply Lemma 8 and show that joint y_i cannot be the first inside.

If a convex hull edge disappears, then two adjacent triangles become coplanar. By the pigeonhole principle, two of the vertices of that quadrilateral are from the same chain, which implies that a middle bar x_iy_i is on the convex hull. But as long as x_iy_i pierces its triangle, it cannot be on the convex hull. We show below that the triangle piercing is invariant.

First, however, we argue that end points w_i and z_i never enter the hull, by establishing that the hull diameter is less than three as long as 1) and 2) hold.

LEMMA 10. *If the diameter of the convex hull $\mathcal{Q}$ is ≥ 3, then either $\mathcal{Q}$ contains a joint, or a middle bar lies on the boundary of $\mathcal{Q}$.*

PROOF. If the hull diameter is three or more, introduce two planes perpendicular to the diameter segment that cut it into three equal pieces. These planes cut the hull $\mathcal{Q}$ into three pieces; by the pigeonhole principle, either the first or the third piece contains (the interior of) only one middle bar x_iy_i. If both joints of this bar are on the convex hull, then this bar lies on the hull because the defining plane separates these joints from the remaining. □

Thus, the first failures must be piercing conditions, possibly accompanied by an edge (but not a vertex) disappearing from the hull. Without loss of generality, we consider that among the first piercing conditions to fail is one for a bar on chain 1. In preparation for finding a contradiction, we draw the projections of relevant bars from the perspectives of joints y_1 and x_1 in Figure 5, just before any piercing condition fails.

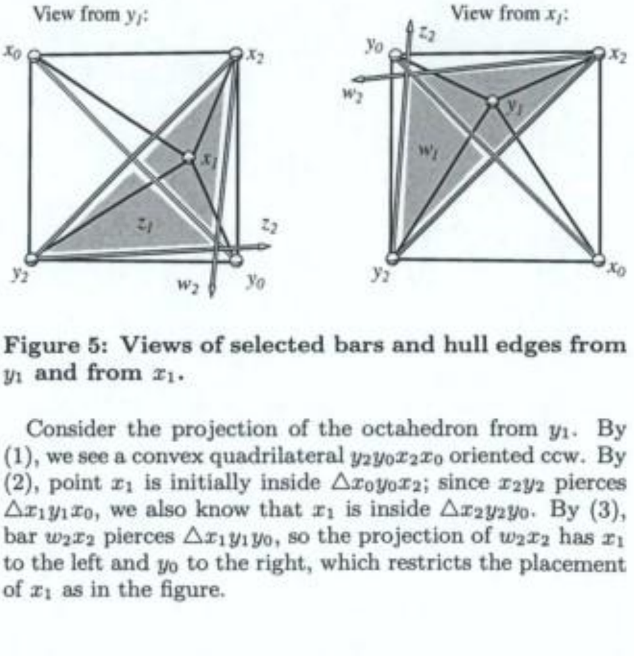


Figure 5: Views of selected bars and hull edges from y_1 and from x_1.

Consider the projection of the octahedron from y_1. By (1), we see a convex quadrilateral $y_2y_0x_2x_0$ oriented ccw. By (2), point x_1 is initially inside $\triangle x_0y_0x_2$; since x_2y_2 pierces $\triangle x_1y_1x_0$, we also know that x_1 is inside $\triangle x_2y_2y_0$. By (3), bar w_2x_2 pierces $\triangle x_1y_1y_0$, so the projection of w_2x_2 has x_1 to the left and y_0 to the right, which restricts the placement of x_1 as in the figure.

Now, suppose that the condition that x_1y_1 pierces $\triangle x_0y_0x_2$ is among the first to fail—that is, one or more of its five orientation determinants become zero. We show that each case contradicts a known property. (We do one case analysis in detail to as illustration.) We know that $[x_1x_0y_0x_2] > 0$ and $[x_0y_0x_2y_1] > 0$, since the triangle is strictly inside the convex hull and both vertices x_1 and y_1 are on the boundary of $\mathcal{Q}$. Thus, it follows from Lemma 7 that $[x_1y_1x_0y_0]$ can become zero only if bars x_1y_1 and x_0y_0 are touching. Bracket $[x_1y_1y_0x_2]$ can become zero only if the projection of x_2w_2 has moved to be disjoint from the projection of x_1y_0, meaning that the piercing condition for x_2w_2 has previously failed. Finally, $[x_1y_1x_2x_0]$ can become zero only if the condition that x_2y_2 pierces $\triangle x_1y_1x_0$ has previously failed. This establishes that the piercing condition for x_1y_1 cannot be among the first to fail.

We make a similar argument in the projection from y_1 for the piercing conditions for y_1z_1. By (4), point z_1 projects to the left of x_2y_2 and y_0x_0. Because y_2z_2 pierces $\triangle x_1y_1y_0$, the projection of y_2z_2 has y_0 to the right and x_1 to the left; the orientation determinant also says that, in projection, z_1 is to the left of y_2z_2. Thus, z_1 is restricted to the shaded region. Since y_1z_1 goes through the hull, Lemma 7 implies that the y_1z_1 will touch the bars x_0y_0, x_2y_2, or y_2z_2 if their corresponding brackets go to zero. Thus, z_1 can leave the shaded region only by touching a bar or by a previous failure of a piercing condition. Notice that the points in the shaded region satisfy all the conditions imposed upon y_1z_1 in (4).

The argument for x_1w_1 is similar and establishes that there can be no first failure of piercing conditions. This completes the proof of Theorem 9.

3.2 A 3-chain and 4-chain can interlock

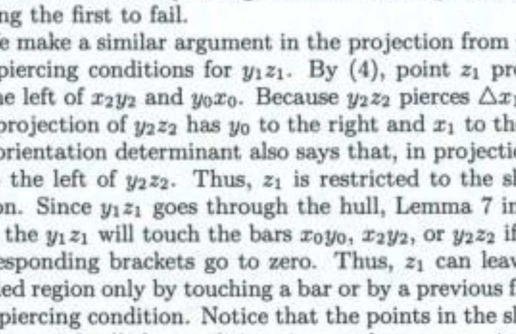


Figure 6: An example showing a locked 3-chain and 4-chain. Added lines show that the convex hull of joints is a bi-pyramid.

THEOREM 11. *Open flexible 3- and 4-chains can interlock.*

PROOF. Figure 6 depicts the core of two linked chains, $ABCDE$ and $wxyz$, where bars between joints have unit length and end bars have length greater than $BC + CD + xy = 3$. We analyze the convex hull of the flexible joints $\mathcal{Q} = CH(B, C, D, x, y)$ as the points move.

In the initial embedding of Figure 6, we make several observations that we will show are invariants. Recall that a statement that, for example, xy pierces $\triangle DCB$ is shorthand for saying that five orientation determinants are positive: $[xDCB]$, $[DCBy]$, $[xyDC]$, $[xyCB]$, and $[xyBD]$.

(1): Bar xy pierces $\triangle DCB$. Equivalently, the hull $\mathcal{Q}$ is a bi-pyramid, with edges to B, C, and D in ccw order around x and cw order around y.

(2): End bar DE pierces $\triangle Byx$ and hull face $\triangle BCx$ and makes positive orientation determinant $[DExw]$.

(3): End bar BA pierces $\triangle Cyx$ and hull face $\triangle DCy$ and makes positive determinants $[BAzy]$ and $[BAxw]$.

(4): End bar xw pierces $\triangle DCB$ and hull face $\triangle CBy$ and makes positive determinant $[wxyz]$.

(5): End bar yz pierces $\triangle BCD$ and hull face $\triangle BCx$.

Any motion that separates these chains must change the convex hull $\mathcal{Q}$ and invalidate observation (1), so some set of observations must be first to fail. We show by finding contradictions that none of these can be among the first, establishing that there is no separating motion. Unfortunately, this configuration has no symmetries to cut down on the number of cases.

To begin, we apply Lemma 8 to argue that the first event cannot include x or y vanishing inside the hull $\mathcal{Q}$. Consider x first. Since xw and xy pierce $\triangle DCB$, both bars intersect tetrahedron $CH(B, C, D, E)$. Since $[DEyx]$ and $[DExw]$ are positive, we can apply Lemma 8 to show that x cannot vanish into tetrahedron $CH(B, C, D, E)$ without some other hull change occurring. But vanishing into $CH(B, C, D, E)$ would be necessary before x could vanish into hull $\mathcal{Q}$. Similarly, yx and yz pierce $\triangle BCD$ and straddle BA, so Lemma 8 implies that the point y cannot vanish into the tetrahedron $CH(A, B, C, D)$ unless $\mathcal{Q}$ has already changed.

Next, we show that (1) cannot be among the first conditions to fail; that xy must remain inside the hull. Since we know that x and y remain on opposites sides of the plane BCD, we can most easily to argue about orientation determinants in 3D by considering projections from the perspectives of one of the joints, as illustrated in Figure 7. Consider

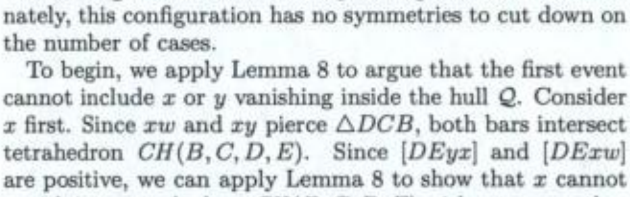


Figure 7: Projections of the linkage of Figure 6 from x, D, B, and y.

the view from x, where we see y inside a ccw-oriented triangle $\triangle BCD$. By condition (2), DE pierces $\triangle BCx$, and $[DEyx]$ is positive (from DE piercing $\triangle Byx$); these further restrict y to lie in a triangle formed by the projections of bars BC, CD and DE. By Lemma 7, the projection of y cannot reach the projections of BC or CD without causing bars to intersect. Nor can it reach reach BD without causing bars xy and DE to intersect unless there has been a previous failure of DE to pierce $\triangle BCx$, violating condition (2). Thus, condition (1) cannot be among the first condition to fail.

As long as the hull keeps its structure, we can make an argument like that of Lemma 10 to show that the diameter of the hull is at most three, which implies that end vertices never enter the hull. For end bar piercing conditions, therefore, we can continue to consider projections from joints, without worrying that a joint or end vertex will disappear inside the hull.

To see that condition (2) cannot be among the first to fail, consider the view from D, where we see a convex quadrilateral $xCyB$ whose diagonals are bars that restrict the point that is the projection of DE. By (4) and (1), bars xw and xy pierce $\triangle DCB$, so there is a triangle formed by projections of bars xw, xy, and BC that contains the projection of E. For this point to leave the projection of $\triangle Byx$ or $\triangle BCx$ or change the sign of $[DExw]$, bar DE would intersect bars xw, xy, or BC inside $\mathcal{Q}$, or the condition of (4) that xw pierces $\triangle DCB$ would have previously failed.

For condition (3), we have a similar case in the view from B. If the projection of A were to leave the projection of $\triangle Cyx$ or $\triangle DCy$, bar BA would intersect bar xy, yz, or CD, or there would have been a previous failure of condition (5), that yz pierces $\triangle BCD$.

For the piercing conditions of (4), it is sufficient to establish that xw always pierces $\triangle CBy$, because as long as it is satisfied and (1) xy pierces $\triangle DCB$, we automatically have xw piercing $\triangle DCB$. We must also establish that $[wxyz] > 0$ as points move. Consider once again the view from x. Bar xw projects to a point in a region bounded by the projections of CB, zy, BA, and DE as long as (4) bar xw pierces $\triangle CBy$, (5) bar yz pierces $\triangle BCx$ and satisfies $[wxyz] > 0$, (3) bar BA pierces $\triangle Cyx$, and (2) $[DExw] > 0$. Since xw cannot intersect bars CB, zy, BA, or DE, for the projection of w to leave $\triangle CBy$ or cause $[wxyz]$ to become negative, a piercing condition from (5) or (3) must have previously failed. Thus condition (4) cannot be among the first to fail.

For (5), consider the view from y. As long as xy pierces $\triangle DCB$, bar yz piercing $\triangle BCx$ is the more restrictive condition. Bar yz projects to a point in a triangle bounded by projections of AB, CB, and xw, since (3) AB pierces $\triangle Cyx$ and (4) xw pierces $\triangle CBy$. (In this case, we cannot use the condition $[DExw] > 0$, since the projection of point E could lie inside $\triangle CBx$.) Since yz cannot intersect bars AB, BC, or xw, the only way to leave $\triangle BCx$ would be after a previous failure of piercing conditions from (3) or (4).

Since no event can occur among the first events, we know that any motion will preserve the triangles of the convex hull $\mathcal{Q}$, and that the chains remain interlocked. $\square$

3.3 2-rigid + 3-rigid can interlock

The remaining subsections investigate interlocking configurations with restricted motion.

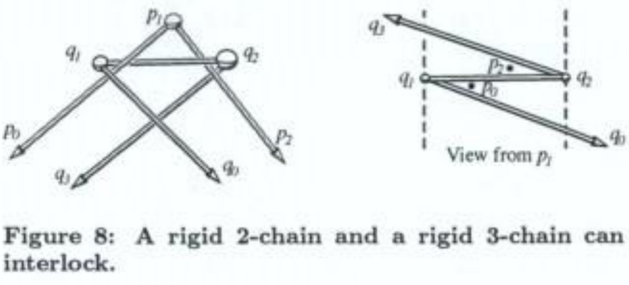


Figure 8: A rigid 2-chain and a rigid 3-chain can interlock.

THEOREM 12. *A rigid 2-chain can interlock with a rigid 3-chain.*

PROOF. The starting configuration is as shown in Fig. 8. For the two chains $P = (p_0, p_1, p_2)$, and $Q = (q_0, q_1, q_2, q_3)$, we assume that point $q_1 = (0,0,0)$, point $q_2 = (1,0,0)$, bar q_0q_1 goes through the point $q_0' = (1,-1,-1)$, bar q_2q_3 goes through the point $q_3' = (0,1,-1)$, and all end bars have length L. The vertex angle at p_1 is $\pi/2 < \beta < \pi$. Draw a central projection of the configuration onto the xy plane from viewpoint p_1, as in Figure 8.

In the starting configuration, both bars of P intersect $\mathcal{T} = CH(q_0', q_1, q_2, q_3')$, and during any separating motion, those bars must cease intersecting $\mathcal{T}$. The diameter of $\mathcal{T}$ is less than 3, so if $L > 3/\tan\beta$, we know that if p_0 or p_2 enter $\mathcal{T}$, then one of the end bars of P will have already left $\mathcal{T}$. So during any separating motion, one of p_0 or p_2 will have to cross one of the dotted lines in the projection shown in Figure 8. Note that before the motion starts, the dot product of the planar vectors in the projection $p_0p_2 \cdot q_1q_2 > 0$, and as soon as one of p_0 or p_2 intersects one of the dotted lines in the projection, $p_0p_2 \cdot q_1q_2 < 0$. Since this is a continuous motion, at some instant we have $p_0p_2 \cdot q_1q_2 = 0$; that is, the plane containing 2-chain P is perpendicular to q_1q_2.

Consider the intersections of Q with the plane containing P at that instant. The intersection with q_1q_2 is at $(y,z) = (0,0)$, the intersection with q_0q_1 lies on the segment joining $(0,0)$ to $(-1,-1)$, and the intersection with q_2q_3 lies on the segment joining $(0,0)$ to $(1,-1)$. Thus, the support line of p_0p_1 would have to be below $(-1,-1)$ and above $(0,0)$ and the support line of p_1p_2 would have to be above $(0,0)$ and below $(1,-1)$. But this would imply that $\beta < \pi$ which contradicts the fact that P is rigid. $\square$

3.4 2-rigid + 4-flexible can interlock

Consider the rigid 2-chain $P = (p_0, p_1, p_2)$ and flexible 4-chain $Q = (q_0, \dots, q_4)$ shown in Fig. 9. The lengths of the internal edges $k_1 = q_1q_2$, and $k_2 = q_2q_3$ are unity, and the length of all end bars is set to some large value L to be determined later. Let $\mathcal{T}$ be the tetrahedron with vertices $\{p_1, q_1, q_2, q_3\}$. We show:

LEMMA 13. *Starting from the configuration portrayed at the left of Fig. 9, consider any motion where none of the vertices p_0, p_2, q_0 or q_4 ever enter the tetrahedron $\mathcal{T}$. Then at all times, the edges p_0p_1 and p_1p_2 both intersect triangle $q_1q_2q_3$.*

PROOF. Along with the conclusion stated in the lemma, we will show that a few other conditions remain true at all times during the motion:

$$\begin{array}{ll} [q_0q_1q_2q_3] < 0, & [q_1q_2q_3q_4] > 0 \\ [p_0p_1q_0q_1] < 0, & [p_0p_1q_iq_{i+1}] > 0 \text{ for } i = 1,2,3 \\ [p_1p_2q_3q_4] > 0, & [p_2p_1q_iq_{i+1}] > 0 \text{ for } i = 0,1,2 \end{array}$$

and the edges p_0p_1 and p_1p_2 intersect triangle $\triangle q_1q_2q_3$, q_0q_1 intersects $\triangle p_1q_2q_3$ and q_3q_4 intersects $\triangle p_1q_1q_2$. We prove this by showing that none of these conditions can be the first one to become false. Note that these conditions also imply that p_1 remains above the plane containing $\triangle q_1q_2q_3$.

First consider all determinants involving p_0p_1 or p_1p_2. For this, we project the configuration from the viewpoint p_1 onto the $\triangle q_1q_2q_3$ as in the middle of Fig. 9. Let q_0' be the intersection of the edge q_0q_1 and the triangle $\triangle p_1q_2q_3$, q_0' projects to the intersection point of the projections of the edges q_0q_1 and q_2q_3. Likewise, let q_4' be the intersection of the edge q_3q_4 and the triangle $\triangle p_1q_1q_2$; point q_4' projects to the intersection point of the projections of the edges q_3q_4 and q_1q_2. Also, let r be the projection of the intersection between the projections of the edges q_0q_1 and q_3q_4; point r is the projection of points on those two edges that lie inside $\mathcal{T}$.

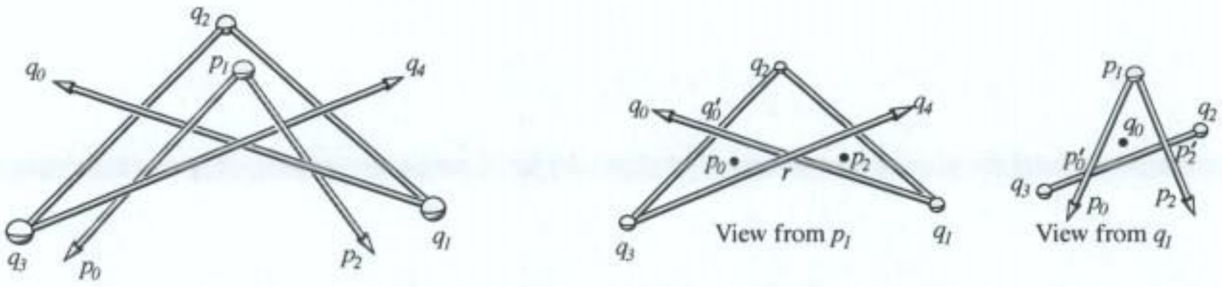


Figure 9: A rigid 2-chain P and a flexible 4-chain Q, with views from vertices p_1 and q_1.

In the projection, p_0 becomes a point lying inside the triangle $\triangle rq_0'q_3$. The three edges of this triangle are the projection of portions of edges completely contained in $\mathcal{T}$, and p_0 is not contained in $\mathcal{T}$, so none of the determinants involving p_0p_1 can change sign as the first violated condition without involving an edge crossing. The same argument can be made about edge p_1p_2 and triangle $\triangle rq_4'q_1$. The same projection also shows that the edges p_0p_1 and p_1p_2 will not stop intersecting triangle $\triangle q_1q_2q_3$ before some determinant involving one of these two edges changes sign.

For the events involving q_0q_1, we project the configuration from the viewpoint q_1 onto the $\triangle p_1q_2q_3$ as in the right of Fig. 9. Let p_0' be the intersection of p_0p_1 and $\triangle q_1q_2q_3$, p_0' projects to the intersection point of the projections of the edges p_0p_1 and q_2q_3. Let p_2' be the intersection of p_1p_2 and $\triangle q_1q_2q_3$, p_2' projects to the intersection point of the projections of the edges p_1p_2 and q_2q_3. In the projection, q_0 becomes a point lying inside the triangle $\triangle p_0'p_1p_2'$. The three edges of this triangle are the projection of portions of edges completely contained in $\mathcal{T}$, and q_0 is not contained in $\mathcal{T}$, so none of the determinants involving q_0q_1 can change sign as the first violated condition without involving an edge crossing. The same projection also shows that q_0q_1 will not stop intersecting triangle $\triangle p_1q_2q_3$ before some determinant involving q_0q_1 changes sign. The events involving q_3q_4 can be treated in the same manner, and so none of the events can occur first. $\square$

THEOREM 14. *Given any angle $0 < \beta < \pi$, there is an interlocked configuration of a 2-chain with a 4-chain, if the vertex angle of the 2-chain is restricted to stay $\geq \beta$ during the entire motion.*

PROOF. Consider the configuration shown at the left of Fig. 9. We show that the length L of all 4 end bars can be made large enough so that the configuration is interlocked. By the previous lemma, in order to unlock P and Q, point p_0 or p_2 must enter tetrahedron $\mathcal{T}$ through $\triangle q_1q_2q_3$. At the time one of these endpoints, say p_0, enters $\triangle q_1q_2q_3$, p_1p_2 still intersects $\triangle q_1q_2q_3$. But the closest point to p_0 on p_1p_2 is at distance $L\sin\beta$. Since the diameter of the triangle $\triangle q_1q_2q_3$ is less than 2, the configuration will be locked if $L\sin\beta > 2$, or $L > 2/\sin\beta$. $\square$

3.5 2-flexible + 5-rigid can interlock

THEOREM 15. *A flexible 2-chain can interlock with a rigid 5-chain.*

PROOF. We can build this configuration with the coordinates of Figure 10 and check that initially it has positive orientation determinants $[p_1p_2q_iq_{i+1}]$, for $i \in \{0,1,2\}$, and $[p_0p_1q_iq_{i+1}]$, for $i \in \{2,3,4\}$. The four planes $q_iq_{i+1}q_{i+2}$ for $i \in \{0,1,2,3\}$ define a tetrahedron τ, shown dotted in Figure 10, that contains p_1. We can calculate the coordinates s and t, as in the figure, so tetrahedron $\tau = CH(q_2, q_3, s, t)$.

In fact, p_1 cannot leave τ without causing bars of the two chains to intersect. Consider the view from p_1. Ends p_0 and p_2 project to points that are contained in triangles that are projections of bars of Q. These projected triangles are invariant as long as p_1 is in τ: Because the planes $q_0q_2q_3$ and $q_5q_3q_2$ completely contain τ, the end bars of Q project onto q_2q_3 until two edges of a projected triangle becomes collinear, which occurs only if p_1 reaches a face of τ. But this would also force an intersection in the projection between an end bar of P and a bar of Q. Since the length of the end bars of P is > 9, and the greatest distance of τ from a point of the projected triangle is $|sq_1| = 6$, the bars do intersect, as promised. $\square$

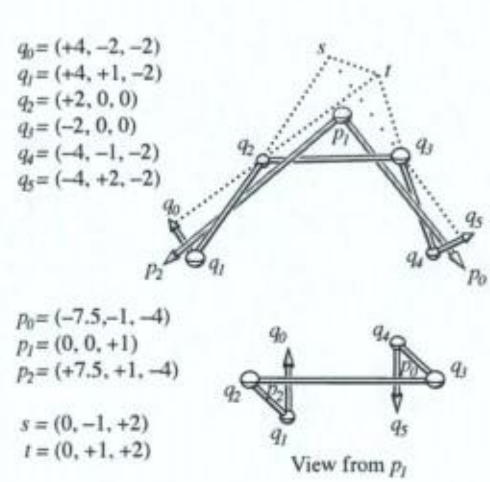


Figure 10: A flexible 2-chain and a rigid 5-chain can interlock.

3.6 3-rigid + 3-flexible can interlock

As shown in Section 2.1, two flexible 3-chains cannot interlock. To obtain a locked configuration for two 3-chains, we could restrict the motion of the chains in several ways. To make these ways precise, consider a 3-chain with vertices p_0, p_1, p_2, and p_3, and define

- the *vertex angle at* p_i, for $i = 1, 2$, which is the angle $\angle p_{i-1}p_ip_{i+1}$, and
- the *dihedral angle* of the 3-chain, which is the angle between the orthogonal projections of p_0p_1 and p_2p_3 onto a plane perpendicular to p_1p_2.

In a flexible chain, these angles are completely unrestricted. For a revolute chain, the vertex angles cannot change during the motion. We will prove that two 3-chains can be locked if:

- The sum of the two vertex angles for each chain is bounded from above by some angle $\alpha < \pi$, or
- Each of the three angles of one of the chains is bounded from below by some angle $\beta > 0$, the other chain being completely flexible.

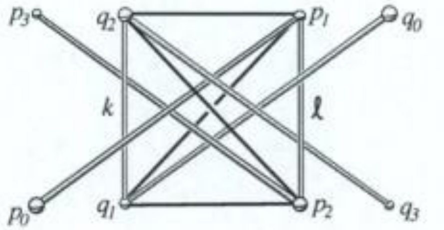


Figure 11: Two 3-chains that interlock if the joints are restricted.

Consider the 3-chains $P = (p_0, \ldots, p_3)$ and $Q = (q_0, \ldots, q_3)$ shown in Fig. 11. The lengths of middle edges $\ell = p_1p_2$ and $k = q_1q_2$ are unity, and the length of all end bars is set to some large value L to be determined later. Let $\mathcal{T}$ be the tetrahedron with vertices $\{p_1, p_2, q_1, q_2\}$. We first show:

LEMMA 16. *Starting from the configuration of Fig. 11, consider any motion where none of the vertices p_0, p_3, q_0 or q_3 ever enter the tetrahedron $\mathcal{T}$, then at all times,*

$$[p_ip_{i+1}q_jq_{j+1}]\left\{\begin{array}{l} < 0 \text{ for } i = j = 1 \\ > 0 \text{ otherwise,}\end{array}\right. \tag{1}$$

and the end bar starting at each vertex of $\mathcal{T}$ intersects the opposite facet of the tetrahedron.

PROOF. It can be verified that expression (1) is true at the starting configuration. Consider the first occurrence of an event that might cause (1) to become false. To consider $[p_0p_1q_jq_{j+1}]$ for $j = 0, 1, 2$, we project the inside of $\mathcal{T}$ from vertex p_1. This is illustrated in Figure 12.

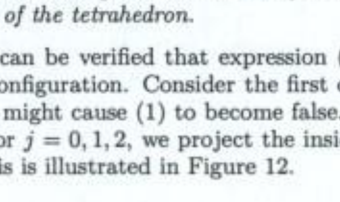


Figure 12: View from p_1

Point p_1 sees the triangle $q_1q_2p_2$ containing p_0. The segment q_0q_1 intersects $\triangle p_1p_2q_2$ and thus intersects p_2q_2 in the projection, and the segment q_2q_3 intersects $\triangle p_1p_2q_1$ and so intersects p_2q_1 in the projection. Because p_0 is actually the projection of p_0p_1, the possible projections of p_0p_1 are bounded by the segments q_0q_1, q_1q_2, and q_2q_3. All the other cases are symmetric to this one except $[p_1p_2q_1q_2]$. But this corresponds to the segments ℓ and k becoming coplanar and $\mathcal{T}$ becoming empty. But this cannot happen before one of the other events. $\square$

THEOREM 17. *Given any angle $0 < \beta < \pi$, there is an interlocked configuration of two 3-chains where the dihedral angle and both vertex angles of the first chain are $\geq \beta$ during any motion and the other chain is unrestricted.*

PROOF. By Lemma 16, the dihedral angle of P is at most the angle θ between $\triangle p_1p_2q_1$ and $\triangle p_1p_2q_2$ (and thus $\theta \geq \beta$) as long as p_0, p_3, q_0 and q_3 stay out of $\mathcal{T}$. The restriction on the vertex angles of P also imply that one of the angles $\angle p_1p_2q_1$ and $\angle p_1p_2q_2$ is at least β, and the same for the angles $\angle p_2p_1q_1$ and $\angle p_2p_1q_2$. Since ℓ and k are both of length 1, then if the longest distance between any two points in $\mathcal{T}$ is D, then p_1q_1, p_1q_2, p_2q_1, and p_2q_2 are all of length $\geq D - 2$. Along with the restrictions on the angles of P, this implies that $(D - 2)\sin\beta \leq 1$ as long as p_0, p_3, q_0 and q_3 stay out of $\mathcal{T}$. Thus if we set the length L of the end bars larger than $2 + 1/\sin\beta$, p_0, p_3, q_0 and q_3 will never enter $\mathcal{T}$, and the configuration is locked. $\square$

COROLLARY 18. *A rigid 3-chain and a flexible 3-chain can interlock.*

3.7 3-revolute + 3-revolute can interlock

In this subsection we consider 3-chains of Figure 11 as revolute chains, and consider the cones obtained by rotating each of the end bars around the middle edge. We will need a new lemma:

LEMMA 19. *In any motion starting from the configuration of Fig. 11, the four cones defined by the chains P and Q have a non-empty intersection as long as none of the vertices p_0, p_3, q_0, or q_3 enter the tetrahedron $\mathcal{T}$.*

PROOF. Using Lemma 16, we claim that the end bars of one of the chains have to intersect both cones of the other chain. To see this, observe that if, say, bar p_0p_1 does not intersect the cone at q_1, then $[p_0p_1q_0q_1]$ and $[p_0p_1q_1q_2]$ have opposite signs (because q_0q_1 is inside the cone), which contradicts lemma 16. Pick a point $\hat{q}$ at the intersection of the boundary of the two cones of Q, such that $\overrightarrow{\hat{q}q_1}$ and $\overrightarrow{q_2\hat{q}}$ have a positive orientation with p_0p_1 and p_2p_3. This implies that bars p_0p_1 and p_1p_2 both intersect the triangle $q_1q_2\hat{q}$. Construct $\hat{p}$ the same way, and notice that the triangles $p_1p_2\hat{p}$ and $q_1q_2\hat{q}$ intersect. Since the triangles are subsets of the cone intersections of their chains, this completes the proof. $\square$

THEOREM 20. *Given any angle $0 < \alpha < \pi$, there is an interlocked configuration of two 3-chains where the sum of the two vertex angles of each chain stays $\leq \alpha$ during any motion (and the dihedral angles are unrestricted).*

PROOF. Let R_i, for $i = 1, 2$, be the union, over all possible pairs of vertex angles with sum $\leq \alpha$, of the intersections of the two cones of C_i. Note that R_i is contained in a sphere of radius $(\cot((\pi-\alpha)/2)+1)/2$ centered at the midpoint of its middle bar. By Lemma 19, we know that R_1 and R_2 intersect, and so do the spheres that contain them, as long as the conditions of lemma 16 are satisfied. So if we set the length L of the end bars larger than $\cot((\pi-\alpha)/2)+2$, then vertices p_0, p_3, q_0 and q_3 will never enter $\mathcal{T}$, and the configuration is interlocked. $\square$

COROLLARY 21. *Two revolute 3-chains can interlock.*

4. CONCLUSION

We have settled the majority of the problems for small interlocked chains. Two problems that would complete Table 1 remain open, as well as other questions that we find interesting:

1. What is the smallest k for which a flexible k-chain can interlock with a flexible 2-chain? We believe that $6 \leq k \leq 11$.
2. What is the smallest k for which a revolute k-chain can interlock with a flexible 2-chain? Does cutting one-third of the vertices of a flexible chain suffice to separate the pieces? Corollary 3 says one-half suffices, but our results do not immediately lead to a better bound.
3. What are the interlocking configurations for sets of three or more chains with restricted motions? For example, we conjecture that a revolute 3-chain and two rigid 2-chains can interlock.
4. What is the complexity of deciding whether given chains are interlocked?

Acknowledgments

We thank Andrea Mantler for her careful reading of the proofs. Work in this area was initiated in the Waterloo algorithmic open-problem session, and continued at CCCG 2000, with contributions by Therese Biedl, Hamish Carr, Eowyn Čenek, Timothy Chan, Beenish Chaudry, Martin Demaine, Rudolf Fleischer, John Iacono, Anna Lubiw, Veronica Morales, Katherine Sinclair, Geetika Tewari, and Mingwei Wang. The last author wishes to thank Sarah Snoeyink for pipe cleaners and coffee stirrers.

5. REFERENCES

[1] T. Biedl, E. Demaine, M. Demaine, S. Lazard, A. Lubiw, J. O'Rourke, M. Overmars, S. Robbins, I. Streinu, G. Toussaint, and S. Whitesides. Locked and unlocked polygonal chains in 3D. In *Proc. 10th ACM-SIAM Sympos. Discrete Algorithms*, pages 866–867, January 1999. Full version LANL arXive cs.CG/9910009.

[2] J. Canny. A new algebraic method for robot motion planning and real geometry. In *Proceedings of the 28th Annual Symposium on Foundations of Computer Science*, pages 39–48, Los Angeles, California, October 1987.

[3] J. Canny. Some algebraic and geometric computations in PSPACE. In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, pages 460–469, Chicago, Illinois, May 1988.

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

[5] N. G. de Bruijn. *Nieuw Archief voor Wiskunde*, 2:67, 1954. Problems 17 and 18. Answers in *Wiskundige Opgaven met de oplossingen*, 20:19–20, 1955.

[6] E. D. Demaine and J. O'Rourke. Open problems from CCCG'99. In *Proc. 12th Canad. Conf. Comput. Geom.*, pages 269–272, August 2000.

[7] Erik Demaine, Stefan Langerman, Joseph O'Rourke, and Jack Snoeyink. Interlocked closed and open linkages with few joints. Submitted. Preliminary version by DLO'R in *CCCG*, 2001.

[8] M. Soss. *Geometric and Computational Aspects of Molecular Reconfiguration*. PhD thesis, McGill University School of Computer Science, 2001.

[9] M. Soss and G. T. Toussaint. Geometric and computational aspects of polymer reconfiguration. *J. Math. Chemistry*, 27(4):303–318, 2000.

A 2-CHAIN CAN INTERLOCK WITH A k-CHAIN

JULIE GLASS, STEFAN LANGERMAN, JOSEPH O'ROURKE, JACK SNOEYINK, AND JIANYUAN K. ZHONG

ABSTRACT. One of the open problems posed in [3] is: what is the minimal number k such that an open, flexible k-chain can interlock with a flexible 2-chain? In this paper, we establish the assumption behind this problem, that there is indeed some k that achieves interlocking. We prove that a flexible 2-chain can interlock with a flexible, open 16-chain.

1. INTRODUCTION

A *polygonal chain* (or just *chain*) is a linkage of rigid bars (line segments, edges) connected at their endpoints (joints, vertices), which forms a simple path (an *open chain*) or a simple cycle (a *closed chain*). A *folding* of a chain is any reconfiguration obtained by moving the vertices so that the lengths of edges are preserved and the edges do not intersect or pass through one another. The vertices act as universal joints, so these are *flexible chains*. If a collection of chains cannot be separated by foldings, the chains are said to be *interlocked*.

Interlocking of polygonal chains was studied in [4, 3], establishing a number of results regarding which collection of chains can and cannot interlock. One of the open problems posed in [3] asked for the minimal k such that a flexible open k-chain can interlock with a flexible 2-chain. An unmentioned assumption behind this open problem is that there is some k that achieves interlocking. It is this question we address here, showing that $k = 16$ suffices.

It was conjectured in [3] that the minimal k satisfies $6 \leq k \leq 11$. This conjecture was based on a construction of an 11-chain that likely does interlock with a 2-chain. We employ some ideas from this construction in the example described here, but for a 16-chain. Our main contribution is a proof that $k = 16$ suffices. It appears that using more bars makes it easier to obtain a formal proof of interlockedness.

Results from [3] include:

(1) Two open 3-chains cannot interlock.
(2) No collection of 2-chains can interlock.
(3) A flexible open 3-chain can interlock with a flexible open 4-chain.

This third result is crucial to the construction we present, which establishes our main theorem, that a 2-chain can interlock a 16-chain (Theorem 1 below.)

2. Idea of Proof

We first sketch the main idea of the proof. If we could build a rigid trapezoid with small rings at its four vertices (T_1, T_2, T_3, T_4), this could interlock with a 2-chain, as illustrated in Figure 1(a). For then pulling vertex v of the 2-chain away from the trapezoid would necessarily diminish the half apex angle α, and pushing v down toward the trapezoid would increase α. But the only slack provided for α is that determined by the diameter of the rings. We make as our subgoal, then, building such a trapezoid.

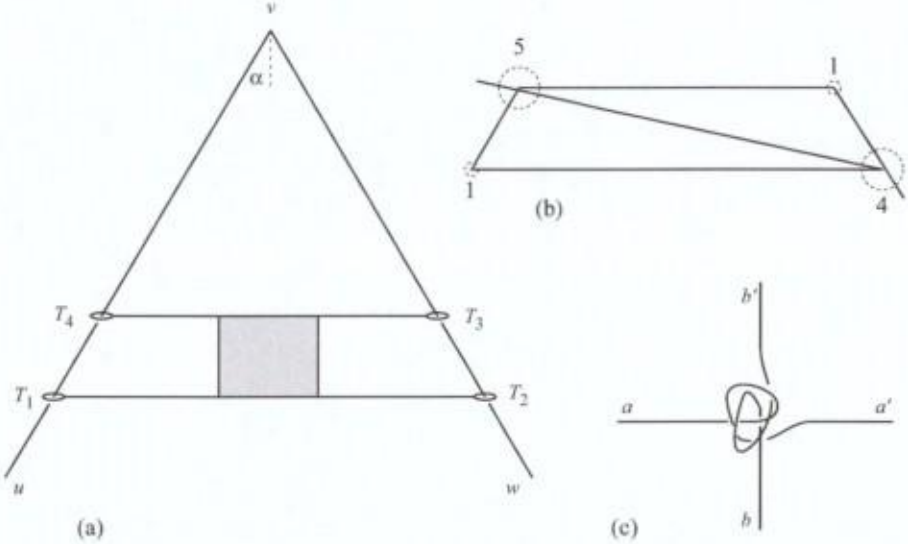


FIGURE 1. (a) A rigid trapezoid with rings would interlock with a 2-chain; (b) An open chain that simulates a rigid trapezoid; (b) Fixing a crossing of aa' with bb'.

We can construct a trapezoid with four links, and rigidify it with two crossing diagonal links. In fact, only one diagonal is necessary to rigidify a trapezoid in the plane, but clearly a single diagonal leaves the freedom to fold along that diagonal in 3D. This freedom will be removed by the interlocked 2-chain, however, so a single diagonal suffices. To create this rigidified trapezoid with a single open chain, we need to employ 5 links, as shown in Figure 1(b). But this will only be rigid if the links that meet at the two vertices incident to the diagonal are truly "pinned" there. In general we want to take one subchain aa' and pin its crossing with another subchain bb' to some small region of space. See Figure 1(c) for the idea.

This pinning can be achieved by the "3/4-tangle" interlocking from [3], result (3) above. So the idea is replace the two critical crossings with a small copy of this configuration. This can be accomplished with 7 links per 3/4-tangle, but sharing with the incident incoming and outgoing trapezoid links potentially reduces the number of links needed per tangle. We have achieved 5 links at one tangle and 4 at the other. The other two vertices of the trapezoid need to simulate the rings in Figure 1(a), and this can be accomplished with one extra link per vertex. Together with the 5 links for the main trapezoid skeleton, we employ a total of $5 + (5 + 4 + 1 + 1) = 16$ links.

3. A 2-CHAIN CAN INTERLOCK AN OPEN 16-CHAIN

3.1. **Open flexible 3- and 4-chains can interlock.** It was proved that open flexible 3- and 4-chains can interlock in [3]. The construction, which we call a *3/4-tangle*, is repeated in Figure 2.

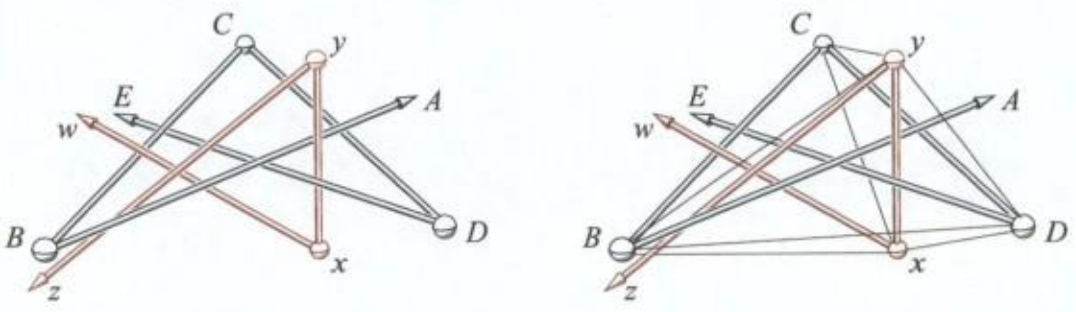


FIGURE 2. Fig. 6 from [3].

It was proved in Theorem 11 of [3] that the convex hull $CH(B, C, D, x, y)$ of the joints B, C, D, x, and y does not change.

We first establish bounds on how far the vertices of the construction can move. Let $BC = CD = xy = 1$ unit, and end bars $AB = DE = xw = yz = 3$ units.

Lemma 1. *Let P be the midpoint of xy. Then in any folding of the interlocked 3- and 4- chain: (1) The distance between P and the endpoints w, z of the 3-chain can be no more than* 3.5 *units, (2) The distance between P and joints B, C, D, x, and y can be no more than* 2.5 *units, and (3) The distance between P and the endpoints A, E of the 4-chain can be no more than* 5.5 *units.*

Proof. (1) Since P is the midpoint of bar xy, x and y are exactly 0.5 units away from P. The joints w, x and P form a triangle, by the triangle inequality $Pw < Px + xw = 0.5 + 3 = 3.5$ units; similarly, $Pz < 3.5$.

(2) We now prove that the distance between P and the joints B, C, D, x, and y can be no more than 2.5 units. In the convex hull $CH(B, C, D, x, y)$, bar xy pierces $\triangle BCD$, where B and D can be imagined to be connected by a rubber band, then $BD < BC + CD = 2$. We observe that: (i) any two points inside $\triangle BCD$ or on the boundary BC, CD, BD are less than 2 units apart, and (ii) the distance between the midpoint P and the plane determined by B, C, D must be less than 0.5 units. From the fact that bar xy pierces $\triangle BCD$, the distance between P and any point on or inside $\triangle BCD$ is less than $2 + 0.5 = 2.5$ units. Since P is the midpoint of bar xy, x

and y are exactly 0.5 units away from P. Therefore, P and the joints B, C, D, x, and y can be no more than 2.5 units as claimed.

(3) Finally, by the triangle inequality $PA < PB + AB < 2.5 + 3 = 5.5$ units; similarly, $PE < 5.5$. □

For $\epsilon > 0$, choosing $BC = CD = xy = \frac{1}{6}\epsilon$, and end bars $Ab = DE = xw = yz = \frac{1}{2}\epsilon$ yields the following:

Corollary 1. *In the above interlocked 3- and 4-chains, let P be the midpoint of xy, then all joints B, C, D, x, y and endpoints A, E, w, z stay inside the ϵ-ball centered at P.*

3.2. **A 2-chain can interlock an open 16-chain.** Take two 3/4-tangles, where all joints and end points of the pair stay within an ϵ-ball centered at the midpoint of the middle link of the 3-chain. Position the tangles as two of the "vertices" of a trapezoid with the links arranged as shown in Figure 3. This design follows Figure 1(b) in spirit, but varies the connections at the diagonal endpoints to increase link sharing. The lower right vertex achieves maximum sharing, in that all three incident trapezoid edges are shared with links of the 3/4-tangle. The upper left vertex shares two incident links. We extend the first and last links of the trapezoid chain to be very long so that the end vertices of the chain are well exterior to any of the ϵ-balls.

3.2.1. *2-chain Through Trapezoid Jag Corners.* Call the simple structure at the other two corners *jag loops*. These corners also can be assured to remain in an ϵ-ball simply by making the extra link length ϵ. Thus we have that all corners of the trapezoid stay within ϵ-balls.

We first argue that the jag loop "grips" the 2-chain link through it, under the assumption of near rigidity of the trapezoid. Let (u, v, w) be the 2-link chain, and let (a, b, c, d) be the vertices constituting a 1-link jag at a corner of the trapezoid. The short link of the jag is bc. The near-rigidity of the trapezoid permits us to take ab to be roughly horizontal (the base of the trapezoid) and cd to be roughly at angle θ with respect to the base (the angle at a base corner of the trapezoid). The link uv is nearly parallel to de, and is woven through the jag as illustrated in Figure 4. The words "roughly" and "nearly" here are intended as shorthand for "approaches, as $\epsilon \to 0$."

Lemma 2. *The plane containing $\triangle abc$ continues to separate v above from u below (where "above" is determined by the counterclockwise ordering of a, b, c) under all nonintersecting foldings of the chains.*

Proof. We argue that uv continues to properly pierce $\triangle abc$ under all foldings, from which it follows that the initial separating property is maintained. The overall structure of the trapezoid prevents uv from moving directly through $\triangle abc$: neither v nor u can get close to the triangle. So the only way the piercing could end is if uv passes through a side of $\triangle abc$. Two of these sides—ab and bc—are links, and avoiding intersection prevents passage through those. Thus uv would have to pass through ac,

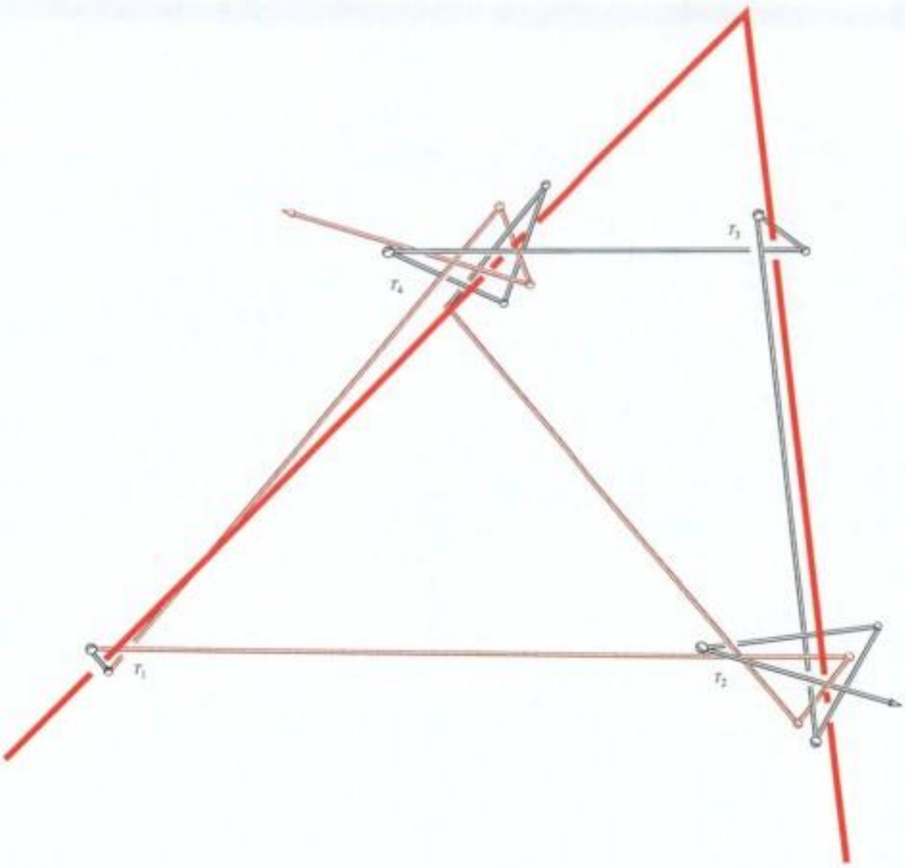


FIGURE 3. An open 16-chain forming a nearly rigid trapezoid.

which is not a link. However, to do this, we now argue it would have to pass through the link cd.

The gap between ab and cd is at most $|bc| = \epsilon$. uv must pass through this gap to "escape" and pass through the segment ac. Because $|uv| \gg \epsilon$, uv must turn "sideways" to pass through it. More precisely, let Q be a plane parallel to ab and cd and midway between them, i.e., Q passes through the midpoint of the gap. uv must align to lie nearly in Q to pass through the gap. Because uv is on the "wrong side" of ab, there are only two ways uv can reach Q: either to align roughly parallel to ab, or to align roughly parallel to cd. In either case, it would then be possible to pass uv through the gap, by keeping it close to the long link to which it is nearly parallel. However, the first alignment places uv at an angle near θ with respect to cd; but it must be nearly parallel to cd. The second alignment requires flipping uv around so that u is above v in the view shown in the figure, in order to get on the other side of ab. But this then makes uv approximately antiparallel to cd, rather than nearly

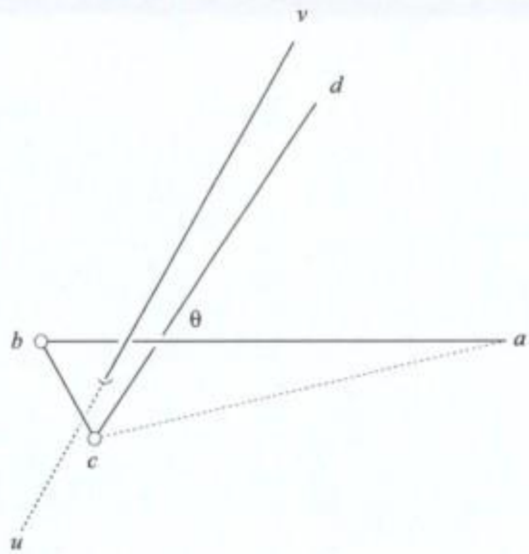


FIGURE 4. A 1-link "jag."

parallel as it must be. Thus the only escape route is impossible, and uv maintains its piercing of $\triangle abc$. □

Corollary 2. *The 1-link jag interlocks with uv, under the constraints imposed by the nearly rigid trapezoid.*

3.2.2. *2-chain Through Trapezoid Tangle Corners.* Next we argue that the link uv can thread through the corner T_4 of the trapezoid so that it is "gripped" by the 3/4-tangle there. Note that the (T_1, T_4) trapezoid link connects to the 3-chain at T_4, which is itself just a jag loop. But uv cannot thread properly through both jag loops on either end of the (T_1, T_4) link. So instead we thread uv through the 4-chain at T_4.

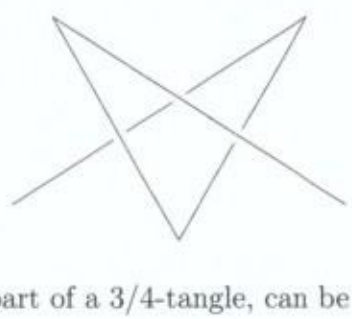


FIGURE 5. A 4-chain, part of a 3/4-tangle, can be viewed as two jag loops.

Now, a 4-chain can be viewed as two jag loops; see Figure 5. Moreover, the 4-chain and 3-chain participating in a 3/4-tangle can be viewed as each lying nearly in planes that are twisted with respect to one another. So we chooose to twist the 4-chain at

T_4 so that uv threads properly through one of its two jag loops. Similarly, the link vw threads from the jag at T_3 through the 3-chain at T_2 (and not through the 4-chain to which (T_4, T_2) is connected).

Applying Corollary 2 to guarantee interlocking yields:

Lemma 3. *The 2-chain links, when threaded as just described, are interlocked with the 3/4-tangles, under the constraints of a nearly rigid trapezoid.*

We should mention that the foregoing argument would be unnecessary if we had instead used a 2-link jags at T_1 and T_3, which would give freedom to position the jag to permit piercing the tangles however desired (and which would lead to an 18-link interlocking chain).

Finally, there is more than enough flexibility in the design to ensure that uv and vw can indeed share the same 2-chain apex v.

3.2.3. *Apex v Cannot Move Far.* Thus the 2-chain (u, v, w) cannot slide free of any of the trapezoid corners unless one of its vertices enters the ϵ-ball containing the corner. We argue below that this cannot occur. We start with a simple preliminary lemma.

Lemma 4. *When ϵ is sufficiently small, a line piercing two disks of radius ϵ can angularly deviate from the line connecting the disk centers at most $\delta \leq 2\epsilon/L$, where L is the distance between the disk centers.*

Proof. Figure 6 illustrates the largest angle δ, $(\frac{1}{2})L \sin\delta = \epsilon$, so $\sin\delta = 2\epsilon/L$, and the claim follows from the fact that $\lim_{x\to 0} \dfrac{\sin x}{x} = 1$. □

Let the trapezoid have base of length $2B$, side length L, and base angle θ. Let the triangle determined by the trapezoid have height h and half-angle α at the apex, so $\tan\theta = h/B$, or $h = B\tan\theta$. See Figure 6(b). The following lemma captures the key constraint on motion of the 2-link.

Lemma 5. *If the sides of the trapezoid pass through the ϵ-disks illustrated, then the height of the triangle approaches h as $\epsilon \to 0$.*

Proof. h_{min} occurs with a triangle apex angle of $\alpha + \delta$ and a base angle of $\theta - \delta$. Let b be the amount by which B is lengthened. In $\triangle xyz$, $b = \frac{\epsilon}{\sin(\theta-\delta)}$, and in $\triangle xYZ$, $\tan(\theta - \delta) = \frac{h_{min}}{B+b}$. Thus we have that

$$h_{min} = \left(B + \frac{\epsilon}{\sin(\theta-\delta)}\right)\tan(\theta-\delta) = B\tan(\theta-\delta) + \frac{\epsilon}{\cos(\theta-\delta)}.$$

Thus h_{min} is continuous near $\epsilon = 0$. Also, if $\epsilon \to 0$ then $\delta \to 0$ since $\delta \leq \frac{2\epsilon}{L}$. Therefore $\lim_{\epsilon\to 0} h_{min} = B\tan\theta = h$ since $\tan(\theta) = h/B$.

For h_{max} all the signs reverse to yield that $\lim_{\epsilon\to 0} h_{max} = B\tan\theta = h$.

We conclude that the height of the triangle approaches h as ϵ approaches 0 as desired. □

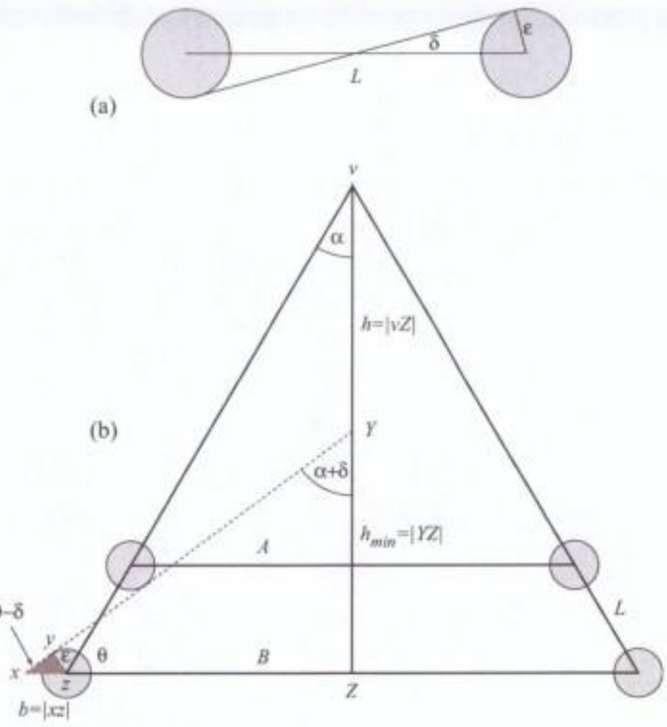


FIGURE 6. Trapezoid Lemma: (a) Line through two disks deviates at most δ; (b) Trapezoid structure, with h_{min} computation illustrated.

3.2.4. *Main Theorem.* We connect 3D to 2D via the plane determined by the 2-link in the proof of the main theorem below.

Theorem 1. *The 2-link chain is interlocked with the 16-link trapezoid chain.*

Proof. Let H be the plane containing the 2-link chain. We know that the links of the 2-chain must pass through ϵ-balls around the four vertices of the trapezoid. H meets these balls in disks each of radius $\leq \epsilon$. The Trapezoid Lemma shows that the height of the triangle approaches h as ϵ approaches 0. Thus, by choosing ϵ small enough, we limit the amount that the apex v of the 2-link chain can be separated from or pushed toward the trapezoid to any desired amount.

We previously established (in Corollary 2 and Lemma 3) that the 2-chain links are interlocked with the 3/4-tangles and jag loops through which they pass, under the assumption that the trapezoid is nearly rigid. The near-rigidity of the trapezoid could only be destroyed by a 2-chain link escaping from one of the jag loops through which it is threaded. But up until the time of this first escape, the trapezoid is nearly rigid; and so there can be no first escape.

Thus, choosing ϵ small enough to prevent any of the vertices of the 2-link chain from entering the ϵ-balls ensures that the 2-link chain is interlocked with the trapezoid chain. □

4. Discussion

We do not believe that $k = 16$ is minimal. We have designed two different 11-chains both of which appear to interlock with a 2-chain. However, both are based on a triangular skeleton rather than on a trapezoidal skeleton, and place the apex v of the 2-chain close to the 11-chain. It seems it will require a different proof technique to establish interlocking, for the simplicity of the proof presented here relies on the vertices of the 2-chain remaining far from the entangling chain.

Another direction to explore is closed chains, for which it is reasonable to expect fewer links. Replacing the 3/4-tangles with "knitting needles" configurations [2][1] produces a closed chain that appears interlocked, but we have not determined the minimum number of links that can achieve this.

Acknowledgement. We thank Erik Demaine for discussions throughout this work. We thank the participants of the DIMACS Reconnect Workshop held at St. Mary's College in July 2004 for helpful discussions. JOR acknowledges support from NSF DTS award DUE-0123154.

References

[1] T. Biedl, E. Demaine, M. Demaine, S. Lazard, A. Lubiw, J. O'Rourke, M. Overmars, S. Robbins, I. Streinu, G. Toussaint, and S. Whitesides. *Locked and unlocked polygonal chains in 3D.* Discrete Comput. Geom., 26(3):269–282, 2001.

[2] J. Cantarella and H. Johnston, *Nontrivial embeddings of polygonal intervals and unknots in 3-space*, Journal of Knot theory and Its Ramifications, 7(8): 1027–1039, 1998.

[3] E. D. Demaine, S. Langerman, J. O'Rourke, and J. Snoeyink, *Interlocked Open Linkages with Few Joints*, Proc. 18th ACM Sympos. Comput. Geom., 189–198, 2002.

[4] E. D. Demaine, S. Langerman, J. O'Rourke, and J. Snoeyink, *Interlocked open and closed linkages with few joints*, Comp. Geom. Theory Appl., 26(1): 37–45, 2003.

DEPARTMENT OF MATHEMATICS & COMPUTER SCIENCE, CALIFORNIA STATE UNIVERSITY HAYWARD, HAYWARD, CA 94542
E-mail address: jglass@csuhayward.edu

UNIVERSITÉ LIBRE DE BRUXELLES, DÉPARTEMENT D'INFORMATIQUE,, ULB CP212, BRUXELLES, BELGIUM
E-mail address: Stefan.Langerman@ulb.ac.be

DEPARTMENT OF COMPUTER SCIENCE, SMITH COLLEGE, NORTHAMPTON, MA 01063
E-mail address: orourke@cs.smith.edu

DEPT. COMPUT. SCI., UNIV. NORTH CAROLINA,, CHAPEL HILL, NC 27599–3175, USA.
E-mail address: snoeyink@cs.unc.edu

DEPARTMENT OF MATHEMATICS & STATISTICS, CALIFORNIA STATE UNIVERSITY SACRAMENTO, 6000 J STREET, SACRAMENTO, CA 95819
E-mail address: kzhong@csus.edu

Flat-State Connectivity of Linkages under Dihedral Motions*

Greg Aloupis[1], Erik D. Demaine[2], Vida Dujmović[1], Jeff Erickson[3], Stefan Langerman[1], Henk Meijer[4], Joseph O'Rourke[5], Mark Overmars[6], Michael Soss[7], Ileana Streinu[5], and Godfried T. Toussaint[1]

[1] School of Computer Science, McGill University
`{athens, vida, sl, godfried}@uni.cs.mcgill.ca`
[2] Laboratory for Computer Science, Massachusetts Institute of Technology
`edemaine@mit.edu`
[3] Dept. of Computer Science, University of Illinois at Urbana-Champaign
`jeffe@cs.uiuc.edu`
[4] Dept. of Computing and Information Science, Queen's University
`henk@cs.queensu.ca`
[5] Dept. of Computer Science, Smith College
`{streinu, orourke}@cs.smith.edu`
[6] Dept. of Information and Computing Sciences, Utrecht University
`markov@cs.uu.nl`
[7] Chemical Computing Group Inc.
`soss@chemcomp.com`

Abstract. We explore which classes of linkages have the property that each pair of their flat states—that is, their embeddings in $\mathbb{R}^2$ without self-intersection—can be connected by a continuous dihedral motion that avoids self-intersection throughout. Dihedral motions preserve all angles between pairs of incident edges, which is most natural for protein models. Our positive results include proofs that open chains with nonacute angles are flat-state connected, as are closed orthogonal unit-length chains. Among our negative results is an example of an orthogonal graph linkage that is flat-state disconnected. Several additional results are obtained for other restricted classes of linkages. Many open problems are posed.

1 Introduction

Motivation: Locked Chains. There has been considerable research on reconfiguration of polygonal chains in 2D and 3D while preserving edge lengths and avoiding self-intersection. Much of this work is on the problem of which classes of chains can *lock* in the sense that they cannot be reconfigured to straight or convex configurations. In 3D, it is known that some chains can lock [4], but the exact class of chains that lock has not been delimited [3]. In 2D, no chains can lock [5,13]. All of these results concern chains with *universal joints.*

* Research initiated at the 17th Winter Workshop on Computational Geometry, Bellairs Research Institute of McGill University, Feb. 1–8, 2002.

P. Bose and P. Morin (Eds.): ISAAC 2002, LNCS 2518, pp. 369–380, 2002.

Motivation: Protein Folding. The backbone of a protein can be modeled as a polygonal chain, but the joints are not universal; rather the bonds between residues form a nearly fixed angle in space. The study of such *fixed-angle* chains was initiated in [11], and this paper can be viewed as a continuation of that study. Although most protein molecules are linear polymers, modeled by open polygonal chains, others are rings (closed polygons) or star and dendritic polymers (trees) [12,6].

The polymer physics community has studied the statistics of "self-avoiding walks" [9,10,15], i.e., non-self-intersecting configurations, often constrained to the integer lattice. To generate these walks, they consider transformations of one configuration to another, such as "pivots" [7] or "wiggling" [8]. Usually these transformations are not considered true molecular movements, often permitting self-intersection during the motion, and perhaps are better viewed as string edits.

In contrast, this paper maintains the geometric integrity of the chain throughout the transformation, to more closely model the protein folding process. We focus primarily on transformations between planar configurations.

Fixed-angle linkages. Before describing our results, we introduce some definitions. A *(general) linkage* is a graph with fixed lengths assigned to each edge. The focus of this paper is *fixed-angle linkages*, which are linkages with, in addition, a fixed angle assigned between each pair of incident edges. We use the term *linkage* to include both general and fixed-angle linkages.

A *configuration* or *realization* of a general linkage is a positioning of the linkage in $\mathbb{R}^3$ (an assignment of a point in $\mathbb{R}^3$ to each vertex) achieving the specified edge lengths. The *configuration space* of a linkage is the set of all its configurations. To match physical reality, of special interest are *non-self-intersecting configurations* or *embeddings* in which no two nonincident edges share a common point. The *free space* of a linkage is the set of all its embeddings, i.e., the subset of configuration space for which the linkage does not "collide" with itself.

A configuration of a fixed-angle linkage must additionally respect the specified angles. The definitions of configuration space, embedding, and free space are the same. A *reconfiguration* or *motion* or *folding* of a linkage is a continuum of configurations. Motions of fixed-angle linkages are distinguished as *dihedral motions*.

Dihedral motions. A dihedral motion can be "factored" into *local dihedral motions* or *edge spins* [11] about individual edges of the linkage. Let $e = (v_1, v_2)$ be an edge for which there is another edge e_i incident to each endpoint v_i. Let Π_i be the plane through e and e_i. A *(local) dihedral motion about* e changes the dihedral angle between the planes Π_1 and Π_2 while preserving the angles between each pair of edges incident to the same endpoint of e. See Fig. 1. The edges incident to a common vertex in a fixed-angle linkage are moved rigidly by a dihedral motion. In particular,

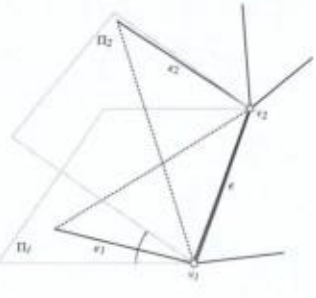


Fig. 1. A local dihedral motion (spin) about edge e.

if the edges are coplanar, they remain coplanar.[1] If we view e and $e_1 \in \Pi_1$ as fixed, then a dihedral motion spins e_2 about e.

Flat-state connectivity. A *flat state* of a linkage is an embedding of the linkage into $\mathbb{R}^2$ without self-intersection. A linkage X is *flat-state connected* if, for each pair of its (distinct, i.e., incongruent) flat states X_1 and X_2, there is a dihedral motion from X_1 to X_2 that stays within the free space throughout. In general this dihedral motion alters the linkage to nonflat embeddings in $\mathbb{R}^3$ intermediate between the two flat states. If a linkage X is not flat-state connected, we say it is *flat-state disconnected*.

Flat-state disconnection could occur for two reasons. It could be that there are two flat states X_1 and X_2 which are in different components of free space but the same component of configuration space. Or it could be that the two flat states are in different components of configuration space. The former reason is the more interesting situation for our investigation; currently we have no nontrivial examples of the latter possibility.

Results. The main goal of this paper is to delimit the class of linkages that are flat-state connected. Our results apply to various restricted classes of linkages, which are specified by a number of constraints, both topological and geometric. The topological classes of linkages we explore include general graphs, trees, chains (paths), both open and closed, and sets of chains. We sometimes restrict all link lengths to be the same, a constraint of interest in the context of protein backbones; we call these *unit-length* linkages. We consider a variety of restrictions on the angles of a fixed-angle linkage, where the angle between two incident links is the smaller of the two angles between them within their common plane. A chain *has a monotone state* if it has a flat state in which it forms a monotone chain in the plane. For sets of chains in a flat state, we *pin* each chain at one of the end links, keeping its position fixed in the plane.

In some cases we restrict the motions of a linkage in one of two ways. First, we may enforce that only certain edges permit local dihedral motion, in which case we call the linkage *partially rigid*. (Such a restriction also constrains the flat states that we can hope to connect, slightly modifying the definition of flat-state connected.) Second, we may restrict the motion to consist of a sequence of 180° edge spins, so that each move returns the linkage to the plane. Most of our examples of flat-state disconnected linkages are either partially rigid or restricted to 180° edge spins.

With the above definitions, we can present our results succinctly in Table 1.

---

[1] Our definition of "dihedral motion" includes rigid motions of the entire linkage, which could be considered unnatural because a rigid motion has no local dihedral motions. However, including rigid motions among dihedral motions does not change our results. For a linkage of a single connected component, we can modulo out rigid motions; and for multiple connected components, we always pin vertices to prevent rigid motions.

Table 1. Summary of results. The '—' means no restriction of the type indicated in the column heading. Entries marked '?' are open problems

Constraints on Fixed-Angle Linkage				Flat-state
Connectivity	Angles	Lengths	Motions	connectivity
Open chain	—	—	—	?
	has a monotone state		—	?
	nonacute	—	—	Connected
	equal acute	—	—	Connected [2]
	each in $(60°, 90°]$	unit	—	Connected [2]
	—	—	$180°$ edge spins	Disconnected
	orthogonal	—	$180°$ edge spins	Connected
Set of chains, each	orthogonal	—	—	Connected
pinned at one end	orthogonal	—	partially rigid	Disconnected
Closed chain	—	—	—	?
	nonacute	—	—	?
	orthogonal	—	—	?
	orthogonal	unit	—	Connected
Tree	—	—	—	?
	orthogonal	—	—	?
	orthogonal	—	partially rigid	Disconnected
Graph	orthogonal	—	—	Disconnected

2 Flat-State Disconnection

It may help to start with negative results, as it is not immediately clear how a linkage could be flat-state disconnected. Several of our examples revolve around the same idea, which can be achieved under several models. We start with partially rigid orthogonal trees, and then modify the example for other classes of linkages.

2.1 Partially Rigid Orthogonal Tree

An *orthogonal tree* is a tree linkage such that every pair of incident links meet at a multiple of $90°$. *Partial rigidity* specifies that only certain edges permit dihedral motions. Note that the focus of a dihedral motion is an edge, not the joint vertex.

Fig. 2(a–b) shows two incongruent flat states of the same orthogonal tree; we'll call the flat states $X_{(a)}$ and $X_{(b)}$. All but four edges of the tree are frozen, the four incident to the central degree-4 root vertex x. Call the 4-link branch of the tree containing a the *a-branch*, and similarly for the others. Label the vertices of the a-branch (a, a_1, a_2, a_3), and similarly for the other branches.

We observe three properties of the example. First, as mentioned previously, fixed-angle linkages have the property that all links incident to a particular vertex remain coplanar throughout all dihedral motions. In Fig. 2, this means that $\{x, a, b, c, d\}$ remain coplanar; and we view this as the plane Π of the flat

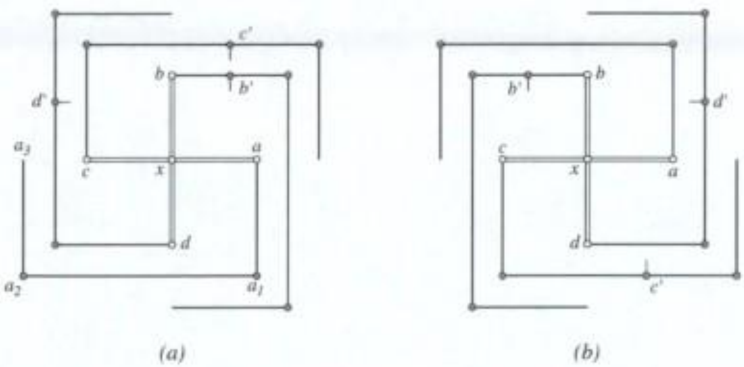


Fig. 2. Two flat states of a partially rigid orthogonal tree. The four open edges are the only ones not rigid, permitting dihedral motions.

states under consideration. Note that, for example, a rotation of a about bd would maintain the $90°$ angles between all edges adjacent consecutively around x, but would alter the $180°$ angle between xa and xc, and thus is not a fixed-angle motion.

Second, the short links, or "pins," incident to vertices b', c', and d' must remain coplanar with their branch, because they are rigid. For example, the b' pin must remain coplanar with xb, for otherwise the rigid edge bb' would twist.

Third, $X_{(a)}$ and $X_{(b)}$ do indeed represent incongruent flat states of the same linkage. The purpose of the b' pin is to ensure that its relation to (say) the c' pin in the two states is not the same. Without the b' pin, a flat state congruent to $X_{(b)}$ could be obtained by a rigid motion of the entire linkage, flipping it upside-down. It is clear that state $X_{(b)}$ can be obtained from state $X_{(a)}$ by rotating the a-branch $180°$ about xa, and similarly for the other branches. Thus the two flat states are in the same component of configuration space. We now show that they are in different components of the free space.

Theorem 1. *The two flat states in Fig. 2 of an orthogonal partially rigid fixed-angle tree cannot be reached by dihedral motions that avoid crossing links.*

Proof: Each of the four branches of the tree must be rotated $180°$ to achieve state $X_{(b)}$. We first argue that two opposite branches cannot rotate to the same side of the Π-plane, either both above or both below. Without loss of generality, assume both the a- and the c-branches rotate above Π. Then, as illustrated in Fig. 3, vertex a_1 must hit a point on the c_1c_2 edge, for the length aa_1 is the same as the distance from a to c_1c_2.

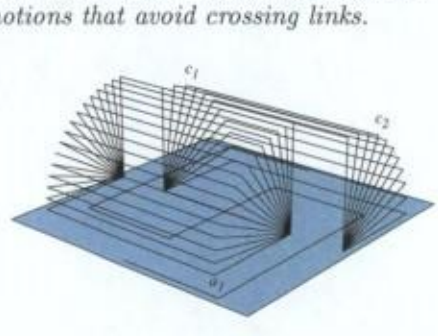


Fig. 3. The a- and c-branches collide when rotated above.

Now we argue that two adjacent branches cannot rotate to the same side of Π. Consider the a- and b-branches, again without loss of generality. As it is more difficult to identify an exact pair of points on the two branches that must collide, we instead employ a topological argument. Connect a shallow rope R from a to a_3 underneath Π, and a rope S from b to b_3 that passes below R. See Fig. 4. In $X_{(a)}$, the two closed loops

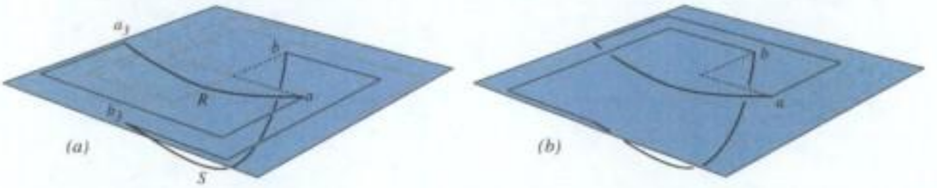


Fig. 4. With the additions of the ropes R and S underneath, the a-chain is not linked with the b-chain in (a), but is linked in (b).

$A = (R, a, a_1, a_2, a_3)$ and $B = (S, b, b_1, b_2, b_3)$ are unlinked. But in $X_{(b)}$, A and B are topologically linked. Therefore, it is not possible for the a- and b-branches to rotate above Π without passing through one another.

By the pigeon-hole principle, at least two branches must rotate though Π. Whether these branches are opposite or adjacent, a collision is forced. □

2.2 Orthogonal Graphs and Partially Rigid Pinned Chains.

We can convert the partially rigid tree in Fig. 2 to a completely flexible graph by using extra "braces" to effectively force the partial rigidity. We can also convert the tree into four partially rigid chains, each pinned at one endpoint near the central degree-4 vertex. Thus we obtain the following two results:

Corollary 1 *The described orthogonal fixed-angle linkage has two flat states that are not connnected by dihedral motions that avoid crossing links.*

Corollary 2 *The four orthogonal partially rigid fixed-angle pinned chains corresponding to Figure 2 are not connected by dihedral motions that avoid crossing links.*

3 Nonacute Open Chains

We now turn to positive results, starting with the simplest and perhaps most elegant case of a single open chain with nonacute angles. After introducing notation, we consider two algorithms establishing flat-state connectivity, in Sections 3.1 and 3.2.

An abstract polygonal chain $\mathcal{C}$ of n links is defined by its fixed sequence of link lengths, $(\ell_1, \ldots, \ell_n)$, and whether it is open or closed. For a fixed-angle chain, the $n-1$ or n angles α_i between adjacent links are also fixed. A realization C of a chain is specified by the position of its $n+1$ vertices: $v_0, v_1, \ldots, v_n$. If the chain is closed, $v_n = v_0$. The links or edges of the chain are $e_i = (v_{i-1}, v_i)$, $i = 1, \ldots, n$, so that the vector along the ith link is $v_i - v_{i-1}$. The plane in which a flat state C is embedded is called Π or the xy-plane.

3.1 Lifting One Link at a Time

The idea behind the first (unrestricted) algorithm is to lift the links of the chain one-by-one into a monotone chain in a vertical plane. Once we reach this *canonical state*, we can reverse and concatenate motions to reach any flat state from any other.

We begin by describing the case of orthogonal chains, as illustrated in Fig. 5, and the algorithm will generalize to arbitrary nonacute chains. The invariant at the beginning of each step i of the algorithm is that we have lifted the chain $e_1, \ldots, e_i$ into a monotone chain in a vertical plane, while the rest of the chain $e_{i+1}, \ldots, e_n$ remains where it began in the xy-plane. Initially, $i = 0$ and the lifted chain contains no links, and we simply lift the first link e_1 to vertical by a $90°$ edge spin around the second link e_2. For general i, we first spin the lifted chain around its last (vertical) link e_i so that the vertical plane contains the next link to lift, e_{i+1}, and so that the chain $e_1, \ldots, e_{i+1}$ is monotone. Then we pick up e_{i+1} by a $90°$ edge spin around e_{i+2}. Throughout, the lifted chain remains monotone and contained in the positive-z halfspace, so we avoid self-intersection.

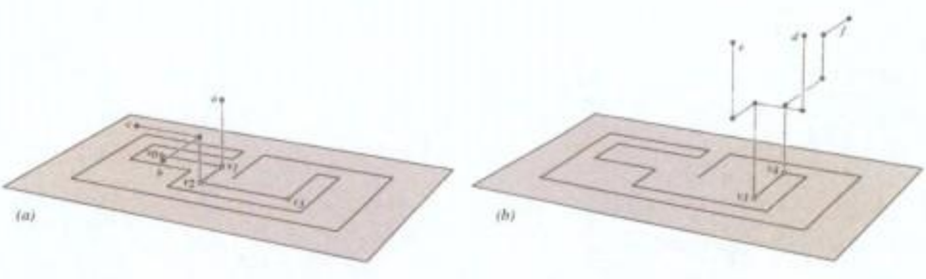


Fig. 5. Picking up a planar orthogonal chain into a monotone canonical state. (a) Lifting edges $e_1 = (v_0, v_1)$ and $e_2 = (v_1, v_2)$: a, b, c. (b) Lifting edges e_3 and e_4: d, e, f.

Nonacute chains behave similarly to orthogonal chains, in particular, the canonical state is monotone, although it may no longer alternate between left and right turns. Now there may be multiple monotone states, and we must choose the state that is monotone in the z dimension. The key property is that, as the chain $e_1, \ldots, e_i$ rotates about e_{i+1}, the chain remains monotone in the z direction, so it does not penetrate the xy-plane.

This algorithm proves the following result:

Theorem 2. *Any nonacute fixed-angle chain is flat-state connected.*

3.2 Lifting Two Links at a Time

The algorithm above makes at most 2 edge spins per link pickup, for a total of $2n$ edge spins to reach the canonical state, or $4n$ edge spins to reach an arbitrary flat state from any other. This bound is tight within an additive constant.

We can reduce the number of edge spins to $1.5n$ to reach the canonical state, or $3n$ to reach an arbitrary flat state, by lifting two edges in each step as follows. As before, in the beginning of each step, we spin the lifted chain $e_1, \ldots, e_i$ about the last link e_i to orient it to be coplanar and monotone with the next link e_{i+1}. Now we spin by $90°$ the lifted chain and the next two links e_{i+1} and e_{i+2} about the following link e_{i+3}, bringing e_{i+1} and e_{i+2} into a vertical plane, and tilting the lifted chain $e_1, \ldots, e_i$ down to a horizontal plane (parallel to the xy-plane) at the top. Then we spin the old chain $e_1, \ldots, e_i$ by $90°$ around e_{i+1}, placing it back into a vertical plane, indeed the same vertical plane containing e_{i+1} and e_{i+2}, so that the new chain $e_1, \ldots, e_{i+2}$ becomes coplanar and monotone. We thus add two links to the lifted chain after at most three motions, proving the $1.5n$ upper bound; this bound is also tight up to an additive constant.

Corollary 3 *Any nonacute fixed-angle chain with n links can be reconfigured between two given flat states in at most $3n$ edge spins.*

4 Multiple Pinned Orthogonal Open Chains

In this section we prove that any collection of open, orthogonal chains, each with one edge pinned to the xy-plane, can be reconfigured to a canonical form, establishing that such chain collections are flat-state connected. We also require a "general position" assumption: no two vertices from different chains have a common x- or y-coordinate. Let C_i, $i = 1, \ldots, k$, be the collection of chains in the xy-plane. Each has its first edge pinned, i.e., v_0 and v_1 have fixed coordinates in the plane; but dihedral motion about this first edge is still possible (so the edge is not frozen). Call an edge parallel to the x-axis an *x-edge*, and similarly for *y-edge* and *z-edge*. The canonical form requires each chain to be a staircase in a plane parallel to the z-axis and containing its first (pinned) edge. If the first chain edge is a y-edge, the staircase is in a *yz quarter plane* in the halfspace $z > 0$ above xy; if the first chain edge is an x-edge, the staircase is in an *xz quarter plane* in the halfspace $z < 0$ below xy.

The algorithm processes independently the chains that are destined above or below the xy-plane, and keeps them on their target sides of the xy-plane, so there is no possibility of interference between the two types of chains. So henceforth we will concentrate on the chains C_i whose first edge is a y-edge, with the goal of lifting each chain C_i into a staircase S_i in a yz quarter plane. At an intermediate state, the staircase S_i is the portion of the lifted chain above the xy-plane, and C_i the portion remaining in the xy-plane. The *pivot edge* of the staircase is its first edge, which is a z-edge. Let $(\ldots, c_i, b_i, a_i)$ be the last three vertices of the chain C_i. Let a_i have coordinates (a_x, a_y); we'll use analogous notation for b_i and c_i. Vertex a_i at the foot of a staircase is its *base vertex* and the last edge of the chain, (b_i, a_i), is the staircase's *base edge*.

After each step of the algorithm, two invariants are reestablished:

1. All staircases for all chains are in (parallel) yz quarter planes;
2. The base edge for every staircase is a y-edge, i.e., is in the plane of the staircase.

We will call these two conditions the *Induction Hypothesis.*

The main idea of the algorithm is to pick up two consecutive edges of one chain, which then ensures that the next edge of that chain is a y-edge. The chain is chosen arbitrarily. To simplify the presentation, we assume without loss of generality that c_i is to the right of b_i. First, the staircases whose pivot's x-coordinates lie in the range $[b_x, c_x]$ are reoriented to avoid crossing above the (b_i, c_i) edge.

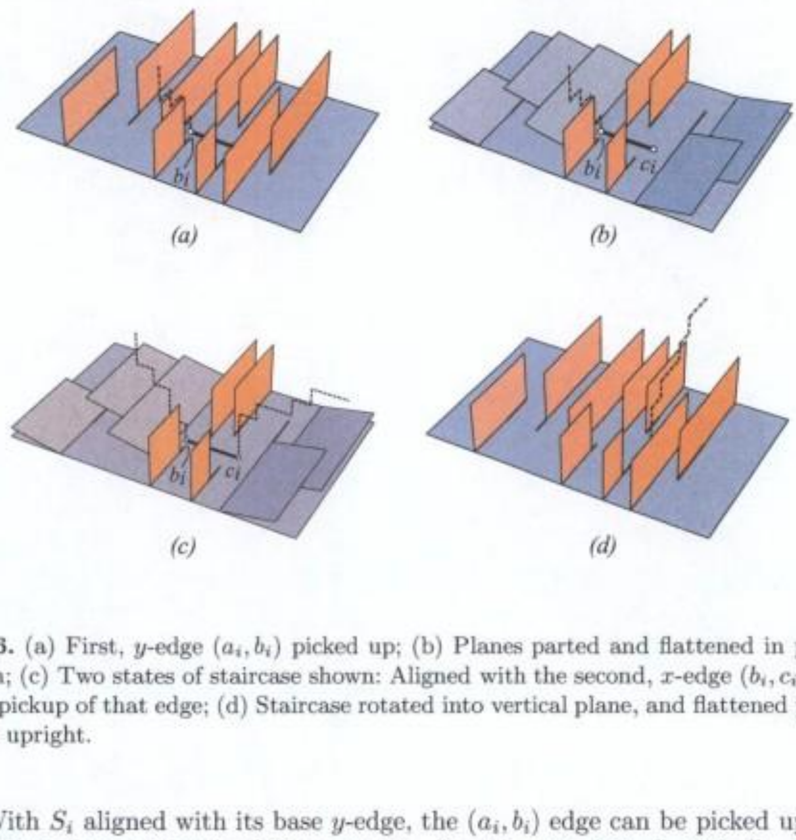


Fig. 6. (a) First, y-edge (a_i, b_i) picked up; (b) Planes parted and flattened in preparation; (c) Two states of staircase shown: Aligned with the second, x-edge (b_i, c_i), and after pickup of that edge; (d) Staircase rotated into vertical plane, and flattened planes made upright.

With S_i aligned with its base y-edge, the (a_i, b_i) edge can be picked up into a vertical plane without collision; see Fig. 6a. We now align S_i with (b_i, c_i), by "parting" the planes at b_x toward the left, laying all planes left of b_x down toward $-x$ (Fig. 6b), and then rotating S_i to be horizontal. Now we pick up (b_i, c_i) into a xz quarter plane, after laying down all planes right of c_x; see Fig. 6(c). Finally, reorient the xz-plane to be vertical and then restore all tilted planes to their yz orientation. We have reestablished the Induction Hypothesis. See Fig. 6(d).

Repeating this process eventually lifts every chain into parallel vertical planes, leaving only the first (pinned) y-edge of each chain in the xy-plane.

5 Unit Orthogonal Closed Chains

Our only algorithm for flat-state connectivity of closed chains is specialized to unit-length orthogonal closed chains. Despite the specialization, it is one of the most complex algorithms, and will only be mentioned in this abstract. We use orthogonally convex polygons as a canonical form, justified by the first lemma:

Lemma 1. *Let C and D be two orthogonally convex embeddings of a unit-length orthogonal closed chain with n vertices. There is a sequence of edge spins that transforms C into D.*

The more difficult half is establishing the following:

Lemma 2. *Let C be a flat state of a unit-length orthogonal closed chain with n vertices. There is a sequence of edge spins that transforms C into an orthogonally convex embedding.*

These lemmas establish the following theorem:

Theorem 3. *Any unit-length orthogonal closed chain is flat-state connected.*

6 180° Edge Spins

A natural restriction on dihedral motions is that the motion decomposes into a sequence of moves, each ending with the chain back in the xy-plane—in other words, 180° edge spins. This restriction is analogous to Erdős flips in the context of locked chains [14,1,3]. In this context, we can provide sharper negative results—general open chains can be flat-state disconnected—and slightly weaker positive results—orthogonal open chains are flat-state connected.

6.1 Restricted Flat-State Disconnection of Open Chains

We begin by illustrating the difficulty in reconfiguring open chains by 180° edge spins; see Fig. 7. Spinning about edge 1 does nothing; spinning about edge 2

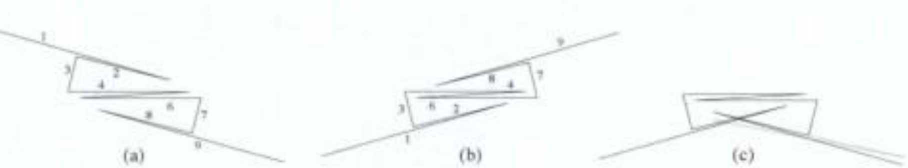


Fig. 7. (a–b) Two flat states of a chain that cannot reach each other via a sequence of 180° edge spins. (c) Attempt at spinning about edge 4.

causes edges 1 and 3 to cross; spinning about edge 3 makes no important change to the flat state; spinning about edge 4 causes edges 2 and 8 to cross as shown in Fig. 7(c); spinning about edge 5 causes edges 4 and 6 to cross (in particular); and the remaining cases are symmetric. This case analysis establishes the following theorem:

Theorem 4. *The two incongruent flat states in Fig. 7(a–b) of a fixed-angle open chain cannot be reached by a sequence of* $180°$ *edge spins that avoid crossing links.*

6.2 Restricted Flat-State Connection of Orthogonal Open Chains

The main approach for proving flat-state connectivity of orthogonal chains is outlined in two figures: spin around a convex-hull edge if one exists (Fig. 8), and otherwise decompose the chain into a monotone (staircase) part and an inner part, and spin around a convex-hull edge of the inner part (Fig. 9). Such spins avoid collisions because of the empty infinite strips $R(e_1)$, $R(e_2)$, ... through the edges of the monotone part of the chain. In Fig. 9, the monotone portion of the chain is e_1, e_2, e_3, which terminates with the first edge e_3 that does not have an entire empty strip $R(e_3)$. Each spin of either type makes the chain more monotone in the sense of turning an edge whose endpoints turn in the same direction into an edge whose endpoints turn in opposite directions; hence, the number of spins is at most n. Using a balanced-tree structure to maintain information about recursive subchains, each step can be executed in $O(\log n)$ time, for a total of $O(n \log n)$ time. In addition, we show how the algorithm can be modified to keep the chain in the nonnegative-x halfspace with one vertex pinned against the $x = 0$ plane.

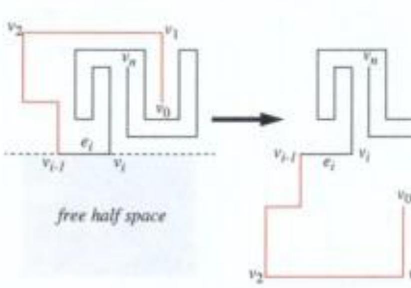


Fig. 8. A dihedral rotation about a convex-hull edge resolves a violation of the canonical form.

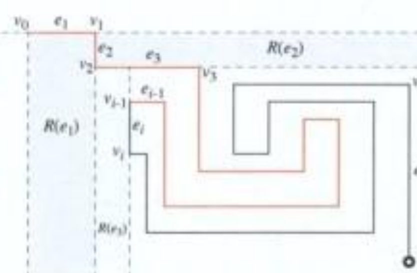


Fig. 9. Determining the chain e_1, e_2, ..., e_{i-1} that can be rotated about e_i.

Theorem 5. *Orthogonal chains are flat-state connected even via restricted sequences of* $180°$ *spins that keep the chain in the nonnegative-*x *halfspace with one vertex pinned at* $x = 0$. *The sequence of* $O(n)$ *spins can be computed in* $O(n \log n)$ *time.*

7 Conclusion and Open Problems

See Table 1 for several open problems. In particular, these three classes of chains seem most interesting, with the first being the main open problem:

1. Open chains (no restrictions).
2. Open chains with a monotone flat state.
3. Orthogonal trees (all joints flexible).

References

1. Oswin Aichholzer, Carmen Cortés, Erik D. Demaine, Vida Dujmović, Jeff Erickson, Henk Meijer, Mark Overmars, Belén Palop, Suneeta Ramaswami, and Godfried T. Toussaint. Flipturning polygons. In *Proc. Japan Conf. Discrete Comput. Geom.*, Lecture Notes in Computer Science, Tokyo, Japan, November 2000. To appear in *Discrete Comput. Geom.*
2. Greg Aloupis, Erik D. Demaine, Henk Meijer, Joseph O'Rourke, Ileana Streinu, and Godfried Toussaint. Flat-state connectedness of fixed-angle chains: Special acute chains. In *Proc. 14th Canad. Conf. Comput. Geom.*, August 2002.
3. T. Biedl, E. Demaine, M. Demaine, S. Lazard, A. Lubiw, J. O'Rourke, M. Overmars, S. Robbins, I. Streinu, G. Toussaint, and S. Whitesides. Locked and unlocked polygonal chains in 3D. *Discrete Comput. Geom.*, 26(3):269–282, 2001.
4. J. Cantarella and H. Johnston. Nontrivial embeddings of polygonal intervals and unknots in 3-space. *J. Knot Theory Ramifications*, 7:1027–1039, 1998.
5. R. Connelly, E. D. Demaine, and G. Rote. Straightening polygonal arcs and convexifying polygonal cycles. In *Proc. 41st Annu. IEEE Sympos. Found. Comput. Sci.*, pages 432–442. IEEE, November 2000.
6. Maxim D. Frank-Kamenetskii. *Unravelling DNA*. Addison-Wesley, 1997.
7. M. Lal. Monte Carlo computer simulations of chain molecules. *Molecular Physics*, 17:57–64, 1969.
8. B. MacDonald, N. Jan, D. L. Hunter, and M. O. Steinitz. Polymer conformations through wiggling. *Journal of Physics A: Mathematical and General Physics*, 18:2627–2631, 1985.
9. N. Madras, A. Orlitsky, and L. A. Shepp. Monte Carlo generation of self-avoiding walks with fixed endpoints and fixed length. *Journal of Statistical Physics*, 58:159–183, 1990.
10. Neal Madras and Gordon Slade. *The Self-Avoiding Walk*. Birkhäuser, 1993.
11. M. Soss and G. T. Toussaint. Geometric and computational aspects of polymer reconfiguration. *J. Math. Chemistry*, 27(4):303–318, 2000.
12. C. E. Soteros and S. G. Whittington. Polygons and stars in a slit geometry. *Journal of Physics A: Mathematical and General Physics*, 21:L857–L861, 1988.
13. I. Streinu. A combinatorial approach to planar non-colliding robot arm motion planning. In *Proc. 41st Annu. IEEE Sympos. Found. Comput. Sci.* IEEE, November 2000. 443–453.
14. G. T. Toussaint. The Erdős-Nagy theorem and its ramifications. In *Proc. 11th Canad. Conf. Comput. Geom.*, pages 9–12, 1999.
15. S. G. Whittington. Self-avoiding walks with geometrical constraints. *Journal of Statistical Physics*, 30(2):449–456, 1983.

Algorithmica (2006) 44: 167–181
DOI: 10.1007/s00453-005-1205-7

Algorithmica

Geometric Restrictions on Producible Polygonal Protein Chains[1]

Erik D. Demaine,[2] Stefan Langerman,[3] and Joseph O'Rourke[4]

Abstract. Fixed-angle polygonal chains in three dimensions serve as an interesting model of protein backbones. Here we consider such chains produced inside a "machine" modeled crudely as a cone, and examine the constraints this model places on the producible chains. We call this notion *producible*, and prove as our main result that a chain whose maximum turn angle is α is producible in a cone of half-angle $\geq \alpha$ if and only if the chain is flattenable, that is, the chain can be reconfigured without self-intersection to lie flat in a plane. This result establishes that two seemingly disparate classes of chains are in fact identical. Along the way, we discover that all producible configurations of a chain can be moved to a canonical configuration resembling a helix. One consequence is an algorithm that reconfigures between any two flat states of a "nonacute chain" in $O(n)$ "moves," improving the $O(n^2)$-move algorithm in [ADD$^+$].

Finally, we prove that the producible chains are rare in the following technical sense. A random chain of n links is defined by drawing the lengths and angles from any "regular" (e.g., uniform) distribution on any subset of the possible values. A random configuration of a chain embeds into $\mathbb{R}^3$ by in addition drawing the dihedral angles from any regular distribution. If a class of chains has a locked configuration (and no nontrivial class is known to avoid locked configurations), then the probability that a random configuration of a random chain is producible approaches zero geometrically as $n \to \infty$.

Key Words. Polygonal chains, Locked chains, Fixed-angle chains, Flattenable chains, Protein folding, Protein backbone.

1. Introduction. The backbone of a protein molecule may be modeled as a three-dimensional polygonal chain, with joints representing residues and fixed-length links (edges) representing bonds. The joints are not universal; rather successive bonds form nearly fixed angles in space. The motions at the joints are then called *dihedral* motions. The study of such *fixed-angle* chains was initiated in [ST] and continued in [ADM$^+$] and [ADD$^+$]. These papers identified *flat states* of a chain—embeddings into a plane without self-intersection—as geometrically interesting. A chain that can reconfigure in $\mathbb{R}^3$ via dihedral motions between any two of its flat states is called *flat-state connected*. A chain that has a flat state but is in a configuration that cannot reach that

[1] The first author was supported by NSF CAREER Award CCF-0347776 and DOE Grant DE-FG02-04ER25647. The third author was supported by NSF Distinguished Teaching Scholars Award DUE-0123154.
[2] MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge, MA 02139, USA. edemaine@mit.edu.
[3] Chercheur Qualifié du FNRS. Département d'informatique, Université Libre de Bruxelles, ULB CP212, Bruxelles, Belgium. Stefan.Langerman@ulb.ac.be.
[4] Department of Computer Science, Smith College, Northampton, MA 01063, USA. orourke@cs.smith.edu.

Received February 23, 2004. Communicated by N. Katoh.
Online publication December 30, 2005.

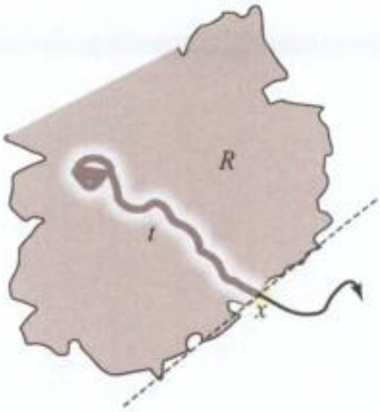


Fig. 1. The ribosome R in crosssection. The protein is created in tunnel t and emerges at x.

state (via dihedral motions, without self-intersection) is called *unflattenable* or simply *locked*.[5]

We look here at a particularly simple but natural constraint on the "production" of a fixed-angle chain. Our inspiration derives from the ribosome, which is the "machine" that creates protein chains in biological cells. Figure 1 shows a schematic cross section of a ribosome and its exit tunnel, based on a model developed by Nissen et al. [NHB$^+$]. We consider a very simple geometric model that roughly captures the exit point x of the ribosome: the chain is produced inside a cone of some half-angle β, with the chain emerging through the cone's apex. See Figure 2. This constraint immediately implies that the maximum turn angle α in the produced chain is at most 2β. We consider the somewhat stronger condition that $\alpha \leq \beta$. These conditions are consistent with our analogy to the ribosome, where the cone is roughly a half-plane (half-angle $\beta = 90^\circ$) and the chain has obtuse angles around 110° (turn angle $\alpha = 70^\circ$).

We show in Section 3 that this simple constraint guarantees that all producible chains are flattenable and furthermore mutually reachable. There are several interesting aspects to this result. First, we are naturally led in our proof to a canonical form, called α-CCC, which bears a resemblance to the helical form preferred by many proteins. Second, we show in Section 5 that long "random" chains are locked with probability approaching 1, implying that producible protein chains are rather special. Third, we show in Section 4 that if we strengthen the production model to allow producing chain turn angles of more than 2β, then locked chains can be produced. This example shows the importance of our condition that $\alpha \leq \beta$ (or a similar condition such as $\alpha \leq 2\beta$).

2. Definitions

2.1. *Chains and Motions*. The fixed-angle polygonal chain P has $n+1$ vertices $V = \langle v_0, \ldots, v_n \rangle$ and is specified by the fixed turn angle θ_i at each vertex v_i, $i = 1, \ldots, n-1$,

[5] In fact, this definition is slightly more specific than the usual notion of "locked," which says that there are two arbitrary configurations of the linkage that are mutually unreachable.

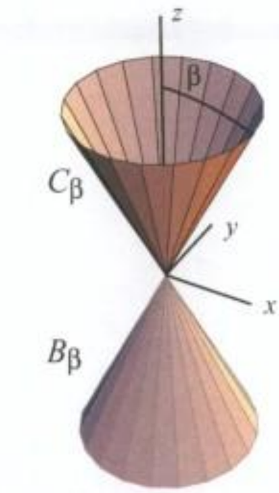


Fig. 2. The chain is produced in cone C_β, and emerges at the origin into the complementary cone B_β below the xy-plane.

and by the edge length d_i between v_i and v_{i+1}, $i = 0, \ldots, n-1$. When all angles $\theta_i \leq \alpha$ for some $0 < \alpha < \pi$, P is called a *($\leq \alpha$)-chain.*[6] We write $P[i, j]$, $i \leq j$, for the polygonal subchain composed of vertices $v_i, \ldots, v_j$.

A *configuration* $Q = \langle q_0, \ldots, q_n \rangle$ of the chain P (see Figure 3) is an embedding of P into $\mathbb{R}^3$, i.e., a mapping of each vertex v_i to a point $q_i \in \mathbb{R}^3$, satisfying the constraints that the angle between vectors $q_{i-1}q_i$ and q_iq_{i+1} is θ_i, and the distance between q_i and q_{i+1} is d_i. The points q_i and q_{i+1} are connected by a straight line segment e_i. Thus, a configuration can be specified by the position of e_0 and dihedral angles δ_i, $i = 1, \ldots, n-2$, where δ_i is the angle between planes $e_{i-1}e_i$ and e_ie_{i+1}. The configuration is *simple* if no two nonadjacent segments intersect.

A *motion* $M = \langle m_0, \ldots, m_n \rangle$ of a chain P is a list of $n+1$ continuous functions $m_i\colon [0, \infty] \to \mathbb{R}^3$, $i = 0, \ldots, n$, such that $M(t) = \langle m_0(t), \ldots, m_n(t) \rangle$ is a configuration of P for all $t \in [0, \infty]$. The motion is said to be *simple* if all such configurations $M(t)$

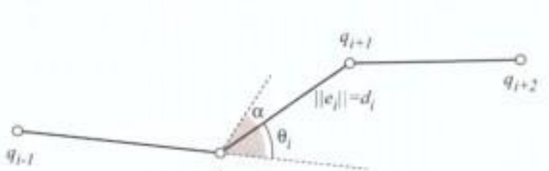


Fig. 3. Notation for a configuration Q.

[6] Other work [ADM⁺], [ADD⁺] focuses on the angle between adjacent edges, which for us is $\pi - \alpha$. Thus "nonacute chains" in that work corresponds to ($\leq \pi/2$)-chains here. Our use of the turn angle is more in consonance with cone production.

are simple. We normally assume that the motion is *finite* in the sense that, after some time T, M becomes independent of t.

2.2. *Chain Production.* As mentioned above, our model is that the chain is produced inside an infinite open cone C_β with apex at the origin, axis on the z-axis, and half-angle (angle to the positive z-axis) β; see Figure 2. In fact the production happens in the closure $\overline{C}_\beta$ of the cone (the cone plus its surface). Vertices and edges are produced sequentially over time inside the cone $\overline{C}_\beta$ and eventually exit through the origin. The production process maintains the invariant that at most one link, the last link produced, is (partially) inside the cone; once a link is fully outside the cone it must remain so. The last produced link must constantly touch the origin, with one half of the segment inside the cone and the other half outside the cone. The rest of the chain can move freely as long as it stays simple and never meets the cone C_β.

More precisely, at time $t_0 = 0$, the machine creates q_0 at the apex of C_β, q_1 inside $\overline{C}_\beta$, and the segment e_0 connecting them; see Figure 4. In general, at time t_i, vertex q_i reaches the origin, and q_{i+1} and e_i are created at arbitrary locations inside the cone $\overline{C}_\beta$. The vertex q_i stays in $\overline{C}_\beta$ between times t_{i-1} and t_i, and stays outside C_β after time t_i. In total there are $n + 1$ critical times satisfying $0 = t_0 < t_1 < \cdots < t_n$.

Formally, a β-*production* F is a set of $n+1$ continuous functions $f_i\colon [t_{i-1}, \infty] \to \mathbb{R}^3$, $i = 0, \ldots, n$, such that, for all $t \in [t_{j-1}, t_j]$, $f_j(t) \in \overline{C}_\beta$, $F(t) = \langle f_0(t), \ldots, f_j(t)\rangle$ is a simple configuration of the segment e_{j-1} is incident to the origin, and no segment e_i intersects C_β, $i < j-1$. A configuration Q is β-*producible* if there exists a β-production F with $F(\infty) = Q$. We say that a configuration is $(\geq \alpha)$-producible if it is β-producible for some $\beta \geq \alpha$.

One consequence of this model is that, as the last link produced exits the cone $\overline{C}_\beta$, it must enter what we call the *complementary cone* $\overline{B}_\beta$. For $\beta \leq \pi/2$ (a convex cone C_β), the complementary cone $\overline{B}_\beta$ is the mirror image of $\overline{C}_\beta$ with respect to the xy-plane. For $\beta \geq \pi/2$ (a reflex cone C_β), the complementary cone B_β is the region of space exterior

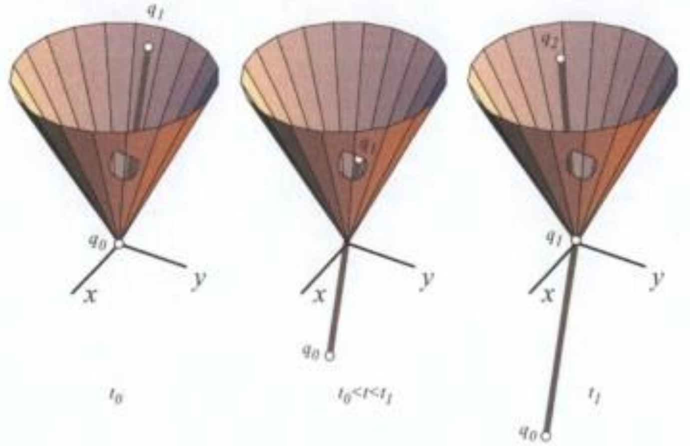


Fig. 4. Production of e_0 and e_1 during $t \in [t_0, t_1]$.

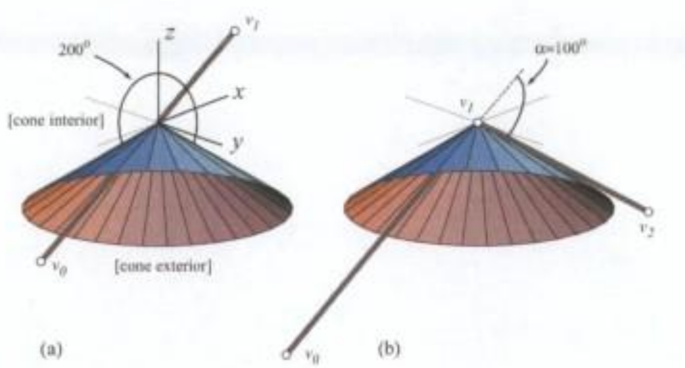


Fig. 5. Production in cone of $\beta > \pi/2$. Here $\beta = 100^\circ$, so that the full cone angle is 200°. The viewpoint is under the xy-plane. (a) e_0 exits to the exterior of the cone during $t \in [t_0, t_1)$. (b) e_1 is created at $t = t_1$ inside the cone, forming, in this instance, a turn angle of 100°.

to C_β. (This region is smaller than the mirror image of $\overline{C}_\beta$ in this case.) Figure 5 shows an example of production when $\beta \geq \pi/2$.

This complementary cone restricts the achievable turn angles in the producible chains:

LEMMA 1. *To produce a chain whose maximum turn angle is α using a cone C_β, we must have $\alpha/2 \leq \beta \leq \pi - \alpha/2$.*

PROOF. Suppose $\theta_i = \alpha$. At time t_i, when q_{i+1} is created inside the cone, q_i is at the apex, and q_{i-1} is outside. Because we stipulate continuous motion, q_{i-1} must be inside the cone $\overline{B}_\beta$ below the xy-plane, for it must have been there throughout $t \in [t_{i-1}, t_i)$. For the same reason, q_{i+1} must be in the mirror image of $\overline{B}_\beta$ with respect to the xy-plane, because e_i is just about to enter $\overline{B}_\beta$. The cone $\overline{B}_\beta$ and its mirror image each form an angle $\min(\beta, \pi - \beta)$ with the z-axis, so in order for e_{i-1} and e_i to fit those cones, $\alpha/2 \leq \min(\beta, \pi - \beta)$. □

Note that arbitrarily sharp turn angles can be produced in a cone $C_{\pi/2}$, which might be viewed as a half-space with a pinhole exit at the origin.

We prove that there exists a simple motion between any two β-producible configurations of the same chain, and that all such configurations are flattenable. Next we define the notion of a "simple" motion.

2.3. *Complexity of a Motion.* There are of course many ways to define the complexity of a motion M. As a first approximation, we could assume that each dihedral angle $\delta_i^M(t)$ of the segment e_i is a piecewise-linear function of time t, and the complexity $T(M)$ of the motion M is the total number of linear pieces over all functions $\delta_i^M(t)$. That is, $T(M) = \sum_{i=1}^{n-2} T(\delta_i^M)$, where $T(\delta_i^M)$ is the number of linear pieces in the function δ_i^M. Unfortunately, this definition is not acceptable, as it restricts the range of possible motions M. The definition can be generalized to allow arbitrary functions $\delta_i^M(t)$, given some corresponding measure of complexity $T(\delta_i^M)$, with the added restriction that for every time range $t \in [r, s]$ during which $\delta_i^M(t)$ is a linear function, that time range contributes at most 1 to the complexity $T(\delta_i^M)$. For example, if $\delta_i^M(t)$ is a piecewise-polynomial function, $T(\delta_i^M)$ could be defined as the sum of the degrees of the polynomial pieces; or more generally $T(\delta_i^M(t))$ might measure the number of inflection points or monotonic pieces of $\delta_i^M(t)$.

The complexity of a production F can be defined in an analogous way, where $\delta_i^F(t)$ is defined only for the time range $t \geq t_{i+1}$. The resulting value will only account for the dihedral motions outside the cone C_β. We still need to add the complexity of the movement of point $f_{i+1}(t)$ before it exits the cone for all i, i.e., at time $t \in [t_i, t_{i+1})$. If we assume that the chain exits the cone at a constant rate, we only need to consider the vector $u^F(t) = (0, f_{i+1}(t))$ for $t \in [t_i, t_{i+1})$, described in polar coordinates by the angle $\rho^F(t)$ of $u^F(t)$ with the z-axis, and the angle $\gamma^F(t)$ of the projection of $u^F(t)$ onto the xy-plane with the x-axis. The complexity will be expressed by $T(\gamma^F)$ and $T(\rho^F)$, with the restriction that $T(\rho^F)$ be at least the number of connected components in $\{t\colon \rho^F(t) = 0\}$. For example, the number of pieces in a piecewise-linear function, or the sum of degrees in a piecewise-polynomial function, would qualify. We further impose on $T(\gamma^F)$ and $T(\rho^F)$ the same restriction as for $T(\delta_i^F)$. The total complexity of the production is then $T(F) = \sum_{i=1}^{n-2} T(\delta_i^F) + T(\rho^F) + T(\gamma^F)$.

3. Producible $\equiv$ Flattenable. Key to our main theorem is showing that every $(\geq \alpha)$-producible configuration of a $(\leq a)$-chain can be moved to a canonical configuration, and therefore to every other $(\geq \alpha)$-producible configuration of that chain.

3.1. *Canonical Configuration.* We begin by defining the canonical configuration of $(\leq a)$-chains, called the *α-cone canonical configuration* or *α-CCC*. To understand the constraints of a configuration Q better, consider normalizing all edge vectors $q_i q_{i+1}$ to unit vectors $u_i = (q_{i+1} - q_i)/\|q_{i+1} - q_i\|$ which lie on the unit sphere. The α-CCC is constructed to have the property that all such vectors lie along a circle of radius $\alpha/2$ on that sphere. In other words, the vectors u_i lie on the boundary of a cone with half-angle $\alpha/2$.

To ease the description, we use the cone $\overline{C}_{\alpha/2}$ (not C_α) to define α-CCC, but note that the cone and the chain could be rotated and translated. By convention, we place u_0 on the boundary of $\overline{C}_{\alpha/2}$ in the positive quadrant of the yz-plane. Because Q is a configuration of P, the angle between u_{i-1} and u_i is θ_i and so, on the sphere, u_i lies on the circle of radius θ_i centered at u_{i-1}. Because $\theta_i \leq \alpha$, this circle intersects the boundary of $\overline{C}_{\alpha/2}$. We set u_i to be the first intersection counterclockwise from u_{i-1} on the boundary of $\overline{C}_{\alpha/2}$ (where counterclockwise is viewed from the origin). See Figure 6 for an example.

The position of the u_i's on the unit sphere as described above, along with the position of q_0, uniquely determine the position of the α-CCC of the chain. Because the u_i vectors all have positive z coordinates, we know that the resulting configuration is simple. See Figure 7. We can also show that the α-CCC is completely contained in $\overline{C}_{\alpha/2}$:

LEMMA 2. *If all unit edge vectors u_i are contained in a cone $\overline{C}_\beta$ for some half-angle $\beta > 0$, then the configuration Q is inside $q_0 + \overline{C}_\beta$, the cone translated so its apex is at*

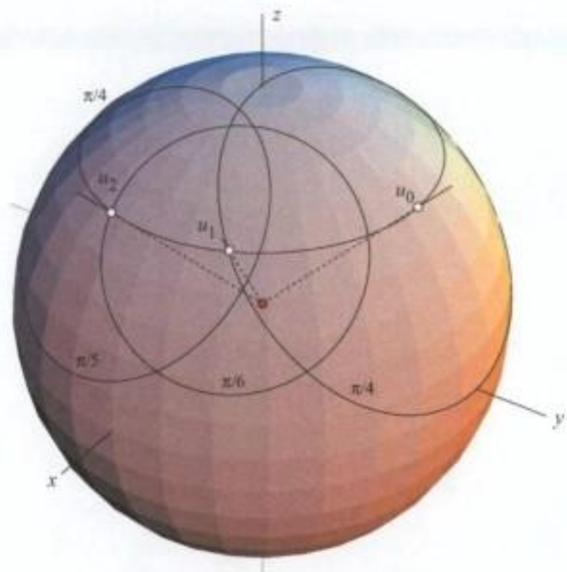


Fig. 6. u_0 lies on the cone $C_{\pi/4}$. $(\theta_1, \theta_2, \theta_3) = (\pi/4, \pi/6, \pi/5)$, respectively.

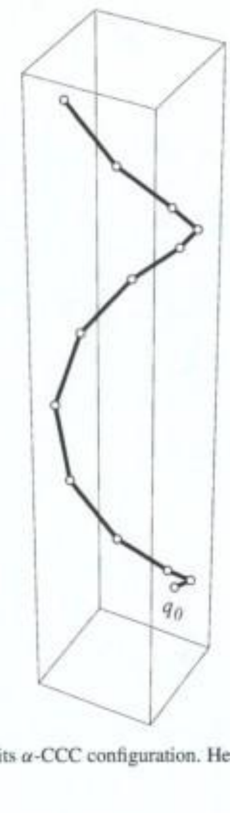


Fig. 7. A chain in its α-CCC configuration. Here $\theta_i = \pi/4$ for all i.

q_0. Furthermore, if $u_0 \neq u_1$, then only the first bar of the chain can touch the boundary of $q_0 + \overline{C}_\beta$.

PROOF. The proof is by induction on n. The claim holds for the one-point chain $Q[n, n]$. Assume $Q[1, n]$ is contained in a cone with apex q_1. Now q_1 is in the cone with apex q_0, so the cone with apex at q_1 is contained in the one with apex at q_0. Furthermore, the boundary of these cones intersect only if q_1 is on the boundary of $q_0 + \overline{C}_\beta$, and in that case the intersection is contained in the line of support q_0q_1. □

In the α-CCC, u_i is always different from u_{i+1}.

3.2. *Canonicalization.* Next we show how to find a motion from any $(\geq \alpha)$-producible configuration of a $(\leq \alpha)$-chain to the corresponding α-CCC.

THEOREM 1. *If a configuration Q of a $(\leq \alpha)$-chain P is $(\geq a)$-producible by a production F, then there is a motion M from Q to the α-CCC, with $T(M) \leq T(F) + 3n$.*

PROOF. Suppose that Q is β-producible for $\beta \geq \alpha$, and that F is a β-production with $F(\infty) = Q$. By scaling time appropriately, we can arrange that $t_i = i$, and the configuration freezes at time $n + 1$, i.e., $F(t) = F(n + 1)$ for $t > n + 1$.

We construct a motion M from Q to the α-CCC, constructed inside $\overline{C}_\beta$. A key idea in our construction is to play the production movements backwards. More precisely, for all $i = 0, \ldots, n$, we define $m_i(t) = f_i(n + 1 - t)$ for the (reverse) time interval $t \in [0, n + 2 - i]$. (Beyond reverse time $n + 2 - i$, the original production time is less than $n + 1 - (n + 2 - i) = i - 1$ and thus f_i is no longer defined.) To complete the construction, we just have to define $m_i(t)$ for $t > n + 2 - i$, that is, the motion of the part of the chain that has already re-entered the cone $\overline{C}_\beta$.

During the time interval $(n - i, n + 1 - i)$, the edge e_i is entering the cone $\overline{C}_\beta$ through the origin, $P[0, i]$ is outside C_β, and $P[i + 1, n]$ is inside C_β. We maintain the invariant that $P[i, n]$ is in α-CCC, contained in a cone $\overline{C}_{\alpha/2}$ translated and rotated to some position $\overline{C}'_{\alpha/2}$. See Figure 8. So the dihedral angle of e_j does not change for $j > i$, i.e., $P[i + 1, n]$ is held rigid. Because $P[0, i]$ moves freely outside of C_β according to the reversed movements of the β-production, we can only control the dihedral angle of e_i in order to maintain that $\overline{C}'_{\alpha/2}$ (and so $P[i + 1, n]$) stays inside $\overline{C}_\beta$.

Again, consider the vectors u_j. The invariant means that all u_j, $j = i, \ldots, n - 1$, touch the boundary of some circle σ of radius $\alpha/2$ on the unit sphere centered on the apex of the cone, and σ must be inside $\overline{C}_\beta$. For any position u_i, we place σ so that its center is on the great arc between u_i and u_{+z}, where u_{+z} is the unit vector along the the z-axis. This implies that u_i is the farthest point from u_{+z} on σ and since, by the production constraints, u_i is in $\overline{C}_\beta$, σ is in $\overline{C}_\beta$ as well and the invariant is satisfied. As long as $u_i \neq u_{+z}$, this position of σ is unique and the resulting motion is continuous because the production is continuous. When $u_i = u_{+z}$, a discontinuity might be introduced, but these discontinuities can easily be removed by stretching the moment of time at which a discontinuity occurs and filling in a continuous motion between the two desired states.

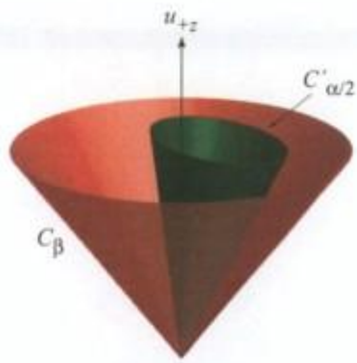


Fig. 8. Cone $C_{\alpha/2}$ is nested inside C_β. The diameter of the former is no more than the radius of the latter.

At time $t = n + 1 - i$, vertex i enters $\overline{C}_\beta$ and the invariant needs to be restored for the next phase. At that time, the vector u_{i-1} lies in $\overline{C}_\beta$, and u_i is on a circle τ of radius θ_i centered at u_{i-1}. Let σ' be the desired new position for σ, that is, the circle whose radius is $\alpha/2$, and whose center is on the great arc between u_{i-1} and u_{+z}. We know that σ' and τ intersect and all intersections are inside $\overline{C}_\beta$ because σ' is in $\overline{C}_\beta$. See Figure 9(a). We first move u_i along τ to the first intersection between σ' and τ counterclockwise from u_{i-1} on σ' (Figure 9(b)) by changing the dihedral angle of e_{i-1}, and simultaneously moving σ accordingly as described above by changing the dihedral angle of e_i. This can be done while maintaining the invariant because the intersection of τ and $\overline{C}_\beta$ is connected. We then rotate σ about u_i to the position σ' (Figure 9(c)) by changing the dihedral angle of e_i. This motion can be done while maintaining the invariant because the set of dihedral angles of e_i for which σ is in $\overline{C}_\beta$ is connected.

The complexity of all dihedral motions outside of C_β is $\sum_{i=1}^{n-2} T(\delta_i^F)$. The dihedral motions of e_i during times $t \in (n - i, n + 1 - i)$ mirror exactly $\gamma^F(n + 1 - t)$, except at discontinuities, which correspond to times for which $u_i = u_{+z}$, which is exactly when

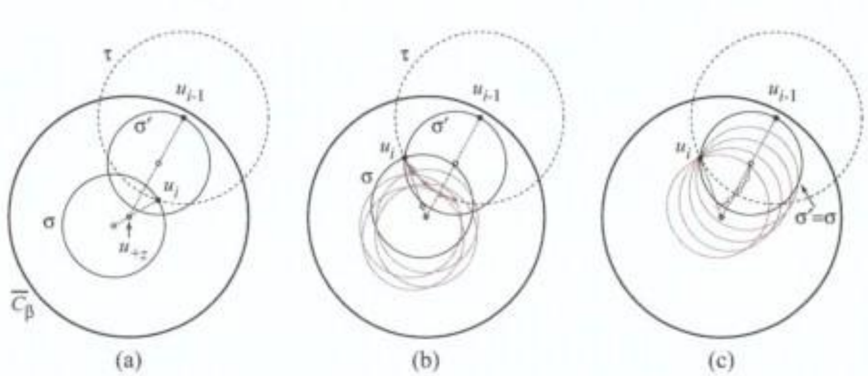


Fig. 9. Restoring the invariant. View looking down u_{+z}. (a) σ and σ' are both radius $\alpha/2$, determined by $\overline{C}_{\alpha/2}$, which moves inside $\overline{C}_\beta$, centered on u_{+z}. τ is of radius θ_i. (b) u_i walks to the counterclockwise point of $\sigma' \cap \tau$. (c) σ is rotated about u_i. Here $\alpha/2 = 30^\circ < \theta_i = 50^\circ < \alpha = 60^\circ < \beta = 70^\circ$.

$\rho^F(n+1-t) = 0$, so the total complexity of these dihedral motions is bounded by $T(\rho^F) + T(\gamma^F)$. Finally, whenever a vertex attains the apex of the cone, we perform three dihedral rotations (linear functions of time) to restore the invariant. Summing it all, we obtain $T(M) \le \sum_{i=1}^{n-2} T(\delta_i^F) + T(\rho^F) + T(\gamma^F) + 3n = T(F) + 3n$. $\square$

COROLLARY 1. *For any two simple ($\ge \alpha$)-producible configurations Q_1 and Q_2 of a common ($\le \alpha$)-chain, with respective productions F_1 and F_2, there is a simple motion M from Q_1 to Q_2—that is, $M(0) = Q_1$ and $M(\infty) = Q_2$—for which $T(M) \le T(F_1) + T(F_2) + 6n$.*

PROOF. Because Q_1 and Q_2 are ($\ge \alpha$)-producible, the previous theorem gives us two motions M_1 and M_2 with $M_1(0) = Q_1$, $M_1(\infty) = \alpha$-CCC, $M_2(0) = Q_2$, and $M_2(\infty) = \alpha$-CCC. By rescaling time, we can arrange that $M_1(t) = M_2(t) = \alpha$-CCC for t beyond some time T. Then define $M(t) = M_1(t)$ for $0 \le t \le T$, $M(t) = M_2(2T-t)$ for $T < t \le 2T$, and $M(t) = Q_2$ for $t > 2T$. $\square$

LEMMA 3. *An α-CCC of a ($\le \alpha$)-chain is β-producible for any $\alpha/2 \le \beta \le \pi - \alpha/2$. The complexity of the production is at most $2n - 1$.*

PROOF. Let Q be a α-CCC positioned in $\overline{C}_{\alpha/2}$ with q_0 at the origin. Let $q(t)$ be the point at distance t from q_0 along Q. The position $F(t)$ of the produced portion of the α-CCC at time t is Q translated so that $q(t)$ is at the origin and deleting all the edges of Q completely inside $C_{\alpha/2}$. By Lemma 2, all edges of $F(t)$ except for the edge containing the origin are contained in the cone $B_{\alpha/2}$. F is thus a valid β-production for any $\alpha/2 \le \beta \le \pi - \alpha/2$. The β-production does not use any dihedral rotation so $T(\delta_i^F) \le 1$, $\rho^F(t) = \alpha/2$ for all t so $T(\rho^F) \le 1$, and γ^F is constant for every edge, so $T(\gamma^F) \le n$ $\square$

COROLLARY 2. *If a configuration Q of a ($\le \alpha$)-chain has a β-production F for some $\beta \ge \alpha$, then it has a β'-production F' for all $\alpha/2 \le \beta' \le \pi - \alpha/2$ and $T(F') \le T(F) + 5n + 1$.*

PROOF. Using Theorem 1, let M be the motion from Q to an α-CCC, and let M' be the reverse motion from the α-CCC to Q. Let R be the sum of the edge lengths of the chain. The production F' first produces an α-CCC in $B_{\alpha/2}$ using Lemma 3. The α-CCC is then translated by a distance $R/\sin\alpha/2$ in the negative direction along the z-axis. At this point, the sphere centered at q_n and of radius R does not intersect the outside of $B_{\alpha/2}$. Keeping q_n fixed, we perform the motion M' to obtain configuration Q. $\square$

3.3. *Connection to Flat States.* Finally, we relate flat configurations to productions and prove our main result that flattenability is equivalent to producibility.

LEMMA 4. *All flat configurations of a ($\le \alpha$)-chain have a β-production F for any β satisfying $\alpha \le \beta \le \pi/2$. Furthermore, $T(F) \le n$.*

PROOF. Assume the configuration is in the xy-plane. Any such flat configuration can be created using the following process. First, draw e_0 in the xy-plane. Then, for all consecutive edges e_i, create e_i in the vertical plane through e_{i-1} at angle θ_{i-1} with the xy-plane, then rotate it to the desired position in the xy-plane by moving the dihedral angle of e_{i-1}. During the creation and motion of e_i, it is possible to enclose it in some continuously moving cone C of half-angle β whose interior never intersects the xy-plane: at the creation of e_i, C is tangent to the xy-plane on the support line of e_{i-1} and with its apex at p_i, and thus contains e_i. During the rotation of e_i, e_i will eventually touch the boundary of C. We then move C along with e_i so that both e_i and the xy-plane are tangent to C. When e_i reaches the xy-plane, we translate C along e_i until its apex is p_{i+1}. Viewing the construction relative to C and placing C on C_β gives the desired β-production. □

COROLLARY 3. *$(\leq \pi/2)$-chains are flat-state connected. The motion between any two flat configurations uses at most $8n$ dihedral motions.*

PROOF. Consider two flat configurations Q and Q' of a $(\leq \pi/2)$-chain. By Lemma 4, Q and Q' are both $(\pi/2)$-producible, and so, by Corollary 1, there exists a motion M such that $M(0) = Q$ and $M(+\infty) = Q'$. □

COROLLARY 4. *All α-producible configurations of $(\leq \alpha)$-chains are flattenable, provided $\alpha \leq \pi/2$. For a production F, the flattening motion M has complexity $T(M) \leq T(F) + 7n$.*

PROOF. Consider an α-producible configuration Q of an $(\leq \alpha)$-chain P. Because $\alpha \leq \pi/2$, the chain P also has a flat configuration Q' [ADD$^+$]. By Lemma 4, Q' is producible, and so by Corollary 1, there exists a motion M such that $M(0) = Q$ and $M(+\infty) = Q'$. The bound on $T(M)$ is by composition of the bounds in Lemma 4 and Corollary 1. □

We note that the restriction in our results to $\alpha \leq \pi/2$ accords with the generally obtuse (about 110°) protein bond angles, which correspond to turn angles α of about 70°.

4. A More Powerful Machine. We now show that our result does not hold without the assumption $\alpha \leq \beta$, under a somewhat stronger model of production that also breaks the $\alpha \leq 2\beta$ claim of Lemma 1.

The stronger model of production separates the creation of the next vertex v_{i+1} from the moment that the previous vertex v_i reaches the origin. Specifically, we suppose that v_{i+1} is not created at t_i, but rather imagine the time instant t_i to be stretched into a positive-length interval $[t_i, t_i']$, allowing time for $v_i v_{i-1}$ to rotate exterior to the cone prior to the creation of v_{i+1} (at time t_i'). This flexibility removes the connection in Lemma 1 between the half-angle β of the cone and the turn angles α produced, permitting chains of large turn angle from any cone. Indeed, the sequence of motions depicted in

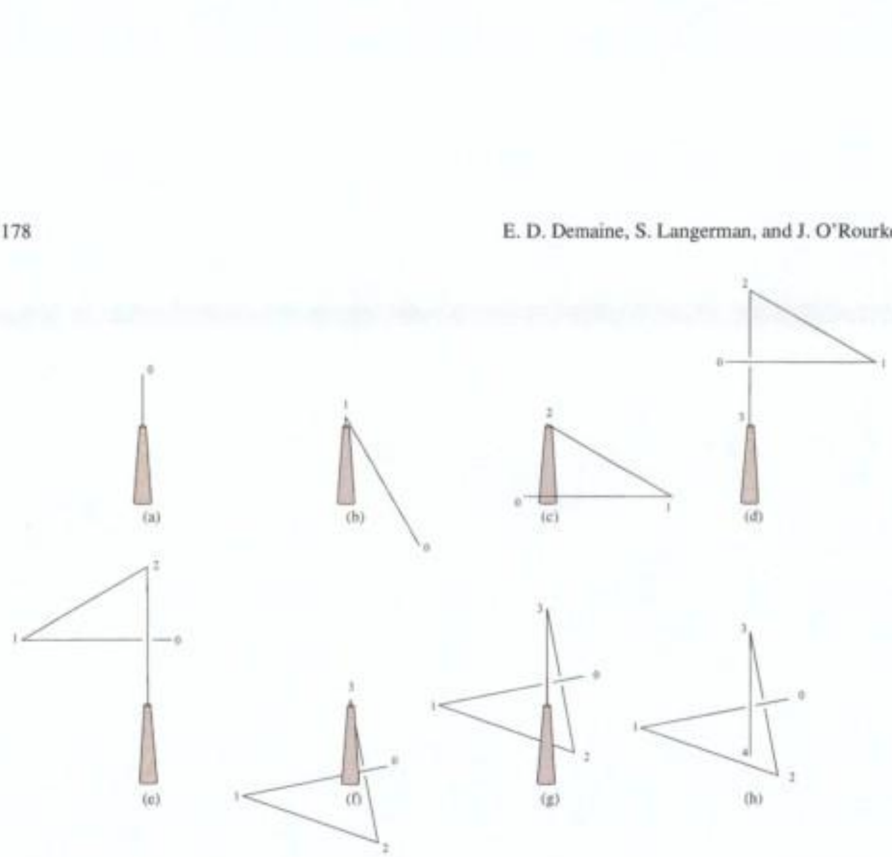


Fig. 10. Production of a locked chain under a model that permits large turning angles to be created. For clarity, the cone is reflected to aim upward. (a) $e_0 = (q_0, q_1)$ emerges; (b) turn at q_1; (c) turn at q_2 and dihedral motion at q_1 places e_1 in front of cone; (d) e_2 nearly fully produced; (e) chain spun about e_2 (or viewpoint changed); (f) rotation at q_3 away from viewer places chain behind cone; (g) e_3 emerges; (h) final locked chain shown loose; the turn angle θ_3 at q_3 can be made arbitrarily close to π.

Figure 10 exploits this large-angle freedom to emit a 4-link fixed-angle chain that is locked.

5. Random Chains. This section proves that the producible/flattenable configurations are a vanishingly small subset of all possible configurations of a chain, for almost any chain. Essentially, the results below say that if there is one configuration of one chain in a class that is unflattenable, then a randomly chosen configuration of a randomly chosen chain from that class is unflattenable with probability approaching 1 geometrically as the number of links in the chain grows. Furthermore, this result holds for any "reasonable" probability distribution on chains and their configurations.

To define probability distributions, it is useful to embed chains and their configurations into Euclidean space. A chain $P = \langle \theta_1, \ldots, \theta_{n-1}; d_0, \ldots, d_{n-1} \rangle \in [0, \pi/2]^{n-1} \times [0, \infty)^n$ is specified by its turn angles θ_i and edge lengths d_i. A configuration $Q = \langle \delta_1, \ldots, \delta_{n-2} \rangle \in [0, 2\pi)^{n-2}$ of P is specified by its dihedral angles. We also need to be precise about our use of the term "unflattenable" for chains versus configurations. A simple configuration Q is *unflattenable* or simply *locked* if it cannot reach a flat configuration; a chain P is *lockable* if it has a locked configuration.

We consider the following general model of random chains of size n. Call a probability distribution *regular* if it has positive probability on any positive-measure subset of some open set called the *domain*, and has zero probability density outside that domain.[7] For Euclidean d-space $\mathbb{R}^d$, a probability distribution is regular if it has positive probability on any positive-radius ball inside the domain. Uniform distributions are always regular.

For chains of k links, we emphasize the regular probability distribution $\mathcal{P}_k^{\Theta,\mathcal{D}}$ obtained by drawing each turn angle θ_i independently from a regular distribution Θ, and drawing each edge length d_i independently from a regular distribution $\mathcal{D}$. Similarly, for not-necessarily-simple configurations of a fixed chain P, we emphasize the regular probability distribution obtained by drawing each dihedral angle δ_i independently from a regular distribution Δ. We can modify this probability distribution to have a domain of all simple configurations of P instead of all configurations of P, by zeroing out the probability density of nonsimple configurations, and rescaling so that the total probability is 1. The resulting distribution is denoted $\mathcal{Q}^{P,\Delta}$, and it is regular because of the following well-known property:

LEMMA 5. *The subspace of simple configurations of a chain P is open.*

PROOF. Consider the space $[0, 2\pi)^{k-2}$ of all configurations of P. The simplicity of a configuration Q of P can be expressed by the $O(k^2)$ constraints that no two nonadjacent segments intersect. These (semi-algebraic) constraints are all of the form $g(Q) < 0$ where $g(Q) = g(\delta_1, \ldots, \delta_{k-2})$ is a multinomial of a constant number of terms in $\sin(\delta_i)$ and $\cos(\delta_i)$. Each constraint defines an open set in the configuration space. The conjunction of the constraints corresponds to the intersection of these finitely many sets, which is open. □

First we show that individual locked examples immediately lead to positive probabilities of being locked. The next lemma establishes this property for configurations of chains, and the following lemma establishes it for chains.

LEMMA 6. *For any regular probability distribution $\mathcal{Q}$ on simple configurations of a lockable chain P, if there is a locked simple configuration in the domain of $\mathcal{Q}$, then the probability of a random simple configuration Q of P being locked is at least a constant $c > 0$.*

PROOF. Let Q' be a locked simple configuration in the domain of $\mathcal{Q}$. Let C be the component of the space of simple configurations containing Q', and let D be the intersection of C and the domain of $\mathcal{Q}$. Because C is open and thus D is open, there exists a constant $\varepsilon > 0$ such that the ball B of radius ε centered at Q' is contained in D, and all $Q'' \in B$ are locked as well. Choose c to be the probability of choosing a configuration in B, which is positive by regularity. □

LEMMA 7. *For any regular probability distribution $\mathcal{P}$ on chains, if there is a lockable chain in the domain of $\mathcal{P}$, then the probability of a random chain P being lockable is at least a constant $\rho > 0$.*

[7] A closely related but more specific notion of regular probability distributions in one dimension was introduced by Willard [Wil] in his extensions to interpolation search.

PROOF. Consider the space of all chains and configurations of those chains, $\mathcal{C} = [0, \pi/2]^{n-1} \times [0, \infty)^n \times [0, 2\pi)^{n-2}$. As described in Lemma 5, the constraint that a particular configuration is locked can be phrased as a set of open semi-algebraic constraints, except now the constraints depend on all $3n - 3$ variables (not just the dihedral angles). Intersecting all these open semi-algebraic sets results in a subspace $\mathcal{L} \subset \mathcal{C}$ of all locked configurations of all chains. Projecting this open set down to $\mathcal{L}' \subseteq [0, \pi/2]^{n-1} \times [0, \infty)^n$ by dropping the dihedral angles results in another open semi-algebraic set, because open semi-algebraic sets are closed under projection.

Now let P' be a lockable chain in the domain of $\mathcal{P}$, let C be the component of $\mathcal{L}'$ containing P', and let D be the intersection of C and the domain of $\mathcal{P}$. Because C, and thus D, is open, there is a constant $\varepsilon > 0$ such that the ball B of radius ε centered at P' is contained in D, and all $P'' \in B$ are lockable. Choose ρ to be the probability of choosing a chain in B, which is positive by regularity. □

Next we show that these positive-probability examples of being locked lead to increasing high probabilities of being locked as we consider larger chains.

THEOREM 2. *Let P_n be a random chain drawn from the regular distribution $\mathcal{P}_n^{\Theta,\mathcal{D}}$. If there is a lockable chain in the domain of $\mathcal{P}_n^{\Theta,\mathcal{D}}$ for at least one value of n, then*

$$\lim_{n\to\infty} Pr[P_n \text{ is lockable}] = 1.$$

$\lim_{n\to\infty} Pr[P_n \text{ is lockable}] = 1$. *Furthermore, if Q_n is a random simple configuration drawn from the regular distribution $\mathcal{Q}^{P_n,\Delta}$, then*

$$\lim_{n\to\infty} Pr[Q_n \text{ is flattenable}] = \lim_{n\to\infty} Pr[Q_n \text{ is producible}] = 0.$$

$\lim_{n\to\infty} Pr[Q_n \text{ is flattenable}] = \lim_{n\to\infty} Pr[Q_n \text{ is producible}] = 0$. *Both limits converge geometrically.*

PROOF. Suppose there is a lockable chain of k links. By Lemma 7, $\Pr[P_k \text{ is lockable}] > \rho > 0$. Break P_n into $\lfloor n/k \rfloor$ subchains of length k. Each of these subchains is chosen independently from $\mathcal{P}_k^{\Theta,\mathcal{D}}$ and is not lockable with probability $< 1 - \rho$. Now P_n is lockable (in particular) if any of the subchains are lockable, so the probability that P_n is not lockable is $< (1-\rho)^{\lfloor n/k \rfloor}$ which approaches 0 geometrically as n grows. Likewise, by Lemma 6, the probability that Q_k is locked is $> c\rho$ for some constant $0 < c < 1$, and so the probability that Q_n is flattenable is $< (1 - c\rho)^{\lfloor n/k \rfloor}$ which approaches 0 geometrically as n grows. □

Thus, producible configurations of chains become rare as soon as one chain in the domain of the distribution is lockable. The locked "knitting needles" example of [CJ] and [BDD+] can be built with chains satisfying $\alpha \le \pi/2$ by replacing the acute-angled universal joints with obtuse, fixed-angled chains of very short links. Thus for any regular distribution including such examples in its domain, we know that configurations of $(\le \alpha)$-chains are rarely producible for the case we have considered, $\alpha \le \pi/2$. We do not know of any nontrivial regular probability distribution $\mathcal{P}_n^{\Theta,\mathcal{D}}$ whose domain has no lockable chains. In particular, for equilateral (all edge-lengths equal) fixed-angle chains, it is not known whether angle restrictions can prevent the existence of locked configurations. As protein backbones are nearly equilateral, it is of particular interest to answer this question.

Future directions for research include resolving the locked question just mentioned, incorporating the short side-chains that jut from the protein backbone, and more realistically modeling the ribosome structure.

Acknowledgements. Much of this work was completed at the Workshop on Geometric Aspects of Molecular Reconfiguration organized by Godfried Toussaint at the Bellairs Research Institute of McGill University in Barbados, February 2002. We appreciate the helpful discussions with the other participants: Greg Aloupis, Prosenjit Bose, David Bremner, Vida Dujmović, Herbert Edelsbrunner, Jeff Erickson, Ferran Hurtado, Henk Meijer, Pat Morin, Mark Overmars, Suneeta Ramaswami, Ileana Streinu, Godfried Toussaint, and especially Yusu Wang.

References

[ADD⁺] G. Aloupis, E. Demaine, V. Dujmović, J. Erickson, S. Langerman, H. Meijer, I. Streinu, J. O'Rourke, M. Overmars, M. Soss, and G. Toussaint. Flat-state connectivity of linkages under dihedral motions. In *Proc. 13th Annu. Internat. Sympos. Alg. Comput.*, volume 2518 of *Lecture Notes in Computer Science*, pages 369–380. Springer, Berlin, 2002.

[ADM⁺] G. Aloupis, E. D. Demaine, H. Meijer, J. O'Rourke, I. Streinu, and G. Toussaint. Flat-state connectedness of fixed-angle chains: special acute chains. In *Proc. 14th Canad. Conf. Comput. Geom.*, pages 27–30, 2002.

[BDD⁺] T. Biedl, E. Demaine, M. Demaine, S. Lazard, A. Lubiw, J. O'Rourke, M. Overmars, S. Robbins, I. Streinu, G. Toussaint, and S. Whitesides. Locked and unlocked polygonal chains in three dimensions. *Discrete Comput. Geom.*, 26(3):269–282, 2001.

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

[NHB⁺] P. Nissen, J. Hansen, N. Ban, B. Moore, and T. A. Steitz. The structural basis of ribosome activity in peptide bond synthesis. *Science*, 289:920–930, August 2000.

[ST] M. Soss and G. T. Toussaint. Geometric and computational aspects of polymer reconfiguration. *J. Math. Chem.*, 27(4):303–318, 2000.

[Wil] D. E. Willard. Searching unindexed and nonuniformly generated files in log log N time. *SIAM J. Comput.*, 14:1013–1029, 1985.

Computing Signed Permutations of Polygons *

Greg Aloupis † Prosenjit Bose‡ Erik D. Demaine§
Stefan Langerman† Henk Meijer¶ Mark Overmars‖
Godfried T. Toussaint†

Abstract

Given a planar polygon (or chain) with a list of edges $\{e_1, e_2, e_3, \ldots, e_{n-1}, e_n\}$, we examine the effect of several operations that *permute* this edge list, resulting in the formation of a new polygon. The main operations that we consider are: *reversals* which involve inverting the order of a sublist, *transpositions* which involve interchanging subchains (sublists), and *edge-swaps* which are a special case and involve interchanging two consecutive edges. Using these permuting operations, we explore the complexity of performing certain actions, such as convexifying a given polygon or obtaining its mirror image. When each edge of the given polygon has also been assigned a *direction* we say that the polygon is *signed*. In this case any edge involved in a reversal changes direction. The complexity of some problems varies depending on whether a polygon is signed or unsigned. An additional restriction in many cases is that polygons remain simple after every permutation.

1 Introduction

Much focus has been placed recently on the problem of sorting a permutation of n integers by *reversals* [HP98, BP95]. As one might guess, a single reversal

*This research was partly funded by NSERC and FCAR.

†School of Computer Science, McGill University. {athens,sl,godfried}@uni.cs.mcgill.ca

‡Carleton University. jit@scs.carleton.ca

§MIT Laboratory for Computer Science, 200 Technology Square, Cambridge, MA 02139, USA, edemaine@mit.edu

¶Queens University. henk@cs.queensu.ca

‖University of Utrecht. markov@cs.uu.nl

is applied to a consecutive set of these integers and the result is that their order is inverted. The key problem that arises is determining the minimum number of reversals necessary to sort a given permutation. This number is called the *reversal distance* of the permutation. A variation of this problem involves *signed* permutations [BMY01]. In this case any integer affected by a given reversal also changes *parity*.

Each of these interesting combinatorial problems has its roots in bioinformatics and molecular biology [HP98, HP96, BP95, CJM$^+$00, BMY01]. Specifically, genomes have been modeled as linear or cyclic sequences, where each element in a sequence is a *block* of smaller elements that are never separated. A popular model for mutation involves inverting parts of these sequences. In order to determine the number of such mutations needed to transform one genome to another, one may compute the reversal distance of the associated permutations. An extension of this model is to consider the direction of each *block*. This leads to the study of signed permutations. We illustrate signed inversions in Figure 1 which has been modified from [BP95].

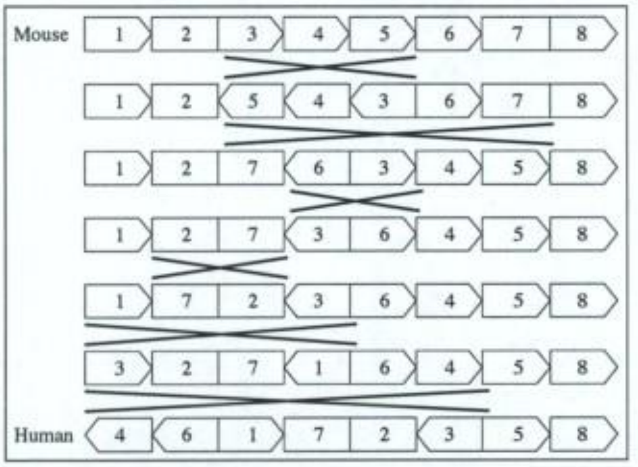


Figure 1: A most parsimonious evolutionary scenario for the transformation of human into mouse chromosome assuming that the X chromosome evolves solely by inversions [BP95]. Each block represents a conserved linkage group of genes. Reversal distance is equal to six.

The problem of computing *transposition distance* also stems from bioinformatics. In this case, a transposition involves exchanging two disjoint sets of consecutive integers in a permutation. Computing reversal distance has been shown to be NP-hard [Cap99] for unsigned permutations, but for the signed version a linear time algorithm exists [BMY01]. Computing transposition distance is of unknown complexity [CJM+00]. The reader may also be interested in [BHK01, BP98, EEK+01].

In this paper we extend the ideas mentioned above from one dimension to two. Instead of considering permutations of integers, we consider permutations of *edges* which form polygons or chains. We define operations such as *edge-swaps*, *reversals* and *transpositions*, in analogy to $\mathbb{R}^1$. We introduce the notions of *signed permutations of polygons and chains.* These concepts give rise to a wide range of problems to be solved.

2 Definitions

First we introduce the notion of a *signed polygon* or *signed permutation* of a polygon. Any polygon P may be described by a list of edges $\{e_1, e_2, e_3, \ldots, e_{n-1}, e_n\}$. A signed polygon is no different, except that each edge is also assigned a direction. The same holds for chains. This is a generalization of the notion of parity that is used in $\mathbb{R}^1$. If the directions of all edges are consistent as we traverse a polygon or chain, then this polygon or chain is *oriented.* In Figure 2 we illustrate some signed polygons and chains.

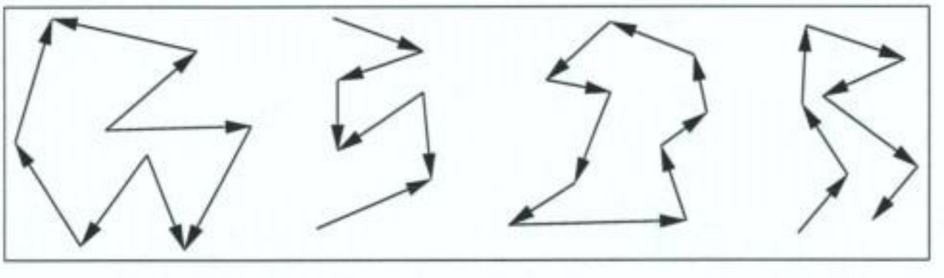


Figure 2: From left to right, a signed polygon, signed chain, oriented polygon, oriented chain.

Without loss of generality, suppose that we are dealing with an oriented polygon. A *transposition* of two edges A and B involves interchanging their positions so that the resulting polygon remains oriented. This is illustrated in

Figure 3. If A and B are consecutive, this operation is defined as an *edge-swap* or plainly *swap* (Figure 4a). It is not difficult to see that entire subchains may also be transposed. A *single-edge* transposition involves transposing an edge with an empty subchain. One may also think of this operation as a transposition between the single edge and one of its neighboring subchains (Figure 4b).

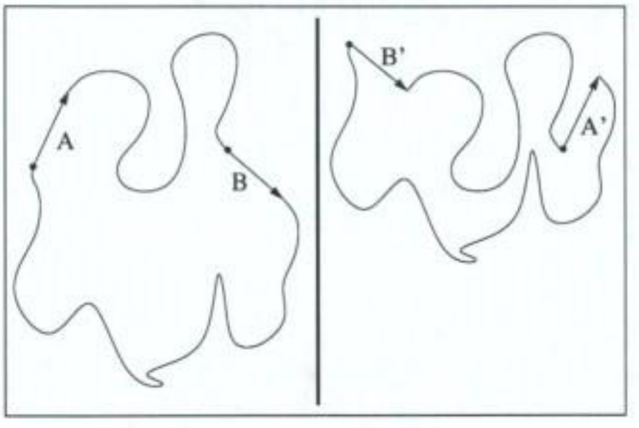


Figure 3: Transposing two edges A and B.

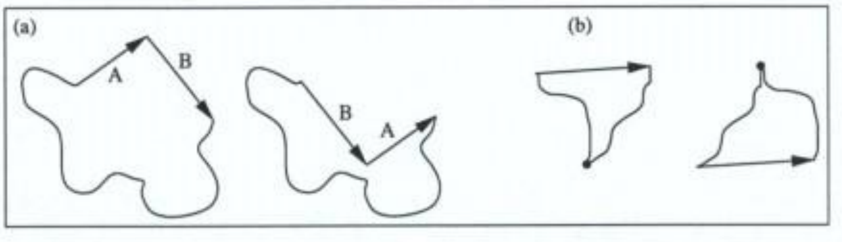


Figure 4: (a) An edge swap. (b) A single-edge transposition.

A *reversal* of a subchain belonging to a polygon involves inverting the order of the edges in the subchain. Geometrically this rotates the subchain rigidly in the plane by an angle of π so that its endpoints are placed exactly at each other's original location. For unsigned polygons this operation appears identical to the *flipturn* operation introduced by Joss and Shannon [GZ01]. However, here we allow reversals to take place on any subchain, not only on pockets. For signed polygons the direction of each edge involved in the reversal is switched, as is done for parity in $\mathbb{R}^1$. In Figure 5 we illustrate a reversal of subchain $\{e_i, \ldots, e_j\}$ for a signed (initially oriented) polygon. One

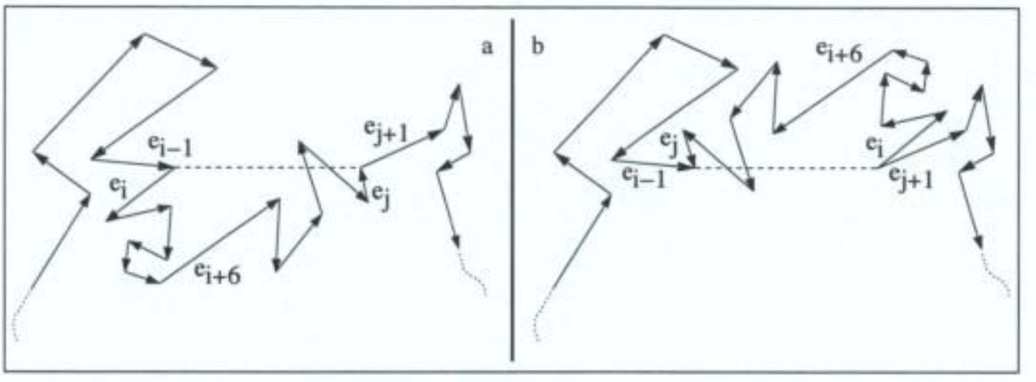


Figure 5: Reversing a subchain of a signed polygon.

can see that for unsigned polygons, an edge-swap is merely a transposition *or* a reversal of two consecutive edges. For signed polygons there is a difference in the resulting direction of each edge.

Each of the operations above results in the same shape when used on a polygon, regardless of the directions of its edges. In other words, to compute how a polygon changes shape, one can imagine that it is oriented. However, for chains alternate definitions exist. For example consider the oriented chain in Figure 6. We may choose to perform a reversal on edges (A, B, C) in at least two ways. One way (shown on top) is identical to what is done for polygons. This is convenient but also means that the endpoints of the chain will never move. A second way (shown at the bottom) is to preserve orientation. This may allow the chain to form more interesting configurations. We use the latter definition in Theorem 3.6 in the next section.

3 Permuting Polygons

Scott [Sco82] has shown that precisely two permutations of an edge list form oriented convex polygons, and these have maximal area. It is also known that if the longest edge of a polygon has unit length, this polygon may be permuted to fit into a circle of radius $\sqrt{5}$ [GY79].

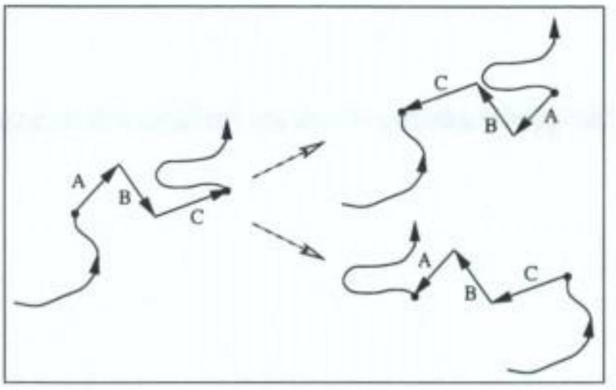


Figure 6: Two ways that a reversal may be defined on a chain.

For the remainder of this section we present our results concerning permutations of polygons or chains. We impose the restriction that simplicity must be maintained at all times, unless mentioned otherwise.

In Figure 7a we show a polygon which does not admit any edge-swaps. Examples such as this one may be extended easily to create any n-gon which will not admit edge-swaps. In Figure 7b we show a polygon which does not admit single-edge transpositions, with the exception of a few edge-swaps for some edges that are almost collinear. These transpositions cannot change the basic shape. Thus we see that sometimes local permutations will not be sufficient to achieve desired reconfigurations.

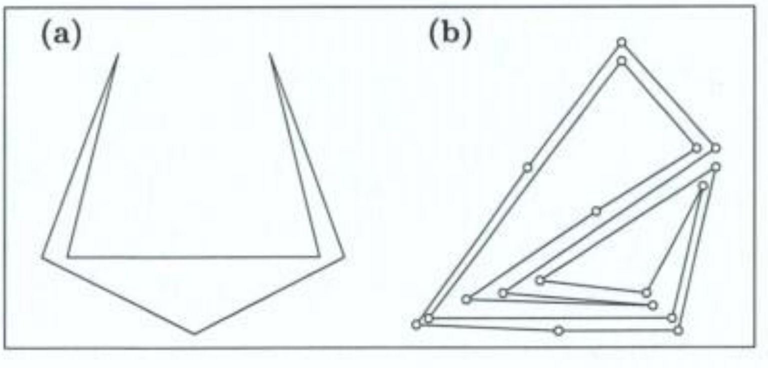


Figure 7: Polygon (a) does not admit edge-swaps. Polygon (b) does not admit single-edge transpositions.

Theorem 3.1 *A simple polygon may be convexified with $O(n^2)$ reversals while maintaining simplicity after each reversal.*

Proof: This result holds for the more restricted reversal operation of flipturns [ABC+00]. □

Theorem 3.2 *A star-shaped polygon can be convexified with $O(n^2)$ edge-swaps while maintaining star-shapedness after each edge-swap, and this bound is tight in the worst case.*

Proof: Let k be a point in the kernel and without loss of generality suppose that the polygon is oriented clockwise. If the polygon is not convex, there must exist two successive edges $\overrightarrow{ab}$ and $\overrightarrow{bc}$ which form a left hand turn (see Figure 8a). Since the polygon is star-shaped, b is the only vertex in

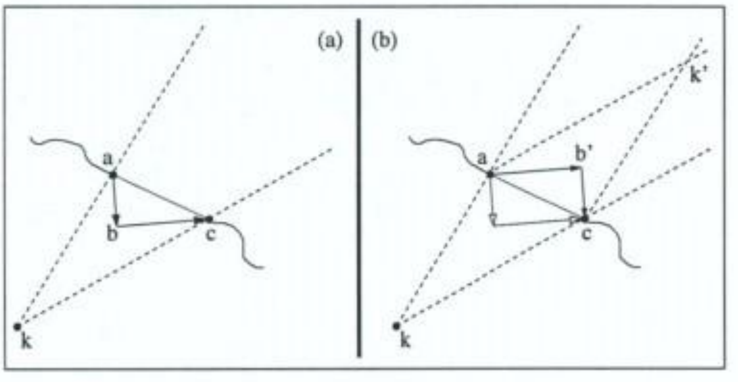


Figure 8: Using edge-swaps to convexify a star-shaped polygon.

the cone formed by the half-lines ka and kc. If we edge-swap $\overrightarrow{ab}$ and $\overrightarrow{bc}$, we obtain the configuration shown in Figure 8b. The new position of b (shown as b') must be somewhere in the triangle (a, c, k'). The swapped edges are still visible from k, and they do not interfere with the other edges of the polygon. Thus the polygon remains star-shaped. Furthermore any point in the kernel remains in the kernel and any point in the polygon remains in the polygon.

Every edge e may be found only within a halfplane determined by a line parallel to e that passes through k. Now suppose that two edges, $\overrightarrow{ab}$ and $\overrightarrow{cd}$ form a right hand turn. This means that b and c coincide as shown

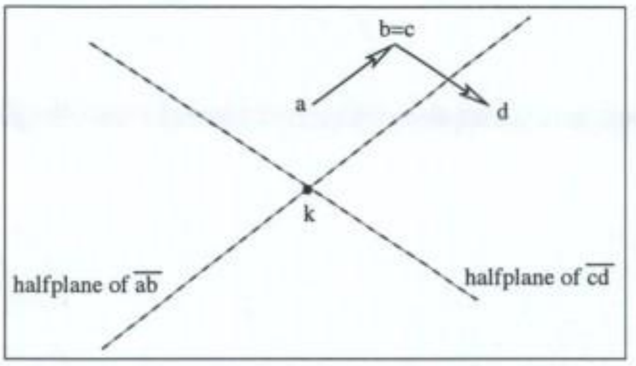


Figure 9: Any pair of edges may be swapped at most once.

in Figure 9. It is impossible to move these edges within their respective halfplanes and into a left hand turn without obstructing visibility from k to either b or c. Thus once a pair of edges forming a left hand turn are swapped, they will never form a left hand turn again. The polygon will become convex only when there are no swaps to be made on left hand turns. Since any pair of edges may be swapped at most once, $O(n^2)$ swaps suffice to convexify a star-shaped polygon. In Figure 10 we show that this bound is tight. Every edge e_i $(2 \leq i \leq n-2)$ must be swapped with edges $e_1, \ldots, e_{i-1}$ for the polygon to become convex. □

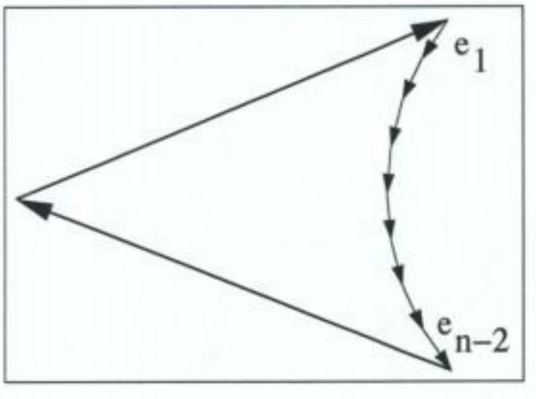


Figure 10: A star-shaped polygon which requires $\Omega(n^2)$ edge-swaps to become convex.

For the following two theorems we do not enforce simplicity.

Theorem 3.3 *Determining whether a signed polygon may be permuted using transpositions so that its shape is rotated by an angle of π takes $\Theta(n \log n)$ time in the algebraic decision tree model of computation.*

Proof: Since only transpositions are allowed, each edge of the polygon must have its *opposite* also present in the polygon. This property also suffices if we do not impose the restriction of maintaining simplicity at all times. By "opposite" we mean an edge with the same angle, but opposite direction. For example, in Figure 11 edges a, b, c, d, e of the polygon on the left are matched by edges $\bar{a}, \bar{b}, \bar{c}, \bar{d}, \bar{e}$. This means that the shape of this polygon may be rotated by an angle of π (shown on right) with appropriate transpositions.

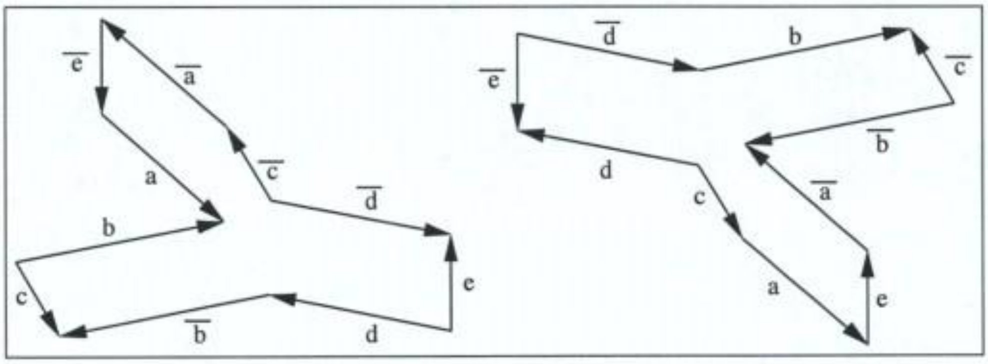


Figure 11: Left: a signed polygon for which every edge is matched by an "opposite". Right: a permutation of the polygon with the same shape rotated by an angle of π.

If we translate every edge to the origin (so that they are directed away from the origin), we obtain a set of n points. The shape of the given polygon can be rotated if and only if every such point has a reflection through the origin. This can be determined in $O(n \log n)$ time with a radial sort, and the matching lower bound is obtained by a reduction from *Set Equality* (see [Ead88]). □

Theorem 3.4 *Determining whether we can obtain the mirror image of a signed polygon using transpositions takes $\Theta(n \log n)$ time in the algebraic decision tree model of computation.*

Proof: In order to be able to obtain a mirror image, there must exist an axis through which every edge has its reflection present (allowing translation). For example consider the polygon on the left in Figure 12. If we take a vertical line as an axis of symmetry, then edges d and j are reflections of each other. The same holds for pairs (b, h) and (f, k). Vertical edges do not need a matching edge. If such an axis exists, then a mirror image of the polygon can be obtained using transpositions. As in Theorem 3.3 we can place every edge at the origin to obtain a set of n points. The symmetries of this point set may be found using the Knuth-Morris-Pratt string matching algorithm [KMP77]. The overall time complexity is $O(n \log n)$. This is pointed out by Eades [Ead88] who also mentions that such reflection tests have $\Omega(n \log n)$ lower bounds on fixed degree decision tree machines. □

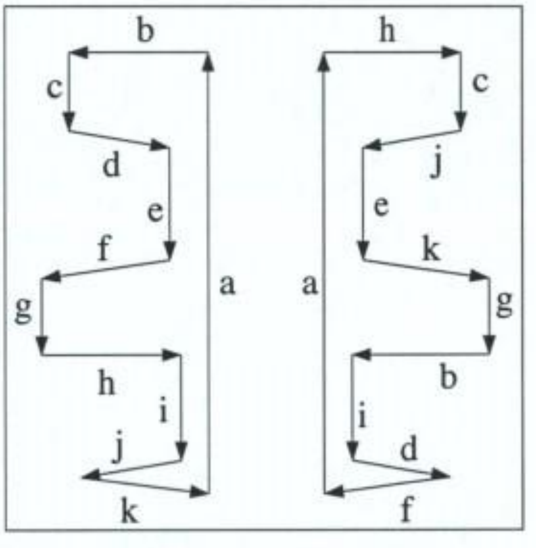


Figure 12: Two polygons that are mirror images and have different permutations of the same edge list.

Theorem 3.5 *Given an oriented polygon P and a rectangle R, deciding whether P can be permuted by transpositions into an oriented polygon P' that can be drawn inside R is (weakly) NP-complete.* [1]

[1] This result also holds if P' is to be placed inside a strip or circle, instead of a rectangle.

Proof: Consider an integer partition problem with $S = \{a_0, a_1, \ldots, a_{n-1}\}$ and $a_i > 0$ for all i. Let $A = \sum_{a \in S} a/2$. Deciding whether there is a subset S' of S with $\sum_{a \in S'} a = A$ is (weakly) NP-hard. Consider the following polygon P. Denote the edges of P in counter-clockwise order by $\{e_0, e_1, \ldots, e_{2n+3}\}$. The edges with even indices are parallel to the x-axis; the edges with odd indices are parallel to the y-axis. Let ϵ be a positive number less than one. The edge e_0 has length $a_0 + \epsilon$. Edges e_i for $i = 2, 4, 6, \ldots, 2n-2$ have length $a_{i/2}$. Edge e_{2n} has length A. Edge e_{2n+2} has length $A + \epsilon$. Edges e_i for $i = 1, 3, 5, \ldots, 2n+1$ have length 1. Edge e_{2n+3} has length $n+1$.

We also assign directions to the edges, so that the edges form a counter-clockwise traversal of P. All edges of length 1 go up. The edges e_i for $i = 0, 2, 4, \ldots, 2n-2$ go from left to right. The few remaining edges go down and right to left, as illustrated in Figure 13(a) with n=7.

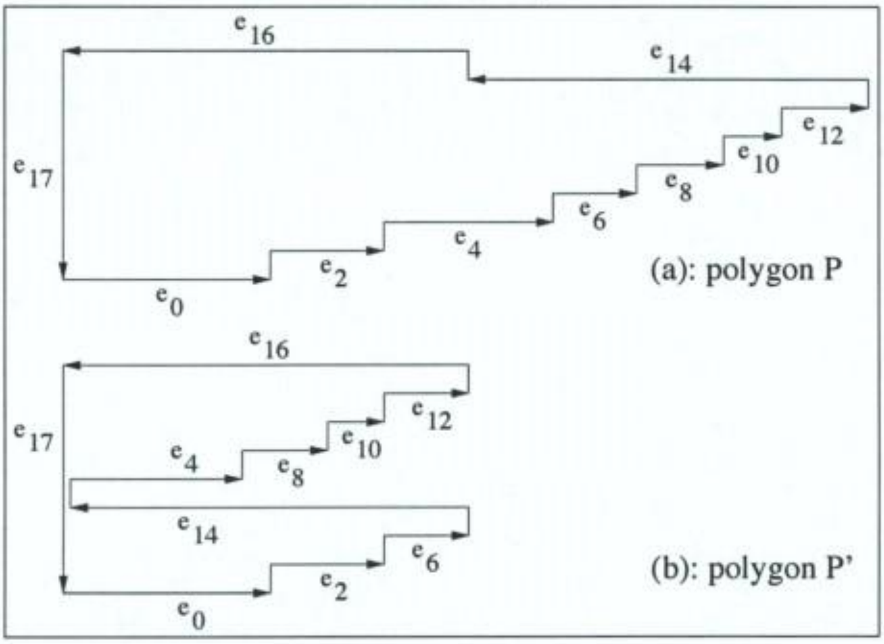


Figure 13: Polygons P and P' with 18 vertices.

Let R be a rectangle of size $A + \epsilon$ by $n + 1$. W.l.o.g assume that R has $(-\epsilon, 0)$ and $(A, n+1)$ as its left-bottom and right-top corner. Suppose P can be permuted into a polygon P' that can be drawn in the rectangle R. Again w.l.o.g assume that e_{2n+2} of P' lies along the top side of R and e_{2n+3}

along the left side of R. This implies that the left and right endpoints of e_{2n} are $(0, y)$ and (A, y) for some value of y with $1 \leq y \leq n$. Moreover the edge e_0 lies below e_{2n}. The edges form a counter-clockwise traversal of P'. Since e_{2n} has a direction that goes from right to left, the horizontal edges above e_{2n} connect the left endpoint of e_{2n} with the right endpoint of e_{2n+2}, so their lengths must add up to A. Therefore the partition has a solution if and only if P can be permuted into a polygon P' that fits in R. Figure 13(b) shows a permutation of the polygon in Figure 13(a) that fits in rectangle R. $\square$

Theorem 3.6 *The maximum endpoint distance over all permutations of an oriented chain may be computed in $O(n \log n)$ time.*

Proof: Fix one endpoint at the origin. Endpoint distance depends only on the direction of each edge. If we knew the direction in which to position the second endpoint, it would be a simple matter to select the direction of each edge in order to maximize the distance. Position two vectors at the origin for each edge, representing its possible directions. Sort the vectors radially and compute the sum of all vectors in one halfplane determined by a line ℓ through the origin. This represents the maximum distance in a direction perpendicular to ℓ. By rotating ℓ and updating the vector sum whenever a vector enters or exits the rotating halfplane, we obtain the endpoint distance over all directions. The time complexity is dominated by the sorting step, so the entire procedure takes $O(n \log n)$ time using $O(n)$ space. $\square$

Acknowledgements

Much of this work was completed at the *Workshop on Geometric Aspects of Molecular Reconfiguration* organized by Godfried Toussaint at the Bellairs Research Institute of McGill University in Barbados, February 2002. We appreciate the helpful discussions with the other participants: David Bremner, Vida Dujmovic, Herbert Edelsbrunner, Jeff Erickson, Ferran Hurtado, Patrick Morin, Joseph O'Rourke, Suneeta Ramaswami, Ileana Streinu, Yusu Wang. We also appreciate the comments of an anonymous referee.

References

[ABC$^+$00] Hee-Kap Ahn, Prosenjit Bose, Jurek Czyzowicz, Nicolas Hanusse, Evangelos Kranakis, and Pat Morin. Flipping your lid. *Geombinatorics*, X(2):57–63, 2000.

[BHK01] Piotr Berman, Sridhar Hannenhalli, and Marek Karpinski. 1.375-approximation algorithm for sorting by reversals. Technical Report DIMACS TR:2001-41, 2001.

[BMY01] David A. Bader, Bernard M.E. Moret, and Mi Yan. A linear time algorithm for computing inversion distance between signed permutations with an experimental study. *Journal of Computational Biology*, 8(5):483–491, 2001.

[BP95] Vineet Bafna and Pavel A. Pevzner. Sorting by reversals: Genome rearrangements in plant organelles and evolutionary history of X chromosome. *Mol. Biol. Evol.*, 12(2):239–246, 1995.

[BP98] Vineet Bafna and Pavel A. Pevzner. Sorting by transpositions. *SIAM J. Discrete Math.*, 11(2):224–240, May 1998.

[Cap99] Alberto Caprara. Sorting permutations by reversals and Eulerian cycle decompositions. *SIAM J. Discrete Math.*, 12(1):91–110, 1999.

[CJM$^+$00] Mary E. Cosner, Robert K. Jansen, Bernard M.E. Moret, Linda A. Raubeson, Li-San Wang, Tandy Warnow, and Stacia Wyman. A new fast heuristic for computing the breakpoint phylogeny and experimental phylogenetic analyses of real and synthetic data. In *Proc. 8th Int'l Conf. on Intelligent Systems for Molecular Biology ISMB-2000, San Diego*, pages 104–115, 2000.

[Ead88] Peter Eades. Symmetry finding algorithms. In *Computational Morphology: A Computational Geometric Approach to the Analysis of Form, Godfried T. Toussaint (ed.), North-Holland.* 1988.

[EEK$^+$01] Henrik Eriksson, Kimmo Eriksson, Johan Karlander, Lars Svensson, and Johan Wästlund. Sorting a bridge hand. *Discrete Math.*, 241(1–3):289–300, October 2001.

[GY79] Lidija I. Golovina and Isaak M. Yaglom. *Induction in Geometry*. Mir Publishers, Moscow, 1979.

[GZ01] Branko Grunbaum and Joseph Zaks. Convexification of polygons by flips and by flipturns. *Discrete Mathematics*, 241:333–342, 2001.

[HP96] Sridhar Hannenhalli and Pavel A. Pevzner. To cut or not to cut (applications of comparative physical maps in molecular evolution). In *Seventh Anuual ACM-SIAM Symposium on Discrete Algorithms*, pages 304–313, 1996.

[HP98] Sridhar Hannenhalli and Pavel A. Pevzner. Transforming cabbage into turnip (polynomial algorithm for sorting signed permutations by reversals). In *Proc. 27th Annual ACM Symposium on the Theory of Computing*, pages 178–189, 1998.

[KMP77] Donald E. Knuth, James H. Morris, and Vaughan R. Pratt. Fast pattern matching in strings. *SIAM Journal of Computing*, 6:322–350, 1977.

[Sco82] Paul R. Scott. On a family of polygons. *Mathematics Magazine*, 55(2):104–105, March 1982.

Vertex Pops and Popturns

Greg Aloupis* Brad Ballinger† Prosenjit Bose‡ Mirela Damian§ Erik D. Demaine¶
Martin L. Demaine¶ Robin Flatland‖ Ferran Hurtado** Stefan Langerman*†† Joseph O'Rourke‡‡
Perouz Taslakian§§ Godfried Toussaint§§

1 Introduction

This paper considers transformations of a planar polygon P according to two types of operations. A *vertex pop* (or a *pop*) reflects a vertex v_i, $i \in \{1, \ldots, n\}$, across the line through the two adjacent vertices v_{i-1} and v_{i+1} (where index arithmetic is modulo n). A *popturn* rotates v_i in the plane by 180° about the midpoint of the line segment $v_{i-1}v_{i+1}$. Pops and popturns are moves similar to "Erdős pocket flips" and "flipturns" [5, 3] in that they preserve the lengths of the polygon edges.

Our goal in this paper is to study which polygons can be convexified by a series of pops or popturns, under various intersection restrictions and definitional variants.

We distinguish between three types of polygons. A *simple polygon* is non-self-intersecting, in that edges intersect only at common endpoints. A polygon is *weakly simple* if its boundary does not "properly cross" itself. Finally, a general *polygon* may be self-intersecting with proper crossings. Pops and popturns can easily introduce weak or proper crossings, so the latter two classes are often more natural to study.

We also focus on two subclasses of polygons. In an *orthogonal polygon*, adjacent edges meet at right angles. In an *equilateral* or *unit polygon*, all edge lengths are equal, say, to 1. In equilateral polygons, pops and popturns become identical operations.

We will see that a vertex pop can create a *hairpin vertex* (or a *pin*): a vertex v_i whose incident edges overlap collinearly. If also $v_{i-1}=v_{i+1}$ (which arises naturally in unit polygons), then the reflection line for a pop of v_i is not determined. Whether to allow a pop of such a pin, and if so, how to define it, leads to many interesting variations, detailed in Section 3 below.

*Dépt. d'Informatique, Univ. Libre de Bruxelles
†Davis School for Independent Study, Davis, CA, USA
‡School of Comput. Sci., Carleton Univ., Ottawa, Canada
§Dept. Comput. Sci., Villanova Univ., USA
¶CSAIL, MIT, USA
‖Dept. Comput. Sci., Siena College, USA
**Dept. Mat. Aplicada II, Univ. Politécnica de Catalunya
††Chercheur qualifié du F.R.S.-FNRS
‡‡Dept. Comput. Sci., Smith College, USA
§§School of Comput. Sci., McGill Univ., Canada

Polygons	Moves	Convexifiable?
arbitrary	popturns	yes, always
simple	popturns	yes iff no purse
weakly simple, unit, orthogonal	pops+180° rot. or pops+untwists	yes, always

Table 1: Summary of our results.

Our results. Table 1 lists our results. If crossings are permitted, it remains unresolved whether every polygon can be convexified via vertex pops, but we show popturns suffice. Restricting to simple polygons, it is known that every star-shaped polygon can be convexified by popturns [1, Thm. 3.2]. We characterize precisely the class of polygons that can be convexified by simple popturns: those without a "purse." Our final result is specialized to unit orthogonal polygons, which can be reconfigured under various hairpin move restrictions.

2 Popturns

The polygon P with clockwise vertices $(v_1, v_2, \ldots, v_n)$ can be seen as a cyclic sequence of rooted vectors $(e_1, \ldots, e_n)$, where $e_i = (v_{i-1}, v_i)$. A sequence of vectors is *simple* if they form a simple polygon, and *clockwise* if each vector has the interior of P on its right side. In the following, we will use the terms *sequence of vectors* and *polygon* interchangeably. A popturn then corresponds to swapping two adjacent vectors in their cyclic ordering. We call the popturn *weakly simple* if the resulting polygon is simple. The two vectors to be popturned and their images form a parallelogram. We call the popturn *simple* if this parallelogram does not intersect or contain P. This is the case if and only if the resulting polygon is simple and clockwise.

If we permit crossings, popturns can convexify by simulating bubble sort on edge directions, where each adjacent swap corresponds to a popturn; see [2, p. 32].

Theorem 1 *Any polygon of n vertices can be convexified (permitting crossings) by a sequence of at most $\frac{1}{2}\binom{n}{2}$ popturns.*

In the remainder of this section, we concentrate on simple polygons and simple popturns. The *turning angle* $\tau_i = \tau_{v_i}$ at vertex v_i is the clockwise angle between the vectors e_i and e_{i+1} ($-\pi < \tau_i < \pi$), and the *total turning angle* $\tau_{i,j} = \tau_{e_i,e_j}$ between edges e_i and e_j is $\sum_{l=i}^{j-1} \tau_l$. Because the polygon is closed, simple, and clockwise, $\tau_{i,j} + \tau_{j,i} = 2\pi$. Notice that the total turning angle τ_{e_i,e_j} between two edges, e_i and e_j does not change after a simple popturn unless e_i and e_j are adjacent, i.e., $j = i+1$ or $j = i-1$, and the popturn is performed at their common vertex. Consider, for example, Fig. 1(a), in which a popturn at v_3 reorders the sequence of vectors $\{\ldots, e_1, e_2, e_3, e_4, \ldots\}$ to $\{\ldots, e_1, e_3, e_2, e_4, \ldots\}$. Only two edge-turning angles change: $\tau_{e_2,e_3} = \frac{1}{4}\pi$ becomes $\tau'_{e_2,e_3} = (2+\frac{1}{4})\pi$, wrapping around the entire polygon; meanwhile, $\tau_{e_3,e_2} = 1\frac{3}{4}\pi$ becomes $\tau'_{e_3,e_2} = -\frac{1}{4}\pi$.

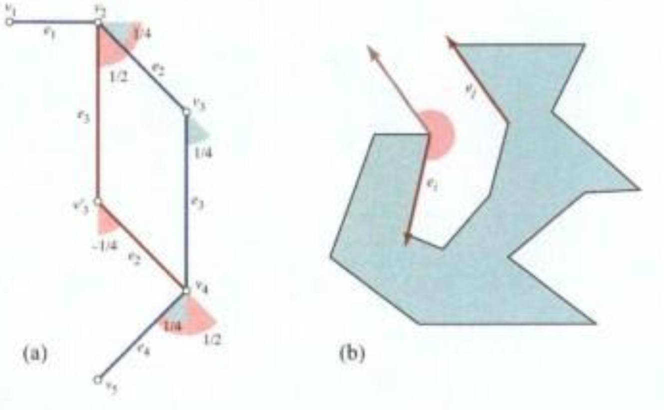


Figure 1: (a) Popturn at v_3 of $\{e_2, e_3\}$. (b) Purse $e_i, \ldots, e_j$.

A *purse* is a (cyclic) subsequence $e_i, \ldots, e_j$ such that $\tau_{i,j} \leq -\pi$; see Fig. 1(b). We show the following:

Lemma 2 *If $e_i, \ldots, e_j$ is a purse, then e_i and e_j can never be made adjacent by any sequence of simple popturns, and $\tau_{i,j}$ is constant.*

Proof: As stated previously, τ_{e_i,e_j} will be affected by a popturn only if e_i and e_j are adjacent, i.e., $j = i+1$ or $j = i-1$. In the first case, $\tau_{i,j} = \tau_{i,i+1}$ must be strictly between $-\pi$ and π. In the second case, $\tau_{i,j} = \tau_{j+1,j} = 2\pi - \tau_{j,j+1}$, which is strictly between π and 3π. However, purse $e_i, \ldots, e_j$ has $\tau_{i,j} \leq -\pi$, meeting neither case. Before e_i and e_j become adjacent, $\tau_{i,j}$ must change, but before $\tau_{i,j}$ can change, e_i and e_j must become adjacent. Thus e_i and e_j will never become adjacent. □

A vertex v_i is *reflex* if $\tau_i < 0$. A popturn at a reflex vertex is called a *reflex popturn*.

Lemma 3 *Given a simple clockwise polygon, if the popturn at a reflex vertex v_i is not simple, then the polygon has a purse.*

Proof: Let v_i' be the position of v_i after the popturn. If the popturn at v_i is not simple, then the parallelogram $v_{i-1}v_iv_{i+1}v_i'$ intersects P. Suppose that P has no purse. It follows that edge e_{i+2} is outside of the parallelogram. Then by the Jordan curve theorem, there is a proper intersection between the boundary of P and $v_{i+1}v_i'$ or $v_i'v_{i-1}$. Assume by symmetry that there is such an intersection on the edge $v_{i+1}v_i'$ and let q be the first proper intersection encountered while walking from v_{i+1} to v_i'; see Fig. 2.

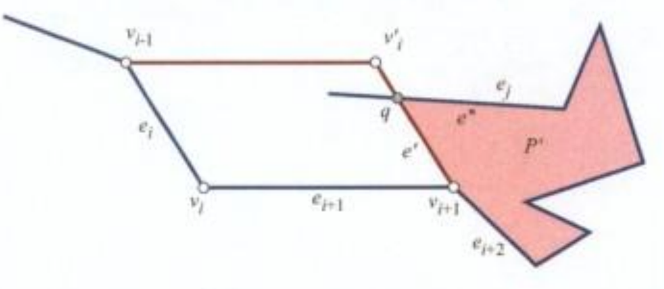


Figure 2: Proof of Lem. 3.

Let P' be a counterclockwise polygon formed by taking the portion of P between v_{i+1} and q and a vector e' from q to v_{i+1}. Let e^* be the vector preceding e' in P'. The polygon P' is closed, simple, and counterclockwise. Thus the total turning angle $\tau_{e',e^*} + \tau_{e^*,e'} = -2\pi$. But the vector e^* is part of a vector of P, say e_j, and e_i is parallel to e'; thus, $\tau_{i,j} = \tau_{e',e^*}$. Finally, e^* and e' are adjacent, so $\tau_{e^*,e'}$ must be strictly between $-\pi$ and π, and $\tau_{i,j} = \tau_{e',e^*} = -\tau_{e^*,e'} - 2\pi < -\pi$. Thus $e_i, \ldots, e_j$ is a purse. □

Theorem 4 *A simple polygon P can be convexified by a finite sequence of simple popturns if and only if P contains no purse.*

Proof: If P contains a purse $e_i, \ldots, e_j$, then by definition, $\tau_{i,j} < -\pi$ and by Lem. 2, e_i and e_j will never become adjacent, which implies that the value of $\tau_{i,j}$ will remain the same after any sequence of simple popturns. In a clockwise convex polygon, the total turning angle between every pair of edges is non-negative. This implies P can never become convex after any sequence of simple popturns.

Note that applying any sequence of popturns to P will result in a polygon which is a permutation of the original vectors. If P contains no purse, then by Lem. 3, the popturn at any reflex vertex is simple. Such a popturn will increase the area of the polygon, so the same permutation of vectors will never be repeated. Since the number of different permutations of vectors is finite, any sequence of reflex popturns will have to be finite as well. At the end of such a maximal sequence, no reflex vertex remains and the polygon is convex. □

Now we can more precisely bound the number of popturns needed to convexify a polygon:

Lemma 5 *Let P be a polygon that has no purse. Any maximal sequence of reflex popturns will convexify an n-gon P after exactly $|\{(i,j)|\tau_{i,j} < 0\}| \leq \binom{n}{2}$ popturns.*

The situation is significantly more complex in the case of weakly simple popturns. In the full version we prove:

Theorem 6 *Deciding if a polygon can be convexified by a sequence of weakly simple popturns is NP-Hard.*

3 Unit Orthogonal Polygons

When restricted to simple pops, even the 12-vertex polygon in Fig. 3a cannot be convexified. Here we loosen that restriction and allow hairpin vertices. A hairpin vertex v_i in a unit polygon has $v_{i-1}=v_{i+1}$, which leaves a pop of v_i undefined. We feel it is natural to define the

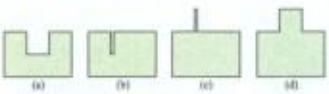


Figure 3: (a) A unit polygon that cannot be convexified by pure pops. (b,c,d) Convexifying by pin popping.

pop of a pin v_i when $v_{i-1} = v_{i+1}$ as the reflection across the line L perpendicular to the pin edges and through their common endpoint. This permits convexifying the previous example; see Fig. 3(b-d). Through an extension of the argument in Thm. 4, we can show that pops together with pin pops still do not suffice to convexify all unit polygons while remaining weakly simple—again, the polygon must have no purse. But rather than detail this argument, we turn instead to positive results.

3.1 Pin-move Extensions

There are three natural pin-move extensions: rotating a pin 90°, rotating 180°, or "untwisting" a pin. The first is related to the work of Dumitrescu and Pach [4], in that their "coin moves" can be simulated in certain contexts with the help of 90° pin rotations. However, we do not pursue this connection, and only observe that 90° rotations are subsumed by 180° rotations. We next show that the second two pin movements, and therefore the first, permit convexifying any unit orthogonal polygon while remaining weakly simple.

Let P be a unit orthgonal polygon. We define a U-shaped boundary piece $(v_i, v_{i+1}, ..., v_{j-1}, v_j)$ to be a *cup* if v_{i+1} and v_{j-1} are both reflex or both convex and $v_{i+1}, \ldots v_{j-1}$ are collinear. The line segment v_iv_j is the cup *lid*. A cup is *open* if no piece of ∂P lies along its lid. A *horizontal cup* (or H-cup) is an upright or upside down U-shape; a *vertical cup* (or V-cup) is a U-shape on its side.

Our reconfiguration algorithm converts P to a canonical form by moving pins around ∂P. If $[v_{i-1}, v_i, v_{i+1}]$ is a pin, call v_i its *tip* and $v_{i-1} = v_{i+1}$ its *base*. We distinguish two types of pins. A *flat* pin has the tip vertex v_i coincident with either v_{i-2} or v_{i+2}; see Figs. 4a, 4b. A *barb* pin has a tip vertex v_i distinct from both v_{i-2} and v_{i+2}; see Figs. 4c, 4d.

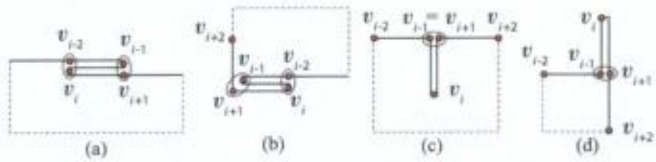


Figure 4: (a,b) Flat pins (c,d) Barb pins.

Now we relax the condition that pops preserve simplicity of the polygon, and allow for simple pin "twists" in a small neighborhood around their base point. A *twisted* pin (e.g. Fig. 5b) is the result of pop(v_i) applied in the following two conditions: (i) $[v_{i-1}, v_i, v_{i+1}]$ is a simple (untwisted) flat pin, and (ii) v_iv_{i+1} and $v_{i+1}v_{i+2}$ are orthogonal (cf. Figs. 4b, 5a). Once a pin becomes twisted, we immediately untwist it (cf. Fig. 5c). Note that our pop operations apply on the simple polygon obtained by separating the pin base into two points within an epsilon-disk of the base, as illustrated in the pin drawings. Although pin untwisting may seem like "cheating," in fact the operation is quite natural, for the coincidence of v_{i-1} with v_{i+1} means that Figs. 5b and 5c are geometrically identical. Although it may appear

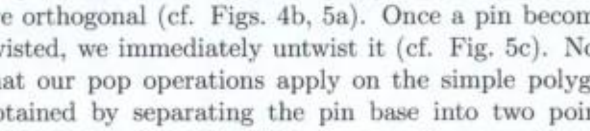


Figure 5: (a) Initial pin $[v_{i-1}, v_i, v_{i+1}]$ (b) Pin twisted after pop(v_i) (c) Untwisted pin.

from Figs. 5a, 5c that the result of popping/untwisting a pin is the same as rotating the pin 180° about its reflex base point, the pop and the 180°-turn operations are not always identical. Nevertheless, we show that they are equivalent in the sense that the composite operations (see Sec 3.3) used by our algorithm can be defined in terms of either pop/untwist pin operations or pop/180°-turn pin rotations.

3.2 Canonical Form

Let P be a polygon with $2x$ horizontal edges and $2y$ vertical edges. The *canonical form* of P is a rectangle of length x and height y. It is used as an intermediate stage in reconfiguring P into another polygon with a same number of horizontal and vertical edges.

3.3 Composite Operations

We define three composite operations used by the reconfiguration algorithm. Each can be implemented using pop/untwist operations or pop/180°-turn operations.
SLIDE(Π): Moves the pin Π one lattice edge cw around the boundary. See Fig. 6.

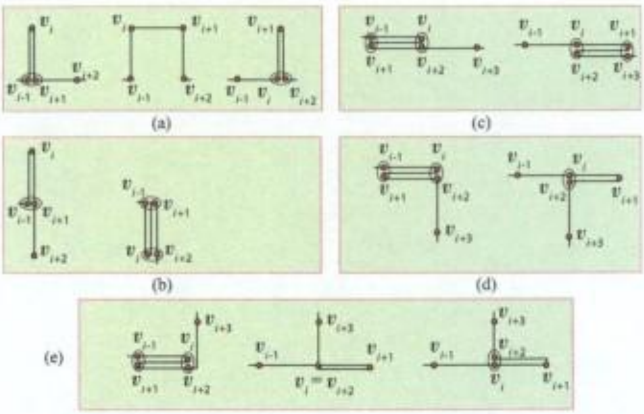


Figure 6: (a, b) Sliding a barb pin (c, d, e) Sliding a flat pin.

WALK(Π, c): Applies a sequence of SLIDE operations to walk the pin Π cw along ∂P until its base coincides with corner point c.
POPSWEEP(Π, c): Here Π is an outward pointing barb pin whose base vertex b is connected to vertex c by a straight boundary segment. This operation pops all vertices on the boundary segment, starting with b.

3.4 Converting P To Canonical Form

Let $T = \ell r$ be the leftmost among the topmost maximal horizontal sections of ∂P, with ℓ (r) the left (right) endpoint of T. The algorithm uses the composite operations to convert P into a canonical rectangle R that has its lower-right corner at r (see Fig. 7d). Initially, R is degenerate and coincides with line segment $[\ell, r]$.

The algorithm repeatedly creates a pin Π and walks it around ∂P to the top left corner t_ℓ of R (initially $t_\ell = \ell$), where it uses a POPSWEEP operation to expand Π into a new (top) row or (left) column of R. A pin is created by popping all base vertices of an open cup, which always exists (Lem. 7). E.g., in Fig. 7d, popping base vertex b_2 of cup (a_2, b_2, c_2, d_2) creates a pin (see Fig. 7e). The cup must be open, for otherwise popping the base vertices results in ∂P touching along non-pin edges. In the first iteration, the algorithm uses a pin Π corresponding to an upright open H-cup; this ensures that, once it reaches ℓ, Π expands into a row extending from ℓ to r, turning R into a one-row rectangle ($\ell = b_\ell, t_\ell, t_r, r$). Figs. 7a-7g show this for two pins.

Lemma 7 *If P is not a rectangle, it has at least one open cup in the halfplane H bounded above by T.*

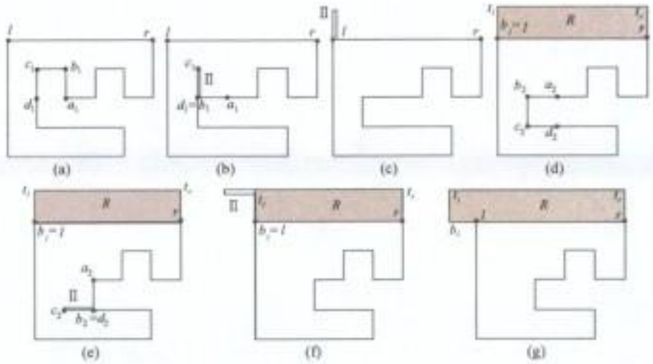


Figure 7: (a) P (b) H-cup (a_1, b_1, c_1, d_1) turned into pin Π by pop(b_1) (c) Pin Π after WALK (d) Rectangle R after POPSWEEP (e-g) Same steps for next pin.

Proof: If P is not a rectangle, then the reconfiguration is not complete and some part of ∂P lies in the interior of H. Therefore, P has at least one upright horizontal cup in H, namely the H-cup with a lowest horizontal edge as base. Of all upright horizontal cups in H, let $C = (v_i, v_{i+1}, \ldots v_{j-1}, v_j)$ be one with a highest base. Assume for the sake of contradiction that C is not open. Then its lid contains some (maximal) horizontal section $v_k, \ldots, v_{k+s}$ of ∂P. Assume w.l.o.g. that $k \geq j$. If v_k and v_j coincide, then $(v_{j-2}, v_{j-1}, v_j = v_k, v_{k+1})$ is an open V-cup in H and the proof is finished, and similarly if v_{k+s} and v_i coincide. So assume $k \neq j$ and $k + s \neq i$. Then $v_k, \ldots, v_{k+s}$ must be the base of an upright H-cup, call it D. Simple arguments show that D lies in H and is higher than C, a contradiction. □

Theorem 8 *The described algorithm transforms P into a rectangle in $O(n^3)$ pop operations.*

Acknowledgments. We thank the other participants of the 2007 *Workshop on Reconfiguration* at the Bellairs Research Institute of McGill University.

References

[1] G. Aloupis, P. Bose, E. D. Demaine, S. Langerman, H. Meijer, M. Overmars, and G. T. Toussaint. Computing signed permutations of polygons. In *Proc. 14th Canad. Conf. Comput. Geom.*, pages 68–71, August 2002.

[2] B. Ballinger. *Length-Preserving Transformations on Polygons.* PhD thesis, Univ. Calif. Davis, 2003. `http://www.math.ucdavis.edu/~brad/`.

[3] E. D. Demaine, B. Gassend, J. O'Rourke, and G. T. Toussaint. Polygons flip finitely... right? In *Proc. Snowbird Conference Discrete and Computational Geometry: Twenty Years Later.* AMS, 2007. To appear.

[4] A. Dumitrescu and J. Pach. Pushing squares around. In *20th Annu. ACM Sympos. Comput. Geom.*, pages 116–123, 2004.

[5] G. T. Toussaint. The Erdős-Nagy theorem and its ramifications. *Comput. Geom. Theory Appl.*, 31(3):219–236, 2005.

Edge-Unfolding Nested Polyhedral Bands *

Greg Aloupis [a,1], Erik D. Demaine [b,2], Stefan Langerman [c,3], Pat Morin [d,1], Joseph O'Rourke [e,4], Ileana Streinu [e,5] and Godfried Toussaint [a,1]

[a] *School of Computer Science, McGill University*

[b] *MIT Computer Science and Artificial Intelligence Laboratory*

[c] *Département d'Informatique, Université Libre de Bruxelles*

[d] *School of Computer Science, Carleton University*

[e] *Department of Computer Science, Smith College*

Abstract

A *band* is the intersection of the surface of a convex polyhedron with the space between two parallel planes, as long as this space does not contain any vertices of the polyhedron. The intersection of the planes and the polyhedron produces two convex polygons. If one of these polygons contains the other in the projection orthogonal to the parallel planes, then the band is *nested*. We prove that all nested bands can be *unfolded*, by cutting along exactly one edge and folding continuously to place all faces of the band into a plane, without intersection.

Key words: polyhedra, folding, slice curves

* A preliminary version of this paper appeared in *Proceedings of the 16th Canadian Conference on Computational Geometry*, August 2004, pages 60–63.

Email addresses: `athens@cs.mcgill.ca` (Greg Aloupis), `edemaine@mit.edu` (Erik D. Demaine), `stefan.langerman@ulb.ac.be` (Stefan Langerman), `morin@scs.carleton.ca` (Pat Morin), `orourke@cs.smith.edu` (Joseph O'Rourke), `streinu@cs.smith.edu` (Ileana Streinu), `godfried@cs.mcgill.ca` (Godfried Toussaint).

[1] Supported in part by NSERC.

[2] Supported in part by NSF CAREER award CCF-0347776 and DOE grant DE-FG02-04ER25647.

[3] Chercheur qualifié du FNRS.

[4] Supported in part by NSF grant DUE-0123154.

[5] Supported in part by NSF grant CCF-0430990.

1 Introduction

It has long been an unsolved problem to determine whether every polyhedron may be cut along edges and unfolded flat to a single, non-overlapping polygon [12,9,7,6]. An interesting special case emerged in the late 1990s: [6] can the *band* of surface of a convex polyhedron enclosed between parallel planes, and containing no polyhedron vertices, be unfolded without overlap by cutting an appropriate single edge? A band and its associated polyhedron are illustrated in Figure 1.

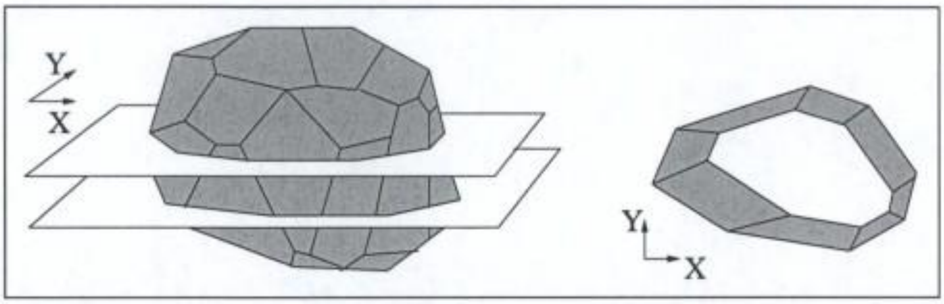


Fig. 1. A polyhedron cut by two parallel planes, and the projection of the resulting band onto the xy plane.

This band forms the side faces of what is known as a *prismatoid* (the convex hull of two parallel convex polygons in $\mathbb{R}^3$) but the band unfolding question ignores the top and bottom faces of the prismatoid. An example was found (by E. Demaine and A. Lubiw) that shows how flattened bands can end up overlapping if a "bad" edge is chosen to cut; see Figure 2.

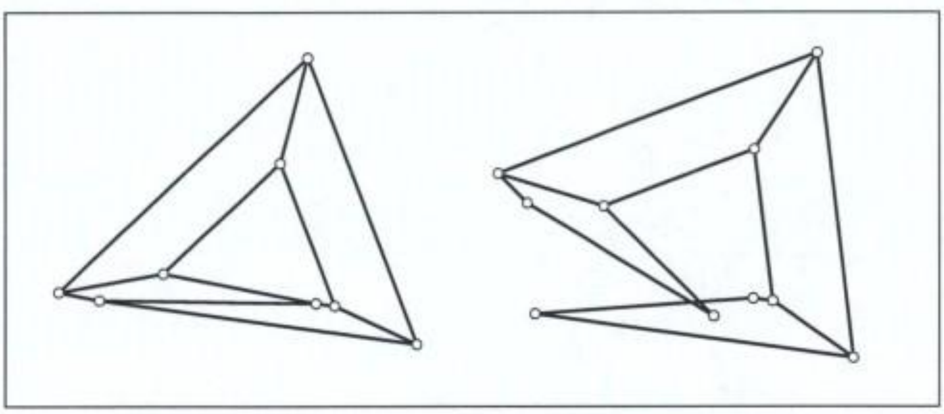


Fig. 2. Projection of a band that self-intersects when cut along the wrong edge and unfolded. Left: original band. Edges at the bottom are nearly collinear. Right: self-intersecting unfolding.

Band-like constructs have been studied before. Bhattacharya and Rosenfeld [3] define a polygonal *ribbon* as a finite sequence of polygons, not necessarily coplanar, such that each pair of successive polygons intersects exactly in a

[6] Posed by E. Demaine, M. Demaine, A. Lubiw, and J. O'Rourke, 1998.

common side. Triangular and rectangular ribbons (both open and closed) have also been studied. Arteca and Mezey [2] deal with continuous ribbons. Simple bands can be used as linkages to transfer mechanical motion, as pointed out by Cundy and Rollett [5]. Open and closed rings of rigid panels connected by hinges have also been considered in robotics as another model for robot arms with revolute joints. For example, their singularities are well understood mathematically [4]. As a special case of the more general *panel-and-hinge* structures studied in rigidity theory, they are relevant to protein modeling [13]. In all these instances, almost no attention was paid to questions regarding their non-self-intersecting states or their self-collision-avoiding motions.

There is one unfolding result that is particularly relevant to our problem, which may be interpreted as unfolding infinitely thin bands. This result states that a *slice curve*, the intersection of a plane with a convex polyhedron, develops (unfolds) in the plane without overlap [8,10]. This result holds regardless of where the curve is cut. Thus, both the top and the bottom boundary of any band (and in fact any slice curve between) cannot self-intersect after a band has been flattened. So overlap can only occur from interaction with the cut edge, as in Figure 2.

Here we will prove that a particular type of band can be unfolded by explicitly identifying an edge to be cut. A band is *nested* if projecting the top boundary A orthogonally onto the plane of the bottom boundary B results in a polygon nested inside B. For example, the band in Figure 1 is nested. Intuitively, we might expect to obtain a nested band if both parallel planes cut the polyhedron near its "top". We prove that all nested bands can be unfolded. Our proof provides more than non-overlap in the final planar state: it ensures non-intersection throughout a continuous unfolding motion.

2 Bands

We first define bands more formally and analyze their combinatorial and geometric structure, without regard to unfolding.

Let P be the surface of a convex polyhedron with no coplanar faces. Let $z_0, z_1, \ldots, z_m$ denote the sorted z coordinates of the vertices of P. Pick two z coordinates z_A and z_B that fall strictly between two consecutive vertices z_i and z_{i+1}, and suppose that z_A is above z_B: $z_i < z_B < z_A < z_{i+1}$. The *band* determined by P, z_A, and z_B is the intersection of P's surface with the horizontal slab of points whose z coordinates satisfy $z_B \leq z \leq z_A$.

The band is a polyhedral surface with two components of boundary, called A and B. Specifically, we define A as the *top (polygonal) chain* of the band, i.e., the intersection of P's surface with the plane $z = z_A$, and B is the *bottom chain*, corresponding to the plane $z = z_B$. Both chains A and B are convex polygons in their respective horizontal planes, being slice curves of a convex polyhedral surface P. All vertices of the band are vertices of either A or B.

Every vertex of the band is incident to exactly three edges: two along the chain A or B containing the vertex, and the third connecting to the other chain. This third edge, called a *hinge*, is part of an edge of the original polyhedron P connecting a vertex of P with z coordinate less than z_B to a vertex of P with z coordinate greater than z_A. The hinge from each vertex of the band defines a perfect matching between vertices of the top chain A and vertices of the bottom chain B. This matching is consistent with the cyclic orders of A and B in the sense that, if vertex a_i of A is paired with vertex b_i of B, then the vertex a_{i+1} clockwise around A from a_i is paired with the vertex b_{i+1} clockwise around B from b_i. This correspondence defines a consistent clockwise labeling of the vertices $a_0, a_1, \ldots, a_{n-1}$ of A and the vertices $b_0, b_1, \ldots, b_{n-1}$ of B, unique up to a common cyclic shift. [7]

Each face of the band is a quadrilateral spanned by two adjacent vertices a_i and a_{i+1} on the top chain A and their corresponding vertices b_i and b_{i+1} on the bottom chain B. This facial structure follows from the edge structure of the band. Each face is planar because it corresponds to a portion of a face of the original polyhedron P. Because edges $a_i a_{i+1}$ and $b_i b_{i+1}$ lie in a common plane as well as in parallel horizontal planes, the edges themselves must be parallel. Thus every face of the band is in fact a trapezoid, with parallel top and bottom edges.

3 Nested Bands

Next we analyze the geometric structure of nested bands in particular, still without regard to unfolding.

A band is *nested* if the orthogonal projection of A into the xy plane is strictly contained inside the orthogonal projection of B into the xy plane. (Of course, a band is just as nested if instead B's projection is contained inside A's projection, but in that case we just reflect the band through the xy plane.)

Nested bands have a particularly simple structure when projected into the xy plane. As with all bands, each face projects to a trapezoid. The unique property of a nested band is that none of its edges cross in projection. This property follows because the projected edges are a subset of a triangulation of

[7] Throughout this paper, indices are taken modulo n.

the projections of A and B, which themselves do not intersect by the nested property. (In non-nested bands, edges of A intersect edges of B in projection.) Thus the projected trapezoidal faces of the band form a planar decomposition of the region of the xy plane interior to the projection of B and exterior to the projection of A. When dealing with projections, we will refer to A (B) as the *inner (outer) chain.*

In the xy projection, the *normal cone* of a vertex a_i of A (or more generally any convex polygon) is the closed convex region between the two exterior rays that start at a_i and are perpendicular to the incident edges $a_{i-1}a_i$ and a_ia_{i+1} respectively. See Figure 3. The two rays forming this cone decompose the local exterior of A around a_i into three regions: left (counterclockwise), inside, and right (clockwise) of the normal cone.

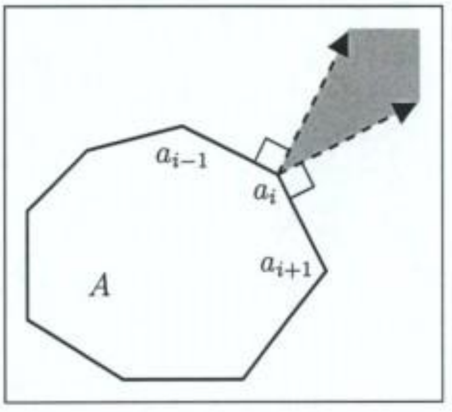


Fig. 3. The normal cone of a vertex a_i.

Lemma 1 *In the xy projection of a nested band, not all hinges a_ib_i can be to the right (or all to the left) of the normal cones of their inner endpoint a_i.*

PROOF. The following proof refers exclusively to the xy projection. Suppose by symmetry that all hinges are clockwise (right), or on the right border, of their respective normal cones on the inner chain A. For each i, define T_i to be the trapezoid with vertices $a_{i-1}, a_i, b_i, b_{i-1}$, and let h_i denote its height, i.e., the distance between the opposite parallel edges $a_{i-1}a_i$ and $b_{i-1}b_i$. See Figure 4. Because a_ib_i is right of the perpendicular at a_i to a_ia_{i+1}, and because the interior angle at b_i is convex, the convex angle $a_ib_ib_{i-1}$ is less than the convex angle $b_ia_ia_{i+1}$. Thus, the height h_i of T_i is less than the height h_{i+1} of the clockwise next trapezoid T_{i+1}. Applying this argument to every T_i, we obtain a cycle of strict inequalities $h_0 < h_1 < \cdots < h_{n-1} < h_0$, which is a contradiction. □

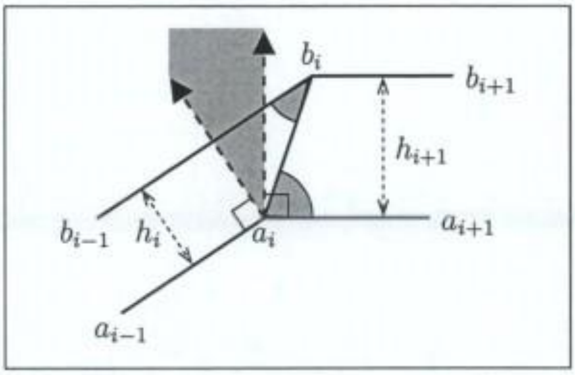


Fig. 4. If the hinge $a_i b_i$ is right of the normal cone at a_i, then the top shaded angle is less than the bottom shaded angle, so $h_i < h_{i+1}$.

4 Opening Convex Chains

Before we study the unfolding of bands, we first study what happens when opening a convex closed chain (polygon) by cutting it at some vertex a_i and increasing all other internal angles.

We introduce some basic notation and terminology for a convex closed chain; refer to Figure 5(a). Given a clockwise-oriented convex closed chain $A = \langle a_0, a_1, \ldots, a_{n-1} \rangle$ in the plane, the *interior angle* α_j at a vertex a_j, $0 \le j \le n-1$, is the angle $a_{j-1}a_ja_{j+1}$ located on the right side of the chain. Let $\tau_j = \pi - \alpha_j$ be the *turn angle* at a_j, which is positive (to the right) because of the clockwise orientation of A. Let θ_j be the counterclockwise angle of the vector $a_j - a_{j-1}$ from the positive x axis. If $a_i - a_{i-1}$ is fixed along the positive x axis, then for a chain with all right turns, we have $\theta_i = 0$, $\theta_{i-1} = \tau_{i-1}$, and in general,

$$\theta_{i-k} = \sum_{j=i-k}^{i-1} \tau_j. \tag{1}$$

An *opening* of a convex closed chain A at a_i is a motion $A'(t)$ that cuts the chain at a_i, holds the edge $a_{i-1}a_i$ fixed, and monotonically increases all other interior angles. See Figure 5(b). More precisely, an opening of A at a_i consists of a nonstrictly increasing function $\delta_j : [0,1] \to [0, \tau_j]$, with $\delta_j(0) = 0$, for each $j \ne i$. For any $t \in [0,1]$, the opened chain $A'(t) = \langle a^*(t), a'_{i+1}(t), a'_{i+2}(t), \ldots, a'_{n-1}(t), a'_0(t), a'_1(t), \ldots, a'_i(t) \rangle$ at time t is obtained from A by fixing $a'_i(t) = a_i$, fixing $a'_{i-1}(t) = a_{i-1}$, and opening each interior angle α_j, $j \ne i$, to $\alpha'_j(t) = \alpha_j + \delta_j(t)$. The opening separates two copies of a_i; we call the stationary copy a_i and the moving copy $a^*(t)$. Because $\delta_j(0) = 0$, the opening motion starts at $A'(0) = A$. Because $\delta_j(t)$ is nonstrictly increasing, the interior angles $\alpha'_j(t)$ only open with t. Because $\delta_j(t) \le \tau_j$, the interior angle $\alpha'_j(t)$ remains at most $\alpha_j + \tau_j = \pi$, so the opened chain $A'(t)$ has only right turns. Thus these chains

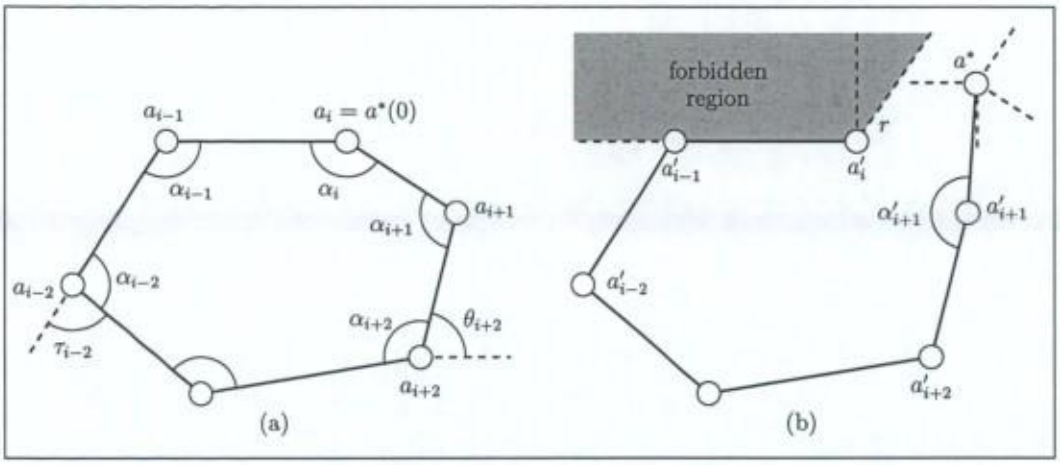


Fig. 5. (a) A convex closed chain A, and (b) an opening of α_{i+1}.

$A'(t)$ can use the same definitions of interior angle $\alpha'_j(t)$, turn angle $\tau'_j(t)$, and counterclockwise angle $\theta'_j(t)$ at a vertex $a'_j(t)$, $j \neq i$, and the analog of Equation (1) still holds.

Lemma 2 *During any opening $A'(t)$ of a convex closed chain A at a_i, every edge $a'_k(t)a'_{k+1}(t)$ turns clockwise in the sense that the vector $a'_{k+1}(t) - a'_k(t)$ rotates only clockwise as t increases; in particular, $a^*(t)a_{i+1}(t)$ turns clockwise.*

PROOF. The transformation of an edge $a_{k-1}a_k$ of A to $a'_{k-1}(t)a'_k(t)$ induced by the opening at time t can be expressed as a composition of rotations, rotating clockwise by $\delta_j(t)$ around each vertex a_j for $j = k, k+1, \dots, i-1$. In particular, the vector $a'_k(t) - a'_{k-1}(t)$ is a rotation of $a_k - a_{k-1}$ clockwise by $\sum_{j=k}^{i-1} \delta_j(t)$. Because $\delta_i(t) \geq 0$ and $\delta_i(t)$ only increases with t, $a_{k+1}(t) - a_k(t)$ rotates only clockwise as t increases. □

Lemma 3 *During any opening $A'(t)$ of a convex closed chain A at a_i, the Euclidean distance between any two vertices $a'_j(t)$ and $a'_k(t)$ only increases with t.*

PROOF. Cauchy's arm lemma [8,10] states that opening the interior angles $\alpha_1, \alpha_2, \dots, \alpha_{n-1}$ of a convex open chain $a_0, a_1, \dots, a_n$ nonstrictly increases the Euclidean distance between the endpoints a_0 and a_n. The lemma follows from applying Cauchy's arm lemma to the chain $a_j, a_{j+1}, \dots, a_k$ or the chain $a_k, a_{k+1}, \dots, a_j$, whichever excludes the missing edge a_ia^*. □

We define three classes of shapes that an open chain $A' = \langle a^*, a'_{i+1}, a'_{i+2}, \dots, a'_{n-1}, a'_0, a'_1, \dots, a'_i \rangle$ with only right turns may have: convex, weakly convex, and spiral. Refer to Figure 6. The chain A' is *convex* if joining the endpoints a'_i and a^* with a closing segment yields a convex polygon. The chain A' is *weakly convex* if joining the endpoints a'_i and a^* with a segment yields a nonconvex simple polygon with no exterior angles smaller than $\pi/2$. Such a weakly convex chain is called *R-weakly convex* or *L-weakly convex* depending on which endpoint is on the hull: if a'_i is on the hull, then the chain is L-weakly convex; if a^* is on the hull, then the chain is R-weakly convex. If the chain A' is neither convex nor weakly convex, then it is a *spiral*.

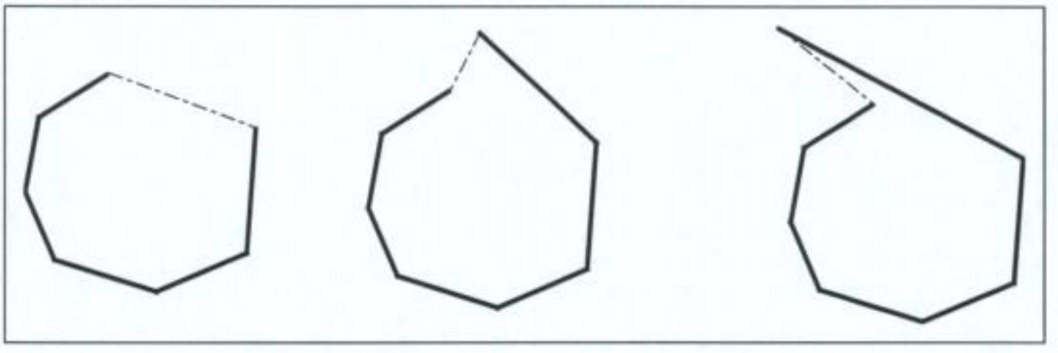


Fig. 6. Types of chains, from left to right: convex, R-weakly convex, spiral. Endpoints are joined by dashed line segments.

Lemma 4 *During any opening $A'(t)$ of a convex closed chain A at a_i, $A'(t)$ remains convex or weakly convex, and the endpoint $a^*(t)$ remains outside the normal cone of a_i.*

PROOF. Define the *forbidden region* to be the normal cone of a_i unioned with the quarter-plane above the horizontal ray emanating leftward from a_i; see Figure 5. Initially, no vertex a_j is inside the forbidden region. By Lemma 3, no vertex $a'_j(t)$ can cross an edge $a'_{k-1}(t)a'_k(t)$, for to cross the edge, $a'_j(t)$ would have to approach one of the edge's endpoints. In particular, no vertex $a'_j(t)$ can cross the edge $a_{i-1}a_i$. Because the opened chain $A'(t)$ has only right turns, the only way for a vertex $a'_j(t)$ of the chain to enter the forbidden region is for $a^*(t)$ to cross the ray r emanating from a_i normal to a_ia_{i+1}. Such penetration is possible only when $a^*(t)$ is above or on the horizontal line through the edge $a_{i-1}a_i$, so we consider values of t for which this is the case.

We claim that, for such values of t, the direction of the edge $a^*(t)a'_{i+1}(t)$ remains in the clockwise range from the direction of a_ia_{i+1} to the horizontal leftward direction. By Lemma 2, the edge turns clockwise from its original direction of a_ia_{i+1}. If the direction were ever to reach horizontal leftward, it would be impossible to connect $a'_{i+1}(t)$ to a_{i-1} by only turning right and using a total turn angle less than 2π. (Turn angles only decrease while opening, and the initial total turn angle excluding a_i is less than 2π.) The vertices $a'_j(t)$ thus remain in the clockwise wedge around $a^*(t)$ from the direction of a_ia_{i+1} to horizontal leftward. These vertices are the possible centers of a clockwise rotation affecting $a^*(t)$. The resulting instantaneous direction of motion of $a^*(t)$ is thus in the clockwise range from the direction of the normal ray r to vertical downwards (the previous cone of directions rotated clockwise by $\pi/2$). Furthermore, in the case of instantaneous motion along the direction of r, the actual motion of $a^*(t)$ is clockwise of the direction of r. Therefore, $a^*(t)$ moves away from the ray r for these values of t, so it could never cross r. □

Lemma 5 *Let $A'(t)$ and $A''(t)$ be openings of a convex closed chain A at a_{i+1} and at a_i, respectively, with the same angle-opening $\delta_j(t)$ functions for $j \neq i, i+1$. If $A'(t)$ is R-weakly convex, then $A''(t)$ cannot be L-weakly convex.*

PROOF. Because the lemma concerns only a single time t, we omit the t argument. We apply a series of transformations that transform A' into A''; refer to Figure 7. Because A' is R-weakly convex, a^* must be in the upper-right quadrant of a'_{i+1}. Now we make a new cut at a'_i, and translate the entire opened chain, except the fixed edge a_ia_{i+1}, so that a^* re-attaches to a'_{i+1}. We let a''_i denote the translated copy of a'_i, and let $a^{**}a''_{i+1}$ denote the original fixed edge. Now a''_i must be in the lower-left quadrant of a^{**}.

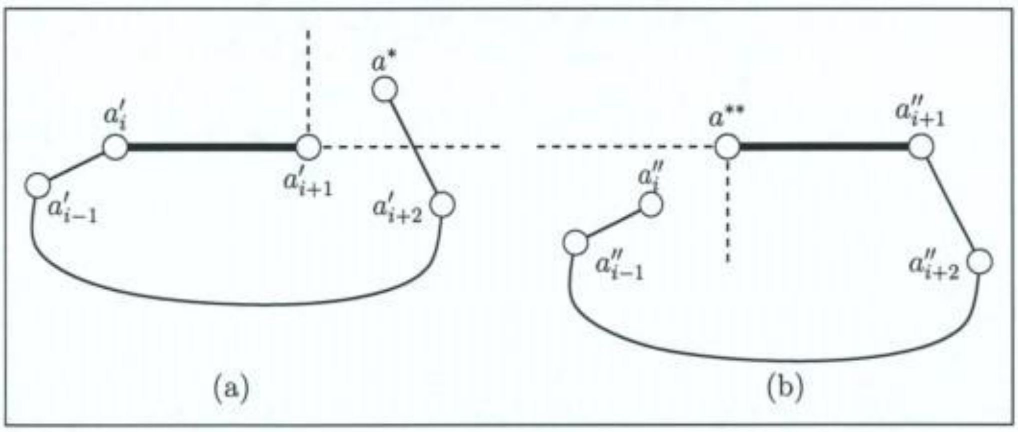


Fig. 7. (a) An opened chain A'. (b) Translating part of the chain to switch the cut vertex. This is a new opened chain A'' except that the angle α_{i+1} is not yet opened.

Now we have a new opened chain, except that we have not taken care of the opening of angles α''_i and α''_{i+1}. Because A' opened the angle at a'_i by rotating the chain that we merely translated, and a''_i no longer has an angle to open, we must rotate the translated chain to return it to the original orientation. This rotation is counterclockwise, because the opening rotation at a'_i was clockwise. Next, because a''_{i+1} (previously a^*) has an angle not present in A', we must open that angle by again rotating the entire translated chain. Again the rotation is counterclockwise to open a''_{i+1}. (Technically, we should also rotate the entire chain to make $a''_{i-1}a''_i$ horizontal, but this does not change weak convexity.) During these counterclockwise rotations, a''_i might cross into the lower-right quadrant of a^{**}, but a''_i cannot cross into the upper-left quadrant of a^{**}. Therefore cutting at a_i cannot produce an L-weakly convex chain. □

5 Unfolding Nested Bands

Having completed our study of unfolding cut chains, we now return the original problem of unfolding bands. Our results on chains help understand the motions of the top boundary A and the bottom B of the band. The rest of our study focuses on the cut edge, which can cause intersection as in Figure 2 if we are not careful.

After cutting a single hinge, a *flattening motion* is a continuous motion during which each face moves rigidly but remains connected to each adjacent face via their common hinge, and the final configuration is planar. If no intersection occurs during the motion, then this motion is a *continuous unfolding.* If the resulting configuration is non-self-intersecting, but intersection occurs during the motion, then we call the motion an *instantaneous unfolding* and the resulting configuration an *unfolded state.* Thus in Figure 2 we would say that the band has been flattened, but because it self-intersects it has not unfolded. These notions can be defined precisely by specifying rigid motions of the faces as functions of time that satisfy the connectivity constraints, similar to openings of chains.

We now describe the particular flattening motion that will lead to our unfolding, though it requires some effort to prove non-intersection, particularly of the final state. The flattening motion is based on *squeezing* together the two parallel planes $z = z_A$ and $z = z_B$ that contain A and B, keeping the planes parallel and keeping each chain on its respective plane. At time $t \in [0, 1]$, the squeezing motion reduces the vertical separation between the two parallel planes down to $(1 - t)(z_A - z_B)$, that is, it linearly interpolates the separation from the original $z_A - z_B$ down to 0.

The squeezing uniquely determines the hinge dihedral angles necessary to keep the vertices of the band on their respective moving planes (assuming exactly one edge of the band has been cut). See Figure 8 for an example of the projected motion. For nested bands, the motion increases the interior angle at every vertex of each chain in projection. This property can be seen by examining any two adjacent faces that are being "squeezed". Both faces rotate continuously to become more horizontal. If we forced one of the faces to keep its vertices in the parallel planes, but allow the second face to only follow this motion rigidly (i.e., the dihedral angle at the hinge remains fixed), then the edges of the second face would no longer be on the horizontal planes. To compensate, the second face must perform a (dihedral) rotation about the hinge. In fact, the interior angle at the hinge must increase (flatten), causing the interior angles of the chains to increase (open). Because the interior angle at a vertex of a nested band can open only to π, the opening chain will always have only right turns. Thus we can apply the analysis of opening chains from

Section 4. For example, Lemma 4 tells us that the opening chains never become spirals, so in particular never self-intersect while flattening (a fact already known from the slice-curve result of [8,10]).

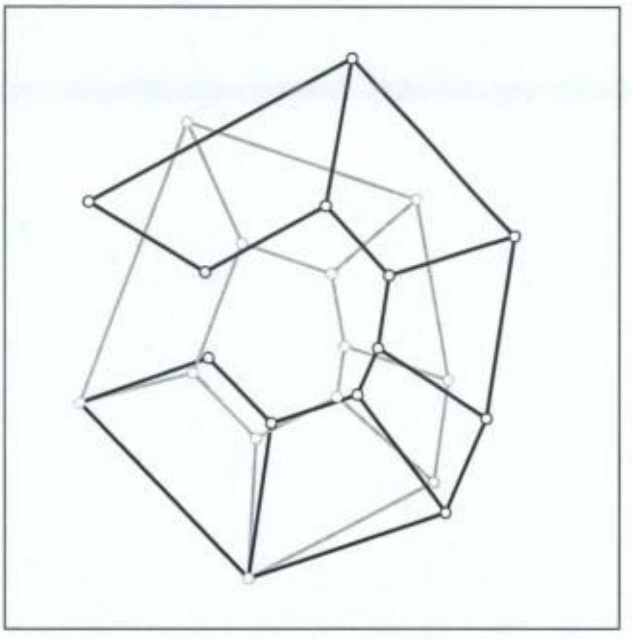


Fig. 8. A view from above of a nested band during a squeezing motion. The original configuration has a lighter shade. For each trapezoid, the height increases and its parallel edges rotate clockwise relative to their original positions.

As the parallel planes squeeze together, each band face remains a trapezoid in the projection. Edges a_ia_{i+1} and b_ib_{i+1} remain parallel and retain their original lengths throughout. Hinge projections lengthen as the band is squeezed, which causes the trapezoid angles to change. Because b_i and b_{i+1} move orthogonally away from a_ia_{i+1}, acute trapezoid angles increase toward $\pi/2$ and obtuse angles decrease toward $\pi/2$.

The goal of this section is to show that the band does not self-intersect if we cut a specific hinge. We mention that self-intersection of the band in 3D implies self-intersection in projection, so it suffices to prove that there is no self-intersection in projection to establish that there is no self-intersection in 3D.

Suppose that we cut hinge a_ib_i and hold $a_{i-1}a_i$ fixed along the x axis in the positive direction. The motion separates two copies of a_i; we call the stationary one a_i, and call the moving one a^*, as in Figure 5. Correspondingly, for the outer chain, the direction of $b_{i-1}b_i$ remains fixed (it moves away from $a_{i-1}a_i$ because the trapezoid enlarges in projection, but remains parallel), and b^* is a "moving" endpoint. Thus the cut hinge is split into edges a_ib_i and a^*b^*. See Figure 9.

Call a chain A *safe* if it is either convex, or it is R-weakly convex and the

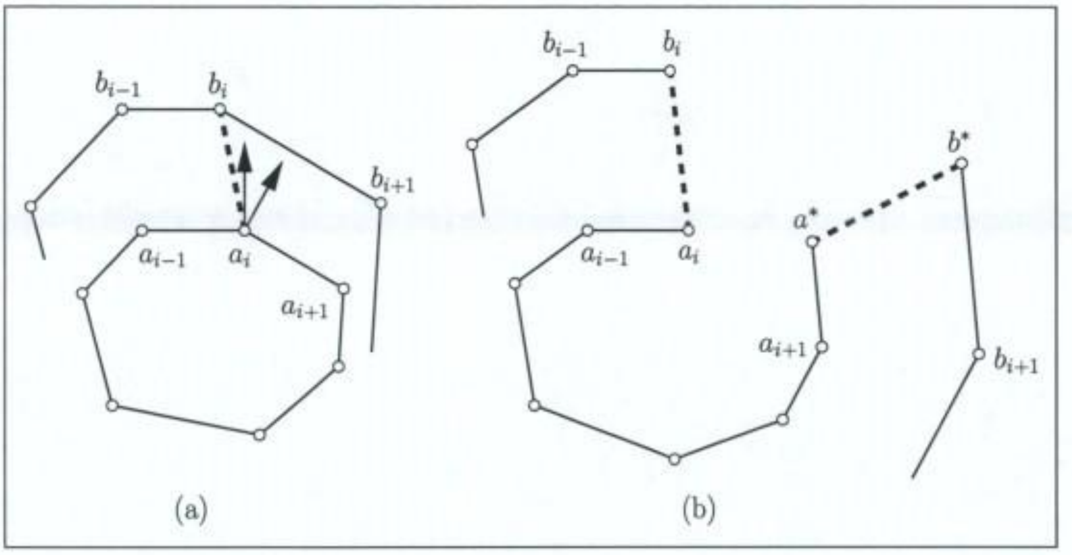


Fig. 9. (a) Projection of the inner convex chain A and part of the outer chain B. Hinge a_ib_i and the normal cone of vertex a_i are shown. (b) The result of cutting at a_ib_i and flattening.

hinge a_ib_i is left of or in the normal cone at a_i, or it is L-weakly convex and a_ib_i is right of or in the normal cone at a_i. An opening of the band is *safe* if the opened inner chain A is safe. See Figure 10. We will prove that safe openings of the band never self-intersect, i.e., are unfoldings. Then we will prove that there is always a suitable hinge a_ib_i that leads to a safe opening.

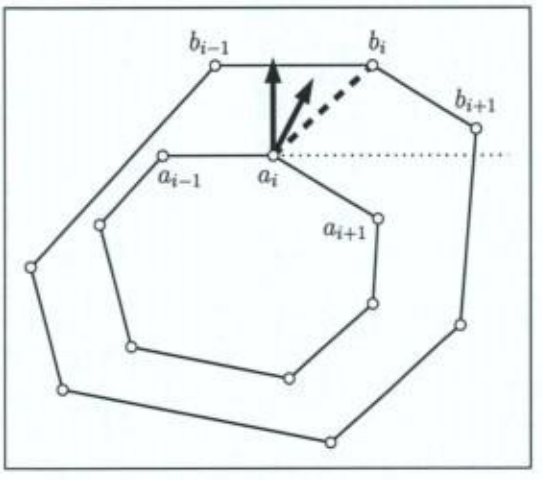


Fig. 10. After cutting at a_i, the inner chain will become R-weakly convex if a^* ends up above the line determined by $a_{i-1}a_i$ (dotted). In this case, the cut is labeled *safe* if hinge a_ib_i (dashed) is left of or in the normal cone at a_i (which is not the case in this figure).

Our next lemma covers an opened band by a clockwise-turning family of rays emanating from the inner chain A, dependent only on the cut edges and not on the outer chain B. This covering will allow us to prove nonoverlap of the opened band—in fact, an infinite version of the band with no bounding outer chain—in certain cases using the nonoverlap of A.

Lemma 6 *For any safe opening of the band, there is a function r assigning a ray $r(p)$ from each point p on the chain A such that*

(1) r is a continuous function;
(2) the direction of $r(p)$ rotates only clockwise as p moves along A from a^* to a_i;
(3) the total turn angle made by $r(p)$ as p travels along A from a^* to a_i is at most 2π;
(4) the ray $r(p)$ is locally exterior to the polygon formed by A and the edge $a_i a^*$; and
(5) the ray $r(a_i)$ passes through b_i, and the ray $r(a^*)$ passes through b^*.

(Only Property 4 requires safeness.)

PROOF. First we assign $r(a_j)$ for each vertex a_j. We set $r(a_i)$ to the ray from a_i passing through b_i, and set $r(a^*)$ be the ray from a^* passing through b^*. Thus we obtain Property 5. For each $j \neq i$, let $[u_j, w_j]$ denote the clockwise range of directions of rays that are left of the two incident edges $a_{j-1}a_j$ and $a_j a_{j+1}$ (and hence locally exterior to A). We set the direction of $r(a_j)$, $j \neq i$, according to three cases:

(1) If the direction of $r(a_i)$ is in the clockwise range $[u_j, w_j]$, then we set the direction of $r(a_j)$ to the direction of $r(a_i)$.
(2) Otherwise, if the direction of $r(a^*)$ is in the clockwise range $[u_j, w_j]$, then we set the direction of $r(a_j)$ to the direction of $r(a^*)$.
(3) Otherwise, we set the direction of $r(a_j)$ to the direction in the middle of the range $[u_j, w_j]$, i.e., $r(a_j)$ is the angular bisector of the exterior (nonconvex) angle at a_j.

Finally, we make r a continuous function over points on A by linearly interpolating the direction from $r(a_{j-1})$ to $r(a_j)$ for points along the edge $a_{j-1}a_j$, keeping the rays left of the edge. Thus we obtain Property 1.

Next we show Property 2 for the points along any edge $a_{j-1}a_j$. We split into three cases. If $r(a_{j-1})$ and $r(a_j)$ are exterior angular bisectors of a_{j-1} and a_j, respectively, then the claim follows because the exterior angles are nonconvex, so $r(a_{j-1})$ is left of the edge normal (at a_{j-1}), while $r(a_j)$ is right of the edge normal (at a_j). If $r(a_{j-1})$ has the same direction as $r(a^*)$, then $r(a_{j-1})$ must be strictly left of the line from a_{j+1} to a_j (in direction), while $r(a_j)$ is nonstrictly right of this line, so the claim follows. The case when $r(a_j)$ has the same direction as $r(a_i)$ is symmetric. Thus we obtain Property 2.

Next we show Property 3. Along each edge $a_{j-1}a_j$ of A for which $r(a_{j-1})$ and $r(a_j)$ are angular bisectors, the ray turns $\frac{1}{2}(\tau_{j-1}+\tau_j)$: $\frac{1}{2}\tau_{j-1}$ turn from $r(a_{j-1})$ to a normal to $a_{j-1}a_j$, and $\frac{1}{2}\tau_j$ turn from that normal to $r(a_j)$. Thus the total turn caused by such edges is at most $\frac{1}{2}\sum_{j\neq i,i+1}(\tau_{j-1}+\tau_j) = \sum_j \tau_j - \tau_i - \frac{1}{2}(\tau_{i-1}+\tau_{i+1})$. In the original chain A before opening, the total turn angle $\sum_j \tau_j$ is 2π, and opening the chain only decreases the turn angle τ_j at each vertex a_j, so $\sum_j \tau_j$ remains at most 2π. Thus the total turn of ray from being normal to $a^* a_{i+1}$ to being normal to $a_{i-1}a_i$, visiting the angular bisectors of a_j, $j \neq i$, in between, is at most $2\pi - \tau_i$. If the projected trapezoid angle at a_i ($\angle a_{i-1}a_i b_i$) is acute, then this total turn has already accounted for reaching (in fact, going beyond) the direction of ray $r(a_i)$; if the angle is obtuse, however, then we must also add the clockwise angle from the normal of $a_{i-1}a_i$ to $r(a_i)$ to the total turn. Similarly, if the projected trapezoid angle at a^* ($\angle b^* a^* a_{i+1}$) is obtuse, then we must add the clockwise angle from $r(a^*)$ to the normal of $a^* a_{i+1}$ to the total turn. Before the opening, the sum of these two clockwise angles is τ_i, and the flattening of the trapezoids only decreases these projected angles. Thus, the additional turn remains at most τ_i. The total turn angle of the rays is therefore at most 2π, proving Property 3.

Finally we show Property 4. The property holds along any edge $a_{j-1}a_j$ of A, with respect to that edge, because rays $r(a_{j-1})$ and $r(a_j)$ are both chosen to be left of the edge $a_{j-1}a_j$, and because by Property 2, $r(a_j)$ is clockwise of $r(a_{j-1})$ in the halfplane left of $a_{j-1}a_j$. It remains to show Property 4 at a_i and a^* with respect to the closing edge $a_i a^*$. Assume without loss of generality that A is either convex or R-weakly convex. (Otherwise, imagine opening from the other side, swapping the roles of a_i and a^*.) In either case, $a^* a_{i+1}$ is an edge of the convex hull of A. Because the incident projected trapezoid of the band is left of this edge, $a^* b^*$ and hence $r(a^*)$ are left of this edge. Thus $r(a^*)$ is exterior to A. For convex chains, the same argument shows that $r(a_i)$ is left of the edge $a_{i-1}a_i$ and hence exterior to A, completing the proof in this case. Now consider R-weakly convex chains. By safeness, $a_i b_i$ and hence $r(a_i)$ is left of or in the normal cone at a_i. By Lemma 4, a^* is right of this normal cone. Hence, $r(a_i)$ is locally outward with respect to the edge $a_i a^*$. Therefore, in all cases, we have Property 4. □

Lemma 7 *For any ray assignment r on a safe opened chain A satisfying Properties 1–4 of Lemma 6, no two rays $r(p)$ and $r(q)$ intersect for two points $p \neq q$ of A.* [8]

PROOF. Consider any two points p and q on A, and assume by symmetry that q appears after p in the clockwise order around A. Let ℓ be the directed line from p to q. For the rays $r(p)$ and $r(q)$ to intersect, they have to be on the same side of ℓ.

Suppose first that $r(p)$ and $r(q)$ are both right of ℓ, as in Figure 11(a). For

[8] Thereby avoiding total protonic reversal [11].

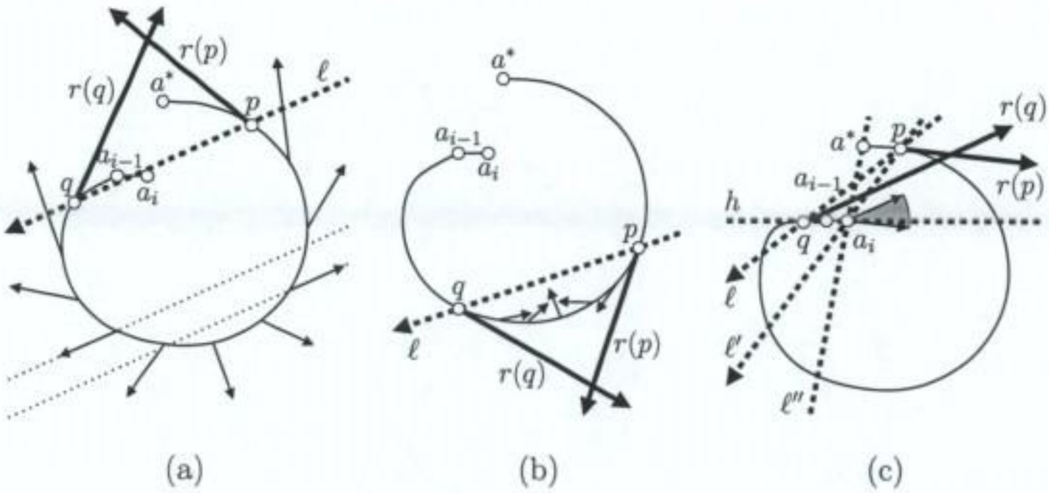


Fig. 11. Three cases of rays $r(p)$ and $r(q)$ attempting to cross.

these rays to intersect, $r(q)$ must be clockwise of $r(p)$ in the halfplane right of ℓ. As we move a point x from p to q clockwise around A, $r(x)$ must rotate continuously clockwise by Properties 1 and 2 of Lemma 6. During this motion, $r(x)$ sweeps the clockwise angle from $r(p)$ to the reverse direction of ℓ, then it sweeps the π clockwise angle from the reverse direction of ℓ to the forward direction of ℓ, and finally it sweeps the clockwise angle from the forward direction of ℓ to $r(q)$. If $r(p)$ is counterclockwise of $r(q)$ in the halfplane right of ℓ, the first and last angle must overlap, summing to more than π, and hence $r(x)$ must sweep an angle more than 2π during x's motion, contradicting Property 3 of Lemma 6. Therefore $r(p)$ and $r(q)$ cannot intersect right of ℓ.

It remains to consider the case when both $r(p)$ and $r(q)$ are left of ℓ. For these rays to intersect, $r(p)$ must be clockwise of $r(q)$ in the halfplane left of ℓ. As a first subcase, suppose that the entire subchain of A from p to q is nonstrictly left of ℓ, as in Figure 11(b); in particular, this subcase happens when A is convex. As in the previous case, if we move a point x from p to q clockwise around A, $r(x)$ must rotate continuously clockwise by Properties 1 and 2 of Lemma 6. If $r(p)$ is clockwise of $r(q)$ in the halfplane left of ℓ, then $r(x)$ must at some point locally enter the polygon, contradicting Property 4 of Lemma 6. Hence $r(p)$ and $r(q)$ cannot intersect in this subcase.

We are left with the subcase when A is weakly convex and the subchain of A between p and q is at some point right of ℓ, as in Figure 11(c). This last property implies that ℓ intersects A between p and q. Assume without loss of generality that A is R-weakly convex. and thus a^* is above the horizontal line h through $a_{i-1}a_i$. (Otherwise, imagine opening from the other side, swapping the roles of a_i and a^*.) Now h partitions the chain A into two convex subchains, where the subchain above h precedes the subchain below h in the clockwise order of A. For ℓ to intersect A between p and q, p and q must be on opposite sides of h, and by the clockwise ordering, p must be above h and q must be below (or on) h. By R-weak convexity, both a^* and p are in the upper-right quadrant from a_i. In particular, the line ℓ'' through the closing edge $a_i a^*$ and the line ℓ' through a_i and p both have positive slope. Now ℓ' partitions the portion of A clockwise after p into two convex subchains, and if we direct ℓ' from p to a_i, the subchain nonstrictly left of ℓ' contains p. For ℓ to intersect A between p and q, q must be on the subchain right of ℓ', which is in the lower-left quadrant of a_i. Hence, the slope of ℓ must be positive and at most the slope of ℓ'. (Note that the slope of a line does not depend on the line's orientation.) Furthermore, the slope of ℓ' is at most the slope of ℓ''. By Properties 1, 2, and 3 of Lemma 6, the direction of $r(a_i)$ must be in the clockwise range from the direction of $r(q)$ to the direction of $r(p)$. In particular, this cone of directions is in the halfplane left of ℓ. By the slope arguments above, this cone is contained in the nonconvex clockwise wedge from the ray starting at a_i through a^* to the horizontal leftward ray starting at a_i. But then $r(a_i)$ locally enters the polygon, contradicting Property 4 of Lemma 6. □

Lemma 8 *For any ray assignment r on a safe opened chain A satisfying Properties 1–5 of Lemma 6, the union of rays $r(p)$ over all points p on A covers the opened band.*

PROOF. The chains A and B, together with the hinges $a_i b_i$ and $a^* b^*$, define a bounded but possibly self-intersecting polygon, namely, the opened band. For each point p on A, let $b(p)$ denote the first point of the boundary of this polygon that is intersected by the ray $r(p)$. By boundedness of the polygon, the ray $r(p)$ must exit the band. By Property 4 of Lemma 6, $r(p)$ cannot immediately exit at p; and by Lemma 7, $r(p)$ cannot exit by intersecting A at any other point q because then $r(p)$ would intersect $r(q)$. By Lemma 7, $r(p)$ cannot exit by intersecting either of the hinges, because then it would intersect $r(a_i)$ or $r(a^*)$. Thus, $r(p)$ must exit the polygon by intersecting B at some point $b(p)$.

By Property 1 of Lemma 6, $b(p)$ varies continuously along B. By Lemma 7, $b(p) \neq b(q)$ for any two points $p \neq q$ of A. By Lemma 5, $b(a_i) = b_i$ and $b(a^*) = b^*$. Thus, as we vary p along A from a^* to a_i, $b(p)$ varies continuously and monotonically along B from b^* to b_i. At any point p during this motion, the ray $r(p)$ covers the segment $pb(p)$ contained by the band. These segments define a *ruling* of the band, starting at $a^* b^*$, ending at $a_i b_i$, and in between moving along the two other boundary chains A and B.

The consequence is that the continuum of segments $pb(p)$, and hence the containing rays $r(p)$, cover the band. This consequence can be seen perhaps more clearly by dividing the ruling at the finitely many key times when p is a vertex of A or $b(p)$ is a vertex of B. Then we effectively divide the problem into the regions of time between these key times, where we simply have a linear ruling of a quadrangle. □

Combining Lemmas 7 and 8, we obtain the following important consequence:

Corollary 9 *Any safe opening of a band does not self-intersect.*

Now we turn to proving that a safe opening always exists. By Lemma 1, there is a vertex a_k whose hinge is counterclockwise of the normal cone at a_k, while the hinge at a_{k+1} is clockwise of its respective cone. For the cuts at both vertices to produce unsafe inner chains, cutting at a_k must produce an L-weakly convex chain, while cutting at a_{k+1} must produce an R-weakly convex chain. See Figure 12.

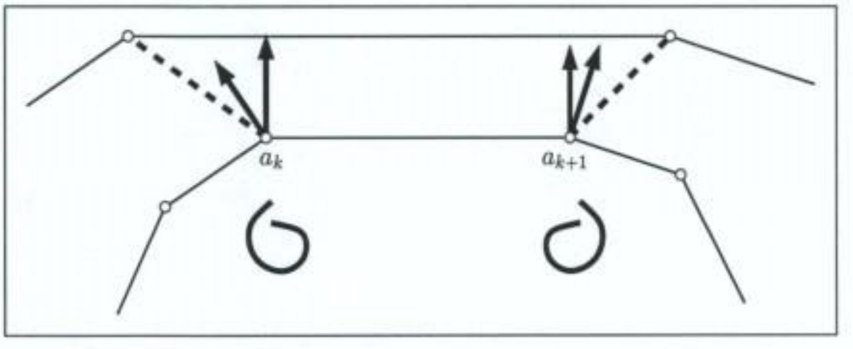


Fig. 12. Two successive vertices, a_k and a_{k+1}, whose cuts produce different weakly convex chains (indicated by the curves below the vertices).

But by Lemma 5, this situation is impossible. Thus, we can always find a suitable vertex to cut so that the inner chain opens to a safe position, which by Corollary 9 implies that we can always find an edge to cut along so that a nested band has an unfolded state. This completes the proof of our main result:

Theorem 10 *Every nested band has an unfolded state.*

The nonintersection of the final state turns out to be the main challenge for our unfolding motion, and we can use it to establish non-intersection throughout:

Theorem 11 *Every nested band has a continuous unfolding motion.*

PROOF. The squeezing motion that we have defined has the property that all the points with the same original height have the same new height at any time t during the squeezing motion, and vice versa for $t < 1$. To see this, parameterize a point p on the band by its original height z_p divided by the height z of the original band. After partially squeezing the band to height z_S, the new height of p will be $z_S(z_p/z)$.

Now, suppose that two points p and q intersected at some time $t < 1$ during the squeezing motion. At this time, the points have the same height, so at their original positions at time 0, p and q must also have the same height h. We can view the motion of p and q as the development of a slice curve $z = h$. But by the results of [8,10], p and q can never intersect.

We conclude that no intersection can occur until the final flattened configuration of the band, which is a singularity where the above arguments do not apply. By Theorem 10, there is a cut that produces an unfolded state. Therefore, by making the same cut and applying the squeezing motion, we obtain a continuous unfolding of the band. □

6 Remarks

We note that another natural continuous unfolding motion exists, consisting of $n-1$ peeling moves. After cutting a hinge that produces an unfolded state, we begin by performing a dihedral rotation about its neighboring hinge, so that two trapezoids become coplanar. Subsequent moves are simple dihedral rotations about successive hinges, and each step adds one more trapezoid to the coplanar subset. Because this motion is not necessary for our results on nested bands, a detailed proof of its correctness is omitted. We mention it, though, because follow-on work establishes that this motion unfolds non-nested bands, even those that contain polyhedron vertices on their boundaries [1].

Even with it established that arbitrary bands can be unfolded without overlap, it remains interesting to see whether this can lead to a non-overlapping unfolding of prismatoids, including the top and bottom faces. It is natural to hope that these faces could be nestled on opposite sides of the unfolded band, but we do not know how to ensure non-overlap.

Acknowledgments

This work was initiated at the 19th Bellairs Winter Workshop on Computational Geometry held January 30–February 6, 2004 at Bellairs Research Institute in Barbados. We thank the other participants of that workshop—Prosenjit Bose, Mirela Damian, Vida Dujmović, Ferran Hurtado, John Iacono, Erin Mcleish, Henk Meijer, Perouz Taslakian, Sue Whitesides, and David Wood—for helpful discussions and contributing to a fun and creative atmosphere. We also thank Martin Demaine for helpful discussions.

References

[1] Greg Aloupis. *Reconfigurations of Polygonal Structures.* Ph.D. Thesis, McGill University, March 2005.

[2] Gustavo A. Arteca and Paul G. Mezey. A method for the characterization of foldings in protein ribbon models. *Journal of Molecular Graphics*, 8:66–80, June 1990.

[3] Prabir Bhattacharya and Azriel Rosenfeld. Polygonal ribbons in two and three dimensions. *Pattern Recognition*, 28(5):769–779, 1995.

[4] Ciprian Borcea and Ileana Streinu. Singularities of hinge structures. In *Effective Methods in Algebraic Geometry*, Porto Conte, Sardinia, May 2005.

[5] H. Martyn Cundy and Arthur P. Rollett. *Mathematical Models.* Oxford University Press, 1961.

[6] Erik D. Demaine and Joseph O'Rourke. *Geometric Folding Algorithms: Linkages, Origami, and Polyhedra.* Cambridge University Press, to appear in 2006 or 2007.

[7] Erik D. Demaine and Joseph O'Rourke. A survey of folding and unfolding in computational geometry. in *Combinatorial and Computational Geometry*, edited by Jacob E. Goodman, Janos Pach and Emo Welzl, Mathematical Sciences Research Institute Publications, volume 52, August 2005, pages 167-211, Cambridge University Press.

[8] Joseph O'Rourke. On the development of the intersection of a plane with a polytope. *Comput. Geom. Theory Appl.*, 24(1):3–10, 2003.

[9] Joseph O'Rourke. Folding and unfolding in computational geometry. In *Discrete Comput. Geom.*, volume 1763 of *Lecture Notes Comput. Sci.*, pages 258–266. Springer-Verlag, 2000. Papers from the Japan Conf. Discrete Comput. Geom., Tokyo, Dec. 1998.

[10] Joseph O'Rourke. An extension of Cauchy's arm lemma with application to curve development. In *Discrete Comput. Geom.*, volume 2098 of *Lecture Notes Comput. Sci.*, pages 280–291. Springer-Verlag, 2001. Papers from the Japan Conf. Discrete Comput. Geom., Tokyo, Nov. 2000.

[11] Ivan Reitman (director), and Dan Akroyd and Harold Ramis (writers). *Ghost Busters.* Colombia Pictures, 1984.

[12] G. C. Shephard. Convex polytopes with convex nets. *Math. Proc. Cambridge Philos. Soc.*, 78:389–403, 1975.

[13] Walter Whiteley. Rigidity of molecular structures: generic and geometric analysis. In *Rigidity Theory and Applications*, M. F. Thorpe and D. M. Duxbury, eds., 1999, Kluwer Academic, pp. 21-46.

GRID VERTEX-UNFOLDING ORTHOSTACKS*

ERIK D. DEMAINE[†]

MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar St., Cambridge, MA 02139, USA

edemaine@mit.edu

JOHN IACONO[‡]

Department of Computer and Information Science, Polytechnic University, 5 MetroTech Center, Brooklyn, NY 11201, USA

http://john.poly.edu

STEFAN LANGERMAN[§]

Département d'informatique, Université Libre de Bruxelles, ULB CP212, 50 Av. F.D. Roosevelt, 1050 Brussels, Belgium

Stefan.Langerman@ulb.ac.be

Received June 1, 2005
Revised October 20, 2006
Communicated by Godfried Toussaint

Biedl et al. [1] presented an algorithm for unfolding orthostacks into one piece without overlap by using arbitrary cuts along the surface. They conjectured that orthostacks could be unfolded using cuts that lie in a plane orthogonal to a coordinate axis and containing a vertex of the orthostack. We prove the existence of a vertex unfolding using only such cuts.

Keywords: Edge unfolding; orthogonal polyhedra; cutting; folding.

1. Introduction

A long-standing open question is whether every convex polyhedron can be *edge unfolded*—cut along some of its edges and unfolded into a single planar piece without overlap [12,11,7,10]. A related open question asks whether every polyhedron[a] (not

*A preliminary version of this paper appeared in *Revised Selected Papers from the Japan Conference on Discrete and Computational Geometry*, Tokyo, Oct. 2004, LNCS 3742, 2005, pages 76–82.

[†]Research supported in part by NSF grants CCF-0347776, OISE-0334653, and CCF-0430849, and by DOE grant DE-FG02-04ER25647.

[‡]Research supported in part by NSF grants OISE-0334653 and CCF-0430849.

[§]Chercheur qualifié du FNRS.

[a]A *polyhedron (without boundary)* is an embedded connected polyhedral complex without boundary, i.e., a connected set of polygons in Euclidean 3-space such that (1) every two polygons meet

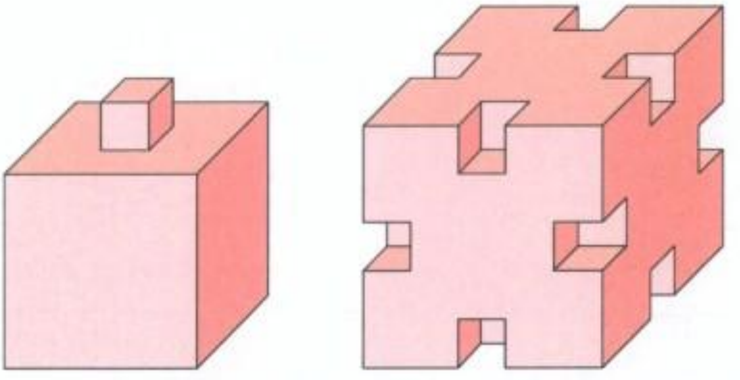


Fig. 1. These orthostacks are not edge-unfoldable [1]. The first one is also not vertex-unfoldable.

necessarily convex but forming a closed surface) can be *generally unfolded*—cut along its surface (not just along edges) and unfolded into a single planar piece without overlap. Biedl et al. [1] made partial progress on both of these problems in the context of *orthostacks*. An orthostack is an orthogonal polyhedron[b] of which every horizontal planar slice not including a horizontal face is a single simple (orthogonal) polygon. Biedl et al. showed that not all orthostacks can be edge unfolded (see Figure 1), but that all orthostacks can be generally unfolded. In their general unfoldings, all cuts are parallel to coordinate axes, but many of the cuts do not lie in coordinate planes that contain polyhedron vertices. Given the lack of pure edge unfoldings, the closest analog we can hope for with (nonconvex) orthostacks is to find *grid unfoldings* in which every cut is in a coordinate plane that contains a polyhedron vertex. In other words, a grid unfolding is an edge unfolding of the refined ("gridded") polyhedron in which we slice along every coordinate plane containing a polyhedron vertex. Biedl et al. [1] asked whether all orthostacks can be grid unfolded.

We make partial progress on this problem by showing that every orthostack can be *grid vertex-unfolded*, i.e., cut along some of the grid lines and unfolded into a vertex-connected planar piece without overlap. Vertex unfoldings were introduced in [8,9]; the difference from edge unfoldings is that faces can remain connected along single points (vertices) instead of having to be connected along whole edges. As before, a vertex unfolding must be a single planar piece without overlap. In fact, our vertex unfoldings consist of a single path of polygons, with consecutive polygons connected together at common vertices. Furthermore, as argued in [8,9], connections

at either a common vertex, a common edge, or not at all; (2) every edge is incident to exactly two polygons; and (3) every vertex is incident to exactly a topological disk of polygons, with only cyclically adjacent polygons sharing an edge. Note that a polyhedron is treated as a surface throughout this paper.

[b]An *orthogonal polyhedron* is a polyhedron (without boundary) in which every face is perpendicular to a coordinate axis. This definition implies that every face is an orthogonal polygon.

through a vertex never need to cross: for four incident faces A, B, C, D in cyclic order around a vertex v if a vertex unfolding connects A to C and B to D both via v, we can uncross the connection and keep the unfolding a single path by making different connections through v. Our unfolding places faces orthogonally into the plane: all edges of the unfolded faces are parallel to a coordinate axis. (This property is not forced by gridness in vertex unfoldings.) Our unfolding may, however, place faces so as to touch along boundary edges; we guarantee nonoverlap only of polygon interiors.

Our use of grid refinement seems to be necessary for vertex-unfolding, because the box-on-box example in Figure 1(left) has no vertex-unfolding if we are allowed to cut only along edges. It remains open whether there is such an example requiring grid cuts for a vertex-unfolding, but where every face has no holes (i.e., is homeomorphic to a disk).

Since the conference version of this paper, Damian et al. [5] generalized our techniques to grid vertex-unfold all orthogonal polyhedra of genus zero. Also, by further axis-parallel refinement of an orthogonal polyhedron beyond the grid, they have shown how to edge-unfold "orthostacks with orthogonally convex slabs" [6], "Manhattan towers" [3], "well-separated orthotrees" [2], and general orthogonal polyhedra [4]. The last case requires an exponential amount of refinement, making the two special cases of interest.

2. Grid Vertex Unfolding

Given an orthostack K, let $z_0 < z_1 < \cdots < z_n$ be the distinct z coordinates of vertices of K. Refer to Figure 2. Subdivide the faces of K by cutting along every plane perpendicular to a coordinate axis that passes through a vertex of K. This subdivision *rectangulates* K We use the term *rectangle* to refer to one element of this facial subdivision, while *face* refers to a maximal edge-connected set of coplanar rectangles. (Thus faces can have holes, but at most one in an orthostack.) We use *up* and *down* to refer to the z dimension, and use *left* and *right* to refer to the x dimension.

2.1. *Rectangle Categorization*

We partition the rectangles of K into several categories. After this categorization, the description of the unfolding layout is not difficult.

For $i = 0, 1, \ldots, n-1$, define the *i-band* to be the set of vertical rectangles (i.e., that lie in an xz plane or in a yz plane) whose z coordinates are between z_i and z_{i+1}. By the definition of rectangles, all of the rectangles of an i-band have the same extent in the z dimension, namely, $[z_i, z_{i+1}]$. By the definition of an orthostack, each i-band is connected, forming the boundary of an extruded simple orthogonal polygon.

For $i = 0, 1, \ldots, n$, we define the *i-faces* to be the faces of K in the horizontal plane $z = z_i$. As we have defined them, an i-face has several properties. It may

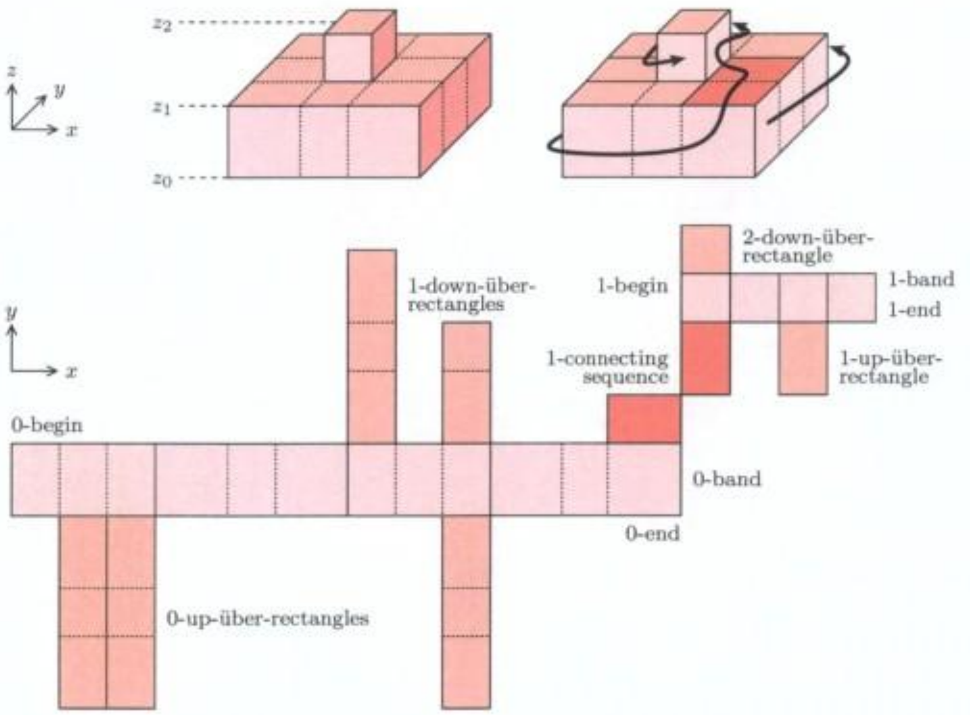


Fig. 2. Top-left: A rectangulated orthostack K with three distinct z coordinates z_0, z_1, z_2. Top-right: Categorization into i-band rectangles (light), i-über rectangles (medium), and i-connecting rectangles (dark); and the tour visiting i-band and i-connecting rectangles. Bottom: The resulting unfolding.

have the interior of K above or below it (but not both). The perimeter of the i-face (both perimeters if the i-face has a hole) has a nonempty intersection with the $(i-1)$-band, provided $i > 0$, and with the i-band, provided $i < n$. (If an i-face f is incident to only the i-band, then all edges of f must be incident to vertical faces above $z = z_i$, which form a cycle of faces in the i-band, so by connectivity of the i-band no other i-face can be incident to the i-band; also, by connectivity of the polyhedron, there cannot be another i-face meeting only the $(i-1)$-band; so f must be the bottom face of the polyhedron. Similarly, an i-face incident to only the $(i-1)$-band must be the top face of the polyhedron.)

We also need the notions of the "begin rectangle" and "end rectangle" of the i-band. Choose the 0-*band begin rectangle* to be an arbitrary rectangle of the 0-band. For $i \geq 0$, define the i-*band end rectangle* to be the rectangle of the i-band that is adjacent to the i-band begin rectangle in the clockwise direction as viewed from $+z$. For $i \geq 1$, define the i-*connecting face* to be the i-face that shares an edge with the $(i-1)$-band end rectangle, if such a face exists. Thus, the i-connecting face does not exist if and only if the $(i-1)$-band end rectangle shares an edge with the i-band. For $i \geq 1$, choose the i-*band begin rectangle* to be one of the rectangles of the i-band that shares an edge with the i-connecting face, if it exists, or else the rectangle of the i-band that shares an edge with the $(i-1)$-band end rectangle. The *i-band interior rectangles* are rectangles of the i-band that are neither the begin rectangle nor the end rectangle.

Define the *i-connecting sequence* to be an arbitrarily chosen edge-connected sequence of rectangles in the i-connecting face, if it exists, starting at the rectangle that shares an edge with the $(i-1)$-band end rectangle and ending at the rectangle that shares an edge with the i-band begin rectangle. This sequence is chosen to contain the fewest rectangles possible (a shortest path in the dual graph on the rectangles in the i-connecting face), in order to prevent the path from looping around an island and thereby isolating interior portions of the i-face. If the i-connecting face does not exist, the i-connecting sequence is the empty sequence. The rectangles in the i-connecting sequence are called *i-connecting rectangles*; all other rectangles of the i-faces are called *normal rectangles*.

We now merge all normal rectangles with their normal neighbors in the x dimension. Call the resultant rectangular regions *über-rectangles*. Thus i-faces are partitioned into the i-connecting rectangles and the i-über-rectangles. Every i-über-rectangle is connected to the perimeter of an i-face; otherwise, the rectangles that compose it could be used to construct a shorter i-connecting path. Thus, every i-über-rectangle shares an edge with either the $(i-1)$-band or the i-band (or both). Define an *i-up-über-rectangle* to be an über-rectangle that is incident to the i-band and an *i-down-über-rectangle* to be an über-rectangle that is incident to the $(i-1)$-band. If an über-rectangle is incident to both, we classify it arbitrarily.

Thus we have partitioned K into i-band begin rectangles, i-band end rectangles, i-band interior rectangles, i-up-über-rectangles, i-down-über-rectangles, and i-connecting rectangles. We now proceed to a description of the unfolding.

2.2. *Unfolding Algorithm*

Our unfolding of an orthostack consists of several components strung together at distinguished rectangles called *anchors*. Specifically, there are two types of components, i-main components and i-connecting components, both of which are anchored at two rectangles, a begin rectangle and an end rectangle. The *i-main component* consists of the entire i-band (the i-band begin rectangle, the i-band interior rectangles, and the i-band end rectangle), the $(i+1)$-down-über-rectangles, and the i-up-über-rectangles. The *i-connecting component* consists of the $(i-1)$-band end rectangle, the i-connecting rectangles (if any), and the i-band begin rectangle. It serves to connect the $(i-1)$-main component and the i-main component (at the $(i-1)$-band end rectangle and the i-band begin rectangle, respectively).

To ensure that components do not overlap each other, we enforce that the components are *anchored* in the following sense. A component is anchored at anchor rectangles R and S if, in the unfolded layout of the component, no rectangles are in the hatched region of Figure 3. More precisely, every rectangle is strictly right of

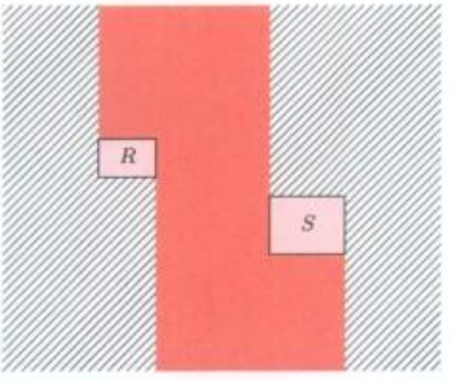


Fig. 3. A component anchored at R and S must avoid the hatched regions, remaining within the shaded region.

R and strictly left of S, or directly above R, or directly below S.

We can combine two anchored components with a common anchor while avoiding overlap. More precisely, given a component C anchored at anchors R and S, and another component C' anchored at S and T with the same orientation of S, we can combine the two unfolded layouts by translating C' so that the two copies of S coincide (with matching orientations). The conditions on the rectangles in the two components C and C' guarantee nonoverlap of the combined unfolded layout. To guarantee the matching orientations of anchors, we enforce that the positive z direction of every vertical (i-band) rectangle becomes the positive y direction in the planar unfolding.

We edge-unfold the i-main component by leaving one edge attached between the über-rectangles of the component (arbitrarily, if there is a choice), and cutting along all of the other edges of the über-rectangles. As shown in Figure 4, the layout induced by this edge unfolding consists of a central horizontal rectangular strip, which contains all i-band rectangles, and has the $(i+1)$-down-über-rectangles connected to the top of this strip, and the i-up-über-rectangles connected to the bottom of this strip. The leftmost rectangle of this strip is the i-band begin rectangle, and the rightmost rectangle of the strip is the i-band end rectangle. There is nothing below the leftmost rectangle or above the rightmost rectangle because these vacant locations are where the connecting rectangles are attached, and connecting rectangles are not über-rectangles. (In the special cases $i = 0$ and $i = n$, there can be an über-rectangle below the leftmost rectangle and above the rightmost rectangle, respectively, but in these cases, we can choose to attach the über-rectangle at its opposite edge.) Therefore the edge unfolding of the i-main component is anchored at the i-band begin and end rectangles.

We vertex-unfold the i-connecting component by a sequence of modifications to the edge-unfolding of the rectangles in the component. Let $R_0, R_1, \dots, R_k$ denote these rectangles in connected order, where R_0 is the $(i-1)$-band end rectangle and

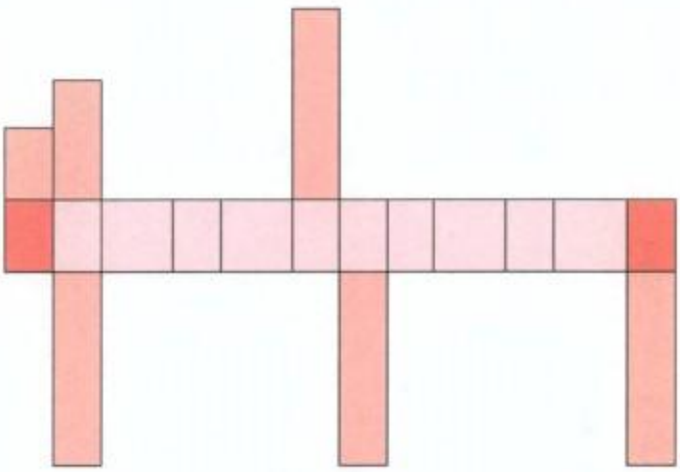


Fig. 4. An example of an unfolded i-main component. The dark rectangles are the i-band begin rectangle (left) and i-band end rectangle (right). They are connected by the remainder of the i-band (light). Above the i-band are the $(i+1)$-down-über-rectangles and below are the i-up-über-rectangles (medium). This example is a possible outcome for the 0-main component of Figure 2.

R_k is the i-band begin rectangle. The i-connecting rectangles $R_1, R_2, \ldots, R_{k-1}$ all come from an i-face, so they were planar even before the edge unfolding. The $(i-1)$-band end rectangle R_0 is adjacent to R_1 along the edge originally in the positive z direction; we rotate the edge-unfolding so that this edge is the top edge of R_0, with R_1 stacked above. Now for $2 \leq j < k$, assume that $R_0, R_1, \ldots, R_{j-1}$ have been placed, and R_{j-1} and R_j remain connected at a common edge which is not the left edge of R_{j-1}. There are three cases, depending on whether R_j shares the top, bottom, or right edge of R_{j-1}; see Figure 5. In the third case, we do nothing; in the first two cases, we vertex-unfold R_j by $90°$ around the right endpoint of the shared edge. After this step, R_{j+1} lies in one of the dark shaded squares, sharing R_j's top, bottom, or right edge, so the induction proceeds. We handle the i-band begin rectangle R_k differently to guarantee the proper orientation. Again there are three cases, depending on whether R_k shares the top, bottom, or right edge of R_{k-1}; see Figure 6. The shared edge corresponds the edge of R_k in the negative z direction, so in each case we vertex-unfold if necessary to make that edge the bottom edge in the unfolding. In the end, each rectangle R_j is strictly right of the previous rectangles, except R_k which might be on top of R_{k-1}. Thus, the anchored unfolding of the i-connecting component does not self-intersect.

By combining the anchored unfoldings of the 0-main component, the 1-connecting component, the 1-main component, etc., the $(n-1)$-main component, the $(n-1)$-connecting component, and the n-main component, we obtain the desired vertex unfolding:

Theorem 1. *Every orthostack can be grid vertex-unfolded.*

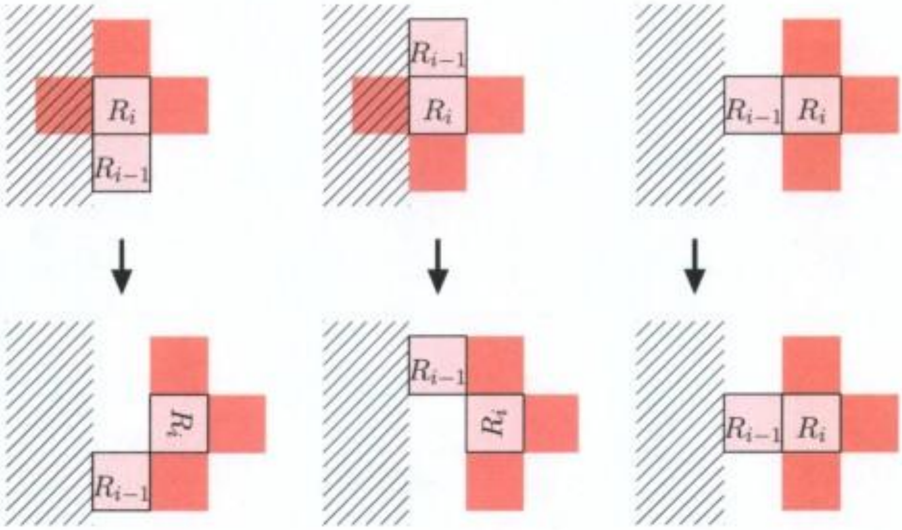


Fig. 5. How to vertex-unfold R_i after $R_0, R_1, \ldots, R_{i-1}$ have been placed (all but the last of which are in the hatched region). There are three cases, from left to right: R_i above, R_i below, and R_i to the right. In all cases, R_{i+1} is in one of the dark shaded regions, which is never left of R_i after vertex-unfolding. The illustrated unfoldings work no matter what are the sizes of the rectangles.

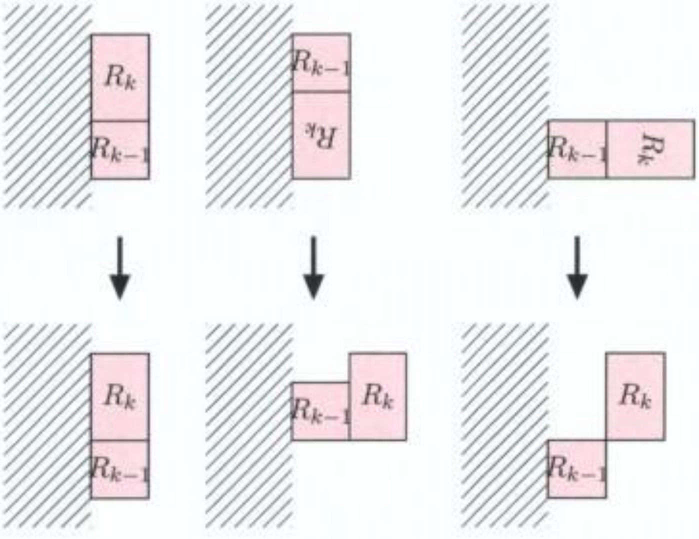


Fig. 6. How to vertex-unfold the last rectangle R_k after $R_0, R_1, \ldots, R_{k-1}$ have been placed (all but the last of which are in the hatched region). There are three cases, from left to right: R_k above, R_k below, and R_k to the right. In all cases, we must orient R_k so that the edge opposite R_{k-1} is on top. The illustrated unfoldings work no matter what are the sizes of the rectangles.

The construction leads to an algorithm whose running time is linear in the number of rectangles, which is at most quadratic in the combinatorial complexity of the polyhedron.

Acknowledgments

This work was initated while the authors visited McGill University's Computational Geometry Lab. We thank Mirela Damian and Joseph O'Rourke for helpful discussions. We also thank Koichi Hirata, Joseph O'Rourke, and an anonymous referee for helpful comments on the paper.

References

1. T. Biedl, E. Demaine, M. Demaine, A. Lubiw, M. Overmars, J. O'Rourke, S. Robbins, and S. Whitesides. Unfolding some classes of orthogonal polyhedra. In *Proceedings of the 10th Canadian Conference on Computational Geometry*, Montréal, Canada, August 1998.
2. M. Damian, R. Flatland, H. Meijer, and J. O'Rourke. Unfolding well-separated orthotrees. In *Abstracts from the 15th Annual Fall Workshop on Computational Geometry*, Philadelphia, PA, November 2005.
3. M. Damian, R. Flatland, and J. O'Rourke. Unfolding manhattan towers. In *Proceedings of the 17th Canadian Conference on Computational Geometry*, pages 211–214, Windsor, Canada, August 2005.
4. M. Damian, R. Flatland, and J. O'Rourke. Epsilon-unfolding orthogonal polyhedra. Technical Report 082, Smith College, February 2006. arXiv:cs.CG/0602095.
5. M. Damian, R. Flatland, and J. O'Rourke. Grid vertex-unfolding orthogonal polyhedra. In *Proceedings of the 23rd Annual Symposium on Theoretical Aspects of Computer Science*, volume 3884 of *Lecture Notes in Computer Science*, pages 264–276, Marseille, France, February 2006. arXiv:cs.3013175.
6. M. Damian and H. Meijer. Edge-unfolding orthostacks with orthogonally convex slabs. In *Abstracts from the 14th Annual Fall Workshop on Computational Geometry*, pages 20–21, Cambridge, MA, November 2004.
7. E. D. Demaine. Folding and unfolding linkages, paper, and polyhedra. In *Revised Papers from the Japan Conference on Discrete and Computational Geometry*, volume 2098 of *Lecture Notes in Computer Science*, pages 113–124, Tokyo, Japan, November 2000.
8. E. D. Demaine, D. Eppstein, J. Erickson, G. W. Hart, and J. O'Rourke. Vertex-unfolding of simplicial manifolds. In *Proceedings of the 18th Annual ACM Symposium on Computational Geometry*, pages 237–243, Barcelona, Spain, June 2002.
9. E. D. Demaine, D. Eppstein, J. Erickson, G. W. Hart, and J. O'Rourke. Vertex-unfolding of simplicial manifolds. In *Discrete Geometry: In Honor of W. Kuperberg's 60th Birthday*, pages 215–228. Marcer Dekker Inc., 2003.
10. E. D. Demaine and J. O'Rourke. A survey of folding and unfolding in computational geometry. In J. E. Goodman, J. Pach, and E. Welzl, editors, *Discrete and Computational Geometry*, Mathematical Sciences Research Institute Publications. Cambridge University Press, 2005. To appear.
11. J. O'Rourke. Folding and unfolding in computational geometry. In *Revised Papers from the Japan Conference on Discrete and Computational Geometry*, volume 1763 of *Lecture Notes in Computer Science*, pages 258–266, Tokyo, Japan, December 1998.

12. G. C. Shephard. Convex polytopes with convex nets. *Mathematical Proceedings of the Cambridge Philosophical Society*, 78:389–403, 1975.

Wrapping the Mozartkugel

Erik D. Demaine* Martin L. Demaine* John Iacono† Stefan Langerman‡

Abstract

We study wrappings of the unit sphere by a piece of paper (or, perhaps more accurately, a piece of foil). Such wrappings differ from standard origami because they require infinitely many infinitesimally small "folds" in order to transform the flat sheet into a positive-curvature sphere. Our goal is to find shapes that have small area even when expanded to shapes that tile the plane. We characterize the smallest square that wraps the unit sphere, show that a 0.1% smaller equilateral triangle suffices, and find a 20% smaller shape that still tiles the plane.

Keywords: chocolate, marzipan, praline, nougat

1 Introduction

The Mozartkugel ("Mozart sphere") [9, 8] is a famous fine Austrian confectionery: a sphere with marzipan in its core, encased in nougat or praline cream, and coated with dark chocolate. It was invented in 1890 by Paul Fürst in Salzburg (where Wolfgang Amadeus Mozart was born), six years after he founded his confectionery company, Fürst. Fürst (the company) still to this day makes Mozartkugeln by hand, about 1.4 million per year, under the name "Original Salzburger Mozartkugel" [6]. At the 1905 Paris Exhibition, Paul Fürst received a gold medal for the Mozartkugel.

Many other companies now make similar Mozartkugeln, but Mirabell is the market leader with their "Echte (Genuine) Salzburger Mozartkugeln" [7]. Over 1.5 billion have been made, about 90 million per year, originally by hand but now by industrial methods, and Mirabell claims their product to be the only Mozartkugel that is perfectly spherical. They are also the only Mozartkugel to be taken into outer space, by the first Austrian astronaut Franz Viehböck as a gift to the Russian cosmonauts on the MIR space station. Despite industrial techniques, each Mozartkugel still takes about 2.5 hours to make.

Although most of a Mozartkugel is edible, each sphere is individually wrapped in a square of aluminum foil. To minimize the amount of this wasted, inedible material, it is natural to study the smallest piece of foil that can wrap a unit sphere. Because the pieces will be cut from a large sheet of foil, we would also like the unfolded shape to tile the plane.

We formalize this practical problem in the next section; the main difficulty is to allow a continuum of infinitesimal folds to curve the paper, a feature not normally modeled by mathematical origami. We then study wrappings by squares and equilateral triangles, and show that the latter leads to a small (0.1%) savings, which may prove significant on the many millions of Mozartkugel consumed each year. Even better, if we allow wrapping by arbitrary shapes that tile the plane, we show how to achieve a 20% savings. In addition to direct savings in material costs for Mozartkugel manufacturers, the reduced material usage also indirectly cuts down on CO_2 emissions, and therefore partially solves the global-warming problem and consequently the little-reported but equally important chocolate-melting problem.

2 Wrapping Problem

In standard mathematical origami [4, 5], a *piece of paper* is a two-dimensional manifold (usually flat), and a *folding* is an isometric mapping of this piece of paper into Euclidean 3-space. Here *isometric* means that distances are preserved, as measured by shortest paths on the piece of paper before and after mapping via the folding.

But there is no isometric folding of a square into a sphere: isometric folding preserves curvature. Therefore we define a new, less restrictive type of folding that allows changing curvature but still prevents stretching of the material. Namely, a *wrapping* is a continuous contractive mapping of a piece of paper into Euclidean 3-space. Here *contractive* means that every distance either decreases or stays the same, as measured by shortest paths on the piece of paper before and after mapping via the folding. This definition effectively assumes that the length contraction can be achieved by continuous infinitesimal pleating.

We can model one family of wrappings by expressing which distances are preserved isometrically. An optimal wrapping should be isometric along some path, for otherwise we could uniformly scale the entire wrapping and make a larger object. We call a path *stretched* if the wrapping is isometric along it. A *stretched wrapping* has the property that every point is covered by some stretched path. Such a wrapping can be specified by a set of stretched paths whose

*MIT, {edemaine,mdemaine}@mit.edu

†Polytechnic University, http://john.poly.edu/

‡Chercheur qualifié du FNRS, Université Libre de Bruxelles, stefan.langerman@ulb.ac.be

union covers the entire piece of paper. Although not all such specifications are valid—we need to check that all other paths are contractive—the specification does uniquely determine a wrapping. We specify all of our wrappings in this way, under the belief that stretched wrappings are generally the most efficient.

A special case of stretched wrapping is when the stretched paths consist of the shortest paths from one point x to every other point y. In this case, we are rolling geodesics in the piece of paper onto geodesics of the target surface. This situation corresponds to continuous unfoldings of smooth polyhedra as considered by Benbernou, Cahn, and O'Rourke [1]. Although perhaps the most natural kind of wrapping, this special case is too restrictive for our purposes, as it essentially forces the sphere to be wrapped by a disk of radius π, for those geodesics to reach around to the pole opposite x. We will show how to wrap with far less paper than this disk of area π^3.

Note that, if we start with an arbitrarily long and narrow rectangle, we can wrap the sphere using paper area arbitrarily close to the surface area 4π of the sphere [3]. This wrapping is not very practical, however; in particular, it makes it difficult to make a nondistorted logo on the surface of the sphere.

The only other known optimal wrapping result (where no contraction is necessary) is wrapping a unit cube with a square [2].

3 Petal Wrapping

Our wrappings are based on the following *k-petal wrapping*. On the sphere we first construct k stretched paths $p_1, p_2, \ldots, p_k$ from the south pole to the north pole, dividing the 2π angle around each pole into k equal parts of $2\pi/k$. To each path p_i we assign an "orange peel" with apex angles $2\pi/k$, centered on the path p_i and bounded by the Voronoi diagram of p_{i-1}, p_i, p_{i+1}. These orange peels partition the surface of the sphere into k equal pieces.

Then we construct a continuum of stretched paths to cover each orange peel. Specifically, for every point q along each path p_i, we construct two stretched paths emanating from q, proceeding along geodesics perpendicular to p_i in both directions, and stopping at the boundary of p_i's orange peel.

These stretched paths cover every point of the sphere (covering boundary points twice). It remains to find a suitable piece of paper that wraps according to these stretched paths. The main challenge is to unfold the half of an orange peel left of a path p_i. Then we can easily glue the two halves together along the (straight) unfolded path p_i, and finally join the resulting petals at the unfolded south pole.

To unroll half of a petal, we parameterize as shown in Figure 1. Here $B = \pi/k$ is the half-petal angle; c is a given amount that we traverse along the center path p_i starting at the south-pole endpoint; $A = \pi/2$ specifies that we turn perpendicular from that point; and b is the distance that we travel in that direction. Our goal is to determine b in terms of c.

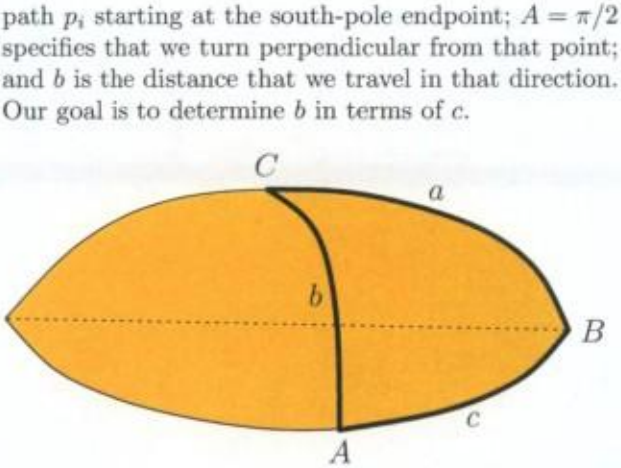


Figure 1: Half of a petal, labeled in preparation for spherical trigonometry.

By the spherical law of cosines,

$$\cos C = -\cos A \cos B + \sin A \sin B \cos c.$$

Now $\cos A = \cos(\pi/2) = 0$ and $\sin A = \sin(\pi/2) = 1$, so this equation simplifies to $\cos C = \sin B \cos c$. Hence, $\sin C = \sqrt{1 - \sin^2 B \cos^2 c}$. By the spherical law of sines,

$$\frac{\sin B}{\sin b} = \frac{\sin C}{\sin c}.$$

Substituting $\sin C = \sqrt{1 - \sin^2 B \cos^2 c}$, we obtain

$$\frac{\sin B}{\sin b} = \frac{\sqrt{1 - \sin^2 B \cos^2 c}}{\sin c},$$

i.e.,

$$\sin b = \frac{\sin B \sin c}{\sqrt{1 - \sin^2 B \cos^2 c}}.$$

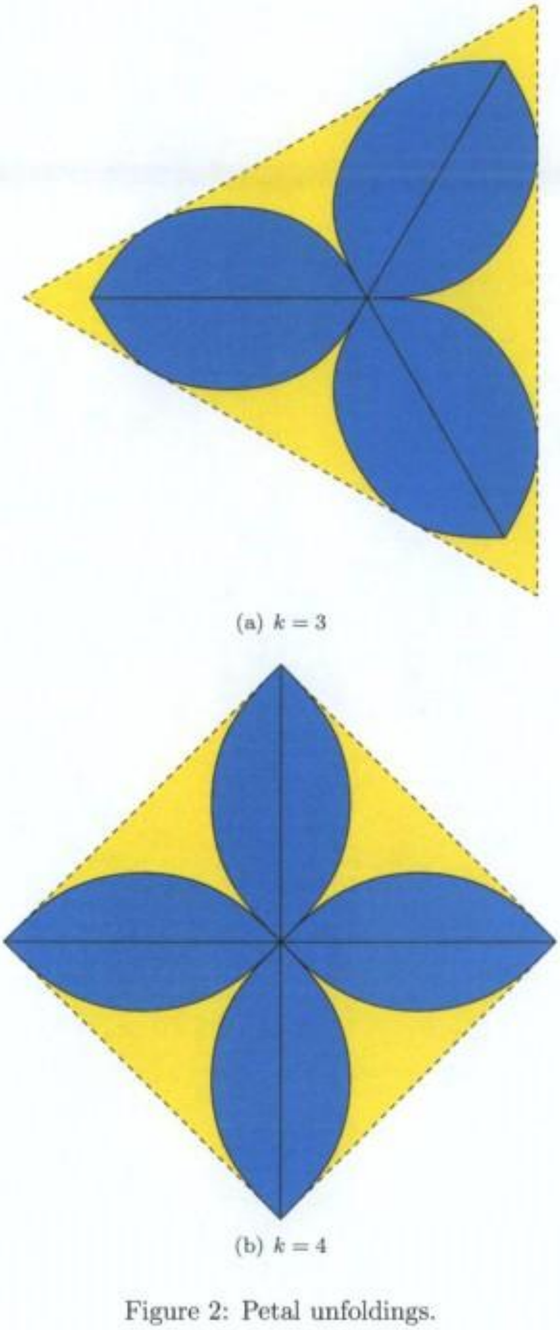
Taking arccos of both sides, we determine the value of b in terms of the parameter c and the known quantity $B = \pi/k$.

Figure 2 shows two examples of the resulting petal unfolding, with $k = 3$ and $k = 4$.

4 Square Wrapping

The angle at the tip of the petals can be computed by taking the derivative $\partial b/\partial c$ at $c = 0$. For $k = 4$, this derivative is 1 which implies a half angle of $\pi/4$. Because the petals are convex, the convex hull of the petal unfolding for $k = 4$ is exactly the square of diagonal 2π. No smaller square could wrap the unit sphere because the length of the path connecting the center of the square to the point mapped to the antipodal point must have length at least π. This square has area $2\pi^2$.

Note that the same area is attainable by a rectangle of dimensions $2\pi \times \pi$: draw one path p around the



(a) $k=3$

(b) $k=4$

Figure 2: Petal unfoldings.

equator of the sphere and cover the sphere by a continuum of stretched paths perpendicular to p emanating from every point of p until the north and the south pole of the sphere. The same rectangle is also exactly a 2-petal unfolding. Interestingly, the area of this rectangle wrapping is also 2π. The Echte Salzburger Mozartkugel is wrapped by Mirabell using the same rectangle (expanded a bit to ensure overlap) but with a slightly different folding.

5 Triangle Wrapping

For $k=3$, the angle at the tip of the petals can be computed similarly to obtain $2\pi/3$, which is natural as the three petals meet at the north pole, their angles summing to 2π. However, the convex hull of the 3-petal unfolding is not a triangle. We compute its smallest enclosing equilateral triangle. The supporting lines of the triangle will be each tangent to two of the petals. The tangent point on the petal can be computed by finding the point (c,b) on its boundary that maximizes the direction $(-\cos(\pi/3), \sin(\pi/3))$. Plugging this into the previous equations, we obtain

$$c = \arccos\left(\tfrac{\sqrt{57}}{6} - \tfrac{1}{2}\right) \approx 0.710086.$$

This implies that the supporting line is at a distance

$$\tfrac{\pi}{2} - \tfrac{1}{2}\arccos\left(\tfrac{\sqrt{57}}{6} - \tfrac{1}{2}\right) + \tfrac{\sqrt{3}}{2}\arcsin\left(\frac{\sqrt{\sqrt{57}-5}}{\sqrt{\sqrt{57}-3}}\right) \approx 0.620190\pi$$

from the center. The area of the inscribing equilateral triangle is therefore $3h^2\tan(\pi/6) \approx 1.998626\,\pi^2$, about 0.1% less than the $2\pi^2$ area of the smallest wrapping square.

6 Tiling

Instead of expanding the petal unfoldings to tilable regular polygons, we can pack the petal unfoldings directly and expand them just to fill the extra space. Figure 3 shows an even better tiling resulting from the 3-petal unfolding. A quick computation shows that only about $1.6\,\pi^2$ area of paper is required for each wrapping, a substantial improvement.

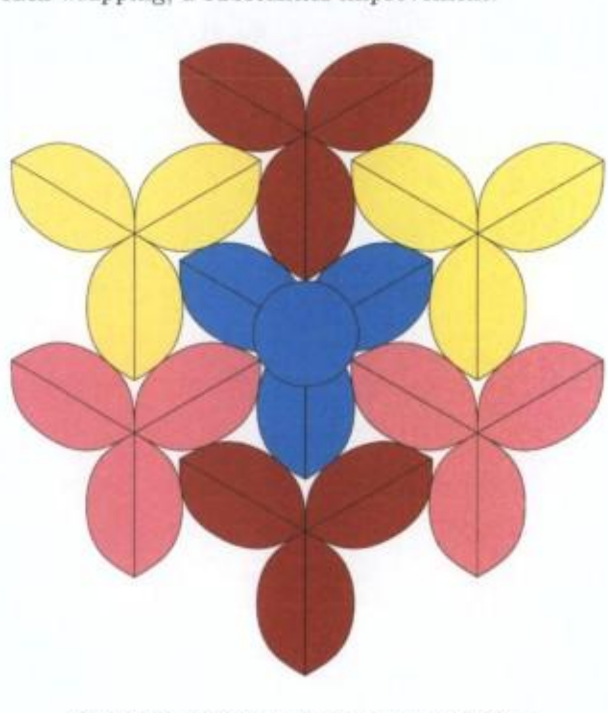


Figure 3: Packing the 3-petal unfolding.

7 Conclusion

This paper initiates a new research direction in the area of *computational confectionery*. We leave as open problems the study of wrapping other geometric confectioneries, or further improving our wrappings of the Mozartkugel. In particular, what is the optimal convex shape that can wrap a unit sphere? What is the optimal shape that also tiles the plane? What about smooth surfaces other than the sphere?

Acknowledgements

The authors thank Luc Devroye, Vida Dujmović, and Dania El-Khechen for helpful discussions, the anonymous referees for their comments, and the Café du Monde for their inspirational beignets.

References

[1] Nadia Benbernou, Patricia Cahn, and Joseph O'Rourke. Unfolding smooth prismatoids. In *14th Annual Fall Workshop on Computational Geometry*, Cambridge, MA, 2004. arXiv:cs.CG/0407063.

[2] Michael L. Catalano-Johnson and Daniel Loeb. Problem 10716: A cubical gift. *American Mathematical Monthly*, 108(1):81–82, January 2001. Posed in volume 106, 1999, page 167.

[3] Erik D. Demaine, Martin L. Demaine, and Joseph S. B. Mitchell. Folding flat silhouettes and wrapping polyhedral packages: New results in computational origami. *Computational Geometry: Theory and Applications*, 16(1):3–21, 2000.

[4] Erik D. Demaine, Satyan L. Devadoss, Joseph S. B. Mitchell, and Joseph O'Rourke. Continuous foldability of polygonal paper. In *Proceedings of the 16th Canadian Conference on Computational Geometry*, pages 64–67, Montréal, Canada, August 2004.

[5] Erik D. Demaine and Joseph O'Rourke. *Geometric Folding Algorithms: Linkages, Origami, and Polyhedra*. Cambridge University Press, to appear.

[6] Fürst. Original salzburger mozartkugel. http://www.original-mozartkugel.com/.

[7] Mirabell. The brand. http://www.mozartkugel.at/.

[8] Die freie Enzyklopädie Wikipedia. Mozartkugel. http://de.wikipedia.org/wiki/Mozartkugel.

[9] The Free Encyclopedia Wikipedia. Mozartkugel. http://en.wikipedia.org/wiki/Mozartkugel.