Contemporary Mathematics

All Polygons Flip Finitely... Right?

Erik D. Demaine, Blaise Gassend, Joseph O'Rourke, and Godfried T. Toussaint

Abstract. Every simple planar polygon can undergo only a finite number of pocket flips before becoming convex. Since Erdős posed this finiteness as an open problem in 1935, several independent purported proofs have been published. However, we uncover a plethora of errors, gaps, and omissions in these arguments, leaving only two proofs without flaws and no proof that is fully detailed. Fortunately, the result remains true, and we provide a new, simple (and correct) proof. In addition, our proof handles nonsimple polygons with no vertices of turn angle 180°, establishing a new result and opening several new directions.

1. Introduction

Pocket flipping. Given a simple polygon in the plane, a *pocket* is a maximal connected region exterior to the polygon and interior to the convex hull. Provided the polygon is not convex, at least one such a pocket exists; see Figure 1(a). The boundary of such a pocket consists of one edge of the convex hull, called the *pocket lid*, and a subchain of the polygon, called the *pocket subchain. Flipping* a pocket transforms the polygon by reflecting the pocket subchain through the line extending the pocket lid, keeping the rest of the polygon fixed. Such an operation is called a *pocket flip* or simply a *flip*. Equivalently, if we view the polygon as a linkage where the edges are rigid bars and the vertices are universal joints, a pocket flip can be implemented by rotating the pocket subchain 180° around the axis through the pocket lid; see Figure 1(b). (Here the motion takes place in 3D, but after each flip the polygon remains planar.)

As long as the polygon remains nonconvex, we can pick an arbitrary pocket and flip it. Because the line extending the pocket lid is a line of support of the polygon (meeting the convex hull's boundary but not its interior), and the pocket subchain flips to the other side of that line, the resulting polygon always remains

2000 *Mathematics Subject Classification.* Primary 52C10; Secondary 52A10, 52-02.

Key words and phrases. pocket, convex hull, Erdős flip, limit, convexifying.

A preliminary version of this paper appeared at the *18th Canadian Conference on Computational Geometry* [**DGOT06**].

First author supported by NSF grant CCF-0347776, DOE grant DE-FG02-04ER25647, and AFOSR grant FA9550-07-1-0538.

Third author supported by NSF award DUE-0123154.

Fourth author supported by NSERC and FQNRT.

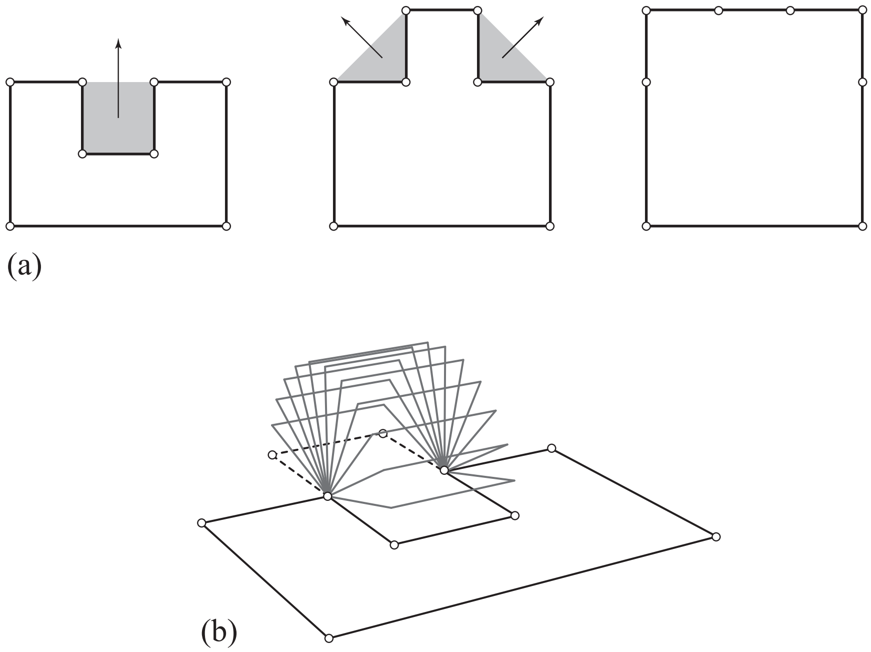


Figure 1. (a) Flipping the (shaded) pockets of a polygon until convexification. (b) First flip viewed as a 180° rotation.

simple. What happens if we repeat the pocket-flipping process? In the example of Figure 1(a), the polygon convexifies after a sequence of three flips. But could the process go on forever, flipping smaller and smaller pockets and only convexifying in the limit? This paper is about proving that every flipping sequence is finite:

Theorem 1.1. *Every simple polygon convexifies after finitely many flips, no matter which pocket is chosen at each step.*

History. This theorem has a surprisingly intricate history, with multiple independent discoveries of the problem and the theorem. This history has been largely already reported in the surveys by Grünbaum [**Grü95**] and Toussaint [**Tou99, Tou05**], each of which also provides their own proof of Theorem 1.1. The primary purpose of this paper is to uncover some new surprises in this intricate history: many of the purported proofs of the theorem, including these two latest, are in fact incorrect or incomplete.

The earliest known reference to pocket flipping is a problem proposal in the *American Mathematical Monthly* by Paul Erdős [**Erd35**]. For historical context, this publication was a year after Erdős obtained his Ph.D., left Hungary, and became a post-doctoral fellow at the University of Manchester; he was probably age 22. Erdős asked a slightly different question:[1] simultaneously flip all pockets of the

[1]His exact phrasing of the problem, which is the entirety of [**Erd35**], is as follows: "Given any simple polygon P which is not convex, draw the smallest convex polygon P' which contains P. This convex polygon P' will contain the area P and certain additional areas. Reflect each of these additional areas with respect to the corresponding added side, thus obtaining a new polygon P_1. If P_1 is not convex, repeat the process, obtaining a polygon P_2. Prove that after a finite number of such steps a polygon P_n will be obtained which is convex." [**Erd35**]

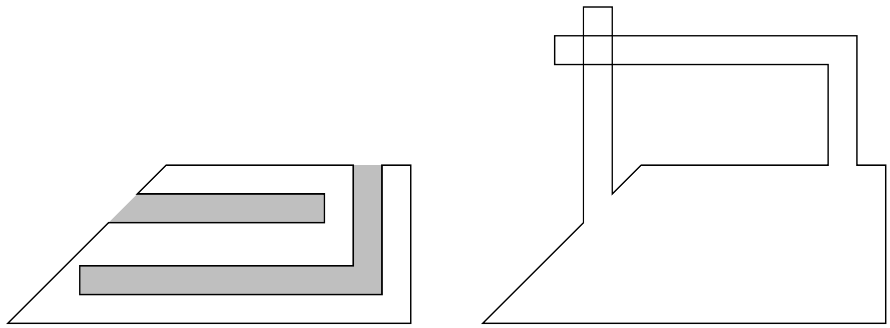


Figure 2. Flipping all pockets simultaneously can transform a simple polygon into a nonsimple polygon.

polygon, and repeat this process until the polygon is convex. The trouble with this version of the problem is that simultaneously flipping all pockets can make the polygon self-intersecting (nonsimple); see Figure 2. While the original notion of pocket flipping is ill-defined for nonsimple polygons, there is a natural generalization of pocket flipping in this scenario, which we will consider in Section 4. Interestingly, in this scenario, Erdős's original problem of simultaneous pocket flips remains unsolved.

Other than the original formulation [**Erd35**], the literature has considered the version where we flip a single pocket, recompute the convex hull and pockets, and repeat. The first known paper to consider this precise problem is a follow-up four years later in the *American Mathematical Monthly* by Béla de Sz.-Nagy [**dSN39**]. In one elegant page, Nagy points out the issue with simultaneous pocket flips, using an example similar to Figure 2, and then attempts to prove Theorem 1.1. Unfortunately, we found that the attempted proof makes a fatal mistake in one superficially believable sentence arguing that the limit of any potentially infinite flip sequence is convex; see Section 3.1 for the details. (Amusingly, the one incorrect sentence uses the word "obviously.") The proofs by Grünbaum [**Grü95**] and Toussaint [**Tou99, Tou05**] are both modifications/simplifications of Nagy's "proof." Interestingly, both replace the one incorrect sentence with a different argument, but we found that argument to be either unjustified (in the case of [**Grü95**]) or incorrect for a different reason (in the case of [**Tou99, Tou05**]).

The surveys of Grünbaum [**Grü95**] and Toussaint [**Tou99, Tou05**] uncover four other purported proofs of Theorem 1.1; refer to Table 1. Three proofs are in Russian. The first two, by Reshetnyak [**Res57**] and Yusupov [**Yus57**], seem to be discovered independently of each other and of Erdős and Nagy. Reshetnyak's proof seems to be the first correct proof of Theorem 1.1. In contrast, we show Yusupov's argument to be incorrect, in particular making one of the mistakes of Toussaint's argument. The next proof, by Bing and Kazarinoff [**KB59, BK61, Kaz61a**], was developed just two years later; the proof was also translated into English in Kazarinoff's *Analytic Inequalities* book [**Kaz61a**]. The latter is the only correct proof of Theorem 1.1 to appear in English (prior to the present paper). Interestingly, the Russian version [**BK61**] mentions Reshetnyak's proof as well as Nagy's work, stating in particular that Nagy's argument is incorrect. Unfortunately, they do not justify their claim, so it went relatively unnoticed, aside from a mention

Reference		Genesis	Flaws, omissions, comments
Nagy [**dSN39**]	§3.1	Erdős [**Erd35**]	**Flawed:** $C^k \not\subseteq P^{k+1}$.
Reshetnyak [**Res57**]	§3.2	independent[a]	**Correct** though somewhat imprecise.
Yusupov [**Yus57**]	§3.3	independent[a]	**Flawed:** P^* might have pockets, and only some vertices might flatten.
Bing & Kazarinoff [**KB59, BK61, Kaz61a**]	§3.4	Erdős [**Erd35**], Nagy [**dSN39**], Reshetnyak [**Res57**]	**Correct** though somewhat terse. Claims Nagy's proof is incorrect. False conjecture: $2n$ flips suffice.
Wegner [**Weg93**]	§3.6	Kaluza [**Kal81**]	**Flawed:** Area increase can be small.
Grünbaum [**Grü95**]	§3.7	all of above	**Omission:** Why P^* is convex. Based on Nagy's argument. Requires specific flip sequence.
Toussaint [**Tou99, Tou05**]	§3.8	all of above	**Flawed:** P^* might have pockets. Based on Nagy's argument.

Table 1. A chronological summary of the purported proofs of Theorem 1.1.
[a]These papers have no bibliography and make no explicit references to prior work.

by Grünbaum [**Grü95**]. The last proof is by Wegner [**Weg93**], who learned of the problem from an independent posing by Kaluza [**Kal81**]. This proof differs substantially from all other proofs, using several sophisticated mathematical tools; unfortunately, it too makes a critical mistake. In Section 3, we describe all of these arguments and discuss their weaknesses, gaps, and errors.

Our story. The genesis of this paper is a graduate class at MIT on "Folding and Unfolding in Computational Geometry," taught by the first author and taken by the second author. The fifth lecture, on Wednesday, September 22, 2004, presented Theorem 1.1 and the latest argument [**Tou05**]. The first author was a little hesitant about one step in the proof, and mentioned the worry; by the end of the presentation, one of the students found a reason why that step was indeed flawed. Over the weekend, the first, third, and fourth authors discussed the error, and decided to go back to the classic original argument by Nagy [**dSN39**], after whom the Erdős-Nagy Theorem is named. Nagy's argument differed precisely in the troublesome step, and seemed to offer a perfect replacement. What a beautiful argument, and from 1939 no less! So the first author presented it in the next class, on Monday, September 27, 2004. But by the end of the presentation, several students raised their hands and questioned this new step. Soon we had a counterexample. Suddenly the beautiful argument was also wrong—a 65-year old error! Two proofs dead in two weeks! Was the "Erdős-Nagy Theorem" even true? Fortunately, when the flaw in [**Tou05**] was found, the second author had found an idea for a fix, and after the flaw in [**dSN39**] was found, this idea was quickly solidified into a full proof. Its presentation in the next lecture, Wednesday, September 29, went without a hitch, and it appears in the conference version of this paper [**DGOT06**] and in a book by the first and third authors [**DO07**].

We then decided to review the various other purported proofs of the theorem, and found even more errors, but also discovered that our fixed proof was essentially identical to the proof of Bing and Kazarinoff [**KB59, BK61, Kaz61a**]; the latter was only more terse. So this proof feels quite "natural"—but it is not the only one, as we discovered the correct proof by Reshetnyak [**Res57**] which employs a different set of clever ideas. However, these seem to be the only two completely correct proofs in the literature, surprisingly few compared to the five flawed proofs.

Nonsimple polygons. More recently, we considered the generalized scenario in which the polygon is not necessarily simple. Here flips can be defined in terms of subchains instead of pockets. Specifically, consider any two distinct vertices of the polygon intersecting a line of support (either two overlapping vertices on the convex hull, or two distinct points on an edge of the convex hull). These vertices divide the polygon into two subchains. If neither of these chains lies collinearly along the line of support, a *pocket flip* corresponds to reflecting one of these subchains through the line of support. For simple polygons, this definition is equivalent to standard pocket flips. But does every nonsimple polygon convexify after a finite number of flips, no matter how those flips are chosen? Figure 3 shows an example.

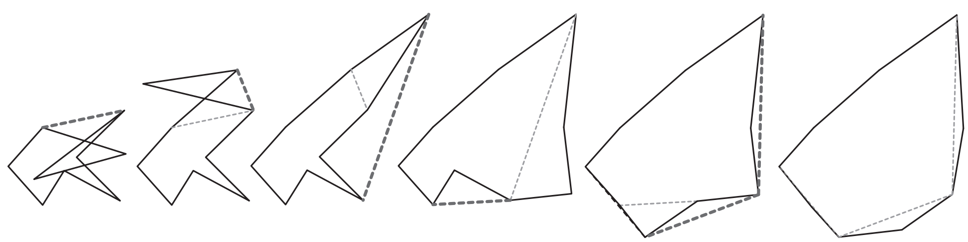


Figure 3. Flipping a nonsimple polygon to convexity. Lids are dashed, dark before the flip and light afterward.

Tantalizingly, the Russian paper by Bing and Kazarinoff [**BK61**] claims, without proof: "Observation. The polygon need not be a simple polygon." However, this claim remains unproved in the literature, and indeed we show it to be false under its most natural interpretation. On the positive side, Wegner [**Weg00**], Grünbaum and Zaks [**GZ98**], and Toussaint [**Tou99, Tou05**] all proved a slightly weaker theorem: there exists a finite sequence of flips that convexifies a given nonsimple polygon. (Here the notion of "convexification" is somewhat weaker than usual; see Section 4.2.) Toussaint's flip sequence is faster to compute than Grünbaum and Zaks's—$\Theta(n)$ time per operation instead of $\Omega(n^2)$—but what about arbitrary flip sequences as in Erdős's original problem? This question was partially resolved in an unpublished proof [**BCC$^+$01**] which uses Theorem 1.1 as a black box and argues that flipping a nonsimple polygon monotonically decreases the number of crossings until some time, after which the polygon behaves essentially the same as a simple polygon. However, this proof handles only "proper crossings" at single points in the relative interiors of edges; it does not handle arbitrary polygons, e.g., with edges that overlap along a segment.

We show that the finiteness of arbitrary flip sequences depends on the existence of *hairpin* vertices: vertices whose two incident edges overlap, forming a turn angle of 180°. When there are hairpin vertices, we show in Section 4.4 that a poorly chosen flip sequence can go on forever in the worst case. Without hairpin vertices, however, we obtain the desired positive result in Section 4.3:

THEOREM 1.2. *Every polygon, not necessarily simple but having no hairpin vertices, convexifies after finitely many flips, no matter which flip is chosen at each step.*

Our proof of this theorem also serves as a new, self-contained, and fully detailed proof of Theorem 1.1.

On errors in geometry. Given the plethora of errors concerning the geometric problem considered in this paper, the question naturally arises whether this is an isolated event. So it seems appropriate to add a word about the frequency of mistakes in geometric research in general, and its role in the discovery process. Indeed, psychologists, mathematicians, and physicists have been interested for over a hundred years in the nature of discovery, in mathematics in general [**Had45**], and geometry in particular [**Ein83, Lak76, Mac06, Pap80, Tou93, ST03**]. For example, Lakatos [**Lak76**] traces nearly two hundred years of history concerning the errors made in the evolution of the definition of polyhedra. More recently, Toussaint [**Tou93**] traces nearly two thousand years of history concerning errors made in proving Euclid's second proposition, and then argues that Euclid's original proof had no errors at all.

The literature suggests that there is a greater frequency of errors in geometry than in other fields of mathematics. One reason for these errors, including the errors on the polygon-flipping problem considered here, seems to boil down to this fact: geometry often deals with a space of infinite size where there are an infinite number of cases (sometimes unwittingly created by the mathematician) that must be reduced by the human mind into a finite number of cases (or none at all), with the aid of our visual system which is fraught with its own visual illusion traps. This situation appears to be different from many other branches of mathematics, such as combinatorics or algebra, that do not have a strong visual component and that often need to consider only a finite number of cases.

For a concrete example of this issue, consider the original pocket-flipping problem formulation posed by Erdős. Polygons with their pockets have an infinite number of possible shapes. Erdős probably tried some examples to conclude that, in all infinitely many cases, simultaneously flipping several pockets yields a new simple polygon (so that we may "repeat the process"). Nagy clearly explored more "unusual" shapes in this infinite space, leading him to a counterexample along the lines of Figure 2.

Geometric infinite-case analyses are exacerbated by the visual thinking component inherent in geometry, which has a tendency to derail logical thinking. In [**Pap80**] and [**ST03**], the authors distinguish between two types of thinking: *logico-mathematical* and *kinesthetic*. The latter type is also referred to as *body-syntonic*, a term used by Papert [**Pap80**] for a similar notion, and makes use of kinesthetic heuristics. A *kinesthetic heuristic* is where cognition, understanding, and learning take place through perceptible results of dynamic manipulation of objects to support useful insights on the problem being studied. Kinesthetic heuristics emphasize the experimental aspects in front of logico-mathematical deductions as a primary tool of understanding. More powerful kinesthetic heuristics tend to create stronger geometric illusions and therefore call for extra care in logical thinking to determine whether the resulting proofs of a theorem are correct. Flipping a pocket of a simple polygon is a perfect example of such a kinesthetic heuristic.

For a more detailed discussion of this topic, see [**Tou93, ST03**].

2. Notation

We begin the technical part of this paper by introducing some notation used throughout our descriptions of old and new arguments about pocket flipping.

Let $P = P^0 = \langle v_0, v_1, \ldots, v_{n-1}\rangle$ denote the initial polygon and its vertices in order, forming the n edges $v_0v_1, v_1v_2, \ldots, v_{n-2}v_{n-1}, v_{n-1}v_0$. Let $P^k = \langle v_0^k, v_1^k, \ldots, v_{n-1}^k\rangle$ denote the resulting "descendant" polygon after k arbitrary pocket flips; if P^k is convex for some k, then we define $P^k = P^{k+1} = P^{k+2} = \cdots$. Let C^k denote the convex hull of P^k. When we talk about convergence, it is always with respect to $k \to \infty$. When the limit of P^k exists, we denote it by P^*, its vertices by v_i^*, etc.

The *(directed) turn angle* $\tau_i^k \in (-180°, 180°]$ at vertex v_i^k is the signed angle between the two vectors before and after v_i^k: the angle that turns $v_i^k - v_{i-1}^k$ to $v_{i+1}^k - v_i^k$. Call a vertex v_i^k *flat* if its turn angle τ_i^k is 0, and *pointed* otherwise. If the absolute value of the turn angle τ_i^k has a limit, then we call vertex v_i *asymptotically flat* if the limit angle is zero and *asymptotically pointed* otherwise.

We use $\|x - y\|$ to denote the Euclidean distance between two points x and y, or equivalently, the Euclidean length of the vector $x - y$,

3. Previous Arguments

Next we turn to the purported proofs of Theorem 1.1, that simple polygons always flip finitely, as summarized in Table 1. We describe the arguments in chronological order along with any errors or omissions they make.

3.1. Nagy. The first claimed proof of Theorem 1.1, published by Béla de Sz.-Nagy in 1939 [**dSN39**], is brilliant in overall design, but unfortunately has a fatal flaw that has gone largely undetected until now. Previously, the Russian paper by Bing and Kazarinoff [**BK61**] (but not the English book [**Kaz61a**] or talk abstract [**KB59**]) remarked that "The proof of this theorem, given by B. Sz. Nagy, is incorrect." Grünbaum [**Grü95**] noticed this claim, but pointed out that "there is no basis for this claim." Whether Bing and Kazarinoff found the (same) flaw is unclear, but at the least, Nagy's argument was widely believed until now.

Nagy's argument consists of four main steps:

(1) The sequence P^k converges to a limit P^*.
(2) The limit P^* is convex.
(3) Asymptotically pointed vertices converge in finite time.
(4) The sequence P^k converges in finite time.

The flaw is in Step 2, where Nagy claims that $P^0 \subseteq C^0 \subseteq P^1 \subseteq C^1 \subseteq \cdots$.[2] This claim implies that P^k and C^k converge to the same, necessarily convex limit. As illustrated in Figure 4, however, the claim is incorrect. When there are multiple pockets to choose from, $C^k \not\subseteq P^{k+1}$.

Despite most later arguments being based on Nagy's, this flaw seems unique to Nagy's argument. Many later arguments use the other, correct steps of Nagy's argument, to which we now turn.

In Step 1, Nagy observes that the perimeter of P^k is constant. Because $P^0 \subseteq P^1 \subseteq P^2 \subseteq \cdots$ (a true half of the false claim), the fixed perimeter bounds the

[2] Specifically, he states: "Each polygon in the sequence P^0, C^0, P^1, C^1, P^2, C^2, $\cdots$ contains obviously the foregoing ones in its interior." [**dSN39**].

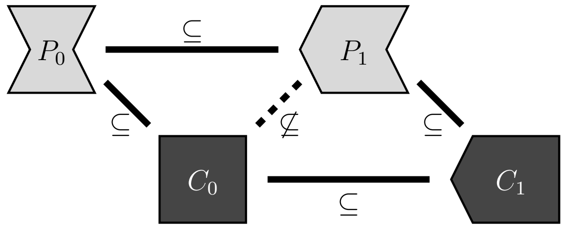


Figure 4. Nagy's error: $P^0 \subseteq C^0 \not\subseteq P^1 \subseteq C^1$.

possible reach of vertices in P^k. Thus Nagy concludes that each v_i^k has a point of accumulation. Then he observes that, for any point x inside or on the boundary of P^k, $\|x - v_i^m\| \leq \|x - v_i^{m+1}\|$ for all $m \geq k$. This nondecreasing-distance property holds because the line extending the pocket lid of this flip is the Voronoi diagram of v_i^m and v_i^{m+1}, and because the Voronoi cell of v_i^m contains P and thus x; however, Nagy simply asserts the property. Nagy applies the property to conclude that $\|v_i^m - v_i^p\| \leq \|v_i^m - v_i^{p+1}\|$ for all $p \geq m$, which prevents the existence of multiple points of accumulation, thus proving convergence.

To prove Step 3, Nagy uses an argument illustrated in Figure 5 to show that asymptotically pointed vertices of the limit polygon converge in finite time. This argument is easy to illustrate, but requires some care to justify in detail, while Nagy's presentation is terse.

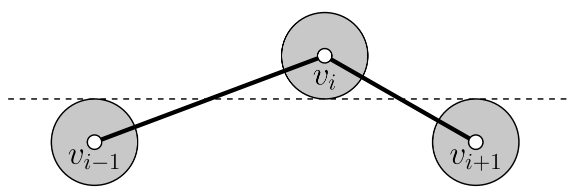


Figure 5. For an asymptotically pointed vertex v_i^k, once all the vertices are within a small enough disk around their limit, there is a line that separates the disk of v_i^k from all the other disks. Thus v_i^k subsequently remains on the convex hull of P^k and cannot be flipped again.

Finally, in Step 4, once all the asymptotically pointed vertices have converged, no more flips are possible, because they would cause the convex hull to increase beyond its limit, namely, the convex hull of the converged locations of the asymptotically pointed vertices.

3.2. Reshetnyak. In 1957,[3] a paper in Russian by Reshetnyak [**Res57**] rediscovers the problem and gives a proof that is somewhat hand-wavy, but quite original and correct. It consists of two main steps:

(1) Vertices that move an infinite number of times are asymptotically flat.
(2) Once the asymptotically pointed vertices have converged, the asymptotically flat vertices cannot move.

[3]The publication states that the paper was "Submitted to the editors 24 January 1956."

Thus, Reshetnyak's Steps 1 and 2 essentially establish Nagy's Steps 3 and 4 directly, without establishing that the polygons actually have a limit, and thus bypassing Nagy's Step 1 and troublesome Step 2.

Step 1 considers a vertex v_j that moves infinitely many times, and uses Nagy's "constant perimeter" argument to extract an infinite sequence $n_1, n_2, \ldots$ such that $v_{j-1}^{n_m}$, $v_j^{n_m}$, $v_{j+1}^{n_m}$, $v_{j-1}^{n_m+1}$, $v_j^{n_m+1}$, and $v_{j+1}^{n_m+1}$ all converge as $m \to \infty$, and such that $v_j^{n_m} \neq v_j^{n_m+1}$ for all m. The latter property implies that $v_j^{n_m}$ and $v_j^{n_m+1}$ are reflections of each other through a line of support, call it L^{n_m}, and similarly for vertices v_{j-1} and v_{j+1}. A further subsequence extraction leads to a sequence $v_j^{n'_m}$ for which $L^{n'_m}$ converges to a limit L^*. The limit of $v_j^{n'_m}$ must lie on L^* because of the reflection property, and likewise for v_{j-1} and v_{j+1}, so the turn angle $\tau_j^{n'_m}$ must converge to either 0 or 180°. By Nagy's "nondecreasing distances" argument, the distance between any two polygon vertices is nondecreasing, so the absolute turn angles are nonincreasing, and therefore the sequence of turn angles $\tau_j^{n'_m}$ must converge to 0.

Step 2 can be seen as a more involved version of our Lemma 4.2 described in Section 4.3. Reshetnyak considers an arbitrary line L of support of P^{k_0} at the time k_0 at which all asymptotically pointed vertices have converged. He shows that, for all $k \geq k_0$, no vertex of P^k can lie on the opposite side of L from P^{k_0}. Specifically, if some vertex were to go some distance $h > 0$ on the other side of L, then there would always be a vertex of distance at least h on that side of L. But by Step 1, for large enough k, the turn angles of the vertices on that side of L get to within ε of flat. This straightening leads to an upper bound of εpn on h, where p is the perimeter of P and n is the number of vertices. Because this upper bound holds for all $\varepsilon > 0$, h must in fact be zero, a contradiction.

Reshetnyak's argument in Step 2 is a bit of a "proof by picture," and therefore lacks precision, but in the end it is correct. As such, Reshetnyak's proof is likely the first correct proof of Theorem 1.1.

Reshetnyak's paper has no references. Indeed, the author states that "No other proof of this theorem is known to us." The only indication of where the problem might have arisen is that the "result would help in solving some extremal problems." This statement is motivated later in the literature by Bing and Kazarinoff; see Section 3.4 below.

3.3. Yusupov. Perhaps coincidentally, another 1957 paper in Russian, by Yusupov [**Yus57**], rediscovers the problem and claims a proof of Theorem 1.1. Yusupov's argument follows roughly the same outline as Nagy's argument, although most steps are argued differently, so the similarity seems to be coincidence. Unfortunately, Yusupov's argument has two distinct flaws: a different flaw from Nagy's in Step 2, and a new flaw in Step 3. To our knowledge, neither flaw was noticed previously.

For Step 1, Yusupov argues that the polygons have a limit if they are treated as sets. Here he uses that the sets are "monotonically nondecreasing" ($P^0 \subseteq P^1 \subseteq P^2 \subseteq \cdots$), and that the sets are bounded by the usual perimeter argument. This argument does not imply that any particular vertex converges, but this stronger claim seems unnecessary (though implicitly assumed) in the rest of Yusupov's argument.

For Step 2, Yusupov simply claims that "The limit polygon is convex, because, otherwise, a part of it would admit a . . . [pocket flip] and the polygon would not be the limit." In other words, any limit polygon must have all possible flips already made. This reasoning is incorrect in general. For some intuition why, imagine that there are two portions of the polygon that each can flip infinitely often (hypothetically, of course). If we choose an infinite flip sequence that visits pockets in just one of those portions, then the other portion never gets flipped, so the resulting limit polygon is nonconvex. There are specific flip sequences that would provably avoid this problem (for example, visiting the pockets in round-robin order; see also Grünbaum's proof in Section 3.7), but this would necessarily weaken the claim to convexification after finitely many properly chosen flips.

For Step 3, Yusupov uses the lemma that, if the shortest edge length is $\ell_{\min}$, then a flip increases the area of the polygon by at least $\frac{1}{2}\ell_{\min}^2 \sin\tau$ where τ is the turn angle of "one of the flipped vertices." Although unjustified in the paper, this lemma is true: triangulate the pocket, take a triangle ("ear") with two polygon edges as sides, use the $\frac{1}{2}\ell_1\ell_2 \sin\theta$ area formula for a triangle with side lengths ℓ_1 and ℓ_2 and incident angle θ, and observe that $\sin\theta = \sin(180° - \tau) = \sin\tau$. By the polygon limit argument of Step 1, the polygon area is bounded, so the area increase per flip must converge to 0. Also, Yusupov observes that absolute turn angles only decrease by flips, as they change only when a vertex is an endpoint of the pocket lid. Yusupov then claims that the turn angles of vertices flipped infinitely often must therefore converge to 0. This argument is flawed: all we know is that some turn angle in each infinitely flipped subchain must converge to 0 (corresponding to an ear), but we have little control over which turn angle.[4]

If we assume Steps 2 and 3, then Step 4 is not difficult, as in Nagy's argument. At least three vertices stop moving after finitely many flips, because the limit polygon cannot be flat (as it contains P^0), so it must have at least three nonflat vertices. (Like the general form of Step 3, it seems difficult to deduce just these three converging vertices directly from the lemma above.) Then the infinitely flipped vertices between these converging vertices must be asymptotically flat (again by Step 3), but then they must actually be flat once the converging vertices converge, just to reach the desired distance. Thus ends Yusupov's argument.

Like Reshetnyak, Yusupov's paper has no references, and the two contemporaries seem unaware of each other. Yusupov's paper also lacks any context for the problem, jumping directly into definition, theorem, and attempted proof.

3.4. Bing and Kazarinoff. At the AMS Annual Meeting in 1960, Kazarinoff and Bing [**KB59**] presented the pocket-flipping problem and a solution. In 1961, full proofs appear in a paper by Bing and Kazarinoff [**BK61**] and also in Kazarinoff's *Analytic Inequalities* book [**Kaz61a**].

Bing and Kazarinoff's proof, as described in both [**BK61**] and [**Kaz61a**], has no missing steps, and suffers only from being terse. In fact, the journal editor of [**BK61**], presumably having found the proof too terse, added some explanatory footnotes, without which the proof would indeed be rather incomplete. The proof consists of three main steps:

(1) The sequence P^k converges to a limit P^*.
(2) Pointed vertices of the convex hull of P^* converge in finite time.

[4] Another possible interpretation of the lemma is that τ measures one of the turn angles at an endpoint of the pocket lid. This alternative makes the use of the lemma more plausible, but makes the lemma false, by examples like those in Figure 7.

(3) The sequence P^k converges in finite time. (Same idea as Nagy.)

For Step 1, Bing and Kazarinoff use Nagy's "constant perimeter" and "nondecreasing distances" arguments to conclude that, for x interior to C^0, the sequence $\|x - v_i^k\|$ is bounded and nondecreasing, and thus it converges. Applying this argument for three noncollinear points x_1, x_2, and x_3 shows that each v_i^k converges to the unique intersection of three circles.

In Step 2, they argue that, because P^k converges, the interior angles of its vertices must also converge. Thus, any vertex that converges to a pointed vertex in the convex hull of P^* has an interior angle less than $180°$ after a finite number of steps. Because a vertex moves only when it is flipped, and a flip changes an interior angle α into the angle $360° - \alpha$, the vertex can no longer move.

As mentioned by Grünbaum [**Grü95**], "In all three publications by Bing and Kazarinoff it is conjectured that the convexification of every polygon with n sides is achieved after at most $2n$ flips."[5] Kazarinoff's book [**Kaz61a**] goes on to ask: "Can you prove or disprove this conjecture? Paul Erdős did." Grünbaum [**Grü95**] writes: "I am not aware of the reason for this statement, and I do not know what Erdős did in this context; there appears to be no further mention of the convexification question in Erdős' writings after [**Erd35**]." In any case, the conjecture has since been shown to be false, by Joss and Shannon; see Section 3.5.

Bing and Kazarinoff end their Russian paper [**BK61**] with two revelations: (1) "Observation: the polygon need not be a simple polygon." (2) "The proof of this theorem, given by B. Sz. Nagy, is incorrect." They do not provide further explanation for either claim. Now we know that Nagy's "proof" is indeed incorrect, but it remains a mystery what exactly Bing and Kazarinoff had in mind. Bing and Kazarinoff were also aware of Reshetnyak's proof, stating in an early footnote that "A somewhat different proof of the same theorem is presented by Yu. G. Reshetnyak" (including a full reference to [**Res57**]). As for nonsimple polygons, we will see in Section 4 that so far only less-general forms of this claim have been proved [**GZ01, Tou05, Weg00, BCC$^+$01**], and that the most general interpretation of this claim is in fact false. In particular, Bing and Kazarinoff's proof fails in Step 1 for nonsimple polygons, where it is no longer guaranteed that $P^k \subseteq P^{k+1}$, so we cannot easily choose three noncollinear points that remain interior to the polygon.

Another book by Kazarinoff, *Geometric Inequalities* [**Kaz61b**], mentions Theorem 1.1 without proof. What makes this reference interesting is that Kazarinoff uses Theorem 1.1 in the proof of another theorem, on isoperimetry: given any n-gon whose edge lengths are not all equal, one can construct another n-gon with the same perimeter, with all sides of equal length, and with larger area. His proof of this theorem first convexifies the polygon using pocket flips, mentioning Theorem 1.1 without proof. Then he re-orders the edge-length sequence by reflecting a two-edge chain v_{i-1}, v_i, v_{i+1} through the perpendicular bisector of v_{i-1} and v_{i+1}. This reflection operation enables him to bring two different-length edges next to each other

[5] The situation is actually a little more subtle than this: the talk abstract [**KB59**] phrases the conjecture as follows: "For fixed n, N [the number of flips] is bounded by at least $2n$ for all P and all choices of r_m's [the flips]." We assume that the phrase "at most" was intended instead of "at least," given the phrasings in [**BK61, Kaz61a**], but it is plausible that something more subtle was intended.

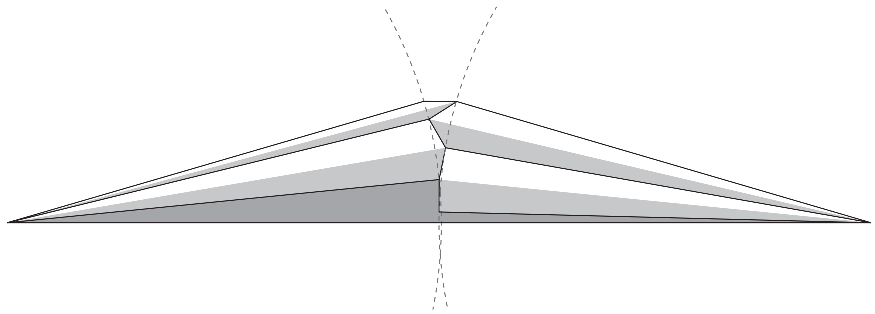


Figure 6. Quadrangles similar to the one shown at the bottom can require arbitrarily many flips to convexify.

and perform a local improvement. The initial convexification is necessary to ensure that the reflection operation preserves simplicity of the polygon.[6]

This application seems to be the motivation for Bing and Kazarinoff's statement in [**BK61**] that the "theorem allows us to reduce a host of extremal problems for simple polygons to simpler problems on convex polygons," as well as Reshetnyak's similar earlier statement.

3.5. Joss and Shannon. In 1973, two students of Grünbaum at the University of Washington, R. R. Joss and R. W. Shannon, worked on pocket flipping but did not publish their results. Grünbaum [**Grü95**] gives an account of the unfortunate circumstances surrounding this event. They found a counterexample to the conjecture of Bing and Kazarinoff, though they were unaware of the conjecture. Specifically, they showed that, given any positive integer k, there exist simple polygons of constant size (even quadrilaterals) that cannot be convexified with fewer than k flips. Figure 6 shows their counterexample. See [**Grü95, Tou05**] for more details about their work and their story.

3.6. Wegner. In 1981, Kaluza [**Kal81**], apparently unaware of the previous work, posed the pocket-flipping problem again and asked whether the number of flips could be bounded as a function of the number of polygon vertices. In 1993, Bernd Wegner [**Weg93**] took up Kaluza's challenge and claims solutions to both problems again. His proof of convexification in a finite number of flips is quite different from the others, but his example of unboundedness is the same as that of Joss and Shannon. More recently, Wegner has extended his work to polygons on the sphere [**Weg96a, Weg96b, Weg96c**] and to more general curves [**Weg96c, Weg99, Weg00**].

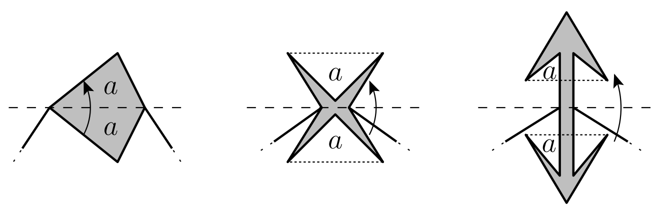
Wegner's argument [**Weg93**] is certainly the most intricate of the proofs we have seen. His argument is very technical—for example, using convergence results

[6] In fact, Kazarinoff first proposes a modified reflection operation that works as described at convex vertices v_i, while for reflex vertices v_i, it first reflects the chain through the line $v_{i-1}v_{i+1}$ and then reflects through the perpendicular bisector. Interestingly, this operation is essentially a simple form of "flipturns," previously thought to have been invented by Joss and Shannon [**Grü95**]. Then Kazarinoff notes that these reflection operations can lead to self-intersection in a nonconvex polygon, so he introduces pocket flipping for initial convexification.

from the theory of convex bodies—and difficult to summarize. To his credit, Wegner carefully details his reasoning, unlike many other authors. Unfortunately, his argument is flawed.

Wegner's approach most closely resembles Yusupov's. For example, both observe that the absolute turn angles are monotonically nondecreasing. The exclusive use of absolute turn angles makes Wegner's argument stand out, but prevents the use of angles to show convergence in finite time as in Bing and Kazarinoff's proof.

Instead, Wegner introduces the area A^k of P^k and tries to show that, after a finite number of flips, performing an additional flip would cause A^k to exceed the area A^* of the limit polygon P^*. He lower-bounds the increase in area during a flip that moves any vertex v_i^k by considering the area a of the triangle $v_{i-1}^k v_i^k v_{i+1}^k$. Wegner argues that, during such a flip, A^k will increase by at least $2a$, and uses this fact to force A^k beyond its limit. Figure 7(a) shows the presumed motivation for this claim, a simple reflex vertex v_i^k. Unfortunately, Figure 7(b–c) shows examples of convex and reflex vertices where the area A^k increases by an arbitrarily small amount relative to a.[7] To our knowledge, this flaw was not noticed previously.



(a) (b) (c)

Figure 7. Flipping a simple reflex vertex increases the polygon area by twice the area a of the incident triangle (a), but this property is not true of a convex vertex (b), nor for a more complicated reflex vertex (c).

3.7. Grünbaum. Branko Grünbaum [**Grü95**] described some of the intricate history of this problem following the appearance of Nagy's paper [**dSN39**], uncovering the aforementioned rediscoveries of the theorem. He also provided his own argument, "essentially the one" by Nagy, but somewhat more terse. One main difference is that, at each step, he flips the pocket that has maximum area (if there is more than one pocket to choose from). Therefore Grünbaum [**Grü95**] actually argues a weaker theorem: there exists a (well-chosen) sequence of flips that convexifies after finitely many flips.

Grünbaum's argument has a similar four-step structure to Nagy's:

(1) A subsequence of the sequence P^k converges to a convex limit.
(2) The whole sequence converges.
(3) Asymptotically pointed vertices converge in finite time. (Same as Nagy.)
(4) The sequence converges in finite time. (Same as Nagy.)

For Step 1, Grünbaum invokes Nagy's "constant perimeter" argument to show that a subsequence converges. He then claims that "Due to the choices of the

[7] We originally thought [**DGOT06**] that this flaw arose only with convex vertices, which would be easy to fix, because a convex vertex becomes reflex after one flip, so the next time it moves, Wegner's argument would indeed apply. Then we found Figure 7(c).

exposed pairs [pocket lids] as maximizing the area, the polygon P^* is convex," without further explanation. We view this unjustified claim as a gap in the proof, because the convexity of P^* has been a stumbling block in most claimed proofs of the theorem (particularly Nagy's and Toussaint's).

In Step 2, Grünbaum invokes a nondecreasing-distance argument: "since each flip either increases or leaves unchanged the distance from a vertex to any point inside the polygon, it follows that P^* is, in fact, the limit of the complete sequence of polygons P_i, without the need to select a convergent subsequence." We view this argument as also requiring more justification: either it requires the selection of three noncollinear interior points from which to measure distances, as in Bing and Kazarinoff's proof, or it requires concluding the nondecrease of $\|v_i^m - v_i^p\|$, as in Nagy's proof.

3.8. Toussaint. Motivated by the desire to present a simple, clear, elementary, and pedagogical proof of such a beautiful theorem, Toussaint [**Tou99**] presented a more detailed and readable argument in 1999. He combined Bing and Kazarinoff's approach to proving the convergence of P^k with Nagy's approach of proving that convergence occurs in finite time.

The original argument that appeared in [**Tou99**] uses one instead of three noncollinear points x_1, x_2, and x_3 to conclude that the vertices v_i^k converge. However, without further justification, it is possible that v_i^k circles around x and thus has multiple accumulation points. Because Toussaint's argument is explicit in the details, this is a clear oversight. This led the first and third authors of this paper to point out the problem, and propose the three-point solution (at the time, unaware of Bing and Kazarinoff's proof). This correction appeared in the journal version of Toussaint's argument [**Tou05**].

Unfortunately, both arguments [**Tou99, Tou05**] make an invalid deduction for establishing the convexity of the limit polygon P^* in general: "we note that the limit polygon ... must be convex, for otherwise, being a simple polygon, another flip would alter its shape contradicting that it is the limit polygon." This flaw is coincidentally the same as one of the errors made by Yusupov. (At the time, however, Toussaint was unaware of Yusupov's proof.) As described in Section 3.3, this flaw is easy to repair for specific flip sequences, but this would weaken the claim.

4. New Proof, for Nonsimple Polygons

We now turn to present a new proof that polygons always convexify after finitely many flips. Our initial motivation for such a proof was to have one fully detailed, correct, and accessible proof of Theorem 1.1, whereas previously the only correct proof in English was [**Kaz61a**]. This motivation led to the proof in the conference version of this paper [**DGOT06**], which was essentially a somewhat more detailed version of the proof by Bing and Kazarinoff [**BK61, Kaz61a**]. Instead of presenting this proof, however, we present a new proof that establishes a stronger theorem, allowing certain kinds of nonsimple (self-crossing) polygons. To our knowledge, our Theorem 1.2 is a new result, which has only been approached up to now.

4.1. Previous Results for Nonsimple Polygons. Wegner [**Weg00**], Grünbaum and Zaks [**GZ01**], and Toussaint [**Tou05**] each show that, for any polygon (possibly nonsimple), there exists a particular way of selecting flips so that the polygon convexifies in a finite number of steps.

Wegner [**Weg00**] deals with the case of nonsimple polygons but concentrates on proving the result for curves more general than polygons. For the case of polygons [**Weg00**, Section 3], he provides only a rough outline. He assumes throughout that all self-intersections are transversal, i.e., that there are only a finite number of intersection points. But he adds that the general case of "self-intersections with equal tangents" can be handled also, but "the arguments are quite lengthy" and omitted. He states in the section on curves that it is obvious that the flips must be chosen in a special way. Otherwise, flipping with arbitrary choices, one may enter an infinite sequence of iterations never ending at a simple closed curve. However, his arguments do not appear to apply to polygons. Finally, he generalizes the results to polygons and curves embedded on other surfaces such as the sphere.

Grünbaum and Zaks [**GZ01**] handle arbitrary nonsimple polygons, putting great care into precise definitions of pocket flips and weakened forms of convexity that are required when vertices lie on nonincident edges. They introduce a potential function, the sum of pairwise distances between polygon vertices, and at each step choose the flip that maximizes the increase in this potential. Then they prove finiteness of this flipping process with a new argument. Computing the potential function requires $\Theta(n^2)$ time, so determining the next flip to perform requires $\Omega(n^2)$ time.

Toussaint [**Tou99, Tou05**] handles arbitrary nonsimple polygons as well. His construction is a clean inductive argument that uses Theorem 1.1 purely as a black box, applied to a specifically constructed flip sequence. The construction is also easier to compute, requiring only $\Theta(n)$ time to determine the next flip (the same time required to execute a flip).

Another, unpublished approach to the nonsimple case is by Biedl et al. [**BCC$^+$01**]. In this inductive proof, the nonsimple case with a finite number of crossings is reduced to the simple case by a black-box reduction. The first observation is that the number of crossing pairs of edges never increases by flips. If all the crossings get removed, we are done. Otherwise, the polygon can be transformed by replacing each crossing with a small construction as in Figure 8. This transformation does not change the effect of flips that are applied to the polygon (because we assume the crossing is never removed), so by induction the transformed polygon convexifies in a finite number of steps. This argument applies to any flip sequence, but it only allows proper crossings that occur away from vertices.

In contrast to these results, our proof works for any flip sequence and for any polygon that has no hairpin vertices, i.e., no turn angle of 180°. The hairpin limitation is necessary in the worst case: some polygons with hairpin vertices can flip an infinite number of times, as we will see in Section 4.4.

4.2. Definitions. Before proceeding to our proof, we define the generalized notions of pocket flips and convexity.

As in Section 2, a polygon is defined by a finite circular sequence of vertices $v_0, v_1, \ldots, v_{n-1}$. Such a polygon may now self-intersect in a variety of ways: two nonadjacent edges might properly cross, a vertex could lie on the relative interior of an edge, two nonadjacent vertices might coincide, and/or two edges could overlap

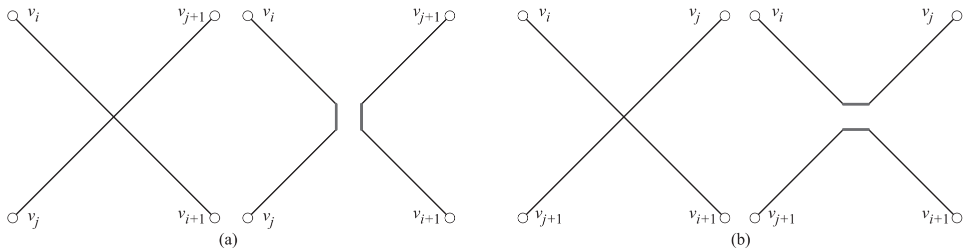


Figure 8. Replacing crossing edges with noncrossing edges.

collinearly. We may, however, assume without loss of generality that no two vertices adjacent in the sequence are identical, i.e., the polygon has no zero-length edges. Such edges can be added or removed without affecting pocket flips.

A *pocket lid* of a polygon P is a triple (v_i, v_j, L), where v_i and v_j are vertices of P, $i \neq j$, and L is a line through v_i and v_j that is a line of support for P. Moreover, v_i and v_j split P into two subchains, $v_i, v_{i+1}, \ldots, v_j$ and $v_j, v_{j+1}, \ldots, v_i$, neither of which may be entirely contained in the line L. We permit v_i and v_j to be overlapping. In this degenerate case, there may be an infinite family of pocket lids through the same two vertices v_i and v_j, as illustrated in Figure 9. All of these possibilities need to be considered in our proof of Theorem 1.2.

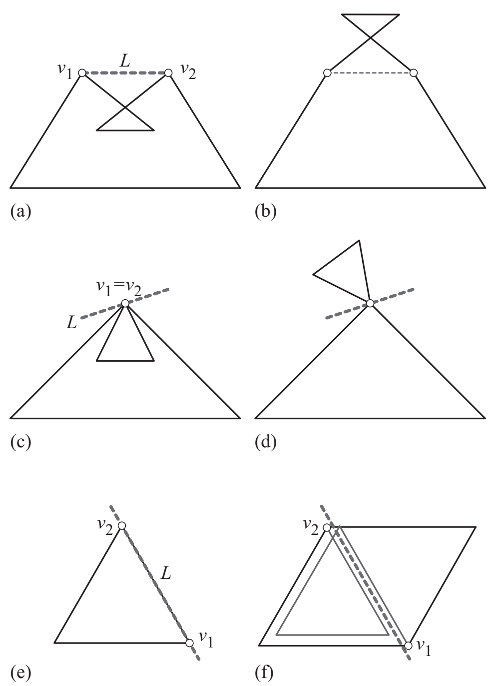


Figure 9. (a, b) Pocket lid and flip. (c) A degenerate pocket lid with $v_1 = v_2$. (d) After one of the possible flips. (e) A twice-cycled triangle forming a hexagon. (f) Details of flip of one subchain delimited by v_1 and v_2, with overlapping edges separated for clarity.

A *pocket flip* of a polygon P around pocket lid (v_i, v_j, L) is the polygon obtained by reflecting one of the subchains of P defined by v_i and v_j through line L. The requirement that each subchain not be contained in L implies that a pocket flip is never a trivial motion (doing nothing or globally reflecting the whole polygon).

As in Theorem 1.1, a polygon P is *convex* when it is simple and contained in the boundary of the convex hull of P. (In Grünbaum and Zaks's terminology [**GZ01**], this notion of convexity is called "weakly convex" because it allows flat vertices.)

Grünbaum and Zaks's results [**GZ01**] use an even weaker notion of convexity, called *exposed*, in which the polygon edges are contained in the boundary of the convex hull of P and each vertex of the convex hull contains only one vertex of P. We can avoid this weaker notion in the case of no hairpin vertices, but we will consider it when discussing possible generalizations.

4.3. Proof. Our proof of Theorem 1.2 divides into two major parts:

PART 1. Every polygon, not necessarily simple but having no hairpin vertices, admits only finitely many flips, no matter which flip is chosen at each step.

PART 2. Any polygon P having no hairpin vertices and no valid flips is convex.

We start with two lemmas that will be used in the proof of Part 1. The first lemma gives us a tool to compare the relative orientation of two turn angles using just inter-vertex distances. A turn angle has the same sign as the signed area of the triangle formed by the vertex and its two neighbors (positive if oriented counterclockwise, negative if oriented clockwise). Inconveniently, the sign of this area cannot be computed using just the inter-vertex distances in the triangle, as these distances are the same for the triangle and its mirror image. Fortunately, given two triangles and given the inter-vertex distances between all six vertices, we can calculate the relative signs of the signed areas (and hence the turn angles) of the two triangles.

LEMMA 4.1. *Given two triangles ABC and DEF, the product of their signed areas is a continuous function of just the inter-vertex distances.*

Proof: First, dot products of vectors connecting vertices are continuous functions of the inter-vertex distances: for arbitrary A, B, C, D,

$$\begin{aligned}
& 2\,(A-B)\cdot(C-D) \\
= \;& 2\,A\cdot C - 2\,B\cdot C - 2\,A\cdot D + 2\,B\cdot D \\
= \;& -\;(A\cdot A - 2\,A\cdot C + C\cdot C) \;+\; (B\cdot B - 2\,B\cdot C + C\cdot C) \\
& +\;(A\cdot A - 2\,A\cdot D + D\cdot D) \;-\; (B\cdot B - 2\,B\cdot D + D\cdot D) \\
= \;& -\;(A-C)\cdot(A-C) \;+\; (B-C)\cdot(B-C) \\
& +\;(A-D)\cdot(A-D) \;-\; (B-D)\cdot(B-D) \\
= \;& -\|A-C\|^2 + \|B-C\|^2 + \|A-D\|^2 - \|B-D\|^2.
\end{aligned}$$

Second, introducing vectors $b = B - A$, $c = C - A$, $e = E - D$, and $f = F - D$, the product of the signed triangle areas is proportional to (one quarter of)

$$(b \times c)\cdot(e \times f),$$

where $\times$ denotes the 3D cross product and $\cdot$ denotes the dot product. Lagrange's identity (also known as the quadruple vector product [**Wei06**]) shows that this expression is the same as

$$(b\cdot e)(c\cdot f) - (b\cdot f)(c\cdot e),$$

which can be computed from inter-vertex distances using the dot-product formula above. The purely algebraic expression for calculating signed area products is clearly continuous with respect to the inter-vertex distances. □

The second lemma says that, essentially, if a chain spans a fixed-length gap while its angles flatten, then in fact it is already completely straight.

LEMMA 4.2. *Let $\langle v_p^k, v_{p+1}^k, \ldots, v_q^k \rangle$, $k \geq c$, be an infinite sequence of configurations of a polygonal chain such that (1) v_p^k and v_q^k are a fixed distance d apart, and (2) the intermediate turn angles converge to zero: $\tau_i^k \to 0$ as $k \to \infty$, for $i = p+1, p+2, \ldots, q-1$. Then, in fact, all the configurations in the sequence are identical, with the intermediate vertices lined up along the segment from v_p^k to v_q^k with zero turn angles.*

Proof: We introduce a reference frame centered at v_p^k and with the x axis directed towards v_{p+1}^k, as shown in Figure 10. The directed angle β_i^k between the x axis and the edge $v_i^k v_{i+1}^k$ is given by $\beta_i^k = \sum_{j=p+1}^{i} \tau_j^k$, where we stipulate that β_p^k is zero. Let ℓ_i denote the (fixed) edge length $\|v_i^k - v_{i+1}^k\|$. This allows us to calculate the coordinates of v_q^k as $\left(\sum_{i=p}^{q-1} \ell_i \cos \beta_i^k,\ \sum_{i=p}^{q-1} \ell_i \sin \beta_i^k\right)$. The Pythagorean Theorem now tells us that the distance d between v_p^k and v_q^k satisfies $d^2 = \left(\sum_{i=p}^{q-1} \ell_i \cos \beta_i^k\right)^2 + \left(\sum_{i=p}^{q-1} \ell_i \sin \beta_i^k\right)^2$. As $k \to \infty$, $\tau_i^k \to 0$, so $\beta_i^k \to 0$ also, and we obtain that $d = \sum_{i=p}^{q-1} \ell_i$. Thus the chain is just long enough to span the distance d, so for all $k \geq c$, the vertices between v_p^k and v_q^k are lined up between the two endpoints, and the angles are flat all along. □

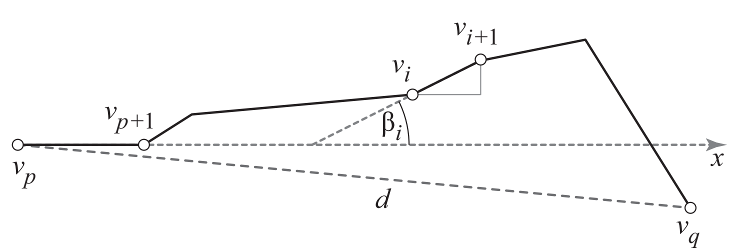


Figure 10. A "flat" chain with endpoints v_p and v_q at a fixed distance d.

We now turn to the proof of Part 1, which consists of four main steps:

(1) Distances between vertices converge, are nondecreasing, and only increase for vertices on different subchains.
(2) The absolute turn angles $|\tau_i^k|$ are nonincreasing and converge.
(3) After a finite number of steps, asymptotically pointed vertices no longer move relative to each other.
(4) After a finite number of steps, asymptotically flat vertices no longer move relative to the asymptotically pointed vertices.

Proof of Part 1: Assume for contradiction that there is an infinite sequence of pocket flips that can be successively applied to a polygon $P = P^0$, leading to the polygon sequence $P^0, P^1, P^2, \ldots$.

Step 1. First we introduce the distance matrix $\mathbf{D}^k$ with entries $\mathbf{D}^k_{i,j} = \|v^k_i - v^k_j\|$, and show that this matrix has a limit $\mathbf{D}^*$. We claim that each entry $\mathbf{D}^k_{i,j}$ is a nondecreasing function of k. Moreover, $\mathbf{D}^k_{i,j}$ is bounded by the perimeter of the polygon P (which equals the perimeter of P^k). Thus it follows that $\mathbf{D}^k$ converges to a limit $\mathbf{D}^*$.

To see the claim, consider an arbitrary pocket flip. Two vertices v^k_i and v^k_j can be either on the same subchain or on different subchains. If they are on the same subchain, or if one of them is an endpoint of both subchains, their relative distance $\mathbf{D}^k_{i,j}$ is unchanged by the flip. Now suppose that they are interior to opposite subchains so that, say, v^k_i reflects across the pocket's line of support L to v^{k+1}_i, while v^{k+1}_j coincides with v^k_j. If either v^k_i or v^k_j is on L, then the distance is still unchanged. Otherwise, v^k_i and v^k_j are strictly on the same side of L, because L is a line of support of P^k. Refer to Figure 11. Because L is the perpendicular bisector of the segment $v^k_i v^{k+1}_i$, i.e., the Voronoi diagram of v^k_i and v^{k+1}_i, we conclude that v^k_j is closer to v^k_i than to v^{k+1}_i, i.e., $\mathbf{D}^{k+1}_{i,j} > \mathbf{D}^k_{i,j}$.

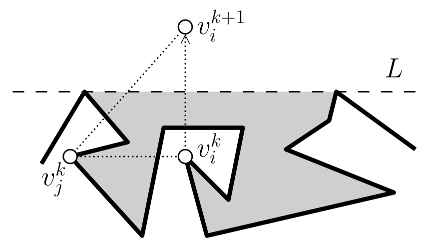


Figure 11. The distance from v_j to v_i increases during a flip.

Step 2. Next we analyze the absolute turn angles $|\tau^k_i|$. Because $\mathbf{D}^k_{i-1,i+1}$ is monotonically nondecreasing, $|\tau^k_i|$ is monotonically nonincreasing. Being also nonnegative, $|\tau^k_i|$ converges. It now makes sense to label each vertex as either asymptotically flat or asymptotically pointed. Moreover, because P^0 has no hairpin vertices, i.e., $|\tau^0_i| < 180°$, no descendant polygon P^k has hairpin vertices either.

Step 3. Next we show that, after a finite number of pocket flips, all asymptotically pointed vertices are in the same pocket subchain of every flip, and thus no longer move relative to each other.

First we introduce the signed-area-product matrix $\mathbf{T}^k$ where entry $\mathbf{T}^k_{i,j}$ is the product of the signed areas of the triangles $v^k_{i-1}v^k_i v^k_{i+1}$ and $v^k_{j-1}v^k_j v^k_{j+1}$. By Lemma 4.1, $\mathbf{T}^k$ is a continuous function of $\mathbf{D}^k$. Thus it has a limit $\mathbf{T}^*$, so there exists a k_0 such that, for any $k \geq k_0$, all entries of $\mathbf{T}^k$ with nonzero limits already have the same sign as in $\mathbf{T}^*$.

Now consider two asymptotically pointed vertices v^k_i and v^k_j. Then $\mathbf{T}^k_{i,j}$ must have a nonzero limit just like the two signed areas in the product (because the vertices are not hairpins). If v^k_i and v^k_j are interior to opposite subchains in a pocket flip, then one of the areas in $\mathbf{T}^k_{i,j}$ changes sign while the other does not, leading to $\mathbf{T}^{k+1}_{i,j} = -\mathbf{T}^k_{i,j}$. For $k \geq k_0$, $\mathbf{T}^k_{i,j}$ cannot change sign, so we must conclude that

v_i^k and v_i^k are on the same subchain in the kth flip. Thus, for $k \geq k_0$, there is no longer any relative motion between asymptotically pointed vertices.

Step 4. Finally we show that, for $k \geq k_0$, asymptotically flat vertices do not move relative to the asymptotically pointed vertices. Consider a maximal chain $v_{p+1}^k, v_{p+2}^k, \ldots, v_{q-1}^k$ of asymptotically flat vertices, surrounded by asymptotically pointed vertices v_p^k and v_q^k. (Such asymptotically pointed vertices exist because every polygon, and hence P^*, has at least two pointed vertices.) By Step 3, v_p^k and v_q^k do not move relative to each other for $k \geq k_0$. By Lemma 4.2, $\tau_{p+1}^k = \tau_{p+2}^k = \cdots = \tau_{q-1}^k = 0$ for $k \geq k_0$, and thus the asymptotically flat vertices are immobile relative to the asymptotically pointed vertices. Therefore, for $k \geq k_0$, the polygon undergoes only rigid motion, which contradicts the assumption that it undergoes an infinite sequence of pocket flips. $\square$

All that remains is to prove Part 2: polygons without hairpins and without flips are already convexified.

Proof of Part 2: Consider a polygon P with no pocket lids and no hairpin vertices. We assume that P has no flat vertices because flat vertices make no difference to the existence of pockets. We prove that P is identical to the boundary H of its convex hull by showing (1) each vertex of H is collocated with exactly one vertex of P, (2) every edge of H has a corresponding edge in P, and (3) every vertex of P is collocated with a vertex of H. These properties imply that P is identical to H, and hence convex, because P and H have the same vertices by (1) and (3), and they have the same edges by (2).

To show (1), we suppose for contradiction that there exist vertices v_i and v_j of P collocated at a vertex h_k of H. We claim that (v_i, v_j, L), where L is any line of support that intersects H only at h_k, is then a valid pocket lid. Indeed, the pocket subchains defined by v_i and v_j are contained in the convex hull, and must leave h_k because there are no zero-length edges. Thus neither subchain can be contained in L. We conclude that each vertex of H is collocated with exactly one vertex of P.

To show (2), we consider vertices v_i and v_j at two consecutive corners of H, and the line of support L connecting them. These vertices define two pocket subchains, one of which has to be contained in L to avoid creating a pocket. That subchain is contained in a line and has no hairpins or flat angles, so the subchain must be just one edge. Thus every edge of H is an edge of P.

To show (3), suppose for contradiction that there is a vertex v_i of P that is not vertex of H. We can pick i so that $v_{i+1} = h_k$ is a corner of H. Because P has no hairpins, H has at least three vertices, so h_k has two distinct adjacent vertices in H. By (1) and (2), these vertices must be collocated with v_i and v_{i+2}, which thus are vertices of H, contradicting the choice of i. Therefore every vertex of P is collocated with a vertex of H. $\square$

To summarize, starting from any polygon without hairpin vertices, Part 1 limits any flip sequence to be finite. Once we run out of flips, Part 2 guarantees the polygon is convex. Thus any polygon without hairpins convexifies in a finite number of pocket flips, regardless of the sequence of flips, concluding the proof of Theorem 1.2.

4.4. Polygons with Hairpins. Our new result, Theorem 1.2, forbids hairpin vertices. Here we explain this exclusion and discuss the nuances of how the definitions treat hairpins.

First, can Part 2 be strengthened to allow hairpin vertices? Figure 12 shows a nonconvex polygon with two hairpin vertices. Our definition of pocket lid forbids either chain from lying entirely along the line of support, so this polygon has no pocket lid. Forbidding collinear chains is also necessary: flipping one of the "zero-area pockets" in this or any other polygon does not change the polygon except possibly by a global reflection, leading to an infinite sequence of flips. To handle this issue, Grünbaum and Zaks [**GZ01**] replace convexity with the notion of exposed polygons, which are contained in the boundary of their convex hull, and for which each vertex of the convex hull corresponds to only one polygon vertex. The polygon in Figure 12 is already exposed. More generally, Part 2 can allow hairpin vertices if we replace "convex" with "exposed." (This fact is implicit in [**GZ01**].)

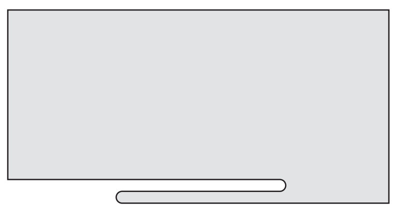


Figure 12. A polygon with two hairpin vertices. All three bottom horizontal segments are collinear. No flips are possible, but this polygon is not convex.

Second, can Part 1 be extended to allow hairpin vertices? Unfortunately, Figure 13 shows that the answer is NO. In this degenerate quadrilateral, choosing a shallow angle ε between L and ab allows any decreasing sequence of turn angles to be generated by successive pocket flips. Indeed, our definition of flips allows the lid to form an angle of $\varepsilon = 0$, causing a cyclic flip sequence. Thus some polygons flip infinitely under poor choices of the pocket lids. Grünbaum and Zaks [**GZ01**] avoid this problem by selecting a particular sequence of flips that convexifies the polygon.

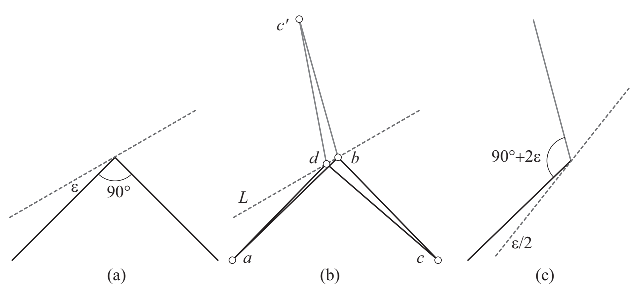


Figure 13. (a) Degenerate quadrilateral $abcd$ before, angle 90°; (b) Details: vertices b and d coincide; L makes angle ε with ab. (c) After flip, quadrilateral angle $90°+2\varepsilon$. Successive flips tilt L at angles $\varepsilon/2, \varepsilon/4, \ldots$.

Although excluding hairpins is necessary in the worst case, there are polygons with "safe" hairpins that nevertheless always flip finitely. Figure 14 illustrates the difference between an "unsafe" and a "safe" hairpin attached to the same polygon. In both cases, *abcde* is a sequence of vertices containing a hairpin *bcd*. In (a), choosing L poorly leads to a new polygon that again permits a similar flip, leading to an infinite flip sequence. However, any flip in (b) prevents further "access" to the coincident vertices b and d, so the hairpin is removed in a future flip. This idea

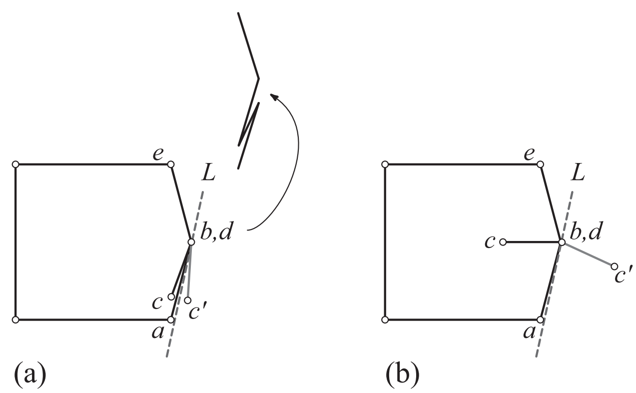


Figure 14. (a) An unsafe hairpin *bcd*. Flipping across a line L nearly parallel to *ab* produces a new hairpin $bc'd$ which can again be flipped across a line nearly parallel to bc'. (b) A safe hairpin *bcd*. Flipping across L leads to removal of the hairpin at the next flip.

leads to our first open problem below.

5. Open Problems

Safe hairpins. Characterize which polygons with hairpins always flip finitely, under our definition of pocket flips. Although we can show that only hairpins attached to the rest of the polygon by two coincident vertices could lead to infinite sequences of flips, distinguishing precisely between "safe" and "unsafe" hairpin configurations remains open.

Definition of pocket flips. We have crafted our definition of a pocket lid and pocket flip to be the most general subject to avoiding certain degenerate pitfalls illustrated in Section 4.4. However, it is possible that polygons always flip finitely under some narrower definition of a flip, e.g., restricting the line of support to bisect the range of possible lines of support when it is ambiguous.

Simultaneous flips. Perhaps the most intriguing challenge is to revisit the original problem posed by Erdős, in which all pockets are flipped simultaneously. With our extension to nonsimple polygons, Figure 2 is no longer an impediment. However, in the nonsimple case, it might not be possible to flip all pocket lids, because some edge might be flipped by two different pocket flips. (This situation happens, for example, in the next step of Figure 2.) We are therefore limited to applying some maximal set of compatible pockets flips. It remains open to specify which set of pockets should be flipped together, and to characterize the polygons that flip finitely under this specification.

The counterexample in Figure 15 shows a nonsimple polygon for which flipping a maximal set of pockets leaves the polygon unchanged up to a reflection. This example is somewhat degenerate because both pocket flips share the same line of support. We do not know whether there are any infinitely flipping examples with two different lines of support, or where the line of support does not contain all unflipped edges. Moreover, were such examples to exist, we do not know whether they are reachable by a sequence of pocket flips from a simple polygon. Thus Erdős's original problem, suitably modified, is still very much alive.

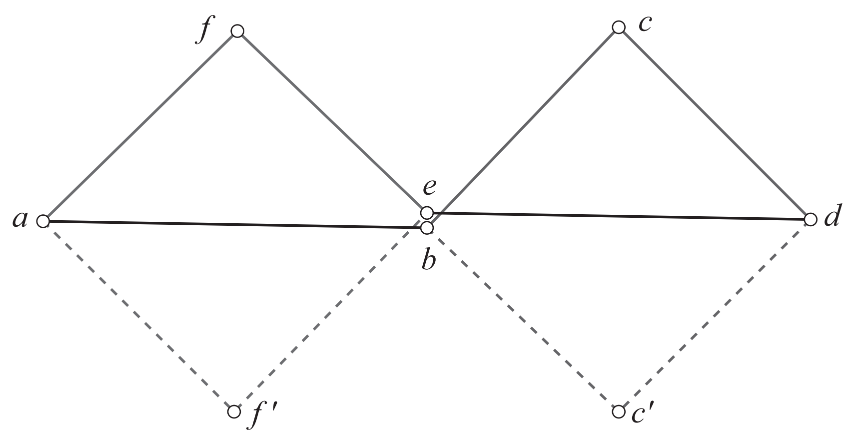


Figure 15. A polygon *abcdef* that is unchanged other than a reflection by applying two simultaneous flips. Vertices b and e coincide, and $\{a, b, d, e\}$ are collinear. The flipped vertex positions are denoted c' and f'.

Curiously, Grünbaum and Zaks [**GZ01**] mention as a comment that, in the context of simultaneous flips, "it is possible to establish an affirmative solution to Erdős's problem, even without restriction to simple polygons." However, this claim remains unjustified.

Bounding flips. In closing, we repeat an open problem posed by Mark Overmars in 1998: is there a reasonable upper bound on the number of flips admitted by a given polygon? For example, can such an upper bound be computed in polynomial time (thus forbidding an explicit execution of the flip sequence)? Is there an upper bound in terms of the number n of vertices and the ratio r between the largest and smallest distances between a vertex and an edge? A pseudopolynomial bound (polynomial in n and r) would be particularly interesting.

Acknowledgments

We would like to thank Martin Demaine, Kathleen Dickey, and Paula Mickevich for helping track down the Russian papers, and Alexandr Andoni for a mathematically accurate translation thereof. We also thank Branko Grünbaum and Bernd Wegner for correspondence and for providing some references on this problem. Finally, we thank Valérie Leblanc for help in proofreading.

References

[BCC+01] Therese Biedl, Timothy Chan, Christian Cumbaa, Erik D. Demaine, Martin L. Demaine, MohammadTaghi Hajiaghayi, Graeme Kemkes, and Ming wei Wang, *University of Waterloo Algorithmic Problem Session*, Unpublished discussions, May 28 2001.

[BK61] R. H. Bing and Nicholas D. Kazarinoff, *On the finiteness of the number of reflections that change a nonconvex plane polygon into a convex one*, Matematicheskoe Prosveshchenie **6** (1961), 205–207, In Russian.

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

[DO07] Erik D. Demaine and Joseph O'Rourke, *Geometric folding algorithms: Linkages, origami, polyhedra*, Cambridge University Press, 2007.

[dSN39] Béla de Sz.-Nagy, *Solution of problem 3763*, American Mathematical Monthly **46** (1939), 176–177.

[Ein83] Albert Einstein, *Geometry and experience*, Sidelights on Relativity, Dover, 1983, Expanded address given to Prussian Academy of Sciences, Berlin, January 27, 1921.

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

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

[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.

[GZ01] ———, *Convexification of polygons by flips and by flipturns*, Discrete Mathematics **241** (2001), no. 1–3, 333–342.

[Had45] Jacques Hadamard, *The psychology of invention in the mathematical field*, Princeton University Press, Princeton, 1945.

[Kal81] Theo Kaluza, *Problem 2: Konvexieren von Polygonen*, Mathematische Semesterberichte **28** (1981), 153–154.

[Kaz61a] Nicholas D. Kazarinoff, *Analytic inequalities*, Holt, Rinehart and Winston, 1961, Dover reprint published in 2003.

[Kaz61b] ———, *Geometric inequalities*, The Mathematical Association of America, Yale University, 1961.

[KB59] N. D. Kazarinoff and R. H. Bing, *A finite number of reflections render a nonconvex plane polygon convex*, Notices of the American Mathematical Society **6** (1959), 834.

[Lak76] Imre Lakatos, *Proofs and refutations: The logic of mathematical discovery*, Cambridge University Press, Cambridge, 1976.

[Mac06] Ernst Mach, *Space and geometry*, The Open Court Publishing Co., Chicago, 1906.

[Pap80] Seymour Papert, *Mindstorms: Children, computers, and powerful ideas*, Basic Books Inc., New York, 1980.

[Res57] Yu. G. Reshetnyak, *On a method of transforming a nonconvex polygonal line into a convex one*, Uspehi Mat. Nauk **12** (1957), no. 3, 189–191, In Russian.

[ST03] J. Antonio Sellares and Godfried T. Toussaint, *On the role of kinesthetic thinking in computational geometry*, Journal of Mathematical Education in Science and Technology **34** (2003), no. 2, 219–237.

[Tou93] Godfried T. Toussaint, *A new look at Euclid's second proposition*, The Mathematical Intelligencer **15** (1993), no. 3, 12–23.

[Tou99] ———, *The Erdős-Nagy theorem and its ramifications*, Proceedings of the 11th Canadian Conference on Computational Geometry, August 1999, Vancouver, Canada, pp. 9–12.

[Tou05] ———, *The Erdős-Nagy theorem and its ramifications*, Computational Geometry: Theory and Applications **31** (2005), no. 3, 219–236.

[Weg93] Bernd Wegner, *Partial inflation of closed polygons in the plane*, Contributions to Algebra and Geometry **34** (1993), no. 1, 77–85.

[Weg96a] ———, *Chord-stretching convexifications of spherical polygons*, Proceedings of the Geometry and Topology Day, May 1996, Coimbra, Portugal, pp. 1–8.

[Weg96b] ———, *Chord-stretching convexifications of spherical polygons*, Textos de Matemática **Coimbra, Série B** (1996), no. 10, 42–49.

[Weg96c] ———, *Convexifications of spherical curves*, Proceedings of the Fourth Congress on Geometry, 1996, Thessaloniki, Grece, pp. 424–430.

[Weg99] ———, *Chord-stretched convex versions of planar curves with self-intersections*, Technical report, Department of Mathematics, Technical University of Berlin, 1999.

[Weg00] ———, *Chord-streched convex versions of planar curves with self-intersections*, Rendiconti del Circolo Matematico di Palermo, Serie II, Supplemento **65** (2000), no. 2, 193–202, Proceedings of the 3rd International Conference on Stochastic Geometry, Convex Bodies and Empirical Measures.

[Wei06] Eric W. Weisstein, *Vector quadruple product*, 2006, http://mathworld.wolfram.com/VectorQuadrupleProduct.html.

[Yus57] A. Ya. Yusupov, *Odno svoistvo odnosvyaznykh nevypuklykh mnogoygolnikov (a property of simply connected nonconvex polygons, in russian)*, Uchenye Zapiski Bukharskogo Gosudarstvenogo Pedagogiceskogo Instituta (1957), 101–103.

MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar St., Cambridge, MA 02139, USA
E-mail address: `edemaine@mit.edu`

MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar St., Cambridge, MA 02139, USA
E-mail address: `gassend@alum.mit.edu`

Department of Computer Science, Smith College, Northampton, MA 01063, USA
E-mail address: `orourke@cs.smith.edu`

School of Computer Science, McGill University, Montréal, Québec, Canada
E-mail address: `godfried@cs.mcgill.ca`