Continuous Foldability of Polygonal Paper

Erik D. Demaine* Satyan L. Devadoss† Joseph S. B. Mitchell‡ Joseph O'Rourke§

Abstract. We prove that any given well-behaved folded state of a piece of paper can be reached via a continuous folding process starting from the unfolded paper and ending with the folded state. The argument is an extension of that originally presented in [DM01].

1 Introduction. In defining an "origami" or "folding" of a piece of paper, there is a distinction between specifying the geometry of the final folded state (a single folding, e.g., an origami crane) and specifying a continuous folding motion from the unfolded sheet to the final folded state (an entire animation of foldings). It is conceivable that some folded state exists, but that the piece of paper could not reach that state via a continuous folding process, e.g., the state could only be reached by passing portions of the paper through itself, or by cutting and regluing.

Our main result is that in fact every well-behaved folded state of a simple polygonal piece of paper can be reached by a continuous folding motion, and so the entire configuration space of all well-behaved folded states of a piece of paper is connected. As a consequence, other results that define foldings with specific properties need not distinguish between folded states and continuous folding motions, and can use the more convenient specification of a single folded state.

The same result as ours was established in [DM01] for the special case of a rectangular piece of paper and a folded state having a flat patch. Here we extend the result to an arbitrary simple polygonal piece of paper and to any well-behaved, possibly entirely curved, folded state, in addition to formalizing definitions and adding detail to the proof.

2 Definitions. We believe that research in mathematical origami has been somewhat hampered by lack of clear, formal foundation, so we devote a relatively lengthy section to this topic before turning to the proofs. At a high level, our definitions generalize Justin's definition of flat folded states [Jus94].

*MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar St., Cambridge, MA 02139, USA, edemaine@mit.edu. Partially supported by NSF CAREER award CCF-0347776.

†Department of Mathematics & Statistics, Williams College, Williamstown, MA 01267, USA, satyan.devadoss@williams.edu. Partially supported by NSF CARGO grant DMS-0310354.

‡Department of Applied Mathematics and Statistics, University of Stony Brook, Stony Brook, NY 11794-3600, USA, jsbm@ams.sunysb.edu. Partially supported by NSF (CCR-0098172, ACI-0328930), NASA Ames (NAG2-1620), Metron Aviation, and the U.S.-Israel Binational Science Foundation (No. 2000160).

§Department of Computer Science, Smith College, Northampton, MA 01063, USA, orourke@cs.smith.edu. Supported by NSF Distinguished Teaching Scholars award DUE-0123154.

2.1 Folded States: Overview. A *piece of paper* P is a closed set defined by a simple planar polygon (i.e., the interior and boundary of that polygon). A *folded state* (f, λ) of P is an isometric function $f : P \to \mathbb{R}^3$ mapping P into Euclidean 3-space, together with a partial function $\lambda : P^2 \to \{-1, +1\}$ specifying the local "stacking order" of pairs of points in contact. The pair (f, λ) must satisfy several conditions, detailed below. First, f must be *isometric* in the sense that the intrinsic geodesic (shortest-path) distance between any two points of P is the same when measured either on P or on the folded-state geometry f. Thus the paper does not stretch or shrink when mapped by the folded state. One consequence of being isometric is that f must be *continuous*, meaning that the folded state does not tear the paper. Second, λ must be *symmetric*, in the sense that it consistently assigns the order of q with respect to p and the order of p with respect to q, *transitive*, so that we obtain a consistent total order on several points in contact, and *consistent*, in the sense that it assigns the same ordering to nearby pairs of points in contact. Third, (f, λ) must be *noncrossing* in the sense that the paper does not cross itself when mapped by the folded state.

2.2 Well-Behaved Folded States. We place a piecewise-smoothness restriction on the geometry f of the folded state. Specifically, f is *well-behaved of order k* if it is piecewise-C^k, i.e., P can be decomposed into a finite number of open sets $R_1, R_2, \ldots, R_m \subset P$, with $\cup_i \overline{R_i} = P$, such that f has continuous derivatives up to order k on each R_i, and the boundary of each R_i consists of finitely many C^k curves. For most of the results in this paper, we need only well-behavedness of order 1, so that we can define a tangent plane at every interior point, but for one proof we need well-behavedness of order 2, so we assume this property of f from now on. We call all boundary points of $\cup_i R_i$ *crease points*.

2.3 Folded States: Isometry. A folded state (f, λ) is *isometric* in the sense that, for any two points $p, q \in P$, the geodesic distance between p and q is the same when measured on either P or the folded-state geometry f. (The isometry condition is independent of λ.) The *geodesic distance* between p and q on P is the length of a shortest path: $\inf \{\text{arclength}\, C \mid \text{curve } C : [0,1] \to P \text{ with } C(0) = p, C(1) = q\}$, where $\text{arclength}\, C$ is defined as usual as $\int_{t=0}^{1} \| \frac{d}{dt} C(t) \| dt$, and $\| \cdot \|$ denotes the Euclidean norm. The *geodesic distance* between p and q on the folded-state geometry f is the length of a shortest path, where length is measured after mapping the curve onto the surface by composing with f: $\inf \{\text{arclength}(f \circ C) \mid \text{curve } C : [0,1] \to P \text{ with } C(0) = p, C(1) = q\}$. Note that even if f folds C on top of itself, this definition captures the length appropriately.

2.4 Folded States: Order. For two distinct noncrease points $p \neq q$ of P mapped into contact by f and having neighborhoods that are mapped into contact by f, $\lambda(p, q) \in \{+1, -1\}$ defines a "stacking order" on p and q. The partial function λ is undefined at (p, q) in all other cases. More precisely, $\lambda(p, q)$ is defined for two points $p, q \in P$ precisely when (a) $p \neq q$, (b) p and q are noncrease points of f, (c) $f(p) = f(q)$, and (d) there are neighborhoods N_p of p and N_q of q (in P) such that $f(N_p) = f(N_q)$. In particular, $\lambda(p, q)$ is defined precisely when $\lambda(q, p)$ is defined.

Intuitively, $\lambda(p, q)$ specifies whether q is above $(+1)$ or below (-1) p with respect to the surface normal of f at p. Note that f does not provide this topological information because $f(N_p) = f(N_q)$; we need to separately keep track of the ordering relation between such points in contact.

λ must satisfy three conditions:

Symmetry. Let $\mathbf{n}(p)$ denote the surface normal vector of f at a noncrease point $p \in P$. Intuitively, the direction of this surface normal specifies which side of the surface was originally the top side of the piece of paper. The symmetry condition constrains any two points $p, q \in P$ for which $\lambda(p, q)$ is defined. If $\mathbf{n}(p) = \mathbf{n}(q)$, i.e., the neighborhoods N_p and N_q are oriented the same, then $\lambda(p, q) = -\lambda(q, p)$. Otherwise, $\mathbf{n}(p) = -\mathbf{n}(q)$, i.e., the neighborhoods are oriented oppositely, and then $\lambda(p, q) = \lambda(q, p)$. See Figure 1.

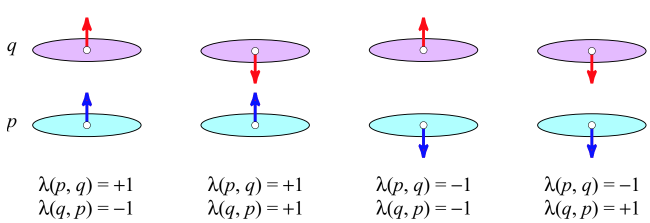


Figure 1: Symmetry of λ.

Transitivity. The transitivity condition constrains the relationship among three points $p, q, r \in P$ in contact. If $\lambda(p, q)$ and $\lambda(q, r)$ are defined, and $\lambda(q, p) = -\lambda(q, r)$, then $\lambda(p, r)$ is defined and $\lambda(p, r) = \lambda(p, q)$. Intuitively, the condition $\lambda(q, p) = -\lambda(q, r)$ specifies that p and r are on opposite sides of q, and the consequence $\lambda(p, r) = \lambda(p, q)$ specifies that r is related to p in the same way as q. See Figure 2.

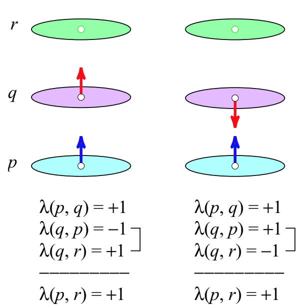


Figure 2: Transitivity of λ.

Consistency. The consistency condition constrains any two points $p, q \in P$ for which $\lambda(p, q)$ is defined. For any connected neighborhoods N_p of p and N_q of q for which $f(N_p) = f(N_q)$, and for any pair of points $\tilde{p} \in N_p$ and $\tilde{q} \in N_q$ for which $f(\tilde{p}) = f(\tilde{q})$, we have $\lambda(\tilde{p}, \tilde{q}) = \lambda(p, q)$. Intuitively, this condition specifies that the entire region of contact surrounding p and q is consistently ordered.

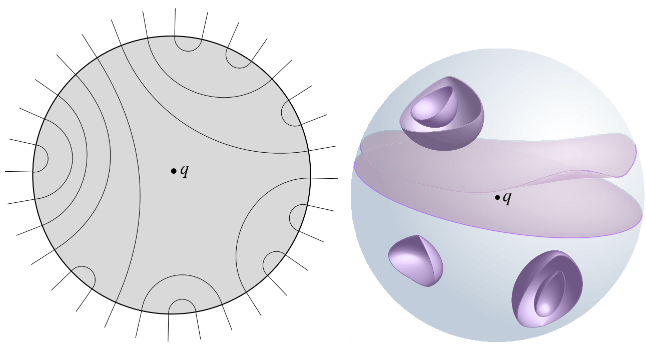
2.5 Folded States: Noncrossing. Intuitively, the noncrossing condition enforces that portions of paper that come into contact geometrically do not properly cross. When the contact between layers occurs in a two-dimensional region (open set), λ arms us with additional order information to disambiguate the geometry. When the contact occurs in a zero- or one-dimensional region (non-open set), the geometry itself is sufficient to disambiguate the ordering.

1D. We start with the definition of the noncrossing condition in the case of folding a one-dimensional piece of paper P (a line segment, or equivalently an interval of $\mathbb{R}$) into $\mathbb{R}^2$. For each point q in $\mathbb{R}^2$, we constrain the local behavior of the folded-state image $f(P)$ around the point q. The idea is to look at portions of paper that come to this point, and ensure that connections between these portions at this point do not cross each other. The main issue here is how to define the notion of a connection.

Consider a point $p \in P$ for which $f(p) = q$. The local behavior of f near p can be characterized, even when p is a crease point, by measuring the unit direction vector from $f(p)$ to $f(p + \delta)$, and the unit direction vector from $f(p)$ to $f(p - \delta)$, as $\delta \to 0$. By the analogous well-behavedness assumption for 1D that there are finitely many crease points between which f is C^1, these limiting directions are well-defined. If f is C^1 at p, then in fact the two directions are negations of each other; in general they correspond to the left and right derivatives of f at p with the left one negated. The two direction vectors can be mapped to two points on the unit circle $\hat{C}$.

We view the interior of the unit circle $\hat{C}$ as an infinitesimal expansion of the behavior at q. The two points on $\hat{C}$ corresponding to p serve as connections between this local behavior and the rest of the folded-state image away from q. We require that the local behavior within $\hat{C}$ connects these two points by a curve, corresponding to an infinitesimal stretching of the point p of paper. Considering all points $p \in P$ for which $f(p) = q$, we obtain a collection of pairs of points on the unit circle $\hat{C}$, where each pair must be connected by a curve within $\hat{C}$. See Figure 4(a). The noncrossing condition requires that these curve connections can be made without crossings; equivalently, there cannot be four points p_1, p_2, p_3, p_4 in cyclic order around $\hat{C}$ such that both (p_1, p_3) and (p_2, p_4) appear as pairs in the collection.

One detail remains: we may obtain multiple copies of the same point on the unit circle $\hat{C}$, and the noncrossing condition requires that these points be distinctly ordered (disambiguated) around the circle. Without loss of generality, suppose that, for $i \in \{1, 2\}$, $p_i \in P$, $f(p_i) = q$, and the unit direction vector from $f(p_i)$ to $f(p_i + \delta)$ approaches v as $\delta \to 0$. If $f((p_1, p_1 + \varepsilon)) \neq f((p_2, p_2 + \varepsilon))$ for all $\varepsilon > 0$, then the local divergence from v of the directions from $f(p_1)$ to $f(p_1 + \varepsilon)$ and from $f(p_2)$ to $f(p_2 + \varepsilon)$ (for small enough ε)



(a) 1D piece of paper. (b) 2D piece of paper.

Figure 4: Illustration of local noncrossing behavior around a point q (the center of the circle or sphere).

specifies a geometric ordering on the two copies of v around the unit circle $\hat{C}$. Otherwise, $f((p_1, p_1+\varepsilon)) = f((p_2, p_2+\varepsilon))$ for some $\varepsilon > 0$, in which case $\lambda(p_1+x, p_2+x)$ provides a consistent value of $+1$ or -1 for all $x \in (0, \varepsilon)$. In this case we use the λ value to determine the order of the two copies of v around the unit circle: as $x \to 0$, $\mathbf{n}(p_1+x)$ approaches one of the two unit tangent vectors to the unit circle $\hat{C}$ at v (see Figure 3), and $\lambda(p_1+x, p_2+x)$ specifies whether p_2's copy of v should be in that direction ($+1$) or in the opposite direction (-1) from p_1's copy of v. By transitivity of λ, this definition provides a consistent total order among all copies of a point v with defined pairwise λ values, which in turn are totally ordered according to the geometry.

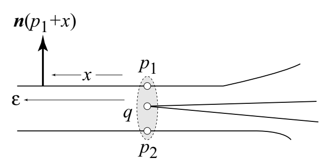


Figure 3: $f(p_1) = f(p_2) = q$ and $f((p_1, p_1+\varepsilon)) = f((p_2, p_2+\varepsilon))$ for some $\varepsilon > 0$.

2D. Finally, we define the noncrossing condition for folded states of a two-dimensional piece of paper P folded in $\mathbb{R}^3$. As before we constrain the local behavior of the folded-state image $f(P)$ around each point q in $\mathbb{R}^3$. This local behavior is characterized by, for each point $p \in P$ for which $f(p) = q$, the unit direction vectors from $f(p)$ to $f(p+\varepsilon v)$ as $\varepsilon \to 0$ over all unit vectors v in $\mathbb{R}^2$. (These unit direction vectors are the normalized directional derivatives of f at p.) For each p interior to P, these vectors give us a closed curve on the unit sphere $\hat{S}$; and each p on the boundary of P gives us an open curve on $\hat{S}$. Each closed curve can be parameterized as a function from the unit circle $\hat{C}$ of directions in $\mathbb{R}^2$ (corresponding to v) to points on the unit sphere $\hat{S}$; similarly, each open curve can be parameterized as a function from the unit interval $[0, 1]$. Thus the picture on the sphere $\hat{S}$ can be viewed as the union of finitely many folded states of 1D pieces of paper, except that some pieces of paper are circles instead of line segments and the folding space is a sphere instead of the plane; see Figure 4(b). Our definition of 1D noncrossing condition trivially generalizes to this multicomponent circular/spherical scenario. The 2D noncrossing condition at q is exactly the 1D noncrossing condition on these folded states applied to every point on the unit sphere $\hat{S}$.

2.6 Folding Motions. Let Fold_P denote the set of all folded states (f, λ) of P. A *folding motion* is a continuous function $M : [0, 1] \to \mathrm{Fold}_P$, where the argument $t \in [0, 1]$ represents time. Let F and Λ denote corresponding functions from $[0, 1]$ that return f and λ, respectively: $M(t) = (F(t), \Lambda(t))$. Continuity of M with respect to t consists of two parts: continuity of F and continuity of Λ.

Time continuity of geometry. Continuity of F is defined in the usual way: for every $\varepsilon > 0$, there is a $\delta > 0$ such that $|t_1 - t_2| < \delta$ implies $d(F(t_1), F(t_2)) < \varepsilon$. However, for this definition to make sense, we need a metric d on the geometric component f of folded states.

For two such folded-state geometries f_1 and f_2, define their *distance* $d(f_1, f_2)$ by $d(f_1, f_2) = \sup_{p \in P} \|f_1(p) - f_2(p)\|$. Thus we measure distance as the maximum Euclidean displacement of a point in P when comparing how that point is mapped by the two folded-state geometries.

Time continuity of order. Continuity of Λ constrains any two points $p, q \in P$ and time $t \in [0, 1]$ for which $\Lambda(t)(p, q)$ is defined. We consider the possible departure of q from p as time increases; the possible arrival of q at p is symmetric (e.g., by reversing time). Let $\mathbf{N}(t)(p)$ denote the surface normal vector of $F(t)$ at a noncrease point $p \in P$. Then we have one of the following two conditions:

1. Departure case: $\left(\frac{d}{dt}F(t)(q) - \frac{d}{dt}F(t)(p)\right) \cdot \mathbf{N}(t)(p)$ is strictly positive if $\Lambda(t)(p, q) = +1$ and is strictly negative if $\Lambda(t)(p, q) = -1$. The derivative is taken with respect to time intervals on the positive side of t, i.e., $(t, t+\varepsilon)$. Intuitively, this condition ensures that q instantaneously departs on the correct side of p, as specified by Λ, in the sense that the the relative motion vector of q with respect to p has a correctly signed dot product with the normal vector at p.

2. Contact case: For every $\varepsilon > 0$, there is a $\delta > 0$ such that, for every $\Delta t \in [0, \delta]$, there is a point q' within a disk of radius ε centered at q for which $\Lambda(t + \Delta t)(p, q')$ is defined and $\Lambda(t + \Delta t)(p, q') = \Lambda(t)(p, q)$. See Figure 5. Intuitively, there is a point q' arbitrarily close to q such that p comes in contact with q' after a sufficiently small motion, and the ordering of that contact is the same as that of p and q at time t.

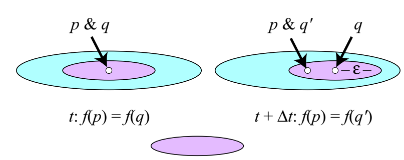
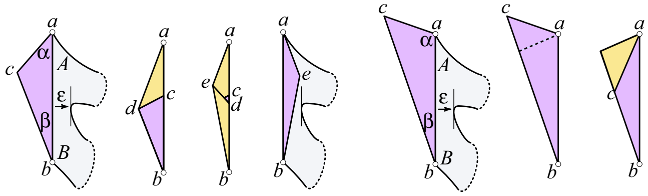


Figure 5: After time Δt, a point q' within ε of q remains in contact with p.

3 Rolling between Flat Folded States. We now proceed to the proof of our main result. The first part claims that we can "roll up" P into a triangle. This motion will use only flat folded states. A folded state (f, λ) is *flat* if the third (z)

coordinate of f is always zero. The *silhouette* of a flat folded state f is the image $f(P)$ of the folded-state geometry.

Lemma 1 *Let $T \subseteq P$ be a triangle that does not intersect any diagonal of some triangulation of the simple polygonal piece of paper P. Then there is a continuous folding of P from the unfolded state into a flat state whose silhouette is congruent to T, such that the intermediate folded states are flat, have finitely many creases, and are nested by subset over time.*



(a) α and β both acute. (b) α not acute.

Figure 6: Illustration of the proof of Lemma 1.

Proof sketch. We repeatedly remove an ear not containing T from the triangulation of P, by continuously rolling the ear onto itself until it fits within P as shown in Figure 6. ☐

4 Unfurling onto the Target Folded State.

We are now prepared for the main theorem:

Theorem 2 *If (f, λ) is a folded state of a simple polygonal piece of paper P that is well-behaved of order 2, then there is a continuous folding motion of P into (f, λ).*

Figure 7 provides a précise of the proof.

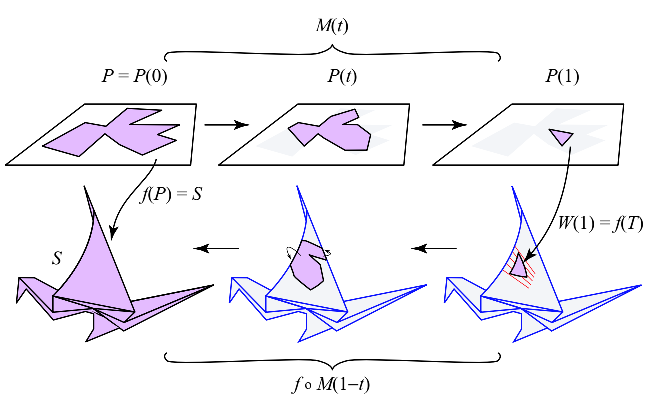


Figure 7: The construction of a continuous folding motion of P into f (not to scale). $S = f(P)$ is the image of the folded state. W is the continuous folding motion that wraps T onto its home $f(T)$ on S. M is the motion that takes P to a flat folded state T within the plane. (The origami bird is based on a design by L. Zamiatina at http://documents.wolfram.com/v4/MainBook/G.2.28.html.)

Proof. Let (f, λ) be a folded state with image $S = f(P)$. Fix some triangulation of P; f maps its diagonals to curves on S. We now locate a triangle T in P (not necessarily a triangle of the triangulation), mapping to a patch $f(T)$ on S, that satisfies these properties:

1. The interior of T avoids all triangulation diagonals.
2. The interior of T avoids all crease points.
3. There is a direction in which the orthogonal projection of $f(T)$ is non-self-overlapping.

It is easy to achieve the first two properties by selecting a suitably small triangle in P. Any such patch is a developable surface, and "torsal ruled," which means that it may be swept out by lines generated by a well-behaved curve [PW01, p. 328]. That the patch is C^2 suffices to ensure that a small enough piece will have a non-overlapping projection. The ruling of the patch can be used to obtain a a continuous folding motion $W(t)$ that wraps the flat triangle T onto this patch $f(T)$ of S. For example, one could bend the ruling lines of the ruled surface $f(T)$, interpolating from a straight segment to the generating curve of the ruling.

Now we apply Lemma 1 to obtain a continuous rolling motion M from P to T, with each $M(t) = (F(t), \Lambda(t))$ flat, $F(0)(P) = P$, and $F(1)(P) = T$. If we then apply the motion $\tilde{W}(t) = (W(t) \circ F(1), \Lambda(1))$, we bring the multilayer flat folding of P from T to $f(T)$.

The last step of the construction is to "unravel" $f(T)$ onto S. One can imagine S as a virtual scaffold, as depicted in Figure 7. The unraveling of $f(T)$ reverses the motion $M(t)$ by considering $M(1-t)$ for $t \in [0, 1]$, but rather than progressing through the continuum of flat states, the motion unfurls on the surface S. Thus, at each time t, we are composing the folded state (f, λ) with $M(1-t)$. The geometry of this composition is simply $f \circ F(1-t)$. The subset-nesting property from Lemma 1 ensures that f is applied only within its domain P. The ordering $\lambda_t(p, q)$ is defined as $\Lambda(1-t)(p, q)$ when that is defined and as $\lambda(F(1-t)(p), F(1-t)(q))$ when that is defined. (Note that at most one of these two alternates is defined; if $\Lambda(1-t)(p, q)$ is defined, then $F(1-t)(p) = F(1-t)(q)$, so $\lambda(F(1-t)(p), F(1-t)(q))$ is undefined.) The noncrossing of this composed state follows from the noncrossing of both (f, λ) and $M(1-t)$. At the end we have continuously folded P into (f, λ). ☐

Corollary 3 *The configuration space of well-behaved folded states is connected.*

Bounding the number of steps (and even defining what constitutes a "step") in the motion from the flat P to the folded state f remains for future consideration.

Acknowledgments. We thank Piyush Kumar and Alan Saalfeld for early discussions that inspired the main result of this paper.

References.

[DM01] E. D. Demaine and J. S. B. Mitchell. Reaching folded states of a rectangular piece of paper. In *Proc. 13th Canadian Conf. Comput. Geom.*, pp. 73–75, 2001.

[Jus94] J. Justin. Towards a mathematical theory of origami. In *Proc. 2nd Internat. Meeting of Origami Science and Scientific Origami*, pp. 15–29, 1994.

[PW01] H. Pottmann and J. Wallner. *Computational Line Geometry*. Springer-Verlag, 2001.