Discrete Comput Geom 8:219–250 (1992)

Discrete & Computational Geometry

Nonoverlap of the Star Unfolding*

Boris Aronov[1] and Joseph O'Rourke[2]

[1] Computer Science Department, Polytechnic University,
Brooklyn, NY 11201, USA
aronov@ziggy.poly.edu

[2] Department of Computer Science, Smith College,
Northampton, MA 01063, USA
orourke@sophia.smith.edu

Abstract. The star unfolding of a convex polytope with respect to a point x on its surface is obtained by cutting the surface along the shortest paths from x to every vertex, and flattening the surface on the plane. We establish two main properties of the star unfolding:

1. It does not self-overlap: it is a simple polygon.
2. The ridge tree in the unfolding, which is the locus of points with more than one shortest path from x, is precisely the Voronoi diagram of the images of x, restricted to the unfolding.

These two properties permit conceptual simplification of several algorithms concerned with shortest paths on polytopes, and sometimes a worst-case complexity improvement as well:

- The construction of the ridge tree (in preparation for shortest-path queries, for instance) can be achieved by an especially simple $O(n^2)$ algorithm. This is no worst-case complexity improvement, but a considerable simplification nonetheless.
- The exact set of all shortest-path "edge sequences" on a polytope can be found by an algorithm considerably simpler than was known previously, with a time improvement of roughly a factor of n over the old bound of $O(n^7 \log n)$.
- The geodesic diameter of a polygon can be found in $O(n^9 \log n)$ time, an improvement of the previous best $O(n^{10})$ algorithm.

Our results suggest conjectures on "unfoldings" of general convex surfaces.

* Part of the research for this paper was carried out while the first author was at the DIMACS Center, Rutgers University (Center for Discrete Mathematics and Theoretical Computer Science), a National Science Foundation Science and Technology Center—NSF-STC88-09648. The second author's research was supported by NSF Grant CCR-882194.

1. Introduction

A new way of organizing the set of all shortest paths from a fixed point x on the surface $\mathscr{P}$ of a (convex) polytope was introduced by Agarwal *et al.* in [AAOS1] and by Chen and Han in [CH], independently and simultaneously.[1] The main idea already appeared in Aleksandrov's work 40 years ago, although he used it only to show that $\mathscr{P}$ can be triangulated.[2] We follow [AAOS1] and refer to this structure as the *start unfolding* of a polytope, so called because of its "star-like" appearance.[3] The star unfolding may be obtained by cutting the polytope along the shortest paths from x to each vertex of $\mathscr{P}$, and flattening the surface on the plane. The star unfolding contrasts with the *source unfolding* [SS], which simply lays out all shortest paths around the source x. In comparison, the star unfolding arranges the paths around their destinations, the ends opposite x. These notions are made precise in Section 1.1.

The star unfolding has proven to be a useful structure for algorithms that involve shortest paths, as detailed in [AAOS1] and [CH]. However, an unfortunate complication was left unresolved in both of these papers: it was not known whether the star unfolding might overlap in a planar layout. This uncertainty forced the algorithms to be unpleasantly complex. The first result of this paper is that indeed the star unfolding does not overlap (Theorem 9.1).

The second result is that the "ridge tree," the locus of points with more than one shortest path from the source, is precisely the Voronoi diagram of the source images in the star unfolding, restricted to the unfolding (Theorem 10.2). This relationship was suspected by researchers, but never established. An illustration is shown in Figs. 1 and 2.[4]

Together these results both conceptually simplify previous algorithms, and in several instances improve the worst-case time complexity as well. In particular, algorithms for constructing the ridge tree, for finding shortest-path edge sequences, and for computing the geodesic diameter of a polytope are all improved. These consequences are discussed briefly in Section 12; details will appear in [AAOS2].

1.1. *Definitions and Basic Properties*

In this section we give formal definitions of the star unfolding and the ridge tree, taken largely from [AAOS1]. Consider a convex polytope in $\mathbb{R}^3$ with n vertices; let $\mathscr{P}$ denote its *surface*. We reserve the term *corners* to refer to vertices of $\mathscr{P}$.

---

[1] During final revisions we learned that Rasch [R] also defined a notion equivalent to the star unfolding.

[2] See p. 171 of [A2] and p. 226 of [A1]. Curiously, Aleksandrov says in [A1] that "Of course [the star unfolding] may self-overlap when unfolded."

[3] See Appendix 2 for several examples. Note that the star unfolding is not necessarily a star-shaped polygon!

[4] The star unfoldings in this figure and those in Fig. 20 were produced with code written by Julie Dibiase and Stacia Wyman at Smith College. The ridge tree was computed by code written primarily by Susan Weller at Johns Hopkins University. The image of the polytope was produced by Darcy Barrant and Jay Greco of Smith College.

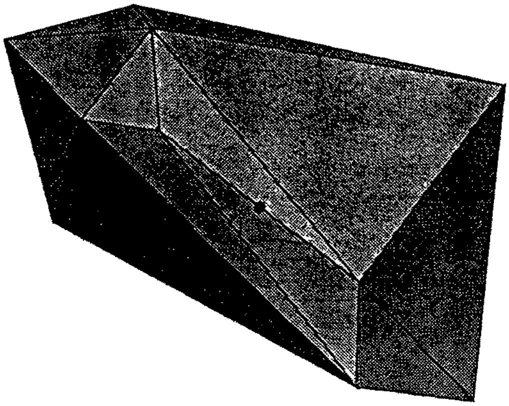

Fig. 1. A polytope of 11 corners; x is marked.

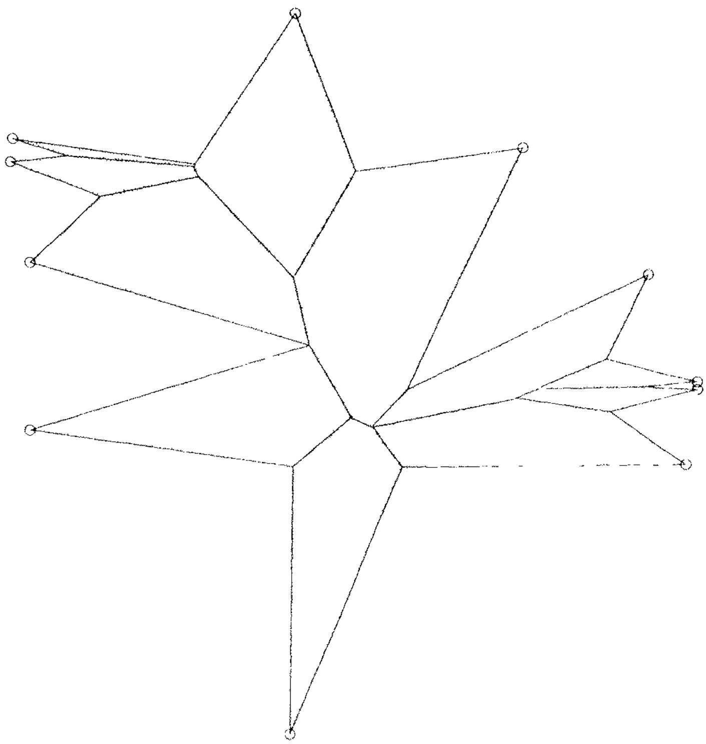

Fig. 2. The star unfolding of the polytope in Fig. 1 with respect to x, with the ridge tree shown. (The scale of Fig. 1 is not maintained.)

1.1.1. Ridge Trees. Given a point x on $\mathscr{P}$, $y \in \mathscr{P}$ is a *ridge point with respect to* x if there are two or more distinct shortest paths between x and y. To simplify our discussion we assume that x does not lie at a corner and has a unique shortest path to each corner. Then ridge points with respect to x form a *ridge tree* T_x embedded on $\mathscr{P}$,[5] whose leaves are corners of $\mathscr{P}$, and whose internal vertices have degree at least three and correspond to points of $\mathscr{P}$ with three or more distinct shortest paths to x [SS]. We define a *ridge* as a maximal connected subset of T_x consisting of points with exactly two distinct shortest paths to x, and containing no corners of $\mathscr{P}$. These are the "edges" of T_x. Ridges are (open) shortest paths [AAOS1]. A *ridge vertex* is a point of the ridge tree incident to more than one ridge (and therefore is of degree three or more). Additionally we consider each corner a vertex of the ridge tree. Under the above assumptions on x each corner has exactly one incident ridge.

1.1.2. Star Unfolding. Let $x \in \mathscr{P}$ be a noncorner point, so that there is a unique shortest path connecting x to each corner of $\mathscr{P}$. These paths are called *cuts* and are composed of *cut points*. The cuts together with edges of $\mathscr{P}$ induce a convex decomposition of $\mathscr{P}$, which we treat as a surface $\mathscr{P}_x$ of a polytope. It is geometrically identical to $\mathscr{P}$, but combinatorially different.

Now form a two-dimensional complex from the faces of $\mathscr{P}_x$ as follows. The cells of the complex are the faces of $\mathscr{P}_x$, each a compact convex polygon. For each pair of adjacent faces of $\mathscr{P}_x$ sharing an edge e of $\mathscr{P}_x$, which is a portion of an edge of $\mathscr{P}$, topologically identify the two faces along e. We define the *star unfolding* S_x as the resulting two-dimensional complex, endowed with its natural intrinsic metric [AZ]. Note that we do not include in our definition any reference to unfolding or flattening. We assume that the complex carries with it labeling information consistent with $\mathscr{P}_x$. Its polygonal boundary ∂S_x consists entirely of edges originating from cuts. It is shown in [AAOS1] that S_x is topologically equivalent to a closed disk.

We think of S_x as laid out in the plane with adjacent faces placed on opposite sides of the line containing their shared edge. The essence of Theorem 9.1 is that nonadjacent faces in such a layout do not overlap either.

1.1.3. Folding and Unfolding Maps. For $y \in S_x$, let $F(y)$ be the unique point $p \in \mathscr{P}$ corresponding to y. F can be viewed as a folding function, mapping each point x in S_x to the single point p up to which it folds. Let $U = F^{-1}$ be the unfolding map, which maps $p \in \mathscr{P}$ to $U(p)$, the set of points in S_x that derive from p. We say that the points in $U(p)$ are *images* of p. Thus $U(p)$ for a point p not on a cut is a single point, $U(x)$ is a set of n distinct points in S_x, a noncorner point $y \in \mathscr{P}$ distinct from x and lying on a cut has exactly two images in S_x, and the corners of $\mathscr{P}$ map to single points. A "segment" in S_x is a connected object that maps to a line segment when S_x is unfolded in the plane. More formally, a connected curve $s \subset S_x$

[5] For smooth surfaces (Riemannian manifolds), the ridge tree is known as the "cut locus" [K].

is a *segment* in S_x if its preimage $F(s)$ is a geodesic on $\mathscr{P}$. In particular, ∂S_x is a cycle of $2n$ segments. In addition, for a point $y \in \mathscr{P}$, any shortest path π from x to y maps to a segment $\pi^* \subset S_x$ connecting an element of $U(y)$ to an element of $U(x)$ [AAOS1].

In [AAOS1] care was taken to distinguish objects on $\mathscr{P}$ and in S_x. Here we are intentionally less careful, to take advantage of the notational simplification gained from the natural correspondence between a set $Q \subseteq \mathscr{P}$ and $U(Q) \subseteq S_x$: unless confusion is possible, we call both Q.

1.1.4. Source Images. Let $X = U(x) = \{x_1, x_2, \ldots, x_n\}$ be the set of source images in S_x. We label the source images and the corners so that they appear as $p_1 x_1 p_2 x_2 \cdots x_{i-1} p_i x_i p_{i+1} x_{i+1} \cdots p_n x_1$ in counterclockwise order around ∂S_x. Note that, with this convention, shortest paths to corners emanate in the cyclic order $\pi(x, p_1), \ldots, \pi(x, p_n)$, clockwise around x on $\mathscr{P}$. We adopt this as the standard ordering of the corners.

1.1.5. Peels. Let a *peel* be the closure of a connected component of the set obtained by removing from $\mathscr{P}$ both the ridge tree T_x and the cuts. A peel is isometric to a convex polygon [SS]. Each peel's boundary consists of x, the shortest paths to two consecutive corners of $\mathscr{P}$, p_i and p_{i+1}, and the unique path in T_x connecting p_i to p_{i+1}. A peel can be thought of as the collection of all the shortest paths emanating from x "between" $\pi(x, p_i)$ and $\pi(x, p_{i+1})$.

1.2. Key Ideas

Both main theorems are proved by induction on the number of corners. There are three key ideas to their proofs.

First, the reduction from n to $n - 1$ corners is chosen to occur in a particular part of the ridge tree, a spot that is shown to always exist.

Second, a powerful theorem of Aleksandrov is used to show that the reduction indeed results in a polytope, to which the induction hypothesis then applies.

Finally, the induction hypotheses are stronger than the bare statements of nonoverlap and the indicated Voronoi property: for both theorems we prove additional structural properties of the unfolding to establish the results.

1.3. Outline

The next section establishes a lemma about ridge trees that identifies the area where the reduction is made. Section 3 then details the reduction. Section 4 describes Aleksandrov's theorem, and Section 5 works out the consequences for the star unfolding. The basis of the induction proofs is explored in Section 6. Key geometric properties of the reduction are established in Section 7. All the material up to this point is used in common for the two main theorems.

Section 8 introduces structural constraints on the star unfolding, and in Section 9 the nonoverlap theorem is proved. The proof of the Voronoi property is given in Section 10.

Extensions to smooth surfaces and algorithmic consequences are discussed briefly in Sections 11 and 12, respectively.

2. Tree Lemmas

This section establishes a simple property of ridge trees (Lemma 2.4), which is used to identify the location on the polytope where the reduction is effected.[6] The notion of "curvature" is used throughout this paper. The *curvature* at a corner p of $\mathscr{P}$ is 2π minus the sum of the face angles incident to p. The curvature of every corner is strictly between 0 and 2π. We use α_i to represent the curvature at p_i. All curvature on a polytope is concentrated at the corners.

Lemma 2.1 (Descartes). *The sum of the curvatures at the vertices of $\mathscr{P}$ is 4π.*

A tree is called *cubic* if every internal node has degree three.

Lemma 2.2. *If T is a cubic tree of $n \geq 4$ leaves, then there are at least two internal nodes in T, each of which is incident to two leaves.*

Proof.[7] Let T' be the subtree of T consisting of just the internal degree-3 nodes; equivalently, T' is obtained by removing all leaves from T. Let a and b be end nodes of a longest path (greatest number of arcs) in T'. It cannot be that $a = b$, for then T could have only $n = 3$ leaves. The nodes a and b have degree 1 in T', and clearly satisfy the claim of the lemma. □

Lemma 2.3. *Any cubic ridge tree T_x contains a ridge vertex adjacent to two corners of $\mathscr{P}$ whose sum of curvatures is no more than 2π.*

Proof. Since any polytope has at least four vertices, T_x satisfies the conditions of Lemma 2.2. For each of the two ridge vertices whose existence is guaranteed by that lemma, consider the sum of curvatures of its two leaf neighbors. If both sums exceed 2π, then the total curvature of $\mathscr{P}$ exceeds 4π, in contradiction to Lemma 2.1. □

Although ridge trees are generically cubic, not all ridge trees are, and we must extend Lemma 2.3 to noncubic trees.

Lemma 2.4. *Any ridge tree T_x contains a ridge vertex adjacent to two consecutive corners of $\mathscr{P}$, whose sum of curvatures is no more than 2π. For a polytope with $n > 4$*

[6] The reader may skip to the statement of Lemma 2.4 without significant loss of continuity.

[7] We thank Joseph Malkevitch for suggesting this proof.

vertices, the sum is strictly less than 2π; *for* $n = 4$, *the curvatures might sum to exactly* 2π.

Proof. Replace every ridge vertex v of degree $k > 3$ with a rooted binary tree B whose leaves are the k nodes adjacent to v, and whose arcs not incident to a leaf are "pseudoridges" of zero length. B is made as full as necessary to cover the k adjacencies, and the leaves are ordered consistent with their circular ordering about v. Now remove the root of B and connect its two children by another pseudoridge arc. Now we have replaced v by a cubic subtree.

Applying this procedure to every ridge vertex whose degree exceeds three produces a cubic tree T' with the same leaf nodes. Apply Lemma 2.3 to T' and obtain nodes v'_1 and v'_2. If v'_i is a true ridge vertex of T_x, it satisfies the conditions of the lemma. If on the other hand v'_i is a pseudovertex introduced in the expansion, it is the expansion of a true ridge vertex v, and the two leaf neighbors of v'_i are necessarily consecutive corners satisfying the lemma.

Finally we turn to the $n = 4$ case. Since the curvature of every vertex is strictly greater than zero, it is only possible to have both pairs of corners summing to exactly 2π when $n = 4$; otherwise Lemma 2.1 would be violated. The regular tetrahedron shows that indeed it is possible for both pairs to sum to exactly 2π. $\square$

The fact that the sum can be exactly 2π when $n = 4$ necessitates special arguments in the base cases of the induction proofs of the two main theorems.

3. Reduction

Let v be the ridge vertex adjacent to the two consecutive corners p_i and p_{i+1}, guaranteed by Lemma 2.4 to have curvatures totaling at most 2π. Make a planar layout of the portion of S_x containing the peels for x_{i-1}, x_i, and x_{i+1}. These three peels meet at v, and do not overlap, because each peel is convex and occupies a disjoint angular wedge emanating from v. See Fig. 3(a). The reduction that permits us to use the induction hypothesis replaces the two corners p_i and p_{i+1} of $\mathscr{P}$ with a new corner p'; eventually we will show this produces a new polytope of $n - 1$ corners $\mathscr{P}'$. We now describe the reduction.

We define $R \subset S_x$ to be the simple polygon $(v, x_{i-1}, p_i, x_i, p_{i+1}, x_{i+1})$, a hexagon that is contained in the union of the three peels discussed above. This region is shaded in Fig. 3(a). R denotes the corresponding region on $\mathscr{P}$ as well. We excise R from the complex S_x, and replace it with a region R', which is the planar quadrilateral $(v, x_{i-1}, p', x_{i+1})$. Let $\angle abc$ denote the angle at b contained counterclockwise between the rays ba and bc. The corner point p' is placed on the bisector of $\angle x_{i-1}vx_{i+1}$ so that its external angle (i.e., its curvature) is the sum of the curvatures at p_i and p_{i+1}: $\alpha' = \alpha_i + \alpha_{i+1}$. See Fig. 3(b).

Lemma 3.1. *There is a point* p' *on the ray bisecting* $\angle z_{i-1}vx_{i+1}$, *whose external angle is* $\alpha_i + \alpha_{i+1}$, *unless* $\alpha_i + \alpha_{i+1} = 2\pi$ *and* $n = 4$.

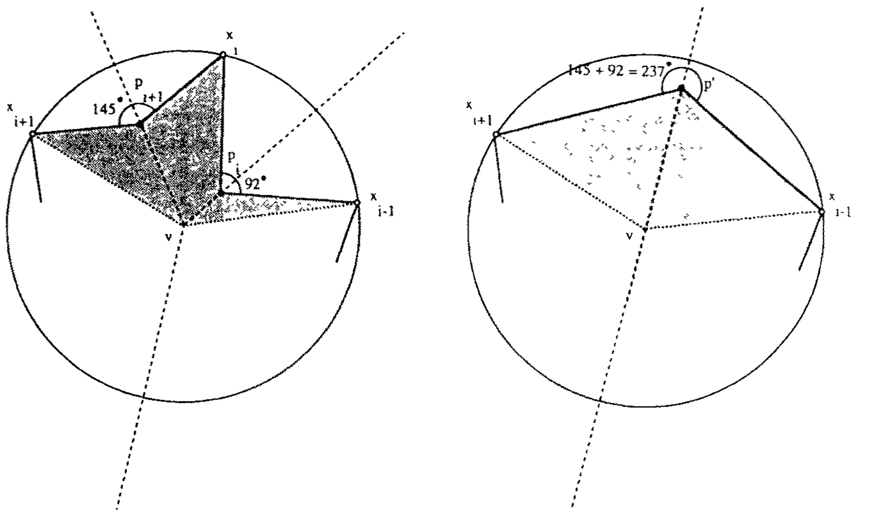

Fig. 3. The reduction: replacement of two corners (a) by one (b).

Proof. The minimum value of α_i occurs when p_i coincides with v; and similarly for α_{i+1}. Therefore $\alpha_i + \alpha_{i+1}$ is larger than $\angle x_{i-1} v x_{i+1}$. On the other hand, by Lemma 2.4, $\alpha_i + \alpha_{i+1} \leq 2\pi$. As p' is moved along the bisector from v to ∞, the exterior angle varies continuously between these two extremes. Therefore there must be some location where p' achieves the precise curvature sum, as long as $\alpha_i + \alpha_{i+1}$ is strictly less than 2π. This is guaranteed for $n > 4$ by Lemma 2.4. □

A second illustration of the reduction is shown in Fig. 4.

This lemma demonstrates that the region R' is well defined. Replacing R by R' produces a new complex $S'_x = (S_x - R) \cup R'$, which has $n - 1$ "corners." The key

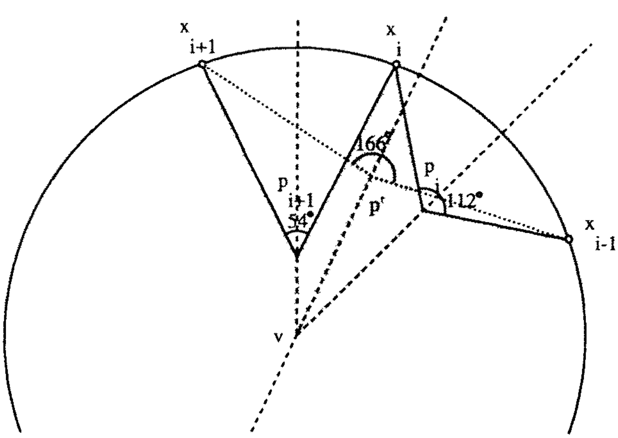

Fig. 4. The reduction, shown with R and R' superimposed.

to the success of the induction proof is to show that this complex corresponds to a (unique) convex polytope $\mathscr{P}'$ and moreover S'_x is its star unfolding. This is by no means obvious, but fortunately it is a corollary of a beautiful theorem of Aleksandrov, which we describe in the next section.

4. Aleksandrov's Theorem

Definition 4.1. A *net* [A2, p. 44] is a complex of polygons with edges topologically identified, such that:

1. Identified edges have the same length.
2. There is a path from every polygon to every other.
3. Every edge of a polygon is identified with at most one edge of another polygon.

Theorem 4.2 (Aleksandrov). "*Every net that is homeomorphic to a sphere and whose angle sum at every vertex is* $\leq 2\pi$, *corresponds to a closed convex polyhedron*" [A2, p. 169].

Other formulations of Theorem 4.2 are cited in Appendix 1.

The star unfolding S_x, with the identification of the two images of cuts from x to each corner, is a net homeomorphic to a sphere, obviously corresponding to the polytope $\mathscr{P}$ from which it is derived.

Lemma 4.3. *Aleksandrov's theorem applies to* S'_x.

Proof. We first argue that S'_x is a net. Since S'_x is obtained by replacing R by R' in S_x, it is clear that S'_x is connected, and each polygon edge is identified with at most one other. Thus we need only check the length condition.

The length condition is satisfied, since $|x_{i-1}p'| = |x_{i+1}p'|$, as p' is on the bisector between these source images. All other edges of S'_x are "inherited" from S_x in pairs.

Next we must argue that the angle sum at every vertex does not exceed 2π; that the net is homeomorphic to a sphere is clear.

Because the curvature at p' is $\leq 2\pi$, the angle condition that is imposed on nets is satisfied at all "corners." We must also show that the angle condition holds at x itself.

The sum of the interior angles of a simple k-gon is $\pi(k-2)$. We compute this sum for both R and R', which are simple polygons of six and four vertices, respectively. Let τ_j be the (positive) interior angle at x_j in R, and let τ'_j be the corresponding angle in R'. Finally, let β be the interior angle at v, common to both R and R'. Then the two sums are

$$\tau_{i-1} + (2\pi - \alpha_i) + \tau_i + (2\pi - \alpha_{i+1}) + \tau_{i+1} + \beta = 4\pi$$

and

$$\tau'_{i-1} + (2\pi - \alpha') + \tau'_{i+1} + \beta = 2\pi.$$

Using $\alpha' = \alpha_{i-1} + \alpha_i$, we have that $\tau_{i-1} + \tau_i + \tau_{i+1} = \tau'_{i-1} + \tau'_{i+1}$. Thus the sum of the angles incident to x has not changed as a result of the reduction. □

5. Reduced Star Unfolding

By Lemma 4.3 and Theorem 4.2, S'_x folds to a polytope $\mathscr{P}'$, to which the induction hypothesis applies. Now we concentrate on the transformation from $\mathscr{P}'$ to $\mathscr{P}$ as represented in Fig. 5: the region R' is cut out and replaced by R, the reverse of the reduction discussed in Section 3. The goal of this section is to show that S'_x is precisely the star unfolding of $\mathscr{P}'$. Namely, the star unfolding of $\mathscr{P}'$ is exactly the same as S_x, the unfolding of $\mathscr{P}$, except for the regions R and R' cut and pasted.[8] This permits us to reason entirely with the unfoldings.

We use the notation $\mathscr{P} - R$ to represent the surface of P with the region R removed; since $R \subset S_x$, this is a shorthand for $\mathscr{P} - F(R)$. We should point out that our argument does not depend on the three-dimensional geometry of $\mathscr{P}'$ with respect to that of $\mathscr{P}$. It is clear that in general the dihedral angles on $\mathscr{P}' - R'$ differ from the corresponding angles on $\mathscr{P} - R$. However, these angles play no role in our proofs, since shortest paths depend only on the intrinsic metric.

Lemma 5.1. *If $\pi'(x, y)$ is a shortest path on $\mathscr{P}'$, lying wholly within $\mathscr{P}' - R'$, then the same path is shortest on $\mathscr{P}$.*

Proof. First note that because π' avoids R', which is the only region that differs between $\mathscr{P}'$ and $\mathscr{P}$, there is a path corresponding to π' on $\mathscr{P}$, which we continue to call π'.

Suppose in contradiction to the lemma that $\pi'(x, y)$ is not shortest on $\mathscr{P}$. Then

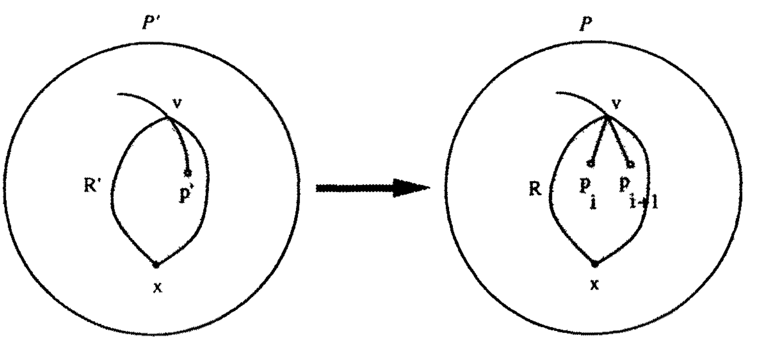

Fig. 5. The reduction reversed, viewed on the polytope surface.

[8] The reader may skip to Lemma 5.4 and its corollaries without significant loss of continuity.

since only region R is different on $\mathscr{P}$, it must be the case that any shortest path $\pi(x, y)$ on $\mathscr{P}$ must cross R on its way to y. Not only must π cross R, it must be a path emanating from the source image x_i in S_x; for shortest paths from x_{i-1} and x_{i+1} to points outside R never meet the interior of R (see Fig. 3). Thus π must be in x_i's peel. By construction (Section 3), this peel is wholly contained within R, but the destination y is not in R—a contradiction. □

Lemma 5.2. *The paths that comprise the boundary of R' in $\mathscr{P}'$ are shortest paths.*

Proof. Let the paths be π'_1 and π'_2. These paths have exact correspondents on $\mathscr{P}$, where they form the boundary of R. On $\mathscr{P}$ they are shortest paths by construction (Section 3).

Suppose π'_1 is not a shortest path from x to v on $\mathscr{P}'$, in contradiction to the claim of the lemma. Then there is a shortest path π' on $\mathscr{P}'$ from x to v that is shorter than π'_1. We distinguish several cases and derive a contradiction for each:

1. $\pi' \subseteq R'$. In a layout of R' (see Fig. 3), this path must unfold to a straight-line segment and thus coincide with $x_{i\pm 1}v$, i.e., correspond to π'_1 or π'_2.
2. $\pi' \subseteq \mathscr{P}' - R'$. Then there is a corresponding path π on $\mathscr{P}$ which has the same length as π' and thus is shorter than π_1, contradicting the choice of π'_1.
3. π' is in $\mathscr{P}' - R'$ in the vicinity of x, but crosses π'_1 or π'_2 into R' before reaching v. Let π' cross π'_1 at $y \neq v$. Then $y \in T_x$ on $\mathscr{P}$, since initial segments of π' and π_1 correspond to distinct shortest paths to y on $\mathscr{P}$. However, then π_1 is a shortest path on $\mathscr{P}$ that passes through and beyond a point of the ridge tree, contradicting the property that a shortest path never extends past a ridge point [SS].
4. π' is interior to R' in the vicinity of x, but crosses π'_1 or π'_2 out to $\mathscr{P}' - R'$ before reaching v. In a layout of R' (again see Fig. 3), this corresponds to a path from (say) x_{i-1} through $y \in x_{i+1}v$. (Note that a path in the layout from x_{i-1} to any point $y \in x_{i-1}v$ must coincide with $x_{i-1}v$ and thus not enter the interior of R'.) However, this path cannot be shortest within R', as it crosses the (x_{i-1}, x_{i+1}) bisector, so $|x_{i+1}y| < |x_{i-1}y|$, contradicting the assumption that π' is shortest on $\mathscr{P}'$. □

Lemma 5.3. *If $y \in \mathscr{P}' - R'$, a shortest path $\pi'(x, y)$ lies wholly within $\mathscr{P}' - R'$. If y lies in the interior of R', any shortest path $\pi'(x, y)$ lies in R'.*

Proof. Suppose to the contrary that $y \in \mathscr{P}' - R'$ and π' intersects the interior of R'. Then it must intersect $\partial R'$. So it must either cross one of the shortest paths forming the boundary of R' in $\mathscr{P}'$ or pass through the ridge vertex v, but two shortest paths from x cannot cross and a shortest path never extends past a ridge point [SS]. The other case is handled similarly. □

Lemma 5.4. *S'_x is the star unfolding of $\mathscr{P}'$.*

Proof. It is sufficient to show that S'_x can be obtained from $\mathscr{P}'$ by cutting along the shortest paths from x to every corner.

Indeed, the boundary of S'_x is formed by the twin images $x_{j-1}p_j$ and p_jx_j, $j \neq i$, $i+1$, of the shortest paths from x to corners of $\mathscr{P}$, and two segments $x_{i-1}p'$ and $p'x_{i+1}$. As $p_j \in \mathscr{P}' - R'$, for $j \neq i$, $i+1$, by Lemma 5.1 $x_{j-1}p_j$ and p_jx_j correspond to the shortest path from x to p_j on $\mathscr{P}'$. Lemma 5.3, on the other hand, implies that $x_{i-1}p'$ and $p'x_{i+1}$ are the two images of the shortest path from x to p'. Thus S'_x is indeed obtained from $\mathscr{P}'$ by cutting along the shortest paths from x to all corners. □

Corollary 5.5. *The ridge trees are the same in S_x and S'_x outside the regions that differ between these two unfoldings*: $T'_x - R' = T_x - R$.

Proof. Since the entire shortest path structure is the same outside R and R' by Lemmas 5.1 and 5.3, the locus of points with two or more shortest paths from x is the same. □

Corollary 5.6. *In $\mathscr{P}'$, x is not a ridge point of any corner of S'_x.*

Proof. By Lemmas 5.1 and 5.3, the number of shortest paths to p_j, $j \neq i$, $i+1$, does not change between $\mathscr{P}$ and $\mathscr{P}'$. Thus x is not a ridge point for any of these corners, since, by assumption, x has this property in $\mathscr{P}$. It only remains to check p'. However, by Lemma 5.3, there is a unique shortest path to p'. □

This permits us to assume "nonridgeness" inductively.

5.1. *More Notation*

Lemma 5.4 permits the following view of the reduction, which we adopt in the remainder of the paper. S_x and S'_x differ only in the replacement of two corners and one source image in S_x, by one corner in S'_x. If we lay S_x and S'_x on top of one another in the plane, the $n-1$ source images that they share will coincide. We therefore use the same labels for these sources:

$$x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n;$$

and for the common corners:

$$p_1, p_2, \ldots, p_{i-1}, p_{i+2}, \ldots, p_n.$$

In what follows x_i always refers to the source image of S_x removed by the reduction, and p_i and p_{i+1} refer to the two corners removed; p' is used to denote the corner added to S'_x. In general, primes denote quantities of S'_x.

5.2. *Example*

Figure 6(a) shows an unfolding of a square pyramid, with x at the midpoint of one of the base square's edges. Figure 6(b) shows the star unfolding, and a region

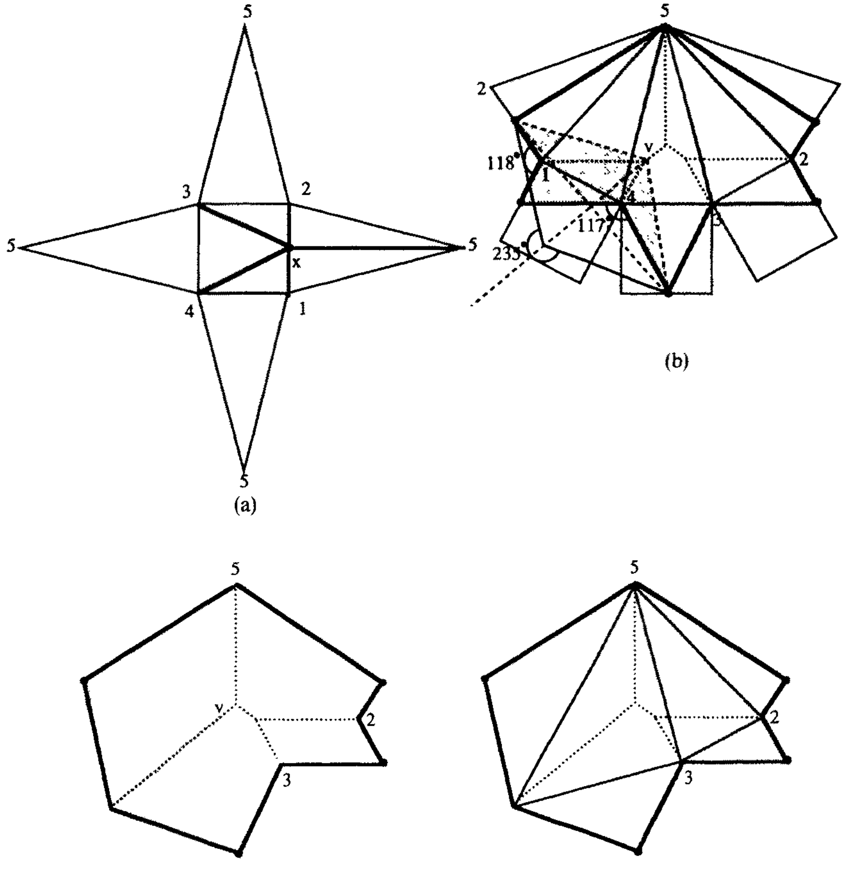

Fig. 6. The star unfolding of a pyramid reduced to a tetrahedron.

R identified for the reduction step of the induction. Figure 6(c) shows the unfolding after R is replaced by R'. If Fig. 6(d), which is Fig. 6(c) redrawn, is folded along the lines shown, the result is a convex polyhedron (a tetrahedron), as guaranteed by Aleksandrov's theorem. If the reduction is applied to Fig. 6(d), the base case of the induction is reached, a doubly covered triangle. All three star unfoldings produced in this reduction process are shown in Fig. 7.

6. Induction Basis

6.1. Generic Case: Doubly Covered Triangle ($n = 3$)

Each reduction step reduces n, the number of vertices, by one. The "generic" basis of the induction is $n = 3$, when the star unfolding is a hexagon: three corners and

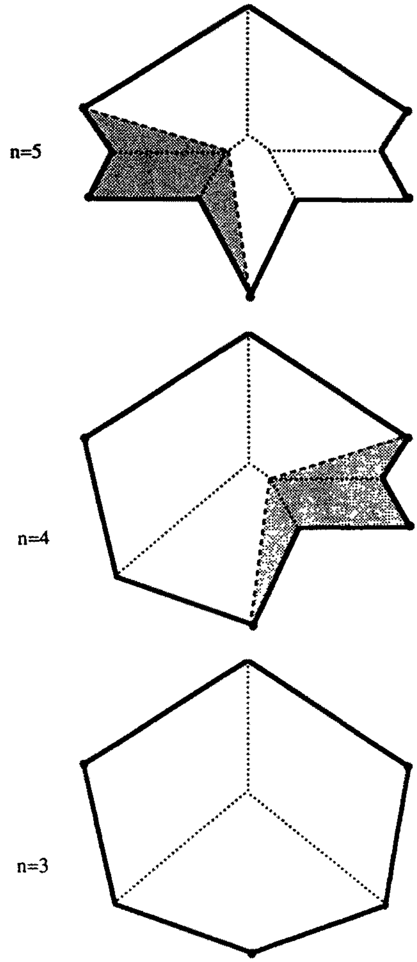

Fig. 7. Three star unfoldings from the pyramid.

three source images. An example is shown in Fig. 7. The corresponding polytope is a flat, "doubly covered" triangle with x on one side, a degenerate case permitted by Aleksandrov's theorem. Although this doubly covered triangle has zero volume, it behaves as the surface of any other convex polytope.

6.2. Special Case: Special Tetrahedron ($n = 4$)

In the special case when $n = 4$ and the pair of vertices guaranteed by Lemma 2.4 have curvature sum exactly 2π, the reduction Lemma 3.1 does not apply (p' would have to be on the bisector "at infinity"), and the base case is a tetrahedron. Although the reduction fails, there is a sense in which it can be carried out

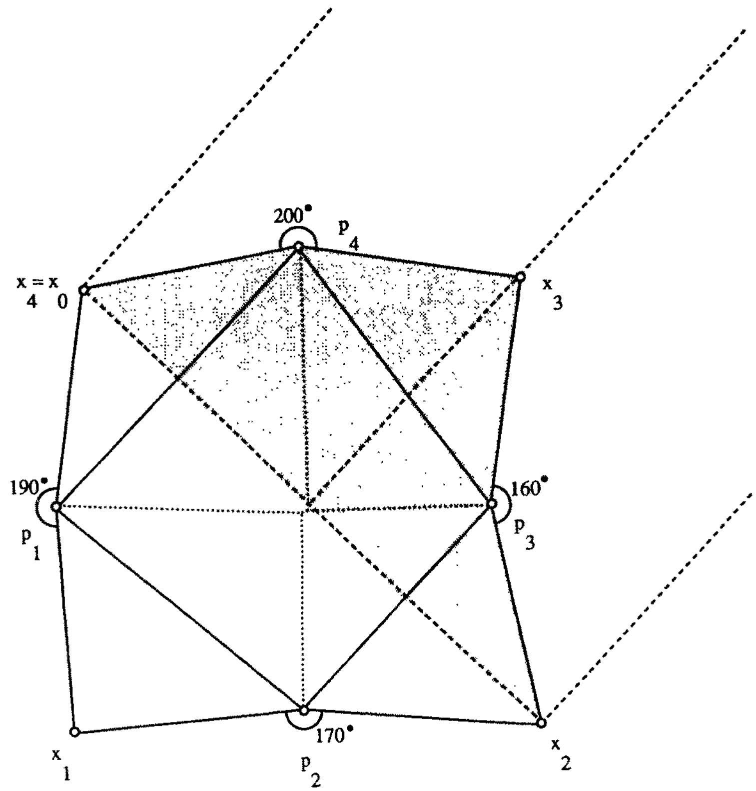

Fig. 8. $n = 4$, angle sums equal to 2π.

nevertheless, and we proceed in this section to demonstrate this in order to facilitate establishing the bases of the induction proofs.

Figure 8 shows an example of a star unfolding for $n = 4$, with $\alpha_1 + \alpha_2 = \alpha_3 + \alpha_4 = 2\pi$. If we choose to reduce p_3 and p_4 via Lemma 3.1, the region R is as shaded in the figure. Applying the reduction sends p' out to infinity along the dashed (x_2, x_4) bisector. The result is an unbounded hexagon S'_x, with two adjacent parallel edges: $x_2 p'$ and $x_4 p'$. Unfortunately this unbounded figure falls outside the purview of Aleksandrov's theorem (Theorem 4.2), and so we cannot claim that S'_x is the unfolding of a polytope. However, it is the unfolding of a doubly covered unbounded "triangle."

Lemma 6.1. *An unbounded hexagon that results from applying the reduction to S_x for $n = 4$ with $\alpha_i + a_{i+1} = 2\pi$, is an unfolding of a doubly covered unbounded triangle, one with one bounded edge and two parallel unbounded edges.*

Proof. Orient the two unbounded edges of the hexagon vertically downward, as shown in Fig. 9. (In this figure p_1 and p_2 are depicted as both lying above $x_2 x_4$, although one or the other could be below. The subsequent geometric argument

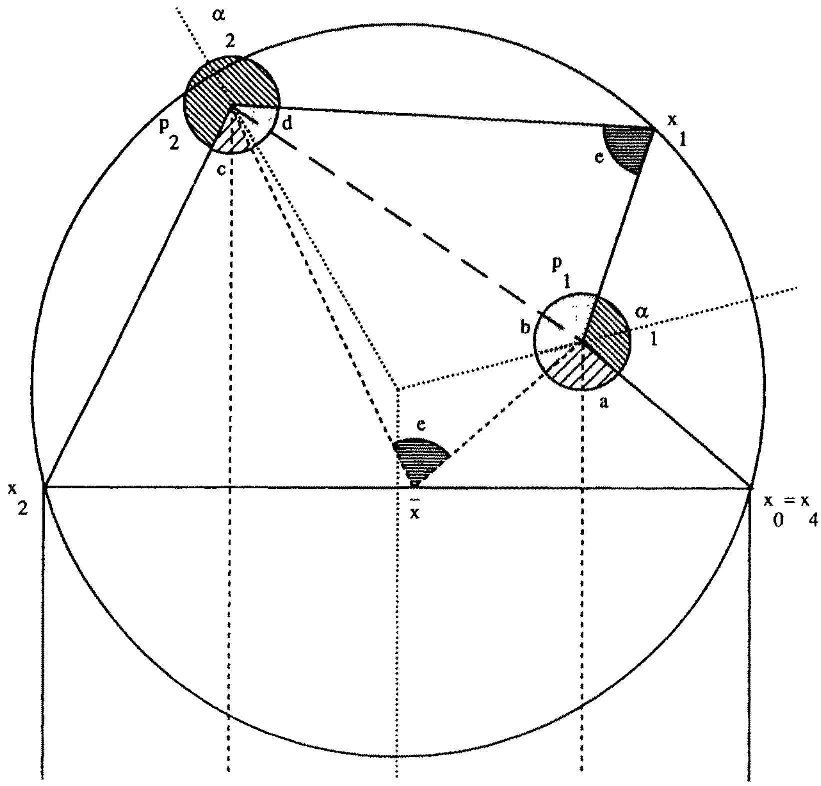

Fig. 9. Unbounded hexagon folds to an unbounded doubly covered triangle.

is not altered.) Relabel x_4 to be x_0. By assumption, $\alpha_1 + a_2 = 2\pi$. Let $\alpha_1 \leq \alpha_2$ without loss of generality. We show that folding x_1 over the p_1p_2 fold line sends it to a point $\bar{x}$ on the segment x_0x_2, and that folding x_0 over the vertical line through p_1, and x_2 over the vertical line through p_2, maps these points to the same point $\bar{x}$. This show that the folds produce a doubly covered infinite triangle with edges p_1p_2 and the two vertical rays from p_1 and p_2 downward.

Reflect the $x_1p_1p_2$ triangle through the line containing p_1 and p_2, obtaining the point $\bar{x}$ as shown in Fig. 9. Let e, b, e, and d be the four angles of the quadrilateral $(\bar{x}, p_1, x_1, p_2)$ as shown in the figure. The angle at x_1 is the same as that at $\bar{x}$ by reflection. We have that $b + d + 2e = 2\pi$.

Let the angles surrounding p_1 be α_1, b, and a as shown in the figure, and around p_2 be α_2, c, and d. We have $\alpha_1 + b + a = 2\pi$ and $\alpha_2 + d + c = 2\pi$. From $\alpha_1 + \alpha_2 = 2\pi$, it follows that $a + b + c + d = 2\pi$.

We now compute the angle $\angle x_0\bar{x}x_2$. Note that the triangles $(x_0, p_1, \bar{x})$ and $(x_2, p_2, \bar{x})$ are both isosceles. Thus

$$\angle x_0\bar{x}x_2 = e + (\pi - a)/2 + (\pi - c)/2.$$

(Here the angle $\pi - a$ is negative if p_1 is below x_0x_2, and $\pi - c$ is negative if p_2 is below.) Substituting $e = \pi - (b + d)/2$ reduces this to $2\pi - (a + b + c + d)/2$ which is π. This shows precisely what we claimed above: $\bar{x}$ lies on x_0x_2, and reflecting x_0 and x_2 about the vertical lines through p_1 and p_2 (respectively) maps them to $\bar{x}$. □

With this lemma available we may use $n = 3$ as the only base of the induction proofs, with the understanding that the doubly covered triangle may be unbounded in the sense above.

7. Reduction Geometry

In this section we establish a crucial geometric lemma concerning the relative angles of edges in R' and R. This relationship derives ultimately from the fact that the curvature α' at p' is the sum of the curvatures α_i and α_{i+1}.

Lemma 7.1 (Reduction Angles). *In the reduction $S_x \Rightarrow S'_x$, edge $x_{i-1}p'$ of S'_x is "exterior" to edge $x_{i-1}p_i$ of S_x, and edge $x_{i+1}p'$ is "exterior" to edge $x_{i+1}p_{i+1}$, in the sense that*

$$\angle p'x_{i-1}v \geq \angle p_i x_{i-1}v,$$

$$\angle vx_{i+1}p' \geq \angle vx_{i+1}p_{i+1}.$$

Proof. See Fig. 4: the dashed line bounding R' is exterior to R in the vicinity of $x_{i\pm 1}$.

The proof is by simple plane geometry, but the argument is complex enough to require new notation. Refer to Fig. 10(a). Place the origin of the coordinate system at v. Let the three sources x_{i-1}, x_i, and x_{i+1} be at angles $a = 0$, b, and c measured counterclockwise from the horizontal axis. Then p_i is at angle $b/2$, p_{i+1} is at angle $(b + c)/2$, and p' is at angle $c/2$. Observe that $b/2 < c/2 < (b + c)/2$. Let the curvature at p_i and p_{i+1} be α and γ, respectively. We first claim that the ray $x_{i-1}p_i$ crosses the $c/2$ bisector, as does the ray $x_{i+1}p_{i+1}$. This is proved in Lemma 7.2 below. Assume now without loss of generality that the ray $x_{i-1}p_i$ intersects the $c/2$ bisector at a point q further away from v than does the ray $x_{i+1}p_{i+1}$. Let $\beta = \angle x_{i-1}qx_{i+1}$. Our goal is to show that $\beta \leq \alpha + \gamma$, because this will show that p', which achieves an exterior angle of $\alpha + \gamma$ by construction (Lemma 3.1), lies further out on the $c/2$ bisector than does q. Since q was determined by the larger angle at x_{i-1} and x_{i+1}, the lemma will be established.

Consider the shaded triangle in Fig. 10(a). Its three angles are $\alpha/2$, $(c - b)/2$, and $\pi - \beta/2$. Summing to π and solving for β yields $\beta = (c - b) + \alpha$. However, now note that the minimum value of γ is $c - b$, when p_{i+1} is at v; and we argue in the proof of Lemma 7.2 below that $v \neq p_{i+1}$. Therefore $\beta < \gamma + \alpha$.

This argument works identically even if $\alpha > \pi$, as shown in Fig. 10(b), or if $\angle x_{i-1}vx_{i+1} > \pi$, as shown in Fig. 10(c), or both (figure omitted). We cannot have both $\alpha > \pi$ and $\gamma > \pi$, since their sum is at most 2π. □

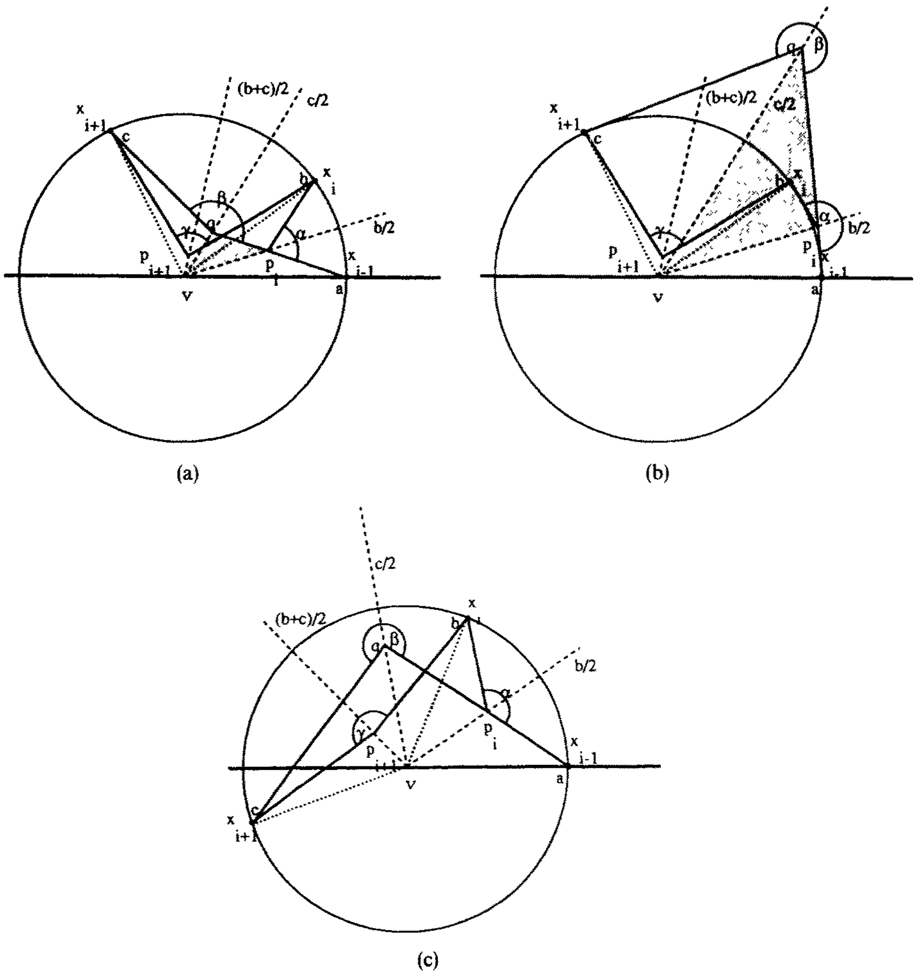

Fig. 10. Reduction angles.

Lemma 7.2. *In the notation used in the proof of Lemma 7.1, rays $x_{i-1}p_i$ and $x_{i+1}p_{i+1}$ must cross the $c/2$ bisector.*

Proof. By symmetry, it is sufficient to consider only one of the rays. Suppose to the contrary that $x_{i-1}p_i$ is parallel to the $c/2$ bisector, as illustrated in Fig. 11. As already observed, $c - b$ is the smallest possible value of γ, the curvature at p_{i+1}. Note that it can never be achieved, as that would require p_{i+1} to coincide with ridge vertex v, which would mean that x is a ridge point with respect to p_{i+1}, a situation we have explicitly excluded. Consider the quadrilateral (v, x_{i-1}, p_i, x_i), shown shaded in the figure. Its four angles sum to 2π:

$$b + (\pi - c/2) + (2\pi - \alpha) + (\pi - c/2) = 2\pi.$$

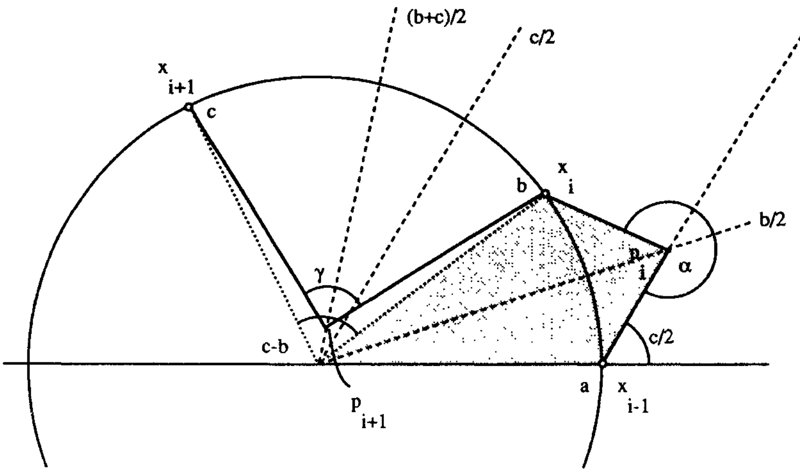

Fig. 11. The $x_{i-1}p_i$ ray cannot be parallel to the $c/2$ bisector.

This simplifies to $\alpha + c - b = 2\pi$, but, since $\gamma > \gamma_m = c - b$, $\alpha + \gamma > 2\pi$. Thus, in order for the ray to just miss the $c/2$ bisector, we must have the sum of the curvatures at p_i and p_{i+1} exceed 2π, contradicting the choice of p_i, p_{i+1}. If $x_{i-1}p_i$ is not parallel to the $c/2$ bisector and misses it, then we derive $\alpha + \gamma_m > 2\pi$, and the same conclusion follows. □

8. Sectors

Examination of Fig. 4 shows that in the $S'_x \Rightarrow S_x$ transition, R may extend beyond R', which presents a fundamental difficulty for a proof of nonoverlap of S_x from the nonoverlap of S'_x: nonoverlap of S'_x does not suffice—we need something stronger. The required stronger condition is provided by a structural geometric constraint on the shape of the star unfolding, which we phrase in terms of circle sectors that lie just outside ∂S_x.

8.1. *Definition of Sectors*

We now define a region of the plane associated with each corner of a layout of S_x. The definition does not assume that S_x does not overlap, as it only depends on the positions of x_{j-1}, p_j, and x_j in the layout.

Define the *sector* s_j associated with p_j as the closed sector of the disk centered on p_j bounded by the radii p_jx_{j-1} and p_jx_j, and exterior to S_x near p_j. See Fig. 12. The sectors for the unfolding of the pyramid shown in Fig. 6 are depicted in Fig. 13. We will see that the sector interiors are pairwise disjoint and exterior to S_x.

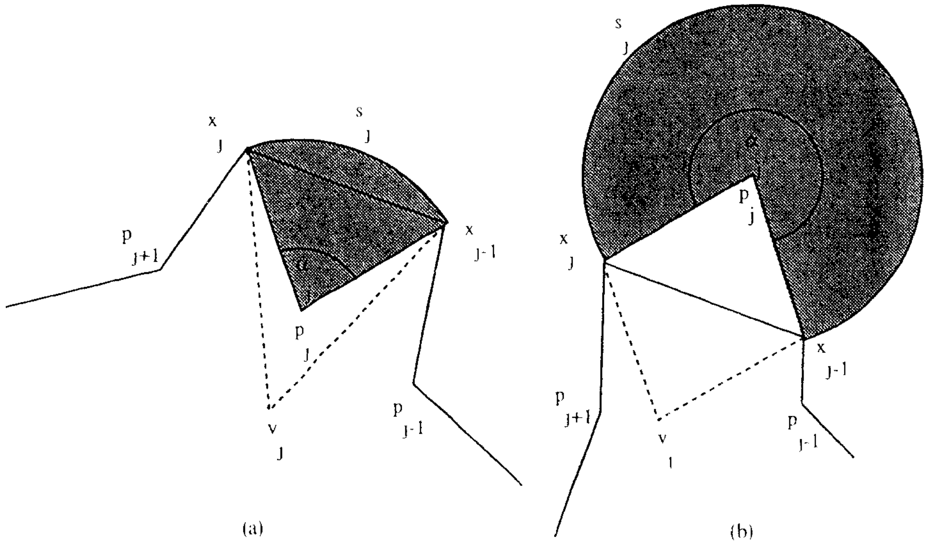

Fig. 12. Definition of sectors: (a) $\alpha_j < \pi$; (b) $\alpha_j > \pi$.

8.2. Reduction Notation

We use induction based on the reduction described in Section 3. Assuming various induction hypotheses for the reduced S'_x, we are attempting to establish the hypotheses for S_x. In the $S'_x \Rightarrow S_x$ transition, we consider S'_x to be "old" and S_x to be "new," and use these adjectives liberally. Thus S_x includes a new source x_i inserted between x_{i-1} and x_{i+1}. We use primes to indicate an old quantity in S_x

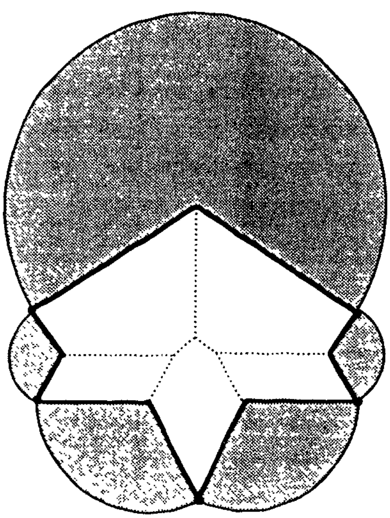

Fig. 13. Sectors for the pyramid unfolding (Figs. 6 and 7)

that is altered in the transition to S_x. For instance, v' is the ridge-tree vertex adjacent to corner p' and the sector associated with p' is s'.

8.3. Sectors Nested

The key property of sectors is that the reduction implies a "nesting" of sectors in a certain sense, as illustrated in Fig. 14. We will see in the next section that this nesting implies that the sector interiors are pairwise disjoint and lie outside S_x. In preparation, we show that adjacent sectors do not overlap in the vicinity of their point of adjacency:

Lemma 8.1. *The interiors of adjacent sectors are disjoint in a neighborhood of their shared source point* x_j.

Proof. The lemma follows from two facts:

(1) The interior angle at x_j in S_x is no greater than π.
(2) The interior angle of a sector incident to x_j has measure $\pi/2$.

Claim (1) follows because the source x must lie in some convex face of $\mathscr{P}$, and so the shortest paths to the vertices of this face already cut up the angles about x into pieces no larger than π. Claim (2) is by the definition of a sector: the arc is orthogonal to the circle radii $p_j x_j$ and $p_j x_{j+1}$. □

We now establish that in a planar layout of the regions involved in the reduction, the new regions do not overlap and are nested in the old regions.

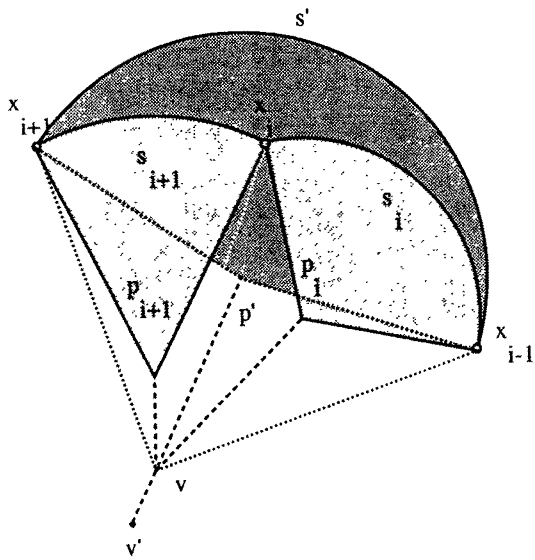

Fig. 14. Nesting of sectors.

Lemma 8.2 (Sector Nesting). *If S'_x does not overlap, then in the $S'_x \Rightarrow S_x$ transition, R, s_i, and s_{i+1} do not overlap each other, and $R \cup s_i \cup s_{i+1} \subset R' \cup s'$.*

Proof. Draw R, s_i, and s_{i+1} in the plane on top of $R' \cup s'$; the latter does not overlap by assumption. Recall that R is the hexagon $(v, x_{i-1}, p_i, x_i, p_{i+1}, x_{i+1})$ and R' is the quadrilateral $(v, x_{i-1}, p', x_{i+1})$ (see Fig. 14). Since R and R' have identical "inner" boundaries $x_{i-1}v \cup vx_{i+1}$, we only need to show that the "outer" boundary of $s_i \cup s_{i+1}$ falls inside the outer boundary of s' to establish the nesting. This follows from the reduction angles lemma, Lemma 7.1. As $\angle p'x_{i-1}v \geq \angle p_i x_{i-1} v$, the normal to $x_{i-1}p'$, which is tangent to s', falls outside the normal to $x_{i-1}p_i$, which is tangent to s_i. The same is true at x_{i+1}. Thus the boundary arc of s_i incident to x_{i-1}, and the boundary arc of s_{i+1} incident to x_{i+1}, both fall inside s' in the vicinity of x_{i-1} and x_{i+1}, respectively. Both of these arcs end at x_i. It therefore only remains to show that x_i falls inside the outer boundary of s'.

Recall that x_{i-1}, x_i, and x_{i+1} all fall on a circle C centered on v (by the definition of the reduction). Because p' by construction falls on the ray bisecting $\angle x_{i-1}vx_{i+1}$, C is inside $R' \cup s'$ between x_{i-1} and x_{i+1}. Therefore x_i falls inside the same region, and nesting is established.

This in conjunction with Lemma 8.1 shows that the boundary arcs of s_i and s_{i+1} are curves in the plane disjoint except at x_i. Since s_i and s_{i+1} are locally exterior to R by construction, R, s_i, and s_{i+1} do not overlap. □

9. Nonoverlap

Let $Q_x = S_x \cup (\bigcup_j s_j)$ be the "complex" consisting of the star unfolding with the sectors glued in at their common edges.

Theorem 9.1 (Nonoverlap). *The star unfolding augmented by the sectors, Q_x, does not overlap: S_x does not overlap itself, the sectors do not overlap each other, and the sectors do not overlap S_x.*

Proof. The proof is by induction.

Basis. As discussed in Section 6, the basis is a doubly covered triangle, $n = 3$, although we must consider both bounded and unbounded cases. We first discuss bounded triangles. Clearly, S_x itself does not overlap in the bounded case, for it is the union of three peels glued together at the single ridge vertex. Each sector is clearly exterior to S_x, and every pair of the three sectors are adjacent to one another, so Lemma 8.1 shows that the sectors do not overlap in the vicinity of their shared source images. Finally, it is easy to see that rays from $\bar{x}$ through x_j partition the plane into three regions each containing one sector interior, where $\bar{x}$ is the image on the plane of the common point to which each x_j maps if folded over the segment $p_j p_{j+1}$. See Fig. 15. Note that $\angle p_j \bar{x} x_j = \angle \bar{x} x_j p_j < \pi/2$, the sector boundary ∂s_j is orthogonal to $x_j p_j$ at x_j, and similarly ∂s_{j+1} is orthogonal to $x_j p_{j+1}$; hence the ray $\bar{x}x_j$ separates s_j and s_{j+1}. Thus the sectors do not overlap.

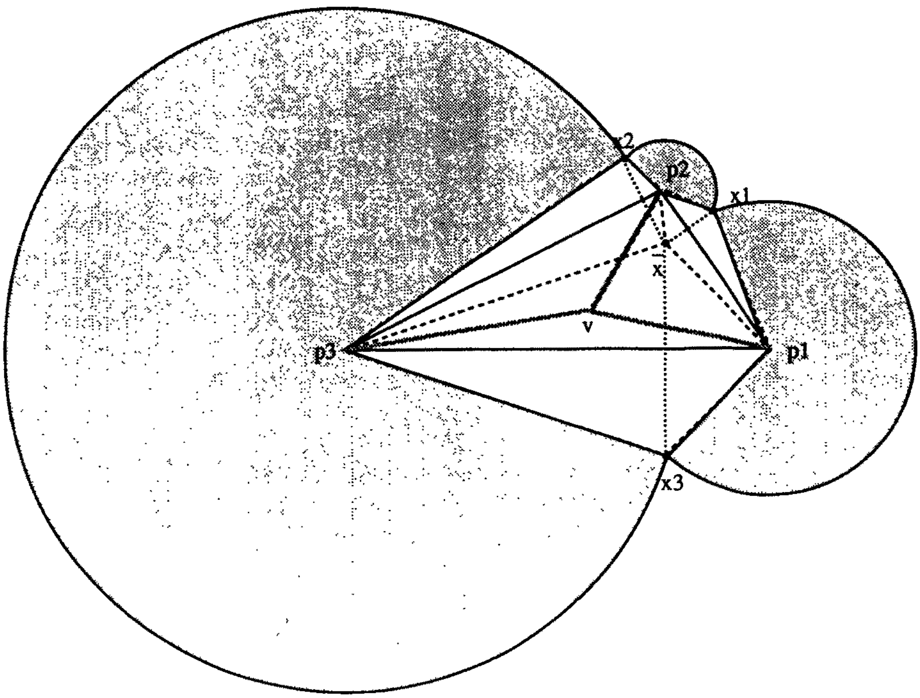

Fig. 15. Sectors in the base case.

For unbounded triangles, let p_1 and p_2 be the two corners as in Fig. 16. Define s' to be the third (unbounded) sector: it is the half-plane to the left of the directed line through x_0x_2, minus the vertical strip between x_0 and x_2. Just as in the bounded case, the rays from $\bar{x}$ through x_j, $j = 0, 1, 2$, separate the interiors of s_1, s_2, and s', the only difference being that $\angle \bar{x}x_0p' = \angle p'x_0\bar{x} = \pi/2$, and the rays through x_0 and x_2 lie along $\partial s'$.

General Step. Assume $Q' = S'_x \cup (\bigcup_j s_j)$ does not overlap by induction. This means, in particular, that $R' \cup s'$, which is just a subset of Q', does not overlap with $Q' - (R' \cup s')$. Now, by sector nesting (Lemma 8.2), $R \cup s_i \cup s_{i+1} \subset R' \cup s'$, so none of the changes made in the $S'_x \Rightarrow S_x$ transition cause overlap with $Q' - (R' \cup s')$. The portion added, $R \cup s_i \cup s_{i+1}$, does not overlap by Lemma 8.2. Therefore

$$Q_x = [Q' - (R' \cup s')] \cup (R \cup s_i \cup s_{i+1})$$

does not overlap. □

In particular, we have shown that S_x is a simple polygon.

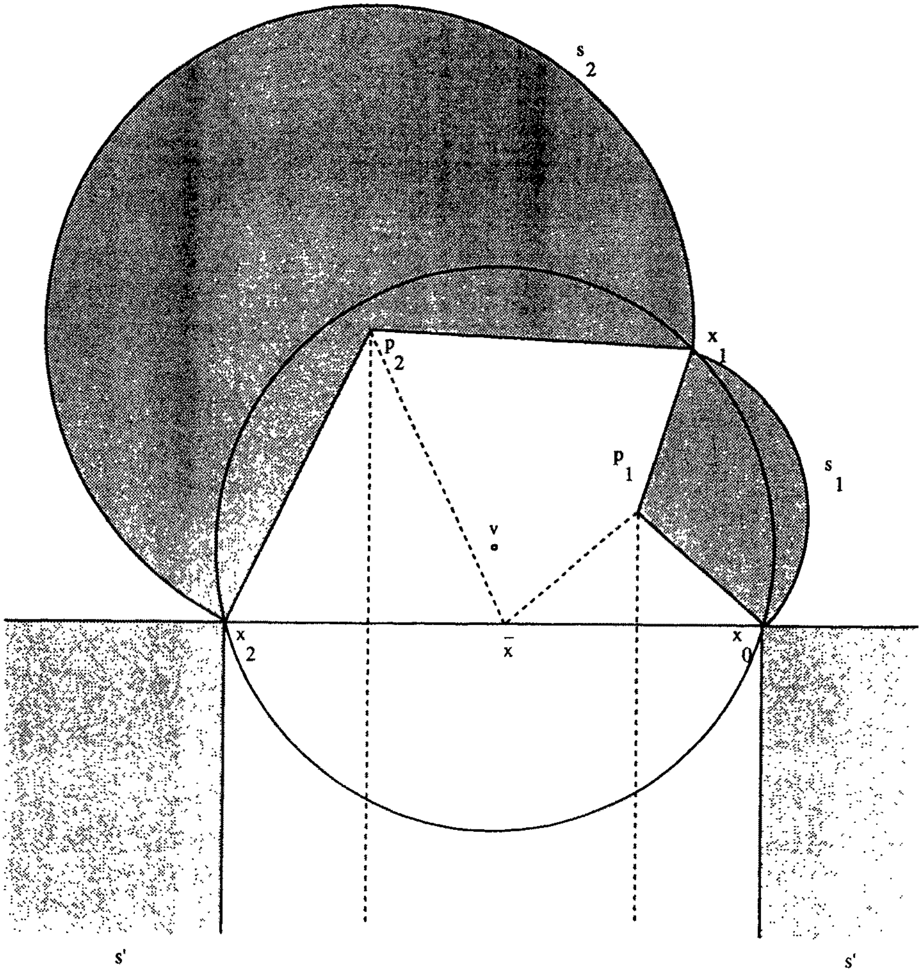

Fig. 16. Sectors in the special $n = 4$ case.

10. The Voronoi Property

We prove in this section that the ridge tree is a subset of the Voronoi diagram of the source images. Recall that X is the set of source images in the unfolding. Let $\mathscr{V}(X)$ be the Voronoi diagram of X, viewed as a set of points in a layout of S_x in the plane. We prove that $T_x = \mathscr{V}(X) \cap S_x$. We establish this by showing that a certain collection of "Voronoi disks" are empty of source images. Let D_y be the open disk centered on a point $y \in S_x$ with radius equal to the shortest path distance from x to y. We call D_y a *Voronoi disk*. The proof has the following outline:

(1) Q_x (S_x augmented by the sectors) *contains the union* of the Voronoi disks D_y for all ridge points $y \in T_x$.

(2) This containment implies that the Voronoi disks of all ridge points are *empty of source images.*
(3) This implies that the Voronoi disk D_y of *any point* $y \in S_x$ is empty of source images. Moreover, among points in S_x, only ridge points have more than one source image on the boundary of their Voronoi disk.
(4) The emptiness of the disks in turn implies the Voronoi property.

Steps (2)–(4) of the proof are easy, and we dispense with them prior to launching into the more difficult step (1).

(2) Suppose Q_x contains the Voronoi disks for all ridge points. The source images lie on the boundary of S_x, and the exterior arc bounding sector s_j begins and terminates at consecutive source images. As Q_x does not self-overlap (Theorem 9.1), the sources are on the boundary of Q_x. The emptiness of the disks follows immediately, as they are all open and contained in Q_x.

(3) Assume that the Voronoi disk of every ridge point is free of source images. Let $y \in S_x - T_x$. Suppose that y lies in the peel of x_j. By extending the shortest path $\pi(x, y)$ past y we obtain a point $z \in T_x$ with the property that all of $\pi(x, z)$ lies in the same peel. By assumption, D_z is free of source images and, by construction, x_j lies on the boundary ∂D_z of D_z. By definition of a Voronoi disk, D_y has radius $|yx_j|$ and thus lies inside D_z; moreover, $\partial D_z \cap \partial D_y = \{x_j\}$. So D_y is empty and its boundary contains exactly one source point, as claimed.

(4) Suppose now that the Voronoi disk for each point of S_x is empty of source images and no point of S_x outside T_x has more than one source image on the boundary of its Voronoi disk. This immediately implies that $T_x = \mathscr{V}(X) \cap S_x$, as $\mathscr{V}(X)$ is by definition the collection of points y in the plane for which the largest open disk centered at y and free of points of X touches two or more points of X.

The essence of the Voronoi property then reduces to (1) above, which we prove via induction based on the reduction used in the nonoverlap proof.

Lemma 10.1. *Q_x, the star unfolding augmented by the sectors, includes the union of all Voronoi disks for ridge points*:

$$\bigcup_{y \in T_x} D_y \subset Q_x.$$

Proof. The proof is by induction, using the standard reduction described in Section 3.

Basis. Again we partition the basis into the bounded and unbounded cases. The bounded case is illustrated in Fig. 15. The Voronoi disk for the single ridge vertex v passes through x_1, x_2, x_3. The disks for the corners determine the sectors. The disks corresponding to points lying along the ridge from v to a corner p_i are contained in $D_v \cup D_{p_i}$. It only remains to show that D_v and D_{p_i} lie in Q_x. The arc of ∂D_v between x_{i-1} and x_i must fall inside the arc ∂s_i of the sector s_i between the same source images, because the center of D_v is interior to the center of s_i. Therefore ∂D_v is never exterior to Q_x, so $D_v \subset Q_x$. Since part of D_{p_i} is covered by sector s_i, we need only show that $\partial D_{p_i} - s_i$ is inside Q_x. However, it is clear that D_v covers that portion of ∂D_{p_i}, and we just showed that D_v is inside Q_x; so $D_{p_i} \subset Q_x$.

The unbounded case, illustrated in Fig. 16, is handled similarly. In particular, $D_{p_i} \subset Q_x$ for $i = 1, 2$ for the same reasons as above, and $D_{p'}$, the unbounded Voronoi disk associated with the corner p' "at infinity," is the half-plane below x_0x_2, and so is covered by Q_x. Finally, ∂D_v lies inside the two bounded sectors and in the vertical strip between x_0 and x_2, and so $D_v \subset Q_x$.

General Step. Assume by the induction hypothesis that Q'_x includes the Voronoi disks for all ridge points in T'_x. We aim to show the same property holds true for Q_x, which is formed by removing $R' \cup s'$ from Q'_x, and adding $R \cup s_i \cup s_{i+1}$. Let us call a disk D_y "old" if $y \in T'_x - R'$ and "new" if $y \in T_x \cap R$. We divide the proof into showing that both the old disks and the new disks are included in Q_x. By Corollary 5.5, the old and the new disks together comprise all the relevant disks.

Old Disks. Notice that by Lemma 5.3, an old disk has the same radius in S_x as in S'_x, because for points in $\mathscr{P}' - R'$ the distance to x does not change.

Since S_x only differs from S'_x in the R and R' regions, $Q_x = Q'_x - \Delta$, where $\Delta = (R' \cup s') - (R \cup s_i \cup s_{i+1})$ (see Lemma 8.2). This region is illustrated in Fig. 17. So our goal is to show that no old disk intersects Δ; for if one did, it would not be contained in Q_x. We approach this by partitioning the plane into regions, considering old disks with centers in the various regions, and showing for each region that no disk could intersect Δ.

To define the partition, first draw the directed line $L = vp'$, as shown in Fig. 17. We prove the property for disk centers lying to one side of this line, say the side containing x_{i-1}; the other side is analogous. To this side of L, we partition the plane into four regions, closed along L. Let M be the ray $p'x_{i-1}$. The four regions are:

1. R',
2. s',
3. A: left of M, excluding $R' \cup s'$,
4. B: right of M, excluding $R' \cup s'$.

Disks whose centers are in R' are not old disks by definition. Disk centers in s' cannot lie on T'_x, since $T'_x \subset S'_x$, and s' is exterior to S'_x by Theorem 9.1. Any disk D_y with $y \in A$ could only intersect Δ by intersecting the arc boundary of s'. This violates nonoverlap, since $D_y \subset Q'_x$ by the induction hypothesis.

The remaining case is a disk D_y with $y \in B$. We claim D_y could only intersect Δ by including $x_{i\pm 1}$, again contradicting the induction hypothesis. Let $C_y = \partial D_y$ and $C_v = \partial D_v$. This claim is proved by examining the relationship between C_y and C_v. Recall that C_v has x_{i-1}, x_i, x_{i+1} on its boundary, and so Δ lies just outside it. Let α be the arc of C_v from x_{i-1} counterclockwise to x_{i+1}. See Fig. 17.

1. $D_y \subseteq D_v$. Then D_y does not intersect Δ.
2. D_y intersects Δ, but C_y does not intesect α. Then it must be that $D_y \supset \alpha$, and therefore D_y includes x_{i-1} and x_{i+1}.
3. D_y intersects Δ, and C_y intersects α. Consider two further cases.
   (a) C_y intersects α once. Then D_y must include either x_{i-1} or x_{i+1}. (This is the case illustrated in Fig. 17.)

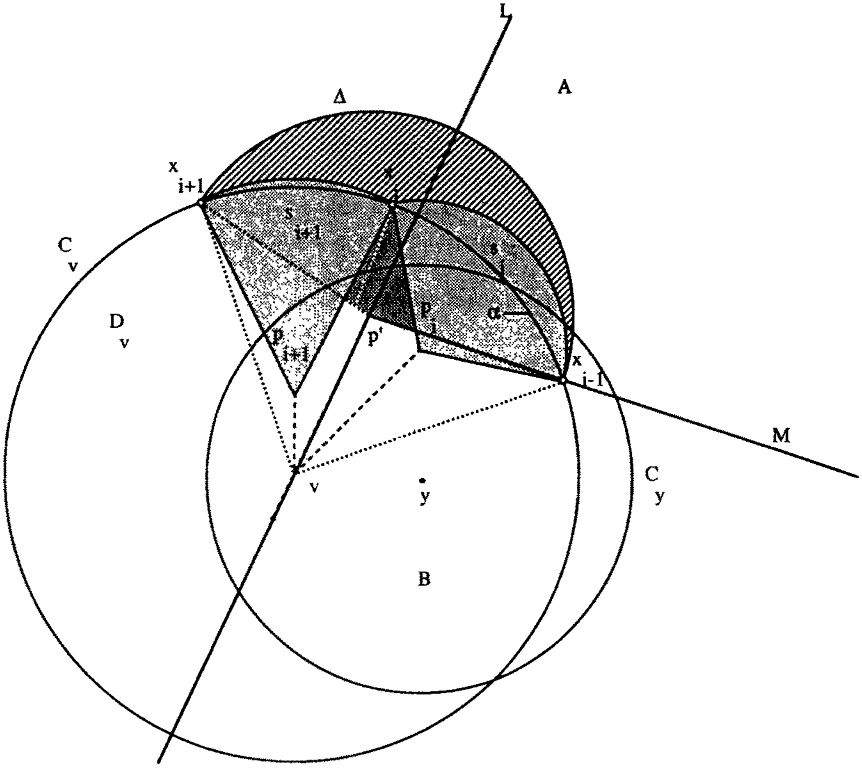

Fig. 17. Old disks are contained in Q_x.

(b) C_y intersects α twice. If the radius of C_y is larger than that of C_v, D_y includes both x_{i-1} and x_{i+1}. So it must be that the radius of C_y is smaller than that of C_v. However, now notice that y must lie in the wedge $\angle x_{i-1}vx_{i+1}$, and that B lies outside this wedge by definition.

New Disks. The new disks are those whose centers are on the two ridges vp_i and vp_{i+1}. We only consider the former ridge, as the latter is symmetric. For any $y \in vp_i$ it is clear that $D_y \subset D_{p_i} \cup D_v$. Since D_v is an old disk, $D_v \subset Q_x$. It remains to prove that $D_{p_i} \subset Q_x$. Part of the boundary of Q_x is an arc of D_{p_i} and D_{p_i} lies on the correct side of that boundary. So it remains to see that $\partial D_{p_i} - s_i$ is inside Q_x. It is, as it is inside $D_v \subset Q_x$. Thus we have shown that $D_y \subset Q_x$. □

Finally we may claim the second main result of this paper:

Theorem 10.2 (Voronoi Property). *The ridge tree is the portion of the Voronoi diagram of the source images that lies inside the star unfolding*:

$$T_x = \mathscr{V}(X) \cap S_x.$$

11. General Convex Surfaces

There is every reason to expect that our main theorems hold true for arbitrary convex surfaces as well as for polytopes. In this section we define analogs for the key geometrical concepts used in the theorems and formulate several conjectures.

The analog of the ridge tree is the *cut locus* [K]: the locus of points with two or more distinct shortest paths to x. The cut locus for a point x on a sphere is a single point, antipodal to x. The cut locus for a point on the rim of a cylinder is sketched in Fig. 18.

Clearly, the star unfolding cannot be defined via cuts to vertices. We choose to define it from a development of the cut locus. A curve on a smooth surface is *developed* in the plane by rolling the surface without slippage so that the curve is the point of contact: the points of contact in the plane constitute the development of the curve. This can be generalized to define the development of the cut locus.

Conjecture 11.1. *The cut locus develops in the plane without self-intersection.*

That the ridge tree develops or unfolds without self-intersection is a consequence of nonoverlap, Theorem 9.1.

We now define the star unfolding of a surface. First, develop the cut locus. Second, from each point y of the cut locus, draw segments in the plane corresponding to all the shortest paths from the source x that are incident to y. Draw each segment to have the length of the corresponding shortest path, and to make the same angle at the point y with the cut locus, as it does on the surface of $\mathscr{P}$. The star unfolding is this particular layout of all the shortest paths from x on $\mathscr{P}$. An example is shown in Fig. 19, which depicts the developed cut locus from Fig. 18, and a number of segments out to images of x.

Conjecture 11.2. *The star unfolding of a smooth surface is a simple closed region of the plane, whose boundary is the locus of all source images.*

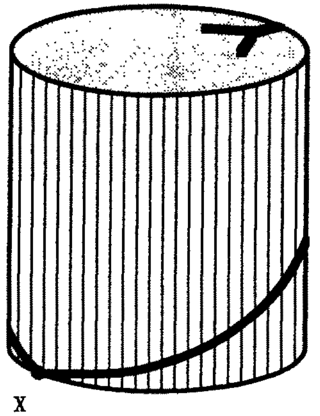

Fig. 18. The cut locus on a cylinder, radius $r = 1$, height $h = 2$, with respect to a point x on the rim.

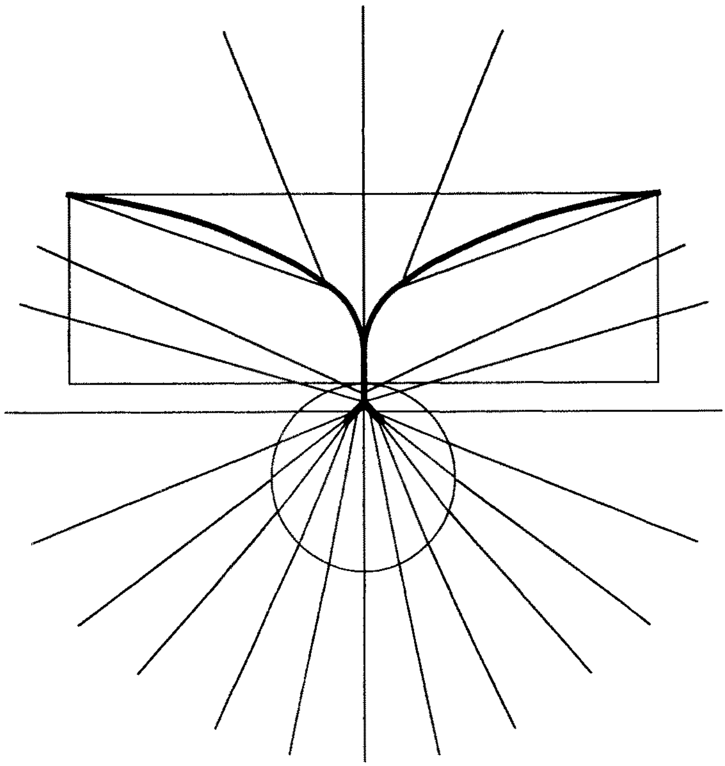

Fig. 19. Star unfolding of the cylinder shown in Fig. 18. The dark curves are the cut locus, the straight lines various shortest paths to x.

This is the generalization of Theorem 9.1.

Finally we conjecture the analog of the Voronoi property, Theorem 10.2:

Conjecture 11.3. *The developed cut locus is the medial axis of the locus of the source images.*

The "medial" or "symmetric" axis of a Jordan curve is the locus of centers of interior disks that meet the curve in more than one point [L].

12. Algorithmic Consequences

The primary consequence of our results is that it is now an easy matter to construct the ridge tree, formerly an object of formidable conceptual complexity: find shortest paths to all corners, build the star unfolding in the plane, and compute the conventional Voronoi diagram of the set of source images.[9] In particular, our

[9] This is how Fig. 2 was produced.

results now justify Chen and Han's simple and efficient quadratic algorithm for single-source shortest path queries [CH].

Second, in [AAOS2], an algorithm is presented for computing the exact set of edge sequences in $O(n^7 \log n)$ time. An edge sequence is the list of edges crossed by a shortest path; they are used for finding shortest paths amidst polyhedra [SS]. A major factor in the algorithm's time complexity is the number of combinatorial changes the ridge tree may undergo as the source moves along a straight line without crossing a ridge of any corner. The only bound proved in [AAOS1] was $O(n^4)$. However, knowing by Theorem 10.2 that the ridge tree is actually a subgraph of a Voronoi diagram, we may obtain an $O(n^3)$ bound on the number of changes using lower-envelope theory. This observation simplifies the algorithm and its analysis, and we believe it will decrease the time complexity by a factor of $O(n)$; this work is still in progress [AAOS2].

Third, the $O(n^{10})$ algorithm of [AAOS1] for computing the "geodesic diameter" of a polytope (the maximum possible separation between two points on its surface) may be improved by our results in two ways. At the center of $O(n^9)$ iterations in that algorithm is a linear-time calculation to disambiguate possible overlap of the star unfolding, and an $O(n)$ visibility calculation. The first is obviated by our nonoverlap theorem (Theorem 9.1) and the second by the Voronoi property (Theorem 10.2). The result is an $O(n^9 \log n)$ algorithm for the diameter.

These algorithmic consequences will be developed in [AAOS2].

Acknowledgments

We thank Pankaj K. Agarwal and Catherine Schevon, our coauthors on [AAOS1] and [AAOS2], whose ideas have contributed significantly to this paper. We thank Joseph Malkevitch for several suggestions, and the referees for insightful comments. The second author also thanks Robert Currier, Vladimir Oliker, and Andrejs Treibergs for pointers to the differential geometry literature.

Appendix 1

We cite two theorems equivalent to Aleksandrov's Theorem 4.2. The first is from Pogorelov's book [P], in a chapter in which he summarizes Aleksandrov's result in more modern language (and in an English translation).

Theorem A.1. "*Any convex polyhedral metric given on a sphere or on a manifold homeomorphic to a sphere, is realizable as a closed convex polyhedron* (*possibly degenerating into a doubly covered plane polygon*)" [P, p. 20].

A *convex polyhedral metric* is a two-dimensional manifold, each of whose points has a neighborhood isometric to a circular cone (which may degenerate into a plane). That Aleksandrov's nets are convex polyhedral metrics is a consequence of the "gluing theorem" [P, p. 33]. This is a general theorem which, when applied to our special case, says that if one forms a net in Aleksandrov's sense, such that the face angles at each corner sum to $\leq 2\pi$, then the resulting complex has an intrinsic metric with positive curvature. Since gluing polygons with this angle restriction guarantees that every point will have a neighborhood isometric to a cone, these nets define convex polyhedral metrics. Then Theorems A.1 and 4.2 can be seen to be equivalent.

A second exposition of Aleksandrov's theorem may be found in Buseman's book [B]. His phrasing of the theorem is as follows.

Theorem A.2. "*A polyhedral metric with non-[negative]*[10] *curvature on the sphere can be realized as one, and up to motions only one (possibly degenerate), polyhedron*" [B, p. 128].

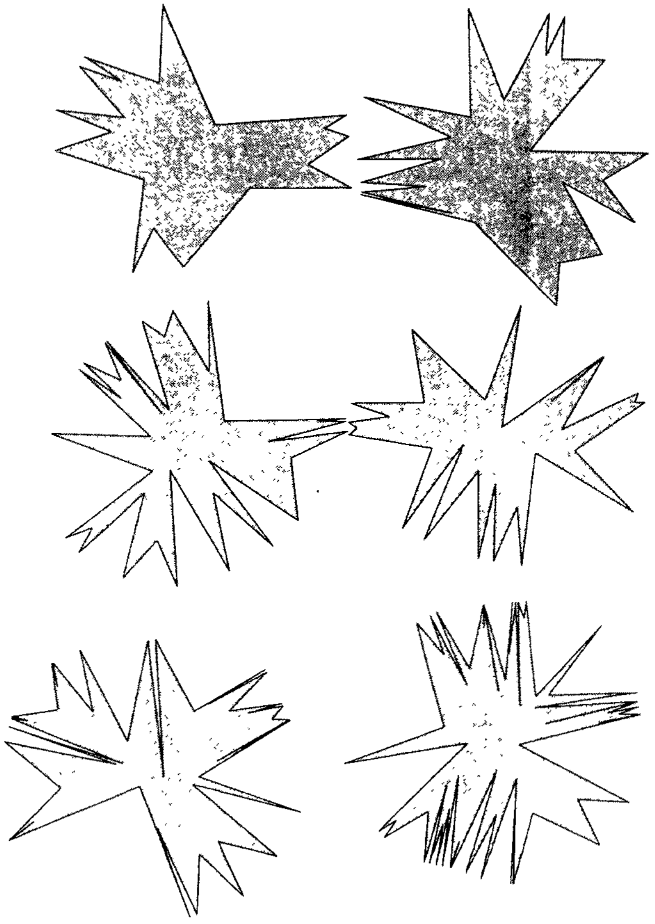

Fig. 20. Star unfoldings of six randomly generated polytopes.

[10] He actually writes "non-positive," but as the proof makes clear, this is a typographical error.

Appendix 2

Some additional examples of star unfoldings are shown in Fig. 20.

References

[AAOS1] P. K. Agarwal, B. Aronov, J. O'Rourke, and C. Schevon. Star unfolding of a polytope with applications. In J. R. Gilbert and R. Karlsson, editors, *Proc. of 2nd Scandanavian Workshop on Algorithm Theory*, pages 251–263. Lecture Notes in Computer Science, Vol. 447. Springer-Verlag, Berlin, 1990.

[AAOS2] P. K. Agarwal, B. Aronov, J. O'Rourke, and C. Schevon. The star unfolding of a polytope. Technical report, Smith College, 1992. Full version of [AAOS1]; currently under revision.

[A1] A. D. Aleksandrov. *Die Innere Geometrie der Konvexen Flächen*. Mathematische Lehrbücher und Monographien. Akademie-Verlag, Berlin, 1955.

[A2] A. D. Aleksandrov. *Konvexe Polyeder*. Mathematische Lehrbücher und Monographien. Akademie-Verlag, Berlin, 1958.

[AZ] A. D. Aleksandrov and V. A. Zalgaller. *Intrinsic Geometry of Surfaces*. American Mathematical Society, Providence, RI, 1967. Translation of the 1962 Russian original.

[B] H. Buseman. *Convex Surfaces*. Wiley-Interscience, New York, 1958.

[CH] J. Chen and Y. Han. Shortest paths on a polyhedron. In *Proc. of 6th ACM Symposium on Computational Geometry*, pages 360–369, 1990.

[H] M. Henle. *A Combinatorial Introduction to Topology*. Freeman, San Francisco, CA, 1979.

[K] S. Kobayashi. On conjugate and cut loci. In S. S. Chern, editor, *Studies in Global Geometry and Analysis*, pages 96–122. Mathematical Association of America, Washington, DC, 1967.

[L] D. T. Lee. Medial axis transformation of a planar shape. *IEEE Trans. Pattern Anal. Mach. Intell.*, 4:363–369, 1982.

[P] A. V. Pogorelov. *Extrinsic Geometry of Convex Surfaces*. Translations of Mathematical Monographs, Vol. 35. American Mathematical Society, Providence, RI, 1973.

[R] R. Rasch. Shortest paths along a convex polyhedron. Manuscript, Berlin, 1990.

[SS] M. Sharir and A. Schorr. On shortest paths in polyhedral spaces. *SIAM J. Comput.*, 15:193–215, 1986.

Received March 11, 1991.