Unfolding Some Classes of Orthogonal Polyhedra*

Therese Biedl† Erik Demaine‡§ Martin Demaine‡ Anna Lubiw‡§
Mark Overmars¶ Joseph O'Rourke‖** Steve Robbins†† Sue Whitesides†§††

Abstract

In this paper, we study unfoldings of orthogonal polyhedra. More precisely, we define two special classes of orthogonal polyhedra, *orthostacks* and *orthotubes*, and show how to generate unfoldings by cutting faces, such that the resulting surfaces can be flattened into a single connected polygon.

1 Introduction

An *unfolding* of a polyhedron is a cutting of the polyhedron's surface so that the surface can be flattened into a single simple polygon that does not overlap itself. Finding unfoldings is a classic problem. Not only does it have connections with origami, but it also has potential industrial applications, because an unfolding allows one to construct a desired physical shape by folding and gluing a polygonal cut-out of a stiff sheet of material.

Unfoldings can be classified by whether all cuts are along edges of the polyhedron (*edge cuts*), or whether cuts across faces are allowed as well. It is an open problem to determine whether every convex polyhedron has an unfolding with only edge cuts [CFG93, pp. 73–75][Fuk97][She75]. Without this restriction there are at least two different methods for finding an unfolding of a given convex polyhedron [AO92].

Unfolding nonconvex polyhedra seems relatively unexplored. For example, it is open whether there is a nonconvex polyhedron with convex faces that has no unfolding, even if only edge cuts are allowed. Here we examine the case of an *orthogonal polyhedron*, i.e., a simple three-dimensional polyhedron each face of which is perpendicular to one of the coordinate axes.

We will look at two classes of orthogonal polyhedra, "orthostacks" and "orthotubes", defined formally in Sections 2 and 3. Intuitively, they can be viewed as what are called *generalized cylinders* in the computer graphics community: a curve along which a cross-section is swept.

*The research reported here was initiated at the International Workshop on Wrapping and Folding, co-organized by Anna Lubiw and Sue Whitesides, at the Bellairs Research Institute of McGill University, January 31 - February 6, 1998. We would like to thank Hazel Everett, Sylvain Lazard, Ileana Streinu, and Godfried Toussaint for useful discussions.

†School of Computer Science, McGill University, 3480 University Street #318, Montreal, Quebec H3A 2A7, Canada. {therese,stever,sue}@cs.mcgill.ca

‡Department of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. {eddemaine, mldemaine, alubiw}@waterloo.ca.

§Supported by NSERC.

¶Department of Computer Science, Utrecht University, 3508 Utrecht, The Netherlands. markov@cs.ruu.nl.

‖Department of Computer Science, Smith College, Northampton, MA 01063, USA. orourke@cs.smith.edu.

**Supported by NSF.

††Supported by FCAR.

This curve gives these polyhedra a somewhat "linear" nature, which is what we exploit in our unfoldings.

In the case of orthotubes, the curve of the generalized cylinder is an arbitrary non-self-intersecting orthogonal curve, and the cross-section is a rectangle that changes only near bends of the curve. Orthotubes can "corkscrew" through space, and, because we allow the curve to be closed, they can even form cycles and knots. See Figure 8.

Orthostacks, on the other hand, are much more restricted in the choice of the curve, but much more general in the choice of the cross-section. For an orthostack, the curve is parallel to the x-axis. The cross-section consists of simple connected orthogonal polygon that change only at finitely many values, though not so as to disconnect the orthostacks. In other words, orthostacks are formed by "stacking up" extrusions of simple connected orthogonal polygons. In particular, they contain the class of all orthogonally convex polyhedra.

We prove that all orthostacks and orthotubes can be unfolded. Throughout this paper, we consider only unfoldings without overlap. Our cuts are not always along edges; indeed, we will provide examples where this is not possible.

2 Unfolding Orthostacks

2.1 Definitions

We need the following notation: An *x-plane* is a plane that is perpendicular to the x-axis; *y-planes* and *z-planes* are defined similarly. An *x-face* is a face of the orthogonal polyhedron that is part of an x-plane; *y-faces* and *z-faces* are defined similarly. An *x-edge* is an edge of the orthogonal polyhedron that is parallel to the x-axis; *y-edges* and *z-edges* are defined similarly. An *x-cross-section C_x* of a polyhedron is the intersection of the polyhedron with an x-plane.

We define an *orthostack* as an orthogonal polyhedron every x-cross-section of which is a simple connected polygon.

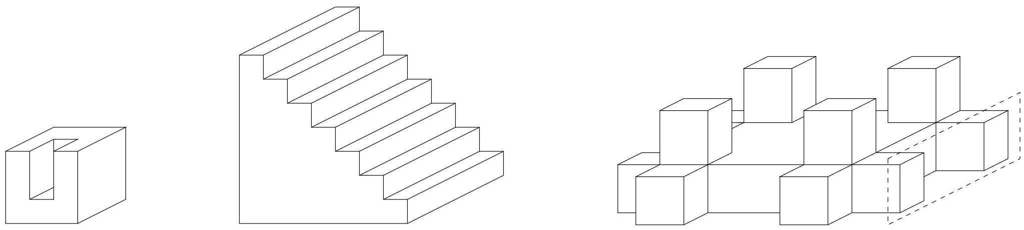


Figure 1: The two orthogonal polyhedra to the left are orthostacks. The one to the right is not because the cross-section at the indicated plane has two connected components.

As an x-plane is swept from smaller to larger values of its x-coordinate, the cross-section of an orthostack changes at only finitely many x-values. Denote these values $x_0 < x_1 < \ldots < x_k$. Thus, for all values $x_{i-1} < x_\alpha < x_i$ the cross-section with x-coordinate x_α is the same; we will call this cross-section C_i. By definition of an orthostack, C_i is a simple connected polygon. For later ease of notation, define $C_0 = C_{k+1} = \emptyset$. The *ith band* is the intersection of the surface of the orthostack with the region $\{(x_\alpha, y, z) : x_{i-1} < x_\alpha < x_i\}$. See Figure 2.

2.2 Outline of the Unfolding

The basic idea for unfolding an orthostack is to cut it into the bands defined above, cut each band to make a strip, and to spread these strips out in the plane in order of increasing x-coordinate.

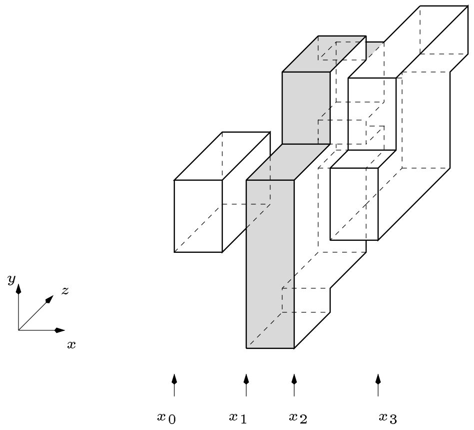


Figure 2: An orthostack and its cross-section x_i values. The visible part of the second band is shown shaded.

Each band consists of parts of y-faces and z-faces, and all bands together cover all y-faces and all z-faces. Thus, only the x-faces are missing, and we attach these faces between the strips in such a way that everything remains overlap-free and connected.

As we will see, the place to cut the ith band to form a strip depends on the previous and the next band. Because we cannot necessarily meet the requirements imposed by the previous and the next band simultaneously, we will first cut each band in two with an x-plane. More precisely, for $1 \leq i \leq n$, let the *left ith band* L_i be the intersection of the surface of the orthostack with the region $\{(x_\alpha, y, z) : x_{i-1} < x_\alpha < \frac{1}{2}(x_{i-1} + x_i)\}$, and let the *right ith band* R_i be the intersection of the surface of the orthostack with the region $\{(x_\alpha, y, z) : \frac{1}{2}(x_{i-1} + x_i) \leq x_\alpha < x_i\}$. For later ease of notation, define $R_0 = L_{k+1} = \emptyset$.

Recall that C_i is the cross-section of the orthostack with a plane through $x_i - \epsilon$, for a small constant $\epsilon > 0$, and thus it is the shadow in x-direction of the band R_i. Similarly, C_{i+1} is the shadow of the band L_{i+1}. For $0 \leq i \leq k$, define $D_i = C_i \oplus C_{i+1}$, the symmetric difference of polygons C_i and C_{i+1}. Note that D_i is nonempty because we only take cross-sections at changes. The connected pieces of D_i correspond to the x-faces with x-coordinate x_i. Such an x-face looks "to the right" (i.e., toward $+x$) if the corresponding piece of D_i is inside C_i and outside C_{i+1}, and "to the left" (toward $-x$) if it is outside C_i and inside C_{i+1}. See Figure 3.

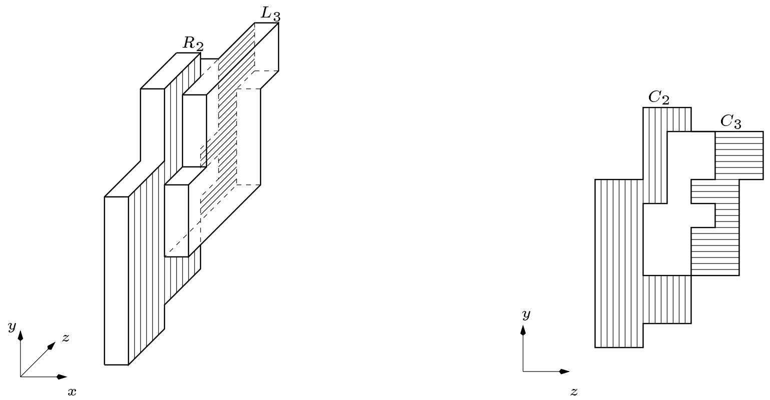


Figure 3: The transition between R_2 and L_3 of the above example, shown once from the front and once from right. The shaded area is D_2.

Note that $\bigcup_{i=0}^{k}(R_i \cup D_i \cup L_{i+1})$ covers all faces of the orthostack. Thus, to unfold an orthostack, it suffices to show: (1) how to form one connected component of $R_i \cup D_i \cup L_{i+1}$ by cutting each band into a rectangular strip, cutting D_i into pieces, and attaching them without overlap at appropriate places on the strips with one of the pieces connecting the two strips; and (2) how to connect these $k+1$ components without overlap.

2.3 Attaching D_i to R_i and L_{i+1}

In the following, fix an index $i \neq 0, k$. We cut the symmetric difference D_i into a collection of rectangles $P_1, \ldots, P_l$, by extending every y-edge of D_i in both directions until it hits a z-edge. See Figure 4.

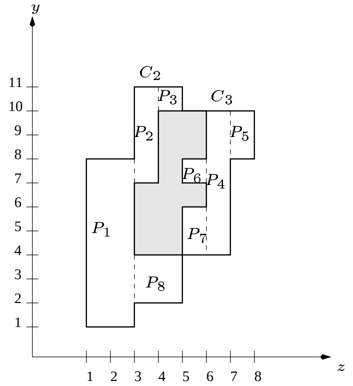


Figure 4: D_i is partitioned into rectangles $P_1, \ldots, P_l$ by cutting along extensions of y-edges. $I = C_i \cap C_{i+1}$ is shown shaded.

To connect the bands R_i and L_{i+1}, we proceed as follows. Let I be the intersection of C_i and C_{i+1}; this intersection is non-empty because the orthostack is connected. Let e be a z-edge of I with the largest y-coordinate. In Figure 4, e is the edge between endpoints $(4, 10)$ and $(6, 10)$.

Edge e belongs to the boundary of C_i or C_{i+1} or both. We apply one of two methods to connect the bands, depending on whether e belongs to both boundaries or only one. If, as in Figure 4, part of e belongs to both boundaries, and part of e belongs to only one boundary, then either of the following methods can be employed.

Method (1): If part of e belongs to both C_i and C_{i+1}, then let $[z_{min}, z_{max}]$ be a maximal interval of overlap between C_i and C_{i+1}. Since C_i and C_{i+1} are the shadows of R_i and L_{i+1}, respectively, it follows that in this interval, the two bands R_i and L_{i+1} run in parallel and attach to each other. In Figure 4, we have $z_{min} = 5$ and $z_{max} = 6$.

Method (2): If not all of e belongs to both C_i and C_{i+1}, then e is incident to the symmetric difference D_i, and therefore to at least one of the rectangles, say to P_j. This rectangle will be called the *connecting bridge*. The top and bottom of P_j are parts of edges of C_i and/or C_{i+1}. Because e is the topmost edge of I, the bottom of P_j is incident to I, and the top of P_j is not incident to I; therefore the top and the bottom of P_j cannot be edges of the same polygon. Since C_i and C_{i+1} are the shadows of R_i and L_{i+1}, respectively, it follows that either the bottom of P_j attaches to R_i and the top of P_j attaches to L_{i+1}, or, vice versa, the bottom of P_j attaches to L_{i+1}, and the top of P_j attaches to R_i. Let z_{min} and z_{max} be the minimum and maximum z-coordinate of P_j, respectively. In Figure 4, we have $P_j = P_3$, $z_{min} = 4$ and $z_{max} = 5$.

We will treat the first method as a special case of the second method, by considering the interval $[z_{min}, z_{max}]$ as a zero-height rectangle P_j. Cut R_i by extending the y-edge that forms the z_{max}-boundary of P_j across R_i. Cut L_{i+1} by extending the y-edge that forms the z_{min}-boundary of P_j across L_{i+1}. Each cut band now unfolds into a rectangular strip; place this strip in the xy-plane with unchanged x-coordinates such that the inside of the strip points towards the $+z$-direction. Abusing notation, we call these flattened strips R_i and L_{i+1} as well. Rectangle P_j attaches to the flattened strip R_i at the top (y_{max}) end, and to the flattened strip L_{i+1} at the bottom (y_{min}) end. Thus, after shifting one of the strips in y-direction, we can glue P_j to both strips, connecting them.

The half-infinite rectangle to the right ($+x$) of R_i and the half-infinite rectangle to the left ($-x$) of L_{i+1} will be used to place all other pieces of D_i. More precisely, each rectangle $P_1, \ldots, P_l$ (except the connecting bridge P_j) is glued to the band at its y_{max} end. The half-infinite rectangle from both bands is empty and each piece is a rectangle, therefore no intersections are possible. See Figure 5.

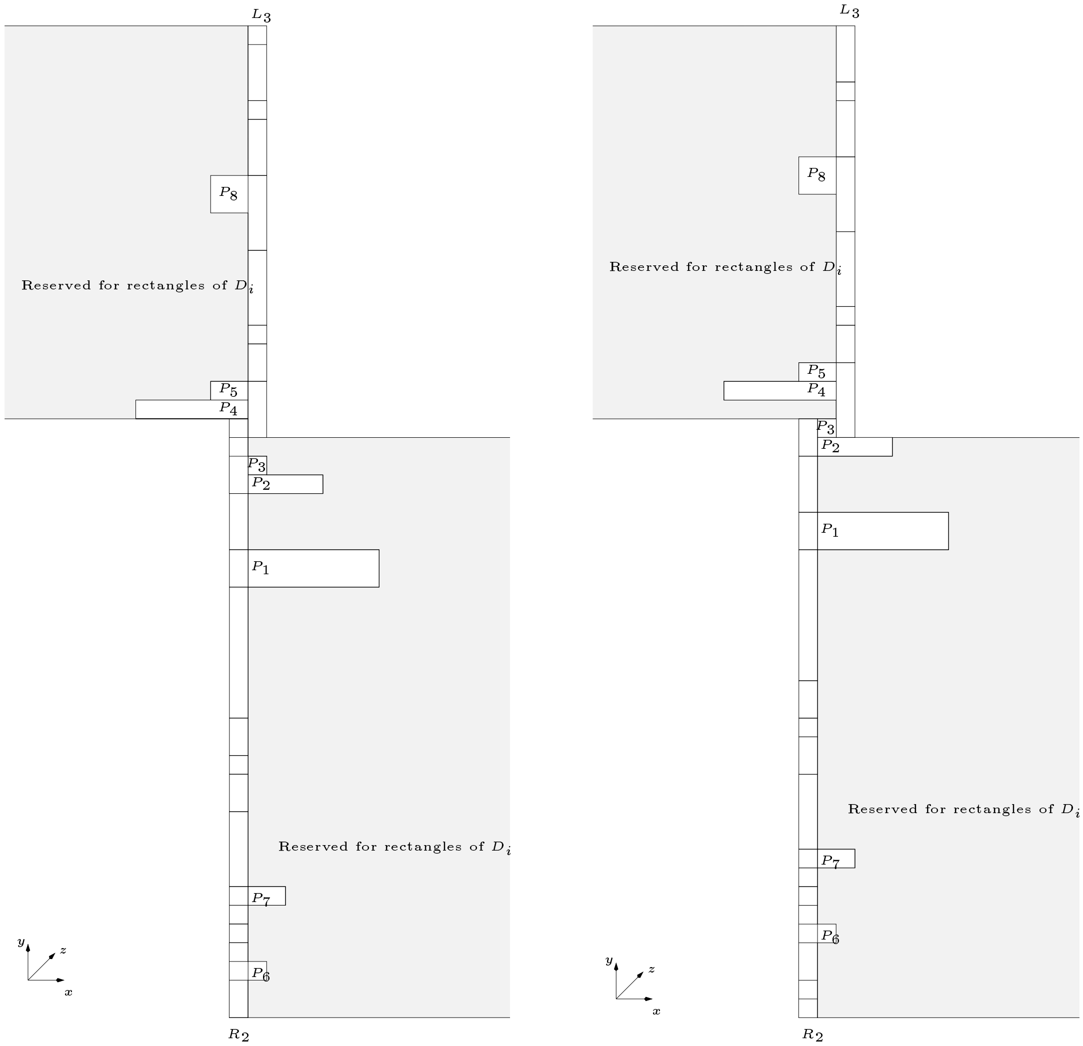


Figure 5: Splitting the symmetric difference, and attaching it to the strips. We show this using Method (1) on the left, and Method (2) on the right.

Thus, we have shown how to unfold $R_i \cup D_i \cup L_{i+1}$, $1 \leq i \leq k-1$. We can handle $D_0 \cup L_1$ and $R_k \cup D_k$ similarly. We cut L_1 at the same place where we cut R_1, cut D_0 into rectangles, and attach these rectangles in the half-infinite rectangle to the left of L_1. We cut R_k at the same place where we cut L_k, cut D_k into rectangles, and attach these rectangles in the half-infinite rectangle to the right of R_k.

2.4 Connecting the layouts of $R_i \cup D_i \cup L_{i+1}$

All that remains to show is how to attach band L_i to band R_i, $i = 0, \ldots, k$. The bands L_i and R_i have been cut at possibly different places to form strips. If the cuts are actually at the same place, then the strips can simply be rejoined. If the cuts are at different places, there are two ways to rejoin the strips, with R_i staggered above or below L_i, respectively. We choose the former for all $i = 1, \ldots, k$, lining up the top of the strip L_i with the appropriate place in strip R_i. See Figure 6, particularly the middle pair of strips.

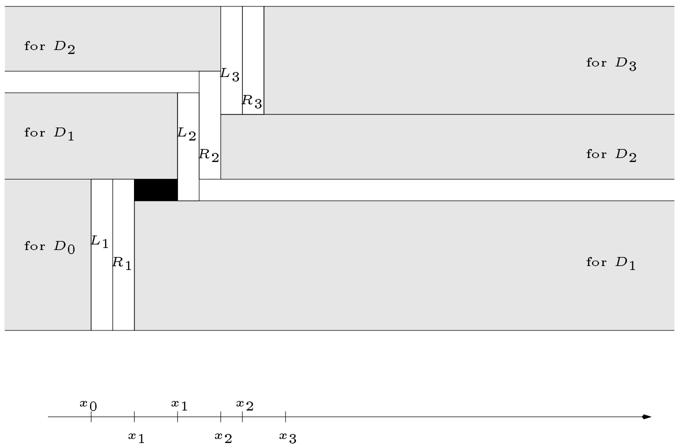


Figure 6: We glue the pieces together such that the L_i is not above R_i for $1 \leq i \leq k$. The black rectangle is a connecting bridge. (The figure is not to scale with the previous example.)

Thus we have achieved an unfolding of the orthostack. To estimate the complexity of this unfolding, we would like to obtain a bound on the number of vertices in the resulting polygon. Every symmetric difference D_i contains at least one rectangle, so we can charge the up to 8 vertices of $R_i \cup L_{i+1}$ to this rectangle. Thus the number of vertices of the unfolded polygon is $O(r)$, where r is the total number of rectangles. All the rectangles obtained by partitioning D_i lie in the x-plane with x-coordinate x_i, and each rectangle P_k is incident to at least one vertex v of the orthostack in that plane (because P_k's border contains an edge extension cut, and each edge includes two vertices). Charge each rectangle P_k to one of the vertices v on its boundary. No vertex v can be incident to more than two rectangles (cf. Figure 4), and so no vertex receives more than two charges. Thus the number of rectangles from D_i is $O(n_i)$, where n_i is the number of vertices of the orthostack in the x-plane with x-coordinate x_i. The total number of vertices of the polygon therefore is $O(r) = \sum_i O(n_i) = O(n)$.

Thus, we have proved the following theorem:

Theorem 1 *The surface of any orthostack may be cut and unfolded flat to a planar orthogonal simple polygon. Each cut is perpendicular to one of the coordinate axes. For an orthostack of n vertices, the polygon has $O(n)$ vertices.*

It would be a reasonable restriction to require that all cuts lie in a plane perpendicular to a coordinate axis and containing at least one vertex of the polyhedron. However, our construction splits the ith band into L_i and R_i with a cut that violates this restriction. We leave as an open problem finding an unfolding under this more stringent condition.

2.5 Unfolding with edge cuts is impossible

In our construction, we used cuts that are not edge cuts. We now show that this is necessary for the orthostacks in Figure 7, which by Theorem 1 can be unfolded if we allow cuts across faces.

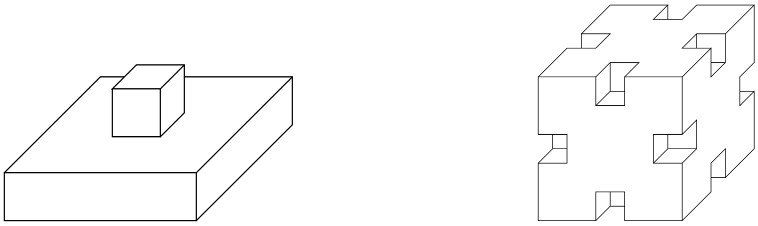


Figure 7: Two orthostacks that cannot be unfolded with edge cuts.

First, consider the left polyhedron of Figure 7. In an unfolding with only edge cuts, all faces are left intact. Therefore, the unfolding of the little cube must lie inside the hole of the annulus that is the top face of the big box. This is impossible, because the hole does not have enough area for the surface of the small cube. Hence, there is no unfolding with edge cuts only.

A more intriguing example is the right polyhedron of Figure 7, which is a cube with small "bites" taken out of every edge. Here, every face is a simple orthogonal polygon without holes. Nevertheless, this polyhedron has no unfolding with edge cuts only. An unfolding using only edge cuts provides a tree of the faces. Consider two of the large faces that are closest in this tree. They must either be joined directly, or via the faces of the bite between them. The first possibility leaves no area for the faces of the bite, and the second one either leads to overlap between faces or again leaves no area for the bite. So both possibilities are ruled out, or in other words, there is no unfolding with edge cuts.

An even more interesting example would be an orthogonal polyhedron where all faces are orthogonally convex, and nevertheless there is no unfolding with only edge cuts. We leave this as an open problem.

3 Unfolding Orthotubes

An *orthotube* is defined as a (possibly cyclic) sequence of axis-parallel *blocks* (i.e., rectangular boxes) $B_0, \ldots, B_{k-1}$ such that for $i = 0, \ldots, k-1$, $B_i \cap B_{i+1}$ is a 2-dimensional face of both B_i and B_{i+1}, and such that $B_i \cap B_j$, $j \neq i-1, i+1$, is either empty, or a 1-dimensional or 0-dimensional face contained in $B_i \cap B_{i-1}$ or $B_i \cap B_{i+1}$. Here, we define $B_{-1} = B_k = \emptyset$ if the orthotube is not cyclic, and $B_{-1} = B_{k-1}$, $B_k = B_0$ otherwise. See Figure 8 for some examples of orthotubes. Notice that neither the left hand nor the middle orthotube, an "orthogonalized trefoil knot", is an orthostack.

As we will show now, any orthotube has an unfolding. As opposed to orthostacks, we will only use cuts that lie in a plane perpendicular to a coordinate axis and containing at least one vertex of the polyhedron.

Consider an orthotube consisting of the blocks $B_0, \ldots, B_{k-1}$. We will assume that the orthotube is cyclic; the other case will be treated later. Let the *surface of a block* B_i be the union of all 2-dimensional faces that are common to B_i and the orthotube. With this definition, the surface of block B_i is the surface of a block with two faces missing; we call these missing faces *holes*. There are two classes of blocks: the *tube blocks*, where the two holes are on opposite sides, and the *corner blocks*, where this is not the case.

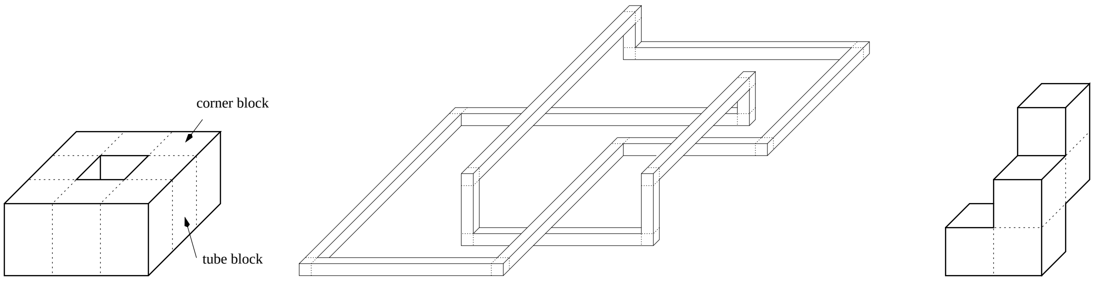


Figure 8: Examples of orthotubes.

Each tube block B_i is unfolded by cutting its surface along one of the four edges that is not incident to a hole, the block then unfolds into a rectangle. See Figure 9.

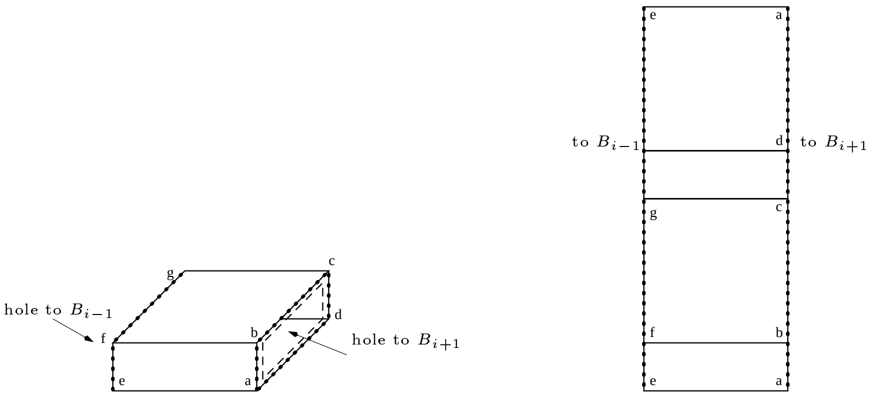


Figure 9: The unfolding of a tube block B_i. Edges that are shared with B_{i-1} or B_{i+1} are shown with underlying dots. The right figure shows the outside of the tube block. All folds of the unfolding are mountain folds.

We use two of the possible unfoldings of a corner block B_i (the choice between these two will be made according to the unfolding of B_{i-1}, detailed below). To unfold a corner block B_i, we cut along two edges of its surface as shown in Figure 10; the surface of B_i then folds flat.

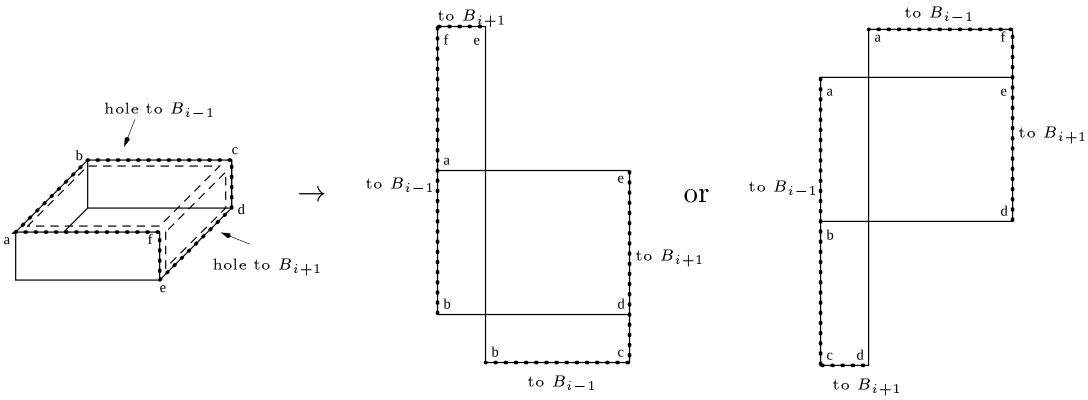


Figure 10: Two unfoldings of a corner block. Edges that are shared with B_{i-1} or B_{i+1} are shown with underlying dots. The second and third figure show the outside of the corner block. All folds of the unfolding are mountain folds.

Rotate the unfoldings of each block B_i such that at least two edges that it has in common with B_{i-1} are on the left side, and such that at least two edges that it has in common with B_{i+1} are on the right side, see Figures 9 and 10.

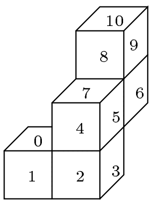
Now we show by induction on i how the unfolding of B_i can be connected to the unfolding of $B_0 \cup \ldots \cup B_{i-1}$. We maintain the induction hypothesis that in the unfolding of $B_0 \cup \ldots \cup B_{i-1}$, there are at least two edges between B_{i-1} and B_i that are on the rightmost boundary of the current polygon. We start with an unfolding of B_0 (choose an arbitrary one if B_0 is a corner block); the induction hypothesis is then satisfied regardless of whether B_0 is a tube block or a corner block.

Now assume that the unfoldings of $B_0, \ldots, B_{i-1}$ have been combined to a polygon P. Two edges that are common to B_{i-1} and B_i are on the rightmost boundary of P. However, one of these edges may not be part of the surface of B_i despite being an edge of B_i; namely, it may be part of both holes of the surface of B_i. (See for example the edge between face 7 and 8 in Figure 11, which does not belong to the surface of the block with face 6.) But there exists at most one such edge, which leaves at least one edge e that belongs to the surface of B_i and that is on the rightmost boundary of P (if there is more than one, then we choose e among them arbitrarily).

If B_i is a tube block, then the unfolding of B_i will contain e on the left side. If B_i is a corner block, then at least one of the unfoldings of B_i depicted in Figure 10 will contain e on the left side; choose an unfolding of B_i such that this is the case. Now attach this unfolding of B_i, using edge e, on the rightmost boundary of P. This will not result in overlap, because e is on the leftmost boundary of the unfolding of B_i. At least two edges between B_i and B_{i+1} are now on the rightmost boundary of the polygon, so the induction hypothesis is satisfied.

All that remains to show is how to handle an acyclic orthotube. In this case, the first and last block have only one hole. Temporarily remove the face opposite to this hole, and unfold B_0 and B_{k-1} as if they were tube blocks. Then attach the missing face at an arbitrary edge of B_0 or B_{k-1}, respectively.

In Figure 11, we show the complete example of an orthotube and its unfolding.



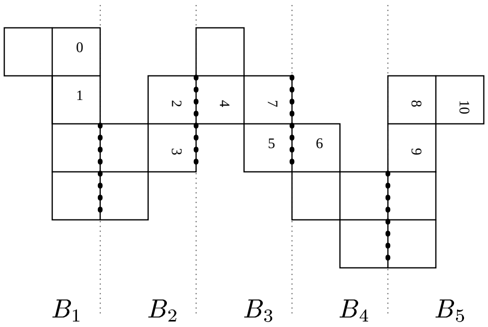


Figure 11: An orthotube and its unfolding. The edges used to connect blocks are shown with underlying dots. The space for each block is indicated with thin dotted lines. The right figure shows the outside of the orthotube block. All folds of the unfolding are mountain folds.

Because every cut has been applied in a plane that contains a vertex, it follows immediately that the number of vertices of the resulting unfolding is proportional to the number of vertices of the orthotube.

As mentioned, the unfolding of each block has only mountain-folds. The edges used to glue together two blocks are not folds; hence the resulting unfolding has only mountain-folds.

Theorem 2 *The surface of any orthotube may be cut and unfolded flat to a planar orthogonal simple polygon. Each cut is perpendicular to one of the coordinate axes and in a plane that contains at least one vertex of the orthotube. For an orthotube of n vertices, the polygon has $O(n)$ vertices. The unfolding uses only mountain-folds.*

4 Conclusion

In this paper, we studied the problem of finding unfoldings of some nonconvex polyhedra. In particular, we studied two classes of orthogonal polyhedra, called orthostacks and orthotubes, and showed that for these classes an unfolding is always possible, provided cuts across faces are allowed.

Several open problems have been pointed out throughout the paper. The most important ones are the following:

1. Can an orthostack be unfolded using only cuts that lie in a plane perpendicular to a coordinate axis and containing at least one vertex of the orthostack?
2. A natural extension to orthotubes, which form a path or cycle of blocks, is the class of orthotrees, i.e., trees of blocks. Can all orthotrees be unfolded?
3. Can an arbitrary orthogonal polyhedron be unfolded?
4. Can an orthogonal polyhedron, all faces of which are orthogonally convex polygons, be unfolded using only edge cuts?

Finally, unfoldings of other classes of nonconvex polyhedra remain to be investigated.

References

[AO92] B. Aronov and J. O'Rourke. Nonoverlap of the star unfolding. *Discrete Comput. Geom.*, 8:219–250, 1992.

[CFG93] H. T. Croft, K. J. Falconer, and R. K. Guy. *Unsolved Problems in Geometry*. Springer-Verlag, New York, 1993.

[Fuk97] Komei Fukuda. Strange unfoldings of convex polytopes, Web-page, 1997. See http://www.ifor.math.ethz.ch/staff/fukuda/unfold_home/unfold_open.html.

[She75] G. C. Shephard. Convex polytopes with convex nets. *Math. Proc. Camb. Phil. Soc.*, 78:389–403, 1975.