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, i.e., how to cut along edges and across faces so that we can then flatten the surface into a single simple polygon.

1 Introduction

An old question in computational geometry is whether every convex polyhedron has a cutting along the edges of the polyhedron such that the surface of the polyhedron can be folded flat without overlap. (See Section 2 for precise definitions.) The answer to this question was always assumed to be "yes," and for example Albrecht Dürer (1471-1528) gives such cuttings for the regular solids and the Achimedian solids in the fourth book of his painter's manual [Dür25]. The question was stated explicitly first in [She75], and to this day an answer remains to be found. Interest in this problem was revived recently by K. Fukuda; and his web page [Fuk97] states a number of conjectures how to find such an unfolding, and counter-examples to the same conjectures. If the problem is relaxed in the sense that cuts are allowed not only along edges but also across faces, then there are at least two different methods for finding an unfolding of a given convex polyhedron [AO92].

Unfolding nonconvex polyhedra seems relatively unexplored. To advance this area, we examine in this paper 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. Any orthogonal polygon that is not an axis-parallel box is nonconvex.

We will look at two classes of orthogonal polyhedra, "orthostacks" and "orthotubes," defined formally in Sections 3 and 4. 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. 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 10. Still, we can show that an unfolding always exists, even for closed knots.

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

†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.

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

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 can be any simple orthogonal polygon that changes only finitely many times, though not so as to disconnect the orthostack. In other words, orthostacks are formed by "stacking up" extrusions of simple orthogonal polygons. In particular, they contain the class of all orthogonally convex polyhedra.

We prove that all orthostacks and all orthotubes can be unfolded. Our cuts to achieve this unfolding are not always along edges; indeed, we will provide examples where this is not possible.

2 Definitions

A *k-D set* is a set of points in k-dimensional space, i.e. a set $\mathcal{S} \subseteq \mathcal{R}^k$. As a writing convention, we will use calligraphic letters for 3-D sets, bold-face letters for 3-D sets the interior of which is empty (such as the boundary of a 3-D set), and non-bold letters for 2-D sets.

[Need to define "projection", or whatever we want to call it, here or for the orthostacks.]

2.1 Polygons and Polyhedra

A *(closed) polygonal curve* is a set of n line-segments (called *edges*) (v_i, v_{i+1}), $i = 0, \ldots, n-1$ in the plane such that $v_n = v_0$. The points $v_1, \ldots, v_n$ are called *vertices* of the polygonal curve. A *simple polygon* is a closed bounded 2-D set whose interior is connected and whose boundary is a polygonal curve. A *polygon* (possibly with holes) is a closed bounded 2-D set whose interior is connected, and whose boundary is the union of finitely many disjoint polygonal curves. A 3-D set that lies entirely within a plane will be called a *3-D polygon* if its projection onto that plane is a polygon. An *open polygon* is a set the closure of which is a polygon.

A *polyhedron* is a bounded 3-D set $\mathcal{P}$ such that the following conditions hold:[1]

- The *surface* $\partial\mathcal{P}$ is the union of finitely many 3-D polygons called *faces*.
- If two faces intersect, then only in line segments that are edges for both faces.
- An *edge* of the polyhedron, i.e., an edge of one of the faces, belongs to exactly two faces of the polyhedron.
- If v is a *vertex* of the polyhedron, i.e., a vertex of one of the faces, then the faces incident to v can be sorted into one circuit. More precisely, if $\mathbf{F}_{i_0}, \ldots, \mathbf{F}_{i_{k-1}}$ are the faces for which v is a vertex, then (after suitable renaming) for $j = 0, \ldots, k-1$ the intersection of $\mathbf{F}_{i_j}$ and $\mathbf{F}_{i_{j+1}}$ contains an edge of $\mathcal{P}$ with one endpoint at v.

A polyhedron $\mathcal{P}$ is called *without cavities* if the surface of $\mathcal{P}$ is connected. In this paper, we will only study polyhedra without cavities.

A *polyhedral surface* $\mathbf{P}$ is a 3D-set that is a subset of the surface of some polyhedron $\mathcal{P}$; it is called *closed* if $\mathbf{P} = \partial\mathcal{P}$, and *open* otherwise. A vertex/edge/face of a polyhedral surface $\mathbf{P}$ is a point/line segment/3D-polygon contained in $\mathbf{P}$ that is a vertex/edge/face of some polyhedron $\mathcal{P}$ with $\mathbf{P} \subseteq \partial\mathcal{P}$.

[1] In fact, many different definitions of polyhedra exist in the community. Our definition is based on Coxeter [Cox63].

2.2 Cuttings and Unfoldings

Intuitively, an unfolding of a polyhedron is a simple polygon that is obtained by cutting the surface of the polyhedron in a suitable way, and then flattening it. A more precise definition is given in the following.

A *cutting* of a polyhedral surface $\mathbf{P}$ is a finite set of closed line segments that lie in $\partial\mathbf{P}$. The *pieces* of the cutting are the connected components of $\mathbf{P}\backslash(\bigcup_{i=1}^{m}\mathbf{e}_i \cup \bigcup_{j=1}^{k}\mathbf{l}_j)$, where $\mathbf{e}_1, \ldots, \mathbf{e}_m$ are the edges of $\mathbf{P}$, and $\mathbf{l}_1, \ldots, \mathbf{l}_k$ are the line segments of the cutting. We use these pieces (which are open 3-D polygons) to define the polygon that results from flattening the cut surface in a precise way.

We say that two pieces $\mathbf{P}_i$, $\mathbf{P}_j$ of the cutting are *adjacent* if $\overline{\mathbf{P}_i} \cap \overline{\mathbf{P}_j}$ contains a positive-length line segment that is not part of a line segment of the cutting. A cutting is called a *tree cutting* if the adjacency relationships between pieces of the cutting form a tree. Given a tree cutting of $\mathbf{P}$, we *flatten* $\mathbf{P}$ as follows: Start with an arbitrary piece $\mathbf{P}_1$ of the cutting, and translate and rotate it into the xy-plane, with the side of $\mathbf{P}_1$ that was on the outside of $\mathbf{P}$ looking towards $+z$. We call this polygon in the xy-plane the *polygon P_1 corresponding to* $\mathbf{P}_1$. As long as not all pieces have been added, find one missing piece $\mathbf{P}_j$ that is adjacent to one piece $\mathbf{P}_i$ that has been added already; such a $\mathbf{P}_j$ exists, and only one $\mathbf{P}_i$ can exist for $\mathbf{P}_j$ because the adjacencies form a tree. Let P_i, P_j be the polygons corresponding to $\mathbf{P}_i, \mathbf{P}_j$ in the xy-plane. Translate and rotate P_j such that the intersection of P_i and P_j corresponds exactly to the intersection of $\mathbf{P}_i$ and $\mathbf{P}_j$. Note that it is possible that two pieces of the cutting may overlap in this flattening. We are interested in obtaining flattenings where this is not the case, and therefore proceed to our main definition:

Definition 1 *An* unfolding *of a polyhedral surface* $\mathbf{P}$ *is a polygon that is the flattening of some tree cutting* $\{\mathbf{l}_1, \ldots, \mathbf{l}_k\}$ *of* $\mathbf{P}$*. The line segments* $\mathbf{l}_1, \ldots, \mathbf{l}_k$ *are called the* cuts of the unfolding*. An* unfolding of a polyhedron $\mathcal{P}$ *is an unfolding of* $\partial\mathcal{P}$*.*

There are two aspects to finding an unfolding of a polyhedron. The first aspect is to find the cuts. The second aspect is to find an algorithm how, given the cuts, one should proceed in actually flattening the surface without separating the pieces first and reattaching them again. The problem here is that in applications the surface of the polyhedron is typically made from stiff material that should not be bent or intersect itself while being flattened. This second problem, which we refer to as a *rigid unfolding*, is a generalization of the problem of straightening linkages without intersection. It is known that not all linkages in 3D can be straightened [BDD$^+$99], and it is conjectured that the decision problem of whether a linkage in 3D can be straightened is NP-complete.

In this paper, we study only the first aspect, i.e., we give algorithms to find for some classes of polyhedra a tree cutting such that the flattening is a simple polygon.

2.3 Orthogonality

An *orthogonal polygon* is a polygon in which each pair of adjacent edges form an angle that is a multiple of $\pi/2$. An *orthogonal polyhedron* is one in which any pair of edges with a common endpoint meet at an angle that is a multiple of $\pi/2$.

We use the term *x-plane* to denote a plane orthogonal to the x-axis. An *x-face* is a face lying in an x-plane. An *x-line* is a line parallel to the x-axis. An *x-edge* is an edge contained in an x-line, and an *x-cut* is a cut lying in an x-line. The terms are defined analogously for "y-" and "z-". An *orthogonal cut* is a cut that is either an x-cut, y-cut or z-cut, and an *orthogonal plane* is a plane that is either an x-plane, y-plane or z-plane.

3 Unfolding Orthostacks

3.1 Definitions

Informally, orthostacks are stacks of extrusions of orthogonal polygons. The precise definition is as follows:

Let $z_0 < \ldots < z_s$ be arbitrary numbers. Let $S_1, \ldots, S_s$ be orthogonal polygons without holes such that for $i = 1, \ldots, s-1$ the intersection of S_i and S_{i+1} is non-empty and non-degenerate, i.e., if $I = S_i \cap S_{i+1}$ then $I \neq \emptyset$ and the closure of the interior of I is again I. For $i = 1, \ldots, s$, let $\mathbf{E}_i$ be the *extrusion* of S_i between z_{i-1} and z_i, i.e., $\mathbf{E}_i = \{(x, y, z) : (x, y) \in S_i, z_{i-1} \leq z \leq z_i\}$. The *orthostack with respect to* $S_1, \ldots, S_s$ *and* $z_0, \ldots, z_s$ is then $\bigcup_{i=1}^{n} \mathbf{E}_i$. Another way to look at an orthostack is to view it as an orthogonal polyhedron every intersection of which with a z-plane is at most one orthogonal polygon without holes.

The *ith band* $\mathbf{B}_i$ is the extrusion of ∂S_i between z_{i-1} and z_i. See Figure 1. Band $\mathbf{B}_i$ thus consists of the x-faces and the y-faces of $\mathbf{E}_i$, and therefore $\bigcup_{i=1}^{s} \mathbf{B}_i$ is exactly the union of all x-faces and y-faces of the orthostack.

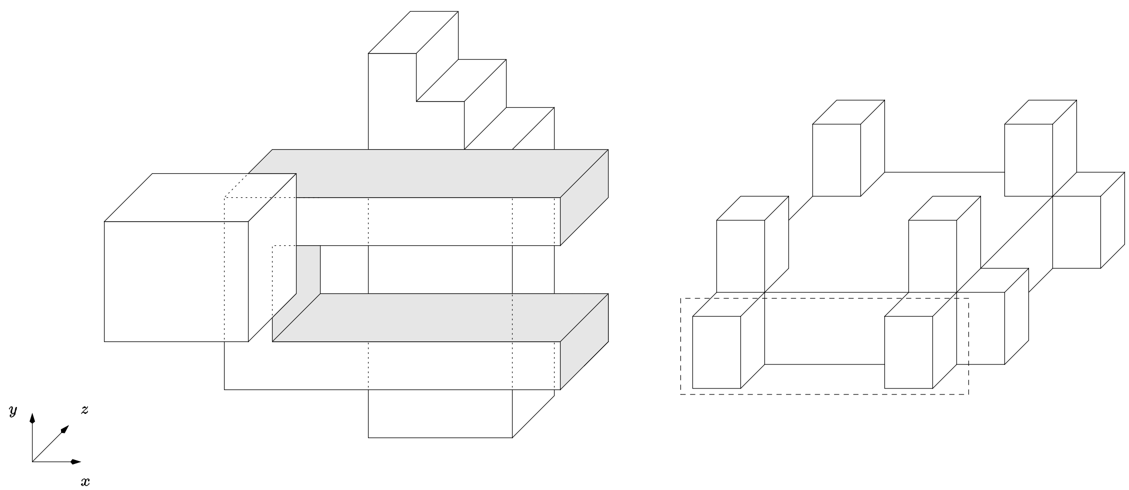

Figure 1: The left picture shows an orthostack, with the visible parts of $\mathbf{B}_2$ shaded. The right picture shows a polyhedron that is not an orthostack, because the intersection with the indicated plane consists of two polygons.

3.2 Outline of the Unfolding

Let an orthostack defined by $S_1, \ldots, S_s$ and $z_0 < \ldots < z_s$ be given. We assume that s is minimal, i.e., $S_i \neq S_{i+1}$ for $i = 1, \ldots, s-1$. For later ease of notation, define $S_0 = S_{s+1} = \emptyset$.

The basic idea for unfolding the orthostack is to cut its surface into the bands $\mathbf{B}_1, \ldots . \mathbf{B}_s$ defined above, cut each band to make a strip, and to spread these strips out in the plane in order of increasing z-coordinate. Since the bands cover all x-faces and all y-faces, only the z-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 $\mathbf{B}_i$ depends both on the previous band $\mathbf{B}_{i-1}$ and on the next band $\mathbf{B}_{i+1}$ for $1 < i < s$. Because we cannot necessarily meet the requirements imposed by the previous and the next band simultaneously, we will first cut $\mathbf{B}_i$ in two halves with a z-plane. More precisely, for $1 \leq i \leq n$, let $\mathbf{L}_i$ be the extrusion of ∂S_i between z_{i-1} and $\frac{1}{2}(z_{i-1} + z_i)$, and let

$\mathbf{R}_i$ be the extrusion of ∂S_i between $\frac{1}{2}(z_{i-1} + z_i)$ and z_i. We call $\mathbf{L}_i$ and $\mathbf{R}_i$ *half-bands*; note that $\mathbf{L}_i \cup \mathbf{R}_i = \mathbf{B}_i$. For later ease of notation, define $\mathbf{R}_0 = \mathbf{L}_{s+1} = \emptyset$. We will unfold each half-band separately, and reattach the half-bands appropriately later.

For the part of the unfolding that contains the z-faces, we need the following definitions illustrated in Figure 2: For $0 \leq i \leq s$, let D_i^+ be the closure of $S_i - S_{i+1}$, and let $\mathbf{D}_i^+$ be the projection of D_i^+ into the $(z = z_i)$-plane. Thus, $\mathbf{D}_i^+$ is the union of all z-faces with coordinate z_i that look "away from the viewer" (i.e., toward $+z$). For $0 \leq i \leq s$, let D_i^- be the closure of $S_{i+1} - S_i$, and let $\mathbf{D}_i^-$ be the projection of D_i^- into the $(z = z_i)$-plane. Thus, $\mathbf{D}_i^-$ is the union of all z-faces with coordinate z_i that look "towards the viewer" (i.e., toward $-z$). Define $\mathbf{D}_i = \mathbf{D}_i^+ \cup \mathbf{D}_i^-$; which thus is the projection of the symmetric difference of S_i and S_{i+1} into the $(z = z_i)$-plane. Note that $\mathbf{D}_i$ has a non-empty interior, because s was minimal, and therefore $S_i \neq S_{i+1}$.

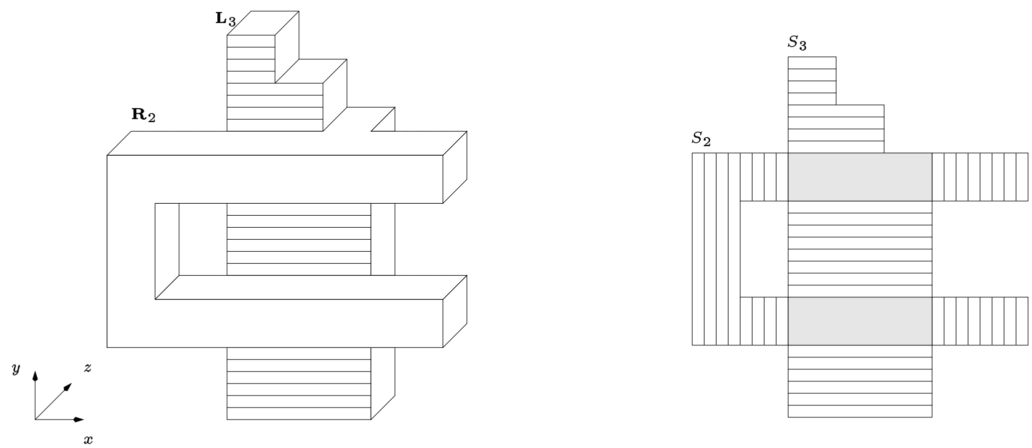

Figure 2: $\mathbf{D}_2$ in the example of Figure 1; $\mathbf{D}_2^+$ is marked by y-lines, $\mathbf{D}_2^-$ is marked by x-lines. In the right picture, we show the projection onto the xy-plane.The shaded area marks $S_2 \cap S_3$.

Note that $\bigcup_{i=0}^{s}(\mathbf{R}_i \cup \mathbf{D}_i \cup \mathbf{L}_{i+1})$ is exactly the union of all faces of the orthostack. Thus, to unfold the surface of an orthostack, we proceed as follows: (1) For $i = 0, \ldots, s$, we unfold $\mathbf{R}_i \cup \mathbf{D}_i \cup \mathbf{L}_{i+1}$, $i = 0, \ldots, s$, by cutting each half-band into a rectangular strip, connecting the strips directly or using a rectangle $\mathbf{C}_i^*$ from $\mathbf{D}_i$, cutting the remainder of $\mathbf{D}_i$ into rectangles, and attaching these rectangles without overlap at appropriate places on the strips. (2) For $i = 0, \ldots, s$, we show how to reattach the unfoldings of $\mathbf{R_i} \cup \mathbf{D}_i \cup \mathbf{L}_{i+1}$ to each other without overlap.

Details of these two steps will be given in the next two subsections. Figure 3 gives an overview of the final unfolding.

3.3 Attaching $\mathbf{D}_i$ to $\mathbf{R}_i$ and $\mathbf{L}_{i+1}$

In this section, we study how to unfold $\mathbf{R}_i \cup \mathbf{D}_i \cup \mathbf{L}_{i+1}$ for one fixed $i \neq 0, s$. (The treatment is similar, and even easier, if $i \in \{0, s\}$, because then one of $\mathbf{R}_i$ and $\mathbf{L}_{i+1}$ is empty.) The problem here is where to cut the bands $\mathbf{R}_i$ and $\mathbf{L}_{i+1}$ to guarantee that they can be connected. This connection will be done using a piece $\mathbf{C}^*$ of $\mathbf{D}_i$, or directly if $\mathbf{R}_i$ and $\mathbf{L}_{i+1}$ are connected to each other. That this is always possible is shown in the following lemma.

Lemma 3.1 *If $0 < i < s$, then there exists an open rectangle $\mathbf{C}$ in $\mathbf{D}_i$ such that one x-edge of $\mathbf{C}$ belongs to $\mathbf{R}_i$ and the other x-edge of $\mathbf{C}$ belongs to $\mathbf{L}_i$. More precisely, there exist values $x_1 < x_2$ and $y_1 \leq y_2$ such that $\mathbf{C} = \{(x, y, z_i) : x_1 < x < x_2, y_1 < y < y_2\} \subset \mathbf{D}_i$, and such that*

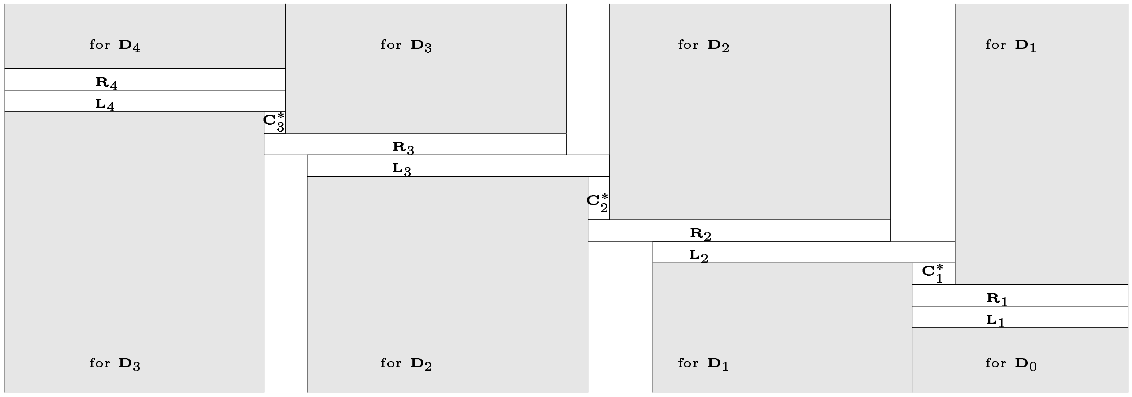

Figure 3: Overview of the unfolding.

- $e_1 = \{(x, y_1) : x_1 < x < x_2\} \subset \partial S_i$, $e_2 = \{(x, y_2) : x_1 < x < x_2\} \subset \partial S_{i+1}$, *or*
- $e_1 = \{(x, y_1) : x_1 < x < x_2\} \subset \partial S_{i+1}$, $e_2 = \{(x, y_2) : x_1 < x < x_2\} \subset \partial S_i$.

Proof: Let I be the intersection of S_i and S_{i+1}; this intersection is non-empty and non-degenerate by definition of an orthostack. Let y_1 be the largest y-coordinate of a point in I, and let e be an x-edge of I which has y-coordinate y_1. In Figure 4, we have $y_1 = 7$ and e is the edge between endpoints $(3, 7)$ and $(6, 7)$.

Edge e belongs to ∂S_i or ∂S_{i+1} or both. We apply one of two methods to find x_1, x_2, y_2, 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 ∂S_i and ∂S_{i+1}, then let $[x_1, x_2]$ be a maximal interval of $e \cap \partial S_i \cap \partial S_{i+1}$, i.e., a maximal interval of overlap between ∂S_i and ∂S_{i+1} within e. In Figure 4, which illustrates this method, we have $x_1 = 5$ and $x_2 = 6$. Also, set $y_2 = y_1$. Note that in this case $\mathbf{C}$ is empty. Also, $e_1 = e_2 \subseteq \partial S_i \cap \partial S_{i+1}$, so the claim holds.

Method (2): If some part of e belongs to only one of ∂S_i and ∂S_{i+1}, then let (x_0, x_2) be a maximal open interval of $e \cap (\partial S_i \backslash \partial S_{i+1} \cup \partial S_{i+1} \backslash \partial S_i)$, i.e., a maximal interval of non-overlap between ∂S_i and ∂S_{i+1} within e. In Figure 5, which illustrates this method, we have $x_0 = 3$ and $x_2 = 5$. Let $j \in \{i, i+1\}$ be such that e belongs to ∂S_j, and let k be such that $\{j, k\} = \{i, i+1\}$.

Let x_1 be minimal with $x_0 \leq x_1 < x_2$ such that no y-edge of S_i or S_{i+1} has x-coordinate in the interval (x_1, x_2); x_1 exists because there are only finitely many y-edges. In Figure 5, we have $x_1 = 4$.

Fix an arbitrary value $x^* \in (x_1, x_2)$. Let $y_2 > y_1$ be the smallest value such that (x^*, y_2) belongs to $\partial S_i \cup \partial S_{i+1}$. This exists because $(x^*, y_1) \in I$, and because (x^*, y_1) belongs to ∂S_j but not to ∂S_k. Also, y_2 is independent of the choice of x^*, because the y-coordinate of the next boundary can change only if there is a y-edge of S_i or S_{i+1}, but no such edge exists in the interval (x_1, x_2). We observe the following:

- The points $\{(x^*, y) : y > y_1\}$ do not belong to I by definition of y_1.
- Because (x^*, y_1) belongs to I, and belongs to ∂S_j but not to ∂S_k, the points $\{(x^*, y) : y_1 < y < y_2\}$ belong to $S_k \backslash S_j$. In particular therefore, the points $\{(x^*, y, z_i) : y_1 < y < y_2\}$ belong to $\mathbf{D}_i$.

- Point (x^*, y_2) cannot belong to to ∂S_j, because otherwise $(x^*, y_2) \in I$, which contradicts the definition of y_1.
- Therefore, $(x^*, y_2) \in \partial S_k$ by definition of y_2.

Since these claims hold for any $x_1 < x^* < x_2$, it follows that all points in $\mathbf{C}$ belong to $\mathbf{D}_i$. Also, e_1 belongs to ∂S_j and e_2 belongs to ∂S_k, which proves the claim. □

Let $\mathbf{C}^*$ be the rectangle $\{(x, y, z_i) : x_1 \leq x \leq x_2, y_1 \leq y \leq y_2\}$. Note that $\mathbf{C}^*$ is possibly degenerate, i.e., a x-segment, if $y_1 = y_2$. Also, either the $(y = y_1)$-edge of $\mathbf{C}^*$ belongs to $\mathbf{R}_i$ and the $(y = y_2)$-edge of $\mathbf{C}^*$ belongs to $\mathbf{L}_{i+1}$, or vice versa. If $\mathbf{C}^*$ is non-degenerate, then it is the closure of set $\mathbf{C}$ of the previous lemma, and therefore belongs to $\mathbf{D}_i$ because $\mathbf{D}_i$ is closed. In either case, we call rectangle $\mathbf{C}^*$ the *connecting bridge*, because it will be used to connect the two half-bands $\mathbf{R}_i$ and $\mathbf{L}_{i+1}$ (or, if the rectangle is degenerate, the two half-bands will directly attach to each other).

This connecting bridge determines where the half-bands $\mathbf{R}_i$ and $\mathbf{L}_{i+1}$ are cut into strips. Specifically, we cut $\mathbf{R}_i$ by extending the $(x = x_1)$-edge of $\mathbf{C}^*$, and $\mathbf{L}_{i+1}$ by extending the $(x = x_2)$-edge of $\mathbf{C}^*$. More precisely, if $j \in \{1, 2\}$ is such that $\{(x, y_j) : x_1 \leq x \leq x_2\} \subset \partial S_i$, then add the z-segment $\{(x_1, y_j, z) : \frac{1}{2}(z_{i-1} + z_i) \leq z \leq z_i\}$ to the cutting; this segment is on $\mathbf{R}_i$. If k is such that $\{j, k\} = \{1, 2\}$, then by Lemma 3.1 $\{(x, y_k) : x_1 \leq x \leq x_2\} \subset \partial S_{i+1}$, and we add the z-segment $\{(x_2, y_k, z) : z_i \leq z \leq \frac{1}{2}(z_i + z_{i+1})\}$ to the cutting; this segment is on $\mathbf{L}_{i+1}$. These cuts are demonstrated in Figures 4 and 5.

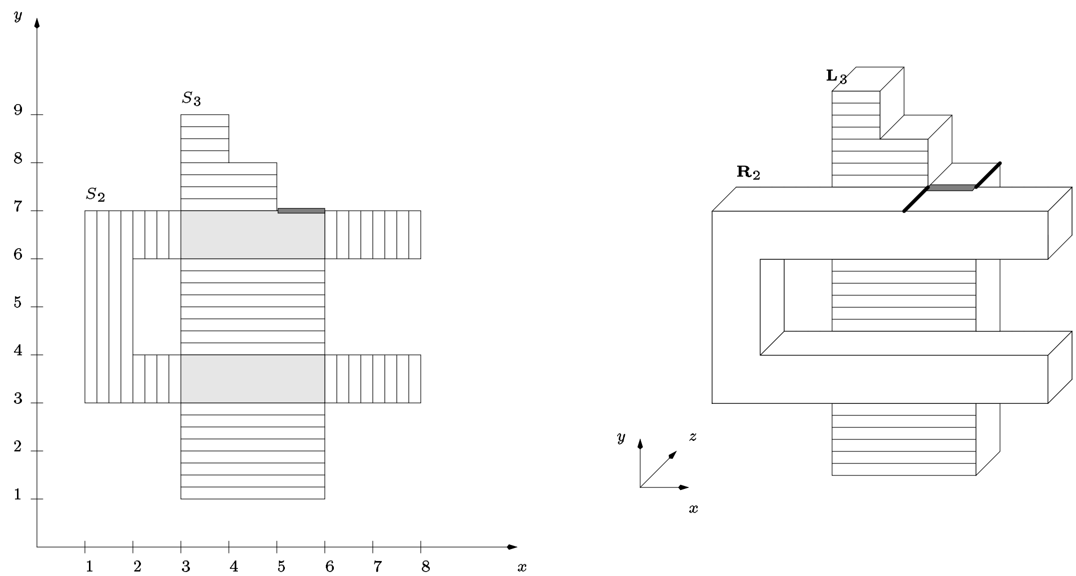

Figure 4: How to find a connecting bridge (shown darkly shaded) with Method (1), i.e., if the edge belongs to both ∂S_i and ∂S_{i+1}. In this case, the connecting bridge is degenerate, which means that the two half-bands directly attach to each other. We also show how to extend the boundaries of the connecting bridge into cuts for the two half-bands.

Now we turn to the issue of how to attach the missing z-faces, i.e., the faces in $\mathbf{D}_i \backslash \mathbf{C}^*$. We cut $\mathbf{D}_i \backslash \mathbf{C}^*$ into a collection of rectangles $\mathbf{P}_1, \ldots, \mathbf{P}_l$, by extending every y-edge of $\mathbf{D}_i$ in both directions until it hits an x-edge. These rectangles $\mathbf{P}_1, \ldots, \mathbf{P}_l$ lie in the $(z = z_i)$-plane. See Figure 6.

For $j = 1, \ldots, l$, we add to the cutting all edges of $\mathbf{P}_j$ except the top $(+y)$ edge. The top edge of $\mathbf{P}_j$ remains attached to one of the half-bands $\mathbf{R}_i$ and $\mathbf{L}_{i+1}$. If $\mathbf{C}^*$ is non-degenerate, then we

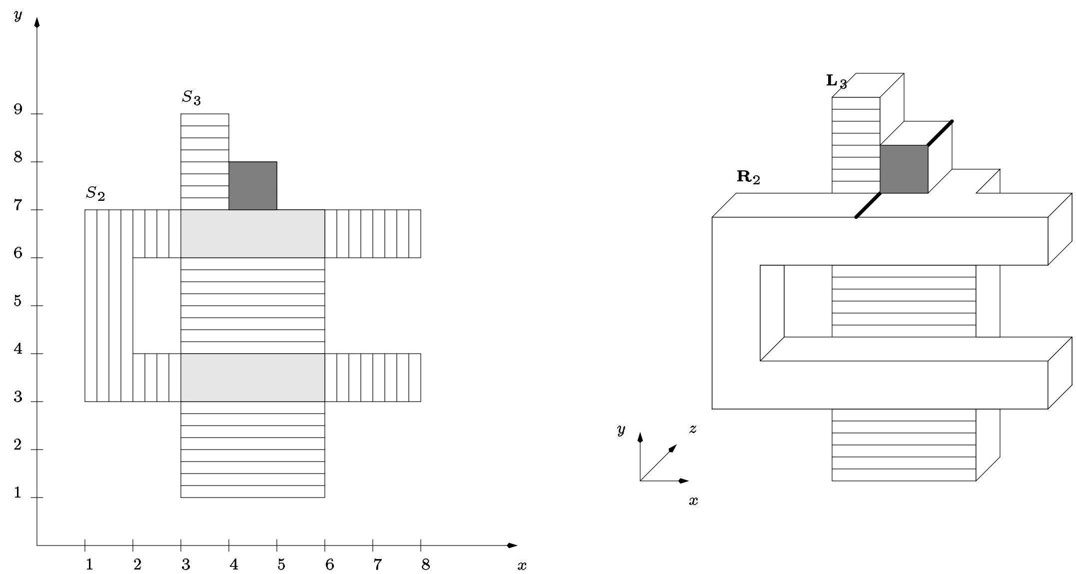

Figure 5: How to find a connecting bridge (shown darkly shaded) with Method (2), i.e., if the edge e belongs to only one of ∂S_i and ∂S_{i+1} (shown here is $e \in \partial S_2$). In this case, the connecting bridge is non-degenerate and belongs to $\mathbf{D}_i$. We also show how to extend the boundaries of the connecting bridge into cuts for the two half-bands.

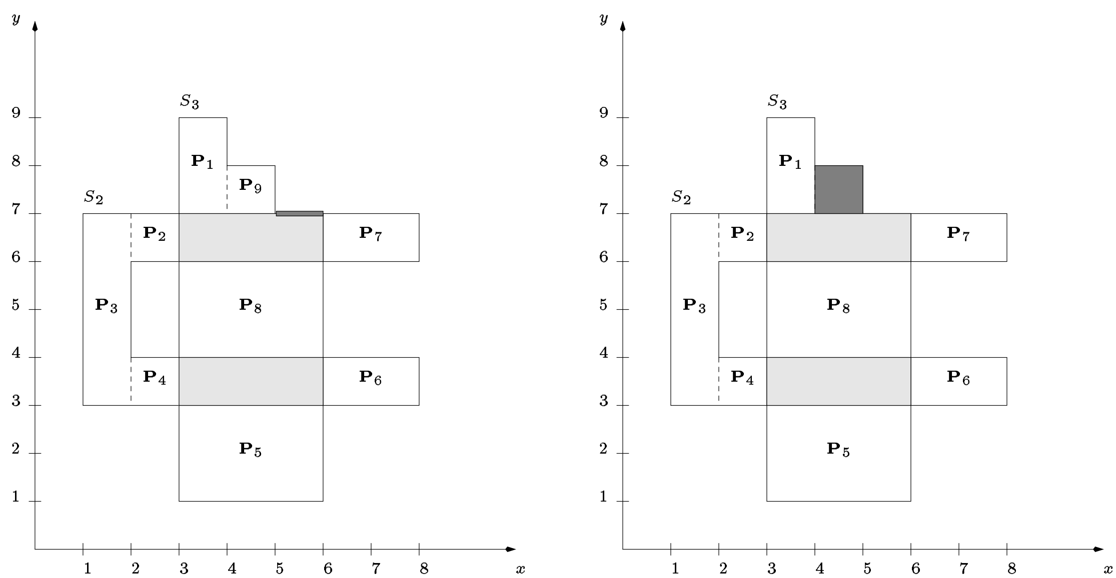

Figure 6: $\mathbf{D}_i \backslash \mathbf{C}^*$ is partitioned into rectangles $\mathbf{P}_1, \ldots, \mathbf{P}_l$ by cutting along extensions of y-edges. We show the splitting of rectangles both when $\mathbf{C}^*$ is degenerate and when it is non-degenerate.

also add to the cutting the two y-edges, but not the x-edges, of $\mathbf{C}^*$. Finally, we add to the cutting all line segments that are common to $\mathbf{R}_i$ and $\mathbf{L}_{i+1}$ and that do not belong to $\mathbf{C}^*$.

The resulting cutting is a tree-cutting of $\mathbf{R}_i \cup \mathbf{D}_i \cup \mathbf{L}_{i+1}$, because the adjacencies of the half-bands form paths, these paths are adjacent to each other if $\mathbf{C}^*$ is degenerate and both adjacent to $\mathbf{C}^*$ otherwise, and the rectangles $\mathbf{P}_1, \ldots, \mathbf{P}_l$ attach to exactly one face of a half-band. The complete set of cuts for $\mathbf{R}_i \cup \mathbf{D}_i \cup \mathbf{L}_{i+1}$ is shown in Figure 7.

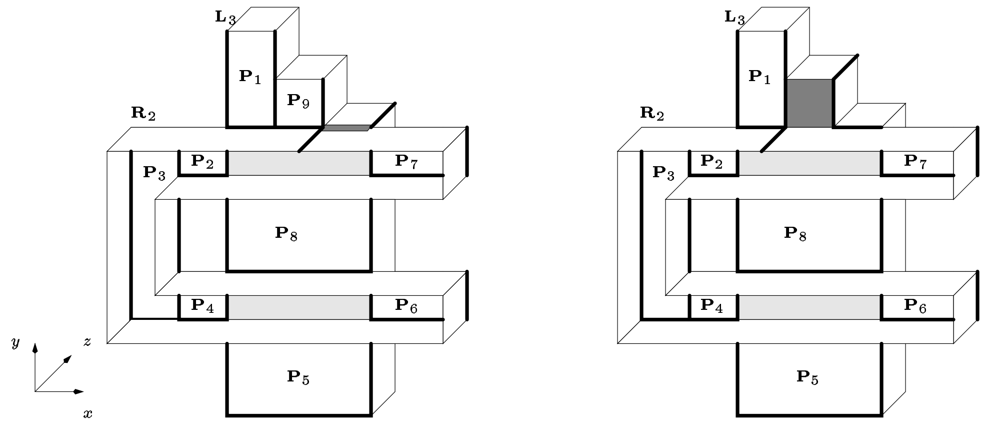

Figure 7: The cutting of $\mathbf{R}_i \cup \mathbf{D}_i \cup \mathbf{L}_{i+1}$, both when $\mathbf{C}^*$ is degenerate and when it is not.

We have to show that the flattening of this tree-cutting does not overlap itself. Consider Figure 8 and study first R_i, the flattening of $\mathbf{R}_i$. This is obtained by flattening a path of rectangular faces, all of them of the same width in z-dimension, and connected to each other at their z-edges. Thus, R_i is a rectangular strip; place this strip in the xy-plane such that the inside of the strip points towards the $+z$-direction. Similarly, the flattening L_{i+1} of $\mathbf{L}_{i+1}$ is a strip, and we place it in the same fashion.

Recall that rectangle $\mathbf{C}^*$ consisted of the points $\{(x, y, z_i) : x_1 \leq x \leq x_2, y_1 \leq y \leq y_2\}$. Translate L_{i+1} such that $\mathbf{C}^*$, interpreted as a 2D-polygon C^*, can be connected to both R_i and L_{i+1}; thus shift L_{i+1} such that the bottom $(-y)$ end of L_{i+1} is exactly $y_2 - y_1$ above the top $(+y)$ end of R_{i+1}, and the right $(+x)$ end of L_{i+1} is exactly $x_2 - x_1$ to the right of the left $(-x)$ end of R_{i+1}. Then C^* fits exactly between R_i and L_{i+1}.

The half-infinite rectangle above R_i and the half-infinite rectangle below L_{i+1} will be used to place all other pieces of $\mathbf{D}_i$. More precisely, each rectangle $\mathbf{P}_1, \ldots, \mathbf{P}_l$ attaches to one half-band at its top edge. The half-infinite rectangle from both half-bands is empty and each piece is a rectangle, therefore no intersections are possible. See Figure 8.

Thus, we have shown how to unfold $\mathbf{R}_i \cup \mathbf{D}_i \cup \mathbf{L}_{i+1}$, $0 < i < s$. We can handle $\mathbf{D}_0 \cup \mathbf{L}_1$ and $\mathbf{R}_s \cup \mathbf{D}_s$ similarly. We cut $\mathbf{L}_1$ by extending the cut of $\mathbf{R}_1$, cut $\mathbf{D}_0$ into rectangles, and attach these rectangles in the half-infinite rectangle below $\mathbf{L}_1$. We cut $\mathbf{R}_s$ by extending the cut of $\mathbf{L}_s$, cut $\mathbf{D}_s$ into rectangles, and attach these rectangles in the half-infinite rectangle above $\mathbf{R}_s$.

3.4 Connecting the layouts of $\mathbf{R}_i \cup \mathbf{D}_i \cup \mathbf{L}_{i+1}$

All that remains to show is how to attach half-band $\mathbf{L}_i$ to half-band $\mathbf{R}_i$, $i = 0, \ldots, s$. The half-bands $\mathbf{L}_i$ and $\mathbf{R}_i$ 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 to the left or to the right of L_i, respectively. We choose the former for all $i = 1, \ldots, s$, lining up the top of the strip R_i with the appropriate place in strip L_i. See Figure 3, particularly the two middle pair of strips. The flattenings of all triplets thus combined do not intersect.

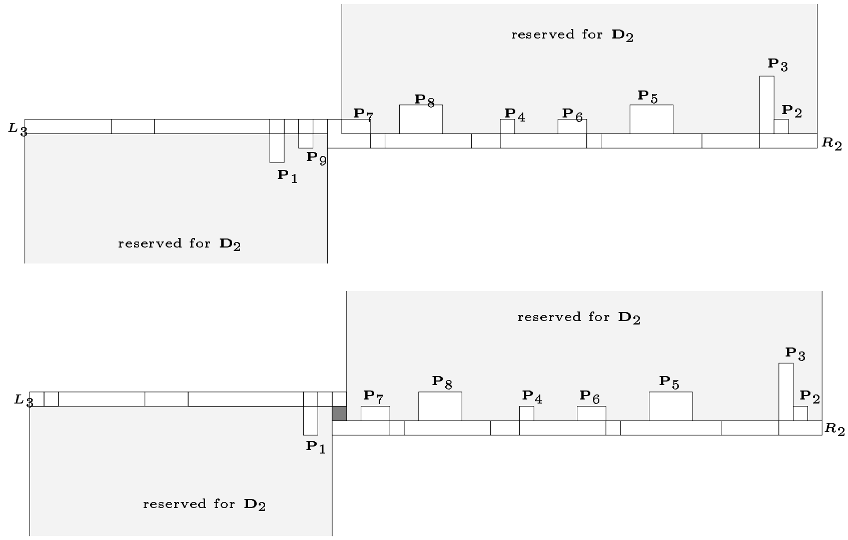

Figure 8: Splitting $\mathbf{D}_i$ into rectangles and attaching them to the strips. We show this for a degenerate connecting bridge on the top, and for a non-degenerate connecting bridge on the bottom.

3.5 Complexity of the unfolding

Thus we have found 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 relative to the number of vertices of the orthostack. Every symmetric difference $\mathbf{D}_i$ is non-empty and thus contains at least one rectangle (which is possibly the connecting bridge). So we can charge the up to 8 vertices of $R_i \cup L_{i+1}$ to one rectangle of $\mathbf{D}_i$, and the number of vertices of the unfolded polygon is $O(r)$, where r is the total number of rectangles obtained from $\mathbf{D}_0, \ldots, \mathbf{D}_s$. All the rectangles obtained by partitioning $\mathbf{D}_i$ lie in the $(z = z_i)$-plane, and each rectangle $\mathbf{P}_k$ is incident to at least one vertex v of the orthostack in the $(z = z_i)$-plane, because $\mathbf{P}_k$'s y-edges are obtained by extending y-edges of the orthostack. Charge each rectangle $\mathbf{P}_k$ to one of the vertices v on the boundary of $\mathbf{P}_k$. No vertex v can be incident to more than two rectangles (cf. Figure 6), and so no vertex receives more than two charges. Thus the number of rectangles from $\mathbf{D}_i$ is $O(n_i)$, where n_i is the number of vertices of the orthostack in the $(z = z_i)$-plane. 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 *Any orthostack of n vertices has an unfolding with $O(n)$ vertices such that all cuts are orthogonal.*

It would be a reasonable restriction to require that all cuts lie in an orthogonal plane that contains at least one vertex of the polyhedron. However, our construction splits band $\mathbf{B}_i$ into half-bands $\mathbf{L}_i$ and $\mathbf{R}_i$ with a cut that violates this restriction. We leave as an open problem finding an unfolding under this more stringent condition.

3.6 Unfolding with edge cuts is impossible

In our construction, we used cuts that are not edge cuts, i.e., that are not along an edge of the polyhedron (at least not if we represent the polyhedron with the minimal possible number of faces).

We now show that this is necessary for the orthostacks in Figure 9, which by Theorem 1 can be unfolded if we allow cuts across faces.

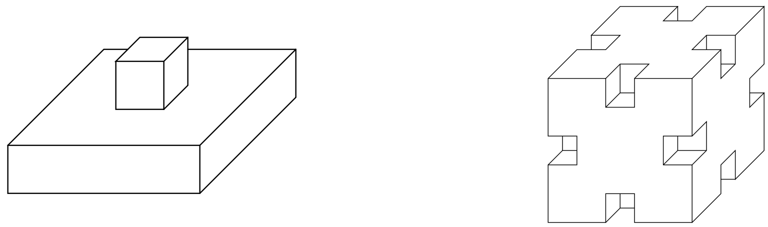

Figure 9: Two orthostacks that cannot be unfolded with edge cuts.

First, consider the left polyhedron of Figure 9. In any cutting with only edge cuts, all faces are left intact, i.e., are pieces of the cutting. 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 small cube's surface area exceeds the area of the hole. Hence, there is no unfolding with edge cuts only.

A more intriguing example is the right polyhedron of Figure 9, which is a cube with small "bites" taken out of every edge. Here, every face is an orthogonal polygon without holes. Nevertheless, this polyhedron has no unfolding with edge cuts only. For assume there exists a tree cutting with only edge cuts such that the flattening is a simple polygon. The pieces of this cutting are the faces. Consider two of the large faces that are closest in the tree of adjacencies of the cutting. They must either be joined directly, or via the faces of the bite between them. The first possibility does not leave enough area for the faces of the bite, and the second one either leads to overlap between faces or again leaves insufficient 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. Very recently, it was shown that there exists a (non-orthogonal) non-convex polyhedron with convex faces that has no unfolding with only edge cuts [BDEK99]. No generalization of this result to orthogonal polyhedra is apparent, and we leave this as an open problem.

4 Unfolding Orthotubes

In the previous section, we showed how to unfold any orthostack. The class of orthostacks contains a wide variety of orthogonal polyhedra, but there are also many orthogonal polyhedra which are not orthostacks. Our next step therefore was to search for orthogonal polyhedra that are not orthostacks and either to show that they cannot be unfolded or to find unfolding algorithms.

It remains open whether there exists an orthogonal polyhedron that cannot be unfolded, but our search for such a polyhedron led us to consider the *orthogonalized trefoil knot* (see the right picture of Figure 10), because the trefoil knot has been used for other impossibility results as well [BDD$^+$99]. As it turns out, the orthogonalized trefoil knot *can* be unfolded, and the algorithm to find this unfolding led us to a large class of orthogonal polyhedra, the orthotubes, that can be unfolded, and that are radically different from orthostacks.

Informally, an orthotube is the union of *blocks* (i.e., axis-parallel rectangular boxes) $\mathcal{B}_0, \ldots, \mathcal{B}_{k-1}$ such that $\mathcal{B}_i$ attaches to $\mathcal{B}_{i+1}$ for $i = 0, \ldots, k-2$ (this will be made more precisely below). We have two types of orthotubes: *cyclic orthotubes* for which $\mathcal{B}_{k-1}$ attaches in turn to $\mathcal{B}_0$, and *acyclic orthotubes*, for which this is not the case. To define an orthotube precisely, set $\mathcal{B}_{-1} = \mathcal{B}_k = \emptyset$ to obtain an acyclic orthotube, and $\mathcal{B}_{-1} = \mathcal{B}_{k-1}$, $\mathcal{B}_k = \mathcal{B}_0$ to obtain a cyclic orthotube.

The precise definition is then as follows: An *orthotube* is the union of blocks $\mathcal{B}_0, \ldots, \mathcal{B}_{k-1}$ such that for $i = 0, \ldots, k-1$, $\mathcal{B}_i \cap \mathcal{B}_{i+1}$ is a 2-dimensional face of both $\mathcal{B}_i$ and $\mathcal{B}_{i+1}$, and such that $\mathcal{B}_i \cap \mathcal{B}_j$, $j \neq i-1, i+1$, is either empty, or a vertex or an edge of $\mathcal{B}_i \cap \mathcal{B}_{i-1}$ or $\mathcal{B}_i \cap \mathcal{B}_{i+1}$. See Figure 10 for some examples of cyclic orthotubes, neither of which is an orthostack. As we will show now, any orthotube has an unfolding.

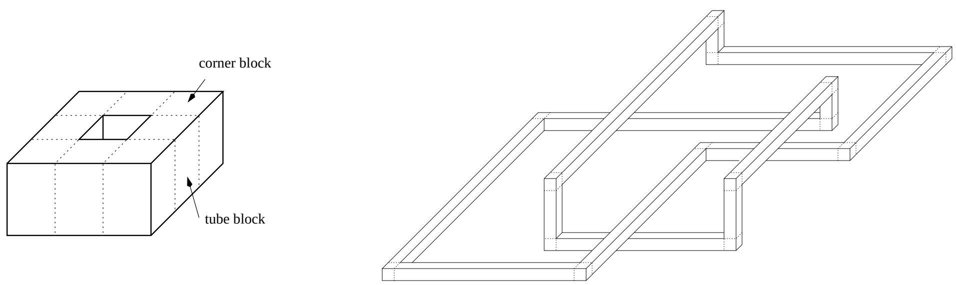

Figure 10: Examples of orthotubes.

Consider an orthotube $\mathcal{P}$ consisting of the blocks $\mathcal{B}_0, \ldots, \mathcal{B}_{k-1}$. We assume that the blocks are chosen maximally, i.e., $\mathcal{P}$ cannot be described using less than k blocks. We also assume for now that the orthotube is cyclic; the other case will be treated later. In the following, all additions are modulo k. Let the *wrapping* $\mathbf{B}_i$ *of a block* $\mathcal{B}_i$ be $\mathcal{B}_i \cap \partial\mathbf{P}$, i.e., the union of all faces of $\mathcal{B}_i$ that are on the surface of $\mathcal{P}$. Notice that the union of the wrappings of all blocks is the surface of the orthotube.

The wrapping $\mathbf{B}_i$ of block $\mathcal{B}_i$ is the surface of $\mathcal{B}_i$ with two faces missing, namely, the faces that correspond to $\mathcal{B}_i \cap \mathcal{B}_{i-1}$ and $\mathcal{B}_i \cap \mathcal{B}_{i+1}$. We call these missing faces the *holes* of $\mathcal{B}_i$. There are two classes of blocks: the *tube blocks*, where the two holes are parallel, and the *corner blocks*, where this is not the case. See Figure 10 for an example. Abusing notation, we will use the words hole, tube block and corner block also for the wrapping $\mathbf{B}_i$ of a block $\mathcal{B}_i$ when they really apply to $\mathcal{B}_i$.

We explain in the following first the options for unfolding one wrapping $\mathbf{B}_i$, and then show how to combine these unfoldings. If $\mathbf{B}_i$ is a tube block, then we unfold it by cutting along an edge of $\mathbf{B}_i$ that is incident to a hole of $\mathbf{B}_i$. There are four such edges, and the choice between them is arbitrary. For this cut, the wrapping then unfolds into a rectangle, see Figure 11.

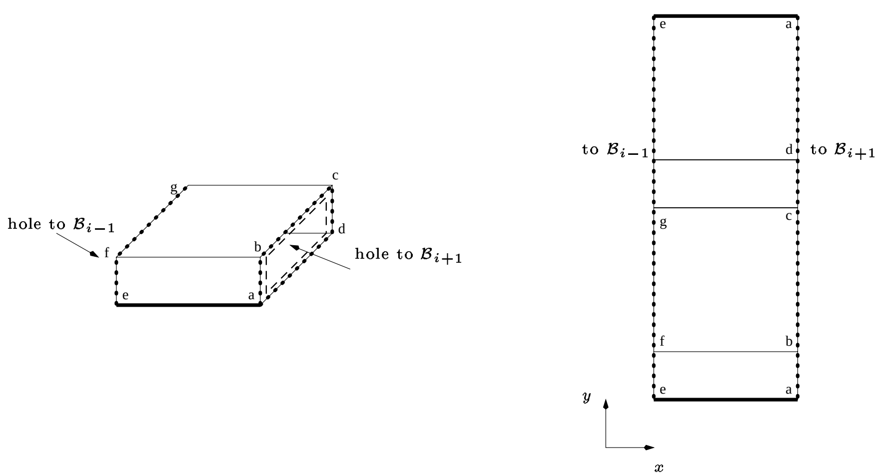

Figure 11: The unfolding of the wrapping of a tube block $\mathcal{B}_i$. Edges that are shared with $\mathcal{B}_{i-1}$ or $\mathcal{B}_{i+1}$ are shown with underlying dots. The cut is shown with a thick line.

If $\mathbf{B}_i$ is the wrapping of a corner block, then we use one of two possible unfoldings for $\mathbf{B}_i$. There are four edges of $\mathbf{B}_i$ for which exactly one endpoint is incident to a hole of $\mathbf{B}_i$. Of these four edges, we cut along two edges that are not parallel. This gives two possible cuts, and the choice between them is not arbitrary, but depends on the unfolding of $\mathbf{B}_{i-1}$ as will be explained below. With either sets of cuts, $\mathbf{B}_i$ then unfolds to a simple polygon; see Figure 12.

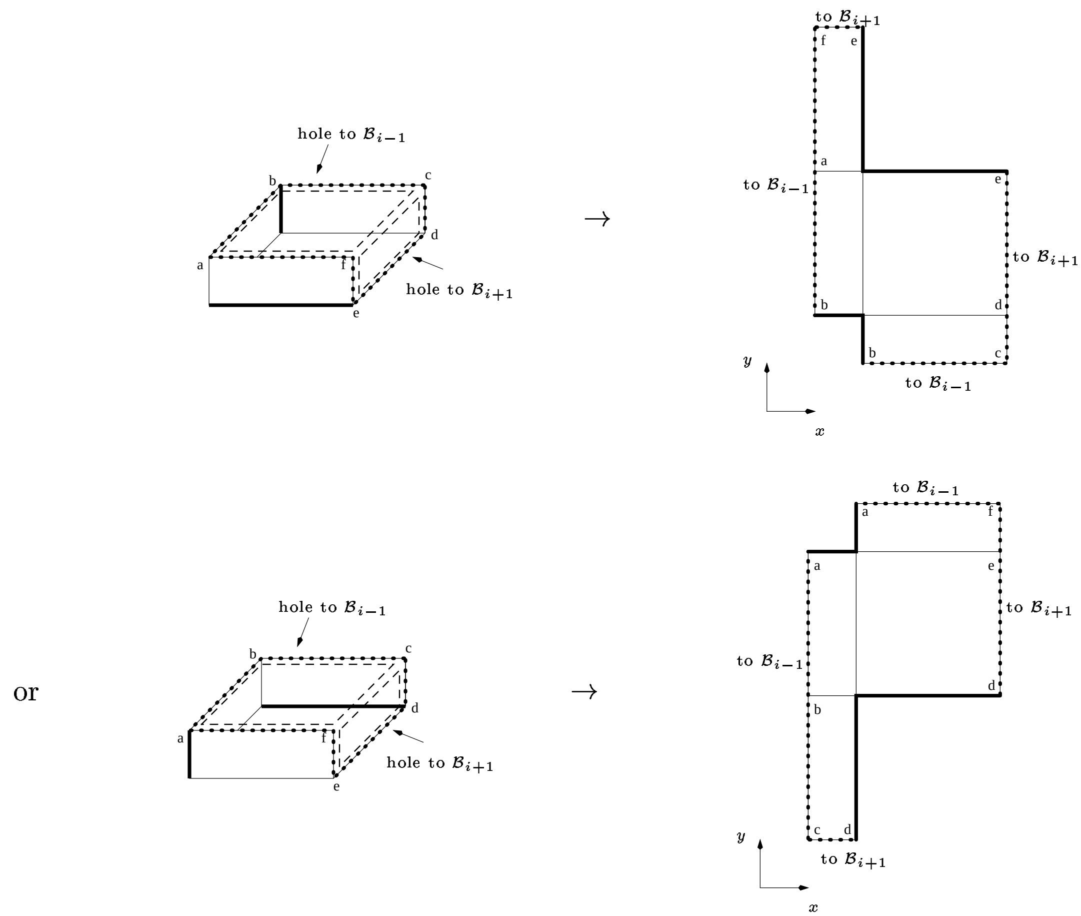

Figure 12: The unfolding of the wrapping of a corner block $\mathcal{B}_i$. Edges that are shared with $\mathcal{B}_{i-1}$ or $\mathcal{B}_{i+1}$ are shown with underlying dots. The cuts are shown with thick lines.

For later use, we note that after a suitable rotation, the unfoldings of a corner block have the following properties: (1) For each unfolding, there are two edges common to $\mathcal{B}_i$ and $\mathcal{B}_{i-1}$ on the leftmost $(-x)$ boundary of the unfolding, and two edges common to $\mathcal{B}_i$ and $\mathcal{B}_{i+1}$ on the rightmost $(+x)$ boundary of the unfolding. (2) Any edge common to $\mathcal{B}_i$ and $\mathcal{B}_{i-1}$ is on the leftmost boundary of at least one of the two unfoldings.

Now we show by induction on i how $\bigcup_{j=0}^{i} \mathbf{B}_j$ can be unfolded. We maintain the induction hypothesis that at least two edges common to $\mathcal{B}_{i-1}$ and $\mathcal{B}_i$ are on the rightmost $(+x)$ boundary of the unfolding of $\mathbf{B}_0 \cup \ldots \cup \mathbf{B}_{i-1}$.

We start with an unfolding B_0 of $\mathbf{B}_0$, choosing an arbitrary one if $\mathbf{B}_0$ is a corner block. Rotate B_0 such that at least two edges common to $\mathcal{B}_0$ and $\mathcal{B}_1$ are on the right $(+x)$ side of B_0; this is always possible as demonstrated in Figures 11 and 12. The induction hypothesis is then satisfied.

Now assume that $\mathbf{B}_0 \cup \ldots \cup \mathbf{B}_{i-1}$ has been unfolded, and the rightmost $(+x)$ boundary of its unfolding P contains two edges e_1, e_2 that are common to $\mathcal{B}_{i-1}$ and $\mathcal{B}_i$. At least one of these edges, say e_1, is also an edge of $\mathbf{B}_i$, because the only edge of $\mathcal{B}_i$ that could fail to be an edge of $\mathbf{B}_i$ is the edge of a corner block that is incident to both holes; there is at most one such edge.

If $\mathcal{B}_i$ is a tube block, then create an unfolding B_i of $\mathbf{B}_i$ as described above. Rotate B_i such that the edges common to $\mathcal{B}_i$ and $\mathcal{B}_{i-1}$ are on the left $(-x)$ side of B_i; see Figure 11. In particular therefore e_1 is on the left side of B_i, and all four edges common to $\mathcal{B}_i$ and $\mathcal{B}_{i+1}$ are on the right side of B_i.

If $\mathcal{B}_i$ is a corner block, then among the two possible unfoldings of $\mathbf{B}_i$ described above, choose an unfolding B_i such that after suitable rotation e_1 is on the leftmost boundary of B_i and two edges common to $\mathcal{B}_i$ and $\mathcal{B}_{i+1}$ are on the rightmost boundary of B_i; this is always possible as observed above.

In either case, attach B_i to the unfolding P of $\mathbf{B}_0 \cup \ldots \cup \mathbf{B}_{i-1}$ using edge e_1. In other words, translate B_i such that the two locations of e_1 in P and B_i coincide. This will not result in overlap, because e_1 is on the leftmost boundary of B_i and on the rightmost boundary of P. At least two edges common to $\mathcal{B}_i$ and $\mathcal{B}_{i+1}$ are on the rightmost boundary of B_i, and so now on the rightmost boundary of $P \cup B_i$, and the induction hypothesis is satisfied.

This proves that any cyclic orthotube can be unfolded, so all that remains to show is how to handle an acyclic orthotube. In this case, the first block $\mathcal{B}_0$ and the last block $\mathcal{B}_{k-1}$ have only one hole each. Let $\mathbf{F}_0$ be the face opposite to the hole of $\mathcal{B}_0$ and let $\mathbf{F}_{k-1}$ be the face opposite to the hole of $\mathcal{B}_{k-1}$. Then $\mathbf{B}_0 \backslash \mathbf{F}_0$ and $\mathbf{B}_{k-1} \backslash \mathbf{F}_{k-1}$ are the wrappings of tube blocks, thus we can unfold $\bigcup_{i=0}^{k-1} \mathbf{B}_i \backslash \mathbf{F}_0 \backslash \mathbf{F}_{k-1}$ as described before. Denote this unfolding by P. The edges of $\mathbf{F}_0$ are on the leftmost $(-x)$ boundary of P by construction, hence we can attach $\mathbf{F}_0$, interpreted as a 2D-polygon, at any of these edges to P without creating overlap. The edges of $\mathbf{F}_{k-1}$ are on the rightmost $(+x)$ boundary of P by construction, hence we can attach $\mathbf{F}_{k-1}$, interpreted as a 2D-polygon, at any of these edges to P without creating overlap. Thus we obtain an unfolding of an acyclic orthotube.

In Figure 13, we show the complete example of an orthotube and its unfolding.

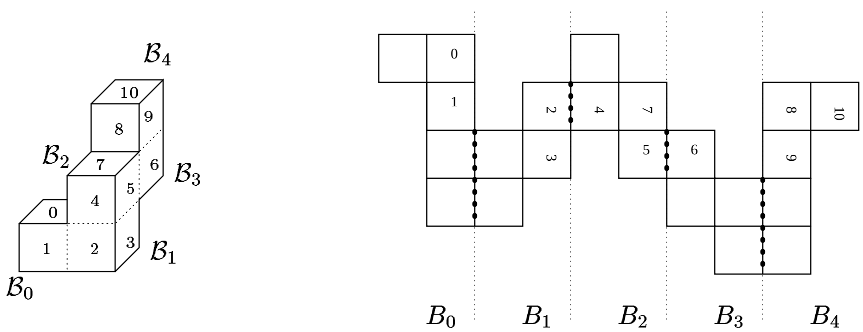

Figure 13: An orthotube and its unfolding. The edges used for connecting blocks are shown with underlying thick dots.

Each cut in the unfolding of an orthotube is along the edge of a block. Unfortunately, an edge of a block is not necessarily an edge of the orthotube, so this is not an unfolding with edge cuts only. But on the positive side, because we chose the minimum possible number of blocks to describe the orthotube, every face of a block is incident to a vertex of the orthotube. Every cut therefore lies in an orthogonal plane containing a vertex of the orthotube. Also, every face of a block has 4 edges, and every vertex of the orthotube is incident to at most 6 faces of blocks, which implies that the number of cuts, and hence the complexity of the unfolding, is proportional to the number of vertices of the orthotube.

Theorem 2 *Any orthotube with n vertices has an unfolding with $O(n)$ vertices such that all cuts* *are in an orthogonal plane containing a vertex of the orthotube.*

Our unfoldings of orthotubes have another interesting property: In the origami-community, one distinguishes between *mountain-folds*, which are folds that bend the two pieces attached to it "away" from the viewer, and *valley-folds*, which are folds that bend to two pieces attached to it "toward" the viewer. Studying the unfoldings of a tube block and a corner block reveals that, presuming the outside of the block is turned "toward" the viewer, we have only mountain-folds. Because the edges used to glue together the blocks together are not folds, this implies that our unfoldings of orthotubes have only mountain-folds.

5 Conclusion

In this paper, we studied the problem of finding unfoldings of two classes of orthogonal polyhedra, orthostacks and orthotubes. We showed that any such orthogonal polyhedron has an unfolding. Moreover, the complexity of the unfolding is proportional to the complexity of the orthogonal polyhedron. To our knowledge, this is the first result on unfolding any class of non-convex polyhedra.

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

1. Does every orthogonal polyhedron have an unfolding?

   We suspect that the answer to this question is no. We have discovered an open orthogonal polyhedral surface that cannot be unfolded. However, a generalization of this construction to a closed polyhedral surface remains to be found.

2. What other classes of orthogonal polyhedra have an unfolding?

   We found a super-class to orthostacks and orthotubes, called *orthocylinders*, that can be unfolded. An orthocylinder is an orthotube where each tube block is replaced by an orthostack such that the leftmost and rightmost face of the orthostack are exactly the holes of the tube block. Such an orthocylinder can be unfolded by unfolding each orthostack separately, and combining these unfoldings with the unfoldings of the corner blocks as for orthotubes.

   For which other classes can we find unfoldings? For example, a natural extension to an orthotube, which is a path or cycle of blocks, is an *orthotree*, i.e., a tree of blocks. Can all orthotrees be unfolded?

3. Define a *rigid unfolding* as one that permits flattening while (a) keeping all faces of the cutting rigid and (b) avoiding self intersection of the faces. It was shown that our cutting of the orthogonalized trefoil knot has no rigid unfolding [BLS99]. Does our cutting of an orthostack have a rigid unfolding? If not, is there a different cutting of an orthostack the flattening of which is simple, and that has a rigid unfolding?

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

Acknowledgments

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. The research was supported by NSERC, FCAR and NSF.

References

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

[BDD+99] T. Biedl, E. Demaine, M. Demaine, S. Lazard, A. Lubiw, J. O'Rourke, M. Overmars, S. Robbins, I. Streinu, G. Toussaint, and S. Whitesides. Locked and unlocked polygonal chains in 3D. In *SIAM-ACM Conference on Discrete Algorithms*, pages 866–867, 1999.

[BDEK99] Marshall Bern, Erik D. Demaine, David Eppstein, and Eric Kuo. Ununfoldable polyhedra. In *11th Canadian Conference on Computational Geometry*, 1999. To appear.

[BLS99] T. Biedl, A. Lubiw, and J. Sun. When can a net fold to a polygon? In *11th Canadian Conference on Computational Geometry*, 1999. To appear.

[Cox63] H.S.M. Coxeter. *Regular Polytopes.* The Macmillan Company, 1963.

[Dür25] Albrecht Dürer. *Unterweysung der Messung mit dem Zirckel uñ richtscheyt in Linien ebnen unnd ganzen corporen.* Nürnberg, 1525. Translated and with a commentary by Walter L. Strauss in *The Painter's Manual*, Abaris Books, New York, 1977.

[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.