Theoretical Computer Science 289 (2002) 897–917

Theoretical Computer Science

www.elsevier.com/locate/tcs

Embedding problems for paths with direction constrained edges ☆

Giuseppe Di Battista[a], Giuseppe Liotta[b], Anna Lubiw[c], Sue Whitesides[d,*]

[a] *Dipartimento di Informatica ed Automazione, Università di Roma Tre, Roma, Italy*
[b] *Dipartimento di Ingegneria Elettronica e dell'Informazione, Università di Perugia, Perugia, Italy*
[c] *Department of Computer Science, University of Waterloo, Waterloo, Canada*
[d] *School of Computer Science, McGill University, Montreal, Que., Canada*

Received November 2001

Abstract

We determine the reachability properties of the embeddings in R^3 of a directed path, in the graph-theoretic sense, whose edges have each been assigned a desired direction (*East*, *West*, *North*, *South*, *Up*, or *Down*) but no length. We ask which points of R^3 can be reached by the terminus of an embedding of such a path, by choosing appropriate positive lengths for the edges, if the embedded path starts at the origin, does not intersect itself, and respects the directions pre-assigned to its edges. This problem arises in the context of extending planar graph embedding techniques and VLSI rectilinear layout techniques from 2D to 3D. We give a combinatorial characterization of reachability that yields linear time recognition and layout algorithms. Finally, we extend our characterization to R^d, $d > 3$. © 2002 Published by Elsevier Science B.V.

Keywords: Shape; Embedding; Path; Rectilinear layout

☆ Research partially supported by operating grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada, by the project "Algorithms for Large Data Sets: Science and Engineering" of the Italian Ministry of University and Scientific and Technological Research (MURST 40%), and by the project "Geometria Computazionale Robusta con Applicazioni alla Grafica ed a CAD" of the Italian National Research Council (CNR).

* Corresponding author.

E-mail addresses: gdb@dia.uniroma3.it (G.D. Battista), liotta@diei.unipg.it (G. Liotta), alubiw@uwaterloo.ca (A. Lubiw), sue@cs.mcgill.ca (S. Whitesides).

0304-3975/02/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S0304-3975(01)00408-X

1. Introduction

Consider a directed, polygonal curve in 3D consisting of n axis-parallel straight line segments of positive lengths, and define its *shape* to be the sequence σ of n direction labels *East*, *West*, *North*, *South*, *Up*, or *Down* determined by the directions of the n segments of the curve. Clearly, every such curve has a unique shape.

On the other hand, one may start with a shape σ, and produce curves that have that shape by specifying coordinates for the initial endpoint of the curve and assigning to each element in σ a positive length. Such curves are called *realizations* of σ. Of course, it may happen that such a realization, while having the desired shape, intersects itself. As will be seen shortly, self-intersections can easily be avoided by choosing lengths with some care, provided that σ contains no element (say, for example, *East*) that is immediately followed by its opposite direction (*West*). What will concern us here is whether a given shape has a realization satisfying a given reachability constraint: not only should the realization be a curve that does not intersect itself, but also, the curve should start at the origin and terminate at a specified point in space.

To formalize this notion of shape realization subject to a reachability constraint, consider a directed, graph theoretic path P, and specify a *shape* for this purely combinatorial object P by giving an ordered sequence σ of n labels *East*, *West*, *North*, *South*, *Up*, or *Down*, one for each of the n edges of P. Each label in the sequence specifies a direction for the corresponding directed edge of P when that edge is realized as a straight line segment in 3D. A *realization* of this pair P,σ of combinatorial objects as an orthogonal polygonal curve in 3D is specified by giving a start point for the curve and an assignment of positive lengths for the edges of the curve.

Given a point p in 3D and a shape σ for a directed, graph theoretic path P, the 3D *shape reachability problem* is to find an assignment (if one exists) of positive lengths to the directed edges of P so that the resulting realization of P,σ is an orthogonal, polygonal curve that

- ends at p if it starts at the origin;
- is *simple*, that is, does not intersect itself; and
- satisfies the direction constraints on its segments as specified by σ.

The 3D shape reachability problem has instances that do not admit solution. For example, consider a point p that lies in the octant determined by the three directions *East*, *Up*, and *North* and consider the shape σ given by the sequence of labels $UWDESWN$, where U stands for *Up*, W stands for *West*, and so on. Shape σ cannot be the shape of any simple, orthogonal, polygonal realization of a path that starts at the origin and terminates at p, even though σ contains an E label, a U label, and an N label (see Fig. 1).

Of course the notion of shape extends to arbitrary dimensional space. The shape of a directed, orthogonal polygonal curve in d dimensions is a sequence of n direction labels of the form $X_1^+, X_1^-, \ldots, X_d^+, X_d^-$. Because, the 3D setting is our main interest, we will mainly be describing our results there, proving in Section 4.1 some lemmas in 2D needed to obtain the 3D results. We will use the direction labels E, W, N, W in

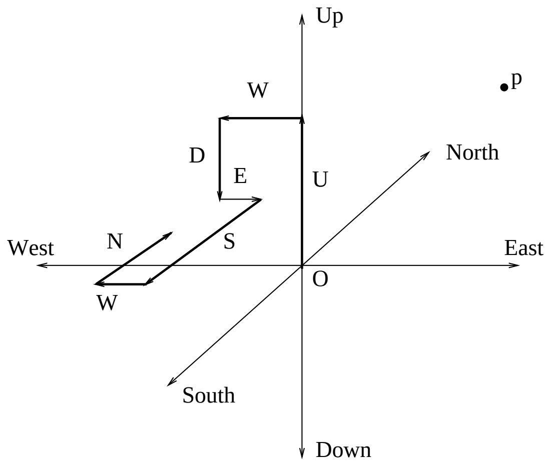

Fig. 1. No path of shape $\sigma = UWDESWN$ can start at the origin O and reach a point p in the octant determined by the three directions *East*, *Up*, and *North*.

2D and the direction labels E, W, N, S, U, D in 3D, rather than the notation X_1^+, X_1^-, etc. We will use the latter notation in Section 5, when we generalize our characterization from R^3 to R^d.

We regard a point in R^d as having $2d$ coordinates: the X_i^+ coordinate and the X_i^- coordinate have opposite signs but agree in absolute value. Thus, in R^3, for example, a point that is 3 units *East* of the origin has E coordinate 3 and W coordinate -3.

In general, we omit the specification of the dimension in front of terms such as "shape", as the dimension of the ambient space should be clear from the context.

The main result of this paper is a combinatorial characterization, for arbitrary dimension d, of those shape reachability problem instances that admit solution. We also give recognition and embedding algorithms based on this characterization. In fixed dimension such as 3D, these algorithms are linear. When the algorithms must handle inputs of arbitrary dimension, then the running time grows superexponentially with the dimension.

Various 2D versions of our 3D shape problem have been studied in several papers, including [12,14]. This kind of problem has also been considered for non-orthogonal polygons and for some graphs drawn with non-orthogonal edges (see [6,10,13]).

A basic technique for 2D orthogonal layout in VLSI and in graph drawing is to give a "shape" for a graph, that is, an assignment of direction labels E, W, N, and S to the edges, and then to determine lengths for the edges so that the layout is non-crossing [14]. The graph must have maximum degree less than or equal to 4. In the VLSI context, each vertex represents either a corner of a bounding box containing components, or a connection pin on the side of such a box. The edges at a vertex connect it to its two neighbors (pins or box corners) on the bounding box and, in the case of a pin vertex, to a pin on another box.

The well-known *topology-shape-metrics approach* [4] for constructing a rectilinear embedding of a planar graph consists of three main steps, called *planarization*, *orthogonalization*, and *compaction*. The planarization step determines an embedding, i.e., the face cycles, for the graph in the plane. The orthogonalization step then specifies for each edge (u,v) a shape that the orthogonal polygonal line representing (u,v) is to have in the final embedding. For example, (u,v) could be labelled $NESNE$, which would say "starting from u first go North, then go East, etc.". Finally, the compaction step computes the final embedding, giving coordinates to vertices and bends.

The push to manufacture at ever smaller scales (see, e.g., [7] for one view of the possibilities) means that the VLSI 2D layout problems of the past will be replaced by a host of 3D graph embedding problems, arising from new 3D manufacturing technologies. Furthermore, there is interest in 3D layout of graphs for visualization purposes, and some experimental evidence that such layouts can be useful [15].

While the literature on 3D orthogonal embeddings of graphs is quite rich (see, e.g. [1,2,3,5,8,9,11,16,17,18]), the extension of the topology-shape-metrics approach to 3D remains, as far as we know, to be carried out. The 3D shape reachability results we present here are an essential pre-requisite for such a program.

2. Overview of the main results for 3D

In order to state our characterization result for 3D precisely, we introduce the concepts of a *flat* and of a *canonical sequence* in a shape σ. The definitions here are given for the context of three dimensions; we will extend them to higher dimensions in Section 5, where we generalize our characterization. As these are key concepts for our work, it is important to understand them clearly from the outset. However, we do not expect that it will be clear at this point why these concepts are so useful.

A *flat* of σ is a consecutive subsequence σ' of σ such that σ' contains at least two elements and is maximal with respect to the property that its labels come from the union of two oppositely directed pairs of directions, i.e., either from the set $\{N,S,E,W\}$, or from the set $\{N,S,U,D\}$, or from the set $\{U,D,E,W\}$. Thus, any realization of the shape σ' must consist of segments that lie on the same axis-aligned plane. Since there are at least two segments, they cannot be colinear. For example, the shape $\sigma = UWDESWN$ of the path drawn in Fig. 1 contains two flats: $F_1 = UWDE$, and $F_2 = ESWN$. Observe that two consecutive flats of σ share a label and that they must be drawn on perpendicular planes. For example, the last label, E, of F_1 coincides with the first label, E, of F_2. (Observe that the notion of "flat" also makes sense in 2D, where any shape of length $n \geqslant 2$ consists of exactly one flat.)

A not necessarily consecutive subsequence $\tau \subseteq \sigma$, where τ consists of $k \geqslant 1$ elements, is a *canonical sequence* provided that:

- any two elements of τ indicate mutually orthogonal directions (hence $1 \leqslant k \leqslant 3$ in 3D); and
- if a flat F of σ contains one or more labels of τ, then $\tau \cap F$ forms a consecutive subsequence of σ.

Note that the following facts follow immediately from the definition of canonical sequence: every single element of a shape is a canonical sequence; any two consecutive elements of a shape form a canonical sequence; and two non-consecutive orthogonal elements of a shape form a canonical sequence if and only if they do not belong to the same flat.

The *type* of a canonical sequence τ is given by the (unordered) set of direction labels it contains.

For example, the shape $\sigma = UWDESWN$ of the path drawn in Fig. 1 contains two canonical sequences of type $\{U,N,W\}$. To obtain one of the canonical sequences of this type, consider the first two elements of σ, which are a U and a W. These elements belong to the same flat, F_1, where they appear consecutively; they do not belong to F_2. The last element of σ is an N, which belongs to flat F_2 but not to flat F_1. Thus, the conditions for a canonical sequence of length $k=3$ and type $\{U,N,W\}$ are satisfied by the subsequence of σ consisting of its first, second, and last elements. To obtain the other canonical sequence of type $\{U,N,W\}$, take the first, next-to-last, and last elements of σ. Note, however, that σ does not contain a canonical sequence of type $\{U,E,N\}$: these labels occur in unique positions in σ, and in particular, the unique U and the unique E both belong to the same flat F_1, but they are not consecutive in σ.

The essence of our combinatorial characterization of solvable shape reachability instances is given below. For concreteness, the characterization is given with respect to the UNE octant. As explained in the next section, the results for other octants can be obtained by a suitable permutation of the direction labels.

Theorem 1. *Let σ be a shape and let p be a point in the UNE octant. Then there exists a simple, orthogonal, polygonal curve of shape σ that starts at the origin and that terminates at p if and only if σ contains a canonical sequence of type $\{U,N,E\}$.*

In other words, p can be reached, starting from the origin, by a simple orthogonal polygonal curve of shape σ if and only if it is possible to choose from σ a particular U element $\bar{U}$, a particular N element $\bar{N}$, and a particular E element $\bar{E}$, not necessarily in that order, such that if any flat of σ contains two of the elements $\bar{U},\bar{N},\bar{E}$, then these two elements are consecutive in σ.

We believe that the necessity of this condition is far from obvious. On the other hand, the proof of the sufficiency of the condition will be given by a construction for which we can offer the following intuition. Imagine that the three particular elements of a canonical sequence are realized as very long line segments, connected by sequences of shorter segments. A construction based on this idea would create a path that could reach a point p in the UNE octant. The fact that two of these long segments must be adjacent if they belong to the same flat offers hope that the construction can produce a simple path; by contrast, if two long segments with orthogonal direction labels belonged to the same flat but were not adjacent, then, depending on the construction, they might cross each other.

The rest of this paper is organized as follows. Preliminaries are in Section 3. As a basic tool for obtaining the results in this paper, we developed a theory of shape reachability in 2D. Our results on the shape reachabilty problem both in 2D and in 3D are presented in Section 4. An extension of our results to higher dimensional spaces can be found in Section 5. We conclude with some open problems in Section 6.

3. Preliminaries

We regard each coordinate axis in the standard 3D coordinate system as consisting of the origin plus two open semi-axes, directed away from the origin and labelled with a pair of opposite direction labels from the set $\{N, W, S, E, U, D\}$. A triple XYZ of distinct unordered direction labels no two of which are opposite specifies the *XYZ octant*. Note that unless stated otherwise, we consider octants to be open sets. Similarly, a pair XY of distinct orthogonal direction labels specifies the *XY quadrant* in 2D or 3D. Finally, a direction label X specifies the *X semi-axis*, which consists of those points that are positive multiples of the unit vector in the X direction. For short, we call a semi-axis an axis. Thus, 3D space is partitioned into eight octants, twelve quadrants, six axes, and the origin.

Let σ be a shape consisting of n elements or labels.[1] An *embedding* or *drawing* of σ, denoted as $\Gamma(\sigma)$, is a *non-self-intersecting*, directed, orthogonal polygonal curve consisting of n segments such that the kth segment ($k = 1, \ldots, n$) of $\Gamma(\sigma)$ has positive length and has the orientation specified by the kth label of σ. Unless stated otherwise, we assume that the start of the directed curve $\Gamma(\sigma)$ lies at the origin of the reference system.

Assumptions. Since, we are interested in shapes that admit embeddings, we assume from now on that shapes do not contain adjacent labels that are oppositely directed. Furthermore, we assume that shapes do not contain adjacent labels that are identical, since the reachability properties of shapes with adjacent, identical labels are not changed by coalescing such labels. Finally, to avoid lengthy special case handling, we assume that the point p to be reached in an instance of the shape reachability problem lies in an octant in 3D (and in the analogous full-dimensional subspace in the general case).

Remark. Let $\phi(\,)$ be a permutation of the six direction labels that maps opposite pairs of labels to possibly different opposite pairs (for example, ϕ might map N, S, E, W, U, D to E, W, N, S, D, U, respectively). Note that $\phi(\,)$ defines a linear transformation of 3D space that is not necessarily a rotation or reflexion. Nevertheless, this transformation defines a bijection between embeddings of σ and embeddings of $\phi(\sigma)$. Here, $\phi(\sigma)$ denotes the sequence of labels obtained by applying ϕ to the labels of σ. For concreteness, we often state our results and proofs referring to some given octant, quadrant, or axis where points of embeddings of σ can lie. However, the results can also be stated

[1] Since the elements of a shape have direction labels as values, we often refer to the elements as labels.

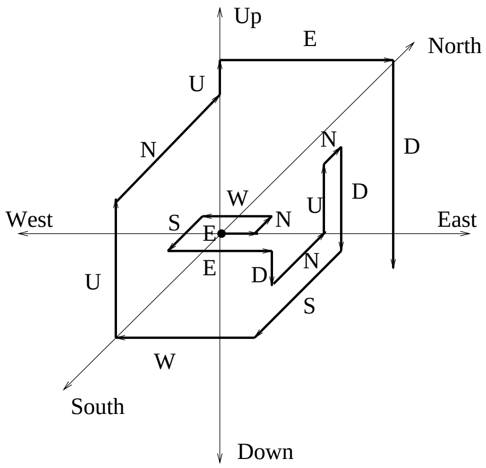

Fig. 2. The expanding drawing for $\sigma = ENWSEDNUNDSWUNUED$. The drawing starts at the origin, and each segment travels one unit farther in its direction than any preceding segment has gone.

with respect to any other octant, quadrant, or axis since they are preserved under the $\phi(\ )$ transformation. This remark generalizes to higher dimensions.

Definition. A drawing $\Gamma(\sigma)$ of a shape σ is an *expanding drawing* if each segment travels farther in its direction than the extreme points, with respect to that direction, of the previous segments of $\Gamma(\sigma)$.

Definition. A drawing $\Gamma(\sigma)$ of a shape σ is *extensible* if its last segment can be replaced by one of arbitrary length without causing collisions with $\Gamma(\sigma)$.

Expanding drawings are useful because they provide a simple way to make an extensible drawing (Fig. 2).

Lemma 1. *Every shape σ admits an expanding, extensible drawing*

Proof. We must prove that expanding drawings do not intersect themselves and that they are extensible. Suppose $\Gamma(\sigma)$ contains n segments, and let B_0 denote the point at which it starts. Clearly, the lemma holds for $n \leqslant 2$, so assume $n > 2$. For $i \geqslant 1$, let B_i denote the axis-aligned bounding box for the first i segments. Let s_i denote the segment corresponding to the ith element of σ. Since the tip of s_i protrudes from B_{i-1}, and since s_{i+1} is orthogonal to s_i, the infinite line determined by s_{i+1} does not intersect B_{i-1}. Thus, s_{i+1} can be made any length, and so $\Gamma(\sigma)$ has no self-intersections and is extensible. $\square$

Assumption. To be concrete, we assume from now on that unless stated otherwise, each segment of an expanding drawing travels one unit beyond the bounding box of the previous segments.

In the next sections, we shall characterize the reachability properties of shapes by means of canonical sequences. We will sometimes use special notation to distinguish the labels of a shape σ that have been chosen to belong to a canonical sequence. For example, we say that $\{\bar{U}, \bar{N}, \bar{E}\}$ is a canonical sequence in σ meaning not only that it has type $\{U, N, E\}$ but also that the U label chosen to be in the canonical sequence is a particular element $\bar{U}$ of σ, and so on. We also use symbols such as $\hat{U}, \hat{N}, \hat{E}$. Indeed, when it is necessary to refer to two different elements having the same value such as U, we may use both $\hat{U}$ and $\bar{U}$ in the same discussion or proof. Since particular elements of σ are indicated by this kind of special notation, the order of these particular elements in σ and in the canonical sequence is determined. This order is not necessarily the same as the order in which the elements may appear inside the curly brackets for set notation.

4. The shape reachability problem

We say that a drawing $\Gamma(\sigma)$ of a shape σ *reaches point* p if $\Gamma(\sigma)$ terminates at p when it starts at the origin. A shape σ *reaches a point* p if it admits a drawing $\Gamma(\sigma)$ that reaches p. A shape σ *reaches a set of points* if it reaches each point in the set. For example, shapes NEU, NUE, and $NWUE$ all reach the UNE octant; that is, for each point of the UNE octant, each of these shapes has a drawing that terminates at that point.

Lemma 2. *Let σ be a shape. If σ reaches a point p in an axis, quadrant or octant, then it reaches that entire axis, quadrant or octant, respectively.*

Proof. Let $\Gamma(\sigma)$ be a drawing of σ that reaches p. Then for any other point q in the same axis, quadrant or octant as p, we can construct a drawing of σ that reaches q by suitably scaling $\Gamma(\sigma)$. For example, if p and q are in the NE quadrant, then we multiply the lengths of the segments of $\Gamma(\sigma)$ that are associated with N or S labels by a constant equal to the ratio of the length of the orthogonal projection of q to the N axis to the length of the orthogonal projection of p to the N axis. Lengths of the segments of $\Gamma(\sigma)$ that are associated with the labels E or W are scaled separately in an analogous way. $\square$

Note that if a shape reaches the UNE octant, then it must contain a U label, an N label and an E label. Notice, however, that the converse is not always true. As previously observed, the shape $UWDESWN$ does not reach the UNE octant, even though it contains a U label, an E label, and an N label.

In the rest of this section, we investigate the problem of determining whether a shape can reach a given portion of 2D or 3D space. We distinguish between shapes that consist of only one flat and shapes that have more than one flat. A shape of the first type is a 2D shape, while a shape of the second type is a 3D shape. Clearly, a canonical sequence for a 2D shape can contain at most two elements, and if there are two, they must be adjacent.

4.1. Reachability in 2D

In this section we answer the following question: Given a point p in a quadrant and a 2D shape σ, can σ reach p? By Lemma 2, this question can be answered by characterizing when σ can reach the quadrant containing p.

We start by investigating a special type of reachability. Let σ be a shape in 2D whose labels belong to $\{N,S,E,W\}$, and let $\Gamma(\sigma)$ be a drawing of σ that reaches the *NE* quadrant. We say that $\Gamma(\sigma)$ reaches the *NE* quadrant *for the first time with label X* if X is the direction associated with the first segment of $\Gamma(\sigma)$ entering the *NE* quadrant when walking along $\Gamma(\sigma)$ starting at the origin.

Lemma 3. *Let $\Gamma(\sigma)$ be a drawing of a 2D shape $\sigma = \sigma' E$ such that $\Gamma(\sigma)$ reaches the NE quadrant for the first time with its last label, E. Then σ contains the subsequence NE, where E immediately follows N.*

Proof. If $\Gamma(\sigma)$ has two edges, then $\sigma = NE$. Assume, for $k \geqslant 3$, that the statement is true for shapes with $k-1$ elements and that σ has k elements. The next-to-last edge of $\Gamma(\sigma)$ must be directed either N or S because its last edge is directed E. In the first case, we are done. For the second case, let p denote the starting point of the next-to-last edge. This point p must lie above the horizontal line l through the last edge, since by assumption, the next-to-last edge is directed S. Note that in fact, p either lies on the N axis or in the *WN* quadrant. In travelling from the origin to p, drawing $\Gamma(\sigma)$ must cross l in the *NW* quadrant, since this portion of the drawing can contain no points of the *NE* quadrant and cannot cross l on the N axis. Let q denote the initial endpoint of the first (closed) edge of $\Gamma(\sigma)$ to reach or cross l. Clearly, q is not at the origin, and hence is not the initial endpoint of the first edge of $\Gamma(\sigma)$. Also, q is clearly *SW* relative to the terminal endpoint of $\Gamma(\sigma)$. Applying the induction hypothesis to the portion of the drawing from q to the terminal endpoint of $\Gamma(\sigma)$, implies that the strict subsequence of σ that corresponds to this portion of the drawing, and hence σ itself, contains the consecutive subsequence *NE*. □

Lemma 4. *Let $\sigma = \sigma'\sigma''$ be a 2D shape such that σ' reaches the NE quadrant. Then σ reaches the NE quadrant.*

Proof. Let $\Gamma(\sigma')$ be a drawing for σ' that reaches the *NE* quadrant. Append to $\Gamma(\sigma')$ one by one segments having the directions specified by σ'', using the rule that each new segment should travel half-way to the closest obstacle in front of it, if there is one; the new segment should travel one unit if there is no obstacle in front of it. Here, the coordinate axes as well as the segments are regarded as obstacles. Clearly, $\Gamma(\sigma'')$ does not leave the *NE* quadrant, and hence $\Gamma(\sigma)$ reaches the *NE* quadrant. □

We are now ready to consider the more general problem of quadrant reachability.

Theorem 2. *A 2D shape σ reaches the NE quadrant if and only if σ contains a canonical sequence of type $\{N,E\}$.*

Proof. We prove first that if σ contains a canonical sequence τ of type $\{N, E\}$, then it reaches the NE quadrant. First, assume $\sigma = \sigma_1 NE \sigma_2$, where σ_1 or σ_2 or both may be the empty string. By Lemma 1, σ_1 has an expanding, extensible drawing $\Gamma(\sigma_1)$. Extend $\Gamma(\sigma_1)$ to a drawing for σ as follows. Append to the terminal point of $\Gamma(\sigma_1)$ a segment oriented N that is so long that (i) its terminal point projects to the N axis and (ii) the resulting drawing is still an expanding drawing. Because the new drawing is expanding, we can append to its terminal point a segment oriented E that is long enough to enter the NE quadrant. Thus, $\sigma_1 NE \subseteq \sigma$ reaches the NE quadrant, and hence by Lemma 4, so does σ. The case $\sigma = \sigma_1 EN \sigma_2$, where NE has been replaced by EN, is handled analogously.

Suppose now that $\sigma = X_1, \ldots, X_n$ reaches the NE quadrant, and let $\Gamma(\sigma)$ be a drawing of σ that terminates at a point in the NE quadrant. Suppose the first segment of $\Gamma(\sigma)$ to intersect the NE quadrant is the ith segment, whose associated direction is X_i. Let σ' denote the initial sequence $X_1, \ldots, X_i$ of σ. Then $\Gamma(\sigma)$ contains a drawing $\Gamma(\sigma')$ of σ'. If $X_i = E$, then by Lemma 3, σ' and hence σ contains the consecutive subsequence NE. Similarly, if $X_i = N$, then σ contains the consecutive subsequence EN. □

Based on the results above and on Lemma 2 it is straightforward to design a linear time algorithm for deciding whether a 2D shape σ with n labels can reach a point p in some quadrant, and if so, for constructing an embedding that reaches p. Namely, suppose p is a point of the NE quadrant. By Theorem 2, p can be reached if and only if σ contains NE or EN consecutively, which can be determined by a linear-time scan of σ. Also, observe that the construction in the proof of sufficiency in Theorem 2 computes the coordinates of the endpoints of the segments of a drawing and requires $O(n)$ time if the real RAM model of computation is adopted. Since the lengths of some segments might require $\Theta(\lg n)$ bits to record, the running time becomes $O(n \lg n)$ for a Turing machine model.

Theorem 3. *Let σ be a 2D shape with n labels, and let p be a point of a quadrant. There exists an algorithm that decides whether σ reaches p and that computes an embedding for σ that reaches p when such an embedding exists. The algorithm runs in $O(n)$ time in the real RAM model of computation.*

4.2. Reachability in 3D

In this section we answer the following question: Given a point p in an octant and a 3D shape σ, can σ reach p? Again, by Lemma 2, this question can be answered by characterizing when σ can reach a given octant. To this aim, we introduce the key concept of a *doubly extensible drawing*. This concept will later be used for the study of shape reachability.

A *doubly extensible* drawing is a drawing in which the first and last segments can be replaced by arbitrarily long segments without creating any intersections within that drawing. Fig. 3 shows an example of a doubly extensible drawing.

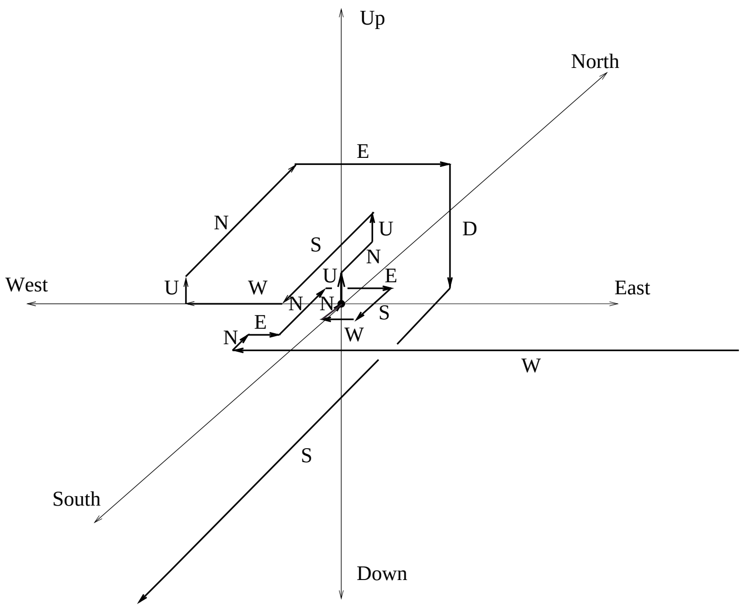

Fig. 3. A doubly expanding drawing for $\sigma = WNENESWNUNUSWUNEDS$.

In the constructions of the lemmas that follow, we use the *reverse* shape σ^r of a shape σ; σ^r is obtained by listing the labels of σ in reverse order and then replacing each label with the label that corresponds to the opposite direction. For example, if $\sigma = WNENESWN$, then listing the elements in reverse order gives $NWSENENW$; replacing each of these labels by its opposite then gives $\sigma^r = SENWSWSE$.

Lemma 5. *Let σ be a shape with n labels such that σ either contains exactly two labels or contains at least two flats. Then σ has a doubly extensible drawing that can be computed in* $\mathrm{O}(n)$ *time in the real RAM model of computation.*

Proof. If σ consists of exactly two labels, then since these labels are orthogonal, the statement of the lemma is clearly true.

Now suppose σ contains at least two flats, and denote the first flat by σ_1. Let m denote the length of σ_1. Thus, $\sigma = \sigma_1\sigma_2$, where σ_2 contains the remaining elements of σ. Thus, the length of σ_2 is $n-m>0$. In Fig. 3, $\sigma_1 = WNENESWN$ and $\sigma_2 = UNUSWUNEDS$.

Before giving in detail the construction for a doubly extensible drawing for σ, we first give an overview. (See Fig. 3.)

The construction of a doubly extensible drawing for σ will resemble the construction in the proof of Lemma 1, but with the following difference. Before, segments were created in the same order in which they appeared in σ. This time, they will be created in a different order. An artist following our construction would place a pen at the origin, draw σ_1 in reverse by creating an expanding drawing of σ_1^r, then lift the pen and return it to the origin, and finally draw σ_2 in the forward direction to form a complete drawing for σ.

Now we describe the construction of a doubly extensible drawing in more detail. To obtain the drawing of σ_1, by Lemma 1, make an expanding drawing $\Gamma(\sigma_1^r)$ for σ_1^r, placing the start point of $\Gamma(\sigma_1^r)$ at the origin and making each segment extend the frontier in its direction of travel by one unit. In the example given by the figure, $\sigma_1^r = SENWSWSE$. To obtain a drawing for σ_1, take the start point of $\Gamma(\sigma_1^r)$ as the endpoint for $\Gamma(\sigma_1)$, and take the endpoint of $\Gamma(\sigma_1^r)$ as the start point for $\Gamma(\sigma_1)$. Thus, the drawing of σ_1 *terminates* at the origin, and the start point of the drawing for σ is *away* from the origin.

Note that the first label of σ_2 is orthogonal to the flat of σ_1. (In the figure, the first label of σ_2 is U.) Starting at the origin, draw a directed segment of unit length for this label. Hence, the segment leaves the plane containing $\Gamma(\sigma_1)$, and the bounding box of the drawing made thus far becomes truly 3D. Continue creating segments for the labels of σ_2 one by one in such a way that each new segment is given a length that increases the extent of the bounding box of the current drawing (which includes $\Gamma(\sigma_1)$) by one unit in the direction of the new segment. This completes the description of the construction.

In order to prove that the computed drawing is doubly extensible, it must be checked that no self-intersections have been created, and that furthermore, the segments that correspond to the first and last elements of σ may be made arbitrarily long without creating intersections.

Let s_i denote the segment corresponding to the ith element of σ, and let d_i denote the ith segment to be *drawn*. Thus, $d_1 = s_m$, since s_m is drawn first, and $d_j = s_j$ for all $j > m$. Let B_0 denote the point at the origin, and let B_j denote the bounding box of $d_1, \dots, d_j$.

For $1 < j \leqslant n$, segment d_j protrudes beyond B_{j-1}. Clearly, $\Gamma(\sigma_1)$ does not intersect either itself, or d_{m+1}, d_{m+2}. For $m+2 \leqslant j \leqslant n$, segment d_j is attached to the end of d_{j-1} that protrudes from B_{j-2}. Hence, $\Gamma(\sigma)$ does not intersect itself, and $s_n = d_n$ may be replaced by an arbitrarily long segment without causing a collision.

It now remains to be shown that segment $s_1 = d_m$ may be replaced by an aribitrarily long segment without creating a collisions, even with a possibly extended s_n.

For a contradiction, suppose extending s_1 causes an intersection with s_j. Clearly, $j > m$ since $\Gamma(\sigma_1^r)$ is extensible, so $s_j = d_j$. Also, j is clearly not equal to $m+1$, $m+2$. Now consider the bounding box B_{j-2}, which contains s_1, and the extension s_1^+ of s_1 from B_{j-2}. The tip of segment s_{j-1} protrudes from B_{j-2} and then sends the segment s_j in an orthogonal direction into s_1^+. Thus, s_{j-2} is parallel to s_1^+, and s_1^+, s_{j-2}, s_{j-1} lie in a common plane. This plane does not contain the origin, since s_1 does not lie on a line through the origin.

Using the same notation for segments and their corresponding elements of σ, let $d_k = s_k$ be the first element in the flat containing s_{j-1}, s_j. Clearly, $k > m+2$ and s_k is orthogonal to s_1^+. Thus, the tip of s_{k-1} protrudes from B_{k-2} in some direction X, so the common X coordinate value of all points on $s_k, \dots, s_j$ is greater than the common X coordinate value of all points on s_1^+, contradicting that s_j intersects s_1^+.

Finally, observe that the drawing algorithm described above is based on a linear scan of σ; while the lengths of some segments might require $\Theta(\lg n)$ bits to record, the time complexity is $O(n)$ for the real RAM model of computation. □

We call the doubly extensible drawing constructed in the proof above a *doubly expanding* drawing.

The next lemma gives a sufficient condition for octant reachability; the proof is based on connecting together singly and doubly expanding drawings for subsequences of shapes and gives rise to a linear time embedding algorithm.

Lemma 6. *Let σ be a* 3D *shape that contains a canonical sequence of type* $\{U,N,E\}$. *Then σ reaches the UNE octant.*

Proof. Assume the hypothesis holds, and let $\hat{U}$, $\hat{N}$, and $\hat{E}$ denote the direction labels U, N, and E of the canonical sequence. Without loss of generality, we may assume that $\hat{U}$, $\hat{N}$ and $\hat{E}$ appear in σ in the order $\hat{U},\hat{E},\hat{N}$ and hence that $\sigma = \sigma_1\hat{U}\sigma_2\hat{E}\sigma_3\hat{N}\sigma_4$, where σ_i $(i = 1,\ldots,4)$ may be empty.

We give an algorithm that takes as input any point p in the UNE octant and computes in linear time an embedding of σ that terminates at p. The algorithm creates separate coordinate systems, called "local coordinate" systems, for various parts of the drawing and then later positions these local coordinate systems into one global coordinate system. Thus, the origin of each local coordinate system will eventually be assigned coordinates in the global coordinate system. This will determine the global coordinates of the various pieces of the drawing.

Compute an extensible drawing (for example, an expanding drawing) for $\sigma_1\hat{U}$, and store the coordinate information for the start point of the edge for $\hat{U}$ in the local coordinate system for $\sigma_1\hat{U}$ as a vector whose tail is at the local origin and whose tip is at the start point of the edge for $\hat{U}$. Compute the bounding box, possibly degenerate, for the portion of $\Gamma(\sigma_1\hat{U})$ corresponding to σ_1. If σ_1 is empty, this bounding box is simply a point, the local origin.

Compute a doubly extensible drawing (for example, a doubly expanding drawing) for $\hat{U}\sigma_2\hat{E}$, as given by the construction in the proof of Lemma 5. Compute the coordinates of the vector whose tail is at the terminal point of the edge for $\hat{U}$ and whose tip is at the start point of the edge for $\hat{E}$. Compute the bounding box, possibly degenerate, for the portion of the doubly expanding drawing corresponding to σ_2. Process $\hat{E}\sigma_3\hat{N}$ similarly.

Finally, compute a drawing $\Gamma(\hat{N}\sigma_4)$ for $\hat{N}\sigma_4$ that *terminates* at the origin of its local coordinate system by computing an extensible (for example, expanding) drawing for $(\hat{N}\sigma_4)^r$ that *starts* at the origin of the local coordinate system. Compute the coordinates of the vector whose tail is at the terminus of the edge of $\hat{N}$ in $\Gamma(\hat{N}\sigma_4)$ and whose tip is at the origin. Compute the bounding box for the portion of the doubly extensible drawing that corresponds to σ_4.

It now remains to connect these drawings together by assigning new lengths to $\hat{U}$, $\hat{E}$, and $\hat{N}$ and placing the local coordinate systems and their drawings into a global coordinate system.

Place the origin of the drawing for σ_1 at the global origin. Compute a length $l_{\hat{U}}$ for $\hat{U}$ such that a vector of length $l_{\hat{U}}$ in the U direction, when added to the first of the four vectors computed above, results in a vector whose U coordinate is positive.

Similarly, compute lengths for $l_{\hat{E}}$ and $l_{\hat{N}}$. Compute by simple arithmetic on known vertex coordinates any increase in these lengths needed to ensure that the bounding boxes computed above are well separated in the final drawing.

Finally, compute three scale factors, one for each pair of opposite labels, and apply them to the lengths of the edges with these labels so that the drawing terminates at the desired point p. Since the drawing undergoes a linear transformation, it remains crossing free.

The correctness of the algorithm follows from Lemma 5 and the separation of the bounding boxes. □

Note that before the application of the scale factors in the proof just given, the construction had yielded a drawing whose segments had integer lengths and whose bounding box was of length $O(n)$ in each dimension, where n is the number of elements of σ.

The next lemma proves that the condition of Lemma 6 is also necessary for octant reachability.

Lemma 7. *Let σ be a 3D shape that reaches the UNE octant. Then σ contains a canonical sequence of type $\{U,N,E\}$.*

Proof. First, we begin by proving a weaker form of this:

If σ is a 3D shape that can reach the open *NE* quarter-space (i.e., the union of the *UNE* octant, the *DNE* octant, and the *NE* quadrant), then there is a canonical sequence of type $\{N,E\}$ in σ.

To prove this, consider a drawing $\Gamma(\sigma)$ that reaches the open *NE* quarter-space. Let (a,b) be the first segment of Γ that enters this quarter-space. Suppose, without loss of generality, that (a,b) has direction E. Since this E segment lies in the N half-space, there must be a previous N segment, (c,d), that enters this half-space. If the chosen E and N segments do not share a flat, we are done since the labels of (a,b) and of (c,d) form a canonical sequence in σ. So assume that they do share a flat. Let Π be the *NSEW* plane containing the flat. Note that b is in the (open) *NE* quadrant of Π. We claim that c is a point in the closed *SW* quadrant of Π. This is because d, which is in the (open) N half-plane but not in the (open) *NE* quadrant, must lie in the (open) *NW* quadrant or on the N axis. Imagine moving the origin to c, and let σ' denote the subsequence of σ corresponding to the portion of $\Gamma(\sigma)$ from c to b. Let $\Gamma(\sigma')$ denote the drawing of σ' inherited from $\Gamma(\sigma)$. Since $\Gamma(\sigma')$ travels to the *NE* quadrant, by Lemma 3 one can find a consecutive pair *NE* or *EN* in σ' and hence also in σ. This completes the proof of the preliminary result.

To obtain a proof of the original statement, consider a drawing $\Gamma(\sigma)$ that reaches the *ENU* octant. Let (p,q) be the first segment of $\Gamma(\sigma)$ that enters this octant, and suppose without loss of generality that (p,q) has direction U. Since this U segment lives in the *NE* quarter-space, there must be a previous segment, (r,s), that enters this quarter-space. Assume without loss of generality that (r,s) is an N segment. We remark that:

- Since s cannot be in the *ENU* octant (otherwise segment (p,q) would not be the segment entering this octant), s is a point in the *END* octant or the *NE* quadrant.
- Since (r,s) is an N segment, r is a point in the *ESD* octant, or the *SE* quadrant or the *DE* quadrant.

Let σ' denote the portion of σ that corresponds to the portion of $\Gamma(\sigma)$ from the origin to s, and let $\Gamma(\sigma')$ denote the drawing of σ' inherited from $\Gamma(\sigma)$. Since $\Gamma(\sigma')$ reaches the *EN* quarter-space, by the preliminary result just proved there is a canonical sequence $\{\bar{E},\bar{N}\}$ in σ'. Let σ'' denote the portion of σ that corresponds to the portion of $\Gamma(\sigma)$ from r to q, and let $\Gamma(\sigma'')$ denote the drawing of σ'' inherited from $\Gamma(\sigma)$. Imagine moving the origin to r. Since $\Gamma(\sigma'')$ reaches the *UN* quarter-space, by the preliminary result just proved there is a canonical sequence $\{\hat{U},\hat{N}\}$ in σ''.

Note that $\Gamma(\sigma')$ and $\Gamma(\sigma'')$ share only one segment, namely (r,s). We can therefore partition σ as follows: $\sigma=\sigma_1\ell_{r,s}\sigma_2$, where $\ell_{r,s}$ is the N label associated with segment (r,s); hence, $\sigma'=\sigma_1\ell_{r,s}$ and $\sigma''=\ell_{r,s}\sigma_2$. To summarize, σ contains a canonical sequence $\{\bar{E},\bar{N}\}$ followed by $\ell_{r,s}$ followed by a canonical sequence $\{\hat{U},\hat{N}\}$, where $\ell_{r,s}$ might be equal to $\bar{N}$ or to $\hat{N}$. In order to show that σ has a canonical sequence of type $\{U,N,E\}$, we start by observing that $\bar{E}$ and $\hat{U}$ belong to different flats of σ, for if they were in the same flat it would also contain the intervening $\ell_{r,s}=N$, which is impossible.

We construct a canonical sequence of type $\{U,N,E\}$ by considering the following (not mutually exclusive) cases:

- Label $\bar{N}$ precedes $\bar{E}$ in σ. In this case $\{\bar{E},\bar{N},\hat{U}\}$ is a canonical sequence appearing in the order $\bar{N},\bar{E},\hat{U}$.
- Label $\hat{N}$ follows $\hat{U}$ in σ. In this case $\{\bar{E},\hat{N},\hat{U}\}$ is a canonical sequence appearing in the order $\bar{E},\hat{U},\hat{N}$.
- The labels appear in the order $\bar{E},\bar{N},\hat{N},\hat{U}$. If $\bar{N}$ and $\hat{U}$ are in different flats, then $\bar{E},\bar{N},\hat{U}$ form a canonical sequence. Otherwise, if $\bar{E}$ and $\hat{N}$ are in different flats then $\bar{E},\hat{N},\hat{U}$ form a canonical sequence. If neither of these cases applies then $\bar{N}$ and $\hat{U}$ share a flat and $\bar{E}$ and $\hat{N}$ share a flat; thus all four labels lie in one flat, which is impossible. □

Lemmas 6 and 7 can be summarized as follows.

Theorem 4. *Let σ be a 3D shape. Shape σ reaches the UNE octant if and only if it contains a canonical sequence of type $\{U,N,E\}$.*

Theorem 4 and Lemma 2 yield a linear time algorithm for deciding whether a 3D shape σ with n labels can reach a point in an octant. Given a point p that can be reached, the drawing algorithm in the proof of Lemma 6, which computes the coordinates of the endpoints of the segments, requires $O(n)$ time if the real RAM model of computation is adopted. Since the lengths of some segments might require $\Theta(\lg n)$ bits to record, the running time becomes $O(n\lg n)$ for a Turing machine model.

Theorem 5. *Let σ be a 3D shape with n labels, and let p be a point of an octant. There exists an algorithm that decides whether σ reaches p and that computes an embedding for σ that reaches p when such an embedding exists. The algorithm runs in $O(n)$ time in the real RAM model of computation.*

Proof. First, we show a linear time method for testing whether σ reaches the *UNE* octant. The methods for the other octants are analogous.

It follows from Theorem 4 that σ reaches the *UNE* octant if and only if σ contains a subsequence X_i, X_j, X_k of elements such that $X_i X_j X_k$ is one of the six permutations of the labels U, N, E and such that if two of X_i, X_j, X_k are in the same flat, then they are adjacent. Since no flat of σ can contain all three of the labels U, N and E, such a subsequence X_i, X_j, X_k must satisfy the condition that either

(i) no flat of σ contains two of X_i, X_j, and X_k, or
(ii) $j = i + 1$ or $j = k - 1$.

We now describe an algorithm that tests for the existence of such a subsequence X_i, X_j, X_k by making several scans through σ. For simplicity of explanation, we do not combine the scans as this does not affect the asymptotic analysis. As a preliminary step, scan σ to ensure that each of U, N, and E occurs at least once in σ. If not, then σ cannot reach the *UNE* octant.

Next, consider the permutation *UNE* of $\{U, N, E\}$, and look for a subsequence X_i, X_j, X_k satisfying either (i) or (ii) above and such that $X_i = U$, $X_j = N$, and $X_k = E$. The remaining five permutations will be processed in the same manner.

To test for X_i, X_j, X_k satisfying (i), observe that if such a subsequence exists, and if X_i is not the first occurrence of U in σ, then (i) still holds if X_i is replaced by the first occurrence of U in σ. Hence, let $\hat{X}_i$ denote the first occurrence of U. Likewise, (i) still holds if X_j is replaced by the first occurrence of N in a flat following the ones that contain X_i. Let $\hat{X}_j$ denote this occurrence of N. Finally, (i) stillholds if X_k is replaced by the first occurrence of E in a flat after the ones containing X_j. Let $\hat{X}_k$ denote this occurrence.

Scan for $\hat{X}_i$, $\hat{X}_j$ and $\hat{X}_k$. Shape σ contains a subsequence U, N, E satisfying (i) if and only if the scan successfully finds $\hat{X}_i$, $\hat{X}_j$ and $\hat{X}_k$.

If the scan is not successful, then test for the existence of a subsequence X_i, X_j, X_k satisfying (ii). Such a subsequence must either satisfy $j = i + 1$ or $j = k - 1$. If such a subsequence exists such that $j = i + 1$, then (ii) still holds if X_i and X_j are replaced by the first occurrence of the subsequence *UN* such that N immediately follows U. Let $\hat{X}_i \hat{X}_{i+1}$ denote this occurrence. Furthermore, (ii) still holds if X_k is replaced by the first occurrence of E in a flat following the one containing $\hat{X}_i \hat{X}_{i+1}$. Let $\hat{X}_k$ denote this occurrence.

Scan for $\hat{X}_i \hat{X}_{i+1} \hat{X}_k$. Shape σ contains a subsequence U, N, E such that N immediately follows U if and only if the scan successfully finds $\hat{X}_i$, $\hat{X}_{i+1}$ and $\hat{X}_k$.

If this scan is not successful, to complete the search for a subsequence U, N, E satisfying (ii), scan for a subsequence $\hat{X}_i, \hat{X}_{k-1}, \hat{X}_k$, where $\hat{X}_i$ is the first occurrence of U in σ, and where $\hat{X}_{k-1} \hat{X}_k$ is the first occurrence of the subsequence *NE* such that E immediately follows N in a flat following the ones that contain $\hat{X}_i$.

If this last scan is not successful, then σ does not contain a subsequence U, N, E satisfying either (i) or (ii).

In a similar manner, test the other five permutations of U, N, E for one that corresponds to a subsequence of σ that satisfies either (i) or (ii).

Shape σ reaches the UNE octant if and only if one of these tests is successful. Each of these six tests involves a small, constant number of scans that can be done in $O(n)$ time and space within the real RAM model of computation. A successful test produces the labels $\{\hat{U}, \hat{N}, \hat{E}\}$ used by the linear time algorithm in the proof of Lemma 6. This algorithm then produces in linear time an embedding that terminates at p. $\square$

5. Higher dimensions

Our characterization of shapes that can reach a given point generalizes to arbitrary dimension. The definition of a flat remains the same: it is a maximal 2D object. The intuition for why the notion of a flat is also a key concept for the arbitrary dimensional case is the following. Two segments that intersect determine a plane. If they are segments that we wish to make long in a drawing, and they do not belong to the same flat, then some perturbation should remove the intersection. If they lie in the same flat, we cannot necessarily make them long without intersection. The condition that two canonical segments in the same flat must be adjacent means that they can be made long without colliding with each another.

Of course, in the context of R^d, a canonical sequence should be allowed to contain as many as d orthogonal direction vectors. As the intuition given above suggests, the definiton of canonical sequence is otherwise unchanged.

Definition. Let σ be a shape containing labels from d pairs of opposite direction labels. A not necessarily consecutive subsequence $\tau \subseteq \sigma$, where τ consists of $k \geqslant 1$ elements, is a *canonical sequence* provided that:

- any two elements of τ indicate mutually orthogonal directions (hence $1 \leqslant k \leqslant d$); and
- if a flat F of σ contains one or more labels of τ, then $\tau \cap F$ forms a consecutive subsequence of σ.

Here we adopt the usual notation for dimension greater than three, regarding a point as a d-tuple of coordinate values, called the $X_1, \ldots, X_d$ coordinates, respectively. We now refer to the full, two-way infinite coordinate axes as the $X_1^{\pm}, \ldots, X_d^{\pm}$ axes. The positive semi-infinite axes are denoted $X_1^+, \ldots, X_d^+$. The set of points of R^d whose coordinates are all positive is called the $X_1^+, \ldots, X_d^+$ space. As before, for concreteness, we assume that the point we wish to reach lies in this space.

Theorem 6. *A shape σ reaches the $X_1^+, \ldots, X_d^+$ space if and only if σ contains a canonical sequence of type $\{X_1^+, \ldots, X_d^+\}$.*

Proof. The construction for the proof of the sufficiency of the condition is exactly as for the $d=3$ case: doubly-expanding drawings can be made to connect long segments realizing the canonical labels as in the proof of Lemma 6.

To prove the necessity of the condition for dimension $d>3$, consider a drawing $\Gamma(\sigma)$ that reaches the $X_1^+,\dots,X_d^+$ space. Let (a_1,b_1) be the first segment of $\Gamma(\sigma)$ to have a tip with any positive coordinate. Permute the indices of the axes if necessary so that the X_1 coordinate of b_1 is positive. Note that b_1 has no other positive-valued coordinate. This is because travelling along a segment whose direction label is $X_i^\pm$ changes only the X_i coordinate. Thus (a_1,b_1) must have direction label X_1^+, since travelling along (a_1,b_1) changes the X_1 coordinate from negative or 0 to positive value; all other coordinates of points on (a_1,b_1) must have 0 or negative value.

For $2\leqslant i\leqslant d$, let (a_i,b_i) be the first segment to have a tip with positive values for coordinates $X_1,\dots,X_{i-1}$ plus an additional coordinate with positive value. By the observation made above for b_1, this additional positive-valued coordinate is unique. Permute the indices of the $X_i^\pm,\dots,X_d^\pm$ axes if necessary so that the additional positive-valued coordinate is the X_i coordinate.

We remark that:

- b_{i-1} has positive values for its $X_1,\dots,X_{i-1}$ coordinates and negative or 0 values for its remaining coordinates.
- Since (a_{i-1},b_{i-1}) has direction label X_{i-1}^+, point a_{i-1} has positive values for its $X_1,\dots,X_{i-2}$ coordinates if any, and negative or 0 values for its remaining coordinates.

For $1\leqslant i\leqslant d$, denote by σ_i the subsequence of σ that corresponds to the portion of $\Gamma(\sigma)$ from the origin to b_i, and let $\Gamma(\sigma_i)$ denote the portion of $\Gamma(\sigma)$ from the origin to b_i.

We prove by induction that for $1\leqslant i\leqslant d$, shape σ_i, and hence σ, contains a canonical sequence of type $\{X_1^+,\dots,X_i^+\}$. The statement holds trivially for $i=1$. Suppose the statement holds for all values less than some $i>1$, and let us prove that the statement holds for i.

Since $\Gamma(\sigma_{i-1})$ reaches a point b_{i-1} whose first $i-1$ coordinates are all positive, then by the induction hypothesis there is a canonical sequence of type $\{X_1^+,\dots,X_{i-1}^+\}$ in σ_{i-1}. Let $\{\bar{X}_1^+,\dots,\bar{X}_{i-1}^+\}$ in σ_{i-1} be such a sequence.

If the labels of σ corresponding to (a_{i-1},b_{i-1}) and (a_i,b_i) do not lie in the same flat, then adding the label $\bar{X}_i^+$ corresponding to (a_i,b_i) to the canonical sequence for σ_{i-1} gives a canonical sequence for σ_i.

If the elements of σ corresponding to (a_{i-1},b_{i-1}) and (a_i,b_i) *do* lie in a common flat, this flat must be an $X_{i-1}^\pm X_i^\pm$ flat.

Let σ'' denote the subsequence of σ corresponding to the portion of $\Gamma(\sigma)$ from a_{i-1} to b_i, and let $\Gamma(\sigma'')$ denote the drawing of σ'' inherited from $\Gamma(\sigma)$. Since (a_{i-1},b_{i-1}) and (a_i,b_i) share a flat, Theorem 2 concerning quadrant reachability in 2D applies: σ'' must contain a canonical sequence $\{\hat{X}_{i-1}^+,\hat{X}_i^+\}$ of type $\{X_{i-1}^+,X_i^+\}$. This is because b_i has positive X_{i-1} and X_i coordinates whereas a_{i-1} has 0 or negative values for its X_{i-1} and X_i coordinates. If $i=2$, the argument thus far provides a complete proof that σ_2 contains a canonical sequence of type $\{X_1^+,X_2^+\}$.

Now we complete the proof for the case $i>2$. Note that $\Gamma(\sigma_{i-1})$ and $\Gamma(\sigma'')$ share only one segment, namely (a_{i-1},b_{i-1}). We can therefore partition σ_i as follows: $\sigma=\sigma_1 \ell_{(a_{i-1},b_{i-1})}\sigma_2$, where $\ell_{(a_{i-1},b_{i-1})}$ is the X_{i-1}^+ label associated with segment (a_{i-1},b_{i-1}); hence, $\sigma_{i-1}=\sigma_1\ell_{(a_{i-1},b_{i-1})}$ and $\sigma''=\ell_{(a_{i-1},b_{i-1})}\sigma_2$.

To summarize, σ_i contains a canonical sequence $\{\bar{X}_1^+,\dots,\bar{X}_{i-1}^+\}$, followed by $\ell_{(a_{i-1},b_{i-1})}$, followed by a canonical sequence $\{\hat{X}_{i-1}^+,\hat{X}_i^+\}$, where $\ell_{(a_{i-1},b_{i-1})}$ might be equal to $\bar{X}_{i-1}^+$, or to $\hat{X}_{i-1}^+$, or to both, but not to $\bar{X}_i^+$.

In order to show that σ_i has a canonical sequence of type $\{X_1^+,\dots,X_i^+\}$, we first observe that none of the $\bar{X}_1^+,\dots,\bar{X}_{i-2}^+$ shares a flat with $\hat{X}_i^+$; otherwise, the shared flat would also contain the intervening label $\ell_{(a_{i-1},b_{i-1})}=X_{i-1}^+$, which is impossible.

We construct a canonical sequence of type $\{X_1^+,\dots,X_i^+\}$ for σ_i by considering the following (not mutually exclusive) cases:

- Among the canonical labels for σ_{i-1}, label $\bar{X}_{i-1}^+$ is not the last to occur in σ_{i-1}. In this case $\{\bar{X}_1^+,\dots,\bar{X}_{i-1}^+,\hat{X}_i^+\}$ is a canonical sequence for σ_i, with some canonical label $\bar{X}_j$ occuring between $\bar{X}_{i-1}^+$ and $\hat{X}_i^+$ in σ_i.
- Label $\hat{X}_{i-1}^+$ follows $\hat{X}_i^+$ in σ. In this case $\{\bar{X}_1^+,\dots,\bar{X}_{i-2}^+,\hat{X}_{i-1}^+,\hat{X}_i^+\}$ is a canonical sequence for σ_i; in σ_i, label $\hat{X}_{i-1}^+$ is preceded by $\hat{X}_i^+$, which is preceded by the remaining canonical labels in some order.
- The order of the labels in σ_i is such that $\hat{X}_i^+$ is preceded by $\hat{X}_{i-1}^+$, which is equal to or preceded by $\bar{X}_{i-1}^+$, which is preceded by $\bar{X}_j^+$ for some $j<i-1$, which is preceded by any remaining canonical labels in some order. If $\bar{X}_{i-1}^+$ and $\hat{X}_i^+$ are in different flats, then $\bar{X}_1^+,\dots,\bar{X}_{i-1}^+,\hat{X}_i^+$ form a canonical sequence. Otherwise, if $\bar{X}_j^+$ and $\hat{X}_{i-1}^+$ are in different flats, then $\bar{X}_1^+,\dots,\bar{X}_j^+,\hat{X}_{i-1}^+,\hat{X}_i^+$ form a canonical sequence. If neither of these cases applies, then $\bar{X}_{i-1}^+$ and $\hat{X}_i^+$ share a flat and $\bar{X}_j^+$ and $\hat{X}_{i-1}^+$ share a flat; thus all four of these labels, at least three of which are distinct, lie in one flat, which is impossible.

This completes the proof. □

6. Conclusion and open problems

We have given a combinatorial characterization, together with linear time recognition and embedding algorithms, for those instances of the 3D shape reachability problem that admit solutions. Our constructive embedding algorithms produce drawings that, before scaling to reach particular points, have segments of positive integer lengths and that lie in bounding boxes of length O(n) on each side. Our results may enable the extension to 3D of the classical *topology-shape-metrics* approach to 2D rectilinear layout problems.

We have also generalized our 3D characterization to arbitrary dimension d.

Several issues remain open. We mention some that in our opinion are interesting.

1. Study the reachability problem under the additional constraint that the lengths of some of the segments are given.

2. Study the reachability problem in the presence of obstacles.
3. Study the reachability problem for paths that must be embedded in a fixed grid and that are to reach specified points in this grid.
4. Study the reachability problem for a larger set of direction vectors.
5. Our recognition algorithm generalizes to arbitrary dimension d. A naive generalization would yield an algorithm whose running time is worse than exponential in d, as it would treat as separate cases the $d!$ possible ways in which the canonical labels could appear in a canonical sequence. Find a recognition algorithm with running time as low as possible.

In the case of item 3, it is no longer true that a shape can reach a point if and only if it can reach all the other points in the same octant. One can no longer construct a drawing and then scale it so that its terminus is located at some specified point p, as scaling would move some edges off grid lines.

Note also that for drawings in a fixed grid, bounding box volume becomes an issue. While expanding and doubly expanding drawings provide simple constructions, they may waste space.

References

[1] T.C. Biedl, Heuristics for 3D-orthogonal graph drawings, in: Proc. 4th Twente Workshop on Graphs and Combinatorial Optimization, 1995, University of Twente, Enschede, The Netherlands, 7–9 June 1995, pp. 41–44.

[2] T. Biedl, T. Shermer, S. Wismath, S. Whitesides, Orthogonal 3-D graph drawing, J. Graph Algorithms Appl. 3 (4) (1999) 63–79.

[3] R.F. Cohen, P. Eades, T. Lin, F. Ruskey, Three-dimensional graph drawing, Algorithmica 17 (2) (1997) 199–208.

[4] G. Di Battista, P. Eades, R. Tamassia, I. Tollis, Graph Drawing, Prentice-Hall, Englewoods Cliffs, NJ, 1999.

[5] G. Di Battista, M. Patrignani, F. Vargiu, A split and push approach to 3D orthogonal drawing, J. Graph Algorithms Appl. 4(3) (2000) 105–133, in: S.H. Whitesides, G. Liotta (Eds.), Special Issue on Selected Papers from the 1998 Symposium on Graph Drawing.

[6] G. Di Battista, L. Vismara, Angles of planar triangular graphs, SIAM J. Discrete Math. 9 (3) (1996) 349–359.

[7] K.E. Drexler, Nanosystems Molecular Machinery, Manufacturing and Computation, Wiley, New York, 1992.

[8] P. Eades, C. Stirk, S. Whitesides, The techniques of Kolmogorov and Bardzin for three dimensional orthogonal graph drawings, Inform. Process. Lett. 60 (1996) 97–103.

[9] P. Eades, A. Symvonis, S. Whitesides, Three dimensional orthogonal graph drawing algorithms, Discrete Appl. Math. 103 (2000) 55–87.

[10] A. Garg, New results on drawing angle graphs, Comput. Geom. Theory Appl. 9 (1–2) (1998) 43–82, in: G. Di Battista, R. Tamassia (Eds.), Special Issue on Geometric Representations of Graphs.

[11] A. Papakostas, I.G. Tollis, Algorithms for incremental orthogonal graph drawing in three dimensions, J. Graph Algorithms Appl. 3 (4) (1999) 81–115.

[12] R. Tamassia, On embedding a graph in the grid with the minimum number of bends, SIAM J. Comput. 16 (3) (1987) 421–444.

[13] V. Vijayan, Geometry of planar graphs with angles, in: Proc. 2nd Ann. ACM Symp. Comput. Geom. 1986, Yorktown Heights NY, 2–4 June 1986, pp. 116–124.

[14] G. Vijayan, A. Wigderson, Rectilinear graphs and their embeddings, SIAM J. Comput. 14 (1985) 355–372.

[15] C. Ware, G. Franck, Viewing a graph in a virtual reality display is three times as good as a 2D diagram, in: Proc. IEEE Conf. on Visual Languages, St. Louis, Missouri, USA, October 1994, pp. 182–183.
[16] D.R. Wood, Three-dimensional orthogonal graph drawing, Ph.D. Thesis, School of Computer Science and Software Engineering, Monash University, 2000.
[17] D.R. Wood, An algorithm for three-dimensional orthogonal graph drawing, in: Graph Drawing (Proc. GD '98), Lecture Notes in Computer Science, Vol. 1547, Springer, Berlin, 1998, pp. 332–346.
[18] D.R. Wood, Multi-dimensional orthogonal graph drawing with small boxes, in: Graph Drawing (Proc. GD '99), Lecture Notes in Computer Science, Vol. 1731, Springer, Berlin, 1999, pp. 311–322.