Discrete Comput Geom 38:615–637 (2007)
DOI: 10.1007/s00454-007-1353-4

Discrete & Computational
Geometry
© 2007 Springer Science+Business Media, Inc.

Plane Embeddings of Planar Graph Metrics*

MohammadHossein Bateni,[1] Erik D. Demaine,[2] MohammadTaghi Hajiaghayi,[2,3] and Mohammad Moharrami[1]

[1]Department of Computer Engineering, Sharif University of Technology,
Azadi St., Tehren, Iran
{bateni,mmoharrami}@gmail.com

[2]Computer Science and Artificial Intelligence Laboratory, MIT,
32 Vassar St., Cambridge, MA 02139, USA
{edemaine,hajiagha}@mit.edu

[3]Department of Computer Science, Carnegie Mellon University,
4109 Wean Hall, Pittsburgh, PA 15213, USA

Abstract. Embedding metrics into constant-dimensional geometric spaces, such as the Euclidean plane, is relatively poorly understood. Motivated by applications in visualization, ad-hoc networks, and molecular reconstruction, we consider the natural problem of embedding shortest-path metrics of unweighted planar graphs (*planar graph metrics*) into the Euclidean plane. It is known that, in the special case of shortest-path metrics of trees, embedding into the plane requires $\Theta(\sqrt{n})$ distortion in the worst case [M1], [BMMV], and surprisingly, this worst-case upper bound provides the best known approximation algorithm for minimizing distortion. We answer an open question posed in this work and highlighted by Matoušek [M3] by proving that some planar graph metrics require $\Omega(n^{2/3})$ distortion in any embedding into the plane, proving the first separation between these two types of graph metrics. We also prove that some planar graph metrics require $\Omega(n)$ distortion in any crossing-free straight-line embedding into the plane, suggesting a separation between low-distortion plane embedding and the well-studied notion of crossing-free straight-line planar drawings. Finally, on the upper-bound side, we prove that all outerplanar graph metrics can be embedded into the plane with $O(\sqrt{n})$ distortion, generalizing the previous results on trees (both the worst-case bound and the approximation algorithm) and building techniques for handling cycles in plane embeddings of graph metrics.

* A preliminary version of this paper appeared in *Proceedings of the 22nd Annual ACM Symposium on Computational Geometry*, June 2006, pages 197–206. The second and third authors were supported in part by the NSF under Grant Number ITR ANI-0205445. The third author was also supported in part by the Institute for Theoretical Physics and Mathematics (IPM) under Grant Number CS1383-2-02.

1. Introduction

Metric embedding is a mathematical study arising out of the classic field of multidimensional scaling, originally motivated by such applications as visualization, compression, clustering, and nearest-neighbor searching [S1], [S2], [K1], [K2], [W], and more recently finding applications in geometric reconstruction of ad-hoc wireless sensor networks [ČHH], [SRB], [PCB] and molecular structure of proteins [BKL], [CH], [H]. Roughly speaking, the goal is to embed a given metric (matrix of pairwise distances among n points) into a target space while minimizing the maximum additive or multiplicative error, called *distortion*, introduced in the distances.[1] Of particular interest in many of these applications is embedding into low-dimensional geometric spaces, typically Euclidean. For example, in visualization, the natural target spaces are two- and three-dimensional Euclidean space, for display on an LCD panel or a holographic display. Similarly, in ad-hoc wireless sensor networks and molecule structure of proteins, two- and three-dimensional Euclidean spaces are the natural spaces inhabited by these objects, so embedding into these spaces corresponds to reconstructing the original object.

Yet despite persistent efforts by many researchers leading to many recent results about embedding—see, e.g., [IM] for a survey—we remain in the dark about most aspects of embedding into low-dimensional geometric spaces. Even when the target space is the one-dimensional line, little is known. For example, when the given metric is the shortest-path metric of an (unweighted) tree, the best known approximation factor for multiplicative distortion is $\tilde{O}(n^{1/3})$ (improving on the $O(n^{1/2})$-approximation for general graphs) [BDG$^+$], and it is unknown whether it is possible to achieve a factor of $n^{o(1)}$; see [BCIS]. (On the other hand, additive distortion is less interesting in this context: there is an $O(1)$-approximation for embedding a general metric into the line [HIL].)

Even less is known about embedding into the two-dimensional plane. For additive distortion and the Euclidean plane, the only $o(n)$-approximation algorithm runs in pseudo-quasipolynomial time [BDHI]. For additive distortion and the ℓ_1 plane, there is a polynomial-time $O(1)$-approximation [Bă]. However, for the most natural case of multiplicative distortion, nothing beyond a trivial $O(n)$-approximation is known for embedding general metrics into any ℓ_s plane. (The trivial $O(n)$-approximation follows because any metric can be embedded with $O(n)$ distortion even into the line [M1].) Essentially the only result is by Babilon et al. [BMMV]: an $O(\sqrt{n})$-approximation for embedding the shortest-path metric of an (unweighted) tree into the Euclidean plane. In fact, this approximation result comes trivially from a worst-case bound: all such metrics can be embedded into the Euclidean plane with multiplicative distortion $O(\sqrt{n})$ (improving on an earlier bound by Gupta [G]), and thus such an embedding is an $O(\sqrt{n})$-approximation on the optimal multiplicative distortion. This bound is tight in the worst case for trees (or even stars), but can we generalize beyond trees?

One of the major open questions about embedding into the Euclidean plane, posed in [M3] and [BMMV] and addressed in this paper, is whether there is a worst-case bound on

[1] More precisely, given a metric M on a point set V and a target space of d-dimensional ℓ_s space, the goal is to find a mapping $f\colon V \to \mathbb{R}^d$ to either approximately minimize *additive distortion* $\max_{v,w\in V} |\, \|f(v) - f(w)\|_s - M[v, w]\,|$, or approximately minimize *multiplicative distortion* $\max_{v,w\in V} \|f(v) - f(w)\|_s / M[v, w]$ subject to $\|f(v) - f(w)\|_s \geq M[v, w]$ (noncontractiveness). (In d-dimensional ℓ_s space, distances and lengths are measured according to the ℓ_s norm $\|(x_1, \ldots, x_d)\|_s = \sqrt[s]{x_1^s + \cdots + x_d^s}$.)

distortion better than $O(n)$ for shortest-path metrics of all (unweighted) *planar* graphs, in particular whether there is an $O(\sqrt{n})$ bound like trees. It has been known for over 15 years that the shortest-path metrics of some nonplanar graphs, such as K_5 with each edge subdivided $n/10$ times, requires multiplicative distortion $\Theta(n)$ [M1], [IM]. However, for shortest-path metrics of (unweighted) planar graphs, which we call *planar graph metrics*, no worst-case lower bound better than $\Omega(\sqrt{n})$ is known. One of the results of this paper is a stronger, more complicated lower bound of $\Omega(n^{2/3})$ for the shortest-path metrics of a family of planar graphs called "globe graphs". Interestingly, this graph family has not been considered before (at least in the context of embedding) and the lower-bound argument uses topological graph theory, in contrast to the standard packing arguments in some embedding lower bounds.

Planar graph metrics are a natural family of metrics to embed: many structures of interest, such as communication networks and traffic networks, are planar (or nearly planar) graphs, and low-distortion embedding of such structures is important, e.g., for visualization. There is also a wealth of knowledge about *crossing-free straight-line drawings* of planar graphs into the plane, originating with Tutte's embedding theorem [T] and studied more recently in the context of graph drawing (see, e.g., [TBET]). It is natural to consider to what extent such embeddings can preserve approximate distances as well. Another result of this paper is that, in fact, $\Omega(n)$ distortion can be required if we restrict attention to crossing-free straight-line embeddings of planar graph metrics, even when the graph has bounded treewidth.

One intriguing aspect of the problem of embedding planar graph metrics into the plane is that, in the context of embeddings, planar graphs usually behave the same as graphs excluding any fixed minor. For example, the same upper bound of $O(\sqrt{\log n})$ distortion is known for embedding into Euclidean space [R], [M2], the same upper bound of $O(1)$ distortion is known for embedding into $O(\log n)$-dimensional ℓ_∞ space [KLMN], and the same upper bound of $O(1)$ distortion is conjectured for embedding into ℓ_1 space [GNRS]. (All of these bounds require a superconstant number of dimensions.) While we prove in this paper an $\Omega(n^{2/3})$ lower bound on the distortion required to embed a planar graph metric into a plane, a sublinear upper bound is still quite possible. In contrast, the subdivided K_5 graph mentioned above excludes a fixed minor (e.g., $K_{3,3}$), yet its shortest-path metric requires $\Omega(n)$ distortion when embedded into the plane. Therefore, any sublinear upper bound on distortion for planar graph metrics will need substantially different techniques compared with previous approaches, which seem to apply equally well to planar graphs and graphs excluding a fixed minor.

Toward this goal, we develop new techniques for embedding into the plane with sublinear distortion. Specifically, we prove an $O(\sqrt{n})$ upper bound on distortion for the shortest-path metrics of (unweighted) *outerplanar graphs*, i.e., graphs that can be drawn in the plane without crossings and with all vertices on the outer face. Interestingly, this result does not take the obvious approach of embedding the dual tree of an outerplanar graph; rather, it uses a breadth-first-search tree. The main challenge is in handling the cycles in the graph.

Our Results. In more detail, the main results of this paper are as follows.

On the lower-bound side, we prove that some planar graph metrics require $\Omega(n^{2/3})$ distortion in any embedding into the plane, and prove that some planar graph metrics require $\Omega(n)$ distortion in any crossing-free straight-line embedding into the plane. These

results imply that embedding planar graph metrics into the plane with sublinear distortion is different from planar drawings of planar graphs, suggesting that we cannot apply the well-studied tools from that domain. In fact, our $\Omega(n^{2/3})$ lower bound builds on the $\Omega(n)$ crossing-free lower bound, extending the structure of the underlying planar graph substantially to force a "planar-like" drawing. We use topological graph theory to reason about which vertices must be embedded "interior" to cycles in the graph, even though these cycles may embed as curves with many self-crossings and crossings of each other. To help cope with this difficulty, we use a combinatorial approach to eliminate some of the crossings.

On the upper-bound side, we prove that every outerplanar graph metric can be embedded into the plane with $O(\sqrt{n})$ distortion. In particular, this result improves the best approximation factor for minimum-distortion embedding of outerplanar graph metrics from $O(n)$ to $O(\sqrt{n})$, which is the only nontrivial such result for multiplicative distortion into the Euclidean plane other than trees. The result uses the following ideas. First, we decompose the graph into a "tree of cycles", cutting apart faces so that they do not share edges. We cannot use the dual tree itself: the main challenge is to handle the cycles in the graph correctly. Second, we perform a breadth-first search (BFS) in this graph, partitioning the graph into a BFS tree and several nontree edges. Third, we use a modification of the $O(\sqrt{n})$-distortion embedding for trees so as to embed the BFS tree without stretching the nontree edges too much. We modify the input to the algorithm, augmenting the BFS tree with extra paths in key locations to force desired gaps in the output, and we modify the output of the algorithm, shifting vertices to close unwanted gaps in the output.

2. Lower Bound for Planar Graph Embedding

In this section we prove a separation between embedding trees and embedding planar graphs into the plane. Namely, we prove that embedding some planar graphs with n vertices into the plane requires distortion $\Omega(n^{2/3})$, whereas trees can be embedded into the plane with distortion $O(n^{1/2})$. Along the way, we prove that every crossing-free straight-line embedding of some planar graphs with n vertices into the plane has distortion $\Omega(n)$.

Throughout this section, we use the odd-even notion of "inside" and "outside" of self-crossing polygons. Namely, for a point v not on the boundary of a (potentially self-crossing) polygon P, we count the number of times that an arbitrary half-infinite ray emanating from v properly crosses the polygon P; we do not count when the ray "grazes" the polygon (with the polygon locally on one side of the ray). If this count is odd, the point v is *inside* the polygon P; otherwise, v is *outside* P. Also, let $|P|$ denote the number of vertices of polygon P.

Lemma 1. *For any point v inside a (potentially self-crossing) polygon P in the plane, if the distance from v to every vertex of P is at least r, then there is an edge of P of length $\Omega(r/|P|)$.*

Proof. Let $u_1, u_2, \ldots, u_{|P|}$ be the vertices of P in counterclockwise order. For each edge of P, we consider the absolute *subtended angle* of this edge at v, as shown in Fig. 1.

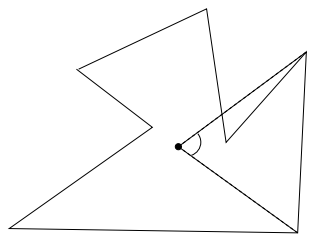

Fig. 1. The subtended angles of an edge at a point.

In any polygon, these subtended angles sum to at least 360° (larger because of taking absolute values, and if the winding number of P is larger than 1). Thus, at least one edge (u_i, u_{i+1}) has subtended angle at least $360°/|P|$ (and at most 180° by definition). In the triangle v, u_i, u_{i+1}, because $\|u_i - v\|$ and $\|u_{i+1} - v\|$ are both at least r, and the angle between vu_i and vu_{i+1} is at least $360°/|P|$, $\|u_i - u_{i+1}\|$ is $\Omega(r/|P|)$. □

Lemma 2. *Any crossing-free topological embedding of K_4 into the plane* (*drawing each edge as a curve*) *places one vertex inside the closed-curve embedding of the cycle induced by the other three vertices.*

Proof. The class of graphs excluding K_4 as a minor is the class of *series–parallel graphs*, i.e., graphs of treewidth at most 2; see Theorem 17 of [Bo]. This graph class includes as a subclass all outerplanar graphs. Thus K_4 is not outerplanar, so any crossing-free topological embedding into the plane must place some vertex not on the outside face. The shortest cycle in K_4 is of length 3, so in fact three of the vertices must form the outside face, and the remaining fourth vertex must be interior to this face. □

Our result about crossing-free straight-line planar embeddings is based on graphs that inspire our main result later.

Theorem 3. *There are planar graphs for which every crossing-free straight-line embedding into the plane has distortion $\Omega(n)$.*

Proof. Consider the graph G obtained by starting with K_4 and attaching a path of length n to each vertex of K_4. Any crossing-free straight-line embedding of G in particular embeds K_4 without crossings, so by Lemma 2, one vertex v of K_4 is inside the cycle C connecting the other three vertices of K_4. Any crossing-free straight-line embedding of G must therefore place the path attached to v completely inside C. Let u denote the endpoint of this path (other than v). The graph distance from u to every vertex of C is at least n, so in any expansive embedding of G, the Euclidean distance between u and every vertex of C is at least n. By Lemma 1 with $P = C$, an edge of C (which is an edge of K_4) has length $\Omega(n/|C|) = \Omega(n)$. Therefore, the distortion of any crossing-free straight-line embedding of G is $\Omega(n)$. □

Our main result is based on the *globe graph*, defined as follows; see Fig. 2. On a sphere (the globe), draw $n^{1/3}$ longitude and $n^{1/3}$ latitude lines. These lines define a preliminary

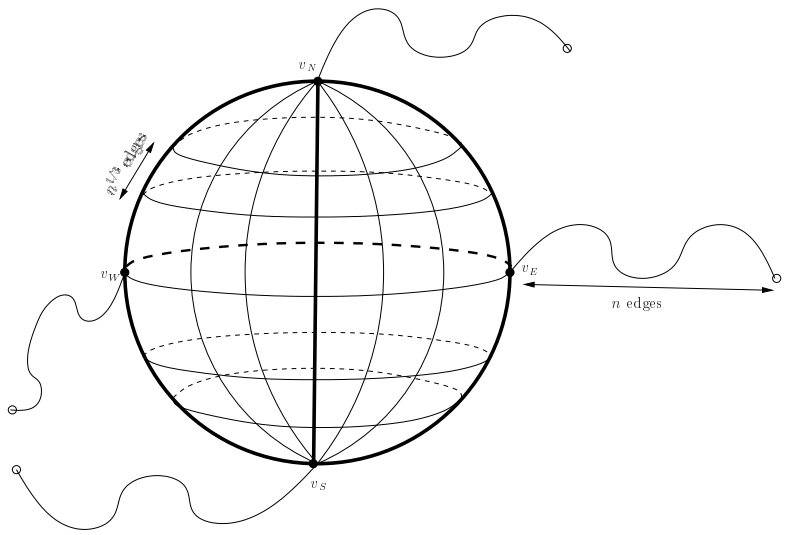

Fig. 2. The globe graph G_n.

graph, whose vertices are the intersections of the lines (including the two poles) and whose edges are portions of the lines. We subdivide each of these edges into a path of $n^{1/3}$ edges. We attach a path of length n to each of the two poles v_{N} and v_{S}, and to each of two antipodal points v_{W} and v_{E} on the equator. The resulting globe graph G_n is planar and has $\Theta(n)$ vertices.

In the spherical embedding of the globe graph shown in Fig. 2, each face is called a *country*.[2] Thus each country is a cycle of length $\Theta(n^{1/3})$.

The four vertices v_{N}, v_{S}, v_{W}, and v_{E}, together with the three longitudes at 0° and $\pm 90^\circ$ and the rear half of the equator, define a subdivided K_4 as a subgraph of G_n. This subgraph is shown in Fig. 2 with heavy edges. The distance between any two vertices in the subdivided K_4 is within a constant factor of the distance between those vertices in the globe graph G_n.

Lemma 4. *If two graph edges cross in a straight-line embedding of a graph into the plane, and the graph distance between each endpoint of the first edge and each endpoint of the second edge (four pairs) is at least s, then the distortion of the embedding is $\Omega(s)$.*

Proof. Consider the quadrilateral formed by the endpoints of the two edges. In any expansive embedding of the graph into the plane, the Euclidean length of each side of the quadrilateral is at least s (the graph distance between the two edge endpoints). By the triangle inequality, at least one diagonal of the quadrilateral (i.e., one of the two edges) also has Euclidean length at least s. Therefore, the distortion of any such embedding is $\Omega(s)$. □

[2] We use the term "country" instead of "face" in order to refer to the same cycles when considering not-necessarily-planar embeddings of the globe graph.

Lemma 5. *Any $o(n^{2/3})$-distortion straight-line embedding of the subdivided K_4 into the plane puts one vertex of K_4 inside the cycle (polygon) connecting the other three vertices of K_4.*

Proof. Call a vertex of K_4 *good* if it is inside the cycle connecting the other three vertices of K_4. If the straight-line embedding of the subdivided K_4 is crossing-free, then by Lemma 2, there must be a good vertex. Otherwise, we show how to modify the embedding into an embedding of K_4, drawing each edge as a curve, in such a way that eliminates all crossings while preserving the parity of the number of good vertices of K_4. Thus, if there is an odd number of good vertices in the new embedding, then there was a good vertex in the original embedding too.

Throughout the modification of the embedding, we maintain a *labeling* of some portions of these curves. Namely, in the original embedding, we label every edge of the subdivided K_4 whose endpoints have shortest-path distance at most $\frac{1}{3}n^{2/3}$ from vertex v_d, $d \in \{\mathrm{N}, \mathrm{S}, \mathrm{E}, \mathrm{W}\}$, with label d. (Every point has at most one label.) As we modify the embedding of K_4, every point of a curve that remains in the embedding keeps the same label that we originally assigned it.

This labeling has several properties with respect to the original straight-line embedding, and as we shall see, with respect to the modified embeddings we create. The first property is simple, while the latter three are implications of Lemma 4, which tells us that, in the original straight-line embedding of the subdivided K_4 with distortion $o(n^{2/3})$, two endpoints of two crossing edges must have graph distance $o(n^{2/3})$. The four properties are as follows. First, for any point x labeled a along the curve representing an edge $\{v_a, v_b\}$ of K_4, the subcurve from v_a to x is entirely labeled a. Second, for any three distinct vertices v_a, v_b, v_c of K_4, the edge $\{v_a, v_b\}$ embeds to a curve that does not cross any portion of a curve labeled c. Third, two nonincident edges of K_4 cannot cross. Fourth, two incident edges of K_4 sharing an endpoint v_a can intersect only along the portions labeled a. Because we only remove portions of curves and/or change curves locally around removed crossings, we maintain all four of these properties as invariants of the modified embedding.

Whenever an edge of K_4, say $\{v_\mathrm{N}, v_\mathrm{S}\}$, is embedded as a self-crossing curve, we remove a self-crossing as follows; see Fig. 3. Consider starting from one endpoint v_N of

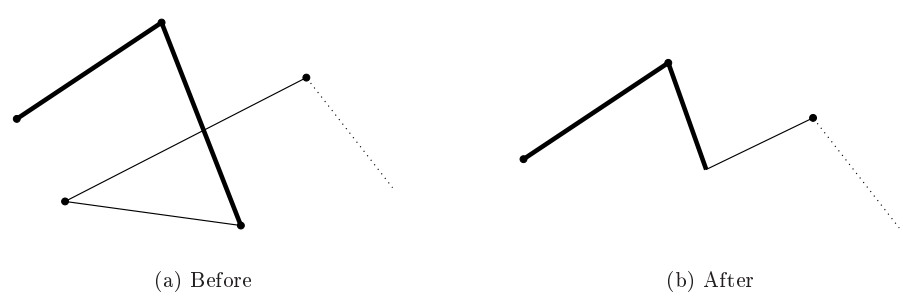

Fig. 3. Eliminating self-intersections. Bold lines indicate a hypothetical labeling of part of the curve.

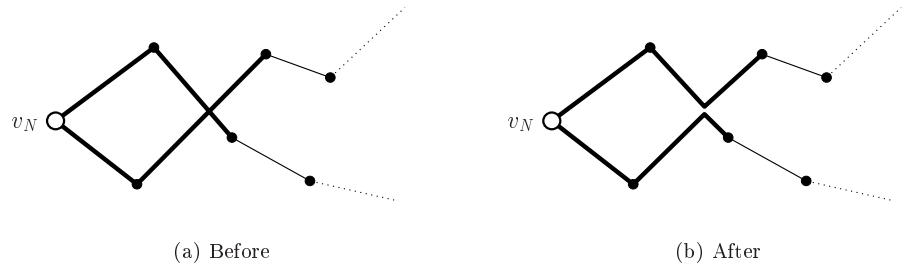

Fig. 4. Eliminating edge intersections. Bold lines indicate portions of the curve labeled v_N.

the curve and following the curve until reaching a point q that has already been visited. Thus q decomposes the curve into three pieces: a curve from v_N to q, a closed curve starting and ending at q, and a curve from q to v_S. We remove the closed curve from the embedding of K_4. This removal affects only the embedding of the edge $\{v_N, v_S\}$ of K_4, which is not involved in the definition of whether v_N or v_S is good, so the goodness of v_N and v_S is preserved. On the other hand, the nonincident edge $\{v_W, v_E\}$ of K_4 cannot cross the edge $\{v_N, v_S\}$ (before or after the removal), so v_W is inside the removed closed curve if and only if v_E is inside the removed closed curve. Thus, the goodness of v_W and v_E either both remain the same or both change as a result of the removal. Hence, the removal preserves the parity of the number of good vertices of K_4.

Whenever the edges of K_4 do not have self-intersections (i.e., the previous modification has been applied fully), yet two (incident) edges, say $\{v_N, v_S\}$ and $\{v_N, v_W\}$, embed to curves that cross each other at a point x, then we remove the crossing as follows; see Fig. 4. The point x divides the edge $\{v_N, v_S\}$ into two curves, one of which connects v_N to x. Similarly, the edge $\{v_N, v_W\}$ gives us a curve from v_N to x. We swap these two curves from v_N to x, switching the identity of to which edge each curve belongs. This change affects the embedding of only these two edges, which are not involved in the definition of whether v_N is good, so the goodness of v_N is preserved. Also, because we did not remove any portions of curves connecting v_N, v_S, and v_W from the embedding—we only changed how they were connected—we preserve the number of crossings with a ray emanating from v_E, and thus we preserve its parity and whether v_E is good. Finally, we claim that the goodness of v_S and v_W either both change or both stay the same. By our invariants, the point x is labeled N on both curves from v_N to x, and therefore the entire curves from v_N to x are labeled N. Therefore, the edge $\{v_S, v_W\}$ must embed to a curve that does not cross either of these curves, so v_S and v_W must both be either inside or outside of the closed curve formed by the two curves from v_N to x. In the former case the goodness of v_S and v_W both change, and in the latter case they both stay the same.

By repeatedly applying each of the two rules above for removing crossings, giving preference to the removal of self-crossings, we eventually remove all crossings in the embedding of K_4. Removing a crossing between two different edges of K_4 can introduce new self-crossings, but every operation decreases the total number of crossings, so the process terminates. As we have argued, the parity of the number of good vertices of K_4

has remained the same, any crossing-free straight-line embedding has a good vertex, and therefore we obtain the desired good vertex in the original embedding. □

By symmetry, assume that vertex v_N is embedded inside the cycle (polygon) C in the subdivided K_4 connecting v_S, v_W, v_E. Let u denote the endpoint of the path attached to v_N (other than v_N itself). By Lemma 4, u must also be inside the cycle connecting v_S, v_W, v_E in any $o(n^{2/3})$-distortion embedding, because every edge of the chain attached to v_N has distance $\Omega(n^{2/3})$ to every edge of the cycle.

Lemma 6. *In any $o(n^{2/3})$-distortion embedding of the globe graph, the endpoint u is embedded inside a country of the globe graph.*

Proof. If we draw the globe graph on a sphere as in Fig. 2, then the cycle C divides the sphere into two regions: inside and outside. We define the *inside region* to be the region containing v_N, which is roughly three-quarters of the sphere; the *outside region* is the other region, the back lower quarter of the sphere. We define an *inside country* to be a country (face) of the globe graph that is within the inside region (viewed on the sphere). An edge of an inside country is *boundary* if it is an edge of the cycle C.

By the definition of u being inside C, a half-infinite ray emanating from u properly crosses C an odd number of times. Let S denote the sum, over all inside countries, of the number of times that the ray properly crosses that country in the embedding. This sum decomposes into two parts: the number of crossed boundary edges of inside countries and the number of crossed nonboundary edges of inside countries. Every crossed nonboundary edge is shared by exactly two inside countries and therefore is counted exactly twice in the sum S. Hence, the parity of S is the same as the parity of the number of crossed nonboundary edges (i.e., crossed edges of C), which we know to be odd. Thus, at least one term in the sum S is odd, i.e., at least one inside country is crossed an odd number of times by the ray. Therefore, u is embedded inside this country. □

Theorem 7. *Every embedding of the globe graph into the plane has distortion $\Omega(n^{2/3})$.*

Proof. If there were an embedding with distortion $o(n^{2/3})$, then by Lemma 6, the endpoint u of one of the attached paths would be embedded inside a country of the globe graph. The distance of u to the vertices of the country is at least n, while the country is a cycle of $\Theta(n^{1/3})$ vertices. By Lemma 1, the distortion of the embedding must be $\Omega(n^{2/3})$, a contradiction. □

The same result as Theorem 7 can be proved with a mathematically more natural graph instead of the cartographically natural globe graph. Namely, take a tetrahedron and repeatedly refine each triangle into four subtriangles until each edge of the tetrahedron is refined into $\Theta(n^{1/3})$ edges. Then we refine each edge further into a path of $n^{1/3}$ edges, and attach a path of length n to each corner of the tetrahedron.

3. Outerplanar Embedding

In this section we give an algorithm for embedding outerplanar graph metrics into the plane with $O(\sqrt{n})$ distortion. Our algorithm consists of five main steps:

1. Modify the graph so that no two faces[3] share an edge, and so that the number of vertices of every face is congruent to 1 modulo 4, at the expense of $O(1)$ distortion.
2. Build the BFS tree rooted at an arbitrary vertex r, conceptually removing the *nontree edges*.
3. Add extra "branches" (paths) in between subtrees of the tree in order to change the shape of the embedding of the BFS tree (to come in Step 4) to have a desired horizontal gap between the two subtrees.
4. Embed the resulting tree using Babilon et al.'s tree embedding algorithm.
5. Shift the upper half of the left side of each face to the right to compensate (if needed) for the nontree edges' improvements to shortest paths.

Step 1. We begin with the transformation of the graph into a "tree of cycles":

Lemma 8. *Given an outerplanar graph G, we can compute an outerplanar graph G' such that $V(G') \supseteq V(G)$, no two faces of G' share an edge, the number of vertices on every face of G' is congruent to* 1 *modulo* 4, *and any embedding of the shortest-path metric of G' into the plane induces an embedding of the shortest-path metric of G with an extra multiplicative factor of $O(1)$ in distortion. Also, G' has $O(n)$ vertices.*

Proof. To satisfy the distortion constraint, it suffices to construct a suitable graph G' such that, for any two vertices u, v of G, the shortest-path distance between u and v in G' is within a factor of $O(1)$ of their distance in G. The dual of G is a forest consisting of one tree for every biconnected component of G. It suffices to consider each biconnected component separately, and then glue the modified components together at the vertices with common labels.

The modification of a biconnected component proceeds recursively. To initialize, we set e to any edge of G incident to only one face, and we color the edge e and its endpoints red. (Throughout this proof, we use an edge to represent its embedded open curve; thus we are in fact coloring all points along this curve.) In general, we proceed as follows. Let f be the unique face incident to e. For each edge $e' \neq e$ of face f, we double the edge e' and duplicate one of its nonred endpoints u, as shown in Fig. 5. More precisely, the edge e' splits the outer face into two pieces, and the edges incident to u get split accordingly into two pieces, one for each copy of u. Because $e' \neq e$, at least one of the endpoints u of e' is different from the red endpoints of e; we split u into two vertices, one with the same label u, and the other with duplicate label u'. We color each of the duplicated edges and their endpoints red. Afterward, we conceptually remove the face f, and we recurse in each biconnected component that has more than one face. Note that, inductively, each biconnected component has exactly one red edge, incident to exactly one face.

[3] Throughout this section we use the term "face" to refer to bounded faces only, excluding the outside face.

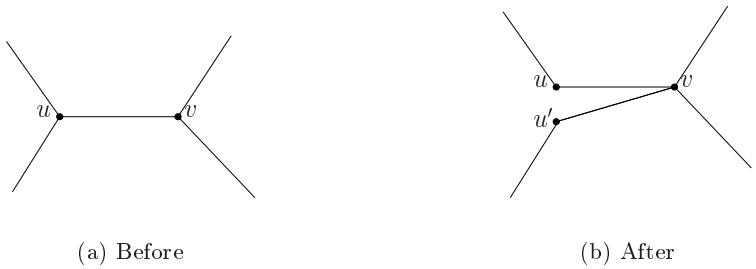

Fig. 5. Eliminating edge neighborings.

Because every duplicated vertex is immediately colored red, and no red vertices are duplicated, the number of vertices in the resulting graph G' is $O(n)$. Also, the shortest-path distance between any vertex u of G and its duplicate u' is exactly 2 in G'. Thus, any shortest path in G' between distinct vertices in G is at most a constant factor longer than the shortest path in G, so we obtain the desired distortion bound.

Finally, we can modify G' to guarantee that every face has length congruent to 1 modulo 4. For each face in G' of length congruent to $j \neq 1$ modulo 4, $2 \leq j \leq 4$, we subdivide one of the edges into a path of $6 - j$ edges. Now every face has length congruent to 1 modulo 4. Each edge is subdivided $O(1)$ times, so the total length of any path, and the size of the whole graph, increases by at most a constant factor. □

From now on, we work on G' instead of G. We fix a combinatorial outerplanar embedding of G', so that cycles correspond to faces.

Step 2. Next we compute a BFS tree of G' (in the primal). For simplicity, we assign the root of the breadth-first search to be a degree-2 vertex of G'. We view this root vertex as being at the bottom of the embedding. Define the *height* of a vertex to be its level in the BFS tree, i.e., its tree distance to the root. We view the y coordinates of the embedding as being proportional to height; together with the combinatorial planar embedding of G', this view defines a notion of "left" and "right" between vertices at the same height. See Fig. 6.

The BFS tree decomposes the edges of G' into *tree edges* and *nontree edges*. Each nontree edge completes a face of the outerplanar graph G' together with only tree edges, because the faces of G' are the biconnected components of G'. Indeed, there is exactly one nontree edge per face f. Because faces of G' have odd length by Lemma 8, the endpoints of the nontree edge of f have equal height, namely, the maximum height among all vertices on the face f. In Fig. 6 we draw the nontree edge as an apex at the top of each face. On the other hand, the *base vertex* of a face is the vertex of minimum height, which is where the breadth-first search first enters the face. The base vertex and the nontree edge divide the tree edges of the face into a *left side* and a *right side*, of equal size, each starting at the base vertex and ending at an endpoint of the nontree edge. We consider the base vertex to belong to both the left and right sides unless specified otherwise.

At the end, the endpoints of each nontree edge will need to be brought closer together. Intuitively, no vertex is pulled in two directions by two different nontree edges: every

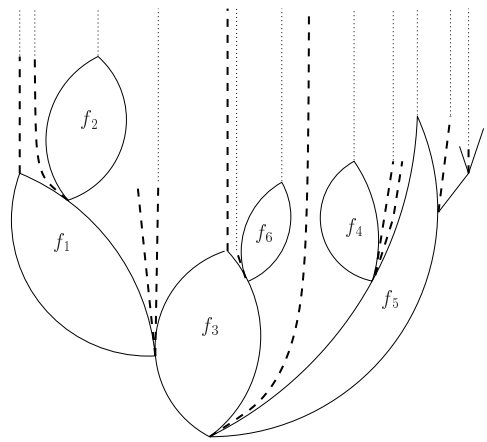

Fig. 6. The graph G'' with extra branches, before trimming (dotted) and after trimming (dashed).

vertex is incident to at most one nontree edge, because a vertex is the base of all but at most one incident face, and the endpoints of the nontree edge of a face have strictly larger height than the base of the face.

Step 3. We now define a further modification G'' of the graph. The modification is based on attaching new child paths, called *extra branches*, of certain lengths to certain nodes of the BFS tree. Extra branches are always added in pairs, and each pair of extra branches is effectively treated as a unit. There are three types of pairs of extra branches that we add.

Extra branches of the first and second types are created as follows. For each vertex u with $k \geq 2$ children in the BFS tree, we attach $k - 1$ pairs of extra branches, one pair between every two consecutive children of u in the BFS tree. We first try attaching the pair of branches at u, and embedding the branches to enter the region between the two children. However, if this would cause the branches to enter a face f of the graph G', then we instead attach the two branches at the two endpoints of the nontree edge of f, one branch at each endpoint, embedded between any children of the two endpoints. Branches attached to u are of the *first type*; branches attached to the nontree edge of f are of the *second type*. In either case the *origin* of each branch is the vertex u (even though the branch may not be attached to u).

Next we define the *height* of the extra branches of the first and second types, i.e., the height of the top vertex of each such branch. Initially, we set all of the heights to be n; then we trim some of the branches in order to guarantee that the number of vertices in G'' is only a constant factor larger than the number of vertices in G'. We also preserve the property that there is at least one extra vertex between any two nonextra vertices at the same height not belonging to the same face. The trimming algorithm proceeds as follows: while there are two extra vertices at the same height and of the same type, with no original vertices or extra vertices of the same type in between, and not coming from the same pair of extra branches, remove the extra vertex that originates from a higher

vertex, breaking ties according to horizontal order with left having the highest priority. Note that extra branches of the second type are given priority according to the height of their origin at the bottom of the face, not the height of the endpoints of the nontree edge.

We add pairs of extra branches of the third type wherever a vertex u is on the left side of a face f (excluding the base vertex) and has at least one child strictly left of f. In this case, we attach a pair of extra branches to u in between the face f and the children of u strictly left of f. The height of these branches is defined to be the minimum of two heights: (1) the maximum height among descendants of the children of u strictly left of f, and (2) the maximum height among vertices strictly right of f that are descendants of vertices on the right half of the face f. One vertex may have attached to it pairs of extra branches of multiple types; in this case, we arbitrarily assign the left-to-right ordering of the pairs in the combinatorial embedding. (In fact, we could also simply remove the shorter pairs of extra branches.)

Figure 6 shows an example with six faces. The extra branches after trimming are drawn as dashed lines; the dotted extensions show the original branches before trimming. For instance, there are two pairs of extra branches originating from the junction between f_1 and f_3. As drawn, the left extra branch is of the third type, and the right extra branch is of the first type (and has been trimmed).

The extra branches preserve many aspects of the graph. We never add a branch inside a face, so the faces are preserved. The extra branches do not change the distances between vertices in G'. Furthermore, the resulting graph G'' is not much larger than G':

Lemma 9. *The graph G'' with extra branches has size $O(n)$.*

Proof. We prove that the number of extra vertices is linear separately for each type. For extra branches of the first or second type separately, if we remove one branch from each pair of extra branches, then by the definition of trimming, each extra vertex is either the leftmost at its height or it has a unique nonextra vertex immediately to its left. Thus we can charge the two vertices from a pair of extra branches either to the height or to the unique nonextra vertex. For extra branches of the third type, if we remove one branch from each pair of extra branches, then we claim that each extra vertex has a unique nonextra vertex immediately to its left. This claim follows from the definition of the height of an extra branch of the third type: any potentially conflicting horizontally adjacent branches of the third type must be above one of the two heights whose minimum we take. Thus we can charge to the nonextra vertex. Because the maximum height and the number of nonextra vertices is $O(n)$, the total number of extra vertices is $O(n)$. □

Step 4. We apply the Babilon et al. tree embedding algorithm [BMMV] to the BFS tree of the graph G'' (i.e., the BFS tree of G' augmented with the extra branches). Intuitively, the algorithm assigns an angular wedge to each subtree proportionally according to the number of vertices in the subtree, and assigns y coordinates according to the height of each node plus a complicated perturbation. In particular, the algorithm preserves the specified combinatorial embedding of the BFS tree of G''. Precisely, the algorithm

assigns to each vertex v the coordinates $(x'(v), y(v))$ as follows:

$$\begin{aligned} y(v) &= 2\sqrt{n}\, h(v) + [\ell(v) \bmod \sqrt{n}], \\ x'(v) &= \frac{k(v)}{\sqrt{n}}. \end{aligned}$$

The functions $h(v)$, $\ell(v)$, and $k(v)$ are defined as follows. First, $h(v)$ denotes the height of v. Let π_v denote the BFS path from the root to v, i.e., the path of ancestors of v. These paths define a partial order on the vertices: $u \prec v$ if $u \notin \pi_v$, $v \notin \pi_u$, and π_u goes to the left of π_v at the vertex at which π_u and π_v branch. Now $\ell(v)$ denotes the number of predecessors of v in this partial order: $\ell(v) = |\{u \in V \mid u \prec v\}|$. Define $\mathrm{sgn}_v(u)$ to be 0 if $u \in \pi_v$ or $v \in \pi_u$, $+1$ if $u \prec v$, and -1 if $v \prec u$. Let $a_v(u)$ denote the tree distance from v to the nearest common ancestor of u and v, i.e., the vertex at which π_u and π_v branch. Finally, $k(v) = \sum_{u \in V} \mathrm{sgn}_v(u) a_v(u)$.

Theorem 10 [BMMV, Theorem 1]. *Babilon et al.'s tree embedding* $(x'(v), y(v))$ *embeds the BFS tree with* $O(\sqrt{n})$ *distortion*: *for a constant* $c_0 > 0$, *the Euclidean distance between any two vertices is between* (a) c_0 *times their tree distance and* (b) $O(\sqrt{n})$ *times their tree distance*.

Before we describe our modifications to this embedding, we prove a few basic facts about it:

Lemma 11. *If p is the parent of node w in the BFS tree, then $k(w) - k(p) = g(w)$, where $g(v) = \sum_{u \in V} \mathrm{sgn}_v(u)$ is the number of vertices to v's left minus the number of vertices to v's right.*

Proof. Let $c(v) = |\{u \mid v \in \pi_u\}|$ be the number of descendants of v (including v itself). By the definition of $k(v)$, there are two ways in which $k(w)$ differs from $k(p)$. First, for every vertex u with $\mathrm{sgn}_p(u) \neq 0$, $a_w(u) = a_p(u) + 1$, for an overall increase in the summation by $g(p)$. Second, the $c(p) - c(w)$ descendants of p that are not descendants of w are newly counted in the summation; each such descendant u has lowest common ancestor p with w, so $a_w(u) = 1$, leading to a contribution of -1 if $u \prec w$ and $+1$ if $w \prec u$. Thus the total such contribution is $g(w) - g(p)$, i.e., the number of descendants of p left of w minus the number of descendants of p right of w. Therefore the overall difference $k(w) - k(p)$ is $g(w)$. □

Lemma 12. *In Babilon et al.'s tree embedding algorithm, the difference in x coordinates between two vertices of the same height is no less than the difference in x between their parents.*

Proof. Suppose two vertices w and z have the same height. Assume by symmetry that w is to the left of z. Let p be w's parent and let q be z's parent. Assume that $p \neq q$; otherwise, $x'(q) - x'(p) = 0$ and the lemma follows trivially. Thus, p is to the left of q. We analyze the difference in x coordinates between w and its parent p, i.e., $x'(w) - x'(p) = (k(w) - k(p))/\sqrt{n}$. It suffices to consider the difference $k(w) - k(p)$,

ignoring the factor of $\sqrt{n}$. By Lemma 11, $k(w) - k(p) = g(w)$ and $k(z) - k(q) = g(z)$. Thus $[k(z)-k(w)]-[k(q)-k(p)] = [k(z)-k(q)]-[k(w)-k(p)] = g(z)-g(w)$, which is nonnegative because $w \prec z$. Therefore $k(z)-k(w) \geq k(q)-k(p)$, so $x'(z)-x'(w) = (k(z) - k(w))/\sqrt{n} \geq (k(q) - k(p))/\sqrt{n} = x'(q) - x'(p)$ as desired. □

Lemma 13. *In Babilon et al.'s tree embedding algorithm, the difference in x coordinates between two vertices of the same height and on the same face of G'' is a concave function of their height.*

Proof. Consider two vertices w and z with the same height on face f, with w on the left branch of f and z on the right branch of f. Let p be w's parent and let q be z's parent. We analyze the change in horizontal distance as we move upward, i.e., from the parents p and q to the children w and z. Again it suffices to consider the difference in k values, ignoring the factor of $\sqrt{n}$. By Lemma 11, and as computed in the proof of Lemma 12, $[k(z) - k(w)] - [k(q) - k(p)] = [k(z) - k(q)] - [k(w) - k(p)] = g(z) - g(w)$. Thus, the horizontal distance increases by $g(z) - g(w)$ when we go up. Now $g(z)$ decreases as we move up, because z follows its leftmost child branch, while $g(w)$ increases as we move up, because w follows its rightmost child branch. Therefore the change in horizontal distance decreases as we go up, i.e., the horizontal distance has a negative second derivative and is thus concave. □

Step 5. Because of nontree edges, shortest paths in the BFS tree may be much larger than in G''. To compensate for the lack of nontree edges in tree embedding algorithm, we "shift" part of each face. Intuitively, the tree embedding algorithm pushes apart the two sides to match the shortest paths in the BFS tree between the vertices of the sides. On the other hand, the existence of the nontree edge between its endpoints decreases the shortest path between the vertices in the upper half of the two sides, so much that the endpoints have distance 1 from each other. Roughly speaking, if the endpoints of the nontree edge of a face are farther than $c_1\sqrt{n}$, for a suitable constant c_1, we shift the upper half of the left side to the right, so that the endpoints have distance $\Theta(\sqrt{n})$ at the end. Figure 7 shows an example.

More precisely, we put a shift value $s(u)$ on any vertex u in the upper half of the left side of a face f. Thus, the distance in G'' from u to the base vertex of f is at least $\lfloor |f|/4 \rfloor$ where $|f|$ is the number of vertices on face f. (The face f is well-defined because a vertex u can be in the upper left part of at most one face.) Let v denote the vertex on the right side of f with the same height as u, and let p and q be the parents of u and v, respectively. Let g and t be the vertices midway along the left and right sides, respectively, of the face f (of distance exactly $\lfloor |f|/4 \rfloor$ from the base vertex of f). If $x'(t) - x'(g) \leq c_1\sqrt{n}$, we define $s(u) = 0$. Otherwise, we define

$$s(u) = [x'(p) - x'(u)] - [x'(q) - x'(v)] + \frac{x'(t) - x'(g) - c_1\sqrt{n}}{\lfloor |f|/4 \rfloor},$$

where $x'(v)$ is the x coordinate assigned to v by Babilon et al.'s algorithm as defined above. Note that, by Lemma 12, $x'(p) - x'(u) + x'(v) - x'(q) \geq 0$, and thus $s(u) \geq 0$. For any vertex u that is not on the upper half of the left side of any face, we define $s(u) = 0$.

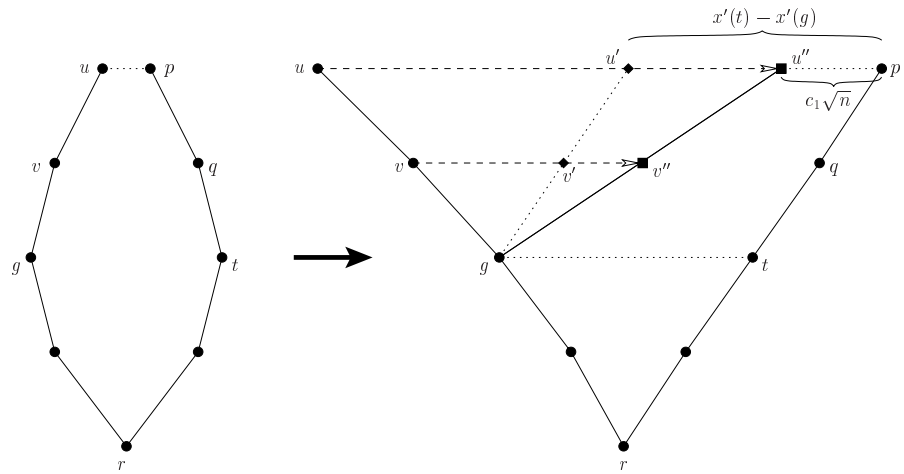

Fig. 7. The shifting process for a nine-vertex face of the outerplanar graph G'', shown on the left, with nontree edge $\{u, p\}$. The tree embedding algorithm places u and v as shown on the right, and the shifting process moves them to u'' and v'', respectively. The idea is first to move the upper-left quarter of the face so that the distances at each level between the left and right vertices of the face is the same. In this example the distance between g and t is equal to the distance between u' and p which is equal to the distance between v' and q. Second, we move these vertices to the positions u'' and v'', respectively, by moving each proportional to its height, so that the distance at the top of the face (between u'' and p) becomes $c_1\sqrt{n}$.

The shift value $s(u)$ applies to u and all of its descendants. Therefore, our formula for the embedding is as follows:

$$y(v) = 2\sqrt{n}\,h(v) + [\ell(v) \bmod \sqrt{n}],$$
$$x(v) = \frac{k(v)}{\sqrt{n}} + \sum_{u \in \pi_v} s(u).$$

In other words, we first make the upper half of the left side of the face roughly parallel to the upper half of the right side of the face, and then slant the left side of the upper half of the face to the right by having a linear growth in shift amount as we proceed up the left side.

Lemma 14. *For every vertex u, $s(u) = O(\sqrt{n})$.*

Proof. By Lemma 11, $|k(u) - k(p)| = O(n)$. Thus, $x'(p) - x'(u) = (k(p) - k(u))/\sqrt{n} = O(\sqrt{n})$, and similarly, $x'(v) - x'(q) = O(\sqrt{n})$. Because the distance in the BFS tree between g and t is less than $|f|$, and, by Theorem 10(b), we have $|x'(g) - x'(t)| = O(|f|\sqrt{n})$. Therefore, $(|x'(g) - x'(t)| - c_1\sqrt{n})/\lfloor |f|/4 \rfloor = O(\sqrt{n})$, so by summing all terms, we obtain that $s(u) = O(\sqrt{n})$. □

Distortion Analysis. We prove that the distortion of our embedding is $O(\sqrt{n})$ in two parts: the least any distance is contracted is $\Omega(1)$, and the most any distance is expanded is $O(\sqrt{n})$.

Lemma 15. *The shifting process preserves the horizontal ordering of vertices with the same height. Furthermore, two vertices of the same height that got closer during the shifting process remain at a horizontal distance of at least* $c_1\sqrt{n}$.

Proof. Suppose that u and v are at the same height and that u is to the left of v in the tree. Let p be the nearest ancestor of u with positive shift value $s(p)$, and let f be the face causing the shift. Let q be the ancestor of v at the same height as p. Assume that $p \neq q$, because otherwise u is not shifted relative to p and any shifting of v only separates u and v, so there is nothing to prove. Thus, p is strictly to the left of q. Let r be the vertex on the right side of face f with the same height as p and q. Thus, r is strictly right of p and (nonstrictly) left of q in the tree.

We claim that, after shifting, p is to the left of r by a horizontal distance of at least $c_1\sqrt{n}$. Let d denote the graph distance between p and r in G'', and suppose that g and t are the vertices midway along the left and right sides, respectively, of the face f. Because this face was shifted, $x'(t) - x'(g) > c_1\sqrt{n}$. Thus,

$$\begin{aligned}
x(r) - x(p) &= x'(r) - x'(p) - \sum_{r \in f \cap \pi_p} s(r) \\
&= x'(r) - x'(p) - [x'(g) - x'(p)] + [x'(t) - x'(r)] \\
&\quad - \sum_{r \in f \cap (\pi_p \setminus \pi_g)} \frac{x'(t) - x'(g) - c_1\sqrt{n}}{\lfloor |f|/4 \rfloor} \\
&= x'(t) - x'(g) - (\lfloor |f|/4 \rfloor - d/2) \frac{x'(t) - x'(g) - c_1\sqrt{n}}{\lfloor |f|/4 \rfloor} \\
&= \left[1 - \frac{\lfloor |f|/4 \rfloor - d/2}{\lfloor |f|/4 \rfloor}\right][x'(t) - x'(g)] + \frac{\lfloor |f|/4 \rfloor - d/2}{\lfloor |f|/4 \rfloor} c_1\sqrt{n} \\
&= \frac{d/2}{\lfloor |f|/4 \rfloor}[x'(t) - x'(g)] + \left(1 - \frac{d/2}{\lfloor |f|/4 \rfloor}\right) c_1\sqrt{n} \\
&\geq \frac{d/2}{\lfloor |f|/4 \rfloor} c_1\sqrt{n} + \left(1 - \frac{d/2}{\lfloor |f|/4 \rfloor}\right) c_1\sqrt{n} \\
&= c_1\sqrt{n}.
\end{aligned}$$

Observe that r either has a strictly smaller height than u, or it is to the right of $p = u$. Thus, we can apply induction to the vertices r and q to conclude that r remains to the left of q after shifting. Hence, p remains to the left of q (in addition to r) by a horizontal distance of at least $c_1\sqrt{n}$. By repeated application of Lemma 12, $x'(v) - x'(u) \geq x'(q) - x'(p)$. Because u is shifted the same amount as p and v is shifted at least as much as q, $x(v) - x(u) \geq x(q) - x(p)$, which is at least $c_1\sqrt{n}$. Therefore, u remains to the left of v by a horizontal distance of at least $c_1\sqrt{n}$. □

Lemma 16. *For sufficiently large* c_1, *the Euclidean distance between two nonextra vertices of the same height is at least a constant factor times their graph distance.*

Proof. Consider any two nonextra vertices u and v of the same height, and assume by symmetry that u is to the left of v. (By Lemma 15, the notion of "left" is the same before and after shifting.) Assume without loss of generality that u got closer to v during the shifting process; otherwise, the lemma follows from Theorem 10(a).

If u and v are on a common face f, then they are in the upper half of a face that was shifted. Let d denote the graph distance between u and v in G'', and suppose that g and t are the vertices midway along the left and right sides, respectively, of the face f. By Theorem 10(a), the Euclidean distance between g and t in the tree embedding is at least $c_0\lfloor|f|/2\rfloor$. Furthermore, their vertical distance $|y(g) - y(t)|$ is at most $\sqrt{n}$, and because this face was shifted, by Lemma 15, their horizontal distance $x'(t) - x'(g)$ is at least $c_1\sqrt{n}$. Therefore, $x'(t) - x'(g) \geq (c_1 - 1)c_0\lfloor|f|/2\rfloor$. Now we can prove that $x(v) - x(u) = \Omega(d)$ and therefore that the Euclidean distance between u and v in our embedding is $\Omega(d)$:

$$\begin{aligned}
x(v) - x(u) &= x'(v) - x'(u) - \sum_{r \in f \cap \pi_u} s(r) \\
&= x'(v) - x'(u) - [x'(g) - x'(u)] + [x'(t) - x'(v)] \\
&\quad - \sum_{r \in f \cap (\pi_u \setminus \pi_g)} \frac{x'(t) - x'(g) - c_1\sqrt{n}}{\lfloor|f|/4\rfloor} \\
&= x'(t) - x'(g) - \sum_{r \in f \cap (\pi_u \setminus \pi_g)} \frac{x'(t) - x'(g) - c_1\sqrt{n}}{\lfloor|f|/4\rfloor} \\
&\geq x'(t) - x'(g) - \sum_{r \in f \cap (\pi_u \setminus \pi_g)} \frac{x'(t) - x'(g)}{\lfloor|f|/4\rfloor} \\
&= \left[1 - \frac{\lfloor|f|/4\rfloor - d/2}{\lfloor|f|/4\rfloor}\right][x'(t) - x'(g)] \\
&= \frac{d/2}{\lfloor|f|/4\rfloor}[x'(t) - x'(g)] \\
&\geq \frac{d/2}{\lfloor|f|/4\rfloor}(c_1 - 1)c_0\lfloor|f|/2\rfloor \\
&= (c_1 - 1)c_0 d,
\end{aligned}$$

which is $\Omega(d)$ for sufficiently large c_1.

Now we prove the claim when u and v are not on a common face f. Let p be the nearest common ancestor of u and v. We distinguish two cases. In Case 1, p is the base vertex of a face f such that the children of p that are ancestors of u and v are precisely the two children on the left and right sides, respectively, of f. In Case 2, there is no such face f.

Case 2 is the easier case. Because the children of p that are ancestors of u and v are not on the two sides of any face, there is a pair of extra branches of the first type attached to p in between u and v. Take the leftmost such pair of extra branches (which survives the longest). The height of these branches is at least $h(u) = h(v)$ because this pair of

branches is the dominant pair (in terms of priority) among ancestors of p in between u and v. Let w be a vertex in this pair of branches at the same height as u and v. By Theorem 10(a), in the tree embedding, the Euclidean distance between w and v is at least c_0 times their tree distance, which equals their graph distance (because the tree is a BFS tree and w is on its own path from the nearest common ancestor p), which equals the graph distance from u to v. Because w is not shifted relative to p, and any shifting of v relative to p brings w and v farther apart, shifting preserves this property that the Euclidean distance between w and v is at least c_0 times the graph distance from u to v. By Lemma 15, u is to the left of w and thus the horizontal distance between u and v is at least the horizontal distance between w and v. Hence the Euclidean distance between u and v is at least the Euclidean distance between w and v minus $2\sqrt{n}$ (to account for possible vertical variations), which is at least c_0 times the graph distance from u to v minus $2\sqrt{n}$. By Lemma 15, because u shifted toward v, their horizontal distance is also at least $c_1\sqrt{n}$. By setting c_1 large enough, we obtain that the Euclidean distance between u and v is at least a constant factor times the graph distance from u to v.

Now we consider Case 1. Let q be the highest vertex (on the left side) of the face f that is an ancestor of u. Because we are in Case 1, $q \neq p$, so there is a pair of extra branches of the third type attached to q. By construction of the height of branches of the third type, the height of these branches is at least $h(u) = h(v)$. Let r be a vertex in this pair of branches that has the same height as u. We distinguish two subcases according to whether the highest vertex of f is strictly lower than u. In each subcase we argue that $|x(r)-x(v)| = \Omega(\mu_{rv}) - O(\sqrt{n})$ where μ_{rv} denotes the graph distance between r and v. The construction of the branches ensures that $\mu_{rv} = \mu_{uv}$. By Lemma 15, r is between u and v, so $|x(u) - x(v)| \geq |x(r) - x(v)|$. Thus we can conclude that $|x(u) - x(v)| = \Omega(\mu_{rv}) - O(\sqrt{n}) = \Omega(\mu_{uv}) - O(\sqrt{n})$. As argued above, $|x(u) - x(v)| \geq c_1\sqrt{n}$. Therefore, for sufficiently large c_1, we obtain that $|x(u) - x(v)| = \Omega(\mu_{uv})$ as desired.

In the first subcase the highest vertex of f is not strictly lower than u (intuitively, the line segment connecting r and v crosses the face f). Let w and z be the vertices on the left and right side, respectively, of f with the same height as u and v. Then we have

$$\begin{aligned} |x(u) - x(v)| &\geq |x(r) - x(w)| + |x(w) - x(z)| + |x(z) - x(v)| \\ &\geq \Omega(\mu_{rw} + \mu_{wz} + \mu_{zv}) - O(\sqrt{n}) \\ &\geq \Omega(\mu_{rv}) - O(\sqrt{n}). \end{aligned}$$

The first inequality follows from the horizontal ordering given by Lemma 15. The second inequality follows from Theorem 10(a); because r does not shift relative to q, and w shifting relative to q could only increase $|x(r) - x(w)|$; by the proof above for the case of two vertices on the same face; and because similarly z and v can shift only apart from each other. The third inequality follows from the triangle inequality.

In the second subcase the highest vertex of f is strictly lower than u. Consider the pair of extra branches of the second type attached to the endpoints of the nontree edge of f. As usual, these branches have height at least $h(u) = h(v)$ because this is a dominant pair of branches (in terms of priority) among ancestors of p in between u and v. (Thus, intuitively, the line segment connecting r and v crosses the extra branches attached to the nontree edge of f.) Let w and z be the vertices of the left and right branches in the

pair, respectively, with the same height as u. Similar to the previous subcase, we have

$$\begin{aligned} |x(u) - x(v)| &\geq |x(r) - x(w)| + |x(z) - x(v)| \\ &\geq \Omega(\mu_{rw} + \mu_{zv}) - O(\sqrt{n}) \\ &\geq \Omega(\mu_{rw} + \mu_{wv}) - O(\sqrt{n}) \\ &\geq \Omega(\mu_{rv}) - O(\sqrt{n}). \end{aligned}$$

The first, second, and fourth inequalities follow for the same reasons as the previous subcase. The third inequality follows because, by the construction of extra branches, $|\mu_{zv} - \mu_{wv}| \leq 1$. □

Lemma 17. *For sufficiently large constants c_1 and c_2, two vertices at different heights with graph distance at least $c_2\sqrt{n}$ have Euclidean distance at least a constant factor times their graph distance.*

Proof. Suppose that vertices u and v have graph distance at least $c_2\sqrt{n}$. If the height difference between u and v is $\Omega(\sqrt{n})$, then the difference in y coordinates is $\Omega(n)$, which is enough for any graph distance. Therefore, we can assume that $|h(u) - h(v)| \leq \varepsilon\sqrt{n}$ for any constant $\varepsilon > 0$. Assume by symmetry that $h(u) < h(v)$. Let w be the ancestor of v at the same height at u.

We claim that the descendants of w (in particular, v) embed to lie within a cone with apex at v and whose sides are within θ of vertical for a constant $\theta < 90°$. This claim is true in the original Babilon et al. embedding: the overall cone for the root node has a vertical axis and angle of at most 45° from this axis, and all subtree cones are contained in the original [BMMV, proof of Claim 1]. Now we bound the effect of the shifting process. Consider any two adjacent vertices in the graph with heights differing by 1. Their vertical distance is $\Theta(\sqrt{n})$, so by the 45° cone property of the original tree embedding, their horizontal distance in the original tree embedding is $O(\sqrt{n})$. Also, by Lemma 14, every shift value is $O(\sqrt{n})$. Thus, their horizontal distance in our embedding is $O(\sqrt{n})$, at most a constant factor times the vertical distance. Because this holds for all edges of the graph that span adjacent heights, the entire subtree of w's descendants must lie in a cone with a constant lower bound on the absolute slope of the two sides.

Because w is an ancestor of v in the BFS tree, the length μ_{vw} of the shortest path from v to w is $h(v) - h(w)$, which we assumed to be at most $\varepsilon\sqrt{n}$. By the triangle inequality, $\mu_{uv} \leq \mu_{uw} + \mu_{vw}$. The left-hand side is large—$\mu_{uv} \geq c_2\sqrt{n}$—yet the second term on the right-hand side is small—$\mu_{vw} \leq \varepsilon\sqrt{n}$. Thus, $\mu_{uw} \geq (1 - \varepsilon)\mu_{uv} \geq (1 - \varepsilon)c_2\sqrt{n}$.

By Lemma 16, the Euclidean distance between u and w is $\Omega(\mu_{uw})$, and we have just shown that $\mu_{uw} \geq (1 - \varepsilon)c_2\sqrt{n}$. By construction, the difference in y coordinates between u and w is at most $\sqrt{n}$. Thus, by choosing c_2 large enough, we obtain that the horizontal distance between u and w is at least any desired constant times $\sqrt{n}$, while the vertical distance is at most $\sqrt{n}$. Hence, for c_2 sufficiently large, u is outside w's cone, and furthermore, the absolute slope of the line segment from u to w is strictly less than the absolute slope of the sides of the cone. Therefore, the Euclidean distance from u to any point in the cone is at least a constant factor times the Euclidean length of the line

segment from u to w, which we showed is $\Omega(\mu_{uw}) = \Omega(\mu_{uv})$. In particular, v is in w's cone, so the Euclidean distance from u to v is $\Omega(\mu_{uv})$. □

Lemma 18. *The expansion in our embedding between any pair of nonextra vertices is* $O(\sqrt{n})$.

Proof. First we claim that every edge of the graph G' has Euclidean length $O(\sqrt{n})$. Each graph edge is either a tree edge or a nontree edge. The tree edges had $O(\sqrt{n})$ distortion in the original tree embedding by Theorem 10(b), and, by Lemma 14, the shifting process changes the Euclidean length of each edge by at most an additive $O(\sqrt{n})$. Thus, the tree edges still have Euclidean length $O(\sqrt{n})$.

Now we prove the claim for nontree edges. Let w and z be the left and right endpoints, respectively, of the nontree edge of face f. If $x'(z) - x'(w) > 2c_1\sqrt{n}$, then by the concavity in Lemma 13, the points g and t midway along the left and right sides, respectively, of the face f satisfy $x'(t) - x'(g) > c_1\sqrt{n}$. Thus, in this case, the face f is shifted. After shifting, we have

$$\begin{aligned}
x(z) - x(w) &= x'(z) - x'(w) - \sum_{r \in f \cap \pi_w} s(r) \\
&= x'(z) - x'(w) - [x'(g) - x'(w)] + [x'(t) - x'(z)] \\
&\quad - \sum_{r \in f \cap (\pi_w \setminus \pi_g)} \frac{x'(t) - x'(g) - c_1\sqrt{n}}{\lfloor |f|/4 \rfloor} \\
&= x'(t) - x'(g) - (\lfloor |f|/4 \rfloor) \frac{x'(t) - x'(g) - c_1\sqrt{n}}{\lfloor |f|/4 \rfloor} \\
&= c_1\sqrt{n}.
\end{aligned}$$

Hence, if the nontree edge $\{w, z\}$ is shifted, which is forced when $x'(z) - x'(w) > 2c_1\sqrt{n}$, then $x(z) - x(w) = c_1\sqrt{n}$. Otherwise, we have $x(z) - x(w) = x'(z) - x'(w) \le 2c_1\sqrt{n}$. Also, $|y(w) - y(z)| \le \sqrt{n}$. Therefore, every edge of G' has length $O(\sqrt{n})$.

By the triangle inequality, the distortion of the embedding of G' is $O(\sqrt{n})$, because the bound holds for each edge of a shortest path. By Lemma 8, the distortion of the embedding of the original graph G is $O(\sqrt{n})$. □

We conclude with the main result of this section:

Theorem 19. *Every outerplanar graph G can be embedded into the plane with distortion* $O(\sqrt{n})$.

Proof. We claim that the Euclidean distance between any two vertices u and v in the constructed embedding is at least a constant factor times their graph distance. If u and v have the same height, Lemma 16 proves the claim. Otherwise, their graph distance is either at least $c_2\sqrt{n}$ or smaller. In the first case, Lemma 17 proves the claim. In the second case, the Euclidean distance between u and v is at least their distance in y coordinates, which by construction is at least $\sqrt{n}$, proving the claim.

On the other hand, Lemma 18 proves that the Euclidean distance between any two vertices u and v in the constructed embedding is at most $O(\sqrt{n})$ times their graph distance. Therefore, scaling the embedding by a constant multiplicative factor yields a noncontractive embedding with expansion $O(\sqrt{n})$ as desired. □

4. Open Problems

The main open problem is to obtain tight worst-case bounds on the distortion required to embed planar graph metrics into the plane. To this end, it would be a natural step to consider bounded-outerplanarity graphs (which generalize the notion of outerplanar graphs), series–parallel graphs (which are slightly more general than outerplanar graphs), and bounded-treewidth graphs (which generalize both of these classes). Note that the series–parallel graph known as the diamond graph, used previously in embedding lower bounds [BC], [LN], [NR], can be embedded into the plane with $O(\sqrt{n}\log n)$ distortion.

We also emphasize the open problem from [BMMV] and [M1] of embedding the shortest-path metrics of *weighted* trees and planar graphs into the plane. Matoušek [M1] conjectures an upper bound of $O(\sqrt{n})$ for trees, but no $o(n)$ upper bound is known. It may also be easier to prove an $\omega(n^{2/3})$ worst-case lower bound for weighted planar graph metrics.

Acknowledgments

We thank Anastasios Sidiropoulos for helpful discussions. We also thank the anonymous referees for their helpful comments on the paper.

References

[Bă] Mihai Bădoiu. Approximation algorithm for embedding metrics into a two-dimensional space. In *Proceedings of the* 14*th Annual ACM–SIAM Symposium on Discrete Algorithms*, pages 434–443, 2003.

[BC] Bo Brinkman and Moses Charikar. On the impossibility of dimension reduction in ℓ_1. In *Proceedings of the* 44*th Symposium on Foundations of Computer Science*, pages 514–523, October 2003.

[BCIS] Mihai Bădoiu, Julia Chuzhoy, Piotr Indyk, and Anastasios Sidiropoulos. Low-distortion embeddings of general metrics into the line. In *Proceedings of the* 37*th Annual ACM Symposium on Theory of Computing*, pages 225–233, 2005.

[BDG+] Mihai Bădoiu, Kedar Dhamdhere, Anupam Gupta, Yuri Rabinovich, Harald Räcke, R. Ravi, and Anastasios Sidiropoulos. Approximation algorithms for low-distortion embeddings into low-dimensional spaces. In *Proceedings of the* 16*th Annual ACM–SIAM Symposium on Discrete Algorithms*, pages 119–128, 2005.

[BDHI] Mihai Bădoiu, Erik D. Demaine, MohammadTaghi Hajiaghayi, and Piotr Indyk. Low-dimensional embedding with extra information. *Discrete & Computational Geometry*, 36(4):609–632.

[BKL] Bonnie Berger, Jon Kleinberg, and Tom Leighton. Reconstructing a three-dimensional model with arbitrary errors. *Journal of the ACM*, 46(2):212–235, 1999.

[BMMV] Robert Babilon, Jirí Matoušek, Jana Maxová, and Pavel Valtr. Low-distortion embeddings of trees. *Journal of Graph Algorithms and Applications*, 7(4):399–409, 2003.

[Bo] Hans L. Bodlaender. A partial k-arboretum of graphs with bounded treewidth. *Theoretical Computer Science*, 209(1–2):1–45, 1998.

[CH] G. M. Crippen and T. F. Havel. *Distance Geometry and Molecular Conformation*. Wiley, New York, 1988.

[ČHH] Srdan Čapkun, Maher Hamdi, and Jean-Pierre Hubaux. GPS-free positioning in mobile ad-hoc networks. In *Proceedings of the* 34*th Hawaii International Conference on System Sciences*, pages 3481–3490, January 2001.

[G] Anupam Gupta. Embedding tree metrics into low-dimensional euclidean spaces. *Discrete & Computational Geometry*, 24(1):105–116, 2000.

[GNRS] Anupam Gupta, Ilan Newman, Yuri Rabinovich, and Alistair Sinclair. Cuts, trees and l_1-embeddings of graphs. *Combinatorica*, 24(2):233–269, 2004.

[H] B. Hendrickson. The molecule problem: exploiting structure in global optimization. *SIAM Journal on Optimization*, 5:835–857, 1995.

[HIL] Johan Håstad, Lars Ivansson, and Jens Lagergren. Fitting points on the real line and its application to RH mapping. *Journal of Algorithms*, 49(1):42–62, October 2003.

[IM] Piotr Indyk and Jiři Matoušek. Low-distortion embeddings of finite metric spaces. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, chapter 8, pages 177–196. CRC Press, Boca Raton, FL, second edition, 2004.

[K1] Joseph B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964.

[K2] Joseph B. Kruskal. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29:115–129, 1964.

[KLMN] Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured descent: a new embedding method for finite metrics. *Geometric and Functional Analysis*, 15(4):839–858, 2005.

[LN] James R. Lee and Assaf Naor. Embedding the diamond graph in L_p and dimension reduction in L_1. *Geometric and Functional Analysis*, 14(4):745–747, 2004.

[M1] Jiří Matoušek. Bi-Lipschitz embeddings into low-dimensional Euclidean spaces. *Commentationes Mathematicae Universitatis Carolinae*, 31(3):589–600, 1990.

[M2] Jiří Matoušek. *Lectures on Discrete Geometry*, chapter 15. Springer, New York, 2002.

[M3] Jiří Matoušek. Problem 2.6: Planar-graph metrics into $\mathbb{R}^2$. In *Open Problems from the Workshop on Discrete Metric Spaces and their Algorithmic Applications*. Haifa, Israel, March 2002. `http://kam.mff.cuni.cz/~matousek/haifaop.ps`.

[NR] Ilan Newman and Yuri Rabinovich. A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete & Computational Geometry*, 29(1):77–81, 2003.

[PCB] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The Cricket location-support system. In *Proceedings of* 6*th Annual International Conference on Mobile Computing and Networking*, pages 32–43, Boston, MA, August 2000.

[R] Satish Rao. Small distortion and volume preserving embeddings for planar and euclidean metrics. *Proceedings of the* 15*th Annual ACM Symposium on Computational Geometry*, pages 300–306, 1999.

[S1] R. N. Shepard. The analysis of proximities: multidimensional scaling with an unknown distance function 1. *Psychometrika*, 27:125–140, 1962.

[S2] R. N. Shepard. The analysis of proximities: multidimensional scaling with an unknown distance function 2. *Psychometrika*, 27:216–246, 1962.

[SRB] C. Savarese, J. Rabaey, and J. Beutel. Locationing in distributed ad-hoc wireless sensor networks. In *Proceedings of the International Conference on Acoustics*, *Speech*, *and Signal Processing*, pages 2037–2040, Salt Lake City, UT, May 2001.

[T] W. T. Tutte. How to draw a graph. *Proceedings of the London Mathematical Society* (3), 13:743–767, 1963.

[TBET] Ioannis G. Tollis, Giuseppe Di Battista, Peter Eades, and Roberto Tamassia. *Graph Drawing*: *Algorithms for the Visualization of Graphs*. Prentice-Hall, Englewood Cliffs, NJ, 1998.

[W] Working Group on Algorithms for Multidimensional Scaling. Algorithms for multidimensional scaling. `http://dimacs.rutgers.edu/SpecialYears/2001_Data/Algorithms/MDSdescription.html`.

Received August 11, 2007, *and in revised form March* 27, 2007. *Online publication August* 30, 2007.