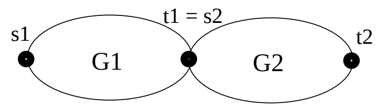
15-859(J): Metric Embeddings and Methods **Lecturer:** Anupam Gupta
Topic: Embeddability results on series-parallel graphs **Date:** November 18, 2003
Scribe: Ryan Williams

23.1 Series-Parallel Graphs

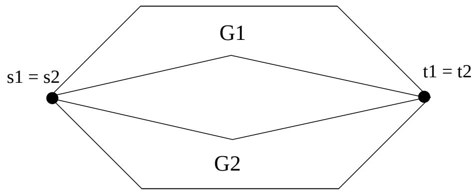
In the previous lecture, we saw results concerning the embeddability of outer-planar graphs into ℓ_1 and distributions of trees. In this lecture we will focus on embeddings for the class of series-parallel graphs. Here are three different ways to define and view these graphs.

Inductive definition: It will be convenient for us to distinguish two nodes in each series-parallel graph, generally called the *terminals* s and t. The class SP of series-parallel graphs are defined as:

1. The graph with the single edge (s,t) is in SP, with distinguished nodes s and t.
2. If $G_1 = (V_1, E_1), G_2 = (V_2, E_2) \in \mathsf{SP}$, with distinguished nodes s_1, t_1 and s_2, t_2 respectively, then
   (a) $G_1 \oplus G_2 \in \mathsf{SP}$ (with distinguished nodes $s = s_1, t = t_2$). The graph $G_1 \oplus G_2$ is defined by taking $V = V_1 \cup V_2$, $E = E_1 \cup E_2$, then merging s_2 and t_2 into a supernode, as the following figure suggests.



   (b) $G_1 \mid\mid G_2 \in \mathsf{SP}$ (with distinguished nodes $s = s_1 = s_2, t = t_1 = t_2$). The graph $G_1 \mid\mid G_2$ is defined by taking $V = V_1 \cup V_2$, $E = E_1 \cup E_2$, then merging s_1 and s_2, and merging t_1 and t_2.



Minor-exclusion definition: Recall planar graphs may be characterized as those which exclude K_5 and $K_{3,3}$ as a minor, and outerplanar graphs exclude $K_{2,3}$ and K_4. Similarly, the class of series-parallel graphs SP are those which exclude K_4 as a minor. (We will not prove this in class.)

Another inductive definition: This specification actually builds up a class containing the 2-connected series-parallel graphs ($\mathsf{2\text{-}conn\text{-}SP}$), but not necessarily the entire class SP. Recall that this is sufficient for our purposes, since an embedding of all 2-connected series-parallel graphs into ℓ_1 can be easily extended (via gate-sums) to embed all series-parallel graphs into ℓ_1.

To generate graphs in the class $\mathsf{2\text{-}conn\text{-}SP}$ inductively, we permit the cycle to be in $\mathsf{2\text{-}conn\text{-}SP}$; this gives us the base case. Furthermore, for any graph G in $\mathsf{2\text{-}conn\text{-}SP}$ with an edge $e = \{u, v\}$, the graph G' obtained by attaching a new path from u to v in G is also in $\mathsf{2\text{-}conn\text{-}SP}$. It is not difficult to see that all 2-connected series-parallel graphs may be built up in this way.

23.2 Results for the next two lectures

We showed in previous lectures that outer-planar graphs could be isometrically embedded into ℓ_1, and could be embedded into a distribution of dominating trees with distortion 8. In contrast to these, we will present an $\Omega(\log n)$ distortion lower bound for SP into distributions of trees, and a distortion-16 embedding for SP into ℓ_1. Let $\mathsf{distrib(Trees)}$ be the class of possible distributions on dominating trees.

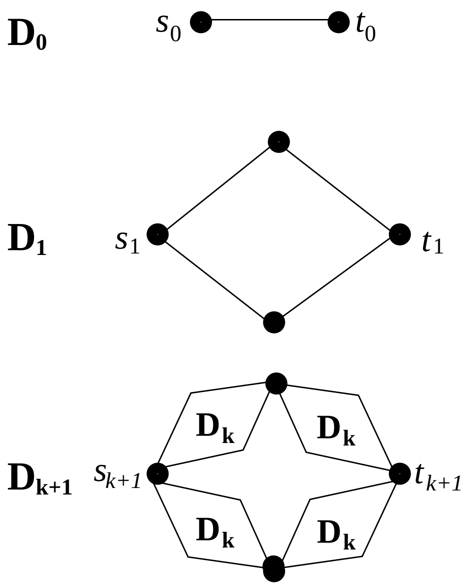
Theorem 23.2.1 $\mathsf{SP} \xhookrightarrow{\Omega(\log n)} \mathsf{distrib(Trees)}$.

Theorem 23.2.2 $\mathsf{SP} \xhookrightarrow{16} \ell_1$.

In this lecture, we will prove Theorem 23.2.1, and will outline the strategy for Theorem 23.2.2.

23.3 Lower bound for distributions of trees

To prove the lower bound of Theorem 23.2.1, we need to present an infinite collection of series-parallel graphs such that any embedding of these graphs into a tree distribution has high distortion. We will use the *diamond graphs*, which are defined inductively according to these pictures.



Two copies of $\mathbf{D_k}$ are connected to each other in series, and then two such series combinations are connected in parallel to give us $\mathbf{D_{k+1}}$. Let us note a few easy-to-prove observations concerning these $\mathbf{D_i}$ and their s_k, t_k pairs.

Observation 23.3.1 *For non-negative k:*

- $|E(\mathbf{D_k})| = 4^k$.
- $|V(\mathbf{D_k})| = \Theta(4^k)$.
- $d(s_k, t_k) = 2^k$ *(in $\mathbf{D_k}$).*
- *The min-cut value between s_k and t_k in $\mathbf{D_k}$ is 2^k. Hence there are 2^k edge-disjoint paths from s_k to t_k in $\mathbf{D_k}$; pairing these up allows us to obtain 2^{k-1}* simple *cycles of length 2^{k+1}.*

Theorem 23.2.1 follows immediately from the next claim about embeddings of diamond graphs.

Claim 23.3.2 *Every distribution of dominating trees incurs distortion $\Omega(k)$ on $\mathbf{D_k}$.*

Proof of 23.3.2: We want to show that for all distributions $\mathcal{D}$ over dominating trees on n nodes, there are nodes u, v such that

$$E_{T \in \mathcal{D}}[d_T(u, v)] \geq \Omega(\log n) \cdot d_{\mathbf{D_k}}(u, v).$$

It will suffice to prove that for any tree T with $d_T(u, v) \geq d_{\mathbf{D_k}}(u, v)$ for all $u, v \in V(\mathbf{D_k})$, the sum of all edge weights in T is large. More precisely, we want to show that

$$\sum_{\{u,v\} \in E(\mathbf{D_k})} d_T(u, v) \geq \Omega(k) \sum_{\{u,v\} \in E(\mathbf{D_k})} d_{\mathbf{D_k}}(u, v) = \Omega(k 4^k). \tag{23.3.1}$$

Using the observations above, this will imply for all distributions $\mathcal{D}$ over dominating trees on 4^k nodes,

$$\sum_{\{u,v\} \in E(\mathbf{D_k})} \left(\frac{1}{|E(\mathbf{D_k})|} \right) d_T(u, v) \geq \Omega(k).$$

That is, if we uniformly chose edges $\{u, v\}$ of $\mathbf{D_k}$ at random, the expected distortion of $\{u, v\}$ by any dominating tree T is at least $\Omega(k)$. We now devote our attention to proving inequality 23.3.1. We will draw upon the known lower bound for embedding the n-node unit-weight cycle (C_n) into an arbitrary tree.

Fact 23.3.3 (Rabinovich and Raz (1998)) *For every tree T, every embedding from C_n into T has distortion at least $n/3 - 1$.*

Let T be a dominating tree for $\mathbf{D_k}$. Using T, we will color edges of $\mathbf{D_k}$ with colors from a palette of size k, according to the rule: color edge $\{u, v\} \in E(\mathbf{D_k})$ with the ith color, if $d_T(u, v) \geq 2^{i+1}/3 - 1$. Observe that some edges have many colors, and every edge has at least one color. Define

$$E_i := \{e \in E(\mathbf{D_k}) \ : \ e \text{ has color } i\}.$$

What is the cardinality of these E_i sets? We claim they are quite large.

Claim 23.3.4 *$|E_i| \geq 2^{2k-1-i}$.*

Assuming this holds, we can prove inequality 23.3.1. By definition of the coloring scheme and $\sum_{i=0}^{j} 2^i = 2^j - 1$, for all $\{u, v\} \in E(\mathbf{D_k})$ we have

$$d_T(u, v) \geq \max_{j:\{u,v\} \in E_j} (2^{j+1}/3 - 1)$$

and hence, using that $\sum_{j' \leq j} 2^{j'+1} \leq 2 \cdot 2^{j+1}$,

$$(d_T(u, v) + 1) \geq \frac{1}{2} \sum_{j:\{u,v\} \in E_j} 2^{j+1}/3$$

Thus

$$\begin{aligned} \sum_{\{u,v\} \in E(\mathbf{D_k})} [d_T(u, v) + 1] &\geq \frac{1}{2} \sum_{\{u,v\} \in E(\mathbf{D_k})} \ \sum_{j:\{u,v\} \in E_j} 2^{j+1}/3 = \frac{1}{3} \sum_{j=1}^{k} |E_j| \cdot 2^j \\ &\geq \frac{1}{3} \sum_{j=1}^{k} 2^{2k-1-j} \cdot 2^j = \frac{1}{3} \sum_{j=1}^{k} 2^{2k-1} = \frac{1}{6} k 4^k, \end{aligned}$$

and hence

$$\sum_{\{u,v\} \in E(\mathbf{D_k})} d_T(u, v) \geq (k/6) 4^k - |E(\mathbf{D_k})| = (k/6 - 1)\, 4^k = \Omega(k 4^k),$$

proving Theorem 23.3.2, and hence Theorem 23.2.1. ■

Proof of Claim 23.3.4: We want to show that $|E_j| \geq 2^{2k-1-j}$. Let us start with the claim that $|E_k| \geq 2^{k-1}$. To see this, first note that any simple cycle C of $\mathbf{D_k}$ that goes through both s_k and t_k has length 2^{k+1}. The above fact implies that C is embedded into this T with distortion at least $2^{k+1}/3 - 1$, therefore at least one edge e in C has color k, so $e \in E_k$. Recall the minimum s_k-t_k cut in $\mathbf{D_k}$ is 2^k; this implies there are 2^{k-1} *edge-disjoint* cycles C containing both s_k and t_k. Each such cycle C has an edge e with color k, therefore $|E_k| \geq 2^{k-1}$.

To prove $|E_{k-1}| \geq 4 \cdot 2^{k-2}$, think of the four copies of $\mathbf{D_{k-1}}$ that make up $\mathbf{D_k}$. Using essentially identical reasoning to the above case, each $\mathbf{D_{k-1}}$ subgraph has at least 2^{k-2} edges with color $k-1$. The edges of each copy are disjoint, so $|E_{k-1}| \geq 4 \cdot 2^{k-2}$.

For general $j = 1, \ldots, k$, there are 4^{k-j} disjoint copies of $\mathbf{D_j}$ in $\mathbf{D_k}$, each having at least 2^{j-1} edges of color j. Hence $|E_j| \geq 4^{k-j} \cdot 2^{j-1} = 2^{2k-2j+j-1} = 2^{2k-j-1}$, and the claim holds. ■

23.4 Embedding series-parallel graphs into ℓ_1: a starting point

We now begin our foray into Theorem 23.2.2, *i.e.* the existence of an embedding from SP into ℓ_1 with distortion at most 16. We will not finish the description of this embedding entirely in this lecture, but shall prove some intermediate results leading up to the result.

First, let's consider how the diamond graphs (the "hard" series-parallel graphs from the previous proofs) might embed into ℓ_1 without much distortion. To this end, recall the $\mathbf{D_k}$ have a very nice structure, in that every simple s_k-t_k path has the same length. Let's make a definition for this kind of graph. It will turn out that these graphs can be embedded fairly well into ℓ_1.

Definition 23.4.1 *A* bundle *is a graph in* SP *with distinguished nodes s and t such that all simple s-t paths have equal length. Let* Bundles $\subseteq$ SP *be the subclass of bundles.*

Theorem 23.4.2 Bundles $\xhookrightarrow{2}$ ℓ_1.

Proof: Let B be a bundle with distinguished s, t nodes. Our approach will be to choose a random s-t cut (S, T) of the vertices of B, and define $\phi(u, v) := d_B(s, t) \cdot \delta_{(S,T)}(u, v)$ where $\delta_{(S,T)}(u, v)$ is the indicator function for (S, T) (*i.e.* equal to 1 if u and v are separated by (S, T), and 0 otherwise). The distribution of random cuts should have the property that

$$E[\phi(u, v)] = E[\delta_{(S,T)}(u, v)] \cdot d_G(s, t) = Pr[u \text{ and } v \text{ are separated by } (S, T)] \cdot d_B(s, t) \approx d_B(u, v).$$

Such a μ will also be an embedding into ℓ_1, since non-negative linear combinations of cut metrics are ℓ_1-metrics. Note that since the cut has $s \in S$ and $t \in T$ by definition,

$$E[\phi(s, t)] = Pr[s \text{ and } t \text{ are separated in } (S, T)] \cdot d_B(s, t) = d_B(s, t).$$

Picking the random cut: We will pick the random cut (S, T) of B in terms of the inductive definition of SP given earlier.

1. If B is the single edge $\{s, t\}$, then $S := \{s\}$ and $T := \{t\}$.
2. If $B = B_1 \mid\mid B_2$ for bundles B_1 and B_2, take the (inductively chosen) random cuts (S_i, T_i) for B_i (with $i = 1, 2$) and set $S := S_1 \cup S_2$, $T := T_1 \cup T_2$.
3. If $B = B_1 \oplus B_2$ for B_1 and B_2 (with distinguished nodes s_1, t_1 and s_2, t_2, respectively), then set $S := S_1$ with probability $\frac{d_{B_1}(s_1,t_1)}{d_{B_1}(s_1,t_1)+d_{B_2}(s_2,t_2)}$, and $S := V(B_1) \cup S_2$ otherwise. (T is simply $[V(B_1) \cup V(B_2)] \setminus S$.)

   That is, we either choose a cut where T contains all of B_2 plus a little of B_1, or we choose a cut where S contains all of B_1 and a little of B_2. The probability of choosing one over the other is biased by the path lengths for the two bundles.

The distortion bound of 2 for ϕ follows immediately from the next claim, and in particular from (23.4.4).

Claim 23.4.3 *For all $u, v \in V(B)$,*

$$Pr[s, v \text{ separated in } (S,T)] = \frac{d_B(s,v)}{d_B(s,t)}; \tag{23.4.2}$$

$$Pr[v, t \text{ separated in } (S,T)] = \frac{d_B(v,t)}{d_B(s,t)}; \tag{23.4.3}$$

$$\frac{d_B(u,v)}{2d_B(s,t)} \leq Pr[u, v \text{ separated in } (S,T)] \leq \frac{d_B(u,v)}{d_B(s,t)}. \tag{23.4.4}$$

The proof of the claim is by "decompositional" induction on B (it may also be thought of as induction on the number of edges of B). More precisely, we argue under the assumption that B_1 and B_2 obey the claim that B obeys the claim (where $B = B_1 \oplus B_2$ or $B = B_1 \mid\mid B_2$).

Base Case: B is the single $\{s,t\}$ edge; here, $|E| = 1$, and $1/2 \leq d_B(s,t) = 1 \leq 1$.

Inductive step: assume the claim holds for B_1 and B_2 where either $B = B_1 \oplus B_2$ or $B = B_1 \mid\mid B_2$.

1. $B = B_1 \oplus B_2$.
   If both $u, v \in V(B_1)$, then by the definition of the process above,

   $$Pr[u, v \text{ separated in } (S,T)] = Pr[u, v \text{ separated in } (S_1, T_1)] \cdot \frac{d_{B_1}(s_1,t_1)}{d_{B_1}(s_1,t_1) + d_{B_2}(s_2,t_2)}$$

   But plugging in the probability for u and v to be separated in B_1 from the inductive hypothesis proves the claim. E.g., to prove (23.4.4), note that

   $$\frac{d_{B_1}(u,v)}{2d_{B_1}(s,t)} \leq Pr[u, v \text{ separated in } (S_1,T_1)] \leq \frac{d_{B_1}(u,v)}{d_{B_1}(s,t)}$$

   by induction. Plugging this into the above expression gives us (23.4.4); similar calculations give us (23.4.2) and (23.4.3). Finally, an identical argument proves the claim for $u, v \in V(B_2)$.
   If $u \in V(B_1)$ and $v \in V(B_2)$, observe that

   $$\begin{aligned} Pr[u, v \text{ separated in } (S,T)] &= \frac{d_{B_1}(s_1,t_1)}{d_{B_1}(s_1,t_1) + d_{B_2}(s_2,t_2)} \cdot Pr[u, t_1 \text{ separated in } (S_1,T_1)] \\ &+ \frac{d_{B_2}(s_2,t_2)}{d_{B_1}(s_1,t_1) + d_{B_2}(s_2,t_2)} \cdot Pr[s_2, v \text{ separated in } (S_2,T_2)]. \end{aligned}$$

   However, we can now use (23.4.2) and (23.4.3) inductively to compute the probabilities on the right hand side, and this gives us

   $$\begin{aligned} &Pr[u, v \text{ separated in } (S,T)] \\ &\quad= \frac{d_{B_1}(s_1,t_1)}{d_{B_1}(s_1,t_1) + d_{B_2}(s_2,t_2)} \cdot \frac{d_{B_1}(u,t_1)}{d_{B_1}(s_t,t_1)} + \frac{d_{B_2}(s_2,t_2)}{d_{B_1}(s_1,t_1) + d_{B_2}(s_2,t_2)} \cdot \frac{d_{B_2}(v,s_2)}{d_{B_2}(s_2,t_2)} \\ &\quad= \frac{d_B(u,v)}{d_B(s,t)}, \end{aligned}$$

   which proves (a stronger version of) the claim. (Note that this really proves that if u and v lie on a path from s to t, the embedding maintains their distance exactly.)

2. $B = B_1 \parallel B_2$. Since the terminals for B_1, B_2 and B are the same, note that (23.4.2) and (23.4.3) both follow directly from the induction hypothesis; hence it suffices for us to show (23.4.4).

   Firstly, if u and v are in the same B_i, the claim readily follows by the induction hypothesis. So assume $u \in B_1$ and $v \in B_2$, and a shortest path from u to v in B goes through $s = s_1 = s_2$ (it either goes through this node, or $t = t_1 = t_2$, so this will be without loss of generality). We have

$$\begin{aligned} Pr[u, v \text{ separated in } (S,T)] &= Pr[u \in S_1] \cdot Pr[v \in T_2] + Pr[u \in T_1] \cdot Pr[v \in S_2] \\ &= Pr[u, t_1 \text{ separated in } (S_1, T_1)] \cdot Pr[v, s_2 \text{ separated in } (S_2, T_2)] \\ &\quad + \; Pr[u, s_1 \text{ separated in } (S_1, T_1)] \cdot Pr[v, t_2 \text{ separated in } (S_2, T_2)]. \qquad (23.4.5) \end{aligned}$$

   As before, we will argue that each of these probabilities are bounded above and below by induction, and hence the overall probability for u and v is bounded above and below appropriately.

   Let $\alpha = d_{B_1}(u, s_1)$, and $\beta = d_{B_2}(v, s_2)$. By definition of $B_1 \parallel B_2$, we have that $d_B(s,t) = d_{B_1}(s_1, t_1) = d_{B_2}(s_2, t_2)$ (since B is a bundle), and by our assumption, $d_B(u,v) = \alpha + \beta \leq 2d_B(s,t) - (\alpha + \beta)$. We therefore have (omitting some algebra here)

$$\frac{d_B(u,v)}{2d_B(s,t)} = \frac{1}{2}\frac{\alpha+\beta}{d_B(s,t)} \leq \frac{\alpha}{d_B(s,t)} \cdot \frac{(d_B(s,t)-\beta)}{d_B(s,t)} + \frac{\beta}{d_B(s,t)} \cdot \frac{(d_B(s,t)-\alpha)}{d_B(s,t)} \leq \frac{\alpha+\beta}{d_B(s,t)} = \frac{d_B(u,v)}{d_B(s,t)},$$

   and as the reader might expect, the middle expression is precisely $Pr[u, v \text{ separated in } (S,T)]$ from equation (23.4.5).

■

23.4.1 Embedding all SP graphs

The overall idea of the embedding from SP into ℓ_1 is to preprocess a given graph G into another graph $\hat{G}$ incurring a distortion of 2. This graph $\hat{G}$ is such that when we randomly delete edges from $\hat{G}$ (by a process similar to that for outerplanar graphs), the subgraph H obtained is one whose 2-connected components are all bundles; furthermore, H maintains distances in $\hat{G}$ to within a factor of 4 in expectation. Finally, the 2-connected components of H are then mapped into ℓ_1 using the above distortion 2 embedding. This sequence of steps will be fully covered in the next lecture.

References

Gupta, A., Newman, I., Rabinovich, Y., and Sinclair, A. (1999). Cuts, trees and ℓ_1-embeddings of graphs. In *40th Annual Symposium on Foundations of Computer Science (New York, 1999)*, pages 399–408. IEEE Computer Soc., Los Alamitos, CA.

Rabinovich, Y. and Raz, R. (1998). Lower bounds on the distortion of embedding finite metric spaces in graphs. *Discrete Comput. Geom.*, 19(1):79–94.