Discrete metric spaces

Victor Chepoi

LIF, Université d'Aix-Marseille II

Purpose of this minicourse: to present the most important classical and recent results about metric spaces.

Part 0: Preliminaries and definitions

Part I: Isometric embeddings into l_2-, l_∞, and l_1-spaces

Part II: Median spaces, hyperconvex spaces, CAT(0) spaces

Part III: Approximate isometric embeddings into l_1 and l_2

Part IV: Graph classes defined by metric properties

#1 Definitions I: metric spaces

metric space: (X,d), where X is a set and
$d: X\times X \longrightarrow \mathbb{R}_+$ a function called distance
such that

(i) $d(x,y)=d(y,x)$ $\forall x,y\in X$;

(ii) $d(x,y)=0 \Leftrightarrow x=y$;

(iii) $d(x,y)\leq d(x,z)+d(z,y)$
(triangle inequality)

Examples:

(i) norm metrics (Minkowski metrics):

$(\mathbb{R}^m, d_{\|\cdot\|})$, where $d_{\|\cdot\|}(x,y)=\|x-y\|$ $\forall x,y\in\mathbb{R}^m$

ℓ_p-metrics $(p\geq 1)$:

$$d_{\ell_p}(x,y)=\left(\sum_{k=1}^{m}|x_k-y_k|^p\right)^{\frac{1}{p}}, \qquad x=(x_1,\dots,x_m),\ y=(y_1,\dots,y_m)$$

ℓ_p^m denotes the metric space $(\mathbb{R}^m, d_{\ell_p})$

Three basic host ℓ_p-metric spaces:

ℓ_2^m $\quad d_{\ell_2}(x,y)=\left(\sum_{k=1}^{m}|x_k-y_k|^2\right)^{\frac{1}{2}}$ – Euclidean distance

ℓ_1^m $\quad d_{\ell_1}(x,y)=\sum_{k=1}^{m}|x_k-y_k|$ – ℓ_1-distance or rectilinear distance

ℓ_∞^m $\quad d_{\ell_\infty}(x,y)=\max\{|x_k-y_k| : 1\leq k\leq m\}$ – Chebyshev distance

Definitions I: metric spaces (cont.)

(ii) Hamming distance d_H:

$$d_H(x,y) = |\{i \in \{1,\dots,m\} : x_i \neq y_i\}| \text{ for } \forall x,y \in \mathbb{R}^m.$$

For binary vectors $x,y \in \{0,1\}^m$, the Hamming distance $d_H(x,y)$, the l_1-distance $d_{l_1}(x,y)$, and the graph distance $d(x,y)$ of the m-cube coincide.

(iii) standard graph distance d_G:

$G = (V, E)$ - connected, not necessarily finite, undirected and unweighted graph

length of a path – number of edges in this path

$d_G(x,y) =$ the length of a shortest path connecting two vertices x, y of G

(iv) finite metric spaces:

(V_n, d), where $V_n = \{1, \dots, n\}$ (for convenience) and d is a metric on V_n

Distance matrix: the $n \times n$ symmetric matrix D whose (i,j)-th entry is $d(i,j)$ for all $i,j \in V_n$.

Metric cone: set $E_n := \{ij : i,j \in V_n, i \neq j\}$; then any distance d on V_n can be viewed as a vector $(d_{ij})_{1 \leq i < j \leq n} \in \mathbb{R}^{E_n}$.

The triangle inequalities $d(i,j)+d(i,k)\geq d(k,j)$ for all $i,j,k\in V_n$ define a convex cone in the space $\mathbb{R}^{E_n}$ called the semimetric cone and denoted by MET_n.

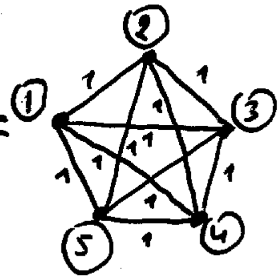
Particular finite metric spaces:

(i) planar metrics (V_n, d)

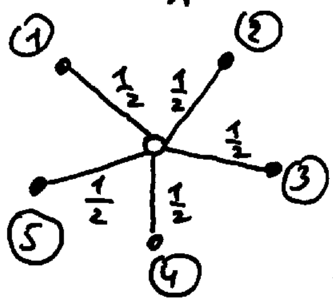
(ii) tree metrics (V_n, d)

one can construct a planar graph $G=(V,E)$ or a tree $T=(V,E)$ such that $V_n \subseteq V$ and $d = d_G|_{V_n}$

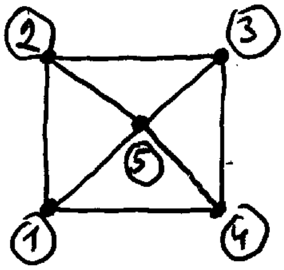
$$D_n = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & & \ddots & & 1 \\ \vdots & & & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} =$$



$=$



$$D = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 \\ & 0 & 1 & 2 & 1 \\ & & 0 & 1 & 1 \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} =$$



not a tree metric

Definitions II: isometric embeddings

A metric space (X, d) is isometrically embeddable into a metric space (X', d') if there exists a mapping φ (the isometric embedding) from X to X' such that

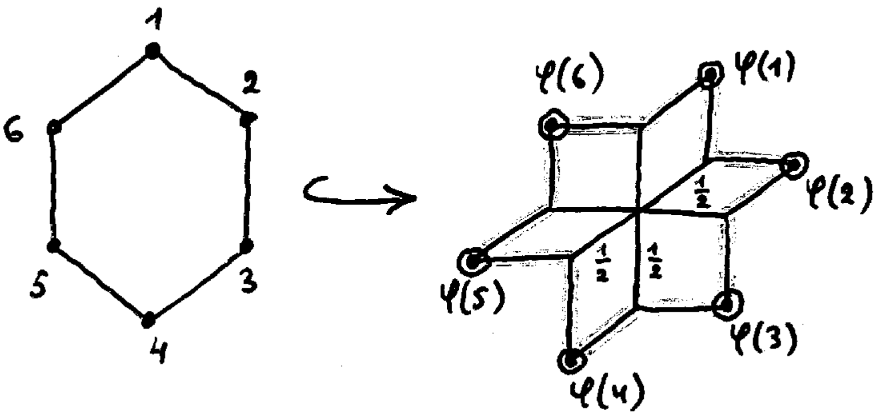
$$d(x,y) = d'(\varphi(x), \varphi(y)) \quad \forall\, x, y \in X$$

Then (X, d) is said to be an isometric subspace of (X', d').

(X, d) is ℓ_p-embeddable if (X, d) is isometrically embeddable into the space ℓ_p^m for some $m \geq 1$.

Example: the ℓ_1-embedding of the metric of the 6-cycle

$$D = \begin{pmatrix} 0 & 1 & 2 & 3 & 2 & 1 \\ & 0 & 1 & 2 & 3 & 2 \\ & & 0 & 1 & 2 & 3 \\ & & & 0 & 1 & 2 \\ & & & & 0 & 1 \\ & & & & & \end{pmatrix}$$



A metric space (X, d) is distortion β embeddable into a metric space (X', d') if there exists a mapping φ (the β-embedding) from X to X' such that

$$d(x,y) \leq d'(\varphi(x), \varphi(y)) \leq \beta \cdot d(x,y) \quad \forall\, x, y \in X.$$

(or equivalently – for ℓ_p-embeddings – if there exists a mapping $\psi: X \to X'$ such that

$$\frac{1}{\beta} \cdot d(x,y) \leq d'(\psi(x), \psi(y)) \leq d(x,y) \quad \forall\, x, y \in X)$$

Definitions II: isometric embeddings

The host space (X', d') is said to have order of congruence p if, for every metric space (X, d),

$$(Y, d) \text{ embeds isometrically into } (X', d') \text{ for every } Y \subset X,\ |Y| \leq p$$
$$\Downarrow$$
$$(X, d) \text{ embeds into } (X', d')$$

and p is the smallest such integer (possibly infinite).

Questions

For a given host space or class of host spaces:

(i) is the order of congruence finite?

(ii) is the order of congruence bounded by a constant?

(iii) can the isometric subspaces be effectively characterized?

(iv) is the decision question "Is an input finite metric space isometrically embeddable into the host space?" polynomial or NP-complete?

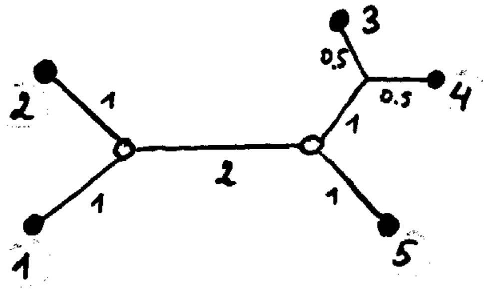
(v) find small distortion embeddings of finite metric spaces into the host space(s).

(vi) find small distortion embeddings into host spaces of small dimension.

Tree metrics

Tree metric a finite metric space isometrically embeddable into a weighted tree

$$D = \begin{pmatrix} 0 & 2 & 4.5 & 4.5 & 4 \\ & 0 & 4.5 & 4.5 & 4 \\ & & 0 & 1 & 2.5 \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$



Theorem (Zaretskii, 1965; Buneman, 1971) For a finite metric space (V_n, d) the following conditions are equivalent:

(i) (V_n, d) is a tree metric;

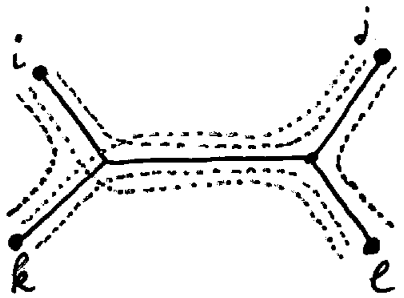
(ii) every quadruplet of V_n isometrically embeds into a tree;

(iii) (V_n, d) satisfies the following four-point condition:

$$d(i,j)+d(\ell,k) \le \max\{d(i,\ell)+d(j,k),\, d(i,k)+d(j,\ell)\}.$$

Idea of the proof:

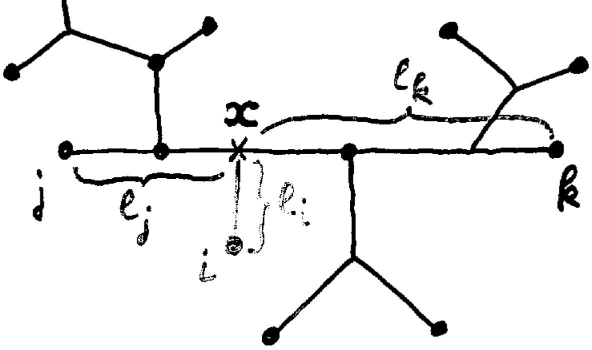
$\Rightarrow$



$\Leftarrow$ By induction. For a current point i find the points $j<i$ and $k<i$ // ℓ_i minimizing $\frac{1}{2}(d(i,j)+d(i,k)-d(j,k))$

On the path of the current tree pick the point x such that $d(x,j)=\frac{1}{2}(d(j,i)+d(j,k)-d(i,k))=\ell_j$

To the current tree add the edge xi of length ℓ_i so that i becomes a leaf and x an inner vertex.



Isometric embeddings into ℓ_p-spaces

Theorem 1 (Bretagnolle, Dacunha-Castelle, Krivine, 1966)
A metric space (X,d) is L_p-embeddable if and only if every finite subspace of (X,d) is L_p-embeddable.

Theorem 2 (Malitz, Malitz, 1992) Let $p, m \geq 1$ be integers. Then a metric space (X,d) is ℓ_p^m-embeddable if and only if every finite subspace of (X,d) is ℓ_p^m-embeddable.

Theorem 3 (Fréchet) Any n point metric space (V_n, d) can be isometrically embedded into ℓ_∞^n.

Proof: For each point $i \in V_n$ define a cooridate $\varphi_i : V_n \to \mathbb{R}_+$ by setting $\varphi_i(j) = d(i,j)$ and let $\varphi(j) = (\varphi_1(j), \dots, \varphi_n(j))$. We claim that φ is an isometric embedding into ℓ_∞^n.

Indeed:

$$\begin{aligned}
\|\varphi(j) - \varphi(k)\|_\infty &= \max_{i \in V} |\varphi_i(j) - \varphi_i(k)| \\
&= \max_{i \in V} |d(i,j) - d(i,k)| \\
&\leq d(j,k) \text{ by triangle inequality.}
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\|\varphi(j) - \varphi(k)\|_\infty &\geq |\varphi_k(j) - \varphi_k(k)| \\
&= |d(j,k) - d(k,k)| \\
&= d(j,k) \qquad \square
\end{aligned}$$

Theorem 4 (Menger, 1928) Given $m \geq 1$, a metric space (X, d) can be isometrically embedded in ℓ_2^m if and only if for every $Y \subseteq X$ with $|Y| = m+3$, (Y, d) can be isometrically embedded in ℓ_2^m.

Proof sketch:

Stronger assertion: (X, d) can be isometrically embedded in ℓ_2^m and not into ℓ_2^{m-1} if and only if there exist a subset $Y = \{x_0, x_1, \dots, x_m\}$ of X such that

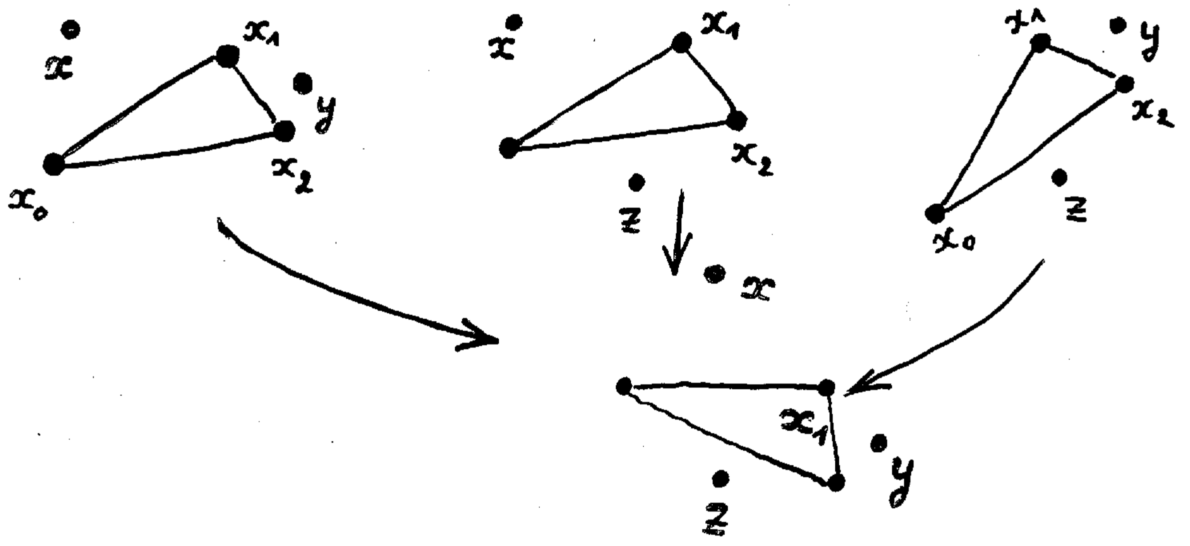
(i) (Y, d) can be embedded into ℓ_2^m but not in ℓ_2^{m-1};

(ii) for every $x, y \in X$, the metric space $(Y \cup \{x, y\}, d)$ can be isometrically embedded in ℓ_2^m.

Notice that:

(i) if φ is an isometric embedding of Y in ℓ_2^m, then $\varphi(Y)$ has full affine rank $m+1$;

(ii) if φ', φ'' are two isometric embeddings of Y in ℓ_2^m, then we can find an orthogonal transformation mapping every $\varphi'(x_i)$ into $\varphi''(x_i)$, $x_i \in Y$. donc, une isométrie



Question: Given a finite metric space (V_n, d), can we check in polynomial time if (V_n, d) embeds in ℓ_2?

Answer: Let φ be a mapping from V_n to ℓ_2.
Let $\varphi(i) = v_i$ and assume that $\varphi(1) = v_1 = \vec{0}$.

Then

$$\|v_i - v_j\|^2 = d_{ij}^2 \quad \forall i,j \in V \iff$$

$$\|v_i\|^2 + \|v_j\|^2 - 2\langle v_i, v_j\rangle = d_{ij}^2 \iff$$

$$\langle v_i, v_j\rangle = \frac{1}{2}\left(d_{1i}^2 + d_{1j}^2 - d_{ij}^2\right),$$

because $\|v_i\|^2 = \|v_i - v_1\|^2 = d_{1i}^2$, $\|v_j\|^2 = \|v_j - v_1\|^2 = d_{1j}^2$

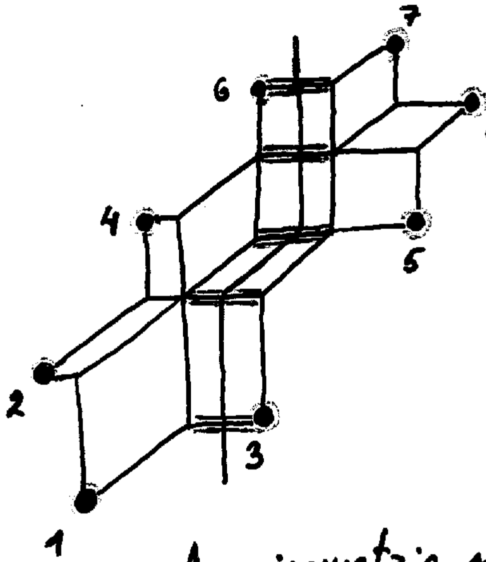
Denote $A_{ij} = \frac{1}{2}\left(d_{1i}^2 + d_{1j}^2 - d_{ij}^2\right)$ and consider the matrix A. Then φ is an isometric embedding in ℓ_2 if and only if $A_{ij} = \langle v_i, v_j\rangle$, i.e. A can be written as $B^T B$, or, in other words, A is a positive semi-definite matrix (Schoenberg, 1938)

recognition of positive semi-definiteness can be done in polynomial time using an algorithm based on Gaussian elimination.

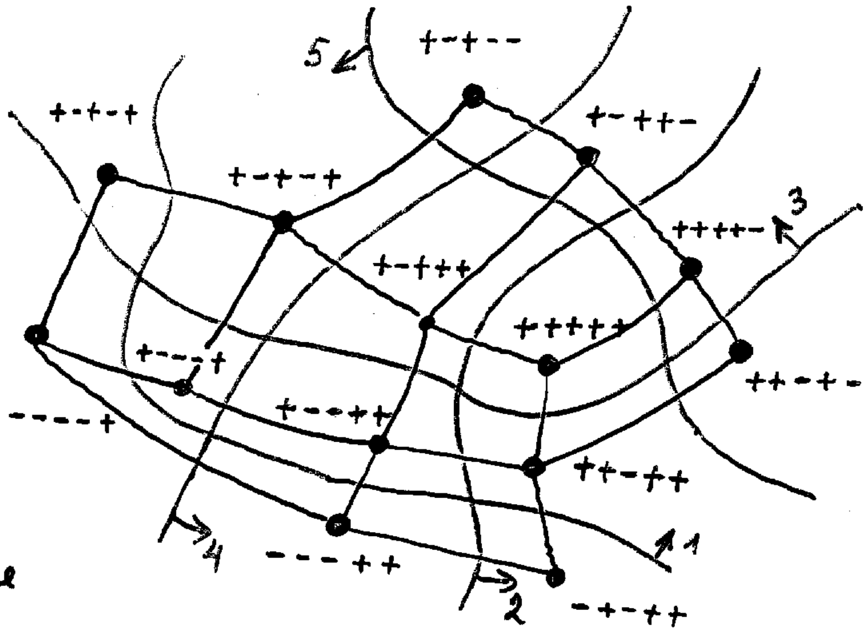
Isometric embeddings into ℓ_1

Theorem 5 (Karzanov, 1986) Deciding if a finite metric space (V, d) ℓ_1-embeds is NP-complete.

Reduction from MAX CUT



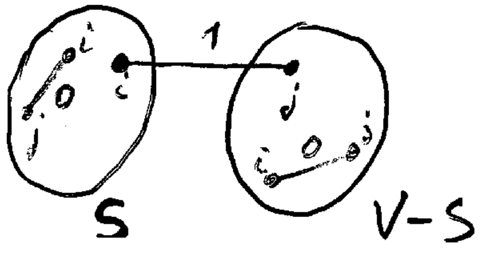
An isometric subspace of ℓ_1^3



An arrangement of five pseudoline and the isometric embedding of its graph of regions into the cube $\{\pm 1\}^5$.

<u>Cut semimetric</u>: for $S \subseteq V_n = \{1, \dots, n\}$

$$\delta(S)_{ij} = \begin{cases} 1 & \text{if } |S \cap \{i,j\}| = 1 \\ 0 & \text{otherwise} \end{cases}$$



Theorem 6: For a finite metric space (V_n, d) the following conditions are equivalent:

(i) (V_n, d) is ℓ_1-embeddable, i.e., there exist n vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ for some m such that $\|u_i - u_j\|_1 = d_{ij}$ $\forall i,j \in V_n$;

(ii) $d = \sum_{S \subseteq V_n} \lambda_S \delta(S)$ for some nonnegative λ_S, i.e. d belongs to the cut cone $CUT_n = \{\sum_{S \subseteq V_n} \lambda_S \delta(S) : \lambda_S \geq 0 \ \forall S \subseteq V_n\}$

(iii) there exist a measure space $(\Omega, \mathcal{A}, \mu)$ and events $A_1, \ldots, A_n \in \mathcal{A}$ such that $d_{ij} = \mu(A_i \Delta A_j)$ $\forall i,j \in V_n$.

Remark: In case of embedding into hypercubes or Hamming metric spaces, λ_S is a nonnegative integer and $d_{ij} = |A_i \Delta A_j|$.

Remark (Ball, 1990): The dimension m of $\mathbb{R}^m$ in (i) can be as large as $\frac{(n-3)(n-2)}{2}$ for $n \geq 4$! Fichet (1988) showed that $m \leq \frac{n(n-1)}{2} - 1$.

Remark: The difficulty to find an ℓ_1-embedding via (ii) consists in finding the cuts $(S, V_n - S)$ such that $\lambda_S > 0$ (by Caratheodory's theorem their number is $\leq \frac{n(n-1)}{2}$), i.e. those cuts which define the embedding.

<u>Theorem 7</u> (Bandelt & Chepoi, 1995) A metric space (X, d) isometrically embeds into $(\mathbb{R}^2, d_{\ell_1})$ if and only if (Y, d) embeds for any $Y \subseteq X$, $|Y| \leq 6$.

The idea of the proof will be given below.

<u>Remark</u>: From Theorem 7 follows that the congruence orders of ℓ_1^2 and ℓ_∞^2 is 6.

<u>Remark</u> (Bandelt, Chepoi, Laurent, 1997) The congruence order of ℓ_1^m is at least m^2 for $m \geq 3$ odd and at least $m^2 - 1$ for $m \geq 4$ even.

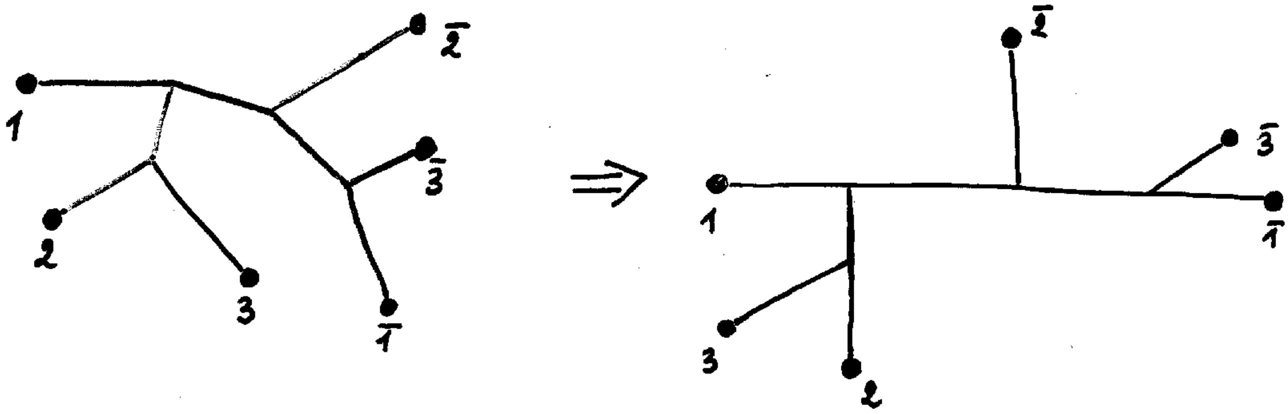
<u>Open question</u>: Find the congruence order of ℓ_1^3. Is it finite? (It is only known to be ≥ 10).

<u>Remark</u> (Jeff Erikson, 2004) The congruence order of ℓ_∞^3 is not bounded!

<u>Open question</u>: What is the largest set of an equilateral set of ℓ_1^m? It is conjectured to be $2m$, but it is known to be $\leq m \log_2 m$ (Alon, Pudlak, 2003)

Examples of ℓ_1-embeddable spaces

(i) Tree metrics (folklore): any tree with m leaves isometrically embeds into $\ell_1^{\lfloor \frac{m}{2} \rfloor}$.



(ii) Spherical metrics (Kelly, 1970)

$S_m = \{ x \in \mathbb{R}^{m+1} : \sum_{i=1}^{m+1} x_i^2 = 1 \}$ – m-dimensional unit sphere

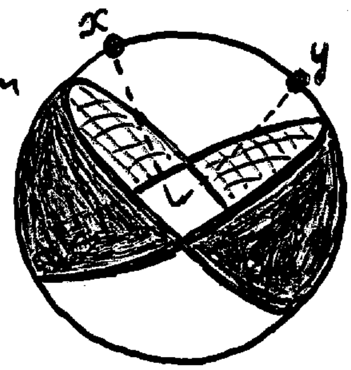
$d_S(x,y) := \arccos(x^T y) \quad \forall x,y \in S_m$ – spherical distance

the geodesic distance on the sphere S_m between the points x and y (great circle metric)

For $x \in S_m$, let $H(x) = \{ y \in S_m : d_S(x,y) \le \frac{\pi}{2} \}$ be the hemisphere containing x.

Consider the measure μ on S_m defined by

$$\mu(A) = \frac{vol(A)}{vol(S_m)} \quad \text{for } A \subseteq S_m$$



Theorem 8 (Kelly, 1970)

$$\mu(H(x) \Delta H(y)) = \frac{1}{\pi} \arccos(x^T y) = \frac{1}{\pi} d_S(x,y)$$
$$\forall x,y \in S_m$$

(iii) ℓ_2-distances

Theorem 9 (Schoenberg, 1935, Kelly, 1975) For a finite metric space (V_n, d), d is isometrically ℓ_2-embeddable implies that d is isometrically ℓ_1-embeddable.

Idea of the proof: Let $u_1, \dots, u_n \in \mathbb{R}^m$, $d_{ij} = \|u_i - u_j\|_2 \ \forall i,j$

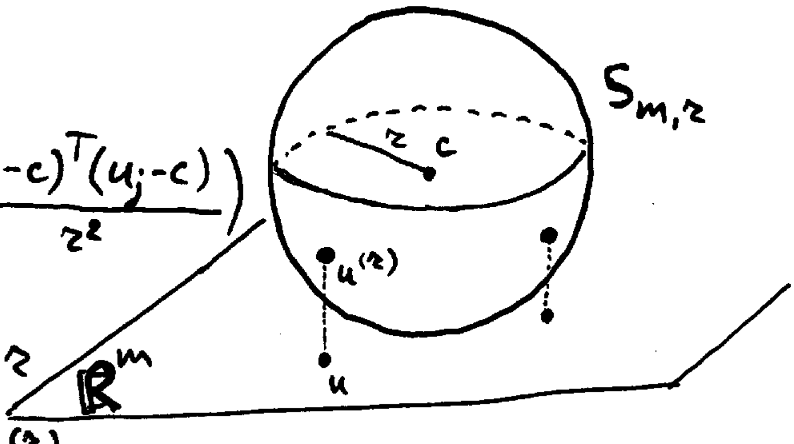
Express d as a limit of distances $d^{(r)}$ $(r \to \infty)$, where each $(V_n, d^{(r)})$ can be isometrically embeddable into the sphere $S_{m,r}$ of radius r and apply Theorem 8.

Let $S_{m,r}$ be the sphere of $\mathbb{R}^{m+1}$ of radius r and center $c = (0, \dots, 0, r)$. Lift every $u \in \mathbb{R}^m$ with $\|u\|_2 \leq r$ to a point $u^{(r)} \in S_{m,r}$ by setting $u^{(r)} := (u, r - \sqrt{r^2 - (\|u\|_2)^2})$

Let $r \geq \max_{i=1}^n \|u_i\|_2$.

Set $d^{(r)}(i,j) = r \cdot \arccos\left(\frac{(u_i - c)^T (u_j - c)}{r^2}\right)$

the spherical distance in $S_{m,r}$ between the points $u_i^{(r)}$ and $u_j^{(r)}$



$\lim_{r \to \infty} d^{(r)}(i,j) = \|u_i - u_j\|_2$ because

$$d^{(r)}(i,j) \approx r \arccos\left(1 - \frac{(\|u_i - u_j\|_2)^2}{2r^2}\right) \approx \|u_i - u_j\|_2$$

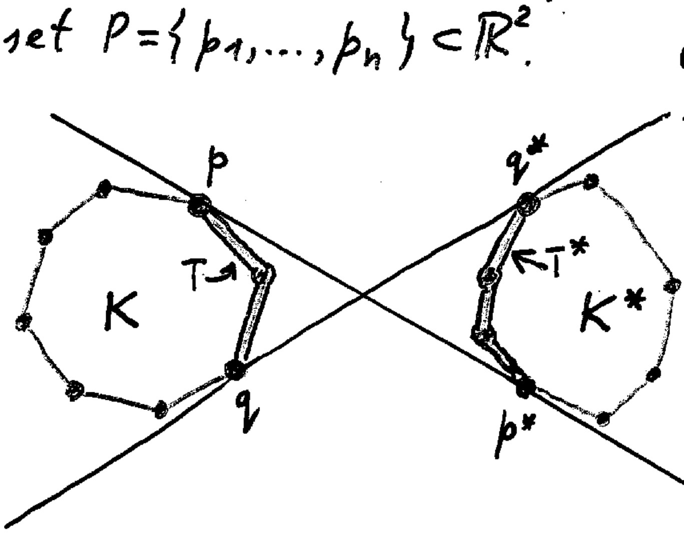
(iv) *projective metrics in the plane*

A continuous metric d on $\mathbb{R}^2$ is called a projective metric if it satisfies $d(x,z) = d(x,y) + d(y,z)$ for any collinear points x, y, z lying in that order on a common line

The following theorem proved independently by Alexander'78 and Ambartzumian'77 gives a simple solution to the Hilbert's fourth problem in the plane.

Theorem 10 (Alexander, 1978; Ambartzumian, 1977) Let d be a projective metric on $\mathbb{R}^2$. Then $(\mathbb{R}^2, d)$ is L_1-embeddable, namely there exists a positive Borel measure μ on the lines of $\mathbb{R}^m$ satisfying $2d(x,y) = \mu([[x,y]])$ for $x, y \in \mathbb{R}^2$, where $[[x,y]]$ denotes the set of lines crossing the segment $[x,y]$

The main step in its proof is to define explicitly the measure on lines crossing the segments $[p_i, p_j]$ for any finite set $P = \{p_1, \dots, p_n\} \subset \mathbb{R}^2$.



$\sigma(K, K^*) = d(p, p^*) + d(q, q^*) - d(T) - d(T^*)$, where $d(T)$, $d(T^*)$ are the perimeters (with respect to d) of the chains T and T^*.

For any line ℓ, denote by K the convex hull of points of P left from ℓ, and by K^* the convex hull of points of P right from ℓ.

Set $d(\ell) = \sigma(K, K^*)$

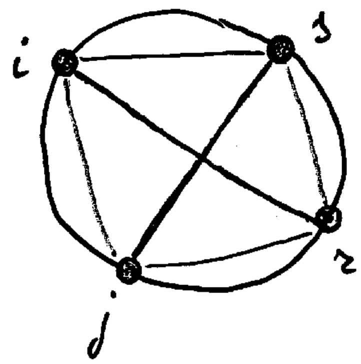
Theorem 11: $\forall p_i, p_j \in P$, $2d(p_i, p_j) = \sum \{ d(\ell) : \ell \cap [p_i, p_j] \neq \emptyset \}$.

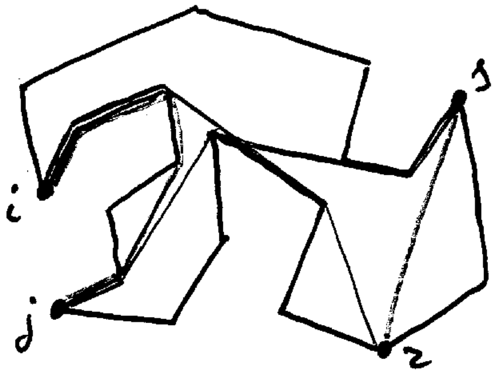
(V) Kalmanson distances

A distance d on V_n is called a Kalmanson distance there exists a (circular) ordering $1, \dots, n$ of the points of V_n such that

$$\max\{d_{ij} + d_{rs}, d_{is} + d_{jr}\} \leq d_{ir} + d_{js}$$

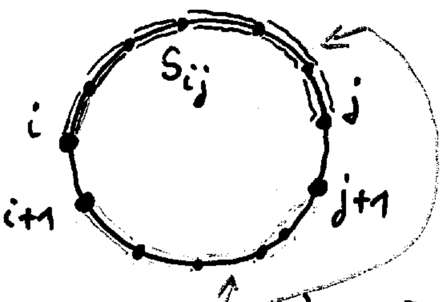
for all $i < j < r < s$ in the circular order





<u>Theorem 12</u> (Chepoi & Fichet, 1996) A distance d is a Kalmanson distance if and only if it is circular decomposable. In particular, Kalmanson distances are ℓ_1-embeddable.

<u>Idea of the proof</u>:



Circular cut $\{S_{ij}, \bar{S}_{ij}\}$

$$\alpha_{ij} = d_{ij+1} + d_{i+1j} - d_{ij} - d_{i+1\,j+1}$$

Notice that $\alpha_{ij} \geq 0$ for any circular order compatible with a Kalmanson distance d, but in general we can have $\alpha_{ij} < 0$

<u>Lemma</u>: For any finite metric space (V_n, d) and any circular ordering of V_n, we have

$$2\, d(u,v) = \sum\{\alpha_{ij} : \{S_{ij}, \bar{S}_{ij}\} \text{ separates } u \text{ and } v\} \quad \forall u, v \in$$

(vi) Totally decomposable metrics

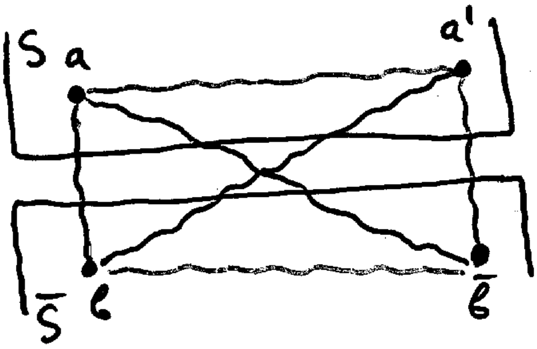
Recall that d is l_1-embeddable iff $d = \sum_{S \subseteq V_n} \lambda_S \delta(S)$.

In general, (i) it is difficult to find $\lambda_S \geq 0 \ \forall S \subseteq V_n$;
(ii) the l_1-decomposition is not unique;
(iii) there is an exponential number of cuts $(S, \bar{S})$ participating in the decomposition.

Bandelt & Dress (1992) a canonical decomposition of every finite metric into a sum of $O(n^2)$ cut metrics plus a residue.

For a cut $\{S, \bar{S}\}$ of V_n define its isolation index by

$$\alpha_S = \frac{1}{2} \min_{\substack{a,a' \in S \\ b,\bar{b} \in \bar{S}}} \Big\{ \max \{ d(a,b) + d(a',b'), d(a,b') + d(a',b), d(a,a') + d(b,b') \} - d(a,a') - d(b,b') \Big\}$$



$\{S, \bar{S}\}$ is a d-cut if $\alpha_S > 0$

Theorem 13 (Bandelt & Dress, 1992) For any metric d on V_n

$$d = \sum_{\substack{\{S,\bar{S}\} \text{ is} \\ \text{a d-cut}}} \alpha_S \delta(S) + d'$$

prime residue

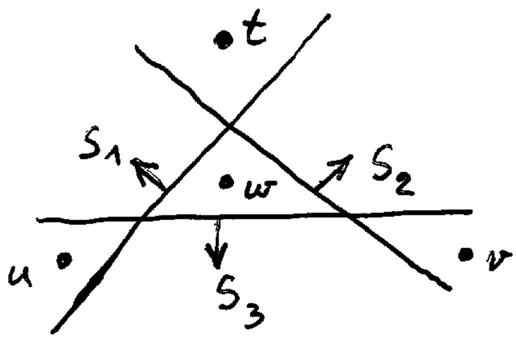
The metric d is called totally decomposable if $d' = 0$, i.e. if $d = \sum_{\{S,\bar{S}\}\text{-d-cut}} \alpha_S \delta(S)$.

Bandelt & Dress (1992) established the following properties of totally decomposable metrics and d-cuts:

(i) there exists only $O(n^2)$ d-cuts (in fact, their (0,1)-vectors are linearly independent);

(ii) the d-cuts can be constructed in polynomial time;

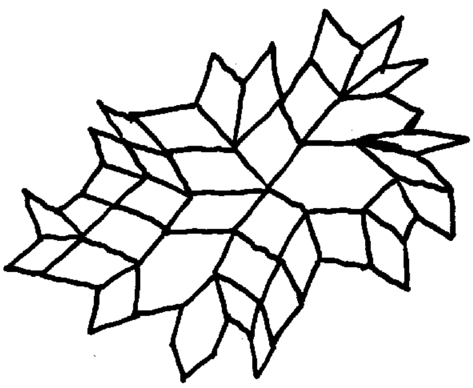
(iii) the d-cuts are weakly compatible, i.e. for any three d-cuts $\{S_1, \bar{S}_1\}, \{S_2, \bar{S}_2\}, \{S_3, \bar{S}_3\}$ if $S_1 \cap S_2 \cap S_3 \neq \emptyset$ implies $\bar{S}_1 \cap \bar{S}_2 \cap \bar{S}_3 = \bar{S}_i \cap \bar{S}_j$ for some i, j.

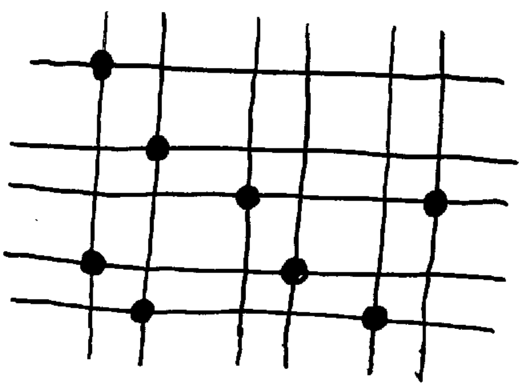


not weakly compatible

Theorem 14 (Bandelt & Dress, 1992) A finite metric space (V_n, d) is totally decomposable if and only if every 5-point subspace of V_n is totally decomposable.

Examples of totally decomposable metrics: Kalmanson metrics, tree metrics, finite subspaces of $(\mathbb{R}^2, d_{\ell_1})$ and Cartesian products of two trees, metric of some plane graphs, ...





<u>Theorem 7</u> (Bandelt & Chepoi, 1995) A metric space (X,d) isometrically embeds into $(\mathbb{R}^2, d_{\ell_1})$ if and only if (Y,d) embeds for any $Y \subseteq X$, $|Y| \leq 6$.

<u>Idea of the proof</u>:

(i) By compactness theorem (Theorem 2) it suffices to consider finite X;

(ii) Every finite isometric subspace of $(\mathbb{R}^2, d_{\ell_1})$ is totally decomposably, so using Theorem 14 one can test if (X,d) is totally decomposable by inspecting all (Y,d) $Y \subseteq X$, $|Y| \leq 5$.

(iii) So (X,d) is totally decomposable, and the d-cuts should define the embedding into $\mathbb{R}^2$. For this, the d-cuts should be represented into two chains:

$$S_1' \subseteq S_2' \subseteq \ldots \subseteq S_p', \quad \bar{S}_1' \supseteq \bar{S}_2' \supseteq \ldots \supseteq \bar{S}_p'$$

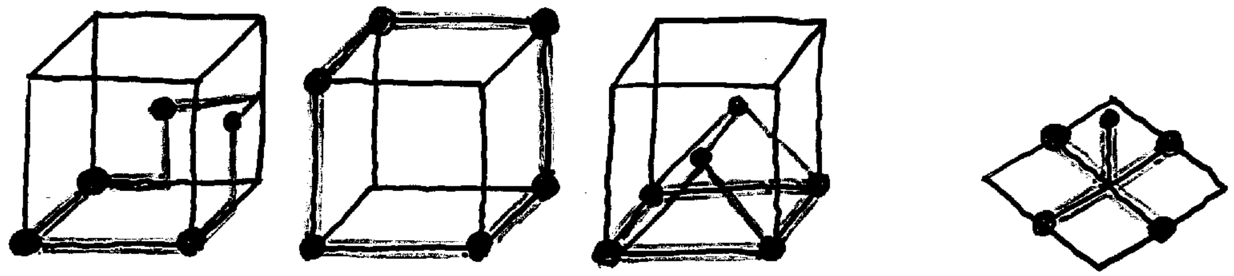
$$S_1'' \subseteq S_2'' \subseteq \ldots \subseteq S_q'', \quad \bar{S}_1'' \supseteq \bar{S}_2'' \supseteq \ldots \supseteq \bar{S}_q''.$$

To establish this, we show the following theorem

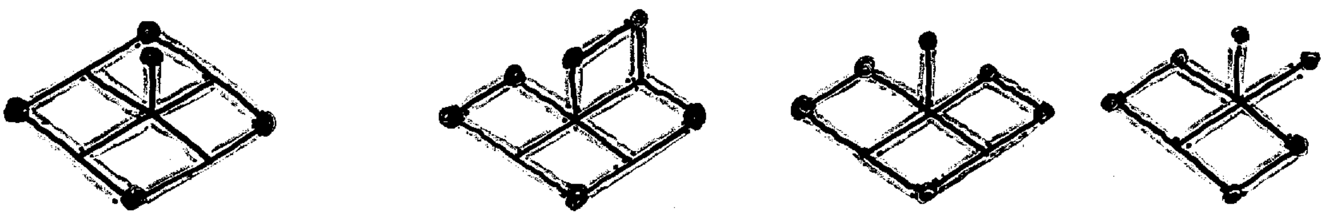
<u>Theorem</u> A totally decomposable space (X,d) embeds in $(\mathbb{R}^2, d_{\ell_1})$ if and only if for any d-cuts $\{S_1, \bar{S}_1\}, \ldots \{S_k, \bar{S}_k\}$ ($k \leq 5$ the ordered set of halfs $S_1, \ldots, S_k, \bar{S}_1, \ldots, \bar{S}_k$ has at most four minimal members (equivalently, the d-cuts can be partitionned into two chains if and only if any subset of at most 5 d-cuts can be partitionned).

(iv) Assume that $\forall\, Y \subset X$, $|Y| \leq 6$, we have that (Y,d) is embeddable into $(\mathbb{R}^2, d_{e_1})$, however X contains $k \leq$ d-cuts $\{S_1, \bar{S}_1\}, \dots, \{S_k, \bar{S}_k\}$ that violate the condition of previous Theorem.

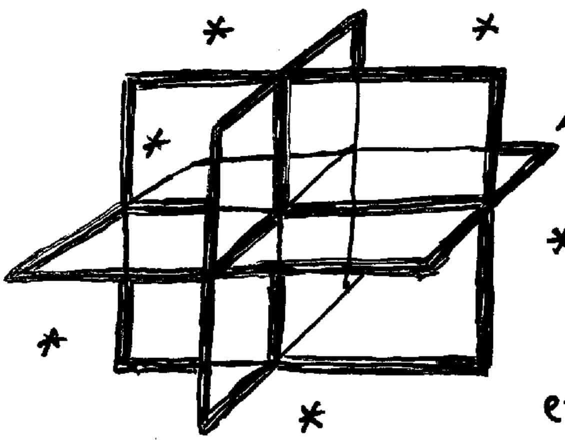
We consider the cases $k=3$, $k=4$, $k=5$, and in each case we derive a 6-point subspace of X which is not embeddable into $(\mathbb{R}^2, d_{e_1})$. Those critical subspaces (the minors) are:



C_5 $\qquad$ C_6 $\qquad$ $K_3 \times K_2$ $\qquad$ K_5



For example, if $k=3$, then $S_1, S_2, S_3, \bar{S}_1, \bar{S}_2, \bar{S}_3$ are all minimal by inclusion. Then $S_i \cap S_j \neq \emptyset$, $S_i \cap \bar{S}_j \neq \emptyset$ $\forall i,j$ and we get C_6



If the intersection of every two non-complementary halfspace intersect X, then

* then there is a point of X in every orthant except maybe two opposite orthant

#2 Low-distortion embedding of finite metric spaces into ℓ_1, ℓ_2, and ℓ_∞

A mapping $\varphi: X \to X'$, where (X,d) and (X',d') are metric spaces, has distortion at most β, or is called a β-embedding, where $\beta \geq 1$, if there is an $c \in (0,+\infty)$ such that for all $x,y \in X$

$$c\, d(x,y) \leq d'(\varphi(x),\varphi(y)) \leq c\beta\, d(x,y).$$

If X' is a normed space, we usually require $c = \frac{1}{\beta}$ or $c=1$.

If $c=1$, we obtain

$$d(x,y) \leq d'(\varphi(x),\varphi(y)) \leq \beta\, d(x,y). \quad \text{(dilation)}$$

If $c = \frac{1}{\beta}$, we obtain

$$\frac{1}{\beta}\, d(x,y) \leq d'(\varphi(x),\varphi(y)) \leq d(x,y) \quad \text{(contraction)}$$

Mappings with bounded distortion are also called bi-Lipschitz mappings.

(i) Bourgain's, Matousek's, and Rao's low-distortion embeddings theorems;

(ii) The Johnson-Lindenstrauss flattening lemma;

(iii) Probabilistic embeddings into tree-metrics;

(iv) Applications of embeddings;

(v) Proof of Bourgain's theorem.

<u>Theorem 15</u> (Bourgain, 1985) Any n-point metric space (V_n, d) can be embedded into ℓ_1^m (in fact, into every ℓ_p) with distortion $O(\log n)$ and dimension $m = O(\log^2 n)$

<u>Idea of the proof</u> (more details below):

Set $\ell = \lfloor \log_2 n \rfloor$ and $q = \lfloor C \log n \rfloor$ (C is a suitable constant)

Consider an embedding into $\ell_1^{\ell q}$ with coordinates indexed by $i = 1, \dots, \ell$ and $j = 1, \dots, q$:

for each i, j, select a subset $A_{ij} \subseteq V_n$ by putting each $x \in V_n$ into A_{ij} with probability $\frac{1}{2^i}$, all random choices being mutually independent.

Set $\varphi(x)_{ij} = d(x, A_{ij}) := \min\{d(x,a) : a \in A_{ij}\}$

and

$$\varphi(x) = (\varphi(x)_{11}, \dots, \varphi(x)_{1q}, \dots, \varphi(x)_{\ell 1}, \dots, \varphi(x)_{\ell q})$$

Then $\varphi : V_n \to \ell_1^{\ell q}$ is an $O(\log n)$-distortion ℓ_1-embedding with propability at least $1/2$.

A_{*1} A_{*2} A_{*3}

<u>Remark</u>: The dimension m of ℓ_1^m in the original Bourgain' proof was exponential. It has been reduced to $O(\log^2 n)$ by Linial, London, and Rabinovich (1995) using Chernoff bounds.

Theorem 16 (Matoušek, 1996) For an integer $b > 0$ set $\beta = 2b-1$. Then any n-point metric space (V_n, d) can be embedded into ℓ_∞^m with distortion β, where $m = O(b n^{\frac{1}{b}} \log n)$

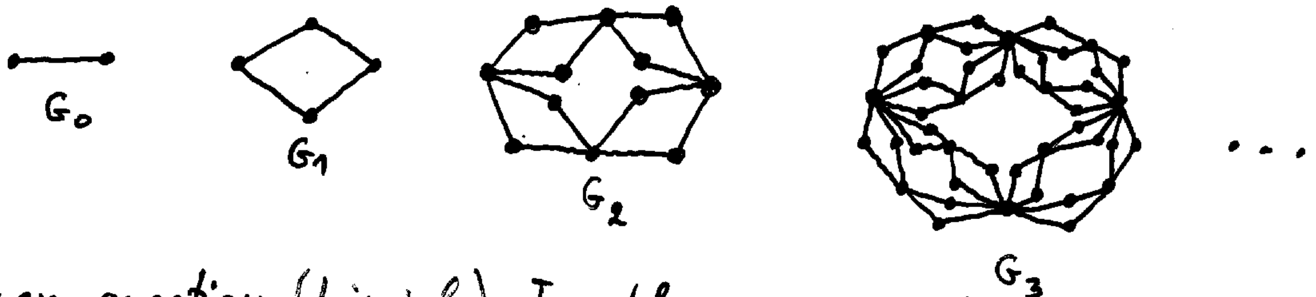
Remark: Recall that (V_n, d) embeds into ℓ_∞^n without any distortion according to Frechet's theorem.

Remark: For the special case $\beta = O(\log n)$, Matousek's result implies that (V_n, d) embeds into $\ell_\infty^{O(\log^2 n)}$ with $O(\log n)$ distortion.

Planar metric: A finite metric space (V_n, d) is planar if (V_n, d) isometrically embeds (without any distortion!) into a planar graph.

Theorem 17 (Rao, 1999) Any n-point planar metric can be embedded into ℓ_2 with distortion $O(\sqrt{\log n})$.

Remark: Newman & Rabinovich (2002) established that the $O(\sqrt{\log n})$ bound is sharp for the diamond graph (known also as Laakso's fractal)



Open question (Linial) Is there a constant C such that any planar metric embeds into ℓ_1 with distortion $\leq C$?

Open question (Linial, Rabinovich) Characterize planar metrics. In particular, given a metric, can one decide in polynomial time whether it is a planar metric?

Doubling dimension of a metric space (Assouad, 1983)

The doubling constant of a metric space (X,d) is the smallest value λ such that every ball in X can be covered by λ balls of half the radius.

The doubling dimension of (X,d) is $\log_2 \lambda$.

The doubling dimension of m-dimensional ℓ_p space is roughly m.

Theorem 18 (Assouad, 1983) If the doubling dimension of a metric space (X,d) is bounded, then for any $0<\alpha<1$ the metric space (X, d^{α}) (the snowflaked version of d) embeds into ℓ_2^m with distortion β, where m and β depend only on the doubling dimension of X.

Remark: Assouad (1983) conjectured that Theorem 18 holds even when $\alpha=1$, but Semmes (1996) disproved this conjecture. Gupta, Krauthgamer, Lee (2003) established that the result holds for trees with doubling dimension $< \infty$.

Edit (or Levenstein) distance: Σ – a finite alphabet, Σ^* all strings with symbols from Σ.
For $s, s' \in \Sigma^*$, $d_E(s,s')$ - the minimum number of edit operations (insertion, deletion, substitution) transforming s into s'.

Open question (Indyk) Is there a constant c such that the metric space (Σ^*, d_E) embeds into ℓ_1 with distortion $\leq c$?

Theorem 19 (*Johnson-Lindenstrauss flattening lemma, 1984*)
Let X be an n-point set in a Euclidean space, and let $\varepsilon \in (0,1]$ be given. Then X can be embedded into $\ell_2^{O(\varepsilon^{-2}\log n)}$ with distortion $(1+\varepsilon)$.

Remark: Theorem 19 can be viewed as a dimensionality reduction result: a set of points in a high-dimensional space is mapped to a space with low dimension, while (approximately) preserving important characteristics of the pointset.

Idea of the proof: Set $m := \frac{200 \ln n}{\varepsilon^2}$ and assume $m<n$ (otherwise, there is nothing to prove).

Let L be a random m-dimensional linear subspace of ℓ_2^n.

Let $p: \ell_2^n \to L$ be the orthogonal projection onto L.

Claim: For any two distinct points $x,y \in \ell_2^n$, the condition

$$(1-\tfrac{\varepsilon}{3})\mu \|x-y\|_2 \leq \|p(x)-p(y)\|_2 \leq (1+\tfrac{\varepsilon}{3})\mu \|x-y\|_2 \qquad (*)$$

is violated with probability at most n^{-2}.

Since $|X|=n$ and $\frac{n(n-1)}{2} < n^2$ pairs of distinct $x,y \in X$, there exists some L such that $(*)$ holds for all $x,y \in X$.

In this case, the mapping $p: X \to L$ has distortion $\beta \leq \frac{1+\varepsilon/3}{1-\varepsilon/3} < 1+\varepsilon$ for $\varepsilon \leq 1$.

The value of μ in previous claim is defined by:

<u>Lemma</u> (concentration of the length of the projection)
For a unit vector $x \in S^{n-1}$, let $f(x) = \sqrt{x_1^2 + \ldots + x_m^2}$ be the length of the projection of x on the subspace L_0 spanned by the first m coordinates. Then $f(x)$ is sharply concentrated around a suitable number $\mu = \mu(n,m)$

$$\Pr[f(x) \geq \mu + t] \leq 2e^{-t^2 n/2} \text{ and } \Pr[f(x) \leq \mu - t] \leq 2e^{-t^2 n/2},$$

where $\Pr$ is the uniform probability measure of the sphere S^n

<u>Theorem 20</u> (Brinkman, Charikar, 2003; see Lee & Naor, 2004 for a very short proof) There exists an n-point subset $X \subseteq \ell_1$ such that for any $\beta > 1$, if X embeds into ℓ_1^m with distortion β, then $m \geq n^{\Omega(1/\beta^2)}$. In other words, the dimensionality reduction is impossible in ℓ_1-metrics.

<u>Idea of the proof of Lee & Naor:</u>
Let G_k be the kth diamond graph with all edges of length 2^{-k}.

(i) G_k can be embedded with constant distortion into ℓ_1.
(ii) using simple counting and the "short diagonal lemma" it is shown that for every $1 < p \leq 2$, any embedding of G_k into ℓ_p incurs distortion $\geq \sqrt{1 + (p-1)k}$
(iii) ℓ_1^m is $O(1)$-isomorphic to ℓ_p^m, when $p = 1 + \frac{1}{\log m}$

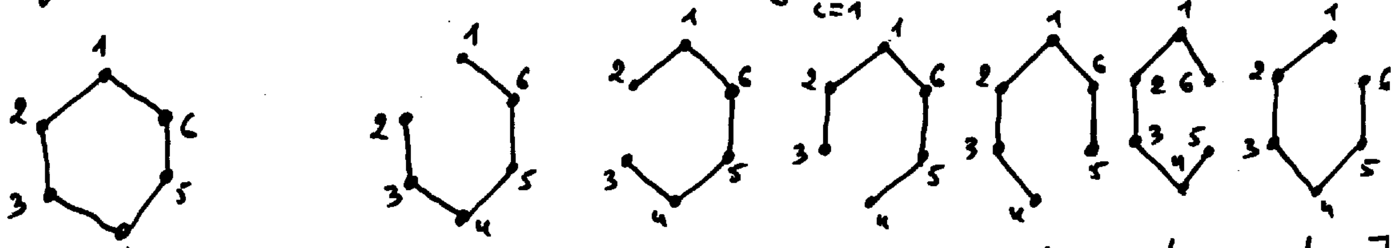
Dominating metric: let d, d' be metrics on the same set X. Then d' dominates d if $d'(x,y) \geq d(x,y)$ for all $x,y \in X$

Let $\mathcal{D}$ be a finite family of metrics on the same set X and let $\mathcal{P}$ be a probability distribution over $\mathcal{D}$. Then $(\mathcal{D}, \mathcal{P})$ β-probabilistically approximates a metric d on X if

(i) every metric $d_i \in \mathcal{D}$ dominates d;

(ii) $\forall x,y \in X$, $E_{d_i \in \mathcal{D}}[d_i(x,y)] = \sum_{i=1}^{k} p_i d_i(x,y) \leq \beta\, d(x,y)$.

Example: Let d be the graph metric of the cycle C_n. Let d_i be the graph metric of the path obtained by removing the i-th edge of C_n. The $\frac{1}{6}\sum_{i=1}^{6} d_i \leq 3d$



$$3 \times d \geq \frac{1}{6}\left[\, d_1 + d_2 + d_3 + d_4 + d_5 + d_6 \,\right]$$

Theorem 21 (Fakcharoenphol, Rao, Talwar, 2003) Let (X,d) be a finite metric space. Then there exists a set $\mathcal{D} = \{d_1, \dots, d_k\}$ of tree metrics on X and a probability distribution $\mathcal{P}$ over $\mathcal{D}$ such that $(\mathcal{D}, \mathcal{P})$ $O(\log n)$-probabilistically approximates d, where $n = |X|$. If d is a metric of a weighted graph G, then the tree metrics can be chosen to be spanning trees of G.

Remark: Theorem 21 improves on previous result of Bartal (1998).

Theorem 15 (Bourgain, 1985) $(V_n, d) \xhookrightarrow{O(\log n)} \ell_1^{O(\log^2 n)}$

Proof: First consider the following one-dimensional embedding of (V_n, d): pick $S \subseteq V_n$ and $\forall v \in V_n$ set

$$\sigma(v) = \min\{d(v,s) : s \in S\}.$$

Lemma 1: $|\sigma(u) - \sigma(v)| \leq d(u,v) \quad \forall u,v \in V_n$

Proof: Let s_1 and s_2 be the closest vertices of S to u and v, resp. Assume w.l.o.g. that $d(s_1,u) \leq d(s_2,v)$. Then

$$|\sigma(u) - \sigma(v)| = d(s_2,v) - d(s_1,u) \leq d(s_1,v) - d(s_1,u) \leq d(u,v),$$

the last inequality follows from triangle inequality $\square$

Now, pick ℓ subset of V_n, $S_1, \dots, S_\ell$, and define the ith coordinate of $v \in V_n$ to be $\sigma_i(v) = \min_{s \in S_i} d(s,v)/\ell$. Let $\ell = \log_2 n + 1$; for each $2 \leq i \leq \ell$, set S_i is formed by picking each vertex of V_n with probability $1/2^i$.

From Lemma 1 we conclude that

$$\|\sigma(u) - \sigma(v)\|_1 = \sum_{i=1}^{\ell} |\sigma_i(u) - \sigma_i(v)| \leq d(u,v).$$

Now we will ensure that a single distance $d(u,v)$ is not overshrunk. For this, we consider the expected contribution of set S_i $E[|\sigma_i(u) - \sigma_i(v)|]$ to the ℓ_1-distance between u and v.

Let $B(x,r) = \{v \in V : d(x,v) \leq r\}$ denote the ball of radius r around x.

Lemma 2: If for some choice of $r_1 \geq r_2 \geq 0$ and constant c,

$$\Pr[(S_i \cap B(u, r_1) = \emptyset) \text{ and } (S_i \cap B(v, r_2) \neq \emptyset)] \geq c,$$

then the expected contribution of S_i is $\geq c(r_1 - r_2)/\ell$.

Proof: Under the event described, $d(u, S_i) \geq r_1$ and $d(v, S_i) \leq r_2$. Then $\sigma_i(u) \geq r_1/\ell$ and $\sigma_i(v) \leq r_2/\ell$. Therefore,

$$|\sigma_i(u) - \sigma_i(v)| \geq (r_1 - r_2)/\ell,$$

thus the expected contribution of S_i is $\geq c(r_1 - r_2)/\ell$. □

For each set S_i we will define r_1 and r_2 such that the statement of Lemma 2 holds.

Lemma 3: For $1 \leq t \leq \ell - 1$, let A and B be disjoint subsets of V_n such that $|A| < 2^t$ and $|B| \geq 2^{t-1}$. Form set S by picking each vertex of V_n independently with probability $p = 1/2^{t+1}$. Then,

$$\Pr[(S \cap A = \emptyset) \text{ and } (S \cap B \neq \emptyset)] \geq \frac{1}{2}(1 - e^{-\frac{1}{4}})$$

Proof: $\Pr[S \cap A = \emptyset] = (1-p)^{|A|} \geq (1 - p|A|) \geq \frac{1}{2}$

$\Pr[S \cap B = \emptyset] = (1-p)^{|B|} \leq e^{-p|B|} \leq e^{-\frac{1}{4}}$ (we used $1 - x \leq e^{-x}$)

whence

$$\Pr[S \cap B \neq \emptyset] = 1 - (1-p)^{|B|} \geq 1 - e^{-\frac{1}{4}}.$$

Since $A \cap B = \emptyset$, the events $[S \cap A = \emptyset]$ and $[S \cap B \neq \emptyset]$ are independent, thus the desired probability is $\geq \frac{1}{2}(1 - e^{-\frac{1}{4}})$. □

Set $c = \frac{1}{2}(1 - e^{-\frac{1}{4}})$.

For $0 \le t \le \ell - 1 = [\log_2 n]$, define

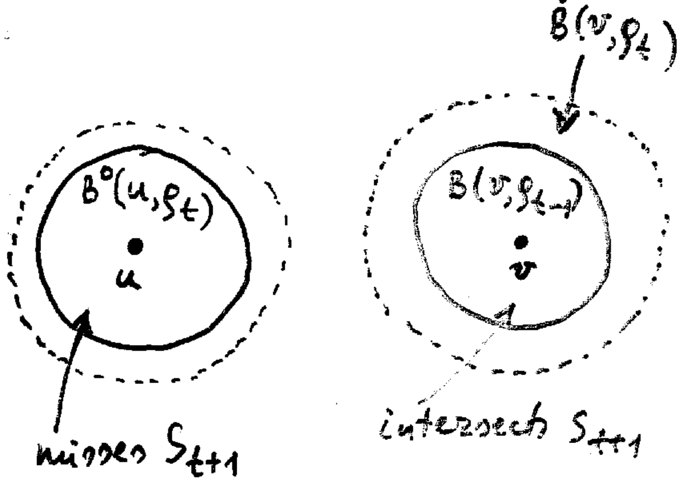
$$\rho_t = \min\{\rho \ge 0 : |B(u,\rho)| \ge 2^t \text{ and } |B(v,\rho)| \ge 2^t\}.$$

Let $\hat{t} = \max\{t : \rho_t < d(u,v)/2\}$; clearly $\hat{t} \le \ell - 2$.

Let $B^\circ(x,r) = \{s \in V : d(x,s) < r\}$ – the open ball.

Lemma 4: For $1 \le t \le \hat{t}$, the expected contribution of S_{t+1} is at ~~most~~ least $c \cdot \frac{\rho_t - \rho_{t-1}}{\ell}$. For $t = \hat{t}+1$, the expected contribution of S_{t+1} is at most $\frac{c}{\ell}\left(\frac{d(u,v)}{2} - \rho_{t-1}\right)$.

Proof: We will prove only the first assertion, i.e. $1 \le t \le \hat{t}$. By definition of ρ_t, at least one of the open balls $B^\circ(u,\rho_t)$, $B^\circ(v,\rho_t)$ contains fewer than 2^t vertices. Assume w.l.o.g. $|B^\circ(u,\rho_t)| < 2^t$. By definition, $|B(v,\rho_{t-1})| \ge 2^{t-1}$. Since $\rho_{t-1} < \rho_t < d(u,v)/2$, the two sets $B^\circ(u,\rho_t)$ and $B(v,\rho_{t-1})$ are disjoint. By Lemma 3, the probability that S_{t+1} is disjoint from $B^\circ(u,\rho_t)$ and intersects $B(v,\rho_{t-1})$ is at least c. Since $B^\circ(u,\rho_t)$ is a ball centered at u and radius $<\rho_t$ the assertion follows from Lemma 2.1



Lemma 5: The expected contribution of all sets $S_2, \dots, S_\ell$ is at ~~most~~ least $\frac{c}{2\ell} d(u,v)$.

Proof: By Lemma 4, the expected contribution of all sets $S_2, \dots S_\ell$ is at least the following telescoping sum:

$$\frac{c}{\ell}\left[(\rho_1 - \rho_0) + (\rho_2 - \rho_1) + \dots + \left(\frac{d(u,v)}{2} - \rho_{\hat{t}}\right)\right] = \frac{c}{2\ell} d(u,v). \quad \square$$

<u>Lemma 6</u>: $\Pr\left[\text{contribution of all sets is} \geq \frac{c\,d(u,v)}{2\ell}\right] \geq \frac{c/2}{1-c/2}$

<u>Proof</u>: follows from Lemma 5.

<u>Chernoff bound</u>: let $X_1, \ldots, X_N$ be independent Bernoulli trials with $\Pr[X_i = 1] = p$ and let $X = \sum_{i=1}^{N} X_i$ ($E[X] = Np$). Then for $0 < \varepsilon \leq 1$,

$$\Pr[X < (1-\varepsilon)Np] < e^{-\frac{\varepsilon^2 Np}{2}}.$$

Pick sets $S_2, \ldots, S_\ell$ using probabilities specified above, independently $N = O(\log n)$ times each. Call the sets so obtained S_{ij}, $1 \leq i \leq \ell$, $1 \leq j \leq N$. Consider the $\ell \cdot N = O(\log^2 n)$ dimensional embedding of (V_n, d) with respect to these $\ell \cdot N$ sets.

<u>Lemma 7</u>: $\Pr\left[\|\sigma(u) - \sigma(v)\|_1 \geq \frac{p\,c\,d(u,v)}{4\ell}\right] \geq 1 - \frac{1}{2n^2}$ $\left(p = \frac{c}{2-c}\right)$

<u>Proof</u>: Think of picking sets $S_2, \ldots, S_\ell$ once as a single Bernoulli trial (thus we have N such trials). A trial succeed if the contribution of all sets is $\geq \frac{c\,d(u,v)}{2\ell}$; the probability of success is $\geq p = \frac{c}{2-c}$ by Lemma 6.

Using Chernoff bound with $\varepsilon = \frac{1}{2}$, the probability that at most $Np/2$ of these trials succeed is $\leq e^{Np/8} \leq \frac{1}{2n^2}$ for $N = O(\log n)$. If at least $Np/2$ trials succeeds the ℓ_1-distance between $\sigma(u)$ and $\sigma(v)$ will be $\geq \frac{p\,c\,d(u,v)}{4\ell} = \frac{d(u,v)}{O(\log n)}$.

Adding the error prob. for all $\frac{n(n-1)}{2}$ pairs:

<u>Theorem</u>: With probability $\geq \frac{1}{2}$ this $O(\log^2 n)$ dimensional embedding has distortion $O(\log n)$

#3 Graph classes defined by distance properties

Purpose: introduce and characterize main graph classes interesting from the metric point of view and related to l_1, l_2, l_∞-metrics and Hamming distance

Classes of graphs:

(i) median graphs;

(ii) Helly graphs;

(iii) bridged graphs;

(iv) weakly median graphs;

(v) isometric subgraphs of hypercubes and Hamming graphs

(vi) l_1-graphs;

(vii) superconnected (lopsided) set systems and graphs;

(viii) basis graphs of matroids and Δ-matroids.

Main generalizations of l_2, l_∞, and l_1:

(1) $l_2 \longrightarrow$ CAT(0) metric spaces;

(2) $l_\infty \longrightarrow$ hyperconvex (injective) metric spaces;

(3) $l_1 \longrightarrow$ median metric spaces.

Bridged graphs can be viewed as discrete analogous of CAT(0) spaces

Helly graphs are discrete analogous of hyperconvex

Graph classes defined by distance properties

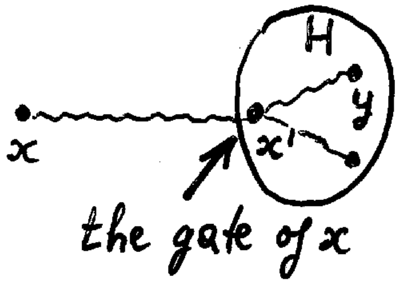
$G=(V,E)$ - connected not necessarily finite, undirected and unweighted graph endowed with the standard graph distance $d(u,v):=d_G(u,v)$

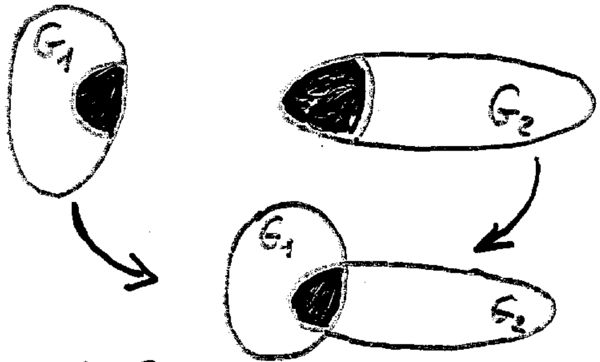
Interval $I(u,v)=\{x\in V:\ d(u,v)=d(u,x)+d(x,v)\}$

Convex set $S\subseteq V$: $I(u,v)\subseteq S\quad \forall\, u,v\in S$

Halfspace $H\subseteq V$: convex set with a convex complement $V-H$

Gated subgraph (set) H: $\forall\, x\notin H\ \exists\, x'\in H: x'\in I(x,y)\ \forall y\in H$



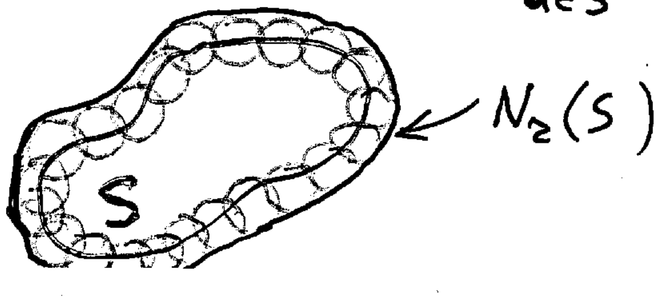


Gated amalgam G of G_1 and G_2

Ball (or r-neighborhood) $B(u,r)=N_r(u)=\{x\in V: d(u,x)\le r\}$

r-neighborhood of a set S:

$$N_r(S)=\{x\in V: d(x,S)\le r\}=\bigcup_{u\in S} N_r(u)$$



Definitions (cont.)

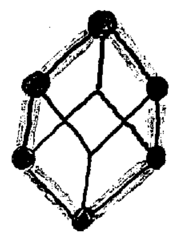
isometric subgraph: an induced subgraph $H=(Y,F)$ of a graph $G=(X,E)$ such that $d_H(u,v)=d_G(u,v)\ \forall u,v\in Y$

isometric embedding $\varphi: H \to G$: $\forall u,v\in Y,\ d_G(\varphi(u),\varphi(v)) = d_H(u,v)$

scale k embedding $\varphi: H\to G$: $\forall u,v\in Y,\ d_G(\varphi(u),\varphi(v)) = k\, d_H(u,v)$

retract: a subgraph $H=(Y,F)$ of $G=(X,E)$ such that there exists an idempotent nonexpansive mapping ψ from G to H, i.e $\psi(y)=y\ \forall y\in Y$ and $d_G(\psi(x),\psi(y))\le d_G(x,y)\ \forall x,y\in X$

Remark: retracts are isometric subgraphs of the host graph, but not the converse:
the 6-cycle C_6 is an isometric subgraph but not a retract of the 3-cube Q_3



Remark: in previous definitions, the host graph $G=(X,E)$ can be replaced by an arbitrary metric space (X,d)

Main host spaces for isometric embedding of graphs

geometric: ℓ_1- and ℓ_∞-spaces

graphic: hypercubes, Hamming graphs, half-cubes, Johnson graphs, ℓ_1- and ℓ_∞-grids

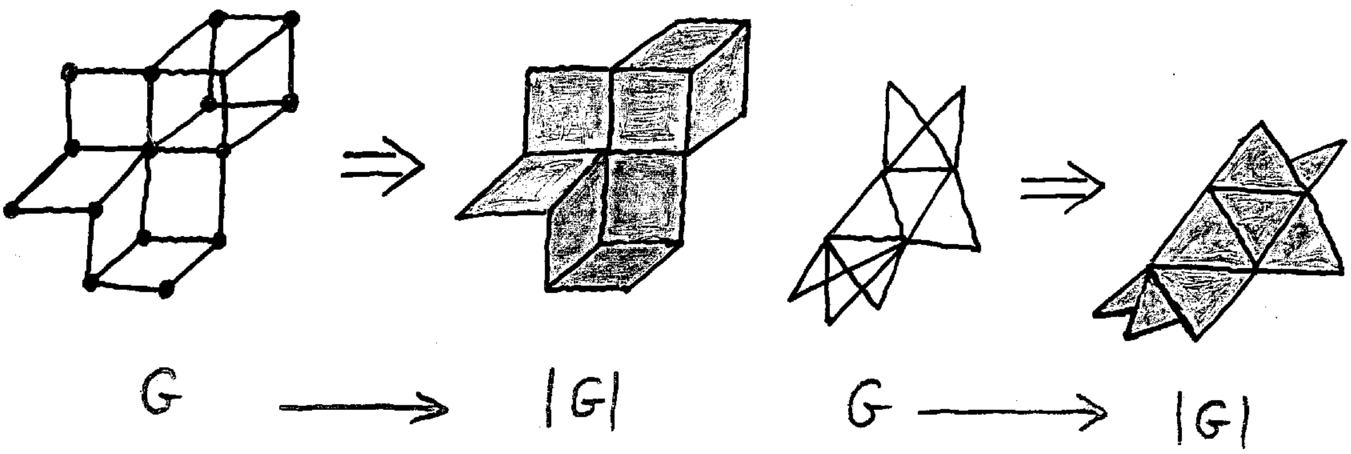
Definitions (cont.)

How to derive cell complexes from graphs?

(a) cubical complexes: replace every graphic cube by a unit solid cube;

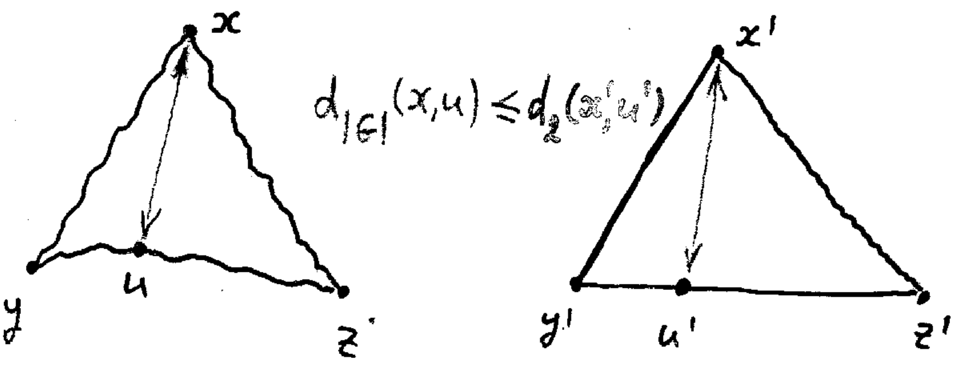
(b) simplicial complexes: replace every clique (complete subgraph) by a simplex;

(c) cell complexes from planar graphs: replace every interior face by a regular polygon with unit side.



Remark: $|G|$ can be endowed with an intrinsic ℓ_1-, ℓ_2-, or ℓ_∞-metric.

CAT(0) complexes: geodesic triangles in $|G|$ are thinner than the comparison euclidean triangle



CAT(0) complexes have many nice properties, some of which characterize them:

(i) any two points can be joined by a unique geodesic (shortest path);

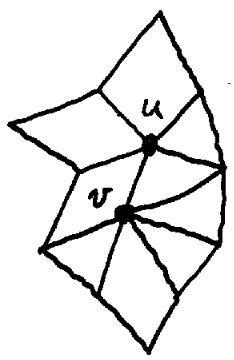
(ii) r-neighborhoods of convex sets are convex;

(iii) do not contain isometrically embedded cycles;

(iv) if α and β are geodesics in $|G|$, then the function $f: [0,1] \to |G|$ given by $f(t) = d(\alpha(t), \beta(t))$ is convex;

(v) global nonpositive curvature.

For our case (c), the condition (v) can be read as: the sum of angles around any interior vertex is at least 2π.



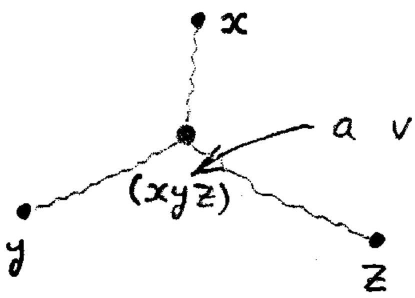
$$\Sigma(u) = 3 \times \frac{\pi}{2} + 2 \times \frac{\pi}{3} > 2\pi$$

$$\Sigma(v) = 2 \times \frac{\pi}{2} + 4 \times \frac{\pi}{3} > 2\pi$$

For more details on CAT(0) spaces see the book by Bridson and Haefliger.

Median graphs

median

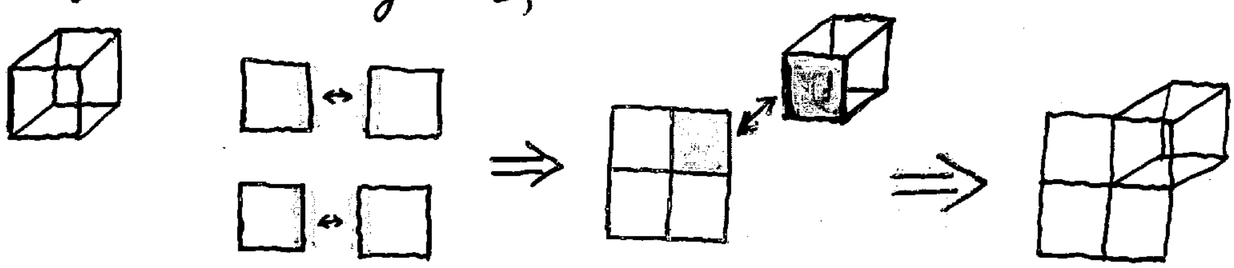


a vertex in $I(x,y) \cap I(y,z) \cap I(z,x)$

median graphs: graphs in which every triplet x, y, z has a unique median denoted (xyz)

Characterizations of median graphs:

(i) (Isbell) Median graphs are precisely the graphs which are obtained from cubes via successive gated amalgams;



resulting median graph

(ii) (Bandelt) Median graphs are precisely the retracts of hypercubes;

(iii) (Schaefer) Median graphs are precisely the connected components of solutions of 2SAT instances;

Median graphs (cont.)

(iv) (Avann) The median operator of a median graph satisfies the following equations:

(1) $(aab) = a$ (majority)

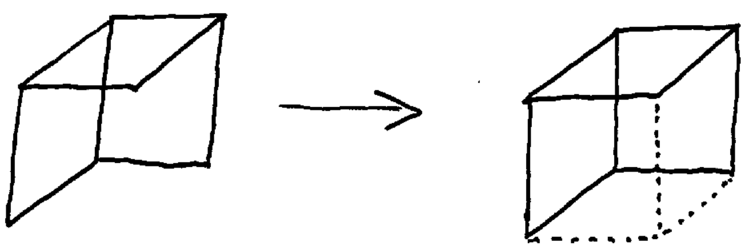
(2) $(\sigma(a)\sigma(b)\sigma(c)) = (abc)$ $\forall$ permutation σ (symmetry)

(3) $((abc)dc) = (a(bcd)c)$ (associativity)

Conversely, every ternary algebra satisfying (1), (2), and (3) comes from a median graph;

(v) ($\overset{\text{Roller}}{\overline{\text{Ch.}}}$) G is a median graph if and only if the cubical complex $|G|$ is CAT(0);

(vi) (Gromov) a cubical complex $|G|$ is CAT(0) if and only if $|G|$ is simply connected and satisfies the following combinatorial condition: if three $(k+2)$-cubes intersect in a k-cube and pairwise intersect in $(k+1)$-cubes, then they are contained in a $(k+3)$-cube;



Other properties:

(vii) (van de Vel) $(|G|, \ell_1)$ is an ℓ_1-subspace;

(viii) (Mai & Tang) $(|G|, \ell_\infty)$ is an absolute retract, i.e. a retract of every space in which it embeds isometrically

Bridged graphs

bridged graph: a graph in which every isometric cycle has length 3

Characterizations of bridged graphs:

(i) (Ch. & Soltan, Farber & Jamison) Bridged graphs are precisely the graphs in which the neighborhoods $N_r(S)$ of convex sets S are convex;

(ii) (Chepoi) G is bridged if and only if the simplicial complex $|G|$ is simply connected and for every vertex v, $N_1(v)$ does not contain induced 4-cycles and 5-cycles;

(iii) (Anstee & Farber) Bridged graphs are precisely the dismantlable graphs without induced 4- and 5-cycles.
(Chepoi) The dismantling scheme is provided by BFS.

Dismantling scheme: ordering $v_1, \ldots, v_n$ of vertices of G such that $\forall v_i \; \exists v_j \in N_1(v_i), j > i$ such that all neighbors $v_k, k > i$, of v_i are also neighbors of v_j.

Examples: (a) Chordal graphs

(b) graphs for which the simplicial complex $|G|$ is 2-dimensional and CAT(0);

(c) planar graphs in which all inner faces are triangles and all inner vertices have degree ≥ 6

Hyperconvex spaces and Helly graphs

Hyperconvex space: a geodesic (Menger-convex) metric space in which every family of pairwise intersecting balls has a point in common (Helly property).

Helly graphs: the graphs in which the balls have the Helly property.

Remark: Helly graphs are the discrete analogies of hyperconvex spaces.

Theorem (Aronszajn, Panitchpakdi, 1959)

(i) Hyperconvex spaces are exactly the absolute retracts in the category of metric spaces, or equivalently, they are the retracts of l_∞-spaces;

(ii) Helly graphs are exactly the absolute retracts in the category of (reflexive) graphs.

Examples of Helly graphs: l_∞-grid, and, more generally take a median graph G and replace every maximal cube by a clique; the resulting graph G^∇ is Helly.

Theorem (Isbell, 1964; Dress, 1984) For every finite metric space (V_n, d) there exists the smallest hyperconvex space containing (V_n, d) as an isometric subspace (the tight span or the injective hull of d);

The same holds for graphs.

Weakly median graphs

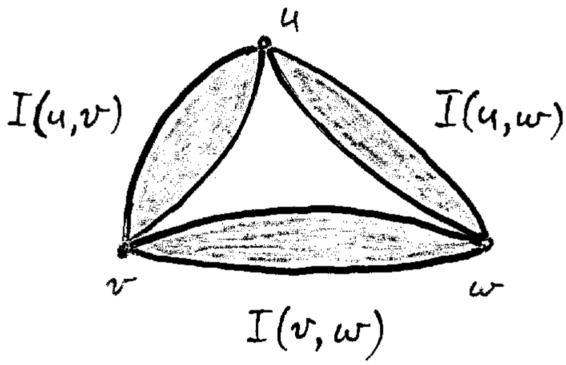
<u>Question:</u> How to extend notions like "median" and "median graph"?

metric triangle uvw:

$I(u,v) \cap I(v,w) = \{v\}$

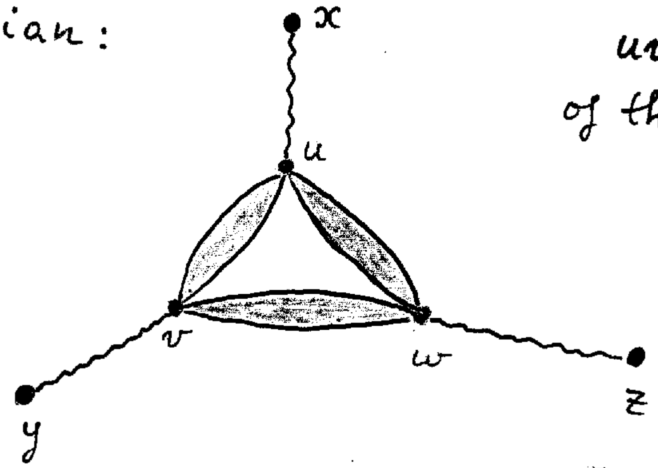
$I(v,w) \cap I(w,u) = \{w\}$

$I(w,u) \cap I(u,v) = \{u\}$



strongly equilateral metric triangle uvw: $d(u,x) \equiv \text{const} \quad \forall x \in I(v,w)$

quasi-median:



uvw is a quasi-median of the triplet x, y, z

Remark: every triplet of vertices admits at least one quasi-median

apex: u is called an apex of x with respect to y, z and is denoted by (xyz)

Analogously are defined the apices (yxz) and (zxy)

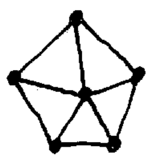
Weakly median graphs (cont.)

weakly modular graphs: graphs in which all metric triangles are strongly equilateral;

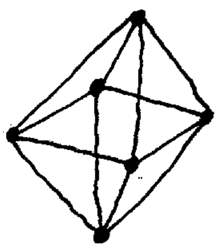
weakly median graphs: weakly modular graphs in which every triplet of vertices has a unique quasi-median

Characterisation:

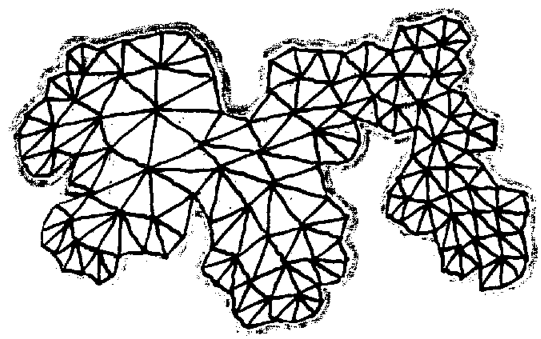
(i) (Bandelt & Ch.) Finite weakly median graphs are precisely the graphs obtained by successive applications of gated amalgamations from Cartesian products of the following prime graphs: 5-wheels, subhyperoctahedra, and two-connected plane graphs such that all inner faces are triangles and all inner vertices have degrees ≥ 6.



5-wheel



3-octahedron



bridged triangulation

(ii) (Bandelt & Ch.) Every finite weakly median graph is a retract of a Cartesian product of prime weakly median graphs and vice versa.

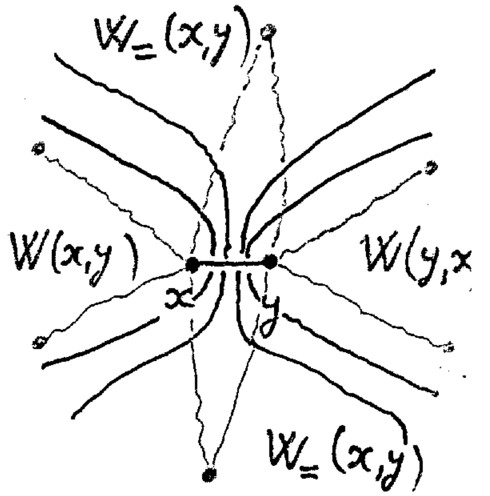
(iii) (Bandelt & Ch.) Every weakly median graph G is l_1-embeddable. G has a scale 2 embedding in a hypercube iff it does not contain an induced K_6 minus an edge.

(iv) (Bandelt & Ch.) Apex algebras of weakly median graphs are characterized by a set of 5 axioms among discrete ternary algebras

Isometric subgraphs of hypercubes and Hamming graphs

For an edge xy of a graph G set:

$$W(x,y) = \{z : d(x,z) < d(y,z)\}$$
$$W(y,x) = \{z : d(y,z) < d(x,z)\}$$
$$W_=(x,y) = \{z : d(x,z) = d(y,z)\}$$



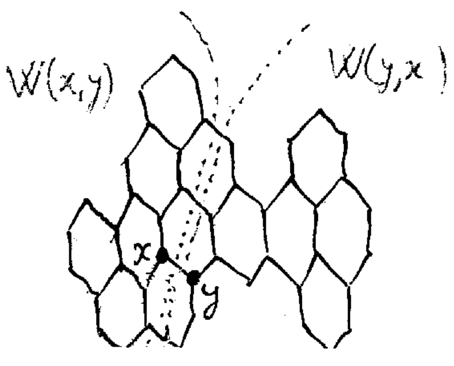
Djokovic: G is isometrically embeddable into a hypercube iff it is bipartite and for every edge xy the sets $W(x,y)$ and $W(y,x)$ are convex (i.e., they are complementary halfspaces).

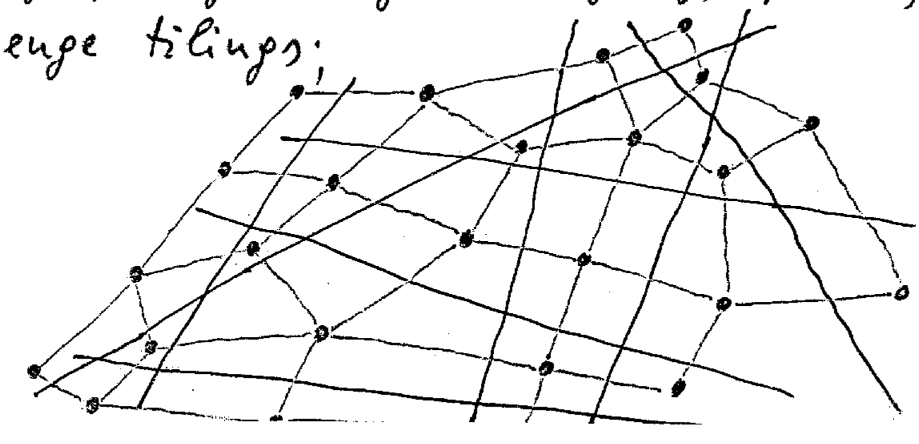
Ch. (answering a question by Winkler): G is isometrically embeddable into a Hamming graph (Cartesian product of complete graphs) iff for every edge xy the sets $W(x,y)$, $W(y,x)$, $W(x,y) \cup W_=(x,y)$, and $W(y,x) \cup W_=(x,y)$ are convex.

Examples: (i) benzenoids: planar graphs in which all inner faces are hexagons and all inner vertices have degree 3;

(ii) tope graphs of arrangements of hyperplanes;

(iii) lozenge tilings;





l_1-graphs

<u>Remark</u>: G is an l_1-graph iff it admits a scale embedding into a hypercube.

Ch., Deza, Grishukhin: A planar graph G is an l_1-graph iff it admits a scale 2 embedding into a hypercube (i.e., an isometric embedding into a halfcube).

Shpectorov: G is an l_1-graph iff it admits an isometric embedding into a Cartesian product of halfcubes and octahedra.

<u>Remark</u>: Shpectorov's result yields a polynomial recognition of l_1-graphs (in contrast to l_1-metrics)

<u>Question</u>: Provide a Djokovic-like characterization of isometric subgraphs of halfcubes.

<u>Some classes of planar l_1-graphs</u> (Chepoi, Dragan, Vaxès)

(4,4)-graphs, i.e. plane graphs in which all inner faces have length ≥ 4 and all inner vertices have degree ≥ 4;

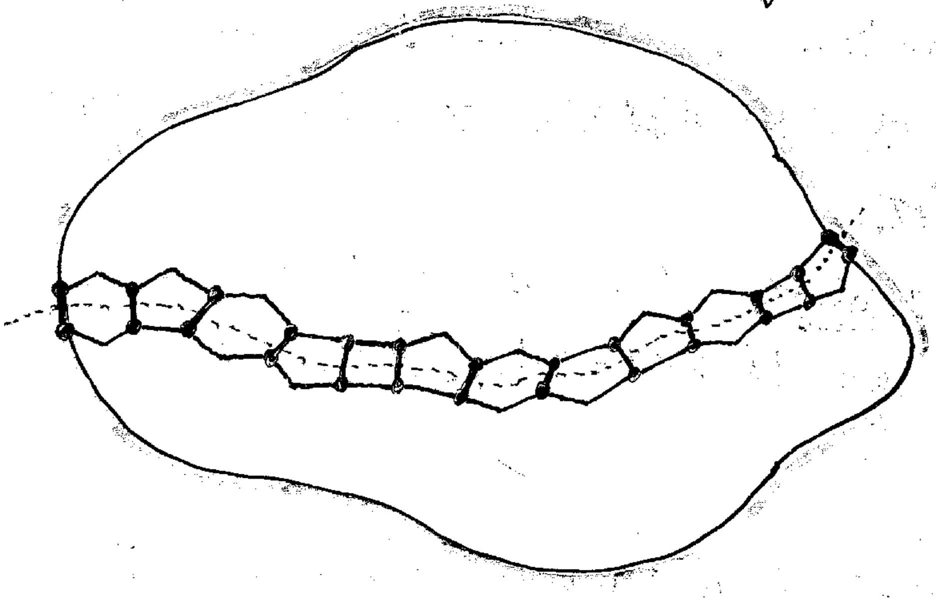
(6,3)-graphs, i.e. plane graphs in which all inner faces have length ≥ 6 and all inner vertices have degree ≥ 3;

(3,6)-graphs, i.e. plane graphs in which all inner faces have length ≥ 3 and all inner vertices have degree ≥ 6.

<u>Remark</u>: For every planar graph G of type $(4,4)$, $(3,6)$, or $(6,3)$, the cell complex $|G|$ is CAT(0).

<u>Remark</u>: It turn out that the planar graphs of types $(4,4)$, $(3,6)$, and $(6,3)$ have been investigated in combinatorial group theory, in particular by R. Lyndon who established the following maximality principle: if S is a subgraph of G bounded by a simple cycle ∂S and v is a vertex of S, then all furthest from v vertices of S are located on ∂S.

<u>Idea of the l_1-embedding</u>: use the alternating cuts of G



(i) the union of faces cut by an alternating cut is a strip consisting by the edges of the cut and two paths whose lengths differ by at most 1;

(ii) any alternating cut split the vertices of G into two convex sets

(iii) via every edge of G pass two alternating cuts.

Superconnected subsets of hypercubes

(following Bandelt, Ch., Dress & Koolen)

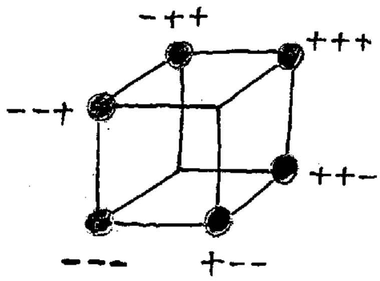
$\mathfrak{S} \subseteq \{\pm 1\}^I$: set of maps from a finite set I to $\{\pm 1\}$;

For $Y \subseteq I$, let

$$\mathfrak{S}_Y := \{t \in \{\pm 1\}^{I-Y} : \text{some extension } s \in \{\pm 1\}^I \text{ of } t \text{ belongs to } \mathfrak{S}\}$$

$$\mathfrak{S}^Y := \{t \in \{\pm 1\}^{I-Y} : \text{every extension } s \in \{\pm 1\}^I \text{ of } t \text{ belongs to } \mathfrak{S}\}$$

<u>Example</u>: $I = \{1,2,3\}$



$\mathfrak{S}_{\{3\}} = \{\{--\}, \{-+\}, \{+-\}, \{++\}\}$

$\mathfrak{S}^{\{3\}} = \{\{--\}, \{++\}\}$

$\mathfrak{S}_{\{1,2\}} = \{\{-\}, \{+\}\}$, $\mathfrak{S}^{\{1,2\}} = \emptyset$

Two ways to derive an abstract simplicial complex from $\mathfrak{S}$:

$$\overline{\mathcal{K}}(\mathfrak{S}) := \{Y \subseteq I : \mathfrak{S}_{I-Y} = \{\pm 1\}^Y\}$$

$$\underline{\mathcal{K}}(\mathfrak{S}) := \{Y \subseteq I : \mathfrak{S}^Y \neq \emptyset\}$$

In previous example,

$$\overline{\mathcal{K}}(\mathfrak{S}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\}$$

$$\underline{\mathcal{K}}(\mathfrak{S}) = \{\emptyset, \{1\}, \{2\}, \{3\}\}$$

Superconnected subsets of hypercubes (cont.)

$\forall\, \mathcal{S} \subseteq \{\pm 1\}^I$, $\quad \# \underline{\mathcal{X}}(\mathcal{S}) \le \# \mathcal{S} \le \# \overline{\mathcal{X}}(\mathcal{S})$ holds

Bandelt, Ch, Dress, Koolen: For a set $\mathcal{S} \subseteq \{\pm 1\}^I$ of sign maps the following conditions are equivalent:

(i) superconnectivity: $\mathcal{S}^Y$ is connected for all $Y \subseteq I$;

(ii) superisometry: $\mathcal{S}^Y$ is isometric for all $Y \subseteq I$;

(iii) commutativity: $(\mathcal{S}^Y)_Z = (\mathcal{S}_Z)^Y$ for all disjoint subsets Y, Z of I;

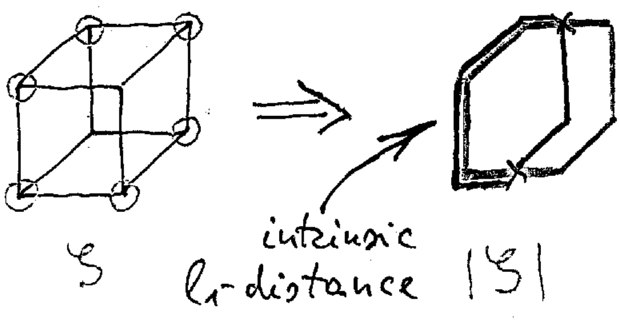
(iv) ampleness I: $\# \mathcal{S} = \# \overline{\mathcal{X}}(\mathcal{S})$;

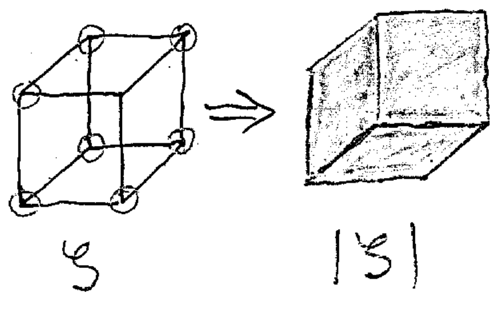
(v) ampleness II: $\underline{\mathcal{X}}(\mathcal{S}) = \overline{\mathcal{X}}(\mathcal{S})$;

(vi) $\mathcal{S}$ is isometric and both $\mathcal{S}^e$ and $\mathcal{S}_e$ are super-connected for some $e \in I$;

(vii) $\mathcal{S}$ is connected, and $\mathcal{S}^e$ is superconnected for every $e \in I$;

(viii) l_1-isometry: the cubical complex $|\mathcal{S}|$ endowed with the intrinsic l_1-metric is an isometric subspace of $(\mathbb{R}^I, \|\cdot\|_1)$.





Superconnected subsets of hypercubes (cont.)

Examples: (i) vertex-sets of median graphs;

(ii) signed maps of convex sets of antimatroids (convex geometries);

(iii) maximum set systems of a given Vapnik-Chervonenskis dimension;

(iv) lozenge tilings;

(v) signed maps of regions of simple affine arrangements of hyperplanes;

(vi) (J. Lawrence) signed maps of orthants intersecting a given convex (in the usual sense) set.

Remark: Superconnected sets are equivalent to lopsided sets introduced and characterized in a different way by J. Lawrence.

Open question: Is it true that for every proper nonempty superconnected subset $\mathcal{S} \subset \{\pm 1\}^I$ there exist $s \in \mathcal{S}$ and $t \in \{\pm 1\}^I \setminus \mathcal{S}$ such that $\mathcal{S} \cup \{t\}$ and $\mathcal{S} \setminus \{s\}$ are superconnected?

Basis graphs of matroids and even Δ-matroids

matroid: a collection $\mathcal{B}$ of subsets of a finite set I, called bases, which satisfy the following exchange property;

(EP) for all $A, B \in \mathcal{B}$ and $i \in A \setminus B$ there exists $j \in B \setminus A$ such that $A \setminus \{i\} \cup \{j\} \in \mathcal{B}$.

The base $A \setminus \{i\} \cup \{j\}$ is obtained from A by an elementary exchange;

basis graph $G = G(\mathcal{B})$ of a matroid $\mathcal{B}$ is the graph whose vertices are the bases of $\mathcal{B}$ and edges are the pairs A, B of bases differing by an elementary exchange

<u>Remark</u>: Basis graphs faithfully represent their matroids

<u>Remark</u>: Since all bases of a matroid $\mathcal{B}$ have the same cardinality, (EP) implies that $G(\mathcal{B})$ is an isometric subgraph of a Johnson graph (one slice of a halfcube).

<u>Remark</u>: A characterization of basis graphs of matroids emloying distance properties was provided by S. Maurer. We simplified and generalized this result to basis graphs of even Δ-matroids

Basis graphs of matroids and even Δ-matroids (cont.)

Δ-matroid (Bouchet; Chandrasekaran & Kabadi; Dress & Havel)
a collection $\mathcal{B}$ of subset of a finite set I, called bases, not necessarily equicardinal, satisfying the following symmetric exchange property:

(SEP) for all $A, B \in \mathcal{B}$ and $i \in A \Delta B$, there exists $j \in B \Delta A$ such that $A \Delta \{i,j\} \in \mathcal{B}$.

The base $A \Delta \{i,j\}$ is obtained from A by an elementary exchange;

even Δ-matroid: all bases have the same cardinality modulo 2

basis graph $G = G(\mathcal{B})$ of an even Δ-matroid $\mathcal{B}$: the graph whose vertices are the bases of $\mathcal{B}$ and edges are the pairs A, B of bases differing by a single exchange, i.e. $|A \Delta B| = 2$.

Axiom (SEP) implies that $G(\mathcal{B})$ is an isometric subgraph the halfcube, i.e. $|A \Delta B| = 2 d_{G(\mathcal{B})}(A,B) \ \forall A, B \in \mathcal{B}$

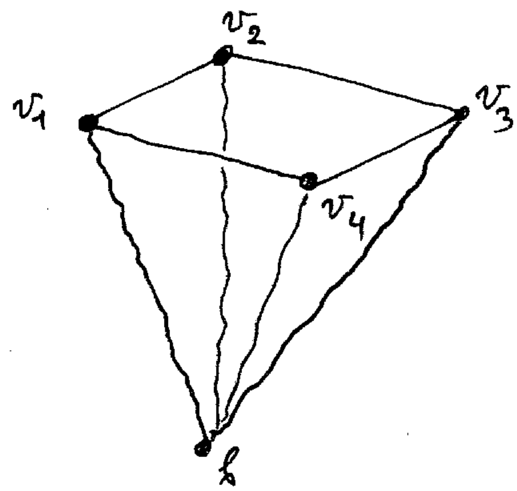
interval condition

(ICm) if $d(u,v) = 2$, then $I(u,v)$ contain an induced 4-cycle and itself is an induced subgraph of the m-octahedron

positioning condition

(PC) for each vertex b and each induced 4-cycle $v_1 v_2 v_3 v_4$
$$d(b, v_1) + d(b, v_3) = d(b, v_2) + d(b, v_4)$$

Basis graphs of matroids and even Δ-matroids (cont.)



$$d(v_1,b)+d(v_3,b)=d(v_2,b)+d(v_4,b)$$

Ch.: G is a basis graph of an even Δ-matroid iff it satisfies the positioning condition (PC), the interval condition (IC4), and the neighborhood $N(b)$ of some vertex is the line graph of some graph Γ.

Maurer: G is a basis graph of a matroid iff it satisfies the positioning condition (PC), the interval condition (IC3) and the neighborhood $N(b)$ of some vertex is the line graph of some bipartite graph $\Gamma=(A\dot\cup B,F)$.

Idea of the proof: define a mapping $\varphi: V\to 2^I$ in the following way:

(a) $\varphi(b)=\emptyset$;

(b) $\forall x\in N(b)$ encodes some edge ij of Γ; put $\varphi(x)=\{i,j\}$;

(c) $\forall v\notin N(b)\cup\{b\}$, let $\varphi(v)=\bigcup\{\varphi(x): x\in I(v,b)\cap N(b)\}$.

Properties of φ: φ is injective; all sets $\varphi(v)$ have even cardinality; φ is an isometric embedding of G into a halfcube. This implies that $\mathcal{B}_\varphi=\{\varphi(v): v\in V\}$ is an even Δ-matroid. If Γ is bipartite with $I=A\dot\cup B$, then $\mathcal{B}_\varphi\Delta A=\{\varphi(v)\Delta A: v\in V\}$ is a matroid of rank $|A|$.