

Embeddings of Graphs in Euclidean Spaces

J. Reiterman, V. Rödl, and E. Šiňajová

Department of Mathematics FJFI, Technical University of Prague, Husova 5,
 110 00 Prague 1, Czechoslovakia

Abstract. The dimension of a graph $G = (V, E)$ is the minimum number d such that there exists a representation $x \rightarrow \bar{x} \in R^d$ ($x \in V$) and a threshold t such that $xy \in E$ iff $\bar{x}\bar{y} \geq t$. We prove that $d(G) \leq n - \chi(G)$ and $d(G) \leq n - \sqrt{n}$ where $n = |V|$ and $\chi(G)$ is the chromatic number of G ; we present upper bounds for the dimension of graphs with a large girth and we show that the complement of a forest can be represented by unit vectors in R^d . We prove that $d(G) \geq \frac{1}{15}n$ for most graphs and that there are 3-regular graphs with $d(G) \geq c \log n / \log \log n$.

Introduction

We consider the geometric dimension of graphs introduced in [7]. Let $G = (V, E)$ be a finite graph without loops and multiple edges. The dimension $d(G)$ of G is the minimum number d such that G admits a representation in R^d ; a representation consists of an assignment $x \rightarrow \bar{x} \in R^d$ ($x \in V$) and a threshold $t \in R$ such that for every couple $x, y \in V$, $x \neq y$,

$$xy \in E \quad \text{iff} \quad \bar{x}\bar{y} \geq t,$$

where $\bar{x}\bar{y}$ is the scalar product of vectors \bar{x}, \bar{y} .

In [7] we also considered a related notion—the spherical dimension $sd(G)$ of a graph G , which is defined analogously but with an additional condition $\|\bar{x}\| = 1$ for all $x \in V$. Clearly, $sd(G) \geq d(G)$ for every graph G .

The notion of spherical dimension is closely related to that of sphericity which was introduced in [4]. The sphericity is defined as the minimum number d such that there exists an assignment $x \rightarrow \bar{x} \in R^d$ and a threshold $\rho \geq 0$ such that

$$xy \in E \quad \text{iff} \quad \|\bar{x} - \bar{y}\| \leq \rho.$$

In [7] it is noted that

$$\text{sd}(G) - 1 \leq \text{sph}(G) \leq \text{sd}(G)$$

for every graph G .

Let us list some bounds for $d(G)$, $\text{sd}(G)$, and $\text{sph}(G)$:

- (1) $d(G) \leq n - \alpha(G)$ ($n = |V|$, $\alpha(G)$ is the size of a maximal independent set in G).
- (2) $\text{sph}(G) \leq n - \omega(G)$ (where $\omega(G)$ is the size of a maximal clique in G).
- (3) $\text{sd}(G) \leq cd \log(d)$ if the complement \bar{G} of G has maximum degree $\leq d$.
- (4) $d(T) \leq 3$ if T is a tree.
- (5) $\text{sph}(T) \leq \frac{1}{2}(k+1) \log_k |T|$ if T is a tree with degree $\leq k$.
- (6) $\text{sph}(T) \leq 108 \log |T|$ if T is a tree.
- (7) $\text{sph}(F) \leq 8 \lceil \log |F| \rceil$ if F is a forest.
- (8) $c_1 n / \log n \leq \text{sph}(Q_n) \leq c_2 n / \log n$ where Q_n is the graph of the n -dimensional cube.
- (9) $\text{sd}(G) \leq cd^2 \log n$ if G has maximum degree $\leq d$.
- (10) $\text{sph}(G_1 + G_2 + \dots + G_m) \leq 2(n-1)$ where $n = \max V(G_i)$.
- (11) $d(\bigcup_i G_i) \leq \max d(G_i) + 1$ where the G_i 's are disjoint.
- (12) $d(G \cup K_A) \leq d(G) + 1$ where $A \subset V$ and K_A is a clique on A .
- (13) $\text{sph}(G) < 12(2c-1)^2 \log |G|$ for $\lambda_{\min} \geq -c$ ($c \geq 2$) and $|G| > [12(2c-1)^2 \log |G|]^2$, where λ_{\min} is the minimal eigenvalue of G .
- (14) $\text{sph}(L(G)) < 108 \log m$ where $m = |E|$ and $m > (108 \log m)^2$.
- (15) $\text{sph}(G) \geq \log \alpha(G) / (\log(2r(G)+1))$ where $r(G)$ is the radius of G .

For these and related further results see [1]-[4] and [7].

The aim of this paper is to present some other upper and lower results for $d(G)$. In Section 1 we prove that $d(G) \leq n - \chi(G)$ where $\chi(G)$ is the chromatic number of G which, together with (1) above, yields $d(G) \leq n - \sqrt{n}$. In Section 2 we consider the dimension of forests and graphs with a large girth. In Section 3 we deal with lower bounds for $d(G)$. We prove that most of the graphs on n vertices have $d(G) \geq n/15$ for sufficiently large n . We also prove that, contrary to (3) above, there are graphs G_n with maximum degree ≤ 3 but with $d(G_n) > \log n / 18 \log \log n$.

1. General Upper Bounds

1.1. Theorem. *For every graph on n vertices,*

$$d(G) \leq n - \chi(G).$$

Proof. 1. The set V of vertices of $G = (V, E)$ with $\chi(G) = \chi$ can be written as

$$V = K \cup A_{k+1} \cup A_{k+2} \cup \dots \cup A_\chi,$$

where K is a clique of size k , each A_i has size ≥ 2 , and A_i is a maximal independent set in the induced subgraph G_i on $K \cup A_{k+1} \cup \dots \cup A_i$, $k+1 \leq i \leq \chi$.

Indeed, let $A_1 \cup \dots \cup A_\chi$ be the decomposition of V corresponding to a coloring of G such that $|A_1| \leq |A_2| \leq \dots \leq |A_\chi|$. If A_χ is not maximal independent in G_χ , we move suitable vertices of $A_1 \cup \dots \cup A_{\chi-1}$ to A_χ to make it such. Then we proceed in the same way with $A_{\chi-1}, A_{\chi-2}, \dots, A_1$. Thus if $k+1$ is the least integer with $|A_{k+1}| \geq 2$ then $K = A_1 \cup \dots \cup A_k$ is a clique of size k .

2. Let $|A_{k+1}| = d+1$. Then G_{k+1} can be represented in R^d with a threshold $t < 0$. Indeed, consider the vertices $\bar{z}_0, \dots, \bar{z}_d$ of the regular simplex in R^d centered at the origin with

$$\bar{z}_i \bar{z}_j = -1 \quad (i \neq j), \quad \bar{z}_i^2 = d.$$

Let $K = \{x_1, \dots, x_k\}$, $A_{k+1} = \{u_0, \dots, u_d\}$. Put

$$\bar{x}_i = \sum_{r \in I_i} \bar{z}_r \quad (i = 1, \dots, k),$$

where

$$I_i = \{r | x_i u_r \in E\}$$

and

$$\bar{u}_j = 2d^2 \bar{z}_j \quad (j = 0, \dots, d).$$

We have:

- (a) $\bar{x}_i \bar{x}_j = \sum \{\bar{z}_r \bar{z}_s | r \in I_i, s \in I_j\} \geq -d^2$.
- (b) $\bar{u}_i \bar{u}_j = -4d^4 < -d^2$ for $i \neq j$.
- (c) If $x_i u_j \in E$ then

$$\begin{aligned} \bar{x}_i \bar{u}_j &= 2d^2 \sum \{\bar{z}_r \bar{z}_j | r \in I_i\} \\ &= 2d^3 + 2d^2 \sum \{\bar{z}_r \bar{z}_j | r \in I_i - \{j\}\} \geq 2d^3 - 2d^3 = 0. \end{aligned}$$

- (d) If $x_i u_j \notin E$ then $\bar{x}_i \bar{u}_j = -2d^2 |I_i| < -d^2$ for $I_i \neq \emptyset$ because of the maximality of A_{k+1} .

Thus vectors \bar{x}_i, \bar{u}_j form a representation of G_{k+1} in R^d with the threshold $t = -d^2$.

3. Suppose G_h can be represented in R^d with a threshold $t < 0$. Then G_{h+1} can be represented in R^{d+s} , where $s+1 = |A_{h+1}|$, with a negative threshold. Indeed, let x_1, \dots, x_m be vertices of G_h and u_0, \dots, u_s vertices of A_{h+1} . Let $\bar{x}_1, \dots, \bar{x}_m$ form the representation of G_h in R^d with a threshold $t < 0$. Thus there is $t_1 < t$ such that

$$\bar{x}_i \bar{x}_j \geq t \quad \text{if } x_i x_j \in E \quad \text{and} \quad \bar{x}_i \bar{x}_j \leq t_1 \quad \text{if } x_i x_j \notin E.$$

Again, choose vectors $\bar{z}_0, \dots, \bar{z}_s$ in R^s with

$$\bar{z}_i \bar{z}_j = -1 \quad \text{for } i \neq j, \quad \bar{z}_i^2 = s.$$

Put

$$\begin{aligned} \tilde{x}_i &= \left(\bar{x}_i, \varepsilon \sum_{r \in I_i} \bar{z}_r \right) \in R^{d+s} \quad (i = 1, \dots, m), \\ \tilde{u}_j &= (\bar{0}, \alpha \bar{z}_j) \in R^{d+s} \quad (j = 0, \dots, s), \end{aligned}$$

where I_i is defined as in 2 above and $\alpha > 0$, $\varepsilon > 0$ will be specified later. Then we have:

(a) If $x_i, x_j \in E$ then

$$\tilde{x}_i \tilde{x}_j = \bar{x}_i \bar{x}_j + \varepsilon^2 \sum \{ \bar{z}_r \bar{z}_u \mid r \in I_i, u \in I_j \} \geq t - \varepsilon^2 s^2.$$

(b) If $x_i, x_j \notin E$ then

$$\tilde{x}_i \tilde{x}_j \leq t_1 + \varepsilon^2 \sum_{r=0}^s \bar{z}_r^2 \leq \varepsilon^2 s(s+1) + t_1.$$

(c) If $x_i, u_j \in E$ then

$$\begin{aligned} \tilde{x}_i \tilde{u}_j &= \alpha \varepsilon \sum \{ \bar{z}_r \bar{z}_j \mid r \in I_i \} \\ &= \alpha \varepsilon \bar{z}_j^2 + \alpha \varepsilon \sum \{ \bar{z}_r \bar{z}_j \mid r \in I_i - \{j\} \} \geq \alpha \varepsilon s - \alpha \varepsilon s = 0. \end{aligned}$$

(d) If $x_i, u_j \notin E$ then

$$\tilde{x}_i \tilde{u}_j = -\alpha \varepsilon |I_i| \leq -\alpha \varepsilon$$

because the inclusion maximality of A_{h+1} in G_{h+1} forces $|I_i| > 0$.

(e) $\tilde{u}_i \tilde{u}_j = -\alpha^2$ for $i \neq j$.

We are going to show that α, ε can be chosen such that vectors \tilde{x}_i, \tilde{u}_j form a representation of G_{h+1} with any threshold t' where $t_1 < t' < t < 0$. Indeed, first choose $\varepsilon > 0$ sufficiently small such that

$$\begin{aligned} t - \varepsilon^2 s^2 &> t' && \text{(see (a)),} \\ t_1 + \varepsilon^2 s(s+1) &< t' && \text{(see (b)).} \end{aligned}$$

Then choose α sufficiently large such that

$$\begin{aligned} -\alpha \varepsilon &< t' && \text{(see (d)),} \\ -\alpha^2 &< t' && \text{(see (e)).} \end{aligned}$$

4. It follows from 2 and 3 above that by induction $G = G_\chi$ can be represented in R^d where

$$\begin{aligned} d &= (|A_{k+1}| - 1) + (|A_{k+2}| - 1) + \dots + (|A_\chi| - 1) \\ &= |V - K| - (\chi - k) = n - \chi. \end{aligned}$$

This concludes the proof of the theorem. □

1.2. Corollary. *For every graph G on n vertices,*

$$d(G) \leq n - \sqrt{n}.$$

Proof. In [6] the threshold dimension $\theta(G)$ of a graph G is introduced and shown to satisfy $\theta(G) \leq n - \alpha(G)$ where $\alpha(G)$ is the size of a maximal independent set in G . Following [7], $d(G) \leq \theta(G)$ for every graph G . Thus $d(G) \leq n - \alpha(G)$. As $\alpha(G) \cdot \chi(G) \geq n$, we have $\max(\alpha(G), \chi(G)) \geq \sqrt{n}$, hence

$$d(G) \leq \min(n - \alpha(G), n - \chi(G)) \leq n - \sqrt{n}. \quad \square$$

2. Upper Bounds for Graphs Without Cycles

2.1. Proposition. *If G is a forest then $d(G) \leq 3$; the representation can be chosen to have a positive threshold.*

Proof. This is proved in [7] for G a tree; in more detail, for every $\epsilon > 0$, a tree $G = (V, E)$ can be represented in R^3 with the threshold $t = 1$ such that vectors \bar{x} representing vertices $x \in V$ satisfy $1 < \|\bar{x}\| < 1 + \epsilon$. Then $\text{diam}\{\bar{x} | x \in V\} \rightarrow 0$. Thus, as the tree G is connected, its representation can be found to be placed in an arbitrarily small ball with center on the unit sphere S in R^3 . Now, if G is a forest with components G_1, \dots, G_m , then choose distinct points c_1, \dots, c_m and balls B_i with centers c_i , each with a radius r . Represent G_i in B_i ; if r is small enough, this yields a representation of G in R^3 with the threshold $t = 1$.

2.2. Proposition. *If G is a forest that does not contain the tree T_0 below as an induced subgraph then $d(G) \leq 2$ where the representation in R^2 can be chosen to have a positive threshold t .*

Proof. In fact, this is proved in [7] for $G = (V, E)$ with $t' = 1 - \epsilon$ ($0 < \epsilon < 1$) where ϵ can be arbitrarily small, and vectors $\bar{x} \in R^2$ representing vertices of V satisfy $1 \geq \|\bar{x}\| \geq 1 - \epsilon$. Thus the representation can be constructed with $\text{diam}\{\bar{x} | x \in V\}$ arbitrarily small. Using the same argument as in the preceding proposition, this can be used to represent any forest without induced T_0 in R^2 with a positive threshold (Fig. 1).

2.3. Proposition. *A graph G on n vertices can be represented in R^d where:*

- I. $d \leq \frac{2}{7}n + 1$ if G has a girth at least 7,
- II. $d \leq \frac{1}{3}n + \frac{2}{3}$ if G has a girth at least 6,
- III. $d \leq \frac{2}{3}n + \frac{3}{3}$ if G has a girth at least 5.

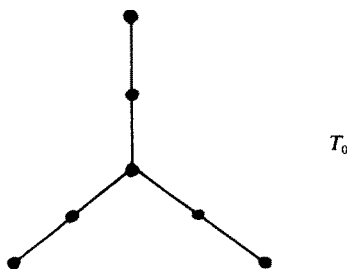


Fig. 1

Remark. Proposition 2.3 yields linear upper bounds for $d(G)$ for G without short cycles, which, however, can be further improved (to appear in a subsequent paper). Using a probabilistic approach we can, for example, show an upper bound of the form $c_1 \rho (\log n)^{c_2}$ for any C_4 -free graph with edge density ρ . As it is well known that $\rho \leq \sqrt{n}$ for any C_4 -free graph, this clearly improves Proposition 2.3 for $n \geq n(c_1, c_2)$. To prove Proposition 2.3 we need the following.

Lemma. *Let a graph $G = (V, E)$ contain a set C of vertices spanning one of the graphs C_5, C_6, C_7, C_8, P_7 such that each vertex in $V - C$ is adjacent to at most one of the vertices of C . Let the induced subgraph G' of G on $V - C$ be represented in R^d with a positive threshold. Then G can be represented in R^{d+2} with a positive threshold.*

Proof. Let $C = \{v_0, \dots, v_{r-1}\}$ span C_r , $5 \leq r \leq 8$. Put

$$\bar{z}_i = \left(\cos \frac{2\pi i}{r}, \sin \frac{2\pi i}{r} \right) \quad (i = 0, \dots, r-1).$$

Then

$$\bar{z}_i^2 = 1, \quad \bar{z}_i \bar{z}_{i+1} = \cos \frac{2\pi}{r} \geq \cos \frac{2\pi}{5} > 0.$$

$$\bar{z}_i \bar{z}_j \leq 0 \quad \text{for } i \neq j \pm 1 \pmod r, \quad i \neq j.$$

Let $x \rightarrow \bar{x} \in R^d$ be a representation of G' in R^d with a positive threshold t , without loss of generality, $t = 1$. Then, if $\alpha > C$ is sufficiently small, vectors

$$\tilde{x} = (\bar{x}, \alpha \bar{z}_i) \in R^{d+2} \quad (x \in V - C, v_i x \in E),$$

$$\tilde{x} = (\bar{x}, \bar{0}) \in R^{d+2} \quad (x \in V - C, v_i x \notin E \text{ for no } i)$$

form a representation of G' in R^{d+2} with a threshold t' , where

$$\cos \frac{2\pi}{r} < t' < 1.$$

Put

$$\tilde{v}_i = (\bar{0}, \bar{z}_i / \alpha).$$

Then, for sufficiently small $\alpha > 0$,

$$\tilde{v}_i \tilde{v}_{i+1} = \frac{1}{\alpha^2} \cos \frac{2\pi}{r} > 1,$$

$$\tilde{v}_i \tilde{v}_j \leq 0 \quad \text{if } i \neq j \pm 1 \pmod r, \quad i \neq j,$$

$$\tilde{x} \tilde{v}_i = 1 > t' \quad \text{if } x \in V - C \text{ is adjacent to } v_i,$$

$$\tilde{x} \tilde{v}_j \leq \cos \frac{2\pi}{r} < t' \quad \text{if } x \in V - C \text{ is not adjacent to } v_j.$$

This proves that vectors \tilde{x} ($x \in V - C$) and \tilde{v}_i ($i = 0, \dots, r - 1$) form a representation of G in R^{d+2} with the threshold t' .

If C spans P_7 , add a new vertex x_0 to G to be adjacent just to an endpoint of P_7 ; the resulting graph G' contains a copy of C_8 on $C \cup \{x_0\}$ and hence G' can be represented in R^{d+2} with a positive threshold by the preceding part of the proof.

Proof of Proposition 2.3. We shall prove by induction on n that G has a representation in R^d with a positive threshold where d is bounded by $\frac{2}{7}n + 1$ or $\frac{1}{3}n + \frac{2}{3}$ or $\frac{2}{5}n + \frac{3}{5}$ in cases I, II, and III, respectively. This is certainly true if $n \leq 3$. Let $n > 3$.

(a) Let G be a forest. Then $d(G) \leq 3 \leq \frac{2}{7}n + 1 \leq \frac{1}{3}n + \frac{2}{3} \leq \frac{2}{5}n + \frac{3}{5}$ if $n \geq 7$, see 2.1 above. If $4 \leq n < 7$ then $d(G) \leq 2 \leq \frac{1}{3}n + \frac{2}{3} \leq \frac{2}{7}n + 1 \leq \frac{2}{5}n + \frac{3}{5}$ by 2.2.

(b) Let $G = (V, E)$ not be a forest. Let D be the shortest cycle in G of length k . If $k > 8$, choose $C \subset D$ spanning P_7 ; if $k \leq 8$, let $C = D$. Due to the minimality of D , every vertex of the induced subgraph G' on $V - C$ is adjacent to at most one vertex of C . If $k \geq 7$ or $k = 6$ or $k = 5$ then G' has $\leq n - 7$ or $n - 6$ or $n - 5$ vertices, respectively. By the induction assumption, G can be represented in $R^{d'}$ where

$$d' \leq 2(n - 7)/7 + 1$$

or

$$d' \leq (n - 6)/3 + \frac{2}{3}$$

or

$$d' \leq 2(n - 5)/5 + \frac{3}{5},$$

respectively, with a positive threshold. By the lemma, G can be represented in R^d with a positive threshold where $d = d' + 2$, hence $d \leq \frac{2}{7}n + 1$, or $d \leq \frac{1}{3}n + \frac{2}{3}$ or $d \leq \frac{2}{5}n + \frac{3}{5}$, respectively.

2.4. The following is an essential improvement of a result of Frankl and Maehara [2] who proved

$$sd(G) \leq 8 \log n$$

if \bar{G} is a forest. Let us remark that $sd(G)$ is unbounded for trees. It was proved in [7] that

$$sd(B_n) \geq c \log n / \log \log n,$$

where B_n is the "complete binary" tree on $n = 2^s - 1$ vertices with s levels.

Proposition. *If the complement \bar{G} of G is a forest then*

$$sd(G) \leq 6.$$

Proof. Let $G = (V, E)$ and \bar{G} be a forest. In each component K_i of \bar{G} , choose a vertex r_i to be a root of the tree K_i . This defines decomposition of each K_i into levels. Also, neighbors of each vertex x that is not a root consist of a father and of a family of sons. Let V_1 be the set of all vertices on odd levels (thus all roots r_i are in V_1) and $V_2 = V - V_1$. Then V_2 can be written as $V_2 = \{x_1, \dots, x_n\}$ in such a way that sons of each $y \in V_1$ form a segment of the form

$$x_j, x_{j+1}, \dots, x_k.$$

Choose reals α_i ,

$$0 < \alpha_1 < \frac{\pi}{n} < \alpha_2 < \frac{2\pi}{n} < \dots < \frac{(n-1)\pi}{n} < \alpha_n < \pi.$$

Fix a $y \in V_1$ with sons as above and with the father x_r , define a complex polynomial g_y ,

$$\begin{aligned} g_y(z) &= \left(z - \exp\left(i \frac{\pi(j-1)}{n}\right) \right) \left(z - \exp\left(i \frac{k\pi}{n}\right) \right) \exp\left(-i \frac{\pi(j+k-1)}{2n}\right) \\ &\quad \times \left(z - \exp\left(i \frac{\pi(r-1)}{n}\right) \right) \left(z - \exp\left(i \frac{\pi r}{n}\right) \right) \exp\left(-i \frac{\pi(2r-1)}{2n}\right). \end{aligned}$$

We can verify that the coefficients of g_y are of the form

$$g_y(z) = c_2 + c_1 z + c_0 z^2 + \bar{c}_1 z^3 + \bar{c}_2 z^4, \quad c_0 \in \mathbf{R}.$$

Put $b_i^* = 2 \operatorname{Re} c_i$, $a_i^* = 2 \operatorname{Im} c_i$ ($i = 1, 2$), and $d_0^* = c_0$. For a real α , define $f_y(\alpha) = g_y(\exp i\alpha) \cdot \exp(-2i\alpha)$. Then

$$\begin{aligned} f_y(\alpha) &= c_2 e^{-2i\alpha} + c_1 e^{-i\alpha} + c_0 + \bar{c}_1 e^{i\alpha} + \bar{c}_2 e^{2i\alpha} \\ &= d_0^* + a_1^* \sin \alpha + b_1^* \cos \alpha + a_2^* \sin 2\alpha + b_2^* \cos 2\alpha. \end{aligned}$$

Thus f_y is real valued, $f_y(\alpha) = (d_0^*, a_1^*, b_1^*, a_2^*, b_2^*) \cdot (1, \sin \alpha, \cos \alpha, \sin 2\alpha, \cos 2\alpha)$. As the only roots of g_y are $\exp(i(\pi(j-1)/n))$, $\exp(i(\pi k/n))$, $\exp(i(\pi(r-1)/n))$, $\exp(i(\pi r/n))$ we see that the only roots of f_y are $\pi(j-1)/n$, $\pi k/n$, $\pi(r-1)/n$, and $\pi r/n$. Hence f_y has the same sign at α_j , α_{j+1} , \dots , α_{k-1} , α_k , and α_r , and the opposite sign at the other α_i 's. Put

$$\begin{aligned} \varepsilon &= -\operatorname{sgn} f_y(\alpha_r) \sqrt{3} / \sqrt{d_0^{*2} + a_1^{*2} + b_1^{*2} + a_2^{*2} + b_2^{*2}}, \\ h_y &= \varepsilon f_y, \quad d_0^y = \varepsilon d_0^*, \quad a_i^y = \varepsilon a_i^*, \quad b_i^y = \varepsilon b_i^* \quad (i = 1, 2). \end{aligned}$$

Then $h_y(\alpha_i) < 0$ for $j \leq i \leq k$ and for $i = r$ and $h_y(\alpha_i) > 0$ otherwise.

Put

$$\begin{aligned} \bar{y} &= (d_0^y, a_1^y, b_1^y, a_2^y, b_2^y, 2), \\ \bar{x}_i &= (1, \sin \alpha_i, \cos \alpha_i, \sin 2\alpha_i, \cos 2\alpha_i, -2). \end{aligned}$$

Then

$$\begin{aligned} \bar{x}_i \bar{y} &= h_y(\alpha_i) - 4 < -4 \quad \text{for } j \leq i \leq k, \quad i = r, \quad \text{i.e., for } x_i y \in E, \\ \bar{x}_i \bar{y} &= h_y(\alpha_i) - 4 > -4 \quad \text{otherwise,} \\ \bar{x}_i \bar{x}_j &> -1 + 4 > -4 \quad \text{for all } i, j, \\ \|\bar{x}_i\| &= \|\bar{y}\| = \sqrt{7}. \end{aligned}$$

Using the same construction for all $y \in V_1$, we also see that, for $y_1, y_2 \in V_1$,

$$\bar{y}_1 \bar{y}_2 > -3 + 4 > -4.$$

Thus we have a representation of G in R^6 by vectors with the same norm $\sqrt{7}$ and with the threshold $t = -4$ which concludes the proof. □

3. Lower Bounds

3.1. We shall prove that most graphs on n vertices have dimension $\geq n/15 - 1$. However, the only graphs of that dimension we explicitly know are those containing K_r for $r > n/15$ [7].

Theorem. *Let t_n be the total number of graphs G on $n \geq 38$ vertices. Then at least $(1 - 1/n)t_n$ of them have*

$$d(G) \geq n/15 - 1.$$

Proof. 1. Let G_1, G_2, \dots, G_{t_n} be a list of all graphs on the set $V = \{x_1, \dots, x_n\}$ of vertices where $d(G_1) \leq d(G_2) \leq \dots \leq d(G_{t_n})$. Let $r = \lfloor t_n/n \rfloor$, $d = d(G_{r+1}) + 1$. Then G_1, \dots, G_r can be represented in R^d with the threshold 1; in fact, if $x_i \rightarrow \bar{x}_i \in R^{d-1}$ is a representation in R^{d-1} with a threshold $t < 1$ (or $t \geq 1$) then $x_i \rightarrow (\bar{x}_i, \sqrt{1-t})$ or $(x_i \rightarrow (\bar{x}_i/t, 0), \text{ resp.})$ is a representation in R^d with the threshold 1. Also, using small changes of coordinates we can modify the representations in such a way that $\bar{x}_i \bar{x}_j > 1$ rather than $\bar{x}_i \bar{x}_j \geq 1$ if $x_i x_j$ is an edge, and that vectors $\bar{x}_1, \dots, \bar{x}_d$ representing vertices x_1, \dots, x_d are linearly independent. Finally, if $\tilde{x}_i = (\tilde{x}_{i1}, \dots, \tilde{x}_{id}) \in R^d$ is the vector of coordinates of x_i with respect to the orthonormal basis obtained from the base $\bar{x}_1, \dots, \bar{x}_d$ by the Gramm-Schmidt orthonormalization process, then $\bar{x}_i \bar{x}_j = \tilde{x}_i \tilde{x}_j$; moreover, obviously $\tilde{x}_{i,i+1} = \tilde{x}_{i,i+2} = \dots = \tilde{x}_{id} = 0$ for $i = 1, \dots, d-1$. To summarize our consideration, we have representations $x_i \rightarrow \tilde{x}_i^G$ of graphs $G = G_1, \dots, G_r$ in R^d such that

$$\begin{aligned} \tilde{x}_i^G \tilde{x}_j^G &> 1 \quad \text{if } x_i x_j \text{ is an edge in } G, \\ \tilde{x}_i^G \tilde{x}_j^G &< 1 \quad \text{if } x_i x_j \text{ is a nonedge in } G, \\ \tilde{x}_{ij}^G &= 0 \quad \text{if } 1 \leq i < j \leq d. \end{aligned}$$

2. Define polynomials P_{ij} ($1 \leq i < j \leq n$) in $(nd - (d(d-1)/2))$ variables $x_{11}, x_{21}, x_{22}, x_{31}, x_{32}, x_{33}, \dots, x_{d1}, \dots, x_{dd}, x_{d+1,1}, \dots, x_{d+1,d}, \dots, x_{n1}, \dots, x_{nd}$ by

$$P_{ij}(x_{uv})_{u,v} = \sum_{k=1}^d x_{ik}x_{jk} - 1$$

(where $x_{ik} = 0$ if $1 \leq i < k \leq d$ and analogously for x_{jk}). Then for $\tilde{x}^G = (\tilde{x}_{uv}^G)_{u,v} \in \mathbb{R}^N$ (where $N = nd - (d(d-1)/2)$) we have

$$P_{ij}(\tilde{x}^G) > 0 \quad \text{if } x_i x_j \text{ is an edge in } G,$$

$$P_{ij}(\tilde{x}^G) < 0 \quad \text{if } x_i x_j \text{ is a nonedge in } G.$$

3. Following a result of Warren [5], given arbitrary polynomials P_1, \dots, P_M on N variables of degree $\leq D$ where $M \geq N$, the total number of sign sequences $(\text{sgn } P_1(x), \dots, \text{sgn } P_M(x))$ that consist of terms $+1, -1$ does not exceed $(4eDM/N)^N$.

4. Let us apply this estimate to our polynomials P_{ij} ; as distinct graphs $G \neq G'$ induce distinct sign sequences of P_{ij} at \tilde{x}^G and $\tilde{x}^{G'}$, we have

$$r \leq (4eDM/N)^N.$$

Here $r = 2^{\binom{n}{2}}/n$, $D = 2$, $M = \binom{n}{2}$, and $N = nd - (d(d-1)/2)$. Hence

$$r \leq \left(8e \binom{n}{2} / N \right)^N \leq \left(\frac{4e}{N/n^2} \right)^N,$$

$$r^{1/n^2} \leq \left(\frac{4e}{y} \right)^y \quad \text{where } y = \frac{N}{n^2} = \frac{nd - (d(d-1)/2)}{n^2} \in (0, 1),$$

$$\frac{\ln r}{n^2} \leq y^{\ln(4e/y)}. \tag{*}$$

We have

$$\frac{\ln r}{n^2} = \frac{(\ln 2) \binom{n}{2} - \ln n}{n^2} = \frac{\ln 2}{2} - \frac{\ln 2}{2n} - \frac{\ln n}{n^2} \geq \frac{\ln 2}{2} - \frac{1}{85.9}$$

for $n \geq 38$. The numerical solution of the equation $y \ln(4e/y) - (\ln 2)2 + 1/85.9 = 0$ on $(0, 1)$ is $y \approx 1/15.26175$, and the function $y \ln(4e/y)$ is increasing on $(0, 1)$. Hence (*) implies $y \geq 1/15.262$ for $n \geq 38$. Putting $x = d/n$ we have

$$y = \frac{d}{n} - \frac{1}{2} \left(\frac{d}{n} \right)^2 + \frac{1}{2n} \left(\frac{d}{n} \right) = x - \frac{1}{2}x^2 + \frac{1}{2n}x \leq x - \frac{1}{2}x^2 + \frac{1}{16}x$$

for $n \geq 38$. Thus

$$\frac{77}{76}x - \frac{1}{2}x^2 \geq \frac{1}{15.262}.$$

A numerical solution then gives $x \geq 1/14.95 > \frac{1}{15}$, hence $d > n/15$. Then $d(G) > n/15 - 1$ for all $G = G_{r+1}, G_{r+2}, \dots, G_{t_n}$ which concludes the proof. \square

Remark. The constant $\frac{1}{15}$ can be improved for large n but not to $\frac{1}{14}$ using our method.

3.2. In [7] we proved that if the complement \bar{G} of a graph G has degree bounded by d then $d(G) \leq 2d \log(8d)$, independently of the number of vertices of G . We show that this is not the case if G has bounded degree:

Theorem. *If n is sufficiently large then there exists a bipartite 3-regular graph G on $2n$ vertices with*

$$d(G) \geq \frac{\log n}{18 \log \log n}.$$

To prove the theorem we state two auxiliary lemmas.

Lemma 1. *For sufficiently large n there exists a bipartite graph $G = (V, E)$ on $2n$ vertices such that:*

- (i) G is 3-regular.
- (ii) For every partition $V = A \cup B$, $|A| = |B| = n$ there exists a matching M , $|M| = cn \geq n/34$ with e having one endpoint in A and another in B for every $e \in M$.
- (iii) G does not contain induced cycles of length $\leq 2t_0$ where $t_0 = \frac{1}{8} \log n$.

Proof. We will proceed by random construction. Let $V = X \cup Y$ and $|X| = |Y| = n$. Let $\pi_1, \pi_2, \pi_3: X \rightarrow Y$ be three random bijections, each taken with probability $1/n!$ independently of the choice of the others. Let $G = (V, E)$ be the “random graph” with edge set $E = \{\{x, \pi_i(x)\}; x \in X, i = 1, 2, 3\}$. Thus the probability of a particular graph G equals $\text{pg}(G)/(n!)^3$ where $\text{pg}(G)$ denotes the number of ways that G can be written as a union (not necessarily disjoint) of three perfect matchings.

Claim 1. *Denote by p_n the probability that G is 3-regular. Then*

$$\lim_{n \rightarrow \infty} p_n \geq 1/e^3.$$

Proof. The permutation π_1 can be chosen in $n!$ ways. If π_1 is fixed, then $n!$ $(1 - 1/1! + 1/2! - \dots + (-1)^n/n!)$ of permutations π_2 has the property that $\pi_2(x) \neq \pi_1(x)$ for all $x \in X$, i.e., that edges $x\pi_i(x)$ ($x \in X, i = 1, 2$) form a 2-regular graph H . If π_2 is such a permutation, consider the $n \times n$ matrix $(a_{ij})_{i \in X, j \in Y}$ where

$$a_{ij} = 0 \quad \text{if } ij \text{ is an edge of } H,$$

$$a_{ij} = 1/(n-2) \quad \text{otherwise.}$$

Then $\sum_{i \in X} a_{ij} = 1$ for every $j \in Y$ and $\sum_{j \in Y} a_{ij} = 1$ for every $i \in X$.

Hence (a_{ij}) is a doubly stochastic matrix. By [8] and [9], for the permanent

$$P = \sum_{\pi \in S_n} a_{1\pi(1)} a_{2\pi(2)} \dots a_{n\pi(n)}$$

of (a_{ij}) we have $P \geq n!/n^n$. On the other hand, $P = q(1/(n-2))^n$ where q is the number of nonzero summands in P . As the nonzero summands clearly correspond to these π such that π_1, π_2, π induce a 3-regular graph, it follows $q \geq n!/(1-2/n)^n$.

Hence the number of triples π_1, π_2, π_3 inducing a 3-regular graph is $\geq n! n! (1 - 1/1! + 1/2! - \dots + (-1)^n/n!) n! (1 - 2/n)^n$. Then

$$p_n \geq \left(1 - \frac{1}{1!} + \frac{1}{2!} - \dots + \frac{(-1)^n}{n!}\right) \left(1 - \frac{2}{n}\right)^n \rightarrow 1/e^3$$

which concludes the proof of Claim 1. □

Claim 2. Denote by q_n the probability that G satisfies (ii) with $c = \frac{1}{33}$. Then $q_n \rightarrow 1$ for $n \rightarrow \infty$.

Proof. Let $V = A \cup B$ be a fixed partition, $|A| = |B| = n$, $a = |X \cap A|$, $b = |X - A|$, and $a \geq b$. Then $b = |Y \cap A|$, $a = |Y - A|$, and $a + b = n$. Set $c = \frac{1}{33}$,

$$S_i^A = \{(x, y) | x \in X \cap Y, y \in Y - A, \pi_i(x) = y\}.$$

We have

$$\text{Prob}\{|S_i^A| < cn\} \leq \frac{\sum_{0 \leq j < cn} \binom{a}{j} \binom{b}{a-j}}{\binom{a+b}{a}}. \tag{1}$$

The right-hand side is nonzero only if $b \geq a - j$ for some j , i.e., only if $a - cn \leq b$ which yields

$$a \leq \lfloor (n/2)(1+c) \rfloor. \tag{2}$$

The right-hand side of (1) may be further bounded from above by

$$\frac{\sum_{0 \leq j < cn} \binom{a}{j}^2}{\binom{a+b}{a}} \leq \frac{\binom{a}{\lceil cn \rceil}^2}{\binom{n}{a}} \leq \frac{\left(\frac{\lfloor (n/2)(1+c) \rfloor}{\lceil cn \rceil} \right)^2}{\binom{n}{\lfloor (n/2)(1+c) \rfloor}} \leq 2^{-n\varphi(c) + K \log n},$$

where $\varphi(c) = 2c \log c + \frac{1}{2}(1-c) \log((1-c)/2) - \frac{3}{2}(1+c) \log((1+c)/2)$ and K is independent of n . As $\varphi(\frac{1}{33}) > \frac{2}{3}$,

$$\text{Prob}\{|S_i^A| < cn\} \leq 2^{-(2/3)n}$$

for sufficiently large n . Then $\text{Prob}\{|S_i^A| < cn \mid i = 1, 2, 3\} \leq 2^{-2n}$. The number of all partitions in question is $\binom{2n}{n}$. Thus the probability that $|S_i^A| < cn$ ($i = 1, 2, 3$) for some of them is $\leq \binom{2n}{n} 2^{-2n} \rightarrow 0$. □

Claim 3. Denote by r_n the probability that G has the following property: after deleting $\leq \sqrt{n}$ edges from G , the graph “ G minus deleted edges” does not contain any cycle of length $\leq 2t_0$ where $t_0 = \frac{1}{8} \log n$. Then $r_n \geq 0.99$ holds for sufficiently large n .

Proof. Fix a set C of $2t$ pairs xy ($x \in X, y \in Y$) forming a cycle, where $t \leq t_0$. We will estimate the probability $r_n(C)$ that G contains C as a cycle. Let π_1, π_2, π_3 be bijections inducing $G = (V, E)$ and let E_i be the set of edges induced by π_i ($i = 1, 2, 3$). Set $M_i = C \cap E_i$; clearly, M_i is a matching. By [10, p. 129] we have $2^{2t} + 2$ ways to split C into three matchings M_1, M_2, M_3 . For each of them we have $a = (n - t_1)!(n - t_2)!(n - t_3)!$ triples π_1, π_2, π_3 with $E_i \cap C = M_i$ ($i = 1, 2, 3$), where $|M_i| = t_i$ ($i = 1, 2, 3$).

Hence

$$r_n(C) \leq (2^{2t} + 2) \cdot \text{Max} \frac{(n - t_1)!(n - t_2)!(n - t_3)!}{(n!)^3} \leq \frac{2^{2t+1}}{n^{2t}}$$

for sufficiently large n (as $t_i \ll \sqrt{n}$). The number of cycles of length $2t$ formed by pairs $xy, x \in X, y \in Y$, is

$$\binom{n}{t}^2 t!(t-1)! \leq \frac{n^{2t}}{t}$$

Thus the expected number of cycles of length $\leq 2t_0$ contained in G is at most

$$\sum_{t \leq t_0} \frac{2^{2t+1}}{t} < \frac{\sqrt{n}}{100}$$

Hence we infer that the probability ≥ 0.99 , G contains at most \sqrt{n} cycles of length $\leq 2t_0$. Choosing an edge from each of these cycles we get the required set of edges.

Now we are ready to finish the proof of the lemma. Due to Claims 1–3, for sufficiently large n there exists a bipartite graph $G = (V, E)$ on $2n$ vertices that is 3-regular, has the partition property (ii), and only $\leq \sqrt{n} \cdot 2t_0$ of its edges is contained in induced cycles of length $\leq 2t_0$.

We modify G to get a new graph satisfying (i)–(iii) of the lemma. Consider an edge xy of G that is contained in a cycle C of length $\leq 2t_0$. Choose another edge $x'y'$ such that the distance from x to x' in G is $\geq 2t_0 + 1$ and that $x'y'$ is contained in no cycle of length $\leq 2t_0$; such $x'y'$ does exist, for at most $\frac{1}{8}n \log n$ vertices are contained in cycles of length $\leq 2t_0$ and at most $3(2^{2t_0+1} - 1) \cdot \sqrt{n}$ vertices $x' \neq x$ are joined with x by a path of length $\leq 2t_0 + 1$. Delete edges $xy, x'y'$ from G and add xy', yx' to G . This does destroy the cycle C but does not create any new cycle of length $\leq 2t_0$; moreover, the resulting graph remains 3-regular. Repeating this procedure at most \sqrt{n} times (once for each short cycle) we get a 3-regular graph without cycles of length $\leq 2t_0$. Also, the new graph satisfies property (ii) with $|M| \geq \frac{1}{33}n - \sqrt{n}$, i.e., $|M| \geq n/34$ for sufficiently large n .

Lemma 2. *The graph $G = (V, E)$ from Lemma 1 has the following property: for every function $f: V \rightarrow \mathbb{R}$ there is an induced subtree T of G such that:*

- (i) $\text{diam } T \leq 2t_0$.
- (ii) $|L| \geq n^{1/17}$, where L is the set of leaves of T .
- (iii) $f(x) \geq f(y)$ whenever x is a leaf and y is a nonleaf in T .

Proof. For every $x \in V$, let L_x be the set of all vertices of V where the distance from x is $t_1 = \lfloor t_0/2 \rfloor - 1$. As G does not contain cycles of length $\leq 2t_0$ and G is 3-regular, L_x has cardinality

$$S = 3 \cdot 2^{t_1-1} < \sqrt{n}.$$

Write V as $V = \{v_1, \dots, v_{2n}\}$ where $f(v_1) \geq f(v_2) \geq \dots \geq f(v_{2n})$. Choose j such that for $P = \{v_1, \dots, v_j\}$ and for $B = \bigcup_{x \in P} L_x$ we have

$$n - \sqrt{n} \leq |B| \leq n$$

and put $A = V - B$. Choose $A' \subset A, B' \supset B$ with $|A'| = |B'| = n$. By Lemma 1, condition (ii), there exists a matching M joining A' with B' with $|M| \geq \frac{1}{34}n$. Let A_1 be the set of vertices x in A' such that $xy \in M$ for some $y \in B$. As $|B' - B| < \sqrt{n}$ we have $|A_1| \geq \frac{1}{34}n - \sqrt{n}$ and for $A_2 = A - A_1$ we have $|A_2| = |A| - |A_1| \leq \frac{33}{34}n + 2\sqrt{n} \leq \frac{34}{33}n$ for sufficiently large n .

- (a) We claim that there exists $x_0 \in A$ such that

$$|L_{x_0} - A_2| \geq \frac{S}{35}.$$

Indeed, otherwise $|L_x \cap A_2| > \frac{34}{35}S$ for every $x \in A$. Denote by N the number of pairs (x, y) with $x \in A$, $y \in A_2$, and $y \in L_x$ (equivalently, $x \in L_y$). Then, on one hand,

$$N = \sum_{x \in A} |L_x \cap A_2| > |A| \frac{34}{35} S \geq \frac{34}{35} nS$$

and, on the other hand,

$$N = \sum_{y \in A_2} |L_y \cap A| \leq |A_2| S \leq \frac{34}{35} nS,$$

a contradiction.

(b) Let $L_{x_0} - A_2 = \{x_1, \dots, x_m\}$, $m \geq S/35$. For every $i \in \{1, \dots, m\}$, we have a path P_i of length t_1 from x_0 to x_i and a path Q_i of length $\leq t_1 + 1$ from x_i to some $z_i \in P$ (the existence of Q_i follows by the definition of B if $x_i \in B$; if $x_i \in A_1$ we use an M -edge $x_i y$ with $y \in B$ and a path of length $= t_1$ from y to a vertex in P). As the x_i 's are pairwise distinct and G does not contain cycles of length $2t_0 \geq 4t_1 + 2$, the z_i 's are pairwise distinct, too, and the union of P_i 's and Q_i 's is an induced subtree T of G . Assuming, in addition, that z_i is the only P -vertex on Q_i , the tree T satisfies (iii) due to the definition of P . We have $\text{diam } T \leq 4t_1 + 2 \leq 2t_0$ and $|L| = m \geq S/35 = \frac{3}{35} 2^{\lfloor t_0/2 \rfloor - 2} \geq 2^{(1/17) \log n} = n^{1/17}$ for sufficiently large n . □

Proof of the Theorem. Let $G = (V, E)$ be the 3-regular bipartite graph on $2n$ vertices from Lemma 2; let $x \rightarrow \bar{x}$ ($x \in V$) be its representation in R^d with $t = 1$. Set $f(x) = \|\bar{x}\|$ for $x \in V$. Let T be the induced subtree of Lemma 2. Thus we have $\gamma > 0$ such that $\|\bar{x}\| \geq \gamma$ or $\|\bar{x}\| \leq \gamma$ if x is a leaf or a nonleaf in T , respectively. Let x_0 be a nonleaf such that the distance from x_0 to each of the leaves is at most t_0 . For every leaf x , let f_x and g_x be vertices on the path from x_0 to x where the distance from x is 1 and 2, respectively.

Let L_1 be a maximal set of leaves with the property

$$x, y \in L_1, \quad x \neq y \text{ implies } g_x \neq g_y \text{ and } g_x g_y \notin E.$$

It follows easily from the 3-regularity of G and the maximality of L_1 that $|L_1| \geq |L|/8 \geq n^{1/17}/8$. For every $x \in L_1$ set $\tilde{x} = \alpha \bar{x} + (1 - \alpha) \bar{g}_x$ where $\alpha \in \langle 0, 1 \rangle$ is chosen so that $\|\tilde{x}\| = \gamma$; such α does exist because $\|\bar{x}\| = \|\bar{x}\| \geq \gamma$ for $\alpha = 0$ and $\|\tilde{x}\| = \|\bar{g}_x\| \leq \gamma$ for $\alpha = 1$; for $x \in L_1$ set $\tilde{x} = \bar{x}$. Then for $x, y \in L_1$, $x \neq y$,

$$\begin{aligned} \tilde{x}\tilde{y} &= [\alpha \bar{x} + (1 - \alpha) \bar{g}_x][\beta \bar{y} + (1 - \beta) \bar{g}_y] < 1, \\ \tilde{x}\tilde{f}_x &= [\alpha \bar{x} + (1 - \alpha) \bar{g}_x] \bar{f}_x \geq 1. \end{aligned}$$

For every $x \in L_1$, let K_x be the ball with center \tilde{x} and radius $r = \sqrt{2\gamma^2 - 2}/2$. We have for $x, y \in L_1$

$$\|\tilde{x} - \tilde{y}\| = (\tilde{x}^2 - 2\tilde{x}\tilde{y} + \tilde{y}^2)^{1/2} > (2\gamma^2 - 2)^{1/2} = 2r.$$

Hence the balls K_x are pairwise disjoint. On the other hand, if $x_0, x_1, \dots, x_t = x$ is a path from x_0 to $x \in L$, then

$$\|\tilde{x}_i - \tilde{x}_{i+1}\| = (\tilde{x}_i^2 - 2\tilde{x}_i\tilde{x}_{i+1} + \tilde{x}_{i+1}^2)^{1/2} \leq (2\gamma^2 - 2)^{1/2} = 2r,$$

hence balls K_x are contained in the ball with center x_0 and radius $\leq (2t_0 + 1)r$. Then

$$|L_1| V_d r^d \leq V_d [(2t_0 + 1)r]^d,$$

where V_d is the volume of the unit ball in R^d . It follows

$$n^{1/17}/8 \leq (2t_0 + 1)^d,$$

and thus

$$d \geq \frac{\log(n^{1/17}/8)}{\log(2t_0 + 1)} \geq \frac{\frac{1}{17} \log n - 3}{\log(\log n) - 2} \geq \frac{\log n}{18 \log(\log n)}$$

for sufficiently large n . □

References

1. P. Frankl, H. Maehara, On the contact dimension of graphs, to appear.
2. P. Frankl, H. Maehara, The Johnson-Lindenstrauss lemma and the sphericity of some graphs, to appear.
3. H. Maehara, On the sphericity for the join of many graphs, *Discrete Math.* **49** (1984), 311-313.
4. H. Maehara, Space graphs and sphericity, *Discrete Appl. Math.* **7** (1984), 55-64.
5. H. E. Warren, Lower bounds of approximation by nonlinear manifolds.
6. V. Chvátal, P. L. Hammer, Aggregation of inequalities in integer programming, *Ann. Discrete Math.* **1** (1977), 145-162.
7. J. Reiterman, V. Rödl, E. Šiňajová, Geometrical embeddings of graphs, *Discrete Math.*, to appear.
8. G. P. Egoryčev, *Solution of the van der Waerden Problem for Permanents*, IFSO 13M, Akad. Nauk SSSR.
9. D. I. Falikman, A proof of the van der Waerden conjecture on the permanent of a doubly stochastic matrix, *Mat. Zametki* **29** (1981), 931-938.
10. J. A. Bondy, U. S. R. Murty, *Graph Theory with Applications*, Macmillan, 1978.

Received December 8, 1986, and in revised form July 10, 1987.