

# Explicit Ramsey graphs and orthonormal labelings

Noga Alon \*

Submitted: August 22, 1994; Accepted October 29, 1994

## Abstract

We describe an explicit construction of triangle-free graphs with no independent sets of size  $m$  and with  $\Omega(m^{3/2})$  vertices, improving a sequence of previous constructions by various authors. As a byproduct we show that the maximum possible value of the Lovász  $\theta$ -function of a graph on  $n$  vertices with no independent set of size 3 is  $\Theta(n^{1/3})$ , slightly improving a result of Kashin and Konyagin who showed that this maximum is at least  $\Omega(n^{1/3}/\log n)$  and at most  $O(n^{1/3})$ . Our results imply that the maximum possible Euclidean norm of a sum of  $n$  unit vectors in  $R^n$ , so that among any three of them some two are orthogonal, is  $\Theta(n^{2/3})$ .

## 1 Introduction

Let  $R(3, m)$  denote the maximum number of vertices of a triangle-free graph whose independence number is at most  $m$ . The problem of determining or estimating  $R(3, m)$  is a well studied Ramsey type problem. Ajtai, Komlós and Szemerédi proved in [1] that  $R(3, m) \leq O(m^2/\log m)$ , (see also [17] for an estimate with a better constant). Improving a result of Erdős, who showed in [7] that  $R(3, m) \geq \Omega((m/\log m)^2)$ , (see also [18], [13] or [4] for a simpler proof), Kim [10] proved, very recently, that the upper bound is tight, up to a constant factor, that is:  $R(3, m) = \Theta(m^2/\log m)$ . His proof, as well as that of Erdős, is probabilistic, and does not supply any explicit construction of such a graph. The problem of finding an explicit construction of triangle-free graphs of independence number  $m$  and many vertices has also received a considerable amount of attention. Erdős [8] gave an explicit construction of such graphs with

$$\Omega(m^{(2 \log 2)/3(\log 3 - \log 2)}) = \Omega(m^{1.13})$$

vertices. This has been improved by Cleve and Dagum [6], and improved further by Chung, Cleve and Dagum in [5], where the authors present a construction with

$$\Omega(m^{\log 6/\log 4}) = \Omega(m^{1.29})$$

---

\*AT & T Bell Labs, Murray Hill, NJ 07974, USA and Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Research supported in part by a United States Israel BSF Grant.

vertices. The best known explicit construction is given in [2], where the number of vertices is  $\Omega(m^{4/3})$ .

Here we improve this bound and describe an explicit construction of triangle free graphs with independence numbers  $m$  and  $\Omega(m^{3/2})$  vertices. Our graphs are Cayley graphs and their construction is based on some of the properties of certain Dual BCH error-correcting codes. The bound on their independence numbers follows from an estimate of their Lovász  $\theta$ -function. This fascinating function, introduced by Lovász in [14], can be defined as follows. If  $G = (V, E)$  is a graph, an *orthonormal labeling* of  $G$  is a family  $(b_v)_{v \in V}$  of unit vectors in an Euclidean space so that if  $u$  and  $v$  are distinct non-adjacent vertices, then  $b_u^t b_v = 0$ , that is,  $b_u$  and  $b_v$  are orthogonal. The  $\theta$ -number  $\theta(G)$  is the minimum, over all orthonormal labelings  $b_v$  of  $G$  and over all unit vectors  $c$ , of

$$\max_{v \in V} \frac{1}{(c^t b_v)^2}.$$

It is known (and easy; see [14]) that the independence number of  $G$  does not exceed  $\theta(G)$ . The graphs  $G_n$  we construct here are triangle free graphs on  $n$  vertices satisfying  $\theta(G_n) = \Theta(n^{2/3})$ , and hence the independence number of  $G_n$  is at most  $O(n^{2/3})$ .

The construction and the properties of the  $\theta$ -function settle a geometric problem posed by Lovász and partially solved by Kashin and Konyagin [12], [9]. Let  $\Delta_n$  denote the maximum possible value of the Euclidean norm  $\|\sum_{i=1}^n u_i\|$  of the sum of  $n$  unit vectors  $u_1, \dots, u_n$  in  $R^n$ , so that among any three of them some two are orthogonal. Motivated by the study of the  $\theta$ -function, Lovász raised the problem of determining the order of magnitude of  $\Delta_n$ . In [12] it is shown that  $\Delta_n \leq O(n^{2/3})$  and in [9] it is proved that this is nearly tight, namely that  $\Delta_n \geq \Omega(n^{2/3}/(\log n)^{1/2})$ . Here we show that the upper bound is tight up to a constant factor, that is:

$$\Delta_n = \Theta(n^{2/3}).$$

The rest of this note is organized as follows. In Section 2 we construct our graphs and estimate their  $\theta$ -numbers and their independence numbers. The resulting lower bound for  $\Delta_n$  is described in Section 3. Our method in these sections combines the ideas of [9] with those in [2]. The final Section 4 contains some concluding remarks.

## 2 The graphs

For a positive integer  $k$ , let  $F_k = GF(2^k)$  denote the finite field with  $2^k$  elements. The elements of  $F_k$  are represented, as usual, by binary vectors of length  $k$ . If  $a, b$  and  $c$  are three such vectors, let  $(a, b, c)$  denote their concatenation, i.e., the binary vector of length  $3k$  whose coordinates are those of  $a$ , followed by those of  $b$  and those of  $c$ . Suppose  $k$  is not divisible by 3 and put  $n = 2^{3k}$ . Let  $W_0$  be the set of all nonzero elements  $\alpha \in F_k$  so that the leftmost bit in the binary representation of  $\alpha^7$  is 0, and let  $W_1$  be the set of all nonzero elements  $\alpha \in F_k$  for which the leftmost bit of  $\alpha^7$  is

1. Since 3 does not divide  $k$ , 7 does not divide  $2^k - 1$  and hence  $|W_0| = 2^{k-1} - 1$  and  $|W_1| = 2^{k-1}$ , as when  $\alpha$  ranges over all nonzero elements of  $F_k$  so does  $\alpha^7$ .

Let  $G_n$  be the graph whose vertices are all  $n = 2^{3k}$  binary vectors of length  $3k$ , where two vectors  $u$  and  $v$  are adjacent if and only if there exist  $w_0 \in W_0$  and  $w_1 \in W_1$  so that  $u + v = (w_0, w_0^3, w_0^5) + (w_1, w_1^3, w_1^5)$ , where here the powers are computed in the field  $F_k$  and the addition is addition modulo 2. Note that  $G_n$  is the Cayley graph of the additive group  $(Z_2)^{3k}$  with respect to the generating set  $S = U_0 + U_1 = \{u_0 + u_1 : u_0 \in U_0, u_1 \in U_1\}$ , where  $U_0 = \{(w_0, w_0^3, w_0^5) : w_0 \in W_0\}$ , and  $U_1$  is defined similarly. The following theorem summarizes some of the properties of the graphs  $G_n$ .

**Theorem 2.1** *If  $k$  is not divisible by 3 and  $n = 2^{3k}$  then  $G_n$  is a  $d_n = 2^{k-1}(2^{k-1} - 1)$ -regular graph on  $n = 2^{3k}$  vertices with the following properties.*

1.  $G_n$  is triangle-free.
2. Every eigenvalue  $\mu$  of  $G_n$ , besides the largest, satisfies

$$-9 \cdot 2^k - 3 \cdot 2^{k/2} - 1/4 \leq \mu \leq 4 \cdot 2^k + 2 \cdot 2^{k/2} + 1/4.$$

3. The  $\theta$ -function of  $G_n$  satisfies

$$\theta(G_n) \leq n \frac{36 \cdot 2^k + 12 \cdot 2^{k/2} + 1}{2^k(2^k - 2) + 36 \cdot 2^k + 12 \cdot 2^{k/2} + 1} \leq (36 + o(1))n^{2/3},$$

where here the  $o(1)$  term tends to 0 as  $n$  tends to infinity.

**Proof.** The graph  $G_n$  is the Cayley graph of  $Z_2^{3k}$  with respect to the generating set  $S = S_n = U_0 + U_1$ , where  $U_i$  are defined as above.

Let  $A_0$  be the  $3k$  by  $2^{k-1} - 1$  binary matrix whose columns are all vectors of  $U_0$ , and let  $A_1$  be the  $3k$  by  $2^{k-1}$  matrix whose columns are all vectors of  $U_1$ . Let  $A = [A_0, A_1]$  be the  $3k$  by  $2^k - 1$  matrix whose columns are all those of  $A_0$  and those of  $A_1$ . This matrix is the parity check matrix of a binary BCH-code of designed distance 7 (see, e.g., [16], Chapter 9), and hence every set of six columns of  $A$  is linearly independent over  $GF(2)$ . In particular, all the sums  $(u_0 + u_1)_{u_0 \in U_0, u_1 \in U_1}$  are distinct and hence  $|S_n| = |U_0||U_1|$ . It follows that  $G_n$  has  $2^{3k}$  vertices and it is  $|S_n| = 2^{k-1}(2^{k-1} - 1)$  regular.

The fact that  $G_n$  is triangle-free is equivalent to the fact that the sum (modulo 2) of any set of 3 elements of  $S_n$  is not the zero-vector. Let  $u_0 + u_1, u'_0 + u'_1$  and  $u''_0 + u''_1$  be three distinct elements of  $S_n$ , where  $u_0, u'_0, u''_0 \in U_0$  and  $u_1, u'_1, u''_1 \in U_1$ . If the sum (modulo 2) of these six vectors is zero then, since every six columns of  $A$  are linearly independent, every vector must appear an even number of times in the sequence  $(u_0, u'_0, u''_0, u_1, u'_1, u''_1)$ . However, since  $U_0$  and  $U_1$  are disjoint this implies that every vector must appear an even number of times in the sequence  $(u_0, u'_0, u''_0)$  and this is clearly impossible. This proves part 1 of the theorem.

In order to prove part 2 we argue as follows. Recall that the eigenvalues of Cayley graphs of abelian groups can be computed easily in terms of the characters of the group. This result, described in, e.g., [15], implies that the eigenvalues of the graph  $G_n$  are all the numbers

$$\sum_{s \in S_n} \chi(s),$$

where  $\chi$  is a character of  $Z_2^{3k}$ . By the definition of  $S_n$ , these eigenvalues are precisely all the numbers

$$\left( \sum_{u_0 \in U_0} \chi(u_0) \right) \left( \sum_{u_1 \in U_1} \chi(u_1) \right).$$

It follows that these eigenvalues can be expressed in terms of the Hamming weights of the linear combinations (over  $GF(2)$ ) of the rows of the matrices  $A_0$  and  $A_1$  as follows. Each linear combination of the rows of  $A$  of Hamming weight  $x + y$ , where the Hamming weight of its projection on the columns of  $A_0$  is  $x$  and the weight of its projection on the columns of  $A_1$  is  $y$ , corresponds to the eigenvalue

$$(2^{k-1} - 1 - 2x)(2^{k-1} - 2y).$$

Our objective is thus to bound these quantities.

The linear combinations of the rows of  $A$  are simply all words of the code whose generating matrix is  $A$ , which is the dual of the BCH-code whose parity-check matrix is  $A$ . It is known (see [16], pages 280-281) that the Carlitz-Uchiyama bound implies that the Hamming weight  $x + y$  of each non-zero codeword of this dual code satisfies

$$2^{k-1} - 2^{1+k/2} \leq x + y \leq 2^{k-1} + 2^{1+k/2}. \tag{1}$$

Let  $p$  denote the characteristic vector of  $W_1$ , that is, the binary vector indexed by the non-zero elements of  $F_k$  which has a 1 in each coordinate indexed by a member of  $W_1$  and a 0 in each coordinate indexed by a member of  $W_0$ . Note that the sum (modulo 2) of  $p$  and any linear combination of the rows of  $A$  is a non-zero codeword in the dual of the BCH-code with designed distance 9. Therefore, by the Carlitz-Uchiyama bound, the Hamming weight of the sum of  $p$  with the linear combination considered above, which is  $x + (2^{k-1} - y)$ , satisfies

$$2^{k-1} - 3 \cdot 2^{k/2} \leq x + 2^{k-1} - y \leq 2^{k-1} + 3 \cdot 2^{k/2}. \tag{2}$$

Since for any two reals  $a$  and  $b$ ,

$$-\left(\frac{a-b}{2}\right)^2 \leq ab \leq \left(\frac{a+b}{2}\right)^2$$

we conclude from (1) that

$$(2^{k-1} - 1 - 2x)(2^{k-1} - 2y) \leq \frac{(2^k - 1 - 2(x+y))^2}{4} \leq 4 \cdot 2^k + 2 \cdot 2^{k/2} + 1/4.$$

Similarly, (2) implies that

$$(2^{k-1} - 1 - 2x)(2^{k-1} - 2y) \geq -\frac{(1 + 2(x - y))^2}{4} \geq -9 \cdot 2^k - 3 \cdot 2^{k/2} - 1/4.$$

This completes the proof of part 2 of the theorem.

Part 3 follows from part 2 together with Theorem 9 of [14] which asserts that for  $d$ -regular graphs  $G$  with eigenvalues  $d = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ ,

$$\theta(G) \leq \frac{-n\lambda_n}{\lambda_1 - \lambda_n}.$$

It is worth noting that the fact that the right hand side in the last inequality bounds the independence number of  $G$  is due to A. J. Hoffman.  $\square$

Since the independence number of each graph  $G$  does not exceed  $\theta(G)$  the following result follows.

**Corollary 2.2** *If  $k$  is not divisible by 3 and  $n = 2^{3k}$ , then the graph  $G_n$  is a triangle-free graph with independence number at most  $(36 + o(1))n^{2/3}$ .  $\square$*

Let  $G_n$  be one of the graphs above and let  $\overline{G}_n$  denote its complement. Since  $G_n$  is a Cayley graph, Theorem 8 in [14] implies that  $\theta(\overline{G}_n)\theta(G_n) = n$  and hence, by Theorem 2.1,  $\theta(\overline{G}_n) \geq (1 + o(1))\frac{1}{36}n^{1/3}$ .

In [9] it is proved (in a somewhat disguised form), that for any graph  $H$  with  $n$  vertices and no independent set of size 3,  $\theta(H) \leq 2^{2/3}n^{1/3}$ . (See also [3] for an extension). Since  $\overline{G}_n$  has no independent set of size 3 and since for every graph  $H$ ,  $\theta(H)\theta(\overline{H}) \geq n$  (see Corollary 2 of [14]) the following result follows.

**Corollary 2.3** *If  $k$  is not divisible by 3 and  $n = 2^{3k}$ , then  $\theta(G_n) = \Theta(n^{2/3})$  and  $\theta(\overline{G}_n) = \Theta(n^{1/3})$ . Therefore, the minimum possible value of the  $\theta$ -number of a triangle-free graph on  $n$  vertices is  $\Theta(n^{2/3})$  and the maximum possible value of the  $\theta$ -number of an  $n$ -vertex graph with no independent set of size 3 is  $\Theta(n^{1/3})$ .*

### 3 Nearly orthogonal systems of vectors

A system of  $n$  unit vectors  $u_1, \dots, u_n$  in  $R^n$  is called *nearly orthogonal* if any set of three vectors of the system contains an orthogonal pair. Let  $\Delta_n$  denote the maximum possible value of the Euclidean norm  $\|\sum_{i=1}^n u_i\|$ , where the maximum is taken over all systems  $u_1, \dots, u_n$  of nearly orthogonal vectors. Lovász raised the problem of determining the order of magnitude of  $\Delta_n$ . Konyagin showed in [12] that  $\Delta_n \leq O(n^{2/3})$  and that

$$\Delta_n \geq \Omega(n^{4/3 - \log 3/2 \log 2}) \geq \Omega(n^{0.54}).$$

The lower bound was improved by Kashin and Konyagin in [9], where it is shown that

$$\Delta_n \geq \Omega(n^{2/3}/(\log n)^{1/2})$$

The following theorem asserts that the upper bound is tight up to a constant factor.

**Theorem 3.1** *There exists an absolute positive constant  $a$  so that for every  $n$*

$$\Delta_n \geq an^{2/3}.$$

Thus,  $\Delta_n = \Theta(n^{2/3})$ .

**Proof.** It clearly suffices to prove the lower bound for values of  $n$  of the form  $n = 2^{3k}$ , where  $k$  is an integer and 3 does not divide  $k$ . Fix such an  $n$ , let  $G = G_n = (V, E)$  be the graph constructed in the previous section and define  $\theta = \theta(G)$ . By Theorem 2.1,  $\theta \leq (36 + o(1))n^{2/3}$ . By the definition of  $\theta$  there exists an orthonormal labeling  $(b_v)_{v \in V}$  of  $G$  and a unit vector  $c$  so that  $(c^t b_v)^2 \geq 1/\theta$  for every  $v \in V$ . Therefore, the norm of the projection of each  $b_v$  on  $c$  is at least  $1/\sqrt{\theta}$  and by assigning appropriate signs to the vectors  $b_v$  we can ensure that all these projections are in the same direction. With this choice of signs, the norm of the projection of  $\sum_{v \in V} b_v$  on  $c$  is at least  $n/\sqrt{\theta}$ , implying that

$$\left\| \sum_{v \in V} b_v \right\| \geq n/\sqrt{\theta} \geq \left(\frac{1}{6} - o(1)\right)n^{2/3}.$$

Note that since the vectors  $b_v$  form an orthonormal labeling of  $G$ , which is triangle-free, among any three of them there are some two which are orthogonal. This implies that  $(b_v)_{v \in V}$  is a nearly orthogonal system and shows that for every  $n = 2^{3k}$  as above

$$\Delta_n \geq \left(\frac{1}{6} - o(1)\right)n^{2/3},$$

completing the proof of the theorem.  $\square$

## 4 Concluding remarks

The method applied here for explicit constructions of triangle-free graphs with small independence numbers cannot yield asymptotically better constructions. This is because the independence number is bounded here by bounding the  $\theta$ -number which, by Corollary 2.3, cannot be smaller than  $\Theta(n^{2/3})$  for any triangle-free graph on  $n$  vertices.

Some of the results of [9] can be extended. In a forthcoming paper with N. Kahale [3] we show that for every  $k \geq 3$  and every graph  $H$  on  $n$  vertices with no independent set of size  $k$ ,

$$\theta(H) \leq Mn^{1-2/k}, \tag{3}$$

for some absolute positive constant  $M$ . It is not known if this is tight for  $k > 3$ . Combining this with some of the properties of the  $\theta$ -function, this can be used to show that for every  $k \geq 3$  and any system of  $n$  unit vectors  $u_1, \dots, u_n$  in  $R^n$  so that among any  $k$  of them some two are orthogonal, the inequality

$$\left\| \sum_{i=1}^n u_i \right\| \leq O(n^{1-1/k})$$

holds. This is also not known to be tight for  $k > 3$ . Lovász (cf. [11]) conjectured that there exists an absolute constant  $c$  so that for every graph  $H$  on  $n$  vertices and no independent set of size  $k$ ,

$$\theta(H) \leq ck\sqrt{n}.$$

Note that this conjecture, if true, would imply that the estimate (3) above is *not* tight for all fixed  $k > 4$ .

**Acknowledgment** I would like to thank Nabil Kahale for helpful comments and Rob Calderbank for fruitful suggestions that improved the presentation significantly.

## References

- [1] M. Ajtai, J. Komlós and E. Szemerédi, *A note on Ramsey numbers*, J. Combinatorial Theory Ser. A 29 (1980), 354-360.
- [2] N. Alon, *Tough Ramsey graphs without short cycles*, to appear.
- [3] N. Alon and N. Kahale, in preparation.
- [4] N. Alon and J. H. Spencer, **The Probabilistic Method**, Wiley, 1991.
- [5] F. R. K. Chung, R. Cleve and P. Dagum, *A note on constructive lower bounds for the Ramsey numbers  $R(3, t)$* , J. Combinatorial Theory Ser. B 57 (1993), 150-155.
- [6] R. Cleve and P. Dagum, *A constructive  $\Omega(t^{1.26})$  lower bound for the Ramsey number  $R(3, t)$* , Inter. Comp. Sci. Inst. Tech. Rep. TR-89-009, 1989.
- [7] P. Erdős, *Graph Theory and Probability, II*, Canad. J. Math. 13 (1961), 346-352.
- [8] P. Erdős, *On the construction of certain graphs*, J. Combinatorial Theory 17 (1966), 149-153.
- [9] B. S. Kashin and S. V. Konyagin, *On systems of vectors in a Hilbert space*, Trudy Mat. Inst. imeni V. A. Steklova 157 (1981), 64-67. English translation in: Proc. of the Steklov Institute of Mathematics (AMS 1983), 67-70.
- [10] J. H. Kim, *The Ramsey number  $R(3, t)$  has order of magnitude  $t^2/\log t$* , to appear.

- [11] D. E. Knuth, *The sandwich theorem*, Electronic Journal of Combinatorics A1 (1994), 48pp.
- [12] S. V. Konyagin, *Systems of vectors in Euclidean space and an extremal problem for polynomials*, Mat. Zametki 29 (1981), 63-74. English translation in: Mathematical Notes of the Academy of the USSR 29 (1981), 33-39.
- [13] M. Krivelevich, *Bounding Ramsey numbers through large deviation inequalities*, to appear.
- [14] L. Lovász, *On the Shannon capacity of a graph*, IEEE Transactions on Information Theory IT-25, (1979), 1-7.
- [15] L. Lovász, **Combinatorial Problems and Exercises**, North Holland, Amsterdam, 1979, Problem 11.8.
- [16] F. J. MacWilliams and N. J. A. Sloane, **The Theory of Error-Correcting Codes**, North Holland, Amsterdam, 1977.
- [17] J. B. Shearer, *A note on the independence number of a triangle-free graph*, Discrete Math. 46 (1983), 83-87.
- [18] J. Spencer, *Asymptotic lower bounds for Ramsey functions*, Discrete Math. 20 (1977), 69-76.