ELSEVIER

On a problem of Duke–Erdős–Rödl on cycle-connected subgraphs

Jacob Fox [a,1], Benny Sudakov [b,2]

[a] *Department of Mathematics, Princeton University, Princeton, NJ, USA*
[b] *Department of Mathematics, UCLA, Los Angeles, CA 90095, USA*

Abstract

In this short note, we prove that for $\beta < 1/5$ every graph G with n vertices and $n^{2-\beta}$ edges contains a subgraph G' with at least $cn^{2-2\beta}$ edges such that every pair of edges in G' lie together on a cycle of length at most 8. Moreover edges in G' which share a vertex lie together on a cycle of length at most 6. This result is best possible up to the constant factor and settles a conjecture of Duke, Erdős, and Rödl.
© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Let $\mathcal{H}$ be a fixed collection of graphs. A graph G is $\mathcal{H}$-*connected* if every pair of edges of G is contained in a subgraph H of G, where H is a member of $\mathcal{H}$. For example, if $\mathcal{H}$ is the collection of all paths, then, ignoring isolated vertices, $\mathcal{H}$-connectedness is equivalent to connectedness. If $\mathcal{H}$ consists of all paths of length at most d, then each $\mathcal{H}$-connected graph has a diameter at most d, while every graph of diameter d is $\mathcal{H}$-connected for $\mathcal{H}$ the collection of all paths of length at most $d + 2$. So $\mathcal{H}$-connectedness naturally extends basic notions of connectivity.

The definition of $\mathcal{H}$-connectedness was introduced by Duke, Erdős, and Rödl, who initiated the study of this notion in a series of four papers [5–8]. A graph is C_{2k}-*connected* if it

is $\mathcal{H}$-connected where $\mathcal{H}$ consists of all even-length cycles of length at most $2k$. The question studied by Duke, Erdős, and Rödl was to determine the maximum number of edges in a C_{2k}-connected subgraph that one can find in every graph with n vertices and m edges as a function of k, n, and m. The following problem was considered to be one of the main open problems in this area. It was first posed by Duke, Erdős, and Rödl [6] in 1984, and discussed in the two subsequent papers [7,8]. It also appears in the book *Erdős on Graphs* by Chung and Graham [4].

Problem 1.1. Is it true that there are constants c, $\beta_0 > 0$ such that for all $0 \leqslant \beta \leqslant \beta_0$ the following holds. Every graph G with n vertices and $n^{2-\beta}$ edges contains a subgraph G' with $cn^{2-2\beta}$ edges such that every two edges of G' lie together on a cycle of length at most eight?

It is easy to see that such a result would be best possible up to a multiplicative constant. Indeed, the bound on the number of edges in G' is tight when G is a collection of n^β disjoint complete graphs of size roughly $n^{1-\beta}$. In [6], Duke, Erdős, and Rödl obtained a weaker result which proves that the assertion of Problem 1.1 is correct if one allows the cycle length to be at most twelve instead of at most eight. They also showed how to find in G a C_6-connected (and hence also C_8-connected) subgraph G' with at least $cn^{2-3\beta}$ edges.

The analogue of Problem 1.1 for graphs of constant density was solved in [7]. In that paper, the authors proved that for each fixed $d > 0$, every graph G with n vertices and at least dn^2 edges has a subgraph G' on $(1 + o(1))d^2n^2$ edges such that every pair of edges of G' lie together on a cycle of length at most eight. Unfortunately, the proof of Duke, Erdős, and Rödl uses Szemerédi's regularity lemma and consequently gives nothing when d tends to zero.

Duke, Erdős, and Rödl [6] also asked whether Problem 1.1 holds in the stronger form, where the subgraph G' has the additional property that edges sharing a vertex lie together on a cycle of length at most 6. Motivated by this question, we call a graph *strongly C_{2k}-connected* if it is C_{2k}-connected and every pair of edges sharing a vertex lie together on a cycle of length at most $2k - 2$. In this note, we settle Problem 1.1 in its strengthened form for all $\beta < 1/5$.

Theorem 1.2. *For $0 < \beta < 1/5$ and sufficiently large n, every graph G on n vertices and at least $n^{2-\beta}$ edges has a strongly C_8-connected subgraph G' with at least $\frac{1}{64}n^{2-2\beta}$ edges.*

Our proof combines combinatorial ideas together with a probabilistic argument which may be called *dependent random choice*. Early versions of this technique were developed in the papers [10,11,14]. Later, variants were discovered and applied to a large variety of extremal problems (see, e.g., [1,9,12,15,16]). In the concluding remarks, we show how the same proof can be used to obtain a variant of the main graph theoretic lemma which is used in the proof of the celebrated Balog–Szemerédi–Gowers theorem. Hence, we wonder if our result might have new applications in Additive Combinatorics.

2. Proof of Theorem 1.2

Let $\beta < 1/5$, $k = n^\beta$ and let G be a graph with n vertices and at least n^2/k edges. Since $\beta < 1/5$ and n is sufficiently large, we may assume that $n > 2^{20}k^5$. Delete vertices of minimum degree one by one until the remaining induced subgraph G_1 of G has minimum degree at least $\frac{n}{2k}$.

Since the number of vertices deleted in this process is at most n, we have that the number of remaining edges in G_1 is at least

$$e(G_1) \geqslant e(G) - n \cdot \frac{n}{2k} \geqslant \frac{n^2}{k} - \frac{n^2}{2k} = \frac{n^2}{2k}.$$

Let H be the maximum bipartite subgraph of G_1, and let A and B denote the vertex classes of H. Without loss of generality we can assume that $|B| \leqslant |A|$. For a vertex $x \in H$ denote by $d_H(x)$ its degree, i.e., the number of vertices adjacent to x in H. By maximality of H, the degree of every vertex in H is at least half of its degree in G_1 and the number of edges in H is at least half of the number of edges in G_1. Hence the minimum degree in H is at least $\frac{n}{4k}$ and H has at least $\frac{n^2}{4k}$ edges. For two vertices $x_1, x_2 \in H$ define the *common neighborhood* $N_H(x_1, x_2)$ to be the set of vertices of H adjacent to both x_1 and x_2 and the *codegree* $d_H(x_1, x_2)$ to be the size $|N_H(x_1, x_2)|$. We will later use the following simple fact.

Lemma 2.1. *If every pair of vertices in a subset $X \subset A$ have codegree in H at most $\frac{n}{32k^2}$, then $|X| < 8k$.*

Proof. Suppose for contradiction that there is a subset $X = \{x_1, \ldots, x_{8k}\}$ such that every pair of vertices in it have codegree in H at most $\frac{n}{32k^2}$. By the Bonferroni inequality (inclusion–exclusion principle), the number of vertices of B adjacent to at least one of $x_1, \ldots, x_{8k}$ is at least

$$\sum_{1 \leqslant i \leqslant 8k} d_H(x_i) - \sum_{1 \leqslant i < j \leqslant 8k} d_H(x_i, x_j) \geqslant 8k \frac{n}{4k} - \binom{8k}{2} \frac{n}{32k^2} > n.$$

Therefore, the size of B is larger than the total number of vertices n. This contradiction completes the proof. $\quad\square$

Define an auxiliary graph Γ on A where two vertices in Γ are adjacent if their codegree in H is at least $\frac{n}{32k^2}$. Then the previous lemma simply states that Γ has no independent set of size $8k$. Let Γ' be an induced subgraph of Γ on $v \geqslant 16k$ vertices. If the number of edges in Γ' is at most $\frac{v^2}{32k}$, then its average degree is at most $\frac{v}{16k}$. Therefore, by Turán's theorem [18], it has an independent set of size at least $v/(\frac{v}{16k} + 1) \geqslant 8k$, which contradicts Lemma 2.1. Thus we have the following claim.

Lemma 2.2. *Every induced subgraph of Γ on $v \geqslant 16k$ vertices has more than $\frac{v^2}{32k}$ edges.*

In particular, in any induced subgraph Γ_1 of Γ, there are at most $\frac{n}{2^{12}k^4}$ vertices of degree at most $\frac{n}{2^{16}k^5}$. Otherwise, the subgraph $\Gamma' \subset \Gamma_1$ induced by the vertices of degree at most $\frac{n}{2^{16}k^5}$ has $v \geqslant \frac{n}{2^{12}k^4} \geqslant 16k$ vertices and has at most

$$\frac{1}{2} v \cdot \frac{n}{2^{16}k^5} = \frac{1}{32k} v \cdot \frac{n}{2^{12}k^4} \leqslant \frac{v^2}{32k}$$

edges, contradicting the previous lemma.

We say that a vertex $w \in A$ is *bad with respect to a pair* $\{u, v\}$ of vertices of B if $w \in N_H(u, v)$ and w has degree at most $\frac{n}{2^{16}k^5}$ in the induced subgraph $\Gamma[N_H(u, v)]$ of the auxiliary graph Γ.

Lemma 2.3. *Let u, v be two vertices in B. Pick a vertex w in A uniformly at random. Let $\mathcal{E}$ be the event that w is bad with respect to the pair $\{u, v\}$. The probability of event $\mathcal{E}$ is at most $\frac{n}{2^{12}k^4|A|}$.*

Proof. Let t denote the cardinality of $N_H(u, v)$. The probability that $w \in N_H(u, v)$ is given by $|N_H(u, v)|/|A| = t/|A|$. Since, by the discussion in the previous paragraph, at most $\frac{n}{2^{12}k^4}$ vertices in $N_H(u, v)$ have degree at most $\frac{n}{2^{16}k^5}$ in the induced subgraph $\Gamma[N_H(u, v)]$ of Γ, then the probability that a vertex picked uniformly at random from $N_H(u, v)$ has degree at most $\frac{n}{2^{16}k^5}$ in $\Gamma[N_H(u, v)]$ is at most $\frac{n}{2^{12}k^4}\frac{1}{t}$. Hence, the probability of the event $\mathcal{E}$ satisfies

$$\mathbb{P}[\mathcal{E}] = \mathbb{P}\big[w \in N_H(u, v)\big] \cdot \mathbb{P}\big[w \text{ is bad} \mid w \in N_H(u, v)\big]$$
$$\leqslant \frac{t}{|A|} \cdot \frac{n}{2^{12}k^4}\frac{1}{t} = \frac{n}{2^{12}k^4|A|}. \qquad \square$$

Pick a vertex $w \in A$ uniformly at random. Let Y be the random variable counting the number of pairs $\{u, v\}$ in B such that w is bad with respect to $\{u, v\}$. Since there are $\binom{|B|}{2}$ pairs of elements of B and $|B| \leqslant |A| \leqslant n$, then by Lemma 2.3, we have

$$\mathbb{E}[Y] \leqslant \binom{|B|}{2} \frac{n}{2^{12}k^4|A|} < \frac{n^2}{2^{13}k^4}.$$

Hence there exist a choice of w such that the number of pairs $\{u, v\}$ in B for which w is bad is less than $\frac{n^2}{2^{13}k^4}$. Pick such a w and delete all vertices from A that have fewer than $\frac{n}{32k^2}$ neighbors in $N_H(w)$. That is, delete those vertices in A that are not adjacent to w in the auxiliary graph Γ. Let A' be the remaining subset of A.

Delete one by one vertices v from $N_H(w)$ for which there are at least $\frac{n}{2^7k^2}$ vertices u in the remaining set such that w is bad for $\{u, v\}$. Since w is bad only for at most $\frac{n^2}{2^{13}k^4}$ pairs, it is easy to see that we deleted at most

$$\frac{n^2/(2^{13}k^4)}{n/(2^7k^2)} = \frac{n}{2^6k^2} \tag{1}$$

vertices. Denote the remaining subset of $N_H(w)$ by B'. Note that $|B'| \geqslant |N_H(w)| - \frac{n}{2^6k^2} = \frac{n}{4k} - \frac{n}{2^6k^2} \geqslant \frac{n}{5k}$. By the definition of B', we have that for every $v \in B'$, there are fewer than $\frac{n}{2^7k^2}$ vertices $u \in B$ such that w is bad for pair $\{u, v\}$. Let G' be the bipartite subgraph of H induced by $A' \cup B'$. We will show that this graph satisfies the assertion of Theorem 1.2. The next lemma summarizes several important properties of G'.

Lemma 2.4.

(i) *The degree in G' of every vertex in A' is at least $\frac{n}{2^6k^2}$.*
(ii) *For every vertex $v \in B'$ there are fewer than $\frac{n}{2^7k^2}$ vertices $u \in B'$ such that $\{v, u\}$ have less than $\frac{n}{2^{16}k^5}$ common neighbors in A'.*
(iii) *The number of edges in G' is at least $\frac{n^2}{2^6k^2}$.*

Proof. (i) Recall that to obtain A' we removed from A all vertices of small degree in $N_H(w)$. Thus the vertices in A' all have degree at least $\frac{n}{32k^2}$ in $N_H(w)$. Also by (1), we deleted at most

$\frac{n}{2^6 k^2}$ vertices from $N_H(w)$ to obtain B'. Therefore, all vertices from A' still have at least $\frac{n}{2^6 k^2}$ remaining neighbors in B'.

(ii) Let $\{v, u\}$ be a pair of vertices in B' for which w is good. By definition, this means that there are at least $\frac{n}{2^{16} k^5}$ vertices z in A such that z is a common neighbor of $\{v, u\}$ and the codegree of z and w is at least $\frac{n}{32 k^2}$. All these vertices z have high degree in $N_H(w)$ and were not deleted when we constructed A'. This implies that there are at least $\frac{n}{2^{16} k^5}$ common neighbors of pair $\{v, u\}$ in A'. To conclude the proof of this part note that by our construction for every vertex $v \in B'$ there are less than $\frac{n}{2^7 k^2}$ vertices $u \in B'$ such that w is bad for $\{v, u\}$.

(iii) Since the minimum degree in H is at least $\frac{n}{4k}$, we have that $|N_H(w)| \geqslant \frac{n}{4k}$ and the number of edges between $N_H(w)$ and A is at least $\frac{n}{4k}|N_H(w)| \geqslant \frac{n^2}{16 k^2}$. Since the vertices we deleted from A all have degree at most $\frac{n}{32 k^2}$ in $N_H(w)$, the total number of remaining edges between A' and $N_H(w)$ is at least

$$\frac{n^2}{16 k^2} - \frac{n}{32 k^2}|A| \geqslant \frac{n^2}{16 k^2} - \frac{n^2}{32 k^2} = \frac{n^2}{32 k^2}.$$

By (1), the number of edges between A' and $N_B(w) \setminus B'$ is at most

$$|A| |N_H(w) \setminus B'| \leqslant n \cdot \frac{n}{2^6 k^2} \leqslant \frac{n^2}{2^6 k^2}.$$

Hence, the number of edges between A' and B', which is the number of edges of G', is at least $\frac{n^2}{32 k^2} - \frac{n^2}{2^6 k^2} = \frac{n^2}{2^6 k^2}$. $\square$

Having finished all the necessary preparation we are now ready to complete the proof of Theorem 1.2. Recall that $n > 2^{20} k^5$ and let $(a, b), (a', b') \in A' \times B'$ be two edges of G'.

Case 1. (a, b) and (a', b') do not share a vertex. By properties (i) and (ii) of Lemma 2.4, there are at least $d_{G'}(a) - \frac{n}{2^7 k^2} \geqslant \frac{n}{2^6 k^2} - \frac{n}{2^7 k^2} = \frac{n}{2^7 k^2}$ neighbors b_1 of a such that b' and b_1 have at least $\frac{n}{2^{16} k^5} > 4$ common neighbors in A'. Fix any such $b_1 \neq b$ and let a_1 be a common neighbor of $\{b', b_1\}$ which is different from a, a'. Similarly, we can pick a neighbor b_2 of a' different from b, b', b_1 such that b and b_2 have at least $\frac{n}{2^{16} k^5} > 4$ common neighbors in A'. Let a_2 be a common neighbor of $\{b, b_2\}$ which is distinct from a, a', a_1. Then $a, b_1, a_1, b', a', b_2, a_2, b, a$ form an 8-cycle which contains edges $(a, b), (a', b')$.

Case 2. $a = a'$. Let a_1 be a neighbor of b with $a_1 \neq a'$. (Note that by property (ii) of the previous lemma the degree of b is at least $\frac{n}{2^{16} k^5} > 4$.) Then, as in the previous case, we have that there is a neighbor b_1 of a_1 different from b, b' such that b' and b_1 have at least $\frac{n}{2^{16} k^5} > 4$ common neighbors in A'. Let a_2 be a common neighbor of $\{b', b_1\}$ which is distinct from a, a_1. Then $a, b, a_1, b_1, a_2, b', a$ form a 6-cycle which contains edges $(a, b), (a, b')$.

Case 3. $b = b'$. Let b_1 be a neighbor of a with $b_1 \neq b$. Then again, as in Case 1, there is neighbor b_2 of a' different from b, b_1 such that b_1 and b_2 have at least $\frac{n}{2^{16} k^5} > 4$ common neighbors in A'. Let a_2 be a common neighbor of $\{b_1, b_2\}$ which is distinct from a, a'. Then $a, b, a', b_2, a_2, b_1, a$ form a 6-cycle which contains edges $(a, b), (a', b)$. $\square$

3. Concluding remarks

- We suspect that the approach which was used to settle Problem 1.1 with some changes might work also for values of β larger than $1/5$. On the other hand, since for our proof it is crucial to have vertices with large codegree, it surely fails if $\beta \geqslant 1/2$. It would be very interesting to determine all values of β for which Problem 1.1 have a positive answer. For every β that is sufficiently close to 1 there are graphs with $n^{2-\beta}$ edges and no 8-cycle (see, e.g., [3]). Clearly, for such β the answer to this problem is negative.
- Duke, Erdős, and Rödl [6] showed that for $0 < \beta < 1/2$ and a graph G on n vertices and at least $n^{2-\beta}$ edges, there is a C_6-connected subgraph G' on at least $cn^{2-3\beta}$ edges, and this result is tight up to the multiplicative constant c. However, it is still open whether this result can be strengthened to show that every graph G with n vertices and $n^{2-\beta}$ edges has a strongly C_6-connected subgraph G' with at least $cn^{2-3\beta}$ edges. Duke, Erdős, and Rödl [6] proved that such a graph G will have a strongly C_6-connected subgraph G' with at least $cn^{2-5\beta}$ edges.
- The Balog–Szemerédi–Gowers theorem is a very useful tool in Additive Combinatorics. For example, it is an important ingredient in Gowers' proof [10] of Szemerédi's theorem on arithmetic progressions in dense sets. For detailed discussion and more applications of this theorem see, e.g., the books by Nathanson [13] and by Tao and Vu [17].
  Let A and B be sets of integers. The *sumset* $A + B$ is defined to be the collection of sums $a + b$ with $a \in A$, $b \in B$. For a bipartite graph $G = (A, B; E)$, the *partial sumset* $A +_G B$ is defined to be the collection of sums $a + b$ with $(a, b) \in E(G)$. The Balog–Szemerédi theorem [2] states that for A and B sets of n integers with $|E(G)| \geqslant n^2/k$ and $|A +_G B| \leqslant cn$ for some k and c, there are $A' \subset A$ and $B' \subset B$ such that $|A'|, |B'| \geqslant n/K$ and $|A' + B'| \leqslant Cn$, where K and C only depend on k and c. The original proof of this theorem gave a poor bound on K and C in terms of k and c. Gowers [10] discovered a new proof in which K and C can be taken to be polynomials in k and c.
  One can deduce the Balog–Szemerédi–Gowers theorem rather quickly from a graph-theoretic lemma proved by Sudakov, Szemerédi, and Vu [16], which essentially says that for every dense bipartite graph $G = (A, B; E)$ with $|A| = |B|$, there are linear-sized subsets $A' \subset A$ and $B' \subset B$ such that for every pair $(a, b) \in A' \times B'$, there are a quadratic number of paths in G of length three between a and b. The following theorem strengthens this graph-theoretic lemma, showing that the paths of length three can be taken to lie entirely within subgraph of G induced by $A' \cup B'$.

Theorem 3.1. *For every bipartite graph $G = (A, B; E)$ with $n \geqslant 2^{18}k^5$ vertices and $|E| \geqslant n^2/k$ edges, there are subsets $A' \subset A$ and $B' \subset B$ such that the subgraph G' of G induced by $A' \cup B'$ has at least $\frac{n^2}{2^6k^2}$ edges and for every $a \in A'$ and $b \in B'$, there are at least $\frac{n^2}{2^{24}k^7}$ paths between a and b in G' of length three.*

Proof. The proof follows easily from Lemma 2.4. By properties (i) and (ii) of this lemma, there are at least $d_{G'}(a) - \frac{n}{2^7k^2} \geqslant \frac{n}{2^6k^2} - \frac{n}{2^7k^2} = \frac{n}{2^7k^2}$ neighbors b_1 of a such that pair $\{b, b_1\}$ have at least $\frac{n}{2^{16}k^5}$ common neighbor in A'. For any such $b_1 \neq b$ and any common neighbor $a_1 \neq a$ we have a path of length three a, b_1, a_1, b. The number of such paths is clearly at least $(\frac{n}{2^7k^2} - 1)(\frac{n}{2^{16}k^5} - 1) \geqslant \frac{n^2}{2^{24}k^7}$. $\square$

We wonder if this theorem might have new applications in Additive Combinatorics.

Acknowledgment

References

[1] N. Alon, M. Krivelevich, B. Sudakov, Turán numbers of bipartite graphs and related Ramsey-type questions, Combin. Probab. Comput. 12 (2003) 477–494.

[2] A. Balog, E. Szemerédi, A statistical theorem of set addition, Combinatorica 14 (1994) 263–268.

[3] C.T. Benson, Minimal regular graphs of girths eight and twelve, Canad. J. Math. 18 (1966) 1091–1094.

[4] F. Chung, R. Graham, Erdős on Graphs: His Legacy of Unsolved Problems, AK Peters, Ltd., Wellesley, MA, 1998.

[5] R.A. Duke, P. Erdős, Subgraphs in which each pair of edges lies in a short common cycle, in: Proceedings of the Thirteenth Southeastern Conference on Combinatorics, Graph Theory and Computing, Boca Raton, FL, 1982, in: Congr. Numer., vol. 35, 1982, pp. 253–260.

[6] R.A. Duke, P. Erdős, V. Rödl, More results on subgraphs with many short cycles, in: Proceedings of the Fifteenth Southeastern Conference on Combinatorics, Graph Theory and Computing, Baton Rouge, LA, 1984, in: Congr. Numer., vol. 43, 1984, pp. 295–300.

[7] R.A. Duke, P. Erdős, V. Rödl, Extremal problems for cycle-connected graphs, in: Proceedings of the Twenty-Second Southeastern Conference on Combinatorics, Graph Theory, and Computing, Baton Rouge, LA, 1991, in: Congr. Numer., vol. 83, 1991, pp. 147–151.

[8] R.A. Duke, P. Erdős, V. Rödl, Cycle-connected graphs. Topological, algebraical and combinatorial structures, Discrete Math. 108 (1992) 261–278.

[9] R.A. Duke, P. Erdős, V. Rödl, On large intersecting subfamilies of uniform setfamilies, Random Structures Algorithms 23 (2003) 351–356.

[10] W.T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four, Geom. Funct. Anal. 8 (1998) 529–551.

[11] A. Kostochka, V. Rödl, On graphs with small Ramsey numbers, J. Graph Theory 37 (2001) 198–204.

[12] A. Kostochka, B. Sudakov, On Ramsey numbers of sparse graphs, Combin. Probab. Comput. 12 (2003) 627–641.

[13] M.B. Nathanson, Additive Number Theory. Inverse Problems and the Geometry of Sumsets, Grad. Texts in Math., vol. 165, Springer-Verlag, New York, 1996.

[14] B. Sudakov, Few remarks on the Ramsey–Turán-type problems, J. Combin. Theory Ser. B 88 (2003) 99–106.

[15] B. Sudakov, Large K_r-free subgraphs in K_s-free graphs and some other Ramsey-type problems, Random Structures Algorithms 26 (2005) 253–265.

[16] B. Sudakov, E. Szemerédi, V.H. Vu, On a question of Erdős and Moser, Duke Math. J. 129 (2005) 129–155.

[17] T. Tao, V.H. Vu, Additive Combinatorics, Cambridge University Press, Cambridge, 2006.

[18] P. Turán, On a problem in extremal graph theory, Math. Fiz. Lapok 48 (1941) 436–452 (in Hungarian).