The sum of degrees in cliques

Béla Bollobás[*†‡] and Vladimir Nikiforov[*]

Submitted: Sep 10, 2004; Accepted: Nov 1, 2005; Published: Nov 7, 2005
Mathematics Subject Classifications: 05C35

Abstract

For every graph G, let

$$\Delta_r(G) = \max\left\{\sum_{u\in R} d(u) : R \text{ is an } r\text{-clique of } G\right\}$$

and let $\Delta_r(n,m)$ be the minimum of $\Delta_r(G)$ taken over all graphs of order n and size m. Write $t_r(n)$ for the size of the r-chromatic Turán graph of order n.

Improving earlier results of Edwards and Faudree, we show that for every $r \geq 2$, if $m \geq t_r(n)$, then

$$\Delta_r(n,m) \geq \frac{2rm}{n}, \tag{1}$$

as conjectured by Bollobás and Erdős.

It is known that inequality (1) fails for $m < t_r(n)$. However, we show that for every $\varepsilon > 0$, there is $\delta > 0$ such that if $m > t_r(n) - \delta n^2$ then

$$\Delta_r(n,m) \geq (1-\varepsilon)\frac{2rm}{n}.$$

1 Introduction

Our notation and terminology are standard (see, e.g. [1]): thus $G(n,m)$ stands for a graph of n vertices and m edges. For a graph G and a vertex $u \in V(G)$, we write $\Gamma(u)$ for the set of vertices adjacent to u and set $d_G(u) = |\Gamma(u)|$; we write $d(u)$ instead of $d_G(u)$ if the graph G is understood. However, somewhat unusually, for $U \subset V(G)$, we set $\widehat{\Gamma}(U) = |\cap_{v\in U}\Gamma(v)|$ and $\widehat{d}(U) = \left|\widehat{\Gamma}(U)\right|$.

We write $T_r(n)$ for the r-chromatic Turán graph on n vertices and $t_r(n)$ for the number of its edges.

[*]Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA
[†]Trinity College, Cambridge CB2 1TQ, UK
[‡]Research supported in part by DARPA grant F33615-01-C-1900.

For every $r \geq 2$ and every graph G, let $\Delta_r(G)$ be the maximum of the sum of degrees of the vertices of an r-clique, as in the abstract. If G has no r-cliques, we set $\Delta_r(G) = 0$. Furthermore, let

$$\Delta_r(n,m) = \min_{G=G(n,m)} \Delta_r(G).$$

Since $T_r(n)$ is a K_{r+1}-free graph, it follows that $\Delta_r(n,m) = 0$ for $m \leq t_{r-1}(n)$. In 1975 Bollobás and Erdős [2] conjectured that for every $r \geq 2$, if $m \geq t_r(n)$, then

$$\Delta_r(n,m) \geq \frac{2rm}{n}. \tag{2}$$

Edwards [3], [4] proved (2) under the weaker condition $m > (r-1)n^2/2r$; he also proved that the conjecture holds for $2 \leq r \leq 8$ and $n \geq r^2$. Later Faudree [7] proved the conjecture for any $r \geq 2$ and $n > r^2(r-1)/4$.

For $t_{r-1}(n) < m < t_r(n)$ the value of $\Delta_r(n,m)$ is essentially unknown even for $r = 3$ (see [5], [6] and [7] for partial results.) A construction due to Erdős and Faudree (see [7], Theorem 2) shows that, for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $t_{r-1}(n) < m < t_r(n) - \delta n^2$ then

$$\Delta_r(n,m) \leq (1-\varepsilon)\frac{2rm}{n}.$$

The construction is determined by two appropriately chosen parameters a and d and represents a complete $(r-1)$-partite graph with $(r-2)$ chromatic classes of size a and a d-regular bipartite graph inserted in the last chromatic class.

In this note we prove a stronger form of (2) for every r and n. Furthermore, we prove that $\Delta_r(n,m)$ is "stable" as m approaches $t_r(n)$. More precisely, for every $\varepsilon > 0$, there is $\delta > 0$ such that if $m > t_r(n) - \delta n^2$ then

$$\Delta_r(n,m) \geq (1-\varepsilon)\frac{2rm}{n}$$

for n sufficiently large.

1.1 Preliminary observations

If $M_1, ..., M_k$ are subsets of a (finite) set V then

$$\left|\cap_{i=1}^{k} M_i\right| \geq \sum_{i=1}^{k} |M_i| - (k-1)|V|. \tag{3}$$

The size $t_r(n)$ of the Turán graph $T_r(n)$ is given by

$$t_r(n) = \frac{r-1}{2r}n^2 - \frac{s}{2}\left(1 - \frac{s}{r}\right).$$

where s is the remainder of n modulo r. Hence,

$$\frac{r-1}{2r}n^2 - \frac{r}{8} \leq t_r(n) \leq \frac{r-1}{2r}n^2. \tag{4}$$

2 A greedy algorithm

In what follows we shall identify a clique with its vertex set.

Faudree [7] introduced the following algorithm $\mathcal{P}$ to construct a clique $\{v_1, ..., v_k\}$ in a graph G:

Step 1: v_1 is a vertex of maximum degree in G;

Step 2: having selected $v_1, ..., v_{i-1}$, if $\widehat{\Gamma}(v_1, ..., v_{i-1}) = \varnothing$ then set $k = i - 1$ and stop $\mathcal{P}$, otherwise $\mathcal{P}$ selects a vertex of maximum degree $v_i \in \widehat{\Gamma}(v_1, ..., v_{i-1})$ and step 2 is repeated again.

Faudree's main reason to introduce this algorithm was to prove Conjecture (2) for n sufficiently large, so he did not study $\mathcal{P}$ in great detail. In this section we shall establish some properties of $\mathcal{P}$ for their own sake. Later, in Section 3, we shall apply these results to prove an extension of (2) for every n.

Note that $\mathcal{P}$ need not construct a unique sequence. Sequences that can be constructed by $\mathcal{P}$ are called *$\mathcal{P}$-sequences;* the definition of $\mathcal{P}$ implies that $\widehat{\Gamma}(v_1...v_k) = \varnothing$ for every $\mathcal{P}$-sequence $v_1, ..., v_k$.

Theorem 1 *Let $r \geq 2$, $n \geq r$ and $m \geq t_r(n)$. Then every graph $G = G(n, m)$ is such that:*

(i) *every $\mathcal{P}$-sequence has at least r terms;*

(ii) *for every $\mathcal{P}$-sequence $v_1, ..., v_r, ...$,*

$$\sum_{i=1}^{r} d(v_i) \geq (r-1)n; \tag{5}$$

(iii) *if equality holds in (5) for some $\mathcal{P}$-sequence $v_1, ..., v_r, ...$ then $m = t_r(n)$.*

Proof Without loss of generality we may assume that $\mathcal{P}$ constructs exactly the vertices $1, ..., k$ and hence $d(1) \geq ... \geq d(k)$.

Proof of (i) and (ii) To prove *(i)* we have to show that $k \geq r$. For every $i = 1, ..., k$, let $M_i = \Gamma(i)$; clearly,

$$\sum_{i=1}^{k} d(i) \leq (q-1)n,$$

since, otherwise, (3) implies that $\widehat{\Gamma}(v_1...v_k) \neq \varnothing$, and so $1, ..., k$ is not a $\mathcal{P}$-sequence, contradicting the choice of k. Suppose $k < r$, and let q be the smallest integer such that the inequality

$$\sum_{i=1}^{h} d(i) > (h-1)n \tag{6}$$

holds for $h = 1, ..., q-1$, while

$$\sum_{i=1}^{q} d(i) \leq (q-1)n. \tag{7}$$

Clearly, $1 < q \leq k$.

Partition $V = \cup_{i=1}^{q} V_i$, so that

$$\begin{aligned} V_1 &= V \backslash \Gamma(1), \\ V_i &= \widehat{\Gamma}([i-1]) \backslash \widehat{\Gamma}([i]) \quad \text{for} \quad i = 2, ..., q-1, \\ V_q &= \widehat{\Gamma}([q-1]). \end{aligned}$$

We have

$$\begin{aligned} 2m &= \sum_{j \in V} d(j) = \sum_{h=1}^{q} \sum_{j \in V_h} d(j) \leq \sum_{i=1}^{q} d(i) |V_i| \\ &= d(1)(n - d(1)) + \sum_{i=2}^{q-1} d(i) \left(\widehat{d}([i-1]) - \widehat{d}([i]) \right) + d(q) \widehat{d}([q-1]) \\ &= d(1) n + \sum_{i=1}^{q-1} \widehat{d}([i]) (d(i+1) - d(i)). \end{aligned} \tag{8}$$

For every $i \in [q-1]$, set $k_i = n - d(i)$ and let $k_q = n - (k_1 + ... + k_{q-1})$. Clearly, $k_i > 0$ for every $i \in [q]$; also, $k_1 + ... + k_q = n$.

Furthermore, for every $h \in [q-2]$, applying (3) with $M_i = \Gamma(i)$, $i \in [h]$, and (6), we see that,

$$\widehat{d}([h]) = \left| \widehat{\Gamma}([h]) \right| \geq \sum_{i=1}^{h} d(i) - (h-1) n = n - \sum_{i=1}^{h} k_i > 0.$$

Hence, by $d(h+1) \leq d(h)$, it follows that

$$\widehat{d}([h]) (d(h+1) - d(h)) \leq \left(n - \sum_{i=1}^{h} k_i \right) (d(h+1) - d(h)). \tag{9}$$

Since, from (7), we have

$$d(q) \leq (q-1) n - \sum_{i=1}^{q-1} d(i) = \sum_{i=1}^{q-1} k_i, \tag{10}$$

in view of (9) with $h = q-1$, it follows that

$$\begin{aligned} \widehat{d}([q-1]) (d(q) - d(q-1)) &\leq \left(n - \sum_{i=1}^{q-1} k_i \right) (d(q) - d(q-1)) \\ &\leq \left(n - \sum_{i=1}^{q-1} k_i \right) \left(\sum_{i=1}^{q-1} k_i - d(q-1) \right). \end{aligned}$$

Recalling (8) and (9), this inequality implies that

$$
\begin{aligned}
2m \leq nd(1) &+ \sum_{h=1}^{q-2}\left(n-\sum_{i=1}^{h}k_i\right)(d(h+1)-d(h)) \\
&+\left(n-\sum_{i=1}^{q-1}k_i\right)\left(\sum_{i=1}^{q-1}k_i-d(q-1)\right).
\end{aligned}
$$

Dividing by 2 and rearranging the right-hand side, we obtain

$$
m \leq \left(n-\sum_{i=1}^{q-1}k_i\right)\left(\sum_{i=1}^{q-1}k_i\right)+\sum_{1\leq i<j\leq q-1}k_ik_j=\sum_{1\leq i<j\leq q}k_ik_j. \tag{11}
$$

Note that

$$
\sum_{1\leq i<j\leq q}k_ik_j=e(K(k_1,...,k_q)).
$$

Given n and $k_1+...+k_q=n$, the value $e(K(k_1,...,k_q))$ attains its maximum if and only if all k_i differ by at most 1, that is to say, when $K(k_1,...,k_q)$ is exactly the Turán graph $T_q(n)$. Hence, the inequality $m \geq t_r(n)$ and (11) imply

$$
t_r(n) \leq m \leq e(K(k_1,...,k_q)) \leq t_q(n). \tag{12}
$$

Since $q<r\leq n$ implies $t_q(n)<t_r(n)$, contradicting (12), the proof of *(i)* is complete.

To prove *(ii)* suppose (5) fails, i.e.,

$$
\sum_{i=1}^{r}d(i)<(r-1)n.
$$

Hence, (10) holds with a strict inequality and so, the proof of (12) gives $t_r(n)<t_r(n)$. This contradiction completes the proof of *(ii)*.

Proof of (iii) Suppose that for some $\mathcal{P}$-sequence $v_1,...,v_r,...$ equality holds in (5). We may and shall assume that $v_1,...,v_r=1,...,r$, i.e.,

$$
\sum_{i=1}^{r}d(i)=(r-1)n.
$$

Following the arguments in the proof of *(i)* and *(ii)*, from (12) we conclude that

$$
t_r(n)\leq m\leq t_r(n).
$$

and this completes the proof. □

3 Degree sums in cliques

In this section we turn to the problem of finding $\Delta_r(n,m)$ for $m \geq t_r(n)$. We shall apply Theorem 1 to prove that every graph $G = G(n,m)$ with $m \geq t_r(n)$ contains an r-clique R with

$$\sum_{i\in R} d(i) \geq \frac{2rm}{n}. \tag{13}$$

As proved by Faudree [7], the required r-clique R may be constructed by the algorithm $\mathcal{P}$. Note that the assertion is trivial for regular graphs; as we shall show, if G is not regular, we may demand strict inequality in (13).

Theorem 2 *Let $r \geq 2$, $n \geq r$, $m \geq t_r(n)$ and let $G = G(n,m)$ be a graph which is not regular. Then there exists a $\mathcal{P}$-sequence $v_1, ..., v_r, ...$ of at least r terms such that*

$$\sum_{i=1}^{r} d(v_i) > \frac{2rm}{n}.$$

Proof Part *(iii)* of Theorem 1 implies that for some $\mathcal{P}$-sequence, say $1, ..., r, ...$, we have

$$\sum_{i=1}^{r} d(i) > (r-1)n.$$

Since $d(i) < n$, we immediately obtain

$$\sum_{i=1}^{s} d(i) > (s-1)n \tag{14}$$

for every $s \in [r]$.

The rest of the proof consists of two parts: In part *(a)* we find an upper bound for m in terms of $\sum_{i=1}^{r} d(i)$ and $\sum_{i=1}^{r} d^2(i)$. Then, in part *(b)*, we prove that

$$\frac{1}{r}\sum_{i=1}^{r} d(i) \geq \frac{2m}{n},$$

and show that if equality holds then G is regular.

(a) Partition the set V into r sets $V = V_1 \cup ... \cup V_r$, where,

$$V_1 = V \backslash \Gamma(1),$$
$$V_i = \widehat{\Gamma}([i-1]) \backslash \widehat{\Gamma}([i]) \text{ for } i = 2, .., r-1,$$
$$V_r = \widehat{\Gamma}([r-1]).$$

We have,

$$2m = \sum_{i\in V} d(i) = \sum_{h=1}^{r}\sum_{j\in V_h} d(j) \le \sum_{i=1}^{r} d(i)\,|V_i|$$
$$= \sum_{i=1}^{r-1} (d(i) - d(r))\,|V_i| + nd(r) \tag{15}$$

Clearly, for every $i \in [r-1]$, from (3), we have

$$\left|\widehat{\Gamma}([i+1])\right| \ge \left|\widehat{\Gamma}([i])\right| + |\Gamma(i+1)| - n = \left|\widehat{\Gamma}([i])\right| + d(i+1) - n$$

and hence, $|V_i| \le n - d(i)$ holds for every $i \in [r-1]$. Estimating $|V_i|$ in (15) we obtain

$$2m \le \sum_{i=1}^{r-1} (d(i) - d(r))(n - d(i)) + nd(r)$$
$$= n\sum_{i=1}^{r} d(i) - \sum_{i=1}^{r} d^2(i) + d(r)\left(\sum_{i=1}^{r} d(i) - n(r-1)\right).$$

(b) Let $S_r = \sum_{i=1}^{r} d(i)$. From $d(r) \le S_r/r$ and Cauchy's inequality we deduce

$$2m \le nS_r - \sum_{i=1}^{r} d^2(i) + \frac{S_r}{r}(S_r - (r-1)n)$$
$$\le nS_r - \frac{1}{r}(S_r)^2 + \frac{S_r}{r}(S_r - (r-1)n) \le \frac{nS_r}{r},$$

and so,

$$\sum_{i=1}^{r} d(i) \ge \frac{2rm}{n}. \tag{16}$$

To complete the proof suppose we have an equality in (16). This implies that

$$\sum_{i=1}^{r} d^2(i) = \frac{1}{r}\left(\sum_{i=1}^{r} d(i)\right)^2$$

and so, $d(1) = ... = d(r)$. Therefore, the maximum degree $d(1)$ equals the average degree $2m/n$, contradicting the assumption that G is not regular. □

Since for every $m \ge t_r(n)$ there is a graph $G = G(n,m)$ whose degrees differ by at most 1, we obtain the following bounds on $\Delta_r(n,m)$.

Corollary 1 *For every $m \ge t_r(n)$*

$$\frac{2rm}{n} \le \Delta_r(n,m) < \frac{2rm}{n} + r.$$

4 Stability of $\Delta_r(n,m)$ as m approaches $t_r(n)$

It is known that inequality (2) is far from being true if $m \leq t_r(n) - \varepsilon n$ for some $\varepsilon > 0$ (e.g., see [7]). However, it turns out that, as m approaches $t_r(n)$, the function $\Delta_r(n,m)$ approaches $2rm/n$. More precisely, the following stability result holds.

Theorem 3 *For every $\varepsilon > 0$ there exist $n_0 = n_0(\varepsilon)$ and $\delta = \delta(\varepsilon) > 0$ such that if $m > t_r(n) - \delta n^2$ then*

$$\Delta_r(n,m) > (1-\varepsilon)\frac{2rm}{n}$$

for all $n > n_0$.

Proof Without loss of generality we may assume that

$$0 < \varepsilon < \frac{2}{r(r+1)}.$$

Set

$$\delta = \delta(\varepsilon) = \frac{1}{32}\varepsilon^2.$$

If $m \geq t_r(n)$, the assertion follows from Theorem 2, hence we may assume that

$$\frac{2rm}{n} < \frac{2rt_r(n)}{n} \leq (r-1)n.$$

Clearly, our theorem follows if we show that $m > t_r(n) - \delta n^2$ implies

$$\Delta_r(n,m) > (1-\varepsilon)(r-1)n \tag{17}$$

for n sufficiently large.

Suppose the graph $G = G(n,m)$ satisfies $m > t_r(n) - \delta n^2$. By (4), if n is large enough,

$$m > t_r(n) - \delta n^2 > \left(\frac{r-1}{2r} - \delta\right)n^2 - \frac{r}{8} \geq \left(\frac{r-1}{2r} - 2\delta\right)n^2. \tag{18}$$

Let $M_\varepsilon \subset V$ be defined as

$$M_\varepsilon = \left\{u : d(u) \leq \left(\frac{r-1}{r} - \frac{\varepsilon}{2}\right)n\right\}.$$

The rest of the proof consists of two parts. In part *(a)* we shall show that $|M_\varepsilon| < \varepsilon n$, and in part *(b)* we shall show that the subgraph induced by $V \backslash M_\varepsilon$ contains an r-clique with large degree sum, proving (17).

(a) Our first goal is to show that $|M_\varepsilon| < \varepsilon n$. Indeed, assume the opposite and select an arbitrary $M' \subset M_\varepsilon$ satisfying

$$\left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)\varepsilon n < |M'| < \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)\varepsilon n. \tag{19}$$

Let G' be the subgraph of G induced by $V \backslash M'$. Then

$$e(G) = e(G') + e(M', V \backslash M') + e(M') \leq e(G') + \sum_{u \in M'} d(u) \qquad (20)$$
$$\leq e(G') + |M'| \left(\frac{r-1}{r} - \frac{\varepsilon}{2} \right) n.$$

Observe that second inequality of (19) implies

$$n - |M'| > (1 - \varepsilon) n.$$

Hence, if

$$e(G') \geq \frac{r-1}{2r} (n - |M'|)^2$$

then, applying Theorem 2 to the graph G', we see that

$$\Delta_r(G) \geq \Delta_r(G') \geq \frac{2re(G')}{n - |M'|} \geq (r-1)(n - |M'|) > (r-1)(1-\varepsilon) n,$$

and (17) follows. Therefore, we may assume

$$e(G') < \frac{r-1}{2r} (n - |M'|)^2.$$

Then, by (18) and (20),

$$\frac{r-1}{2r} (n - |M'|)^2 > e(G') > -|M'| \left(\frac{r-1}{r} - \frac{\varepsilon}{2} \right) n + \left(\frac{r-1}{2r} - 2\delta \right) n^2.$$

Setting $x = |M'| / n$, this shows that

$$\frac{r-1}{2r} (1-x)^2 + x \left(\frac{r-1}{r} - \frac{\varepsilon}{2} \right) - \left(\frac{r-1}{2r} - 2\delta \right) > 0,$$

which implies that

$$x^2 - \varepsilon x + 4\delta > 0.$$

Hence, either

$$|M'| > \left(\frac{\varepsilon - \sqrt{\varepsilon^2 - 16\delta}}{2} \right) n = \left(\frac{1}{2} - \frac{1}{2\sqrt{2}} \right) \varepsilon n$$

or

$$|M'| < \left(\frac{\varepsilon + \sqrt{\varepsilon^2 - 16\delta}}{2} \right) = \left(\frac{1}{2} + \frac{1}{2\sqrt{2}} \right) \varepsilon n,$$

contradicting (19). Therefore, $|M_\varepsilon| < \varepsilon n$, as claimed

(b) Let G_0 be the subgraph of G induced by $V \backslash M_\varepsilon$. By the definition of M_ε, if $u \in V \backslash M_\varepsilon$, then

$$d_G(u) > \left(\frac{r-1}{r} - \frac{\varepsilon}{2} \right) n,$$

and so

$$d_{G_0}(u) > \left(\frac{r-1}{r} - \frac{\varepsilon}{2}\right) n - |M_\varepsilon| > \frac{r-2}{r-1}\left(n - |M_\varepsilon|\right).$$

Hence, by Turán's theorem, G_0 contains an r-clique and, therefore,

$$\Delta_r(G) > r\left(\frac{r-1}{r} - \frac{\varepsilon}{2}\right) n \geq (1-\varepsilon)(r-1)\, n,$$

proving (17) and completing the proof of our theorem. □

Acknowledgement. The authors are grateful to Prof. D. Todorov for pointing out a fallacy in an earlier version of the proof of Theorem 2 and to the referee for his valuable suggestions.

Added on July 1st, 2005. The results of this paper were first presented in a seminar at Memphis University in February, 2002 and also form part of the second author's Ph.D. thesis [10], Ch. 7. The results in Theorems 1 and 2 were reproduced by Khadzhiivanov and Nenov in [8], [9].

References

[1] B. Bollobás, *Modern Graph Theory*, Graduate Texts in Mathematics 184, Springer Verlag, 1998, xiv–394pp.

[2] B. Bollobás and P. Erdős, Unsolved problems, *Proc. Fifth Brit. Comb. Conf. (Univ. Aberdeen, Aberdeen, 1975)*, Winnipeg, Util. Math. Publ., 678–680.

[3] C. Edwards, The largest vertex degree sum for a triangle in a graph, *Bull. Lond. Math. Soc.*, **9** (1977), 203–208.

[4] C. Edwards, Complete subgraphs with largest sum of vertex degrees, *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. I, Colloq. Math. Soc. János Bolyai,* **18**, North-Holland, Amsterdam-New York, 1978, pp. 293–306.

[5] P. Erdős and R. Laskar, On maximum chordal subgraph, Proceedings of the fourteenth Southeastern conference on combinatorics, graph theory and computing (Boca Raton, Fla., 1983). *Congr. Numer.* **39** (1983), 367–373.

[6] G. Fan, Degree sum for a triangle in a graph, *J. Graph Theory* **12** (1988), 249–263.

[7] R. Faudree, Complete subgraphs with large degree sums, *J. Graph Theory* **16** (1992), 327–334.

[8] N. Khadzhiivanov and N. Nenov, Sequences of maximal degree vertices in graphs, *Serdica Math. J.* **30** (2004), 95–102.

[9] N. Khadzhiivanov and N. Nenov, Saturated β-sequences in graphs, *C. R. Acad. Bulgare Sci.* **57** (2004), 49–54.

[10] V. Nikiforov, Stability results in extremal graph theory, PhD thesis, Inst. of Math. and Inform., Bul. Acad. Sci., Sofia.