CONNECTED GRAPHS WITHOUT LONG PATHS

P.N. BALISTER[1], E. GYŐRI[1,2], J. LEHEL[1,3], AND R.H. SCHELP[1]

Abstract. We determine the maximum number of edges in a connected graph with n vertices if it contains no path with $k+1$ vertices. We also determine the extremal graphs.

Dedicated to Miklós Simonovits on his 60th birthday

1. Introduction

A problem, first considered by Erdős and Gallai [2], was to determine the maximum number of edges in any graph on n vertices if it contains no path with $k+1$ vertices. This maximum number, $\mathrm{ext}(n, P_{k+1})$, is called the extremal number for the path P_{k+1}. Erdős and Gallai proved the following theorem, which was one of the earliest extremal results in graph theory.

Theorem 1.1 ([2]). *For every $k \geq 0$, $\mathrm{ext}(n, P_{k+1}) \leq \frac{1}{2}(k-1)n$ with equality if and only if $n = kt$, in which case the extremal graph is $\bigcup_{i=1}^{t} K_k$.*

In 1975 this result was improved by Faudree and Schelp [3], determining $\mathrm{ext}(n, P_{k+1})$ for all $n > k > 0$ as well as the corresponding extremal graphs. This is given by

Theorem 1.2 ([3]). *If G is a graph with $|V(G)| = kt + r$, $0 \leq r < k$, containing no path with $k+1$ vertices then $|E(G)| \leq t\binom{k}{2} + \binom{r}{2}$ with equality if and only if G is either (i) $(\bigcup_{i=1}^{t} K_k) \cup K_r$, or (ii) $(\bigcup_{i=1}^{t-l-1} K_k) \cup (K_{(k-1)/2} + \overline{K}_{(k+1)/2+lk+r})$ for some l, $0 \leq l < t$, when k is odd, $t > 0$, and $r = (k \pm 1)/2$.*

We use $\overline{G}$ to denote the edge complement of a graph G, $G \cup H$ to denote the vertex-disjoint union of graphs G and H, and $G + H$ to denote the join of G and H, defined as $G \cup H$ together with all edges between G and H.

In this paper we consider the extremal problem for P_{k+1} taken over all *connected* graphs. We determine this number as well as the extremal graphs. These extremal graphs are particular examples of graphs of the following form.

Definition. *For $n \geq k > 2s > 0$ let $G_{n,k,s} = (K_{k-2s} \cup \overline{K}_{n-k+s}) + K_s$ (see Figure 1).*

Note that $|E(G_{n,k,s})| = \binom{k-s}{2} + s(n-k+s)$ and since $k > 2s$, $G_{n,k,s}$ contains no P_{k+1}.

Date: January 29, 2005.

[1]Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152-3240.

[2]On leave from the Rényi Institute, Budapest. Research is partially supported by OTKA Grant No. 37204.

[3]On leave from the Computer and Automation Research Institute of the Hungarian Academy of Sciences.

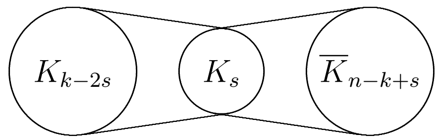

FIGURE 1. The graph $G_{n,k,s}$.

The central result of this paper is:

Theorem 1.3. *Let G be a connected graph on n vertices containing no path on $k+1$ vertices, $n > k \geq 3$. Then $|E(G)|$ is bounded above by the maximum of $\binom{k-1}{2} + (n-k+1)$ and $\binom{\lceil (k+1)/2 \rceil}{2} + \lfloor \frac{k-1}{2} \rfloor (n - \lceil \frac{k+1}{2} \rceil)$. If equality occurs then G is either $G_{n,k,1}$ or $G_{n,k,\lfloor (k-1)/2 \rfloor}$.*

A simple calculation shows that for $k > 4$ the maximum edge count is given by the first expression for $n \leq n_c$, and by the second for $n \geq n_c$, where

$$n_c = \begin{cases} (5k-5)/4 & \text{if } k \text{ is odd;} \\ (5k-2)/4 & \text{if } k \text{ is even.} \end{cases} \tag{1.1}$$

For $k = 3$ or 4 the bounds and extremal graphs are equal for all n.

The extremal numbers, but not the extremal graphs, were previously obtained in [4]. The proof given there uses the extremal number for a cycle of length $\geq k$ in any 2-connected graph by forming the 2-connected graph G' by joining a new vertex to all the vertices of G. This proof does not lend itself to finding the extremal graphs. The proof given here, however, is constructive, giving both the extremal numbers and the extremal graphs.

Note that in the statement of Theorem 1.3, the second class of extremal graphs $G_{n,k,s}$ satisfy $k - 2s = 2$ for k even and $k - 2s = 1$ for k odd. In particular, when k is odd and $n = k + \frac{k-1}{2}$, we have $G_{n,k,\lfloor (k-1)/2 \rfloor} = K_{(k-1)/2} + \overline{K}_k$. Thus Theorem 1.2 shows that Theorem 1.3 holds when k is odd and $n = k + \frac{k-1}{2}$. With this fact it is not difficult to prove the case where k is odd and $n \geq k + \frac{k-1}{2}$. Unfortunately proving the remainder of this theorem is somewhat harder and will require several lemmas. One result that will also be used generalizes a result of Erdős and Gallai [2] and appears in [1].

Lemma 1.4 ([1])**.** *Let G be a graph and for each $v \in V(G)$ let p_v be the number of edges in the longest path in G starting at v. Then $|E(G)| \leq \frac{1}{2} \sum_{v \in V(G)} p_v$, with equality if and only if G is a disjoint union of complete graphs.*

The strategy of the proof of Theorem 1.3 is to take a longest path P in G and divide the vertices into two sets, $V(P)$ and $Y = V(G) \setminus V(P)$. For each $v \in Y$ we bound the number of edges in $G[V(P)]$ as a function of the number of edges from v to P and the length of the longest path in $G[Y]$ starting at v. Combining these bounds for all $v \in Y$ and using Lemma 1.4 then gives the result.

2. Proof of Theorem 1.3

A key part of the proof is to analyze the case when a longest path in G is of length k and misses precisely one vertex, so that $|V(G)| = k+1$. The next two lemmas deal with this situation.

Lemma 2.1. *Suppose S is a set of $s+1$ independent vertices in a graph G of order $k+1$, with the degree of each $w \in S$ at least s in G and suppose $G \setminus S$ is complete and G contains no Hamiltonian path. Then the neighborhoods of each $w \in S$ are the same and $G = G_{k+1,k,s}$.*

Proof. Consider the set of all paths in G starting at a fixed vertex $w_1 \in S$. Of these, consider the set $\mathcal{P}$ of paths that contain the maximal number of vertices of S. Pick a path $P \in \mathcal{P}$ of maximal length. We shall show that if $G \neq G_{k+1,k,s}$ then P is a Hamiltonian path. Suppose first that P does not contain some vertex $w \in S$. Let $u_1, \dots, u_s$ be s neighbors of w.

Case A: One of the vertices u_i occurs in P after all the vertices of S in P.

In this case we can change the path P after u_i, removing subsequent vertices and adding w so that it ends with $u_i w$. This new path contains more vertices of S, contradicting the assumption that $P \in \mathcal{P}$.

Case B: There are neighbors u_i and u_j of w on P that are not separated by a vertex of S on P.

In this case we delete the vertices of the path P between u_i and u_j and insert w so that the path goes $u_i w u_j$. This uses one more vertex of S, contradicting the assumption that $P \in \mathcal{P}$.

Case C: Some vertex u_i does not lie in the path P.

The path P must end with a vertex $w' \in S$ (otherwise we could extend P by adding u_i). Let the path end with the edge uw'. Then $u \in G \setminus S$ since S is an independent set. Now replacing this edge with uu_iw gives a longer path meeting the same number of vertices of S. This contradicts the choice of $P \in \mathcal{P}$.

Since none of the cases A, B, or C occur, at least $s+1$ vertices of S lie on the path P. Hence P contains all the vertices of S. If P ends at a vertex of $G \setminus S$ then we can append the remaining vertices of $G \setminus S$ to get a Hamiltonian path. Similarly, if there are two neighboring vertices uu' of P that both lie in $G \setminus S$ then we can insert the remaining vertices of $G \setminus S$ between them. Otherwise P must be of the form $w_1 u_1 w_2 u_2 \dots w_s u_s w_{s+1}$, $w_i \in S$, $u_i \notin S$. If the neighborhoods of the w_i are not all equal to $\{u_1, \dots, u_s\}$, there must be some $u \in G \setminus S$ not equal to any of the u_i which is a neighbor of some w_j. We can now insert u after w_j in P to get a longer path, contradicting the choice of P.

Hence if P is not a Hamiltonian path then the neighborhoods of every $w_i \in S$ must be exactly $\{u_1, \dots, u_s\}$ and so $G = G_{k+1,k,s}$. □

Define for $s_0 < k/2$ the function

$$h(k, s_0) = \max_{s_0 \le s < k/2} \left\{ \binom{k-s}{2} + s(s+1) \right\}. \tag{2.1}$$

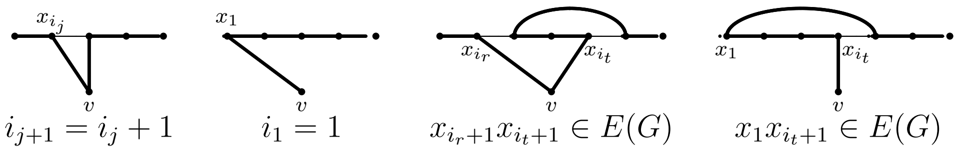

FIGURE 2. Cases for which G has a Hamiltonian path in Lemma 2.2.

This is just the maximum of the edge counts of the graphs $G_{k+1,k,s}$, $s_0 \le s < k/2$, occurring in Lemma 2.1.

The function $\binom{k-s}{2} + s(s+1)$ is convex in s, so

$$h(k, s_0) = \max\left\{\binom{k-s_0}{2} + s_0(s_0+1),\ \binom{k-s_m}{2} + s_m(s_m+1)\right\} \tag{2.2}$$

where $s_m = \lfloor \frac{k-1}{2} \rfloor$. We note for future reference that

$$h(k, s_0) - h(k-1, s_0) \ge \min_{s_0 \le s < (k-1)/2} \{k - s - 1\} = \lceil \tfrac{k}{2} \rceil \tag{2.3}$$

Lemma 2.2. *Suppose G is a connected graph on $k+1$ vertices with no Hamiltonian path, but with a path P with k vertices, and suppose the vertex $v \in V(G) \setminus V(P)$ has degree s_0. Then $|E(G)| \le h(k, s_0)$. If equality holds then $G = G_{k+1,k,s_0}$.*

Proof. The bound $h(k, s_0)$ is monotonic, decreasing with increasing s_0. Hence we may assume that v is a vertex of maximal degree such that $G - v$ has a Hamiltonian path. Let $s \ge s_0$ be the degree of v and $P = (x_1, \dots, x_k)$ a Hamiltonian path of $G - v$. Let the neighbors of v be $x_{i_1}, \dots, x_{i_s}$.

If $i_{j+1} = i_j + 1$ for any j then G contains the Hamiltonian path $(x_1, \dots, x_{i_j}, v, x_{i_{j+1}}, \dots, x_k)$. Similarly if $i_1 = 1$ or $i_s = k$ then G has a Hamiltonian path (see Figure 2). Therefore $i_1 \ge 2$, $i_{j+1} \ge i_j + 2$, and so $k - 1 \ge i_s \ge 2s$. Thus $s \le (k-1)/2$.

If there were an edge of the form $x_{i_r+1}x_{i_t+1}$ for some $t > r$, or an edge of the form $x_1x_{i_t+1}$ then G would have a Hamiltonian path (see Figure 2). Since $vx_1, vx_{i_t+1} \notin E(G)$, the set $\{v, x_1, x_{i_1+1}, \dots, x_{i_s+1}\}$ is an independent set of G of size $s+2$.

We shall now show that, without loss of generality, every $x_i \in \{x_1, x_{i_1+1}, \dots, x_{i_s+1}\}$ has degree d_i either at most s, or more than $\lceil \frac{k}{2} \rceil$ in G. To see this, let G' be the graph obtained from G by deleting the vertex x_i, and adding (if possible) the edge $x_{i-1}x_{i+1}$. Clearly G' has a path of length $k-1$ together with a vertex v of degree s (since $vx_i \notin E(G)$). If G' had a Hamiltonian path using the edge $x_{i-1}x_{i+1}$ then G would have a Hamiltonian path (replace $x_{i-1}x_{i+1}$ with $x_{i-1}x_ix_{i+1}$). On the other hand, if G' had a Hamiltonian path not using the edge $x_{i-1}x_{i+1}$ then $G - x_i$ would have a Hamiltonian path. In this case, the degree of x_i would be at most s by choice of v. Hence either $d_i \le s$ or G' has no Hamiltonian path. If G' has no Hamiltonian path, then by induction on k we can assume $|E(G')| \le h(k-1, s)$. (The base of the induction is the case $k = 2s+1$, where $G - x_i$ always has a Hamiltonian path.) Thus $|E(G)| \le h(k-1, s) + d_i$. If $d_i \le \lceil \frac{k}{2} \rceil$ then by (2.3) $|E(G)| \le h(k, s) \le h(k, s_0)$. Moreover, if we have equality, then $G' = G_{k,k-1,s}$, the edge $x_{i-1}x_{i+1}$ did not need to be added to make G', and G consists of G' with one more vertex x_i. Since G has no P_{k+1}

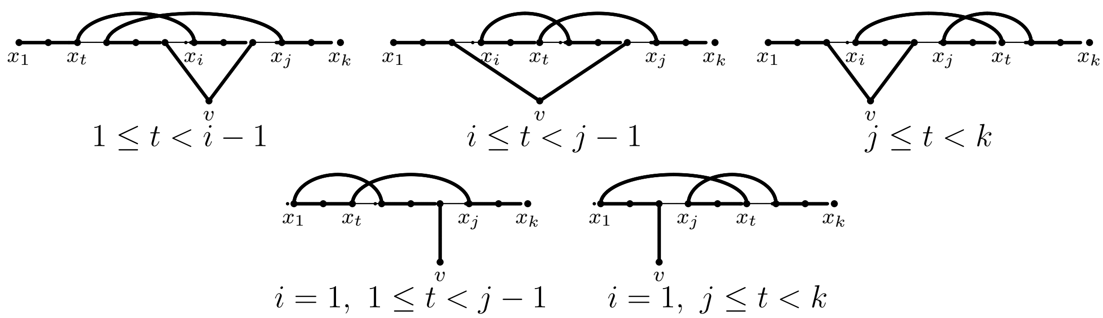

FIGURE 3. Cases for which G has a Hamiltonian path in Lemma 2.2.

it is clear that x_i cannot be joined to the independent set of size $s+1$ in G', thus G is a subgraph of $G_{k+1,k,s}$. Since $|E(G)| = h(k,s)$ we must have $G = G_{k+1,k,s}$ and since all the vertices $x \in G$ for which $G - x$ has a Hamiltonian path have the same degree s, $s = s_0$.

Now we shall show that at most one of these degrees d_i is greater than $\lceil \frac{k}{2} \rceil$. Suppose otherwise and assume $d_i, d_j > \lceil \frac{k}{2} \rceil$ with $i < j$, $i, j \in \{1, i_1+1, \ldots, i_s+1\}$. We shall deal with the case $i > 1$ first. If $1 \le t < i-1$ and $x_t x_i, x_{t+1} x_j \in E(G)$ then G has a Hamiltonian path (see Figure 3). Similarly when $i \le t < j-1$, $x_{t+1}x_i, x_t x_j \in E(G)$, or $j \le t < k$, $x_t x_i, x_{t+1} x_j \in E(G)$. Set

$$A = \{t : x_{t+\delta} x_i \in E(G)\} \qquad B = \{t : x_{t+1-\delta} x_j \in E(G)\}$$

where $\delta = 1$ if $i \le t < j$ and $\delta = 0$ otherwise. Since $i-1, k \notin B$, $j-1 \notin A$, the sets A and B are disjoint subsets of $\{1, \ldots, k\}$. Also, since $x_1 x_j, x_i x_j \notin E(G)$, $|A| = d_i$, $|B| = d_j$. Thus $d_i + d_j \le k$. This contradicts the assumption that $d_i, d_j > \lceil \frac{k}{2} \rceil$.

The case when $i = 1$ is similar. We cannot have $1 \le t < j-1$, $x_{t+1}x_1, x_t x_j \in E(G)$, or $j \le t < k$, $x_t x_1, x_{t+1} x_j \in E(G)$. Thus once again A and B (defined as above) are disjoint subsets of $\{1, \ldots, k\}$ since $k \notin B$ and $j-1 \notin A$. Also $x_1 x_j \notin E(G)$ so $|A| = d_i$, and $|B| = d_j$. Hence $d_i + d_j \le k$, contradicting the assumption that $d_i, d_j > \lceil \frac{k}{2} \rceil$.

Thus there is an independent subset S of $s+1$ elements of the set $\{v, x_1, x_{i_1+1}, \ldots, x_{i_s+1}\}$ all of whose elements have degree at most s in G. The maximum number of edges in G is therefore $\binom{k+1-|S|}{2} + s|S| = \binom{k-s}{2} + s(s+1) \le h(k,s) \le h(k,s_0)$.

For the extremal graph note that the maximum number of edges occurs when $G \setminus S$ is complete and the degree of every $x \in S$ is exactly s. Hence $G = G_{k+1,k,s}$ by Lemma 2.1. Since all the vertices $x \in G$ for which $G - x$ has a Hamiltonian path have the same degree s, $s = s_0$. □

Lemma 2.3. *Let G be a connected graph on n vertices with a path P with k vertices but no path with $k+1$ vertices and set $Y = V(G) \setminus V(P)$. Let $v \in Y$ be a vertex adjacent to $s \ge 1$ vertices of P and assume a longest path Q in Y starting at v has $p+1$ vertices, $p \ge 0$. Then $2s + 2p \le k-1$ and $|E(G[V(P)])| \le \binom{k}{2} - f_1(s,p)$ where*

$$f_1(s,p) = (p+1)(k-p-2) + \binom{s-1}{2} \qquad (s > 0). \tag{2.4}$$

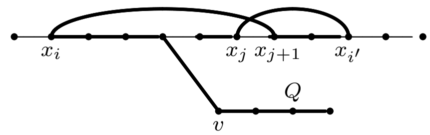

FIGURE 4. Path with $k+1$ vertices in Lemma 2.3 ($p = 3$).

Proof. Label the vertices $P = (x_1, \ldots, x_k)$ and let $x_{i_1}, \ldots, x_{i_s}$ be the vertices adjacent to v. The path $(x_k, \ldots, x_{i_1}, v) \cup Q$ must have at most k vertices, so $i_1 \geq p+2$. Similarly $i_s \leq k - (p+1)$. As in the proof of Lemma 2.2, $i_{r+1} \geq i_r + 2$. In particular $k - 2p - 3 \geq i_s - i_1 \geq 2(s-1)$, so $2s + 2p \leq k - 1$.

For each $i = 1, \ldots, p+1$ consider the pair of vertices x_i and $x_{i'} = x_{i+k-p-1}$. Suppose edges exist from x_i to x_{j+1} and x_j to $x_{i'}$ where $i < j < i' - 1$. Then there is a cycle $(x_i, x_{i+1}, \ldots, x_j, x_{i'}, \ldots, x_{j+1}, x_i)$ of length $k - p$. If $s \geq 1$ then at least one vertex of this cycle is joined to v and hence to the path Q with $p+1$ vertices in Y. This gives a path with $k+1$ vertices (see Figure 4). Hence at least one of the edges $x_i x_{j+1}$ and $x_j x_{i'}$ does not exist in G. Similarly the edges $x_i x_{i'}$, $x_i x_{i'+1}$, $\ldots$, $x_i x_k$, and $x_{i-1} x_{i'}$, $\ldots$, $x_1 x_{i'}$ do not exist. Thus we have a total of $k-2$ edges missing from either x_i or $x_{i'}$.

The total number of missing edges can be calculated as $(p+1)(k-2) - (p+1)p = (p+1)(k-p-2)$ since each pair $\{x_i, x_{i'}\}$ is incident to $k-2$ missing edges and if $i < j$ then at most two edges from $\{x_i, x_{i'}\}$ to $\{x_j, x_{j'}\}$ have been double counted ($x_i x_{j'}$ and $x_{i'} x_j$, but not $x_i x_j$ or $x_{i'} x_{j'}$).

As in Lemma 2.2 the edges $x_{i_r+1} x_{i_t+1}$ are missing in G. For $1 \leq r < t \leq s-1$ these vertices are not among the $\{x_i, x_{i'}\}$ considered above. Hence there are an additional $\binom{s-1}{2}$ missing edges not already counted. Hence we have at least $f_1(s,p) = (p+1)(k-p-2) + \binom{s-1}{2}$ missing edges from $G[V(P)]$. □

The following lemma deals with the case when $s = 0$.

Lemma 2.4. *Let G be a connected graph on n vertices with a path P with k vertices but no path with $k+1$ vertices and set $Y = V(G) \setminus V(P)$. Let $v \in Y$ be a vertex that is not adjacent to any vertex of P and assume a longest path Q in Y starting at v has $p+1$ vertices. Then $p \leq k-3$ and $|E(G[V(P)])| \leq \binom{k}{2} - f_1(0,p)$ where*

$$f_1(0,p) = \tfrac{1}{2}(k-1)(\lceil \tfrac{p}{2} \rceil + 1). \tag{2.5}$$

Proof. Since G is connected, there must be a path from v to P in G. Let v' be the last vertex of this path that lies in Y, so that v' is connected by a path in $G[Y]$ to v and is adjacent to $s \geq 1$ vertices of P. Let $p'+1$ be the number of vertices in the longest path in $G[Y]$ with v' as an endvertex. Then $p' \geq \lceil \frac{p}{2} \rceil$. Thus, by Lemma 2.3, $2p' + 2s \leq k-1$, so $\lceil \frac{p}{2} \rceil \leq p' \leq (k-3)/2$ and $p \leq k-3$. Also by Lemma 2.3, there are at least $f_1(s,p') \geq (p'+1)(k-p'-2)$ edges missing from $G[V(P)]$, but $p' \leq (k-3)/2$ gives $k - p' - 2 \geq (k-1)/2$, so there are at least $\frac{1}{2}(k-1)(p'+1) \geq \frac{1}{2}(k-1)(\lceil \frac{p}{2} \rceil + 1)$ missing edges from $G[V(P)]$. □

Proof of Theorem 1.3. For $k > 2s$ the number of edges in $G_{n,k,s}$ is $\binom{k-s}{2}+s(n-k+s)$, which is strictly increasing with k. Hence we may assume by induction on k that G contains a path with k vertices. For $k = 3$ we note that all connected graphs on $n \geq 4$ vertices contain a P_3.

Let P be a path with k vertices in G and let $Y = V(G) \setminus V(P)$. For each vertex $v \in Y$ let s_v be the number of vertices of P adjacent to v in G and let p_v be the number of edges in the longest path in $G[Y]$ starting at v. By Lemma 1.4 the number of edges in $G[Y]$ is bounded above by $\sum_{v\in Y} p_v/2$. By Lemma 2.2, the number of edges in $G[V(P)]$ is bounded above by $\binom{k}{2} - f_0(s_v)$ where

$$\begin{aligned} f_0(s) &= s + \tbinom{k}{2} - \max\left\{\tbinom{k-s}{2} + s(s+1),\ \tbinom{k-s_m}{2} + s_m(s_m+1) - 1\right\} \\ &= \min\left\{\tfrac{s}{2}(2k-3s-1),\ s + \tfrac{s_m}{2}(2k-3s_m-3)+1\right\}, \end{aligned} \tag{2.6}$$

The extra -1 in the second expression in the max comes from the fact that $|E(G[V(P) \cup \{v\}])| = \binom{k-s_m}{2} + s_m(s_m+1) > \binom{k-s}{2} + s(s+1)$ contradicts the extremal graph given by Lemma 2.2. By Lemma 2.3 or 2.4, the number of edges in $G[V(P)]$ is also bounded above $\binom{k}{2} - f_1(s_v, p_v)$, where

$$f_1(s,p) = \begin{cases} (p+1)(k-p-2) + \binom{s-1}{2} & \text{if } s > 0; \\ \frac{1}{2}(k-1)(\lceil \frac{p}{2} \rceil + 1) & \text{if } s = 0. \end{cases} \tag{2.7}$$

Hence

$$|E(G[V(P)])| \leq \tbinom{k}{2} - f(s_v, p_v), \qquad \text{where } f(s,p) = \max\{f_0(s), f_1(s,p)\}. \tag{2.8}$$

The total number of edges in G is therefore bounded by

$$\tbinom{k}{2} - \max_v f(s_v, p_v) + \sum_{v\in Y}(s_v + p_v/2). \tag{2.9}$$

This in turn is bounded above by the average $\frac{1}{n-k}\sum_{v\in Y} E_v$ of the values

$$E_v = \tbinom{k}{2} - f(s_v, p_v) + (n-k)(s_v + p_v/2). \tag{2.10}$$

This average is a linear function of n with slope given by the average value of $s_v + p_v/2$.

Claim 1. $1 \leq \frac{1}{n-k}\sum_{v\in Y}(s_v + p_v/2) \leq \lfloor (k-1)/2 \rfloor$.

Proof. For the upper bound we have $2s_v + 2p_v \leq k-1$ when $s_v > 0$, so $s_v + p_v/2 \leq s_v + p_v \leq \lfloor (k-1)/2 \rfloor$, and $p_v \leq k-3$ when $s_v = 0$, so $s_v + p_v/2 < \lfloor (k-1)/2 \rfloor$. For the lower bound, the only case when $s_v + p_v/2 < 1$ is when $s_v = 0$ and $p_v = 1$ (G is connected so we cannot have $s_v = p_v = 0$). In this case Y contains a component which is a star at v. At least one of the endvertices, u, of this star must be joined to P so has $p_u, s_u \geq 1$. Any other endvertex of the star must have $p_w \geq 2$. The average value of $s_x + p_x/2$ for this component is therefore at least 1. The claim now follows. □

The statement of the theorem requires us to bound the number of edges of G by the maximum of two linear functions of n, with slopes 1 and $\lfloor (k-1)/2 \rfloor$ respectively. Since we have a linear bound $\frac{1}{n-k}\sum_v E_v$ with slope between these two values, it is enough to prove the bound at the point $n = n_c$. Indeed, we shall show that E_v is bounded above at $n = n_c$ by the expressions in the statement of the theorem for every $v \in Y$ separately.

Claim 2.

$$\binom{k}{2} - f(s,p) + (n_c - k)(s + p/2) \leq \binom{k-1}{2} + (n_c - k + 1), \tag{2.11}$$

and if equality holds, $p = 0$ and $f(s,p) = s + \binom{k}{2} - h(k,s)$.

Proof. Substituting the values of n_c from (1.1) and rearranging gives $f(s,p) \geq g(s,p)$, where

$$g(s,p) = \begin{cases} \frac{1}{8}(k-2)(2s+p+6) & \text{if } k \text{ is even,} \\ \frac{1}{8}(k-5)(2s+p+6)+3 & \text{if } k \text{ is odd.} \end{cases} \tag{2.12}$$

Note that $\frac{1}{8}(k-2)(2s+p+6) \geq \frac{1}{8}(k-5)(2s+p+6)+3$ when $s \geq 1$ or $p \geq 2$.

Case A: $s = 0$.
In this case we use $f(s,p) \geq f_1(0,p) = \frac{1}{2}(k-1)(\lceil\frac{p}{2}\rceil+1)$. Now $p \geq 1$ so $4(\lceil\frac{p}{2}\rceil+1) \geq p+6$. Thus $f(s,p) > g(s,p)$ for all even k and for odd k when $p \geq 2$. For odd k and $p = 1$, $f_1(0,p) = k-1 > \frac{7}{8}(k-5)+3 = g(s,p)$. Thus in all cases we get a strict inequality.

Case B: $s \geq 1$, $p > 0$, $p \geq 2(s-1)/7$.
Use $f(s,p) \geq f_1(s,p) = (p+1)(k-p-2) + \frac{1}{2}(s-1)(s-2)$. Then

$$8(f-g) \geq (k-2)(7p-2s+2) - 8(p+1)p + 4(s-1)(s-2). \tag{2.13}$$

However $7p - 2s + 2 \geq 0$ and $k \geq 2s + 2p + 1$, so

$$\begin{aligned} 8(f-g) &\geq (2p+2s-1)(7p-2s+2) - 8(p+1)p + 4(s-1)(s-2) \\ &= (6p-1)p + 2(s-1)(5p-3) > 0. \end{aligned} \tag{2.14}$$

Hence we have strict inequality $f > g$ in this case also.

Case C: $s \geq 1$, $\binom{k-s}{2} + s(s+1) < \binom{k-s_m}{2} + s_m(s_m+1)$.
Use $f(s,p) \geq f_0(s) = s + \frac{s_m}{2}(2k - 3s_m - 3) + 1$. For even k,

$$\begin{aligned} 8(f-g) &\geq 8s + (k^2 - 2k + 8) - (k-2)(2s+p+6) \\ &= (k-6)(k-2-2s-2p) + (8+(k-10)p). \end{aligned} \tag{2.15}$$

Since $2s + 2p \leq k - 1$ and k is even, $2s + 2p \leq k - 2$. Hence if $k \geq 10$ we get a strict inequality. We also get a strict inequality when $4 \leq k < 10$ and $p = 0$. For $k < 10$ and $p > 0$ we have $s \leq 3$, so $2(s-1)/7 < 1$ and we are done by Case B. If k is odd then

$$\begin{aligned} 8(f-g) &\geq 8s + (k^2 - 4k + 11) - (k-5)(2s+p+6) - 24 \\ &= (k-9)(k-1-2s-2p) + (8+(k-13)p). \end{aligned} \tag{2.16}$$

Since $2s + 2p \leq k - 1$, if $k \geq 13$ we get a strict inequality. If $k < 13$ and $p > 0$ then $s \leq 4$, $2(s-1)/7 < 1$, so we are done using Case B. If $k < 13$ and $p = 0$, we get a strict inequality except for the cases $k = 5, 7$, $s = 1$ where we get equality. However, in both these cases, $f(s,p) = f_0(s) = \frac{s}{2}(2k - 3s - 1) = s + \frac{k}{2} - h(k,s)$.

Case D: $s \geq 1$, $\binom{k-s}{2} + s(s+1) \geq \binom{k-s_m}{2} + s_m(s_m+1)$.
Use $f(s,p) \geq f_0(s) = \frac{s}{2}(2k-3s-1)$. Since $s \geq 1$,

$$\begin{aligned} 8(f-g) &\geq 4s(2k-3s-1) - (k-2)(2s+p+6) \\ &= 6(s-1)(k-2s-2) - p(k-2). \end{aligned} \tag{2.17}$$

If k is odd we also have

$$\begin{aligned} 8(f-g) &\geq 4s(2k-3s-1) - (k-5)(2s+p+6) - 24 \\ &= 6(s-1)(k-2s-1) - p(k-5). \end{aligned} \tag{2.18}$$

Thus if $p = 0$ then we have $f \geq g$ with equality only when $s = 1, \lfloor \frac{k-1}{2} \rfloor$. If $p > 0$ we get a strict inequality unless $p(k-2) \geq 6(s-1)(k-2s-2)$. However, if $\binom{k-s}{2} + s(s+1) \geq \binom{k-s_m}{2} + s_m(s_m+1)$ then $s \leq \lfloor \frac{k+2}{6} \rfloor \leq \frac{5}{12}(k-2)$, so $p \geq s-1$ and we are done by Case B, unless $k = 4$. However in this case we also have $p \geq s-1$ since $s = 1$.

Hence we have a strict inequality for all pairs (s,p) except for some cases when $p = 0$ and $f(s,p) = \frac{s}{2}(2k-3s+1) = s + \frac{k}{2} - h(k,s)$ where we get equality. This completes the proof of Claim 2. □

For the extremal graph, note that we must have equality in Claim 2 for each $v \in Y$, so $p_v = 0$ and $|E(G[V(P) \cup \{v\}])| = h(k, s_v)$ for all $v \in Y$. Thus Y is an independent set and for each $v \in Y$, $G[V(P) \cup \{v\}] = G_{k+1,k,s_v}$ by Lemma 2.2. Hence $G[V(P)] = G_{k,k,s_v}$. Since these graphs are non-isomorphic for different values of s_v (e.g., the minimum degree is s_v), $s_v = s$ is a constant for all $v \in Y$. Let S be the independent set $\overline{K_s}$ in $G_{k,k,s}$. Then $S \cup \{v\}$ is independent for each $v \in Y$ (otherwise G would contain a path with $k+1$ vertices). Applying Lemma 2.1 to $G[V(P) \cup \{v\}]$ we see that the neighborhood of v is the same as the neighborhood of any $w \in S$. In particular it is the same for all $v \in Y$. Thus $G = G_{n,k,s}$. The number of edges is maximized when $s = 1$ or $\lfloor \frac{k-1}{2} \rfloor$, so the result follows. □

References

[1] P.N. Balister, B. Bollobás, O.M. Riordan, and R.H. Schelp, Graphs with large maximum degree containing no odd cycles of a given length *Journal of Combin. Theory B* **87** (2003), 366–373.
[2] P. Erdős and T. Gallai, On maximal paths and circuits in graphs, *Acta Math. Acad. Sci. Hungar.* **10** (1959), 337–356.
[3] R.J. Faudree and R.H. Schelp, Path Ramsey numbers in multicolorings, *Journal of Combin. Theory B* **19** (1975), 150–160.
[4] G.N. Kopylov, On maximal paths and cycles in a graph, *Soviet Math. Dokl.* Vol **18** (1977), 593–596.

Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152-3240
E-mail address, P.N. Balister: `balistep@msci.memphis.edu`
E-mail address, E. Győri: `gyori@renyi.hu`
E-mail address, J. Lehel: `jlehel@memphis.edu`
E-mail address, R.H. Schelp: `schelpr@msci.memphis.edu`