TAIWANESE JOURNAL OF MATHEMATICS
Vol. 22, No. 1, pp. 17–21, February 2018
DOI: 10.11650/tjm/8101

A Short Derivation for Turán Numbers of Paths

Gerard Jennhwa Chang

Abstract. This paper gives a short derivation for a result by Faudree and Schelp that the Turán number $\mathrm{ex}(n; P_{k+1})$ of a path of $k+1$ vertices is equal to $q\binom{k}{2} + \binom{r}{2}$, where $n = qk + r$ and $0 \le r < k$, with the set $\mathrm{EX}(n; P_{k+1})$ of extremal graphs determined.

As said by Bollobás [1] that extremal graph theory, in its strictest sense, is a branch of graph theory developed and loved by Hungarians. In particular, Paul Erdős is an important representative.

Extremal graph theory studies extremal (maximal or minimal) graphs which satisfy a certain property. Extremality can be taken with respect to different graph invariants, such as order, size or girth. More abstractly, it studies how global properties of a graph influence local substructures of the graph. For example, a simple extremal graph theory question is "which acyclic graphs on n vertices have the maximum number of edges?" The extremal graphs for this question are trees on n vertices, which have $n-1$ edges. More generally, a typical question is the following: given a graph property P, an invariant μ and a set of graphs $\mathcal{G}$, we wish to find the minimum value of m such that every graph in $\mathcal{G}$ which has μ larger than m possess property P. In the example above, P is the property of being cyclic, μ is the number of edges in the graph and $\mathcal{G}$ is the set of n-vertex graphs. Thus every graph on n vertices with more than $n-1$ edges must contain a cycle.

Extremal graph theory started in 1941 when Turán determined the maximum number of edges of an n-vertex graph that contains no complete graph K_k of k vertices as a subgraph. Although the special case of $k = 3$ was established by Mantel in 1907, we now usually called this kind of forbidden subgraph problems as Turán-type problems. More precisely, suppose $\mathcal{F}$ is a family of graphs, the *Turán number* $\mathrm{ex}(n; \mathcal{F})$ is the maximum number of edges of a graph of n vertices not containing a subgraph in $\mathcal{F}$. We use $\mathrm{EX}(n, \mathcal{F})$ to denote the set of all graphs of n vertices and $\mathrm{ex}(n; \mathcal{F})$ edges not containing a subgraph in $\mathcal{F}$. For the case of $\mathcal{F} = \{F\}$, we use $\mathrm{ex}(n; F)$ for $\mathrm{ex}(n; \mathcal{F})$ and $\mathrm{EX}(n; F)$ for $\mathrm{EX}(n; \mathcal{F})$.

Received September 29, 2016; Accepted April 25, 2017.
Communicated by Xuding Zhu.
2010 *Mathematics Subject Classification.* 05C35.
Key words and phrases. Turán number, extremal graph, path.
This research was supported in part by the Ministry of Science and Technology under grant NSC101-2115-M-002-005-MY3.

Let $T_{n,k}$ be the complete k-partite graph each of whose partite sets is of size $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$; and let $t_{n,k}$ be the number of edges of $T_{n,k}$. It can be verified that if $n = kq + r$ with $0 \le r < k$, then $t_{n,k} = (1 - 1/k)n^2/2 - r(k-r)/(2k)$. Then Turán's theorem says that $\mathrm{ex}(n; K_{k+1}) = t_{n,k}$ and $\mathrm{EX}(n; K_{k+1}) = \{T_{n,k}\}$.

After Turán's result, various Turán numbers have been studied for different graphs. Unlike the precise value of $\mathrm{ex}(n; K_{k+1})$, most of the results on $\mathrm{ex}(n; \mathcal{F})$ are of asymptotic type. For instance, Erdős and Stone [4] proved that $\mathrm{ex}(n; K_{k+1[s]}) = (1 - 1/k + o_s(1))n^2/2$, where $K_{k+1[s]}$ is the complete $(k+1)$-partite graph ech of whose partite set is of size s. As a consequence, we have Erdős and Simonovits's theorem [3] (now is often called Erdős-Stone-Simonovits Theorem) that if $k = \min_{F \in \mathcal{F}} \chi(F) - 1 > 0$, then $\mathrm{ex}(n; \mathcal{F}) = (1 - 1/k + o_{\mathcal{F}}(1))n^2/2$. For the case of $k = 1$, the above result is of no interest. In fact, Erdős [3] conjectured that for any bipartite graph F, there are constants c and a with $1 < a < 2$ such that $\mathrm{ex}(n; F) \sim cn^a$. This is still open now.

Even for the graph as simple as P_{k+1}, it is not easy to determine $\mathrm{ex}(n; P_{k+1})$. Let $n = kq + r$ with $0 \le r < k$. The graph $G_{n,k} := qK_k \cup K_r$ does not contain P_{k+1} as a subgraph and has $g_{n,k} := q\binom{k}{2} + \binom{r}{2}$ edges. Consequently, $\mathrm{ex}(n; P_{k+1}) \ge g_{n,k}$. In fact, this is an equality. However, the proof is not easy as $G_{n,k}$ is not the only graph in $\mathrm{EX}(n; P_{k+1})$. The first result on this line is Erdős and Gallai's theorem [2] that $\mathrm{ex}(n; P_{k+1}) \le (k-1)n/2$; and if the equality holds, then k is a factor of n and $\mathrm{EX}(n; P_{k+1}) = \{\frac{n}{k} K_k\}$. Notice that there is a gap $(k-r)r/2$ between $(k-1)n/2$ and $g_{n,k}$.

For more than one decade, this was the best result on P_{k+1}, until the set $\mathrm{EX}(n; P_{k+1})$ was completely determined by Faudree and Schelp [5]. Besides $G_{n,k}$, another kind of graph in $\mathrm{EX}(n; P_{k+1})$ is $G_{n,k,\ell} := \ell K_k \cup (K_{(k-1)/2} + \overline{K_{n-\ell k-(k-1)/2}})$, where $k \ge 3$ is odd, $r = (k \pm 1)/2$ and $0 \le \ell < q$. Figure 1 shows $G_{8,5}$ and $G_{8,5,0}$ with $r = (k+1)/2$.

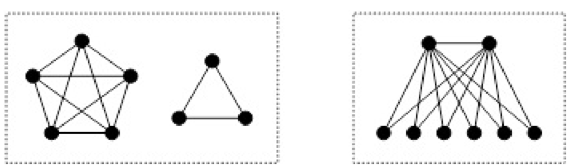

Figure 1: Graphs $G_{8,5}$ and $G_{8,5,0}$.

Faudree and Schelp [5] established the following theorem with a long proof. The purpose of this note is to simplify the proof.

Theorem 1. *Suppose G is a graph with n vertices, where $n = kq + r$ and $0 \le r < k$. If G does not contain P_{k+1} as a subgraph, then $|E(G)| \le g_{n,k}$. Furthermore, the equality holds if and only if $G = G_{n,k}$ or $G = G_{n,k,\ell}$ when $k \ge 3$ is odd, $r = (k \pm 1)/2$ and $0 \le \ell < q$.*

Proof. We shall prove the theorem by induction. The theorem is obvious when $n \le k$ or

$k = 1$. Suppose now $n > k > 1$ and the theorem holds for graphs G' with $n' + k' < n + k$.

Suppose G is not connected, say $G = G_1 \cup G_2$ with each G_i has $n_i = kq_i + r_i$ vertices, where $0 \le r_i < k$. By the induction hypothesis, $|E(G)| = |E(G_1)| + |E(G_2)| \le g_{n_1,r_1} + g_{n_2,r_2} \le g_{n,k}$. Assuming $r_1 \le r_2$, the last inequality follows from that $G_{n,k}$ can be obtained from $G_{n_1,k} \cup G_{n_2,k}$ by moving the vertices one by one from K_{r_1} to K_{r_2} until all vertices of K_{r_1} are removed or K_{r_2} becomes K_k. Notice that the number of edges increases at every movement. Furthermore, if $|E(G)| = g_{n,k}$, then each $|E(G_i)| = g_{n_i,k}$ and $r_1 = 0$, as no movement was done. Then $G_1 = G_{n_1,k} = q_1K_k$, and $G_2 = G_{n_2,k}$ or $G_{n_2,k,\ell}$. Therefore $G = G_{n,k}$ or $G_{n,k,\ell}$. Now we may assume that G is connected.

Suppose G contains no P_k. By the induction hypothesis, $|E(G)| \le g_{n,k-1} < g_{n,k}$. The last inequality follows from that $G_{n,k-1}$ can be obtained from $G_{n,k}$ by moving one vertex from each K_k to a smaller clique. Notice that the number of edges decreases at every movement. Now we may assume that G contains some P_k.

Claim. *If connected graph H has $p > h$ vertices and contains no P_{h+1} but one P_h called $P = (x_1, x_2, \ldots, x_h)$, then $\deg(x_1)+\deg(x_h) \le h-1$ and $H-x_1$ and $H-x_h$ are connected.*

Proof. The inequality $\deg(x_1) + \deg(x_h) \le h - 1$ follows from that for any $1 < j \le h$, at least one of x_1x_j and x_hx_{j-1} is not an edge, for otherwise $(x_1, x_2, \ldots, x_{j-1}, x_h, x_{h-1}, \ldots, x_j, x_1)$ is a C_h. Since $p > h$ and H is connected, this C_h together with some vertex outside it produce a P_{h+1}, a contradiction to the assumption. Finally, since the neighbors of x_1 and x_h are all in P, we have that $H - x_1$ and $H - x_h$ are connected. □

Suppose $n \ge 2k$. Since G contains no P_{k+1} but some P_k. Repeatedly applying the claim k times (the h used may decrease when starts from k), we have k vertices $z_1, z_2, \ldots, z_k$ and $k+1$ connected graphs $G_0 = G$ and $G_i = G_{i-1} - z_i$ for $1 \le i \le k$ such that $\deg_{G_{i-1}}(z_i) \le (k-1)/2$. By the induction hypothesis, $|E(G)| \le k(k-1)/2 + g_{n-k,k} = g_{n,k}$. Furthermore, if $|E(G)| = g_{n,k}$, then $\deg_{G_{i-1}}(z_i) = (k-1)/2$ for $1 \le i \le k$ and $|E(G_k)| = g_{n-k,k}$. As G_k is connected, $G_k = K_k$ or $G_{n-k,k,0}$. For the case of $G_k = K_k$, it together with z_k produces a P_{k+1}, a contradiction. Now suppose $G_k = G_{n-k,k,0}$. Let X be the set of vertices of $K_{(k-1)/2}$ and Y the remaining independent set in $G_{n-k,k,0}$. If every z_i is adjacent to all vertices in X, then $z_1, z_2, \ldots, z_k$ together with $G_{n-k,k,0}$ form $G_{n,k,0}$. Otherwise, there is a minimum indexed z_i adjacent to some vertex in $Y \cup \{z_1, z_2, \ldots, z_{i-1}\}$. Since every vertex of $Y \cup \{z_1, z_2, \ldots, z_{i-1}\}$ is adjacent to all vertices of X, it is the end vertex of a P_k, which together with z_i form a P_{k+1}, a contradiction. Now we may assume that $k < n < 2k$, or equivalently, $n = q + r$ with $0 < r < k$.

Suppose G has some vertex x of degree $\deg(x) \le r - 1$. By the induction hypothesis, $|E(G)| \le |E(G - x)| + r - 1 \le g_{n-1,k} + r - 1 = g_{n,k}$. Furthermore, if $|E(G)| = g_{n,k}$, then $|E(G - x)| = g_{n-k,k}$ and $\deg(x) = r - 1$. In this case, $G - x = G_{n-1,k} = K_k$ or

$G - x = G_{n-1,k,0}$. For the case of $G - x = K_k$, it together with x produces a P_{k+1}, a contradiction. For the case of $G - x = G_{n-1,k,0}$, we have $r - 1 = (k \pm 1)/2$. Same as the proof in the previous paragraph, x can only be adjacent to all vertices in X. So $r - 1 = (k-1)/2$ and then $G = G_{n,k,r}$ with $r = (k+1)/2$. Now we may assume that $\deg(x) \geq r$ for all vertices x in G.

Suppose G contains some C_{k-1} called $C = (x_1, x_2, \ldots, x_{k-1}, x_1)$, and the remaining $r+1$ vertices form a set X. The set X is independent, for otherwise if X has two adjacent vertices x and y, then C together with xy and a shortest path from xy to C produce a P_{k+1}, a contradiction. Let $S = \{x_i \in C : x_i \text{ is adjacent to some vertex in } X\}$ and $S' = \{x_{i-1} \in C : x_i \in S\}$ where $x_0 = x_{k-1}$. Suppose $x_{i-1} \in S'$ is adjacent to $x_{j-1} \in S'$ for some $i < j$. Choose $x \in X$ adjacent to x_i and $y \in X$ adjacent to x_j. Consider $P = (x, x_i, x_{i+1}, \ldots, x_{j-1}, x_{i-1}, x_{i-2}, \ldots, x_j, y)$. For the case of $x \neq y$, P is a P_{k+1}, a contradiction. For the case of $x = y$, P is a C_k which together with some vertex outside it produces a P_{k+1}, a contradiction. Therefore, S' is an independent set and so has no two vertices consecutively in C. Then $S \cap S' = \emptyset$ and $s := |S| = |S'| \leq (k-1)/2$ for which the equality holds only when k is odd. Therefore

$$|E(G)| \leq \binom{k-1}{2} - \binom{s}{2} + s(r+1) \leq \binom{k}{2} + \binom{r}{2} = g_{n,k}.$$

The second inequality follows from that twice the later minus the the former is equal to $2(k-1) - 4s + (s-r)^2 + (s-r)$. Notice that $s \leq (k-1)/2$. Also, as $s - r$ is an integer, $(s-r)^2 + (s-r) \geq 0$ with equality if and only if $r = s$ or $r = s+1$. The desired inequality then follows. Furthermore, if $|E(G)| = g_{n,k}$, then $s = (k-1)/2$ and so k is odd, $S \cup S' = V(C)$, $r = s$ or $r = s+1$, and every vertex in S is adjacent to all other vertices in $V(C) \cup X$. These give that $G = G_{n,k,0}$. Now we may assume that G has no C_{k-1}.

Having all the underlined conditions mentioned above, we now choose a P_k called $Q = (x_1, x_2, \ldots, x_k)$, and let all other r vertices form a set Y. Choose a vertex $y \in Y$ such that $T = \{x_j \in Q : yx_j \in E(G)\}$ has size t of maximum possible. Since G has no C_{k-1}, either $x_2 \notin T$ or $x_{k-1} \notin T$, and by symmetric we may assume that $x_{k-1} \notin T$. For any $x_j \in T$, $x_1 x_{j+1} \notin E(G)$ for otherwise $x_1 x_{j+1} \in E(G)$ would imply that $(y, x_j, x_{j-1}, \ldots, x_1, x_{j+1}, x_{j+2}, \ldots, x_k)$ is a P_{k+1}, a contradiction. Hence there are t such kind of non-edges using x_1 as an end vertex. Besides, there is an extra non-edge $x_1 x_k$, since $x_{k-1} \notin T$.

Since G has no P_{k+1}, all neighbors of x_1 and x_k are in Q. Let the sets $A = \{x_i \in Q : x_1 x_{i+1} \in E(G)\}$ and $B = \{x_i \in Q : x_{i-1} x_k \in E(G)\}$. If there is some $x_i \in A \cap B$, then $(x_1, x_2, \ldots, x_{i-1}, x_k, x_{k-1}, \ldots, x_{i+1}, x_1)$ is a C_{k-1}, a contradiction. Therefore, $A \cap B = \emptyset$ and so $|A \cup B| = |A| + |B| = \deg(x_1) + \deg(x_k) \geq 2r$. For $x_i \in A$, we may rearrange Q into the $P_k : (x_i, x_{i-1}, \ldots, x_1, x_{i+1}, x_{i+2}, \ldots, x_k)$. For $x_i \in B$, we may rearrange Q into the

$P_k : (x_i, x_{i+1}, \ldots, x_k, x_{i-1}, x_{i-2}, \ldots, x_1)$. Hence every $x_i \in A \cup B$ has t non-edges using x_i as an end vertex. Totally, there are at least $(2rt+1)/2$ non-edges between the vertices in Q. Hence

$$|E(G)| \leq \binom{k}{2} - \frac{2rt+1}{2} + rt + \binom{r}{2} < g_{n,k}.$$

The theorem is thus proved. □

References

[1] B. Bollobás, *Extremal Graph Theory*, Dover Publications, Mineola, NY, 2004.

[2] P. Erdős and T. Gallai, *On maximal paths and circuits of graphs*, Acta Math. Acad. Sci. Hungar **10** (1959), 337–356.

[3] P. Erdős and M. Simonovits, *A limit theorem in graph theory*, Studia Sci. Math. Hungar **1** (1966), 51–57.

[4] P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091.

[5] R. J. Faudree and R. H. Schelp, *Path Ramsey numbers in multicolorings*, J. Combinatrial Theory Ser. B **19** (1975), no. 2, 150–160.

Gerard Jennhwa Chang
Department of Mathematics, National Taiwan University, Taipei 10617, Taiwan
and
National Center for Theoretical Sciences, Mathematics Division, Second Floor, Astronomy-Mathematics Building, National Taiwan University, Taipei 10617, Taiwan
E-mail address: `gjchang@math.ntu.edu.tw`