Extremal Graphs with Bounded Densities of Small Subgraphs

Jerrold R. Griggs,[1,†] **Miklós Simonovits,**[2,‡] **and George Rubin Thomas**[3]

[1] DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTH CAROLINA
COLUMBIA, SC 29208 USA
Email address: griggs@math.sc.edu

[2] MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, HUNGARY

[3] 55 MERIDENE CRESCENT
LONDON, ONT N5X 2M1 CANADA

Received August 27, 1996; revised June 8, 1998

Abstract: Let $\mathrm{Ex}(n, k, \mu)$ denote the maximum number of edges of an n-vertex graph in which every subgraph of k vertices has at most μ edges. Here we summarize some known results of the problem of determining $\mathrm{Ex}(n, k, \mu)$, give simple proofs, and find some new estimates and extremal graphs. Besides proving new results, one of our main aims is to show how the classical Turán theory can be applied to such problems. The case $\mu = \binom{k}{2} - 1$ is the famous result of Turán.
 J Graph Theory 29: 185–207, 1998

Keywords: *extremal graphs, Turán's Theorem, Dirac's Theorem*

Correspondence to: Professor Griggs.
†Supported by Grants NSA/MSP/MDA904-92H3053 and -95H1024 and by NSF DMS-9701211.
‡Supported by Hungarian Research Grant OTKA 1909.
 CCC 0364-9024/98/030185-23

1. INTRODUCTION AND NOTATION

We consider undirected graphs G without loops and multiple edges. The set of vertices, the set of edges, and the chromatic number are denoted by $V(G), E(G)$, and $\chi(G)$, respectively. We denote the number of vertices (resp., edges) by $v(G)$ (resp., $e(G)$). The first subscript in the case of graphs indicates the number of vertices, e.g., C_k, P_k are the cycle and path graphs on k vertices. For $X \subseteq V(G), G[X]$ denotes the subgraph induced by X and $e(X)$ denotes the number of edges in it. Let $d_G(v)$ denote the degree of vertex v in G, and put

$$\delta(G) = \min_{v \in V} d_G(v) \text{ and } \Delta(G) = \max_{v \in V} d_G(v).$$

For vertex-disjoint graphs $G^1, \ldots, G^k$, their *product*, $\Pi_{i \leq k} G^i$, is the graph obtained by taking their vertex-disjoint copies and joining x, y when they belong to different G^i's. The product of two graphs G^1, G^2 is also denoted by $G^1 \otimes G^2$. The complement of G is denoted by $\bar{G}$.

Given a family $\mathcal{L}$ of forbidden graphs, what is the maximum number of edges a graph G_n, i.e., a graph on n vertices, can have without containing subgraphs from $\mathcal{L}$? Here ‘‘containing’’ means there is a copy of a member of $\mathcal{L}$, *not necessarily induced*. The maximum is denoted by $\text{ex}(n, \mathcal{L})$ and the $\mathcal{L}$-free graphs attaining this maximum are called *extremal graphs*. The family of extremal graphs is denoted by $\text{EX}(n, \mathcal{L})$.

The case $\mathcal{L} = \{K_k\}$ was solved in 1941 by Turán [34], who showed that the unique optimum is the graph $T_{n,k-1}$ described as follows: The *Turán graph* $T_{n,p}$ on n vertices and p classes is obtained by grouping the vertices as evenly as possible into p parts and joining two vertices by an edge if and only if they belong to different parts. The case $\text{ex}(n, \{K_3\}) = \lfloor n^2/4 \rfloor$ had been proved in 1907 by Mantel [25].

In the 1960s a whole new area, called Extremal Graph Theory, emerged around Turán's Theorem. One aim of this article is to exhibit the strength and usefulness of the general theory through a special interesting class $\mathcal{L}$.

The main question we investigate in this is is the following.

Dirac-type Extremal Problem. *Given the parameters k and μ, and the number of vertices n, determine the maximum number* $\text{Ex}(n, k, \mu)$ *of edges a graph G_n can have if no k-vertex subgraph of G_n has more than μ edges.*

Many people investigated this question, starting with Dirac [5] and Erdős, and continuing with Simonovits [29], B. Stechkin [33], and Abloncy (unpublished). Analogous problems for *hypergraphs* were investigated by Brown, Erdős, and T. Sós [3, 4], where the problems become much more involved, and sometimes extremely deep. One result illustrating this is due to Ruzsa and Szemerédi [26]. For more about Turán-type hypergraph results consult the surveys by Füredi [16] and Sidorenko [27].

Let $\mathcal{L}_{k,\mu}$ be the family of all graphs of k vertices having more than μ edges, so that

$$\mathrm{Ex}(n,k,\mu) = \mathrm{ex}(n,\mathcal{L}_{k,\mu}).$$

For $\mu := \binom{k}{2} - \lambda$, let $\mathcal{L}_{k,-\lambda}$ denote the family of graphs on k vertices with more than $\binom{k}{2} - \lambda$ edges. $\mathrm{EX}(n,k,\mu)$ is the family of extremal graphs for $\mathcal{L}_{k,\mu}$. Let $I(k,\lambda)$ denote the set of graphs in which every subgraph of k vertices has at least λ edges missing. The graphs G_n having maximum number of edges in $I(k,\lambda)$ for a fixed n are just the graphs in $\mathrm{EX}(n,k,\mu)$ for $\mu = \binom{k}{2} - \lambda$.

It is convenient[1] to denote the number of edges in the Turán graph $T_{n,p}$ by the function $t_p(n)$. Then $t_2(n) = \lfloor n^2/4 \rfloor$, and, in general,

$$t_p(n) = \left(1 - \frac{1}{p}\right)\binom{n}{2} + O(n).$$

Dirac's Theorem is a direct strengthening of Turán's Theorem.

Dirac's Theorem. *[5, Thm. 3] For $p \geq 1$, if $e(G_n) > e(T_{n,p})$, then G_n contains a subgraph consisting of K_{p+r+1} with at most r edges missing, for every r such that $0 \leq r \leq p-1$ and $n \geq p+r+1$.*

2. OVERVIEW OF KNOWN AND NEW RESULTS

2.1. Asymptotic Description of Ex(n, k, μ)

(a) The Kővári–T. Sós–Turán Theorem [23] asserts that $\mathrm{ex}(n, K_{a,b}) = O(n^{2-1/a})$. For $\mu < \lfloor k^2/4 \rfloor$ we can apply this result with $a = \lfloor k/2 \rfloor, b = \lceil k/2 \rceil$ to get that

$$\text{if } \mu < \left\lfloor \frac{k^2}{4} \right\rfloor, \quad \text{then} \quad \mathrm{Ex}(n,k,\mu) = O(n^{2-1/\lfloor k/2 \rfloor}) = O(n^{2-(1/\sqrt{\mu})}).$$

In most cases there are better exponents. We mention here only one result of Goldberg and Gurvich [18], when $\mathrm{Ex}(n,k,\mu)$ is linear in n. Consider the smallest case not covered by Dirac's Theorem, $\mathrm{Ex}(n,3,1)$: G contains no two intersecting edges, hence it is uniquely optimal to let G_n consist of $\lfloor \frac{n}{2} \rfloor$ disjoint edges. In general, it is not hard to find the extremum for $0 \leq \mu \leq k-2$ (see [18]). A proof of the corresponding result can be found also in [19], where the structure of the extremal graphs is also determined.

(b) The case $\mu = k-1$ is related to the well-known, difficult, unsolved problem of finding the maximum number of edges in graphs of girth exceeding k. The best

[1] However, we shall also continue to write $e(T_{n,p})$ when we wish to emphasize, not just this number, but its connection to the Turán graph. For we believe extremal graph theory should be made in terms of extremal graphs and extremal structures, and not so much in terms of formulas, whenever this is possible.

upper and lower bounds are due to Bondy and Simonovits [2] and to Lazebnik, Ustimenko, and Woldar [24], resp.,

$$\text{for} \quad k = 2s+1, \quad c_k n^{1+2/(3s-3+a)} < \text{ex}(n, \{C_3, C_4, \ldots, C_k\}) \\ \leq \text{Ex}(n, k, k-1) \leq c_k^* n^{1+1/\lfloor k/2 \rfloor}, \quad (1)$$

where $a = 0$ or 1 according as s is odd or even.[2]

(c) From now on, we assume that $\mu \geq \lfloor k^2/4 \rfloor$. Then $T_{n,2}$ contains no forbidden subgraphs, showing that $\text{Ex}(n, k, \mu) \geq t_2(n) = \lfloor n^2/4 \rfloor$. In this case, we always know the asymptotic behavior as $n \to \infty$.

Erdős and Simonovits (3) showed that, as a consequence of a 1946 result of Erdős and Stone [12], that the order of magnitude of $\text{ex}(n, \mathcal{L})$ depends only on the minimum chromatic number of the excluded subgraphs:

$$\lim_{n\to\infty} \frac{\text{ex}(n, \mathcal{L})}{\binom{n}{2}} = 1 - \frac{1}{p}, \quad (2)$$

where $p = p(\mathcal{L})$ is defined by

$$p(\mathcal{L}) = \min_{L \in \mathcal{L}} \chi(L) - 1. \quad (3)$$

Note that $t_2(k) < t_3(k) < \cdots < t_{k-1}(k) < t_k(k) = \binom{k}{2}$. For fixed k and μ, define $p \geq 2$ by $t_p(k) \leq \mu < t_{p+1}(k)$. Then we have

$$\text{Ex}(n, k, \mu) \geq e(T_{n,p}) = t_p(n).$$

For all graphs $L \in \mathcal{L}_{k,\mu}$, we have $e(L) > e(T_{k,p})$. Since $T_{k,p}$ has the most edges of any p-colorable graph on k vertices, it follows that $\chi(L) > p$. Since $T_{p+1,k} \in \mathcal{L}_{k,\mu}$, we have in (3) that $p(\mathcal{L}_{k,\mu}) = p$. Hence, by Erdős–Stone (2),

$$\text{Ex}(n, k, \mu) \sim \left(1 - \frac{1}{p}\right) \binom{n}{2} \sim e(T_{n,p}), \quad (4)$$

as $n \to \infty$.

2.2. Extremal Graphs for Dirac's Theorem

Let us compare estimates on $\text{Ex}(n, k, \mu)$ with the number of edges in the Turán graph $T_{n,p}$. We call $\text{Ex}(n, k, \mu) - e(T_{n,p})$ the *remainder term*. There are 3 cases:

—$T_{n,p}$ is extremal (i.e., the remainder term is 0);

—the remainder term is positive but has an $O(n)$ upper bound;

—the remainder term is at least n^{1+c} and at most n^{2-c}, for some constant $c \in (0, 1)$.

Dirac's Theorem belongs to the first case. We prove it in the following form.

[2] This is the best asymptotic lower bound for all $s \geq 2, \neq 5$. For $s = 5$, the regular generalized hexagon gives a better bound, $\Omega(n^{1+1/5})$.

Theorem 2.1. *Suppose $n \geq k \geq 2\lambda > 0$. If $G_n \in I(k, \lambda)$, then*

$$e(G_n) \leq t_{k-\lambda}(n).$$

Equality is attained, e.g., when G_n is the Turán graph $T_{n,k-\lambda}$. Our proof, presented in Section 3, is simpler and shorter than Dirac's. It involves an edge-density argument that is equivalent to the method used by Katona, Nemetz, and Simonovits [22] to prove Turán's Theorem. Katona [21] used this method again to investigate 3-graphs. The method was also described by Gessel [17], who explored the solutions to the recurrence, (8) below, generated by this argument.

Following the proof, we discuss more general bounds due to Dirac.

In Section 4 we shall investigate the structure of the extremal graphs in Dirac's Theorem. We will see that the Turán graph is the unique extremal graph for Theorem 2.1 except in the following two cases:

—when $k \geq 2\lambda \geq 4$, and $n = k$.

—when $k = 2\lambda \geq 2$ and $k + 1 \leq n \leq 2k - 2$.

In the first of these two cases, there are always at least two extremal graphs for Theorem 2.1, since any graph on k vertices with λ edges missing will do. The second case is included in Theorem 2.3 below.

Theorem 2.2. *Let $k > 2\lambda > 0$ and $n \geq k + 1$. If $S_n \in I(k, \lambda)$ and $e(S_n) = t_{k-\lambda}(n)$, then $S_n = T_{n,k-\lambda}$.*

Theorem 2.3. *Suppose that $k = 2\lambda > 0$ and $n \geq k + 1$. If $S_n \in I(k, \lambda)$ and $e(S_n) = t_{k-\lambda}(n)$, then S_n is one of the following graphs, depending on n:*

(1) *For $n = 2\lambda + r$ with $1 \leq r \leq \lambda$, $\lambda - r$ components of $\bar{S}_n$ are paths on one or more vertices and the rest, if any, are cycles.*

(2) *For $n = 3\lambda + r$ with $1 \leq r \leq \lambda$, r components of $\bar{S}_n$ are K_4's and the rest, if any, are cycles.*

(3) *For $n \geq 4\lambda + 1$, S_n is the Turán graph $T_{n,k-\lambda}$.*

Theorem 2.2 is due to Dirac. Theorem 2.3 is new, except that Dirac described it for $\lambda = 2$. For $n > n_0(k, \mu)$, one can derive all of the theorems above from the general Erdős–Simonovits structural theorem, Theorem 2.8 below, or from Theorem A.1 of the Appendix. Theorem A.1 is a general result, describing a large family of cases when the remainder term is linear, including *all* the cases of $\mathcal{L}_{k,\mu}$ with linear error terms. Those cases where the extremal graphs are the Turán graphs follow also from Theorem 2.9 below. For $n \geq 2(k - \lambda)$, the part of Theorem 2.2 for which $T_{n,k-\lambda}$ is extremal follows from Theorem 2.11.

For some related results of the second author, see also [32].

2.3. Further Exact Values

The next two theorems extend the inductive arguments of Section 3. We describe all cases (k, μ, p), $\mu \geq e(T_{k,p})$, such that $\text{Ex}(n, k, \mu) = e(T_{n,p}) + O(n)$ as $n \to \infty$. To be meaningful, we need $k > p$ here. Writing $\mu = e(T_{k,p}) + a$, we distinguish these three cases depending on a for given k, p.

(a) For $0 \le a < \lfloor \frac{1}{2} \lceil k/p \rceil \rfloor$, adding $a+1$ independent edges to a largest part of $T_{n,p}$ results in a forbidden graph.

(b) For $\lfloor \frac{1}{2} \lceil k/p \rceil \rfloor \le a < k/p$, adding $a+1$ independent edges to $T_{n,p}$ creates no $L \in \mathcal{L}_{k,\mu}$, but adding a path P_{a+2} does.

(c) For $a \ge k/p$, adding P_{a+2} to the first class of $T_{n,p}$ still creates no forbidden subgraphs.

We remark that the formula above has another form that is more natural:

$$\left\lfloor \frac{1}{2} \left\lceil \frac{k}{p} \right\rceil \right\rfloor = \left\lfloor \frac{k+p-1}{2p} \right\rfloor .$$

Theorems 2.4 and 2.5 below describe Cases (a) and (b), respectively. Case (c) is a prototype of the situation that $\text{Ex}(n,k,\mu) - e(T_{n,p})$ is nonlinear in n. We shall describe it here only superficially, in the paragraph preceding Theorem 2.6, and present a typical case in Theorem 2.7.

We begin with Case (a). We prove this result at the end of Section 3 from our main inductive lemma, Theorem 2.11.

Theorem 2.4. *Suppose $0 \le a < \lfloor \frac{1}{2} \lceil \frac{k}{p} \rceil \rfloor$. Then there exists a threshold $n_0(k,p,a)$ such that*

$$\text{Ex}(n,k,t_p(k)+a) = e(T_{n,p}) + a \textit{ for } n \ge n_0(k,p,a).$$

We denote by $T_{n,p,a}$ the graph obtained from $T_{n,p}$ by putting a independent edges into the largest class of $T_{n,p}$. This is an extremal graph for Theorem 2.4, but there are others. One can distribute the a edges arbitrarily among the classes, and they do not have to be independent. For another example, letting $a_1, a_2 \ge 1$ such that $a_1 + a_2 = a + 1$, we can put a star of a_1 edges into one class of $T_{n,p}$, put a star of a_2 edges into another class, and then delete the edge between the centers of the two stars.

Moving to Case (b), our next theorem asserts that if $a < k/p$, then there exists an extremal graph obtained from $T_{n,p}$ by adding as many edges to it as possible without getting forbidden subgraphs. Recall that for graphs $G^1, \ldots, G^p$, with pairwise disjoint vertex-sets, their product $\prod G^i$ is obtained by joining each vertex of G^i to each vertex of G^j.

Theorem 2.5. *Suppose $\lfloor \frac{1}{2} \lceil \frac{k}{p} \rceil \rfloor \le a \le \lceil \frac{k}{p} \rceil - 2$. Let $\mu = e(T_{k,p}) + a$. Then there exists a threshold $n_0(k,p,a)$ such that for $n \ge n_0(k,p,a)$, there exists an extremal graph S_n for $\text{Ex}(n,k,\mu)$ having product form, $S_n = \prod G^i$, where $|v(G^i) - \frac{n}{p}| \le 1$ for all i; G^1 is the vertex-disjoint union of trees, all but one of which have the same size; and $\sum_{j>1} e(G^j) < a$.*

Using this theorem one can easily get the precise value of $\text{Ex}(n,k,\mu)$ for this range. Applying the Structure Theorem 2.8 (or Theorem 2.9), all extremal graphs S_n can be determined, and this is done implicitly in our proof, which is presented in Section 5.

Remark. A more precise description of the product extremal graphs of Theorem 2.5 is the following. Take a Turán graph $T_{n,p}$. Let its classes be $A_1, \ldots, A_p$. To get a good lower bound in Theorem 2.5, let us try to put as many edges in its first class A_1 as possible. If we put a tree T_γ into A_1 for some $\gamma > a+1$, then we certainly get some $T_{k,p}$ with $\geq a+1$ additional edges: we get a forbidden $L \in \mathcal{L}_{k,\mu}$. Therefore, if we add edges to A_1 so that the resulting graph contains no forbidden subgraphs, then each component has at most $a+1$ vertices. Let $\gamma = \gamma(k,\mu)$ be the maximum for which we can put vertex-independent trees $T_1, \dots, T_j$ of equal order $\gamma \leq \mu$ into A_1 so that (i) the number of vertices not covered is smaller than γ and (ii) the resulting graph S_n^0 contains no $L \in \mathcal{L}_{k,\mu}$. Clearly,

$$e(S_n^0) - e(T_{n,p}) = \left(1 - \frac{1}{\gamma}\right) \frac{n}{p} - O(1).$$

By definition, a k-vertex subgraph $L \subseteq S_n^0$ will have at most μ edges. Let μ' be the maximum number of edges in a k-vertex subgraph of S_n^0. Put $\rho(k,\mu) = \mu - \mu' \geq 0$. Here μ' and $\rho(k,\mu)$ depend only on k and μ, and can be calculated easily. Add $\rho(k,\mu)$ edges to S_n^0 arbitrarily: adding to A_1 is also allowed. The obtained graphs contain no $L \in \mathcal{L}_{k,\mu}$. One can prove that for n large enough, all these graphs are extremal. (There may also be further extremal graphs. To get other extremal graphs, one can slightly adjust the sizes of the trees T_j by diminishing some and increasing others, or we can add slightly more edges elsewhere.)

As for Case (c), if $a \geq k/p$, then $\mathrm{Ex}(n,k,\mu) > e(T_{n,p}) + c_1 n^{1+\gamma}$ for some $\gamma > 0$: One can put a graph of girth exceeding k—described in (1)—into one class of $T_{n,p}$. We shall not give a detailed discussion of this case. Rather, we describe one very typical example: the case $k = 6, \lambda = 4$, i.e., when at least 4 edges are missing from each $G_6 \subseteq G_n$. This is the problem $\mathrm{Ex}(n,6,11)$. First, we recall the Octahedron Theorem, which concerns the exclusion of the octahedron graph $O_6 = K_3(2,2,2)$.

Theorem 2.6. *(Erdős and Simonovits [10]) For $n > n_0$, every graph $S_n \in \mathrm{EX}(n, O_6)$ can be obtained as $S_n = U_m \otimes Z_{n-m}$, for some $U_m \in \mathrm{EX}(m, C_4)$ and some $Z_{n-m} \in \mathrm{EX}(n-m, P_3)$, where $m = n/2 + o(n)$.*

Here Z_{n-m} is the graph of $\lfloor \frac{n-m}{2} \rfloor$ independent edges. The maximum size of a C_4-free graph on m vertices and the extremal graphs are determined by Füredi [14, 15] for infinitely many values of m. However, this is not enough to determine the exact value of m in Theorem 2.6. It seems to be hopeless, since $e(U_m)$ is strongly connected with the existence of some finite geometries.

In Section 6 we shall prove the following result for the $\mathrm{Ex}(n,6,11)$-problem. This theorem and its proof are very similar to the Octahedron Theorem. Here, Z_{n-m} has no edges.

Theorem 2.7. *For $n > n_0$, every graph $S_n \in \mathrm{EX}(n,6,11)$ (i.e., S_n is extremal for $\mathcal{L}_{6,-4}$) can be obtained as $S_n = U_m \otimes Z_{n-m}$, for some $U_m \in \mathrm{EX}(m, \{C_3, C_4\})$ and $Z_{n-m} \in \mathrm{EX}(n-m, P_2)$, where $m = n/2 + o(n)$.*

2.4. General Theory of Turán Problems

In the proofs of Theorems 2.5 and 2.7, we shall use the structural variant of the Erdős–Stone–Simonovits theorem, formulated below. This Structure Theorem asserts that in all cases the structure of extremal graphs is asymptotically the same as the structure of the Turán graph. Recall that $p(\mathcal{L}) = \min_{L\in\mathcal{L}} \chi(L) - 1$.

Theorem 2.8. *(Erdős, Simonovits, [7, 8, 28]) Let S_n be extremal for a family $\mathcal{L}$. Let $p = p(\mathcal{L})$. Then for any $x \in V(S_n), d(x) \geq n - \frac{n}{p} + o(n)$. Further, $V(S_n)$ can be partitioned into p classes $A_1, \ldots, A_p$ with the following properties:*

(a) $|A_i| = \frac{n}{p} + o(n)(i = 1, \ldots, p)$ and for all p-partitions $\sum e(G[A_i])$ is the minimum possible.

(b) For every $\epsilon > 0$, the number of vertices of $G[A_i]$ of degree $\geq \epsilon n$ (the degree counted in $G[A_i]$) is at most Ω_ϵ for some constant Ω_ϵ.

(c) Fix a graph M, and let $\epsilon < \frac{1}{2v(M)}$. Denote by A_i^* the subclass of A_i consisting of the vertices joined to A_i by fewer than ϵn edges. If $M \otimes K_{p-1}(k, \ldots, k)$ contains a forbidden subgraph $L \in \mathcal{L}$, then $M \not\subseteq G[A_i^*]$.

The vertices in (b) will be called *exceptional*. There are $\mathcal{L}$'s where the exceptional vertices play an important role, but in some other cases the main point of the analysis is just to show their nonexistence. In all cases considered in this article, the existence of such vertices can be ruled out. (Exceptional vertices can always be ruled out when $K_{p+1}(1, k, \ldots, k)$ contains some forbidden L. In Dirac-type problems with linear remainder terms, this always holds. The exceptional vertices can also be ruled out in Theorems 2.6 and 2.7, but for completely different reasons.)

In our cases, to determine exactly or estimate the value of $\mathrm{Ex}(n, k, \mu)$, we

(i) first characterize the family $\mathcal{L}_{k,\mu}$,

(ii) next find out which members of $\mathcal{L}_{k,\mu}$ really influence the order of magnitude of $\mathrm{ex}(n, \mathcal{L}_{k,\mu})$, and

(iii) finally apply a known exact theorem or known estimates to some forbidden subfamily $\mathcal{L}^* \subseteq \mathcal{L}_{k,\mu}$ (such as Theorem 2.8).

It is surprising that most phenomena occurring in Turán-type extremal problems do occur already in Dirac-type problems.

There are various general theorems implying Dirac's Theorem relatively easily, assuming that we care only for the large values of n: we want to prove the results only for $n > n_0(k, \mu)$. Among others, it is not too difficult to derive it from Theorem 2.8. Later we will see two inductive proofs. Here we quote a general theorem that easily implies Dirac's Theorem.

Theorem 2.9. *(Simonovits [28], cf. Erdős [6] for $p = 2$) Given a family $\mathcal{L}$ of simple graphs, the following statements are equivalent:*

(i) *For $n > n_0(\mathcal{L}), T_{n,p}$ is an extremal graph.*

(ii) *For $n > n_1(\mathcal{L}), T_{n,p}$ is the only extremal graph.*

(iii) *Every graph $L \in \mathcal{L}$ has chromatic number $> p$ and there is an $L_0 \in \mathcal{L}$ with an edge e for which $\chi(L_0 - e) = p$.*

Proof of Dirac's Theorem 2.1 for $n > n_0(k, \lambda)$. Given $n \geq k \geq 2\lambda > 0$, take $\mathcal{L} = \mathcal{L}_{k,-\lambda}$. Removing an edge decreases the chromatic number by at most one, so for each $L \in \mathcal{L}, \chi(L) \geq \chi(K_k) - (\lambda - 1) > k - \lambda$. Taking $L_0 - e$ to be K_k with λ disjoint edges removed gives $\chi(L_0 - e) = k - \lambda$. Thus, (iii) of Theorem 2.9 holds with $p = k - \lambda$, and it follows by (i) that $e(G_n) \leq t_{k-\lambda}(n)$, if $G_n \in I(k, \lambda)$ and $n > n_0(k, \lambda)$. ∎

Theorem 2.9 has an interesting consequence: If, for $n > n_0(\mathcal{L}), T_{n,p}$ is extremal, then, for $n > n_1(\mathcal{L})$, it is the only extremal graph. Here we give a strengthening of this statement by specifying an $n_1(\mathcal{L})$. It will be proved in Section 4.

Theorem 2.10. *If for all $n > n_0(\mathcal{L}), T_{n,p}$ is extremal, then for all $n > n_0(\mathcal{L}) + 2p + 1, T_{n,p}$ is the only extremal graph.*

As a matter of fact, if $n_1 > n_0 + p + 1$ is a multiple of p, then $n \geq n_1$ is enough.

2.5. Main Induction Lemma

Let p be given and (S_n) be a sequence of graphs obtained from $T_{n,p}$ by adding $a < \frac{n}{2p}$ independent edges to one of its larger classes. Then S_n is *almost regular* in the sense that the minimum degree and the maximum degree differ by at most 1. If one deletes an appropriate vertex $x \in S_n$, then one gets an S_{n-1}. This motivates the following theorem.

Theorem 2.11. *Let $\mathcal{L}$ be a given family of graphs. Let $(S_n)_{n \geq m}$ be a sequence of graphs with the following properties:*

(*A*) *S_n contains no $L \in \mathcal{L}$.*

(*B*) *S_m is extremal for $\mathcal{L}$.*

(*C*) *There exists a vertex $x \in V(S_n)$ of minimum degree such that $S_n - x = S_{n-1}$, for $n > m$.*

(*D*) *$\Delta(S_n) \leq \delta(S_n) + 1$, for $n > m$.*

(*E*) *Each S_n has at least 3 vertices of minimum degree for $n > m$.*

Then

(i) *For every $n \geq m, S_n$ is extremal for $\mathcal{L}$.*

(ii) *for every G_n not containing subgraphs in $\mathcal{L}, \delta(G_n) \leq \delta(S_n)$.*

(iii) *For every extremal graph Q_n for $\mathcal{L}, \delta(Q_n) = \delta(S_n)$. If x is a vertex of minimum degree in Q_n, then $Q_n - x$ is also extremal.*

The Inductive Lemma, Theorem 2.11, will be proved directly in the next section, without using the deeper theorems. Theorem 2.10 will also have an "elementary" proof.

3. MINIMUM DEGREE PROOF OF DIRAC'S THEOREM

Here we prove Theorem 2.1. The proof described below could also be called the average-degree-proof. The basic idea of the proof is that, deleting a vertex of minimum degree from a $G \in I(k, \lambda)$, we get a similar graph on $n-1$ vertices. Since the Turán graphs are almost regular, the number of edges goes down roughly by the same amount as in the Turán graph. So we can use induction on n. The average degree is not necessarily integer, and, if it is not, we should delete any vertex of degree smaller than the average. We need a lemma from [22].

Lemma 3.1.

$$\frac{e(G_n)}{\binom{n}{2}} = \frac{1}{n} \sum_{v \in V} \frac{e(G_n - v)}{\binom{n-1}{2}}. \tag{5}$$

More generally, if $m < n$, then

$$\frac{e(G_n)}{\binom{n}{2}} = \frac{1}{\binom{n}{m}} \sum_{\substack{G^* \subseteq G_n \\ v(G^*) = m}} \frac{e(G^*)}{\binom{m}{2}}, \tag{6}$$

where the summation is taken on the induced m-vertex subgraphs.

Proof. Display (6) follows by observing that every $e \in E(G_n)$ appears in $\binom{n-2}{m-2}$ of the graphs G^*. ∎

One can rewrite (5):

$$e(G_n) \leq \frac{1}{n-2} \sum_{v \in V} e(G_n - v). \tag{7}$$

Proof of Theorem 2.1. It is easy to see that $T_{n,k-\lambda} \in I(k, \lambda)$. To start the proof of the upper bound, one sees that $T_{k,k-\lambda}$ has precisely λ edges missing, so the theorem holds when $n = k$. We use this as the basis for induction on n with fixed $(k, \lambda), k \geq 2\lambda > 0$. Let $n > k$ and assume the theorem holds for $n-1$. If we assume that $G_n \in I(k, \lambda)$, then for all $v, G_n - v \in I(k, \lambda)$, so $e(G_n - v) \leq t_{k-\lambda}(n-1)$, by induction. By (7),

$$e(G_n) \leq \frac{n}{n-2}\, t_{k-\lambda}(n-1).$$

So

$$e(G_n) \leq \left\lfloor \frac{n}{n-2}\, t_{k-\lambda}(n-1) \right\rfloor.$$

The desired bound on $e(G_n)$ follows, provided that

$$t_{k-\lambda}(n) = \left\lfloor \frac{n}{n-2}\, t_{k-\lambda}(n-1) \right\rfloor. \tag{8}$$

Consider $T_{n,p}$ for arbitrary p, where we express n in terms of p by $n = qp + r$ with $1 \leq r \leq p$. Deleting a vertex v from one of the r parts of size $q+1$ leaves a graph isomorphic to $T_{n-1,p}$, while deleting a vertex from one of the $p-r$ parts of size q leaves a p-partite graph with $t_p(n-1)-1$ edges. The second case occurs $q(p-r)$

times. Applying (7) to $T_{n,p}$ and simplifying gives

$$e(T_{n,p}) = \frac{n}{n-2}\, t_p(n-1) - \frac{q(p-r)}{n-2}.$$

Thus, for $n > p+1$, we have $q(p-r) = n-(q+1)r < n-2$, implying

$$e(T_{n,p}) = \left\lfloor \frac{n}{n-2}\, t_p(n-1) \right\rfloor .$$

(Notice that this fails for $n = p+1$.) The desired conclusion (8) now follows, since $n \geq k+1 > (k-\lambda)+1$. ∎

In fact, Dirac [5] proved the following more general bound. If every k-vertex subgraph of G_n has at most $e(T_{k,p}) - \beta$ edges, then $e(G_n) \leq e(T_{n,p}) - \beta$. Indeed, if we add β edges to G_n in an arbitrary way, then the graph obtained satisfies the conditions of Theorem 2.1. When $\beta = 0$, the bound is sharp and is attained by $G_n = T_{n,p}$. However, for $\beta > 0$ and n large, this upper bound is too weak: it is weaker by $\approx \frac{n^2}{2p(p-1)}$ than $e(T_{n,p})$. As mentioned in (4), the Erdős–Stone Theorem implies that

$$\mathrm{Ex}(n,k,t_p(k)) - \beta = \left(1 - \frac{1}{p-1}\right)\binom{n}{2} + o(n^2).$$

The Inductive Proof for the Extremum

Proof of Theorem 2.11. We use induction on n. By (B), S_m is extremal. Assume that $n > m$ and we know that S_{n-1} is extremal. We assumed that S_n contains no forbidden subgraphs. To prove that it is extremal, it is enough to show that if $e(G_n) > e(S_n)$, then G_n contains some forbidden L. One can assume that $e(G_n) = e(S_n) + 1$.

(a) If $\delta(G_n) \leq \delta(S_n)$, then we select a vertex $x \in G_n$ of minimum degree. For $G_{n-1} = G_n - x$ we have, by (C),

$$e(G_{n-1}) = e(G_n) - d(x) > e(S_n) - \delta(S_n) = e(S_{n-1}).$$

Thus, for some $L \in \mathcal{L}$, we have $L \subseteq G_{n-1} \subset G_n$.

(b) The other case is when $\delta(G_n) > \delta(S_n)$. Now, by (E),

$$\sum d_G(x_i) \geq \left(\sum d_S(x_i)\right) + 3,$$

and, therefore, $e(G_n) \geq e(S_n) + 2$. This contradiction completes the proof of (i). Now that we know that (S_n) is a sequence of extremal graphs, (ii) is trivial (from $e(G_n) \leq e(S_n)$) and (iii) immediately follows from the argument of (a) applied to a G_n satisfying $e(G_n) = e(S_n)$. ∎

Corollary 3.1. *Under the conditions of Theorem* 2.11, *if* $Q_n \in \mathrm{EX}(n,\mathcal{L})$ *is an arbitrary extremal graph*, $n > m$, *then*

(i) $\delta(Q_n) = \delta(S_n)$ *and*

(ii) *for every vertex x of minimum degree, $Q_n - x$ is an extremal graph.*

Proof of Theorem 2.4. With $\mu := e(T_{k,p}) + a$, apply Theorem 2.11 to $\mathcal{L}_{k,\mu}$ and to the sequence $(S_n) = (T_{n,p,a})$. Clearly, S_k is extremal for $\mathrm{Ex}(n, k, \mu)$ (though not the only one), and the other conditions of Theorem 2.11 are automatically satisfied. ∎

4. EXTREMAL GRAPHS FOR THEOREM 2.1

Lemma 4.1. *For $n > p$, if $e(G_n) = e(T_{n,p})$ and if $x, y \in V(G_n)$ are two independent vertices such that $G_n - x \simeq G_n - y \simeq T_{n-1,p}$, then $G_n \simeq T_{n,p}$.*

Proof. $T_{n,p}$ can be characterized by saying that it is the unique p-chromatic n-vertex graph with maximum number of edges. So we may assume that $\chi(G_n) > p$, otherwise $G_n = T_{n,p}$, by the uniqueness. $G_n - x = T_{n-1,p}$, hence x is joined to each class of $T_{n-1,p}$ (by $\chi(G_n) > p$), which implies that there is a $K_{p+1} \subseteq G_n$ containing x. This K_{p+1} does not contain y, so $K_{p+1} \subseteq G_n - y = T_{n-1,p}$, a contradiction. ∎

Remarks. (a) There are many other ways to prove this simple but important lemma.

(b) If we drop the condition $e(G_n) = e(T_{n,p})$, then the assertion of Lemma 4.1 will not necessarily be true anymore. For example, take a $T_{n,p}$ and let x, y be two vertices from two distinct larger classes, then delete the edge x, y.

Proof of Theorem 2.10. (a) First we show that for $n = p\ell > n_0(\mathcal{L}) + p + 1$, $T_{n,p}$ is the only extremal graph. Apply Theorem 2.11 with $S_n = T_{n,p}$. Let Q_n be an arbitrary other extremal graph. By Theorem 2.11(iii), $\delta(Q_n) = \delta(T_{n,p})$. Clearly, since $T_{n,p}$ is regular, and $e(T_{n,p}) = e(Q_n)$ and the minimum degrees are the same, therefore Q_n is also regular, of degree $(p-1)\ell$. Delete any $p+1$ vertices $x_1, \dots, x_{p+1}$ from Q_n. If $e_X = e(\{x_1, \dots, x_{p+1}\})$, then

$$e(Q_{n-p-1}) = e(Q_n) - \sum_{i=1}^{p+1} d(x_i) + e_X. \tag{9}$$

Deleting a set Y of $p+1$ *appropriate* vertices of $T_{n,p}$, we get a $T_{n-p-1,p}$. Therefore,

$$e(T_{n-p-1,p}) = e(T_{n,p}) - \sum_{i=1}^{p+1} d_T(y_i) + e_Y. \tag{10}$$

Here $e(Y) = \binom{p+1}{2} - 1$. Since all degrees are the same and

$$e(Q_{n-p-1}) \le t_p(n-p-1) \text{ and } e(Q_n) = e(T_{n,p}),$$

from (9) and (10) we get that $e_X \le e_Y : K_{p+1} \not\subseteq Q_n$. So we may apply the uniqueness part of Turán's Theorem: $Q_n = T_{n,p}$.

(b) Let n^* be the smallest n described in (a), so it is the smallest multiple of p greater than $n_0(\mathcal{L}) + p + 1$. Now we show that if $n > n^*$, then $T_{n,p}$ is the only extremal graph. We use induction on n. Assume that for $n-1$ we know that $T_{n-1,p}$ is the only extremal graph. Let Q_n be an arbitrary extremal graph. By Theorem 2.11(iii), $\delta(Q_n) = \delta(T_{n,p})$. Deleting a vertex x of minimum degree of Q_n, we get again an extremal graph $Q_{n-1} = T_{n-1,p}$. If there are 2 independent vertices of minimum degree in Q_n, then we are home, by Lemma 4.1.

Now let x be any vertex of Q_n of minimum degree. By induction, $Q_n - x = T_{n-1,p}$. Let the classes of this $T_{n-1,p}$ be $A_1, \dots, A_p$, where $A_1, \dots, A_j$ have $q+1$ vertices, $A_{j+1}, \dots, A_p$ have q, $0 \le j < p$. The trivial case $j = 0$ will be left to the reader; assume that $j > 0$. The minimum degrees are equal to $\delta_n := (n-q-1)$ in $T_{n,p}$ and Q_n. The minimum degree of $T_{n-1,p}$ is $\delta_{n-1} = \delta_n - 1$. By the properties of the Turán graph, in a larger class of $T_{n-1,p}$ every vertex is of minimum degree $\delta_n - 1$ and stays of minimum degree δ_n even in $T_{n,p}$. Any vertex $y \in A_i$ for $i = 1, \dots, j$ will be of minimum degree in Q_n if it is joined to x. If, on the other hand, y is not joined to x, then it will have degree $\delta_n - 1 < \delta(T_{n,p})$, a contradiction. So, all the vertices of $A_1, \dots, A_j$ are joined to x.

Since x is not joined to all the vertices, we may assume that there is a y, say, in A_{j+1} *not* joined to x. The degree of y in Q_n is the same as its degree in $Q_n - x = T_{n-1,p}$, i.e., $(n-q-1) = \delta_n$. So y and x are 2 independent vertices of minimum degree, and consequently, $Q_n - y = T_{n-1,p}$ as well. By Lemma 4.1, $Q_n = T_{n,p}$. ∎

Proof of Theorem 2.2. We know by the previous results that the uniqueness holds for $n > n_1(k, \lambda)$ and the only thing missing is that this n_1 is so small.

The proof goes by induction on n with (k, λ) fixed. The real new point is that here we have the "induction basis" for a smaller n_0. For the induction basis, suppose that $n = k+1$. Clearly, no vertex of S_n is on two missing edges. This forces S_n to be a Turán graph: $S_n = T_{n,k-\lambda}$.

Now suppose that $n \ge k+2$. Then part (b) of the proof of Theorem 2.10 works: it uses only Theorem 2.11 and that, for $n-1$, we already know the uniqueness. Thus, $S_n = T_{n,k-\lambda}$. ∎

Proof of Theorem 2.3. It can be checked that every graph described in Theorem 2.3 belongs to $I(2\lambda, \lambda)$ and has $e(S_n) = t_{k-\lambda}(n)$. It remains to show that these are the only graphs.

In case (1), where $n = 2\lambda + r$, the complement $\bar{T}_{n,k-\lambda}$ of the Turán graph consists of rK_3's and $\lambda - rK_2$'s. By Theorem 2.11, $\Delta(\bar{S}_n) = \Delta(\bar{T}_{n,k-\lambda}) = 2$, so $\bar{S}_n$ is a disjoint union of paths and cycles. Since $e(\bar{S}_n) = e(\bar{T}_{n,k-\lambda}) = \lambda + 2r = n - (\lambda - r)$, it must be that $\lambda - r$ of the components are paths.

For $n > 3\lambda$, we proceed by induction on n, having already dealt with $n = 3\lambda$ in case (1). For case (2), with $n = 3\lambda + r$, the graph $\bar{T}_{n,k-\lambda}$ consists of rK_4's and $\lambda - rK_3$'s. By Theorem 2.11, $\Delta(\bar{S}_n) = \Delta(\bar{T}_{n,k-\lambda}) = 3$. Let v be a vertex in S_n of degree $\delta(S_n) = n - 4$. Since $e(S_n - v) = e(T_{n,k-\lambda} - v) = t_{k-\lambda}(n-1)$, then by induction, $\bar{S}_n - v$ consists of $r - 1 K_4$'s and $3(\lambda - r + 1)$ vertices in a disjoint union of cycles. Thus, in $\bar{S}_n$, vertex v has degree 3 and can be adjacent only to vertices of degree 2 in $\bar{S}_n - v$, i.e., to vertices on cycles. Suppose that w is adjacent to v, and let x and y be its neighbors in the cycle for w in $\bar{S}_n - v$. Similarly considering $\bar{S}_n - w$, we find x and y have degree 2, so they have degree 3 in $\bar{S}_n$, so they must be adjacent to v. Next considering $\bar{S}_n - x$, we conclude that x and y are adjacent in $\bar{S}_n$. So we have a K_4 on v, w, x, y, and $\bar{S}_n$ consists of rK_4's and a disjoint union of cycles. This completes case (2). Notice that S_n must be the Turán graph $T_{n,k-\lambda}$ for $n = 4\lambda$ and $4\lambda - 1$.

It remains to consider case (3) with $n \geq 4\lambda + 1$. Define q and r by $n = q\lambda + r, 1 \leq r \leq \lambda$. By the Lemma 4.1, S_n has a vertex v of degree $\delta(S_n) = \delta(T_{n,k-\lambda}) = n - 1 - q$. By induction, $S_n - v = T_{n-1,k-\lambda}$. If $S_n \neq T_{n,k-\lambda}$, then v is adjacent to all λ parts of $T_{n-1,k-\lambda}$ and (at least) twice adjacent to at least $\lambda - 1$ parts. Then we can find 2λ vertices in G_n with just $\lambda - 1$ missing edges, contradicting $S_n \in I(2\lambda, \lambda)$. Hence, $S_n = T_{n,k-\lambda}$, as claimed. ∎

5. CASE OF LINEAR REMAINDER TERMS

We will deduce Theorem 2.5 from the Structure Theorem 2.8. One can relatively easily prove Theorem 2.5 using the results of [30], (Theorem A.1 of the Appendix). However, the proofs in [30] are much more involved, and here we will use only the "cheaper parts" of those proofs.

Claim. Using the notations $\gamma(k,\mu)$ and $\rho(k,\mu)$ defined in the second paragraph following Theorem 2.5, and $\nu = \lceil k/p \rceil$, we have, under the conditions of Theorem 2.5, for $\mu = e(T_{k,p}) + a, \rho(k,\mu) < \lfloor \frac{\nu}{2} \rfloor$.

Proof. The condition on a means that $\gamma(k,\mu) > 1$. We must show that, if the first class of $T_{n,p}$ is filled up with independent edges (or larger blocks), then in the next class we cannot put $\lfloor \nu/2 \rfloor$ independent edges. Indeed, putting as many independent edges as possible into a large class of $T_{k,p}$, and into a small class, we get at least as many edges as by putting a path P_ν into a large class: we get an $L \in \mathcal{L}_{k,\mu}$. So $\rho(k,\mu) < \lfloor \frac{\nu}{2} \rfloor$. ∎

Remark. As a matter of fact, for large k,

$$\rho(k,\mu) < \frac{\nu}{\gamma(k,\mu)} - \frac{\nu}{\gamma(k,\mu)+1} \leq \frac{\nu}{2} - \frac{\nu}{3} = \frac{\nu}{6}.$$

Proof of Theorem 2.5. For simpler formulation of some facts, we introduce $\mathcal{L}^*_{k,\mu}$, which consists of all graphs containing some $L \in \mathcal{L}_{k,\mu}$. Obviously, the extremal problems for $\mathcal{L}^*_{k,\mu}$ and $\mathcal{L}_{k,\mu}$ are the same.

All such proofs include a construction, i.e., a sequence (U_n) of graphs not containing any $L \in \mathcal{L}$ and, therefore, providing the lower bound. Now the graphs described in the second paragraph following Theorem 2.5 yield the lower bound: show that if S_n is an arbitrary extremal graph, then

$$e(S_n) > e(T_{n,p}) + \left(1 - \frac{1}{\gamma(k,\mu)}\right)\frac{n}{p} - kp. \tag{11}$$

We shall use (as in the proof of the claim) that $\gamma(k,\mu) \geq 2$. In other words, *now* arbitrarily many independent edges can be put into the first class of $T_{n,p}$ without getting forbidden subgraphs. Therefore,

$$e(S_n) > e(T_{n,p}) + cn \tag{12}$$

for $c = \frac{1}{3p} > 0$ and n large.

(A) First we show that Theorem 2.8 is applicable. One can easily check that if T_{a+1} is any tree of order $a+1$, then $T_{a+1} \times K_{p-1}(\nu, \ldots, \nu) \in \mathcal{L}^*_{k,\mu}$. In particular, $K_{p+1}(1, k, k, \ldots, k)$ contains some $L \in \mathcal{L}_{k,\mu}$. Hence, we can apply Theorem 2.8 with $p(\mathcal{L}_{k,\mu}) = p$: we can partition $V(S_n)$ into classes A_i so that (a), (b), and (c) of Theorem 2.8 hold.

Notation. Let $N(x)$ denote the neighborhood of a vertex x and, if $x \in A_i$, then let $\alpha(x) := |A_i \cap N(x)|, \beta(x) := |V(S_n) - A_i - N(x)|$. In words, $\beta(x)$ is the number of "missing edges" and $\alpha(x)$ is the number of "extra edges" (compared to the corresponding complete p-partite graph). Put $G^i = G[A_i]$.

(B) We can now easily improve (b) by showing that there are *no exceptional vertices* in A_i. As a matter of fact, there are no vertices joined to A_i by at least k edges. Indeed, let $B_i \subseteq A_i$ be the set of vertices joined to at least ϵn vertices of A_i. By Theorem 2.8, $|B_i| = O(1)$. Each $x \in B_i$ is joined to some $y_1, \ldots, y_k \in A_i - B_i$. All but $o(n)$ vertices of $\cup_{j \neq i} A_j$ are joined to each $y_\ell (\ell = 1, \ldots, k)$. By the choice of the partition (by the minimality of the number of *missing* edges), $|N(x) \cap A_j| > \frac{n}{2p} - o(n)$ if $j \neq i$ and $|N(y) \cap A_j| > \frac{n}{p} - \epsilon n - o(n)$ if $y \in A_i - B_i, j \neq i$. Hence, if $B_i \neq \emptyset$, then we can find a $K_p(k, \ldots, k)$ in the neighborhood of an $x \in B_i$, and, therefore, a $K_{p+1}(1, k, k, \ldots, k) \in \mathcal{L}^*_{k,\mu}$ in S_n, a contradiction. So $B_i = \emptyset$. By (c) of Theorem 2.8, applied to $M = K(1,k), \Delta(G^i) < k$. Thus, $\alpha(x) < k$ and $\beta(x) = o(n)$ for every vertex x. As a matter of fact, we obtained that $\Delta(G^i) \leq a$.

Since for every tree T_{a+1} of $a+1$ vertices, $T_{a+1} \times K_{p-2}(\nu, \ldots, \nu) \in \mathcal{L}^*_{k,\mu}$, hence $G^i := G[A_i]$ contains no trees of order $> a$. Thus, G^i has no connected components of more than a vertices. Furthermore (by a similar argument), G^i has no connected components of a or more edges.

(C) We show that for all but at most one $i \leq p, e(G^i) \leq a$.

(C1) If we add $a+1$ edges to $T_{m,p}$ arbitrarily, $(m \geq k)$, then the resulting graph will contain some $L \in \mathcal{L}_{k,\mu}$.

(C2) Clearly,

$$e(S_n) \leq e(T_{n,p}) + \sum_j e(G^j) - M, \tag{13}$$

if M denotes the number of missing edges. If $e(A_1) > a$, then denote by A_j^* the subset of A_j joined to all the endvertices of these edges $(j > 1)$. By (C1) $e(G[A_j^*]) \leq a$. Since $|A_j - A_j^*| = o(n)$ and $\Delta(G^j) < k$, therefore $e(G^j) = o(n)$. If there were two classes containing $a + 1$ edges, or $M > cn$ holds, then for all $j, e(G_j) = o(n)$, and, therefore,

$$e(S_n) = e(T_{n,p}) + o(n)$$

would follow from (13), contradicting (12). So we may assume that $e(G^i) \leq a$ for $i > 1$ and $M = o(n)$. This implies that

$$e(G^1) = e(S_n) - e(T_{n,p}) - pa > \gamma(k, \mu)\frac{n}{p} - o(n).$$

Therefore, all but $o(n)$ vertices of G^1 are covered by trees in G^1, of size $\gamma(k, \mu)$.

(D) We show that $e(G^i) \leq \rho(k, \mu)$ for each $i > 1$. Indeed, assuming the contrary, we may fix $\rho(k, \mu) + 1$ edges in G^i and delete (at most) $o(n)$ vertices of each $G^j (j \neq i)$ joined to at least one of these edges by a *missing* edge. We can easily find k components of the remaining part of G^1, completely joined to these edges and each having at most $\gamma(k, \mu)$ vertices and at least $\gamma(k, \mu) - 1$ edges. These will provide an $L_k \subseteq S_n$ with $e(L_k) > \mu$ edges, by the definition of $\rho(k, \mu)$, a contradiction.

(E) "Filling in a missing edge (u, v) by an extra edge (a, b)" means below that (a, b) is an edge of some G^i and we delete it, u, v are not joined in S_n, belong to different classes, and we join them. If we fill in the *missing* edges by *extra* edges from $\cup_{i>1} E(G[A_i])$, then the resulting graph is extremal again and is a product. To show this we distinguish two cases.

(E1) If the number of *missing* edges was larger than the number of *extra* edges, then—in a second run—we fill in all the remaining *missing* edges as well. In the resulting graph $S_n^*, e(S_n^*) > e(S_n)$. So there is a subgraph $M_k \subseteq S_n^*$ with $e(M_k) > \mu$. In $S_n^*, A_2, \ldots, A_p$ contain no *extra* edges. Now we apply the so-called symmetrization: for $j = 2, \ldots, p$ we replace the vertices of M_k in A_j by the same number of "typical vertices $w_h \in A_j$," which are joined in S_n to all the vertices of $V(M_k) \cap A_1$, and to the replaced vertices of the other A_i's: we get an $M' \subseteq S_n$ with $e(M') > \mu$, a contradiction.

(E2) In the other case, we have filled in all the *missing* edges: we obtained a product S_n^*. By the Claim, there exists an $M_k \subseteq S_n$ with $||V(M_k) \cap A_i| - \frac{k}{p}| < 1$, containing all the *extra* edges of S_n. In other words, a Turán graph $T_{k,p}$ can be put onto $V(S_n)$, so that it covers all the *extra* edges. Clearly, the number of edges in such an M_k does not increase while filling in the *missing* edges. So, if the resulting S_n^* contained a forbidden subgraph, then the original S_n would also contain one. This contradiction completes the proof. ∎

6. CASE OF SUPERLINEAR REMAINDER TERMS

Here we prove Theorem 2.7. The proof is similar to that of the Octahedron Theorem.

Lemma 6.1. *If U_h contains neither C_3 nor C_4 and $e(W_m) = 0$, then $U_h \otimes W_m \in I(6,4)$: it contains no subgraphs on 6 vertices and 12 edges.* ∎

We shall need a slightly modified version of Lemma 2 of [10] stated below without proof. It applies to $K_2(a,b)$ if $a = 1,2,3$ but we formulate it only for C_4.

Lemma 6.2. (a) *For every $\eta > 0$, there exists a $\vartheta > 0$ such that, if G_m contains neither C_3, nor C_4, and has a vertex x of degree $\geq \eta m$, then*

$$e(G_m) \leq (1-\vartheta)\text{ex}(m, \{C_3, C_4\}).$$

(b) *If, in addition, G_m has a subgraph G^* of $\geq (1-\epsilon)m$ vertices with $e(G^*) \leq Cm$, then*

$$e(G_m) \leq \sqrt{\epsilon}m^{3/2} + Cm.$$

(For related results see also [13].)

Proof of Theorem 2.7. (Sketched) Since O_6, the octahedron graph, is a 6-vertex graph with only 3 missing edges, O_6 is one of the excluded graphs. Since $\chi(O_6) = 3$ and all the other graphs with 6 vertices and 12 edges contain a K_4, Theorem 2.8 can be applied and the proof of the Octahedron Theorem 2.6 can almost be copied. For the sake of completeness we sketch this proof, pointing out those parts where the proofs of Octahedron Theorem and Theorem 2.7 differ.

In the Octahedron Theorem 2.6, we exclude only one graph, the octahedron, and we concentrate on two types of occurrences of it:

$$O_6 = C_4 \otimes \overline{K_2} \text{ and } O_6 \subseteq P_3 \otimes P_3,$$

implying that if $Q \otimes R$ contains no O_6, then neither Q nor R can contain C_4 (unless Q or R is a single vertex); further, if one of them contains a P_3, then the other does not.

In our case, i.e., in the case of $\mathcal{L}_{6,11}$, the above assertions must be satisfied, of course, and in addition, we know that neither one of Q and R can contain K_3 either, and (finally), if, say, Q contains a P_4, then R cannot contain edges at all.

We shall fix a sufficiently small $\epsilon > 0$, say $\epsilon = \frac{1}{10000}$. Let S_n be an extremal graph for $\mathcal{L}_{6,-4}$. By Theorem 2.8, we can partition $V(S_n)$ into two classes A_1 and A_2 of size $\approx \frac{n}{2}$ so that $e(A_1) + e(A_2)$ is the minimum possible. This means that each $x \in A_i$ sends more edges to the other class than to its own one.

(i) Lemma 6.1 provides a lower bound on $e(S_n)$:

$$e(S_n) \geq \max_m\{m(n-m) + \text{ ex}(m, \{C_3, C_4\})\}. \tag{14}$$

It is known that $\mathrm{ex}(m, \{C_3, C_4\}) \geq \frac{m^{3/2}}{2\sqrt{2}} + o(m^{3/2})$.[3] By symmetry, we may assume that $e(A_1) \geq e(A_2)$. So we know that

$$e(A_1) \geq \frac{1}{2}\,\mathrm{ex}(|A_1|, \{C_3, C_4\}) \geq \frac{1}{20} n^{3/2}. \tag{15}$$

(ii) We call the vertices of A_i joined to at least ϵn vertices of their own class *exceptional* and denote their set by B_i. By Theorem 2.8(b), $|B_i| = O(1)$. In this part, we *can ignore* the $O(n)$ edges represented by B_i in $G[A_i], (i = 1, 2)$. Let $A_i^* = A_i - B_i$. Then $G[A_1^*]$ contains neither C_4 nor C_3, otherwise we would have a subgraph on 6 vertices and 12 edges.

Further, $e(A_2^*) = 0$. To prove this we use that $e(P_2 \otimes P_4) = 12$. Thus, if $G[A_2^*]$ contained an edge xy, then $P_4 \not\subseteq G[N(x) \cap N(y) \cap A_1]$ would follow, implying that $e(G[N(x) \cap N(y) \cap A_1]) = O(n)$. By $|A_1 - N(x) - N(y)| < \epsilon n$ and by Lemma 6.2(b),

$$e(G[A_1]) < \sqrt{\epsilon} n^{3/2} + O(n),$$

contradicting (15). (*In the octahedron problem here we allow a* 1*-factor*: exclude only vertices of degree 2 in $G[A_2]$.)

(iii) Now we show that $B_i = \emptyset$. If, indirectly, e.g., $x \in B_1$, then any C_4 or C_3 of $G[A_1]$ containing this x and 2 or 3 further vertices from A_1^* can easily be extended into an $L = K(2, 2, 2) \in \mathcal{L}_{6,-4}$ or into a $C_3 \otimes \overline{C_3} \in \mathcal{L}_{6,-4}$. Thus, the subgraph G_1^{**} spanned by x and A_1^* contains neither C_4 nor C_3. Applying Lemma 6.2(a) and $e(A_2) = O(n)$, we get that, for $m = |A_1|$,

$$e(S_n) < m(n - m) + O(n) + e(G^{**}) < e(S_n) - c_2(\epsilon n)^{3/2},$$

a contradiction. A similar argument shows that $B_2 = \emptyset$, too.

(iv) Now we know that $G[A_1]$ contains neither C_3 nor C_4, and $G[A_2] = 0$. Hence, by (14), each vertex of A_1 is joined to each one of A_2. ∎

7. HOW MANY SUBGRAPHS SHOULD BE EXCLUDED?

In this section we investigate whether or not one excluded subgraph can replace a whole large family of excluded subgraphs. In many cases, one finds that for a given $\mathcal{L}$ there is one appropriately chosen $L^* \in \mathcal{L}$ for which

$$\frac{\mathrm{ex}(n, L^*)}{\mathrm{ex}(n, \mathcal{L})} \to 1 \text{ as } n \to \infty. \tag{16}$$

In some other cases, we have the even stronger

$$\mathrm{ex}(n, L^*) = \mathrm{ex}(n, \mathcal{L}). \tag{17}$$

[3] Erdős conjectures that here equality holds, see below.

The answer to the question if (16) or (17) always holds is not so simple: e.g., for (16) NO if $p(\mathcal{L}) = 1$ and YES if $p(\mathcal{L}) > 1$. We illustrate the situation through some simple examples, which mostly follow from known results.

The case when there is a bipartite $L^* \in \mathcal{L}$ was investigated by Erdős and Simonovits, e.g., in [11]. Let us consider the case of $\mathcal{L}_{4,-2}$: the family of graphs on 4 vertices with ≥ 4 edges. Since $C_4 \in \mathcal{L}_{4,-2}$, and a triangle with one hanging edge is also in $\mathcal{L}_{4,-2}$, one can easily see that

$$\text{ex}(n, \{C_3, C_4\}) = \text{ex}(n, \mathcal{L}_{4,-2}) + O(n).$$

However, to decide if $\text{ex}(n, \{C_3, C_4\}) \approx \frac{1}{2} n^{3/2}$ or $\text{ex}(n, \{C_3, C_4\}) \approx (\frac{n}{2})^{3/2}$ or is somewhere in between seems to be one of the difficult problems in extremal graph theory. According to a famous conjecture of Erdős (see, e.g., [11]),

$$\text{ex}(n, \{C_3, C_4\}) \approx \left(\frac{n}{2}\right)^{3/2},$$

i.e., probably neither (16) nor (17) holds.

The Erdős–Stone–Simonovits theorem [9] immediately implies that in the so-called *nondegenerate* cases, i.e., if $\mathcal{L}$ contains no bipartite graphs, then (16) must always hold. We have seen (Theorem 2.9) that if $T_{n,p}$ is an extremal graph for $\mathcal{L}$, then there is always a graph $L^* \in \mathcal{L}$ such that

$$\text{ex}(n, L^*) = \text{ex}(n, \mathcal{L})$$

for $n > n_0$, i.e., (17) holds.

Now we give an example where there is no such L^*: as a matter of fact, we shall provide two examples. The first, deeper case is that of $\text{Ex}(n, 6, 11)$. We show that

$$\text{ex}(n, L) > \text{Ex}(n, 6, 11) + \frac{n}{4} - o(n), \tag{18}$$

if $v(L) = 6$ and $e(L) = 12$. Indeed, by Theorems 2.6 and 2.7 describing the extremal problems of O_6 and of $\mathcal{L}_{6,11}$, we know that

$$\text{ex}(n, O_6) > \text{Ex}(n, 6, 11) + \frac{n}{4} - o(n). \tag{19}$$

(To be more precise, we know for the case of O_6, that if S_n is extremal for $\mathcal{L}_{6,11}$, then $S_n = U_m \otimes W_{n-m}$ for some U_m not containing C_4, (neither C_3) and for some W_{n-m} with $e(W_{n-m}) = 0$. Now, an easy Lemma of [10] corresponding to Lemma 6.1 asserts that if Q is a graph containing no C_4 and R is another graph containing no P_3, then $O_6 \not\subseteq Q \otimes R$. So, adding $[\frac{n-m}{2}]$ edges to W_{n-m} will not create any O_6 in the product. This proves (19).) Further, if $L \in \mathcal{L}_{6,-4}$ and $L \neq O_6$, then $K_4 \subseteq L$. (As a matter of fact, this is Turán's Theorem for $n = 6$ and K_4.) Therefore,

$$\text{ex}(n, L) \geq \text{ex}(n, K_4) \approx \frac{n^2}{3},$$

completing the proof of (18). (If the Erdős conjecture holds, then $\frac{n}{4}$ can be replaced by $(\frac{1}{2} - \frac{1}{2\sqrt{2}}) n\sqrt{n}$.)

Now we provide another, simpler example where the family $\mathcal{L}_{k,-\lambda}$ cannot be replaced by just one excluded subgraph. Fix an $r \geq 2$ and a $2 \leq a \leq r/2$. Put $k = 2r$ and $\mu = r^2 + a$. If $n > 2r$, then $\text{ex}(n, \mathcal{L}_{k,\mu}) = \lfloor n^2/4 \rfloor + a$ by Theorem 2.11. (For some related results see [28] or [30].) On the other hand, we have the following.

Theorem 7.1. *If $k = 2r, a > 1$ and $\mu = r^2 + a$, then for any $L^* \in \mathcal{L}_{k,\mu}$ one has* $\text{ex}(n, L^*) > \frac{n^2}{4} + \frac{n}{4} + O(1)$.

Proof. Pick an arbitrary $L^* \in \mathcal{L}_{k,\mu}$. If there is no $v \in V(L^*)$ for which $\chi(L^* - v) = 2$, then $Z_n = K_3(1, \lceil \frac{n-1}{2} \rceil, \lfloor \frac{n-1}{2} \rfloor)$ (i.e., the graph obtained from $T_{n-1,2}$ by joining a new vertex w to all its vertices) contains no L^* and $e(Z_n) = \lfloor \frac{n^2}{4} \rfloor + \lfloor \frac{n-1}{2} \rfloor$, proving the assertion.

The other case is when, for some $v \in V(L^*),\ \chi(L^* - v) = 2$. Now let Z_n be the graph obtained from $T_{n,2}$ by adding a 1-factor to the first class of $T_{n,2}$. Now $e(Z_n) \geq \lfloor \frac{n^2}{4} \rfloor + \lfloor \frac{n}{4} \rfloor$ and Z_n contains no L^*. This is easy for large values of k and takes a little work for small values of k. So we are done. ∎

APPENDIX: A GENERAL THEOREM IN THE CASE OF LINEAR ERROR TERMS

A fairly general theorem of Simonovits [30] tells us that if, for some sufficiently large $t, L \in \mathcal{L}$ and $L \subseteq P_t \otimes K_{p-1}(t, \ldots, t)$, then there exist extremal graphs of fairly simple structure. This theorem of [30] also provides a necessary and sufficient condition for having *only* these symmetrical extremal graphs. This theorem is applicable in all the cases when $\text{Ex}(n, k, \mu)$ has a linear remainder (though this is not trivial).

We include these results here, since we feel that the theorem below *best describes* the situation investigated in this article (though we could easily prove our results from Theorem 2.8). To explain this theorem, first we have to define the notion of a family of fairly symmetrical graphs.

Definition A.1. Let T_j for $j = 1, \ldots, q$ be distinct connected subgraphs of G. They are called *symmetrical* if

(i) $V(T_i) \cap V(T_j) = \emptyset$ for $1 \leq i < j \leq q$, and
(ii) there are no edges joining T_i to T_j for $1 \leq i < j \leq q$, and
(iii) there exists an isomorphism $\omega_j : T_1 \to T_j$ such that, for every $x \in T_1, u \in G \setminus \cap_\ell V(T_\ell), x$ is joined to u if and only if $\omega_j(x)$ is joined to u.

Definition A.2. A property $\mathcal{A}$ of graphs will be called a *chromatic condition* if

(i) $G \in \mathcal{A}$ and $H \supset G$ implies $H \in \mathcal{A}$.
(ii) If $\rho = \rho(\mathcal{A})$ is a sufficiently large integer, then the following holds: if $T_1, \ldots T_\rho$ are symmetric subgraphs of an $\mathcal{A}$-graph G, then $G - T_\rho$ is also an $\mathcal{A}$-graph.

To rule out the uninteresting cases, we also assume that there are graphs of property $\mathcal{A}$ and of arbitrarily high girth.

Example. The property $\bar{A}_{k,p}$, that one cannot delete k vertices of G to get a graph of chromatic number at most p, is one of the typical chromatic conditions.

Definition A.3. (Family of symmetrical graphs) $\mathcal{G}(n,r,p)$ is the class of graphs G_n satisfying the following symmetry condition:

(i) It is possible to omit $\leq r$ vertices of G_n so that the remaining graph G^* is a product:

$$G^* = \prod_{d \leq p} G_{m_d}, \text{ where } \left| m_d - \frac{n}{p} \right| \leq r.$$

(ii) For each fixed $1 \leq d \leq p$, there exist connected graphs $H_{d,j} \subseteq G_{m_d}$ (and isomorphisms $\omega_{d,j} : H_{d,1} \to H_{d,j}$) such that $H_{d,j}(j = 1, 2, \cdots)$ are symmetric subgraphs of G_n and G_{m_d} is the union of the graphs $H_{d,j}$.

The vertices described in (i) have degree $> n - n/p + cn$ in the typical cases, for some constant $c > 0$.

Given a family $\mathcal{L}$ of graphs and a chromatic property $\mathcal{A}$, we say G_n is $(\mathcal{L}, \mathcal{A})$-extremal if it has the property $\mathcal{A}$, contains no $L \in \mathcal{L}$, and has maximum number of edges under these conditions.

Theorem A.1. (*Existence of sequences of symmetrical extremal graphs*) *Let* $\chi(L) \geq p+1$ *for every* $L \in \mathcal{L}$ *and* $\chi(L^*) = p+1$. *Let* $v(L^*) = \tau$. *If*

$$L^* \subseteq P^\tau \times K_{p-1}(\tau, \ldots, \tau), \tag{20}$$

then there exists a constant $r = r(\mathcal{L})$ *such that for every* $n, \mathcal{G}(n,r,p)$ *contains an extremal graph for* $\mathcal{L}$. *Furthermore, if there exists an* n_0 *such that for* $n > n_0, \mathcal{G}(n,r,p)$ *contains only one extremal graph, then for sufficiently large values of* n *this is the only extremal graph.*

We have mentioned that in all the cases when the "remainder" term is linear, Theorem A.1 describes the situation completely. The reason for this is that, in those cases, (20) is applicable: the basic forbidden graphs are obtained by putting trees into the classes of some Turán graphs, and putting a path into the first class of a $T_{k,p}$ also yields a forbidden graph.

ACKNOWLEDGMENTS

The authors thank Z. Füredi, I. Gessel, and A. Sidorenko for bringing work related to this project to our attention.

References

[1] B. Bollobás, *Extremal graph theory*, Academic Press, New York (1978).

[2] J. A. Bondy and M. Simonovits, Cycles of even length in graphs, *J. Combin. Theory (ser. B)* **16** (1974), 97–105.

[3] W. G. Brown, P. Erdős, and Vera T. Sós, On the existence of triangulated spheres in 3-graphs and related problems, *Periodica Math. Hung. Acad. Sci*. **3** (1973), 221–228.

[4] W. G. Brown, P. Erdős, and Vera T. Sós, *Some extremal problems on r-graphs, New directions in the theory of graphs*, F. Harary, Ed., Academic Press, New York (1973), 53–63.

[5] G. Dirac, Extensions of Turán's theorem on graphs, *Acta Math. Acad. Sci. Hung*. **14** (1963), 417–422.

[6] P. Erdős, On a theorem of Rademacher–Turán, *Illinois J. of Math*. **6** (1962), 122–127 (reprinted in the *Art of Counting*, MIT Press (1973), 131–136).

[7] P. Erdős, Some recent results on extremal problems in graph theory, *Theory of Graphs*, International Symp. Rome (1966), 117–130.

[8] P. Erdős, On some new inequalities concerning extremal properties of graphs, *Theory of Graphs*, Proc. Coll. Tihany, Hungary, P. Erdős and G. Katona, Eds., Academic Press, New York (1968), 77–81.

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

[10] P. Erdős and M. Simonovits, An extremal graph problem, *Acta Math. Acad. Sci. Hungar.* **22** (1971), 275–282.

[11] P. Erdős and M. Simonovits, Compactness results in extremal graph theory, *Combinatorica* **2** (1982), 275–288.

[12] P. Erdős and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc*. **52** (1946), 1089–1091.

[13] R. J. Faudree and M. Simonovits, On a class of degenerate extremal graph problems, *Combinatorica* **3** (1983), 83–93.

[14] Z. Füredi, Graphs without quadrilaterals, *J. Combin. Theory Ser. B* **34** (1983), 187–190.

[15] Z. Füredi, Quadrilateral-free graphs with maximum number of edges, *Proceedings of the Japan Workshop on Graph Th. and Combinatorics*, Keio University, Yokohama, Japan 1994, pp. 13–22.

[16] Z. Füredi, Turán type problems, *Surveys in Combinatorics*, London Math. Soc. Lecture Note Ser., A. D. Keedwell, Ed., Cambridge Univ. Press (1991), 253–300.

[17] I. Gessel, A recurrence associated with extremal problems, preprint (1989).

[18] A. I. Gol'berg and V. A. Gurvich, On the maximum number of edges for a graph with n vertices in which every subgraph with k vertices has at most l edges, *Soviet Math. Doklady* **35** (1987), 255–260.

[19] J. R. Griggs and G. R. Thomas, Maximum size graphs with k-subgraphs of size at most $k-2$, *Proc. 7th Intern. Conf. on Theory and Applns. of Graphs* (Kalamazoo, 1992), John Wiley and Sons, New York (1995), 1147–1154.

[20] R. K. Guy, Sequences associated with a problem of Turán and other problems, in *Combinatorial Theory and its Applications II*, Proc. Coll. Balatonfüred (1969), North–Holland, Amsterdam (1970), 553–560.

[21] G. O. H. Katona, Sums of vectors and Turán's problem for 3-graphs, *Europ. J. Combin.* **2** (1981), 145–154.

[22] G. O. H. Katona, T. Nemetz, and M. Simonovits, On a problem of Turán in the theory of graphs, *Mat. Lapok* **15** (1964), 228–238 (in Hungarian).

[23] T. Kővári, Vera T. Sós, and P. Turán, On a problem of Zarankiewicz, *Colloquia Math.* **3** (1954), 50–57.

[24] F. Lazebnik, V. A. Ustimenko, and A. J. Woldar, A new series of dense graphs of large girth, *Bull. Amer. Math. Soc*. (New Series) **32** (1995), 73–79.

[25] W. Mantel, Problem 28, *Wiskundige Opgaven* **10** (1907), 60–61.

[26] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, *Proc. Coll. Bolyai János Math. Society* **18,** Combinatorics (Keszthely, 1976), Vol. II., North–Holland, Amsterdam–New York (1978), 939–945.

[27] A. F. Sidorenko, What do we know and what we do not know about Turán numbers, *Graphs and Combinatorics* **11** (1995), 179–199.

[28] M. Simonovits, A method for solving extremal problems in graph theory, *Theory of Graphs, Proc. Coll. Tihany* (1966), P. Erdős and G. Katona, Eds., Academic Press, New York (1968), 279–319.

[29] M. Simonovits, On the structure of extremal graphs, Ph.D. thesis, Hungarian Academy of Sciences (1969), p. 85.

[30] M. Simonovits, Extremal graph problems with symmetrical extremal graphs, additional chromatic conditions, *Discrete Math*. **7** (1974), 349–376.

[31] M. Simonovits, Extremal Graph Theory, *Selected Topics in Graph Theory*, Beineke and Wilson, Eds., Academic Press, London (1983), 161–200.

[32] M. Simonovits, How to solve a Turán type extremal graph problem? (linear decomposition), *The Future of Discrete Mathematics*, DIMACS series. Proc. Conf. DIMATIA (Stirin) 1997, Amer. Math. Soc.

[33] B. Stechkin: see V. I. Baranov and B. Sz. Sztechkin, *Extremal combinatorial problems and their applications*, Nauka, Moscow (1989), (in Russian).

[34] P. Turán, On an extremal problem in graph theory, (in Hungarian) *Mat. Fiz. Lapok* **48** (1941) 436–452; see also P. Turán, On the theory of graphs, *Colloq. Math*. **3** (1954), 19–30, and also [35]. MR15, 976b.

[35] *Collected papers of Paul Turán*, Vols. 1–3, Akadémiai Kiadó, Budapest (1989).