

DIGRAPH EXTREMAL PROBLEMS, HYPERGRAPH EXTREMAL PROBLEMS, AND THE DENSITIES OF GRAPH STRUCTURES*

W.G. BROWN

McGill University, Montréal, Canada

M. SIMONOVITS

Eötvös Loránd University, Budapest, Hungary

Received 24 January 1983

We consider extremal problems 'of Turán type' for r -uniform *ordered* hypergraphs, where multiple oriented edges are permitted up to multiplicity q . With any such ' (r, q) -graph' G^n we associate an r -linear form whose maximum over the standard $(n-1)$ -simplex in \mathbf{R}^n is called the (graph-) density $g(G^n)$ of G^n . If $ex(n, \mathbf{L})$ is the maximum number of oriented hyperedges in an n -vertex (r, q) -graph not containing a member of \mathbf{L} , $\lim_{n \rightarrow \infty} ex(n, \mathbf{L})/n^r$ is called the *extremal density* of \mathbf{L} . Motivated, in part, from results for ordinary graphs, digraphs, and multigraphs, we establish relations between these two notions.

1. Introduction

In this paper we shall investigate Turán-type extremal problems for hypergraphs, and, more generally, for ' r -uniform directed q -hypergraphs'; each hyperedge contains r vertices, the same hyperedge may occur up to q times; even more generally, the edges will usually be *ordered* r -tuples – to generalize extremal problems for digraphs.

Given a family \mathbf{L} of q -hypergraphs (which we call 'prohibited'), $ex(n, \mathbf{L})$ will denote the maximum number of hyperedges (counted with multiplicity) an ordered q -hypergraph may possess, under the condition that it contains no $L \in \mathbf{L}$. Such problems are called 'Turán-type', in deference to the seminal work of P. Turán [20], [21]. In [2], [5] and [6] the present authors and P. Erdős have investigated extremal digraph problems, in [4] extremal multigraph problems. We propose to generalize results of those papers to oriented hypergraphs. We shall consider several different types of graph-theoretical objects:

- ordinary graphs without loops or multiple edges,
- multigraphs – where the multiplicity of each edge is bounded from above by a fixed integer q ,

*This research was supported in part by an Operating Grant of the Natural Sciences and Engineering Council of Canada, held by the first author.

- digraphs – where the multiplicity of each arc (= oriented edge) is bounded from above,
 - hypergraphs – where multiple hyperedges of bounded multiplicity are permitted,
- and – most generally –
- ‘ r -uniform directed q -hypergraphs’.

Definition 1. Let r and q be positive integers. An r -uniform directed q -hypergraph H is a set $V(H)$ of vertices, together with a family $E(H)$ of ordered r -tuples of elements of $V(H)$; an r -tuple with a given order (= ‘orientation’) may occur at most q times. We shall assume that the r -tuples consist of r distinct vertices from $V(H)$, i.e. ‘loops are excluded’.¹

The Fundamental Problem. For positive integers r and q we restrict ourselves to r -uniform directed q -hypergraphs. Given a family \mathbf{L} of such hypergraphs and an integer n , what is $\text{ex}(n, \mathbf{L})$, the maximum number of oriented r -tuples a hypergraph on n vertices can have without containing a member of \mathbf{L} (as an r -uniform directed q -hypergraph)?

Graphs, . . . , r -uniform directed q -hypergraphs will be denoted by capital Latin letters, as G, H, \dots, S ; or by G^n, H^n, \dots, S^n , where an upper index will always indicate the number of vertices.² Given a graph G , $e(G)$ will denote the number of edges, ordered r -tuples, etc., (counted with multiplicity, where applicable); $v(G)$ will denote the number of vertices. We streamline our language, where possible: by *graph* we may mean any one of the objects: graph, digraph, . . . , r -uniform directed q -hypergraph, depending upon the context. Where the parameters r, q , are needed, we may speak of an (r, q) -graph, or an (r, q) -digraph. Similarly, the subobjects will usually be called *subgraphs*; and the word *edge* will denote the appropriate type of subset, ordered where appropriate. The symbol $\text{ex}(n, \mathbf{L})$ will also have to be interpreted from the context. The set of *extremal graphs* – having n vertices, exactly $\text{ex}(n, \mathbf{L})$ edges, and no prohibited subgraph (in \mathbf{L}) – will be denoted³ by $\text{EX}(n, \mathbf{L})$. The requirement that multiplicities be bounded is needed to ensure a finite maximum – to exclude trivial cases, as where all edges are identically situated, and no ‘non-trivial’ subgraphs are present.

Ideally, for a given \mathbf{L} , we wish to determine the structure of all extremal graphs in $\text{EX}(n, \mathbf{L})$. Usually this is unattainable, and we must content ourselves with estimates of the *asymptotic* behavior of $\text{ex}(n, \mathbf{L})$ as $n \rightarrow \infty$. In particular, we wish to study the value⁴ of

$$\lim_{n \rightarrow \infty} \text{ex}(n, \mathbf{L})/n^r. \quad (1.1)$$

¹ But compare Section 9.

² An exceptional use of the superscript occurs in Definition 8.

³ When $\mathbf{L} = \{L\}$, we may write $\text{ex}(n, L)$ and $\text{EX}(n, L)$ in place of $\text{ex}(n, \mathbf{L})$ and $\text{EX}(n, \mathbf{L})$.

⁴ That this limit exists is a consequence of Lemma 2 below, which generalizes, trivially, a theorem of Katona, Nemetz, and Simonovits [15].

Often even this goal is unrealizable, and only upper and lower bounds can be determined. So-called ‘degenerate extremal problems’, where the limit in (1.1) is zero, will not be discussed here.

Our object in the present paper is to generalize certain extremal results of the present authors and P. Erdős. We believe that some of the generalizations which we prove are conceptually simpler than the more specialized results: some of the proofs given below are certainly simpler. Detailed motivation for the theorems generalized herein will be found in the references cited. Section 2 contains preliminaries. In Section 3 we prove a ‘continuity’ theorem, concerning approximation of families \mathbf{L} by finite subfamilies, and state a stronger conjecture (cf. [5, Section 9], [6]). In Section 4 we study graphs containing more than $\text{ex}(n, \mathbf{L})$ edges (cf. [12]). Section 5 is concerned with a general theorem of ‘Erdős–Stone’ type (cf. [10]). Section 6 is devoted to an investigation of the set of limits of form (1.1), and its relation to the set of ‘densities’ of graphs (cf. [6]). In Section 7 we prove an ‘approximation’ theorem, concerning the existence of asymptotically extremal sequences of subgraphs ‘of simple structures’ (cf. [5], [6]). Most of our generalizations will be proved first for (r, q) -digraphs; in Section 8 we discuss a principle for deriving corresponding *unoriented* results. In Section 9 we consider briefly generalizations to hypergraphs with loops.

Multidigraphs have been considered by Katona in [24], where he was primarily interested in continuous versions of Turán-type extremal graph problems.

2. Preliminaries: extremal numbers $\text{ex}(n, \mathbf{L})$; extremal (r, q) -graphs

When \mathbf{L} is a family of ordinary graphs (without loops or multiple edges) the limit in (1.1) is determined by the minimum of the chromatic numbers of the graphs in \mathbf{L} (cf. Erdős and Simonovits [11]). Specifically, if p denotes that minimum, then

$$\lim \text{ex}(n, \mathbf{L}) / \binom{n}{2} = 1 - \frac{1}{p-1} \quad \text{as } n \rightarrow \infty. \quad (2.1)$$

For the cases of digraphs with $q = 1$ or multigraphs with $q = 2$, the results of the present authors and P. Erdős apply (see [2], [3], [5], [6]). But no specific limit theorems similar to (2.1) are known in generality. For hypergraphs with $q > 2$, the situation is yet murkier! In the celebrated problem of Turán [21] one considers ordinary 3-uniform hypergraphs (i.e. $r = 3, q = 1$), and \mathbf{L} has only one member: the ‘complete’ 4-vertex graph with four 3-edges; that problem remains unsolved (cf. Section 9).

Given a family \mathbf{L} of prohibited graphs, what is the structure of the *extremal graphs*? Certain specialized results are known, (for example, for digraphs with $q = 1$ [7]). Most of our results in this area are related to the somewhat broader class of ‘almost extremal’ graphs, containing no prohibited graph and whose number of edges is asymptotically $\text{ex}(n, \mathbf{L})$. More precisely, we define

Definition 2. For a given family \mathbf{L} , an *asymptotically extremal sequence*⁵ $\{S^n\}_{n=1,2,\dots}$ (written briefly as $\{S^n\}$) consists of graphs such that

- (a) S^n contains no (prohibited) L in \mathbf{L} ; and
- (b) $e(S^n) = (1 + o(1)) \text{ex}(n, \mathbf{L})$ as $n \rightarrow \infty$.

In some cases we have succeeded in proving theorems of the following form: for given \mathbf{L} a certain fixed asymptotically extremal sequence $\{S^n\}$ 'of very simple structure' has the property that every extremal graph U^n may be obtained from S^n by adjoining or deleting $o(n^r)$ edges (cf. for example, the work of Erdős and Simonovits [11] for graphs; and the papers of Brown, Erdős and Simonovits [5] for multigraphs). A somewhat weaker general result of this type will be proved below in Theorem 6.

Definition 3. Let r and q be positive integers, and G an (r, q) -digraph with vertex set $V(G) = \{v_1, v_2, \dots, v_m\}$. Let $\mathbf{x} = (x_1, x_2, \dots, x_m)$ be a vector of non-negative integers, and let X_1, X_2, \dots, X_m be disjoint sets containing respectively x_1, x_2, \dots, x_m vertices. An (r, q) -digraph $G(\mathbf{x}) = G(x_1, x_2, \dots, x_m)$ is obtained by replacing each vertex v_i by the set X_i of vertices, and taking the corresponding r -edges. More precisely,

$$V(G(\mathbf{x})) = \bigcup_i X_i,$$

$$E(G(\mathbf{x})) = \{(w_1, w_2, \dots, w_r) : w_i \in X_{i_i} \ (i = 1, 2, \dots, r); \\ (v_{i_1}, v_{i_2}, \dots, v_{i_r}) \in E(G)\}$$

where the multiplicity of (w_1, w_2, \dots, w_r) is defined to be that of $(v_{i_1}, v_{i_2}, \dots, v_{i_r})$.

Definition 4. Let G^m be an (r, q) -digraph. Among all vectors $\mathbf{x} = (x_1, x_2, \dots, x_m)$ for which

$$n = x_1 + x_2 + \dots + x_m, \\ 0 \leq x_i \quad (i = 1, 2, \dots, m)$$

is a partition of n into non-negative integers, those for which the number of edges of $G(\mathbf{x})$ is maximized will be called the *optimal vectors* associated with the corresponding *optimal graph* $G(\mathbf{x})$. Any such optimal graph may be denoted by $G(n)$.

Definition 5. Let m, r, q be positive integers, and $G = G^m$ be an (r, q) -digraph. Let the vector $\mathbf{u} = (u_1, u_2, \dots, u_m)$ range over the standard $(m - 1)$ -simplex in \mathbf{R}^m , i.e. $u_i \geq 0$ ($i = 1, 2, \dots, m$), $\sum_i u_i = 1$. We consider the real multilinear form

$$f_G(\mathbf{u}) = \sum u_{i_1} u_{i_2} \dots u_{i_r}$$

⁵ The sequence is indexed by n , the number of vertices.

summed over all (i_1, i_2, \dots, i_r) such that $(v_{i_1}, v_{i_2}, \dots, v_{i_r})$ is an edge of G , where the multiplicity of a monomial in the sum is equal to the multiplicity of the corresponding edge in the graph. The maximum of $g_G(\mathbf{u})$ is called the *graph-density* or simply the *density* of G , and denoted by $g(G)$; a vector \mathbf{u} for which $f_G(\mathbf{u})$ is maximal is called an *optimum vector*. Where the maximum is attained only in the interior of the $(m - 1)$ -simplex, i.e. with $u_i > 0$ ($i = 1, 2, \dots, m$), we say that G is *dense*.

Remark. The variables in $f_G(\mathbf{u})$ are commutative. Thus the coefficient of $u_{i_1} u_{i_2} \dots u_{i_r}$ is the sum of the multiplicities of all edges that are permutations of $\{v_{i_1}, v_{i_2}, \dots, v_{i_r}\}$: it does not depend on the orientation.

Example. Let $q = 3$, and let G be a 3-uniform hypergraph with $V(G) = \{1, 2, 3, 4\}$, and $E(G) = \{(123), (123), (213), (213), (213), (124)\}$ (where multiplicities have been shown by repetition.) Then, with $\mathbf{u} = (u_1, u_2, u_3, u_4)$, $u_i \geq 0$ ($i = 1, 2, 3, 4$), $f_G(\mathbf{u}) = 5u_1u_2u_3 + u_1u_2u_4$. Since $f_G(\mathbf{u}) \leq f_G(u_1, u_2, u_3 + u_4, 0)$, G is *not dense*.

Lemma 1.⁶ Let G be a fixed (r, q) -graph and let t be a positive integer.

(a) The number of edges of $G\langle te \rangle$ is $t^r e(G)$.

(b) As $n \rightarrow \infty$,

$$e(G(n)) = \{g(G) + O(1/n)\}n^r. \tag{2.2}$$

(c) There exists a constant $c_1 = c_1(r, q)$ such that

$$g(G) - c_1/n < e(G(n))/n^r \leq g(G). \tag{2.3}$$

(d) For any vector \mathbf{x} of positive integers, and any positive integer n , $(G(\mathbf{x}))(n) = G(n)$. Moreover, $g(G(\mathbf{x})) = g(G)$.

(e) If H is a subgraph of G , then $g(H) \leq g(G)$.

(f) Let G^n be a digraph containing a subgraph H^m for which $e(H^m) \geq am^r$. Then $g(G^n) \geq a$. In particular, if $e(G^n) > an^r$, then $g(G^n) > a$.

Proof. (b) Let \mathbf{u} be an optimum vector for $G = G^n$, and define \mathbf{x} by $x_i = \lfloor u_i n \rfloor$ or $\lceil u_i n \rceil$ ($i = 1, 2, \dots, m$) chosen⁷ in some way so that $\sum_i x_i = n$. Then

$$\begin{aligned} x_{i_1} x_{i_2} \dots x_{i_r} &= \prod_j (u_j n + (x_j - u_j n)) \\ &= n^r u_{i_1} u_{i_2} \dots u_{i_r} + O(n^{r-1}) \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Thus, as $n \rightarrow \infty$, $|n^{-r} e(G(\mathbf{x})) - g(G)| = O(1/n)$. Conversely, given an optimal vector $\mathbf{y} = (y_1, y_2, \dots, y_m)$ realizing $G(n)$, define a vector \mathbf{v} by $v_i = y_i/n$ ($i = 1, 2, \dots, m$). Then $g(G) \geq f_G(\mathbf{v}) = n^{-r} e(G(\mathbf{y}))$.

⁶ We number the theorems and lemmas proved in the present paper using arabic numerals, and the results quoted without proof from other sources with Latin letters.

⁷ $\lfloor x \rfloor$ denotes the greatest integer in x , $\lceil x \rceil$ denotes $-\lfloor -x \rfloor$.

(d) The first statement is trivial. The second statement follows from the first by (b).

(f) Let t be any positive integer. Then

$$\begin{aligned} g(G) &\geq g(H^m) \\ &= (mt)^{-r}e(H^m(mt))(1 + O(1/t)) \\ &\geq (mt)^{-r}e(H^m\langle te \rangle)(1 + O(1/t)) \\ &= m^{-r}e(H^m)(1 + O(1/t)) \geq a(1 + O(1/t)) \quad \text{as } t \rightarrow \infty. \end{aligned}$$

Remarks. (1) This approach to extremal graph-theoretic problems via a quadratic form associated with the adjacency matrix was pioneered by T. Motzkin and E. Straus (cf. [16]). Straus⁸ and others have considered possible extensions of the technique to hypergraph extremal problems.

(2) In our studies on digraphs and multigraphs ([2], [5], etc.) we approached certain extremal problems using the vehicle of ‘canonical graph structures’: sequences of graphs whose structure may be represented by a finite number of integer-valued parameters. For a precise description the reader is referred to Section 8 below; cf. also [2], [5].

Definition 6. For fixed r and q the set of attained densities will be denoted by \mathcal{D}_g .

3. Infinite sets of prohibited graphs: continuity and compactness problems

The following result has been proved for ordinary graphs [11], digraphs with $q = 1$ [6, Theorem 3], and multigraphs with $q = 2$ [6, Corollary to Theorem 3]; we conjecture that it holds in general.

Conjecture 1 (Compactness). Let r and q be positive integers, and \mathbf{L} an arbitrary family of (r, q) -graphs. There exists a finite subfamily $\mathbf{L}^* \subseteq \mathbf{L}$ such that

$$ex(n, \mathbf{L}) - ex(n, \mathbf{L}^*) = o(n^r) \quad \text{as } n \rightarrow \infty. \tag{3.1}$$

While Conjecture 1 remains open (with the exceptions mentioned), we are able to prove the following weaker result.

Theorem 1 (Continuity). Let r and q be positive integers and \mathbf{L} an arbitrary family of (r, q) -graphs. To each $\varepsilon > 0$ there exists a finite subfamily $\mathbf{L}_\varepsilon \subseteq \mathbf{L}$ for which

$$ex(n, \mathbf{L}) \leq ex(n, \mathbf{L}_\varepsilon) < ex(n, \mathbf{L}) + \varepsilon n^r \tag{3.2}$$

for n sufficiently large.

⁸ Oral communication, also [23].

Remark. For ordinary graphs the truth of the conjecture is a consequence of the Erdős–Stone–Simonovits theorem (2.1) cited above [11]. For digraphs Theorem 1 was proved in [5]; subsequently, the conjecture was proved [6] only for digraphs with $q = 1$. The proof below is much shorter than our earlier proof of that special case.

Definition 7. The *edge-density* of an (r, q) -graph G^n is defined to be the ratio, $e(G^n)/n(n-1) \cdots (n-r+1)$, i.e. it is the average multiplicity of all possible oriented edges.

We require the following lemma – using an argument of Katona, Nemetz, and Simonovits [15, First Corollary to Theorem 1].

Lemma 2. (a) Let G^n be an arbitrary graph, $m \leq n$; let $h = \binom{n}{m}$. Denote by H_1, \dots, H_h all the spanned (= induced) subgraphs of G^n having exactly m vertices. Then

$$\frac{1}{h} \sum_{i=1}^h \frac{e(H_i)}{m(m-1) \cdots (m-r+1)} = \frac{e(G^n)}{n(n-1) \cdots (n-r+1)} \tag{3.3}$$

or, equivalently,

$$\binom{n}{m}^{-1} \sum_i e(H_i) \binom{m}{r}^{-1} = e(G^n) \binom{n}{r}^{-1}. \tag{3.4}$$

In other words, the average of the edge-densities of the m -vertex subgraphs of G^n is equal to the edge-density of G^n .

(b) The ratio $\text{ex}(n, \mathbf{L})/\binom{n}{r}$ decreases monotonely as n increases.

Proof. (a) Since each r -edge of G^n is counted exactly $\binom{n-r}{m-r}$ times,

$$\sum_i e(H_i) = \binom{n-r}{m-r} e(G^n). \tag{3.5}$$

Then (3.3) follows from the identity

$$\binom{n}{r} \binom{n-r}{m-r} = \binom{n}{m} \binom{m}{r}.$$

(b) We apply (a). Let $m < n$. None of the m -vertex spanned subgraphs H_i of an extremal graph $G^n \in \text{EX}(n, \mathbf{L})$ contains any $L \in \mathbf{L}$, so $e(H_i) \leq \text{ex}(m, \mathbf{L})$. The left side of (3.4) is the average of terms, none of which exceeds $\text{ex}(m, \mathbf{L})/\binom{m}{r}$.

Corollary 1 (to Lemma 2). The sequence $\{\text{ex}(n, \mathbf{L})/\binom{n}{r}\}_{n=1,2,\dots}$ converges (monotonely). The sequence $\{\text{ex}(n, \mathbf{L})/n^r\}_{n=1,2,\dots}$ converges.

Corollary 2 (to Lemma 2). Let m, r be positive integers, $a > 0$. If $\mathbf{L} =$

$\{H^m \mid e(H^m) > a(r^m)\}$, then

$$\text{ex}(n, \mathbf{L}) \leq a \binom{n}{r} \quad \text{for all } n \geq m.$$

Proof of Theorem 1. Let \mathbf{L} be a family of graphs satisfying the hypotheses; for each positive integer k let \mathbf{L}_k denote the family consisting of the members of \mathbf{L} having at most k vertices. Let

$$\gamma_k = \lim_{n \rightarrow \infty} \text{ex}(n, \mathbf{L}_k) / \binom{n}{r} \quad (k = 1, 2, \dots)$$

and

$$\gamma = \lim_{n \rightarrow \infty} \text{ex}(n, \mathbf{L}) / \binom{n}{r}$$

Assume that, for some $\varepsilon > 0$, $\gamma_k > \gamma + \varepsilon$ for all k ; (since $\mathbf{L}_k \subseteq \mathbf{L}$, $\gamma_k \geq \gamma$ for all k). Let S_k^n be an extremal graph in $\text{EX}(n, \mathbf{L}_k)$. By Lemma 2,

$$\begin{aligned} e(S_k^n) &= \text{ex}(n, \mathbf{L}_k) \\ &\geq \gamma_k \binom{n}{r} > (\gamma + \varepsilon) \binom{n}{r} \quad \text{for every } n, k. \end{aligned} \quad (3.6)$$

In particular, taking $k = n$, we have

$$e(S_n^n) > (\gamma + \varepsilon) \binom{n}{r} \quad \text{for all } n. \quad (3.7)$$

The graph S_n^n contains no subgraphs from \mathbf{L}_n ; as it has exactly n vertices, it can contain no member of \mathbf{L} either! Thus

$$\text{ex}(n, \mathbf{L}) \geq e(S_n^n) > (\gamma + \varepsilon) \binom{n}{r}. \quad (3.8)$$

In the limit as $n \rightarrow \infty$ we obtain a contradiction to our definition of γ . We conclude that $\lim \gamma_k = \gamma$ as $k \rightarrow \infty$.

4. 'Supersaturated' graphs

A graph G^n may be considered 'saturated' with respect to a given family \mathbf{L} of prohibited subgraphs if it contains no member of \mathbf{L} , but has the maximum number of edges among graphs with that property – i.e. if it is extremal. When the number of edges exceeds $\text{ex}(n, \mathbf{L})$, we may ask *how many* distinct copies of members of \mathbf{L} are present in G^n . (Of course, the graph must not necessarily be thought of as having been built up from a member of $\text{EX}(n, \mathbf{L})$ through the addition of edges.) A corpus of results on 'supersaturated' graphs exists for ordinary graphs [18]. Erdős and Simonovits [12], also Simonovits [18], have investigated properties of

'supersaturated' hypergraphs. The main theorem below – which will be applied in our proof of Theorem 3 – is in that *genre*.

Theorem 2. *Let \mathbf{L} be an arbitrary family of (r, q) -hypergraphs, and let $\gamma = \lim \text{ex}(n, \mathbf{L})/n^r$ as $n \rightarrow \infty$. Let $\varepsilon > 0$. There exists a constant $c_2 = c_2(\mathbf{L}, \varepsilon)$ such that, if $e(G^n) > (\gamma + \varepsilon)n^r$ and if n is sufficiently large, then there exists some $L^1 \in \mathbf{L}$ such that G^n contains at least $c_2 n^1$ copies of L^1 .*

Proof. (This theorem was proved for *one undirected r -uniform l -hypergraph* by Erdős and Simonovits [12].) Assume $\gamma + \varepsilon < q$. By Theorem 1 there exists a finite subfamily $\mathbf{L}^* = \mathbf{L}_{e/\beta} \subseteq \mathbf{L}$ such that

$$\text{ex}(m, \mathbf{L}^*) < (\gamma + \frac{1}{2}\varepsilon)m^r + o(m^r) \quad \text{as } m \rightarrow \infty.$$

Thus

$$\text{ex}(m, \mathbf{L}^*) < (\gamma + \frac{1}{2}\varepsilon)m(m-1) \cdots (m-r+1) \quad \text{if } m > m_0.$$

Assume that $e(G^n) > (\gamma + \varepsilon)n^r$. Let $a(m, n)$ be the number of spanned subgraphs H^m of G^n on exactly m vertices and such that $e(H^m) > (\gamma + \frac{1}{2}\varepsilon)m(m-1) \cdots (m-r+1)$. We may apply (3.4), where no summand on the left exceeds $qr!$, to show that $a(m, n) > c_3 n^m$ for some $c_3 = c_3(m) > 0$ and for n sufficiently large (with respect to m). Now fix $m > m_0$: then at least $c_3 n^m$ of the H^m must each contain an $L \in \mathbf{L}^*$, though not necessarily the same L . The finiteness of \mathbf{L}^* ensures that some one L^1 is contained in at least $c_2 n^m$ of these subgraphs. None of these copies of L^1 could be counted more than $\binom{n-1}{m-1}$ times. Thus the number of distinct copies of L^1 in G^n is at least

$$c_2 n^m / \binom{n-1}{m-1} > c_2 n^1.$$

5. A theorem of 'Erdős–Stone' type

A celebrated theorem of Erdős and Stone [8], subsequently refined by many authors, relates the extremal numbers of (ordinary) complete k -graphs K_k to those of $K_k \langle te \rangle$ for positive integers t . In general, one may consider, for *any* graph G and positive integer t , the graph $G \langle te \rangle$ obtained through replacement of each vertex by t independent vertices. The Erdős–Stone Theorem [8] states that, for any t ,

$$\text{ex}(n, K_k \langle te \rangle) = \text{ex}(n, K_k) + o(n^2) \quad \text{as } n \rightarrow \infty.$$

This surprising result implies the 'Erdős–Stone–Simonovits' Theorem [10], of wide applicability.

Erdős generalized the Erdős–Stone Theorem in two stages, Theorem A and Theorem B below. We shall continue the generalization – to (r, q) -hypergraphs.

Definition 8. Let r, k, t be positive integers. $K_k^{(r)}$ will denote the *unoriented* (sic) ‘complete’ r -uniform² hypergraph having k vertices and, as edges, all $\binom{k}{r}$ possible r -tuples of those vertices. By $K_{k,t}^{(r)}$ we shall denote the r -graph $K_k^{(r)}(\langle te \rangle) = K_k^{(r)}(t, t, \dots, t)$ having kt vertices in k classes of t : the edges are all $\binom{k}{r}t^r$ selections of one vertex from each of the k classes. (The superscript $^{(r)}$ may be suppressed.)

Theorem A⁶ [9, Theorem 1]. *Let n, r, t be positive integers. There exists a constant c (independent of n, r, t) such that*

$$n^{r-ct^{r+1}} < \text{ex}(n, K_{n,t}^{(r)}) \leq n^{r-t^{r+1}}.$$

This he subsequently generalized in

Theorem B [10]. *Let k, r, t be fixed positive integers. Then*

$$\text{ex}(n, K_{k,t}^{(r)}) = \text{ex}(n, K_k^{(r)}) + o(n^r) \quad \text{as } n \rightarrow \infty.$$

We shall apply these results and our Theorem 2 to prove the following generalization of the preceding to (r, g) -graphs.

Theorem 3. *Let \mathbf{L} be an arbitrary family of ‘prohibited’ graphs, and $f: \mathbf{L} \rightarrow \mathbb{N}$ a mapping into the natural numbers. Let $\mathbf{L}^f = \{L \langle f(L)e \rangle : L \in \mathbf{L}\}$. Then*

$$\text{ex}(n, \mathbf{L}^f) = \text{ex}(n, \mathbf{L}) + o(n^r) \quad \text{as } n \rightarrow \infty.$$

Proof. Since $L \subseteq L \langle te \rangle$ for each $L \in \mathbf{L}$ and any positive integer t ,

$$\lim_{n \rightarrow \infty} \text{ex}(n, \mathbf{L}) \leq \lim_{n \rightarrow \infty} \text{ex}(n, \mathbf{L}^f).$$

We proceed to prove the opposite inequality. Define $\gamma = \lim_{n \rightarrow \infty} \text{ex}(n, \mathbf{L})/n^r$. Let $\varepsilon > 0$ be given. By Theorem 1 there exists a finite subfamily $\mathbf{L}^* \subseteq \mathbf{L}$ such that

$$\text{ex}(n, \mathbf{L}^*) < (\gamma + \frac{1}{2}\varepsilon)n^r + o(n^r) \quad \text{as } n \rightarrow \infty.$$

Let $t = \text{Max}\{f(L) : L \in \mathbf{L}^*\}$, and let G^n be given, with $e(G^n) > (\gamma + \varepsilon)n^r$. Then, by Theorem 2, G^n contains at least $c_2(\mathbf{L}^*, \varepsilon)n^l$ copies of some $L^1 \in \mathbf{L}^*$. Define an *unordered* (sic) l -uniform l -hypergraph M on the vertex set $V(G^n)$: the edges are precisely the vertex sets of the copies of L^1 . For n sufficiently large⁹ with respect to T, ε , Theorem A ensures the existence of l disjoint sets each of T vertices, A_1, A_2, \dots, A_l , such that G^n contains all T^l graphs of structure L each having precisely one vertex in each of A_1, A_2, \dots, A_l . Let v_1, v_2, \dots, v_l be some enumeration of the vertices of L^1 . The T^l embeddings of these vertices into distinct A_i ($i = 1, 2, \dots, l$) induce permutations of $\{1, 2, \dots, l\}$. Some permutation occurs at least $T^l/l!$ times: we now consider the l -uniform ordered hypergraph defined on the lT vertices of $\bigcup_i A_i$ by the copies of L associated with *this* permutation. For T

⁹ It suffices to take $\log T > 2l^l \log l$ and $n > c_2(L^*, \varepsilon)^{-T^{l-1}}$.

sufficiently large⁹ with respect to l and t , a second application of Theorem A shows that there must exist an $L^1\langle te \rangle$, hence an $L^1\langle f(L^1)e \rangle$. Thus

$$\limsup_{n \rightarrow \infty} \frac{\text{ex}(n, \mathbf{L}^f)}{n^r} \leq \gamma.$$

6. The set of attained densities

For ordinary graphs the possible densities are of the form $\frac{1}{2}(1-1/p)$, ($p = 1, 2, \dots$) [16]. The set of these values coincides with the set of limits of form (1.1) [11] with $r = 2$. For digraphs with $q = 1$ we have investigated properties of the sets of these densities and limits and have proved [6] (cf. [2, Conjecture 2*]), that the densities form a *well-ordered* set, but a number of questions remain even for digraphs with $q > 1$; an analogous situation holds for multigraphs. We prove below a general inclusion theorem for (r, q) -hypergraphs, then state a conjecture, and show (in Theorem 5) that it has several equivalent forms.

Definition 9. For any family \mathbf{L} , $\lim \text{ex}(n, \mathbf{L})/n^r$ as $n \rightarrow \infty$ is called an *extremal*¹⁰ density. The set of extremal densities (for a fixed class of objects, and fixed r and q), will be denoted by \mathcal{D}_e . (Compare Definition 6.)

Theorem 4. $\mathcal{D}_g \subseteq \mathcal{D}_e$. Moreover, \mathcal{D}_g is dense in \mathcal{D}_e .

Our proof of Theorem 4 will require the following lemma, which we state without proof.

Lemma 3. Let γ be a positive real number, $0 < \gamma \leq q$, and let m be a positive integer. Define

$$\mathbf{L}_{m,\gamma} = \{L = L^m : e(L^m) > \gamma m(m-1) \cdots (m-r+1)\}. \tag{6.1}$$

Then

- (i) $\text{ex}(n, \mathbf{L}_{m,\gamma}) \leq \gamma n(n-1) \cdots (n-r+1) \leq \gamma n^r$ for all $n \geq m$.
- (ii) $\lim \text{ex}(n, \mathbf{L}_{m,\gamma})/n^r \leq \gamma$ as $n \rightarrow \infty$.
- (iii) For any $L \in \mathbf{L}_{m,\gamma}$, $g(L) > \gamma m(m-1) \cdots (m-r+1)/m^r$.

Proof of Theorem 4. (A) Let $\gamma \in \mathcal{D}_g$. There exists a graph G for which $\gamma = g(G)$. Let \mathbf{L} be the family of all graphs H for which $g(H) > \gamma$. We propose to prove that $\text{ex}(n, \mathbf{L})/n^r \rightarrow \gamma$ as $n \rightarrow \infty$. For any vector \mathbf{x} , $G\langle \mathbf{x} \rangle$ cannot contain a subgraph of density greater than γ (Lemma 1); hence

$$\text{ex}(n, \mathbf{L}) \geq e(G(n)) = (1 + O(1/n))n^r \gamma \quad \text{as } n \rightarrow \infty.$$

¹⁰ We have thus introduced three 'densities' the *graph-density* (Definition 5), the *edge-density* (Definition 7), and the *extremal density* (of a family, Definition 9). Compare [24, Lemma 2].

Suppose now that $H^n \in EX(n, \mathbf{L})$, $(n = 1, 2, \dots)$. Then, since $H^n \notin \mathbf{L}$, $g(H^n) \leq \gamma$. Applying Lemma 1(f), we may conclude that $e(H^n) \leq \gamma n^r$, i.e. $ex(n, \mathbf{L}) \leq \gamma^r$. Thus $\lim_{n \rightarrow \infty} ex(n, \mathbf{L})/n^r = \gamma$, so $\gamma \in \mathcal{D}_e$.

(B) Let \mathbf{L} be a family of prohibited graphs such that $ex(n, \mathbf{L})/n^r \rightarrow \gamma$ as $n \rightarrow \infty$. We shall prove that γ is the limit of a sequence of graph densities. (By virtue of Conjecture 2 below, we seek a sequence of graphs whose densities approach γ from below.) Let $\{S^n\}$ be a sequence of extremal digraphs for \mathbf{L} ($n = 1, 2, \dots$). Let $\varepsilon > 0$ be given, sufficiently small, and let m be an integer such that $m(m-1) \cdots (m-r+1)/m^r > 1 - \varepsilon$. We consider the family $\mathbf{L}' = \mathbf{L}_{m, \gamma(1-\varepsilon)}$ of Lemma 3, and apply Theorem 3 with the constant function $f_t(L) = t$ for all $L \in \mathbf{L}'$. For n sufficiently large, all S^n will have $e(S^n) \geq \gamma(1 - \frac{1}{2}\varepsilon)n^r$ and will contain $L \langle te \rangle$ for some $L \in \mathbf{L}'$. The preceding is true for any t . Hence some L_1 in \mathbf{L}' has the property that $L_1(t) \subseteq S^n$ for arbitrary large t and $n = n(t)$. Since S^n contains no L in \mathbf{L} , $e(L_1(t)) \leq ex(t, \mathbf{L})$ for all t . Hence, by (2.2), $g(L_1) \leq \gamma$. But $g(L_1) \geq \gamma(1 - \varepsilon)$, (by definition). It follows that γ is the limit of a sequence¹¹ of graph densities.

Conjecture 2. (a) \mathcal{D}_e is well ordered (under the usual ordering of the reals).

(b) \mathcal{D}_g is well ordered (under the usual ordering of the reals).

(c) (Compactness) For every infinite family \mathbf{L} there exists a finite subfamily $\mathbf{L}^* \subseteq \mathbf{L}$ for which

$$ex(n, \mathbf{L}^*) - ex(n, \mathbf{L}) = o(n^r) \quad \text{as } n \rightarrow \infty. \tag{6.2}$$

Theorem 5. For fixed r and q , conditions (a), (b), and (c) of Conjecture 2 are equivalent.

Remark. It has been shown in [6] that digraphs with $q = 1$ have properties (a), (b), (c) above.

Proof of Theorem 5. (A) By Theorem 4, \mathcal{D}_g is contained in \mathcal{D}_e and dense in \mathcal{D}_e : (a) and (b) are equivalent.

(B) Let us assume that \mathcal{D}_g and \mathcal{D}_e are well ordered. Let \mathbf{L} be an arbitrary family of prohibited graphs. As in the proof of Theorem 1, we denote by \mathbf{L}_k the subfamily of graphs in \mathbf{L} having at most k vertices; and by γ_k and γ the limits of $ex(n, \mathbf{L}_k)/\binom{n}{r}$ and $ex(n, \mathbf{L})/\binom{n}{r}$ as $n \rightarrow \infty$, ($k = 1, 2, \dots$). The sequence γ_k is monotonely decreasing since $ex(n, \mathbf{L}_k)$ is monotonely decreasing in k . Furthermore, Theorem 1 ensures that $\gamma_k \rightarrow \gamma$ as $n \rightarrow \infty$. As the set $\{\gamma_k : k \in \mathbf{N}\}$ is well ordered, $\gamma_k = \gamma$ for k sufficiently large. With such a k take $\mathbf{L}^* = \mathbf{L}_k$. Thus (a) implies (c).

(C) Assume now that (c) holds, but that (a) does not. Let $\{G_k\}_{k=1,2,\dots}$ be a sequence of graphs such that

$$g(G_k) > \gamma \quad (k = 1, 2, \dots) \tag{6.3}$$

¹¹ We do not claim that it is the limit point of the set \mathcal{D}_g . It may be an isolated point in \mathcal{D}_g as well.

and

$$g(G_i) \rightarrow \gamma \text{ as } i \rightarrow \infty. \tag{6.4}$$

Define \mathbf{L} to be the family of all graphs having density greater than γ (thus $G_i \in \mathbf{L}$ for all i), and let $\gamma^* = \lim_{n \rightarrow \infty} \text{ex}(n, \mathbf{L})/n^r$. We prove that $\gamma^* \leq \gamma$. If S^n is an extremal graph in $\text{EX}(n, \mathbf{L})$, then, by Lemma 1(f), $g(S^n) \geq e(S^n)/n^r = \text{ex}(n, \mathbf{L})/n^r$; in the limit, $\gamma \geq \gamma^*$. If (c) is true, there exists a finite subfamily \mathbf{L}^* of \mathbf{L} such that (6.2) holds. Take a $G_k \notin \mathbf{L}^*$. By (6.3) and Lemma 1, for sufficiently large n , $G_k(n)$ must contain some L_i in \mathbf{L}^* . Hence, by Lemma 1,

$$g(G_k) = g(G_k(n)) \geq g(L_i) > \gamma. \tag{6.5}$$

Since \mathbf{L}^* is finite, there is some i for which (6.5) holds for infinitely many k . Hence $\gamma \geq g(G_i) > \gamma$, which is a contradiction. We conclude that (c) implies (a).

Remark. In the cases of ordinary graphs [11], and of oriented (2, 1)-graphs and unoriented (2, 2)-graphs [6], we know that $\mathcal{D}_g = \mathcal{D}_e$. The general question, however, remains open.

7. An ‘approximation’ theorem

For a given family \mathbf{L} of prohibited graphs, our ideal objective would be to determine the family $\text{EX}(n, \mathbf{L})$ of *extremal* graphs. That being usually unattainable, we enquire as to the structure of *asymptotically extremal* sequences. For graphs, digraphs with $q = 1$, and multigraphs with $q = 2$, we have proved the *existence* (cf. Section 9 below) of asymptotically extremal sequences of an ‘easily describable structure’, ([2, Theorem 1]). For (r, q) -graphs in general we prove below a somewhat weaker theorem.

Theorem 6 (Approximation Theorem). *Let \mathbf{L} be a given family of prohibited graphs, and let $\epsilon > 0$. There must exist a graph G such that $G(n)$ contains no $L \in \mathbf{L}$ and $e(G(n)) > \text{ex}(n, \mathbf{L}) - \epsilon n^r$ for every n sufficiently large.*

Proof. Let $\lim_{n \rightarrow \infty} \text{ex}(n, \mathbf{L}) = \gamma$. Let $\{S^n\}$ be an extremal sequence for \mathbf{L} . As in paragraph (B) of the proof of Theorem 4, there exists, for m sufficiently large, $G \in \mathbf{L}_{m, \gamma - \epsilon/2}$ for which $G(n) \subseteq S^{n'}$ for infinitely many n and $n' = n'(n)$. $G(n)$ contains no $L \in \mathbf{L}$ and $g(G) > \gamma - \frac{1}{2}\epsilon$.

Example. Let H consist of all members of a sequence of unoriented 3-graphs $\{H^n\}$, defined recursively as follows: H^1, H^2, H^3 all have no triples. If H^n has been defined for all $n < N$, H^N is formed by taking three disjoint 3-graphs, $H^{[N/3]}$, $H^{[(N+1)/3]}$, $H^{[(N+2)/3]}$, and adjoining as new edges all triples having precisely one vertex in each of the three ‘summands’. Then it can be seen that $\lim_{n \rightarrow \infty} e(H^n)/n^3 = 1/24$. There exists no graph G such that $H^n = G(n)$ for all

large n . (For, in such sequences of optimal G -graphs the ratio |maximum independent set of vertices|/ n tends to a positive limit, whereas in the present case it tends to zero.) Nor could the sequence $\{H^n\}$ be obtained from a sequence $\{G(n)\}$ by adding or deleting $o(n^3)$ edges. However, it can be shown that these graphs are extremal for *some* family \mathbf{L} . For such an \mathbf{L} and $\varepsilon > 0$, there exists an integer k such that $G = H^{3k}$ satisfies the conditions of Theorem 6.

8. Unoriented r -uniform q -hypergraphs

Let the set-valued operator \mathfrak{D} applied to an *unoriented* (r, q) -graph map it onto the set of all *oriented* r -uniform q -hypergraphs obtainable through independent orientations of each of its r -edges. Then we have the following

Lemma 4 ('Transfer Principle'). *Let \mathbf{M} be an arbitrary family of r -uniform unoriented q -hypergraphs. Then $\mathfrak{D}\text{EX}(n, \mathbf{M}) = \text{EX}(n, \mathfrak{D}\mathbf{M})$.*

This permits the passage from the results of this paper, stated for the *oriented* case, to the unoriented. In certain cases (e.g. proof of theorem 1) the steps in the proof themselves have to be checked to determine whether the family \mathbf{L}_ε constructed is 'symmetric', i.e. whether it is the image under \mathfrak{D} of a family of multigraphs. The proofs of Theorem 1 through Theorem 6 can all be seen to have this property.

9. (r, q) -graphs with loops permitted

Portions of the theory of this paper carry over without significant change when we permit loops to occur. An 'edge' of such a general (r, q) -graph G^n will be any one of the n^r points in the cartesian product $(V(G))^r$. It is particularly useful to permit such loops if we wish to obtain extremal sequences or asymptotically extremal sequences of form $\{G(n)\}_n$: by excluding loops we would have to exclude many important cases. We may generalize Definition 3 and Definition 4 in the obvious way to define for any (r, q) -graph G^n with loops permitted and any non-negative integer vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ a graph $G(\mathbf{x})$. The orientation of edges is well defined in terms of that of the edges of G , except insofar as edges having more than one vertex from the same class X_i . This construction is also meaningful when G is unoriented.

We may also wish to permit loops in G but not in the graphs $G(\mathbf{x})$. We have the following familiar example.

Example. Let G be the *unoriented* $(3, 1)$ -graph defined on the set $\{u, v, w\}$

with edgese¹² $\{uuv, vvw, wwu, uvw\}$. Turán conjectured that the graphs $G(n)$ (without loops) are extremal for the 3-uniform complete 4-graph, $K_4^{(3)}$, (that is, the complete unoriented (3, 1)-graph with $V(K_4^{(3)}) = \{a, b, c, d\}$, and $E(K_4^{(3)}) = \{abc, abd, acd, bcd\}$. (Other graphs G' are known [1] such that $e(G'(n)) = e(G(n))$ where $G'(n)$ does not contain $K_4^{(3)}$.)

A sequence $\{S^n\}$ of digraphs which may be interpreted as being of form $G(x)$ for fixed G may be called *canonical*. This concept bears fruit particularly where the graph G is permitted to have loops. In our papers on digraphs and multi-graphs [2], [5], [6], we exploited a restriction of this concept in *matrix digraphs*, which are canonical sequences of optimal graphs $G(n)$ where a particular orientation is placed on the edges connecting more than one (i.e. two) vertices of a class X_i . For such matrix digraphs we have been able to prove an 'inverse' theorem to Theorem 6 [5, Theorem 1, Theorem 3].

References

- [1] W.G. Brown, On an open problem of Paul Turán concerning 3-graphs, *Studies in Pure Mathematics, to the Memory of Paul Turán* (Birkhäuser Verlag, Zurich; Akadémiai Kiadó, Budapest) 91–93.
- [2] W.G. Brown, P. Erdős and M. Simonovits, Extremal problems for directed graphs, *J. Combin. Theory (B)* 15 (1973) 77–93. MR 52#7952.
- [3] W.G. Brown, P. Erdős and M. Simonovits, Multigraph extremal problems, *Colloques Internationaux CNRS, 260: Problèmes Combinatoires et Théorie des Graphes, Proc. Colloq. Orsay, 1976* (Editions du CNRS, Paris, 1978) 63–66.
- [4] W.G. Brown, P. Erdős and M. Simonovits, Multigraph extremal problems, Preprint (1975), about 180 pages.
- [5] W.G. Brown, P. Erdős and M. Simonovits, Inverse extremal digraph problems, *Colloq. Math. Soc. János Bolyai m* (North-Holland, Amsterdam, 198n).
- [6] W.G. Brown, P. Erdős and M. Simonovits, Algorithmic solution of extremal digraph problems, Preprint, McGill University (1982).
- [7] W.G. Brown and F. Harary, Extremal digraphs, in: P. Erdős et al., eds., *Combinatorial Theory and its Applications, Vol. I, Coll. Math. Soc. J. Bolyai 4* (North-Holland, Amsterdam, 1970) 135–198. MR 45#8576.
- [8] P. Erdős and A.H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* 52 (1946) 1087–1091. MR 8, 333b.
- [9] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* 2 (1964) 174–181. MR 32#1134.
- [10] P. Erdős, On some extremal problems on r -graphs, *Discrete Math.* 1 (1971) 1–6. MR 45#6656.
- [11] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* 1 (1966) 51–57. MR 34#5702.
- [12] P. Erdős and M. Simonovits, Oversaturated graphs and hypergraphs, *Combinatorica*, to appear.
- [13] P. Erdős, some recent results on extremal problems in graph theory (Results), *Theory of Graphs: International Symposium, Rome, 1966* (Gordon & Breach, New York; Dunod, Paris, 1967) 118–123. MR 37#2634.
- [14] P. Erdős, On some new inequalities concerning extremal properties of graphs, *Theory of Graphs, Proc. Colloq. Tihany, 1966* (Academic Press, New York–London; Akadémiai Kiadó, Budapest, 1968) 77–81. MR 38#1026.

¹² For convenience we represent the edges as words in the alphabet of vertices.

- [15] Gy. Katona, T. Nemetz and M. Simonovits, Ujabb bizonyítás a Turán-féle gráftétele, és megjegyzések bizonyos általánosításaira (On a problem of Turán in the theory of graphs), *Matematikai Lapok* 15 (1964) 228–238. MR 30#2483.
- [16] T.S. Motzkin and E.G. Straus, Maxima for graphs and a new proof of a theorem of Turán, *Canad. J. Math.* 17 (1965) 533–540. ME 31#89.
- [17] M. Simonovits, A method for solving problems in graph theory, stability problems, *Theory of Graphs*, Proc. Colloq. Tihany, 1966 (Academic Press, New York–London; Akadémiai Kiadó, Budapest, 1968) 279–319. MR 38#2056.
- [18] M. Simonovits, Extremal graph problems, degenerate extremal problems, and supersaturated graphs, Proc. Silver Jubilee Conf. Univ. of Waterloo, 1982.
- [19] V.T. Sós, P. Erdős and W.G. Brown, On the existence of triangulated spheres in 3-graphs, and related problems, *Periodica Math. Hung* 3 (1973) 221–228. MR 48#2003.
- [20] P. Turán, Egy gráfelméleti szélsőérték feladatáról, *Mat. Fiz. Lapok* 48 (1941) 436–452. MR 8, 284j.
- [21] P. Turán, On the theory of graphs, *Colloq. Math.* 3 (1954) 19–30. MR 15, 976b.
- [22] P. Turán, Research Problems, *Magyar Tud. Akad. Kutató Int. Közl.* 6 (1961) 417–423. MR 31#2107.
- [23] V.T. Sós and E.G. Straus, Extremals of functions on graphs, with applications to graphs and hypergraphs, *J. Combin. Theory (B)* 32 (1982) 246–257.
- [24] Gy. Katona, Continuous versions of some extremal hypergraph problems, *Colloq. Math. Soc. Janos Bolyai* 18, Combinatorics, Proc. Fifth Hungarian Colloq. (Keszthely, 1976) II, 653–678. MR 80e = 05071.