

CHAPTER 4: EXTREMAL SET THEORY

Thomas Britz

Classical Theorems

Let \mathcal{A} be a family of subsets of $[n]$. Given imposed constraints, we can ask how big \mathcal{A} can be, how large it must be, and what structure \mathcal{A} must have if it is largest or smallest possible. These are questions typically asked in extremal set theory. This chapter presents a selection of nice and pretty answers.

Theorem

If $A \cap B \neq \emptyset$ for all $A, B \in \mathcal{A}$, then $|\mathcal{A}| \leq 2^{n-1}$.
There is equality if and only if \mathcal{A} contains A or A^C for each $A \subseteq [n]$.

Proof. If $A \in \mathcal{A}$ for some $A \subseteq [n]$, then $A^C \notin \mathcal{A}$. Hence, $|\mathcal{A}| \leq \frac{1}{2}|\mathcal{P}([n])| = \frac{1}{2}2^n = 2^{n-1}$.
If $|\mathcal{A}| < 2^{n-1}$, then $A \notin \mathcal{A}$ and $A^C \notin \mathcal{A}$ for some set A . If $A \cap B \neq \emptyset$ for all $B \in \mathcal{A}$, then add A to \mathcal{A} .
If $A \cap B = \emptyset$ for some $B \in \mathcal{A}$, then $B \subset A^C$. Then $C \cap A^C \supseteq C \cap B \neq \emptyset$ for all $C \in \mathcal{A}$; add A^C to \mathcal{A} .
Keep adding sets or their complement in this way until \mathcal{A} is a family that contains 2^{n-1} sets, namely one of the sets A and A^C for each $A \subseteq [n]$. \square

Example. A *star* is a family whose elements are the sets containing a fixed $i \in [n]$. Each such family achieves a bound in the theorem above. These are not the only families to achieve the bound, as the following example with $n = 3$ shows:

$$\mathcal{A} = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, [3]\} = \binom{[3]}{2} \cup \{[3]\}.$$

Theorem (Seymour 1973; many others)

If $A \cap B \neq \emptyset$ and $A \cup B \neq [n]$ for all $A, B \in \mathcal{A}$, then $|\mathcal{A}| \leq 2^{n-2}$.

We have seen (and partially proven) the following theorem in Chapter 2.

Sperner's Theorem (1928)

If $A \not\subseteq B$ for all $A, B \in \mathcal{A}$, then $|\mathcal{A}| \leq \binom{n}{\lceil n/2 \rceil}$.

Proof. Let \mathcal{C} be the set of maximal chains $\emptyset \subseteq \{a\} \subseteq \dots \subseteq [n]$ in \mathcal{B}_n .

Note that $|\mathcal{C}| = n!$ and $|\mathcal{A} \cap C| \leq 1$ for each $C \in \mathcal{C}$.

Also, each given set $A \subseteq [n]$ is an element of $|A|!(n - |A|)!$ maximal chains. Then

$$\begin{aligned} 1 = \frac{|\mathcal{C}|}{n!} &= \sum_{C \in \mathcal{C}} \frac{1}{n!} \geq \sum_{C \in \mathcal{C}} \sum_{A \in \mathcal{A} \cap C} \frac{1}{n!} \\ &= \sum_{A \in \mathcal{A}} \sum_{A \in C \in \mathcal{C}} \frac{1}{n!} = \sum_{A \in \mathcal{A}} \frac{|A|!(n - |A|)!}{n!} = \sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} \geq \sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{\lceil n/2 \rceil}} = \frac{|\mathcal{A}|}{\binom{n}{\lceil n/2 \rceil}}. \quad \square \end{aligned}$$

Remark. There is equality in Sperner's Theorem if and only if $\mathcal{A} = \binom{[n]}{\lfloor n/2 \rfloor}$.

The Erdős-Ko-Rado Theorem (1961)

If $|A| = k \leq \frac{n}{2}$ and $A \cap B = \emptyset$ for all $A, B \in \mathcal{A}$, then $|\mathcal{A}| \leq \binom{n-1}{k-1}$.

Proof. Write the numbers $1, \dots, n$ on a circle and define $\mathcal{F} := \{F_1, \dots, F_n\}$, where F_i is the k -sized interval $\{i, \dots, i+k-1\} \bmod n$. Each set F_i overlaps $2(k-1)$ other sets, namely the pair of sets F_{i-k+j}, F_{i+j} . These two sets are disjoint, so at most one of these two can belong to \mathcal{A} . Hence, $|\mathcal{F} \cap \mathcal{A}| \leq 1 + (k-1) = k$.

This is also true for permuted circle numbers, giving a new family \mathcal{F}^π . For $A \in \mathcal{A}$, $F \in \mathcal{F}$, we have $A = \pi(F)$ for $k!(n-k)!$ permutations $\pi \in \Sigma_n$. Hence,

$$\begin{aligned} |\mathcal{A}| &= \sum_{A \in \mathcal{A}} \sum_{F \in \mathcal{F}} \frac{1}{n} = \sum_{A \in \mathcal{A}} \sum_{F \in \mathcal{F}} \sum_{\substack{\pi \in \Sigma_n: \\ A = \pi(F)}} \frac{1}{n \cdot k!(n-k)!} \\ &= \sum_{\pi \in \Sigma_n} \sum_{A \in \mathcal{A}} \sum_{\substack{F \in \mathcal{F}: \\ A = \pi(F)}} \frac{1}{n \cdot k!(n-k)!} \\ &= \sum_{\pi \in \Sigma_n} \frac{|\mathcal{A} \cap \mathcal{F}^\pi|}{n \cdot k!(n-k)!} \leq \sum_{\pi \in \Sigma_n} \frac{k}{n \cdot k!(n-k)!} = \frac{n!k}{n \cdot k!(n-k)!} = \binom{n-1}{k-1}. \quad \square \end{aligned}$$

Theorem (Schönheim 1971, Kleitman & Spencer 1973)

If $|A| \leq k \leq \frac{n}{2}$, $A \not\subset B$ and $A \cap B \neq \emptyset$ for all $A, B \in \mathcal{A}$, then $|\mathcal{A}| < \binom{n-1}{k-1}$.

Proof. If $|A| = k$ for all $A \in \mathcal{A}$, then apply the Erdős-Ko-Rado Theorem.

Otherwise, consider the sets A_1, \dots, A_s of the smallest size $\ell := |A_i| \leq k-1$.

Define $\mathcal{B} := \{\mathbf{B}_1, \dots, \mathbf{B}_s\}$, where $\mathbf{B}_i := \{B \in \binom{[n]}{\ell+1} : A_i \subset B\} = \{A_i \cup \{x\} : x \in [n] - A_i\}$.

Note that $|\mathbf{B}_i| = n - \ell \geq 2k - \ell \geq \ell + 2$ and the $|\{i : A_i \subset B\}| \leq \ell + 1$. For each $J \subseteq [s]$,

$$|\mathcal{B}(J)| = \sum_{B \in \mathcal{B}(J)} \frac{\ell+1}{\ell+1} \geq \sum_{B \in \mathcal{B}(J)} \sum_{\substack{j \in J: \\ A_j \subset B}} \frac{1}{\ell+1} = \sum_{j \in J} \sum_{B \in \mathbf{B}_j} \frac{1}{\ell+1} \geq \sum_{j \in J} \frac{\ell+2}{\ell+1} > |J|.$$

By Hall's Theorem, there are distinct B_i with $A_i \subset B_i \in \mathbf{B}_i$ for all $i \in [s]$. Replace each A_i by B_i and use induction, to find a new family \mathcal{A}' with $|\mathcal{A}'| = |\mathcal{A}|$ and $|A| = k$ for all $A \in \mathcal{A}'$ as before. \square

Fischer's Inequality (1940; Majumdar 1953)

If, for some fixed $k \neq 0$, $|A \cap B| = k$ for all $A, B \in \mathcal{A}$, then $|\mathcal{A}| \leq n$.

Proof. If $|A| = k$ for some $A \in \mathcal{A}$, then $A \subseteq B$ for all $B \in \mathcal{A}$, and all remainders $B - A \subseteq [n] - A$ are disjoint, so $|\mathcal{A}| - 1 \leq n - k$, and $|\mathcal{A}| \leq n + 1 - k \leq n$. Suppose then that $|A| \geq k+1$ for all $A \in \mathcal{A}$.

For each $A \in \mathcal{A}$, define the *incidence vector* \mathbf{v}_A by

$$(\mathbf{v}_A)_i := \begin{cases} 1 & i \in A \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$\mathbf{v}_A \cdot \mathbf{v}_B = |A \cap B| = \begin{cases} |A| \geq k+1 & A = B, \\ k & A \neq B. \end{cases}$$

Suppose that $\sum_A \lambda_A \mathbf{v}_A = 0$. Then

$$\begin{aligned} 0 &= \left(\sum_{A \in \mathcal{A}} \lambda_A \mathbf{v}_A \right) \cdot \left(\sum_{A \in \mathcal{A}} \lambda_A \mathbf{v}_A \right) = \sum_{A \in \mathcal{A}} \lambda_A^2 |A| + \sum_{\substack{A, B \in \mathcal{A}: \\ A \neq B}} \lambda_A \lambda_B k \geq \sum_{A \in \mathcal{A}} \lambda_A^2 (k+1) + \sum_{\substack{A, B \in \mathcal{A}: \\ A \neq B}} \lambda_A \lambda_B k \\ &= k \left(\sum_{A \in \mathcal{A}} \lambda_A \right)^2 + \sum_{A \in \mathcal{A}} \lambda_A^2 \geq \sum_{A \in \mathcal{A}} \lambda_A^2 \geq 0. \end{aligned}$$

Hence, $\lambda_A = 0$ for all $A \in \mathcal{A}$, so vectors $\mathbf{v}_A \in \mathbb{R}^n$ are linearly independent. Hence, $|\mathcal{A}| \leq n$. \square

Theorem (Bollobás 1986)

If, for some fixed $k \neq \frac{n+1}{2}$, $|A \triangle B| = k$ for all $A, B \in \mathcal{A}$, then $|\mathcal{A}| \leq n$.

The proof is similar to that of Fischer's Inequality, but uses a ± 1 incidence vector.

Definition. Let \mathcal{A} be a family of subsets of $[n]$.

A subfamily $\mathcal{F} \subseteq \mathcal{A}$ is a *sunflower* with p petals if $|\mathcal{F}| = p$ and $F \cap G = Y$ and petal $F - Y$ is nonempty, for all $F \in \mathcal{F}$ and a fixed set Y .

The Sunflower Lemma (Erdős & Rado 1960)

If $|A| = k$ for all $A \in \mathcal{A}$ and $|\mathcal{A}| > k!(p-1)^k$, then \mathcal{A} contains a sunflower with p petals.

Proof. The theorem is clearly true is $k = 0, 1$. Assume that the theorem is true for $k-1$. Consider a maximal subfamily $\mathcal{F} \subseteq \mathcal{A}$ of pairwise disjoint sets. If $|\mathcal{F}| \geq p$, then \mathcal{F} is a sunflower with p petals.

Suppose that $|\mathcal{F}| \leq p-1$ and set $B := \bigcup_{A \in \mathcal{F}} A$. By the maximality of \mathcal{F} , B intersects each set $A \in \mathcal{A}$. By the Pigeonhole Principle, some element $x \in B$ is contained in at least

$$\frac{|\mathcal{A}|}{|B|} \geq \frac{|\mathcal{A}|}{k(p-1)} > \frac{k!(p-1)^k}{k(p-1)} = (k-1)!(p-1)^{k-1}$$

of the sets of \mathcal{A} . By assumption, $\mathcal{A}' := \{A - \{x\} : x \in A \in \mathcal{A}\}$ has a sunflower \mathcal{G} with p petals.

Then \mathcal{A} has the sunflower $\mathcal{F} := \{A \cup \{x\} : A \in \mathcal{G}\}$ with p petals. \square

Bondy's Theorem (1972)

If $|\mathcal{A}| \leq n$, then the sets $A - \{x\}$ ($A \in \mathcal{A}$) are distinct for some element x .

Proof. Assume that the theorem is not true; then $|\mathcal{A}| \geq 2$.

Set $\mathcal{A}_D := \{D \cap A : A \in \mathcal{A}\}$ and $\mathcal{D} := \{D \subseteq [n] : |\mathcal{A}_D| \geq |D| + 1\}$.

Choose $d \in A \triangle B$ for some distinct $A, B \in \mathcal{A}$. Then

$$|\mathcal{A}_{\{d\}}| \geq |\{\emptyset, \{d\}\}| = 2 = |\{d\}| + 1$$

so $\{d\} \in \mathcal{D}$ and $\mathcal{D} \neq \emptyset$. Choose a maximal set D in \mathcal{D} . Then by assumption, $|\mathcal{A}_D| \leq |\mathcal{A}| - 1$.

By the Pigeonhole Principle, $D \cap A = D \cap B$ for some distinct $A, B \in \mathcal{A}$.

Choose $e \in A \triangle B$ and define $E := D \cup \{e\}$. Then

$$|\mathcal{A}_E| = |\mathcal{A}_D| + 1 \geq (|D| + 1) + 1 = |E| + 1.$$

Hence $E \in \mathcal{D}$, contradicting the maximality of D . \square

Definition. The *complete k -partite graph* $K_{n_1, \dots, n_k} = (V, E)$ has vertices and edges as follows:

$$V = \bigcup_{i \in [k]} V_i \quad \text{with} \quad |V_i| = n_i \quad \text{and} \quad E = \binom{V}{2} - \bigcup_{i \in [k]} \binom{V_i}{2}.$$

Note that if $n_1 = \dots = n_k$ and $|V| = n$, then

$$|E| = \binom{n}{2} - k \binom{\frac{n}{k}}{2} = \frac{1}{2}n(n-1) - \frac{1}{2}k \frac{n}{k} \left(\frac{n}{k} - 1 \right) = \left(1 - \frac{1}{k} \right) \frac{n^2}{2}.$$

Turán's Theorem (1941)

Each simple graph $G = (V, E)$ with $|V| = n$ and $|E| > \left(1 - \frac{1}{p-1} \right) \frac{n^2}{2}$ contains K^p .

Proof. We prove the contrapositive: if G is K_p -free, then $|E| \leq \left(1 - \frac{1}{p-1} \right) \frac{n^2}{2}$. We use induction on n . If $n < p$, then

$$|E| \leq \binom{n}{2} = \left(1 - \frac{1}{n} \right) \frac{n^2}{2} \geq \left(1 - \frac{1}{p-1} \right) \frac{n^2}{2}.$$

Suppose that G is a simple graph on n vertices with no K_p subgraph. Add edges to G until none can be added without creating a K_p subgraph. Then G must contain a K_{p-1} subgraph, say with vertices A .

Set $B := V - A$. Each vertex of B has at most $p-2$ edges to vertices in A .

By induction, there are at most $\left(1 - \frac{1}{p-1} \right) \frac{(n - (p-1))^2}{2}$ edges between the vertices B .

Also, there are $\binom{p-1}{2}$ edges between the vertices of A . Thus,

$$|E| \leq (n - (p-1))(p-2) + \left(1 - \frac{1}{p-1} \right) \frac{(n - (p-1))^2}{2} + \binom{p-1}{2} = \left(1 - \frac{1}{p-1} \right) \frac{n^2}{2}. \quad \square$$

As a special case ($p = 3$) of Turán's Theorem, we obtain the following theorem, first proved in Chapter 1.

Mantel's Theorem (1907)

If $G = (V, E)$ is a simple graph on n vertices with $|E| > \frac{n^2}{4}$ edges, then G has at least one triangle.

Sperner Applications

In this section, we will see unexpected and beautiful applications of extremal set theory.

There are many such applications but we will merely consider applications of two of Sperner's results, the first of which is Sperner's Theorem that we have seen earlier in the chapter and also in Chapter 2.

Theorem (Erdős 1945)

Let $a_1, \dots, a_n \in \mathbb{R}$ satisfy $|a_i| \geq 1$ for all $i \in [n]$. The number of sums $\sum_{i \in [n]} \epsilon_i a_i$ with $\epsilon_i = \pm 1$ that lie in some open interval of length 2 is at most $\binom{n}{\lceil n/2 \rceil}$.

Proof. By changing the signs of ϵ_i , we may assume that $a_i \geq 1$ for each i . Consider the sums $\sum_{i \in [n]} \epsilon_i a_i$ that lie in a given interval I with $|I| = 2$. Each such sum S determines bijectively a subsets $J = \{i \in [n] : \epsilon_i = 2\}$. Assume that $J' \subset J$ for associated sums $S, S' \in I$. Then

$$S - S' = \sum_{i \in [n]} \epsilon_i a_i - \sum_{i \in [n]} \epsilon'_i a_i = \left(\sum_{i \in J} a_i - \sum_{i \in [n] - J} a_i \right) - \left(\sum_{i \in J'} a_i - \sum_{i \in [n] - J'} a_i \right) = 2 \sum_{i \in J - J'} a_i \geq 2,$$

a contradiction, so $J' \not\subset J$. By Sperner's Theorem, the number of sums $S \in I$ is at most $\binom{n}{\lceil n/2 \rceil}$. \square

Kleitman generalised this to arbitrary dimension and multiple regions.

Kleitman's Theorem (simple, 1970)

Let $a_1, \dots, a_n \in \mathbb{R}^d$ satisfy $|a_i| \geq 1$ for all $i \in [n]$. The number of sums $\sum_{i \in [n]} \epsilon_i a_i$ with $\epsilon_i = \pm 1$ that lie in some open interval ball of radius 1 is at most $\binom{n}{\lceil n/2 \rceil}$.

The next two application rely on the following simple but technical combinatorial lemma.

Sperner's Lemma (1928)

Consider a triangulation of a triangle with vertices V_1, V_2, V_3 .

If the vertices V in the triangulation have a colouring $\chi : V \rightarrow [3]$ so that

- $\chi(V_i) = i$ for each $i \in [3]$, and
- $\chi(v) \in \{i, j\}$ for perimeter vertices $v \in V$ between V_i, V_j ,

then at least on triangle is tri-coloured.

Proof. Draw part of the dual graph of this triangulation, namely by placing a vertex in each (triangular) region and drawing an edge between two vertices if and only if this edge passes between a vertex of colour 1 and a vertex of colour 2. The pendent vertices of this graph lie in tri-coloured triangles, and each such triangle contains a pendent vertex. Each vertex within the triangle has either degree 1 or 2; removing cycles and isolated paths yields a tree, namely a number of paths that are vertex-disjoint in one vertex, namely the vertex outside of the triangle, where these paths meet up. There must be an odd number of such paths, having to cross the 1-2 coloured triangle side. Hence, there is an odd number of tri-coloured triangles. In particular, there must be at least one. \square

Brouwer's Fixed Point Theorem (1912)

Each continuous function $f : B_n \rightarrow B_n$ on a ball $B_n \subset \mathbb{R}^n$ has a fixed point.

Proof. The triangle $\triangle \subset \mathbb{R}^3$ with vertices e_1, e_2, e_3 is homeomorphic to B_2 . Assume that $f : \triangle \rightarrow \triangle$ is a continuous function without a fixed point. Let $\delta(\mathcal{T})$ be the maximal length of an edge in a triangulation \mathcal{T} . Form a sequence of triangulations $\mathcal{T}_1, \mathcal{T}_2, \dots$ for which $\delta(\mathcal{T}_k) \rightarrow 0$. Colour the vertices of each \mathcal{T}_k by $\lambda(v) = \min\{oi : (f(v) - v)_i < 0\}$. This is well-defined: $f(v) - v$ lies in the plane $x + y + z = 0$ and is nonzero by assumption, so at least one coordinate is negative (resp. positive).

It is not hard to check that λ is a Sperner colouring; for instance, $\lambda(e_i) = i$. By Sperner's Lemma, each triangulation \mathcal{T}_k contains a tri-coloured triangle $\{v^{k1}, v^{k2}, v^{k3}\}$ with $\lambda(v^{ki}) = i$.

Since \triangle is compact, $\{v^{k1}\}_k$ has a converging subsequence with limit v . Since $\delta(\mathcal{T}_k) \rightarrow 0$, the corresponding subsequences of $\{v^{k2}\}_k$ and $\{v^{k3}\}_k$ must also converge to v . Note that $f(v^{ki})_i < v^{ki}_i$ for all k . By continuity, $f(v)_i \leq v_i$ for all i , which is a contradiction. \square

Another well-known application of Sperner's Lemma is the following theorem.

Monsky's Theorem (1970)

A square cannot be dissected by an odd number of equal-sized triangles.

References

- [1] M. Aigner, *Combinatorial Theory*, Springer-Verlag, New York, 1979.
- [2] M. Aigner and G.M. Ziegler, *Proofs from The Book*, 4th edition, Springer-Verlag, Berlin, 2010.
- [3] I. Anderson, *Combinatorics of Finite Sets*, Oxford University Press, New York, 1987.
- [4] B. Bollobás, *Combinatorics. Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability*, Cambridge University Press, Cambridge, 1986.
- [5] R.L. Graham, M. Grötschel, and L. Lovász (eds.), *Handbook of Combinatorics. I–II*, North-Holland, Amsterdam, 1995.
- [6] S. Jukna, *Extremal Combinatorics. With Applications in Computer Science*, 2nd ed., Springer, Heidelberg, 2011.
- [7] J.H. van Lint and R.M. Wilson, *A Course in Combinatorics*, Cambridge University Press, 1992.