Chapter 11

Extremal Problems

In this chapter, we study extremal problems for families of sets and subsets. We seek the largest or smallest structure with certain properties, or extreme values of a parameter over a class of structures.

11.1. Forced Subgraphs

In Ramsey’s Theorem, we $k$-color $\binom{[n]}{q}$ for some large $n$ and find that some $q$-set is homogeneous. That is, we find a monochromatic copy of the complete $t$-uniform hypergraph with $q$ vertices, written $K_q^{(t)}$. Given $n$ and $q$ with $n \geq q$, we can also force a copy of $K_q^{(t)}$ by having enough edges in a $t$-uniform hypergraph on $n$ vertices; the maximum number of edges avoiding $K_q^{(t)}$ is Turán’s Problem.

For $t \geq 3$, the problem is notoriously difficult; Füredi [1991] and Keevash [2011] give extensive surveys. We do not even know the asymptotic answer for $K_4^{(3)}$ (see Exercise 30). The “flag algebra” method pioneered by Razborov [2007, 2010] has improved many of the bounds; Falgas-Ravry–Vaughan [2013] summarizes the results. Hence we focus on $t = 2$, where the central question started the field of extremal graph theory.

TURÁN’S THEOREM

The traditional phrasing asks for the maximum number of edges in an $n$-vertex graph having no $(r + 1)$-clique. This was solved for $r = 2$ by Mantel [1907] and for general $r$ by Turán [1941]. We already proved the theorem as Theorem 5.2.11; here we take a more quantitative look and give another short proof.

11.1.1. Definition. The Turán graph $T_{n,r}$ is the complete $r$-partite graph with $n$ vertices having $b$ parts of size $a + 1$ and $r - b$ parts of size $a$, where $a = \lfloor n/r \rfloor$ and $b = n - ra$. Let $t_r(n) = |E(T_{n,r})|$. Turán proved that $T_{n,r}$ is the unique largest $n$-vertex graph not containing $K_{r+1}$. The proof by Erdős in Theorem 5.2.11 uses induction on $r$ and the Degree-Sum Formula, showing that for every graph $G$ not containing $K_{r+1}$, there is an
r-partite graph $H$ with $d_H(v) \geq d_G(v)$ for all $v \in V(G)$. It also uses that $T_{n,r}$ is the unique largest $r$-partite $n$-vertex graph (if part-sizes differ by more than 1, shifting a vertex from a larger part to a smallest part gains edges). Here we note the asymptotic value of the number of edges.

11.1.2. Proposition. $t_r(n) = (1 - \frac{1}{r}) \frac{n^2}{2} - O(n)$.

Proof: With part-sizes $x_1, \ldots, x_r$, there are $\sum_{i<j} x_i x_j$ edges. With $\sum x_i = n$ fixed and all $x_i$ real, the sum is maximized when all $x_i$ equal $n/r$ (averaging any two increases the value). The maximum is thus $\left(\frac{n}{r}\right)\frac{n^2}{2}$, which proves $t_r(n) \leq (1 - \frac{1}{r}) \frac{n^2}{2}$.

For the asymptotic optimality of the bound, note that each part-size exceeds $n/r - 1$. Counting the edges by pairs of parts thus yields

$$t_r(n) > \frac{r(r-1)}{2} \left(\frac{n-r}{r}\right)^2 = \frac{1}{2} \left(1 - \frac{1}{r}\right) \left(n^2 - 2nr + r^2\right) = \left(1 - \frac{1}{r}\right) \frac{n^2}{2} - O(n).$$

Turán’s Theorem has many proofs. Turán’s original proof used induction on $n$ (Exercise 7), different from the Erdős proof. The proof in Exercise 8 uses a continuous optimization problem. Theorem 14.1.15 gives a probabilistic proof. Six proofs appear in Aigner [1995], and the five proofs in Aigner–Ziegler [1999] include the proof we present here.

11.1.3. Theorem. (Turán’s Theorem; Turán [1941]) Among $n$-vertex graphs with no $(r+1)$-clique, the unique largest graph is $T_{n,r}$.

Proof: (Zykov [1949]) Since $T_{n,r}$ is the unique largest $r$-partite graph, it suffices to show that any largest graph $G$ not containing $K_{r+1}$ is a complete multipartite graph, which means it does not have $K_2 + K_1$ as an induced subgraph. If $G$ has such a subgraph, with vertex set $\{u, v, w\}$ and $vw \in E(G)$, then we find a larger graph not containing $K_{r+1}$.

If $d(u) < d(v)$, then we replace $u$ with a new vertex $v'$ having the same neighbors as $v$, as shown below. Since $v'$ and $v$ are not adjacent, we did not create an $(r+1)$-clique. Since $d(u) < d(v) = d(v')$, the new graph has more edges.

By symmetry in $v$ and $w$, we may assume $d(u) \geq \max\{d(v), d(w)\}$. Now we replace both $v$ and $w$ by two new copies of $u$, as shown below. Again we have create no larger clique. We lose only $d(v) + d(w) - 1$ edges, since $vw$ is counted twice, and we gain $2d(u)$ edges, so again the new graph is larger.

In fact, for $2 \leq k \leq r$, the Turán graph $T_{n,r}$ is the unique $n$-vertex graph without $K_{r+1}$ that has the most $k$-cliques (Zykov [1949]). This generalization was rediscovered by Erdős [1962] and by Sauer [1971].

An $n$-vertex graph with more than $\frac{n^2}{2} - 1$ edges contains $K_r$. More generally, how many $r$-cliques are forced by $m$ edges on $n$ vertices?
11.1.4. Definition. Let \( k_r(G) \) denote the number of \( r \)-cliques in \( G \).

Let \( G \) be an \( n \)-vertex graph with \( m \) edges. Moon–Moser [1962a] gave a lower bound for \( k_r(G) \) in terms of \( n \) and \( m \). Goodman [1959] minimized \( k_3(G) + k_3(\overline{G}) \) in terms of \( n \). Lovász [1972d] obtained both results via inclusion-exclusion. One would expect a random \( n \)-vertex graph and its complement to have approximately \( 2 \cdot \frac{1}{3}(\binom{n}{3}) \) triangles; surprisingly, the minimum asymptotically equals the average. The proof uses a counting formula that has other applications.

11.1.5. Lemma. If a graph \( G \) has \( n \) vertices and \( m \) edges, then \( k_3(G) + k_3(\overline{G}) \) equals each formula below, where \( d(v) \) is the degree of \( v \) in \( G \).

\[
\begin{align*}
\text{(a)} \quad & \binom{n}{3} - \frac{1}{2} \sum_{v \in V(G)} d(v)[n - 1 - d(v)] & \text{(Sauvée [1961])} \\
\text{(b)} \quad & \binom{n}{3} - (n - 2)m + \sum_{v \in V(G)} \binom{d(v)}{2} & \text{(Goodman [1959])}
\end{align*}
\]

Proof: (Sauvée) View \( G \) and \( \overline{G} \) as a 2-coloring of \( E(K_n) \). Edges of \( G \) have one color and edges of \( \overline{G} \) the other. We count monochromatic triangles.

Assign each pair of incident edges in \( K_n \) weight 2 if they have the same color, weight –1 if not. A vertex triple inducing a monochromatic triangle contributes 6 to the total weight; other triples contribute 0. Hence the sum of the weights of all pairs of incident edges is 6 times the number of monochromatic triangles.

Over the pairs of edges incident to \( v \), the weights sum to

\[
2\binom{d(v)}{2} + 2\binom{n - 1 - d(v)}{2} - d(v)[n - 1 - d(v)].
\]

Using \( \binom{k}{2} + \binom{m - k}{2} = \binom{m}{2} - k(m - k) \), the contribution from \( v \) becomes \( 2\binom{n - 1}{2} - 3d(v)[n - 1 - d(v)] \). Summing over \( v \) and dividing by 6 yields (a).

To obtain (b) from (a), replace \( n - 1 \) with \( n - 2 + 1 \), apply the Degree-sum Formula, and collect terms.

11.1.6. Theorem. (Goodman [1959]) An \( n \)-vertex graph and its complement together have at least \( n(n - 1)(n - 5)/24 \) triangles, sharp when \( n \equiv 1 \pmod{4} \).

Proof: (Sauvée [1961]) To minimize formula (a) of Lemma 11.1.5, we maximize the subtracted terms. This is achieved by setting each \( d(v) \) to \( (n-1)/2 \). The formula becomes \( \frac{n^3}{2} - \frac{3(n-1)^3}{4} \), which simplifies to the claimed lower bound.

By Lemma 11.1.5, \( k_3(G) + k_3(\overline{G}) \) depends only on the vertex degrees, not the choice of edges. Equality in the bound holds if and only if every vertex has degree \( (n-1)/2 \). This can happen only for \( n \) odd and \( (n-1)/2 \) even, so \( n = 4k + 1 \) and \( G \) is \( 2k \)-regular. When \( n \equiv 1 \pmod{4} \), this is achieved by a regular self-complementary graph, such as the graph obtained by adding one vertex adjacent to the low-degree vertices in the near-regular self-complementary graph \( P_4[K_k, K_k, K_k, K_k] \). For other congruence classes, the bound can be improved slightly (Exercise 26).

We return to the counting of cliques in the graph \( G \) itself. For \( r = 3 \), Lemma 11.1.5 provides a lower bound.

11.1.7. Corollary. (Moon–Moser [1962a]) A graph \( G \) with \( n \) vertices and \( m \) edges has at least \( \frac{n^3}{24}(4m - n^2) \) triangles.
Proof: Write the formula of Lemma 11.1.5b in terms of $G$, letting $m' = |E(G)|$ and $d'(v) = d(G(v))$. Note that $3k_3(G) \leq \sum (d'(v)/2)$, since incident edges lie in at most one triangle. Subtracting this upper bound on $k_3(G)$ from Lemma 11.1.5b yields

$$k_3(G) \geq \left( \frac{n^3}{3} \right) - (n - 2)m' + \frac{2}{3} \sum (d'(v)/2).$$

Replace $\sum (d'(v)/2)$ with the lower bound $n(2m'/n)$, replace $m'$ with $(n^2/4) - m$, and simplify (Exercise 1) to obtain $k_3(G) \geq \frac{2n}{3}(4m - n^2)$.

11.1.8. Remark. Lower bounds. For $q + 1 \geq r > p$, the Turán graph $T_{n,q}$ has the fewest $r$-cliques among $n$-vertex graphs $G$ with $k_p(G) = k_p(T_{n,q})$. Bollobás [1976a] proved that linear interpolation yields lower bounds on $k_r(G)$ for intermediate values of $k_p(G)$ (see Bollobás [1978, pp. 297–301]). When $p = 2$ and $r = 3$ and $m = |E(G)|$, interpolation improves Corollary 11.1.7 for $m \in [\frac{n^2}{2}, \frac{3n^2}{4}]$. Since $K_{n/2,n/2}$ has $n^2/4$ edges and no triangles, while $K_{n^2/3,n^2/3}$ has $n^2/3$ edges and $n^2/27$ triangles, interpolation yields at least $\frac{3}{4}(4m - n^2)$ triangles. This improves Corollary 11.1.7, since $\frac{3}{4} > \frac{2n}{3}$ when $m < n^2/3$.

Just above $t_r(n)$ (with $n^2/4 + 1$ edges), Turán’s Theorem guarantees one triangle, the Moon–Moser bound guarantees $n/3$, and interpolation guarantees $4n/9$. In fact, $G$ has at least $\lceil n/2 \rceil$ triangles (Rademacher; see Erdős [1955], in Hebrew). This is sharp, by adding one edge to $T_{n,2}$.

Adding $q$ edges inside one part of $T_{n,2}$ (without forming triangles) creates only $q\lceil n/2 \rceil$ triangles. Erdős [1962] proved that this minimizes $k_3(G)$ when $q < cn$, for some constant $c$. When $q = n/2$, adding a $(k + 1)$-cycle to $K_{k+1,k+1}$ produces only $(n/2)(n/2) - 1$ triangles. For large $n$, Lovász–Simonovits [1976] proved Erdős’ conjecture that $c = 1/2$ is best, plus similar results for larger complete graphs.

Mubayi [2010] greatly generalized the results. For an $(r + 1)$-chromatic graph $F$ having an edge $e$ such that $\chi(F - e) = r$, there is a constant $c_F$ such that for $1 \leq q \leq c_F n$, every $n$-vertex graph with $t_r(n) + q$ edges has at least $qs$ copies of $F$, where $s$ is the minimum number of copies of $F$ formed by adding one edge to $T_{n,r}$. This is sharp for odd cycles and asymptotically sharp in general. The tool is the Graph Removal Lemma, generalizing the Triangle Removal Lemma (Lemma 11.1.22) obtained from the Szemerédi Regularity Lemma (Theorem 11.1.13).

The problem was also studied for larger numbers of edges: we want to minimize $k_r(G)$ when $G$ is an $n$-vertex graph with at least $\gamma n^2$ edges, where $\gamma > \frac{1}{2}(1 - \frac{1}{r-1})$. Lovász–Simonovits [1983] conjectured that the minimizing graph is a complete $r$-partite graph where all the parts (except one smaller) have the same size. This was proved for $r = 3$ by Razborov [2008], for $r = 4$ by Nikiforov [2011], and in general by Reiher [2016].

**ERDŐS–STONE THEOREM**

Just as the Ramsey problem extends by seeking a monochromatic copy of any target graph, so the Turán problem extends by asking how many edges force a given graph. When $K_q$ is forced, every $q$-vertex graph is forced, but fewer edges may force sparser subgraphs.
11.1.11. Theorem. (Erdoes–Simonovits [1966]) If $F$ is an $(r + 1)$-chromatic graph, then \( \lim_{n \to \infty} \text{ex}(n; F)n^{-2} = \frac{1}{2}(1 - \frac{1}{r})\).

**Proof:** Since \( T_{n,r} \) is $r$-partite, \( F \not\subseteq T_{n,r} \) when \( \chi(F) > r \). Thus \( \text{ex}(n; F) \geq t_r(n) = \frac{1}{2}(1 - \frac{1}{r})n^2 - O(n) \). For the upper bound, let \( s \) be the maximum size of a color class in some proper \( (r + 1) \)-coloring of \( F \). In an \( n \)-vertex graph with \( \frac{1}{2}(1 - \frac{1}{r} + c)n^2 \) edges, where \( n \) is sufficiently large in terms of \( c \), the Erdoes–Stone Theorem guarantees the appearance of \( K_{r+1}[s] \) and thus also \( F \). Thus the ratio \( \text{ex}(n; F)/n^2 \) can be brought down as close to \( \frac{1}{2}(1 - \frac{1}{r}) \) as desired by making \( n \) sufficiently large. 

For bipartite graphs, Theorem 11.1.11 gives only \( \text{ex}(n; F) \in o(n^2) \). Counting arguments yield better bounds; in fact, \( \text{ex}(n; C_4) \in O(n^{3/2}) \). Because the optimal constructions involve designs and projective planes, we postpone discussion of \( \text{ex}(n; C_4) \) to Chapter 13. Simonovits [1968] pioneered a method for studying Turan numbers (also called extremal numbers). These have been studied for various bipartite graphs and for the extension to families of graphs; Furedi–Simonovits [2013] provides an extensive survey.

Of most interest is \( \text{ex}(n; C_{2k}) \). Bondy–Simonovits [1974] proved \( \text{ex}(n; C_{2k}) < 90kn^{1+1/k} \) (see Exercise 14.2.12 for a lower bound). The constant was improved by Verstraete [2000] and then Pikhurko [2012]. More recently, Bukh–Jiang [2017] proved \( \text{ex}(n; C_{2k}) \leq 80\sqrt{k} \log k n^{1+1/k} + 10k^2n \) when \( k \geq 2 \) and \( n \geq (2k)^{8k^2} \).

Consider now the Erdos–Stone Theorem. Exercise 31 requests a direct proof. Here we use a tool that has become enormously important in proving asymptotic results in combinatorics. It is the Szemeredi Regularity Lemma, developed originally to prove the conjecture of Erdos and Turan that sets of integers with positive density contain arbitrarily long arithmetic progressions.

Roughly speaking, the Regularity Lemma states that every sufficiently large graph contains within it a multipartite subgraph with large parts such that almost all the bipartite subgraphs induced by pairs of parts look fairly random. This is useful when seeking a particular subgraph, because in randomly generated graphs with a fixed density of edges (see Chapter 14) the number of occurrences of a particular subgraph is almost always close to its expected value.
The induced subgraphs of fixed size in a “fairly random” bipartite graph have about the same edge density, if they are not too small. Hence we restrict the edge density only for subgraphs with at least a fraction \( \varepsilon \) of the vertices from each part. Making \( \varepsilon \) smaller is more restrictive, yielding more “regular” behavior.

11.1.12. Definition. Given disjoint vertex sets \( A \) and \( B \) in a graph \( G \), the subgraph generated by \((A,B)\) is the \( A,B \)-bipartite graph with edge set \([A,B]\). Its density, denoted \( \rho(A,B) \), is \( \frac{|A| \cdot |B|}{|V(G)|} \). The pair \((A,B)\) is \( \varepsilon \)-regular if \( |\rho(A',B') - \rho(A,B)| < \varepsilon \) whenever \( A' \subseteq A \) and \( B' \subseteq B \) with \( |A'| \geq \varepsilon |A| \) and \( |B'| \geq \varepsilon |B| \). An equipartition of \( G \) is a partition of \( V(G) \) into parts whose sizes differ by at most 1.

When we partition \( V(G) \) into singleton sets, every pair is \( \varepsilon \)-regular, but applications require larger parts. We need equipartitions with not too many parts such that almost all the pairs of parts are \( \varepsilon \)-regular. The Regularity Lemma guarantees such a suitable partition in a precise way.

11.1.13. Theorem. (Regularity Lemma; Szemerédi [1978]). Given \( \varepsilon, l > 0 \), there exist constants \( M, N \in \mathbb{N} \) such that every graph with at least \( N \) vertices has an equipartition with \( k \) parts for some \( k \in [l, M] \) such that fewer than \( \varepsilon k^2 \) pairs of parts fail to be \( \varepsilon \)-regular.

Though Szemerédi [1978] is the seminal paper, already the idea was used in Szemerédi [1975]. Szemerédi [2015] describes a number of variants of the lemma. We will prove the one stated below.

11.1.14. Definition. For an \( n \)-vertex graph \( G \), an \( \varepsilon, k \)-partition is a partition \( V_0, \ldots, V_k \) of \( V(G) \) such that \( |V_0| \leq \varepsilon n \) (\( V_0 \) may be empty) and \( |V_1| = \cdots = |V_k| \); call \( V_0 \) the exceptional part. An \( \varepsilon, k \)-partition is an \( \varepsilon \)-regular partition if fewer than \( \varepsilon k^2 \) pairs of nonexceptional parts fail to be \( \varepsilon \)-regular.

11.1.15. Theorem. (alternate Regularity Lemma). For \( \varepsilon, l > 0 \), there are constants \( M, N \in \mathbb{N} \) such that every graph with at least \( N \) vertices has an \( \varepsilon \)-regular \( \varepsilon, k \)-partition for some \( k \) with \( l \leq k \leq M \).

Before proving the Regularity Lemma, we use it to prove the Erdős–Stone Theorem and striking results in graph Ramsey theory and additive combinatorics. Applications of the Regularity Lemma avoid technical detail by capturing the standard technical argument in an “Embedding Lemma”. We form a reduced graph \( R \) from the partition of \( G \) provided by the Regularity Lemma. The vertices of \( R \) are the (nonexceptional) parts of the equipartition, adjacent when they form an \( \varepsilon \)-regular pair with density more than some parameter \( d \). With \( d > \varepsilon \), the Embedding Lemma implies that any specified subgraph \( H \) of \( R[s] \) that has small enough maximum degree occurs as a subgraph of \( G \), for suitable \( s \).

For the Erdős–Stone Theorem, the Regularity Lemma provides a reduced graph with enough edges so that Turán’s Theorem forces \( K_{r+1} \) in \( R \), and then the Embedding Lemma yields \( K_{r+1}[s] \subseteq G \). That is, we apply the Embedding Lemma with \( R = K_{r+1} \) and \( H = R[s] \).
11.1.16. Theorem. (Embedding Lemma) Let $R$ be a graph, and fix $m, s \in \mathbb{N}$.

Given $d, \varepsilon$ with $d > \varepsilon > 0$, let $G$ be a subgraph of $R[m]$ in which each pair of parts corresponding to an edge of $R$ is an $\varepsilon$-regular pair with density at least $d$. Let $H$ be a subgraph of $R[s]$ with $n$ vertices and maximum degree $D$. If $\varepsilon \leq \varepsilon'$ and $s - 1 \leq \varepsilon'm$, where $\varepsilon' = (d - \varepsilon)D/(D + 2)$, then $H \subseteq G$. In fact, $G$ contains more than $(\varepsilon'm)^n$ copies of $H$ (as labeled subgraphs).

Proof: For each vertex $v \in V(R)$, let $R_v$ be the corresponding $s$-set in $R[s]$, and let $A_v$ be the corresponding $m$-set in $G$. Let $V(H) = x_1, \ldots, x_n$. In $V(G)$ we will find representatives $y_1, \ldots, y_n$ of $x_1, \ldots, x_n$ such that $x_ix_j \in E(H)$ implies $y_iy_j \in E(G)$. Furthermore, if $x_i \in R_v$, then $y_i \in A_v$.

We pick $y_1, \ldots, y_n$ in order. At time $t$, we will choose $y_t$ and update (for $j > t$) the set $B_j$ of candidates for $y_j$. Initially $B_j = A_v$, where $x_j \in R_v$. The set $B_j$ shrinks when we choose a vertex $y_t$ that $y_j$ needs as a neighbor. For each $j$ with $j \geq t$, let $Y_j = \{y_i : t < i < j \text{ and } x_iy_j \in E(H)\}$; these already-chosen vertices are required to be neighbors of the vertex we choose as $y_j$. Since we want $y_j \in A_v$, the set $B_j$ of candidates for $y_j$ is $\{y \in A_v : Y_j \subseteq N_G(y)\}$. Letting $\alpha = d - \varepsilon$, we will make our choices to guarantee that when we are ready to pick $y_t$, 

$$|B_j| \geq \alpha^{|Y_j|m} \quad \text{for } j \geq t. \quad (\ast)$$

When we are choosing $y_t$, we do not know which vertices will be chosen to represent the neighbors of $x_j$ in $\{x_{t+1}, \ldots, x_{j-1}\}$. The vertex eventually chosen as $y_j$ must be adjacent to all of them. For this reason, we initially preserve many candidates for $y_j$. As more neighbors of $x_j$ are chosen, the number of candidates we keep decreases. When $|Y_j| = D$, one candidate suffices.

If $x_j \notin N_H(x_i)$, then $Y_j$ and $B_j$ do not change when $y_t$ is chosen. To maintain $(\ast)$, it suffices to choose $y_t$ from $B_t$ so that $y_t$ has at least $\alpha |B_t|$ neighbors in $B_j$ for each $j$ with $j > t$ and $x_j \in N_H(x_i)$.

For such $j$, define $u$ and $v$ by $x_t \in R_u$ and $x_j \in R_v$. The pair $(A_u, A_v)$ is $\varepsilon$-regular in $G$, by definition. We have $B_t \subseteq A_u$ and $B_j \subseteq A_v$. Since the hypotheses guarantee $\alpha^D \geq \varepsilon$, by $(\ast)$ we have $|B_j| \geq \varepsilon |A_v|$.

Let $B$ be the set of vertices in $B_t$ that do not have at least $\alpha |B_t|$ neighbors in $B_j$. Thus $\rho(B, B_j) < \alpha$. If $|B| \geq cm$, then $\varepsilon$-regularity of $(A_u, A_v)$ guarantees $\rho(B, B_j) \geq d - \varepsilon = \alpha$, a contradiction.

Hence all but $cm$ vertices of $B_t$ have at least $\alpha |B_t|$ neighbors in $B_j$. Since there are at most $D$ values of $y_t$ to consider, we discard at most $Dcm$ vertices of $B_t$ in this way. We also discard vertices of $B_j$ already chosen to represent vertices of $H$ in $R_v$; there are at most $s - 1$ of these. There remain at least $(\alpha^D - Dcm) - (s - 1)$ vertices in $B_t$; choose one to be $y_t$.

Since at each $t$ there are at least $(\alpha^D - D\varepsilon)m - (s - 1)$ choices for $y_t$, there are at least $[(\alpha^D - D\varepsilon)m - (s - 1)]^n$ labeled copies of $H$ in $G$. From $\varepsilon' = \alpha^D/(D + 2)$, we have $\alpha^D = 2\varepsilon' + D\varepsilon' \geq 2\varepsilon' + D\varepsilon$. Thus $2\varepsilon' \leq \alpha^D - D\varepsilon$. Since also $s - 1 \leq \varepsilon'm$, we have $(\alpha^D - D\varepsilon)m - (s - 1) \geq \varepsilon'm$.

The Embedding Lemma is used to find small graphs in a large graph $G$. (Komlós–Sárközy–Szemerédi [1997] proved a more difficult extension called the Blow-up Lemma to find spanning subgraphs.) To complete the proof of the Erdős–Stone Theorem, it now suffices to find a copy of $K_{r+1}$ in the reduced graph that comes from the partition of $G$ provided by the Regularity Lemma.
Proof of Erdős–Stone Theorem (Theorem 11.1.10): Fix $r, s \in \mathbb{N}$ and a positive constant $c$. We want to prove that if $n$ is sufficiently large, then an $n$-vertex graph with more than $t_r(n) + cn^2$ edges contains $K_{r+1}[s]$. We may assume $c < \frac{1}{2}$. 

Choose $l$ with $l > \max\{r, \frac{1}{2r}\}$. Choose $d$ with $0 < \frac{d}{2} < c - \frac{1}{2l} < \frac{1}{2}$. Choose $\varepsilon$ small enough so that $\varepsilon \leq \frac{(d-\varepsilon)^2}{r^2}$ and $\varepsilon < \frac{1}{4}(c - \frac{d}{2} - \frac{1}{l})$. Let $M$ and $N$ be the constants in terms of $\varepsilon$ and $l$ needed to apply the Regularity Lemma. Choose $n$ so that $n > \max\{N, M(s-1)/\varepsilon'\}$, where $\varepsilon' = \frac{(d-\varepsilon)^2}{r^2}$.

Let $G$ be an $n$-vertex graph with more than $t_r(n) + cn^2$ edges. By the alternative form of the Regularity Lemma, $G$ has an $\varepsilon$-regular $\varepsilon$, $k$-partition $V_0, \ldots, V_k$, where $l \leq k \leq M$. Let $m = |V_1| = \cdots = |V_k|$.

The reduced graph $R$ with vertex set $v_1, \ldots, v_n$ has $v_iv_j \in E(R)$ if and only if the pair $(V_i, V_j)$ is $\varepsilon$-regular with density at least $d$. Showing that $|E(R)| > t_r(k)$ forces $K_{r+1} \subseteq R$. The Embedding Lemma then implies $K_{r+1}[s] \subseteq G$. Since $\Delta(K_{r+1}[s]) = rs$, applying it requires $\varepsilon \leq \varepsilon'$ and $s - 1 \leq \varepsilon'm$, which were guaranteed by the choice of $\varepsilon$ and then $n$.

The “bad” edges of $G$ are those incident to $V_0$, within $V_i$ for $i \geq 1$, or joining $V_i$ and $V_j$ when $(V_i, V_j)$ is not $\varepsilon$-regular or is $\varepsilon$-regular with density less than $d$. To show that $R$ has enough edges, we show that the number of bad edges in $G$ is small enough that more than $t_r(k)$ pairs with density at least $d$ are needed to accommodate the good edges. Note that $n - \varepsilon n \leq km$ yields $n \leq (1 + \varepsilon)km$, and hence $n^2 < 2\varepsilon k^2 m^2$, since $\varepsilon < 1/4$. Of course, $n^2 > k^2 m^2$.

Since $|V_0| \leq \varepsilon n$, the number of edges incident to $V_0$ is less than $\varepsilon n^2$, which is less than $2\varepsilon k^2 m^2$. Since at most $\varepsilon k^2 m^2$ edges of $G$ lie in such pairs. Fewer than $\frac{1}{2} k^2 m^2$ pairs have density less than $d$, and hence at most $\frac{1}{2} k^2 dm^2$ edges lie in such pairs. Each $G[V_i]$ has at most $\frac{1}{2} m^2$ edges, totaling at most $\frac{1}{2} m^2 k$ edges. Finally, each edge of $R$ arises from a pair with density at least $d$; even when the density is 1, it contributes at most $m^2$ edges to $G$. Thus

$$|E(G)| \leq 3\varepsilon k^2 m^2 + \frac{1}{2} d k^2 m^2 + \frac{1}{2} km^2 + |E(R)| m^2.$$ 

Solving for $|E(R)|$, further substitution yields the desired bound.

$$|E(R)| \geq \frac{|E(G)| - 3\varepsilon k^2 m^2 - \frac{1}{2} d k^2 m^2 - \frac{1}{2} km^2}{m^2}$$

$$\geq k^2 \left( \frac{t_r(n) + cn^2}{k^2 m^2} - 3\varepsilon - \frac{d}{2} - \frac{1}{2k} \right)$$

$$\geq k^2 \left( \frac{t_r(n)}{n^2} + c - 3\varepsilon - \frac{d}{2} - \frac{1}{2l} \right)$$

$$> \frac{1}{2} k^2 (1 - 1/r) \geq t_r(k)$$

Hence $K_{r+1} \subseteq R$, and the Embedding Lemma gives $K_{r+1}[s] \subseteq G$.

11.17.* Remark. As we have noted, the Erdős–Stone Theorem has proofs avoiding these tools. The theorem also can be made more precise. We fixed $c$ and $s$ and showed that $K_{r+1}[s]$ is forced by $t_r(n) + cn^2$ edges when $n$ is large enough. Instead, we can fix $c$ and $n$ and ask how large $s$ can be guaranteed as a function of $n$ when forcing $K_{r+1}[s]$. If it grows with $n$, then the Erdős–Stone Theorem follows. A lower bound on the growth of $s$ is a quantitative strengthening.
Bohobás–Erdős [1973] proved that the guaranteed $s$ depends logarithmically on $n$ (with $r$ and $c$ fixed). Several papers improving the leading coefficient culminated in the result of Ishigami [2002] that every $n$-vertex graph with more than $(1 - 1/r + c)n^2/2$ edges contains $K_{r+1}[s]$ with $s \geq \left\lfloor \log_{r+1} n \right\rfloor$ (using Regularity). Bohobás–Kohayakawa [1994] showed that one can also guarantee that one part has close to linear size. Nikiforov [2008] gave a short proof of a more general theorem yielding both logarithmic growth of $s$ and near-linear size of one part.

\section*{LINEAR RAMSEY FOR BOUNDED DEGREE}

Our second application of the Regularity Lemma is from graph Ramsey theory. The Ramsey number $R(G, G)$ may grow exponentially in the number of vertices of $G$, such as when $G = K_n$. In some families of relatively sparse graphs, the Ramsey number grows at most linearly in the number of vertices! Most famously, one such family is the family of graphs with bounded maximum degree.

\begin{theorem}\textup{(Chvátal–Rödl–Szemerédi–Trotter [1983])} \label{thm:linearRamsey}
For $d \in \mathbb{N}$, there is a constant $c_d$ such that $R(G, G) \leq c_d |V(G)|$ whenever $\Delta(G) = d$.
\end{theorem}

\begin{proof}
Let $l = 3 \max \{3^d, R(d+1, d+1)\}$, and let $\varepsilon = l^{-1}$. The Regularity Lemma provides constants $M, N \in \mathbb{N}$ such that every graph with at least $N$ vertices has an $\varepsilon$-regular $\varepsilon$, $k$-partition for some $k$ with $l \leq k \leq M$. Let $\varepsilon' = \min \{N, M/\varepsilon\}'$, where $\varepsilon' = (1/2 - \varepsilon)^{d}/(d + 2)$. Now fix a graph $G$ with maximum degree $d$ and vertices $x_1, \ldots, x_n$; we prove that $R(G, G) \leq cn$.

Consider a red/blue coloring of $E(K_n)$. Let $H$ and $\overline{H}$ be the subgraphs in red and blue, respectively. Via Steps 1 and 2 below, we find $d + 1$ large sets such that every pair of them is $\varepsilon$-regular and has high density in the same color. In Step 3, the Embedding Lemma allows us to find a copy of $G$ among the edges of that color.

**Step 1:** Some $\varepsilon$-regular partition of $V(H)$ has at least $1/3$ parts with all pairs $\varepsilon$-regular in $H$ and $\overline{H}$. For a pair $(A, B)$ of vertex subsets, the densities in $H$ and $\overline{H}$ sum to $1$. Hence $(A, B)$ is $\varepsilon$-regular in $H$ if and only if it is $\varepsilon$-regular in $\overline{H}$, so we consider only $H$. By the choice of $c$, there is an $\varepsilon$-regular $\varepsilon$, $k$-partition $A_0, \ldots, A_k$ of $H$, where $l \leq k \leq M$. All but $\varepsilon k^2$ pairs are $\varepsilon$-regular. Let $H^*$ be the graph with vertex set $[k]$ putting $i$ and $j$ adjacent if and only if $(A_i, A_j)$ is $\varepsilon$-regular in $H$.

By Turán’s Theorem, a $k$-vertex graph with no $(t+1)$-clique has fewer than $(1 - \frac{1}{k})k^2$ edges. Since $|E(H^*)| \geq \left\lfloor \frac{k}{2} \right\rfloor - \varepsilon k^2$, we find a clique of size at least $\frac{1}{4} k$ equal to $\frac{1}{k}$ parts in $\{A_1, \ldots, A_k\}$ among which every pair is $\varepsilon$-regular.

**Step 2:** This $\varepsilon$-regular partition has $d + 1$ parts with all pairs $\varepsilon$-regular of density at least $1/2$ in the same color ($H$ or $\overline{H}$). We color the edges of the $\frac{1}{2}$ clique in $H^*$ found in Step 1, using red if $\rho_H(A_i, A_j) \geq \frac{1}{2}$ and blue if $\rho_H(A_i, A_j) < \frac{1}{2}$. Since $\frac{1}{2} \geq R(d+1, d+1)$, Ramsey’s Theorem yields a monochromatic copy of $K_{d+1}$ in $H^*$; by symmetry, we may assume it is in red. In the original coloring of $E(K_n)$, each pair among these parts is $\varepsilon$-regular with density at least $\frac{1}{4}$ in the red graph.

**Step 3:** The given 2-coloring of $E(K_n)$ has a monochromatic copy of $G$. We now have $d + 1$ vertex sets in $K_n$, all of size at least $n(1 - \varepsilon)/k$, such that each pair is $\varepsilon$-regular with density at least $1/2$ in the red graph, $H$. Let $G'$ be the subgraph of $H$ induced by these parts. In the language of the Embedding Lemma
(Theorem 11.1.16), the reduced graph $R$ of $G'$ is $K_{d+1}$. Since $\chi(G) \leq \Delta(G) + 1 \leq d + 1$, the graph $G$ is a subgraph of $R[n]$. If the Embedding Lemma applies, then $G'$ contains $G$, as desired. It is important that the requirements to apply the Embedding Lemma are only in terms of $\Delta(G)$, not the number of vertices. Since the size of the parts is at most $cn/k$, we need $\varepsilon \leq \varepsilon'$ and $n - 1 \leq \varepsilon'cn/k$, where $\varepsilon' = (\frac{1}{2} - \varepsilon^d)/(d + 2)$. Since we define $\varepsilon$ to be less than $(1/3)^d$ and defined $c$ so that $\varepsilon'c/k \geq \varepsilon'c/M \geq 1$, the hypotheses hold and the proof is complete.

11.1.19.* Remark. Unfortunately, the threshold for the number of vertices at which the Regularity Lemma applies is huge. The consequence is that the constant $c_d$ provided by the proof of Theorem 11.1.18 grows like an exponential tower with height $d$. By using a variant of the Regularity Lemma, Eaton [1998] reduced the constant to doubly-exponential growth ($2^{2^{O(d)}}$).

Graham–Rödl–Ruciński [2000] avoided the Regularity Lemma and reduced the constant to $2^{O(d(d \log d)^2)}$. For bipartite graphs, they further improved the bound to $8(8d)^d$ and showed that $c_d$ is at least exponential in $d$. Like Ramsey’s Theorem, the Regularity Lemma is a powerful tool to prove the existence of bounds, but accurate bounds generally require more detailed direct arguments.

Theorem 11.1.18 was conjectured by Burr–Erdős [1975], who also conjectured the same claim (with a larger constant) for the larger family of $d$-degenerate graphs. Kostochka–Rödl [2001, 2004] proved bounds that are quadratic in the number of vertices. Finally, C. Lee [2017] proved the conjecture, showing the existence of a constant $c$ such that every $d$-degenerate graph $H$ with $\chi(H) = k$ and $|V(H)| \geq 2^{d^2d^4}$ has Ramsey number at most $2^{d^{2d^4}}|V(H)|$.

ROTH’S THEOREM

We give one more application of the Regularity Lemma. Szemerédi developed the Regularity Lemma to prove the long-standing conjecture of Erdős–Turán [1936] that subsets of the integers with positive density contain long arithmetic progressions. More precisely, subsets of $[n]$ containing no $k$-term arithmetic progression have size at most $o(n)$ (strengthening van der Waerden’s Theorem).

Roth [1953] proved this for $k = 3$; Szemerédi [1969] proved it for $k = 4$. Szemerédi’s Theorem (Szemerédi [1975]) used a bipartite version of the Regularity Lemma to prove it for all $k$. A later proof by Gowers [2001] using Fourier analysis avoids regularity and gives much better numerical bounds. We use regularity to prove the case $k = 3$, following Palmer [2015]; see also Szemerédi [2015]. Roth proved that subsets of $[n]$ with size at most $O(n/\log \log n)$ contain a 3-term progression; Behrend [1946] constructed a set of size $n/e^{c\sqrt{n}}$ that does not.

We begin with a standard computation, useful also in some of the exercises.

11.1.20. Lemma. (Degree-Density Lemma) In a graph $G$, let $(A, B)$ be an $\varepsilon$-regular pair with density $d$. Given a set $Y \subseteq B$ with $|Y| \geq \varepsilon |B|$, the number of vertices in $A$ having at most $(d - \varepsilon)|Y|$ neighbors in $Y$ is less than $\varepsilon |A|$.

Proof: Let $X$ be the set of vertices in $A$ having at most $(d - \varepsilon)|Y|$ neighbors in $Y$. By direct counting, $\rho(X, Y) \leq d - \varepsilon$. If $|X| \geq \varepsilon |A|$, then $\varepsilon$-regularity of $(A, B)$ requires $\rho(X, Y) > d - \varepsilon$. Hence $|X| < \varepsilon |A|$.
11.1.21. Lemma. (Triangle-Counting Lemma) Let $A$, $B$, $C$ be a partition of the vertices of a graph $G$ such that each pair is $\varepsilon$-regular. Let $\alpha$, $\beta$, $\gamma$ be the densities of the pairs $(A, B)$, $(B, C)$, $(C, A)$, respectively. If $\alpha$, $\beta$, $\gamma \geq 2\varepsilon$, then the number of triangles having a vertex in each class is at least

$$(1 - 2\varepsilon)(\alpha - \varepsilon)(\beta - \varepsilon)(\gamma - \varepsilon)|A||B||C|.$$ 

**Proof:** By Lemma 11.1.20, fewer than $\varepsilon|A|$ vertices in $A$ have at most $(\alpha - \varepsilon)|B|$ neighbors in $B$, and fewer than $\varepsilon|A|$ vertices in $A$ have at most $(\gamma - \varepsilon)|C|$ neighbors in $C$. Deleting both such sets from $A$ leaves at least $(1 - 2\varepsilon)|A|$ vertices in $A$ having more than $(\alpha - \varepsilon)|B|$ neighbors in $B$ and $(\gamma - \varepsilon)|C|$ neighbors in $C$.

For such a vertex $x$, let $B' = N(x) \cap B$ and $C' = N(x) \cap C$. Since $\beta \geq 2\varepsilon$, by $\varepsilon$-regularity we have $\rho(B', C') \geq \beta - \varepsilon$, so $|B', C'| \geq (\beta - \varepsilon)|B'||C'|$. With at least $(1 - 2\varepsilon)|A|$ choices for $x$ and at least $(\beta - \varepsilon)(\alpha - \varepsilon)(\gamma - \varepsilon)|B||C|$ desired triangles for each such $x$, the claim follows.

The Triangle-Counting Lemma is a special case of the Embedding Lemma, finding many copies of a fixed subgraph given a partition with all pairs $\varepsilon$-regular. We next show that a graph with few triangles can be made triangle-free by removing few edges. A more recent proof and extension by Fox [2011] gives better bounds and avoids the Regularity Lemma (see also Conlon–Fox [2013]). The more general Graph Removal Lemma states that when $F$ is an $k$-vertex graph, all copies of $F$ in any $n$-vertex graph $H$ containing $o(n^k)$ copies of $F$ can be destroyed by deleting $o(n^2)$ edges of $H$. Rödl–Schacht [2009] generalizes to hypergraphs.

11.1.22. Lemma. (Triangle-Removal Lemma; Ruzsa–Szemerédi [1978]) For $\varepsilon > 0$, there exists a positive number $\delta$ such that for sufficiently large $n$, any $n$-vertex graph $G$ in which at least $\varepsilon n^2$ edges must be deleted to break all triangles has at least $\delta n^3$ triangles.

**Proof:** Since a complete graph has only $\binom{n}{2}$ edges, we may assume $\varepsilon < 1/2$. Choose $\varepsilon'$ and $l$ with $\varepsilon' < \varepsilon/3$ and $l > 1/\varepsilon'$. By the Regularity Lemma, there exist integers $M$ and $N$ such that for $n > N$ we are guaranteed an equipartition $\{V_1, \ldots, V_k\}$ of $V(G)$ with $l \leq k \leq M$ such that all but $\varepsilon'k^2$ pairs of parts are $\varepsilon'$-regular.

With this respect to this partition, we first remove a set of edges that hits all triangles except a certain type. We remove all edges that (1) lie inside parts, (2) join parts that do not form an $\varepsilon'$-regular pair, or (3) join parts forming an $\varepsilon'$-regular pair with density less than $2\varepsilon'$. The number of edges of type (1) is at most $k\binom{n/k}{2}$, which is less than $\varepsilon'n^2$ since $k > 1/\varepsilon'$. There are fewer than $\varepsilon'k^2(n/k)^2$ edges of type (2) and fewer than $\binom{k}{2}2\varepsilon'|n/k|^2$ of type (3), both less than $\varepsilon'n^2$. In total, we have removed fewer than $\varepsilon' n^2$ edges.

By the hypothesis on $G$, at least one triangle $T$ remains. By the types of edges removed, $T$ has its vertices in three distinct parts such that each pair among these three is $\varepsilon'$-regular with density at least $2\varepsilon'$. Also, each part has size at least $\lceil n/k \rceil$, which is at least $n/M$. By the Triangle Counting Lemma, among these three parts the number of triangles remaining is at least $(1 - 2\varepsilon')(\varepsilon')^3(n/M)^3$. Thus the desired conclusion holds with $\delta = (1 - 2\varepsilon')(\varepsilon'/M)^3$.

11.1.23. Theorem. (Roth’s Theorem; Roth [1953]) A subset of $[n]$ containing no 3-term arithmetic progression has size at most $o(n)$. 


Proof: (Ruzsa–Szemerédi [1978]) The meaning of the statement is that for each positive \( \varepsilon \), for large enough \( n \) every set \( S \subseteq [n] \) with \( |S| = \varepsilon n \) contains a 3-term arithmetic progression; this is what we must prove.

We construct a graph \( G \) with vertex set \( A \cup B \cup C \), where \( A = \{a_1, \ldots, a_n\} \), \( B = \{b_1, \ldots, b_{2n}\} \), and \( C = \{c_1, \ldots, c_{3n}\} \). For \( s \in S \), \( j \in n \), add to \( E(G) \) the triangle with vertices \( \{a_j, b_{j+s}, c_{j+2s}\} \). Let \( T \) be this set of triangles.

The \( \varepsilon n^2 \) triangles in \( T \) are edge-disjoint; hence it takes at least \( \varepsilon n^2 \) edges to hit all the triangles in \( G \). By the Triangle-Removal Lemma, \( G \) has at least \( \delta n^3 \) triangles (for some constant \( \delta \)). For large enough \( n \), we have \( \delta n^3 > \varepsilon n^2 \), and hence \( G \) has a triangle \( T \) not in \( T \); it uses edges from distinct triangles in \( T \).

We gave \( G \) no edges within \( A \), \( B \), or \( C \). Let \( V(T) = \{a_j, b_{j+s}, c_{j+2s}\} \). By the construction of \( G \), these three edges require \( s \in S \) and \( t - s \in S \), and also \( t = 2s' \) for some \( s' \in S \). Also, \( s \) and \( s' \) are distinct, since \( T \notin T \). Since \( 2s' - s = t - s \in S \), the numbers \( s, s', 2s' - s \) form an arithmetic progression in \( S \).

PROOF OF THE REGULARITY LEMMA

It remains to prove the Regularity Lemma. For this we will need several lemmas about the behavior of densities of pairs of vertex sets.

11.1.24. Definition. For an \( A, B \)-bigraph \( G \) with density \( d \), let \( f(G) = |A||B|d^2 \).

For a partition \( \Pi \) of \( V(G) \) into \( V_1, \ldots, V_k \), let \( G_{i,j} \) be the subgraph of \( G \) generated by \( (V_i, V_j) \), and let \( f(G, \Pi) = \sum_{i,j} f(G_{i,j}) \).

We show first that if \( \Pi' \) refines \( \Pi \), then \( f(G, \Pi) \leq f(G, \Pi') \leq |E(G)| \). Note that \( f(G, \Pi') = |E(G)| \) when the partition consists of singleton sets. The idea is to iteratively refine pairs that are not \( \varepsilon \)-regular. Each refinement increases \( f \), but the process cannot continue forever, since \( f \) is bounded by \( |E(G)| \).

11.1.25. Lemma. Let \( G \) be an \( A, B \)-bigraph with density \( d \), and let \( \{A_1, A_2\} \) partition \( A \), with \( a_i = |A_i| \). If the subgraph \( G_i \) generated by \( (A_i, B) \) has density \( d_i \), then \( f(G_1) + f(G_2) = f(G) + (a_1/a_2)|A||B|(d - d_1)^2 \).

Proof: Let \( a = |A| \) and \( b = |B| \), and let \( m, m_1, m_2 \) be the numbers of edges in \( G, G_1, G_2 \), respectively. Note that \( f(G) = abd^2 = m^2/(ab) = md \). Letting \( x = m_1/m \) and \( \alpha = a_1/a_2 \), we compute

\[
\frac{ba_1a_2}{m^2a}(f(G_1) + f(G_2) - f(G)) = \frac{ba_1a_2}{a_1b} \left( \frac{m_1^2}{a_1b} + \frac{m_2^2}{a_2b} - \frac{m^2}{ab} \right) = \left( \frac{m_1}{m} \right)^2 \frac{a_2}{a} + \left( \frac{m_2}{m} \right)^2 \frac{a_1}{a} - \alpha^2 \frac{a_1a_2}{a^2} = x^2(1 - \alpha) + \left( 1 - x \right)^2 \alpha - \alpha(1 - \alpha) = x^2 - 2x\alpha + \alpha^2 = (x - \alpha)^2 \]

Therefore,

\[
f(G_1) + f(G_2) - f(G) = \frac{m_2^2}{ba_1a_2}(x - \alpha)^2 = \frac{a_1}{a_2} \frac{m_1}{ab} (\frac{a_1}{a} - \frac{m_1}{m})^2 = \frac{a_1}{a_2} \frac{m_1}{ab} (d - d_1)^2 \]

\[ \blacksquare \]
11.1.26. Lemma. Let $G$ be an $A, B$-bigraph with density $d$. Let $\{A_1, A_2\}$ and $\{B_1, B_2\}$ be partitions of $A$ and $B$, with $|A_1| \geq \varepsilon|A|$ and $|B_1| \geq \varepsilon|B|$. Let $f_{i,j} = f(G[A_i \cup B_j])$. If $|\rho(A_1, B_1) - d| \geq \varepsilon$, then $\sum f_{i,j} \geq f(G) + \varepsilon^4 |A_1||B_1|$.

**Proof:** Let $a, b, a_i, b_j$ be the sizes of $A, B, A_i, B_j$, respectively. Also let $d^* = \rho(A_1, B)$ and $d_{i,j} = \rho(A_i, B_j)$. By Lemma 11.1.25,

$$f(G) = f(G[A_1 \cup B]) + f(G[A_2 \cup B]) - \frac{a_1}{a_2}ab(d - d^*)^2 \leq f_{1,1} + f_{1,2} - \frac{b_1}{b_2}a_1b(d - d_{1,1})^2 + f_{2,1} + f_{2,2} - \frac{a_1}{a_2}ab(d - d^*)^2.$$ 

Here we have ignored the gain due to the second subplit.

It now suffices to show $\frac{a_1}{a_2}a(b - d_{1,1})^2 + \frac{b_1}{b_2}a_1b(d - d_{1,1})^2 \geq \varepsilon^4$. Since $b_1 \geq \varepsilon b$ and $a_1 \geq \varepsilon a$, we only need $\frac{a_1}{a_2}\varepsilon(d - d_{1,1})^2 + \frac{b_1}{b_2}(d - d_{1,1})^2 \geq \varepsilon^4$. For fixed $d$ and $d_{1,1}$, we minimize the left side by setting $d^* = \frac{d + d_{1,1}}{1 + \varepsilon}$, since then

$$\frac{a_1}{a_2}\varepsilon(d^* - d_{1,1})^2 + (d - d^*)^2 = 2\varepsilon(d^* - d_{1,1}) - 2(d - d^*) = 0.$$ 

With $d^*$ so chosen, $d^* - d_{1,1} = \frac{d - d_{1,1}}{1 + \varepsilon}$ and $d - d^* = \frac{d_{1,1}}{1 + \varepsilon}$. Using $|d_{1,1} - d| \geq \varepsilon$, we have

$$\frac{\varepsilon}{1 - \varepsilon}\varepsilon(d^* - d_{1,1})^2 + \frac{1}{1 - \varepsilon}(d - d^*)^2 \geq \frac{\varepsilon}{1 - \varepsilon} \left[ \varepsilon \left( \frac{\varepsilon}{1 + \varepsilon} \right)^2 + \left( \frac{\varepsilon}{1 + \varepsilon} \right)^2 \right] = \frac{\varepsilon^4}{(1 + \varepsilon)^2(1 - \varepsilon)} \geq \varepsilon^4.$$ 

11.1.27. Lemma. Let $G$ be an $n$-vertex graph, and let $\Pi$ be an $\varepsilon, k$-partition of $G$, where $0 < \varepsilon < \frac{1}{n}$. If $\Pi$ is not $\varepsilon$-regular and $|V_0| < (\varepsilon - \varepsilon^2)n$, then there is an $\varepsilon, k'$-partition $\Pi'$ of $G$ with $k \leq k' \leq k2^{k+2}/\varepsilon^2$ such that $|V'_0| \leq |V_0| + n/2^k$ and $f(G, \Pi') \geq f(G, \Pi) + \varepsilon^5 n^2/2$.

**Proof:** If $\Pi$ is not $\varepsilon$-regular, then too many pairs of parts are not $\varepsilon$-regular. Let $m$ be their common size. Each bad pair $(V_i, V_j)$ yields sets $X \subseteq V_i$ and $Y \subseteq V_j$ with $|X|, |Y| \geq m\varepsilon$ such that $|\rho(X, Y) - \rho(V_i, V_j)| \geq \varepsilon$. By Lemma 11.1.26, replacing $V_i$ with $\{X, V_i - X\}$ and $V_j$ with $\{Y, V_j - Y\}$ increases $f(G, \Pi)$ by at least $\varepsilon^4 m^2$.

The idea is to capture this gain simultaneously for all the pairs that are not $\varepsilon$-regular. For each pair involving $V_i$ that is not $\varepsilon$-regular, we obtain a partition $\{X, V_i - X\}$ of $V_i$. With fewer than $k$ such pairs, the least common refinement of these partitions has fewer than $2^k$ parts. Apply this to each $V_i$, obtaining a partition $P$ consisting of $V_0$ and at most $k2^k$ other parts, with unequal sizes.

To compare $f(G, P)$ with $f(G, \Pi)$, we group the pairs of parts according to the pairs in $\Pi$ generating them. When $(V_i, V_j)$ is not $\varepsilon$-regular, witnessed by $(X, Y)$, the contributions from subsets of $V_i$ and $V_j$ can be obtained by first splitting $V_i$ and $V_j$ according to $X$ and $Y$ and then continuing the refinement. The gain of at least $\varepsilon^4 m^2$ is followed by further gains we ignore. When computing the contribution for subsets of $V_i$ with subsets of $V_j$, it does not matter that the initial split is different, since we are obtaining a lower bound on the contributions from a set of subpairs disjoint from those arising from $\{V_i, V_j\}$. Since more than $\varepsilon k^2$ pairs of parts in $\Pi$ failed to be $\varepsilon$-regular, we have $f(G, P) \geq f(G, \Pi) + \varepsilon^4 k^2 m^2$. 

485
We still need a partition whose nonexceptional parts have equal size. Let \( m' \) be the desired part-size, to be specified later. Break each part in \( P \) other than \( V_0 \) into blocks of size \( m' \), with fewer than \( m' \) left over. Combine the leftover vertices with \( V_0 \) to form \( V'_0 \); this produces \( \Pi' \) with \( k' \) parts other than \( V'_0 \), where \( k' \leq n/m' \). Note that \( |V'_0| \leq |V_0| + m'k2^{k} \).

Splitting into blocks of size \( m' \) further increases \( f \) (we ignore this gain), but combining the leftovers into \( V'_0 \) reduces \( f \). Combining singletons with \( V_0 \) would lose the most. The contributions of a singleton part with all other parts sum to at most the degree of the singleton vertex. Since \( m'k2^{k} \) is an upper bound on \(|V'_0 - V_0|\), the amount lost by forming \( V'_0 \) is at most \( m'k2^{k}n \).

Since \( km \geq n(1 - \epsilon) \), we now have \( f(G, \Pi') \geq f(G, \Pi) + \epsilon^5n^2(1 - \epsilon)^2 - m'k2^{k}n \). The gain is at least \( \epsilon^6n^2 - [\epsilon^5n^2(2\epsilon - \epsilon^2) + m'k2^{k}n] \). To ensure net gain at least \( \epsilon^5n^2/2 \), we want \( 2\epsilon^6n + m'k2^{k} < \epsilon^5n/2 \). With \( \epsilon < 1/5 \), it suffices to choose \( m' \) so that \( m'k2^{k} \leq \epsilon^6n/2 \). Set \( m' = \lceil \epsilon^6n/(k2^{k+1}) \rceil \). Now \( \Pi' \) is an \( \epsilon, k' \)-partition with \( k' \) bounded by a bit more than \( k2^{k+1}/\epsilon^6 \). Doubling this value suffices. Finally, note that the exceptional part has gained at most \( \epsilon^6n/2 \) vertices.

The simple formula for the gain per split is why we keep equal size for the nonexceptional parts. Later the exceptional part can be distributed to the others.

**Proof of the Szemerédi Regularity Lemma (as Theorem 11.1.15):** Given positive constants \( \epsilon \) and \( l \), we seek positive integers \( M \) and \( N \) such that every graph with at least \( N \) vertices has an \( \epsilon \)-regular \( \epsilon \), \( k \)-partition for some \( k \in [l, M] \).

To apply Lemma 11.1.27, we need \( \epsilon < \frac{1}{4} \). Since an \( \epsilon \)-regular \( \epsilon \), \( k \)-partition is also an \( \epsilon' \)-regular \( \epsilon' \), \( k \)-partition when \( \epsilon' > \epsilon \), we may assume \( \epsilon < \frac{1}{2} \). Similarly, if the claim is true for a given value of \( l \), it remains true for smaller values, so we may assume \( l > 2/\epsilon \).

Given such \( \epsilon \) and \( l \), define \( M \) and \( N \) as follows. Let \( L_0 = l \) and \( L_{i+1} = \lceil L_i2^{L_i/\epsilon^6} \rceil \). Let \( M = L_{\lfloor \log \epsilon \rfloor} \), and choose \( N \) with \( N > M/l/\epsilon^6 \). Let \( G \) be an \( n \)-vertex graph, where \( n \geq N \). Since \( l > 2/\epsilon \), we can form an \( \epsilon \), \( l \)-partition of \( G \) by breaking \( V(G) \) arbitrarily into \( l \) parts of size \( \lceil n/l \rceil \), with fewer than \( \epsilon n/2 \) vertices left over for the exceptional part.

Since \( |E(G)| < n^2/2 \), fewer than \( \epsilon^{-5} \) iterations of the refinement procedure in Lemma 11.1.27 produce an \( \epsilon \)-regular \( \epsilon \), \( k \)-partition of \( G \). The initial exceptional set has fewer than \( \epsilon n/2 \) vertices, and the exceptional set gains fewer than \( \epsilon n/2 \) with each iteration. With fewer than \( \epsilon^{-5} \) iterations, the final exceptional set has at most \( \epsilon n/2 \) vertices. Also, the final partition has at most \( M \) parts.

**11.1.28.* Remark.** Since the upper bound on the number of parts in Lemma 11.1.27 is \( k2^{k+1}/\epsilon^6 \), iterating \( \epsilon^{-5} \) times makes \( N \) an exponential tower with height \( \epsilon^{-5} \). Thus the Regularity Lemma applies only to enormous graphs.

We have given only a few applications of Regularity; there are many others. Komlós–Simonovits [1996] is an influential early survey; others include Komlós–Shokoufandeh–Simonovits–Szemerédi [2002], Kohayakawa–Rödl [2003], Rödl–Schacht [2010]. Kühn–Osthus [2009] explains the use of the regularity/blow-up method in general. The original lemma gives useful results only for dense graphs; Kohayakawa [1997] developed an analogue for sparse graphs (see also Kohayakawa–Rödl [2003], Gerke–Steger [2005], and Scott [2010]). Frankl–Rödl
EXERCISES 11.1

11.1.1. (−) Complete the computation in Corollary 11.1.7.

11.1.2. (−) Find the minimum size of a maximal triangle-free \(n\)-vertex graph.

11.1.3. (−) Determine \(\text{ex}(n; P_3)\) and \(\text{ex}(n; P_4)\).

11.1.4. Determine the maximum size of \(n\)-vertex graphs of the following types.
(a) Graphs having an independent set of size \(k\).
(b) Graphs having \(k\) components.

11.1.5. Let \(G\) be a connected graph having neither graph below as an induced subgraph. Prove that \(G\) is a complete multipartite graph.

11.1.6. Prove that every \(n\)-vertex graph not containing \(K_{r+1}\) has at most \((1 - \frac{1}{r})\frac{n^2}{2}\) edges.

11.1.7. (c) Turán’s proof of Turán’s Theorem. Recall that \(t_r(n) = |E(T_n, r)|\).
(a) Prove that a maximal graph with no \((r+1)\)-clique has an \(r\)-clique.
(b) For \(n \geq r\), prove that \(t_r(n) = \left(\binom{r}{2}\right)\left(\left\lfloor \frac{n-r}{r-1}\right\rfloor + r\right)\).
(c) Use parts (a) and (b) to prove Turán’s Theorem by induction on \(n\), including the uniqueness of graphs achieving the bound. (Turán [1941])

11.1.8. Let \(S\) be the set of nonnegative vectors in \(\mathbb{R}^n\) with sum 1. Given a graph \(G\) with vertex set \(\{v_1, \ldots, v_k\}\), let \(f(x) = \sum_{(i,j) \in E(G)} x_i x_j\) for \(x \in S\), and let \(\rho = \max_{x \in S} f(x)\).
(a) Prove that \(f\) is maximized by some vector whose nonzero coordinates correspond to the vertices of a clique. Conclude \(\rho = \frac{k}{2}(1 - 1/\omega(G))\).
(b) Prove that \(\rho\) is attained by a vector \(x\) with all coordinates nonzero if and only if \(G\) is a complete multipartite graph.
(c) Prove Turán’s Theorem for \(K_{r+1}\)-free \(n\)-vertex graphs by induction on \(n - \lfloor n/r \rfloor\), using parts (a) and (b) for the base case. (Motzkin–Straus [1965])

11.1.9. (c) The Turán graph \(T_{n,r}\) with size \(t_r(n)\) is the complete \(r\)-partite graph with \(b\) parts of size \(a + 1\) and \(r - b\) parts of size \(a\), where \(a = \lfloor n/r \rfloor\) and \(b = n - ra\).
(a) Prove that \(t_r(n) = (1 - 1/r)n^2/2 - b(r - b)/(2r)\).
(b) Since \(t_r(n)\) is an integer, part (a) yields \(t_r(n) \leq \left\lfloor (1 - 1/r)n^2/2\right\rfloor\). Determine the smallest value of \(r\) such that strict inequality occurs for some \(n\). For this value of \(r\), determine the values of \(n\) such that \(t_r(n) < \left\lfloor (1 - 1/r)n^2/2\right\rfloor\).

11.1.10. Let \(a = \lfloor n/r \rfloor\). Compare the Turán graph \(T_{n,r}\) with the graph \(K_a + K_{n-a}\) to prove directly that \(t_r(n) = \left(\binom{n-a}{2}\right) + (r-1)(\binom{n-a}{r-1})\).

11.1.11. Given positive integers \(n\) and \(k\), let \(q = \lfloor n/k \rfloor, r = n - qk, s = \lfloor n/(k+1) \rfloor\), and \(t = n - s(k+1)\). Use Turán’s Theorem to prove \(\left(\frac{q}{2}\right)k + rq \geq \left(\frac{r}{2}\right)(k+1) + ts\). (Richter [1993])
11.1.12. Let $G$ be a graph with $n$ vertices and $m$ edges. Determine the best possible lower bounds on $\alpha(G)$ and $\omega(G)$.

11.1.13. Let $S$ be a set of $n$ points in a circular region with radius 1, and consider the distance between any two points in $S$. Prove that the distance exceeds $\sqrt{2}$ for at most $n^2/3$ of the pairs. (Bondy–Murty [1976, p.114])

11.1.14. (*) Let $G$ be a graph with $n$ vertices that has $t(n) - k$ edges and at least one $(r+1)$-clique, where $k \geq 0$. Prove that $G$ has at least $f(n) + 1 - k$ cliques of size $r + 1$, where $f(n) = n - \lceil n/r \rceil - r$. (Hint: Prove that a graph with exactly one $(r+1)$-clique has at most $t(n) - f(n)$ edges.) (Erdős [1964], Moon [1965])

11.1.15. (*) Let $f_3(n)$ be the minimum size of a family $F$ of subsets of $[n]$ such that every subset of $[n]$ with size at most $k$ is the union of two (possibly equal) members of $F$. Note that $f_3(n) = f_3(n) = n + 1$. (Comment: Füredi–Katona [2006] determined also $f_4(n)$.)

(a) Prove $\sqrt{2} \cdot 2^{n/2} - 1/2 \leq f_3(n) \leq 2^{\lceil n/2 \rceil} + 2^{\lceil n/2 \rceil} - 1$.
(b) Prove $f_3(n) = 1 + n + \left(\begin{smallmatrix} n \\ 2 \end{smallmatrix}\right) - \left\lfloor n^2/4 \right\rfloor$.

11.1.16. Extensions of Turán’s Theorem. As usual, let $t_r(n) = |E(T_{n,r})|$.

(a) If $n \geq r + 1 \geq 4$, prove that every $n$-vertex graph with $t_{r-1}(n) + 1$ edges contains $K_{r+1} - e$ as a subgraph. (Hint: Remove a vertex of minimum degree.)
(b) If $n \geq n' \geq r \geq 2$, prove that every $n$-vertex graph with $t_r(n) + p$ edges has an $n'$-vertex subgraph with $t_{r'}(n') + p$ edges, where $p \leq 1$.
(c) If $1 \leq p \leq r - 1$, prove that every $n$-vertex graph with $t_r(n)$ edges other than $T_{r,n}$ has an $(r+p)$-vertex subgraph with more than $(r+p) - p$ edges. (Dirac [1963])

11.1.17. (*) The Turán graph $T_{n,r}$ is regular when $r$ divides $n$. Prove that there is a $k$-regular triangle-free graph on $n$ vertices if and only if $k$ is even and at least $5k/2$ or $n$ is odd and at least $5k/2$, with $k$ even when $n$ is odd. (Comment: Bauer [1983] determined all triples $(r, n, k)$ such that there is a $k$-regular $K_{r+1}$-free graph on $n$ vertices.)

11.1.18. Let $G$ be the Petersen graph. For $n \geq 2$, prove $ex(n, G) \geq 2n - 3 + \left\lfloor (n-2)^2/4 \right\rfloor$. (Comment: Simonovits [1999] proved that equality holds.)

11.1.19. Prove that the minimum number of edges in a connected $n$-vertex graph where every edge lies in a triangle is $3(n-1)/2$ if $n$ is odd, $3n/2 - 1$ if $n$ is even. (Erdős [1988a])

11.1.20. (*) Determine the minimum number of edges in a triangle-free graph on $2n$ vertices whose complement contains no $n$-clique. (Erdős [1988b])

11.1.21. Determine the maximum number of edges in a 10-vertex graph with no 4-cycle. (Bialostocki–Schönheim [1984])

11.1.22. For $n \geq 6$, prove that the maximum number of edges in an $n$-vertex graph not having two edge-disjoint cycles is $n + 3$. (Erdős–Pósa [1962])

11.1.23. (*) Prove that if $|V(G)| \geq 6$ and $\delta(G) \geq 4$, then $G$ has two disjoint cycles. Conclude that for $n \geq 6$ the maximum number of edges in an $n$-vertex graph not having two disjoint cycles is $3n - 6$. (Erdős–Pósa [1962]) (Comment: Corradi–Hajnal [1963] proved more generally that if $|V(G)| \geq 5k$ and $\delta(G) \geq 2k$, then $G$ has $k$ pairwise disjoint cycles.)

11.1.24. Let $G$ be a connected graph with $m$ edges and more than one vertex.

(a) Prove that if $G$ is $P_4$-free, then $\overline{G}$ is disconnected. (Seinsche [1974])
(b) Prove that if $\Delta(G) = D$ and $\overline{G}$ is disconnected, then $m \leq D^2$, with equality only for $K_{D,D}$. (Chung–West [1993])

11.1.25. (*) Let $G$ be a graph on $n$ vertices such that $\overline{G}$ has no triangles. For $n \geq 6$, prove that the minimum possible number of triangles in $G$ is $\left(\begin{smallmatrix} n/3 \\ 3 \end{smallmatrix}\right) + \left(\begin{smallmatrix} n/3 \\ 3 \end{smallmatrix}\right)$. 

Chapter 11: Extremal Problems
11.1.26. For an $n$-vertex graph $G$, prove the bounds below and show that they are sharp.

$$k_3(G) + k_3(G) \geq \begin{cases} 2\binom{n}{3} & \text{if } n = 2l \\ \frac{4}{3}k(k-1)(4k+1) & \text{if } n = 4k + 1 \\ \frac{2}{3}k(k+1)(4k-1) & \text{if } n = 4k + 3 \end{cases}$$

11.1.27. Let $G$ be a self-complementary $n$-vertex graph, and let $k = \lfloor n/4 \rfloor$. Conclude from Exercise 11.1.26 that $G$ has at least $k(k-1)(4k-3)/3$ triangles if $n = 4k$ and that $G$ has at least $k(k-1)(4k+1)/3$ triangles if $n = 4k + 1$.

11.1.28. Let $H_k = K_1 \cup kK_2$. Prove that $\text{ex}(n; H_k) = \lfloor n^2/4 \rfloor + 1$ for $n \geq 5$. (Comment: Erdős–Füredi–Gould–Gunderson [1995] proved $\text{ex}(n; H_k) = \lfloor n^2/4 \rfloor + k^2 - k$ for odd $k$ and $\text{ex}(n; H_k) = \lfloor n^2/4 \rfloor + k^2 - 3k/2$ for even $k$.)

11.1.29. A graph $G$ with vertex degrees $d_1, \ldots, d_n$ is $r$-majorizable if there is an $r$-partite graph $H$ with vertex degrees $d'_1, \ldots, d'_n$ such that $d'_i \geq d_i$ for $1 \leq i \leq n$. By Erdős’ proof of Turán’s Theorem (Theorem 5.2.11), every $K_{r+1}$-free graph is $r$-majorizable.

(a) Construct non-$r$-majorizable $n$-vertex graphs containing only one copy of $K_{r+1}$.

(b) Prove that $G$ is $r$-majorizable if $\Delta(G) \leq n(1 - 1/r)$.

(c) Consider $n, D, r \in \mathbb{N}$ such that $n - 1 \geq r(n-D) + \binom{r}{2}$, where $0 \leq t < r$. Given $n_1, \ldots, n_r$ with sum $n - 1$ such that $n - D = n_1 = \cdots = n_t = n_{r+1} < \cdots < n_r$, construct $G$ from $K_{n_1, \ldots, n_r}$ by adding one vertex adjacent to all vertices in $t$ smallest parts and one vertex in each other part. Prove that $G$ is not $r$-majorizable.

(d) The graph of part (c) has $(n-D)t$ copies of $K_{r+1}$. For $r = 2$, prove that this is the fewest among non-$r$-majorizable $n$-vertex graphs with $\Delta(G) = D$. (Comment: West [1982c] proved this for $r = 2$ and also for $(n, D, r) = (7, 5, 3)$; does it always hold?)

11.1.30. Construct a 3-uniform hypergraph with asymptotically $\frac{2}{3} \binom{n}{3}$ edges that does not contain $K_4^{(3)}$, the complete 3-uniform hypergraph with four vertices. (Comment: Turán [1941] conjectured that this is asymptotically the most edges in such an $n$-vertex hypergraph: that is, $\text{ex}(n; K_4^{(3)})/\binom{n}{3} \to 5/9$. Erdős offered $1000$ for a proof. Chung–Lu [1999] proved an upper bound of .593, and Razborov [2011] lowered it to .561.)

11.1.31. (∗) Erdős–Stone Theorem (Theorem 11.1.10). Fix $\varepsilon \in (0, 1)$, and consider $t, r \in \mathbb{N}$.

(a) Prove that when $n$ is sufficiently large, every $n$-vertex graph with minimum degree at least $(1 - 1/r + \varepsilon)n$ contains $K_{t+1}[r]$. (Hint: Use induction on $r$.) (See Lovász [1993])

(b) Prove $\text{ex}(n; K_{t+1}[r]) \leq (1 - 1/r + \varepsilon)n^2/2$ for sufficiently large $n$. (Erdős–Stone [1946])

11.1.32. An ordered graph is a graph with linear order on the vertices. The interval chromatic number of an ordered graph is the least number of independent sets of consecutive vertices needed to cover the vertices (ranging from 2 to $n$ for $n$-vertex ordered paths).

An ordered graph $G$ avoids an ordered graph $H$ if $H$ does not appear in $G$ via an order-preserving injection on the vertices. Let $\text{ex}(n; H)$ be the maximum number of edges in an $n$-vertex ordered graph that avoids $H$. Use the Erdős–Stone Theorem to prove that if $H$ has interval chromatic number $k$, then $\text{ex}(n; H) = \left(1 - \frac{1}{e^2}\right)\binom{n}{2} + o(n^2)$. (Pach–Tardos [2006])
11.1.33. A graph $G$ is $F$-saturated if it does not contain $F$, but adding any edge creates a copy of $F$. For $n \geq t - 1$, prove that $K_{t-2} \phi R_{n-2}$ has the fewest edges among all $K_t$-saturated graphs with $n$ vertices. (Érdős–Hajnal–Moon [1964])

11.1.34. (⋆) Prove that there is an $n$-vertex $F$-saturated graph (see Exercise 11.1.33 for definition) having at most $(\beta(F) - 1)n + \Delta(F)n/2$ edges, where $\beta(F)$ is the vertex cover number of $F$. (Kászonyi–Tuza [1986]; clarified in Füredi [2007])

11.1.35. Let $\text{sat}(n; F)$ be the minimum number of edges in an $F$-saturated graph with $n$ vertices (see Exercise 11.1.33 for definition).

(a) Determine $\text{sat}(n; K_3)$.

(b) With $P$ being the Petersen graph, show that $\text{sat}(n; P) \leq 4n - 4$.

11.1.36. Given the Regularity Lemma (Theorem 11.1.15), prove that one can instead require an $\epsilon$-regular partition in which each part belongs to at most $\epsilon k^2$ irregular pairs (instead of a total of $\epsilon k^2$ irregular pairs).

11.1.37. Let $G$ be a graph with $2n$ vertices, partitioned into sets $A$ and $B$ of size $n$. Show that if $|[A, B]| \geq (1/2 + \epsilon)n^2$, then $G$ contains $K_4$ or an independent set of size $\epsilon n^2$.

11.1.38. Let $G$ be a graph with $3n$ vertices partitioned into sets $A$, $B$, and $C$, each of size $n$. Show that if each pair among $\{A, B, C\}$ is $\epsilon$-regular and has density more than $2\epsilon$, then $G$ contains $K_4$ or an independent set of size more than $\epsilon^2 n$.

11.1.39. The Slicing Lemma. Let $(A, B)$ be an $\epsilon$-regular pair with density $d$, where $0 < \epsilon \leq \alpha \leq 1/2$. Prove that for any $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq \alpha |A|$ and $|B'| \geq \alpha |B|$, the pair $(A', B')$ is $\epsilon/\alpha$-regular with density greater than $d - \epsilon$.

11.1.40. (⋆) Let $G$ be an $A, B$-bipartite graph with $|A| = |B| = n$.

(a) Prove that if $G$ has no independent set consisting of $\delta(G)$ vertices in $A$ and $\delta(G)$ vertices in $B$, then $G$ has a perfect matching.

(b) Obtain a perfect matching when $(A, B)$ is an $\epsilon$-regular pair and $\delta(G) > \epsilon n$.

11.1.41. (⋆) Given an $\epsilon$-regular pair $(A, B)$ with density $d$ and $|A| = |B| = m$, let a vertex $v$ in $A \cup B$ be good for the pair if it has at least $(d - \epsilon)m$ neighbors in the other part.

(a) Let $x \in A$ and $y \in B$ be fixed good vertices in the pair $(A, B)$ as specified above. Prove that if $d > 5\epsilon$, then when $m$ is sufficiently large there is an $x, y$-path of length at least $(1 - \epsilon - \frac{1}{2}d\epsilon)2m$ that avoids any $K$ specified vertices.

(b) Given $d > 5\epsilon$, let $R$ be the reduced graph of $\epsilon$-regular pairs with density at least $d$ resulting from an $\epsilon$-regular partition of an $n$-vertex graph $G$. Prove that if $R$ has $k$ vertices and some component of $R$ has a matching of size $t$, then $G$ has a cycle of length at least $(1 - \epsilon - \frac{1}{2}d\epsilon)^2 t/2 n$. (See Figaj–Łuczak [2007] for similar results.)

11.1.42. (+) Let $G$ be an $A, B$-bipartite graph with $|A| = |B| = n$ and $\delta(G) \geq dn$. Prove that if $(A, B)$ is an $\epsilon$-regular pair and $d > 3\epsilon + \sqrt{\epsilon}$, then $G$ has a spanning cycle. (Haxell [1997])

11.1.43. Let $(A, B)$ be an $\epsilon$-regular pair with density $d$ and $|A| = |B| = n$.

(a) For $k \in \mathbb{N}$, consider $Y \subseteq B$ such that $(d - \epsilon)^{k-1} |Y| \geq \epsilon |B|$. Prove that the number of $k$-tuples $(x_1, \ldots, x_k)$ of elements of $A$ having at most $(d - \epsilon)^k |Y|$ common neighbors in $Y$ is at most $ke^2 |A|^k$. (Komlós–Simonovits [1996]) (Comment: Lemma 11.1.20 is $k = 1$.)

(b) Prove that at least $\frac{1}{2}(1 - 4\epsilon)(d - \epsilon)^4 n^4$ 4-cycles alternate between $A$ and $B$.

11.1.44. (⋆) 6,3-Theorem (Ruzsa–Szemerédi [1978]).

(a) Let $G$ be an $n$-vertex graph with every edge in exactly one triangle. Use the Triangle-Removal Lemma (Lemma 11.1.22) to prove that $G$ has $o(n^2)$ edges.

(b) Let $G$ be an $n$-vertex graph that decomposes into at most $n$ induced matchings (vertices of distinct edges in such a matching are not adjacent). Prove that $G$ has $o(n^2)$ edges.

(c) Let $H$ be a 3-uniform $n$-vertex hypergraph having no six vertices that induce at least three edges. Prove that $H$ has $o(n^2)$ edges.
11.1.45. (♦) Fix $c \in (0, 1)$. Prove that for some $n \in \mathbb{N}$, any set of at least $cn^2$ points in $[n]^2$ contains three distinct points of the form $(x, y), (x + a, y), (x, y + a)$. (Hint: Use Lemma 11.1.22 on a suitable graph.) (Ajtai–Szemerédi [1974], Solymosi [2003])

11.2. Extremal Set Theory

In this section, we consider extremal problems for families of subsets of $[n]$. We ask for the fewest $(k-1)$-sets covered by a fixed number of $k$-sets, the maximum size of a family with no set contained in another, and largest families of pairwise intersecting subsets. West [1982d], Frankl [1987], and Kleitman [1994] survey these and related topics. See also the books Engel [1997], Jukna [2011], and Gerbner–Patkós [2019]. Frankl–Tokushige [2016] is an authoritative survey.

THE KRUSKAL–KATONA THEOREM

In extensions of the Turán problem, we wanted to choose $m$ edges to have the fewest $r$-cliques. Similarly, we may ask for $m$ vertex sets of size $k$ that contain the fewest sets of size $k - 1$.

11.2.1. Definition. A family of sets or set system is a set whose members are sets. A $k$-uniform family is a family of $k$-sets.

We use “member of family” to avoid confusion with “element of set”. A family of subsets of $[n]$ is the edge set of a hypergraph with vertex set $[n]$, and “$k$-uniform” has the same meaning in both contexts. In Section 10.2 we used hypergraph language to generalize notions from graph theory. Some authors use the term “hypergraph” just for a family of sets without specifying the vertex set.

11.2.2. Definition. The $t$-shadow of a family $F$ is the family of all $t$-sets contained in members of $F$. The shadow $\partial F$ of a $k$-uniform family $F$ is its $(k-1)$-shadow. The shade is the family of all $(k+1)$-sets containing members of $F$.

Among all $k$-uniform families of size $m$, we seek the family $F$ with smallest shadow. Since each $k$-set contributes exactly $k$ sets to the shadow and $|F|$ is fixed, the shadow is minimized when the sets in the shadow arises many times. This can be achieved by confining $F$ to subsets of a small subset of $[n]$. The amazing result is that there exists an ordering of the $k$-sets such that, for each $m$, the first $m$ sets in this ordering form an optimal family.

To confine the initial sets to a small subset of $[n]$, we define an ordering of $k$-sets in which $x$ precedes $y$ if the largest element of $x$ is less than the largest element of $y$. We use notation like $x$ and $y$ for $k$-element subsets to emphasize their interpretation as binary incidence vectors.

11.2.3. Definition. The co-lexicographic or colex ordering on a family of $k$-sets is obtained by putting $x < y$ if $x_i < y_i$ in the highest coordinate where their binary incidence vectors differ.
### 11.2.4. Example. Colex ordering on $\binom{[n]}{k}$. It is convenient to start the indexing at 0 for both the list of vectors and the coordinates of the vectors.

<table>
<thead>
<tr>
<th>index</th>
<th>set</th>
<th>positions of 1s</th>
<th>index</th>
<th>set</th>
<th>positions of 1s</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>11100</td>
<td>012</td>
<td>5</td>
<td>10101</td>
<td>024</td>
</tr>
<tr>
<td>1</td>
<td>11010</td>
<td>013</td>
<td>6</td>
<td>01101</td>
<td>124</td>
</tr>
<tr>
<td>2</td>
<td>10110</td>
<td>023</td>
<td>7</td>
<td>10011</td>
<td>034</td>
</tr>
<tr>
<td>3</td>
<td>01110</td>
<td>123</td>
<td>8</td>
<td>01011</td>
<td>134</td>
</tr>
<tr>
<td>4</td>
<td>11001</td>
<td>014</td>
<td>9</td>
<td>00111</td>
<td>234</td>
</tr>
</tbody>
</table>

To facilitate the proof that the first $m$ sets in this order form an optimal family, we study its shadow.

#### 11.2.5. Lemma. If the vector with index $m$ in the colex ordering on $\binom{[n]}{k}$ has 1 in positions $m_1, \ldots, m_k$, with $m_1 < \cdots < m_k$, then $m = \binom{m_1}{k} + \binom{m_{k-1}}{k-1} + \cdots + \binom{m_1}{1}$.

**Proof:** Let $\sigma$ be the vector with index $m$. Indexing starts with 0, so $m$ counts the vectors preceding $\sigma$. We count them another way. To reach $\sigma$, we must skip all vectors whose $k$th 1 appears before position $m_k$. There are $\binom{m_1}{k}$ of these, since the first position is position 0. Some vectors with the last 1 in position $m_k$ also precede $\sigma$. These begin with $\binom{m_{k-1}}{k-1}$ vectors having $k-1$ 1s in positions before $m_{k-1}$. Continuing, the $j$th term in the summation counts the vectors that appear before $\sigma$ in the ordering and have their last $j-1$ 1s in the same positions as $\sigma$.

#### 11.2.6. Corollary. For $m, k \in \mathbb{N}$, there is a unique expression of $m$ in the form $\binom{m_1}{k} + \binom{m_{k-1}}{k-1} + \cdots + \binom{m_1}{1}$ with $m_k > \cdots > m_i \geq i$.

**Proof:** In the colex ordering on $\binom{[n]}{k}$, appending 0s increases the number of coordinates without changing the position of the first $\binom{[n]}{k}$ vectors. Lemma 11.2.5 thus applies for any $n$ with $\binom{[n]}{k} > m$ to obtain an expression for $m$ using the vector with index $m$ in the ordering. Left-justified 1s in the vector yield $m_j = j - 1$ and contribute 0; we drop those terms from the expression in Lemma 11.2.5 to obtain an expression of the desired form. Two such expansions correspond to distinct vectors in the colex ordering and hence distinct values of $m$.

#### 11.2.7. Definition. The unique representation of $m$ described in Corollary 11.2.6 is the $k$-binomial expansion of the integer $m$. For an integer $m$ with $k$-binomial expansion $m = \binom{m_1}{k} + \binom{m_{k-1}}{k-1} + \cdots + \binom{m_1}{1}$ such that $m_k > m_{k-1} \cdots > m_i \geq i$, let $\partial_k(m) = \binom{m_1}{k-1} + \binom{m_{k-1}}{k-2} + \cdots + \binom{m_{i-1}}{1}$.

#### 11.2.8. Lemma. The shadow of the first $m$ vectors in the colex order on $\binom{[n]}{k}$ consists of the first $\partial_k(m)$ vectors in the colex order on $\binom{[n]}{k-1}$.

**Proof:** We count the shadow using Lemma 11.2.5. The family consisting of all elements of $\binom{[n]}{k}$ whose last 1 appears before position $m_k$ has shadow consisting of all vectors of weight $k-1$ whose last 1 precedes position $m_k$. In general, the $j$-th term in the summation for $\partial_k(m)$ considers all vectors before the indexed vector $v_m$ in $\binom{[n]}{k}$ that have their rightmost $j-1$ ones in the same positions as $v_m$. It counts the portion of their shadow consisting of sets whose last $j-1$ ones are in those same positions. This is the portion of the shadow not counted by earlier terms. Thus the full sum counts the entire shadow exactly. ■
11.2.9. Definition. On a family \( F \) in \( 2^{[n]} \), the \textbf{shift operator} \( \tau_{i,j} \) is defined by

\[
\tau_{i,j}(x) = \begin{cases} 
  x - \{j\} + \{i\} & \text{if } j \in x \text{ and } i \notin x \text{ and } x - \{j\} + \{i\} \notin F, \\
  x & \text{otherwise.}
\end{cases}
\]

Let \( \tau_{i,j}(F) = \{ \tau_{i,j}(x) : x \in F \} \).

11.2.10. Lemma. If \( F \subseteq 2^{[n]} \) and \( i < j \), then \( \partial(\tau_{i,j}(F)) \subseteq \tau_{i,j}(\partial F) \).

\textbf{Proof}: Let \( G = \partial F \). Let \( \tau_F \) and \( \tau_G \) denote \( \tau_{i,j} \) on \( F \) and \( G \), respectively. It suffices to show \( y \in \tau_G(G) \) when \( y \in \partial(\tau_F(F)) \). For \( y \in \partial(\tau_F(F)) \), we have \( y = x' - s \) for some \( s \in x' \in \tau_F(F) \). Note also that \( x' = \tau_F(x) \) for some \( x \in F \).

\textbf{Case 1}: \( x' = x \). In this case \( y \in G \), and it suffices to show \( \tau_G(y) = y \). If not, then \( i \notin y \) and \( j \notin y \). Since \( j \in y \) and \( y = x - s \), we have \( s \notin j \) and \( j \in x \). If \( s = i \), then \( y - j + i = x - j \in G \), so \( \tau_G(y) = y \). If \( s \neq i \), then \( i \notin y \) yields \( i \notin x' \); now \( \tau_F(x) = x \) implies \( x - j + i \in F \). Hence \( y - j + i \in \partial F = G \), so \( \tau_G(y) = y \).

\textbf{Case 2}: \( x' \neq x \). Now \( x \) contains \( j \) but not \( i \), and \( x' = x - j + i \). Since \( y = x' - s \), we have \( j \notin y \), and hence \( \tau_G(y) = y \) if \( y \in G \). If \( i \notin y \), then \( i = s \) and \( y = x - j \in G \). Hence we may assume \( i \in y \) and \( y \notin G \). Let \( \hat{y} = y + j - i \), so now \( \hat{y} \in \partial G \subseteq \partial \). With \( \hat{y} \in G \) and \( \hat{y} - j + i = y \in G \), we have \( \tau_G(\hat{y}) = y \), so \( y \in \tau_G(G) \) as desired. \( \blacksquare \)

11.2.11. Theorem. (Kruskal–Katona Theorem; Kruskal[1963], Katona[1968])

For \( F \subseteq \binom{[n]}{i} \) with \( |F| = m \), the minimum of \( |\partial F| \) occurs when \( F \) consists of the first \( m \) members of the colex ordering on \( \binom{[n]}{k} \), and then \( |\partial F| = \partial_k(m) \).

\textbf{Proof}: (Frankl[1984]) By Lemma 11.2.8, the shadow of the specified family has size \( \partial_k(m) \), so it suffices to show \( |\partial F| \geq \partial_k(m) \) for all \( F \subseteq \binom{[n]}{i} \) with \( |F| = m \).

We use induction on \( k \); trivially the claim holds for a 1-uniform family. For the induction step, we use induction on \( m \). Note that \( m = \binom{k}{i} \) when \( m = 1 \), and indeed then \( |\partial F| = k = \binom{k}{i-1} \). Now consider \( m > 1 \).

Fix \( i \in [n] \). Applying any \( \tau_{i,j} \) does not change the number of sets, and it increases the number of sets containing \( i \) if it produces a change. From \( F \), application of finitely many such operators therefore leads to a family \( F^* \), invariant under all \( \tau_{i,j} \), in which \( i \) is a dominant element, meaning \( x - \{j\} + \{i\} \in F^* \) for all \( j \in x \in F^* \). By Lemma 11.2.10, \( |\partial F^*| \leq |\partial F| \). Hence it suffices to prove \( |\partial F| \geq \partial_k(m) \) in the case where \( F \) has a dominant element \( i \).

From \( F \) define \( F' = \{ x : i \notin x \in F \} \) and \( F^- = \{ x - i : i \in x \in F \} \). Note \( F' \subseteq \binom{[n]}{i-1} \), \( F^- \subseteq \binom{[n]}{i+1} \), and \( |F'| + |F^-| = m \). If \( j \in x \in F' \) and \( i \notin x \), then \( x - j + i \in F \), since \( i \) is dominant. Hence \( x - j \in F^- \) and \( \partial F^* \subseteq F^- \). We conclude that \( \partial F \) consists of \( F^- \) plus the sets formed by adding \( i \) to members of the shadow of \( F^- \).

Let \( \sum_{j=1}^{k} \binom{m-1}{j} \) be the \( k \)-binomial expansion of \( m \), as in Definition 11.2.7. Let \( m' = \sum_{j=1}^{k} \binom{m-1}{j} \). If \( |F^-| < \partial_k(m') \), then \( |F'| = m - |F^-| > m - \partial_k(m') = m' \). Now, since \( |F^*| < |F| = m \), the induction hypothesis on \( m \) yields \( |\partial F'| \geq \partial_k(m') > |F^-| \), which contradicts \( \partial F^* \subseteq F^- \).

Hence we may assume \( |F^-| \geq \partial_k(m') \). In this case the induction hypothesis on \( k \) yields \( |\partial F^-| \geq \partial_{k-1}(m') \). Finally, we compute

\[
|\partial F| \geq |F^-| + |\partial F^-| \geq \partial_k(m') + \sum_{j=1}^{k} \binom{m-1}{j-2} \geq \sum_{j=1}^{k} \binom{m-1}{j-1} = \partial_k(m).
\]

\( \blacksquare \)
Note that the optimal family depends only on \( m \) and \( k \), not \( n \), as long as \( \binom{n}{k} \geq m \). The Kruskal–Katona Theorem is a very precise statement. For many applications, a smoothed version due to Lovász [1979] suffices. It treats \( m \) as a value of the “choose \( k \)” polynomial and can also be proved using shift operators (Exercise 9). Instead, we give a short more recent proof. Exercise 10 requests the characterization of when equality holds.

**11.2.12. Theorem.** (Lovász [1979, Ex. 13.31(b)]) For \( u \geq k \), let \( \binom{u}{k} = \frac{u^k}{k!} \). If \( F \subseteq \binom{[n]}{k} \) with \( |F| = \binom{u}{k} \), then \(|\partial F| \geq \binom{u+1}{k+1}\).

**Proof:** (Keevash [2008]) Note that \( F \) is contained among the \( k \)-sets whose \((k-1)\)-subsets are all present in \( \partial F \). Letting \( r = k - 1 \), it therefore suffices to show that when \( G \) is a family of size \( \binom{u}{r} \) in \( \binom{[n]}{r} \), the number of \((r+1)\)-sets whose \( r \)-sets are all present in \( G \) is at most \( \binom{u+1}{r+1} \).

We use induction on \( r \); the claim is immediate when \( r = 1 \). For \( r > 1 \), we treat \( G \) as the edge set of a hypergraph. We may assume that each \( v \in [n] \) lies in some member of \( G \). Let \( d(v) \) be the number of edges of \( G \) containing \( v \). Let \( G'(v) \) be the family of \((r-1)\)-sets obtained by deleting \( v \) from edges of \( G \). Let \( q(v) \) be the number of \((r+1)\)-sets containing \( v \) whose \( r \)-sets all lie in \( G \).

Note that \( S \cup \{v\} \) is counted by \( q(v) \) if and only if \( S \subseteq G \) and all \((r-1)\)-subsets of \( S \) lie in \( G'(v) \). The first condition implies \( q(v) \leq |G| - d(v) \), while the second bounds \( q(v) \) by the number of \( r \)-sets whose \((r-1)\)-subsets all lie in \( G'(v) \).

We claim that for all \( v \), at least one of these bounds is at most \( \frac{u-r}{r}d(v) \). If \( d(v) \geq \binom{r-1}{r-1} \), then \( \binom{r}{r} - d(v) \leq \frac{u}{r} \binom{r-1}{r-1} - d(v) \leq \frac{u-r}{r}d(v) \). If \( d(v) \leq \binom{r-1}{r-1} \), then define \( u' \) by \( d(v) = \binom{u'-1}{r-1} \); note that \( u' \leq u \). Note also that \( |G'(v)| = d(v) \).

By the induction hypothesis, at most \( \binom{r-1}{r} \) elements of \( \binom{[n]}{r} \) have all their \((r-1)\)-sets in \( G'(v) \). We compute \( \binom{u'-1}{r-1} = \frac{u-r}{r} \binom{u-1}{r-1} \leq \frac{u-r}{r}d(v) \).

Finally, every \((r+1)\)-set whose \( r \)-sets are all present in \( G \) is counted \( r+1 \) times in \( \sum_{v \in [n]} q(v) \). Hence the number of these sets is bounded by \( \frac{u-r}{r(r+1)} \sum_{v \in [n]} d(v) \).

Since the degree sum is \( r|G| \), which equals \( r\binom{u}{r} \), the bound simplifies to \( \binom{u}{r+1} \). \( \blacksquare \)


### ANTICHIANS AND INTERSECTING FAMILIES

We began with the Kruskal–Katona Theorem due to its relation to Turán-type problems, but the most fundamental problem in extremal set theory is to maximize the size of a family in which no member contains another. Sperner [1928] gave the answer. The theorem has many proofs, most of which extend to interesting additional results. Chapter 12 presents some of these extensions.
11.2.13. Definition. An antichain of sets is a family of sets in which no member contains another.

11.2.14. Theorem. (Sperner’s Theorem; Sperner [1928]) The maximum size of an antichain of subsets of \([n]\) is \(\binom{n}{\lfloor n/2 \rfloor}\), achieved only by the family of \([n/2]\)-sets or the family of \([n/2]\]-sets.

**Proof:** Let \(k\) and \(l\) be the sizes of the smallest and largest members of a maximum antichain \(F\). Let \(A = F \cap \binom{[n]}{l}\). Each \(l\)-set has \(l\) sets immediately below it, and each \((l-1)\)-set lies under \(n-l+1\) sets of size \(l\). Thus \(|\partial A| \geq |A| \frac{l}{n-l+1}\). If \(l > (n+1)/2\), then replacing the \(l\)-sets in \(F\) by their shadow yields a larger antichain.

Similarly, the shade of \(F \cap \binom{[n]}{k}\) has size at least \(|A| \frac{n-l}{l+1}\). If \(k < (n-1)/2\), then replacing the \(k\)-sets in \(F\) by their shade yields a larger antichain. Thus \((n-1)/2 \leq k \leq l \leq (n+1)/2\).

When \(n\) is even, we obtain \(k = l\). When \(n\) is odd, let \(A = F \cap \binom{[n]}{(n+1)/2}\). The argument to enlarge \(F\) still works unless \(|\partial A| = |A|\), which requires all sets above \(|\partial A|\) to lie in \(A\). Since one can move from any \(i\)-set to any other by alternately adding and deleting elements, \(|\partial A| = |A|\) occurs only when \(A\) is \(\emptyset\) or \(\binom{[n]}{n/2}\). Thus the sets in \(F\) all have the same size.

We can also restrict intersections of members of our family.

11.2.15. Definition. An intersecting family is a family \(F\) such that \(A, B \in F\) implies \(A \cap B \neq \emptyset\); it is \(t\)-intersecting if always \(|A \cap B| \geq t\). A star is a family of sets having a universal common element; a \(t\)-star is a family sharing \(t\) universal common elements.

This definition extends the graph-theoretic notion of star to the hypergraph context. Under various conditions, the large intersecting families are stars.

11.2.16. Example. Intersecting families. An intersecting family of subsets of \([n]\) has at most \(2^{n-1}\) members, since a set and its complement cannot both appear. The bound is achieved by the star consisting of all sets containing a fixed element.

Other largest intersecting families consist of all sets with more than half the elements, plus (for even \(n\)) one from each complementary pair of sets of size \(n/2\). Katona [1964] generalized this to determine the largest \(t\)-intersecting families (see Ahlswede–Khachatrian [2005] for a variety of proofs).

The family of subsets of \([n]\) containing one element and omitting another is an intersecting family whose complements also form an intersecting family. It is a largest such family, with size \(2^{n-2}\) (Seymour [1973], Schönheim [1974], Anderson [1976], Daykin–Lovasz [1976], Hilton [1976]; see Exercises 28–29.)

We next consider intersecting families of special types. Frankl [1995] surveys many such classical problems. We emphasize the Erdős–Ko–Rado Theorem on \(t\)-intersecting families consisting of an antichain of subsets of size at most \(k\).

11.2.17. Definition. An \(EKR(k, t)\)-family is an antichain \(F\) that is also a \(t\)-intersecting family whose members have size at most \(k\).
We will see that a largest such family uses only \( k \)-sets. The \( k \)-sets containing \( t \) particular elements form a \( t \)-star that is an EKR\((k,t)\) family of size \( \binom{n-t}{k-1} \). When \( n \) is sufficiently large, this is the extremal family. We first consider only the case \( t = 1 \), using a counting argument. In this case we may express the problem as a Turán-type problem; we seek a largest \( k \)-uniform hypergraph not containing the hypergraph consisting of two disjoint edges. Published in 1961, the EKR Theorem was actually proved in 1938, as noted by Erdős [1990].

11.2.18. **Theorem.** (Erdős–Ko–Rado Theorem \( t = 1 \); Erdős–Ko–Rado [1961])

For \( n \geq 2k \), the maximum size of an EKR\((k,1)\)-family is \( \binom{n-1}{k-1} \). For \( n > 2k \), equality holds only for stars.

**Proof:** [Katona [1972b]] If some member of such a family \( F \) has size less than \( k \), then replacing all smallest members by their shade strictly enlarges the family (since \( k \leq n/2 \)) while preserving the conditions. Thus we may assume \( F \subset \binom{n}{k} \).

Given a circular ordering \( \sigma \) of \([n]\), we ask how many members of \( F \) occur in \( \sigma \) as a consecutive segment of elements. Write such a member \( x \) as \( (a_1, \ldots, a_k) \), indexed in order in \( \sigma \). Every \( k \)-element segment that intersects \( x \) has a boundary immediately following some \( a_i \) with \( 1 \leq i < k \). Since \( k \leq n/2 \), two \( k \)-sets having the same boundary within \( x \) cannot intersect, as shown below. Hence at most \( k-1 \) members of \( F \) other than \( x \) appear in \( \sigma \).

Summing over all \((n-1)!\) circular orderings yields at most \((n-1)!k\) appearances of members of \( F \). Since each member of \( F \) appears in \( k!(n-k)! \) such orderings, \[ |F| \leq \frac{(n-1)!k}{k!(n-k)!} = \binom{n-1}{k-1}. \]

Equality requires having exactly \( k \) members of \( F \) in every circular ordering of \([n]\). With \( x \) occurring in \( \sigma \), for \( 1 \leq i < k \) the location between \( a_i \) and \( a_{i+1} \) must be the boundary of one member \( x_j \) of \( F \). If always \( x_i \) extends clockwise (or always counterclockwise) from there, then the \( k \) appearances successively shift by one element. Otherwise, some \( x_{i-1} \) and \( x_i \) extend in opposite directions. Since \( n > 2k \), they intersect only if they both contain \( a_i \). Now \( x_i, \ldots, x_{k-1}, x, x_1, \ldots, x_{i-1} \) successively shift by one element. The argument holds for all orderings.

Let \( b \) be the element before \( a_1 \) in \( \sigma \). Consider all circular orderings \( \sigma' \) such that \( b \) is followed immediately by \( a_1, \ldots, a_{k-1} \) in some order and then \( a_k \). Since \( \{b\} \cup \{a_1, \ldots, a_{k-1}\} \notin F \) but \( x \in F \), all \( k \)-element segments in \( \sigma' \) that contain \( a_k \) must lie in \( F \). Using all such \( \sigma' \), every \( k \)-set containing \( a_k \) belongs to \( F \).

The original proof pushed an EKR\((k,t)\) family toward the claimed extremal family using shift operators.

11.2.19. **Lemma.** If \( F \) is a \( k \)-uniform \( t \)-intersecting family, then also \( \tau_{i,j}(F) \) is \( t \)-intersecting, where \( \tau_{i,j} \) is the shift operator of Definition 11.2.9.
Proof: View \(x, y \in F\) as binary vectors, and let \(x' = \tau_{i,j}(x)\) and \(y' = \tau_{i,j}(y)\). By symmetry, suppose \(i < j\). If \(x\) or \(y\) has 11 or 00 in these positions, or if they both have 10, then \(|x' \cap y'| = |x \cap y|\). If they have 10 and 01, then \(x' \cap y' \supseteq x \cap y\).

This leaves the case where \(x\) and \(y\) both have 01 in these positions. If exactly one of \(x'\) and \(y'\) changes to 10, then possibly \(|x' \cap y'| < t = |x \cap y|\). If \(y' = y\) (by symmetry), then \(z \in F\), where \(z = y - j + i\). Since \(F\) is \(t\)-intersecting, \(x\) and \(z\) have at least \(t\) common elements outside \(\{i, j\}\). Hence this holds also for \(x'\) and \(y'\), since \(z\) and \(y'\) agree outside \(\{i, j\}\).

\[\square\]

11.2.20. Theorem. (Erdős–Ko–Rado Theorem \(t = 1\), revisited) For \(n \geq 2k\), the maximum size of an \(EKR(k, 1)\)-family is \(\binom{n}{k-1}\).

Proof: As in Katona’s proof, we may let \(F\) be a \(k\)-uniform intersecting family. We use induction on \(n + k\). When \(n = 2k\), in \(F\) there is at most one member of each complementary pair, and choosing one from each such pair yields an intersecting family. Since \(\binom{2k-1}{k-1} = \frac{1}{2}\binom{n}{k}\), the bound holds.

Now suppose \(n > 2k\). Since the shift operator \(\tau_{i,j}\) does not change the size of a family, by Lemma 11.2.19 we may assume that \(F\) is unchanged under all \(\tau_{i,j}\) with \(i < j\). Partition \(F\) into subfamilies \(F_1\) and \(F_0\), with \(F_1\) consisting of those members containing element \(n\). Inductively, \(|F_0| \leq \binom{n-2}{k-1}\). Trivially \(|F_1| \leq \binom{n}{k-1}\); we obtain the desired bound if we can improve this bound to \(\binom{n}{k-3}\).

Let \(F' = \{A - n : A \in F_1\}\). By the induction hypothesis, it suffices to show that \(F'\) is an intersecting family. Otherwise, we have \(A, B \in F\) with \(A \cap B = \{n\}\). Since \(n > 2k\), their union omits some other element \(i\). Since \(F\) is invariant under \(\tau_{i,n}\), we have \(A - n + i \in F\). Now \(B\) does not intersect \(A - n + i\), a contradiction. \(\square\)

Frankl [1987] proved via shift operators that \(EKR(k, t)\)-families have size at most \(\binom{n-1}{k-1}\) when \(n\) is sufficiently large. Forbidding \(t\)-stars costs a lot; an \(EKR(k, 1)\)-family with no common element has size at most \(\binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1\) (Hilton–Milner [1967]; see Exercise 14). Bollobás [1986] considered larger \(t\).

11.2.21. Theorem. (Bollobás [1986]) For \(2 \leq t < k\), let \(F\) be a \(t\)-intersecting \(k\)-uniform family of subsets of \([n]\) that is not a \(t\)-star. If \(n > 2k - t\), then

\[|F| \leq k \binom{n-t-1}{k-t-1} + \sum_{j=1}^{t} \binom{t}{j} \binom{k-t}{j} \binom{n-t+j}{k-t+j}.
\]

Proof: Let \(F\) be a maximal such family. Since \(n > 2k - t\), we have \(F \neq \binom{[n]}{k}\). If \(|A \cap B| > t\) for all \(A, B \in F\), then adding a \(k\)-set that differs from a member of \(F\) by one element enlarges the family. Hence there exist \(A, B \in F\) with \(|A \cap B| = t\); let \(T = A \cap B\). Since \(F\) is not a \(t\)-star, we find \(C \in F\) with \(T \not\subseteq C\).

For \(0 \leq j \leq t\), let \(F_j\) be the subfamily of members of \(F\) omitting \(j\) elements of \(T\). For \(j \geq 1\), a member of \(F_j\) must have at least \(j\) elements of \(A\) and \(j\) elements of \(B\) outside \(T\). After picking the \(j\) elements of \(T\) to omit, we thus have

\[|F_j| \leq \binom{t}{j} \binom{k-t}{j} \binom{n-t+j}{k-t+j}.
\]

For \(j = 0\), a member of \(F_0\) contains all of \(T\) and needs at least one element of \(C\) outside \(T\), so \(|F_0| \leq k \binom{n-t-1}{k-t-1}\). Since \(|F| = \sum_{j=0}^{t} |F_j|\), the bound is complete. \(\square\)
11.2.22.* Theorem. For $2 \leq t < k$, let $F$ be a $t$-intersecting $k$-uniform family of subsets of $[n]$. If $n \geq 2tk^3$, then $|F| \leq \binom{n}{k-t}$, with equality only for a $t$-star.

Proof: The claim is trivial for $t$-stars. By Theorem 11.2.21, it suffices to show

$$\sum_{j=2}^{t} \binom{t}{j} \binom{t-j}{j-k} \left( \binom{n-t}{k-t} \right)^2 < \binom{n-t}{k-t} - [k + t(k-t)^2] \binom{n-t}{k-t}.$$ 

With $S_j = \binom{t}{j} \binom{t-j}{j-k} \binom{n-t}{k-t}^2$, we have $S_{j+1} / S_j = \frac{(t-j)^2}{(k-t)} \binom{n-t}{k-t} < (j+1)^3$. Since $\sum_{j=2}^{\infty} (j+1)^3 < 1$, we have $\sum_{j=2}^{t} S_j < 2S_2$, so

$$\sum_{j=2}^{t} S_j < t^2(k-t)^t \left( \frac{n-t-2}{k-t} \right)^2 < \frac{t^2(k-t)^5}{n-t} \binom{n-t}{k-t}.$$ 

Since $\frac{t^2(k-t)^5}{n-t} < \frac{1}{4}$ when $n > 2tk^3$, we obtain $\sum_{j=2}^{t} S_j < \frac{1}{4} \frac{n-t}{k-t} (n-t-1)$, and then

$$\sum_{j=2}^{t} S_j < \frac{1}{4} \frac{n-t}{k-t} (n-t-1)$$

finishes the proof.

At first, “sufficiently large” meant $n \geq t + (k-t)(k^3)$; the threshold in Theorem 11.2.22 is better. Frankl [1976] and Wilson [1984] together showed that the $t$-star of $k$-sets is optimal when $n \geq (t+1)(k-t+1)$, extending Theorem 11.2.18. For smaller $n$, other EKR$(k, t)$-families are larger, such as $\binom{[n]}{\lfloor (n+1)/2 \rfloor}$ when $t = 1$ and $n < 2k$. Ahlswede–Khachatrian [1997] determined the extreme in all cases (see Exercise 24), as conjectured by Frankl [1978].

The Erdős–Ko–Rado Theorem for $t = 1$ follows easily from the Kruskal–Katona Theorem, as shown by Daykin [1974]). The earlier result below uses the same ideas and is stronger (simply set $F = G$).

11.2.23.* Theorem. (Kleitman [1968]) Let $k, l, n$ be positive integers such that $k + l \leq n$. If $F \subset \binom{[n]}{k}$ and $G \subset \binom{[n]}{l}$ satisfy $x \cap y \neq \varnothing$ whenever $x \in F$ and $y \in G$, then $|F| \leq \binom{n}{k-l}$ or $|G| \leq \binom{n}{l-k}$.

Proof: Let $F' \subset \binom{[n]}{n-k}$ consist of the complements of members of $F$. Note that $n-k \geq l$, since $k+l \leq n$. The family $G \cup F'$ is an antichain. Thus the $l$-shadow of $F'$ is disjoint from $G$. By the Kruskal–Katona Theorem, the size of the $l$-shadow of $F'$ is at least $\partial_{l+1} \partial_{l+2} \cdots \partial_{n-k}(F')$. Thus $|G| + \partial_{l+1} \partial_{l+2} \cdots \partial_{n-k}(F') \leq \binom{n}{l}$.

If $|F| \geq \binom{n}{n-k}$, then also $|F'| \geq \binom{n}{n-k}$, and the first term in the $(n-k)$-binomial expansion of $|F'|$ is $\binom{n}{n-k}$. Regardless of what else is present, iteration of $\partial$ yields $\binom{n}{n-k}$ as a lower bound on the second term in the inequality above. Thus $|G| \leq \binom{n}{n-k} = \binom{n}{l-k}$.

CHVÁTAL’S CONJECTURE

Every star is an intersecting family. The Erdős–Ko–Rado Theorem implies that in $2^{[n]}$ with $n \geq 2k$, the star $F$ consisting of all sets with size at most $k$ containing 1 is a largest intersecting family of sets with size at most $k$; it simultaneously achieves equality in the EKR bound for sets of each size. Chvátal conjectured a tantalizing generalization.

11.2.24. Definition. A family $I$ is an ideal if $x \in I$ and $y \subseteq x$ imply $y \in I$. The maximal members of an ideal are its bases. A family has the star property if some largest intersecting subfamily is a star.

11.2.25. Conjecture. (Chvátal [1974]) Every ideal of subsets of $[n]$ has the star property.

The term “ideal” is traditional for this topic and the rest of this chapter. Since “ideal” is used in other ways in other areas of mathematics, “down-set” has grown in popularity as an alternative term without alternative meanings.

As we observed before Definition 11.2.24, the Erdős–Ko–Rado Theorem implies the conjecture for the ideal consisting of all subsets of $[n]$ with size at most $k$, where $k \leq n/2$. Example 11.2.16 proves it for the family of all subsets of $[n]$. Other partial results appear in Chvátal [1974], Kleitman–Magnanti [1974], Berge [1975], and Wang–Wang [1978]; they are surveyed in Kleitman [1979].

We show first that an intersecting family cannot contain more than half of an ideal, because ideals partition into pairs of disjoint subsets. The proof removes some disjoint pairs from the “top” of the ideal and applies induction to the rest.

11.2.26. Theorem. (Berge [1976]) If $I$ is an ideal of subsets of $[n]$, then $I$ or $I - \{\emptyset\}$ can be partitioned into pairs of disjoint sets.

Proof: (Daykin–Hilton–Miklós [1983], simplified by M. Pelsmajer) We use induction on $|I|$, with a trivial basis for size 0 or 1. For $|I| > 1$, choose a maximal set $T' \subseteq [n]$ such that $T'$ is a union of disjoint sets $A, B \in I$. Such a set exists, since the empty set qualifies.

Let $H = \{A \subseteq T : A, T - A \in I\}$. The choice of $T$ yields $H \neq \emptyset$. Observe that if $A \in H$ and $A \subseteq C \in I$, then $T - C \subseteq T - A \in I$. Also, if $C \not\subseteq T$, then $C \cup (T - C)$ is larger than $T$, contradicting the choice of $T$. Hence $C \in H$, by the definition of $H$. We conclude that $I - H$ is an ideal.

By construction, $H$ is a subfamily of $I$ partitioned into pairs of disjoint sets (complements within $T$). Since $I - H$ is an ideal, we complete the desired partition by applying the induction hypothesis to $I - H$. ■
By Theorem 11.2.26, Chvátal’s Conjecture holds under any conditions that guarantee a star with half of $I$, such as when some element belongs to all bases (Exercise 30). Chvátal [1974] proved the conjecture in the case where there is a fixed linear order $L$ on $[n]$ such that $\{a_1, \ldots, a_k\} \in I$ whenever $\{b_1, \ldots, b_k\} \in I$ and $a_i \leq b_i$ for $1 \leq i \leq k$. In this case, the star generated by the initial element in $L$ is a largest intersecting family. It need not contain half of $I$, as shown when $I$ is generated by singleton sets.

Several survey papers (including one by this author) mistakenly stated that what Chvátal [1974] proved was the stronger result by Snevily that we prove next. Snevily’s result is a common generalization of the results of Chvátal and Schönheim. Note the use of $+$ and $-$ for addition and deletion of single elements.

11.2.27. Theorem. (Snevily [1991]) If $I$ is an ideal of subsets of $[n]$, and $x$ is an element of $[n]$ such that $A - a + x \in I$ whenever $a \in A \in I$, then the star generated by $x$ is a largest intersecting family in $I$.

Proof: Consider a minimal ideal $I$ violating the claim; note that $|I| > 1$. The figure below shows various parts of $I$ that we define. Let $B$ be the set of bases of $I$ not containing $x$. Let $F$ be a largest intersecting family in $I$.

Let a free base be a base of $I$ that is not expressible as $B - a + x$ with $a \in B \in B$. If $C$ is a free base, then let $I^* = I - \{C\}$; we claim that $I^*$ is a smaller ideal satisfying the hypothesis. This holds because if $C$ is needed as $A - a + x$ for some $A$, then $A$ is below some $B \in B$, and then $C$ is below $B - a + x$, which is in $I$. This contradicts that $C$ is a base.

By the induction hypothesis, the star $F^*$ generated by $x$ is a maximum intersecting family in $I^*$. Since $F^* \subseteq I$, the minimality of $I$ yields $|F| > |F^*|$. Since $F - \{C\} \subseteq I^*$, we obtain $C \in F$ and $|F| = |F^*| + 1$.

If $x \in C$, then $I$ satisfies the claim, since $F^* \cup \{C\}$ is a star with center $x$ in $I$ and has size $|F|$. Hence we may assume that every free base omits $x$ and belongs to $F$. Thus $B$ is the set of free bases, and $B \subseteq F$.

Now let $I'$ be the ideal generated by $B$. Every member of $I - I'$ lies below a base of the form $B - a + x$ for some $B \in B$ and must contain $x$; otherwise it would lie below $B$ and belong to $I'$.

Let $F' = F \cap I'$. Every member of $F'$ omits $x$, since it lies in $I'$. To show that the star generated by $x$ has size at least $|F|$, we will replace $F'$ in $F$ with $|F'|$ sets outside $F$ that contain $x$. We do this in two steps: first we map the members of $F'$ to their mates in the pairing of $I'$ guaranteed by Theorem 11.2.26, and then we add $x$ to each such mate.

Let $H'$ be the set of mates of members of $F'$, and let $H = \{Y \cup \{x\}: Y \in H'\}$
(it does not matter whether $H'$ is an ideal). Since $F'$ is pairwise intersecting and $Y \in H'$ is disjoint from its mate, $H' \cap F' = \emptyset$. Furthermore, $Y \cup \{x\}$ also is disjoint from the mate of $Y$ in $F'$, so $Y \cup \{x\} \notin F - F'$.

It remains only to show $Y \cup \{x\} \in I$ to obtain $(F - F') \cup H$ as a star generated by $x$ with size $|F|$. Since $B \subseteq F'$ and $F' \cap H' = \emptyset$, we have $Y \notin B$. Thus $Y$ is non-maximal in $I'$ and has the form $A - a$ for some $A \in I' \subset I$. The hypothesis then yields $Y \cup \{x\} \in I$.

Subsequent decades have not seen much work on Chvátal’s Conjecture. We mention Wang [2002], Borg [2011] (a weighted version), Kamat [2011], and Friedgut–Kahn–Kalai–Keller [2018].

In addition to considering families closed under taking subsets, we can also consider families closed under taking unions. This leads to a conjecture at least as notorious as that of Chvátal. A family of sets is union-closed if the union of every two members of the family also belongs to the family.

11.2.28. **Conjecture.** *(Union-Closed Sets Conjecture)* If $F$ is a finite union-closed family of finite sets with $|F| \geq 2$, then some element appears in at least half the members of $F$.

The conjecture is generally attributed to Frankl in 1979 and first appeared in print in 1985. Early special cases were proved by Sarvate–Renaud [1989] (Exercise 32), Poonen [1992], and Johnson–Vaughan [1998]. Bruhn–Schaudt [2015] presents a thorough survey. The fundamental nature of the conjecture is shown by reformulations such as a version in terms of bipartite graphs in Exercise 33.

**SUNFLOWERS**

In the problems we have considered thus far, our notion of “star” was a family of sets having a common element. Now we consider a more restricted notion.

11.2.29. **Definition.** A **Δ-system** or sunflower is a family of sets in which any two members have the same intersection.

In a 2-uniform family (a graph), the sunflower condition reduces to a star or a family of disjoint edges, where the common intersection and pairwise intersection is always the empty set. The original term is Δ-system, but “Δ” does not evoke the meaning in any way, so we will use the term “sunflower”. An early use of this term is in Alon–Lovász [1986].

11.2.30. **Theorem.** *(Sunflower Theorem; Erdős–Rado [1960])* Every $k$-uniform family with more than $k!(s - 1)^k$ members contains a sunflower of size $s$, and there is a family of size $(s - 1)^k$ having no such sunflower.

**Proof:** We use induction on $k$. For $k = 1$, the sets are disjoint, and we have more than $s - 1$ of them.

For $k \geq 2$, let $F'$ be the $k$-uniform family, and let $H$ be a maximal subfamily consisting of disjoint sets. If $|H| \geq s$, then $H$ is the desired sunflower. Otherwise,
let $B$ be the union of the sets in $H$; note that $|B| \leq k(s - 1)$. Every member of $F$ intersects $B$. By the pigeonhole principle, some element of $B$ lies in at least $|F|/|B|$ members of $F$. We compute

$$\frac{|F|}{|B|} > \frac{k!(s - 1)^k}{k(s - 1)} = (k - 1)!(s - 1)^{k-1}.$$ 

Deleting $x$ from the sets containing it yields a $(k - 1)$-uniform family $F'$ with more than $(k - 1)!(s - 1)^{k-1}$ members. By the induction hypothesis, $F'$ contains a sunflower of size $s$. Returning $x$ to its sets yields a sunflower in $F$ of size $s$.

For the construction, let $X_1, \ldots, X_k$ be disjoint sets of size $s - 1$. Let $F$ be the family of all transversals ($k$-sets consisting of one element from each of $X_1, \ldots, X_k$). If $F$ contains a sunflower $A_1, \ldots, A_s$, then some element $x$ belongs to exactly one of $A_1, \ldots, A_s$. By symmetry, we may assume $x \in A_1 \cap X_1$. Since the pairwise intersections are the same, the sets $A_1, \ldots, A_s$ must intersect $X_1$ in distinct elements. Since $|X_1| = s - 1$, this is impossible. 

It is not known whether the bound $k!(s - 1)^k$ in Theorem 11.2.30 is sharp.

11.2.31. **Conjecture.** (Erdős–Rado [1960]) For $k \in \mathbb{N}$ there is a constant $c_k$ such that every $k$-uniform family $F$ with $|F| > c_k^k$ contains a sunflower of size $s$.


11.2.32. **Conjecture.** (Erdős–Szemerédi [1978]) There is a constant $c$, less than 2, such that every family of subsets of $[n]$ containing no sunflower of size 3 has at most $c^n$ sets.

By results of Erdős–Szemerédi [1978], Conjecture 11.2.31 implies Conjecture 11.2.32. Using algebraic methods like those in Chapter 15, Naslund–Sawin [2016] proved Conjecture 11.2.32 using a constant $c$ that is just above $3/2^{2/3}$.

The most famous applications of the Sunflower Theorem (and various modifications) are to complexity theory; see Razborov [1985] and Alon–Boppana [1987]. More recently, there is an application to matrix multiplication (Alon–Shpilka–Umans [2013]). Alon–Pudlak [2001] used the method to prove an explicit lower bound for off-diagonal Ramsey numbers: There is a fixed constant $c$ such that for every fixed $s$ and sufficiently large $m$, the construction produces a graph having neither an $s$-clique nor an independent $m$-set. Thus $R(s, m)$ exceeds the order of the constructed graph, which is $m^c \sqrt{\log s / \log \log s}$.

**EXERCISES 11.2**

11.2.1. (−) Fix $m, k \in \mathbb{N}$. When $n$ is sufficiently large, how can a family of $m$ sets of size $k$ be chosen in $[n]$ to maximize the size of the shadow of the family?

11.2.2. (−) Permutations (in word form) intersect if they agree in some position. Find the maximum size of an intersecting family of permutations of $[n]$. (Deza–Frankl [1977])
11.2.3. \((-\)
Does every maximal intersecting family in \(2^{[n]}\) have size \(2^{n-1}\)?

11.2.4. \((-\)
Let \(I\) be an ideal of subsets of \([n]\). Prove that every element of \([n]\) appears in
at most half the members of \(I\).

11.2.5. \((-\)
Use the Sunflower Theorem to prove that every graph with more than \(2(k-1)^2\)
edges contains a matching of size \(k\) or a star of size \(k\).

11.2.6. Determine the maximum size of a family of subsets of \([n]\) such that no element
belongs to at least two sets in the family and is omitted by at least two sets in the family.

11.2.7. Use the Kruskal–Katona Theorem to prove that the size of the shade of a family
of \(m\) elements of \({n \choose k}\) is minimized by the family consisting of the first \(m\) elements of the
order on \({n \choose k}\) in which \(x < y\) if \(x_i > y_i\) in the highest coordinate where they differ.

11.2.8. \((\odot)\) Let \(G\) be an \(n\)-vertex graph with \(m\) edges. For each \(k \in \mathbb{N}\), determine the
maximum possible number of \(k\)-cliques in \(G\), and construct an example achieving equality.
(Bollobás [1981])

11.2.9. Explain how to modify the shift-operator proof of the Kruskal–Katona Theorem
(Theorem 11.2.11) to give a direct proof of Theorem 11.2.12 that the shadow of a \(k\)-uniform
family with size \(\binom{n}{k}\) has size at least \(\binom{n}{k-1}\). (Frankl [1984])

11.2.10. Let \(F\) be a \(k\)-uniform family of size \(\binom{n}{k}\). Refine the proof of Theorem 11.2.12 to
show that the shadow of \(F\) has size exactly \(\binom{n}{k-1}\) if and only if \(u\) is an integer and \(F\) consists
of all \(k\)-sets in a set of size \(u\). (Lovász [1979])

11.2.11. Prove that the list \(0,\ldots,0,f_k,f_{k+1},\ldots,f_{l},0,\ldots,0\) is realizable as \(f_i = |F \cap \binom{[n]}{i}|\)
for some antichain \(F\) of subsets of \([n]\) if and only if
\[
f_k + \partial_{k+1}(f_{k+1}) + \cdots + \partial_{l-1}(f_{l-1}) + \partial_l(f_l)) \leq \binom{n}{k}.
\]
(Clements [1973], Daykin–Goddrey–Hilton [1974])

11.2.12. Fix \(r \geq 2\). For what values of \(n\) is it possible to color every square in an \(n\)-by-\(n\)
grid with one of \(r\) colors so that, for all \(i, j, k\) between 1 and \(n\) with \(i \neq j\) and \(j \neq k\), the
square in row \(i\) and column \(j\) is assigned a different color from the square in row \(j\)
and column \(k\). (Hint: Use Sperner’s Theorem.) (Prop [1998])

11.2.13. Define a rising antichain to be an antichain of subsets of \([n]\) whose sizes are
distinct. Let two rising antichains be equivalent if one is obtained from the other by comple-
menting all sets and/or relabeling the elements of \([n]\).

(a) Determine the maximum size of a rising antichain of subsets of \([n]\).
(b) Prove that when \(n \geq 5\), the largest rising antichains are all equivalent to each other.
(Bey–Griggs [2002])

11.2.14. For \(n > 2k\), construct an intersecting family of \(k\)-subsets of \([n]\) with size \(\binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1\) that is not a star. (Comment:
Hilton–Milner [1967] proved that none is larger.)

11.2.15. For \(n \geq t \geq 1\), let \(q = [(n + t)/2]\). Prove that the size of the largest \(t\)-intersecting
family in \(2^n\) is \(\sum_{i=q}^{n} \binom{n}{i}\) when \(n + t\) is even and \(\sum_{i=q}^{n} \binom{n}{i} + \binom{n-1}{q-1}\) when \(n + t\) is odd.

11.2.16. Let \(F\) be a family of 4-sets of \([n]\) that pairwise intersect at most twice. Prove the
existence of \(S \subseteq [n]\) with \(|S| \geq \lfloor (6n-6)^{1/3} \rfloor\) that contains no member of \(F\). (Adrian [1991])

11.2.17. Let \(F\) be an intersecting family of subsets of \([n]\) whose members have size at most \(k\).
Prove that \(|F| \leq \sum_{i=1}^{k} \binom{n-1}{i-1}\), with equality only for stars if \(k < n-1\).
11.2.18. (♦) Prove that the largest antichain of subsets of $[n]$ consisting of pairs of complementary sets has size $2\binom{n-1}{\lfloor k/2\rfloor}$. (Bollobás [1973])

11.2.19. Let $\pi$ and $\sigma$ be simple $k$-words from $[n]$ (lists of $k$ distinct elements). Say that $\pi$ and $\sigma$ 
intersect if $\pi_i = \sigma_i$ for some $i$. Prove that the maximum size of a pairwise intersecting family of simple $k$-words 
from $[n]$ is $(n-1)/((n-k)!$. (see Deza–Frankl [1978], Lovász–Nešetřil–Pultr [1980])

11.2.20. Let $F$ be a family of subsets of $[n]$ such that each member has size at least $l$ and each pair of members have at most $k$ common elements, where $l \leq k$. Prove that $|F|$ is
maximized uniquely by $F = \{A \subseteq [n]: l \leq |A| \leq k+1\}$.

11.2.21. Intersecting families and words in trees.
(a) Let $H$ be a non-2-colorable hypergraph whose edges are pairwise intersecting. Prove that the minimum size of a set intersecting all edges in $H$ equals the minimum size of an edge in $H$.
(b) Let $T$ be a complete ternary tree with $k$ levels ($3^k$ leaves). Prove that for any 2-coloring of the leaves of $T$, there is a complete binary subtree ($2^t$ leaves) whose leaves all have
the same color.
(c) For an edge-labeled rooted tree, the associated set is the set of all strings formed by listing the labels along a path from the root to a leaf. Let $T$ be a complete ternary tree with each edge labeled 0 or 1. Prove that the minimum size of the associated set for a complete binary subtree of $T$ is the minimum size of a set containing a word in the associated set of every complete binary subtree. (Milans)

11.2.22. Let $A$ be an intersecting antichain of subsets of $[n]$ with $|x| \leq n/2$ for all $x \in A$. Prove that 
$\sum_{x \in A} (|x|^{-1})^2 \leq 1$. (Hint: Assign weight $|x|^{-1}$ to each $x \in A$ and prove that the 
total weight on members of $A$ appearing as consecutive strings in a single cyclic permutation is at most 1.) (Bollobás [1973], Greene–Katona–Kleitman [1976])

11.2.23. Let $X_1, \ldots, X_n$ be independent 0,1-random variables with probability $p \geq 0.5$ of having value 1. Let $Z$ be a convex combination of $\{X_i\}$; that is, $Z = \sum \alpha_i X_i$, with $\alpha_i \geq 0$ and $\sum \alpha_i = 1$. Prove that $P(Z \geq 0.5) \geq p$ when no subset of $\{\alpha_i\}$ sums to .5. (Hint: Use the
EKR Theorem to bound the number of $k$-subsets of $\{\alpha_i\}$ whose sum exceeds .5. Comment:
The claim holds without the restriction that no subset of $\{\alpha_i\}$ sums to .5, but eliminating that restriction requires a limit argument.) (Liggett [1977])

11.2.24. For $0 \leq r \leq k-t$, let $S_r$ be the family of $k$-subsets of $[n]$ that contain at least $t+r$ of the lowest $t+2r$ elements.
(a) Show that $S_r$ is a $t$-intersecting family.
(b) Prove that $|S_r| = \sum_{i=0}^{r} \binom{n-t-1}{r-i} \binom{n-t-i}{r-i}$.
(c) Prove that $|S_1| > |S_0|$ if and only if $n < (k-t+1)(t+1)$. Thus the $t$-star $S_0$ is not the largest
EKR($k$, $t$)-family when $n < (k-t+1)(t+1)$. (Frankl [1978]) (Comment: Ahlswede–Khachatryan [1997] proved that $S_1$ is the largest EKR($k$, $t$)-family when $(k-t+1)(t+1) \leq n < (k-t+1)(t+2)$. A proof appears in Engel [1997, pp. 50–60].)

11.2.25. Let $F$ and $G$ be ideals in $2 \times 3$. Show that it is not always possible to minimize $|F \cap G|$ by letting $F$ be the first $|F|$ vectors in lexicographic order and letting $G$ be the
lexicographically first $|G|$ vectors when the components are read in the opposite order.
(Daykin–Kleitman–West).

11.2.26. Find the largest $k$ such that every $k$-coloring of the subsets of $[n]$ makes all of
$\{A, B, A \cup B, A \cap B\}$ the same color for some distinct sets $A$ and $B$. Solve the same problem
for $n = 6$ when $A$ and $B$ must be incomparable. (Tomescu [1987])
11.2.27. Let $I$ be an ideal of subsets of $[n]$, and let $I' = \{A: A \in I\}$. Prove that there exists a bijection $f: I \to I'$ such that $A \subseteq f(A)$ for all $A \in I$. (Erdős–Herzog–Schönheim [1970], Marica [1971], Daykin–Hilton–Miklós [1983])

11.2.28. Use Berge’s Theorem (Theorem 11.2.26) to prove the following:
   (a) The average size of the sets in any ideal of subsets of $[n]$ is at most $n/2$.
   (b) The maximum size of an intersecting family of subsets of $[n]$ whose complements also form an intersecting family is $2^{n-2}$.

11.2.29. Let $F$ be an intersecting family of subsets of $[n]$ whose complements also form an intersecting family. Use the shift operators $\tau_{1,f}$ and Theorem 11.2.27 to prove $|F| \leq 2^{n-2}$.

11.2.30. **Special cases of Chvátal’s Conjecture.** (Schönheim [1976])
   (a) Prove the conjecture for ideals whose bases have a common element.
   (b) Prove the conjecture for ideals with two bases.
   (c) Use Theorem 11.2.27 to give a proof without the Erdős–Ko–Rado Theorem for the ideal consisting of all subsets of $[n]$ with size at most $k$.

11.2.31. For a set $B \subseteq [n]$ and a family $F$ of subsets of $[n]$, the **translate** of $F$ by $B$, written $F(B)$, is $\{A \Delta B: A \in F\}$, where $\Delta$ denotes symmetric difference.
   (a) Prove that the translate of an ideal by $(x)$ has the star property. (B. Reiniger)
   (b) Prove that the translate of an ideal $I$ by a set $B$ is an ideal if and only if $B$ is contained in every maximal element of $I$. (P. Wenger)
   (c) Let $I$ be an ideal. Prove that every translate of $I$ by a nonempty set $B$ contains a star at least as large as the largest intersecting family in $I$. (Snevily)

11.2.32. Prove that the Union-Closed Sets Conjecture (Conjecture 11.2.28) is true for families having a member with size at most 2. (Sarvate–Renaud [1989])

11.2.33. (⋄) Prove each statement below equivalent to the Union-Closed Sets Conjecture.
   (a) If a finite family $F$ with $|F| \geq 2$ is closed under taking intersections, then $\bigcup_{A \in F} A$ has an element belonging to at most half of the members of $F$.
   (b) Every nontrivial $X, Y$-bigraph $G$ contains in both $X$ and $Y$ a vertex that lies in at most half of the maximal independent sets of $G$. (Bruhn–Charbit–Telle [2013])

11.2.34. For $n - s + 1 < k \leq n$, prove that the family $\binom{[n]}{k}$ has no sunflower of size $s$.

11.2.35. Given a graph $G$, let $H$ be the graph with vertex set $\binom{V(G)}{k}$ such that vertices $A$ and $B$ are adjacent if and only if $G$ has an edge $uv$ with $u \in A - B$ and $v \in B - A$. Suppose that every induced subgraph of $G$ having at most $sk$ vertices has fewer than $\binom{s}{k}$ edges. Use the Sunflower Theorem to prove $\omega(G) \leq k!(s - 1)^k$.

### 11.3. Matroids

The defining properties of matroids are general enough to occur in many contexts and special enough to yield rich combinatorial structure. Many elegant results from graph theory, linear algebra, and elsewhere generalize in the theory of matroids. These include the greedy algorithm for minimum spanning trees, min-max relations for systems of distinct representatives, dimension properties of vector spaces, and duality properties of planar graphs. When a theorem on a special class holds for all matroids, it immediately yields results for other classes.
Matroids were independently introduced by Whitney [1935] (to study planar graphs) and by Nakasawa [1935] (to study linear dependence; reprinted and translated in Nishimura–Kuroda [2009]). They were reinvented by van der Waerden [1937] generalizing independence in vector spaces, and they also arose (from a structural viewpoint) in the theory of geometric lattices (MacLane [1936]). The first modern textbook on the subject was Welsh [1976], later Oxley [1992, 2011].

In this brief treatment, we focus on the application of matroids to extremal problems, leaving deeper structural aspects and more subtle matroidal aspects of graphs to a more advanced book. We start from the notions of ideals and antichains defined in the preceding section.

In many mathematical contexts, sets that avoid conflict are called “independent”. Inherently, subsets of independent sets (and the empty set) are independent. Thus the family of independent sets is closed under taking subsets. The same family can be specified by its antichain of maximal independent sets, its antichain of minimal non-independent sets, or other aspects we will describe later.

To obtain the elegant behavior of matroids, we need an additional property. A restriction on the family of independent sets can be translated into a corresponding restriction on some other specification of the system. Because we can specify the system in many ways, we have many equivalent definitions of matroids.

HEREDITARY SYSTEMS AND EXAMPLES

Before discussing the added properties that yield matroids, we begin with the more general notion of hereditary systems, familiar from our study of ideals. Given a finite set $E$ of elements, we use $2^E$ to denote the family of subsets of $E$, ordered by inclusion; it has size $2^{|E|}$. We view a nonempty ideal in $2^E$ and the other ways of specifying it as a single structure. A matroid will then be such a structure that satisfies one of various equivalent additional constraints.

11.3.1. Definition. A hereditary system $M$ on $E$ consists of a nonempty ideal $I_M$ in $2^E$ and all ways of specifying $I_M$. The ideal $I_M$ is the family of independent sets of $M$; the other subsets of $E$ are dependent. The bases $B_M$ are the maximal independent sets. The circuits $C_M$ are the minimal dependent sets. The rank function $r_M(X)$ of $M$ assigns to each $X \subseteq E$ a value called its rank, which is the maximum size of a member of $I_M$ contained in $X$.

11.3.2. Example. On the left below we sketch the inclusion order on subsets of $E$, with the full set $E$ at the top and the empty set at the bottom. We sketch the relationships among the independent sets, bases, circuits, and dependent sets of a hereditary system. The bases are the maximal elements of $I$, and the circuits are the minimal elements not in $I$. Always $\varnothing \in I$. If every set is independent, then there is no circuit, but every hereditary system has at least one base.

On the right we obtain a hereditary system from a multigraph with three edges. The independent sets are the acyclic edge sets. The only dependent sets are $\{1, 2\}$ and $\{1, 2, 3\}$, the only circuit is $\{1, 2\}$, and the bases are $\{1, 3\}$ and $\{2, 3\}$. The rank of an independent set is its size. For the dependent sets, we have $r(\{1, 2\}) = 1$ and $r(\{1, 2, 3\}) = 2$. ■
11.3.3. Remark. Aspects of hereditary systems. An aspect of a hereditary system $M$ is a way of specifying it. For example, $M$ can be specified by any of $I_M$, $B_M$, $C_M$, or $r_M$, because each of these determines the others. We have expressed $B_M$, $C_M$, $r_M$ in terms of $I_M$. Conversely, if $B_M$ is given, then $I_M$ consists of the sets contained in members of $B_M$. If $C_M$ is given, then $I_M$ consists of the sets containing no member of $C_M$. If $r_M$ is given, then $I_M$ and the other aspects are found by setting $I_M = \{ X \subseteq E : r_M(X) = |X| \}$. This justifies our view of a hereditary system as a unified structure $(I, B, C, r, \cdots)$. We drop the subscripts on $I, B, C, r, \cdots$ when discussing only one hereditary system. Later we introduce additional aspects that can specify a hereditary system.

Most terminology in matroid theory comes from the motivating contexts that led to the discovery of matroids, particularly graphs and linear algebra. We begin with the fundamental example from graphs.

11.3.4. Definition. The cycle matroid $M(G)$ of a multigraph $G$ is the hereditary system on $E(G)$ whose circuits are the edge sets of cycles of $G$. A hereditary system that can be specified in this way is a graphic matroid.

11.3.5. Example. Bases in a cycle matroid $M(G)$. The graph $K^-_4$, with four vertices and five edges, arises by deleting one edge from $K_4$. Spanning trees in $K^-_4$ have three edges, so every set with more than three edges is dependent. Also the two triangles are dependent. This yields eight dependent sets and 24 independent sets among the subsets of $E(K^-_4)$. There are three minimal dependent sets (the cycles) and eight maximal independent sets (the spanning trees).

For a general multigraph $G$, the bases of the cycle matroid $M(G)$ are the edge sets of the maximal forests in $G$. Each contains a spanning tree of each component of $G$, so they have equal size. Consider $B_1, B_2 \in B$ and $e \in B_1 - B_2$. Deleting $e$ from $B_1$ disconnects a component; since $B_2$ contains a tree spanning the same component of $G$, some edge $f \in B_2 - B_1$ can be added to $B_1 - e$ to reconnect it.

The base exchange property is

If $B_1, B_2 \in B_M$, then for all $e \in B_1 - B_2$ there exists $f \in B_2 - B_1$ such that $B_1 - e + f \in B_M$.

Matroids are the hereditary systems satisfying the base exchange property.
**11.3.6. Remark. Notational conventions.** In this subject, we often discuss modifying a set by adding or deleting a single element. For symmetry and simplicity, we use the symbols + and − instead of ∪ and − for this, and we drop the set brackets on singleton sets.

Using bold I, B, C for families of subsets of E allows using I ∈ I, B ∈ B, C ∈ C, respectively, to denote membership. We use roman I, B, C, R to denote properties, e, f, x, y for elements of E, and X, Y, F for subsets of E.

**11.3.7. Example. Rank in cycle matroids.** Let G be an n-vertex graph. For X ⊆ E(G), let G_X be the spanning subgraph of G with edge set X. In M(G), an independent subset of X is the edge set of a forest in G_X. When G_X has k components, the maximum size of such a forest is n − k. Hence r(X) = n − k. Below we show such a forest Y in X.

If r(X + e) = r(X) for some e ∈ E − X, then the endpoints of e lie in a single component of G_X; adding e does not combine components. If we add two such edges, then again we do not combine components. Therefore, r(X) = r(X + e) = r(X + f) implies r(X) = r(X + e + f).

The (weak) absorption property (name suggested by A. Kézdy) is

If X ⊆ E and e, f ∈ E,
then r(X) = r(X + e) = r(X + f) implies r(X + e + f) = r(X).

Matroids are the hereditary systems satisfying the absorption property.

---

Multigraphs may have loops and multiedges. Analogous terminology in hereditary systems captures the behavior in cycle matroids of elements arising from loops and multiedges in multigraphs.

**11.3.8. Definition.** In a hereditary system, a loop is an element forming a circuit of size 1. Two non-loops are parallel elements if they form a circuit. A hereditary system is simple if it has no loops or parallel elements.

Another motivating context for matroids is linear independence in vector spaces. Matroids can be obtained from matrices.

**11.3.9. Definition.** The vector matroid on a finite set E of vectors in a vector space (over a field K) is the hereditary system whose independent sets are the linearly independent sets of vectors in E. A matroid expressible in this way is a linear matroid (or representable matroid) over K. The column matroid M(A) of a matrix A is the vector matroid defined on its columns.

**11.3.10. Example. Circuits in vector matroids.** Technically, one vector (or multiples of it) may be used to represent distinct elements of E, just as a matrix may
have repeated columns; these yield parallel elements. The circuits are the minimal sets \( \{x_1, \ldots, x_k\} \subseteq E \) such that \( \sum c_i x_i = 0 \) using coefficients not all zero. Minimality forces all \( c_i \neq 0 \).

Let \( C_1 \) and \( C_2 \) be distinct circuits containing \( x \). The equations of dependence for \( C_1 \) and \( C_2 \) let us write \( x \) as a linear combination of \( C_1 - x \) and of \( C_2 - x \). Equating these expressions yields an equation of dependence for \( (C_1 \cup C_2) - x \); thus \( (C_1 \cup C_2) - x \) contains a circuit.

The **weak elimination property** is

If \( C_1 \) and \( C_2 \) are distinct circuits and \( x \in C_1 \cap C_2 \), then another member of \( C_M \) is contained in \( (C_1 \cup C_2) - x \).

Matroids are the hereditary systems satisfying the weak elimination property. The column matroid of the matrix below is also the cycle matroid of \( K_4^{-} \).

\[
\begin{pmatrix}
0 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 \\
1 & 1 & 0 & 0
\end{pmatrix}
\]

The weak elimination property implies that in a matroid the relation of being parallel is transitive, but this does not hold for all hereditary systems.

Another class of matroids important for applications arose much later from families of sets. Edmonds–Fulkerson [1965] and Mirsky–Perfect [1967] independently discovered that these are matroids.

11.3.11. **Definition.** The **transversal matroid** induced by sets \( A_1, \ldots, A_m \) with union \( E \) is the hereditary system on \( E \) in which the independent sets are the systems of distinct representatives (SDR) for subsets of \( \{A_1, \ldots, A_m\} \).

11.3.12. **Remark.** **Transversal matroids and bipartite graphs.** The name “transversal” arises from using this word to mean SDR. However, in the study of optimization problems in hypergraphs, the word “transversal” is used without requiring the representatives to be distinct. We therefore use “SDR” here.

Given \( A_1, \ldots, A_m \) with union \( E \), consider the bipartite incidence graph \( G \). The parts are \( E \) and \( [m] \), with \( e \in E \) adjacent to \( i \in [m] \) if and only if \( e \in A_i \). A set \( X \subseteq E \) is independent in the transversal matroid induced by \( A_1, \ldots, A_m \) on \( E \) if and only if \( X \) is covered by some matching in \( G \).

Also, given an \( E, F \)-bigraph \( G \), we can associate with \( v_i \in F \) the set \( N_G(v_i) \). Letting \( A_i = N_G(v_i) \) expresses \( G \) as the incidence graph of the family of sets. Hence every bipartite graph induces a transversal matroid on each of its parts.

The bipartite graph below yields the transversal matroid of the family \( A = \{\{1, 2\}, \{2, 3, 4\}, \{4, 5\}\} \). This matroid is again \( M(K_4^{-}) \).

\[
\begin{align*}
E & \quad 1 & 2 & 3 & 4 & 5 \\
[m] & \quad \{1, 2\} & \{2, 3, 4\} & \{4, 5\}
\end{align*}
\]
11.3.13. Example. Independent sets in transversal matroids. The symmetric difference of matchings $M$ and $M'$ in a bipartite graph consists of alternating paths and even cycles. The only components with more edges from $M'$ than from $M$ are the $M$-augmenting paths. When $|M'| > |M|$, there must be some such path. For an $M$-augmenting path $P$, replacing $M \cap P$ with $M' \cap P$ yields a matching of size $|M| + 1$ that covers all vertices of $M$ plus the endpoints of $P$.

For independent sets $I_1$ and $I_2$ in the transversal matroid generated by $A_1, \ldots, A_m$, let $M_1$ and $M_2$ be matchings that saturate $I_1$ and $I_2$, respectively, in the associated bipartite graph (below, $M_1$ is bold and $M_2$ is solid). If $|I_2| > |I_1|$, then the matching obtained from $M_1$ via an $M_1$-augmenting path in $M_2 \triangle M_1$ covers $I_1$ plus an element $e \in I_2 - I_1$; this “augments” $I_1$.

For a hereditary system on $E$, the augmentation property is

For distinct $I_1, I_2 \in \mathbf{I}$ with $|I_2| > |I_1|$, there exists $e \in I_2 - I_1$ such that $I_1 + e \in \mathbf{I}$.

Matroids are the hereditary systems with the augmentation property.

11.3.14. Example. Augmentation in cycle matroids. Consider $I_1, I_2 \in \mathbf{I}_{M(G)}$ with $|I_2| > |I_1|$. The spanning subgraph $G_{I_1}$ has $n - |I_1|$ components. Therefore, the forest $I_2$ has some edge with endpoints in two components of $G_{I_1}$. This edge can be added to $I_1$ to form a larger forest. Hence the augmentation property holds.

11.3.15. Example. Weak elimination in cycle matroids. The circuits of $M(G)$ are the edge sets of cycles in $G$. Cycles have even degree at each vertex. For $C_1, C_2 \in \mathbf{C}$, the symmetric difference $C_1 \triangle C_2$ also has even degree at each vertex. If $C_1 \neq C_2$, then $C_1 \triangle C_2$ contains a cycle. This is stronger than weak elimination, since $C_1 \triangle C_2 \subseteq C_1 \cup C_2 - x$. In the picture below, $C_1$ and $C_2$ are face boundaries of length 9 sharing the dashed edges, and $C_1 \triangle C_2$ is the union of two disjoint cycles.
For linear matroids, direct proof of the augmentation or base exchange property uses the fact that $k$ linearly independent vectors cannot all be expressed as linear combinations of a smaller set. On the other hand, since we have verified the weak elimination property for linear matroids, this theorem of linear algebra follows from matroid axiomatics!

11.3.16. Lemma. For the rank function $r$ of a hereditary system on $E$,

(r1) $r(\emptyset) = 0$.

(r2) $r(X) \leq r(X + e) \leq r(X) + 1$ whenever $X \subseteq E$ and $e \in E$.

Proof: From the definition $r(X) = \max \{|Y| : Y \subseteq X, Y \in I\}$, we have $r(\emptyset) = 0$. Because $X + e$ contains every independent subset of $X$, we have $r(X + e) \geq r(X)$. Because the independent subsets of $X + e$ not contained in $X$ consist of $e$ together with an independent subset of $X$, we have $r(X + e) \leq r(X) + 1$.

11.3.17.* Remark. Every nonempty ideal is the family of independent sets of a hereditary system. A family $B$ is the set of bases of some hereditary system if and only if it is a nonempty antichain. A family $C$ is the set of circuits of some hereditary system if and only if it is an antichain of nonempty sets.

Characterizing rank functions of hereditary systems is more subtle. Properties (r1) and (r2) above are used to study matroids but do not suffice to make $r$ the rank function of a hereditary system. (Consider $E = \{1, 2\}$, $r(\emptyset) = r(1) = r(2) = 0$, $r(\{1, 2\}) = 1$; the set $\{1, 2\}$ contains no independent set of size 1, even though it has rank 1.) Another technical condition is needed to characterize rank functions of hereditary systems. Fortunately, we do not need this characterization, because when studying matroids we always start with a hereditary system.

AXIOMATICS OF MATROIDS

We view a matroid as a hereditary system satisfying an additional structural property. A constraint on one aspect of a hereditary system yields corresponding constraints on other aspects; thus we have many equivalent definitions. We can show that a hereditary system is a matroid by verifying any of them, and then we can use them all without additional proof.

Adding an edge to a forest creates at most one cycle. This is the property needed to show that the greedy algorithm works to find minimum spanning trees in weighted graphs. The creation of at most one circuit by adding one element to an independent set is in fact a characterization of matroids, as is the effectiveness of the greedy algorithm itself! Both appear in our list.

Given nonnegative weights on the elements, the greedy algorithm iteratively picks a heaviest element whose addition to those already selected yields an independent set. Rado [1957] (see also Boruvka [1926]) proved that matroids are precisely the hereditary systems such that for each nonnegative weight function, the greedy algorithm selects a heaviest independent set.

11.3.18. Definition. A hereditary system $M$ on $E$ is a matroid if it satisfies any of the following additional properties, where $I$, $B$, $C$, $r$ are the independent sets, bases, circuits, and rank function of $M$. 
I: **augmentation** - if $I_1, I_2 \in \mathbf{I}$ with $|I_2| > |I_1|$, then $I_1 + e \in \mathbf{I}$ for some $e \in I_2 - I_1$.

B: **base exchange** - if $B_1, B_2 \in \mathbf{B}$, then for all $e \in B_1 - B_2$ there exists $f \in B_2 - B_1$ such that $B_1 + e + f \in \mathbf{B}$.

B’: **dual base exchange** - if $B_1, B_2 \in \mathbf{B}$, then for all $e \in B_1 - B_2$ there exists $f \in B_2 - B_1$ such that $B_2 + e - f \in \mathbf{B}$.

A: **weak absorption** - $r(X) = r(X + e) = r(X + f)$ implies $r(X + e + f) = r(X)$ whenever $X \subseteq E$ and $e, f \in E$.

A’: **strong absorption** - if $X, Y \subseteq E$, and $r(X + e) = r(X)$ for all $e \in Y$, then $r(X \cup Y) = r(X)$.

U: **uniformity** - for every $X \subseteq E$, the maximal subsets of $X$ belonging to $\mathbf{I}$ have the same size.

R: **submodularity** - $r(X \cap Y) + r(X \cup Y) \leq r(X) + r(Y)$ when $X, Y \subseteq E$.

C: **weak elimination** - if $C_1, C_2 \in \mathbf{C}$ are distinct with $e \in C_1 \cap C_2$, then $(C_1 \cup C_2) - e$ contains a circuit.

J: **induced circuits** - if $I \in \mathbf{I}$, then $I + e$ contains at most one circuit.

G: **greedy algorithm** - for each nonnegative weight function on $E$, the greedy algorithm selects a maximum-weight independent set.

### 11.3.19. Remark. Submodularity

Submodularity arises naturally in linear algebra. The rank of a set $X \subseteq E$ in a vector matroid is the dimension of the space spanned by $X$. For subspaces $U$ and $V$, the result in linear algebra is $\dim U \cap V + \dim W = \dim U + \dim V$, where $W$ is the space spanned by $U \cup V$. When $X$ and $Y$ are spanning sets of vectors in $U$ and $V$, the dimension of the space spanned by $X \cap Y$ may be less than $\dim U \cap V$. Nevertheless, $\dim U \cap V + \dim W \leq \dim U + \dim V$ is proved naturally for vector spaces using the proof of $U \Rightarrow R$ below. Exercise 30 obtains submodularity directly for cycle matroids.

Various of these properties (plus being a hereditary system) have been used to define matroids. Examples include I (Welsh [1976], Schrijver [2003]), U (Edmonds [1965a,c], Bixby [1981], Nemhauser–Wolsey [1988]), A (Whitney [1935]), C (Tutte [1970]), G (Papadimitriou–Steiglitz [1982]), van der Waerden [1937], Rota [1964], Crapo–Rota [1970], and Aigner [1979] use other conditions.

Many authors list properties of hereditary systems when characterizing some aspect of a matroid. This obscures the essence of the characterization and leads to extra work. Working in the context of hereditary systems is simpler. All properties of hereditary systems are always available; we need not verify them when introducing another aspect. Our chain of implications may seem long, so it is worth noting that augmentation (I) and uniformity (U), for example, are easy to show equivalent, and hence the proof can be given in shorter implication chains.

### 11.3.20. Theorem. For a hereditary system $M$, the conditions defining matroids in Definition 11.3.18 are equivalent.

**Proof:** Property I is often used to show that some hereditary system is a matroid.

$I \Rightarrow B$. Since a smaller base could be augmented from a larger base, bases have equal size. Now consider $B_1, B_2 \in \mathbf{B}$ with $e \in B_1 - B_2$. Since $B_1 - e \in \mathbf{I}$ and $|B_2| > |B_1| - 1$, the desired $f$ exists.

$B \Rightarrow A$. Bases cannot differ in size; otherwise, repeated base exchange yields one base inside another. Now let $X' = X + e + f$ with $r(X) = r(X + e) = r(X + f)$
and \( r(X') \geq r(X) \). Among bases containing largest independent sets in \( X \) and \( X' \), choose \( B \) and \( B' \) with largest intersection. Since \( r(X + e) = r(X + f) = r(X) \),php we have \( e, f \notin B \). Since \( |B' \cap X'| > |B \cap X'| \) and \( |B| = |B'| \), there exists \( x \in B - B' \) with \( x \notin X' \). Base exchange guarantees a base \( B - x + x' \) for some \( x' \in B' - B \). Since \( x \notin X' \), we have \( |B - x + x' \cap X| = r(X) \), but \( |(B - x + x') \cap B' > |B \cap B'| \).

\[ A \Rightarrow A'. \] We use induction on \( |Y - X| \). There is nothing to prove when \( |Y - X| = 1 \). When \( |Y - X| > 1 \), choose \( e, f \in Y - X \) and let \( Y' = Y - e - f \). Applying the induction hypothesis using \( Y' \) or \( Y' + e \) or \( Y' + f \) yields \( r(X) = r(X \cup Y') = r(X \cup Y' + e) = r(X \cup Y' + f) \). Now weak absorption yields \( r(X) = r(X \cup Y) \).

\[ A' \Rightarrow U. \] If \( Y \) is a maximal independent subset of \( X \), then \( r(Y + e) = r(Y) \) for all \( e \in X - Y \). By strong absorption, \( r(X) = r(Y) = |Y| \). Hence all such \( Y \) have the same size, \( r(X) \).

\[ U \Rightarrow R. \] Given \( X, Y \subseteq E \), choose a maximum independent set \( I_1 \) from \( X \cap Y \). By uniformity, \( I_1 \) can be enlarged to a maximum independent subset of \( X \cup Y \); call this \( I_2 \). Consider \( I_2 \cap X \) and \( I_2 \cap Y \); these are independent subsets of \( X \) and \( Y \), and each includes \( I_1 \). Hence

\[ r(X \cap Y) + r(X \cup Y) = |I_1| + |I_2| = |I_2 \cap X| + |I_2 \cap Y| \leq r(X) + r(Y). \]

\[ U \Rightarrow R \]

\[ R \Rightarrow C. \] Consider distinct circuits \( C_1, C_2 \in \mathbf{C} \) with \( e \in C_1 \cap C_2 \). We have \( r(C_1) = |C_1| - 1 \) and \( r(C_2) = |C_2| - 1 \). Also \( r(C_1 \cap C_2) = |C_1 \cap C_2| \), since every proper subset of a circuit is independent. If \( (C_1 \cup C_2) - e \) contains no circuit, then its rank is \( |C_1 \cup C_2| - 1 \), and hence \( r(C_1 \cup C_2) \geq |C_1 \cup C_2| - 1 \). Applying submodularity to \( C_1 \) and \( C_2 \) now yields the contradiction

\[ |C_1 \cap C_2| + |C_1 \cup C_2| - 1 \leq |C_1| + |C_2| - 2. \]

\[ C \Rightarrow J. \] If \( I + e \) contains circuits \( C_1 \) and \( C_2 \) for some \( I \in \mathbf{I} \), then \( C_1 \) and \( C_2 \) both contain \( e \). Now weak elimination guarantees a circuit in \( (C_1 \cup C_2) - e \). However, \( (C_1 \cup C_2) - e \) is independent, being contained in \( I \).

\[ J \Rightarrow B'. \] For distinct \( B_1, B_2 \in \mathbf{B} \) and \( e \in B_1 - B_2 \), by \( J \) there is a unique circuit \( C \) in \( B_2 + e \). Since \( C \not\subseteq B_1 \), there exists \( f \in C - B_1 \), and now \( B_2 + e - f \in \mathbf{I} \). To prove \( B_2 + e - f \) is a base, we show that no two bases have distinct sizes. Otherwise, choose bases \( B_1' \) and \( B_2' \) with largest intersection such that \( |B_1'| < |B_2'| \). Since \( B_1' \not\subseteq B_2' \), there is \( e \in B_1' - B_2' \). As before, we find \( f \in B_2' - B_1' \) with \( B_2' + e - f \in \mathbf{I} \). Any base containing \( B_2' + e - f \) is as large as \( B_2' \) but shares more with \( B_1' \) than \( B_2' \).

\[ B' \Rightarrow G. \] With nonnegative weights, some optimum is a base, and the algorithm chooses a base \( B \). Let \( B^* \) be an optimal base having largest intersection with \( B \). If \( B^* \neq B \), then let \( e \) be the first element chosen for \( B \) that is not in \( B^* \). Dual base exchange yields \( f \in B^* - B \) such that \( B^* + e - f \in \mathbf{B} \). The choice of \( B^* \) yields \( w(f) > w(e) \). Since the algorithm chose \( e \) when \( f \) was available, \( w(e) \geq w(f) \). Hence \( B^* + e - f \) contradicts the choice of \( B^* \), so \( B = B^* \).
Given \( I_1, I_2 \in \mathbf{I} \) with \( k = |I_1| < |I_2| \), we design a weight function so that the greedy algorithm finds the desired augmentation. Let \( w(e) = k + 2 \) for \( e \in I_1 \), and let \( w(e) = k + 1 \) for \( e \in I_2 - I_1 \). Let \( w(e) = 0 \) for \( e \notin I_1 \cup I_2 \). Now \( w(I_2) \geq (k + 1)^2 > k(k + 2) = w(I_1) \), so \( I_1 \) does not have maximum weight. However, the greedy algorithm chooses all of \( I_1 \) before any element of \( I_2 - I_1 \). Because it finds a maximum-weight independent set, after absorbing \( I_1 \) it adds an element of \( I_2 - I_1 \) such that \( I_1 + e \in \mathbf{I} \).

11.3.21. Example. Uniform matroids. The uniform matroid of rank \( k \), denoted \( U_{k,n} \) when \( |E| = n \), is defined by \( \mathbf{I} = \{X \subseteq E: |X| \leq k\} \). This immediately satisfies the base exchange and augmentation properties. The free matroid is the uniform matroid of rank \( |E| \). Few uniform matroids are graphic, and few graphic matroids are uniform (Exercise 27). Neither \( M(K_4^+) \) nor \( M(K_4) \) is uniform.

11.3.22. Example. Partition matroids. The partition matroid on \( E \) induced by a partition of \( E \) into blocks \( E_1, \ldots, E_k \) is defined by \( \mathbf{I} = \{X \subseteq E: |X \cap E_i| \leq 1 \text{ for all } i\} \). Since \( \varnothing \in \mathbf{I} \), and since \( X \in \mathbf{I} \) if and only if its elements lie in distinct blocks, \( \mathbf{I} \) is a nonempty ideal. Given \( I_1, I_2 \in \mathbf{I} \) with \( |I_2| > |I_1| \), the set \( I_2 \) must intersect more blocks than \( I_1 \); an element of \( I_2 \) in a block that \( I_1 \) misses yields the desired augmentation of \( I_1 \). Alternatively, \( r(X) \) is the number of blocks having elements in \( X \); this satisfies the absorption property. Every partition matroid is a transversal matroid; consider the \( E, [k]\)-bigraph consisting of \( k \) stars whose leaf sets are the blocks of the partition.

Given a \( U, V \)-bigraph \( G \), each part induces a partition matroid with ground set \( E(G) \), partitioning the edges by their endpoints in one part. A set \( X \subseteq E(G) \) is a matching in \( G \) if and only if \( X \) is independent both in the partition matroid induced by \( U \) and in the partition matroid induced by \( V \). This is the motivation for the idea of matroid intersection discussed later in this section.

When \( G \) has an odd cycle, \( G \) has no set of vertices whose incident sets partition \( E(G) \); thus \( M(K_4^+) \) is not a partition matroid, for example. In a digraph, however, each edge has a head and a tail, and we can define the head partition matroid and the tail partition matroid using the edge partitions induced by incidences with heads and by incidences with tails. For example, the matroid of Example 11.3.2 arises as the partition matroid on \( E \) induced by \( U \) in the bipartite graph on the left below, as the head partition matroid in the first digraph, and as the tail partition matroid in the second digraph.
DUALITY AND MINORS

Duality in matroids generalizes duality in planar graphs (Section 9.1). Every connected plane graph $G$ has a natural dual graph $G^*$ such that $(G^*)^* = G$. The dual associates a vertex of $G^*$ to each face of $G$ and an edge of $G^*$ to each edge of $G$, joining the faces incident to it. Loops and multiedges can arise in $G^*$ (from cut-edges and pairs of faces sharing more than one boundary edge, respectively), so when discussing graphic matroids we allow the generality of multigraphs.

A set of edges forms a spanning tree in a plane graph $G$ if and only if the duals to the remaining edges form a spanning tree in $G^*$ (Exercise 9.1.25). Hence the bases in the cycle matroid $M(G^*)$ are the complements of the bases in $M(G)$. We define duality for matroids and for hereditary systems to extend these duality properties of planar graphs.

11.3.23. Definition. The dual of a hereditary system $M$ on $E$ is the hereditary system $M^*$ defined by $B_{M^*} = \{ \overline{B} : B \in B_M \}$. We may write $B'$ for $B_{M^*}$ when $M$ is understood; these are the cobases of $M$. Similarly, the circuits of $M^*$ are the cocircuits of $M$, denoted $C^*$, etc.

A set in $E$ is spanning if it contains a base; $S_M$ denotes the family of spanning sets. A set is a hyperplane if it is a maximal set containing no base; $H_M$ denotes the family of hyperplanes.

11.3.24. Remark. If $B$ is a nonempty antichain, then the same holds for $\{ \overline{B} : B \in B \}$ (since complements of incomparable sets are incomparable), so the dual of a hereditary system is a hereditary system. Also it is immediate that $(M^*)^* = M$, which explains the name dual.

A set is independent in $M^*$ if and only if it is contained in a cobase of $M$. Hence $I_{M^*} = \{ \overline{S} : S \in S_M \}$, and similarly $S_{M^*} = \{ \overline{I} : I \in I_M \}$.

A set is a circuit in $M^*$ if and only if it is a minimal set contained in no cobase of $M$. Hence $C_{M^*} = \{ \overline{H} : H \in H_M \}$ and $H_{M^*} = \{ \overline{C} : C \in C_M \}$.

Duality is useful because the dual of a matroid is a matroid. This follows from the dual base exchange property ($B'$). In essence, the statement of the dual base exchange property for $M$ is that of the base exchange property for $M^*$.

11.3.25. Theorem. (Whitney [1935]) The dual of a matroid $M$ is a matroid.

Proof: We have seen that $M^*$ is a hereditary system; now we prove the base exchange property for $M^*$. If $\overline{B_1}, \overline{B_2} \in B^*$ and $e \in \overline{B_1} - \overline{B_2}$, then $B_1, B_2 \in B$, with $e \in B_2 - B_1$. By dual base exchange for $M$, there exists $f \in B_1 - B_2$ such that $B_1 + e - f \in B$. Now $\overline{B_1 - e + f} \in B^*$, proving base exchange for $M^*$.

Computing the rank function of the dual is easy using an alternative notion of the rank of a set. Instead of viewing $r(X)$ as the maximum size of an independent subset of $X$, it can be helpful to view it as the maximum size of the intersection of $X$ with a base.

11.3.26. Proposition. The rank function $r^*$ of the dual of a matroid $M$ on $E$ is given by $r^*(X) = |X| - (r(E) - r(\overline{X}))$. 

Proof: The rank of a set $X$ is the maximum size of its intersections with bases. Choose $B^* \in B^*$ so that $r^*(X) = |X \cap B^*|$, and let $B = B^* \setminus X$. Now $B$ is a base having smallest intersection with $X$ and hence largest intersection with $\overline{X}$. Hence $r(\overline{X}) = |B - X|$. Also $|B| = r(E)$, so $r^*(X) = |X| - |X \cap B| = |X| - (r(E) - r(\overline{X})).$
11.3.30. Definition. Given a hereditary system $M$ on $E$ and a set $F \subseteq E$, the restriction of $M$ to $F$, denoted $M|F$ and obtained by deleting $F$, is the hereditary system defined by $I_{M|F} = \{X \subseteq F : X \in I_M\}$. The contraction of $M$ to $F$, denoted $M.F$ and obtained by contracting $F$, is the hereditary system defined by $S_{M.F} = \{X \subseteq F : X \cup F \in S_M\}$. When $F = E - e$, we write $M - e$ for $M|F$ and $M \cdot e$ for $M.F$. The minors of a hereditary system $M$ are the hereditary systems arising from $M$ by applying deletions and contractions.

From the definitions, $M|F$ and $M.F$ are hereditary systems. The notations $M|F$ and $M.F$ appear (briefly) in Oxley [1992] (pages 22 and 104, respectively). Note the distinction between “.” and “|”; the former emphasizes “contracting to” a specified set by “contracting away” the other elements. We use “−” or “|” when eliminating one specified element and “|” or “.” when specifying the elements that remain. Our notation for $M - e$ and $M \cdot e$ is consistent with our usage for graphs but is nonstandard in the matroid community, where $M \setminus e$ for our $M - e$ and $M/e$ for our $M \cdot e$ are most common.

Defining contraction via spanning sets yields a natural duality.

11.3.31. Proposition. For hereditary systems, restriction and contraction are dual operations: $(M.F)^* = (M^*|F)$ and $(M|F)^* = (M^*.F)$.

Proof: $I_{(M|F)^*} = \{X \subseteq F : F - X \in S_{M.F}\} = \{X \subseteq F : (F - X) \cap F \in S_M\}$

$= \{X \subseteq F : \overline{X} \in S_M\} = \{X \subseteq F : X \in I_{M^*}\} = I_{M^*|F}$.

For the second statement, apply the first to $M^*$ and take duals. ■

As expected, restrictions and contractions of matroids are matroids.

11.3.32. Theorem. Given $F \subseteq E$ and a matroid $M$ on $E$, both $M|F$ and $M.F$ are matroids on $F$. Their bases and rank functions are given by

$$B_{M|F} = \{B \cap F : B \in B_M \text{ and } |B \cap F| = r_M(F)\}$$
$$B_{M.F} = \{B \cap F : B \in B_M \text{ and } |B \cap F| = r_M(F)\}$$

$$r_{M|F}(X) = r_M(X)$$
$$r_{M.F}(X) = r_M(X \cup F) - r_M(F)$$

Proof: The augmentation property from $M$ applies to sets in $I_{M|F}$; thus $M|F$ satisfies the augmentation property and is a matroid. Since $M.F = (M^*|F)^*$, duality implies that $M.F$ is also a matroid.

The expressions for the bases and rank function of $M|F$ follow from the definition of $I_{M|F}$; we do the same for $M.F$ from the definition of $S_{M.F}$. A base of $M.F$ is a minimal set $\hat{B} \subseteq F$ such that $\hat{B} \cup F \in S_M$. Thus $\hat{B} \cup F$ contains a base $B$ of $M$. The minimality of $\hat{B}$ implies that $\hat{B} \subseteq B$ and that $B \cap F$ is a maximal independent subset of $F$ (by uniformity on $\hat{B} \cup F$). Thus $B_{M.F}$ consists of the sets in $F$ whose addition to a maximal independent subset of $F$ yields a base of $M$.

To compute $r_{M.F}(X)$, recall that $r_{M.F}(X) = |B' \cap X|$ for some $B' \in B_{M.F}$; let $Y = B' \cap X$. Also $B' = B \cap F$ for some $B \in B_M$ such that $|B \cap F| = r_M(F)$; let $Z = B \setminus \overline{F}$. Since $B \in B_M$, we have $Y \cup Z \in I_M$. If $Y \cup Z$ is not a maximal independent subset of $X \cup F$, then it augments to a base in $M$ that contains $Z$ and has larger intersection with $X$ than $B'$, contradicting the choice of $B'$. Hence uniformity in $M$ yields $|Y \cup Z| = r_M(X \cup F)$. Now $|Y| = r_M(X \cup F) - r_M(F)$. ■
11.3.33. Remark. The formula for \( r_{MF} \) yields a description of the independent sets: \( X \in 1_{MF} \) if and only if adding \( X \) to \( \bar{F} \) increases the rank by \( |X| \). Note that when \( F \) is an independent set \( \{e\} \) of size 1, we have \( r_{MF}(X) = r_M(X + e) - 1 \). In particular, when we contract a nonloop edge in a graph, the maximum size of a forest among the edges of any set containing that edge decreases by 1.

The duality between deletion and contraction is familiar in plane graphs. Deleting an edge \( e \) in a plane graph \( G \) contracts the corresponding edge in \( G^* \); contracting \( e \) deletes the edge in the dual. The fate of the 4-cycle below illustrates that when a circuit of \( M \) intersects \( F \), its intersection with \( F \) need not be a circuit of \( MF \), even when only one element has been contracted. In fact, the circuits of \( MF \) are the minimal nonempty sets in \( \{C \cap F: C \in C_M\} \) (Exercise 48).

Also, restriction and contraction commute (Exercise 48).

\[ \begin{align*}
\rightarrow \quad & \text{contract} \\
\leftarrow \quad & \text{delete}
\end{align*} \]

11.3.34. Corollary. The behavior of cycle matroids and bond matroids under deletion or contraction of an edge \( e \in E(G) \) is

\[
\begin{align*}
M(G - e) &= M(G) - e & M^*(G - e) &= M^*(G) \cdot e \\
M(G \cdot e) &= M(G) \cdot e & M^*(G \cdot e) &= M^*(G) - e
\end{align*}
\]

Proof: Matroid deletion and contraction are defined so that the statements in the first column describe the behavior of cycle matroids. Using these and Proposition 11.3.31, for the second column we compute

\[
\begin{align*}
M^*(G - e) &= [M(G - e)]^* = [M(G) - e]^* = M^*(G) \cdot e \\
M^*(G \cdot e) &= [M(G \cdot e)]^* = [M(G) \cdot e]^* = M^*(G) - e
\end{align*}
\]

Now we characterize planar graphs using matroids. In Theorem 9.1.11, we proved that a set of edges in a plane graph \( G \) forms a cycle if and only if the corresponding dual edges form a bond in \( G^* \). Thus under the natural bijection from edges to dual edges, the cycle matroid of a plane graph \( G \) is (isomorphic to) the bond matroid of \( G^* \). By Corollary 11.3.28, the bond matroid of a graph \( H \) is \([M(H)]^*\). Applying this to \( G \) and \( G^* \) tells us that the bond matroid of \( G \) is (isomorphic to) the cycle matroid of \( G^* \). In particular, the bond matroid of \( G \) is graphic. Using Kuratowski’s Theorem, we will use this to characterize planarity.

Whitney [1933a] approached this question by defining a non-geometric notion of the dual of a graph. Changing his definition slightly, we say that \( H \) is an abstract dual of \( G \) if there is a bijection \( \phi: E(G) \to E(H) \) such that \( X \subseteq E(G) \) is a bond in \( G \) if and only if \( \phi(X) \) is the edge set of a cycle in \( H \). With this definition, the statement that \( G \) has an abstract dual \( H \) is the same as the statement that the bond matroid of \( G \) is graphic; the bijection \( \phi \) establishes an isomorphism between \( M^*(G) \) and \( M(H) \).

11.3.35. Theorem. \((\text{Whitney [1933a]})\) A graph \( G \) is planar if and only if its bond matroid \( M^*(G) \) is graphic.
Proof: We have observed that planar graphs have abstract duals; this proves necessity. For sufficiency, we first prove that existence of an abstract dual is preserved under deletion and contraction of edges. Suppose that $G$ has an abstract dual $H$, so that $M(H) \cong M^*(G)$. Let $e'$ be the edge of $H$ corresponding to $e$ under the bijection. To prove that $H \cdot e'$ is an abstract dual of $G - e$ and that $H - e'$ is an abstract dual of $G \cdot e$, we use Corollary 11.3.34 to compute

\[
M^*(G - e) = M^*(G) \cdot e \cong M(H) \cdot e' = M(H \cdot e'),
\]

\[
M^*(G \cdot e) = M^*(G) - e \cong M(H) - e' = M(H - e').
\]

By Kuratowski’s Theorem, a nonplanar graph contains a subdivision of $K_5$ or $K_{3,3}$. Hence $K_5$ or $K_{3,3}$ is a minor of it. Since existence of abstract duals is preserved under deletion and contraction, showing that $K_5$ and $K_{3,3}$ have no abstract dual implies that every nonplanar graph has no abstract dual.

If $H$ is an abstract dual of $G$, then also $G$ is an abstract dual of $H$, since $M^*(G) \cong M(H)$ if and only if $M(G) \cong M^*(H)$. If $G$ has girth $g$, then bonds of $H$ have size at least $g$, so $\delta(H) \geq g$. Letting $n = |V(H)|$ and $m = |E(H)|$, we have also $|E(G)| = m$. Thus the degree-sum formula yields $n \leq \lfloor 2m/\delta(H) \rfloor \leq \lfloor 2m/g \rfloor$.

Let $H$ be an abstract dual of $K_5$. Since $K_5$ has girth $3$, $n \leq \lfloor 20/3 \rfloor = 6$. Since all bonds of $K_5$ have four or six edges, all cycles of $H$ have four or six edges, and thus $H$ is a simple bipartite graph. However, no simple bipartite graph with at most six vertices has ten edges.

Let $H$ be an abstract dual of $K_{3,3}$. Since $K_{3,3}$ has girth $4$, $n \leq \lfloor 18/4 \rfloor = 4$. Since all bonds of $K_{3,3}$ have at least three edges, all cycles of $H$ have at least three edges, and thus $H$ is a simple graph. However, no simple graph with at most four vertices has nine edges.

Planar graphs have no $K_5$- or $K_{3,3}$-minor, since planarity is preserved under deletion and contraction of edges. Kuratowski’s Theorem implies that nonplanar graphs do have such minors. This yields the characterization by Wagner [1937]: $K_5$ and $K_{3,3}$ are the minimal forbidden minors for planar graphs.

### THE SPAN FUNCTION

In addition to the notion of “spanning set” as “set containing a base”, another notion of span is suggested by vector spaces. A set $S$ of vectors spans a vector $v$ if $v$ is a linear combination of elements of $S$, which means that $S + v$ contains a dependent set. This notion defines another aspect of hereditary systems.

**11.3.36. Definition.** The **span function** of a hereditary system $M$ on $E$ is the function $\sigma_M$ on $2^E$ defined by $\sigma_M(X) = X \cup \{e \in E : Y \cup e \in C_M$ for some $Y \subseteq X\}$. If $e \in \sigma(X)$, then $X$ spans $e$.

A set in a hereditary system is dependent if and only if it contains a circuit, which by Definition 11.3.36 means that $e \in \sigma(X - e)$ for some $e \in X$. Hence we can find the independent sets (and other aspects of $M$) from the span function via $I_M = \{X \subseteq E : (e \in X) \Rightarrow (e \notin \sigma_M(X - e))\}$. The properties of $\sigma$ that we use in studying matroids are (s1, s2, s3) below.
11.3.37. Proposition. If \( \sigma \) is the span function of a hereditary system on \( E \), and \( X, Y \subseteq E \), then
s1) \( X \subseteq \sigma(X) \) (\( \sigma \) is expansive).
\( s2) Y \subseteq X \) implies \( \sigma(Y) \subseteq \sigma(X) \) (\( \sigma \) is order-preserving).
\( s3) e \notin \sigma(X) \) and \( e \in \sigma(X + f) \) imply \( f \in \sigma(X + e) \) (Steinitz exchange).

Proof: Definition 11.3.36 implies immediately that \( \sigma \) is expansive and order-preserving. If \( e \notin \sigma(X) \), then \( e \notin X \). With \( e \in \sigma(X + f) \), also \( e \) belongs to a circuit \( C \) in \( X + f + e \). With \( e \notin \sigma(X) \), we must have \( f \in C \). This circuit yields \( f \in \sigma(X + e) \), and hence \( \sigma \) satisfies the Steinitz exchange property.

11.3.38. Example. Steinitz exchange in cycle matroids. In the cycle matroid \( M(G) \), the meaning of \( e \notin \sigma(X) \) is that \( X \) (the solid edges below) contains no path between the endpoints of \( e \). If \( e \notin \sigma(X) \) but \( e \in \sigma(X + f) \), then adding \( f \) completes such a path. Adding \( e \) to the path completes a cycle, so also \( f \in \sigma(X + e) \).

\[ e \quad f \]

\[ \begin{array}{c}
\text{Definition 11.3.36 immediately implies that the circuits of a hereditary system satisfy the strong dependence property: } e \in C \text{ implies } e \in \sigma(C - e). \text{ The natural notion that an element is spanned by a set } X \text{ if adding it to } X \text{ does not increase the rank is valid for all hereditary systems.} \\
\end{array} \]

11.3.39. Lemma. In a hereditary system, \( [r(X + e) = r(X)] \Rightarrow e \in \sigma(X) \).

Proof: Let \( Y \) be a maximum independent subset of \( X \). Since \( |Y| = r(X) = r(X + e) \), also \( Y \) is a maximum independent subset of \( X + e \). Hence \( e \) completes a circuit with some subset of \( X \) contained in \( Y \), and \( e \in \sigma(X) \).

In fact, the converse characterizes matroids! We call the converse of Lemma 11.3.39 the “incorporation property”. We leave the proof of these equivalences as an exercise, because we will not need to use them.

11.3.40. Theorem. If \( M \) is a hereditary system, then each condition below is necessary and sufficient for \( M \) to be a matroid.
P: incorporation - \( r(\sigma(X)) = r(X) \) for all \( X \subseteq E \).
S: idempotence - \( \sigma^2(X) = \sigma(X) \) for all \( X \subseteq E \).
T: transitivity of dependence - if \( e \in \sigma(X) \) and \( X \subseteq \sigma(Y) \), then \( e \in \sigma(Y) \).
C’: strong elimination - whenever \( C_1, C_2 \in C \), \( e \in C_1 \cap C_2 \), and \( f \in C_1 \Delta C_2 \), there exists \( C \in C \) such that \( f \in C \subseteq (C_1 \cup C_2) - e \).

Proof: See Exercise 49.

Incorporation, idempotence, and transitivity of dependence are well-known properties of matroids. The equivalence of \( C \) and \( C’ \) for hereditary systems was first proved by Lehman [1964]. Brylawski [1986] mentioned another short way to obtain \( C’ \) from other axioms. Idempotence occurs naturally for linear matroids. The span of a set of vectors contains nothing additional in its own span. This suggests another aspect of hereditary systems.
11.3.41. Definition. The **closed sets** of a hereditary system on \( E \) are the sets \( X \subseteq E \) such that \( \sigma(X) = X \) (also called **flats** or **subspaces**).

In a matroid on \( E \), the sets whose span is \( E \) are the sets containing bases, which is why earlier we called these the spanning sets. Similarly, in a matroid the hyperplanes are the maximal proper closed subsets of \( E \). Both statements fail for hereditary systems in general.

The span function of a matroid is also called its **closure function**. A **closure operator** on \( 2^E \) is an expansive, order-preserving, idempotent function from \( 2^E \) to \( 2^E \). Such an operator is the span function of a matroid if and only if it satisfies Steinitz exchange. Since every hereditary system satisfies Steinitz exchange, our approach to matroids as hereditary systems with an added property is not well suited for studying closure operators. The span function of a hereditary system \( M \) is a closure operator if and only if \( M \) is a matroid. Matroids are developed from lattice theory in MacLane [1936], Rota [1964], and Aigner [1979].

We have not exhausted the relationships between various aspects of hereditary systems and matroids. Brylawski [1986] described the transformations among about a dozen aspects of matroids, calling these maps **cryptomorphisms**.

**MATROID INTERSECTION**

Matroid theory took a great leap forward with the Matroid Intersection and Union Theorems. They provided a unified context for many well-known min-max relations, which became corollaries. We proved some of these in earlier chapters. Yielding a unified proof for many important theorems, the Matroid Intersection Theorem can be considered among the most beautiful theorems of combinatorics.

The Matroid Intersection Theorem is a min-max relation for the maximum size of common independent sets in two matroids on the same set \( E \). The intersection of two matroids is a hereditary system but generally not a matroid. For multiple matroids on a set \( E \), we typically use subscripts to distinguish corresponding aspects, as in \( B_i \) for the bases of \( M_i \), etc. We still use \( \overline{X} \) for the complement of \( X \) within the full set \( E \).

11.3.42. Definition. The **intersection** of hereditary systems \( M_1 \) and \( M_2 \) on \( E \) is the hereditary system whose independent sets are \( \{ X \subseteq E : X \in I_1 \cap I_2 \} \).

11.3.43. Example. Since \( I_1 \) and \( I_2 \) are closed under taking subsets, the common independent sets in two hereditary systems also form a hereditary family.

In a bipartite graph \( G \) with edge set \( E \), each part induces a partition matroid on \( E \). A set of edges forms an independent set in one of these partition matroids if and only if its endpoints in the corresponding part of \( G \) are distinct. A set of edges is independent in both matroids if and only if it is a matching in \( G \).

The hereditary system whose independent sets are the matchings in a bipartite graph is not a matroid. The central edge in a path of length 3 forms an independent set, and the two end edges form a larger one, but the smaller set cannot be augmented from the larger, so augmentation fails. This is why the greedy algorithm does not solve maximum matching in bipartite graphs. ■
11.3.44. Theorem. (Matroid Intersection Theorem; Edmonds [1970]) For matroids $M_1$ and $M_2$ on $E$,

$$\max_{I \in \mathcal{I}_1 \cap \mathcal{I}_2} |I| = \min_{X \subseteq E} (r_1(X) + r_2(X)).$$

Proof: (Woodall; see Seymour [1976]) For the upper bound on $|I|$, consider $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ and $X \subseteq E$. The sets $I \cap X$ and $I \cap \overline{X}$ are also common independent sets, and $|I| = |I \cap X| + \left|I \cap \overline{X}\right| \leq r_1(X) + r_2(X)$. Hence $\max |I| \leq \min \{r_1(X) + r_2(X)\}$.

![Venn diagram of sets X, I, and \overline{X}]

To achieve equality, we use induction on $|E|$; when $|E| = 0$ both sides are 0. If every element of $E$ is a loop in $M_1$ or in $M_2$, then $\max |I| = 0 = r_1(X) + r_2(X)$, where $X$ consists of all loops in $M_1$. Hence we may assume that $|E| > 0$ and that some $e \in E$ is a non-loop in both matroids. Let $F = E - e$, and consider the matroids $M_1|F$, $M_2|F$, $M_1.F$, $M_2.F$. By Theorem 11.3.32, a set $X$ is independent in $M_1.F$ if and only if $X + e$ is independent in $M_1$, and similarly for $M_2.F$.

Let $k = \min_{X \subseteq E} \{r_1(X) + r_2(X)\}$; we seek a common independent $k$-set in $M_1$ and $M_2$. If none exists, then $M_1|F$ and $M_2|F$ have no common independent $k$-set, and $M_1.F$ and $M_2.F$ have no common independent $(k - 1)$-set. By the induction hypothesis and rank formulas (Theorem 11.3.32),

$$r_1(X) + r_2(F - X) \leq k - 1 \quad \text{for some } X \subseteq F,$$

$$r_1(Y + e) - 1 + r_2(F - Y + e) - 1 \leq k - 2 \quad \text{for some } Y \subseteq F.$$

We use $(F - Y) + e = \overline{Y}$ and $F - X = \overline{X} + e$ and sum the two inequalities:

$$r_1(X) + r_2(X + e) + r_1(Y + e) + r_2(Y) \leq 2k - 1.$$

Write $X' = X + e$ and $Y' = Y + e$. We apply submodularity of $r_1$ to $X$ and $X'$ and submodularity of $r_2$ to $\overline{X}$ and $\overline{X'}$. With the preceding inequality, this yields

$$r_1(X \cup Y') + r_1(X \cap Y') + r_2(Y \cup \overline{X}) + r_2(\overline{Y} \cap \overline{X'}) \leq 2k - 1.$$

Since $Y \cap \overline{X'} = \overline{X} \cup \overline{Y'}$ and $Y \cup \overline{X} = \overline{X} \cap \overline{Y'}$ (see Venn diagram below), the left side sums two instances of $r_1(Z) + r_2(Z)$. Hence the hypothesis $k \leq r_1(Z) + r_2(Z)$ for all $Z \subseteq E$ yields $2k \leq 2k - 1$. This contradiction implies that $M_1$ and $M_2$ do have a common independent $k$-set. \blacksquare
We have proved special cases of the Matroid Intersection Theorem by other means. We proved the König–Egerváry Theorem in various ways, and we proved the Ford–Fulkerson characterization of CSDR’s from Menger’s Theorem in Theorem 7.2.15. Given any two matroids on the same set, the Matroid Intersection Theorem guarantees a min-max relation for the maximum size of a common independent set, tells us what the result should be, and provides a proof.

11.3.45. Corollary. (König [1931], Egerváry [1931]) In a bipartite graph, a largest matching and smallest vertex cover have equal size.

Proof: In the partition matroids \( M_1 \) and \( M_2 \) on \( E(G) \) induced by the partite sets \( U_1 \) and \( U_2 \), the common independent sets are the matchings in \( G \).

For \( X \subseteq E \), the number of vertices of \( U_i \) incident to edges of \( X \) is \( r_i(X) \). Hence \( r_1(X) + r_2(\overline{X}) \) is the size of a vertex cover, using \( U_1 \) for \( X \) and \( U_2 \) for \( \overline{X} \). On the other hand, if \( T_1 \cup T_2 \) is a vertex cover with \( T_1 \subseteq U_1 \), and \( X \) is the set of edges incident to \( T_1 \), then \( r_1(X) + r_2(\overline{X}) \leq |T_1 \cup T_2| \). We conclude that \( \min\{r_1(X) + r_2(\overline{X})\} \) is the minimum size of a vertex cover.

With \( \alpha'(G) \) and \( \beta(G) \) denoting the maximum size of a matching and the minimum size of a vertex cover, we obtain

\[
\alpha'(G) = \max\{|I| : I \in I_1 \cap I_2\} = \min\{r_1(X) + r_2(\overline{X})\} = \beta(G).
\]

The next application uses the rank function for transversal matroids.

11.3.46. Example. Transversal matroids (Definition 11.3.11). The transversal matroid induced by subsets \( A_1, \ldots, A_m \) of \( E \) has as its independent sets the partial systems of distinct representatives. Equivalently, these are subsets of \( E \) that can be matched into \( |m| \) in the bipartite incidence graph \( G \) with parts \( E \) and \( |m| \).

If \( N(S) < |S| \) for some \( S \subseteq X \subseteq E \), then every matching in \( G \) leaves at least \( |S| - |N(S)| \) elements of \( X \) uncovered. Ore’s Defect Formula (Corollary 6.1.12) is \( r(X) = \min_{S \subseteq X}|X| - (|S| - |N(S)|) \), which we now rewrite as \( r(X) = \min_{S \subseteq X}[|N(S)| + |X - S|] \).

Ore [1955] gave another useful expression for \( r(X) \). For \( J \subseteq |m| \), let \( A(J) = \bigcup_{i \in J} A_i \); in terms of the graph, \( A(J) = N(J) \) (see figure below). Since \( r(X) = \alpha'(G[X \cup \{m\}]) \), we can write the defect formula in terms of neighborhoods of subsets of \( |m| \) instead of subsets of \( X \) to obtain

\[
r(X) = \min_{J \subseteq |m|}[|A(J) \cap X| + m - |J|]. \tag{*}
\]

The upper bound here holds because for any \( J \subseteq |m| \), at most \( |A(J) \cap X| \) elements of \( X \) can be matched into \( J \), and at most \( m - |J| \) can be matched into \( |m| - J \). The statement that equality holds is equivalent to the König–Egerváry Theorem, since \( (A(J) \cap X) \cup (|m| - J) \) is a vertex cover of \( G[X \cup \{m\}] \), and any minimal vertex cover \( Q \) can be achieved in this way by setting \( J = |m| - Q \). \( \blacksquare \)
11.3.47. Corollary. (Ford–Fulkerson [1958]) Families $A = \{A_1, \ldots, A_m\}$ and $B = \{B_1, \ldots, B_m\}$ have a common system of distinct representatives (CSDR) if and only if, for each $I, J \subseteq [m]$,
\[
\left| \left( \bigcup_{i \in I} A_i \right) \cap \left( \bigcup_{j \in J} B_j \right) \right| \geq |I| + |J| - m.
\]

Proof: Let $E$ be the union of all the sets in $A \cup B$. A common partial transversal is a common independent set in the two transversal matroids $M_1, M_2$ induced on $E$ by $A$ and $B$. A common transversal is a common independent set of size $m$. We need only restate the condition $r_1(X) + r_2(\overline{X}) \geq m$ to find the appropriate condition on the set systems.

The rank formula (9) from Example 11.3.46 yields
\[
r_1(X) + r_2(\overline{X}) = \min_{I \subseteq [m]} \{|A(I) \cap X| - |I| + m\} + \min_{J \subseteq [m]} \{|B(J) \cap \overline{X}| - |J| + m\}.
\]

Hence $r_1(X) + r_2(\overline{X}) \geq m$ for all $X$ if and only if
\[
|A(I) \cap X| + |B(J) \cap \overline{X}| \geq |I| + |J| - m \text{ for all } X \subseteq E \text{ and } I, J \subseteq [m].
\]

Since $|A(I) \cap X| + |B(J) \cap \overline{X}| \geq |A(I) \cap B(J)|$, the Ford–Fulkerson condition is sufficient. To see that it is necessary, let $X$ be a subset of $E$ such that $A(I) - B(J) \subseteq \overline{X}$ and $B(J) - A(I) \subseteq X$. Given $I, J$, consider the contribution by elements of $E$ to the left side. In this case $|A(I) \cap X| + |B(J) \cap \overline{X}| = |A(I) \cap B(J)|$, so the condition on $r_1(X) + r_2(\overline{X})$ shows that the condition on $A(I) \cap B(J)$ is necessary.

11.3.48. Remark. The augmenting path approach to maximum bipartite matching generalizes to matroid intersection. The algorithm yields a common independent set $I$ of maximum size and a set $X$ such that $r_1(X) + r_2(\overline{X}) = |I|$ (see Lawler [1976], Edmonds [1979], Faigle [1987]). Finding a maximum common independent set in three matroids is NP-complete (Exercises 25–26).

Maximum matching in general graphs also extends to a matroid context. A common generalization of matroid intersection and general graph matching is known in different phrasings as the Matroid Matching Problem or the Matroid Parity Problem. It is solvable in polynomial time for linear matroids, in which case it yields a min-max relation (Lovász [1978]). Lovász also proved that it is not solvable in polynomial time for general matroids in the same computational model. The solution technique over linear matroids involves “polymatroids”, a polyhedral generalization of matroids.

Next we consider a related min-max relation for “union” of matroids. In this and other applications of matroid intersection, it can be helpful to restrict the range of the minimization.
11.3.49. **Corollary.** The maximum size of a common independent set in matroids $M_1, M_2$ on $E$ is the minimum of $r_1(X_1) + r_2(X_2)$ over sets $X_1, X_2$ such that $X_1 \cup X_2 = E$ and each $X_i$ is closed in $M_i$.

**Proof:** The incorporation property implies that $r_i(\sigma_i(X)) = r_i(X)$. ■

**MATROID UNION**

The intersection of two matroids is seldom a matroid, but a natural concept of union does yield a matroid. The Matroid Union Theorem states this and gives a useful min-max relation for the rank function. The Matroid Intersection and Union Theorems are equivalent; they can be derived from each other. Welsh [1976] proves the Matroid Union Theorem first; here we obtain it from the Matroid Intersection Theorem.

11.3.50. **Definition.** The **union** $M_1 \cup \cdots \cup M_k$ of hereditary systems $M_1, \ldots, M_k$ on $E$ is the hereditary system $M$ on $E$ defined by $I_M = \{I_1 \cup \cdots \cup I_k : I_i \in I_i\}$. The **direct sum** $M_1 \oplus \cdots \oplus M_k$ of hereditary systems $M_1, \ldots, M_k$ on disjoint sets $E_1, \ldots, E_k$ is the hereditary system $M$ on $E_1 \cup \cdots \cup E_k$ defined by $I_M = \{I_1 \cup \cdots \cup I_k : I_i \in I_i\}$.

Since $I_M$ is a nonempty ideal, $M_1 \cup \cdots \cup M_k$ is indeed a hereditary system. The direct sum $M_1 \oplus \cdots \oplus M_k$ on $E_1, \ldots, E_k$ can be expressed as a matroid union. Let $E' = E_1 \cup \cdots \cup E_k$. For $1 \leq i \leq k$, let $M'_i$ be a copy of $M_i$ defined on $E'$ by letting the elements outside $E_i$ be loops. Now $M_1 \oplus \cdots \oplus M_k = M'_1 \cup \cdots \cup M'_k$.

When each $M_i$ is a uniform matroid, the direct sum is a generalized partition matroid. Here $E_1, \ldots, E_k$ partition $E$, there are positive integers $r_1, \ldots, r_k$, and $X \in I$ if $|X \cap E_i| \leq r_i$. The partition matroids defined earlier arise when all $r_i = 1$.

11.3.51. **Proposition.** Given matroids $M_1, \ldots, M_k$ on disjoint sets $E_1, \ldots, E_k$, the direct sum $M = M_1 \oplus \cdots \oplus M_k$ is a matroid.

**Proof:** Since the $E_1, \ldots, E_k$ are pairwise disjoint, the intersection of any $I \in I$ with each $E_i$ is independent in $M_i$. If $I_1, I_2 \in I$ with $|I_2| > |I_1|$, then $|I_2 \cap E_i| > |I_1 \cap E_i|$ for some $i$. Since both sets are independent in $M_i$, we can augment $I_1 \cap E_i$ from $I_2 \cap E_i$ and therefore $I_1$ from $I_2$. Hence $M_1 \oplus \cdots \oplus M_k$ satisfies the augmentation property. ■

Using a direct sum, we prove that the union of matroids is always a matroid, and we compute the rank function.

11.3.52. **Theorem.** **(Matroid Union Theorem; Edmonds–Fulkerson [1965], Nash-Williams [1966])** If $M_1, \ldots, M_k$ are matroids on $E$ with rank functions $r_1, \ldots, r_k$, then the union $M = M_1 \cup \cdots \cup M_k$ is a matroid with rank function $r(X) = \min_{Y \subseteq X}(|X - Y| + \sum r_i(Y))$.

**Proof:** (following Schrijver [2003]). After proving the formula for the rank function, we will verify the submodularity property to prove that $M$ is a matroid.
First we reduce the computation of the rank function to the computation of \( r(E) \). In the restriction of the hereditary system \( M \) to the set \( X \), we have \( I_{M|X} = \{ Y \subseteq X : Y \in I_M \} \) and \( r_{M|X}(Y) = r_M(Y) \) for \( Y \subseteq X \). Thus \( M|X = \cup_i(M|i|X) \), and applying the formula for the rank of the full union to \( M|X \) yields \( r_M(X) \).

Consider a \( k \)-by-\( |E| \) grid of elements \( E' \) in which the \( j \)-th column \( E_j \) consists of \( k \) copies of the element \( e_j \in E \). We define two matroids \( N_1, N_2 \) on \( E' \) such that the maximum size of a set independent in both \( N_1 \) and \( N_2 \) equals the maximum size of a set independent in \( M \). We then compute \( r_M(E) \) by applying the Matroid Intersection Theorem to \( N_1 \) and \( N_2 \). Let \( M'_i \) be a copy of \( M_i \) defined on the elements \( E' \) of row \( i \) in \( E' \). Let \( N_1 \) be the direct sum matroid \( M'_1 \oplus \cdots \oplus M'_k \), and let \( N_2 \) be the partition matroid induced on \( E' \) by the column partition \( \{ E_j \} \).

Each set \( X_i = I_M \) has a decomposition into disjoint subsets \( X_i \in I_i \), since \( I_i \) is an ideal. Given a decomposition \( \{ X_i \} \) of \( X \in I_M \), let \( X'_i \) be the copy of \( X_i \) in \( E' \). Since \( \{ X_i \} \) are disjoint, \( \cup X'_i \) is independent in \( N_2 \), and \( X_i \in I_i \) implies that \( \cup X'_i \) is also independent in \( N_1 \). From \( X \in I_M \), we have constructed \( \cup X'_i \) of size \( |X| \) in \( I_{N_1} \cap I_{N_2} \). Conversely, any \( X' \in I_{N_1} \cap I_{N_2} \) corresponds to a decomposition of a set in \( I_M \) of size \( |X| \) when the sets \( X' \cap E' \) are transferred back to \( E \), because \( N_2 \) forbids multiple copies of elements.

Hence \( r(E) = \max \{|I| : I \in I_{N_1} \cap I_{N_2} \} \). Let the rank functions of \( N_1, N_2 \) be \( q_1, q_2 \), and let \( r'_j \) be the rank function of the copy \( M'_i \) of \( M_i \) on \( E' \). We have \( q_1(X') = \sum r'_j(X' \cap E') \), and \( q_2(X') \) is the number of elements of \( E \) that have copies in \( X' \). The Matroid Intersection Theorem yields \( r(E) = \min_{X' \subseteq E'} \{ q_1(X') + q_2(E' - X') \} \).

By Corollary 11.3.49, the minimum is achieved by a set \( X' \) such that \( E' - X' \) is closed in \( N_2 \). The closed sets in the partition matroid \( N_2 \) are the sets that contain all or none of the copies of each element—the unions of full columns of \( E' \). Given \( X' \) with \( E' - X' \) closed in \( N_2 \), let \( Y \subseteq E \) be the set of elements whose copies comprise \( X' \). Then \( q_2(E' - X') = |E| - |X| \), and \( X' \) contains all copies of the elements of \( Y \), so \( q_1(X') = \sum r'_j(X' \cap E') = \sum r_i(Y) \). We conclude that \( r(E) = \min_{Y \subseteq E} \{ |E| - |X| + \sum r_i(Y) \} \).

To show that \( M \) is a matroid, we verify submodularity for \( r \). Given \( X, Y \subseteq E \), the formula for \( r \) yields \( U \subseteq X \) and \( V \subseteq Y \) such that
\[
r(X) = |X - U| + \sum r_i(U); \quad r(Y) = |Y - V| + \sum r_i(V).
\]
Since \( U \cup V \subseteq X \cap Y \) and \( U \cup V \subseteq X \cup Y \), we also have
\[
r(X \cap Y) \leq |(X \cap Y) - (U \cap V)| + \sum r_i(U \cap V); \quad r(X \cup Y) \leq |(X \cup Y) - (U \cup V)| + \sum r_i(U \cup V).
\]
After applying the submodularity of each \( r_i \) and the diagram below, these inequalities yield \( r(X \cap Y) + r(X \cup Y) \leq r(X) + r(Y) \).

\[
\begin{tikzpicture}
    \node (X) at (0,0) [circle,draw] {X};
    \node (Y) at (1,0) [circle,draw] {Y};
    \node (U) at (0.5,-0.5) [circle,draw] {U};
    \node (V) at (1.5,-0.5) [circle,draw] {V};
    \draw[->] (X) to node [above] {1} (U);
    \draw[->] (Y) to node [above] {0} (V);
    \draw[->] (X) to node [right] {0} (V);
    \draw[->] (Y) to node [right] {1} (U);
\end{tikzpicture}
\]

\[ |(X \cap Y) - (U \cap V)| + |(X \cup Y) - (U \cup V)| = |X - U| + |Y - V| \]

In applying the Matroid Intersection Theorem, we needed \( N_i \) to be a matroid, which required \( \{M_i\} \) to be matroids. Hence this rank formula does not apply for unions of arbitrary hereditary systems.

The Matroid Union Theorem yields short proofs of formulas for packing and covering problems. In each formula below, the optimal subset is closed, because switching from \( X \) to \( \sigma(X) \) improves the numerator without changing the denominator. The graph corollaries were originally proved by difficult ad hoc arguments.

11.3.53. Corollary. (Matroid Covering Theorem; Edmonds [1965a]) In a loopless matroid \( M \) on \( E \), the minimum number of independent sets whose union is \( E \) is 
\[
\max_{\sigma \neq X \subseteq E} \left[ \frac{|X|}{r(X)} \right].
\]

Proof: Let \( M_1, \ldots, M_k \) be copies of \( M \) on \( E \). The set \( E \) is the union of \( k \) independent sets in \( M \) if and only if \( E \) is independent in \( M' = M_1 \cup \cdots \cup M_k \). By the Matroid Union Theorem, \( r'(E) \geq |E| \) is equivalent to \( |E| - |Y| + \sum r_i(Y) \geq |E| \) for all \( Y \subseteq E \). Since \( r_i(Y) = r(Y) \) for all \( i \), we conclude that \( E \) is the union of \( k \) independent sets if and only if \( kr(Y) \geq |Y| \) for all \( Y \subseteq E \).

11.3.54. Corollary. (Nash-Williams [1964]) The number of forests needed to cover the edges of a graph \( G \) (its arboricity) is 
\[
\max_{H \subseteq G : |V(H)| \geq 2} \left[ \frac{|E(H)|}{|V(H)| - 1} \right].
\]

Proof: This follows immediately by applying Corollary 11.3.53 to \( M(G) \). The best lower bound arises from a connected induced subgraph \( H \) (corresponding to a closed set in \( M(G) \)).

11.3.55. Corollary. (Matroid Packing Theorem; Edmonds [1965b]) Given a matroid \( M \) on \( E \), the maximum number of pairwise disjoint bases equals 
\[
\min_{X : r(X) \leq r(E)} \left[ \frac{|E| - |X|}{r(E) - r(X)} \right].
\]

Proof: The set \( E \) contains \( k \) disjoint bases if and only if \( r'(E) \geq kr(E) \) in the union \( M' \) of \( k \) matroids \( M_1, \ldots, M_k \) that are copies of \( M \) on \( E \). By the Matroid Union Theorem, this requires \( |E| - |Y| + \sum r_i(Y) \geq kr(E) \) for all \( Y \subseteq E \). Since \( r_i(Y) = r(Y) \) for all \( i \), we conclude that \( k \) disjoint bases exist if and only if \( |E| - |Y| \geq k(r(e) - r(Y)) \) for all \( Y \subseteq E \).

11.3.56. Corollary. (Nash-Williams [1961], Tutte [1961]) A graph \( G \) has \( k \) edge-disjoint spanning trees if and only if, for every partition \( P \) of \( V(G) \), at least \( k(|P| - 1) \) edges have endpoints in different parts.
**Proof:** (Edmonds [1965b]) We may assume that $G$ is connected. By applying Corollary 11.3.55 to $M(G)$, we must determine when $|E| - |X| \geq k(r(E) - r(X))$ for each closed set $X$. The closed sets correspond to partitions of $V(G)$ into vertex sets inducing connected subgraphs. For each such partition $V_1, \ldots, V_p$, the corresponding closed set $X$ is $\bigcup E(G[V_i])$ with rank $n - p$. Since $|E| - |X|$ counts the edges between parts of the partition and $r(E) - r(X) \geq p - 1$, the graph has $k$ disjoint spanning trees if and only if the condition holds.

Corollary 11.3.56 implies that every $2k$-edge-connected graph has $k$ edge-disjoint spanning trees, and this is sharp (Exercise 62). Nash-Williams and Tutte proved Corollary 11.3.56 independently, by different methods, a few months apart, and the two papers were published in the same volume of the same journal.

**EXERCISES 11.3**

11.3.1. (−) Let $M$ be the hereditary system on $[4]$ whose bases are $\{1, 2\}$ and $\{3, 4\}$. Show that all the properties listed in Definition 11.3.18 fail for $M$.

11.3.2. (−) For each family $C$ below, determine whether it is the family of circuits of a hereditary system on $[6]$. If it is, determine whether the system is a graphic matroid.
   (a) $C = \{\{1, 2, 3\}, \{3, 4, 5\}, \{5, 6, 1\}\}$.
   (b) $C = \{\{1, 2, 3\}, \{3, 4, 5\}, \{1, 2, 4, 5\}\}$.

11.3.3. (−) Characterize the $G$ graphs that yield matroids as follows:
   (a) The stable sets of $G$ are the independent sets of a matroid on $V(G)$.
   (b) The matchings in $G$ are the independent sets of a matroid on $E(G)$.

11.3.4. (−) Show that the stable sets of a graph need not be the independent sets of a matroid by finding vertex-weighted graphs where the ratio between the set found by the greedy algorithm and the maximum weight of a stable set is arbitrarily large.

11.3.5. (−) Explain how to obtain the rank function of a hereditary system from the bases.

11.3.6. (−) Let $B$ be a base of a matroid. For $e \in \mathcal{B}$, prove that $B - f + e$ is a base if and only if $f$ belongs to the circuit formed by adding $e$ to $B$.

11.3.7. (−) Let $e$ be an element of a circuit $C$ in a matroid. Prove that $C$ is the unique circuit created by adding $e$ to some base.

11.3.8. (−) Prove the following implications directly for hereditary systems.
   (a) The augmentation (I) and uniformity (U) properties are equivalent.
   (b) The uniformity (U) and base exchange (B) properties are equivalent.
   (c) Submodularity (R) implies weak absorption (A).
   (d) Strong absorption (A') implies base exchange (B).
   (e) Augmentation (I) implies weak elimination (C).

11.3.9. (−) Prove that if $r(X) = r(X \cap Y)$ for some $X, Y \subseteq E$ in a matroid on $E$, then $r(X \cup Y) = r(Y)$.

11.3.10. (−) A set of $|E| - r(E)$ circuits of a matroid on $E$ form a fundamental set of circuits if the elements $e_1, \ldots, e_n$ can be ordered so that the last element of $C_i$ is $e_{r(E) + i}$. Prove that every matroid has a fundamental set of circuits. (Whitney [1935])

11.3.11. (−) Describe the circuits of a partition matroid $M$. Use this description to prove directly that partition matroids satisfies the weak elimination property.
11.3.12. (-) Prove that every partition matroid is a transversal matroid.

11.3.13. (-) Determine whether the cycle matroid of $G$ below is a transversal matroid.

```
  a
  /|
  | b
  / |
  | c
  /  |
  d  e
    f
  g
```

11.3.14. (-) Let $B_1$ and $B_2$ be bases of a matroid such that $|B_1 \Delta B_2| = 2$. Prove that there is a unique circuit $C$ such that $B_1 \Delta B_2 \subseteq C \subseteq B_1 \cup B_2$.

11.3.15. (-) Prove that the cycle matroid of a graph $G$ is the column matroid over $\mathbb{Z}_2$ of the vertex-edge incidence matrix of $G$.

11.3.16. (-) Let $B_1$ and $B_2$ be bases of a matroid such that $|B_1 \triangle B_2| = 2$. Prove that there is a unique circuit $C$ such that $B_1 \triangle B_2 \subseteq C \subseteq B_1 \cup B_2$.

11.3.17. (-) Let $M$ be the hereditary system on $E(K_n)$ whose independent sets are the edge sets of planar graphs. Determine whether $M$ is a matroid.

11.3.18. (-) Determine whether a set can be a circuit and a cocircuit in the same matroid.

11.3.19. (-) Let $C$ and $C^*$ be a circuit and a cocircuit in a matroid on $n$ elements. Determine the minimum and maximum possible values of $|C| + |C^*|$.

11.3.20. (-) Let $M$ be a matroid on $E$, and fix $A \subseteq E$. Let $I'$ be the family of sets $X \subseteq E$ such that $X \in I$ and $X \cap A = \emptyset$. Prove that $I'$ is the family of independent sets of a matroid.

11.3.21. (-) Let $r$ and $\sigma$ be the rank function and span function of a matroid. Prove that $r(X) = \min \{ |Y| : Y \subseteq X \text{ and } \sigma(Y) = \sigma(X) \}$.

11.3.22. (-) Prove that a matroid of rank $r$ has at least $2^r$ closed sets. (Lazarson [1957])

11.3.23. (-) Let $G$ be an $n$-vertex graph, and let $E_1, \ldots, E_{n-1}$ be a partition of $E(G)$ into $n-1$ sets. Show that matroids can be used to test whether $G$ has a spanning tree with exactly one edge in each subset $E_i$.

11.3.24. (-) Given matroids $M_1, \ldots, M_k$ on $E$, the Matroid Partition Problem asks whether an input set $X \subseteq E$ partitions into sets $I_1, \ldots, I_k$ with $I_i \in I$. Prove that $X$ is partitionable if and only if $|X - Y| + \sum r(Y) \geq |X|$ for all $Y \subseteq X$.

11.3.25. (-) Use HAMILTONIAN PATH in directed graphs to prove that 3-MATROID INTERSECTION is NP-complete.

11.3.26. (-) Use 3-D MATCHING to prove that 3-MATROID INTERSECTION is NP-complete. Given disjoint sets $V_1, V_2, V_3$ and a family of triples that each consists of one element from each $V_i$, 3-D MATCHING is the problem of finding the maximum number of triples that together use each element at most once. (In this terminology, ordinary bipartite matching is 2-D MATCHING.)

11.3.27. Graphic matroids.
   (a) Determine which graphic matroids are also uniform matroids.
   (b) Determine which graphic matroids are also partition matroids.

11.3.28. (∗) Let $E$ be the edge set of a graph $G$. Say that a set $X \subseteq E$ is weakly acyclic if the spanning subgraph of $G$ with edge set $X$ has at most one cycle. Prove that the weakly acyclic sets are the independent sets of a matroid on $E$. 


11.3.29. Using only linear dependence, prove that vector matroids satisfy the induced circuit property: adding an element to a linearly independent set of vectors creates at most one minimal dependent set.

11.3.30. (♦) Submodularity for cycle matroids. Given a graph $G$, let $k(X)$ be the number of components of the spanning subgraph $G_X$ with edge set $X$. Let $U$ and $V$ be the sets of components of $G_X$ and $G_Y$, respectively. Let $H$ be the $U$, $V$-bigraph with $U$, adjacent to $V$, if $U$, $V$ in $G_X$ and $V$, in $G_Y$ have a common vertex. Relate the numbers of vertices, components, and edges in $H$ to the numbers $k(X)$, $k(Y)$, $k(X \cap Y)$, $k(X \cup Y)$, and conclude directly that the rank function of the cycle matroid of $G$ is submodular. (Aigner [1979])

11.3.31. Let $s$ and $t$ be vertices in a digraph $D$. Let $E = V(D) - \{s, t\}$. For $X \subseteq E$, let $r(X)$ be the number of edges from $s \cup X$ to $X \cup t$. Prove that $r$ is submodular.

11.3.32. (♦) Prove the following implications directly for hereditary systems.
   (a) The base exchange property (B) implies the augmentation property (I).
   (b) The augmentation property (I) implies the absorption property (A).
   (c) The strong absorption property (A') implies the submodularity property (R) (without using uniformity).

11.3.33. (♦) Prove the following implications directly for hereditary systems.
   (a) The base exchange property (B) implies the induced circuit property (J).
   (b) The induced circuit property (J) implies the augmentation property (I).
   (c) The induced circuit property implies the weak elimination property (C).

11.3.34. Without using other characterizations of matroids, prove directly that the base exchange and dual base exchange properties are equivalent.

11.3.35. Prove that a hereditary system is a matroid if and only if it satisfies the following “ultra-weak” augmentation property: If $I_1, I_2 \in I$ with $|I_2| > |I_1|$ and $|I_1 - I_2| = 1$, then $I_1 + e \in I$ for some $e \in I_2 - I_1$. (Chappell)

11.3.36. Given a matroid on a set $E$, let $C_1, \ldots, C_k$ be circuits such that none is contained in the union of the others. Let $X$ be a subset of $E$ with $|X| < k$. Prove that $\bigcup_{i=1}^k C_i - X$ contains a circuit. (Welsh [1979])

11.3.37. (♦) Let $e$ be an element in a matroid $M$. Prove that if $C$ is a circuit in $M \cdot e$, then $C$ or $C + e$ is a circuit in $M$.

11.3.38. A refinement of a matroid $M$ is a matroid $N$ on the same elements such that every circuit of $M$ is a circuit of $N$. Prove that $M$ has a refinement $N$ different from $M$ if and only if no circuit of $M$ has size $r(M) + 1$.

11.3.39. (♦) Let $B_1$ and $B_2$ be bases of a matroid $M$.
   (a) Prove that for each $e \in B_1$, there exists $f \in B_2$ such that $B_1 - e + f$ and $B_2 - f + e$ are bases of $M$. (Brualdi [1969])
   (b) Use the cycle matroid $M(K_4)$ to show that there may be no bijection $\pi$: $B_1 \rightarrow B_2$ such that setting $f = \pi(e)$ satisfies part (a) for all $e \in B_1$.
   (c) Given $X_1 \subseteq B_1$, prove that for some $X_2 \subseteq B_2$ both $(B_1 - X_1) \cup X_2$ and $(B_2 - X_2) \cup X_1$ are bases. (Brylawski [1973], Greene [1973], Woodall [1974], Greene-Magnanti [1975])

11.3.40. (♦) Let $B_1$ and $B_2$ be bases of a matroid $M$. Prove that there is a bijection $\pi$: $B_1 \rightarrow B_2$ such that for each $e \in B_1$, the set $B_2 - \pi(e) + e$ is a base of $M$. (Brualdi [1969]) (Hint: Define a $B_1, B_2$-bigraph by making $e \in B_1$ and $f \in B_2$ adjacent when $B_2 + e - f \in B$.)

11.3.41. (♦) Let $M$ be a matroid on a set $E$, and let $w$: $E \rightarrow \mathbb{N}_0$. Use the greedy algorithm to prove a min-max formula for maximum weighted independent set: $\max_{I \in \mathcal{I}} \sum_{e \in I} w(e) = \min \sum_{e \in \mathcal{C}} r(e)$, where the minimum is over all chains (by inclusion) of sets in $E$ such that each element $e \in E$ appears in at least $w(e)$ sets in the chain (sets may repeat in chains).
11.3.42. (♦) Circuits and cocircuits. Consider a matroid $M$ and its dual $M^*$. 
(a) Dual augmentation property. Given disjoint sets $X$ and $X^*$ with $X \in I$ and $X^* \in I^*$, prove that there are disjoint sets $B$ and $B^*$ with $X \subseteq B \in B$ and $X^* \subseteq B^* \in B^*$. 
(b) Let $e$ be an element of a base $B$. Prove that $M$ has exactly one cocircuit disjoint from $B - e$ and that it contains $e$.
(c) Prove that the cocircuits of $M$ are the minimal nonempty sets $C^*$ such that $|C^* \cap C| \neq 1$ for every $C \in C$. 
(d) For distinct elements $x$ and $y$ of a circuit $C$, prove that there is a cocircuit $C^*$ such that $C^* \cap C = \{x, y\}$. (Minty [1966])

11.3.43. The $k$-truncation $M_k$ of a matroid $M$ is defined by $I_{M_k} = \{X \in I_M : |X| \leq k\}$. 
(a) Prove that $M_k$ is a matroid. 
(b) Prove that a matroid can be covered by $t$ independent sets of size at most $k$ if and only if $\max_{X \in I} \min_{k \leq r(X)} \frac{|X|}{k} \leq t$. 
(c) Prove that a matroid of rank at least $k$ with ground set $E$ has $t$ pairwise disjoint independent sets of size $k$ if and only if $\min_{X \in I} \frac{|X|}{r(X) - k} \geq t$. (Chen–Lai [1996])

11.3.44. The $k$-elongation of a matroid $M$ is the hereditary system $M^k$ whose bases are the spanning sets of $M$ with size $k$. 
(a) Prove that $M^k$ is a matroid. 
(b) Prove that $(M_k)^* = (M^*)^{k-1}$ if $k \leq r(M)$, where $M_k$ is the $k$-truncation of $M$ defined in Exercise 11.3.43. (Welsh [1979])

11.3.45. Prove that a matroid is simple (no circuits of size at most 2) if and only if (1) no element appears in every hyperplane, and (2) from every distinct pair of elements some hyperplane contains exactly one.

11.3.46. Let $M$ be a matroid on $E$. For $F \subseteq E$, prove that $C_{M/F} = \{C \subseteq F : C \in C_M\}$ and that $C_{M,F}$ is the set of minimal nonempty members of $\{C \cap F : C \in C_M\}$. 

11.3.47. By duality and matroid restriction, prove $r_{M,F}(X) = r_M(X \cup \overline{F}) - r_M(\overline{F})$. 

11.3.48. (♦) Prove that any minor of a matroid obtained by restricting and then contracting can also be obtained by contracting and then restricting. In particular, if $M$ is a matroid on $E$ and $Y \subseteq X \subseteq E$, prove that $(M|X).Y = (M[X - Y]).Y$ and $(M.X)|Y = (M[X - Y]).Y$. (Hint: Use the rank function.)

11.3.49. (♦) Prove that the properties in Theorem 11.3.40 (involving the span function) are equivalent and characterize matroids: incorporation (P), idempotence (S), transitivity of dependence (T), and strong elimination (C'). (Hint: Prove $U \Rightarrow P \Rightarrow S \Rightarrow T \Rightarrow C' \Rightarrow C$.)

11.3.50. Prove the following properties of closed sets of a matroid. 
(a) The closed sets are the complements of the unions of cocircuits. 
(b) The intersection of two closed sets is closed. 
(c) The span of a set is the intersection of all closed sets containing it. 
(d) The union of two closed sets need not be a closed set.

11.3.51. (+) Prove directly that in a hereditary system, the weak elimination property implies the strong elimination property, using induction on $|C_1 \cup C_2|$. (Lehman [1964])

11.3.52. Given a matroid $M$ on $E$ and $e \in E$, the Shannon switching game $(M, e)$ played by Spanner and Cutter is as follows. On each round Cutter deletes an element of $E - e$, and then Spanner seizes an element of $E - e$. Spanner wants a set that spans $e$; Cutter aims to prevent this. Prove that Spanner has a winning strategy when there are disjoint subsets $X_1, X_2$ of $E - e$ such that $e \in \sigma(X_1) = \sigma(X_2)$. (Lehman [1964]) (Comment: A lengthy proof using the Matroid Union Theorem (Theorem 11.3.52) shows that this sufficient condition is also necessary. A special case using the cycle matroid for a union of two edge-disjoint trees was marketed by Hasbro under the name “Bridg-It”; see West [2001, p.74].)
11.3.53. (+) Given a matroid \( M \), the base exchange graph \( \beta(M) \) has a vertex for each base of \( M \), with two bases adjacent when their symmetric difference has size 2. Prove that \( \beta(M) \) is Hamiltonian when \( M \) has at least three bases. (Hint: Prove the stronger statement that \( \beta(M) \) has a spanning cycle through any edge) (Holzmann–Harary [1972])

11.3.54. Use the formula \( r(X) = \min_{J \subseteq [m]} |A(J) \cap X| + m - |J| \) for the rank function of the transversal matroid on \( E \) induced by subsets \( A_1, \ldots, A_m \) (Example 11.3.46) to prove directly that the rank function satisfies \( r(\emptyset) = 0 \) and \( r(X) \leq r(X + e) \leq r(X) + 1 \).

11.3.55. (∗) Given a bipartite graph \( G \) with \( E \) being one part, let \( M \) be the transversal matroid on \( E \) whose independent sets are the subsets of \( E \) that can be covered by matchings in \( G \). By Ore’s Defect Formula (see Example 11.3.46), \( r(X) = \min_{S \subseteq X} (|N(S)| + |X - S|) \). Prove directly that \( r \) is submodular.

11.3.56. Prove that restrictions and unions of transversal matroids are transversal matroids, but contractions and duals of transversal matroids need not be.

11.3.57. (∗) Let \( M \) be the transversal matroid on \( E \) induced by subsets \( A_1, \ldots, A_m \). Use the Matroid Union Theorem to prove \( r_M(X) = \min_{Y \subseteq X} |X - Y| + |N(Y)| \).

11.3.58. (∗) Common independent and spanning sets.
(a) For matroids \( M_1 \) and \( M_2 \) on a set \( E \), prove \(|I| + |S| = r_1(E) + r_2(E)\), where \( I \) is a largest common independent set and \( S \) is a smallest common spanning set.
(b) Let \( G \) be an \( n \)-vertex bipartite graph with no isolated vertices. Prove \( a'(G) + \beta'(G) = n \) (Gallai’s Theorem, Exercise 6.1.50).
(c) Let \( G \) be a bipartite graph with no isolated vertices. Without using other results, use part (a) directly to prove \( a(G) = \beta'(G) \) (König’s Other Theorem, Theorem 6.1.16).

11.3.59. (∗) Use the Matroid Intersection Theorem to prove that in every acyclic digraph, the vertices can be covered by at most \( k \) pairwise disjoint paths, where \( k \) is the independence number of the underlying graph. (Comment: This is the special case for acyclic digraphs of the Gallai–Milgram Theorem (Theorem 12.1.21).)

11.3.60. Given matroids \( M_1 \) and \( M_2 \) whose families of spanning sets are \( S_1 \) and \( S_2 \), prove that the matroid \( (M_1^* \cup M_2^*)^* \) is the hereditary system whose spanning sets are \( \{X_1 \cap X_2: X_1 \in S_1, X_2 \in S_2\} \).

11.3.61. Matroid Intersection from Matroid Union.
(a) Without the Matroid Intersection Theorem, prove that the maximum size of a common independent set in matroids \( M_1 \) and \( M_2 \) on \( E \) is \( r_{M_1 \cup M_2}(E) - r_{M_1^*}(E) \).
(b) Prove the Matroid Intersection Theorem by applying Matroid Union to \( M_1 \cup M_2^* \).

11.3.62. (∗) Connectivity and spanning trees.
(a) Let \( F \) be a set of at most \( k \) edges in a 2\( k \)-edge-connected graph \( G \). Use Corollary 11.3.56 to prove that \( G - F \) has \( k \) edge-disjoint spanning trees. (Nash-Williams [1961])
(b) For each \( k \), construct a \((2k - 1)\)-edge-connected graph that does not have \( k \) edge-disjoint spanning trees.

11.3.63. Colored trees and \( b \)-detachments.
(a) Let \( G \) be a connected edge-colored graph with color classes \( E_1, \ldots, E_k \). Prove that \( G \) has a spanning tree with distinct colors if and only if \( G - F \) has at most \( t + 1 \) components whenever \( F \) consists of \( t \) color classes. (Hint: Use the Matroid Intersection Theorem.)
(b) A split replaces a vertex \( x \) with two new vertices \( x_1 \) and \( x_2 \) whose neighborhoods in the new graph partition \( N(x) \). Given \( b: V(G) \to \mathbb{N} \), a \( b \)-detachment of a graph \( G \) arises by performing splits until there are \( b(v) \) copies of each vertex \( v \). Use part (a) to prove that \( G \) has a connected \( b \)-detachment if and only if for all \( U \subseteq V(G) \), graph \( G - U \) has at most \( f(U) + 1 - b(U) \) components, where \( b(U) = \sum_{v \in U} b(v) \) and \( f(U) \) is the number of edges incident to \( U \). (Schrijver [2003, p704]; see Nash-Williams [1985] for a more general result.)