Material

Tibor Szabó
Summer 2010 — Combinatorics

This is a regularly updated version of the material. There will be varying degree of detail (partly depending on the material available in a book).

Lecture 1. (Basic Counting) References: Aigner, L-P-V, Brualdi, vL-W

What do we count?

Sets, functions (injective, surjective, bijective), sequences, vectors

We mostly consider **finite** objects.

Number of subsets of an n-element set

Basic Principles of Counting

- Multiplication Principle ($|K_1 \times \cdots \times K_n| = |K_1| \cdot \cdots \cdot |K_n|$, Ex: Hungarian Identification Number: GYYMMDDXXXC (checksum digit); Generalization: Ex: number of 4-digit numbers with no two identical digit next to each other, number of injections, permutations)

- Addition Principle (in other words: case distinction) (Ex: number of possible passwords (six to eight digits chosen from 26 upper or lower case English letters, 10 numbers, and 6 special characters with at least one special character mandatory))

- Bijection Principle (or "Double Counting") (Reprove number of subsets with encoding subsets into bitstrings)

Basic Problems:

Number of subsets of a given size (number of poker hands (5 out of deck of 52), probability of winning the lottery (5 numbers out of 90)), Binomial Coefficients, Binomial Theorem, Pascal's Triangle, Identities: $\binom{n}{k} = \binom{n}{n-k}$, $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ (Proof both by combinatorial and algebraic (manipulation) way), $2^n = \sum_{i=0}^{n} \binom{n}{i}$, $\sum_{i=0}^{n} (-1)^i \binom{n}{i} = 0$ (proof by substitution to Binomial Theorem and by combinatorics (bijection)). $\sum_{i=0}^{n} \binom{n}{i}^2 = \binom{2n}{n}$ (When finding the correct generalization brings the proof on a plate: $\binom{m+n}{r} = \sum_{i=0}^{r} \binom{m}{i}\binom{n}{r-i}$. (Vandermonde Identity))

Estimating factorial, binomial coefficients.

Stirling's Formula:

$$n! \approx \frac{n^n}{e^n}\sqrt{2\pi n}, \text{ that is, } \frac{n!}{\frac{n^n}{e^n}\sqrt{2\pi n}} \to 1.$$

Always true:

$$\left(\frac{n}{k}\right)^k \le \binom{n}{k} \le \frac{n^k}{k!}$$

Upper bound is asymptotically tight when k is a constant.

Always true:

$$\binom{n}{k} \leq e^k \left(\frac{n}{k}\right)^k,$$

is a very good estimate when $k = o(n)$ and $k = \omega(1)$.
(Definitions of $f = o(g) \Leftrightarrow g = \omega(f)$ and $f = O(g) \Leftrightarrow g = \Omega(f)$ and $f = \Theta(g)$)

Finally, when $k = \alpha n$ with α constant, then $\binom{n}{k} = 2^{H(\alpha)n(1+o(1))}$, where $H(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)$ is the binary entropy function (but is constant when α is constant).

For the middle (which is the largest) binomial coefficient, Stirling's formula gives

$$\binom{n}{n/2} \approx \sqrt{\frac{2}{\pi}} \frac{2^n}{\sqrt{n}}.$$

(In this case $H(1/2) = 1$.)

Anagrams ("Permutations of multisets") Number of permutations of $k_1 + \cdots k_r$ objects: k_1 objects a_1, k_2 objects a_2, etc ... is

$$\frac{(k_1 + \cdots k_r)!}{k_1! \cdot \cdots \cdot k_r!}.$$

Distributing pennies (n indistinguishable pennies to k distinguishable children, so each gets at least one.) Answer: $\binom{n-1}{k-1}$. Proof: put pennies in a row and color red the first penny (from left) given to each child. The first penny (from left) must be colored: it must be given to the first child. $k-1$ other pennies must be colored from the remaining $n-1$.)

Inclusion-Exclusion (for one set (recall passwords which must contain a special character), for two sets $A \cup B = |A| + |B| - |A \cap B|$, for three sets (example with students loving M, CS, PH),
general statement (to express the cardinality of the union of k sets): Applications: number of surjections (Corollaries: identities about binomial coefficients), derangements (probablity of derangement is roughly $1/e$), Euler's φ function)

Proof techniques:

- Pigeonhole Principle (Can we find two Berliners with the same number of hairs? Yes. Indirect Proof! How many points should the final be worth so that there are two students in the class with the same score?)

Lecture 2. (Pigeonhole Principle, Ramsey's Theorem, van der Waerden's Theorem)

References: Aigner and/or L-P-V and G-R-S

"Of three ordinary people two must have the same sex." (Kleitman),
Among 367 people there must be two with the same birthday; Twin-paradox
Generalized Pigeonhole Principle: If there are n pigeons in k pigeonholes then there is a hole with at least $\lceil n/k \rceil$ pigeons. (tool to prove existence, in other words "averaging").
$1 \le a_1, \ldots, a_{n+1} \le 2n \Rightarrow$ there are $i \neq j$ $a_i | a_j$.
Proposition. Each sequence of length n contains a monotone subsequence of length $\sqrt{n}$. (Example: $8, 11, 9, 1, 4, 6, 12, 10, 5, 7$ contains four increasing and one decreasing sequence of length four), (Hint for HW problem about candy-eating child)
Ramsey's Theorem in party of 6, Definition of Ramsey number

$$R(k,l) = \min\{N : \ \forall c : E(K_N) \to \{\texttt{blue},\texttt{red}\} \exists \ \texttt{red} K_k \text{ or } \texttt{blue} K_l \ \}$$

HW: $R(4,3) = 9$
$R(4,4) = 18$

Definition of *Paley-graph* P_p, for primes $p \equiv 1 \pmod 4$ (this congruence assumption is needed only to make the definition of an edge symmetric):

Vertex set $V(P_p) = \mathbb{F}_p$ (field of p elements).
Edge set $E(P_p) = \{xy : x - y \in Q_p\}$, where $Q_p = \{z^2 : z \in \mathbb{F}_p\}$ is the set of *quadratic residues* modulo p.

In a Paley-graph every vertex has $(p-1)/2$ neighbors, since $|Q_p| = (p-1)/2$.

P_5 is the 5-cycle. It does not contain a K_3 and no $\bar{K}_3$. This example, together with the "party of 6"-proposition proves that $R(3,3) = 6$.

HW: P_{17} does not contain a K_4 and no $\bar{K}_4$.

For a *Paley-coloring* just color a pair xy with `red` if $x - y \in Q_p$, otherwise `blue`.

Theorem. $R(k,l) \le R(k, l-1) + R(k-1, l)$

Proof: Take $N = R(k, l-1) + R(k-1, l)$ and an arbitrary `red`/`blue` coloring of $E(K_N)$. Pick an arbitrary vertex $x \in V$.
Case 1: x has at least $R(k-1, l)$ `red` neighbors
Case 2: x has at least $R(k, l-1)$ `blue` neighbors
One of these cases happens. (This is the Even More General Pigeonhole Principle: If there are $n_1 + n_2 + \ldots + n_k - k + 1$ pigeons then there is an index i, such that box i contains at least n_i pigeons. Indirect proof.)
In Case 1: if there is a `red` K_{k-1} among the red neighbors of x, then together with x they form a `red` K_k, done. Otherwise there is a `blue` K_l among the red neighbors of x and we are also done.
Case 2 is analogous: if there is a `blue` K_{l-1} among the `blue` neighbors of x, then

together with x they form a `blue` K_l, done. Otherwise there is a `red` K_k among the `blue` neighbors of x and we are also done. □

Corollary For all $k, l \geq 1$, $R(k,l) \leq \binom{k+l-2}{k-1}$. In particular, $R(k,l)$ exists.

Proof. Induction on $k+l$.

Corollary $R(k,k) \leq 4^k$.

The finiteness of $R(k,l)$ is called *Ramsey's Theorem.* The proof above with the estimate is due to Erdős and Szekeres.

How about a lower bound?

A set of vertices of a graph G is called a *homogenous set* of G if it is a clique or an independent set. A graph with no homogenous set of order k is called a *k-Ramsey graph*.

How good Ramsey-graphs are the Paley-graph? It is not known. Numerical data suggests that the largest clique (and hence also independent set) might be much smaller than the square root of the number of vertices. For example, for $p = 6997$ the clique number is only 17. (Shearer) It is only known that the largest clique and independent set is of the order $\sqrt{p}$. This gives $R(k,k) = \Omega(k^2)$.

But for this we have the more trivial Turán-coloring: Partition $(k-1)^2$ vertices into parts of size $k-1$ and color each edge within parts by `red` and edges between parts with `blue`. The largest m.c. clique has size $k-1$, proving $R(k,k) \geq (k-1)^2+1$. Pretty weak considering that the upper bound is exponential.

Is there something better?

Theorem (Erdős) $R(k,k) \geq \sqrt{2}^k$

Proof. Crude counting; count "bad" two-colorings. Only proves existence.

Nobody is able to *construct* explicitly k-Ramsey-graphs with 1.000001^k vertices. One needs a quite unexpected idea even to construct something better than k^2 vertices. (We will come back to this question later in the semester when we discuss the Linear Algebra method.)

\$ 1000 dollar question: Determine $\lim_{k\to\infty} \sqrt[k]{R(k,k)}$. (Currently it is not even known that this limit exists. (\$ 500))

Application: (HW)

Theorem. Color the integers with r colors. Prove that there are three numbers of the same color, such that one is equal to the sum of the other two.

Another question about patterns in two-colored integers. Is it unavoidable to have a monochromatic (*m.c.*) arithmetic progression of length 3 (a 3-*AP*)?
YES.

Let $c : [N] \to \{\texttt{red}, \texttt{blue}\}$ be an arbitrary two-coloring of the first N integers with no m.c. 3-AP. (We do not specify N now, only at the very end. We work under the assumption that it is "large enough".) In any *block* of five consecutive integers $y, y+1, y+2, y+3, y+4$ we find a triple of integers forming a 3-AP, such that the color of the first two is the same, while the third one (of course) has the opposite color. (If $c(y) = c(y+1)$ then $y, y+1, y+2$ is such a triple, if $c(y) = c(y+2)$ then $y, y+2, y+4$ is such a triple, otherwise $y+1, y+2, y+3$ is such a triple.) Let us consider the first $5 \cdot (2^5+1)$ integers as the union of $2^5 + 1 = 33$ disjoint blocks of fives. Each of these blocks can have one of the 2^5 coloring patterns on it. By the Pigeonhole Principle (PP), two of these 33 blocks have identical coloring pattern. In these two blocks we have two 3-APs: a_1, a_1+d, a_1+2d in the first block and a_2, a_2+d, a_2+2d in the second one, so that $c(a_1) = c(a_1+d) = c(a_2) = c(a_2+d)$ (by symmetry we can assume that this color is `red`) , and $c(a_1+2d) = c(a_2+2d)$ is of the opposite color, that is, `blue`. But then what is the color of $z = a_1 + 2d + 2d'$? (Here we denote by $d' = a_2 - a_1$.) If $c(z)$ is `blue` then $a_1 + 2d, a_2 + 2d, z$ is a `blue` 3-AP (with difference d'). If $c(z)$ is `red` then $a_1, a_2 + d, z$ is a `red` 3-AP (with difference $d' + d$). So we proved that there is a m.c. 3-AP if $N = 5 \cdot (2^5 + 1 + 2^5) = 325$.

Proposition. For any two-coloring of [325] there is a m.c. 3-AP.

Remark. BTW The tight answer is: $N = 9$ integers are enough.

Can we find a m.c. 4-AP if N is even larger enough? YES. How large should N be?

Let $c : [N] \to \{\texttt{red}, \texttt{blue}\}$ be a coloring with no m.c. 4-AP. How large blocks should we consider to use the previous idea? Well, we know that within 325 consecutive integers there are three forming a m.c. 3-AP. The extension of this 3-AP to a 4-AP is within the next 162 integers. So any block of 487 consecutive integers contains a 4-AP $a_1, a_1+d, a_1+2d, a_1+3d$, such that $c(a_1) = c(a_1+d) = c(a_1+2d)$ and (of course) $c(a_1+3d)$ is of the opposite color. IF (and it's a big IF) we were able to find a 3-AP of blocks having the same coloring pattern, we would be DONE. (Indeed: then we would have three 4-APs $a_1, a_1+d, a_1+2d, a_1+3d$, $a_2, a_2+d, a_2+2d, a_2+3d$, $a_3, a_3+d, a_3+2d, a_3+3d$, such that $c(a_1) = c(a_1+d) = c(a_1+2d) = c(a_2) = c(a_2+d) = c(a_2+2d) = c(a_3) = c(a_3+d) = c(a_3+2d)$, say `red`, $c(a_1+3d) = c(a_2+3d) = c(a_3+3d)$ is the opposite color `blue` and a_1, a_2, a_3 forming an 3-AP (say with difference d'). Then the integer $z = a_1 + 3d + 3d'$ will be the fourth member of a m.c. AP (which either starts at a_1 (if its color is `red`) or at $a_1 + 3d$ (if its color is `blue`)).)

So how do we find this 3-AP of blocks having the same coloring pattern?? It was so easy in the previous proof, when we just needed to find two (one can say, a 2-AP of) blocks with the same coloring pattern: we just used the PP. It turns out that we must do the same here except the PP-use is in a bit more complex setting. Consider each block (of length 487) as one entity and each of its 2^{487} possible coloring patterns as one possible color of this "entity" and try to find a m.c. 3-AP in this setup. Hence, it seems that to find a m.c. 4-AP we need to extend first the above Proposition to arbitrary number of colors.

Theorem. For any r there is a number $W = W(3, r)$, such that no matter how we color the first W integers with r colors, there will be a m.c. 3-AP.

Remark. By the above, we can then use this Theorem to find a m.c. 4-AP in any two-coloring of the first $487 \cdot W(3, 2^{487})$ integers. (This is an admittedly weak bound, it would be enough to two-color 35 integers. But while this proof generalizes to arbitrary number of colors and length of APs, the 35 bound is ad hoc.)

Proof of Thm. Induction on r. For the base case we can take $r = 2$ which is just our Proposition. We prove first the statement for $r = 3$ to see better the pattern. Let us have a 3-coloring $c : [N] \to \{\texttt{red}, \texttt{blue}, \texttt{yellow}\}$ with no m.c. 3-AP. In any block of 4 consecutive integers we find two identically colored, so in any block of 7 integers we find a 3-AP a_1, a_1+d, a_1+2d, such that $c(a_1) = c(a_1+d)$ and (of course) $c(a_1 + 2d)$ is different from the color of the other two. Taking $3^7 + 1$ consecutive disjoint blocks of 7 integers, we find two that have identical coloring pattern. Hence there are two arithmetic progressions $a_1, a_1 + d, a_1 + 2d$ and $a_2, a_2 + d, a_2 + 2d$, such that $c(a_1) = c(a_1 + d) = c(a_2) = c(a_2 + d)$, say is `red`, and $c(a_1 + 2d) = c(a_1 + 2d)$ is NOT `red`, say is `blue`. Then the integer $z = a_1 + 2d + 2d'$ (where $a_2 = a_1 + d'$) does not have color `red` (because of the 3-AP $a_1, a_2 + d, z$) and it does not have color `blue` (because of the 3-AP $a_1 + 2d, a_2 + 2d, z$). So z is colored `yellow`.

Hence in any block of $7 \cdot (2 \cdot 3^7 + 1)$ integers we find a_1, d, d' such that $c(a_1) = c(a_1 + d + d')$ is one color, $c(a_1 + 2d) = c(a_1 + 2d + d')$ is another color, and (of course) $c(a_1 + 2d + 2d')$ is the third color.

Let's find two blocks of $7 \cdot (2 \cdot 3^7 + 1)$ integers with identical color pattern. These surely exists if we take $3^{7\cdot(2\cdot3^7+1)} + 1$ blocks. Let the distance of these two identically colored blocks be d''.

In the first block we find a_1, d, d' such that $c(a_1) = c(a_1 + d + d')$, say of color `red`, $c(a_1 + 2d) = c(a_1 + 2d + d')$ is of another color, say `blue`, and $c(a_1 + 2d + 2d')$ is of the third color (in our setup it is assumed to be `yellow`). Since the second block has identical color pattern we also have that $c(a_1 + d'') = c(a_1 + d + d' + d'')$ is `red`, $c(a_1 + 2d + d'') = c(a_1 + 2d + d' + d'')$ is `blue`, and $c(a_1 + 2d + 2d' + d'')$ is `yellow`.

Now depending on the color of the integer $y = a_1 + 2d + 2d' + 2d''$ we have a m.c. 3-AP (the possibilities: in color `red` $a_1, a_1 + d + d' + d'', y$, in color `blue` $a_1 + 2d, a_1 + 2d + d' + d'', y$, and in color `yellow` $a_1 + 2d + 2d', a_1 + 2d + 2d' + d'', y$.) And we are done for $r = 3$ colors. I am sure I made a mistake somewhere with the numbers, but if not then clearly $W(3, 3) \leq (2 \cdot 3^{7\cdot(2\cdot3^7+1)} + 1) \cdot (7 \cdot (2 \cdot 3^7 + 1))$.

Now we just need to iterate this idea and we get a bound on $W(3, r)$, which then we can plug into the formula $487 \cdot W(3, 2^{487})$ to get an upper bound to guarantee a m.c. 4-AP in two-colored sequences.

Hmmm..... The bound is a bit wild.

The general theorem is as follows.

VanderWaerden′sTheorem For any integers $r, k \geq 1$ there is an integer $W = W(k, r)$ so for any r-coloring $c : [W] \to [r]$ of the first W integers there is a m.c. k-AP.

Proof: Analogous to the above, hopefully you got the idea. We will discuss this next time. The bound following from this proof is enoooormous. For a long time there was no primitive recursive upper bound known, until Shelah gave a proof for that. The best known bound today is due to Gowers (who got the Fields-medal partly for his work on this problem (or rather on a more general version of it)) and stands at a five times iterated exponential:

$$W(r, k) \leq 2^{2^{r^{2^{2^{k+9}}}}}.$$

Reference: R.L. Graham, B.L. Rothchild, J.H. Spencer: Ramsey Theory (Section 2.1)

Lecture 3. (Ramsey Theory and Applications, Catalan numbers,)

References: Aigner and/or L-P-V and G-R-S

VanderWaerden'sTheorem For any integers $r, k \geq 1$ $W(r,k) < \infty$

Definition $W(k,r) = \min\{N \in \mathbb{N}:$ for any r-coloring $c : [N] \to [r]$ there is a m.c. k-AP$\}$.

Proof: Let $L(r,k,l)$ be the smallest positive integer N such that for any r-coloring $c : [N] \to [r]$ there is a m.c. $(k+1)$-AP or a set of l crossing m.c. k-APs in l distinct colors.

A *set of l crossing k-APs* is a family of l k-APs with starting elements $a^{(1)}, \ldots, a^{(l)}$, and differences $d_1, \ldots, d_l$, respectively such that the $(k+1)st$ element of each of these APs is the same integer: $a^{(1)} + kd_1 = \ldots = a^{(l)} + kd_l$

We show by induction on k that $L(r,k,l) < \infty$ for every $k \geq 1$ and $r \geq l \geq 1$. Then we are done, since $W(k+1,r) \leq 2L(r,k,r)$.

Base case: $L(r,1,l) = l$ for all $l \leq r$.

Let $k \geq 2$: for $l = 1$, $L(r,k,1) \leq W(r,k) < 2L(r,k-1,r) < \infty$.
For $l \geq 2$, $L(r,k,l) < W(r^{2L(r,k,l-1)}, k)2L(r,k,l-1) < \infty$ $\square$

Proposition (Eszter Klein) Among 5 points in the plane in general position (i.e. no three on a line) there are always at least 4 in convex position.

Happy Ending Problem (Klein) Let $M(n)$ be the smallest number such that among any set of $M(n)$ points in the plane there are at least n in convex position. Is $M(n)$ finite?

$M(3) = 3$, $M(4) = 5$.

(a) there are two things that can happen to four points in general position: they are either in convex position or not.

(b) n points are in convex position iff every four element subset is in convex position. (Proof of "if" statement: take convex hull, if there is point inside, it is also contained in some triangle of an arbitrary triangulation of the convex hull: these are four points in non-convex position.)

(c) among any 5 points there are four which are not in non-convex position

Ramsey framework: (a) is a natural two-coloring of the 4-subsets of the point set (red: "convex 4-gon", blue: "non-convex 4-set")
(b) says that we want a LARGE m.c. subset in color red
(c) says we CANNOT have LARGE (size 5) m.c. subset in color blue.

We need Ramsey's Theorem in a situation when we color 4-sets instead of edges.

Definitions Graph: $G = (V,E)$ on vertex set V with edge set $E \subseteq \binom{V}{2}$
Hypergraph: on vertex set V with edge set $E \subseteq 2^V$
s-uniform hypergraph: if edge set $E \subseteq \binom{V}{s}$
Complete s-uniform hypergraph $K_k^{(s)}$ on k vertices is defined by $E(K_k^{(s)}) = \binom{[k]}{s}$.

Example. graph: 2-uniform hypergraph

Definition *s-uniform Ramsey number* $R^{(s)}(k,l)$ is the smallest integer N such that for any 2-coloring $c : \binom{[N]}{s} \to \{\texttt{red}, \texttt{blue}\}$ there exists a subset $I_r \subseteq [N]$ such that $c(J) = \texttt{red}$ for every $J \in \binom{I_r}{s}$ or there exists a subset $I_b \subseteq [N]$ such that $c(J) = \texttt{blue}$ for every $J \in \binom{I_b}{s}$

Theorem $R^{(s)}(k,l)$ is finite for every $s, k, l \geq 1$

Proof. Induction on s: $R^{(1)}(k,l) = k + l - 1$.
Let $s \geq 2$. Induction on $k + l$.
Base cases: For $k \geq l$, $l < s$, we have $R^{(s)}(k,l) = l$,
for $k \leq l$, $k < s$, we have $R^{(s)}(k,l) = k$,
for $k \geq s$, we have $R^{(s)}(k,s) = k$,
and for $l \geq s$, we have $R^{(s)}(s,l) = l$.

Let $c : \binom{[N]}{s} \to \{\texttt{red}, \texttt{blue}\}$ be a two-coloring of the s-sets. Pick an arbitrary vertex, say $N \in [N]$. Canonical projection of c on the $(s-1)$-sets of $[N-1]$: $c^* : \binom{[N-1]}{s-1} \to \{\texttt{red}, \texttt{blue}\}$ defined by $c^*(A) := c(A \cup \{N\})$.
By induction, there is a "large" subset $J_r \subset [N-1]$ such that every $(s-1)$-subset of J_r is $\texttt{red}$ or there is a "large" subset $J_b \subset [N-1]$ such that every $(s-1)$-subset of J_b is $\texttt{blue}$.
How large should "large" be?

In $|J_r|$ it would be enough to have $R^{(s)}(k-1,l)$ vertices. This would guarantee that either there is a m.c. l-subset in $\texttt{blue}$ or an m.c $(k-1)$-subset in $\texttt{red}$, which together with x would form an m.c. k-subset in $\texttt{red}$. (Remember that we are within J_r!)
The argument for $|J_b|$ is analogous.

If $N - 1 = R^{(s-1)}(R^{(s)}(k-1,l), R^{(s)}(k,l-1))$, then this certainly happens proving the finiteness of $R^{(s)}(k,l)$. □

Corollary $M(n) \leq R^{(4)}(n,5) < \infty$

Remark The best known bounds for $M(n)$ are pretty far from each other:

$$2^{n-2} + 1 \leq M(n) \lessapprox \frac{4^n}{\sqrt{n}}.$$

The lower bound is conjectured to be tight by Erdős and Szekeres. It is proven to be tight for $n = 3, 4, 5, 6$.

Catalan Numbers

How many ways can we triangulate a convex n-gon? T_n

$T_3 = 1, T_4 = 2, T_5 = 5$

How many ways can we fully parenthesize the expression $x_1 \cdot \dots \cdot x_n$ (having n factors)? P_n

$P_2 = 1, P_3 = 2, P_4 = 5$

How many nondecreasing lattice path are there from $(0,0)$ to (n,n) which do not go above the diagonal? C_n

$C_1 = 1, C_2 = 2, C_3 = 5$

Coincidence? NO

Theorem $T_{n+2} = P_{n+1} = C_n$ for every $n \geq 1$.

Proof. Prove recurrence and use induction.

- triangulation: fix an edge e of the convex hull and count according to the third vertex of the triangle e participates in

- parentheses: count according to after which variable the very first left parenthesis is closed

- monotone path: count according to where does the path returns to the diagonal the first time.

$$C_n = \sum_{k=1}^{n} C_{k-1} C_{n-k}.$$

□

Theorem

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Proof 1: "Simplest" proof: "bad path": one that does not stay under the diagonal. Bijection between set of bad paths and set of arbitrary paths going from $(0,0)$ to $(n-1, n+1)$. (Reflect the segment of path between the first touch of $y = x+1$ and (n,n) about the line $y = x+1$.) Subtract number of bad paths from all paths: $\binom{2n}{n} - \binom{2n}{n-1}$.

Proof 2: partition the set of all paths $\mathcal{P}$ from $(0,0)$ to (n,n) into $n+1$ classes $\mathcal{P}_i$ based on "how many vertical edges are there on the path above the diagonal". Bijection between $\mathcal{P}_i$ and $\mathcal{P}_{i-1}$ for any $i, 1 \leq i \leq n$. (For a path P consider the first horizontal edge e that arrives to the diagonal $x = y$ from above. Cut P into three pieces: at the endpoints of e. Assemble new path $f(P)$ by taking the end piece, then edge e, finally the starting piece of P. Then $f(P)$ has one less upward edge above the diagonal: the image of the first such edge is not anymore above the diagonal.
The function also has an inverse: cutting a path at the two endpoints of the last horizontal edge that leaves the diagonal and exchanging the starting and ending piece.) □

Lecture 4. (Graph Theory — Basics)

Lecture 5. (Graph Theory — Extremal Problems, Trees)

Lecture 6. (Graph Theory — Matchings)

Lecture 7. (Graph Theory — Connectivity)

Lecture 8. (Graph Theory — Colorings)

References: West

Lecture 9. (Edge-colorings, Extremal graph theory)

References: Diestel, West

Lecture 10. (Roth's Theorem, Extremal combinatorics)

References: Diestel, G-R-S, vL-W

Lecture 11. (Extremal combinatorics, Linear Algebra method)

References: vL-W, Babai-Frankl

Lecture 12. (Linear Algebra method, Algorithmic method)

References: vL-W, Babai-Frankl

Lecture 13. (Algorithmic method, Menger's Theorem, Baranyai's Theorem, Probabilistic Method)

References: vL-W, Jukna

Graphs – Definition

A graph G is a pair consisting of

- a vertex set $V(G)$, and
- an edge set $E(G) \subseteq \binom{V(G)}{2}$.

If there is no confusion about the underlying graph we often just write $V = V(G)$ and $E = E(G)$.

x and y are the endpoints of edge $e = \{x, y\}$.
They are called adjacent or neighbors.
e is called incident with x and y.

A loop is an edge whose endpoints are equal.
Multiple edges have the same set of endpoints. In the definition of a "graph" we don't allow loops and multiple edges. To emphasize this, we often say "simple graph". When we do want to allow multiple edges or loops, we say multigraph.

Remarks A multigraph might have no multiple edges or loops. Every (simple) graph is a multigraph, but not every multigraph is a (simple) graph.

Special graphs

K_n is the complete graph on n vertices.

$K_{n,m}$ is the complete bipartite graph with partite sets of sizes n and m.

P_n is the path on n vertices

C_n is the cycle on n vertices

Further definitions

The degree of vertex v is the number of edges incident with v. Loops are counted twice.

A set of pairwise adjacent vertices in a graph is called a clique. A set of pairwise non-adjacent vertices in a graph is called an independent set.

A graph G is bipartite if $V(G)$ is the union of two (possibly empty) independent sets of G. These two sets are called the partite sets of G.

The complement $\overline{G}$ of a graph G is a graph with

- vertex set $V(\overline{G}) = V(G)$ and
- edge set $E(\overline{G}) = \binom{V}{2} \setminus E(G)$.

H is a subgraph of G if $V(H) \subseteq V(G)$, $E(H) \subseteq E(G)$. We write $H \subseteq G$. We also say G contains H and write $G \supseteq H$.
For a subset $S \subseteq V(G)$ define $G[S]$, the induced subgraph of G on S: $V(G[S]) = S$ and
$E(G[S]) = \{e \in E(G) :$ both endpoints are in $S\}$.

The Petersen graph

$V(P) = \binom{[5]}{2}$

$E(P) = \{\{A, B\} : A \cap B = \emptyset\}$

Properties.

- each vertex has degree 3 (i.e. P is 3-regular)
- adjacent vertices have no common neighbor
- non-adjacent vertices have exactly one common neighbor

Corollary. The girth of the Petersen graph is 5.

The girth of a graph is the length of its shortest cycle.

Isomorphism of graphs

An isomorphism of G to H is a bijection $f : V(G) \to V(H)$ such that $uv \in E(G)$ iff* $f(u)f(v) \in E(H)$. If there is an isomorphism from G to H, then we say G is isomorphic to H, denoted by $G \cong H$.

Claim. The isomorphism relation is an equivalence relation on the set of all graphs.

An isomorphism class of graphs is an equivalence class of graphs under the isomorphism relation.

Example. What are those graphs for which the adjacency relation is an equivalence relation?

Remark. labeled vs. unlabeled

“unlabeled graph” $\approx$ “isomorphism class”.

Example. What is the number of labeled and unlabeled graphs on n vertices?

*if and only if

Equivalence relation

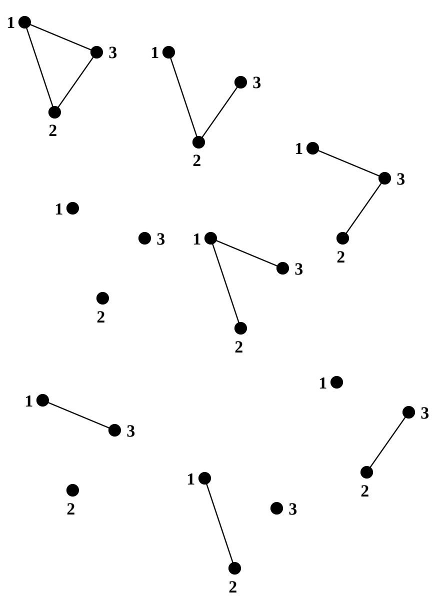
A relation on a set S is a subset of $S \times S$.

A relation R on a set S is an equivalence relation if

1. $(x, x) \in R$ (R is reflexive)
2. $(x, y) \in R$ implies $(y, x) \in R$ (R is symmetric)
3. $(x, y) \in R$ and $(y, z) \in R$ imply $(x, z) \in R$ (R is transitive)

An equivalence relation defines a partition of the base set S into equivalence classes. Elements are in relation iff they are within the same class.

Isomorphism classes



Automorphisms

An automorphism of G is an isomorphism of G to G. A graph G is vertex transitive if for every pair of vertices u, v there is an automorphism that maps u to v.

Examples.

- Automorphisms of P_4
- Automorphisms of $K_{r,s}$
- Automorphisms of Petersen graph.

A decomposition of a graph is a list of subgraphs such that each edge appears in exactly one subgraph in the list.

A graph is self-complementary if it is isomorphic to its complement.

Example. P_4, C_5

Walks, trails, paths, and cycles

A walk is an alternating list $v_0, e_1, v_1, e_2, \ldots, e_k, v_k$ of vertices and edges such that for $1 \leq i \leq k$, the edge e_i has endpoints v_{i-1} and v_i.

A trail is a walk with no repeated edge.

A path is a walk with no repeated vertex.

A u,v-walk, u,v-trail, u,v-path is a walk, trail, path, respectively, with first vertex u and last vertex v.

If $u = v$ then the u,v-walk and u,v-trail is closed. A closed trail (without specifying the first vertex) is a circuit. A circuit with no repeated vertex is called a cycle.

The length of a walk trail, path or cycle is its number of edges.

Connectivity

G is connected, if there is a u,v-path for every pair $u,v \in V(G)$ of vertices.
Otherwise G is disconnected.

Vertex u is connected to vertex v in G if there is a u,v-path. The connection relation on $V(G)$ consists of the ordered pairs (u,v) such that u is connected to v.

Claim. The connection relation is an equivalence relation.

Lemma. Every u,v-walk contains a u,v-path.

The connected components of G are its maximal connected subgraphs (i.e. the equivalence classes of the connection relation).

An isolated vertex is a vertex of degree 0. It is a connected component on its own, called trivial connected component.

Strong Induction

Theorem 1. (Principle of Induction) Let $P(n)$ be a statement with integer parameter n. If the following two conditions hold then $P(n)$ is true for each positive integer n.

1. $P(1)$ is true.
2. For all $n > 1$, “$P(n-1)$ is true” implies “$P(n)$ is true”.

Theorem 2. (Strong Principle of Induction) Let $P(n)$ be a statement with integer parameter n. If the following two conditions hold then $P(n)$ is true for each positive integer n.

1. $P(1)$ is true.
2. For all $n > 1$, “$P(k)$ is true for $1 \leq k < n$” implies “$P(n)$ is true”.

Cutting a graph

A cut-edge or cut-vertex of G is an edge or a vertex whose deletion increases the number of components.

If $M \subseteq E(G)$, then $G - M$ denotes the graph obtained from G by the deletion of the elements of M; $V(G-M) = V(G)$ and $E(G-M) = E(G) \setminus M$.
Similarly, for $S \subseteq V(G)$, $G - S$ obtained from G by the deletion of S and all edges incident with a vertex from S.
For $e \in E(G)$, $G - \{e\}$ is abbreviated by $G - e$.
For $v \in E(G)$, $G - \{v\}$ is abbreviated by $G - v$.

Proposition. An edge e is a cut-edge iff it does not belong to a cycle.

Bipartite graphs

A bipartition of G is a specification of two disjoint independent sets in G whose union is $V(G)$.

Theorem. (König, 1936) A multigraph G is bipartite iff G does not contain an odd cycle.

Proof.

$\Rightarrow$ Easy.

$\Leftarrow$ Fix a vertex $v \in V(G)$. Define sets

$$A := \{w \in V(G) : \exists \text{ an odd } v, w\text{-path } \}$$

$$B := \{w \in V(G) : \exists \text{ an even } v, w\text{-path } \}$$

Prove that A and B form a bipartition.

Lemma. Every closed odd walk contains an odd cycle.

Proof. Strong induction.

Eulerian circuits

A multigraph is Eulerian if it has a closed trail containing all its edges. A multigraph is called even if all of its vertices have even degree.

Theorem. Let G be a connected multigraph. Then

G is Eulerian iff G is even.

Proof.

$\Rightarrow$ Easy.

$\Leftarrow$ (Strong) induction on the number of edges.

Lemma. If every vertex of a multigraph G has degree at least 2, then G contains a cycle.

Proof. Extremality: Consider a maximal path...

Corollary of the proof. Every even multigraph decomposes into cycles.

Eulerian trails

Theorem. A connected graph with exactly $2k$ vertices of odd degree decomposes into $\max\{k, 1\}$ trails.

Proof. Reduce it to the characterization of Eulerian graphs by introducing auxiliary edges.

Example. The “little house” can be drawn with one continous motion.

Remark. The theorem is “best possible”, i.e. a decomposition into *less* than $\max\{k, 1\}$ trails is not possible.

Proof techniques

- (Strong) induction
- Extremality
- Double counting

Neighborhoods and degrees...

The neighborhood of v in G is
$N_G(v) = \{w \in V(G) : vw \in E(G)\}$.
The degree of a vertex v in graph G is
$d_G(v) = |N_G(v)|$.

The maximum degree of G is $\Delta(G) = \max_{v \in V(G)} d(v)$

The minimum degree of G is $\delta(G) = \min_{v \in V(G)} d(v)$

G is regular if $\Delta(G) = \delta(G)$
G is k-regular if the degree of each vertex is k.

The order of graph G is $n(G) = |V(G)|$.
The size of graph G is $e(G) = |E(G)|$.

Double counting and bijections I

Handshaking Lemma. For any graph G,

$$\sum_{v \in V(G)} d(v) = 2e(G).$$

Corollary. Every graph has an even number of vertices of odd degree.
No graph of odd order is regular with odd degree.

Corollary. In a graph G the average degree is $\frac{2e(G)}{n(G)}$ and hence $\delta(G) \leq \frac{2e(G)}{n(G)} \leq \Delta(G)$.

Corollary. A k-regular graph with n vertices has $kn/2$ edges.

Double counting and bijections II

Proposition. Let G be k-regular bipartite graph with partite sets A and B, $k > 0$. Then $|A| = |B|$.

Proof. Double count the edges of G.

Claim. The Petersen graph contains ten 6-cycles.

Proof. Bijection between 6-cycles and claws. (A claw is a $K_{1,3}$.)

Extremal problems— Examples

Proposition 1. If G is an n-vertex graph with at most $n - 2$ edges then G is disconnected.

Proof. By induction on $e(G)$ prove that every graph G has at least $n(G) - e(G)$ components.

A **Question** you always have to ask:
Can we improve on this proposition?

Answer. NO! The same statement is FALSE with $n-1$ in the place of $n - 2$.
Proposition 1 is *best possible*, as shown by P_n.

Proposition 2. If G is an n-vertex graph with at least n edges then G contains a cycle.

Remark. Proposition 2 is also *best possible*, (e.g. P_n).

Proposition 1. + Remark: The minimum value of $e(G)$ over connected graphs is $n - 1$.

Proposition 2. + Remark: The maximum value of $e(G)$ over acyclic (i.e. cycle-free) graphs is $n - 1$.

Extremal problems — More example

Vague description: An extremal problem asks for the maximum or minimum value of a parameter over a class of objects (graphs, in most cases).

Proposition. G is an n-vertex graph with $\delta(G) \geq \lfloor n/2 \rfloor$, then G is connected.

Remark. The above proposition is *best possible*, as shown by $K_{\lfloor n/2 \rfloor} + K_{\lceil n/2 \rceil}$.

Graph $G + H$ is the disjoint union (or sum) of graphs G and H. For an integer m, mG is the graph consisting of m disjoint copies of G.

Prop. + Remark: The maximum value of $\delta(G)$ over disconnected graphs is $\lfloor \frac{n}{2} \rfloor - 1$.

Extremal Problems

graph property	graph parameter	type of extremum	value of extremum
connected	$e(G)$	minimum	$n-1$
acyclic	$e(G)$	maximum	$n-1$
disconnected	$\delta(G)$	maximum	$\left\lfloor \frac{n}{2} \right\rfloor - 1$
K_3-free	$e(G)$	maximum	$\left\lfloor \frac{n^2}{4} \right\rfloor$

Triangle-free subgraphs

Theorem. (Mantel, 1907) The maximum number of edges in an n-vertex triangle-free graph is $\lfloor \frac{n^2}{4} \rfloor$.

Proof.

(i) *There is* a triangle-free graph with $\lfloor \frac{n^2}{4} \rfloor$ edges.

(ii) If G is a triangle-free graph, then $e(G) \leq \lfloor \frac{n^2}{4} \rfloor$.

Proof of (ii) is with extremality. (Look at the neighborhood of a vertex of maximum degree.)

Example of a wrong proof of (ii) by induction.

Bipartite subgraphs

Theorem. Every loopless multigraph G has a bipartite subgraph with at least $e(G)/2$ edges.

Proof # 1. Algorithmic. (Start from an arbitrary bipartition and move over a vertex whose degree in its own part is *more* than its degree in the other part. Iterate. Prove termination. Prove that at termination you have what you want.)

Proof # 2. Extremality. (Consider a bipartite subgraph H with the *maximum number of edges*, prove that $d_H(v) \geq d_G(v)/2$ for every vertex $v \in V(G)$ and use the Handshaking Lemma.)

Remark 1. *Maximum vs. maximal.* Algorithmic proof *not* necessarily ends up in bipartite subgraph with maximum number of edges.

Remark 2. The constant multiplier $\frac{1}{2}$ of $e(G)$ in the Theorem is best possible. *Example:* K_n.

Leaves, trees, forests...

A graph with no cycle is acyclic. An acyclic graph is called a forest.

A connected acyclic graph is a tree.

A leaf (or pendant vertex) is a vertex of degree 1.

A spanning subgraph of G is a subgraph with vertex set $V(G)$.

A spanning tree is a spanning subgraph which is a tree.

Examples. Paths, stars

Properties of trees

Lemma. T is a tree, $n(T) \geq 2 \Rightarrow T$ contains at least two leaves.
Deleting a leaf from a tree produces a tree.

Theorem (Characterization of trees) For an n-vertex graph G, the following are equivalent

1. G is connected and has no cycles.
2. G is connected and has $n-1$ edges.
3. G has $n-1$ edges and no cycles.
4. For each $u, v \in V(G)$, G has exactly one u,v-path.

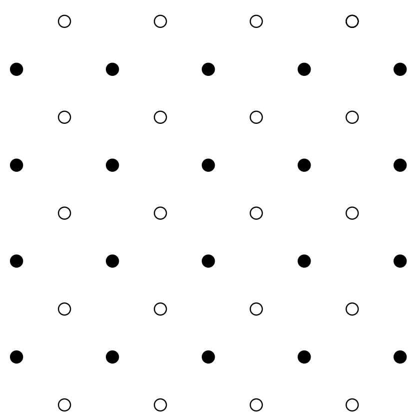
Corollary.

(i) Every edge of a tree is a cut-edge.

(ii) Adding one edge to a tree forms exactly one cycle.

(iii) Every connected graph contains a spanning tree.

Bridg-it* by David Gale



*©1960 by Hassenfeld Bros., Inc. — "Hasbro Toys"

Who wins in Bridg-it?

Theorem. Player 1 has a winning strategy in Bridg-it.

Proof. Strategy Stealing.

Suppose Player 2 has a winning strategy.

Then here is a winning strategy for Player 1:

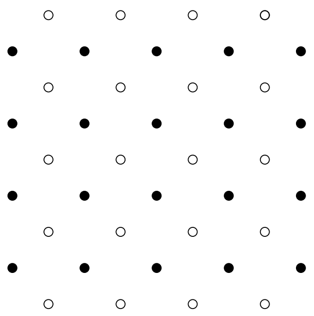
Start with an arbitrary move and then pretend to be Player 2 and play according to Player 2's winning strategy. (Note that playground is symmetric!!) If this strategy calls for the first move of yours, again select an arbitrary edge. Etc...

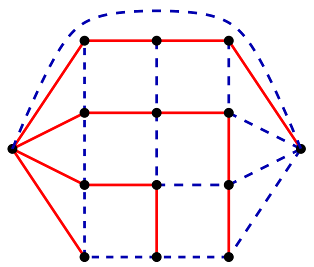
Since you play according to a winning strategy, you win! But we assumed Player 2 also can win $\Rightarrow$ contradiction, since both cannot win.

Good, but HOW ABOUT AN EXPLICIT STRATEGY???*

*In the *divisor-game* strategy-stealing proves the existence of a sure first player win, but NO explicit strategy is known. Similarly for HEX.

An explicit strategy via spanning trees





The game of "Connectivity"

A positional game is played by two players, Maker and Breaker, who alternately take edges of a base graph G. Maker uses a permanent marker, Breaker uses an eraser. Maker wins the positional game "Connectivity" if by the end he occupies a connected subgraph of G. Otherwise Breaker wins.

Theorem. (Lehman, 1964) Suppose Breaker starts the game. If G contains two edge-disjoint spanning tree, then Maker has an explicit winning strategy in "Connectivity".

Proof. Maker maintains two spanning trees T_1 and T_2, such that after each full round,

(i) $E(T_1) \cap E(T_2)$ consists of the edges claimed by Maker,

(ii) $E(T_1) \triangle E(T_2)$ contains only unclaimed edges.

Remark. The other direction of the Theorem is also true.

The tool for Player 1. (i.e. Maker)

Proposition. If T and T' are spanning trees of a connected graph G and $e \in E(T) \setminus E(T')$, then **there is** an edge $e' \in E(T') \setminus E(T)$, such that $T - e + e'$ is a spanning tree of G.

Proposition. If T and T' are spanning trees of a connected graph G and $e \in E(T) \setminus E(T')$, then **there is** an edge $e' \in E(T') \setminus E(T)$, such that $T' + e - e'$ is a spanning tree of G.

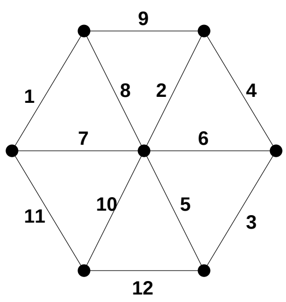
How to build the cheapest road network?

G is a weighted graph if there is a weight function $w : E(G) \to \mathbb{R}$.

Weight $w(H)$ of a subgraph $H \subseteq G$ is defined as

$$w(H) = \sum_{e \in E(H)} w(e).$$

Example:



Kruskal's Algorithm

Kruskal's Algorithm

Input: connected graph G, weight function $w : E(G) \to \mathbb{R}$, $w(e_1) \leq w(e_2) \leq ... \leq w(e_m)$.

Idea: Maintain a spanning forest H of G. At each iteration try to enlarge H by an edge of smallest weight.

Initialization: $V(H) \leftarrow V(G)$, $E(H) \leftarrow \emptyset$, $i \leftarrow 1$

WHILE $i \leq n$
  $e \leftarrow e_i$
  IF e goes between two components of H THEN
    update $H \leftarrow H + e$
    IF H is connected THEN
      stop and return H
  $i \leftarrow i + 1$

Theorem. In a connected weighted graph G, Kruskal's Algorithm constructs a minimum-weight spanning tree.

Proof of correctness of Kruskal's Algorithm

Proof. T is the graph produced by the Algorithm.
$E(T) = \{f_1, \ldots, f_{n-1}\}$ and $w(f_1) \leq \cdots \leq w(f_{n-1})$.

Easy: T is spanning (already at initialization!)
T is a connected (by termination rule) and has no cycle (by iteration rule) $\Rightarrow T$ is a tree.

But **WHY** is T min-weight?

Let T^* be an arbitrary min-weight spanning tree. Let j be the largest index such that $f_1, \ldots, f_j \in E(T^*)$.

If $j = n - 1$, then $T^* = T$. Done.

Proof of Kruskal, cont'd

If $j < n - 1$, then $f_{j+1} \notin E(T^*)$.
There is an edge $e \in E(T^*)$, such that
$T^{**} = T^* - e + f_{j+1}$ is a spanning tree.

(i) $w(T^*) - w(e) + w(f_{j+1}) = w(T^{**}) \geq w(T^*)$
So $w(f_{j+1}) \geq w(e)$.

(ii) Key: When we selected f_{j+1} into T, e was also available. (The addition of e wouldn't have created a cycle, since $f_1, \ldots, f_j, e \in E(T^*)$.)
So $w(f_{j+1}) \leq w(e)$.

Combining: $w(e) = w(f_{j+1})$, i.e. $w(T^{**}) = w(T^*)$.
Thus T^{**} is min-weight spanning tree and it contains a *longer* initial segment of the edges of T, than T^* did.

Repeating this procedure at most $(n - 1)$-times, we transform any min-weight spanning tree into T.

Matchings

A matching is a set of (non-loop) edges with no shared endpoints. The vertices incident to an edge of a matching M are saturated by M, the others are unsaturated. A perfect matching of G is matching which saturates all the vertices.

Examples. $K_{n,m}, K_n$, Petersen graph, Q_k; graphs without perfect matching

A maximal matching cannot be enlarged by adding another edge.

A maximum matching of G is one of maximum size.

Example. Maximum $\neq$ Maximal

Characterization of maximum matchings

Let M be a matching. A path that alternates between edges in M and edges not in M is called an M-alternating path.
An M-alternating path whose endpoints are unsaturated by M is called an M-augmenting path.

Theorem(Berge, 1957) A matching M is a maximum matching of graph G iff G has no M-augmenting path.

Proof. ($\Rightarrow$) Easy.
($\Leftarrow$) Suppose there is no M-augmenting path and let M^* be a matching of maximum size.
What is then $M \triangle M^*$???

> **Lemma** Let M_1 and M_2 be matchings of G. Then each connected component of $M_1 \triangle M_2$ is a path or an even cycle.
>
> For two sets A and B, the symmetric difference is $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

Hall's Condition and consequences

Theorem (Marriage Theorem; Hall, 1935) Let G be a bipartite (multi)graph with partite sets X and Y. Then there is a matching in G saturating X iff $|N(S)| \geq |S|$ for every $S \subseteq X$.

Proof. ($\Rightarrow$) Easy.

($\Leftarrow$) Not *so* easy. Find an M-augmenting path for *any* matching M which does not saturate X.
(Let U be the M-unsaturated vertices in X. Define

$$T := \{y \in Y : \exists\ M\text{-alternating } U,y\text{-path}\},$$
$$S := \{x \in X : \exists\ M\text{-alternating } U,x\text{-path}\}.$$

Unless there is an M-augmenting path, $S \cup U$ **violates** Hall's condition.)

Corollary. (Frobenius (1917)) For $k > 0$, every k-regular bipartite (multi)graph has a perfect matching.

Application: 2-Factors

A factor of a graph is a spanning subgraph. A k-factor is a spanning k-regular subgraph.

Every regular bipartite graph has a 1-factor.

Not every regular graph has a 1-factor.

But...

Theorem. (Petersen, 1891) Every $2k$-regular graph has a 2-factor.

Proof. Use Eulerian cycle of G to create an auxiliary k-regular bipartite graph H, such that a perfect matching in H corresponds to a 2-factor in G.

Graph parameters I

The size of the largest matching (*independent set of edges*) in G is denoted by $\alpha'(G)$.

A vertex cover of G is a set $Q \subseteq V(G)$ that contains at least one endpoint of every edge. (The vertices in Q *cover* $E(G)$).
The size of the smallest vertex cover in G is denoted by $\beta(G)$.

Claim. $\beta(G) \geq \alpha'(G)$.

Certificates

Suppose we knew that in some graph G with 1121 edges on 200 vertices, a particular set of 87 edges is (one of) the largest matching one could find. How could we convince somebody about this?

Once the particluar 87 edges are shown, it is easy to check that they are a matching, indeed.

But why isn't there a matching of size 88? Verifying that none of the $\binom{1121}{88}$ edgesets of size 88 forms a matching could take some time...

If we happen to be so lucky, that we are able to exhibit a vertex cover of size 87, we are saved. It is then reasonable to check that all 1121 edges are covered by the particular set of 87 vertices.

Exhibiting a vertex cover of a certain size **proves** that no larger matching can be found.

Certificate for bipartite graphs — Take 1

1. Correctness of the certificate:

A vertex cover $Q \subseteq V(G)$ is a certificate proving that no matching of G has size larger than $|Q|$.
That is: $\beta(G) \geq \alpha'(G)$, valid for every graph.

2. Existence of optimal certificate for bipartite graphs:

Theorem. (König (1931), Egerváry (1931))
If G is bipartite then $\beta(G) = \alpha'(G)$.

Remarks
1. König's Theorem $\Rightarrow$ For bipartite graphs there always exists a vertex cover proving that a particular matching of maximum size is really maximum.

2. This is NOT the case for general graphs: C_5.

Proof of König's Theorem: For any minimum vertex cover Q, apply Hall's Condition to match $Q \cap X$ into $Y \setminus Q$ and $Q \cap Y$ into $X \setminus Q$.

Arthur and Merlin

A: Show me a pairing, so my 150 knights can merry these 150 ladies!
M: Not possible!
A: Why?
M: Here are these 93 ladies and 58 knights, none of them are willing to merry.
A: Alright, alright ...

A: Seat my 150 knights around the round table, so that neighbors don't fight!
M: Not possible!
A: Why?
M: It will take me forever to explain you.
A: I don't believe you! Into the dungeon!

NP property: can be certified "efficiently"

Example:
"a bipartite graph having a perfect matching" (provide perfect matching)
"a bipartite graph **not** having a perfect matching" (provide vertex cover of size **less** than $n/2$)
"a graph having a Hamilton cycle" (provide Hamilton cycle)

Merlin's Pech: "a graph does **not** have a Hamilton cycle" is not known to be NP

Graph parameters II

The size of the largest independent set in G is denoted by $\alpha(G)$.

The size of the largest matching (*independent set of edges*) in G is denoted by $\alpha'(G)$.

A vertex cover of G is a set $Q \subseteq V(G)$ that contains at least one endpoint of every edge. (The vertices in Q *cover* $E(G)$).
The size of the smallest vertex cover in G is denoted by $\beta(G)$.

Claim. $\beta(G) \geq \alpha'(G)$.

An edge cover of G is a set L of edges such that every vertex of G is incident to some edge in L.
The size of the smallest edge cover in G is denoted by $\beta'(G)$.

Claim. $\beta'(G) \geq \alpha(G)$.

Min-max theorems for bipartite graphs

Theorem. (König (1931), Egerváry (1931)) If G is bipartite then $\beta(G) = \alpha'(G)$.

Lemma. Let G be any graph. $S \subseteq V(G)$ is an independent set iff $\overline{S}$ is a vertex cover.
Hence $\alpha(G) + \beta(G) = n(G)$.

Proof. Easy.

Theorem. (Gallai, 1959) Let G be any graph without isolated vertices. Then $\alpha'(G) + \beta'(G) = n(G)$.

Proof. $\leq$: Take a matching M with $|M| = \alpha'(G)$ and construct an edge cover of size $n(G) - |M|$.
$\geq$: Take an edge cover L with $|L| = \beta'(G)$ and construct matching of size $n(G) - |L|$.

Corollary. (König, 1916) Let G be a bipartite graph with no isolated vertices. Then $\alpha(G) = \beta'(G)$.

Proof. Put together the previous three statements.

How to find a maximum matching in bipartite graphs?

Augmenting Path Algorithm

Input. A bipartite graph G with partite sets X and Y,
a matching M in G,
the set U of unsaturated vertices in X.

Output. EITHER an M-augmenting path OR a certificate (a cover of the same size) that M is maximum.

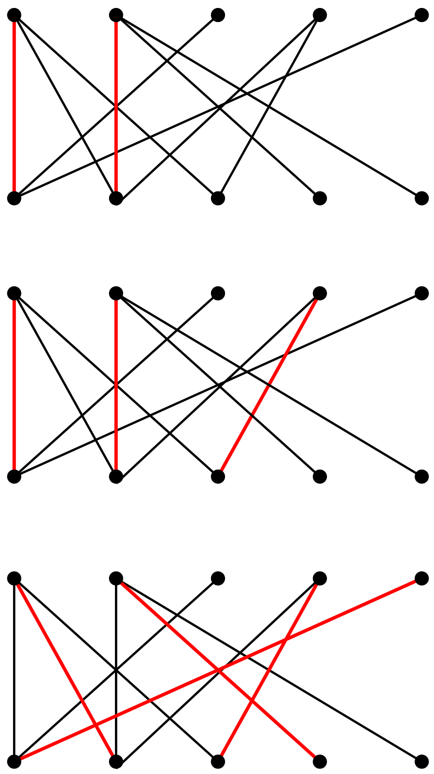
Idea. Explore M-alternating paths from U, letting $S \subseteq X$ and $T \subseteq Y$ be the sets of vertices reached. Mark vertices of S that have been explored for path extensions. As a vertex is reached, record the vertex from which it is reached.

Initialization. $S = U$ and $T = \emptyset$.

Iteration.
IF all vertices in S are marked THEN
  stop and **report** that M is a maximum matching and $T \cup (X \setminus S)$, is a cover of the same size.
ELSE
  select an unmarked $x \in S$ and explore its neighbors $y \in N(x)$, for which $xy \notin M$.
  IF y is unsaturated, THEN
    stop and **report** an M-augmenting path from U to y.
  ELSE
    $\exists w \in X$ with $yw \in M$. **Update**
      $T := T \cup \{y\}$ (y is reached from x),
      $S := S \cup \{w\}$ (w is reached from y).
  After exploring all neighbors of x, mark x and
iterate.

Theorem. Repeatadly applying the Augmenting Path Algorithm to a bipartite graph produces a maximum matching and a minimum vertex cover.

If G has n vertices and m edges, then this algorithm finds a maximum matching in $O(nm)$ time.



Proof of correctness

If Augmenting Path Algorithm does what it supposed to, then after at most $n/2$ application we can produce a maximum matching.
Why does the APA terminate? It touches each edge at most once. Hence running time is $O(nm)$.

What if an M-augmenting path is returned? It is OK, since y is an unsaturated neighbor of $x \in S$, and x can be reached from U on an M-alternating path.

What if the APA returns M as maximum matching and $T \cup (X \setminus S)$ as minimum cover?

Then all edges leaving S were explored, so there is **no edge between** S **and** $Y \setminus T$.

- Hence $T \cup (X \setminus S)$ is indeed a cover.
- $|M| = |T| + |X \setminus S|$ (By selection of S and T.)

If a cover and a matching have the same size in any graph, then they are both optimal.

$$|M| \le \alpha'(G) \le \beta(G) \le |T \cup (X \setminus S)| = |M|.$$

How to find a maximum weight matching in a bipartite graph?

In the maximum weighted matching problem a non-negative weight $w_{i,j}$ is assigned to each edge $x_i y_j$ of $K_{n,n}$ and we seek a perfect matching M to maximize the total weight $w(M) = \sum_{e \in M} w(e)$.

With these weights, a (weighted) cover is a choice of labels $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$, such that $u_i + v_j \ge w_{i,j}$ for all i, j. The cost $c(u, v)$ of a cover (u, v) is $\sum u_i + \sum v_j$. The minimum weighted cover problem is that of finding a cover of minimum cost.

Duality Lemma For a perfect matching M and a weighted cover (u, v) in a bipartite graph G, $c(u, v) \ge w(M)$. Also, $c(u, v) = w(M)$ iff M consists of edges $x_i y_j$ such that $u_i + v_j = w_{i,j}$. In this case, M and (u, v) are both optimal.

The algorithm

The equality subgraph $G_{u,v}$ for a weighted cover (u, v) is the spanning subgraph of $K_{n,n}$ whose edges are the pairs $x_i y_j$ such that $u_i + v_j = w_{i,j}$. In the cover, the excess for i, j is $u_i + v_j - w_{i,j}$.

Hungarian Algorithm

Input. A matrix $(w_{i,j})$ of weights on the edges of $K_{n.n}$ with partite sets X and Y.

Idea. Iteratively adjusting a cover (u, v) until the equality subgraph $G_{u,v}$ has a perfect matching.

Initialization. Let $u_i = \max\{w_{i,j} : j = 1, \ldots, n\}$ and $v_j = 0$.

Iteration.

Form $G_{u,v}$ and find a maximum matching M in it.
IF M is a perfect matching, THEN
  stop and **report** M as a maximum weight matching
  and (u, v) as a minimum cost cover
ELSE
  let Q be a vertex cover of size $|M|$ in $G_{u,v}$.
  $R := X \cap Q$
  $T := Y \cap Q$
  $\epsilon := \min\{u_i + v_j - w_{i,j} : x_i \in X \setminus R, y_j \in Y \setminus T\}$
  Update u and v:
    $u_i := u_i - \epsilon$ if $x_i \in X \setminus R$
    $v_j := v_j + \epsilon$ if $y_j \in T$
Iterate

Theorem The Hungarian Algorithm finds a maximum weight matching and a minimum cost cover.

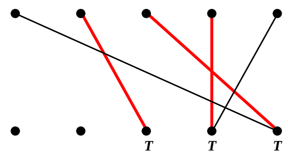
The Assignment Problem — An example

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 7 & 2 \\ 1 & 3 & 4 & 4 & 5 \\ 3 & 6 & 2 & 8 & 7 \\ 4 & 1 & 3 & 5 & 4 \end{pmatrix}$$

Excess Matrix

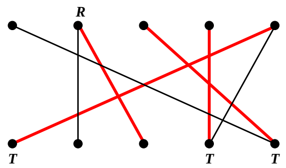
$$\begin{array}{c|ccccc} & 0 & 0 & 0 & 0 & 0 \\ \hline 5 & 4 & 3 & 2 & 1 & 0 \\ 8 & 2 & 1 & 0 & 1 & 6 \\ 5 & 4 & 2 & 1 & 1 & 0 \\ 8 & 5 & 2 & 6 & 0 & 1 \\ 5 & 1 & 4 & 2 & 0 & 1 \\ & & & T & T & T \end{array}$$

Equality Subgraph



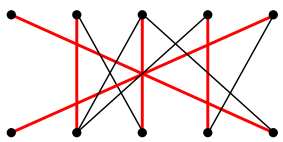
$\epsilon = 1$

$$\begin{array}{c|ccccc|c} & 0 & 0 & 1 & 1 & 1 & \\ \hline 4 & 3 & 2 & 2 & 1 & 0 & \\ 7 & 1 & 0 & 0 & 1 & 6 & R \\ 4 & 3 & 1 & 1 & 1 & 0 & \\ 7 & 4 & 1 & 6 & 0 & 1 & \\ 4 & 0 & 3 & 2 & 0 & 1 & \\ & T & & & T & T & \end{array}$$



$\epsilon = 1$

$$\begin{array}{c|ccccc} & 1 & 0 & 1 & 2 & 2 \\ \hline 3 & 3 & 1 & 1 & 1 & 0 \\ 7 & 2 & 0 & 0 & 2 & 7 \\ 3 & 3 & 0 & 0 & 1 & 0 \\ 6 & 4 & 0 & 5 & 0 & 1 \\ 3 & 0 & 2 & 1 & 0 & 1 \end{array}$$



DONE!!

The Duality Lemma states that if $w(M) = c(u, v)$ for some cover (u, v), then M is maximum weight.

We found a maximum weight matching (transversal). The fact that it is maximum is certified by the indicated cover, which has the same cost:

$$\begin{array}{c|ccccc} & 1 & 0 & 1 & 2 & 2 \\ \hline 3 & 1 & 2 & 3 & 4 & 5 \\ 7 & 6 & 7 & 8 & 7 & 2 \\ 3 & 1 & 3 & 4 & 4 & 5 \\ 6 & 3 & 6 & 2 & 8 & 7 \\ 3 & 4 & 1 & 3 & 5 & 4 \end{array}$$

$$\begin{aligned} w(M) &= 5+7+4+8+4 = 28 = \\ &= 1+0+1+2+2+ \\ &\quad 3+7+3+6+3 = c(u,v) \end{aligned}$$

Hungarian Algorithm — Proof of correctness

Proof. If the algorithm ever terminates and $G_{u,v}$ is the equality subgraph of a (u, v), which is indeed a cover, then M is a m.w.m. and (u, v) is a m.c.c. by Duality Lemma.

Why is (u, v), created by the iteration, a cover?

Let $x_i y_j \in E(K_{n,n})$. Check the four cases.

$x_i \in R,$	$y_j \in Y \setminus T$	$\Rightarrow$	u_i and v_j do not change.
$x_i \in R,$	$y_j \in T$	$\Rightarrow$	u_i does not change v_j increases.
$x_i \in X \setminus R,$	$y_j \in T$	$\Rightarrow$	u_i decreases by ϵ, v_j increases by ϵ.
$x_i \in X \setminus R,$	$y_j \in Y \setminus T$	$\Rightarrow$	$u_i + v_j \geq w_{i,j}$ by definition of ϵ.

Why does the algorithm terminate?

M is a matching in the new $G_{u,v}$ as well. So either
(i) max matching gets larger or
(ii) # of vertices reached from U by M-alternating paths grows. (U is the set of unsaturated vertices of M in X.)

Certificate for bipartite graphs — Take 2

Let G be a bipartite graph with partite sets X and Y.

1. Correctness of the certificate:

A subset $S \subseteq X$ is a certificate proving that the largest matching in G has size at most $|X| - |S| + |N(S)|$.

2. Existence of optimal certificate:

Theorem (Marriage Theorem; Hall, 1935) There is a matching in G saturating X iff $|N(S)| \geq |S|$ for every $S \subseteq X$.

CorollaryThere exists a subset $S \subseteq X$, such that $\alpha'(G) = |X| - |S| + |N(S)|$.

Proof. Homework.

Problem: Certificate makes sense for bipartite graphs only.
Goal: Find a certificate for general graphs.

Matchings in general graphs

An odd component is a connected component with an odd number of vertices. Denote by $o(G)$ the number of odd components of a graph G.

Theorem. (Tutte, 1947) A graph G has a perfect matching iff $o(G - S) \leq |S|$ for every subset $S \subseteq V(G)$.

Proof.

$\Rightarrow$ Easy.

$\Leftarrow$ (Lovász, 1975) Consider a counterexample G with the maximum number of edges.

Claim. $G + xy$ has a perfect matching for any $xy \notin E(G)$.

Proof of Tutte's Theorem — Continued

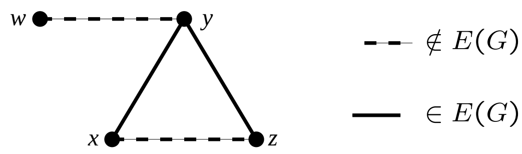
Define $U := \{v \in V(G) : d_G(v) = n(G) - 1\}$

Case 1. $G - U$ consists of disjoint cliques.

Proof: Straightforward to construct a perfect matching of G.

Case 2. $G - U$ is not the disjoint union of cliques.

Proof: Derive the existence of the following subgraph.



Obtain contradiction by constructing a perfect matching M of G using perfect matchings M_1 and M_2 of $G+xz$ and $G + yw$, respectively.

Corollaries

Corollary. (Berge,1958) For a subset $S \subseteq V(G)$ let $d(S) = o(G - S) - |S|$. Then

$$2\alpha'(G) = \min\{n - d(S) : S \subseteq V(G)\}.$$

Proof. $(\leq)$ Easy.
$(\geq)$ Apply Tutte's Theorem to $G \vee K_d$.

Corollary. (Petersen, 1891) Every 3-regular graph with no cut-edge has a perfect matching.

Proof. Check Tutte's condition. Let $S \subseteq V(G)$.
Double-count the number of edges between an S and the odd components of $G - S$.
Observe that between any odd component and S there are at least three edges.

History of maximum matching algorithms

Authors	Year	Order of Running Time
Edmonds*	1965	n^2m
Even-Kariv	1975	$\min\{\sqrt{n}m \log n, n^{2.5}\}$
Micali-Vazirani	1980	$\sqrt{n}m$
Rabin-Vazirani	1989	$n^{\omega+1}$
Mucha-Sankowski	2004	n^{ω}
Harvey	2006	n^{ω}

* In his paper "Paths, Trees, and Flowers" Edmonds defined the notion of polynomial time algorithm

$\omega := \inf\{c :$ two $n \times n$ matrices can be multiplied using $O(n^c)$ arithmetic operations$\}$

Clear: $\omega \geq 2$

Naive algorithm: $\omega \leq 3$

Theorem (Coppersmith-Winograd, 1990) $\omega < 2.38$

Connectivity

A separating set (or vertex cut) of a graph G is a set $S \subseteq V(G)$ such that $G - S$ has more than one component. For $G \neq K_n$, the connectivity of G is $\kappa(G) := min\{|S| : S \text{ is a vertex cut}\}$. By definition, $\kappa(K_n) := n - 1$. A graph G is k-connected if there is no vertex cut of size $k - 1$. (i.e. $\kappa(G) \geq k$)

Examples. $\kappa(K_{n,m}) = \min\{n, m\}$
$\kappa(Q_d) = d$

Extremal problem: What is the minimum number of edges in a k-connected graph?

Theorem. For every n, the minimum number of edges in a k-connected graph is $\lceil kn/2 \rceil$.

Proof:

min $\geq$ $\lceil kn/2 \rceil$, since $k \leq \kappa(G) \leq \delta(G)$
min $\leq$ $\lceil kn/2 \rceil$; Example: Harary graphs $H_{k,n}$.

Edge-connectivity

An edge cut of a multigraph G is an edge-set of the form $[S, \bar{S}]$, with $\emptyset \neq S \neq V(G)$ and $\bar{S} = V(G) \setminus S$.

For $S, T \subseteq V(G)$, $[S, T] := \{xy \in E(G) : x \in S, y \in T\}$.

The edge-connectivity of G is

$\kappa'(G) := \min\{\ |[S, \bar{S}]|\ : [S, \bar{S}] \text{ is an edge cut}\}$.

A graph G is k-edge-connected if there is no edge cut of size $k - 1$ (i.e. $\kappa'(G) \geq k$).

Theorem. (Whitney, 1932) If G is a simple graph, then $\kappa(G) \leq \kappa'(G) \leq \delta(G)$.

Homework. Example of a graph G with $\kappa(G) = k$, $\kappa'(G) = l$, $\delta(G) = m$, for any $0 < k \leq l \leq m$.

Theorem. G is 3-regular $\Rightarrow$ $\kappa(G) = \kappa'(G)$.

Characterization of 2-connected graphs

Theorem. (Whitney,1932) Let G be a graph, $n(G) \geq 3$. Then G is 2-connected iff for every $u, v \in V(G)$ there exist two internally disjoint u, v-paths in G.

Theorem. Let G be a graph with $n(G) \geq 3$. Then the following four statements are equivalent.

(i) G is 2-connected

(ii) For all $x, y \in V(G)$, there are two internally disjoint x, y-path.

(iii) For all $x, y \in V(G)$, there is a cycle through x and y.

(iv) $\delta(G) \geq 1$, and every pair of edges of G lies on a common cycle.

Expansion Lemma. Let G' be a supergraph of a k-connected graph G obtained by adding one vertex to $V(G)$ with at least k neighbors.
Then G' is k-connected as well.

Menger's Theorem

Given $x, y \in V(G)$, a set $S \subseteq V(G) \setminus \{x, y\}$ is an x, y-separator (or an x, y-cut) if $G - S$ has no x, y-path.
A set $\mathcal{P}$ of paths is called pairwise internally disjoint (p.i.d.) if for any two path $P_1, P_2 \in \mathcal{P}$, P_1 and P_2 have no common internal vertices.
Define

$\kappa(x, y) := \min\{|S| : S \text{ is an } x, y\text{-cut,}\}$ and
$\lambda(x, y) := \max\{|\mathcal{P}| : \mathcal{P} \text{ is a set of p.i.d. } x, y\text{-paths}\}$

Local Vertex-Menger Theorem (Menger, 1927) Let $x, y \in V(G)$, such that $xy \notin E(G)$. Then

$$\kappa(x, y) = \lambda(x, y).$$

Corollary (Global Vertex-Menger Theorem) A graph G is k-connected iff for any two vertices $x, y \in V(G)$ there exist k p.i.d. x, y-paths.

Proof: **Lemma.** For every $e \in E(G)$, $\kappa(G - e) \geq \kappa(G) - 1$.

Edge-Menger

Given $x, y \in V(G)$, a set $F \subseteq E(G)$ is an x,y-disconnecting set if $G - F$ has no x,y-path. Define

$$\kappa'(x,y) := \min\{|F| : F \text{ is an } x,y\text{-disconnecting set,}\}$$
$$\lambda'(x,y) := \max\{|\mathcal{P}| : \mathcal{P} \text{ is a set of p.e.d.}^* \ x,y\text{-paths}\}$$

* p.e.d. means pairwise edge-disjoint

Local Edge-Menger Theorem For all $x, y \in V(G)$,

$$\kappa'(x,y) = \lambda'(x,y).$$

Proof. Apply Menger's Theorem for the line graph of G', where $V(G') = V(G) \cup \{s,t\}$ and $E(G') = E(G) \cup \{sx, yt\}$.

The line graph $L(G)$ of a graph G is defined by
$V(L(G)) := E(G)$,
$E(L(G)) := \{ef : e \text{ and } f \text{ share an endpoint}\}$.

Corollary (Global Edge-Menger Theorem) Multigraph G is k-edge-connected iff there is a set of k p.e.d. x,y-paths for any two vertices x and y.

Directed graphs

A directed (multi)graph (or digraph) is a triple consisting of a vertex set $V(G)$, edge set $E(G)$, and a function assigning each edge an ordered pair of vertices.

For an edge $e = (x, y)$, x is the tail of e, y is its head.

By path and cycle in a directed graph we always mean directed path and directed cycle.

A directed graph is weakly connected if the underlying undirected graph is connected; it is strongly connected or strong if there is a u,v-path for any vertex u and any vertex $v \neq u$.

The out-neighborhood of v in G is
$N_G^+(v) = \{w \in V(G) : (v,w) \in E(G)\}$.
The out-degree of v is $d_G^+(v) = |N_G^+(v)|$.

The in-neighborhood of v in G is
$N_G^-(v) = \{w \in V(G) : (w,v) \in E(G)\}$.
The in-degree of v is $d_G^-(v) = |N_G^-(v)|$.

Déjà vu

Directed Handshaking. In a directed multigraph G, we have

$$\sum_{v \in V(G)} d^+(v) = e(G) = \sum_{v \in V(G)} d^-(v).$$

A directed multigraph is Eulerian if it has a directed Eulerian circuit, i.e. a closed directed trail containing all edges.

Theorem. A weakly connected directed multigraph on $n(D) \geq 2$ vertices is Eulerian iff $d^+(v) = d^-(v)$ for each vertex v.

Proof. Similar to the undirected case. Think it over.

Menger's Theorem for directed graphs

Given $x, y \in V(D)$, a set $S \subseteq V(D) \setminus \{x,y\}$ is an x,y-separator (or an x,y-cut) if $D - S$ has no x,y-path.
Define

$$\kappa_D(x,y) := \min\{|S| : S \text{ is an } x,y\text{-cut,}\} \text{ and}$$
$$\lambda_D(x,y) := \max\{|\mathcal{P}| : \mathcal{P} \text{ is a set of p.i.d. } x,y\text{-paths}\}$$

Directed-Local-Vertex-Menger Theorem Let $x, y \in V(D)$, such that $\vec{xy} \notin E(D)$. Then

$$\kappa_D(x,y) = \lambda_D(x,y).$$

Proof. (Aharoni) Let $A = N^+(x)$ and $B = N^-(y)$.

$$D' := D - \{x,y\} - \{\vec{za} : a \in A, z \in V(D)\} - \{\vec{bz} : b \in B, z \in V(D)\}$$

$\mathcal{D}$: family of all A,B-paths in D'.

GOAL: Find a family $\mathcal{P} \subseteq \mathcal{D}$ of pairwise disjoint A, B-paths and a subset $S \subseteq V(D')$ such that
$|S \cap V(P)| \geq 1$ for every $P \in \mathcal{D}$ and
$|S \cap V(P)| = 1$ for every $P \in \mathcal{P}$.

Proving the GOAL is indeed enough. (Think it over)

Proof of GOAL. Define an auxiliary bipartite graph H.

$$\begin{aligned} V(H) &:= \{v^-, v^+ : v \in V(D')\} \\ E(H) &:= \{u^+v^- : \vec{uv} \in E(D')\} \cup \\ &\quad\ \{v^-v^+ : v \in V(D') \setminus A \setminus B\} \end{aligned}$$

By König's Theorem there is a matching M and a vertex-cover C in H, such that $|e \cap C| = 1$ for every $e \in M$.

$$\mathcal{P} := \{x_1 \cdots x_k \in \mathcal{D} : x_i^+ x_{i+1}^- \in M \text{ for } 1 \leq i < k\}.$$
$$S := \{v \in V(D') \ : \ v^+, v^- \in C \text{ or } v^+ \in A^+ \cap C \text{ or } v^- \in B^- \cap C\}.$$

- Any two paths $P_1, P_2 \in \mathcal{P}$ are disjoint.

  $V(P_1) \cap V(P_2) \neq \emptyset$ implies there is $f_1 \in E(P_1)$, $f_2 \in E(P_2)$ such that $f_1 \neq f_2$ and $f_1 \cap f_2 \neq \emptyset$. $P_1, P_2 \in \mathcal{P}$ implies that for any $f_i \in E(P_i)$ either $f_1 = f_2$ or $f_1 \cap f_2 = \emptyset$.

- Any A, B-path $x_0x_1x_2 \cdots x_k$ contains a vertex from S.

  Let i be the largest index such that $x_i^- \in C$. (There is such, unless $x_0^+ \in C$ and $i < k$ unless $x_k^- \in C$)
  Then $x_i^+ \in C$ since $x_i^+ x_{i+1}^-$ must be covered.

- No A, B-path $u_0u_1u_2 \cdots u_k = P \in \mathcal{P}$ contains more than one vertices from S.

  Suppose P does contain more. Let u_i and $u_j \in S \cap V(P)$ such that $u_k \notin S$ for $i < k < j$. Then $u_i^+, u_j^- \in C$ by definition of S. Let k be the largest index, $i < k < j$, such that $u_k^+ \in C$. Then $u_{k+1}^- \in C$ to cover the edge $u_{k+1}^- u_{k+1}^+$. Hence edge $u_k^+ u_{k+1}^- \in M$ is covered twice by C, a contradiction.

Corollaries

Corollary (Directed-Global-Vertex-Menger Theorem) A digraph D is strongly k-connected iff for any two vertices $x, y \in V(D)$ there exist k p.i.d. x, y-paths.

Proof: **Lemma.** For every $e \in E(D)$, $\kappa_D(G-e) \geq \kappa_D(G)-1$.

The proof of the very first, the original Menger Theorem (the Undirected-Local-Vertex version) is

HOMEWORK !!!

Derive implication DLVM $\Rightarrow$ ULVM

Directed Edge-Menger

Given $x, y \in V(D)$, a set $F \subseteq E(D)$ is an x, y-disconnecting set if $D - F$ has no x, y-path. Define

$$\begin{aligned} \kappa'_D(x,y) &:= \min\{|F| : F \text{ is an } x,y\text{-disconnecting set,}\} \\ \lambda'_D(x,y) &:= \max\{|\mathcal{P}| : \mathcal{P} \text{ is a set of p.e.d.}^* \ x,y\text{-paths}\} \end{aligned}$$

* p.e.d. means pairwise edge-disjoint

Directed-Local-Edge-Menger Theorem For all $x, y \in V(D)$,

$$\kappa'_D(x,y) = \lambda'_D(x,y).$$

Proof. Create directed line graph and apply DLVM.

Corollary (Directed-Global-Edge-Menger Theorem) Directed multigraph D is strongly k-edge-connected iff there is a set of k p.e.d.x, y-paths for any two vertices x and y.

Vertex coloring, chromatic number

A k-coloring of a graph G is a labeling $f : V(G) \to S$, where $|S| = k$. The labels are called colors; the vertices of one color form a color class.

A k-coloring is proper if adjacent vertices have different labels. A graph is k-colorable if it has a proper k-coloring.

The chromatic number is

$$\chi(G) := \min\{k : G \text{ is } k\text{-colorable}\}.$$

A graph G is k-chromatic if $\chi(G) = k$. A proper k-coloring of a k-chromatic graph is an optimal coloring.

Examples. K_n, $K_{n,m}$, C_5, Petersen

A graph G is k-color-critical (or k-critical) if $\chi(H) < \chi(G) = k$ for every *proper* subgraph H of G.

Characterization of 1-, 2-, 3-critical graphs.

Lower bounds

Simple lower bounds

$$\chi(G) \geq \omega(G)$$
$$\chi(G) \geq \frac{n(G)}{\alpha(G)}$$

Examples for $\chi(G) \neq \omega(G)$:

- odd cycles of length at least 5,

$$\chi(C_{2k+1}) = 3 > 2 = \omega(C_{2k+1})$$

- complements of odd cycles of order at least 5,

$$\chi(\overline{C}_{2k+1}) = k + 1 > k = \omega(\overline{C}_{2k+1})$$

- random graph $G = G(n, \frac{1}{2})$, almost surely

$$\chi(G) \approx \frac{n}{2 \log n} > 2 \log n \approx \omega(G)$$

Mycielski's Construction

The bound $\chi(G) \geq \omega(G)$ could be arbitrarily bad.

Construction. Given graph G with vertices $v_1, \ldots, v_n$, we define supergraph $M(G)$.

$$V(M(G)) = V(G) \cup \{u_1, \ldots u_n, w\}.$$

$$E(M(G)) = E(G) \cup \{u_i v : v \in N_G(v_i) \cup \{w\}\}.$$

Theorem.

(i) If G is triangle-free, then so is $M(G)$.

(ii) If $\chi(G) = k$, then $\chi(M(G)) = k + 1$.

Forced subdivision

G contains a $K_k \Rightarrow \chi(G) \geq k$

G contains a $K_k \not\Leftarrow \chi(G) \geq k$ (already for $k \geq 3$)

Hajós' Conjecture

G contains a K_k-subdivision $\overset{?}{\Leftarrow} \chi(G) \geq k$

An H-subdivision is a graph obtained from H by successive edge-subdivisions.

Remark. The conjecture is true for $k = 2$ and $k = 3$.

Theorem (Dirac, 1952) Hajós' Conjecture is true for $k = 4$.

Homework. Hajós' Conjecture is false for $k \geq 7$.

Hadwiger's Conjecture

G contains a K_k-*minor* $\overset{?}{\Leftarrow} \chi(G) \geq k$

Proved for $k \leq 6$. Open for $k \geq 7$.

Proof of Dirac's Theorem

Theorem (Dirac, 1952) If $\chi(G) \geq 4$ then G contains a K_4-subdivision.

Proof. Induction on $n(G)$. $n(G) = 4 \Rightarrow G = K_4$.

W.l.o.g. G is 4-critical.

Case 0. $\kappa(G) = 0$ would contradict 4-criticality

Case 1. $\kappa(G) = 1$ would contradict 4-criticality

Case 2. $\kappa(G) = 2$. Let $S = \{x, y\}$ be a cut-set.

$xy \in E(G)$ would contradict 4-criticality

Hence $xy \notin E(G)$.

$\chi(G) \geq 4 \Rightarrow G$ must have an S-lobe H, such that $\chi(H+xy) \geq 4$. Apply induction hypothesis to $H+xy$ and find a K_4-subdivision F in $H+xy$. Then modify F to obtain a K_4-subdivision in G.

Let $S \subseteq V(G)$. An S-lobe of G is an induced subgraph of G whose vertex set consists of S and the vertices of a component of $G - S$.

Proof of Dirac's Theorem— Continued

Case 3. $\kappa(G) \geq 3$. Let $x \in V(G)$. $G - x$ is 2-connected, so contains a cycle C of length at least 3.

Claim. There is an x, C-fan of size 3.

Proof. Add a new vertex u to G connecting it to the vertices of C. By the Expansion Lemma the new graph G' is 3-connected. By Menger's Theorem there exist three p.i.d x, u-paths P_1, P_2, P_3 in G'. □

Given a vertex x and a set U of vertices, and x, U-fan is a set of paths from x to U such that any two of them share only the vertex x.

Fan Lemma. G is k-connected iff $|V(G)| \geq k+1$ and for every choice of $x \in V(G)$ and $U \subseteq V(G)$, $|U| \geq k$, G has an x, U-fan.

Then $C \cup P_1 \cup P_2 \cup P_3 - u$ is K_4-subdivision in G.

Upper bounds

Proposition $\chi(G) \leq \Delta(G) + 1$.

Proof. Algorithmic; Greedy coloring.

A graph G is d-degenerate if every subgraph of G has minimum degree at most d.

Claim. G is d-degenerate iff there is an ordering of the vertices $v_1, \ldots, v_n$, such that $|N(v_i) \cap \{v_1, \ldots, v_{i-1}\}| \leq d$

Proposition. For a d-degenerate G, $\chi(G) \leq d + 1$.
In particular, for every G, $\chi(G) \leq \max_{H \subseteq G} \delta(H) + 1$.

Proof. Greedy coloring.

Brooks' Theorem. (1941) Let G be a connected graph. Then $\chi(G) = \Delta(G) + 1$ iff G is a complete graph or an odd cycle.

Proof. Trickier, but still greedy coloring...

Proof of Brooks' Theorem. Cases.

Case 1. G is not regular.
Let the root be a vertex with degree $< \Delta(G)$.

Case 2. G has a cut-vertex.
Let the root be the cut-vertex.

Assume G is k-regular and $\kappa(G) \geq 2$.

Case 3. $k \leq 2$. Then $G = C_l$ or K_2.

Assume $k \geq 3$. We need a root v_n with nonadjacent neighbors v_1, v_2, such that $G - \{v_1, v_2\}$ is connected. Let x be a vertex of degree less than $n(G) - 1$.

Case 4. $\kappa(G - x) \geq 2$.
Let v_n be a neighbor of x, which has a neighbor y, such that y and x are non-neighbors. Then let $v_1 = x$ and $v_2 = y$.

Case 5. $\kappa(G - x) = 1$.
Then x has a neighbor in every leaf-block of $G - x$. Let $v_n = x$ and v_1, v_2 be two neighbors of x in different leaf blocks of $G - x$.

Block-decomposition of connected graphs

Maximal induced subgraph of G with no cut-vertex is called block of G.

Lemma. Two blocks intersect in at most one vertex.

Proof. If B_1 and B_2 have no cut-vertex and share at least two vertices then $B_1 \cup B_2$ has no cut-vertex either.

The Block/Cut-vertex graph of G is a bipartite graph with vertex set

$$\{B : B \text{ is a block}\} \cup \{v : v \text{ is a cut-vertex}\}.$$

Block B is adjacent to cut-vertex v iff $v \in V(B)$.

Proposition. The Block/Cut-vertex graph of a connected graph is a tree.

Examples for $\chi(G) = \omega(G)$

- cliques, bipartite graphs
- *interval graphs*

  An interval representation of a graph is an assignment of an interval to the vertices of the graph, such that two vertices are adjacent iff the corresponding intervals intersect. A graph having such a representation is called an interval graph.

  Proposition. If G is an interval graph, then

  $$\chi(G) = \omega(G).$$

  Proof. Order vertices according to left endpoints of corresponding intervals and color *greedily*.
- *perfect graphs*

Perfect graphs

Definition (Berge) A graph G is perfect, if $\chi(H) = \omega(H)$ for every induced subgraph $H \subseteq G$.

Conjectures of Berge (1960)
Weak Perfect Graph Conjecture. G is perfect iff $\overline{G}$ is perfect.
Strong Perfect Graph Conjecture. G is perfect iff G does not contain an induced subgraph isomorphic to an odd cycle of order at least 5 or the complement of an odd cycle of order at least 5.

The first conjecture was made into the Weak Perfect Graph Theorem by Lovász (1972)
The second conjecture was made into the Strong Perfect Graph Theorem by Chudnovsky, Robertson, Seymour, Thomas (2002)

Line graphs and edge coloring

A k-edge-coloring of a multigraph G is a labeling $f : E(G) \to S$, where $|S| = k$. The labels are called colors; the edges of one color form a color class. A k-edge-coloring is proper if incident edges have different labels. A multigraph is k-edge-colorable if it has a proper k-edge-coloring.

The edge-chromatic number (or chromatic index) of a loopless multigraph G is

$$\chi'(G) := \min\{k : G \text{ is } k\text{-edge-colorable}\}.$$

A multigraph G is k-edge-chromatic if $\chi'(G) = k$.

Remarks. $\chi'(G) = \chi(L(G))$, so

$$\begin{aligned}\Delta(G) &\le \omega(L(G)) \\ &\le \chi'(G) \le \Delta(L(G)) + 1 \\ &\le 2\Delta(G) - 1\end{aligned}$$

Vizing's Theorem

Example. K_{2n}

Theorem. (König, 1916)
For a bipartite multigraph G, $\chi'(G) = \Delta(G)$

Proposition. $\chi'(Petersen) = 4$.

Theorem. (Vizing, 1964) For a simple graph G,

$$\chi'(G) \le \Delta(G) + 1.$$

Generalization. If the maximum edge-multiplicity in a multigraph G is $\mu(G)$, then $\chi'(G) \le \Delta(G) + \mu(G)$
Example. Fat triangle; $\chi'(G) = \Delta(G) + \mu(G)$.

Proof of Vizing's Theorem (A. Schrijver)

Induction on $n(G)$.

If $n(G) = 1$, then $G = K_1$; the theorem is OK.

Assume $n(G) > 1$. Delete a vertex $v \in V(G)$. By induction $G - v$ is $(\Delta(G) + 1)$-edge-colorable.

Why is G also $(\Delta(G) + 1)$-edge-colorable?

We prove the following

Stronger Statement. Let $k \ge 1$ be an integer. Let $v \in V(G)$, such that

- $d(v) \le k$,
- $d(u) \le k$ for every $u \in N(v)$, and
- $d(u) = k$ for at most one $u \in N(v)$.

Then
$G - v$ is k-edge-colorable $\Rightarrow$ G is k-edge-colorable.

Proof of the Stronger Statement I

Induction on k (!!!)

For $k = 1$ it is OK.

W.l.o.g. $d(u) = k - 1$ for every $u \in N(v)$, except for *exactly one* $w \in N(v)$ for which $d(w) = k$.

Let $f : E(G - v) \to \{1, \ldots, k\}$ be a proper k-edge-coloring of $G - v$, which minimizes*

$$\sum_{i=1}^{k} |X_i|^2.$$

Here $X_i := \{u \in N(v) : u \text{ is missing color } i\}$.

*I.e., we choose the coloring so the $|X_i|$s "as equal as possible".

Proof of the Stronger Statement II

Case 1. There is an i, with $|X_i| = 1$. Say $X_k = \{u\}$.

Let $G' = G - uv - \{xy : f(xy) = k\}$.
Apply the induction hypothesis for G' and $k - 1$.

Case 2. $|X_i| \neq 1$ for every $i = 1, \ldots, k$.

Then

$$\sum_{l=1}^{k} |X_l| = 2d(v) - 1 < 2k.$$

So there are colors i with $|X_i| = 0$ and
j with $|X_j| \geq 3$.

Let $H \subseteq G$ be subgraph spanned by the edges of color i and j.
Switch colors i and j in a component C of H, where $C \cap X_j \neq \emptyset$.
This reduces $\sum_{l=1}^{k} |X_l|^2$, a contradiction. □

Complete k-partite graphs

A graph G is r-partite (or r-colorable) if there is a partition $V_1 \cup \cdots \cup V_r = V(G)$ of the vertex set, such that for every edge its endpoints are in *different* parts of the partition.

G is a complete r-partite graph if there is a partition $V_1 \cup \cdots \cup V_r = V(G)$ of the vertex set, such that $uv \in E(G)$ iff u and v are in *different* parts of the partition. If $|V_i| = n_i$, then G is denoted by $K_{n_1,\ldots,n_r}$.

The Turán graph $T_{n,r}$ is the complete r-partite graph on n vertices whose partite sets differ in size by at most 1. (All partite sets have size $\lceil n/r \rceil$ or $\lfloor n/r \rfloor$.)

Lemma Among r-colorable graphs the Turán graph is the *unique* graph, which has the most number of edges.

Proof. Local change. □

Turán's Theorem

The Turán number $ex(n, H)$ of a graph H is the largest integer m such that there exists an H-free* graph on n vertices with m edges.

Example: Mantel's Theorem states $ex(n, K_3) = \left\lfloor \frac{n^2}{4} \right\rfloor$.

Theorem. (Turán, 1941)

$$ex(n, K_r) = e(T_{n,r-1}) = \left(1 - \frac{1}{r-1}\right)\binom{n}{2} + O(n).$$

Proof. Prove by induction on r that

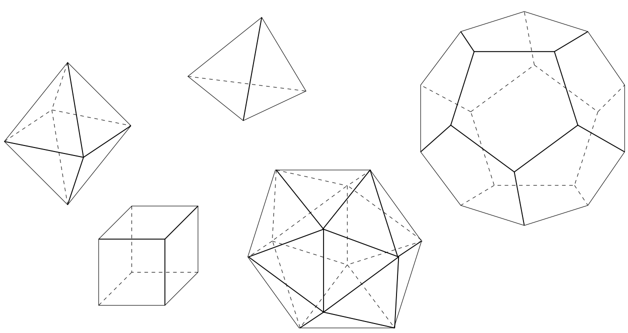
$G \not\supseteq K_r \implies$ there is an $(r-1)$-partite graph H with $V(H) = V(G)$ and $e(H) \geq e(G)$.

Then apply the Lemma to finish the proof. □

*Here H-free means that there is no subgraph isomorphic to H

Turán-type problems

Question. (Turán, 1941) What happens if instead of K_4, which is the graph of the tetrahedron, we forbid the graph of some other platonic polyhedra? How many edges can a graph without an octahedron (or cube, or dodecahedron or icosahedron) have?



The platonic solids

Erdős-Simonovits-Stone Theorem

Theorem. (Erdős-Stone, 1946) For arbitrary fixed integers $r \geq 2$ and $t \geq 1$

$$ex(n, T_{rt,r}) = \left(1 - \frac{1}{r-1}\right)\binom{n}{2} + o(n^2).$$

Corollary. (Erdős-Simonovits, 1966) For any graph H,

$$ex(n, H) = \left(1 - \frac{1}{\chi(H)-1}\right)\binom{n}{2} + o(n^2).$$

Corollaries of the Corollary.

$$\begin{aligned} ex(n, \texttt{octahedron}) &= \frac{n^2}{4} + o(n^2) \\ ex(n, \texttt{dodecahedron}) &= \frac{n^2}{4} + o(n^2) \\ ex(n, \texttt{icosahedron}) &= \frac{n^2}{3} + o(n^2) \\ ex(n, \texttt{cube}) &= o(n^2) \end{aligned}$$

Proof of the Erdős-Simonovits Corollary

Theorem. (Erdős-Stone, 1946) For arbitrary fixed integers $r \geq 2$ and $t \geq 1$

$$ex(n, T_{rt,r}) = \left(1 - \frac{1}{r-1}\right)\binom{n}{2} + o(n^2).$$

Corollary. (Erdős-Simonovits, 1966) For any graph H,

$$ex(n, H) = \left(1 - \frac{1}{\chi(H)-1}\right)\binom{n}{2} + o(n^2).$$

Proof of the Corollary. Let $r = \chi(H)$.

- $\chi(T_{n,r-1}) < \chi(H)$, so $e(T_{n,r-1}) \leq ex(n, H)$.
- $T_{r\alpha,r} \supseteq H$, so $ex(n, T_{r\alpha,r}) \geq ex(n, H)$, where α is a constant depending on H; say $\alpha = \alpha(H)$.

□

The number of edges in a C_4-free graph

Theorem (Erdős, 1938) $ex(n, C_4) = O(n^{3/2})$

Proof. Let G be a C_4-free graph on n vertices.

$C = C(G) :=$ number of $K_{1,2}$ ("cherries") in G.
Doublecount C.

Counting by the midpoint: Every vertex v is the midpoint of exactly $\binom{d(v)}{2}$ cherries. Hence

$$C = \sum_{v \in V}\binom{d(v)}{2}.$$

Counting by the endpoints: Every pair $\{u, w\}$ of vertices form the endpoints of at most one cherry. (Otherwise there is a $C_4 \subseteq G$.) Hence

$$C \leq 1 \cdot \binom{n}{2}.$$

Proof cont'd

Combine and apply Jensen's inequality
(Note that $x \rightarrow \binom{x}{2}$ is a convex function)

$$\binom{n}{2} \geq C \geq \sum_{v \in V}\binom{d(v)}{2} \geq n \cdot \binom{\bar{d}(G)}{2}.$$

$\bar{d}(G) = \frac{1}{n}\sum_{v \in V} d(v)$ is the average degree of G.

$$\frac{n-1}{2} \geq \binom{\bar{d}(G)}{2} \geq \frac{(\bar{d}(G)-1)^2}{2}$$

Hence $\sqrt{n-1} + 1 \geq \bar{d}(G)$. □

Theorem (E. Klein, 1938) $ex(n, C_4) = \Theta(n^{3/2})$
Proof. Homework.

Theorem (Kővári-Sós-Turán, 1954) For $s \geq t \geq 1$

$$ex(n, K_{t,s}) \leq c_s n^{2-\frac{1}{t}}$$

Proof. Homework.

Open problems and Conjectures

Known results.

$$\begin{array}{rcccl} \Omega(n^{3/2}) & \leq & ex(n, Q_3) & \leq & O(n^{8/5}) \\ \Omega(n^{9/8}) & \leq & ex(n, C_8) & \leq & O(n^{5/4}) \\ \Omega(n^{5/3}) & \leq & ex(n, K_{4,4}) & \leq & O(n^{7/4}) \end{array}$$

Conjectures.

$$ex(n, K_{t,s}) = \Theta\left(n^{2-\frac{1}{\min\{t,s\}}}\right)$$ true for $t = 2, 3$ and $s \geq t$ or $t \geq 4$ and $s > (t-1)!$

$$ex(n, C_{2k}) = \Theta\left(n^{1+\frac{1}{k}}\right)$$ true for $k = 2, 3$ and 5

$$ex(n, Q_3) = \Theta\left(n^{\frac{8}{5}}\right)$$

If H is a d-degenerate bipartite graph, then

$$ex(n, H) = O\left(n^{2-\frac{1}{d}}\right).$$

Szemerédi's Regularity Lemma

One of the most important tools in "dense" combinatorics.

Message: every graph G is the approximate union of constantly many random-like bipartite graph. The number of parts depends only on the error of the approximation constant but **not** the size of G!

For disjoint subsets $X, Y \subseteq V$,

$$d(X, Y) := \frac{|E(X, Y)|}{|X| \cdot |Y|}$$

is the density of the pair (X, Y).

A pair (A, B) of disjoint subsets $A, B \subseteq V$ is called ε-regular pair for some $\varepsilon > 0$ if all $X \subseteq A$, and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ satisfy

$$|d(X, Y) - d(A, B)| \leq \varepsilon.$$

Remark Just like in a random bipartite graph...

Szemerédi's Regularity Lemma

A partition $\{V_0, V_1, \ldots, V_k\}$ of V is called an ε-regular partition if

(i) $|V_0| \leq \varepsilon|V|$

(ii) $|V_1| = \cdots = |V_k|$

(iii) all but at most $\varepsilon\binom{k}{2}$ of the pairs (V_i, V_j), with $1 \leq i < j \leq k^2$, are ε-regular

V_0 is the exceptional set

Regularity Lemma (Szemerédi) $\forall\varepsilon > 0$ and -$\forall$ integer $m \geq 1$ $\exists$ integer $M = M(\varepsilon, m)$ such that every graph of order at least m admits an ε-regular partition $\{V_0, V_1, \ldots, V_k\}$ with $m \leq k \leq M$.

Was devised to prove that "dense sets of integers contain an arithmetic progression of arbitrary length".

History of Szemerédi's Theorem

Szemerédi's Theorem (1975) For any integer $k \geq 1$ and $\delta > 0$ there is an integer $N = N(k, \delta)$ such that any subset $S \subseteq \{1, \ldots, N\}$ with $|S| \geq \delta N$ contains an arithmetic progression of length k.

Was conjectured by Erdős and Turán (1936).
Featured problem in mathematics, inspired lots of great new ideas and research in various fields;

- Case of $k = 3$: analytic number theory (Roth, 1953; Fields medal)
- First proof for arbitrary k: combinatorial (Szemerédi, 1975)
- Second proof: ergodic theory (Furstenberg, 1977)
- Third proof: analytic number theory (Gowers, 2001; Fields medal)
- Fourth proof: fully combinatorial (with hypergraphs) (Rödl-Schacht, Gowers, 2007)
- Fifth proof: measure theory (Elek-Szegedy, 2007+)

One of the ingredients in the proof of Green and Tao: "primes contain arbitrary long arithmetic progression"

Proof of the Erdős-Stone Thm

Erdős-Stone Theorem. (Reformulation) For any $\gamma > 0$ and integers $r \geq 2$, $t \geq 1$ there exists an integer $N = N(r,t,\gamma)$, such that any graph G on $n \geq N$ vertices with more than $\left(1 - \frac{1}{r-1}\right)\binom{n}{2} + \gamma n^2$ edges contains $T_{rt,r}$.

Proof strategy:

- Based on an ε-regular partition, build a "regularity graph" R of G. (Regularity Lemma)
- Show that R contains a K_r (Turán's Theorem)
- Show that $K_r \subseteq R \Rightarrow T_{rt,r} \subseteq G$

Regularity graph

Given ε-regular partition $\mathcal{P} = \{V_0, V_1, \ldots, V_k\}$ of G, $m \leq k \leq M(\varepsilon, m)$,
define the regularity graph $R = R(\mathcal{P}, d)$

$V(R) = \{V_1, \ldots, V_k\}$

$V_iV_j \in E(R)$ if (V_i, V_j) is ε-regular pair with density $d(V_i, V_j) \geq d$

Goal Choose ε, m, d such that "most" edges of G go between the sets V_i and V_j with $V_iV_j \in E(R)$

How many edges are not at the "right place"?

# of edges inside V_i: at most $k\binom{n/k}{2} < \frac{n^2}{k} < \frac{n^2}{m}$

# of edges incident to V_0: at most $\varepsilon n \cdot n = \varepsilon n^2$

# of edges between non-regular pairs:
at most $\varepsilon\binom{k}{2}\left(\frac{n}{k}\right)^2 < \varepsilon n^2$

# of edges between pairs of density $< d$:
at most $\binom{k}{2} d \left(\frac{n}{k}\right)^2 \leq dn^2$

Regularity graph contains an r-clique

Conclusion: If ε, m, and d is chosen such that

$$d + 2\varepsilon + \frac{1}{m} < \frac{\gamma}{2}$$

then "most" edges of G go between sets V_i and V_j with $V_iV_j \in E(R)$.
"most" means at least $\left(1 - \frac{1}{r-1}\right)\binom{n}{2} + \frac{\gamma}{2}n^2$

On the other hand: # of edges of G going between sets V_i and V_j with $V_iV_j \in E(R)$:

at most $|E(R)| \cdot \left(\frac{n}{k}\right)^2$

Hence

$$\left(1 - \frac{1}{r-1}\right)\binom{n}{2} + \frac{\gamma}{2}n^2 \leq |E(R)| \cdot \left(\frac{n}{k}\right)^2$$
$$\left(1 - \frac{1}{r-1}\right)\binom{k}{2} + \frac{\gamma}{2}k^2 \leq |E(R)|$$

Choose $m = m(\gamma)$ such that
$ex(m, K_r) \leq \left(1 - \frac{1}{r-1}\right)\binom{m}{2} + \frac{\gamma}{2}m^2$
Then Turán's Theorem $\Rightarrow$ R contains a K_r

Finding $T_{rt,r}$

There are r classes $V_{i_1}, \ldots, V_{i_r}$ such that (V_{i_j}, V_{i_ℓ}) is an ε-regular pair of density at least d, for every $1 \leq j < \ell \leq r$.
Let $\tilde{n} = |V_{i_j}|$. Then $\frac{n}{k} \geq \tilde{n} \geq \frac{1-\epsilon}{k}n$.

We find a $T_{rt,r}$ in $G[V_{i_1} \cup \cdots \cup V_{i_r}]$.

Lemma
Let (A, B) be an ε-regular pair with $d(A,B) \geq d$
Let $Y \subseteq B$ be a subset with $|Y| \geq \varepsilon|B|$.
Then

$$|\{v \in A : d_Y(v) < (d - \varepsilon)|Y|\}| < \varepsilon|A|.$$

Proof. Otherwise the subsets $Y \subseteq B$ and $\{v \in A : d_Y(v) < (d-\varepsilon)|Y|\} \subseteq A$ would contradict the ε-regularity of (A, B). □

For a set $S \subseteq V$ let $\Gamma(S) = \cap_{v \in S} N(v)$ denote the set of common neighbors of the vertices of S.

Finding $T_{rt,r}$

$$(d-\varepsilon)^{t-1}\tilde{n} \geq \varepsilon\tilde{n}$$
$$(r-1)t\varepsilon\tilde{n} \leq \tilde{n} - t$$
$$\Downarrow$$
$$\exists S_1 \subseteq V_1,\ |S_1| = t$$
$$|\Gamma_{V_i}(S_1)| \geq (d-\varepsilon)^t\tilde{n} \quad \text{for } i = 2,3,\ldots,r$$

$$(d-\varepsilon)^{2t-1}\tilde{n} \geq \varepsilon\tilde{n}$$
$$(r-2)t\varepsilon\tilde{n} \leq (d-\varepsilon)^t\tilde{n} - t$$
$$\Downarrow$$
$$\exists S_2 \subseteq V_2,\ |S_2| = t$$
$$|\Gamma_{V_i}(S_1 \cup S_2)| \geq (d-\varepsilon)^{2t}\tilde{n} \quad \text{for } i = 3,\ldots,r$$

.
.
.

$$(d-\varepsilon)^{(r-1)t-1}\tilde{n} \geq \varepsilon\tilde{n}$$
$$t\varepsilon\tilde{n} \leq (d-\varepsilon)^{(r-2)t}\tilde{n} - t$$
$$\Downarrow$$
$$\exists S_{r-1} \subseteq V_{r-1},\ |S_{r-1}| = t$$
$$|\Gamma_{V_r}(\cup_{i=1}^{r-1} S_i)| \geq (d-\varepsilon)^{(r-1)t}\tilde{n}$$

Finding $T_{rt,r}$

$\exists S_r \subseteq N_{V_r}(\cup_{i=1}^{r-1} S_i),\ \ |S_r| = t$
and thus $G[S_1 \cup \cdots \cup S_r]$ contains a $T_{rt,r}$ provided

$$(d-\varepsilon)^{(r-1)t}\tilde{n} \geq t$$

Strongest of the blue conditions:

$$(d-\varepsilon)^{(r-1)t-1} \geq \varepsilon$$

Let's not forget:

$$d + 2\varepsilon + \frac{1}{m} < \frac{\gamma}{2}$$

Choose, for example: $m \geq \frac{6}{\gamma}$ *
$d = \frac{\gamma}{6}$
$\varepsilon = \min\left\{\left(\frac{d}{2}\right)^{t(r-1)}, \frac{1}{t(r-1)}\right\}$

Green conditions are satisfied by choosing a large enough threshold vertex number $N = N(r,t,\gamma)$.

$$r,t,\gamma \ \rightsquigarrow\ m,d,\varepsilon \ \rightsquigarrow\ N$$

*We also needed large m earlier for using Turán's Theorem.

The Erdős-Turán conjecture

A set S of positive integers is k-AP-free if $\{a, a+d, a+2d, \ldots, a+(k-1)d\} \subseteq S$ implies $d = 0$.

$s_k(n) = \max\{|S| : S \subseteq [n] \text{ is } k\text{-AP-free}\}$

How large is $s_k(n)$? Could it be linear in n?

Erdős-Turán Conjecture (Szemerédi's Theorem)
For every constant k, we have

$$s_k(n) = o(n).$$

Construction (Erdős-Turán, 1936)

$$s_3(n) \geq n^{\frac{\log 2}{\log 3}}.$$

$S = \{s : \text{ there is no 2 in the ternary expansion of } s\}$

S is 3-AP-free. For $n = 3^l$, $|S \cap [n]| = 2^l$

Roth's Theorem (1953) $s_3(n) = o(n)$.

Applications of the Regularity Lemma

Removal Lemma For $\forall\gamma > 0\ \exists\delta = \delta(\gamma)$ such that the following holds. Let G be an n-vertex graph such that at least $\gamma\binom{n}{2}$ edges has to be deleted from G to make it triangle-free. Then G has at least $\delta\binom{n}{3}$ triangles.

Proof. Apply Regularity Lemma (Homework).

Roth's Theorem For $\forall\epsilon > 0\ \exists N = N(\epsilon)$ such that for any $n \geq N$ and $S \subseteq [n]$, $|S| \geq \epsilon n$,
there is a three-element arithmetic progression in S.

Proof. Create a tri-partite graph $H = H(S)$ from S.

$$V(H) = \{(i,1) : i \in [n]\} \cup \{(j,2) : j \in [2n]\}$$
$$\cup\{(k,3) : k \in [3n]\}$$

$(i,1)$ and $(j,2)$ are adjacent if $j - i \in S$
$(j,2)$ and $(k,3)$ are adjacent if $k - j \in S$
$(i,1)$ and $(k,3)$ are adjacent if $k - i \in 2S$

Roth's Theorem — Proof cont'd

$(i,1),(i+x,2),(i+2x,3)$ form a triangle for every $i \in [n]$, $x \in S$.
These $|S|n$ triangles are pairwise edge-disjoint.

$\Downarrow$

At least $\epsilon n^2 \geq \frac{\epsilon}{18}\binom{|V(H)|}{2}$ edges must be removed from H to make it triangle-free.

Let $\delta = \delta\left(\frac{\epsilon}{18}\right)$ provided by the Removal Lemma.
There are at least $\delta\binom{|V(H)|}{3}$ triangles in H.

S has no three term arithmetic progression

$\Downarrow$

$\{(i,1),(j,2),(k,3)\}$ is a triangle iff $j-i = k-j \in S$.
Hence the number of triangles in H is equal to $n|S| \leq n^2 < \delta\binom{6n}{3}$, provided $n > N(\epsilon) := \left\lfloor \frac{1}{\delta} \right\rfloor$. □

Behrend's Construction

Construction (Behrend, 1946)

$$s_3(n) \geq n^{1-O\left(\frac{1}{\sqrt{\log n}}\right)}.$$

Construct set of vectors $\bar{a} = (a_0, a_1, \ldots, a_{l-1})$:

$V_k = \{\bar{a} \in \mathbb{Z}^l \; : \; \|\bar{a}\|^2 = k, \; 0 \leq a_i < \frac{d}{2}$ for all $i < q\}$,

where $\|\bar{a}\| = \sqrt{\sum_{i=0}^{l-1} a_i^2}$.

Interpret a vector $\bar{a} \in \{0,1,\ldots,d-1\}^l$ as an integer written in d-ary:

$$n_{\bar{a}} = \sum_{i=0}^{l-1} a_i d^i.$$

Let

$$S_k = \{n_{\bar{a}} : \bar{a} \in V_k\}$$

Claim $S_k \subseteq [d^l]$ is 3-AP-free for every k.

Proof. Assume $n_{\bar{a}} + n_{\bar{b}} = 2n_{\bar{c}}$.
Then $a_i + b_i = 2c_i$ for every $i < l$, because $a_i + b_i$ and $2c_i$ are both $< d$ (so there is no carry-over)
So $\bar{a} + \bar{b} = 2\bar{c}$. But

$$\|2\bar{c}\| = 2\|\bar{c}\| = 2\sqrt{k} = \|\bar{a}\| + \|\bar{b}\| \geq \|\bar{a} + \bar{b}\|,$$

and equality happens only if $\bar{a}$ and $\bar{b}$ are parallel. Since they are of the same length, we conclude $\bar{a} = \bar{b}$. □

Take the *largest* S_k. Bound its size by averaging:

$\bar{a} \in \{0,1,\ldots,d-1\}^l \Rightarrow \|\bar{a}\|^2 < ld^2$,
so there is a k for which

$$|S_k| \geq \frac{|\bigcup_i S_i|}{ld^2} = \frac{(d/2)^l}{ld^2} = \frac{d^{l-2}}{2^l l}$$

For given n, choose $l = \sqrt{\log n}$ and $d = n^{\frac{1}{l}}$.

Hypergraph Turán numbers I – 4-clique

What would be the smallest meaningful clique to generalize Turán's Theorem for in k-uniform hypergraphs with $k > 2$? It is $K_4^{(3)}$.

Construction Let $3|n$. Partition $V_0 \cup V_1 \cup V_2 = [n]$ with $|V_0| = |V_1| = |V_2| = \frac{n}{3}$. Let $\mathcal{H}$ be 3-uniform:
$E(\mathcal{H}) = \{T : |T \cap V_i| = 1 \text{ for all } i = 0, 1, 2\} \cup$
$\{T : |T \cap V_i| = 2, |T \cap V_{i+1}| = 1 \text{ for some } i = 0, 1, 2\}$

Proposition $\mathcal{H}$ contains no copy of $K_4^{(3)}$.

For an k-uniform hypergraph $\mathcal{K}$, let $ex(n, \mathcal{K})$ be the largest number m such that there exists a $\mathcal{K}$-free k-uniform hypergraph on n vertices with m edges.

Consequence $ex(n, K_4^{(3)}) \geq \frac{5}{9}\binom{n}{3}$

Turán's Conjecture ($1000 dollar question)

$$ex(n, K_4^{(3)}) = |E(\mathcal{H})|$$

Remark If conjecture is true, then there are exponentially many extremal constrcutions (Kostochka).

Hypergraph Turán numbers II — Fano plane

Let $\mathcal{F}$ be the 3-uniform hypergraph defined on $V(\mathcal{F}) = [7]$ with $E(\mathcal{F}) = \{123, 345, 561, 174, 376, 572, 246\}$.

Remark $\mathcal{F}$ is called the "Fano plane" (It is the projective plane over the field $\mathbb{F}_2$). Its sets have the nice property that any two of them interesct in exactly 1 element.

A coloring of the vertices of a hypergraph $\mathcal{H}$ is proper if no edge is monochromatic.

Proposition $\mathcal{F}$ is not properly 2-colarable.

Construction Let $\mathcal{H}$ be the 2-colorable hypergraph with the most edges: Partition $V_1 \cup V_2 = [n]$ with $|V_1| = \lfloor \frac{n}{2} \rfloor$ and $|V_2| = \lceil \frac{n}{2} \rceil$.
$E(\mathcal{H}) = \{T \in \binom{[n]}{3} : T \cap V_i \neq \emptyset \text{ for both } i = 1, 2\}$

Claim $\mathcal{H}$ contains no copy of $\mathcal{F}$.

Proof. $\mathcal{F}$ is not 2-colorable. □

Theorem (De Caen-Füredi, Keevash-Sudakov, Füredi-Simonovits, 2006) $ex(n, \mathcal{F}) = |E(\mathcal{H})|$

Posets

$(P, \leq)$ is a poset if the relation $\leq$ on P is

- reflexive ($a \leq a$ for all $a \in P$)
- antisymmetric ($a \leq b$ and $b \leq a \Rightarrow a = b$)
- transitive ($a \leq b$ and $b \leq c \Rightarrow a \leq c$)

a and b are comparable if $a \leq b$ or $b \leq a$. Otherwise a and b are incomparable.

$C \subseteq P$ is a chain if any two elements are comparable.

$A \subseteq P$ is an antichain if no two elements are comparable.

Min-max statement for max-chains

A partition $\mathcal{C} = \{C_1, \ldots, C_l\}$ of P is a chain partition of P if all C_is are chains.

A partition $\mathcal{A} = \{A_1, \ldots A_k\}$ is an antichain partition of P if all A_is are antichains.

Proposition $\max\{|C| : C \text{ is a chain}\} =$
$\min\{|\mathcal{A}| : \mathcal{A} \text{ is an antichain partition of } P\}$

Proof. $\boxed{\leq}$ is immediate.
$\boxed{\geq}$ The set $A = \{x \in P : x \not\leq y \text{ for all } y \in P\}$ of maximum elements forms an antichain, that intersects every maximal chain of P.
So if P has maximum chain size M, then $P \setminus A$ has maximum chain size at most $M - 1$ (in fact equal).
By induction, find a partition of $P \setminus A$ into $M - 1$ antichains and extend it by A to get a partition of P into M antichains. □

Min-max statement for max-antichains

Dilworth's Theorem $\max\{|A| : A \text{ is an antichain}\} = \min\{|\mathcal{C}| : \mathcal{C} \text{ is a chain partition of } P\}$

Proof. (Tverberg) $\boxed{\leq}$ is again immediate.
$\boxed{\geq}$ If there is a chain, that interesects every maximal antichain of P, then we proceed by induction as in the Proposition.
Otherwise let C be a maximal chain, that does not intersect the chain $A = \{a_1, \ldots, a_M\}$ of maximum size M. Let

$$\begin{aligned} A^- &= \{x \in P : x \leq a_i \text{ for some } i\} \\ A^+ &= \{x \in P : x \leq a_i \text{ for some } i\} \end{aligned}$$

- $A^- \cap A^+ = A$ because A is antichain
- $A^- \cup A^+ = P$ because A is maximal.

Apply induction on A^- and on A^+.

For this note that

$$\begin{aligned} A^- \neq P &\Leftarrow \max C \in A^+ \setminus A \Leftarrow C \text{ is maximal} \\ A^+ \neq P &\Leftarrow \min C \in A^- \setminus A \Leftarrow C \text{ is maximal} \end{aligned}$$

Obtain

a chain partition $C_1^-, \ldots, C_M^-$ of A^- and
a chain partition $C_1^+, \ldots, C_M^+$ of A^+, such that
$C_i^- \cap A = \{a_i\} = C_i^+ \cap A$ for all i.

Then $C_1^- \cup C_1^+, \ldots, C_M^- \cup C_M^+$ is a partition of P into M chains. □

Extremal set theory — the classics I

The width of a poset is the size of the largest antichain.

$(2^{[n]}, \subseteq)$ is the Boolean poset.

Sperner's Theorem The width of the Boolean poset is $\binom{n}{\lfloor n/2 \rfloor}$.

Reformulation: How many subsets of $[n]$ can be select if it is forbidden to select two sets such that one is subset of the other?

You can select all $\binom{n}{k}$ subsets of a given size k: they certainly satisfy the property.
$k = \lfloor \frac{n}{2} \rfloor$ maximizes their number.

Sperner's Theorem If $\mathcal{F} \subseteq 2^{[n]}$ is a family of subsets such that for every $A, B \in \mathcal{F}$ we have $A \not\subseteq B$ then

$$|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

Permutation method

Proof. Count permutations $\pi \in S_n$ of $[n]$ which have an initial segment from $\mathcal{F}$. Formally, double-count

$$M = |\{(\pi, F) : \pi \in S_n, F \in \mathcal{F}, F = \{\pi(1), \ldots, \pi(|F|)\}\}|$$

For every $F \in \mathcal{F}$ there are $|F|!(n - |F|)!$ permutations $\pi \in S_n$ with $\{\pi(1), \ldots, \pi(|F|)\} = F$. So

$$M = \sum_{F \in \mathcal{F}} |F|!(n - |F|)!.$$

For every $\pi \in S_n$ there is at most one k such that $\{\pi(1), \ldots, \pi(k)\} \in \mathcal{F}$.

So $M \leq n!$.

Hence

$$\sum_{F \in \mathcal{F}} |F|!(n - |F|)! \leq n!$$

$$1 \geq \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}} \geq \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = |\mathcal{F}| \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$$

Extremal set theory — the classics II

Proposition Let $\mathcal{F} \subseteq 2^{[n]}$ such that any two members of $\mathcal{F}$ have a nonempty intersection. Then

$$|\mathcal{F}| \leq 2^{n-1}.$$

Construction Proposition is best possible: Take all sets containing the element 1.

What if we restrict the sizes of the sets: all members must be of size k.
Taking all sets of size k that contains 1 gives $\binom{n-1}{k-1}$ sets. Is this again best possible?

Theorem (Erdős-Ko-Rado) Let $k, n \in I\!N$, $1 \leq k \leq n/2$. If $\mathcal{F} \subseteq \binom{[n]}{k}$ such that any two members of $\mathcal{F}$ have a nonempty intersection. Then

$$|\mathcal{F}| \leq \binom{n-1}{k-1}.$$

Permutation method II

Proof. (Katona) C_n: set of cyclic permutations of $[n]$.

$$|C_n| = (n-1)!$$

Double-count
$M = |\{(\phi, F) : \phi \in C_n, F \in \mathcal{F}$ is a segment in $\phi\}|$

For $F \in \mathcal{F}$, let $C_F \subseteq C_n$ set of those cyclic permutations that contain F as a segment. $M = \sum_{F \in \mathcal{F}} |C_F|$.

$|C_F| = k!(n-k)! \Longrightarrow |\mathcal{F}|k!(n-k)! = M$.

Claim Every cyclic permutation can contain at most k different $F \in \mathcal{F}$ as a segment.

Claim $\Longrightarrow M \leq |C_n|k = (n-1)!k$.

$$\begin{aligned} |\mathcal{F}|k!(n-k)! &\leq (n-1)!k \\ |\mathcal{F}| &\leq \frac{(n-1)!k}{k!(n-k)!} \end{aligned}$$

Classics III — Sunflowers

A family $\mathcal{F}$ of sets is called k-uniform if every member is a k-elements set.

Family $\mathcal{S}$ is a sunflower (or Δ-system) if $A \cap B = \cap_{F \in \mathcal{S}} F$ for every $A, B \in \mathcal{S}$. The set $\cap_{F \in \mathcal{S}} F$ is called the core of the sunflower and $F \backslash \cap_{F \in \mathcal{S}} F$ are its petals.

Theorem (Erdős-Rado) $\mathcal{F}$ is an l-uniform family and $|\mathcal{F}| \geq 2^l l!$ then $\mathcal{F}$ contains a sunflower with three petals

Construction $X = \{x_1, \ldots, x_l, y_1, \ldots, y_l\}$
Define $\mathcal{F} = \{F \subseteq X : |F \cap \{x_i, y_i\}| = 1$ for every $i\}$.
$\mathcal{F}$ has no sunflower with three petals and $|\mathcal{F}| = 2^l$.

There are better constructions with C^l members where C is some constant > 2. But no superexponential construction is known

The best known upper bound (Kostochka) is slightly below $l!$.

$1000 dollar question: Is there an l-uniform family containing no sunflower with three petals, which has superexponential size (in l)?

Proof. Induction on l. For $l = 1$ we can have at most two one-element subsets.

Let $l > 1$.
There exist a set X of at most $2l$ elements that every $F \in \mathcal{F}$ intersect X (Take two disjoint members of $\mathcal{F}$ if they exist, otherwise take any one member of $\mathcal{F}$.)

$\mathcal{F}_x = \{F \setminus \{x\} : F \in \mathcal{F}, x \in F\}$ is an $(l-1)$-uniform family containing no sunflower with three petals, for every $x \in X$.

By induction $|\mathcal{F}_x| \leq 2^{l-1}(l-1)!$ for every $x \in X$.

Then

$$|\mathcal{F}| \leq \sum_{x \in X} |\mathcal{F}_x| \leq |X| \cdot (2^{l-1}(l-1)!) \leq 2^l l!.$$

Oddtown/Eventown

Eventown: $\mathcal{F} \subseteq [n]$ is an Eventown-family of sets if

- $|F| \equiv 0 \pmod 2$ for all $F \in \mathcal{F}$ and
- $|F_1 \cap F_2| \equiv 0 \pmod 2$ for every $F_1, F_2 \in \mathcal{F}$

How large can $|\mathcal{F}|$ be? As large as $2^{\lfloor n/2 \rfloor}$

Construction. For even n:

$\mathcal{F} = \{F \subseteq [n] : |F \cap \{2i-1, 2i\}|$ is even for all $i \in [\frac{n}{2}]\}$

Oddtown: $\mathcal{F} \subseteq [n]$ is an Oddtown-family of sets if

- $|F| \equiv 1 \pmod 2$ for all $F \in \mathcal{F}$ and
- $|F_1 \cap F_2| \equiv 0 \pmod 2$ for every $F_1 \neq F_2 \in \mathcal{F}$

How large can $|\mathcal{F}|$ be?

Oddtown Theorem The maximum size of an Oddtown-family over $[n]$ is n.

Proof. Let $\mathcal{F} = \{F_1, \ldots, F_m\} \subseteq 2^{[n]}$ be an Oddtown-family.

Let $\mathbf{v_i} \in \{0,1\}^n$ be the characteristic vector of F_i: j^{th} coordinate is 1 if $j \in F_i$, otherwise 0.

Crucial property: $\mathbf{v_i}^T \mathbf{v_j} = |F_i \cap F_j|$

Claim $\mathbf{v_1}, \ldots, \mathbf{v_n}$ is linearly independent over $\mathbb{F}_2$.

Let $\lambda_1 \mathbf{v_1} + \cdots + \lambda_m \mathbf{v_m} = 0$

Then for every i

$$\begin{aligned} 0 &= (\lambda_1 \mathbf{v_1} + \cdots + \lambda_m \mathbf{v_m})^T \mathbf{v_i} \\ &= \lambda_1 \mathbf{v_1}^T \mathbf{v_i} + \cdots + \lambda_i \mathbf{v_i}^T \mathbf{v_i} + \cdots \lambda_m \mathbf{v_m}^T \mathbf{v_i} \\ &= \lambda_i \end{aligned}$$

Since $\mathbf{v_1}, \ldots \mathbf{v_m}$ are linearly independent vectors in an n-dimensional space, $m \leq n$. □

Explicit Ramsey graphs I

A clique or an independent set of a graph G is called a *homogenous set*.

A graph is k-*Ramsey* if it does not contain a homogenous set of order k. (Remark: instead of RED/BLUE-coloring we formulate in terms of edge/non-edge.)

We know that the largest k-Ramsey graph has at least $\sqrt{2}^k$ vertices. ($R(k) \geq \sqrt{2}^k$.)

BUT: can you give one such beast in my hand?

Sure: go over all the $2^{\binom{n}{2}}$ graphs on $n = \sqrt{2}^k$ vertices and check whether their clique number and independence number are below k (Never mind that these are NP-hard problems).
Eventually you'll find a k-Ramsey graph.

Explicit Ramsey graphs II

Why are you not happy with this "construction"?

It takes too much time.

What is then a constructive k-Ramsey graph?

Its adjacency matrix should be constructible in time polynomial in its number of vertices n.
Or even stronger: adjacency of any two vertices should be decidable in time polynomial in $\log n$ (what it takes to write down the label of the two vertices).

Turán construction: $T_{(k-1)^2, k-1}$ has no clique and no independent set of order k.
Has $(k-1)^2$ vertices.

Anything more than quadratic?

The Linear Algebra bound

The key to the Oddtown proof is the following simple observation:

Linear Algebra bound If $\mathbf{v_1}, \ldots, \mathbf{v_m}$ are a set of linearly independent vectors belonging to the span of the vectors $\mathbf{u_1}, \ldots \mathbf{u_k}$, then $m \leq k$.

The Gram matrix $M = (m_{ij})$ of a set of vectors $\mathbf{v_1}, \ldots, \mathbf{v_m}$ is defined by $m_{ij} = \mathbf{v_i}^T \mathbf{v_j}$.

Proposition Vectors $\mathbf{v_1}, \ldots, \mathbf{v_m} \in \mathbb{F}^n$ are linearly independent iff their Gram matrix over $\mathbb{F}$ is nonsingular.

Proof of Oddtown Thm. The Gram matrix of the characteristic vectors of an Oddtown family over $\mathbb{F}_2$ is the identity matrix. Then Apply Linear Algebra bound.$\square$

Construction of Nagy

$V(G) = \binom{[k]}{3}$,
$E(G) = \{AB : |A \cap B| = 1\}$

Theorem G has no homogenous set of order $k+1$.

Proof. An independent set of G is an Oddtown family $\Rightarrow \quad \alpha(G) \leq k$.

For the vertices $C_1, \ldots, C_m$ of a clique, we have that

- $|C_i| = 3$ for all i and
- $|C_i \cap C_j| = 1$ for every $i \neq j$

Hence the Gram matrix of the characteristic vectors is $J_n + 2I_n$ (J_n is the all 1 matrix, I_n is the identity)

M is nonsingular over $I\!R \Rightarrow \mathbf{v_1}, \ldots, \mathbf{v_m} \in R^n$ are linearly independent $\Rightarrow m \leq n$.

So $\omega(G) \leq n$. $\square$

G is a k-Ramsey graph on $\Theta(k^3)$ vertices.

HW: Linear Algebra-free proof (like original)

Construction of Abbott I

What if you want a k-Ramsey graph on k^{100} vertices?

Suppose you got **one** such graph G on k_0^{100} vertices that is k_0-Ramsey. (where k_0 is a constant)

Product graph $K \times H$

$V(K \times H) = V(K) \times V(H)$,
$E(K \times H) = \{(k,h)(k',h') :$ either $h = h', kk' \in E(K)$ or $hh' \in E(H)\}$

Claim $\alpha(G \times H) = \alpha(G) \cdot \alpha(H)$
$\omega(G \times H) = \omega(G) \cdot \omega(H)$

Proof. HW

The powers of G provide constructions of k-Ramsey graphs for an infinite sequence of k:
By Claim G^i is k_i-Ramsey on k_i^{100} vertices where $k_i = k_0^i$.

Given the adjacency relations of G, the adjacency relation in the i^{th} power can be computed very fast.

Construction of Abbott II

So how do we get our "starter" G?

We KNOW there is a k-Ramsey graph on $\sqrt{2}^k$ vertices. So let k_0 be the smallest integer that $\sqrt{2}^{k_0} > k_0^{100}$. Check all graphs on k_0^{100} vertices, one of them is k_0-Ramsey.

How long does this take?

Nothing ... Only CONSTANT time. (since k_0 is a constant depending only on 100.) $\square$

Philosophical question: Is this a "construction"?

Generalized Oddtown

Replacing $\mathbb{F}_2$ in the Oddtown proof with $\mathbb{F}_p$ in the proof immediately gives the

$\mathbf{mod}\,\mathbf{p}$-town Theorem. Let p be a prime number and $\mathcal{F} \subseteq 2^{[n]}$ be a family such that

- $|F| \not\equiv 0 \pmod p$ for all $F \in \mathcal{F}$ and
- $|F_1 \cap F_2| \equiv 0 \pmod p$ for every $F_1 \neq F_2 \in \mathcal{F}$

Then $|\mathcal{F}| \leq n$.

Even More Generalized Oddtown

Theorem ("Nonuniform modular RW-Theorem", Frankl-Wilson, 1981; Deza-Frankl-Singhi, 1983)
Let p be a prime, and L be a set of s integers.
Let $B_1, \ldots, B_m \in 2^{[n]}$ be a family such that

- $|B_i| \notin L \pmod p$
- $|B_i \cap B_j| \in L \pmod p$ for every $i \neq j$.

Then

$$m \leq \sum_{i=0}^{s} \binom{n}{i}.$$

Remark RW stands for Ray-Chaudhuri and Wilson.

Remark Oddtown Theorem: $p = 2$, $L = \{0\}$.
The statement only gives $m \leq n + 1$, but the proof will give $m \leq n$ (because $L = \{0\}$)

Generalizing linear independence of vectors

Let $\mathbb{F}$ be a field and Ω an arbitrary set.Then the set $\mathbb{F}^\Omega = \{f : \Omega \to \mathbb{F}\}$ of functions is a *vector space over* $\mathbb{F}$.

Lemma Let $\Omega \subseteq \mathbb{F}^n$. If $f_1, \ldots, f_m \in \mathbb{F}^\Omega$ and there exist $\mathbf{v_1}, \ldots, \mathbf{v_m} \in \Omega$ such that

- $f_i(\mathbf{v_i}) \neq 0$, and
- $f_i(\mathbf{v_j}) = 0$ for all $j < i$,

then $f_1, \ldots, f_m$ are linearly independent in $\mathbb{F}^\Omega$.

Proof. Suppose $\lambda_1 f_1 + \cdots + \lambda_m f_m = 0$, and let j be the smallest index with $\lambda_j \neq 0$. Substituting $\mathbf{v_j}$ into this function equation we have

$$\underbrace{\lambda_1 f_1(\mathbf{v_j}) + \cdots + \lambda_{j-1} f_{j-1}(\mathbf{v_j})}_{=0,\text{ since } \lambda_i = 0,\, i<j} + \underbrace{\lambda_j f_j(\mathbf{v_j})}_{\neq 0}$$
$$+ \underbrace{\lambda_{j+1} f_{j+1}(\mathbf{v_j}) + \cdots + \lambda_m f_m(\mathbf{v_j})}_{=0,\text{ since } f_i(\mathbf{v_j})=0,\, j<i} = 0,$$

a contradiction. □

Proof of Even More Generalized Oddtown

For each set B_i, let $\mathbf{v_i} \in \mathbb{F}_p$ be its characteristic vector. For $\mathbf{x} = (x_1, \ldots, x_n)$ let

$$f_i(\mathbf{x}) = \prod_{l \in L} (\mathbf{x}^T \mathbf{v_i} - l).$$

Clearly,

$$f_i(\mathbf{v_j}) \begin{cases} \neq 0 & \text{if } i = j \\ = 0 & \text{if } i \neq j \end{cases}$$

So the functions $f_1, \ldots, f_m$ are linearly independent in the subspace they generate in $\mathbb{F}_p[x_1, \ldots, x_m]$. What is the dimension?

Each f_i is the product of s linear functions in n variables. Expanding the parenthesis: f_i is the linear combination of terms of the form $x_1^{s_1} \cdot \cdots \cdot x_n^{s_n}$ with $s_1 + \cdots + s_n = s$?
How many terms like that are there?

Much more than we can afford ...

Multilinearization

We need another trick to reduce the dimension. We use that our vectors (witnessing the linear independence in the Lemma) have only 0 or 1 coordinates.

From f_i define $\tilde{f}_i$ by expanding the product and replacing each power x_i^k by a term x_i for every $k \geq 1$ and i, $1 \leq i \leq m$.

Since $0^k = 0$ and $1^k = 1$ for every $k \geq 1$ we have that $f_i(\mathbf{v_j}) = \tilde{f}_i(\mathbf{v_j})$ for every i, j.

The properties of the functions and vectors remains valid, so the (now) multi**linear** polynomials $\tilde{f}_1, \ldots, \tilde{f}_m$ of total degree s are also linearly independent.

They live in a space spanned by the basic monomials $\prod_{j=1}^{k} x_{i_j}$ of degree at most s. Their number is at most

$$\binom{n}{s} + \binom{n}{s-1} + \cdots + \binom{n}{1} + \binom{n}{0}. \qquad \square$$

The nonmodular version

Theorem (“Nonuniform RW-Theorem”, Frankl-Wilson, 1981) Let L be a set of s integers.
Let $B_1, \ldots, B_m \in 2^{[n]}$ be a family such that $|B_i \cap B_j| \in L$ for every $i \neq j$. Then

$$m \leq \sum_{i=0}^{s} \binom{n}{i}.$$

Proof. Let $|B_1| \leq \cdots \leq |B_m|$ and let $\mathbf{v}_i$ is the characteristic vector of B_i.

We now work over the **reals!**. Define $f_i : \mathbb{R}^n \to \mathbb{R}$:

$$f_i(\mathbf{x}) = \prod_{l \in L, l < |B_i|} (\mathbf{x} \cdot \mathbf{v}_i - l).$$

We have

$$f_i(\mathbf{v_j}) \begin{cases} \neq 0 & \text{if } i = j \\ = 0 & \text{if } i > j \end{cases}$$

Finish similarly as in the Even More Generalized Oddtown Theorem: f_i are linearly independent and so are their multilinearized versions $\tilde{f}_i$. $\square$

Frankl-Wilson graph

$V(G) = \binom{[n]}{p^2-1}$ with $n = p^3 - 1$
$E(G) = \{AB : |A \cap B| \equiv -1 \pmod{p}\}$

Remark $p = 2 \rightsquigarrow$ Nagy graph

No large clique: A clique $B_1, \ldots, B_m$ is a uniform family where every pairwise intersection is in the set $L = \{p-1, 2p-1, \ldots, p^2-p-1\}$, with $|L| = p-1$. By the nonmodular version: $m \leq \sum_{i=0}^{p-1} \binom{n}{i} \approx p^{2p}$

No large independent set: An independent set $A_1, \ldots, A_m$ is a uniform family where every pairwise intersection size $\pmod{p}$ is in the set $L = \{0, 1, \ldots, p-2\}$, with $|L| = p-1$, while $|A_i| = p^2 - 1 \notin L \pmod{p}$. By the EMGOT $m \leq \sum_{i=0}^{p-1} \binom{n}{i} \approx p^{2p}$

Theorem (Frankl-Wilson) The graph G above is k-Ramsey and has a vertex set of order

$$k^{O\left(\frac{\ln k}{\ln \ln k}\right)}.$$

Proof. Calculate (HW)

Chromatic number of the unit-distance graph

G_n is the n-dimensional unit distance graph.

$V(G_n) = I\!R^n$
$E(G_n) = \{\mathbf{xy} : \|\mathbf{x} - \mathbf{y}\| = 1\}$

$1000 dollar question: What is the chromatic number of the plane?
We know $4 \le \chi(G_2) \le 7$. (HW)

Hadwiger-Nelson problem How fast does $\chi(G_n)$ grow?

Claim $\chi(G_n) \le n^{n/2}$. (HW)
$\chi(G_n) \ge n + 1$. (simplex with unit sidelength)

Larman-Rogers (1972) $\chi(G_n) \le const^n$ (HW)
$\chi(G_n) = \Omega(n^2)$

Remark. Clearly, unit-distance plays no special role here. $G_n \cong G_n^\delta$ where G_n^δ is the "δ-distance graph":
$V(G_n^\delta) = I\!R^n$
$E(G_n^\delta) = \{\mathbf{xy} : \|\mathbf{x} - \mathbf{y}\| = \delta\}$

The growth of $\chi(G_n)$ is exponential

Theorem (Frankl-Wilson, 1981) $\chi(G_n) \ge \Omega(1.1^n)$.

Proof. **Goal**: For some distance $\delta > 0$ we find a subgraph $H_n \subseteq G_n^\delta$ with $\alpha(H) \le \frac{|V(H)|}{1.1^n}$.

Key: If $\mathbf{v}_A$ and $\mathbf{v}_B \in I\!R^n$ are the characteristic vectors of sets A and $B \in 2^{[n]}$, then distance of $\mathbf{v}_A$ and $\mathbf{v}_B$ is equal to $\sqrt{|A \triangle B|}$.

If $A, B \in \mathcal{F} \subseteq \binom{[d]}{k}$ are members of a **uniform** family $\mathcal{F}$, then the distance of $\mathbf{v}_A$ and $\mathbf{v}_B$ depends on the intersection size: $\|\mathbf{v}_A - \mathbf{v}_B\| = \sqrt{2(k - |A \cap B|)}$.

an independent set in G_n^δ avoids distance $\delta \rightsquigarrow$
a uniform family avoiding a certain intersection size.

We give a family $\mathcal{F}$ where any subfamily $\mathcal{F}' \subseteq \mathcal{F}$, whose members **avoid** a certain intersection size, is small compared to $|\mathcal{F}|$:

Let $\mathcal{F} := \binom{[4p-1]}{2p-1}$, where p is a prime.
Then pairwise intersection size $p - 1$ is hard to avoid.

Avoiding a certain intersection size

Theorem Let p be a prime number. If $\mathcal{F}' \subseteq \binom{[4p-1]}{2p-1}$ such that for all $A, B \in \mathcal{F}'$ we have $|A \cap B| \ne p - 1$, then

$$|\mathcal{F}| \le 2 \cdot \binom{4p-1}{p-1} < 1.76^n.$$

Proof. Consequence of Generalized Oddtown with $L = \{0, 1, 2, \dots p - 2\}$. □

Let $n = 4p - 1$, $k = 2p - 1$.
Let $\delta = \sqrt{2(2p - 1 - (p - 1))} = \sqrt{2p}$ and define $H \subseteq G_d^\delta$ by $V(H) = \{\mathbf{v}_A : A \in \binom{[n]}{k}\}$.
Then distance δ is hard to avoid in $V(H)$:

$$\alpha(H) \le 1.76^n < \frac{\binom{4p-1}{2p-1}}{1.1^n}.$$

□

Remark Optimizing parameters gives $\chi(G_n) \ge \Omega(1.2^n)$.

Borsuk's Conjecture

"Dead at the age of 60. Died after no apparent signs of illness, unexpectedly, of grave combinatorial causes."
Epitaph of Babai & Frankl for Borsuk's Conjecture

Borsuk's Conjecture (1933) Every set of diameter 1 in $I\!R^d$ can be partitioned into $d + 1$ sets of smaller diameter.

verified for:
– all bodies in dimensions 2 and 3
– centrally symmetric bodies
– bodies with smooth surface

General conjecture is not only dead, but *very dead*.

Theorem (Kahn-Kalai, 1992) Let $f(d)$ denote the minimum number such that every set of diameter 1 in $I\!R^d$ can be partitioned into $f(d)$ pieces of smaller diameter. Then

$$f(d) \ge 1.2^{\sqrt{d}}.$$

Avoiding the **smallest** pairwise intersection

Proof of the Kahn-Kalai Theorem.

Partitioning a pointset $V(H)$ into sets of smaller diameter $\rightsquigarrow$ partitioning the graph $H \subseteq G_d^\delta$ into independent sets in the δ-distance graph where δ is the **largest distance**.

largest distance $\rightsquigarrow$ smallest pairwise intersection

We can already make it hard to avoid **some** intersection size.
How can we make it hard to avoid the **smallest** intersection size?

$n = 4p - 1$, $k = 2p - 1$

$$A \in \binom{[n]}{k} \rightsquigarrow S_A \in \binom{\binom{[n]}{2}}{k(n-k)}$$

$$S_A := \{T \subseteq [n] : |T| = 2, |T \cap A| = 1\}$$

Key:

$$\begin{aligned}|S_A \cap S_B| &= |A \cap B||X \setminus (A \cup B)| + |A \setminus B||B \setminus A| \\ &= k^2 + (n-4k)|A \cap B| + 2|A \cap B|^2\end{aligned}$$

Minimized (for integers) at $|A \cap B| = \left\lfloor \frac{4k-n}{4} \right\rfloor = p - 1$.

Family $\mathcal{S}_\mathcal{F} = \{S_A : A \in \mathcal{F}\} \subseteq 2^{[d]}$ where $d = \binom{n}{2}$.

$|\mathcal{S}_\mathcal{F}| = |\mathcal{F}| = \binom{n}{k} > 1.99^n$

Every subfamily that avoids the smallest intersection has size at most 1.76^n.

So one needs to partition $\mathcal{S}_\mathcal{F}$ into at least $1.1397^n > 1.1^{\sqrt{2d}}$ pieces. □

Proof of Eventown bound

Let $\{F_1, \ldots, F_m\}$ be an Eventown family.
Let $V_\mathcal{F} = \langle \mathbf{v}_1, \ldots, \mathbf{v}_m \rangle \le \mathbb{F}_2^n$ be the linear space spanned by the characteristic vectors.
Let $\mathbf{u}_1, \ldots, \mathbf{u}_k \in V_\mathcal{F}$ be a basis of $V_\mathcal{F}$.

Let $U : \mathbb{F}_2^n \to \mathbb{F}_2^k$ be the linear function defined by $U(\mathbf{x}) = (\mathbf{x} \cdot \mathbf{u}_1, \ldots, \mathbf{x} \cdot \mathbf{u}_k)$.
$\mathbf{u}_1, \ldots, \mathbf{u}_k$ linearly independent, so $\dim im(U) = k$.

Claim $V_\mathcal{F} \subseteq ker(U)$
Proof. Any $\mathbf{x} \in V_\mathcal{F}$ is a linear combination of $\mathbf{v}_i$s, so is any $\mathbf{u}_i$. So

$$\mathbf{x} \cdot \mathbf{u}_i = \sum_{j,k} \alpha_j \beta_k \mathbf{v}_j \cdot \mathbf{v}_k = 0,$$

since by the Eventown rules $\Rightarrow \mathbf{v}_j \cdot \mathbf{v}_l = 0$ for every $1 \le j, l \le m$ □

$k = \dim(V_\mathcal{F}) \le \dim ker(U) = n - k$

$\dim(V_\mathcal{F}) \le \lfloor \frac{n}{2} \rfloor \Rightarrow m \le |V_\mathcal{F}| \le 2^{\lfloor n/2 \rfloor}$. □

Beautiful graphs

Want a graph that is regular and locally looks like tree: the first two levels of neighborhoods of **any** vertex is a tree; furthermore,
that's it: there are no more vertices in the graph.

Formally: r-regular
no C_3 and no C_4
$1 + r + r(r-1) = r^2 + 1$ vertices

When $r = 2 \rightsquigarrow C_5$
When $r = 3 \rightsquigarrow$ Petersen
When $r = 4 \rightsquigarrow$ no such animal (check it)
When $r = 5 \rightsquigarrow$ no such animal (it takes time)

Theorem (Hoffman-Singleton) If there exists an r-regular graph on $r^2 + 1$ vertices with girth five, then $r = 2, 3, 7$, or 57.

When $r = 7 \rightsquigarrow$ Hoffman-Singleton graph
When $r = 57 \rightsquigarrow$???

Adjacency matrices

Let G be a graph, $A = A(G)$ is its adjacency matrix:

$$(A)_{ij} = \begin{cases} 1 & \text{if } ij \in E(G) \\ 0 & \text{if } ij \notin E(G) \end{cases}$$

Claim $A + A(\bar{G}) = J_n - I_n$

Claim G is r-regular then $\mathbf{j} = (1, \dots, 1)$ is an eigenvector of A with eigenvalue r: $A\mathbf{j} = r\mathbf{j}$.

Theorem $(A^k)_{ij} =$ number of i, j-walks of length k in G

Proof. Induction on k.

Principal Axis Theorem If A is real symmetric, then there is an orthogonal basis consisting of eigenvectors.

Adjacency matrices of beautiful graphs

Observations

- $(A^2)_{ii} = r$ for $i \neq j$ (r-regularity).
- $(A^2)_{ij} \leq 1$ for $i \neq j$ (no C_4)
- $(A^2)_{ij} = 0$ for $ij \in E(G)$ (no C_3)
- $(A^2)_{ij} = 1$ for $ij \notin E(G)$
  (*Proof:* Let $ij \notin E(G)$.
  $|V(G)| = r^2 + 1$ and no C_3, no C_4
  $\Rightarrow\ V(G) \setminus (\{i\} \cup N(i)) = \cup_{u \in N(i)} N(u)$
  $\Rightarrow\ j \in N(u)$ for some $u \in N(i)$)

$$A^2 = rI + A(\bar{G}) = (r-1)I + J - A$$

Eigenvalues to the rescue

By the Principal Axis Theorem, there is an orthogonal eigenbasis $\mathbf{j}, \mathbf{v}_1, \dots, \mathbf{v}_{n-1}$. Let $\mathbf{v} \perp \mathbf{j}$ be an arbitrary eigenvector corresponding to eigenvalue λ. Then

$$\begin{aligned} (A^2 + A - (r-1)I - J)\mathbf{v} &= 0 \\ (\lambda^2 + \lambda - (r-1))\mathbf{v} &= 0 \\ \lambda^2 + \lambda - (r-1) &= 0 \end{aligned}$$

$\lambda_{1,2} = \frac{-1 \pm \sqrt{4r-3}}{2} = \frac{-1 \pm s}{2}$, where $s = \sqrt{4r-3}$

Let m_i be the multiplicity of eigenvalue λ_i.

The number of eigenvalues is n, so

$$1 + m_1 + m_2 = n = r^2 + 1.$$

By the invariance of the trace:

$$\begin{aligned} 0 &= \mathrm{Tr}A = r + m_1\lambda_1 + m_2\lambda_2 \\ &= r - \frac{1}{2}(m_1 + m_2) + \frac{s}{2}(m_1 - m_2) \\ &= r - \frac{1}{2}r^2 + \frac{s}{2}(m_1 - m_2) \end{aligned}$$

Cases

Case 1. s is irrational

$\Rightarrow\ m_1 - m_2 = 0\ \Rightarrow\ 0 = r^2 - 2r$

$\Rightarrow\ r = 2\ \rightsquigarrow\ C_5$.

Case 2. s is rational

$\Rightarrow\ s$ is integer

Substitute $r = \frac{s^2+3}{4}$

$$\frac{s^2+3}{4} - \frac{1}{2}\left(\frac{s^2+3}{4}\right)^2 + \frac{s}{2}(m_1 - m_2) = 0$$

$s^4 - 2s^2 + 16(m_1 - m_2)s - 15 = 0\ \Rightarrow\ s|15$

if $s = 1\ \Rightarrow\ r = 1\ \rightsquigarrow\ K_2$

if $s = 3\ \Rightarrow\ r = 3\ \rightsquigarrow$ Petersen graph

if $s = 5\ \Rightarrow\ r = 7\ \rightsquigarrow$ Hoffmann-Singleton graph

if $s = 15\ \Rightarrow\ r = 57\ \rightsquigarrow$?????? $\square$

Algorithmic methods: Baranyai's Theorem

$\chi'(K_n) = n-1$ is saying: $E(K_n)$ can be decomposed into pairwise disjoint perfect matchings.

k-uniform hypergraphs? $E(\mathcal{K}_n^{(k)}) = \binom{[n]}{k}$

Let $k|n$. $\mathcal{S} = \{S_1, \ldots, S_{n/k}\}$ is a "perfect matching in $\mathcal{K}_n^{(k)}$ if $S_i \cap S_j = \emptyset$ for $i \neq j$.

There are many perfect matchings in $\mathcal{K}_n^{(k)}$.

Is there a decomposition of $\binom{[n]}{k}$ into perfect matchings?

Not obvious already for $k = 3$ (Peltesohn, 1936)

$k = 4$ (Bermond)

Theorem (Baranyai, 1973) For every $k|n$, there is a decomposition of $\binom{[n]}{k}$ into perfect matchings.

Network flows

Network (D, s, t, c); D is a directed multigraph,
$s \in V(D)$ is the source, $t \in V(D)$ is the sink,
$c : E(D) \to I\!R^+ \cup \{0\}$ is the capacity.

Flow f is a function, $f : E(D) \to I\!R$

$f^+(v) := \sum_{v \to u} f(vu)$

$f^-(v) := \sum_{u \to v} f(uv)$.

Flow f is feasible if

(i) $f^+(v) = f^-(v)$ for every $v \neq s, t$ (conservation constraints), and

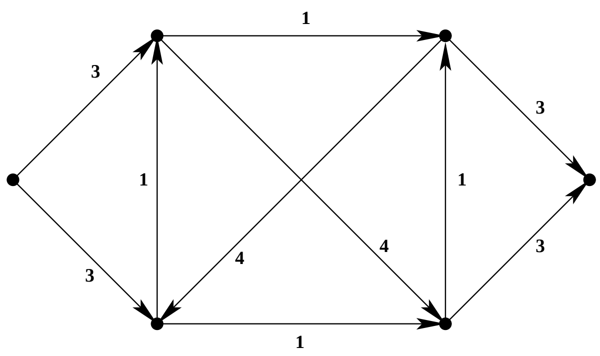
(ii) $0 \leq f(e) \leq c(e)$ for every $e \in E(D)$ (capacity constraints).

value of flow, $val(f) := f^-(t) - f^+(t)$.

maximum flow: feasible flow with maximum value

Example

0-flow



f-augmenting path

G: underlying undirected graph of network D

s,t-path P in G is an f-augmenting path, if
$s = v_0, e_1, v_1, e_2 \ldots v_{k-1}, e_k, v_k = t$ and for every e_i

(i) $f(e_i) < c(e_i)$ provided e_i is "forward edge"

(ii) $f(e_i) > 0$ provided e_i is "backward edge"

Tolerance of P is $\min\{\epsilon(e) : e \in E(P)\}$, where
$\epsilon(e) = c(e) - f(e)$ if e is forward, and
$\epsilon(e) = f(e)$ if e is backward.

Lemma. Let f be feasible and P be an f-augmenting path with tolerance z. Define
$f'(e) := f(e) + z$ if e is forward,
$f'(e) := f(e) - z$ if e is backward.
$f'(e) := f(e)$ if $e \notin E(P)$,
Then f' is feasible with $val(f') = val(f) + z$.

Characterization of maximum flows

Characterization Lemma. Feasible flow f is of maximum value iff there is NO f-augmenting path.

Proof. $\Rightarrow$ Easy.
$\Leftarrow$ Suppose f has no augmenting path.

$S := \{v \in V(D) : \exists$ f-augmenting path from s to $v^*\}$.

Then $t \notin S$ and

$$\sum_{e\in[S,\bar{S}]} c(e) = \sum_{e\in[S,\bar{S}]} f(e) - \sum_{e\in[\bar{S},S]} f(e).$$

We feel, that

(1) $val(f^*) \leq \sum_{e\in[S,\bar{S}]} c(e)$ for any feasible flow f^*, and

(2) $val(f) = \sum_{e\in[Q,\bar{Q}]} f(e) - \sum_{e\in[\bar{Q},Q]} f(e)$, for any $Q \subseteq V(D), s \in Q, t \notin Q$.

Right? Let's see

*some abuse of definition takes place...

The value of feasible flow — Proof of (2)

Lemma If f is any feasible flow, $s \in Q$, $t \notin Q$, then

$$\sum_{e\in[Q,\bar{Q}]} f(e) - \sum_{e\in[\bar{Q},Q]} f(e) = val(f).$$

Proof. By induction on $|\bar{Q}|$. If $|\bar{Q}| = 1$ then $\bar{Q} = \{t\}$ and by definition $f^-(t) - f^+(t) = val(f)$.

Let $|\bar{Q}| \geq 2$ and let $x \in \bar{Q}$, $x \neq t$.
Define $R = Q \cup \{x\}$. Since $|\bar{R}| < |\bar{Q}|$, by induction

$$\begin{aligned} val(f) &= \sum_{e\in[R,\bar{R}]} f(e) - \sum_{e\in[\bar{R},R]} f(e) \\ &= \sum_{e\in[Q,\bar{Q}]} f(e) - \sum_{e\in[\bar{Q},Q]} f(e) + \sum_{u\in Q} f(xu) \\ &\quad - \sum_{u\in Q} f(ux) + \sum_{v\in\bar{R}} f(xv) - \sum_{v\in\bar{R}} f(vx) \\ &= \sum_{e\in[Q,\bar{Q}]} f(e) - \sum_{e\in[\bar{Q},Q]} f(e) + f^+(x) - f^-(x \end{aligned}$$

Remark. $val(f) = f^+(s) - f^-(s)$.

Source/sink cuts — Proof of (1)

Source/sink cut $[S,T] = \{(u,v) \in E(D) : u \in S, v \in T\}$, if $s \in S$ and $t \in T$.

capacity of cut: $cap(S,T) := \sum_{e\in[S,T]} c(e)$.

Lemma. (Weak duality) If f is a feasible flow and $[S,T]$ is a source/sink cut, then

$$val(f) \leq cap(S,T).$$

Proof.

$$\begin{aligned} cap(S,T) &= \sum_{e\in[S,T]} c(e) \\ &\geq \sum_{e\in[S,T]} f(e) \\ &\geq \sum_{e\in[S,T]} f(e) - \sum_{e\in[T,S]} f(e) \\ &= val(f). \end{aligned}$$

Max flow-Min cut Theorem

Max Flow-Min Cut Theorem (Ford-Fulkerson, 1956) Let f be a feasible flow of maximum value and $[S,T]$ be a source/sink cut of minimum capacity. Then

$$val(f) = cap(S,T).$$

Proof. (Corollary to proof of Characterization Lemma) Define

$S := \{v \in V(D) : \exists$ f-augmenting path from s to $v^*\}$.

Since f is maximum, f has no augmenting path. Then $t \in \bar{S}$ and of course $s \in S$.

$$\begin{aligned} cap(S,\bar{S}) &= \sum_{e\in[S,\bar{S}]} c(e) \\ &= \sum_{e\in[S,\bar{S}]} f(e) - \sum_{e\in[\bar{S},S]} f(e) \\ &= val(f). \end{aligned}$$

*some abuse of definition again takes place...

f-augmenting path

G: underlying undirected graph of network D

s,t-path P in G is an f-augmenting path, if
$s = v_0, e_1, v_1, e_2 \dots v_{k-1}, e_k, v_k = t$ and for every e_i

(i) $f(e_i) < c(e_i)$ provided e_i is "forward edge"

(ii) $f(e_i) > 0$ provided e_i is "backward edge"

Tolerance of P is $\min\{\epsilon(e) : e \in E(P)\}$, where
$\epsilon(e) = c(e) - f(e)$ if e is forward, and
$\epsilon(e) = f(e)$ if e is backward.

Lemma. Let f be feasible and P be an f-augmenting path with tolerance z. Define
$f'(e) := f(e) + z$ if e is forward,
$f'(e) := f(e) - z$ if e is backward.
$f'(e) := f(e)$ if $e \notin E(P)$,
Then f' is feasible with $val(f') = val(f) + z$.

Characterization of maximum flows

Characterization Lemma. Feasible flow f is of maximum value iff there is NO f-augmenting path.

Proof. $\Rightarrow$ Easy.
$\Leftarrow$ Suppose f has no augmenting path.

$S := \{v \in V(D) : \exists$ f-augmenting path from s to $v^*\}$.

Then $t \notin S$ and

$$\sum_{e \in [S,\bar{S}]} c(e) = \sum_{e \in [S,\bar{S}]} f(e) - \sum_{e \in [\bar{S},S]} f(e).$$

We feel, that

(1) $val(f^*) \leq \sum_{e \in [S,\bar{S}]} c(e)$ for any feasible flow f^*, and

(2) $val(f) = \sum_{e \in [Q,\bar{Q}]} f(e) - \sum_{e \in [\bar{Q},Q]} f(e)$, for any $Q \subseteq V(D), s \in Q, t \notin Q$.

Right? Let's see

*some abuse of definition takes place...

The value of feasible flow — Proof of (2)

Lemma If f is any feasible flow, $s \in Q$, $t \notin Q$, then

$$\sum_{e \in [Q,\bar{Q}]} f(e) - \sum_{e \in [\bar{Q},Q]} f(e) = val(f).$$

Proof. By induction on $|\bar{Q}|$. If $|\bar{Q}| = 1$ then $\bar{Q} = \{t\}$ and by definition $f^-(t) - f^+(t) = val(f)$.

Let $|\bar{Q}| \geq 2$ and let $x \in \bar{Q}$, $x \neq t$.
Define $R = Q \cup \{x\}$. Since $|\bar{R}| < |\bar{Q}|$, by induction

$$\begin{aligned} val(f) &= \sum_{e \in [R,\bar{R}]} f(e) - \sum_{e \in [\bar{R},R]} f(e) \\ &= \sum_{e \in [Q,\bar{Q}]} f(e) - \sum_{e \in [\bar{Q},Q]} f(e) + \sum_{u \in Q} f(xu) \\ &\quad - \sum_{u \in Q} f(ux) + \sum_{v \in \bar{R}} f(xv) - \sum_{v \in \bar{R}} f(vx) \\ &= \sum_{e \in [Q,\bar{Q}]} f(e) - \sum_{e \in [\bar{Q},Q]} f(e) + f^+(x) - f^-(x) \end{aligned}$$

Remark. $val(f) = f^+(s) - f^-(s)$.

Source/sink cuts — Proof of (1)

Source/sink cut $[S,T] = \{(u,v) \in E(D) : u \in S, v \in T\}$, if $s \in S$ and $t \in T$.

capacity of cut: $cap(S,T) := \sum_{e \in [S,T]} c(e)$.

Lemma. (Weak duality) If f is a feasible flow and $[S,T]$ is a source/sink cut, then

$$val(f) \leq cap(S,T).$$

Proof.

$$\begin{aligned} cap(S,T) &= \sum_{e \in [S,T]} c(e) \\ &\geq \sum_{e \in [S,T]} f(e) \\ &\geq \sum_{e \in [S,T]} f(e) - \sum_{e \in [T,S]} f(e) \\ &= val(f). \end{aligned}$$

Max flow-Min cut Theorem

Max Flow-Min Cut Theorem (Ford-Fulkerson, 1956) Let f be a feasible flow of maximum value and $[S,T]$ be a source/sink cut of minimum capacity. Then

$$val(f) = cap(S,T).$$

Proof. (Corollary to proof of Characterization Lemma) Define

$S := \{v \in V(D) : \exists\ f$-augmenting path from s to $v^*\}$.

Since f is maximum, f has no augmenting path. Then $t \in \bar{S}$ and of course $s \in S$.

$$\begin{aligned} cap(S,\bar{S}) &= \sum_{e\in[S,\bar{S}]} c(e) \\ &= \sum_{e\in[S,\bar{S}]} f(e) - \sum_{e\in[\bar{S},S]} f(e) \\ &= val(f). \end{aligned}$$

*some abuse of definition again takes place...

Directed Edge-Menger

Given $x, y \in V(D)$, a set $F \subseteq E(D)$ is an x,y-disconnecting set if $D - F$ has no x,y-path. Define

$\kappa'_D(x,y) :=$ min$\{|F| : F$ is an x,y-disconnecting set,$\}$
$\lambda'_D(x,y) :=$ max$\{|\mathcal{P}| : \mathcal{P}$ is a set of p.e.d.* x,y-paths$\}$

* p.e.d. means pairwise edge-disjoint

Directed-Local-Edge-Menger Theorem For all $x, y \in V(D)$,

$$\kappa'_D(x,y) = \lambda'_D(x,y).$$

Proof. Build network (D,x,y,c) with $c(e) = 1$ for all $e \in E(D)$. Clearly

- 1-to-1 correspondence between x,y-disconnecting sets and sorce/sink cuts. Hence $\kappa'_D(x,y) = \min cap(S,\bar{S})$.
- each set of p.e.d. path determines a feasible flow. So $\lambda'_D(x,y) \le \max val f$.

But what if there is some clever way to direct differently a flow with **larger** overall value?? This flow then must have fractional values on some of the edges.

Ford-Fulkerson Algorithm

Initialization $f \equiv 0$
WHILE there exists an augmenting path P
  DO augment flow f along P
return f

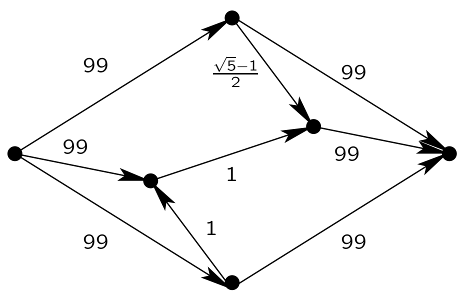
Corollary. (**Integrality Theorem**) If all capacities of a network are integers, then there is a maximum flow assigning integral flow to each edge.
Furthermore, some maximum flow can be partitioned into flows of unit value along path from source to sink.

Running times:

- Basic (careless) Ford-Fulkerson: might not even terminate, flow value might not converge to maximum;
  when capacities are integers, it terminates in time $O(m\,|f^*|)$, where f^* is a maximum flow.
- Edmonds-Karp: chooses a *shortest* augmenting path; runs in $O(nm^2)$

Example

The Max-flow Min-cut Theorem is true for real capacities as well,
BUT our algorithm might fail to find a maximum flow!!!



Example of Zwick (1995)

Remark. The max flow is 199. There is such an unfortunate choice of a sequence of augmenting paths, by which the flow value tends to 3.

Menger's Theorem for directed graphs

Given $x, y \in V(D)$, a set $S \subseteq V(D) \setminus \{x, y\}$ is an x, y-separator (or an x, y-cut) if $D - S$ has no x, y-path.
Define

$\kappa_D(x, y) := \min\{|S| : S$ is an x, y-cut,$\}$ and
$\lambda_D(x, y) := \max\{|\mathcal{P}| : \mathcal{P}$ is a set of p.i.d. x, y-paths$\}$

Directed-Local-Vertex-Menger Theorem Let $x, y \in V(D)$, such that $xy \notin E(D)$. Then

$$\kappa_D(x, y) = \lambda_D(x, y).$$

Proof. We apply the Integrality Theorem for the auxiliary network (D', x^+, y^-, c').

$V(D') := \{v^-, v^+ : v \in V(D)\}$
$E(D') := \{u^+v^- : uv \in E(D)\} \cup \{v^-v^+ : v \in V(D)\}$
$c'(u^+v^-) = \infty^*$ and $c'(v^-v^+) = 1$.

*or rather a very-very large **integer**.

Global Corollaries

A directed graph is weakly connected if the underlying undirected graph is connected; it is strongly connected if there is a directed u, v-path for any vertex u and any vertex $v \neq u$.
Strongly k-edge-connected: after removal of any $k - 1$ edges the digraph remains strongly connected.
Strongly k-connected: after removal of any $k - 1$ vertices the digraph remains stronngly connected.

Corollary (Directed-Global-Edge-Menger Theorem) Directed multigraph D is strongly k-edge-connected iff there is a set of k p.e.d.x, y-paths for any two vertices x and y.

Corollary (Directed-Global-Vertex-Menger Theorem) A digraph D is strongly k-connected iff for any two vertices $x, y \in V(D)$ there exist k p.i.d. x, y-paths.

Proof: **Lemma.** For every $e \in E(D)$, $\kappa_D(G-e) \geq \kappa_D(G)-1$.

Application: Baranyai's Theorem

$\chi'(K_n) = n - 1$ is saying: $E(K_n)$ can be decomposed into pairwise disjoint perfect matchings.

k-uniform hypergraphs? $E(\mathcal{K}_n^{(k)}) = \binom{[n]}{k}$

Let $k|n$. $\mathcal{S} = \{S_1, \ldots, S_{n/k}\}$ is a "perfect matching in $\mathcal{K}_n^{(k)}$ if $S_i \cap S_j = \emptyset$ for $i \neq j$.

There is are many perfect matchings in $\mathcal{K}_n^{(k)}$.
Is there a decomposition of $\binom{[n]}{k}$ into perfect matchings?

Not obvious already for $k = 3$ (Peltesohn, 1936)

$k = 4$ (Bermond)

Theorem (Baranyai, 1973) For every $k|n$, there is a decomposition of $\binom{[n]}{k}$ into perfect matchings.

Proof of Baranyai's Theorem

Induction on the size of the underlying set $[n]$.
NOT the way you would think!!!

We imagine how the $m = \frac{n}{k}$ pairwise disjoint k-sets in each of the $M = \binom{n-1}{k-1} = \binom{n}{k}/m$ "perfect matchings" would develop as we add one by one the elements of $[n]$.

A **multi**set $\mathcal{A}$ is an m-partition of the base set X if $\mathcal{A}$ contains m pairwise disjoint sets whose union is X.

Remarks
An m-partition is a "perfect matching" in the making.
Pairwise disjoint $\Rightarrow$ only $\emptyset$ can occur more than once.

Stronger Statement For every l, $0 \leq l \leq n$ there exists M m-partitions of $[l]$, such that every set S occurs in $\binom{n-l}{k-|S|}$ m-partitions ($\emptyset$ is counted with multiplicity).

Remark For $l = n$ we obtain Baranyai's Theorem since $\binom{0}{k-|S|} = 0$ unless $|S| = k$, when its value is 1.

Proof of Stronger Statement: Induction on l.

$l = 0$: Let all $\mathcal{A}_i$ consists of m copies of $\emptyset$.

$l = 1$: Let all $\mathcal{A}_i$ consists of $m-1$ copies of $\emptyset$ and 1 copy of $\{1\}$.

Let $\mathcal{A}_1, \ldots, \mathcal{A}_M$ be a family of m-partitions of $[l]$ with the required property.
We construct one for $l+1$.

Define a network D:

$V(D) = \{s,t\} \cup \{\mathcal{A}_i : i = 1,\ldots,M\} \cup 2^{[l]}$.

$E(D) = \{s\mathcal{A}_i : i \in [M]\} \cup \{\mathcal{A}_i S : S \in \mathcal{A}_i\}$
$\cup \{St : S \in 2^{[l]}\}$.

Edge $\mathcal{A}_i\emptyset$ has the same multiplicity as $\emptyset$ in $\mathcal{A}_i$.

Capacities: $c(s\mathcal{A}_i) = 1$
$c(\mathcal{A}_i S)$ any positive integer.
$c(St) = \binom{n-l-1}{k-|S|-1}$.

There is flow f of value M:

Flow values: $f(s\mathcal{A}_i) = 1$
$f(\mathcal{A}_i S) = \frac{k-|S|}{n-l}$
$f(St) = \binom{n-l-1}{k-|S|-1}$.

Remark. Edges of type 1 and 3 have maximum flow value.

Claim f is a flow. □

f is clearly maximum ($val(f) = cap(\{s\}, V \setminus \{s\})$).

Integrality Theorem $\Rightarrow$ there is a maximum flow g with integer values. So
$g(s\mathcal{A}_i) = f(s\mathcal{A}_i) = 1$ and
$g(St) = f(St) = \binom{n-l-1}{k-|S|-1}$.

By the conservation constraints at $\mathcal{A}_i$ there exists a unique S_i for each $i = 1,\ldots,M$ such that $g(\mathcal{A}_i S_i) = 1$.

Define m-partitions

$$\mathcal{A}'_i = \mathcal{A}_i \setminus \{S_i\} \cup \{S_i \cup \{l+1\}\}$$

of the set $[l+1]$.

Claim $\{\mathcal{A}'_1, \ldots, \mathcal{A}'_M\}$ is an appropriate family of m-partitions of $[l+1]$.

Proof. Let $T \subseteq [l+1]$.

If $l+1 \in T$, then T occurs in $\mathcal{A}'_i$ iff for $S = T \setminus \{l+1\}$ we have $g(\mathcal{A}_i S) = 1$. By conservation at vertex S:

$$|\{i \in [M] : g(\mathcal{A}_i S) = 1\}| = g(St) = \binom{n-(l+1)}{k-(|S|+1)}.$$

If $l+1 \notin T$, then T occurs in $\mathcal{A}'_i$ iff $T \in \mathcal{A}_i$ and $g(\mathcal{A}_i T) = 0$. The number of these indices i by induction and the above is equal to

$$\binom{n-l}{k-|T|} - \binom{n-(l+1)}{k-(|T|+1)} = \binom{n-(l+1)}{k-|T|}.$$

□