# Graph theory

Eric Shen

Friday, August 28, 2020

The contents of this handout are based on various lectures on graph theory I have heard over the years, including some from MOP and some from Canada/USA Mathcamp.

## Contents


**1. Basic terminology** **3**
- 1.1. Introduction to graphs . . . 3
- 1.2. Dictionary . . . 3

**2. Basic graph facts** **4**
- 2.1. Trees . . . 4
- 2.2. Handshaking lemma . . . 5
- 2.3. Example problems . . . 5

**3. Colorings of graphs** **7**
- 3.1. Introduction to coloring . . . 7
- 3.2. Planar graphs . . . 7
- 3.3. Four-color theorem . . . 8
- 3.4. Example problems . . . 8

**4. Combinatorial constructions** **10**
- 4.1. Basic graph constructions . . . 10
- 4.2. Perfect matchings . . . 11
- 4.3. Corollaries of Hall's theorem . . . 11
- 4.4. A hard walkthrough . . . 12

**5. Extremal graph theory** **13**
- 5.1. Triangle-free bounds . . . 13
- 5.2. Formalization . . . 14
- 5.3. Turán graphs . . . 14
- 5.4. Example problems . . . 15

**6. Application: crossing numbers** **17**
- 6.1. Conventional definitions . . . 17
- 6.2. Some conjectures . . . 17
- 6.3. Crossing number lemma . . . 17
- 6.4. Erdő's unit distance problem . . . 19

**7. Practice problems** **21**

**8. Solutions to walkthroughs** **23**
8.1. Solution 2.6 . . . 23
8.2. Solution 2.7 . . . 23
8.3. Solution 2.8 (HMIC 2020/2) . . . 23
8.4. Solution 3.1 . . . 24
8.5. Solution 3.10 (Six-color theorem) . . . 24
8.6. Solution 3.11 (HMMT 2018 T9) . . . 24
8.7. Solution 3.12 (USA TST 2015/5) . . . 25
8.8. Solution 4.1 . . . 25
8.9. Solution 4.2 . . . 25
8.10. Solution 4.3 . . . 26
8.11. Solution 4.4 . . . 26
8.12. Solution 4.8 (RMM 2020/3) . . . 26
8.13. Solution 5.10 (TSTST 2014/5) . . . 27
8.14. Solution 5.11 (Folklore) . . . 27
8.15. Solution 5.12 (ISL 2004 C8) . . . 28

**A. Even more extremal graph theory** **28**
A.1. E-S-S theorem . . . 28
A.2. The bipartite case . . . 29
A.3. Bounds for trees . . . 30

# §1 Basic terminology

## §1.1 Introduction to graphs

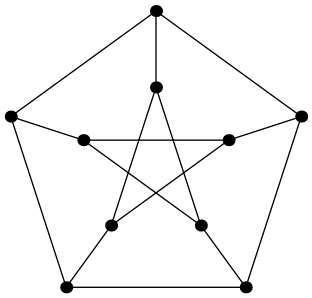
**Definition 1.1** (Graphs)

A **graph** $G = (V, E)$ is a collection of $V$ of **vertices** and $E \subseteq V \times V$ of **edges**. A **simple graph** is one where every edge links a unique pair of distinct vertices.

**Remark.** Beyond this section, all graphs mentioned are simple unless otherwise specified.

When we draw graphs, we think of the edges as connecting pairs of vertices, and represent edges by connecting their endpoints with curves.

Below is the famous **Petersen graph**.

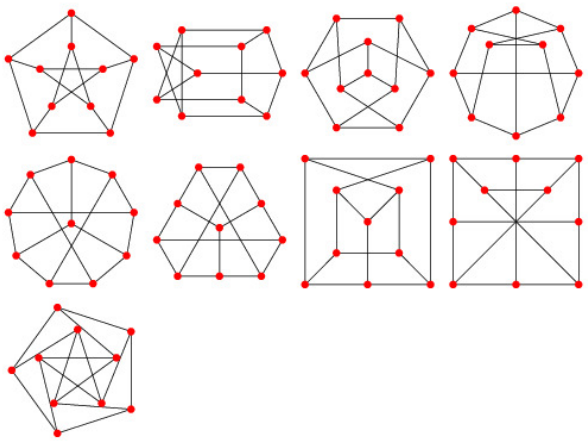


Where we draw the vertices does not matter.

**Definition 1.2** (Isomorphism)

Two (simple) graphs $G$, $H$ with vertex sets $V(G)$, $V(H)$ are **isomorphic**, denoted $G \cong H$, if there is a bijection $f : V(G) \to V(H)$ such two vertices $u$, $v$ are connected by an edge in $G$ if and only if $f(u)$, $f(v)$ are connected by an edge in $H$.

**Question 1.3.** Which of the following graphs are isomorphic to the Petersen graph?



## §1.2 Dictionary

- Two vertices are **adjacent** they are connected by an edge. We also say they are **neighbors**.

- A vertex $v$ is **incident** to an edge $e$ (or vice versa) if $v$ is an endpoint of $e$.
- A graph is **connected** if there is a path between every pair of distinct vertices. Graphs may be split into **connected components**.
- The **open neighborhood** of a vertex $v$ denoted $N(v)$, is the subgraph induced by the set of neighbors of $v$. The **closed neighborhood** includes $v$. A neighborhood is assumed to be open unless otherwise specified.
- A graph is **bipartite** if the vertex set can be partitioned into twos sets $V_1 \sqcup V_2$ such that all edges contain exactly one endpoint in $V_1$ and exactly one endpoint in $V_2$.
- The **chromatic number** of a graph is the minimum number of colors needed to color the vertices so that no two adjacent vertices are colored the same color.
- A **clique** or **complete graph** on $n$ vertices, denoted $K_n$, is the $n$-vertex graph with an edge between any two vertices.
- A **walk** is a sequence of vertices such that consecutive vertices are adjacent. A **path** is a walk with vertices pairwise distinct.
- A **cycle** is a path whose first and last vertices are adjacent.
- The **degree** $\deg v$ of a vertex $v$ is the number of edges incident to $v$.
- A **forest** is a graph with no cycles. A **tree** is a connected forest.[1]
- A graph is **planar** if it is possible to draw it in the plane without any crossing edges.
- A **Hamiltonian path** is a path that contains every vertex. A **Hamiltonian cycle** is a cycle that is a Hamiltonian path.
- A **Eulerian path** is a path that traverses every edge exactly once. A **Eulerian tour** is a cycle that is a Eulerian path.

# §2 Basic graph facts

**Induction is unusually powerful in graph theory.**

## §2.1 Trees

**Theorem 2.1** (Trees)

A connected $n$-vertex graph with $n-1$ edges is a tree.

There are two main ways to define a tree:

(i) A acyclic connected graph.

(ii) Any graph attainable by this process: start with a single vertex, then keep on adding a vertex $v$ and an edge incident to $v$.

To prove Theorem 2.1, we will show that (i) and (ii) are equivalent definitions.

[1] It turns out that the connected components of a forest are all trees, so it is fun to say, "A forest is a collection of trees."

**(ii) implies (i)** Indeed, a graph obeying (ii) is a graph and connected (by construction). We will show it is acyclic.

Assume not, so we have a cycle. Let $v$ be the vertex in the cycle that was added last. Then along with $v$ we must have added two edges, contradiction.

**(i) implies (ii)** A **leaf** is a vertex of degree 1. First a lemma:

**Lemma 2.2**

Every acyclic connected graph has a leaf.

*Proof of lemma.* Evidently no vertex has degree 0, since the graph is connected. Assume for contradiction all degrees are at least two. Start with a vertex, traverse an incident edge, and whenever we arrive at a vertex, traverse another edge incident to that vertex. By infinite Pigeonhole, we eventually end up at a vertex twice, so we have found a cycle. □

Now we can grab a leaf $v_0$ and delete it, along with its incident edge. Then we have fewer cycles, so still acyclic, and the graph is still connected since no path between any two remaining vertices involves $v_0$.

Thus the resulting graph is also acyclic connected, so we can finish by induction.

**Exercise 2.3.** What is the minimal number of edges in a connected graph?

**Exercise 2.4.** For $n \geq 3$, prove that every $n$-vertex graph with exactly $n-3$ edges contains at least three connected components.

## §2.2 Handshaking lemma

In what follows, $V$ is the vertex set of $G$ and $E$ the edge set.

**Theorem 2.5** (Handshake lemma)

Let $G$ be a graph. Then

$$\sum_{v \in V} \deg v = 2|E|.$$

*Proof.* Double counting, left as exercise. □

*Alternate proof of Lemma 2.2.* By Theorem 2.4, the sum of the degrees is $2n-2$. Since all vertices have nonzero degree, there exists a vertex of degree less than two (say, by Pigeonhole). □

## §2.3 Example problems

**Example 2.6** (BAMO 2005/4)

A country has 100 cities, and some pairs of the cities are linked by highways. Every city is reachable from every other city through some sequence of roads. The government would like to make some of the inter-city links into toll roads. Prove that it is possible to do this is such a way that every city has an odd number of toll roads incident to itself. (Assume that highways merge only at cities.)

**Walkthrough.** Say a vertex is *good* if it has an odd number of toll roads, and *bad* otherwise.

**(a)** Pick a random subset of the edges to be toll roads.

**(b)** Prove that the number of bad vertices, so we can pair them up.

**(c)** For every pair of bad toll roads $u$, $v$, come up with a scheme to "toggle" a few toll roads such that $u$, $v$ are no longer bad, and no additional bad vertices are created.

**Example 2.7**

Suppose that an arrow is drawn on each edge of a cube, giving each edge a direction, in such a way that every vertex of the cube has at least one arrow coming out of it and at least one arrow going into it. Prove that under these conditions, it is always possible to find a face of the cube such that the directions of the boundary edges of that face go in a cycle.

**Walkthrough.** I managed to casework bash this once. I do not recommend. Assume for contradiction every vertex has at least one arrow coming in and at least one going out, but no face goes in a cycle

**(a)** Say an angle of a face is *good* if it can be traversed along the drawing arrows, and *bad* otherwise. (So, every face of the cube has four angles, and every vertex of the cube has three angles.)

**(b)** Show that each face has at most two good angles.

**(c)** Show that each vertex of the cube has at least two good angles.

**(d)** Count the number of good angles in two ways. Derive a contradiction therefrom.

**Example 2.8** (HMIC 2020/2)

Some bishops and knights are placed on an infinite chessboard, where each square has side length 1 unit. Suppose that the following conditions hold:

- For each bishop, there exists a knight on the same diagonal as that bishop (there may be another piece between the bishop and the knight).
- For each knight, there exists a bishop that is exactly $\sqrt{5}$ units away from it.
- If any piece is removed from the board, then at least one of the above conditions is no longer satisfied.

If $n$ is the total number of pieces on the board, find all possible values of $n$.

**Walkthrough.** By some experimentation, we can deduce that the answer is $4 \mid n$.

**(a)** Come up with a construction for all $n = 4k$.

**(b)** Consider a directed graph where each node is either a bishop of a knight, and there is an edge from a bishop to a knight if the bishop attacks the knight, and vice versa. (No edges point from a bishop to a bishop, or a knight to a knight.)

**(c)** Verify that all outdegrees are at least 1. Delete some edges from the graph—so it no longer obeys the rule in (b)—such that all outdegrees are exactly 1.

**(d)** Show that all indegrees are exactly 1.

**(e)** Prove that each connected component of the graph is a cycle.

**(f)** Prove that each cycle has length a multiple of four, by checkerboard coloring the chessboard.

# §3 Colorings of graphs

## §3.1 Introduction to coloring

**Example 3.1**

Determine the minimum integer $L$ such that every superposition of two trees on the same vertex set can be vertex-colored using at most $L$ colors.

**Walkthrough.** **(a)** Prove that $L > 2$ by finding a superposition of two trees that can't be colored using two colors.

**(b)** Prove that $L > 2$ by finding a superposition of two trees that can't be colored using three colors. (Hint: four vertices.)

**(c)** Show that any tree can be colored with two vertices.

**(d)** Show that any superposition of two trees can be colored using four vertices. Thus the answer is $L = 4$.

## §3.2 Planar graphs

**Definition 3.2**

**Planar graphs** can be drawn in $\mathbb{R}^2$ with no crossing edges.

**Question 3.3.** Is $K_3$ planar? $K_4$? $K_5$? $K_n$?

**Question 3.4.** Is $K_{2,3}$ planar? $K_{2,n}$? $K_{3,3}$? $K_{m,n}$?

You should find $K_3$, $K_4$, $K_{2,3}$ are planar, but $K_5$, $K_{3,3}$ are not. We will prove these later.

What follows is a bit of abusive notation; we'll use $V$, $E$, $F$ to refer to the number of vertices, edges, and faces (whereas in the past, they referred to the respective sets).

**Theorem 3.5** (Euler characteristic)

For connected planar graphs, $V - E + F = 2$.

**Exercise 3.6.** Try the proof on your own; it's a simple induction. Recall that when inducting in graph theory, we usually look for something to delete.

*Proof.* We are going to induct on $F$, with base case $F = 1$ for trees.

Now take any two faces that share an edge $e$, and delete $e$. Then the resulting graph has the same number of vertices $V' = V$, one fewer face $F' = F - 1$, and one fewer edge $E' = E - 1$. Given $V' - E' + F' = 2$, we have $V - E + F = 2$ as well. □

**Theorem 3.7**

In planar graphs with $V \geq 3$, we have $E \leq 3V - 6$.

*Proof.* Draw $G$ in the plane with no crossings. Then

$$2E = \sum_{\text{faces}} (\#\text{sides}) \geq 3F = 3(2 + E - V) \implies E \leq 3V - 6.$$

□

**Exercise 3.8.** Use Theorem 3.7 to show $K_5$ and $K_{3,3}$ are not planar.

**Remark.** It turns out, $K_5$ are sort of the "building blocks" of non-planar graphs. **Kuratowski's theorem** states that a graph is non-planar if and only if it contains a subgraph that is **homeomorphic** to either $K_5$ or $K_{3,3}$.

Let me explain what this means: a **subdivision** of an edge $u \sim v$ involves deleting the edge $u \sim v$, adding a vertex $w$, and drawing the edges $u \sim w$, $w \sim v$. A **subdivision** of a graph $G$ is a graph resulting from the subdivision of the edges in $G$.

Then $G$ and $H$ are homeomorphic if there some subdivison of $G$ and some subdivison of $H$ are isomorphic.

## §3.3 Four-color theorem

One of the more famous problems in mathematics:

**Theorem 3.9** (Four-color theorem)

Every planar graph can have its vertices colored in at most four colors.

The proof is hard, and the only known proof requires checking thousands of cases by hand (clarification: by computer).

We will prove a weaker version:

**Example 3.10** (Six-color theorem)

Show that every planar graph can have its vertices colored in at most six colors.

**Walkthrough.** Guess what? We're inducting again.

**(a)** We will pick some vertex $v$, delete $v$, color the resulting subgraph in six colors as promised, then add $v$ back. Which vertex $v$ should we delete?

**(b)** Show that every planar graph has a vertex of degree at most 5. (Hint: use a theorem in §3.2.)

**(c)** Show that $v$ can be colored appropriately and finish the proof.

## §3.4 Example problems

**Example 3.11** (HMMT 2018 T9)

Evan has a simple graph $G$ with $v$ vertices and $e$ edges. Show that he can delete at least $\frac{e-v+1}{2}$ edges so that each vertex still has at least half of its original degree.

**Walkthrough.** This problem is short but tricky.

**(a)** Check that we may safely assume $G$ is connected.

**(b)** Try to interpret the quantity $\frac{1}{2}(e-v+1)$. It may help to write it as $\frac{1}{2}[e-(v-1)]$.

**(c)** Motivated by (b), we will actually induct on $e$. To begin the induction, settle the base case $e = v - 1$.

**(d)** Determine a familiar graph such that we can delete half of its edges so that each vertex still has exactly half its original degree.

**(e)** Keep on deleting (d) via induction so that we may assume (d) is no longer a subgraph of $G$.

**(f)** Show that, now that there are no more instances of (d), that all odd cycles are disjoint.

**(g)** The bound is now very weak: conclude the proof by just removing something from each cycle.

**Example 3.12** (USA TST 2015/5)

A tournament is a directed graph for which every (unordered) pair of vertices has a single directed edge from one vertex to the other. Let us define a proper directed-edge-coloring to be an assignment of a color to every (directed) edge, so that for every pair of directed edges $\overrightarrow{uv}$ and $\overrightarrow{vw}$, those two edges are in different colors. Note that it is permissible for $\overrightarrow{uv}$ and $\overrightarrow{uw}$ to be the same color. The directed-edge-chromatic-number of a tournament is defined to be the minimum total number of colors that can be used in order to create a proper directed-edge-coloring. For each $n$, determine the minimum directed-edge-chromatic-number over all tournaments on $n$ vertices.

**Walkthrough.** We first need to use some rough estimates to extract the answer.

**(a)** Let $f(n)$ be the answer. Verify that $f(2) = 1$, $f(3) = 2$, $f(4) = 3$.

**(b)** Show that $f(a+b) \le \max(f(a), f(b)) + 1$. Assuming this bound is sharp, determine the order of magnitude of $f(n)$.

**(c)** To strengthen our confidence in (b), also show that $f(ab) \le f(a) + f(b)$.

**(d)** Now attack from the other direction: prove by induction on $k$ that $f(2^k) \ge f(2^{k-1}) + 1$, and some node has $f(2^k)$ incoming colors.

Now it seems clear the answer is $\lceil \log_2 n \rceil$.

**(e)** This is the magic step: interpret the condition that $\overrightarrow{uv}$, $\overrightarrow{vw}$ are different colors in the following way: for each $i$, every node is either $i$-incoming (meaning accepts incoming edges of color $i$) or $i$-outgoing (meaning emits outgoing edges of color $i$), but not both. It's possible for that neither holds; in that case, it really doesn't matter — just say it is $i$-incoming for convenience.

**(f)** Let each node have a *signature*, a binary string of $k$ bits (where $k$ is the number of colors) such that the $i$th bit equals 0 if the vertex is $i$-incoming and 1 if the vertex is $i$-outgoing.

**(g)** Prove that $n \leq 2^k$.

**(h)** Conversely, prove that if $n \leq 2^k$, then it is possible to assign signatures to each vertex, i.e. such a graph exists.

# §4 Combinatorial constructions

## §4.1 Basic graph constructions

**Example 4.1**

A tournament is a directed graph with an edge between any two vertices. For every $n$, show that there is a tournament on $n$ vertices such that at each vertex, the outdegree and indegree differ by at most 1.

**Walkthrough.** Do this yourself, it is not hard.

**Example 4.2**

For every $n$, show that there is a "half-graph," where every vertex degree is within $\pm 1$ of $\frac{n-1}{2}$, and the total number of edges is as close to $\frac{1}{2}\binom{n}{2}$ as possible.

**Walkthrough.** We will take cases based on $n$ modulo 4.

**Case 1: $n \equiv 1 \pmod 4$.** This is the easy one.

**(a)** For $n \equiv 1 \pmod 4$, come up with a solution in a similar manner to your constructions in Example 4.1

**Case 2: $n$ even.** Slightly trickier.

**(b)** For $n$ even, come up with a graph where every vertex has $\frac{n}{2}$ neighbors.

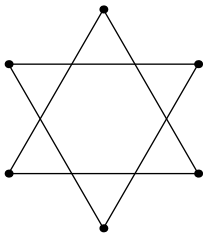
**(c)** There is an excess of edges. Make a few adjustments to drop the total edge count down to an acceptable amount.

**Case 3: $n \equiv 3 \pmod 4$.** This is pretty hard.

**(d)** For $n$ prime, find a way to give each vertex a degree of $\frac{n-3}{2}$ in a similar manner to your construction in part (a).

**(e)** Building off of (d), add an "almost perfect matching" to obtain an acceptable graph for $n$ prime.

**(f)** Now generalize to all $n$ replacing cycles with general **two-factors**: graphs where every vertex has degree two. All cycles are two-factors, but not the other way around; for instance, the following:



## §4.2 Perfect matchings

**Example 4.3**

For every even $n$, there is a partition of the edges of $K_n$ into $n-1$ perfect matchings. For every odd $n$, there is a partition of the edges of $K_n$ into $n$ almost-perfect matchings.

**Walkthrough.** Mainly experimentation with small $n$. See the full solution at the end.

**Example 4.4**

Show that it is possible to partition the set of all subsets of $\{1, 2, \ldots, n\}$ into $\binom{n}{\lfloor n/2 \rfloor}$ lists such taht each list $(S_1, S_2, S_3, \ldots)$ has the property that $S_1 \subset S_2 \subset S_3 \subset \cdots$.

**Walkthrough.** First assume $n$ is even.

**(a)** Draw a **Hasse diagram** on the subsets of $\{1, 2, \ldots, n\}$, where we draw an arrow from $A$ to $B$ whenever $A \subset B$.

**(b)** Using the Hasse diagram, reduce the problem to finding a perfect matching from subsets of size $m+1$ to subsets of size $m$, for $m \geq n/2$. (In this context, it means every vertex of the important group gets a different vertex of the other group.)

Now we introduce Hall's theorem on perfect matchings.

**Theorem 4.5** (Hall's matching theorem)

Let $G$ be a bipartite graph with input set $A$, output set $B$. There exists a perfect matching $f : A \to B$ if and only if for every subset $S \subseteq A$, we have $|N(S)| \geq |S|$.

**(c)** Verify the condition of Hall's theorem and solve the problem for $n$ even.

**(d)** Extend the proof to cover $n$ odd.

## §4.3 Corollaries of Hall's theorem

**Corollary 4.6**

Consider a bipartite graph $G$ with input set $A$, output set $B$, and $|A| \leq |B|$. Suppose that the degrees of all vertices in $A$ are the same, and the degrees of all the vertices in $B$ are the same. Then there is a perfect matching from $A \to B$.

**Corollary 4.7**

Suppose $|A| = |B|$ in the above configuration. If we delete a perfect matching, the resulting graph still has the aforementioned property. Rinse and repeat, and we have a perfect matching partition.

## §4.4 A hard walkthrough

**Example 4.8** (RMM 2020/3)

Let $n \geq 3$ be an integer. In a country there are $n$ airports and $n$ airlines operating two-way flights. For each airline, there is an odd integer $m \geq 3$, and $m$ distinct airports $c_1$, ..., $c_m$, where the flights offered by the airline are exactly those between the following pairs of airports: $c_1$ and $c_2$; $c_2$ and $c_3$; ...; $c_{m-1}$ and $c_m$; $c_m$ and $c_1$.

Prove that there is a closed route consisting of an odd number of flights where no two flights are operated by the same airline.

**Walkthrough.** Before we tackle the RMM problem, let's borrow a powerful tool from computer science:

> Let $G$ be a connected graph. There exists a subgraph $H$ of $G$ such that $H$ is a tree and contains all of $G$'s vertices. We call $H$ a **spanning tree** of $G$.

**(a)** The proof of the existence of spanning trees is another induction. Try it yourself.

**(b)** Verify that we can easily create a spanning tree with the following algorithm: start with a single point in our spanning tree, and keep on selecting vertices connected by a single edge to our spanning tree, and adding that vertex along with its edge.

**(c)** For extra practice, show that in a weighted graph, by greedily adding the vertex whose edge has minimal distance to the spanning tree, we construct the spanning tree with least edge sum. This is **Prim's algorithm** for **minimum spanning trees**.

Now we can work on the RMM problem. We are given a graph of $n$ vertices, and $n$ odd cycles of different colors. The task is to show there is an odd cycle, all of whose edges have distinct colors.

We will try to construct a spanning tree on the $n$ vertices of the graph, where each edge is a distinct color.

**Part I: Constructing the spanning tree.**

**(d)** Begin as you would on (b). Start with an edge (i.e. tree with two vertices) and iteratively add vertices to our tree.

**(e)** Say that at some step in the process, we have a tree on the vertices $v_1, \ldots, v_k$ with $k < n$, and that the $k-1$ edges cover $k-1$ distinct colors. Then there are three cases:

(i) There is another color whose cycle contains some vertex in $\{v_1, \ldots, v_k\}$ but also some vertex in $V \setminus \{v_1, \ldots, v_k\}$.

(ii) There is another color whose cycle is contained entirely in $\{v_1, \ldots, v_k\}$.

(iii) No other colors have cycles that contain an element of $\{v_1, \ldots, v_k\}$.

**(f)** In the case of (i), find a way to increase the size of our spanning tree.

**(g)** In the cases of (ii) or (iii), use the inductive hypothesis to solve the problem.

Now assume (ii), (iii) never happen, so we finish the process and end up with a spanning tree with $n-1$ edges of distinct colors.

**Part II: Using the spanning tree.** We have a spanning tree $n-1$ edges of distinct colors. There is a remaining color, so let $x_1 \sim x_2 \sim \cdots \sim x_m \sim x_1$ (with $m$ odd) be the cycle of the remaining color.

**(h)** For each $i$, let $\ell_i$ be the distance from $x_i$ to $x_{i+1}$ (indices modulo $m$) in the spanning tree. Verify that
$$\ell_1 + \ell_2 + \cdots + \ell_m \quad \text{is even.}$$

**(i)** Use (h) to conclude that for some $i$, the union of the path from $x_i$ to $x_{i+1}$ on the spanning tree with the edge $x_i \sim x_{i+1}$ is an odd cycle all of whose edges are distinct colors. Conclude.

# §5 Extremal graph theory

## §5.1 Triangle-free bounds

**Exercise 5.1.** Suppose I have a graph with five vertices. How many edges are needed to guarantee it contains a triangle?

The answer is seven.

**Theorem 5.2** (Mantel)

If $G$ has $n$ vertices and is triangle-free, then the number of edges is at most $\lfloor n^2/4 \rfloor$. Furthermore, the unique triangle-free graph with $\lfloor n^2/4 \rfloor$ is $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

*Proof.* Let $G$ have $m$ edges. Since $G$ is triangle-free, for any adjacent vertices $x \sim y$, we have $\deg x + \deg y \le n$, since $N(x)$, $N(y)$ must be disjoint.

Summing over all edges,
$$\sum_{x \sim y} (\deg x + \deg y) \le mn.$$
By Jensen's inequality,
$$mn \ge \sum_{x \sim y} (\deg x + \deg y) = \sum_{x \in V} (\deg x)^2 \overset{\text{Jensen}}{\ge} n \left( \sum_{x \in V} \frac{\deg x}{n} \right)^2 = n \left( \frac{2m}{n} \right)^2,$$
which rearranges to $m \le n^2/4$.

We can check the equality case by making everything an equality, but that's annoying and therefore omitted. □

*Alternate proof.* An **independent set** is a set of vertices whose induced subgraph has no edges. Let $\Delta(G)$ be the maximum degree among the vertices of $G$.

Let $A$ is the independent set of maximal size. Since $G$ is triangle-free, $N(x)$ is independent for all $x$, so $\Delta(G) \le |A|$. Let $B = V \setminus A$.

Since $A$ is independent, every edge must touch a vertex in $B$, so
$$m \le \sum_{x \in B} \deg x \le |B| \cdot \Delta(G) \le |B| \cdot |A| \le \left( \frac{|A| + |B|}{2} \right)^2 = \frac{n^2}{4}.$$

□

## §5.2 Formalization

**Definition 5.3**

Let $H$ be a graph with $n \geq |H|$. Then $G$ is **extremal** (for $n$, $H$) if $H \not\subseteq G$ and $G$ has the maximum number of edges possible for any $n$-vertex, $H$-free graph.

**Definition 5.4**

A graph not containing $H$ is **edge maximal** with respect to $H$ if the addition of any edge would create a copy of $H$.

The difference is that edge maximality is sort of a "local maximum." Some graphs may be edge maximal, but not extremal, but all extremal graphs are edge maximal.

**Definition 5.5** (Notation)

Let $\mathrm{ex}(n, H)$ be the maximum number of edges possible in an $n$-vertex graph such that $H$ is not a subgraph.

We've seen that $\mathrm{ex}(n, K_3) = \lfloor n^2/4 \rfloor$ by Mantel's theorem. Our goal will be to compute $\mathrm{ex}(n, K_{r+1})$.

## §5.3 Turán graphs

**Definition 5.6** (Turán graphs)

$T_r(n)$ is the (unique) complete $r$-partite graph on $n \geq r$ vertices where partition classes differ in size by at most one. We let $t_r(n)$ be the number of edges of $T_r(n)$.

**Fact 5.7.** Of complete $k$-partite graphs on $n$ vertices, $T_k(n)$ has the most edges.

*Proof.* AM-GM. □

**Theorem 5.8**

For any $r, n \in \mathbb{N}$,

$$t_r(n) \leq \frac{n^2}{2}\left(1 - \frac{1}{r}\right),$$

with equality iff $r \mid n$.

*Proof for $r \mid n$.* Just compute

$$\#\text{ edges} = \frac{\#\text{ vts} \cdot \deg}{2} = \frac{n}{2}\left(\frac{n}{r}\right)(r-1),$$

which rearranges.

For $r \nmid n$, the proof is rather annoying, so we will omit it. □

**Theorem 5.9** (Turán)

Let $G$ be a graph with $n$ vertices, $m$ edges, and $r \in \mathbb{N}$. If $G$ is $K_{r+1}$-free, then

$$m \le t_r(n) \le \frac{n^2}{2} \le \frac{n^2}{2}\left(1 - \frac{1}{r^2}\right).$$

Furthermore $m = t_{r-1}(n)$ iff $G \cong T_{r-1}(n)$.

In other words, $\operatorname{ex}(n, K_{r+1}) = t_r(n)$.

*Proof of bound.* We will induct on $n$. For the base case $n \le r$, there is nothing to show.

Consider $n > r$. It suffices to only consider $G$ that are edge-maximal wrt. $K_{r+1}$; therefore, $G$ has $K_r$ as a subgraph. Let's name our $K_r$ subgraph $K$, so

$$\begin{aligned} m &\le (\#\text{ in } G - K) + (\#\text{ between } G,\ K) + (\#\text{ in } K) \\ &= t_r(n-r) + (n-r)(r-1) + \binom{r}{2}. \end{aligned}$$

I claim this is just equal to $t_{r-1}(n)$. We will find all three terms in a Turán graph.

For the $\binom{r}{2}$ term, take one vertex from each partition in the $r$-partite graph, thereby forming a $K_r$. Color these vertices red, and the rest of the vertices blue. Then $t_r(n-r)$ counts the number of blue-blue edges and $(n-r)(r-1)$ counts the number of blue-red edges. End proof. □

*Proof of equality case.* We've shown Turán graphs achieve equality, so let's prove equality implies Turán graph.

We can run add on to our above induction. Let the vertices of $K$ be $x_1, \ldots, x_r$. Define $V_i = \{v \in V \mid vx_i \in E\}$. For equality to hold, all $V_i$ are pairwise disjoint, and each is independent. Thus $G$ is an $(r-1)$-partite graph.

End proof since Turán graphs are the unique maimum. □

## §5.4 Example problems

**Example 5.10** (TSTST 2014/5)

Find the greatest positive integer $E$ such that there is an edge-colored graph with 60 vertices and $E$ edges with each edge colored either red or blue such that there is no monochromatic cycle of length either 3 or 5.

**Walkthrough.** **(a)** Identify an $r$ such that each edge of $K_r$ cannot be colored either red or blue such that there is no monochromatic cycle of length 3 nor 5.

**(b)** Show that such a graph as described in the problem is $K_r$-free, and apply Turán's theorem for an upper bound. If done property, the bound should be sharp.

**(c)** Find an equality case (i.e. where the Turán bounds above are sharp) with 4-vertex graphs.

**(d)** Generalize (c) to $4n$-vertex graphs, and conclude.

**Example 5.11** (Folklore)

Show that a graph with $2n$ vertices and $n^2 + 1$ edges has at least $n$ triangles.

**Walkthrough.** We will, again, induct on $n$.

**(a)** Verify the problem for $n \geq 2$.

**(b)** Assume the hypothesis for $n-1$. Show that for graphs $G$ with $2n$ vertices and $n^2+1$ edges, there is at least one triangle $uvw$.

**(c)** Show that $\deg u + \deg v \geq 2n+1$ by the inductive hypothesis.

**(d)** Verify that $\deg u + \deg v + \deg w \geq 3n+2$.

**(e)** Use the Pigeonhole Principle to finish.

**Example 5.12** (ISL 2004 C8)

For a finite graph $G$, let $f(G)$ be the number of triangles and $g(G)$ the number of tetrahedra formed by edges of $G$. Find the least constant $c$ such that

$$g(G)^3 \leq c \cdot f(G)^4$$

for every graph $G$.

**Walkthrough.**

**(a)** The equality case is quite standard. Try to guess the answer.

Let the vertices of $G$ be $v_1, \ldots, v_n$, and let $e_i$ edges contain $v_i$, let $f_i$ triangles contain $v_i$, and let $g_i$ tetrahedrons contain $v_i$. Also let $G$ contain $E$ edges, $F$ triangles, and $G$ tetrahedrons.

**(b)** Verify that $2E = \sum e_i$, $3F = \sum f_i$, $4G = \sum g_i$.

We will first prove an inequality of the form $F^2 \leq c_0 \cdot E^3$.

**(c)** Guess the value of $c_0$ in a similar manner to (a).

**(d)** We want a chain of inequalities of the form

$$3F = \sum_{i=1}^{n} f_i \leq \sum_{i=1}^{n} (\text{something in terms of } e) \leq \cdots \leq \text{constant} \cdot E^{3/2}.$$

That is, we want to bound $f_i \leq$ something in terms of $e$. Take advantage of the degrees of the variables to pick something to fill the blank.

**(e)** Now to extend to $G^3 \leq c \cdot F^4$, we would do something similar. First verify that $g_i^2 \leq c_0 f_i^3$ using (d).

**(f)** Use that in conjunction with $g_i \leq F$ to write a chain of inequalities in the form

$$4G = \sum_{i=1}^{n} g_i \leq \sum_{i=1}^{n} (\text{something in terms of } f) \leq \cdots \leq \text{constant} \cdot F^{4/3}.$$

**(g)** Hopefully your value of $c$ in (f) matches your answer to (a). If it doesn't, let equality hold at all inequalities to rigorously determine the correct value of $c$.

**(h)** Conclude.

# §6 Application: crossing numbers

## §6.1 Conventional definitions

**Definition 6.1** (Drawing)

A **drawing** of a graph in the plane is an assignment of a point in $\mathbb{R}^2$ for each vertex, and a curve in $\mathbb{R}^2$ for each edge, so that

- edges may cross, but at most two edges cross at any point,
- edges don't go through vertices except at endpoints, and
- edges sharing a vertex don't cross.

**Definition 6.2** (Crossing numbers)

The **crossing number** $\operatorname{cr}(G)$ is the minimum number of edge crossings among all drawings of $G$.

Hence,

- a graph $G$ is planar if and only if $\operatorname{cr}(G) = 0$.
- $\operatorname{cr}(K_5) = \operatorname{cr}(K_{3,3}) = 1$.

## §6.2 Some conjectures

**Conjecture 6.3** (Turán, Zarankiewicz)

$$\operatorname{cr}(K_{m,n}) = \left\lfloor \frac{m}{2} \right\rfloor \left\lfloor \frac{m-1}{2} \right\rfloor \left\lfloor \frac{n}{2} \right\rfloor \left\lfloor \frac{n-1}{2} \right\rfloor.$$

**Conjecture 6.4** (Hill)

$$\operatorname{cr}(K_n) = \frac{1}{4} \left\lfloor \frac{n}{2} \right\rfloor \left\lfloor \frac{n-1}{2} \right\rfloor \left\lfloor \frac{n-2}{2} \right\rfloor \left\lfloor \frac{n-3}{2} \right\rfloor.$$

In both conjectures, we know the upper bound. Lower bounds are harder than upper bounds.

## §6.3 Crossing number lemma

**Remark** (Notation conventions henceforth)**.** From now on, $v$ and $e$ will denote number of vertices and number of edges. If $G$ isn't clear, we'll write $v(G)$, $e(G)$.

**Theorem 6.5**

If $G$ is planar, then $e < 3v$.

We've already proven this. (In fact, we've proven $e \le 3v - 6$ for $v \ge 3$.)

**Corollary 6.6**

For all $G$, we have $\operatorname{cr}(G) > e - 3v$.

*Proof.* Fix a drawing with exactly $\operatorname{cr}(G)$ crossings. Form a new graph $G'$ by arbitrarily deleting one edge from each crossing.

Then $v(G') = v(G)$, $e(G') \geq e(G) - \operatorname{cr}(G)$, and $G'$ is planar, so $e(G') < 3v(G')$. Hence

$$\operatorname{cr}(G) \geq e(G) - e(G') > e(G) - 3v(G).$$

□

**Theorem 6.7** (Crossing number lemma, independently by Ajtoi-Chvátal-Newborn-Szemerédi and by Leighton, in 1982–1983)

If $G$ has $e \geq 4v$, then

$$\operatorname{cr}(G) \geq \frac{e^3}{64v^2}.$$

**Remark.** The best constant to replace 64 known is 29. But what really matters is the order of magnitude.

*Proof of Theorem 5.7 by Lovász, Matoušek, Pach, et al.* We will boost Corollary 5.6 to obtain this stronger result, by deleting vertices. But cleverly choosing vertices to delete is difficult, so screw that — delete vertices at random.

We fix some parameter $p \in [0, 1]$. Given $G$, fix a drawing of $G$ with $\operatorname{cr}(G)$ crossings. Let $G_p$ be a random induced subgraph of $G$ gotten by keeping each vertex independently with probability $p$. Then $v(G_p)$, $e(G_p)$, $\operatorname{cr}(G_p)$ are random.

$$\begin{aligned}
\mathbb{E}[v(G_p)] &= p \cdot v(G) \\
\mathbb{E}[e(G_p)] &= p^2 e(G) \\
\mathbb{E}[\operatorname{cr}(G_p)] &\leq p^4 \operatorname{cr}(G).
\end{aligned}$$

Now $\operatorname{cr}(G_p) \geq e(G_p) - 3v(G_p)$, so taking expectation of both sides,

$$\begin{aligned}
\mathbb{E}[\operatorname{cr}(G_p)] &\geq \mathbb{E}[e(G_p)] - 3\mathbb{E}[v(G_p)] \\
p^4 \operatorname{cr}(G) &\geq p^2 e(G) - 3p \cdot v(G) \\
\operatorname{cr}(G) &\geq p^{-2} e - 3p^{-3} v.
\end{aligned}$$

A good choice of $p$ is $4v/e$, for which

$$\operatorname{cr}(G) \geq \frac{e^2}{16v^2} e - 3\frac{e^3}{64v^3} v = \frac{e^3}{64v}.$$

□

**Remark.** Note $e \geq 4v$ is so that $p = 4v/3$ is a valid probability.

## §6.4 Erdő's unit distance problem

**Problem 6.8** (Erdős's unit distance problem, 1946)

For finite $S \subset \mathbb{R}^2$, let $u(S) = \#\{\{x, y\} \subset S : |x - y| = 1\}$ and

$$u(n) = \max_{\substack{S \subset \mathbb{R}^2 \\ |S|=n}} u(S).$$

What is $u(n)$?

Suppose $n$ perfect square. Let $S$ be a $\sqrt{n} \times \sqrt{n}$ unit square grid. For "most" points of $S$, they are at unit distance from four other points of $S$. Here, $u(S) \approx 2n$.

A triangular grid should do better, but we can do even better than that. Let $S_5$ be the same grid, but dilated down by a factor of $1/5$. Why five? Because $5^2 = 3^2 + 4^2$. Then $(0, 0)$ is at unit distance from 12 points of $S_5$, namely

$$(\pm 1, 0), \quad \left(\pm\frac{3}{5}, \pm\frac{4}{5}\right), \quad \left(\pm\frac{4}{5}, \pm\frac{3}{5}\right), \quad (0, \pm 1).$$

Here, $u(S) \approx 6n$.

**Lemma 6.9**

If $k$ is an integer such that $k^2$ can be written as a sum of two perfect squares in $\ell$ ways, let $S_k$ be a grid dilated down by $1/k$. Almost every point will be at unit distance from $\ell$ others, so

$$u(n) \geq u(S_k) \approx \frac{\ell n}{2}.$$

Be careful, though! If $k > \sqrt{2n}$, then diagonal of $S_k$ has length less than one, so we get 0 unit distances. As $k$ gets larger, the edge effect gets more and more important.

So now the question is,

**Question 6.10.** Which $k$, not too big relative to $n$, can be written as the sum of two squares in the most number of ways?

Exists $k$ such that $\ell = n^{c/\log\log n}$. Then $c > 0$ is an absolute constant, and

$$u(n) \geq n^{1+c/\log\log n}.$$

This is the best known lower bound, and it's conjectured this is the best bound.

**Remark.** The function we wrote above grows faster than linear functions in $n$, but slower than $n^{1+\varepsilon}$ for any $\varepsilon > 0$.

**Theorem 6.11** (Spencer, Szemerédi, Trotter '84)

$u(n) \leq Cn^{4/3}$ for some $C > 0$.

*Proof by Székelys.* I will prove $u(n) \leq 8n^{4/3} + n$. Let $S \subset \mathbb{R}^2$ be an arbitrary set of $n$ points.

Draw a unit circle around every point of $S$. Then $2u(S)$ is the number of incidences between $S$ and these circles. Consider the graph $G_0$ with vertex set $S$ and edges between two vertices whenever they are consecutive on a unit circle.

Then

- $v(G_0) = n$ is obvious.
- $e(G_0) = 2u(S)$, since if one circle has $m$ points, it contributes $m$ edges to $G_0$ and also $m$ incidences between circles and points.
- $\text{cr}(G_0) \le 2\binom{n}{2} < n^2$ by looking at the graph already drawn for us. Each pair of circles intersect twice.

A small issue: $G_0$ is not a simple graph, i.e. may be more than one edge between a pair of vertices. There might also be self-loops. But this is all that can go round; there are at most four edges connecting any two vertices.

Let $G$ be gotten from $G_0$ by deleting all this junk: delete all self-loops and keep only one edge between any two vertices.

- $v(G) = v(G_0) = n$ as expected.
- $\text{cr}(G) \le \text{cr}(G_0) < n^2$ as expected.
- $e(G) \ge \frac{1}{4}e(G_0) - n$ since at most four edges between pair of vertices. The $n$ is to subtract out self-loops.

Let's apply the crossing number lemma:

$$n^2 \ge \text{cr}(G) \ge \frac{e(G)^3}{64v(G)^2} \ge \frac{\left(\frac{1}{4}e(G_0) - n\right)^3}{64n^2}.$$

Rearranging,

$$u(S) = \frac{e(G_0)}{2} \le 2\left(4n^{4/3} + n\right) = 8n^{4/3} + n.$$

Remember we needed to assume $e \ge 4v$. If not, then $2u(S) = e(G_0) < 4n$ and $u(S) \le 8n$. $\square$

## §7 Practice problems

**Problem 7.1** (BAMO 2004/3, free 2 points)**.** NASA has proposed populating Mars with 2004 settlements. The only way to get from one settlement to another will be by a connecting tunnel. A bored bureaucrat draws on a map of Mars, randomly placing N tunnels connecting the settlements in such a way that no two settlements have more than one tunnel connecting them. What is the smallest value of N that guarantees that, no matter how the tunnels are drawn, it will be possible to travel between any two settlements?

**Problem 7.2.** Prove that every connected graph with all degrees even has a Eulerian tour.

**Problem 7.3** (MOP 2011 R0.1)**.** Let $n$ be a positive integer. Partition the set $\{0, \dots, 2n-1\}$ into $n$ pairs $\{a_1, b_1\}$, $\{a_2, b_2\}$, ..., $\{a_n, b_n\}$. For $i = 1, \dots, n$ draw a semicircle above the $x$-axis joining $(a_i, 0)$ to $(b_i, 0)$, and draw a semicircle below the $x$-axis joining $(a_i + 1, 0)$ and $(b_i + 1, 0)$.

Suppose that no two of these semicircles intersect except at their endpoints. Prove that the union of the semicircles forms a single closed curve from $(0, 0)$ to $(2n, 0)$.

**Problem 7.4** (ToT 1986)**.** Twenty football teams take part in a tournament. On the first day all the teams play one match. On the second day all the teams play a further match. Prove that after the second day it is possible to select 10 teams, so that no two of them have yet played each other.

**Problem 7.5** (PUMaC 2019/1)**.** A finite graph $G$ is drawn on a blackboard. The following operation is permitted: pick any cycle $C$ of $G$, draw a new vertex $v$, connect it to all vertices of $C$, and finally erase all the edges of $C$. Prove that the operation can only be done a finite number of times.

**Problem 7.6** (China TST 2015/1/3)**.** Fix positive integers $k$, $n$. A candy vending machine has many different colours of candy, where there are $2n$ candies of each colour. A couple of kids each buys from the vending machine 2 candies of different colours. Given that for any $k + 1$ kids there are two kids who have at least one colour of candy in common, find the maximum number of kids.

**Problem 7.7** (ISL 2013 C6)**.** In some country several pairs of cities are connected by direct two-way flights. It is possible to go from any city to any other by a sequence of flights. The distance between two cities is defined to be the least possible number of flights required to go from one of them to the other. It is known that for any city there are at most 100 cities at distance exactly three from it. Prove that there is no city such that more than 2550 other cities have distance exactly four from it.

**Problem 7.8** (ISL 2005 C4)**.** Let $n \geq 3$ be a fixed integer. The edges of the complete graph $K_n$ are labeled with a number from the set $\{1, \dots, r\}$ such that

- each of the $r$ labels is used at least once; and
- in any triangle, two of the edges are labeled with the same number, and this number is greater than the label of the third side.

Find the maximal $r$ for which such a labeling is possible, and determine how many such labelings achieve this maximum.

**Problem 7.9.** If a graph $G$ has average degree $d$, then it has a subgraph $H \subset G$ whose minimum degree is at least $d/2$.

**Problem 7.10** (Japan 1997/3)**.** Let $G$ be a graph with 9 vertices. Suppose given any five points of $G$, there exist at least 2 edges with both endpoints among the five points. What is the minimum possible number of edges in $G$?

**Problem 7.11** (St. Petersburg 1996/4)**.** In a group of several people, some are acquainted with each other and some are not. Every evening, one person invites all of his acquaintances to a party and introduces them to each other. Suppose that after each person has arranged at least one party, some two people are still unacquainted. Prove that they will not be introduced at the next party.

**Problem 7.12** (Canada 2020/5)**.** Let $G$ be a graph with 19998 vertices so that any 9999-vertex subgraph of $G$ has at least 9999 edges. Determine the minimum possible number of edges in $G$.

**Problem 7.13.** Let $G$ be a finite, simple graph with $n \geq 3$ vertices. If $\deg v + \deg w \geq n$ for every pair of non-adjacent vertices $v$ and $w$ of $G$, then $G$ has a Hamiltonian path.

**Problem 7.14** (IMO 2019/3)**.** A social network has 2019 users, some pairs of whom are friends. Whenever user $A$ is friends with user $B$, user $B$ is also friends with user $A$. Events of the following kind may happen repeatedly, one at a time:

> Three users $A$, $B$, and $C$ such that $A$ is friends with both $B$ and $C$, but $B$ and $C$ are not friends, change their friendship statuses such that $B$ and $C$ are now friends, but $A$ is no longer friends with $B$, and no longer friends with $C$. All other friendship statuses are unchanged.

Initially, 1010 users have 1009 friends each, and 1009 users have 1010 friends each. Prove that there exists a sequence of such events after which each user is friends with at most one other user.

# §8 Solutions to walkthroughs

## §8.1 Solution 2.6

Say a vertex is *good* if it has an odd number of toll roads, and *bad* otherwise. Pick a random subset of the edges to be toll roads. By $\sum_v \deg v = 2|E|$, an even number of vertices are good, so an even number of vertices are bad.

Pair up the bad vertices. For every pair of bad toll roads $u$, $v$, take a path between them, and toggle every edge on the path. We can check that $u$, $v$ are now good, and all other vertices in the graph do not change their state. Thus at the end of this process, all vertices are good.

## §8.2 Solution 2.7

We will actually prove the result for general convex polyhedra. The generalization requires Euler's characteristic, stating that $|V| - |E| + |F| = 2$, where $F$ is the set of faces.

Say an angle of a face is *good* if it can be traversed along the drawn arrows, and *bad* otherwise. Let $V$, $E$, $F$ be the sets of vertices, edges, and faces, respectively.

Assume for contradiction every vertex has positive indegree and outdegree, but no face forms a cycle. We rely on two orthogonal estimates, based on parity issues:

- Each face with $S$ sides contains at most $S - 2$ good angles.
- At each vertex of the cube, there are at least two good angles.

Let $I$ be the number of good angles. Observe

$$I \le \sum_{f \in F} (\#\text{ sides} - 2) = 2|E| - 2|F| = 2|V| - 4 < 2|V| \le I,$$

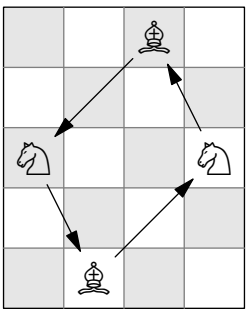
which is absurd.

> **Remark.** The version of the problem with a cube is from Po-Shen Loh's lecture "Graph theory: introduction" at MOP 2019, which is based on his 21-228 Discrete Mathematics course at Carnegie Mellon University.

## §8.3 Solution 2.8 (HMIC 2020/2)

The answer is $4 \mid n$.

To achieve $n = 4k$, take $k$ copies of the following, spaced sufficiently far away from one another so that no two pieces from different arrangements attack each other.



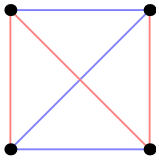
Now we show $4 \mid n$ is necessary. Consider a directed graph where each node is either a bishop or a knight, and there is an edge from a bishop to a knight if the bishop attacks the knight, and vice versa. (No edges point from a bishop to a bishop, or a knight to a knight.)

We are given each node has outdegree at least 1. For each node $v$, arbitrarily delete outgoing edges until the outdegree of $v$ is exactly 1. Then each node has indegree 1, as otherwise some node has indegree 0 and may be deleted. Hence the graph consists of disjoint cycles.

It will suffice to show each cycle has length a multiple of 4. Checkerboard color the chessboard: then for any edge from a bishop to a knight, the two endpoints are the same color, but for any edge from a knight to a bishop, the two endpoints are different colors. Hence the number of edges from a knight to a bishop is even, but this is equal to the number of edges from a bishop to a knight. It follows that the number of edges is a multiple of 4, as desired.

## §8.4 Solution 3.1

The answer is $L = 4$. Here's a superposition of two trees with chromatic number four:



Evidently every tree has chromatic number two.

Color the two trees each with the colors 1, 2, so that each vertex has color $a$ in the first tree and color $b$ in the second tree, with $a, b \in \{1, 2\}$. Then no two neighbors have the same pair $(a, b)$, so we have colored the graph with four colors: $(1, 1)$, $(1, 2)$, $(2, 1)$, $(2, 2)$.

## §8.5 Solution 3.10 (Six-color theorem)

First:

**Lemma**

Every planar graph has a vertex of degree at most 5.

*Proof.* Trivial for $V = 1$, $V = 2$. For $V \geq 3$, we have

$$\sum_v \deg v \leq 2E \leq 6V - 12,$$

so Pigeonhole applies. □

We backwards induct on $V$. Find a vertex of degree at most 5, delete it, color the remaining subgraph, then color $v$ using a color that is not present in the neighborhood of $v$. (Since $\deg v \leq 5$, the neighborhood of $v$ cannot contain all six colors.)

## §8.6 Solution 3.11 (HMMT 2018 T9)

Evidently we may assume $G$ is connected. We will induct on $e$, with base case $e = v - 1$ where $G$ is a tree and we remove no edges. Henceforth $e \geq v$, so there is at least one cycle.

- If there is an even-length cycle, then remove the cycle, apply the inductive hypothesis on the resulting graph, and add back every other edge of the even-length cycle.
- If there are no even-length cycles, then all odd-length cycles are disjoint, so the number of cycles is $e - v + 1$. Simply remove an edge from each cycle.

That is all.

## §8.7 Solution 3.12 (USA TST 2015/5)

The answer is $\lceil \log_2 n \rceil$. Assume such a tournament on $n$ vertices can be colored with $k$ colors. For each color $i \le k$, we may designate each vertex as either *$i$-incoming* or *$i$-outgoing*, meaning the vertex is incident to an incoming edge of color $i$ or an outgoing edge of color $i$, respectively. By hypothesis no vertex may be both $i$-incoming and $i$-outgoing, and if a vertex is not incident to any edge of color $i$, it does not matter whether we designate it as $i$-incoming or $i$-outgoing — for convenience say it is $i$-incoming.

We assign to each vertex a *signature* — a binary string of $k$ bits such that the $i$th bit equals 0 if the vertex is $i$-incoming and 1 if the vertex is $i$-outgoing. Note that:

- no two vertices may have the same signature, since otherwise it is impossible to draw an edge between them;
- if two vertices have different signatures, then it is possible to draw an edge between.

There are $2^k$ possible signatures, so we must have $n \le 2^k$. Conversely if $n \le 2^k$, then we can select a subset of the $2^k$ possible signatures and assign them to the vertices. Between any two vertices draw an edge, thereby forming a valid tournament. Thus the answer is $\lceil \log_2 n \rceil$.

## §8.8 Solution 4.1

Put the $n$ vertices in a circle, so that the (arc) distance between adjacent vertices is 1. Begin by drawing an arrow $u \to v$ whenever the distance from $u$ to $v$ is at most $\lfloor n/2 \rfloor$.

For $n$ odd, this is already sufficient. For $n$ even, arbitrarily draw the diagonals.

## §8.9 Solution 4.2

We take cases modulo 4. Again put the $n$ vertices in a circle.

**Proof for $n \equiv 1 \pmod 4$:** Draw an edge between $u$ and $v$ whenever the distance between $u$ and $v$ is at most $\frac{n-1}{4}$. Then each vertex has degree $\frac{n-1}{2}$ and the total edge count is

$$\#\text{ edges} = n \cdot \frac{n-1}{4} = \frac{1}{2}\binom{n}{2}.$$

**Proof for $n$ even:** First construct the complete bipartite graph $K_{n/2,n/2}$. Each vertex has degree $\frac{n}{2}$, so may remove up to one edge from each vertex.

Remove a total of $\lfloor n/4 \rfloor$, so that the degree is either $\frac{n}{2}$ of $\frac{n-2}{2}$ for each vertex and the total edge count is

$$\#\text{ edges} = \frac{n^2}{4} - \left\lfloor \frac{n}{4} \right\rfloor = \frac{1}{2}\binom{n}{2} + \left\{ \frac{n}{4} \right\}.$$

**Proof for $n \equiv 3 \pmod 4$:** For each $k \le \frac{n-3}{4}$, we can draw a two-factor on the graph of $n$ vertices where each vertex has degree 2 and every edge connects vertices of distance $k$. (This may be a cycle, or a composition of cycles.)

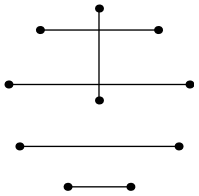
Then, add an almost perfect matching, so each vertex has degree either $\frac{n-3}{2}$ or $\frac{n-1}{2}$ and the edge count is

$$\#\text{ edges} = n \cdot \frac{n-3}{4} + \frac{n-1}{2} = \frac{1}{2}\binom{n}{2} - \frac{1}{2}.$$

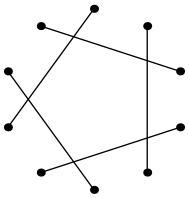
## §8.10 Solution 4.3

Split on cases based on the parity of $n$.

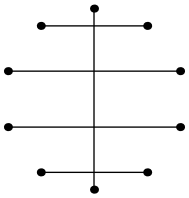
**Proof for $n$ odd:** Rotate the following.



**Proof for $n$ even:** For each odd $i < n/2$, we can construct a perfect matching where each edge has length $i$, e.g. the following:



The remaining $n/2$ matchings are obtained by rotating the following:



## §8.11 Solution 4.4

Imagine a Hasse diagram.

> **Claim.** There is a perfect matching from the set of subsets of size $m+1$ to subsets of size $m$, for $m \geq \lfloor n/2 \rfloor$.

*Proof.* Each subset of size $m+1$ has $m+1$ neighbors, whereas each subset of size $m$ has $n-m$ neighbors. Verify that $n-m \leq m+1$, then apply Hall matching theorem. □

Then for each of the $\binom{n}{\lfloor n/2 \rfloor}$ subsets of size $\lfloor n/2 \rfloor$, apply the perfect matchings to turn each subset into a chain of subsets.

## §8.12 Solution 4.8 (RMM 2020/3)

We proceed by strong induction on $n$, with base case trivial. Let the vertex set be $V$. We will try to construct a spanning tree on the $n$ vertices, where each edge is a distinct color. This will not always be possible, but when it is impossible, we will be able to induct down.

Begin by selecting an edge (i.e. tree with two vertices) $v_1 \sim v_2$. We iteratively add vertices to our tree. Say at some step in the process, we have a tree on the vertices $v_1, \ldots, v_k$, with $k < n$, and that the $k-1$ edges cover $k-1$ distinct colors. There are three cases from here:

(i) We can find another color whose cycle contains some vertex in $\{v_1, \dots, v_k\}$, but also contains some vertex in $V \setminus \{v_1, \dots, v_k\}$. Then there is an edge of this color of the form $v_i \sim w$, where $w \notin \{v_1, \dots, v_k\}$. Add this vertex $w$ to the tree.

(ii) We can find another color whose cycle is contained entirely in $\{v_1, \dots, v_k\}$. Then there are $k$ cycles on the subgraph with vertices $\{v_1, \dots, v_k\}$, so we win by the inductive hypothesis.

(iii) No other colors have cycles that contain an element of $\{v_1, \dots, v_k\}$. In the subgraph consisting of the remaining $n - k$ vertices, there are $n - k$ cycles, so we win by the inductive hypothesis.

If at any point in the process, one of conditions (ii), (iii) holds, then we are already done. Otherwise, we successfully construct a spanning tree with $n - 1$ edges of distinct colors.

Let $x_1 \sim x_2 \sim \cdots \sim x_m \sim x_1$, with $m$ odd, be the cycle formed by the remaining color, and consider the paths from $x_1$ to $x_2$, from $x_2$ to $x_3$, ..., from $x_m$ to $x_1$ in the spanning tree. The sum of their lengths must be even, so one of them, say the path from $x_i$ to $x_{i+1}$ (indices modulo $m$), has even length.

Then union this path with the edge $x_i \sim x_{i+1}$ to construct a cycle of odd size all of whose edges are distinct colors. End proof.

## §8.13 Solution 5.10 (TSTST 2014/5)

The answer is 1350. To achieve this, partition the vertex set into $A$, $B$, $C$, $D$ of 15 vertices each; draw an red edge between any two vertices belonging to $(A, B)$, $(B, C)$, $(C, D)$, $(D, A)$, and blue between vertices belonging to $(A, C)$, $(B, D)$.

For maximality, it is not hard to check there is no $K_5$, so by Turan's theorem,

$$E \le \left(1 - \frac{1}{4}\right) \frac{n^2}{2} = 1350,$$

as desired.

## §8.14 Solution 5.11 (Folklore)

The claim is obvious for $n = 2$. Consider a minimal counterexample on $2n \ge 4$ vertices. By Turán's theorem there is already a triangle.

> **Claim.** If $u$, $v$, $w$ form a triangle, then $\deg u + \deg v \ge 2n + 1$.

*Proof.* If $\deg u + \deg v \le 2n$, remove $u$ and $v$. The resulting graph $G'$ has $2n - 2$ vertices and at least $(n-1)^2 + 1$ edges. By minimality $G'$ contains at least $n - 1$ triangles, so $G$ has $n$ triangles, contradiction. □

Repeating cyclically, we have $\deg u + \deg v + \deg w \ge 3n + \frac{3}{2}$, i.e.

$$\deg u + \deg v + \deg w \ge 3n + 2.$$

But there are $2n - 3$ remaining vertices, so by Pigeonhole, there are at least $n - 1$ additional triangles containing an edge in $\{uv, vw, wu\}$ but is not $uvw$ itself.

## §8.15 Solution 5.12 (ISL 2004 C8)

The answer is $\frac{3}{32}$, attained by arbitrarily large cliques. Let the vertices of $G$ be $v_1$, …, $v_n$, and let $e_i$ edges contain $v_i$, let $f_i$ triangles contain $v_i$, and let $g_i$ tetrahedrons contain $v_i$. Also let $G$ contain $E$ edges, $F$ triangles, and $G$ tetrahedrons, so that $2E_i = \sum e_i$, $3F_i = \sum f_i$, $4G_i = \sum g_i$.

**Claim.** $F^2 \le \frac{2}{9}E^3$.

*Proof.* Note that $f_i \le E$ and $f_i \le \binom{e_i}{2}$. Then

$$3F = \sum_{i=1}^{n} f_i \le \sum_{i=1}^{n} \sqrt{E\binom{e_i}{2}} < \sqrt{\frac{E}{2}} \sum_{i=1}^{n} e_i = \sqrt{2}E^{3/2},$$

and the claim follows. □

Note the trivial bound $g_i \le F$. Consider the subgraph of $G$ comprised of the neighbors of $v_i$ (but not $v_i$). It contains $f_i$ edges and $g_i$ triangles, so $g_i^2 \le \frac{2}{9}f_i^3$. Hence

$$4G = \sum_{i=1}^{n} g_i \le \sum_{i=1}^{n} \sqrt[3]{F \cdot \frac{2}{9}f_i^3} = \sqrt[3]{\frac{2F}{9}} \sum_{i=1}^{n} f_i = 3\sqrt[3]{\frac{2}{9}}F^{4/3},$$

and the desired bound follows.

**Remark.** Repeating the argument presented above gives the following generalization: if $A_t$ is the number of $K_t$'s formed by edges of $G$, then

$$\frac{A_{t+1}^t}{A_t^{t+1}} \le \frac{t!}{(t+1)^t}.$$

# §A Even more extremal graph theory

## §A.1 E-S-S theorem

We will present results on generalizing $K_r$ to $H$.

Say $\chi(H) = r$. THen we already know an $H$-free graph: $T_{r-1}(H)$. This is $H$-free since it can be colored in $r-1$ colors. Thus we already have the bound

$$\operatorname{ex}(n, H) \ge \frac{n^2}{2}\left(1 - \frac{1}{\chi(H)-1}\right).$$

It turns out this bound is **asymptotically optimal** for $n \to \infty$:

**Theorem A.1** (Erdös, Stone, Simonovits)

Let $H$ be a graph. Then

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

The proof is hard, so we won't do it.

## §A.2 The bipartite case

**Theorem A.2**

For every $s \le t$ we have

$$\frac{n^{\left(2-\frac{s-t-2}{st-1}\right)}}{16} \le \operatorname{ex}(n, K_{s,t}) \le \frac{e}{2}(t-1)^{1/s}n^{2-1/s}.$$

**Remark.** In fact there is the stronger upper bound

$$\operatorname{ex}(n, K_{s,t}) \le \left(\frac{1}{2} + o(1)\right)(t-1)^{1/s}n^{2-1/s}.$$

**Upper bound** Let $G$ be edge-maximal and have $n$ vertices, $m$ edges; so $\deg v \ge s-1$ for all $v \in V$.

The idea is to count $K_{1,s}$. There's going to be so many of these that we can take the obvious bound for $K_{s,t}$. So how many $K_{1,s}$'s can we count? Obviously:

$$\#K_{1,s} = \sum_{v \in V} \frac{\deg v}{s}.$$

We will use Jensen's inequality along with this lemma to bound $\#K_{1,s}$:

**Lemma A.3**

For integers $1 \le k \le n$,

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

By Jensen's inequality,

$$\#K_{1,s} = \sum_{v \in V} \binom{\deg v}{s} \ge n\binom{2m/n}{s} \ge n\left(\frac{2m}{ns}\right)^2.$$

Now I claim $\#K_{1,s} \le (t-1)\binom{n}{s}$. The $\binom{n}{s}$ represents the number of sets of $s$ vertices, and each set of $s$ has at most $t-1$ common neighbors. (Else, a $K_{s,t}$ already exists.) We have

$$\#K_{1,s} \le (t-1)\binom{n}{s} \le (t-1)\left(\frac{en}{s}\right)^s.$$

Combining the two inequalities,

$$n\left(\frac{2m}{ns}\right)^s \le (t-1)\left(\frac{en}{s}\right)^s.$$

With some rearranging,

$$m \le \frac{e}{2}(t-1)^{1/s}n^{2-1/s}.$$

**Lower bound** We'll use the probabilistic method. Let $G(n,p)$ be a random graph on $n$ vertices where every edge is chosen with probability $p$. How many edges would you expect?

$$\mathbb{E}[m] = p\binom{n}{2} > \frac{pn^2}{4}.$$

Now I would also like to count the expected number of $K_{s,t}$. We can also do this easily:

$$\mathbb{E}\left[\#K_{s,t}\right] = \binom{n}{s}\binom{n-s}{t}p^{st} \le n^{s+t}p^{st}.$$

By setting $p = \frac{1}{2}n^{-\left(\frac{s+t-2}{st-1}\right)}$, one can verify $\mathbb{E}\left[\#K_{s,t}\right] \le \frac{1}{2}\mathbb{E}[m]$ in the form $n^{s+t}p^{st} \le \frac{1}{8}pn^2$.

Hence

$$\mathbb{E}\left[m - \#K_{s,t}\right] = \mathbb{E}[m] - \mathbb{E}\left[\#K_{s,t}\right] \ge \frac{\mathbb{E}[m]}{2} \ge \frac{pn^2}{8} = \frac{n^{\left(2-\frac{s-t-2}{st-1}\right)}}{16},$$

so there exists a graph $G$ with $m - \#K_{s,t} \ge \frac{1}{16}n^{\text{thing}}$.

Remove an edge from each $K_{s,t}$, leaving a $K_{s,t}$-free graph $G'$ with at least $n^{\text{thing}}$ edges remaining.

## §A.3 Bounds for trees

**Conjecture A.4** (Erdős-Sós conjecture)

Let $T$ be a tree with $t+1$ vertices.

$$\text{ex}(n,T) \le \frac{n(t-1)}{2}.$$

**Remark.** This was supposedly proven in 2015, but it has not yet been written up. Lame asf.

We will prove the subcase:

**Theorem A.5** (Erdős-Gallai)

Let $P$ be a path with $t+1$ vertices.

$$\text{ex}(n,P) \le \frac{n(t-1)}{2}.$$

First, a lemma:

**Lemma A.6**

Every graph has a cycle/path of at least $\delta(G)+1$ vertices, where $\delta$ is the minimum degree.

*Proof Lemma 8.3.* Consider the longest path $x_1, \dots, x_m$. Thus, all neighbors of $x_m$ must lie in the path, so the path contains at least $\delta(G)+1$ vertices. □

**Lemma A.7**

If $G$ is connected, it has a path of length $\min\{n, 2\delta(G)+1\}$.

*Proof Lemma 8.4.* If $G$ has a path of length $n$, done.

Otherwise, we show $G$ has a path of length $2\delta+1$. Consider a maximum length path $x_1, \ldots, x_m$. Suppose $M < 2\delta + 1$.

> **Claim.** We can find an index $1 \le i \le m-1$ such that $x_1 \sim x_{i+1}$ and $x_i \sim x_n$.

*Proof of claim.* There are at least $\delta(G)$ neighbors of $x_1$ among $\{x_2, \ldots, x_m\}$. Analogously among $\{x_1, \ldots, x_{m-1}\}$ are at least $\delta(G)$ neighbors of $x_m$.

Details omitted, but this should follow from Pigeonhole. ■

From this, we have cycle $x_1x_2\cdots x_ix_mx_{m-1}\cdots x_{i+1}x_1$ of length $M$. Since it's the case that $M < n$, there's a vertex $y$ not on the cycle. There should be an index $j$ such that $y$, $x_j$ are connected via a path that does not intersect any other $(x_i)$. Hence we have a longer path starting from $y$ and wrapping around the cycle.

Our initial assumption was incorrect, so it must be the case that $M \ge 2\delta + 1$. Therefore $G$ has a path of length $\min\{n, 2\delta(G)+1\}$. □

*Proof Theorem 8.2 (Erdős-Gallai).* We will induct on the number of vertices in $G$. The base case $n \le t$ is clear, since $G$ must be $P$-free. Hence

$$\operatorname{ex}(n, P) = \frac{n(n-1)}{2} \le \frac{n(t-1)}{2}.$$

Now the inductive step: suppose $G$ has $n > t$ vertices and is $P$-free.

First if $G$ is disconnected, just sum over components using the inductive hypothesis:

$$m = \sum m_i = \sum \frac{n_i(t-1)}{2} = \frac{n(t-1)}{2}.$$

Suppose $G$ is connected, i.e. has a path of length $\min\{n, 2\delta(G)+1\}$ by the lemma.

Now if $\delta(G) \ge \frac{1}{2}t$, then $G$ has a path of length $2\delta + 1$, contradiction. Hence we may assume $\delta(G) < \frac{1}{2}t$, so some vertex $w$ has a small degree, i.e. $\deg w \le \frac{1}{2}(t-1)$. We will remove $w$ from our graph.

Let $G'$ be $G$ with $w$ removed. Then $G'$ has at most $\frac{1}{2}(n-1)(t-1)$ edges. Adding back the $\frac{1}{2}(t-1)$ edges incident to $w$, the number of edges in $G$ is at most

$$\frac{(n-1)(t-1)}{2} + \frac{t-1}{2} = \frac{n(t-1)}{2},$$

end proof. □

An equality case is $r$ copies of $K_t$. Thus this upper bound is optimal for infinitely many values of $n$, i.e. the bound is pretty good.