

# Graph Theory

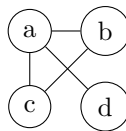
October 31, 2019

## 1 Introduction

### Definitions

Formally, we define a **graph** to be an ordered pair  $G = (V, E)$  where  $E \subseteq V^{(2)} := \{Y \subseteq V : |Y| = 2\}$

e.g.  $V = \{a, b, c, d\}, E = \{\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}\}$



The set  $V(G)$  is the **vertex set** of  $G$  and  $E(G)$  is the **edge set**.

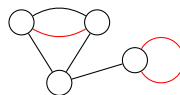
The **order** of  $G$  is  $|V(G)|$ , often written  $|G|$ .

The **size** of  $G$  is  $|E(G)|$ , often written  $e(G)$ .

Note that for the purposes of this course, all graphs will be finite.

This definition that we use precludes:

- Having more than one edge between the same two vertices
- Having an edge with both ends at the same vertex



These 2 red edges are not allowed

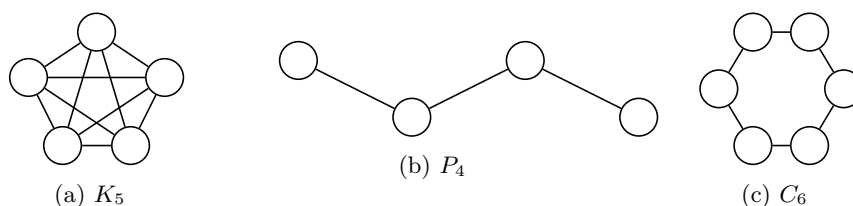
Some authors allow graphs to have these features, and call what we're studying **simple graphs**.

It's easier to write the edge  $\{a, b\}$  as  $ab$ .

If  $ab \in E(G)$ , we say that " $a$  is **adjacent** to  $b$ ", and vice versa. We might write  $ab \in G$ .

Edges containing a vertex  $x$  are said to be **incident** to  $x$  and to each other.

Two graphs  $G, H$  are said to be **isomorphic** if there exists a bijection  $f : V(G) \rightarrow V(H)$  s.t.  $uv \in E(G) \iff f(u)f(v) \in E(H)$ .



There are some canonical graphs:

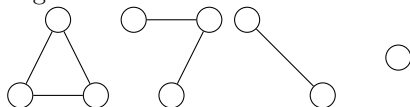
Name	Notation	Order	Size
Empty graph	$E_n$	$n$	0
Complete graph	$K_n$	$n$	$\binom{n}{2}$
Path	$P_n$	$n$	$n - 1$
Circuit	$C_n$	$n$	$n$

A **subgraph** of  $G$  is another graph  $H$  with  $V(H) \subseteq V(G)$ ,  $E(H) \subseteq E(G)$ . (Every graph of order  $\leq n$  is a subgraph of  $K_n$ ).

The subgraph of  $G$  **induced by**  $W \subseteq V(G)$ , written  $G[W]$  is the graph  $(W, E(G) \cap W^{(2)})$ , i.e. the vertex set  $W$  and all edges of  $G$  lying inside  $W$ .

A graph is **connected** if there is a  $u-v$  path (a path from  $u$  to  $v$  in  $G$  for every pair  $u, v \in V(G)$ ). The **components** of  $G$  are the maximal connected subgraphs of  $G$  (induced by the equivalence classes of the relation  $u \leftrightarrow v \iff u = v$  or there is a  $u-v$  path).

E.g.:



has 4 components.

A **forest** is a graph containing no circuit.

A **tree** is a connected forest, i.e. the components of a forest are trees.

Equivalently, a tree is a connected circuit-free graph. This brings us on to our first theorem of the course:

**Theorem 1.1.** *The following are equivalent:*

1.  $G$  is a tree
2.  $G$  is a minimal connected (connected, but the removal of any edge kills connectivity)
3.  $G$  is maximal circuit-free (the addition of any edge creates a circuit)

*Proof.* (a)  $\implies$  (b)

Let  $G$  be connected and circuit free. Suppose  $uv \in G$  and  $G - uv$  is connected. Then there is a  $u-v$  path  $P$  in  $G - uv$ , so  $G$  contains the circuit  $P + uv \nmid (G \text{ circuit free})$

$\therefore G$  is minimal connected.

(b)  $\implies$  (a)

Let  $G$  be minimal connected, but suppose  $G$  has a circuit  $C$ , with  $xy \in C$ . Let  $u, v \in V(G)$ . There is a  $u - v$  path  $P$  in  $G$ .

If  $xy \notin P$ , then  $P$  joins  $u$  to  $v$  in  $G - xy$ .

If  $xy \in P$ , then  $(P \cup C) - xy$  contains a  $u - v$  path in  $G - xy$ .

Hence  $G - xy$  is connected  $\nmid$  ( $G$  **minimal** connected).

$\therefore G$  is circuit free, and connected by hypothesis, hence a tree.

(a)  $\implies$  (c)

Let  $G$  be a tree. If  $uv \notin E(G)$  there is a path  $P$  in  $G$  from  $u$  to  $v$ , so  $P + uv$  is circuit in  $G + uv$ , so  $G$  is maximal circuit free.

(c)  $\implies$  (a)

Let  $G$  be maximal circuit free. Let  $u, v \in V(G)$ . Either  $uv \in E(G)$  or  $G + uv$  has a circuit  $C$ , so  $C - uv$  is a  $u - v$  path in  $G$ , and so  $G$  is connected, and hence a tree.  $\square$

A subgraph  $T$  of a graph  $G$  is said to be **spanning** if  $V(T) = V(G)$ , i.e. the subgraph touches every vertex of  $G$ .

**Corollary 1.2.** *A graph is connected if and only if it has a spanning tree.*

*Proof.*

- $\implies$   $T$  is connected as it is a tree, and is spanning, so all vertices in  $G$  are connected via  $T$ .
- $\Leftarrow$  Remove edges from  $G$  to obtain a minimal connected spanning subgraph  $T$ . By 1.1  $T$  is a tree.

$\square$

The set of **neighbours** of a vertex  $v$  is  $\Gamma(v) = \{w : vw \in E(G)\}$ .

The **degree**  $d(v)$  is the number of neighbours of  $v$ ,  $|\Gamma(v)|$ .

The minimum and maximum degrees in  $G$  are denoted by  $\delta(G)$  and  $\Delta(G)$ .

If  $\delta(G) = \Delta(G) = k$ , i.e.  $d(v) = k \forall v$  we say  $G$  is  **$k$ -regular**. A **regular** graph is one that is  $k$ -regular for some  $k$ . If a graph is 3-regular, it is called **cubic**.

The degrees of the graph written in some order form a **degree sequence** of  $G$ .

**Lemma 1.3** (lemma Lemma).

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

*Proof.* Each edge has 2 vertices at the end, so contributes 2 to the sum of all degrees.  $\square$

A **leaf** is a vertex of degree 1.

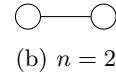
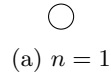
**Theorem 1.4.** *Every tree  $T$  with  $|T| \geq 2$  has at least two leaves.*

*Proof.* We note  $T$  is connected so  $\delta(T) \geq 1$ . Let  $x_1$  be a vertex (a leaf if there is one). Let  $x_1x_2 \dots x_k$  be a maximal path from  $x_1$ .  $T$  is circuit-free so none of  $x_1, \dots, x_{k-2}$  is in  $\Gamma(x_k)$ , and  $x_k$  has no other neighbours except for  $x_{k-1}$ , otherwise this path can be extended, hence  $\Gamma(x_k) = \{x_{k-1}\}$ , so  $x_k$  is a leaf. Repeat the proof starting at  $x_k$  for a second leaf.  $\square$

**Corollary 1.5.** *A tree of order  $n$  has size  $n - 1$*

*Proof.*

This is true for  $n = 1, 2$  by inspection, as there is only 1 tree:



Let  $|T| \geq 3$ . By **1.3**  $T$  has a leaf  $v$ . Then  $T - v$  is certainly circuit free and connected, because the  $x - y$  path in  $T$  uses  $v$  if and only if  $x$  or  $y$  is  $v$ . Hence the corollary follows by induction.  $\square$

**Corollary 1.6.** *The following are equivalent for a graph  $G$  of order  $n$ :*

1.  $G$  is a tree
2.  $G$  is connected and  $e(G) = n - 1$
3.  $G$  is acyclic and  $e(G) = n - 1$

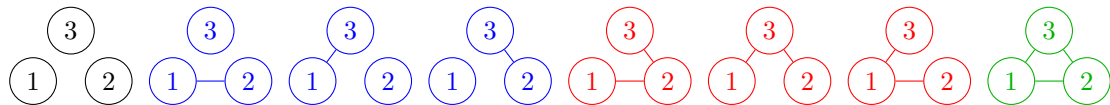
*Proof.*

- (1)  $\implies$  (2), (3): True by definition and **1.5**.
- (2)  $\implies$  (1): By **1.2**  $G$  contains a spanning tree  $T$ . By **1.5**  $e(T) = e(G)$  so  $T = G$
- (3)  $\implies$  (1): Add edges to  $G$  to get a graph  $G'$  maximal acyclic. By **1.1**  $G'$  is a tree, so by **1.5**  $e(G') = n - 1 = e(G)$ , so  $G' = G$ .

$\square$

## How many graphs of order $n$ are there?

Let the vertex set be  $[n] := \{1, 2, \dots, n\}$ . E.g. for  $n = 3$  we have:



There are  $2^{\binom{n}{2}}$  **labelled** graphs of order  $n$  -  $\binom{n}{2}$  potentially edges, each of which we can independently choose to include or not.

However, there are only 4 unlabelled graphs of order 3 - all the graphs of the same colour are essentially the same, just with different labels given to the vertices. In general we must count the number of orbits amongst labelled graphs under the action of a group, using Burnside's Lemma. In fact, it turns out this number is asymptotic to  $2^{\binom{n}{2}}/n!$

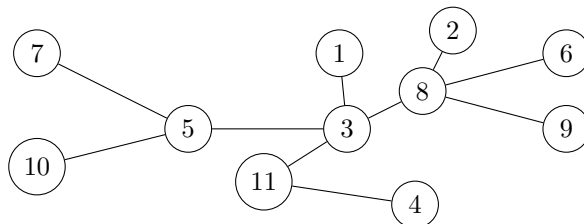
## How many trees of order $n$ are there?

**Theorem 1.7** (Cayley). *There are  $n^{n-2}$  labelled trees of order  $n$ .*

*Proof.* (Prüfer, Non-examinable)

We construct a bijection between trees strings length  $n - 2$  over the alphabet  $n$ .

- Tree  $\rightarrow$  string: Select the lowest leaf, write down its neighbour, remove that leaf, and repeat until only one edge is left.



Here, we would get the string 3, 8, 11, 8, 5, 8, 3, 5, 10. Each vertex is written down  $d(n) - 1$  times

- String  $\rightarrow$  tree: Mark all vertices "unused". Then repeatedly choose the smallest unused vertex not in the string. Join it to the first vertex in the string. Mark it as used. Delete first vertex in the string.

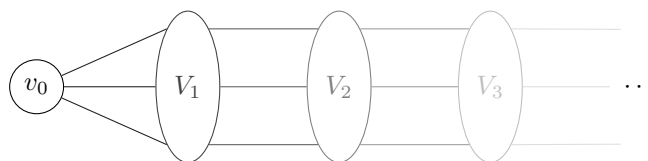
□

A graph is ***r-partite*** if its vertices can be partitioned into  $r$  classes so that no edge lies within a class. In the special case where  $r = 2$  we call the graph ***bipartite***. Remarkably, we can characterise all bipartite graphs.

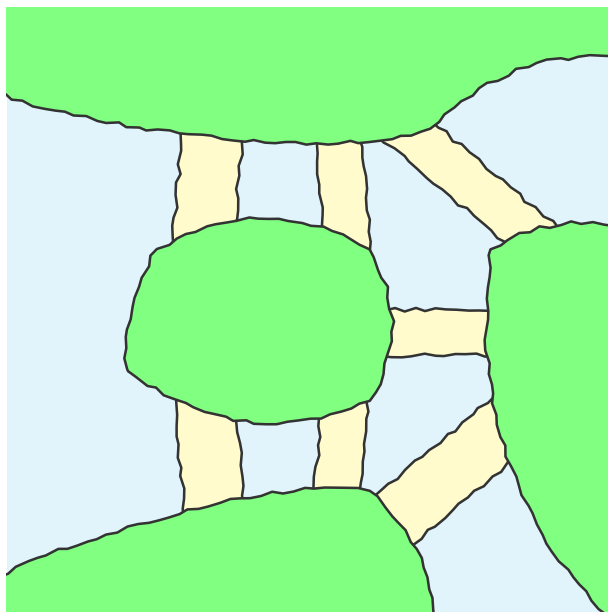
**Theorem 1.8.** *A graph is bipartite if and only if it contains no odd circuit.*

*Proof.* For the "only if" part, in a bipartite graph the vertices of a circuit alternate between the two classes, otherwise an edge would lie in a class. Hence any circuit must be of even length.

For the "if" part, take some graph  $G$ , pick a vertex  $v_0$  and let  $V_i = \{v : d(v_0, v) = i\}$  where  $d(v_0, v)$  is the **distance** of  $v$  from  $v_0$ , i.e. the length of a shortest  $v_0 - v$  path.



If  $G$  is connected,  $V_0, V_1, V_2, \dots$  partition  $V(G)$ . (Note  $V_0 = \{v_0\}$ ). There is no edge between  $V_i$  and  $V_{i+k}$  if  $k \geq 2$  by construction. Suppose  $G$  has no odd circuit. Then  $G$  cannot have an edge inside one of the  $V_i$ : if  $u, v \in V_i$  and  $uv \in G$ , then there is a way of getting from  $v_0$  to itself in an odd number of steps.  $v_0 \xrightarrow{i} u \xrightarrow{1} v \xrightarrow{i} v_0$ . This must contain an odd circuit: let the first path be  $v_0 = a_0, a_1, a_2, \dots, a_{i-1}, u$  and the second be  $v_0 = b_0, b_1, b_2, \dots, b_{i-1}, v$ , and let  $k = \max\{j : a_j = b_j\}$ , so  $0 \leq k < i$ . Then  $a_k, a_{k+1}, \dots, u, v, b_{i-1}, \dots, b_k = a_k$  is an odd cycle. Hence we can let our two classes be  $\cup_i V_{2i}$  and  $\cup_i V_{2i+1}$ , and  $G$  is bipartite. □



The bridges of Königsberg, which inspired this problem

An **Euler cycle** as a “walk” around all the edges of a graph returning to the starting point so that every vertex is visited and every edge is used precisely once.

A graph is **Eulerian** if it is a single vertex or has an Euler cycle.

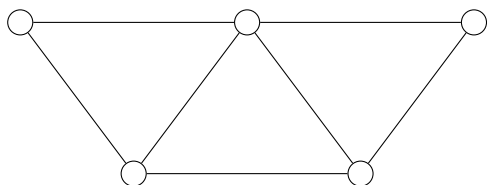
**Theorem 1.9.** *A graph is Eulerian if and only if it is connected and every vertex has even degree*

*Proof.* The forwards direction is clear: an Eulerian graph must be connected, and for every time we enter a vertex along an edge we must also leave that vertex, so each vertex has even degree.

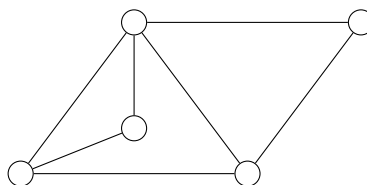
We do the opposite direction by induction on  $e(G)$ . Since  $\delta(G) \geq 2$  by 1.4  $G$  is not a tree and so has a circuit  $C$ . Every vertex of  $G - E(C)$  has even degree, so by induction each component of  $G - E(C)$  is Eulerian. Now set off walking round  $C$ , taking the time to traverse any Euler cycle of  $G - E(C)$  when first encountered. Since  $G$  is connected, this gives an Euler cycle of  $G$ .  $\square$

A graph that can be drawn in the plane without edges crossing is **planar**. A graph drawn in the plane is a **plane graph**. A **face** of a plane graph is a connected region of  $\mathbb{R}^2 \setminus G$ . We define the **size** of a face to be the number of edges it has.

E.g.:



(a) Face sizes 3, 3, 3, 5



(b) Face sizes 3, 3, 4, 4

These two graphs are the same planar graph, but are different plane graphs. Note that we include the outer infinite face. There are no analytical/topological difficulties involved. Any planar graph can be drawn with piecewise linear edges, and in fact with straight line edge (see example sheet 1), so any issues are purely combinatorial.

**Theorem 1.10** (Euler). *Let  $G$  be a connected plane graph with  $n$  vertices,  $m$  edges, and  $f$  faces. Then  $n - m + f = 2$ .*

*Proof.*  $G$  has a spanning tree with  $n - 1$  edges, so  $m \geq n - 1$ . If  $m = n - 1$  then  $G$  is a tree, in which case  $f = 1$  so done. We now proceed by induction on  $m$ . If  $m > n - 1$ , there is a circuit  $C$ . If we remove an edge from this circuit, then we obtain a graph with 1 fewer edge and 1 fewer face, and so by the induction hypothesis,  $n - (m - 1) + (f - 1) = 2 \implies n - m + f = 2$ .  $\square$

A **bridge** (or **isthmus**) is an edge whose removal increases the number of components. Equivalently, it's an edge which is not part of any circuit. In a connected bridgeless plane graph, every edge separates two faces. So if there are  $f_i$  faces of length  $i$ , then  $\sum_i f_i = f$  and  $\sum_i i f_i = 2m$ .

The **girth** of a graph is the length of its shortest circuit.

**Theorem 1.11.** *Let  $G$  be a connected bridgeless planar graph of girth  $g$ , order  $n$ , size  $m$ . Then:*

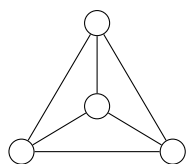
$$m \leq \frac{g}{g-2}(n-2)$$

*In particular, every planar graph of order  $n \geq 3$  has size at most  $3n - 6$  and contains a vertex of degree  $\leq 5$ .*

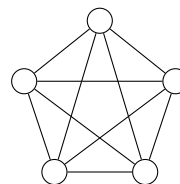
*Proof.* Draw  $G$  in the plane, with  $f_i$  faces of length  $i$ . Then  $2n = \sum_i i f_i \geq g \sum_i f_i = g f$ . By 1.10,  $n - 2 = m - f \geq m - \frac{2}{g}m = m \frac{g-2}{g}$ , establishing the claimed formula.

Now given any planar graph, add edges if necessary to make it connected and bridgeless. Since the girth  $g \geq 3$ ,  $m \leq 3n - 6$ . By the lemma lemma, the average degree is less than 6.  $\square$

## Some Planar Graphs

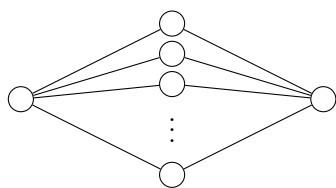


$K_4$  is planar

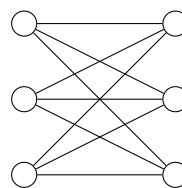


$K_5$  is non-planar, as violates the  $m \leq 3n - 6$

The **complete bipartite** graph  $K_{m,n}$  is bipartite with class sizes  $m$  and  $n$  and all edges between, so  $|K_{m,n}| = m + n$ ,  $e(K_{m,n}) = mn$ , girth = 4 if  $m, n \geq 2$ .



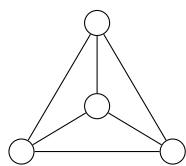
$K_{2,n}$  is planar



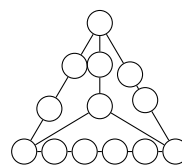
$K_{3,3}$  is non-planar, as violates the  $m \leq 3n - 6$

Any graph of order  $n$  with more than  $3n - 6$  edges is non-planar but there are non-planar graphs with many fewer edges.

Take any graph  $G$ . A **subdivision** or **topological**  $G$  is any graph obtained from  $G$  by replacing its edges with paths whose internal vertices are disjoint. Any such subdivision is denoted by  $TG$ .



$G$



$TG$

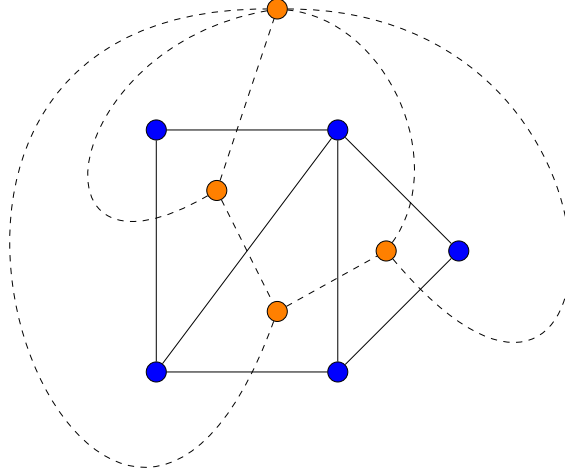
Kuratowski (1930) and Pontryagin (much earlier, unpublished) characterised planar graphs:

**Theorem 1.12** (Kuratowski).  $G$  is planar if and only if it contains no  $TK_5$  and no  $TK_{3,3}$ .

We will not prove this here.

Given a plane graph, we can draw a new graph, the **dual graph**, by placing a dual vertex inside each original face and adding a dual edge for every original edge separating the two faces.





If  $G$  has  $v$  vertices,  $e$  edges,  $f$  faces; dual of  $G$  has  $f$  vertices,  $e$  edges,  $v$  faces. The dual of the dual is the original graph. Note that, as seen in this example, we do allow there to be multiple edges between the same pair of vertices in the dual graph.

## 2 Matchings and Connectivity

Let  $G$  be a bipartite graph with vertex classes  $X, Y$ . A **matching** from  $X$  to  $Y$  is a set of  $|X|$  independent edges (edges that are not incident).

If  $|X| = |Y|$ , such a matching is also a matching of  $Y$  to  $X$ . We call it a **1-factor** (i.e. a **1-regular** spanning subgraph). We can think of this in terms of arranged marriages of things in  $X$  and  $Y$ , but I'm not sure why you would. A matching from  $X$  to  $Y$  does not always exist (for instance, if  $|X| > |Y|$ ), but remarkably, a simple necessary condition is also sufficient.

Firstly, given  $A \subset V(G)$  define  $\Gamma(A) = \cup_{a \in A} \Gamma(a)$ . Then we can say:

**Theorem 2.1** (Hall's Marriage Theorem). *Let  $G$  be a bipartite graph with bipartition  $X, Y$ . Then there is a matching from  $X$  to  $Y$  if and only if:*

$$|\Gamma(A)| \geq |A| \quad \forall A \subseteq X$$

*This condition is known as Hall's condition.*

*Proof.* The “only if” part is immediate. We will prove the “if” part in several ways:

Proof 1:

By induction on  $|X|$ .

Either for all  $\emptyset \neq A \neq X$  the inequality  $|\Gamma(A)| > |A|$  holds. Let  $xy$  be any edge. Then Hall's condition still holds in  $G' = G - x - y$ . Then by induction match  $X - x$  to  $Y - y$ , which together with  $xy$  gives a matching  $X$  to  $Y$ .

Or there exists a “critical set”  $\emptyset \neq B \neq X$  with  $|\Gamma(B)| = |B|$

Let  $G_1 = G[B \cup \Gamma(B)]$ ,  $G_2 = G[(X \setminus B) \cup (Y \setminus \Gamma(B))]$ . Let  $A \subset B$ , then  $\Gamma(A) \subset \Gamma(B)$ . Since Hall's condition holds in  $G$  it holds in  $G_1$ . Let  $A \subset X \setminus B$ . Then  $\Gamma_2(A) := \Gamma(A \cup B) \setminus \Gamma(B)$ , so  $|\Gamma_2(A)| = |\Gamma(A \cup B)| - |\Gamma(B)| \geq |A \cup B| - |B|$  since  $|\Gamma(B)| = |B|$ . So Hall's condition holds in  $G_1$  and  $G_2$ .

**Proof 2** (Rado): Remove edges to get a minimal graph where Hall's Condition holds. If  $d(a) = 1$  for all  $a \in X$ , then what's left is a matching. If not, then some  $a \in X$  has two neighbours  $b_1, b_2$  and sets  $A_1, A_2$  such that  $|\Gamma(A_i)| = |A_i|$ , and  $\Gamma(a) - b_i \subset \Gamma(A_i)$ . So  $|\Gamma(A_1 \cup A_2 \cup \{a\})| = |\Gamma(A_1 \cup A_2)|$ , and  $|\Gamma(A_1 \cup A_2 \cup \{a\})| = |\Gamma(A_1 \cup A_2)| = |\Gamma(A_1) \cup \Gamma(A_2)| = |\Gamma(A_1)| + |\Gamma(A_2)| - |\Gamma(A_1) \cap \Gamma(A_2)| \leq |\Gamma(A_1)| + |\Gamma(A_2)| - |\Gamma(A_1 \cap A_2)| \leq |A_1| + |A_2| - |A_1 \cap A_2| = |A_1 \cup A_2| < |A_1 \cup A_2 \cup \{a\}| \nlessdot$ .  $\square$

**Corollary 2.2** (Defect Form). *Let  $G$  be a bipartite graph with bipartition  $X, Y$  and let  $d \in \mathbb{N}$ . Then  $G$  has  $|X| - d$  independent edges if and only if  $|\Gamma(A)| \geq |A| - d \forall A \subset X$ .*

*Proof.* Introduce  $d$  members of  $Y$  with edges to all of  $X$ . Now Hall's condition holds, so there is a matching in this new graph by **2.1**. Then remove the extra vertices.  $\square$

**Corollary 2.3** (Polyandrous version). *Let  $G$  be a bipartite graph,  $d \in \mathbb{N}$ . Then each woman can have  $d$  husbands (i.e for each member of  $X$  we can assign  $d$  members of  $Y$  so that none of those members of  $Y$  are shared among members of  $X$ ) if and only if  $|\Gamma(A)| \geq d|A|$*

*Proof.* Clone  $X$   $d$  times (so that all the clones have the same edges to  $Y$  as their original). Hall's condition holds for the clones, so we marry the clones to the men, then give the clones' husbands to the original women.  $\square$

Given a family of subsets of  $Y$ , say  $\mathfrak{F} = \{F_1, F_2, \dots, F_n\} \subseteq \mathcal{P}(Y)$ , a **transversal** or **set of distinct representations** is a collection  $\{y_1, \dots, y_n\}$  with  $y_i \in F_i$

**Corollary 2.4.** *There exists a set of distinct representations if and only if  $|\cup_{i \in I} F_i| \geq |I| \forall I \subset [n]$ .*

*Proof.* Construct the bipartite graph with  $X = \mathfrak{F}$  and  $E = \{Fy : y \in F\}$ . Then apply Hall's theorem.  $\square$

A graph  $G$  is  **$k$ -connected** if it has more than  $k$  vertices and  $G - S$  is connected for every subset  $S \subseteq V(G)$  such that  $|S| < k$ . Note that  $k$ -connected *implies*  $k - 1$ -connected and so on.

The **vertex connectivity** of  $G$  is  $\kappa(G) = \max\{k : G \text{ is } k\text{-connected}\}$ . If  $G$  is not complete,  $\kappa(G) = \min\{|S| : G - S \text{ disconnected}\}$ .

Let  $a, b \in V(G)$ ,  $ab \notin E(G)$ . Then we can define their **local connectivity**  $\kappa(a, b; G) = \min\{|S| : S \subseteq V(G) \setminus \{a, b\}, \text{ there is no } a - b \text{ path in } G - S\}$ , i.e. the fewest vertices we can remove to separate  $a$  from  $b$ . Then, if  $G$  is not complete,  $\kappa(G) = \min_{a, b \notin E(G)} \kappa(a, b; G)$ .

There are corresponding edge-connectivities:  $\lambda(G) = \min\{|F| : F \subseteq E(G), G - F \text{ disconnected}\}$ .

Given  $a, b \in V(G)$ , a set  $\{P_1, P_2, \dots, P_t\}$  of  $a - b$  paths is **vertex disjoint** if  $P_i$  and  $P_j$  have no vertex in common except for  $a$  and  $b$ . How large can a set of vertex disjoint  $a - b$  paths be? Certainly, the answer is at most  $\kappa(a, b; G)$ , since any set  $S$  must include at least one vertex from each path in order for  $G - S$  to have no  $a - b$  path.

**Theorem 2.5** (Menger). *Let  $a, b \in V(G)$  and  $ab \notin E(G)$ . Then there exists a set of  $\kappa(a, b; G)$  vertex disjoint  $a - b$  paths, and no larger set.*

*Proof.* We've already noted that there is no larger set. Note also that the theorem is immediate if  $\kappa(a, b; G) = 0$  or 1. If you know the Max Flow-Min Cut theorem, the result is immediate, but we will prove it by hand:

Suppose not, i.e. that there is no set of  $\kappa(a, b; G)$  vertex disjoint  $a - b$  paths for some  $G, a, b$ . Then let  $G, a, b$  be a minimal counterexample (where minimal means firstly over vertices, then over edges). Let  $k = \kappa(a, b; G) \geq 2$ .

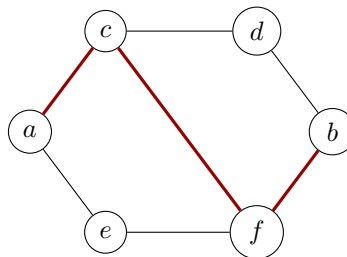
We then say  $S \subseteq V(G) - \{a, b\}$  **separates**  $a$  from  $b$  if there is no  $a - b$  path in  $G - S$ . Suppose there is a separator  $S$  with  $|S| = k$  with  $S \not\subseteq \Gamma(a)$  and  $S \not\subseteq \Gamma(b)$ . Let  $C$  be the component of  $G - S$  containing  $a$ . Form a new graph  $G_a$  by “contracting  $C$  to a point”, i.e. taking  $G - C$  and add a new vertex  $a^*$  joined to all of  $S$ .

Note that every vertex of  $S$  has a neighbour in  $C$ , else a proper subset of  $S$  separates  $a$  from  $b$ . Since  $S \not\subseteq \Gamma(a)$  we have  $e(G_a) < e(G)$ , so the theorem holds for  $G_a$ . Note that any set separating  $a^*$  from  $b$  in  $G_a$  separates  $a$  from  $b$  in  $G$ , so  $\kappa(a^*, b; G_a) \geq k$ , and so by the theorem there is a set of vertex disjoint  $a^* - b$  paths in  $G_a$ , giving a set of  $k$  paths  $P_1, \dots, P_k$  from  $S$  to  $b$  in  $G$ , disjoint except at  $b$ . Similarly, we can define  $G_b$  and  $k$  paths  $Q_1, \dots, Q_k$  from  $S$  to  $a$ , disjoint except at  $a$ . Note that  $P_i$  has no vertex in common with  $Q_j$  except in  $S$ , so we can concatenate them together to get  $k$  vertex-disjoint  $a - b$  paths.

Hence every separator  $S$  of size  $k$ , satisfies  $S \subseteq \Gamma(a)$  or  $S \subseteq \Gamma(b)$ . Now if  $\exists x \in \Gamma(a) \cap \Gamma(b)$  then  $\kappa(a, b; G - x) \geq k - 1$ , so by minimality  $G - x$  has a set of  $k - 1$  vertex disjoint  $a - b$  paths which with  $axb$  gives a set of  $K$  vertex disjoint paths in  $G$ , and so the only case left to consider is  $\Gamma(a) \cap \Gamma(b) = \emptyset$ .

Pick a shortest  $a - b$  path, say  $a, u, v, \dots, b$ . Then  $\{u, v\} \cap \{a, b\} = \emptyset$ . There is no set of  $k$  vertex disjoint  $a - b$  paths in  $G - uv$ , and  $G - uv$  is not a counterexample, so  $\kappa(a, b; G - uv) < k$ . So there is a set  $S' \subset V(G)$ ,  $|S'| < K$  separating  $a$  from  $b$  in  $G - uv$ . Then both  $S' \cup \{u\}$  and  $S' \cup \{v\}$  separate  $a$  from  $b$  in  $G$ , so  $|S' \cup \{u\}| \geq k$ . Moreover,  $|S' \cup \{u\}| \leq k$ , and so  $= k$ , and likewise  $|S' \cup \{v\}| = k$ . Since  $u \notin \Gamma(b)$  we have  $S' \cup \{u\} \subset \Gamma(a)$ . Since  $v \notin \Gamma(a)$  as we took the shortest path, we have  $S' \cup \{v\} \subset \Gamma(b)$ . Then  $S' \subset \Gamma(a) \cap \Gamma(b)$  and  $|S'| = k - 1 > 0$ .  $\square$

NOTE if, say,  $\kappa(a, b; G) = 2$ , then Menger gives a set of 2 vertex disjoint  $a - b$  paths, but it will NOT give an extra path disjoint from some initial chosen path.



In this graph,  $\kappa(a, b; G) = 2$ , by removing say  $\{c, e\}$ , but we cannot get an  $a - b$  path vertex disjoint from the red path.

**Corollary 2.6.** Let  $\kappa(G) \geq k$ , and let  $X, Y \subset V(G)$ ,  $X \cap Y = \emptyset$ ,  $|X| \geq k, |Y| \geq k$ . Then there exists a set of  $k$  disjoint paths from  $X$  to  $Y$ .

*Proof.* From a new graph  $G^*$  by adding a new vertex  $x$  joined to all  $X$  and new  $y$  joined to all of  $Y$ . Then  $\kappa(x, y; G^*) \geq k$ , and so by Menger  $G^*$  has a set of  $k$  vertex disjoint  $x - y$  paths.  $\square$

There is also an edge form of Menger:

**Theorem 2.7** (Menger for edges). *Let  $G$  be a graph and  $a, b \in G$ . Then there's a set of  $\lambda(a, b; G)$  edge-disjoint  $a - b$  paths, and no larger set.*

*Proof.* Either mimic the proof of Menger's theorem, or instead construct the **line graph**  $L(G)$  of  $G$  whose vertex set is  $E(G)$ . Join two vertices  $e, f$  in  $L(G)$  if  $e, f$  are incident edges of  $G$ . Add a new vertex  $a^*$  to  $L(G)$  joined to every vertex of  $L(G)$  which is an edge of  $G$  incident with  $a$ , and  $b^*$  likewise. Call this new graph  $G^*$ .

Then  $\kappa(a^*, b^*; G^*) = \lambda(a, b; G)$ , and so there's a set of  $\lambda(a, b; G)$  vertex disjoint  $a^* - b^*$  paths in  $G^*$ . This contains a set of  $\lambda(a, b; G)$  edge disjoint  $a - b$  paths in  $G$ .  $\square$

### 3 Extremal Graph Theory

A **Hamiltonian circuit** in a graph is a spanning circuit, i.e. a circuit through every vertex. A graph is **Hamiltonian** if it has a Hamiltonian circuit. The problem of determining whether a graph is Hamiltonian is related to the travelling cycle, and it is very likely there is no nice necessary and sufficient condition.

Question: How many edges are needed to guarantee having a Hamiltonian circuit?

Answer: We need more than  $\binom{n-1}{2} + 1$ , i.e. to be nearly complete: consider  $K_{n-1}$  with a single leaf hanging off it.

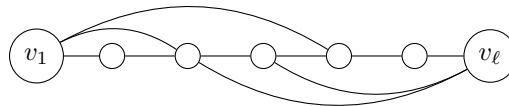
Question: What minimum degree guarantees a Hamiltonian circuit? Note that  $K_{\lceil \frac{n+1}{2} \rceil}$  and  $K_{\lfloor \frac{n+1}{2} \rfloor}$  with a common vertex shows that we need  $\delta(G) \geq \frac{n}{2}$ . So does  $K_{\lfloor \frac{n-1}{2} \rfloor, \lfloor \frac{n+1}{2} \rfloor}$ .

**Theorem 3.1.** *Let  $k \in \mathbb{N}$ , and let  $G$  be a graph of order  $n$  such that  $d(a) + d(b) \geq k$  holds for every pair  $a, b$  of non-adjacent vertices. Then:*

- *If  $k < n$  and  $G$  is connected, then  $G$  has a path of length  $k$ .*
- *If  $k = n$  then  $G$  has a Hamiltonian circuit.*

*Proof.* If  $k = n$  then  $d(a) + d(b) \geq n$ , then  $a, b$  have a common neighbour so in every case  $G$  is connected. We may assume that  $G$  has no Hamiltonian circuit, else in every case we are done, because a Hamiltonian circuit contains paths of all lengths.

Let  $P = v_1 v_2 v_3 \dots v_\ell$  be a path of maximum length. Note that  $G$  has no circuit of length  $\ell$ . Note that  $G$  has no circuit of length  $\ell$  because if  $\ell = n$  then this is a Hamiltonian circuit. Moreover, if  $\ell < n$  we get a path of length  $\ell$  by connectivity. In particular,  $v_1, v_\ell \notin E(G)$  so  $d(v_1) + d(v_\ell) \geq k$ .



Let  $S = \{i : v_1 v_i \in E(G)\}$ , and  $T = \{i : v_{i-1} v_\ell \in E(G)\}$ .

Since  $P$  has maximal length ever neighbour of  $v_1, v_\ell$  lies in  $P$ . So  $|S| = d(v_1)$ ,  $|T| = d(v_\ell)$ . Clearly  $S \cup T \subset \{2, \dots, \ell\}$ . But  $S \cap T = \emptyset$  else there is some  $j \in S \cap T$  such that  $v_1, \dots, v_{j-1}, v_\ell, v_{\ell-1}, \dots, v_j$  is a circuit of length  $\ell$ .

Hence  $\ell - 1 \geq |S \cup T| = |S| + |T| = d(v_1) + d(v_\ell) \geq k$ , giving a path of length  $k$ . If  $k = n$  this is impossible and so there is a Hamiltonian circuit.  $\square$

**Corollary 3.2** (Dirac). *If  $\delta(G) \geq \frac{|G|}{2}$ , then  $G$  is Hamiltonian.*

Note: **3.1** is the best possible we can do if  $\frac{k-1}{2}|n-1|$ .

**Theorem 3.3.** *Let  $G$  be a graph of order  $n$  with no path of length  $k$ . Then  $e(G) \leq (k-1)\frac{n}{2}$ . Equality holds if and only if  $k|n$  and  $G$  is the disjoint union of copies of  $K_k$ .*

*Proof.* We prove this by induction on  $n$ . The cases  $n \leq k$  are trivial. In general, if  $G$  is disconnected this theorem follows immediately by the induction hypothesis on each component. If  $G$  is connected then by **3.1**  $\delta(G) \leq \frac{k-1}{2}$ . Let  $x$  be a vertex of degree  $\leq \frac{k-1}{2}$ . Since  $G$  is connected, it cannot contain  $K_k$ , and so by the induction hypothesis  $e(G-x) < \frac{(k-1)(n-1)}{2}$ .

Thus  $e(G) < (k-1)\frac{n}{2}$   $\square$

This is a nice example of an extremal theorem. It also describes the extremal graphs where equality is attained. We now ask a similar question: “How many edges can a triangle free graph have?”. Note that no bipartite graph has a triangle, so the answer is at least  $e(K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}) = \lfloor \frac{n^2}{4} \rfloor$ .

**Theorem 3.4** (Mantel, 1907). *If  $|G| = n$  and  $G$  has no  $K_3$  then  $e(G) \leq \lfloor \frac{n^2}{4} \rfloor$ .*

*Proof.* By induction on  $n$ . The base cases of  $n \leq 2$  are trivial. We do the induction step in 2 different ways:

1. Let  $ab \in E(G)$ . No vertex of  $G - \{a, b\}$  is joined to more than one of  $a, b$  otherwise we have a triangle. So  $e(G) \leq 1 + (n-2) + e(G - \{a, b\})$ . By the induction hypothesis,  $e(G - \{a, b\}) \leq \lfloor \frac{(n-2)^2}{4} \rfloor = \lfloor \frac{n^2}{4} \rfloor - n + 1$ . So  $e(G) \leq \lfloor \frac{n^2}{4} \rfloor$ .
2. If  $\delta(G) > \lfloor \frac{n}{2} \rfloor$   $G$  has a  $K_3$ . Hence  $\delta(G) \leq \lfloor \frac{n}{2} \rfloor$ . Let  $x$  be a vertex of degree  $\leq \lfloor \frac{n}{2} \rfloor$ . Then  $G-x$  has no  $K_3$ , and so by induction  $e(G-x) \leq \lfloor \frac{(n-1)^2}{4} \rfloor$ . Hence  $e(G) \leq \lfloor \frac{(n-1)^2}{4} \rfloor + \lfloor \frac{n}{2} \rfloor = \lfloor \frac{n^2}{4} \rfloor$ .

$\square$

Note that no  $r$ -partite graph contains  $K_{r+1}$ . What  $r$ -partite graph of order  $n$  has the most edges? It is complete  $r$ -partite, i.e. every vertex is connected to every other vertex in a different class. Moreover, if we have two classes  $X, Y$  with  $|X| \geq |Y| + 2$  then moving a vertex from  $X$  to  $Y$  we gain  $-|Y| + |X| - 1$  edges, which is positive.

The ***r-partite Turán graph of order  $n$***  is the complete  $r$ -partite graph with class sizes differing by  $\leq 1$ , i.e. of sizes  $\lfloor n/r \rfloor$  or  $\lceil n/r \rceil$ . We denote it by  $T_r(n)$  and its number of edges is  $t_r(n)$ .

Note that  $T_2(n) = K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ , and  $t_2(n) = \lfloor n^2/4 \rfloor$ .

To avoid grief with rounding, when calculating with  $t_r(n)$ , the following observations about the graph  $T_r(n)$  are useful:

- \* If  $n \geq r$  then  $T_r(n)$  contains at least one  $K_r$ . Each other vertex has exactly  $r-1$  neighbours in  $K_r$ , and the removal of  $K_r$  gives  $T_r(n-r)$ . Hence  $t_r(n) = \binom{r}{2} + (r-1)(n-r) + t_r(n-r)$ .
- \*\* If we remove a vertex of minimal degree in  $T_r(n)$ , it lies in a largest class and its removal leaves  $T_r(n-1)$ . So  $t_r(n) - \delta(T_r(n)) = t_r(n-1)$ .
- \*\*\*  $\Delta(T_r(n)) \leq \delta(T_r(n)) + 1$  if  $\delta(G) > \delta(T_r(n))$  then  $e(G) > t_r(n)$ , by comparing degree sequences.

**Theorem 3.5** (Turán 1941). *Let  $G$  be a  $K_{r+1}$ -free graph of order  $n$ . Then  $e(G) \leq t_r(n)$  with equality if and only if  $G = T_r(n)$ .*

Note that the case  $r = 2$  gives Mantel's theorem.

*Proof.* By induction on  $n + r$ . If  $n \leq r$ , the theorem is trivial, as  $T_r(n)$  is complete. We assume  $G$  is  $K_{r+1}$ -free,  $e(G) \geq t_r(n)$ , and show  $G = T_r(n)$  in 2 different ways:

1.  $G$  contains  $K_r$  by induction on  $r$ , as  $t_r(n) > t_r(n-1)$ . Pick a  $K_r$  and call it  $K$ . Each vertex of  $G - K$  has at most  $r-1$  neighbours in  $K$ , as otherwise we have a  $K_{r+1}$ . So:

$$e(G) \leq \binom{r}{2} + (r-1)(n-r) + e(G-K)$$

But  $G - K$  is  $K_{r+1}$  free, and so  $E(G - K) \leq t_r(n-r)$  by induction. We thus have  $e(G) = t_r(n)$  and equality holds throughout. So, by induction,  $G - K = T_r(n-r)$  and each vertex of  $G - K$  has exactly  $r-q$  neighbours in  $K$ , i.e. is not joined to exactly one vertex of  $K$ . Notice vertices in different classes of  $G - K$  must miss different vertices of  $K$  (else we have  $K_{r+1}$ ), and each vertex of  $K$  must miss at least one whole class of  $G - K$  (else we have  $K_{r+1}$ ). So  $G = T_r(n)$ .

2. First remove edges to get a subgraph  $G'$  with  $e(G') = t_r(n)$ . From last time, we must have  $\delta(G') \leq \delta(T_r(n))$ . Let  $x$  be a vertex of minimal degree in  $G'$ . Then  $e(G' - x) \geq t_r(n-1)$ . But  $G' - x$  is  $K_{r+1}$ -free, and so by the inductive hypothesis, equality must hold throughout, and  $G' - x = T_r(n-1)$ . Now  $x$  cannot have a neighbour in every class of  $G' - x$ , otherwise we get  $K_{r+1}$ , so  $G'$  is  $r$ -partite. Since  $T_r(n)$  is unique  $r$ -partite of size  $t_r(n)$  we have  $G' = T_r(n)$ . No edge can be added to  $G'$  without  $K_{r+1}$ , so  $G' = G$ .

□

We might ask, for any fixed  $F$ , what is the extremal function:

$$\text{ex}(n, F) = \max\{e(G) : |G| = n, G \text{ doesn't contain } F\}.$$

Mantel showed that  $\text{ex}(n, K_3) = t_2(n)$ . Turán showed  $\text{ex}(n, K_r + 1) = t_r(n)$ , and earlier showed that  $\text{ex}(n, P_{k+1}) \leq (k-1)\frac{n}{2}$ .

In general, it is hard to get an exact answer, but sometimes we can get exact for large  $n$ . e.g.,  $\text{ex}(n, C_5) = t_2(n)$  for  $n \geq 6$ , and  $\text{ex}(n, P) = t_2(n) + n - 2$  for  $n > n_0$  where  $P$  denotes the Petersen graph and  $n_0$  is large. In general however we can't get exact even for  $n$  large, and instead look for an asymptotic value, i.e. can we find  $\lim_{n \rightarrow \infty} \frac{\text{ex}(n, F)}{\binom{n}{2}}$  (Exercise: show this limit exists).

Note that  $t_r(n) = (1 - \frac{1}{r})\binom{n}{2} + O(n)$ , so in the case  $F = K_{r+1}$  this limit is  $1 - \frac{1}{r}$ . The basis of further progress is the remarkable Erdős-Stone theorem, saying that if  $e(G)$  is a bit more than  $t_r(n)$ , then  $G$  contains not just  $K_{r+1}$  but  $K_{r+1}(t) = T_{r+1}((r+1)t)$ , the complete  $(r+1)$ -partite graph with  $t$  vertices in each class. Note that  $K_{r+1}(1) = K_{r+1}$ . This theorem holds for  $n$  “sufficiently large”. we first prove a weaker version with large minimum degree, but it’s the heart of the proof of E-S.

**Lemma 3.6.** *Let  $r, t \geq 1$  be integers  $\geq 0$ , and let  $\epsilon > 0$  be real. If  $n$  is sufficiently large, then every graph  $G$  with  $|G| = n$  and  $\delta(G) \geq (1 - \frac{1}{r} + \epsilon)n$  contains  $K_{r+1}(t)$ .*

*Proof.* Pick some large  $T$  depending on  $r, t, \epsilon$ .  $T(r\epsilon)^t \geq t$  is enough. We will:

1. Show  $G$  contains  $K_r(T)$  which we call  $K$ .
2. There are  $t$  vertices of  $G - K$  joined to  $K$  in such a way as to form  $K_{r+1}(t)$ .

For the first part, we do induction on  $r$ . For  $r = 1$  we just need  $n_1 \geq T$ . For  $r \geq 2$  note that  $\delta(G) \geq \left(1 - \frac{1}{r-1} + \frac{1}{r(r-1)}\right)n$ , just pick  $n_1(r, \epsilon, t) \geq n_1(r-1, \frac{1}{r(r-1)}, T)$ , and so  $G \supset K_r(T)$ . Now choose one of these  $K_r(T)$  and call it  $K$ .

Now for the second part, suppose  $G \not\supset K_{r+1}(t)$ . We estimate in two ways the number of pairs  $(P, Q)$  where  $P$  is a set of  $r$  vertices of  $K$ , one from each class, and  $Q$  is a set of  $t$  vertices of  $G - K$  where everything in  $P$  is adjacent to everything in  $Q$ .

There are  $\leq \binom{n}{t}$  choices for  $Q$ . Given some  $Q$ , since  $G \not\supset K_{r+1}(t)$  there is some class of  $K$  in which  $Q$  has  $\leq t-1$  common neighbours. So there are  $\leq (t-1)T^{r-1}$  sets  $P$  such that  $(P, \text{this } Q)$  is a pair. Hence there are  $\leq (t-1)T^{r-1}\binom{n}{t} = (t-1)T^{r-1}\frac{n^r}{t!} + O(n^{t-1})$ .

On the other hand, given a set  $P$ , each vertex of  $P$  has at most  $(1/r - \epsilon)n$  non-neighbours, so  $P$  has at least  $n - r(1/r - \epsilon)n = r\epsilon n$  common neighbours. Hence there are at least  $\binom{r\epsilon n - |K|}{t}$  pairs  $(\text{this } P, Q)$ . There are  $T^r$  sets  $P$ , so at least  $T^r \binom{r\epsilon n - |K|}{t} = T^r \frac{(t\epsilon)^t}{t!} n^t + O(n^{t-1})$  pairs  $(P, Q)$ . But this is a lower bound exceeding the upper bound for large  $n$ .  $\square$

**Theorem 3.7** (Erdős-Stone, 1946). *Let  $r, t \geq 1$  be integers and  $\epsilon > 0$ . Then if  $n$  is large enough and  $G$  is a graph with  $n$  vertices and  $\geq (1 - 1/r + \epsilon)n$  edges then  $G \supset K_{r+1}(t)$ .*

*Proof.* It’s enough to show that  $G$  contains a large subgraph  $H$  with  $\delta(H) \geq (1 - \frac{1}{r} + \frac{\epsilon}{2})|H|$ . Our argument will give such a  $H$  with  $|H| \geq h = \lfloor \frac{\sqrt{\epsilon n}}{2} \rfloor$ . Hence if we make sure that  $n_0(r, t, \epsilon) \geq \frac{2}{\sqrt{\epsilon}} n_1(r, t, \frac{\epsilon}{2})$  then applying **3.6** to  $H$  gives the required  $K_{r+1}(t)$ . We also choose  $n_0$  so that  $\binom{h+1}{2} > n$  which we can do as  $h = O(n)$ .

If there is no such  $H$  then there is a sequence of subgraphs  $G = G_n \supsetneq G_{n-1} \supsetneq \dots \supsetneq G_h$  where  $|G_j| = j$  and  $G_{j-1}$  is obtained from  $G_j$  by removing a vertex of degree  $\delta(G_j) \leq (1 - \frac{1}{r} + \frac{\epsilon}{2})j$  from  $G_j$ .

Then  $e(G_n) \geq (1 - \frac{1}{r} + \epsilon)\binom{n}{2} - \sum_{j=h+1}^n (1 - \frac{1}{r} + \frac{\epsilon}{2})j > \binom{h}{2}$ , but  $|G_h| = h$  so we have too many vertices.  $\square$

Given  $F$ , the **chromatic number**  $\chi(F)$  is the smallest number of colours such that the vertices can be coloured with no edge joining two vertices of the same colour. This is exactly the smallest  $r$  such that  $F$  is  $r$ -partite.

**Corollary 3.8.**

$$\lim_{n \rightarrow \infty} \frac{\text{ex}(n, F)}{\binom{n}{2}} = 1 - \frac{1}{\chi(F) - 1}$$

*Proof.* Let  $\chi(F) = r + 1$  so that  $r \geq 1$ . Since  $T_r(n)$  is  $r$ -partite it does not contain  $F$ . So  $\text{ex}(n, F) \geq t_r(n) = (1 - \frac{1}{r})\binom{n}{2}$ . Hence the limit is  $\geq 1 - \frac{1}{r}$ .

Now let  $\epsilon > 0$ . If  $|G| = n$  and  $e(G) \geq (1 - \frac{1}{r} + \epsilon)\binom{n}{2}$  then by E-S for large enough  $n$  we have  $G \supset K_{r+1}(|F|) \supset F$ . And so the limit must also be  $\leq 1 - \frac{1}{r} + \epsilon$ , and this holds for all  $\epsilon > 0$ .  $\square$

Note that if  $F$  is bipartite then this says only  $\text{ex}(n, F) = o(n^2)$  whereas for other  $F$  it gives  $\text{ex}(n, F)$  asymptotically). The natural bipartite analogue of Turán is to consider  $F = K_{t,t}$ . Even more natural now is to consider  $z(n, t) = \max\{e(G) : G \text{ } n \times n \text{ bipartite } G \not\supset K_{t,t}\}$  where  $n \times n$  bipartite means  $G$  has two vertex classes  $X, Y$  of order  $n$  and all edges of  $G$  lie between  $X$  and  $Y$ .