Graph Theory - Lecture notes.

D. Yogeshwaran
Indian Statistical Institute, Bangalore.

April 28, 2019

Contents

Index

Chapter 1

Disclaimers and Warnings

- These notes are highly incomplete, not well-written, not well 'latexed' and not meant to be taken seriously like a book.
- The notes might fail on many aesthetic counts and the author has put little effort to ensure that the notes are more pleasing to the eye.
- These are written for the introductory course on graph theory to second year undergraduate students. `http://www.isibang.ac.in/~adean/infsys/database/Bmath/GT.html`
- These are compendium of some of the class material (mostly definitions and result statements) to serve as a pointer to students. I hope to add proof details over the years.
- These are mostly a subset of the material covered in class.
- All figures are taken from various sources and sorry for not acknowledging them all.
- I am thankful to students of the course in the years 2018 and 2019 for pointing out errors, inaccuracies and also listening to my lectures :-)
- Unless mentioned, the proofs or results are all borrowed from some book or the other. Few of them are listed in the bibliography.
- Do not look for any new results or proofs or something else fancy here. Just stuff arranged according to my own teaching convenience.
- I have tried to mention some related research-level questions and open problems at the end of each chapter.
- The main references are [Bollobas 2013, Van Lint and Wilson, West 2001, Diestel 2000].

Chapter 2

Introduction to Graphs

2.1 Definition and some motivating examples

Some notation : $[n] = \{1, \ldots, n\}$, $V \times V$ the usual set product, $\binom{V}{2}$ denote unordered pairs of distinct elements in V.

Definition 2.1.1. *(Graph). A (simple) graph G consists of a finite or countable vertex set $V := V(G)$ and an edge set $E := E(G) \subset \binom{V}{2}$.*

We shall consider only locally-finite graphs i.e., graphs such that vertices occur only in finitely many edge-pairs.

For a vertex set V, we represent edges as $(v, w), v, w \in V$. We also write $v \sim w$ to denote that $(v, w) \in E$. A very common pictorial representation of graphs is as follows : Vertices are represented as points on plane and edges are lines / curves between the two vertices. See Figure 2.1[1]. As an exercise, explicitly define the graphs based on these representations.

See these two talks by Hugo Touchette (`http://www.physics.sun.ac.za/~htouchette/archive/talks/networks-short1.pdf`) and (`http://www.physics.sun.ac.za/~htouchette/archive/talks/complexnetworks1.pdf`) for many examples and applications of graphs in real-life problems. We shall briefly touch upon these in our examples as well.

The following two variants shall be hinted upon in our motivation and some examples but we shall not discuss the same for most of the course.

Remark 2.1.2 (Two variants)**.**

Directed graphs: *These are graphs with directed edges or equivalently the edge-pairs are ordered*

Multi-graphs: *These are graphs with multiple edges between vertices including self-loops.*

[1] All of the figures in these notes are not mine and taken from the internet

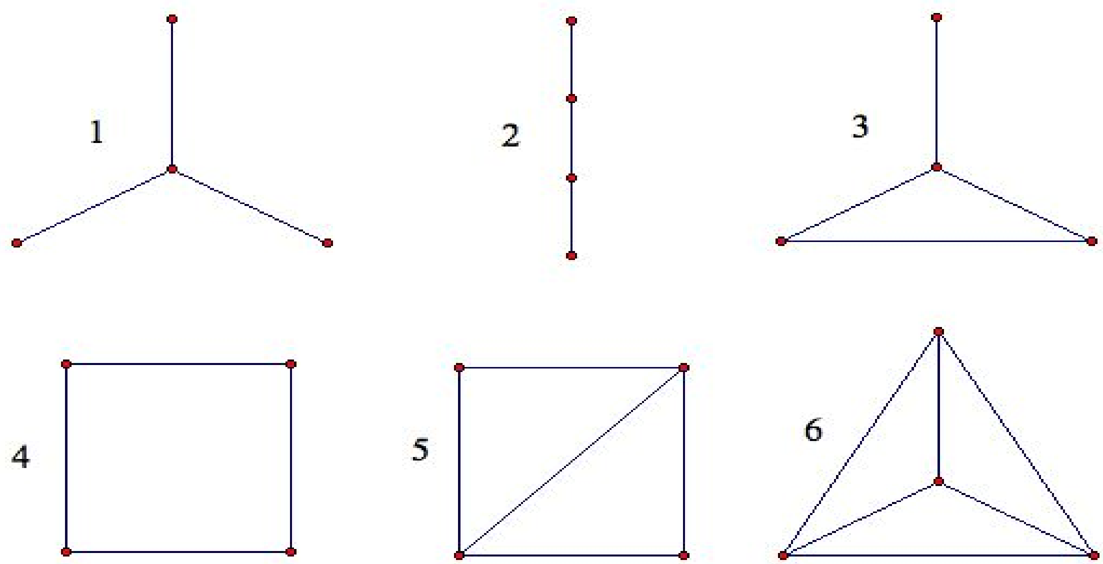


Figure 2.1: Some representations of graphs on vertex sets of cardinalities 3 and 4.

2.1.1 Popular examples of graphs

Let us see some popular examples of graphs.

Example 2.1.3. *(Facebook graph) V is the set of Facebook users and an edge is places between two vertices if they are friends of each other. See Figure 2.1.3*

Example 2.1.4. *(Road networks) V is the set of cities and an edge represents roads/trains/air-routes between the cities. See Figure 2.1.4*

Example 2.1.5. *Collaboration graph V is the set of all mathematicians who have published (say listed on mathscinet) and an edge represents that two mathematicians have collaborated on a paper together.*

Example 2.1.6. *(Complexity of Shakespeare's plays) V represents the characters in a shakespeare play and an edge between two characters means that both appeared in a scene together. See Figures 2.1.6 and 2.1.6 for the graph of Othello and Macbeth. See* `http://www.martingrandjean.ch/network-visualization-shakespeare/` *for more details. Network density is the ratio $|E|/|V|$ where $|.|$ denotes the cardinality of a set.*

More examples of networks abound in biology (`https://en.wikipedia.org/wiki/Biological_network`), chemical reaction networks (`https://en.wikipedia.org/wiki/Chemical_reaction_network_theory`) et al. We shall later give a more historical perspective with some more examples.

Definition 2.1.7. *(Weighted graphs) A graph G with a weight function $w : E \to \mathbb{R}$.*

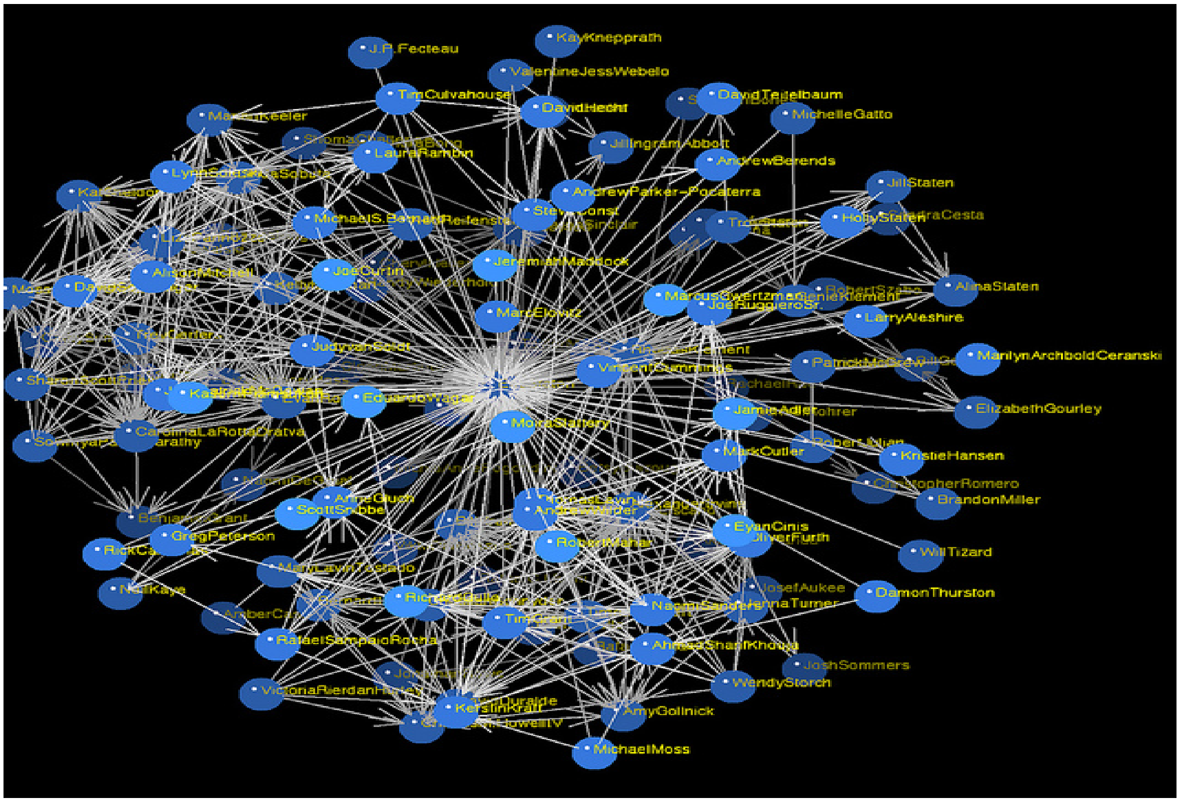


Figure 2.2: A snapshot of facebook graph

Often the weights are non-negative.

Example 2.1.8. *(Traffic Networks) G is the road network with the weight w denoting average traffic in a day. See Figure 2.1.8*

Example 2.1.9. *(Football graph) This is an example of weighted directed graph. Let the vertex set be the* 11 *players in a team and directed edge from i to j represents that player i has passed to player j. Associated with such a directed edge (i, j), a weight $w(i, j)$ that denotes the number of passes from player i to player j. See the football graph from the Spain vs Holland final in 2010 WC in Figure 2.1.9 This illustrates best the appealing visualization offered by graphs. Spain's passing game is very evident in the graph and gives a good way to analyse such effects in sports and other domains.*

2.2 Some history and more motivation

Example 2.2.1. *(Konigsberg Problem. Euler, 1736) The problem was to find a path starting at any point that traverses through all the bridges exactly once and return to the starting point (See figure 2.2.1). After many attempts in vain, Euler showed that this is not possible. This problem is considered the birth of both probability and topology. We shall see Euler's solution later.*

Example 2.2.2. *(Electrical Networks. Kirchoff, 1847) Electrical networks can be represented as weighted directed graphs with current and resistance viewed as weights. This formalism can explain*

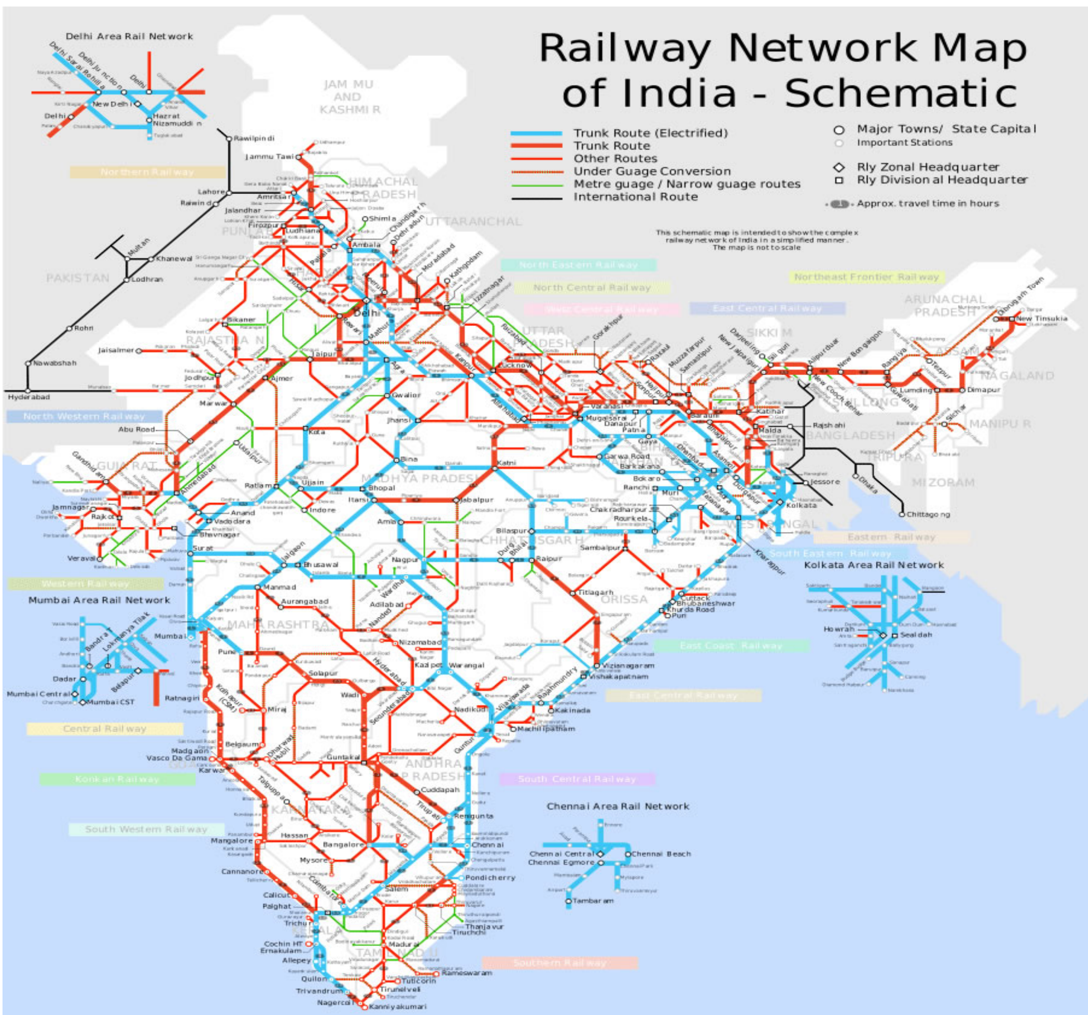


Figure 2.3: Indian railway network

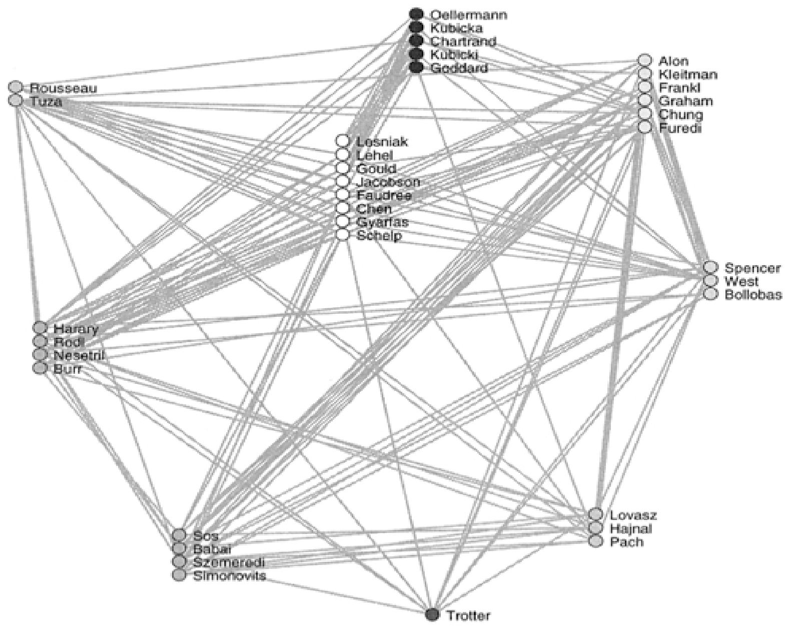


Figure 2.4: Collaboration graph of mathematicians based on mathscinet

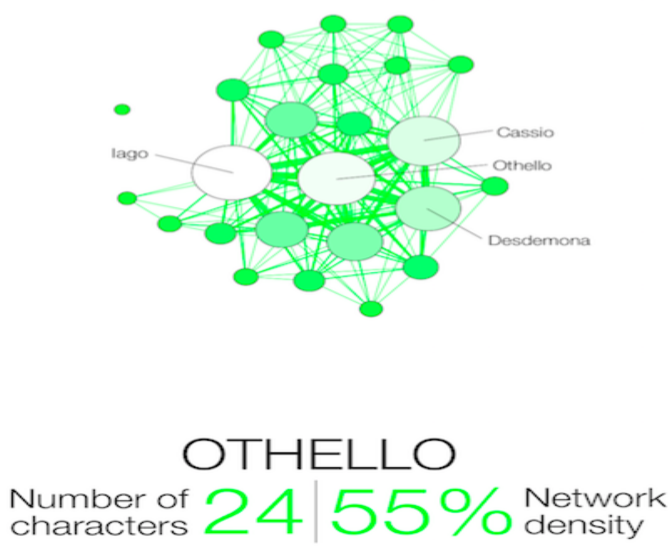


Figure 2.5: Othello characters graph

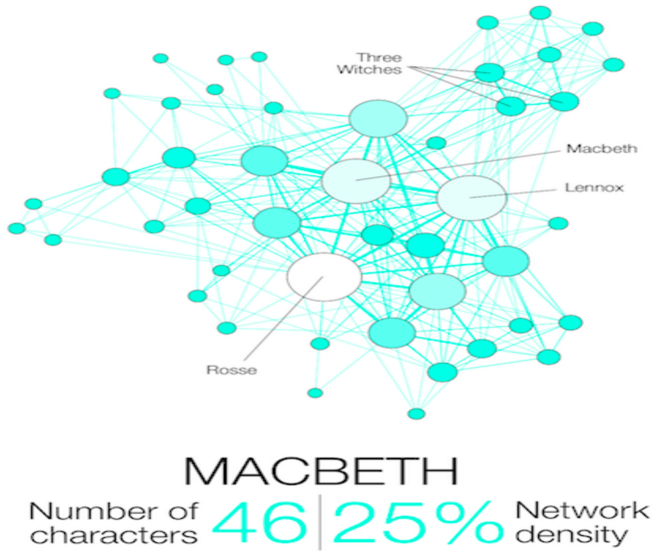


Figure 2.6: Macbeth characters graph

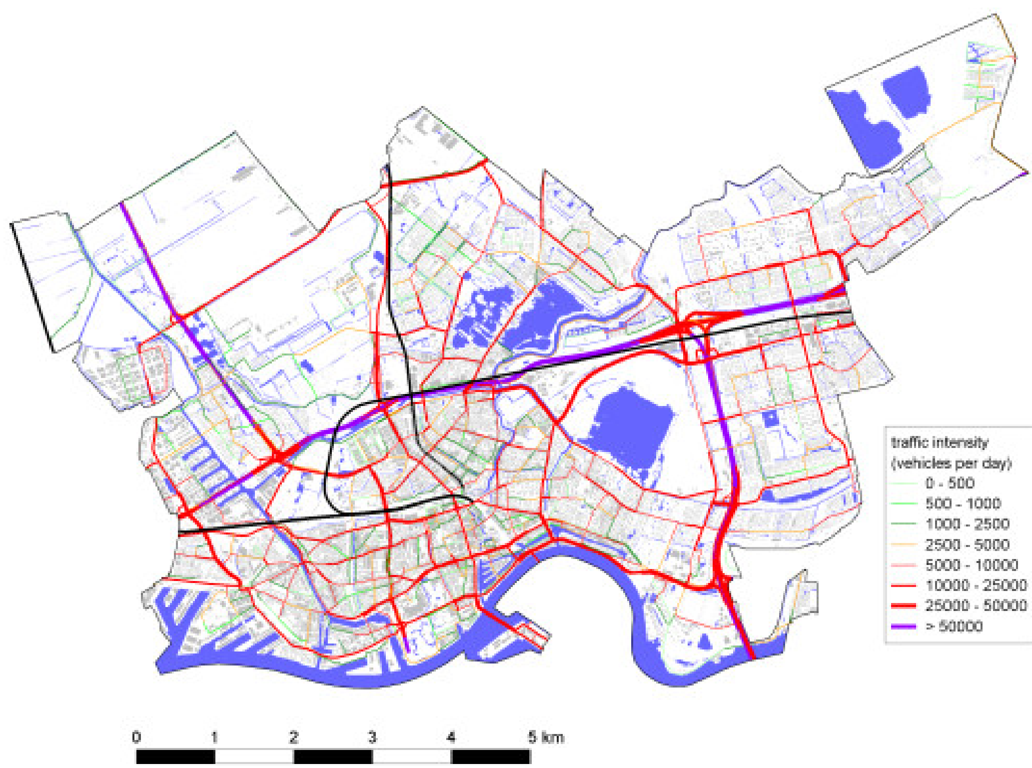


Figure 2.7: A road network with traffic density

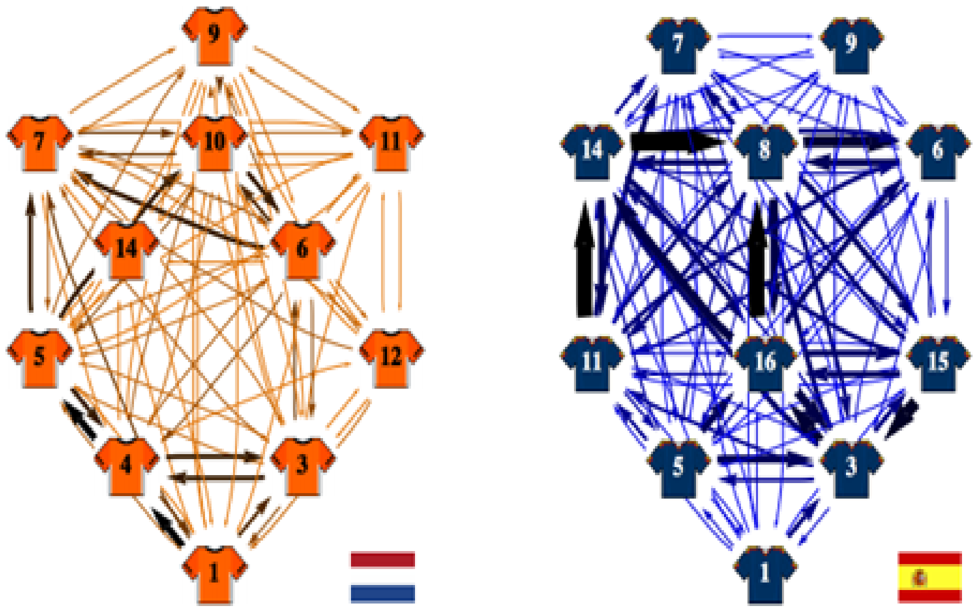


Figure 2.8: Football graph from the Spain vs Holland 2010 WC Final.

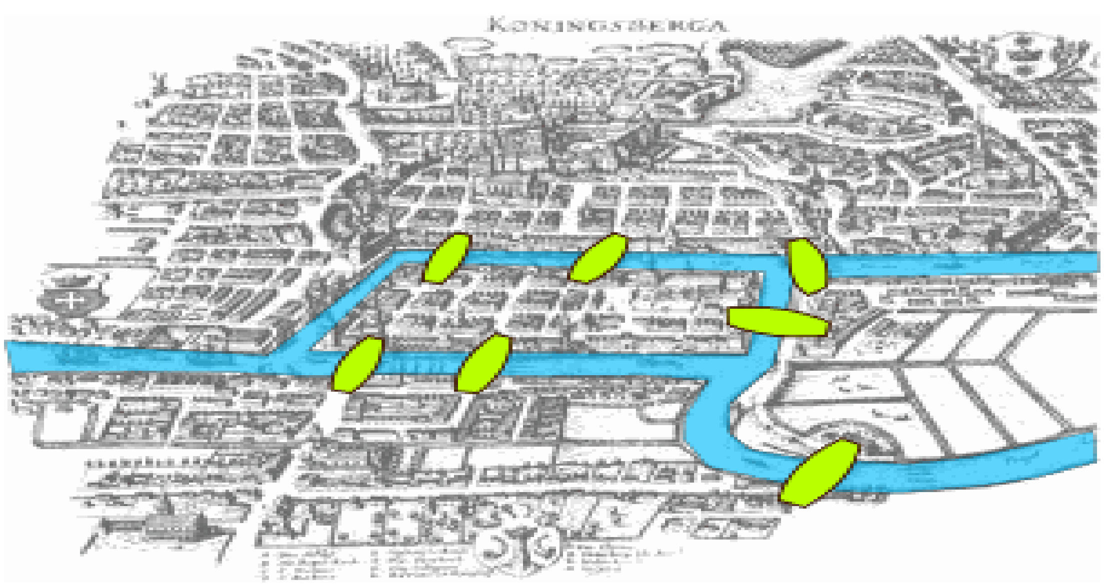


Figure 2.9: Seven bridges of Konigsberg on the river Pregel.

Kirchoff's and Ohm's laws. This connection between graphs and electrical networks is highly useful not only for graph theory and electrical networks but also used in random walks and algebraic graph theory. It is not entirely inaccurate to talk of Kirchoff's formalism also as discrete cohomology.

Example 2.2.3. *(Chemical Isomers. Cayley, 1857) Atoms were represented by vertices and bonds between atoms by an edge. Such a representation was used to understand the structure of molecules. We shall see Cayley's tree enumeration formula which was used in enumeration of number of chemical isomers of a compound. See Figure 2.2.3.*

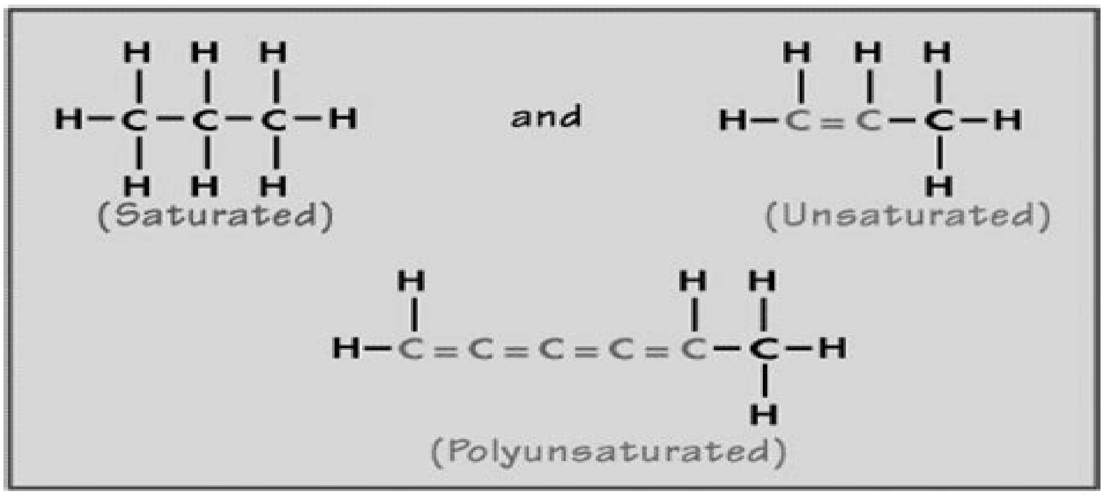


Figure 2.10: Hydrocarbons represented as graphs.

Example 2.2.4. *(Tour of cities. Hamilton, 1859) Given a set of cities and roads between them, find a path that starts at one city, visits all the cities exactly once and returns to the starting city. Can we*

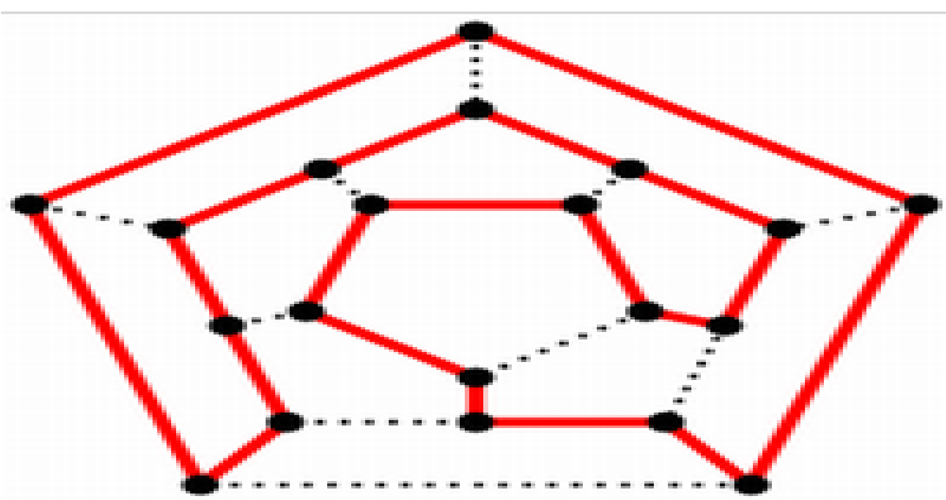


Figure 2.11: Hamilton cycle of a graph

guarantee such a path exists for all road networks ?

Example 2.2.5. *(Four colour theorem) Suppose we take a map of the world and assume that countries are contiguous land masses. Let countries be vertices and edges are drawn between two countries share a boundary. Can we colour countries such that neighbouring countries have different colours ? What is the minimum number of colours required for the same ?*

A more general question, can any graph be drawn as a map i.e., can any graph be drawn on the plane such that edges do not cross each other ?

2.3 Course overview :

Already, we have seen some historic questions that shall be discussed in the course. Now, we shall see something more specific.

2.3.1 Flows, Matchings and Games on Graphs

Example 2.3.1 (Maximum traffic flow)**.** *Consider the traffic network (weighted graphs) and take a starting point (source) and ending point (sink). What is the maximum amount of traffic that can flow from source to the sink in one instance ?*

The solution is very famously known as the max flow-min cut theorem and has many applications.

Example 2.3.2 (Hostel room allocation)**.** *There are n rooms and m students. Each student gives a list of rooms acceptable to them. Can the warden allot rooms to students such that each student gets at least one room in their list ?*

Hall's marriage theorem shall give a deceptively simple sufficient and necessary condition for the same. Hall's marriage theorem shall be proved using max flow-min cut theorem.

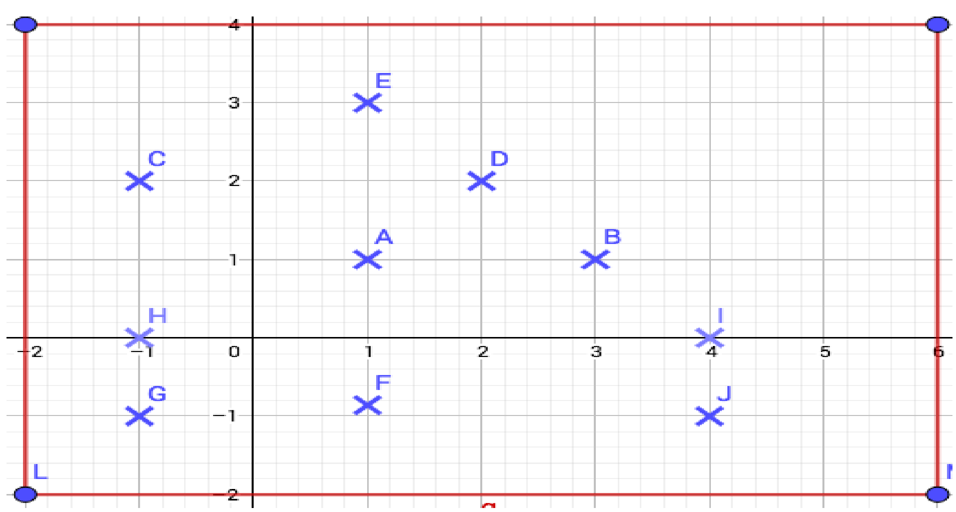


Figure 2.12: An example of a hide and seek scenario.

Example 2.3.3 (Hide and Seek game)**.** *Consider an area with horizontal main roads and vertical cross roads (See the grid in Figure 2.3.3). There are safe houses at certain intersections, marked by crosses in the figure. A robber choses to hide at one of the safe houses. A cop wants to find the main road or the cross road in which the robber stays. What is the best strategy for the cop to succeed ? What is the best strategy for robber to defeat the cop ? We shall exhibit a strategy for both using Hall's marriage theorem.*

2.3.2 Graphs and matrices

A graph G on V can be represented as a $V \times V$ matrix with 1 in the column (v, w) iff (v, w) is an edge in G. One can represent weighted and directed graphs as well. What does rank and nullity mean here ? Are there other matrices that encode the properties of graphs ? This viewpoint shall connect Kirchoff's formalism with cohomology theory.

2.3.3 Random graphs and probabilistic method.

If time permits, we shall mention random graphs. We fix $[n]$ as the vertex set and choose edges at random i.e., at each edge toss a coin and place the edge if it lands heads. What are the properties of these graphs ?

Probabilistic method : Suppose we want to show that there exists graphs with a certain property, let us show that the random graph satisfies the property with positive probability. Thus, the set of graphs satisfying the property is non-empty. This approach was pioneered by Erdös and is still remarkably successful in proving various results.

I shall try to emphasize various mathematical topics (such as cohomology) that show up in their simplest incarnation in graph theory and also, many mathematical tools such as linear algebra, analysis and probability are used to study graphs. A very readable survey on the growing importance of graph theory is [Lovasz 2011].

Chapter 3

The very basics

3.1 Some useful classes of graphs :

Example 3.1.1 (Complete graph)**.** *K_n : $V = [n], E = \binom{V}{2}$.*

Example 3.1.2 (Intersection graph)**.** *Let $S_1, \ldots, S_n$ be subsets of a set S. Define G with $V = [n]$ and $i \sim j$ if $S_i \cap S_j \neq \emptyset$.*

Example 3.1.3 (Delaunay graph)**.** *$P \subset \mathbb{R}^d, d \geq 1$ - finite distinct set of points. For $y \in \mathbb{R}^d, d(y, P) := \min_{x \in P} |x - y|$. Define $C_x := \{y : d(y, P) = |x - y|\}, x \in P$. Delaunay graph is the intersection graph on P with intersecting sets $C_x, x \in P$. C_x is called as the* voronoi cell *of x.*

Example 3.1.4 (Euclidean Lattices)**.** *Let $B_r(x)$ be the closed ball of radius r centered at x. The d-dimensional integer lattice is the intersection graph formed with $\mathbb{Z}^d$ as vertex set and $B_{1/2}(z), z \in \mathbb{Z}^d$ as the intersecting sets. Alternatively, $z \sim z'$ if $\sum_{i=1}^d |z_i - z_i'| = 1$.*

Example 3.1.5 (Cayley graphs)**.** *$H, +$ be a group with a finite set of generators S such that $S = -S$ (*symmetric*). The* Cayley graph *G is defined with vertex set $V = H$ and $x \sim y$ if $x - y \in S$. Since S is symmetric, $y - x \in S$ iff $x - y \in S$ and so abelianness of the group does not matter.*

Exercise 3.1.6. *Show that every graph is an intersection graph.*

Exercise 3.1.7. *Show that Euclidean lattices are Cayley graphs. Find the generators S.*

3.1.1 Some graph constructions

Example 3.1.8 (Bi-partite graph)**.** *Graphs (V, E) such that $V = V_1 \sqcup V_2$ and $E \subset V_1 \times V_2$.*

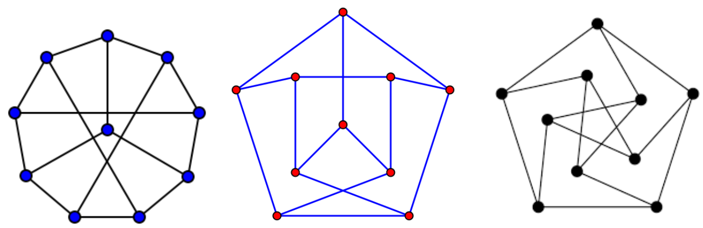
Example 3.1.9 (Complementary graph)**.** *Let $G = (V, E)$ be a graph. The complementary graph is $G^c = (V, \binom{[V]}{2} - E)$.*

Example 3.1.10 (Line graph)**.** *Let $G = (V, E)$ be a graph. The line graph is $L(G)$ with vertex set $V = E$ and $e_1 \sim e_2$ is they are adjacent in G.*

Example 3.1.11 (Petersen graph)**.** *The vertex set of the graph is* $\binom{[5]}{2}$ *and* $\{i,j\} \sim \{k,l\}$ *if* $\{i,j\} \cap \{k,l\} = \emptyset$.

Exercise 3.1.12. *Show that the Petersen graph is the complement of the line graph of* K_5*.*

Exercise 3.1.13. *Are the following three graphs isomorphic to Petersen graph ?*



Exercise 3.1.14. *Find the number of edges in a Line graph* $L(G)$ *in terms of the number of edges in* G *?*

3.2 Some basic notions

We shall assume that all our graphs are finite unless mentioned otherwise. In some examples, we shall illustrate things using infinite graphs but all our results are for finite graphs only.

Fix a graph G with vertex set V and edge set E. We say v, w are neighbours if $v \sim w$.

Definition 3.2.1 (Graph homomorphism and isomorphism)**.** *Suppose* G, H *are two graphs. A function* $\phi : V(G) \to V(H)$ *is said to be a* graph homomorphism *if* $x \sim y$ *implies* $\phi(x) \sim \phi(y)$. ϕ *is said to be an* isomorphism *if* ϕ *is a bijection and* $x \sim y$ *iff* $\phi(x) \sim \phi(y)$ *i.e.,* ϕ, ϕ^{-1} *are graph homomorphisms.* G *and* H *are* isomorphic *(*$G \cong H$*) if there exists an isomorphism between* G *and* H*. An* automorphism *is an isomorphism* $\phi : G \to G$*.*

Exercise 3.2.2. *The set of all automorphisms of* G *is called* $Aut(G)$*. Define a binary operation on* $Aut(G)$ *as followss : For* $g, f \in Aut(G)$, $g.f = g \circ f$ *i.e., the composition operation. Is* $Aut(G)$ *a group ?*

Essentially, G and H are the same graphs upto re-labelling. H is a *subgraph* of G if $V(H) \subset V(G)$ and $E(H) \subset E(G)$. H is an *induced subgraph* of G if H is a subgraph of G and if $v, w \in H$ such that $(v, w) \in E(G)$ then $(v, w) \in E(H)$.

Exercise* 3.2.3. H *is an induced subgraph of* G *iff* H *is the maximal subgraph in* G *with vertex set* $V(H)$*.*

Example 3.2.4 (Trivial homomorphisms)**.** *The identity automorphism is a trivial homomorphism ; If* $H \subset G$*, then the inclusion map from* $V(H)$ *to* $V(G)$ *gives rise to a homomorphism.*

Exercise 3.2.5. *Show that there exists a homomorphism from* G *to* K_2 *iff* G *is bi-partite.*

Exercise* 3.2.6. *Can one characterize classes of graphs G which have homomorphisms to K_k ?*

Exercise* 3.2.7. *Denote by $Hom^*(H,G)$ to be the number of injective homomorphisms from H to G. Let $|V(H)| = k$. Show that*

$$|Hom^*(H,G)| = \sum_{(v_1,\ldots,v_k)\in V(G)^k}^{\neq} \mathbf{1}[H \subset < \{v_1,\ldots,v_k\} >],$$

where $\sum^{\neq}$ denotes that the sum is over distinct elements and $< v_1,\ldots,v_k >$ is the induced subgraph on the vertices $v_1,\ldots,v_k$.

See here `http://www.cs.elte.hu/~lovasz/problems.pdf` for open problems about graph homomorphisms.

Definition 3.2.8 (Some notations)**.** *Let $\mathcal{G}$ be the set of all finite graphs.*

- Graph property : *A set $\mathcal{P} \subset \mathcal{G}$ is said to be a* graph property *if $G_1 \in \mathcal{P}$ and $G_1 \cong G_2$, then $G_2 \in \mathcal{P}$.*
- Graph invariant : *$\phi : \mathcal{G} \to \mathbb{R}$ is a* graph invariant *if $\phi(G_1) = \phi(G_2)$ whenever $G_1 \cong G_2$. Equivalently ϕ is a graph invariant if $\phi^{-1}(r)$ is a graph property for every $r \in \mathbb{R}$.*
- Spanning subgraph : *H is a* spanning subgraph *if $V(H) = V(G)$.*
- Neighbourhood : *If $v \in V$, the neighbourhood of v, $N_v := \{w : w \sim v\}$.*
- Degree : *$d_v := |N_v|$.*
- *v is isolated if $d_v = 0$.*
- Regular graph : *G is d-regular if $d_v = d_w = d$ for all v, w.*
- Minimum degree : *$\delta(G) := \min\{d_v : v \in V\}$.*
- Maximum degree : *$\Delta(G) := \max\{d_v : v \in V\}$.*
- Average degree : *$d(G) := |V|^{-1}\sum_v d_v$.*
- Edge density : *$\epsilon(G) := |E|/|V|$. Show the second equality.*

Prove that $\delta(G) \le d(G) \le \Delta(G)$.

Exercise* 3.2.9. *Which of the graphs in Section 3.1 are regular and what are their average degrees ? Can you compute $Aut(G)$ for these examples ?*

Exercise* 3.2.10. *What can you say about the Minimum degree, maximum degree and average degree of Complementary and Line graphs given those of the original graph ?*

Lemma 3.2.11. *$\sum_v d_v$ is even. Thus the number of odd-degree vertices is even and $d(G) = 2\epsilon(G)$.*

Exercise 3.2.12. *Suppose G is a 3-regular graph on 10 vertices such that any two non-adjacent vertices have exactly one common neighbour. Is G isomorphic to Petersen graph ?*

Definition 3.2.13. *(Path, Walk and Cycle) An ordered set of vertices $P = v_0 \dots v_k$ is said to be* a walk from v_0 to v_k *if $v_i \sim v_{i+1}$. A walk $P = v_0 \dots v_k$ is said to be* a path from v_0 to v_k *if (v_i, v_{i+1}) are distinct for all i. A walk is* closed *if $v_k = v_0$. A path is* simple *(also called as* self-avoiding walk*) if $v_i \neq v_j$ for all $i \neq j$. A walk or path is said to be* closed *if $v_k = v_0$ and else* open. *A closed path is also called as a* circuit. *A circuit $C = v_0 \dots v_{k-1} v_0$ with no repetition of intermediate vertices is called a* cycle *i.e., $v_0, \dots, v_{k-1}$ are distinct.*

For $v \neq w$, we say that v is connected to w (denoted by $v \to w$) if there exists a path from v to w. We shall always assume that $v \to v$. Show that $\to$ induces an equivalence relation. Define the component of v as $C_v := \{w : v \to w\}$. If $v \to w$, then $C_v = C_w$.

Exercise 3.2.14. *For $v \neq w$, show that there exists a walk from v to w iff there exists a path from v to w iff there exists a self-avoiding walk from v to w.*

Exercise* 3.2.15. *Show that $\to$ partitions V into equivalence classes and the equivalence class of v is C_v. Show that C_v is the maximal connected subgraph containing v.*

The equivalence classes induced by $\to$ are called as *connected components* and the number of connected components are denoted by $\beta_0(G)$.

Exercise 3.2.16. *Show that $\beta_0(G) = 1$ iff for all $v \neq w \in G$, $v \to w$.*

We call the graph to be *connected* if $\beta_0(G) = 1$.

Exercise 3.2.17. *Show that $\delta(G), \Delta(G), d(G), \beta_0(G)$ are all graph invariants.*

Exercise 3.2.18. *Let G be a simple graph with at least two vertices. Show that G must contain at least two vertices with the same degree.*

3.3 Graphs as metric spaces

Let G be a connected weighted graph with strictly positive edge-weights i.e., $w(e) > 0$ for all $e \in E$. or un-weighted graphs, set $w \equiv 1$. We now shall view graphs as metric spaces. The weight/length of a path or walk $P = v_0 \dots v_k$ is $w(P) := \sum_{i=0}^{k-1} w((v_i, v_{i+1}))$. Define the distance between two vertices $v \neq w$ as follows :

$$d_G(v, w) := \inf\{w(P) : P \text{ is a path from } v \text{ to } w\}.$$

Set $d_G(v, v) = 0$ for all $v \in V$. Show that $d_G(v, w) = d_G(w, v)$ and further $d_G(v, w) = 0$ implies that $v = w$.

Exercise* 3.3.1 (Graph metric). *Show that (V, d_G) is a metric space.*

Even if G is not connected, we have that d_G satisfies the three axioms of a metric space. Further, define the *diameter of a graph* as

$$diam(G) = \max\{d_G(v, w) : v, w \in G\}.$$

Given a vertex v and $r > 0$, set $B_r(v) := \{w : d_G(v, w) \leq r\}$, the ball of radius r at v. For un-weighted graphs show that $|B_n(v)| \leq \sum_{i=0}^{n} \Delta(G)^i$.

Exercise* 3.3.2. *Let G be the Cayley graph of a free group generated by a finite symmetric set of generators $S = \{s_1, -s_1, \dots, s_n, -s_n\}$. Compute $|B_n(e)|$ for all n.*

Exercise 3.3.3. *Can you compute $|B_n(O)|$ for $\mathbb{L}_d$ for $d \geq 2$?*

For (un-weighted) graphs, let $\Pi_n(v)$ be the set of self-avoiding walks of length exactly n from v. it holds that $|\Pi_n(v)| \leq \Delta(G)(\Delta(G)-1)^n$ where $n \in \mathbb{N}$. Thus for $\mathbb{L}^d$, we have that $|\Pi_n(O))| \leq 2d(2d-1)^n$.

Exercise* 3.3.4. *Is it true that there exists a κ_d such that as $n \to \infty$, we have that $|\Pi_n(O)|^{1/n} \to \kappa_d \in [0, \infty)$?* Hint : Use that $\log |\Pi_n(O)|$ is sub-additive.

3.4 Graphs and matrices : A little peek

We shall now give a brief hint about utility of matrices and linear algebra in answering graph-theoretic questions.

Definition 3.4.1 (Adjacency matrix.). *Let G be a graph on n vertices. For simplicity, assume $V = [n]$. The adjacency matrix $A := A(G) := (A(i,j))_{1 \leq i,j \leq n}$ of a simple finite graph is defined as follows : $A(i,j) = \mathbf{1}[i \sim j]$. The definition can be appropriately extended for multi-graphs.*

A is a symmetric matrix and hence has real eigenvalues.

Lemma 3.4.2. *Let G be a graph on n vertices and A be its adjaceny matrix. Show that $A^l(i,j)$ is the number of walks of length l from i to j.*

Proof. By definition, $A^l(i,j) = \sum_{i_1,\dots,i_{l-1}} A(i,i_1)A(i_1,i_2)\dots A(i_{l-1},j)$ and since A is a $0-1$ valued matrix, we get that $A(i,i_1)A(i_1,i_2)\dots A(i_{l-1},j) \in \{0,1\}$. The proof is complete by noting that $A(i,i_1)A(i_1,i_2)\dots A(i_{l-1},j) = 1$ if $i \sim i_1 \sim i_2 \dots i_{l-1} \sim j$ i.e., $ii_1 \dots i_{l-1}j$ is a walk of length l. $\square$

Exercise* 3.4.3. *Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A. Show that the number of closed walks of length l in G is $\sum_{i=1}^{n} \lambda_i^l$. Here, we count the walk $v_1, \dots, v_{l-1}, v_1$ and $v_2, \dots, v_{l-1}, v_1, v_2$ as distinct walks.*

Exercise* 3.4.4. *Count the number of closed walks of length l in the complete graph K_n.*

Lemma 3.4.5. *Let G be a connected graph on vertex set $[n]$. If $d(i,j) = m$, then $I, A, \ldots, A^m$ are linearly independent.*

Proof. Assume $i \neq j$. Since there is no path from i to j of length less than m, $A^k(i,j) = 0$ for all $k < m$ and $A^m(i,j) > 0$. Thus, if $I, A, \ldots, A^m$ are linearly dependent with coefficients $c_0, \ldots, c_m$, then by the above observation and positivity of entries of A^k for all k, we have that $c_m = 0$ i.e.,$I, A, \ldots, A^{m-1}$ are linearly dependent. Since $d(i,j) = m$, there exists j' such that $d(i,j') = m-1$. Now applying the above argument recursively, we get that $c_0 = c_1 = \ldots = c_m = 0$. □

Corollary 3.4.6. *Let G be a connected graph with k distinct eigenvalues. Then $k > diam(G)$.*

Proof. Let $d = diam(G)$. Recall that *the minimal polynomial* of a matrix A is the monic polynomial Q of least degree such that $Q(A) = 0$. By Lemma 3.4.5, we have that $deg(Q) > d$. The proof is complete by observing that the number of distinct eigenvalues of A is at least $deg(Q)$. □

3.5 Euler's theorem and König's theorem on bi-partite graphs

Definition 3.5.1 (Eulerian graph)**.** *A circuit in a graph visiting every edge exactly once and every vertex is called* an Eulerian circuit *and a graph that has an Eulerian circuit is called* an Eulerian graph.

Exercise* 3.5.2. *Show that a finite graph is Eulerian iff G is connected and is a edge-disjoint union of cycles i.e., $G = C_1 \cup \ldots C_m$ where C_i's are cycles and have no common edges.*

Theorem 3.5.3 (Euler's theorem ; Veblen 1912.)**.** *A finite connected graph is Eulerian iff every vertex has even degree.*

Proof. Since the graph is Eulerian, let $C_1, \ldots, C_m$ be the partition into edge-disjoint cycles. Viewing C_i's as graphs by themselves, observe that $d_v(C_i) = 21[v \in C_i]$ i.e., vertices in C_i have degree 2 and 0 otherwise. Further, we have by edge-disjointness of C_i's that

$$d_v(G) = \sum_i d_v(C_i) = 2\sum_i 1[v \in C_i],$$

and thus proving that every vertex has even degree.

To show the converse, assume that the vertex degrees are all even. Let $x_0x_1\ldots x_l$ be a simple path of maximal length l i.e., there is no simple path of length $> l$. Since $d_{x_0} \geq 2$, there exists a $y \in N(x_0) \setminus \{x_1\}$. Then $yx_0x_1\ldots x_l$ is a simple path of length $l+1$ and this yields a contradiction unless $y = x_i$ for some $1 \leq i \leq l$. Let $x_i = y$. Then $x_0x_1\ldots x_ix_0$ is a cycle, say C_1. Remove C_1 from G and consider $G - C_1$. All vertex degrees in $G - C_1$ are even. So, we can repeat the procedure for a component of $G - C_1$ with at least two vertices and obtain a cycle C_2. Continuing this way, we obtain cycles $C_3, \ldots C_m$ such that G is a disjoint union of $C_1, \ldots, C_m$ and hence G is Eulerian by the claim above. □ □

Exercise* 3.5.4 (Konigsberg problem). *Prove Euler's theorem for multi-graphs and hence show that Konigsberg problem has no solution.*

Exercise* 3.5.5. *Every closed odd length walk contains an odd length cycle.*

Theorem 3.5.6 (König's theorem). *A graph is bi-partite iff it has no odd cycles.*

The König here is the hungarian mathematician Dénes König whose father Gyola König was also a well-known mathematician. Dénes wrote the first book on graph theory.

Proof. **The only if part :** Suppose G is bi-partite with partition $V = V_1 \sqcup V_2$. Let $v_0 v_1 \ldots, v_k v_0$ be a cycle. WLOG[1], suppose $v_o \in V_1$. By bi-partiteness, $v_i \in V_2$ for $1 \leq i \leq k$ and i odd and $v_i \in V_1$ for $1 \leq i \leq k$ and i even. Since $v_k \sim v_0$, $v_k \in V_2$ and hence k is odd. Thus, the length of the cycle $k+1$ is even.

The if part : Let $v_0 \in V$. Set $v \in V_1$ if $d_G(v_0, v)$ is even and else $v \in V_2$. WLOG, suppose $v \sim w$ for $v, w \in V_1$. Then Let P, P' be respectively the shortest paths from v_0 to v, w respectively. Let $P*$ be the inversed parth from v to v_0. Since $v, w \in V_1$, $P*, P'$ have even lengths and so $P'wvP*$ is a closed walk starting at v and has odd length. By Lemma(3.5.5), it has an odd length cycle, a contradiction. Thus there are no edges between vertices in V_1 or V_2 respectively i.e., G is bi-partite with partition $V_1 \sqcup V_2$. □

3.6 ***Some questions***

Question 3.6.1. *Instead of counting $|B_n(0)|$, consider the following counting. Let $f_d(r) = |\{(n_1, \ldots, n_d) \in \mathbb{Z}^d : \sum_{i=1}^d n_i^2 \leq r^2 :\}|$ be the number of lattice points within r radius ball in $\mathbb{R}^d$. Show that $f_d(r)/r^d \to$ a constant. To determine exact asymptotics of $f_2(r)$ is known as the Gauss circle problem.*

Here are two questions from geometric group theory. For more on this fascinating subject, refer to [Clay and Margalit 2017].

Question 3.6.2. *Suppose G is the Cayley graph of a finitely generated countable group such that $n^{-d}|B_n(e)| \to \infty$ for all $d \in (0, \infty)$. Is it true that $n^{-a} \log |B_n(e)| \to \infty$ for some $a > 0$?*

Question 3.6.3. *Can you construct a finitely generated group such that its Cayley graph has the following growth property for some $a < 0.7$?*

$$0 < \liminf_{n \to \infty} n^{-a} \log |B_n(e)| \leq \limsup_{n \to \infty} n^{-a} \log |B_n(e)| < \infty.$$

Now some questions on self-avoiding walks.

Question 3.6.4. *Can you calculate κ_d for any $d \geq 2$?*

[1] Without loss of generality

Note : If you have solutions for any of the above four questions, please consider submitting them here. It was recently shown by Duminil-Copin and Smirnov that $\kappa = \sqrt{2+\sqrt{2}}$ for the hexagonal lattice (see Figure 3.6). In listing problems in IMO which can lead to research problems, Stanislav Smirnov mentions this ([Smirnov 2011]).

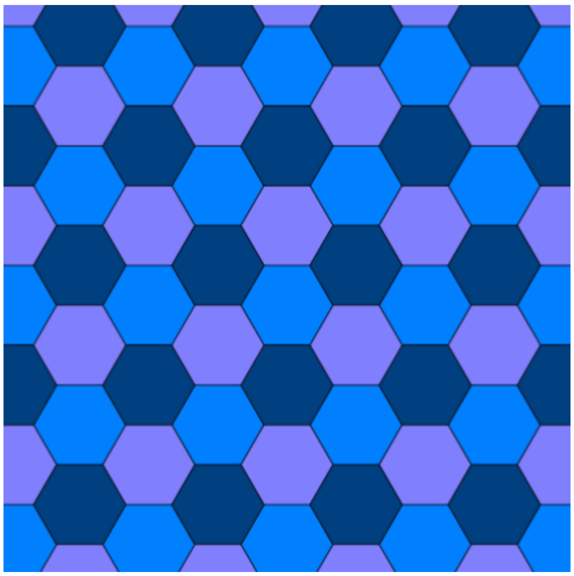


Figure 3.1: Hexagonal lattice

Chapter 4

Spanning Trees

Spanning trees (and minimal spanning trees) are a central object in combinatorial optimization, graph theory and probability. See the end of the chapter for some probabilistic connections.

4.1 Trees and Cayley's theorem

Definition 4.1.1 (Trees). *A graph with no cycles is called a* forest*. A connected forest is called* a tree.

Exercise* 4.1.2. *Show that TFAE*[1] *for a graph G on n vertices.*

1. *G is a forest.*

2. *G has $n - \beta_0(G)$ edges.*

Exercise 4.1.3. *Prove that there exists two vertices of degree* 1 *in a tree.*

Theorem 4.1.4 (Cayley's formula). *The number of labelled trees on n vertices (i.e., spanning trees on K_n) is n^{n-2}*

Proof via Prufer code (H. Prufer (1918)) . We shall construct a bijection

$$P : \mathcal{S}(K_n) := \{\text{spanning trees of } K_n\} \to [n]^{n-2}.$$

Given a tree T on $[n]$, we generate a sequence of trees $T_1, \ldots, T_{n-1}$ inductively as follows : Set $T_1 = T$. Given tree T_i on $n - i + 1$ vertices, let x_i be the least labelled vertex of degree 1 and delete the edge incident on x_i to obtain T_{i+1}. Denote by y_i, the neighbour of x_i. Observe that T_{n-1} is K_2 and the process terminates at T_n when the tree has only one vertex. Now define P as

$$P(T) = (y_1, \ldots, y_{n-2}).$$

[1] the following are equivalent

Clearly P is a map from $\mathcal{S}(K_n)$ to $[n]^{n-2}$. We shall prove that it is a bijection by constructing P^{-1} using induction. For $n = 2$, P is trivially a bijection.

Some observations : (i) $(y_2, \ldots, y_{n-2})$ is the Prufer code of T_2. (ii) The degree $d_i := \sum_{j=1}^{n-2} \mathbf{1}[y_j = i] + 1$ (Try to prove this yourself by induction on n and then see the proof below). (iii) As a consequence we have that $x_k := \min\{i : i \notin \{x_1, \ldots, x_{k-1}, y_k, \ldots, y_{n-2}\}\}$ (again prove by induction on k using (ii)).

Suppose P is a bijection for $n-1$ and let $a = (a_1, \ldots, a_{n-2}) \in [n]^{n-2}$. Now, let $x = \min\{i : i \notin (a_1, \ldots, a_{n-2})\}$. Consider $a' = (a_2, \ldots, a_{n-2})$ and by induction assumption a' is the unique Prufer code of a tree T' on $[n] \setminus \{x\}$ since P is a bijection on $[n] \setminus \{x\}$ by induction assumption. Define $T := T' \cup \{x, a_1\}$. Clearly, T is a tree as $x \notin V(T')$ and $P(T) = a$. If T'' is a tree such that $P(T'') = a$, then by the property (ii) of Prufer code, $[n] \setminus \{a_1, \ldots, a_{n-2}\}$ is nothing but vertices of degree 1. By definition of x, it has the least label among such vertices and thus by construction $(x_1, y_1) = (x, a_1)$. Hence, $P(T_2'') = a'$. By uniqueness of T', we have that $T_2'' = T'$ and hence $T'' = T$. So P^{-1} is well-defined on $[n]^{n-2}$ and P is $1-1$ and onto. Thus P is a bijection.

Trivially (ii) holds for $n = 2$. Assume that it holds for all trees on $n-1$ vertices. Let T be a tree on $[n]$. By (i) and induction hypothesis, for $i \neq x_1$, we have that $d_i(T_2) = \sum_{j=2}^{n-2} \mathbf{1}[y_j = i] + 1$. Since $d_i(T_2) = d_i(T)$ for all $i \neq x_1, y_1$ and $d_{x_1}(T) = 1$, $d_{y_1}(T) = d_{y_1}(T_2) + 1$, (ii) holds for T as well. □

Theorem 4.1.5 (Tree counting theorem)**.** *The number of labelled spanning trees on n vertices with degree sequence $d_1, \ldots, d_n$ is $\binom{n-2}{d_1-1, \ldots d_n-1} = \frac{(n-2)!}{\prod_{i=1}^{n}(d_i-1)!}$ for $n \geq 3$.*

Show that Cayley's theorem follows as a corollary of the tree counting theorem.

Proof. Denote by $t(n; d_1, \ldots, d_n)$ the number of trees with degree sequence $d_1, \ldots, d_n$. The theorem holds for $n = 3$. Assume that it is true for $n-1$. Assume that $d_j = 1$.

Suppose j is joined to i in a tree T, then $T - (i, j)$ is a tree on $[n] - j$ with degree sequence $d_1, \ldots, d_i - 1, \ldots, d_n$. Thus we get that

$$t(n; d_1, \ldots, d_n) = \sum_{i=1, i \neq j}^{n} t(n-1; d_1, \ldots, d_i - 1, \hat{d_j}, \ldots, d_n),$$

where $\hat{d_j}$ means d_j is absent. Note that it is not possible that $d_i = 1$ and for $d_i = 1$, $t(n-1; d_1, \ldots, d_i - 1, \ldots, d_n) = 0$. Now check inductively that $t(n; d_1, \ldots, d_n) = \binom{n-2}{d_1-1, \ldots d_n-1}$. □

Exercise* 4.1.6. *Give an alternative proof of tree counting theorem using Prufer codes.*

Remark 4.1.7 (More proofs of tree counting theorem)**.** *As with any such a fundamental result, there are multiple proofs.*

1. *We shall see a more powerful tree counting argument (known as Kirchoff's matrix tree theorem) using matrix theory later.*

2. *There is another proof via bijection due to Joyal and a recent double counting argument due to Pitman.*

All these five proofs can be found in [Aigner et al. 2010, Chapter 30].

Exercise* 4.1.8. *Show that the following are equivalent. Let G be a graph with k components.*

1. *$T \subset G$ is a spanning forest [2] i.e., a forest such that every component of the forest is a spanning tree of the corresponding compoent.*

2. *T is a 'minimal' spanning subgraph with k components. Here minimality of T means that $T - e$ has $k+1$ components for any $e \in T$.*

3. *T is a maximal subgraph without cycles. Here maximality of T means that there is no $T' \supsetneq T$ such that T' has no cycles.*

Exercise 4.1.9. *Let $H \subset G$ be a subgraph and $e \in G \setminus H$. Then exactly one of the following holds : (1) $\beta_0(H \cup e) = \beta_0(H) - 1$ or (2) $\beta_0(H \cup e) = \beta_0(H)$ and there exists a cycle $C \subset H \cup e$ such that $e \in C$ and $C \setminus e$ is not a cycle in H.*

Definition 4.1.10 (Cut edges and vertices)**.** *An edge e is a cut-edge in a graph G if $\beta_0(G - e) = \beta_0(G) + 1$. A vertex is a cut-vertex in a graph G if $\beta_0(G - v) > \beta_0(G)$.*

Exercise 4.1.11. *An edge is a cut-edge iff it belongs to no cycle.*

Lemma 4.1.12. *TFAE*

1. *G is a forest*

2. *There exists a unique simple path between u to v for all $u \neq v$ which are in the same component.*

3. *Every edge is a cut-edge.*

Proof. Since forest has no cycles, (i) $\Rightarrow$ (iii) follows from the above exercise. Since every edge is a cut-edge, this implies that there are no cycles and hence G is a forest i.e., (iii) $\Rightarrow$ (i). If there exists two distinct simple paths from u to v for $u \neq v$ in the same component, then the union of the paths contains a cycle (Prove this as an exercise !). This contradicts (i) and (iii). Thus, (i) and (iii) both imply (ii). Assuming (ii), it is easy to see that there is no cycle in G and hence both (i) and (ii) hold. $\square$

Lemma 4.1.13 (Insertion property of Spanning Trees)**.** *Suppose T is a spanning tree of G and there exists $e \in G - T$. Then $\exists e' \in T$ such that $T + e - e'$ is also a spanning tree.*

[2] This terminology is a slight abuse of our usual usage of the term 'spanning'.

Proof. For any $e' \in T$, $T+e-e'$ has same number of edges as T. Hence it suffices to show that there exists e' such that $T+e-e'$ does not have a cycle. This can be seen easily as follows : Since T is a spanning tree, $T+e$ contains a cycle. Now remove any edge e' in the cycle and we can see that $T+e-e'$ does not contain a cycle. □

Exercise* 4.1.14 (Deletion property of Spanning Trees)**.** *Suppose T, T' are spanning trees of G and $e \in T - T'$. Then $\exists e' \in T'$ such that $T - e + e'$ is also a spanning tree.*

4.2 Minimal spanning trees

Definition 4.2.1 (Minimal Spanning Tree)**.** *Consider a weighted graph G, w. Given a subgraph $H \subset G$, we define $w(H) := \sum_{e \in H} w(e)$. MST is said to be a* minimal spanning tree *if $w(MST) = \min\{w(T) : T \text{ is a spanning tree}\}$.*

Of course, minimal spanning tree exists if G is a connected graph. A set of edges $S \subset E$ is said to be a *cut*[3] if $\beta_0(G-S) = \beta_0(G)+1$ and for any $S' \subsetneq S$, $\beta_0(G-S') = \beta_0(G)$.

Proposition 4.2.2 (Some properties of MST)**.** *Let G be a connected graph with edge-weights.*

1. Uniqueness : *If $w : E \to \mathbb{R}$ is an injective function, then MST is unique.*

2. Cut property : *If M is a MST and C is a cut in G, then one of the minimal weight edges in C should be in the M.*

3. Cycle property : *If M is a MST and C is a cycle in G, then one of the maximal weight edges in C will not be in M.*

Proof. (i) : Let T_1, T_2 be MSTs such that $T_1 \neq T_2$. Since T_1, T_2 have the same vertex set, there exists an edge in $T_1 \Delta T_2$. Choose the edge e_1 with the least weight and WLOG let $e_1 \in T_1$. Since T_2 is a spanning tree, $T_2 + e_1$ has a cycle C. Since $C \subsetneq T_1$, there exists $e_2 \in C - T_1$ and also $e_2 \in T_1 \Delta T_2$. Thus $w(e_2) > w(e_1)$ as e_1 has the least weight in $T_1 \Delta T_2$. As in the proof of insertion property for spanning trees (Lemma 4.1.13), we can show that $T_2 + e_1 - e_2$ is a tree. Thus, we have that $w(T_2 + e_1 - e_2) < w(T_2)$ and hence contradicting the minimality of T_2.

(ii) : Let $C = \{e_1, \ldots, e_k\}$ in non-decreasing order of weights. If $e_1 := (u,v) \notin M$, then $M \cup e_1$ has a cycle. Since there exists a path in M from u to v and C is a cut, this path must pass through e_i for some $i > 1$. This gives that $M - e_i + e_1$ is a spanning tree(since it has no cycles and has $n-1$ edges). But since M is a MST, $w(e_i) \leq w(e_1)$. By choice of e_1, $w(e_1) \leq w(e_i)$ and so $w(e_i) = w(e_1)$. Thus e_i is a minimal weight edge and $e_i \in M$ as required.

(iii) : One can argue as above for this case too. □

Exercise* 4.2.3. *Prove (iii) in the above proposition.*

[3]Sometimes this is called a minimal cut and a cut is $S \subset E$ if $\beta_0(G-S) > \beta_0(G)$.

4.3 Kruskal and other algorithms

Kruskal's algorithm : Input graph weighted connected Graph $G = (V, E, w)$.

Step 1: Initialize with $D = \emptyset \subset E$ and $M = (V(G), \emptyset)$.

Step 2: Select one of the smallest (in terms of weight) edges in $E - D$. Call it e.

Step 3: If $M \cup e$ does not create a cycle, set $M = M \cup e$.

Step 4: Set $D = D \cup e$. If $|E(M)| < n - 1$ and $D \neq E$, go to Step 2 else to Step 5.

Step 5: Output M.

Theorem 4.3.1. *The output of the Kruskal's algorithm is a minimal spanning tree if G is connected.*

Proof. The proof consists of two steps - Firtly, we will show that the output M is a spanning tree and then show it is of minimum weight.

Step 1 : M is a spanning tree. Clearly M does not contain a cycle. If M has at least two components C_1, C_2, then let e an edge with minimal weight between C_1 and C_2. e exists as G is connected. e would have been added in Step 2 when it was considered. Thus, we derive a contradiction to the fact that M is disconnected. Hence M is connected and acyclic, i.e., a spanning tree.

Step 2 : M has minimum weight. We shall show that the following claim holds by induction : At every stage of the algorithm, there is a minimal spanning tree T such that $M \subset T$.

The above claim suffices because at the last step, the output M is the M at the end of Step 3 and if it is a subgraph of a spanning tree T, then it must be that $M = T$ as M is a spanning tree from Step 1.

Induction argument : The claim holds in the beginning because $M = \emptyset$. Assume that the claim holds at some stage of the algorithm for a M and a T.

Now, suppose that the next chosen edge creates a cycle, then M does not change and the claim still holds.

So, let the next chosen edge not create a cycle and hence M becomes $M + e$. If $e \in T$, we are done. Suppose $e \notin T$. Then $T + e$ has a cycle C. This cycle contains edges which do not belong to M, since e does not form a cycle in M but does in T. Note that, there exists $f \in C - M$ which has not been considered by the algorithms before and hence it must have weight at least as large as e. If f had been considered before e in the algorithm, $M - e + f$ will not contain a cycle (as $M - e + f \subset T$) and so f would have been added to M already.

Then $T - f + e$ is a tree by the insertion property and it has the same or less weight as T. If $w(f) = w(e)$, then $T - f + e$ is a minimum spanning tree containing $M + e$ and again the claim holds. If $w(f) > w(e)$, then $T - f + e$ is a spanning tree of strictly smaller weight than T and leads to a contradiction. □

Prim-Dijkstra-Jarnik's algorithm : Input graph weighted connected Graph $G = (V, E, w)$.

Step 1: Initialize with $M = D = \emptyset$ and $S = \{v\}, T = V - S$ for some $v \in V$.

Step 2: Select one of the smallest (in terms of weight) edges in $E \cap (S \times T)$. Call it e and $M = M \cup e$.

Step 3: Say $e = (v_1, v_2) \in S \times T$ then set $S = S \cup \{v_2\}$ and $T = V - S$.

Step 4: If $T \neq \emptyset$, then go to Step 2 else to Step 5.

Step 5: Output M.

Theorem 4.3.2. *The output of Prim-Dijkstra-Jarnik's algorithm is a minimal spanning tree if G is connected.*

Proof. Denote the output by $\tilde{M}$. The proof of the fact that $\tilde{M}$ is a spanning tree is left as an exercise as the steps were sketched in the class. W e shall only show minimality here.

Claim : As with the proof of Kruskal's algorithm, we shall show that at every step $M \subset M^*$ with the latter being a MST.

The claim is trivially true at the initial step when $M = \emptyset$ and since MST exists. Suppose the claim is true upto some step with M, T, S already determined i.e., $M \subset M^*$, a MST.

Let $e = (u, v)$ be the edge chosen by the Prim's algorithm at this stage and suppose $e \notin M^*$. Then, there exists a path P from u to v in M^*. This implies that there exists an edge $e' \in P \cap (T \times S)$. But since Prim's algorithm selected e instead of e', we have that $w(e) \leq w(e')$. But since $P + e$ is a cycle in $M^* + e$, we can remove any edge from the cycle and still get a spanning tree. Thus, $M^* + e - e'$ is also a spanning tree and since M^* is a MST, we have that $w(e) \geq w(e')$. Thus, we get that $w(e) = w(e')$ and $M + e \subset M^* + e - e'$, a MST. This proves the claim and completes the proof. $\square$

Exercise* 4.3.3. *The multi-set of weights for a minimal spanning tree is unique i.e., for any $s \in \mathbb{R}$ and any two MSTs M, M', we have that $|\{e \in M : w(e) = s\}| = |\{e \in M' : w(e) = s\}|$.*

Exercise 4.3.4 (Minimal Spanning Forests)**.** *Define and suitably extend these notions to minimal spanning forests.*

Exercise* 4.3.5. *Given a connected graph (G, E) with non-negative edge weights w satisfying the following conditions : $w(u, v) = \infty$ if $(u, v) \notin E$ and $w(u, v) > 0$ if $u \sim v$. Fixing a starting vertex x, show that the following algorithm computes the shortest paths from x to all other vertices in G.*

Step 1 : Let $S = \{x\}, t(x) = 0, t(u) = w(u, x)$ for $u \notin S$.

Step 2 : Select a $u \notin S$ such that $t(u) = \min_{z \notin S} t(z)$. Change S to $S \cup \{u\}$ and update $t(z)$ for $z \notin S$ to $\min\{t(z), t(u) + w(u, z)\}$.

Step 3 : Continue Step 2 until $S = V(G)$ and set $d(u, x) = t(u)$ for all u.

Show that $d(u, v) = d_G(u, v)$.

4.4 ***Some questions : Random spanning trees***

Let us consider the following weighted graph : K_n, the complete graph is our graph and the edge weights $\{w(e)\}_{e \in E(K_n)}$ are i.i.d. $U[0,1]$ random variables. Let M_n denote the minimal spanning tree (why is it unique ?). Alan Frieze ([Frieze 1985]) showed that $w(M_n) \to \zeta(3) = \sum_{k=1}^{\infty} k^{-3} = 1.202\ldots$ where ζ is the famed Riemann-zeta function. See [Addario-Berry 2015] for a newer proof.

We will now consider another weighted graph : Let $V_1, \ldots, V_n$ denote i.i.d. uniform points in $[0,1]^d$. Let G_n be the complete graph on $\{V_1, \ldots, V_n\}$ with edge weights being the euclidean distance between the points i.e., $w(V_i, V_j) = |V_i - V_j|$ for $i \neq j$. Let M_n denote the minimal spanning tree (why is it unique ?). It is known that $n^{-(d-1)/d} w(M_n) \to C_d$ for some $C_d \in (0, \infty)$. See the wonderful monograph of [Steele 1997] for a proof of this and various other related results in probabilistic combinatorial optimization such as stochastic travelling salesman problem, minimal matchings etc. Determining more information about the constant C_d is still open. See [Bertsimas 1990, Rhee 1992, Frieze and Pegden 2017] for some progress in this direction.

See the following talk by Lougi Addario-Berry for more on the thriving research on probabilistic aspects of minimal spanning trees - `http://problab.ca/louigi/talks/Msts.pdf`.

Chapter 5

Extremal Graph Theory

The extremal refers to the fact that one is concerned with extremal (maximal or minimal) questions about graphs. A prototypical question is the maximal or minimal number of edges in a graph with a certain property. We have already answered such a question about maximal number of edges for a graph to be a tree. Turan's theorem, which we shall shortly, is considered one of the foundational results in this subject. It is in answering a question in extremal graph theory, Erdös made powerful use of probabilistic ideas and this gave birth to what is now famously known as the probabilistic method. We shall see an illustration of this in the second section.

5.1 Existence of complete subgraphs and Hamiltonian circuits

Lemma 5.1.1. *Every graph has a self-avoiding walk of length $\delta(G)$.*

Proof. WLOG assume $\delta(G) \geq 1$. Let v_0 be a vertex in G. Given $v_i, i < \delta(G)$, we can choose a neighbour v_{i+1} of v_i such that $v_{i+1} \neq v_0, \ldots, v_{i-1}$ as v_i has at least $\delta(G)$ neighbours. Hence we get a self-avoiding walk $v_0 v_1 \ldots v_{\delta(G)}$ as needed. $\square$

Define *girth* of a graph $g(G) := \min\{l(C) : C, \text{a cycle}\}$. Set $g(G) = \infty$ if there exists no cycle. The diameter of a graph if $diam(G) := \max\{d(u, v) : u, v \in V\}$ and

Lemma 5.1.2. *If G has a cycle, $g(G) \leq 2diam(G) + 1$.*

Proof. Assume otherwise. Let $C = v_0 \ldots v_k v_0$ be the cycle of minimal length. Suppose k is even i.e., $l(C) = k + 1$ is odd. Then $d(v_0, v_{k/2}) = k/2$. But by definition of diameter, $k/2 \leq diam(G)$ and so $l(C) = k + 1 \leq 2diam(G) + 1$. If k is odd then $d(v_0, v_{(k+1)/2}) = k + 1/2$ and again we have that $l(C) = k + 1 \leq 2diam(G)$ $\square$

Lemma 5.1.3 (Mantel's theorem, 1907)**.** *If G is a graph on n vertices with no triangle then $|E| \leq \lfloor n^2/4 \rfloor$. Equivalently, if $|E| > n^2/4$, then $g(G) = 3$.*

Proof. Let G have $[n]$ has a vertex set and no triangles. Let z_i be weight of i such that $\sum_i z_i = 1$ and we shall try to maximize $S = \sum_{i \sim j} z_i z_j$. Let $k \nsim l$. Assume that $\sum_{j \sim k} z_j z_k = x z_k$ and $\sum_{j \sim l} z_j z_l = y z_l$. WLOG assume $x \geq y$. Observe that $z_k x + z_l y \leq (z_k + \epsilon)x + (z_l - \epsilon)y$. Thus if $(z_1, \ldots, z_n)$ is a configuration of weights, then $(z_1, \ldots, z_k + z_l, \ldots, 0, \ldots, z_n)$ is a configuration with larger S i.e., we have transferred weight from z_l to z_k. If we repeat the procedure again, we see that it stops when the weights are concentrated on two adjacent vertices. This is because whenever the weights are concentrated on at least three vertices, there are two vertices which are not neighbours as the induced subgraph is not complete. If the two adjacent vertices are i, j then $S = z_i z_j \leq 1/4$.

Let $z_i = n^{-1}$ for all $i \in [n]$. Then the corresponding $S = n^{-2}|E|$ and this is at most $1/4$ by the above argument. Hence the theorem is proved. □

Theorem 5.1.4 (Turan, 1941)**.** *If a simple graph on n vertices has no complete subgraph K_p, then $|E| \leq M(n,p) := \frac{(p-2)n^2 - r(p-1-r)}{2(p-1)}$ where $r \equiv n(\mod p-1)$.*

The above bound can be achieved as follows : Let $S_1, \ldots, S_{p-1}$ be an almost equal partition of V i.e., $S_1, \ldots, S_r$ are subsets of sizze $t+1$ and the rest are of size t for some $t \geq 0$ and $r \geq 1$. Construct the complete multi-partite graph on $S_1, \ldots, S_{p-1}$ such that all the edges between S_i and S_j are present for $i \neq j$ and these are the only edges. This does not have a complete subgraph K_p and the number of edges is $M(n,p)$.

Proof. Let t be such that $n = t(p-1) + r$. We will prove by induction on t. If $t = 0$, then $n = r, M(n,p) = n(n-1)/2$ and the theorem trivially holds as $n \leq p-1$. Now, consider a graph G on n vertices with no K_p subgraph (i.e., a subgraph isomorphic to K_p) and let G have the maximum number of edges subject to these constraints. Hence, G contains a subgraph H isomorphic to K_{p-1}. If not, one can add an edge to G without creating a K_p subgraph and so contradicting its maximality. The vertices $V - H$ are joined to at most $p-2$ vertices in H. Since $|V - H| = n - p + 1 = (t-1)(p-1) + r$ and the induced subgraph $< V - H >$ also does not contain a K_p subgraph, by induction hypothesis, $|E(< V - H >)| \leq M(n-p+1, p)$. Thus, we have that

$$|E(G)| \leq M(n-p+1, p) + (n-p+1)(p-2) + \binom{p-1}{2}$$

and one can easily verify that the RHS is equal to $M(n,p)$. □

See [Aigner et al. 2010, Chapter 36] for more proofs of Turan's theorem.

Exercise* 5.1.5. *If you generalize the argument for Mantel's theorem given in the class, what is the bound you get in Turan's theorem ?*

Theorem 5.1.6. *If a graph G on n vertices has more than $\frac{1}{2}n\sqrt{n-1}$ edges, then G has girth ≤ 4. That is G contains a triangle or quadrilateral.*

Proof. Suppose $g(G) \geq 5$. Let $v_1, \ldots, v_d$ be the neighbours of a vertex v. Since there are no triangles, $v_j \notin N_{v_i}$ for $i \neq j$ and since there are no quadrilaterals, $N_{v_i} \cap N_{v_j} = \{v\}$ for $i \neq j$. Thus, we have that $\cup_{i=1}^d N_{v_i} \setminus \{v\} \cup N_v \cup \{v\} \subset [n]$ and so $\sum_{i=1}^d (d_{v_i} - 1) + d + 1 \leq n$ and hence $\sum_{w \sim v} d_w \leq n - 1$. Thus, we et

$$n(n-1) \geq \sum_v \sum_{w \sim v} d_w = \sum_v d_v^2 \geq n^{-1}(\sum_v d_v)^2 = n^{-1}4|E|^2.$$

□

Theorem 5.1.7. *If G is a graph on n vertices with $n \geq 3$ and $\delta(G) \geq n/2$, then G contains an Hamilton cycle.*

Recall that a *Hamilton circuit* is a simple closed path passing through every vertex exactly once.

Proof. Suppose G is a counterexample to the theorem and G be such a graph with maximal number of edges i.e., addition of an edge to G creates a cycle. Let $v \nsim w$ and hence $G \cup (v, w)$ will contain a Hamilton cycle $v = v_1 v_2 \ldots v_n = w, v$. Thus $v_1 v_2 \ldots v_n$ is a simple path. Define sets $S_v := \{i : v \sim v_{i+1}\}$ and $S_w := \{i : w \sim v_i\}$. Since $\delta(G) \geq n/2$, $|S_v|, |S_w| \geq n/2$ and further $S_v, S_w \subset \{1, \ldots, n-1\}$. Hence $S_v \cap S_w \neq \emptyset$ and assume that $i_0 \in S_v \cap S_w$. Then $v = v_1 v_2 \ldots v_{i_0} w = v_n v_{n-1} \ldots v_{i_0+1} v_1 = v$ is a Hamiltonian circuit in G, contradicting our assumption. □

Where does the argument given in the class fail for $n = 2$?

Exercise* 5.1.8. *1. If $\delta(G) \geq 2$, every graph has a cycle of length at least $\delta(G) + 1$.*

2. *Prove that a graph G with $g(G) = 4$ and $\delta(G) \geq k$ has at least $2k$ vertices. Is $2k$ the best lower bound ?*

3. *Let G be a graph with $g(G) = 5$ and $\delta(G) \geq k$. Show that G has at least $k^2 + 1$ vertices.*

4. *Prove that a k-regular bipartite graph has no cut-edge for $k \geq 2$.*

5. *Show that Petersen graph (defined in Section 3.1) is the largest 3-regular graph with diameter 2.*

6. *Let G be a simple graph on n vertices ($n > 3$) with no vertex of degree $n - 1$. Suppose that for any two vertices of G, there is a unique vertex joined to both of them. If x and y are not adjacent show that $d(x) = d(y)$. Now, show that G is a regular graph.*

7. *Show that a simple graph on n vertices with $\lfloor n^2/4 \rfloor$ edges and no triangles, then it is the complete bipartite graph $K_{k,k}$ if $n = 2k$ or $K_{k,k+1}$ if $n = 2k + 1$.*

8. *If a simple graph on n vertices has e edges, then it has at least $\frac{e}{3n}(4e - n^2)$ triangles.*

9. *Suppose G is a graph with at least one edge. Then there exists a subgraph H such that $\delta(H) > \epsilon(H) \geq \epsilon(G)$ where recall that $\epsilon(G)$ was defined as the edge density of the graph (see Definition 3.2.8).*

5.2 ***Probabilistic Method : An introduction***

This section is not part of syllabus but rather an introduction to one of the most powerful tools in modern day graph theory and combinatorics.

Exercise 5.2.1. *In any graph G on 6 vertices, either $K_3 \subset G$ or $K_3 \subset \bar{G}$.*

Given m, n we define the *Ramsey numbers* $R(m, n)$ as follows :

$$R(m,n) := \inf\{t : \text{For any } G \subset K_t,\ K_m \subset G \text{ or } K_n \subset G^c\}.$$

Note that $R(m,n) = R(n,m)$, $R(m,2) = m$. Our previous lemma gives that $R(3,3) \leq 6$. Ramsey(1929) showed that $R(m,n) < \infty$ for all m, n. Using probabilistic method, Erdös showed that $R(k,k) > \lfloor 2^{k/2} \rfloor$. We will present the proof now.

Let $n = \lfloor 2^{k/2} \rfloor$ and we wish to show that there is a coloring of edges of K_n with blue and red such that there is no monochromatic clique of size k. The set of all possible colorings on K_n is $\Omega_n := E(K_n)^{\{R,B\}}$ i.e., a set of cardinality $2^{\binom{n}{2}}$. The crucial idea in probabilistic method is to pick a random variable X with values in Ω_n such that the probability X is a coloring with no monochromatic k-clique is positive and this trivially implies that the set of colorings with no monochromatic k-clique is non-empty as desired. As it is common, we can view $f \in \Omega_n$ as a function $f : E(K_n) \to \{R, B\}$.

Let $A_n := \{f \in \Omega_n : f \text{ has no monochromatic } k\text{-clique}\}$. Instead of showing $P(X \in A_n) > 0$, we will show that $P(X \in A_n^c) < 1$. We shall shortly see why the required upper bounding is easier. Firstly, note that if $f \notin A_n$, then there exists a subset $S \subset [n]$ and $|S| = k$ such that $f_{E_S} \equiv R$ or $f_{E_S} \equiv B$ where we have used E_S to abbreviate the edge set of $< S >$, the complete subgraph on S. Hence, we have that

$$A_n^c \subset \bigcup_{S \subset [n], |S| = k} (\{f : f_{E_S} \equiv R\} \cup \{f : f_{E_S} \equiv B\}).$$

By the union bound for probability distribution, we have that

$$P(X \notin A_n) \leq \sum_{S \subset [n], |S| = k} (P(X_{E_S} \equiv R) + P(X_{E_S} \equiv B)).$$

The union bound is what makes it easier to upper bound than lower bound. Also, so far we have not used any specific property of the random variable at all. We can now make a 'clever' choice of the random variable to obtain the desired upper bound. This flexibility in making 'clever' choices of random variables to obtain desired bounds is another hallmark of the probabilistic method and lends enormous power to the method.

Now choose $X \in \Omega_n$ as follows : $X(e), e \in E(K_n)$ are i.i.d. random variables with probability such that $P(X(e) = R) = 1/2 = P(X(e) - B)$. Equivalently, $P(X = f) = 2^{-\binom{n}{2}}$ for all $f \in \Omega_n$ i.e., X is uniformly distributed in Ω_n. Now trivially, we have that

$$P(X_{E_S} \equiv R) = P(X_{E_S} \equiv B) = 2^{-\binom{k}{2}}$$

and thus

$$P(X \notin A_n) \leq \binom{n}{k} 2^{1-\binom{k}{2}}.$$

I will leave it as an exercise to verify that for the choice of n we have made $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ as desired. This proves $R(k,k) > n := \lfloor 2^{k/2} \rfloor$.

The defect of probabilistic method as obvious in the above proof is that it is non-constructive i.e., it did not give explicitly a coloring that has a monochromatic k-clique in K_n.

Read [Alon and Spencer 2004, Chapter 1] and [Aigner et al. 2010, Chapter 40] for more illustrative examples on the probabilistic method and if interested, read more of [Alon and Spencer 2004].

Regardless of your liking of the proof via probabilistic method, Ramsey numbers raise many interesting questions themselves. For example, $R(5,5)$ is still unknown. Here is a legendary statement from Erdös attesting to the complexity of Ramsey numbers (see [Spencer 1994, Page 4]) :

" Erdös asks us to imagine an alien force, vastly more powerful than us, landing on Earth and demanding the value of $R(5,5)$ or they will destroy our planet. In that case, he claims, we should marshal all our computers and all our mathematicians and attempt to find the value. But suppose, instead, that they ask for $R(6,6)$. In that case, he believes, we should attempt to destroy the aliens. "

For more on Ramsey numbers, read `https://en.wikipedia.org/wiki/Ramsey's_theorem`.

5.3 ***Some other methods***

Extremal graph theory and combinatorics is fertile ground with interesting ideas from different fields of mathematics. We have indicated the probabilistic method above and we mention two notabl ones : Linear algebra method ([Babai 1992]) and the very recent Polynomial method ([Guth 2016, Grochow 2019]). An excellent overview of all these three methods can be found in [Jukna 2011].

Chapter 6

Matchings, covers and factors.

6.1 Hall's marriage theorem, Koenig's, Gallai's and Berge's theorems

Definition 6.1.1. *A subset M of edges is said to be* independent / matching *if no two edges are incident on any vertex or equivalently, every vertex is contained in at most one edge. A* complete matching *M on a subset $S \subset V$ is a matching that contains all the vertices in S. A perfect matching is a complete matching on G.*

Alternatively one can consider a matching of a graph M as a subgraph of G such that $d_M(v) = 1$ for all $v \in V(M)$. A matching is complete if M is spanning. A vertex v is said to be *saturated* if $v \in M$ and else *unsaturated*. For a subset $S \subset V$, $N(S) = \cup_{v \in S} N(v)$.

Theorem 6.1.2 (Hall's marriage theorem ; Hall, 1935)**.** *Let G be a bi-partite graph with the two vertex sets being V_1, V_2. Then there exists a complete matching on V_1 iff $|N(S)| \geq |S|$ for all $S \subset V_1$.*

We will give a proof via induction now and a later exercise will involve proof using max-flow min-cut theorem. See [Diestel 2000, Section 2.1] for two more proofs.

Proof. Let $|V_1| = k$ and our proof will be by induction on k. If $k = 1$, the proof is trivial.

Let $G = V_1 \cup V_2$ be such that the result holds for any graph with strictly smaller V_1.

Suppose that $|N(S)| \geq |S| + 1$ for all $S \subsetneq V_1$. Then choose $(v, w) \in E \cap V_1 \times V_2$ and consider the induced subgraph $G' :=< V - \{v, w\} >$. Since we have removed only w from V_2 and that $|N(S)| \geq |S| + 1$ for all $S \subsetneq V_1$, we get that $|N(S')| \geq |S'|$ for all $S' \subset V_1 - \{v\}$. Thus there is a complete matching M on $V_1 - \{v\}$ in G' by induction hypothesis and $M \cup (v, w)$ is a complete matching on V_1 in G as desired.

If the above is not true i.e., there exists $A \subset V_1$ such that $N(A) = B$ and $|A| = |B|$. Then, by induction hypothesis, there is a complete matching M_0 on A in the induced subgraph $< A \cup B >$. Trivially, Hall's condition holds i.e., for all $S \subset A$, $|N(S) \cap B| = |N(S)| \geq |S|$. Let $G' := G - < A \cup B >$. Let $S \subset V_1 - A$. Suppose if $|N'(S)| < |S|$ where $N'(S) = N(S) \cap (V_2 - B)$. Then, we have that

$N(S \cup A) = N'(S) \cup B$ and hence $|N(S \cup A)| \leq |N'(S)| + |B| < |S| + |A|$, a contradiction. Hence, G' also satisfies Hall's condition and again by induction hypothesis G' has a complete matching M' on $V_1 - A$. Thus, we have a complete matching $M := M_0 \cup M'$ on V_1 in G. $\square$

Exercise 6.1.3. *For $k > 0$, a k-regular bipartite graph has a perfect matching.*

Proposition 6.1.4. *Let $d \geq 1$. Let G be a bipartite graph on $V_1 \sqcup V_2$ such that $|N(S)| \geq |S| - d$ for all $S \subset V_1$. Then G has a matching with at least $|V_1| - d$ independent edges.*

Proof. Set $V_2' := V_2 \cup [d]$. Define G' with vertex set as $V_1 \sqcup V_2'$ and edge set as $E(G) \cup (V_1 \times [d])$. Then, it is easy to see that Hall's condition is true on G' and hence there is a complete matching M of V_1 in G'. Now, if we remove the edes in M incident on $[d]$, we get a matching with at least $|V_1| - d$ edges as required. $\square$

Exercise* 6.1.5. *Let G be a bi-partite graph with partitions $V_1 = \{x_1, \ldots, x_m\}$ and $V_2 = \{y_1, \ldots, y_n\}$. Then G has a subgraph H such that $d_H(x_i) = d_i, d_H(y_j) \leq 1$, $\forall 1 \leq i \leq m$ and $\forall 1 \leq j \leq n$ iff for all $S \subset V_1$, we have that $|N(S)| \geq \sum_{x_i \in S} d_i$.*

Definition 6.1.6 (Independent sets and covers)**.** *An* independent set of vertices *is $S \subset V$ such that no two vertices in S are adjacent. A subset of vertices $S \subset V$ is a* vertex cover *if every edge in G is incident to at least one vertex in S. An edge cover is a set of edges $E' \subset E$ such that every vertex is contained in at least one edge in E'.*

For the relevance of independent sets to information theory/communication, see Section 6.4.

Definition 6.1.7 (Independence number and cover number)**.**

$$\begin{aligned}
\alpha(G) &= \max\{|S| : S \textit{ independent vertex set}\}.\\
\alpha'(G) &= \max\{|M| : M \textit{ independent edge set}\}.\\
\beta(G) &= \min\{|S| : S \textit{ vertex cover}\}.\\
\beta'(G) &= \min\{|E'| : E' \textit{ edge cover}\}.
\end{aligned}$$

Exercise 6.1.8. *Prove that a graph G is bipartite if and only if every subgraph H of G has an independent set consisting of at least half of V(H).*

We first derive some trivial relations between the four quantities. If M is a maximal matching, then to cover each edge we need distinct vertices and hence the vertex cover should have size at least M. This yields the first inequality below.

$$\alpha'(G) \leq \beta(G) \leq 2\alpha'(G) \;\; ; \;\; \alpha(G) \leq \beta'(G).$$

As for the second inequality, observe that to cover vertices of an independent set, we need distinct edges.

Lemma 6.1.9. *Let G be a graph. $S \subset V$ is an independent set iff S^c is a vertex cover. As a corollary, we get $\alpha(G) + \beta(G) = n = |V|$.*

The above lemma follows trivially from the definitions.

Theorem 6.1.10 (Konig, Egervary, 1931)**.** *For a bi-partite graph, $\alpha'(G) = \beta(G)$.*

Remark 6.1.11 (Equivalent Formulation :)**.** *A bi-partite graph is equivalent to a $0-1$ matrix. Denote the rows by V_1 and columns by V_2. If (v_1, v_2)-entry is a 1 then there is an edge $v_1 \sim v_2$. It is easy to see that given a bi-partite matrix it can be represented as a $0-1$ matrix with rows indexed by V_1 and columns indexed by V_2.*

By a line, *we shall mean either a row or a column. Under the matrix formulation, vertex cover is a set of lines that include all the 1's. An independent set of edges is a collections of 1's such that no two 1's are on the same line.*

Proof. We will show that for a minimal vertex cover Q, there exists a matching of size at least $|Q|$. Partition Q into $A := Q \cap V_1$ and $B := Q \cap V_2$. Let H and H' be induced subgraphs on $A \sqcup (V_2 - B)$ and $(V_1 - A) \sqcup B$ respectively. If we show that there is a complete matching on A in H and a complete matching on B in H', we have a matching of size at least $|A| + |B|$ $(= |Q|)$ in G. Also, note that it suffices to show that there is a complete matching on A in H because we can reverse the roles of A and B apply the same argument to B as well.

Since $A \cup B$ is a vertex cover, there cannot be an edge between $V_1 - A$ and $V_2 - B$. Suppose for some $S \subset A$, we have that $|N_H(S)| < |S|$. Since $N_H(S)$ covers all edges from S that are not incident on B, $Q' := Q - S + N_H(S)$ is also a vertex cover. By choice of S, Q' is a smaller vertex cover than Q contradiciting the minimality of Q. Hence, we have that Hall's condition holds true for A in H. And by the arguments in the previous paragraph, the proof is complete. □

Exercise 6.1.12. *If a graph G does not contain a path of length more than 2, show that it's connected components are all star graphs.*

Theorem 6.1.13 (Gallai, 1959)**.** *If G is a graph without isolated vertices, then $\alpha'(G) + \beta'(G) = n = |V|$.*

Proof. Suppose M is a maximal matching. Then $S = V - V(M)$ is also an independent set. If there are edges between vertices of S, then such edges can be added to M and one can obtain a larger matching. Hence there are no edges between vertices of S and hence it is a independent set. Construct a edge cover as follows : Add all edges in M to Q and for each $v \in S$, add one of its adjacent edges to Q. Thus $|Q| = |M| + |S|$ and since $V(M) \sqcup S = V$, we can derive that

$$\alpha'(G) + \beta'(G) \leq |M| + |Q| = 2|M| + |S| = n.$$

Let Q be a minimal edge cover. Then Q cannot contain a path of length more than 2. Else, by removing the middle edge in a path of length at least 3, we can obtain a smaller edge cover. By the previous exercise, Q is a graph consisting of star components. If $C_1, \ldots, C_k$ are the components of Q, then $V(C_1) \cup \ldots \cup V(C_k) = V$ and $E(C_1) \cup \ldots \cup E(C_k) = Q$. Now choose a matching $M = \{e_1, \ldots, e_k\}$ by selecting one edge from every component $C_1, \ldots, C_k$. Since C_i's are disjoint, M is a matching. Thus, using the fact that each Q is a forest with k components, we can derive that

$$\alpha'(G) + \beta'(G) \geq |M| + |Q| = k + \sum_{i=1}^{k} |E(C_i)| = \sum_{i=1}^{k} |V(C_i)| = n.$$

□

As a corollary, we get König's result : if G is bi-partite graph without isolated vertices, $\alpha(G) = \beta'(G)$.

Definition 6.1.14 (Augmenting path). *Given a matching M, a M-alternating path P is a path such that its edges alternate between M and M^c. A M-augmenting path is a M-alternating path whose end-vertices do not belong to M.*

Theorem 6.1.15 (Berge, 1957). *A matching M in a graph is a maximum matching in G iff G has no M-augmenting path.*

Exercise* 6.1.16. *If M, M' are two matchings, every component of $M \triangle M'$ is a path or an even cycle.*

Proof. Suppose there is an M-augmenting path P. Let $P = v_0 v_1 \ldots v_k$. Since P is M-augmenting, $(v_0, v_1), (v_2, v_3), \ldots, (v_{k-1}, v_k) \notin M$ and $(v_1, v_2), (v_3, v_4), \ldots, (v_{k-2}, v_{k-1}) \in M$. Now, observe that $M' = M - P \cup \{(v_0, v_1), (v_2, v_3), \ldots, (v_{k-1}, v_k)\}$ is a larger matching than M. Hence if M is a maximum matching, there is no M-augmenting path.

Suppose M' is a larger matching than M. We shall construct an M-augmenting path and prove the theorem by contraposition. Let $F = M \triangle M'$. We know by the above exercise that the components of F are paths or even cycles. Since $|M'| > |M|$, there must be a component of F such that M' has more edges in that component than M'. If a component in F is an even cycle, it consists of same number of edges from M and M'. Thus, the component for which M' has more edges must be a path, say $P = v_0 \ldots, v_k$. Since $P \subset F$, we have that P has to be an M-alternating path i.e., $(v_0, v_1) \in M', (v_1, v_2) \in M, \ldots$ or $(v_0, v_1) \in M, (v_1, v_2) \in M', \ldots$. Since $m' := |M' \cap P| > |M \cap P| = m$ and that P is an M-alternating path, we derive that $m' - m = 1$ and $k = 2m + 1$. Further, this implies that $(v_0, v_1), (v_2, v_3), \ldots, (v_{k-1}, v_k)\ inM'$ and $(v_1, v_2), (v_3, v_4), \ldots, (v_{k-2}, v_{k-1}) \in M$ i.e., P is an M-augmenting path. □

6.2 Graph factors and Tutte's theorem

Definition 6.2.1 (Factor of a graph). *Given a graph G, a subgraph H is said to be* a factor *if $V(H) = V(G)$ i.e., spanning. An r-factor is a factor that is r-regular.*

Thus, 1-factors are nothing but perfect matchings.

Theorem 6.2.2 (Petersen, 1891)**.** *Every regular graph of positive even degree has a 2-factor.*

Proof. Let G be any $2k$-regular graph for $k \geq 1$. WLOG, let G be connected. Let $v_1v_2\dots v_nv_0$ be the Eulerian circuit in G. Now, we replace each vertex v by two vertices v^-, v^+ and the edge v_iv_{i+1} by $v_i^-v_{i+1}^+$. Thus we get a new k-regular graph G'. By Exercise 6.1.3, there is a 1-factor in G'. Now, by merging the vertices v^-, v^+ in G', we get a 2-factor in G. □

For a graph G, let $o(G)$ denote the number of odd components of G.

Theorem 6.2.3 (Tutte , 1947)**.** *A graph has a 1-factor iff $o(G - S) \leq |S|$ for all $S \subset V$.*

Proof. (Lovasz, 1975) In a 1-factor, the odd components will have at least one vertex each to be matched to vertices in S and these vertices in S have to be distinct. Thus, $o(G - S) \leq |S|$ for all $S \subset V$ if G has a 1-factor.

Now let G be a graph without 1-factor. We shall show that there is a *bad set* i.e., a set S violating Tutte's condition - $o(G - S) \leq |S|$ for all $S \subset V$. If $G' = G + e$ has no 1-factor, then so does G. Further if G' has a bad set S, then so does G. Hence, we shall assume that G is an edge-maximal graph without 1-factor and we shall find a bad set in G.

Heuristic : If S is a bad set for G, then all components of $G - S$ have to be complete. If a component, say C, is not complete, we can consider $G \cup e$ where e is a missing edge in C. Since $o(G \cup e - S) = o(G - S) > |S|$, by forward implication of the theorem we have that $G \cup e$ is not a 1-factor i.e., violating edge-maximality of G. Further, by the same reasoning, $s \in S$ must be connected to all vertices in $G - S$. We will now show that these two conditions characterize bad sets and then produce such a set.

Let us say a set $S \subset V$ satisfies *condition B* if all components of $G - S$ are complete and all $s \in S$ are connected to all vertices in $G - S$.

Claim : If S satisfies condition B, then either S is bad or $\emptyset$ is bad.

If S is not bad, then we can join one vertex from each of the odd components to S disjointly and then try to pair up the remaining vertices. In every even component of $G - S$, we can pair up the vertices and in every odd component too, we can pair up the vertices not paired to S. We are only left with remaining vertices S' in S where $|S'| = |S| - o(G - S)$ and S' forms a complete subgraph. But since G does not contain a 1-factor, $|S| - o(G - S)$ is odd. Since, there is a complete matching on $V - S'$, $|V - S'|$ is even and so $|V|$ is odd i.e., $\emptyset$ is a bad set.

Now, to show that G has a set S satisfying condition B, let S be the set of vertices such that $s \in S$ is adjacent to every other vertex. If S does not satisfy condition B, then some component of $G - S$ isn't complete. Let $v \nsim w$ in a component and let v, v_1, v_2 be the first three vertices on a shortest path from v to w. The $v \sim v_1, v_1 \nsim v_2$ but $v \nsim v_2$. Since $v_1 \notin S$, there is a u such that $u \nsim v_1$. By maximality of G, there is a 1-factor H_1 in $G + (vv_2)$ and H_2 in $G + (v_1u)$.

Let $P = u \dots u'$ be a maximal path in G starting at u with an edge from H_1 and alternating between edges in H_1 and H_2. If last edge of P is in H_1, by maximality of P this implies that there is no edge of H_2 incident on u' in G. This implies that $u' = b$ as every other vertex has an edge of H_2 incident on it in G. We will then set $C = P + v_1u$. Note that C is a cycle and it is of even length as P starts in H_1 and ends in H_1. If the last edge of P is in H_2, again by maximality of P this implies that there is no edge of H_1 incident on u' in G and as earlier, this means $u' \in \{v, v_2\}$. Then consider $C = P + (v', v_1) + (v_1, u)$. Again, C is an even cycle as P starts in H_1 and ends in H_2. In either case C is an even cycle and $(v_1, u) \in C - E$. But then consider $H' = H_2 - (C \cap H_2) + (C - H_2)$ i.e., on C replace the edges of H_2 by those of $C - H_2$. Now, $H' \subset G$. Since H_2 is a 1-factor, so is H' and thus we get a contradiction. Thus S is a bad set as required. $\square$

The above proof of Lovasz can be found in [Diestel 2000].

Corollary 6.2.4 (Defective version). *A graph G contains a subgraph H with a 1-factor with $|V(H)| \geq |V(G)| - d$ iff $o(G - S) \leq |S| + d$ for all $S \subset V$.*

Proof. Let G contain a subgraph H with a 1-factor with $|V(H)| = |V(G)| - k$ for some $k \leq d$. Consider $G' := G \vee K_k$. Then G' has a 1-factor by considering H along with a matching between $V(G) - V(H)$ and K_k. Let $S \subset V$. Then $o(G' - (S \sqcup [k])) \leq |S| + k$ as G' has a 1-factor. Further, $G' - (S \sqcup [k]) = G - S$ and so $o(G - S) \leq |S| + k \leq |S| + d$.

Let $o(G - S) \leq |S| + d$ for all $S \subset V$ and d is the minimal such number i.e., $d = \max\{o(G - S) - |S| : S \subset V\}$. We will assume $d \geq 1$ else $d = 0$ and Tutte's 1-factor theorem applies. Suppose $d = o(G - S_0) - |S_0|$ for some $S_0 \subset V$. Then the parity of vertices in V is same as that of an odd multiple of $d + |S_0|$ $(= o(G - S_0))$ plus S_0. The parity of the latter is same as an odd multiple of d which is nothing but the parity of d i.e., $|V|$ and d are both even or odd.

Let $G' := G \vee K_d$ and by above argument $|V'|$ is even. We will show that there exists a 1-factor in G' and this will imply that G contains a subgraph H with a 1-factor such that $|V(H)| \geq |V(G)| - d$. To show existence of 1-factor in G' we will verify the Tutte's condition.

Let $S' \subset V \sqcup [d]$. Suppose that $[d] \backslash S' \neq \emptyset$. Then $G' - S'$ is connected and hence $o(G' - S') \leq 1 \leq |S'|$ unless $|S'| = \emptyset$. For $S" = \emptyset$, $o(G' - S') = 0$ as $|V'|$ is even. Suppose $[d] \subset S'$. Let $S' = [d] \sqcup S$ for some $S \subset V$. Hence $G' - S' = G - S$ and so $o(G' - S') = o(G - S) \leq |S| + d = |S'|$. Thus G' satisfies Tutte's condition. $\square$

Definition 6.2.5 (f-factor). *Given a function $f : V \to \mathbb{N} \cup \{0\}$, an f-factor of a graph G is a subgraph H such that $d_H(v) = f(v)$ for all $v \in V$.*

Tutte [1952] showed a necessary and sufficient condition for a graph G to have an f-factor. The proof was by reducing it to a problem of checking for a 1-factor in a certain simple graph.

A graph construction : Given a function f and a graph G with $f \leq d$, we define a graph H as follows : Let $e := d - f$. To construct H replace each vertex v with a bi-partite graph $K_v := K_{d(v), e(v)}$ with vertex set $A(v) \sqcup B(v)$. For edge $(v, w) \in G$, add an edge between a vertex of $A(v)$ and $A(w)$.

Theorem 6.2.6. *G has an f-factor iff $f \leq d$ and the graph H constructed as above has a 1-factor.*

Proof. If G has a f-factor, $e(v)$ vertices in $A(v)$ are unmatched. These can be matched arbitarily with vertices of $B(v)$, giving us a 1-factor of H.

Suppose H has a 1-factor. Remove $B(v)$ and the vertices in $A(v)$ matched to $B(v)$. Now, at each v, $A(v)$ has $f(v)$ vertices remaining. If we merge these $f(v)$ vertices and call them v, we get a f-factor of G. □

Observe that we did not use the fact that G is a simple graph. Now Tutte's 1-factor condition can be translated to a f-factor condition. See[West 2001, Exercise 3.3.29]. An important application is Erdös-Gallai characterization of degree sequences of graphs (see Section 6.5).

Exercise* 6.2.7.

1. *A tree T has a perfect matching iff $o(T-v)=1$ for all $v \in T$.*
2. *Show that $2\alpha'(G) = \min_{S\subset V}\{|V(G)| - d(S)\}$ where $d(S) = o(G-S) - |S|$.*
3. *For any k, show that there are k-regular graphs with no perfect matching.*
4. *If G is k-regular, has even number of vertices and remains connected when any $(k-2)$ edges are deleted, then G has a 1-factor.*
5. *Every 3-regular graph with no cut-edge has a 1-factor.*
6. *Let Q_n be the hypercube graph on $0,1^n$. What are $\alpha(Q_n), \alpha'(Q_n), \beta(Q_n), \beta'(Q_n)$?*
7. *Every 3-regular simple graph with no cut-edge decomposes (i.e., edge-disjoint union) into copies of P4 (the 4-vertex path).*
8. *Let G be a k-regular bipartite graph. Prove that G can be decomposed into r-factors iff r divides k.*
9. *Is it true that every tree has at most one perfect matching ?*
10. *Let T be a tree on n vertices such that $\alpha(T)=k$. Can you determine $\alpha'(T)$ in terms of n,k ?*
11. *Characterize the graphs G for which the following statements hold. Justify your answers.*
    (1) (max. independent set) $\alpha(G)=1$.
    (2) (max. size of matching) $\alpha'(G)=1$.
    (3) (min. vertex cover) $\beta(G)=1$.
    (4) (min. edge cover) $\beta'(G)=1$.

    NOTE : In each of the above, you are required to prove a statement of the form $\ldots(G)=1$ iff G is $\ldots$.

6.3 ***Gale-Shapley Stable marriage/matching algorithm***

In more real-life scenarios, like college admission, students have preferences for colleges (i.e., a ranking) and so do the colleges. This in graph-theoretic terms, corresponds to a complete bi-partite graph with every vertex ranking every other vertex in the opposite partition. *Stability of matching* here would refer to the fact that there is no pair of vertices that would prefer each other over their current pairing i.e., if $u \sim v$ and $w \sim x$ in M with $u, w \in V_1, v, x \in V_2$, then it is not possible that u prefers x to v and x prefers u to w.

Does such a matching for any possible ? And more importantly, can one construct such a matching ?

Gale-Shapley Algorithm : In 1962, David Gale and Lloyd Shapley ([Gale and Shapley 1962]) proposed an algorithm to achieve stable matching and this probably the best known of such algorithms. Along with Alvin Roth, Lloyd Shapley was awarded the Noble prize in economics in 2012 for "for the theory of stable allocations and the practice of market design."

1. Input : n Men, n Women and their preferences.

Step 1 : Each "unengaged" man "proposes" to the highest women who has not yet rejected him.

Step 2 : Each woman agrees to get "engaged" to the "highest proposer" in her list. If previously engaged and current proposer is ranked hgiher, she rejects her the previous engagement. The other proposers are also rejected.

Step 3 : If any "engagement" is nullified, the man becomes "unengaged"

Step 4: Go to Step 1 if there is an "unengaged" man.

See [West 2001, Theorem 3.2.18] for the formal theorem statement and proof of the algorithm.

6.4 ***Shannon rate of communication***

See [Aigner et al. 2010, Chapter 37] for more details. Suppose V is a set of symbols to be communicated over a network. However, some symbols are likely to be confused. How best to choose a set of symbols that cannot be confused ? Place an edge between two symbols that can be confused and form a graph G. Let us call G, the confusion graph. A set of symbols that cannot be confused is an independent set. Best strategy for communication is to choose independent set.

A different question leading to the same answer. V is set of people trying to communicate but two neighbours cannot communicate at the same time. The best strategy is to choose an indepndent set of vertices and they can communicate at the same time.

In the first problem, if we are allowed to communicate over multiple rounds, can we do better ?

Given graphs G, H, define $G \times H$ as the graph with vertex set $V(G) \times V(H)$ and edges $(u_1, u_2) \sim (v_1, v_2)$ if $(u - 1, v_1) \neq (u_2, v_2)$ and $u_i \sim v_i, i = 1, 2$. Then, the confusion graph for strings of length 2 is $G^2 := G \times G$. Similarly, the confusion graph for strings of length n is G^n.

In G, the rate of information per symbol that

6.5 ***Erdös-Gallai Theorem***

A list of non-negative integers $(d_1, \ldots, d_n)$ is said to be *graphic* if these are vertex degrees of a n-vertex simple graph. We present the famous theorem giving a necessary and sufficient condition for existence of simple graphs with a given degree sequence. We have seen such a condition for trees already in Theorem 4.1.5.

Theorem 6.5.1 (Erdös-Gallai Theorem, 1960)**.** *A list of non-negative integers $(d_1, \ldots, d_n)$ in non-increasing order is graphic iff $\sum_i d_i$ is even and for all $1 \le k \le n$, we have that*

$$\sum_{i=1}^{k} d_i \le k(k-1) + \sum_{i=k+1}^{n} \min\{d_i, k\}.$$

We present the recent proof due to [Tripathi 2010]. One another proof using Tutte's f-factor theorem (Theorem 6.2.6) can be found in [West 2001, Exercise 3.3.29]. There is also a simple proof in [Choudum 1986].

We call a graph G on vertices $v_1, \ldots, v_n$, is said to be *a subrealization* of a nonincreasing list $(d_1, \ldots, d_n)$ if $d(v_i) \le d_i$ for all $1 \le i \le n$. For a subrealization, we say r is *the critical index* if $d(v_i) = d_i$ for $1 \le i < r$ and it is the largest such index. The sufficiency part (and the non-trivial one) of the proof proceeds as follows : We start with a subrealization with n vertices and no edges. Assuming $d_1 \ne 0$, we set $r = 1$. When $r \le n$, we do not change the degree of r but try to increase the degree of vertex v_r and reduce the deficiency $d_r - d(v_r)$.

Proof. The necessity is argued as follows : Evenness is trivial. If we look at the sum of the degrees of the largest k vertices, then the edges are either among the k vertices or to outside. The edges among the k vertices are counted twice and hence at most twice that of K_k i.e., $k(k-1)$. The edges from $1, \ldots, k$ to i for $i > k$ is at most $\min\{d_i, k\}$.

Now, we prove the sufficiency as described above : Given critical index r, define $S = \{v_{r+1}, \ldots, v_n\}$. Our construction shall give that S is an independent set at every step. This is true for $r = 1$.

- Case 0 : $v_r \nsim v_i$ for some v_i such that $d(v_i) < d_i$. Add edge (v_i, v_r).

- Case 1 : $v_r \nsim v_i$ for some $i < r$. Since $d(v_i) = d_i \ge d_r > d(v_r)$, there exists $u \sim v_i, u \nsim v_r, u \ne v_r$. If $d_r - d(v_r) \ge 2$, replace (u, v_i) with $(u, v_r), (v_i, v_r)$. If $d_r - d(v_r) = 1$, then since $\sum_i (d_i - d(v_i))$ is even, there is an index $k > r$ such that $d(v_k) < d_k$. Either add edge (v_r, v_k) if the edge is not present and if not replace $(u, v_i), (v_r, v_k)$ with $(u, v_r), (v_i, v_r)$.

- Case 2: $v_1, \ldots, v_{r-1} \in N(v_r)$ and $d(v_k) \ne \min\{r, d_k\}$ for some $k > r$. Since $d(v_k) \le d_k$ and S is independent, so $d(v_k) \le r$. Thus, we have that $d(v_k) < \min\{r, d_k\}$. If $v_r \nsim v_k$, we can use Case

0. If not, since $d(v_k) < r$, there exists $i < r$ such that $v_k \nsim v_i$. Since $d(v_i) = d_i \geq d_r > d(v_r)$, there exists $u \sim v_i, u \nsim v_r, u \neq v_r$. Remove (u, v_i) and add $(u, v_r), (v_i, v_k)$.

- CASE 3: $v_1, \ldots, v_{r-1} \in N(v_r)$ AND $v_i \nsim v_j$ FOR SOME $i < j < r$: Since $d(v_i) \geq d(v_j) > d(v_r)$, we have that there exists $u \sim v_i, u \nsim v_r, u \neq v_r$ and $w \sim v_j, u \nsim v_r, w \neq v_r$. It is possible that $u = w$. If $u, w \in S$, then delete $(u, v_i), (w, v_j)$ and replace with $(v_i, v_j), (u, v_r)$. If $u \notin S$ or $w \notin S$, apply the arguments as in Case 1.

Suppose that none of the above cases apply. Since Case 1 doesn't apply, $v_1, \ldots, v_{r-1} \in N(v_r)$ and since Case 3 also doesn't apply, $v_1, \ldots, v_r$ are all pairwise adjacent. Case 2 also doesn't apply and hence $d(v_k) = \min\{r, d_k\}$ for all $k > r$. Since S is independent, we have that $\sum_{i=1}^{r} d(v_i) = r(r-1) + \sum_{i=r+1}^{n} \min\{r, d_k\}$ and hence $\sum_{i=1}^{r}(d_i - d(v_i)) \leq 0$ by the Erdös-Gallai condition. Thus, we get that $d(v_i) = d_i$ for all $1 \leq i \leq r$ and we have eliminated deficieny at r. We can now increase r by 1 and continue with the procedure as above.

□

6.6 ***Equivalent theorems to Hall's matching theorem and more applications***

There are various powerful theorems in combinatorics and other areas that are equivalent to Hall's marriage theorem. See `https://en.wikipedia.org/wiki/Hall's_marriage_theorem#Logical_equivalences`

We shall see two of them later - Max-flow min-cut theorem and Menger's theorem (see Chapter 7).

Another non-trivial application of Hall's marriage theorem is to show the existence of Haar measure on compact Topological groups. See [Krishnapur 2019, Part 4] for details on this result and also for more about the equivalences mentioned above.

Chapter 7

Flows on networks, vertex and edge connectivity

7.1 Max-flow min-cut theorem

Given a simple graph $G = (V, E)$, we set $\overleftarrow{E}$ to be the set of ordered edges of E whereby we give both the orientations to an edge $e \in E$. In other words, $e \in \overleftarrow{E}$ implies that $e = (x, y) \neq (y, x)$ and we rather denote $-e = (y, x)$ in such a case. Further, we set $e^- = x, e^+ = y$.

Definition 7.1.1. *(Flow) Let $s, t \in V$ (source, target). A flow f from s to t is a function $f : \overleftarrow{E} \to \mathbb{R}$ such that*

1. *(anti-symmetry) $f(e) = -f(-e)$*
2. *(Kirchoff's node law) $(d^* f)(x) := \sum_{e:e^-=x} f(e) = 0$ for all $x \neq \{s, t\}$.*
3. *(Positive output at source :) $(d^* f)(s) \geq 0$.*

Definition 7.1.2. *Let $c : E \to [0, \infty]$ be the capacity function. A flow from s to t is said to satisfy the capacity constraints if $f(e) \leq c(e)$ and we call such a flow feasible.*

One can suitably define in-flow and out-flow at a vertex and show that Kirchoff's node law implies that the in-flow and out-flow are equal at all vertices except the sink. We can further define *value of a flow* $v(f) := d^* f(s)$ and it can be shown using Kirchoff's node law and anti-symmetry that $v(f) = -d^* f(t)$.

Let $S \subset V$. We call a pair (S, S^c) a (s, t)-cut if $s \in S, t \notin S$. Defining $C(X, Y) = \sum_{x \in X, y \in Y} c(x, y)$, $f(X, Y) = \sum_{x \in X, y \in Y} f(x, y)$, we can show that $v(f) = f(S, S^c) \leq C(S, S^c)$ for any (s, t)-cut (S, S^c) and any feasible $s - t$ flow. Thus, we have that $\sup\{v(f) : f \text{ is a } (s, t)\text{-flow}\} \leq \inf_{S: s \in S, t \notin S} C(S, S^c)$.

Theorem 7.1.3 (Max-flow min-cut theorem ; Elias-Feinstein-Shannon and Ford-Fulkerson (1956))**.**

$$\max\{v(f) : f \textit{ is a feasible } (s, t)\textit{-flow}\} = \inf_{S \subset V: s \in S, t \notin S} C(S, S^c).$$

It will soon become clear that Max-flow min-cut theorem is not a single theorem but a class of theorems that can be proven under different frameworks using variants of the ideas we will use to prove the above theorem - Theorem 7.1.3.

Lemma 7.1.4. *If there exists an infinite capcity $s-t$ path, then the max-flow is infinite and so is the min-cut. Else, the min-cut is finite and so if the max-flow.*

Proof. The proof of first path follows by constructing a sequence of flows with increasing strenghts. We will prove the second part alone. Let there be no infinite capcity $s-t$ path. We construct a finite min-cut as follows : Choose an $s-t$ path P_1 and since it is not infinite, there exists e_1 such that $c(e_1) < \infty$. Now repeat this procedure on $G - e_1$ and choose an edge e_2 in a $s-t$ path P_2 such that $c(e_2) < \infty$. Repeatedly, we can choose edges $e_1, \ldots, e_k$ until $G - \{e_1, \ldots, e_k\}$ has no $s-t$ path. Let S be the set of vertices in the component of s in $G - \{e_1, \ldots, e_k\}$ and clearly $t \in S^c$. Since $s-t$ path exists in G, we have that $E \cap (S \times S^c) \subset \{e_1, \ldots, e_k\}$ and hence $C(S, S^c) \leq \sum_{i=1}^k c(e_i) < \infty$. □

Lemma 7.1.5. *If the capacity function is bounded, then*

$$\sup\{v(f) : f \textit{ is a feasible } (s,t)\textit{-flow}\} = \max\{v(f) : f \textit{ is a feasible } (s,t)\textit{-flow}\}.$$

Lemma 7.1.6. *Let f be a $s-t$ flow in a graph and let $P = e_1 \ldots e_k$ be a $s-t$ path i.e., $e_1^- = s, e_k^+ = t$ and $e_i^+ = e_{i+1}^-$ for all $1 \leq i < k$. Then for every $\epsilon > 0$, f' defined as follows is also a flow : $f'(e) := f(e)$ for $e, -e \neq e_1, \ldots, e_k$, $f'(e) = f(e) + \epsilon, e = e_1, \ldots, e_k$, $f(e) = f(e) - \epsilon, e = -e_1, \ldots, -e_k$. Further $v(f') = v(f) + \epsilon$.*

Proof. Clearly anti-symmetry holds and $(d^* f')(s) = \sum_{e:e^-=s} f'(e) = \sum_{e:e^-=s} f(e) + \epsilon \geq 0$. It remains only to show that $(d^* f')(v) = 0$ for all $v \neq s, t$. This holds trivially for all $v \neq e_i^-, i = 2, \ldots, k$. Suppose $v = e_i^-$ for some $1 < i \leq k$. Then

$$\begin{aligned}(d^* f')(v) = \sum_{e:e^-=v} f'(e) &= \sum_{e:e^-=v, e\neq e_{i-1}, e_i} f(e) + f'(e_i) + f'(-e_{i-1}) \\ &= \sum_{e:e^-=v, e\neq e_{i-1}, e_i} f(e) + f(e_i) + f(-e_{i-1}) = (d^* f)(v) = 0.\end{aligned}$$

□

We first present the Ford-Fulkerson which gives the idea of the proof.

Remark 7.1.7 (Ford-Fulkerson Algorithm). *Given a flow f, define the* residual capacity $c_f(u,v) = c(u,v) - f(u,v)$.

Step 1: Set $f \equiv 0$.

*Step 2 : If there is a path P (called f-*augmenting path *from s to t in G such that $c_f(u,v) > 0$ for all edges $(u,v) \in P$, then go to Step 3 else go to Step 6.*

Step 3 : Find $c_f(P) = \min\{c_f(u,v) : (u,v) \in P\}$ *(*residual capacity*).*

Step 4 : For each edge $(u,v) \in P$*, set* $f(u,v) = f(u,v) + c_f(P)$ *and* $f(v,u) = f(v,u) - c_f(P)$*.*

Step 5 : Go back to Step 2.

Step 6 : Output flow f *as the maximal flow.*

Defining S_f *be the set of vertices that can be reached from* s *with a path* P *such that* $c_f(u,v) > 0$ *for all edges* $(u,v) \in P$*, note that Step 2 can also be rephrased as follows : If* $t \in S_f$ *go to Step 3 else go to Step 6.*

Proof. (Theorem 7.1.3) Suppose that the capacity function is bounded. We will prove the theorem under this assumption and then argue the other case.

Let f be the max-flow whose existence is guaranteed by Lemma 7.1.5. Define

$$S = S_f = \{v : \text{there exists a positive residual } s - v \text{ path}\} \cup \{s\}.$$

If $t \notin S$, then $[S, S^c]$ is as $s - t$ cut. Also, if $e \in S \times S^c$, then $c_f(e) = 0$ as otherwise $e^+ \in S$ which is a contradiction. As we argued before and by the zero residular property of edges in $S\ timesS^c$, we have that

$$v(f) = f(S, S^c) = C(S, S^c)$$

and so the theorem is proved as we already have the inequality.

If $t \in S$, then there is a positive residual path P. If we increase the capacity along the path by $c_f(P)$ as in Lemma 7.1.6 and define a new flow f' then $v(f') = v(f) + c_f(P)$. This will contradict the maximality of f if we show that it satisfies the capacity constraints. Trivially, the capacity constraint is satisfied for all $e, -e \notin P$ as the $f'(e) = f(e)$ for all $e, -e \notin P$. If $-e \in P$, then $f'(e) \leq f(e) \leq c(e)$. If $e \in P$, then $f'(e) = f(e) + c_f(p) \leq f(e) + c_f(e) = c(e)$ and so the capacity constraint is satisfied everywhere and the proof is complete.

Now let us assume that the capacity function is unbounded. But since the min-cut is finite, choose a cut $[A, A^c]$ such that $c(A, A^c) < \infty$. Define c' such that $c'(e) = c(e)\mathbf{1}[c(e) < \infty] + c(A, A^c)\mathbf{1}[c(e) = \infty]$. Observe that the min-cut under c' is same as that in c. Further, if a flow is feasible w.r.t c' then it is feasible w.r.t. c as well. Since c' is a bounded capacity function, we have a max-flow f such that $v(f) =$ min-cut under c'. But then, it also holds that $v(f) =$ min-cut under c and f is feasible under c as well. □

Exercise 7.1.8. *The last part of the proof shows that Lemma 7.1.5 holds under the assumption of finite min-cut alone i.e., the assumption of bounded capacity can be relaxed considerably. Give a direct proof of this without using Max-flow Min-cut theorem.*

Theorem 7.1.9. *Assume that the min-cut is finite. F-F Algorithm terminates if the capacities are integral and also gives that if capacities are integral, there is an integral maximal flow.*

Proof. If capacities are integral, the min-cut is integral and so is the max-flow. At evey step the F-F algorithm increases the flow strength by at least one as the residual capacity along any augmenting path is a positive integer . Since the max-flow is finite, in finitely many steps the algorithm outputs the max flow.

The algorithm starts with $f \equiv 0$ and at every step, the flow on any edge is increased/decreased by the residual capcity if the edge is in a f-augmenting path. Since the residual capacity is integral, the flow is always integral and so is the max-flow. □

Example 7.1.10. *The F-F algorithm need not terminate when capacities are non-integral and here is an example. See* `https://en.wikipedia.org/wiki/Ford_Fulkerson_algorithm#Non-terminating_example`

The definition of a flow can be extended to multiple sources and targets. Let $S \subset V, T \subset V$ be the sources and sinks respectively. The definition of flow can be modified by requiring the Kirchoff's node law to hold for all $x \notin S \cup T$ and positive output at S i.e., $\sum_{s \in S}(d^* f)(s) \geq 0$. An $S - T$ cut is $A \subset V$ such that $S \subset A, T \subset A^c$. We again have a max-flow min-cut theorem as follows.

Theorem 7.1.11 (Max-flow min-cut theorem for multiple sources and sinks.)**.**

$$\max\{v(f) : f \textit{ is a feasible } S-T\textit{-flow}\} = \inf_{A \subset V : S \subset A, T \subset A^c} C(A, A^c).$$

Proof. Let us again assume that the min-cut is finite i.e., there is no infinite capacity $S - T$ path. The trivial inequality follows as in the original theorem and also the fact that the supremum is maximum. To argue the equality, instead of repeating the proof, we shall use a reduction. Define G' as follows : $V' = V \cup \{a, b\}, E' = E \cup (a \times S) \cup (T \times b)$. Set the capacity of the new edges to be infinite. Note that any $S - T$ cut in G is a $a - b$ cut in G'. Further, any $a - b$ cut involving edges in $E' \setminus E$ has infinite capacity. Hence the min $a - b$ cut in G' and G are equal.

If f' is a $a - b$ flow in G', let f be the restriction of f' to G. Clearly f is skew-symmetric and satisfies Kirchoff's node law. We verify the last condition as follows : By Kirchoff's node law in G', we have that

$$0 = \sum_{s \in S}(d^* f')(s) = \sum_{e : e^- \in S, e^+ \in V} f(e) + \sum_{e : e^- \in S, e^+ = a} f(e) = \sum_{s \in S}(d^* f)(s) - (d^* f')(a).$$

Thus, $v(f) = \sum_{s \in S}(d^* f)(s) = (d^* f')(a) = v(f') \geq 0$ and so the max-flow in G' is equal to the max-flow in G. Hence the theorem follows from the single-source single-sink max-flow min-cut theorem. □

A more general version of max-flow min-cut theorem is as follows : Let $G = (V, E)$ be a directed graph, $s, t \in V$ and $c : E \to [0, \infty)$ be a capacity constraint function. Then $f : E \to [0, \infty)$ is a $s - t$ *flow satisfying capacity constraint* c if

1. $\sum_{e : e^- = x} f(e) = \sum_{e : e^+ = x} f(e)$ for all $x \neq s, t$. (*conservation of flow / Kirchoff's node law*).

2. For all $e \in E$, $f(e) \leq c(e)$. (*edge capacity constraint*).

3. $|f| := \sum_{e:e^- = s} f(e) - \sum_{e:e^+ = s} f(e) \geq 0$. (*flow strength is non-negative.*)

As before, a subset $S \subset V$ is a directed $s - t$ cut if $s \in S, t \notin S$. Further capcity of a cut is defined as $c(S) := \sum_{e \in E \cap (S \times S^c)} c(e)$.

Exercise* 7.1.12. *(Max-flow min-cut theorem for directed graphs.) Under the notation as above, we have that*

$$\max\{|f| : f \text{ is a } s - t \text{ flow satisfying capacity constraint } c\ \} = \min\{C(S) : S \subset E \text{ is an } s - t \text{ cut}\}.$$

Proof Sketch : An f-augmenting $x - y$ path is a path $P : x = x_0, \ldots, x_k = y$ such that for each $1 \leq i \leq k$ either (i) $c(x_{i-1}, x_i) - f(x_{i-1}, x_i) > 0$ or (ii) $f(x_i, x_{i-1}) > 0$. In simple words, either the 'forward' edges are not of full capacity or the 'backward' edges have positive flow. Define $c_f(x_{i-1}, x_i) = c(x_{i-1}, x_i) - f(x_{i-1}, x_i)$ in Case (i) and $c_f(x_{i-1}, x_i) = f(x_i, x_{i-1})$ in Case (ii). If both cases hold, pick one arbitarily as $c_f(x_{i-1}, x_i)$. As before set $c_f(P) = \min c_f(x_{i-1}, x_i)$. Show that the flow can be increased by adding $c_f(P)$ to the 'forward' edges and subtracting $c_f(P)$ from the 'backward' edges. Now use the proof idea of Theorem 7.1.3 to complete the proof.

A general version of max-flow min-cut theorem is stated in the exercises.

Exercise* 7.1.13. *Use max-flow min-cut theorem for directed graphs to show the undirected version.*

Exercise 7.1.14. *What is the equivalent of Ford-Fulkerson algorithm for flows on directed graphs.*

Exercise* 7.1.15. *Does the F-F algorithm terminate is the capacities are rational ?*

7.2 Vertex and edge connectivity

We say that a graph is k-vertex (resp. k-edge) connected if $|V| > k$ and $G - S$ is connected for any $S \subset V$ (resp. $|V| > 1$ and $S \subset E$) with $|S| < k$. Define $\kappa(G)$ (resp. $\lambda(G)$) is the largest k such that G is k-vertex (resp. k-edge) connected. We assume that $G = \emptyset$ is disconnected and so G is 0-vertex or 0-edge connected iff $G \neq \emptyset$. More easily, G is 1-vertex or 1-edge connected iff G is connected. Also, by the above convention $\lambda(K_1) = \kappa(K_1) = 0$.

Exercise 7.2.1. *Compute $\lambda(G), \kappa(G)$ for the well-known graphs such as complete graph, path graph, cycle graph, trees, Cayley graph, petersen graph et al.*

Theorem 7.2.2 (Whitney (1932a)). *Let $G \neq \emptyset$. Then $\kappa(G) \leq \lambda(G) \leq \delta(G)$.*

Proof. Removing edges on the vertex with minimum degree proves the first inequality. By definition $\kappa(G) \leq n - 1$ where $n = |V|$. Let $[S, S^c] := E \cap (S \times S^c)$ be the smallest edge-cut i.e., $\lambda(G) = |[S, S^c]|$ (see Exercise 7.2.4 as to why the smallest edge-cut will be of this form). Assume that $S, S^c \neq \emptyset$ i.e., G is connected. We will show that $\kappa(G) \leq |[S, S^c]|$.

If $E \cap (S \times S^c) = (S \times S^c)$ (i.e., every vertex in S is adjacent to every vertex in S^c), then $|[S, S^c]| = |S||S^c| = |S|(n - |S|) \geq n - 1 \geq \kappa(G)$.

Else choose $x \in S, y \notin S$ such that $x \nsim y$. Define $T := \{z \in S^c : z \sim x\} \cup \{z \in S - \{x\} : N(z) \cap S^c \neq \emptyset\}$. Every $x - y$ path has to pass through a vertex in T. In other words, there is no $x - y$ path in $G - T$. Thus, we have that $\kappa(G) \leq |T|$. Now we complete the proof by showing that $|T| \leq [S, S^c]$. For this observe, it suffices to observe that

$$\{(x, z) : z \in N(x) \cap S^c\} \cup \{(z, z') : z \in S - \{x\}, N(z) \cap S^c \neq \emptyset\} \subset E \cap (S \times S^c),$$

and the LHS has cardinality at least $|T|$. In other words, we have selected all the edges from x to S^c and for all other $z \in S - \{x\}$, we have selected one edge each to get at least $|T|$ edges. $\square$

7.2.1 Menger's theorem

Similar to Max-flow min-cut theorem, Menger's theorem also can be proven under differing frameworks. We shall prove two of them and leave the rest as exercises.

Theorem 7.2.3. *(Menger's theorem for edge-connectivity on digraphs) Let $s \neq t$ in a di-graph G. Then we have that*

$$\begin{aligned}\lambda(s,t) &:= \max\{k : \textit{there exist } k \textit{ edge-disjoint } s - t \textit{ directed paths }\} \\ &= \min\{|E'| : E' \textit{ is an } s - t \textit{ directed cut.}\} =: C(s,t),\end{aligned}$$

where E' is an $s - t$ directed cut if every directed path from s to t contains at least one edge in E'.

The proof proceeds by first showing that deletion of incoming edges at s and outgoing edges at t do not change the LHS or RHS. Then, we shall show that by setting $c \equiv 1$ on E, the LHS equals the max-flow and the RHS equals the min-cut thereby proving the theorem via max-flow min-cut theorem for digraphs.

Proof. Assume that $\lambda(s,t) > 0$ else the theorem holds trivially.

Observe that $\lambda(s,t), C(s,t)$ remain unchanged if we delete incoming edges at s and outgoing edges at t and so we shall assume that there are no incoming edges at s and outgoing edges at t.

Suppose there are k edge-disjoint paths. Then, any $s - t$ edge-cut E' has to contain at least one edge from each of the disjoint paths and hence $|E'| \geq k$. Thus, we trivially get that $\lambda(s,t) \leq C(s,t)$ as in the Max-flow Min-cut theorem.

Now we construct a directed network on G by assigning capacity 1 to every edge. Since $c \equiv 1$, the max-flow f is integral and further $f \in \{0, 1\}$. Trivially, if $v(f) \geq \lambda(s,t)$ as we can send a unit flow along each of the disjoint paths and strength and flow properties are preserved under addition. Consider the graph $G_f := (V, \{e : f(e) = 1\})$. Since $v(f) \geq 1$, there is a $s - t$ path P_1 in the graph G_f (see Exercise 7.2.5 for a more general claim). Now consider the graph $G_f - P_1$. This is nothing but the graph $G_{f_1} := (V, \{e : f_1(e) = 1\})$ where $f_1(e) = f(e) - \mathbf{1}[e \in P_1]$. Since $e \to \mathbf{1}[e \in P_1]$ is also a flow, by linearity f_1 is also a flow and further $v(f_1) = v(f) - 1$. Now, the same argument as above yields that there is a $s-t$ path P_2 in $G_f - P_1$ if $v(f_1) \geq 1$ and also P_2 is edge-disjoint of P_1. Repeating this argument, we can obtain $v(f)$ many edge-disjoint $s-t$ paths i.e., $\lambda(s,t) = MF(s,t) :=$ max-flow in G.

By our definition of capacity, note that $C(S, S^c) = |E \cap (S \times S^c)|$ for $s \in S, t \notin S$. Since

$$\{E' \subset E : G - E' \text{ disconnects } s,t\} \supset \{E \cap (S \times S^c) : s \in S, t \notin S, S \subset V\},$$

we have that $C(s,t) \leq MC(s,t) :=$ min-cut in G. Thus we have that

$$MF(s,t) = \lambda(s,t) \leq C(s,t) \leq MC(s,t).$$

Now, from the Max-flow Min-cut theorem for directed graphs (Exercise 7.1.12), we have that the inequalities are all equalities and the proof is complete. □

Exercise 7.2.4. *Give a direct proof that $C(s,t) = MC(s,t)$.*

Exercise* 7.2.5. *Prove the following more general version of the claim used in the above proof : If f is a $s-t$ flow with $v(f) > 0$, then there exists a directed $s-t$ path in the graph $G_f := \{e : f(e) > 0\}$.*

$S \subset V$ is an $s-t$ *vertex cut* if any directed path from s to t has a vertex in S.

Theorem 7.2.6. *(Menger's theorem for vertex-connectivity on digraphs) Let $s \neq t$ in a di-graph G such that $(s,t) \notin E$. Then we have that*

$$\max\{k : \textit{there exist } k \textit{ vertex-disjoint } s-t \textit{ directed paths }\} = \min\{|E'| : E' \textit{ is an } s-t \textit{ vertex cut.}\}$$

Proof. One can try to prove a Max-flow Min-cut theorem for vertex capacities and then use the same to prove Menger's theorem. But we will use a graph transformation to reduce the vertex-connectivity case to edge-connectivity case itself.

Consider the following enhanced graph G' : $V' := \{s,t\} \cup \{x_1, x_2 : x \in V - \{s,t\}\}$ and as for the edges E', $s \sim x_1$ if $s \sim x$ in G ; $x_2 \sim t$ if $x \sim t$ in G ; $x_2 \sim y_1$ if $x \sim y$ in G ; $x_1 \sim x_2$ for all x_1, x_2. Set $c(x_1, x_2) = 1$ for all x and $c \equiv \infty$ otherwise.

Note that any directed $s-t$ path in G' is of the form $sa_1a_2b_1b_2 \ldots z_1z_2t$ where $sab \ldots zt$ is the corresponding directed $s-t$ path in G. Further if $P_1, P_2, \ldots, P_k$ are vertex-disjoint $s-t$ paths in G, the corresponding paths in G' are trivially edge-disjoint as well. Conversely, let P_1', P_2' be two edge disjoint paths in G', say $sa_1a_2b_1b_2 \ldots z_1z_2t$ and $sa_1'a_2'b_1'b_2' \ldots z_1'z_2't$ respectively. Then edge-disjointness implies that $(a_1, a_2) \neq (a_1', a_2'), \ldots, (z_1', z_2')$ and similarly (b_1, b_2) and so on. Thus, the corresponding paths P_1, P_2 in G are vertex-disjoint. Thus we have that

$$\lambda'(s,t) := \text{ max edge-disjoint } s-t \text{ paths in } G' = \text{ max vertex-disjoint } s-t \text{ paths in } G =: \kappa(s,t).$$

If $S \subset V - \{s,t\}$ is a $s-t$ vertex cut in G, then $\{(x_1,x_2) : x \in S\}$ is an $s-t$ edge-cut in G'. In other words, if $C'(s,t)$ is the min edge-cut in G' and $VC(s,t)$ is the min vertex-cut in G, then we have from the above inclusion that $C'(s,t) \subset VC(s,t)$. Consider an edge-cut $E'' \subset E'$ in G' such that $E" \subsetneq \{(x_1,x_2) : x \in V - \{s,t\}\}$. Then $C(E'') = \infty$ and cannot be the min edge-cut in G'. Thus the min edge-cut in G' comes from edges in $\{(x_1,x_2) : x \in V - \{s,t\}\}$ which in turn arise via $s-t$ vertex cuts in G. Hence $C'(s,t) = VC(s,t)$.

Since $\lambda'(s,t) = \kappa(s,t)$ and $C'(s,t) = VC(s,t)$, Menger's theorem for vertex connectivity follows from the Max-flow Min-cut theorem for directed graphs (Exercise 7.1.12) or Menger's theorem for edge connectivity (Theorem 7.2.3).

$\square$

Exercise* 7.2.7. *(Menger's theorem for edge connectivity ; Menger (1927)) Let G be a finite undirected graph and u and v two distinct vertices. Then the size of the minimum edge cut for v and v (the minimum (u,v)-edge cut) is equal to the maximum number of pairwise edge-disjoint paths from u to v.*

Exercise* 7.2.8. *(Menger's theorem for vertex connectivity ; Menger (1927)) Let G be a finite undirected graph, and let u and v be nonadjacent vertices in G. Then, the maximum number of pairwise-internally-disjoint (u,v)-paths in G equals the minimum number of vertices from $V(G) - u, v$ whose deletion separates u and v.*

See `https://en.wikipedia.org/wiki/Menger's_theorem` for versions of Menger's theorem and variants.

7.3 ***Some applications***

We can prove Hall's marriage theorem as well as Konig-Egervary theorem using max-flow min-cut or Menger's theorem. See exercises in the next section. Also, see Section 6.6.

The max-flow min-cut theorem can be derived from a more powerful theorem called the strong duality theorem in linear programming (see `https://en.wikipedia.org/wiki/Max-flow_min-cut_theorem#Linear_program_formulation`). This latter theorem for example can be used to prove the Monge-Kantorovich duality theorem in Optimal transport theory (`https://en.wikipedia.org/wiki/Transportation_theory_(mathematics)`).

7.4 Exercises

1. Let G be a graph and $S, T \subset V(G)$ such that $S \cap T = \emptyset$. Formulate an appropriate notion of a vertex cut and prove a version of the vertex form of the Menger's theorem for the following two scenarios.

(a) Consider the maximum number of disjoint paths from S to T such that the paths do not intersect even at S and T.

(b) Consider the maximum number of disjoint paths from S to T such that the paths are allowed to intersect at S or T.

2. Show that edge connectivity and vertex connectivity are equal if $\Delta(G) \leq 3$.

3. If G is a connected graph and for every edge e, there are cycles C_1 and C_2 such that $E(C_1) \cap E(C_2) = \{e\}$ then G is 3-edge connected.

4. What is the vertex and edge connectivity of the Petersen graph ?

5. Let F be a non-empty set of edges in G. Prove that F is an edge-cut of the form $F = E \cap (S \times S^c)$ for some $S \subset V$ in G iff F contains an even number of edges from every cycle C.

6. If G is a k-connected graph and a new graph G' is formed by adding a new vertex y with at least k neighbours in G, then G' is also k-connected.

7. Let G be a graph with at least 3 vertices. Show that G is 2-connected $\Leftrightarrow$ G is connected and has no cut-vertex $\Leftrightarrow$ for all $x, y \in V(G)$ there exists a cycle through x and y $\Leftrightarrow$ $\delta(G) \geq 1$ and every pair of edges lies in a common cyle.

8. Consider a directed graph $G = (V, E)$ and $s, t \in V$. Further assume that there are no incoming edges at s or no out-going edges at t. An *elementary $s - t$ flow* is a flow f which is obtained by assigning a constant positive value a to the edges on a simple directed $s - t$ path and 0 to all other edges. Show that every flow is a sum of elementary flows and a flow of strength 0.

9. Consider a directed graph $G = (V, E)$ and $s, t \in V$. Further assume that there are no incoming edges at s or no out-going edges at t. If f is an integral flow of strength k, show that there exist k directed paths $p_1, \ldots, p_k$ such that for all $e \in E$, $|\{p_i : e \in p_i\}| \leq f(e)$.

10. * (Generalized max-flow min-cut theorem :) Let $G = (V, E)$ be a directed graph, $s, t \in V$ and $c : (V - \{s, t\}) \cup E \to [0, \infty)$ be a capacity constraint function. Then $f : E \to [0, \infty)$ is a $s - t$ *flow satisfying capacity constraint* c if

(a) $\sum_{e:e^- = x} f(e) = \sum_{e:e^+ = x} f(e)$ for all $x \neq s, t$. (*conservation of flow / Kirchoff's node law*).

(b) For all $e \in E$, $f(e) \leq c(e)$. (*edge capacity constraint*).

(c) $\sum_{e:e^+ = x} f(e) \leq c(x)$ for all $x \neq s, t$. (*vertex capacity constraint.*)

(d) $|f| := \sum_{e:e^- = s} f(e) - \sum_{e:e^+ = s} f(e) \geq 0$. (*flow strength is non-negative.*)

A subset $S \subset V \cup E$ is an $s - t$ cut if every path from s to t passes through either a vertex or an edge in S. Further capcity of a cut is defined as $c(S) := \sum_{v \in S} c(v) + \sum_{e \in S} c(e)$.

Show that

$$\max\{|f| : \text{f is a } s-t \text{ flow satisfying capacity constraint } c\ \} = \min\{C(S) : S \subset V \cup E \text{ is an } s-t \text{ cut}\}.$$

11. ***** Can you state the defective version of Hall's marriage theorem as a max flow-min cut theorem ?

Question 7.4.1. *Can you state Tutte's* 1*-factor theorem also as a max-flow min-cut theorem ?*

Chapter 8

Chromatic number and polynomials

8.1 Graph coloring

Definition 8.1.1. *(Coloring of a graph) A graph is k-colorable if* $\exists\ c : V \to [k]$ *such that* $c(u) \neq c(v)$ *if* $(u, v) \in E$. *The* chromatic number $\chi(G)$ *is defined as*

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

Trivially, we have that $cl(G) \leq \chi(G) \leq \Delta(G) + 1$ where $cl(G)$ is the size of the largest clique (complete subgraph) in G. Further, a graph is k-colorable iff we can partition into k independent sets.

Exercise 8.1.2. *G is k-colorable if there exists an homomorphism from G to K_k.*

8.2 Chromatic Polynomials

Let G be a graph on n vertices. Let $P(G, q), q \in \mathbb{N}_*$ be the number of ways of coloring (properly) a graph with q colours. More formally,

$$P(G, q) = |\{c : V \to [q] : c(u) \neq c(v) \forall u \sim v\}|.$$

We will now show that this is a monic polynomial of degree n and other interesting properties.

Let $G - e$ denote the graph with the edge e removed. Let G/e denote the graph with e contracted i.e., if $e = (v, w)$ then we replace the vertices v, w with v' and edges $v \sim x, w \sim y$ with $v' \sim x, v' \sim y$. Observe that G/e can be a multigraph. But this will not matter to us as proper colorings of a multigraph and its corresponding simple graph are the same.

Proposition 8.2.1. *For any edge* $e = (v, w) \in G$,

$$P(G, q) := P(G - e, q) - P(G/e, q), q \in \mathbb{N}.$$

Proof. The proof follows by two simple observations that any proper q-coloring of G is a proper q-coloring of $G-e$ in which v and w receive distinct colours and any proper q-coloring of $G-e$ in which v and w receive same colours is a proper q-coloring of G/e. □

If $E = \emptyset$, trivially $P(G,q) = q^n$. Now using Proposition 8.2.1, we can inductively show that $P(G,q)$ is a polynomial in q. Hence, we shall define $P(G,x), x \in \mathbb{R}$ to be the polynomial such that $P(G,q)$ is the number of proper q-colorings of G for $q \in \mathbb{N}_*$. Here are some very important proeprties.

Lemma 8.2.2. *1. $P(G,x)$ is a monic polynomial of degree n.*

2. $P(G,x)$ is the unique polynomial such that $P(G,q)$ is the number of proper q-colorings of G for $q \in \mathbb{N}$.

3. $\chi(G) = \min\{k \in \mathbb{N} : P(G,k) > 0\}$.

Now, let us assume that $P(G,x) = \sum_{i=1}^n a_i x^i$ (as $a_0 = 0$ trivially). Then we have that

4. $\sum_{i=1}^n a_i = 0$ or $P(G,x) = x^n$.

5. $a_{n-1} = -|E|$.

6. $a_{n-i} = (-1)^i |a_{n-i}|$.

Proof. All the proofs will follow from deletion-contraction principle and induction. For abbreviation, we call this DCI argument in this section.

1. Suppose $P(G-e,x)$ is a monic polynomial of degree n and $P(G/e,x)$ is a monic polynomial of degree $n-1$, then by DC principle $P(G,x)$ is a monic polynomial of degree n. Now the proof can be completed by induction.

2. The proof follows because two polynomials are equal if they agree at $1,\ldots,n,\ldots,$.

3. This follows trivially by the two definitions.

4,5 and 6 : We will apply the DCI argument again and prove (4),(5) and (6) together. The three identities can be verified for $E = \emptyset$ or G has 1 or 2 vertices easily. Let

$$P(G-e,x) = \sum_{i=0}^{n} (-1)^i |b_{n-i}| x^{n-i}, \; P(G/e,x) = \sum_{i=0}^{n-1} (-1)^i |c_{n-1-i}| x^{n-1-i} = \sum_{i=1}^{n} (-1)^{i-1} |c_{n-i}| x^{n-i}.$$

Then by DC principle we have that

$$P(G,x) = |b_n| x^n + \sum_{i=1}^{n} (-1)^i [|b_{n-i}| + |c_{n-i}|] x^{n-i}.$$

As for (4), observe that $\sum_{i=1}^n a_i = P(G,1) =$ no. of proper 1-coloring of G.If $|E| \neq \emptyset$, there exists no proper 1-coloring of G and so $P(G,1) = 0$. Else $P(G,x) = x^n$ as required.

(6) follows easily as $a_{n-i} = (-1)^i[|b_{n-i}|+|c_{n-i}|]$. As for (5), observe by monicity of the chromatic polynomial, induction and DC principle that that

$$|E(G)| = |E(G-e)| + 1 = -b_{n-1} + c_{n-1} = |b_{n-1}| + |c_{n-1}| = -a_{n-1},$$

where in the second equality we have used (6) for b_{n-1}.

□

Lemma 8.2.3. *If G has k-components $G_1, \ldots, G_k$ then*

$$P(G,x) = \prod_{i=1}^{k} P(G_i, x)$$

and further $a_0 = \ldots = a_{k-1} = 0, |a_k| > 0$.

Proof. First equality follows trivially because any combination of q-colorings of each component gives a q-coloring of the full graph G i.e.,

$$P(G,q) = \prod_{i=1}^{k} P(G_i, q), q \in \mathbb{N}_*.$$

Since $deg(P(G_i, x)) \geq 1$, we have that $a_0 = \ldots = a_{k-1} = 0, |a_k| > 0$. □

Theorem 8.2.4. $P(G,x) = \sum_{X \subset E} (-1)^{|X|} x^{\beta_0(X)}$ *where* $\beta_0(X) = \beta_0((V,X))$.

Proof. Enough to show that $P(G,q) = \sum_{X \subset E} (-1)^{|X|} q^{\beta_0(X)}$ for $q \in \mathbb{N}_*$. Let $c : V \to [q]$ be an *improper coloring* if $c(u) = c(v)$ for some $u \sim v$. Let IC be the set of all improper colorings. For $e = (u,v)$ define $B_e := \{c \in IC : c(u) = c(v)\}$. Then we have that

$$\begin{aligned} P(G,q) &= q^n - |IC| = q^n - |\cup_{e \in E} B_e| \\ &= q^n - \sum_{\emptyset \supsetneq X \subset E} (-1)^{|X|-1} |\cap_{e \in X} B_e| \\ &= q^n + \sum_{\emptyset \supsetneq X \subset E} (-1)^{|X|} |\cap_{e \in X} B_e|. \end{aligned}$$

To complete the proof, enough to prove that for $\emptyset \supsetneq X \subset E$, $|\cap_{e \in X} B_e| = q^{\beta_0(X)}$. Suppose $X = \{e_1, \ldots, e_k\}$ where $e_i = (u_i, v_i), 1 \leq i \leq k$. Let $C_1, \ldots, C_m$ be the components in (V,X) i.e., $m = \beta_0(X)$. It is possible that $u_i = u_j$ or v_j. Thus, we have that

$$\begin{aligned} \cap_{e \in X} B_e &= \{c : V \to [q] : c(u_i) = c(v_i), 1 \leq i \leq k\} \\ &= \{c : V \to [q] : c \text{ is a constant on each } C_i, 1 \leq i \leq m\} \end{aligned}$$

The latter is obtained by choosing one color for each component $C_1, \dots, C_m$ and so its cardinality if q^m as required. $\square$

8.3 Exercises

1. Find the chromatic polynomial of complete graph K_n, cycle C_n, wheel graph W_n (i.e., the graph obtained by adding a new vertex to C_n and connecting it to all the n vertices of C_n), path graph P_n and the Petersen graph.

2. Show that the coefficient of x^{n-2} in the chromatic polynomial of a n vertex graph is $\frac{m(m-1)}{2} - T$ where T is the number of triangles and m is the number of edges.

3. Show that a graph with n vertices is a tree iff $P(G, x) = x(x-1)^{n-1}$.

8.4 ***Read's conjecture and Matroids***

For more details, refer to `https://en.wikipedia.org/wiki/Chromatic_polynomial#Properties`, `http://www-math.ucdenver.edu/~wcherowi/courses/m4408/gtln6.htm` and `http://matroidunion.org/?p=1301`.

Question 8.4.1 (Read's conjecture, 1968). *$|a_0|, \dots, |a_n|$ form a log-concave sequence i.e., $|a_i|^2 \geq |a_{i-1}||a_{i+1}|$.*

Verify the conjecture in the cases you can compute the coefficients such as K_n, C_n, P_n et al.

This was solved recently by June Huh (in his 1st year of Ph.D.) in 2012 using ideas from algebraic geometry.

The crucial DC principle holds for more general combinatorial objects called Matroids (`https://en.wikipedia.org/wiki/Matroid`) and analogous to a chromatic polynomial, one can define what is called the characteristic polynomial of a matroid . The characteristic polynomial of a matroid is defined analogously to the identity in Theorem 8.2.4. Read's conjecture was generalized to characteristic polynomial of matroids and known as Rota-Welsh conjecture . This was proved in 2015 by Karim Adiprasito, June Huh, and Eric Katz. An accessible exposition of these ideas can be found in the blog Matt Baker ([Baker 2015]) and a little more technical account in his survey ([Baker 2018]).

Chapter 9

Graphs and matrices

See [Bapat 2010] for recollection of some basic linear algebra facts used.

9.1 Incidence matrix and connectivity

Let $n = |V|, m = |E|$. We shall assign an orientation to every edge $e \in E$ and call $\overrightarrow{E}$ this set of oriented edges. The orientation can be arbitrary.

Definition 9.1.1 (Incidence matrix). *∂_1 is a $n \times m$ matrix with $\partial_1(i,j) = \chi_{e_j}(v_i)$ where $\chi_e(v) = \mathbf{1}[v = e^+] - \mathbf{1}[v = e^-]$ for $e = (e^-, e^+) \in \overrightarrow{E}, v \in V$.*

$$\begin{array}{cc} \partial_1(i,j) & \begin{array}{cccc} e_1 & e_2 & \dots & e_m \end{array} \\ \begin{array}{c} v_1 \\ v_2 \\ \vdots \\ v_n \end{array} & \begin{pmatrix} -1 & 0 & \dots & 0 \\ 1 & -1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{pmatrix} \end{array}$$

We set $\partial_0 = [1, \dots, 1]$ as a $1 \times n$-matrix.

Remark 9.1.2. *We consider the matrices as a real matrices and our matrix algebra will be with respect to $\mathbb{R}$. One can consider any other field $\mathbb{F}$ instead of $\mathbb{R}$ as well.*

Further set $C_j = \{0\}, j \leq -2, C_{-1} = \mathbb{R}$, $C_0 = \{\sum_{i=1}^n a_i v_i : a_i \in \mathbb{R}\}, C_1 == \{\sum_{i=1}^m a_i e_i : a_i \in \mathbb{R}, e_i \in \overrightarrow{E}\}$, where the sums are to be considered as a formal sum. Observe that C_0, C_1 are$\mathbb{R}$- vector spaces and note that $C_0 \cong \mathbb{R}^V, C_1 \cong \mathbb{R}^{\overrightarrow{E}}$. The canonical basis for C_0 is $\{v_1, \dots, v_n\}$ and for C_1 is $\{e_i : e_i \in \overrightarrow{E}\}$. Here, we think of $v_1 = v_1 + 0v_2 + \dots + 0v_n, e_1 = e_1 + 0e_2 + \dots + 0e_m$ and similarly for other v_i's and e_j's.

We shall view ∂_0, ∂_1 as linear maps in the following sense :

$$\partial_0 : C_0 \to C_{-1}, \;\; z = \sum_{i=1}^n a_i v_i \to \sum_i a_i = \partial_0[a_1, \dots, a_n]^t,$$

$$\partial_1 : C_1 \to C_0, \;\; z = \sum_{i=1}^{m} a_i e_i \to \sum_i a_i(e_i^+ - e_i^-).$$

Suppose we represent $\partial_1 z = \sum_{i=1}^n b_i v_i$ then $(b_1, \ldots, b_n)^t = \partial_1[a_1, \ldots, a_m]^t$. Further if the columns of ∂_1 matrix are $C_1, \ldots, C_m$ then $(b_1, \ldots, b_n)^t = \sum_{i=1}^m a_i C_i$.

Assume that $G \neq \emptyset$. We set $B_i = Im(\partial_{i+1}), Z_i = Ker(\partial_i)$.

Remark 9.1.3. *The above abstraction of representing C_0, C_1 as formal sums seems a little unnecessary as we could have also represented C_0, C_1 as functions from $V \to \mathbb{R}$ and $\overrightarrow{E} \to \mathbb{R}$ respectively. However, this representation will be notationally convenient and is borrowed from algebraic topology where one considers richer structures giving rise to further vector spaces such as $C_2, C_3, \ldots$. Further, if we choose $\mathbb{Z}$-coefficients instead of $\mathbb{R}$-coefficients, C_0, C_1 are modules and not vector spaces. Thus, the above representation allows us to carry over similar algebraic operations even in such a case. As an exercise, the algebraically inclined students can try to repeat the proofs with $\mathbb{Z}$-coefficients. See [Edelsbrunner 2010] for an accessible introduction to algebraic topology and [Munkres 2018] for more details.*

Trivially, we have that $B_{-1} = C_{-1} = \mathbb{R}$. Thus by the fundamental theorem of algebra, we have that $r(Z_0) = n - 1$ where by $r(.)$ we denote the rank/dimension of a vector space. It is easy to see that $v_i - v_j \in Z_0, \forall i \neq j$. Further, we have that $v_1 - v_i, i = 2, \ldots, n$ are linearly independent and hence form a basis for Z_0. Observe that for $z = \sum_{i=1}^m b_i e_i$,

$$\partial_0 \partial_1 z = \partial_0 (\sum_{i=1}^{m} b_i(e_i^+ - e_i^-)) = \sum_{i=1}^{m} b_i(\partial_0 e_i^+ - \partial_0 e_i^-) = 0,$$

where in the first and third equalities we have used the definition of ∂_1, ∂_0 respectively and in the second equality, the linearity of ∂_0. Thus, we have that $\partial_0 \partial_1 = 0$ and in other words $B_0 \subset Z_0$. The same can also be deduced by noting that as matrix product $\partial_0 \partial_1$ is nothing but sum of rows of ∂_1 and hence 0.

Now, if we can understand $r(B_0)$ then we can understand all the ranks involved. Observe that B_0 is nothing but the column space of of ∂_1 i.e., denoting the columns of the matrix ∂_1 by $C_1, \ldots, C_m$, we have that

$$B_0 \cong \{\sum_{i=1}^{m} a_i C_i : a_i \in \mathbb{R}\}.$$

Hence $r(B_0)$ is nothing but the maximum number of linearly independent column vectors.

Suppose $e_1, \ldots, e_k$ denote the edges corresponding to the linearly independent columns. Assume that the subgraph $H = e_1 \cup \ldots \cup e_k$ is connected. Let $V(H) = \{v_1, \ldots, v_l\}$. Consider the incidence matrix restricted to $\{v_1, \ldots, v_l\} \times \{e_1, \ldots, e_k\}$. Call it ∂_1'. Since the non-trivial entries of e_i's are on $v_1, \ldots, v_l$, the columns of ∂_1' are also linearly independent and so the column rank of $\partial_1' = k \geq l - 1$. Let $R_1, \ldots, R_l$ be the rows of the matrix ∂_1'. Since every column has exactly one $+1$ and one -1, we have that $\sum_i R_i = 0$ and thus the row rank of $\partial_1' \leq l - 1$. Since the row rank = column rank, we have that $k = l - 1$ i.e., H is a tree.

Thus, we have proved that if $e_1, \ldots, e_k$ denote the edges corresponding to the linearly independent columns, then every component of $e_1 \cup \ldots \cup e_k$ is a tree i.e., $e_1 \cup \ldots \cup e_k$ is a forest or acyclic subgraph. We show the converse and determine the basis for B_0.

Proposition 9.1.4. *Let C_i be the columns corresponding to e_i in ∂_1. Then $e_1 \cup \ldots \cup e_k$ is acyclic iff $C_1, \ldots, C_k$ are linearly independent.*

Proof. We have shown the 'if' part. We will show the 'only if' part. Suppose that $H = e_1 \cup \ldots \cup e_k$ is acyclic. We will show that $C_1, \ldots, C_k$ are linearly independent by induction. The conclusion holds trivially for $k = 1$. Suppose it holds for all $l < k$. WLOG assume that e_k is a leaf edge with $deg_H(e_k^+) = 1$. Thus $e_k^+ \notin e_1 \cup \ldots \cup e_{k-1}$. If $\sum_{i=1}^{k} a_i C_i = 0$ then since $e_k^+ \in e_k$ only, we have that $a_k = 0$. Thus $\sum_{i=1}^{k-1} a_i C_i = 0$ and since $e_1 \cup \ldots \cup e_{k-1}$ are acyclic and hence by induction hypothesis, we have that $C_1, \ldots, C_{k-1}$ are linearly independent i.e., $a_i = 0, 1 \leq i \leq k$. □

From the previous proposition, the following theorem follows easily.

Theorem 9.1.5. *1. Let $e_1 \cup \ldots \cup e_k$ be the (maximal) spanning forest in G. Then*

$$B_0 = \{\sum_{i=1}^{k} a_i \partial_1(e_i) : a_i \in \mathbb{R}\}.$$

2. $r(B_0) = n - \beta_0(G), r(Z_0) - r(B_0) = \beta_0(G) - 1$.

3. G is connected iff $B_0 = Z_0$.

By the rank-nullity theorem, we have that $r(Z_1) = m - n + \beta_0(G)$ which is nothing but the number of excess edges i.e., edges added after forming a (maximal) spanning forest. Let $e_1 e_2 \ldots, e_k$ form a cycle, denoting the reversal of an edge by $\widehat{e}$, set

$$z_c := \sum_{i=1}^{k} \mathbf{1}[e_i \in \overrightarrow{E}]e_i - \mathbf{1}[\widehat{e_i} \in \overrightarrow{e_i}]\widehat{e_i}.$$

Thus, it is easy to verify that $\partial_1 z_c = 0$ and hence $z_c \in Z_1$. We can extend this further.

Theorem 9.1.6. *Let F be a maximal spanning forest in G and $e_1, \ldots, e_l \in \overrightarrow{E} - E(F)$ where $l = m - n + \beta_0(G)$. Let $C_1, \ldots, C_l$ be the cycles in $F \cup e_1, \ldots, F \cup e_l$ respectively. Then $z_{C_1}, \ldots, z_{C_l}$ form a basis for Z_1.*

Remark 9.1.7. *Another challenging exercise : In terms of matrices, the linear transformation ∂_0, ∂_1 represent right multiplication i.e., we have that*

$$\{0\} \stackrel{\partial_2}{\to} C_1 \stackrel{\partial_1}{\to} C_0 \stackrel{\partial_0}{\to} \mathbb{R} \stackrel{\partial_{-1}}{\to} \{0\}.$$

But we could have considered left multiplication and in which case we will have linear transformations δ_i's as follows :

$$\{0\} \overset{\delta_2}{\leftarrow} C_1 \overset{\delta_1}{\leftarrow} C_0 \overset{\delta_0}{\leftarrow} \mathbb{R} \overset{\delta_{-1}}{\leftarrow} \{0\}.$$

Can you compute the ranks and characterize connectivity as we did above with ∂_0, ∂_1's ?

See [Bapat 2010, Chapter 2] for more on incidence matrices.

We shall define the $0-1$ incidence matrix M () as follows : $M(i,j) = 1[v_i \in e_j]$. In other words, M is the unoriented incidence matrix. The following result explains the importance of considering orientations in the incidence matrix.

Theorem 9.1.8. *Let G be a connected graph. Then $r(M) = n-1$ if G is bi-partite and $r(M) = n$ otherwise.*

See [Bapat 2010, Lemma 2.17] for a proof.

9.1.1 Laplacian Matrix

The Laplacian matrix plays a fundamental role in many topics within probability, graph theory, algebraic topology and analysis. See Section 9.4 for references.

Let $D = diag[deg(1), \ldots, deg(n)]$ be the diagonal matrix with entries as the degrees of vertices. Recall that A is the adjacency matrix defined in Definition 3.4.1. The Laplacian matrix is defined as $L = D - A$. Verify that $L = \partial_1\partial_1^T$ using the matrix representation of ∂_1 and its transpose. Observe that even though ∂_1 is oriented, L is unoriented. Further, it is easy to see that if we think of ∂_1 as linear transformations as in the previous sections, we have that

$$L : C_0 \to C_0; : \sum_{i=1}^{n} a_i v_i \mapsto \sum_{i=1}^{n}(a_i deg(v_i) - \sum_{j: v_j \sim v_i} a_j) v_i.$$

For purposes of our analysis, we will view L primarily as a matrix. Firstly observe that $r(L) = r(\partial_1\partial_1^t) \leq r(\partial_1) \wedge r(\partial_1^t) = r(\partial_1)$. Trivially, we have that $Ker\partial_1^t \subset KerL$.

We set the notation $< x, y >= x.y = xy^t = yx^t$ for $x, y \in \mathbb{R}^n$ and $\|x\|^2 =< x, x >= \sum_{i=1}^n x_i^2$. Observe that for $x \in \mathbb{R}^n$, we have that

$$x^t L x = x^t \partial_1 \partial_1^t x =< \partial_1^t x, \partial_1^t x >= \|\partial_1^t x\|^2. \tag{9.1}$$

Hence, if $x \in KerL$ then by the above equation $\partial_1^t x = 0$ i.e., $x \in Ker\partial_1^t$ i.e., $KerL = Ker\partial_1^t$ by our earlier observation. Hence $n(L) := r(KerL) = r(Ker\partial_1^t) = n - r(\partial_1) = \beta_0(G)$. Since $n(L)$ is the multiplicity of the eigenvalue 0, this shows that the multiplicity of 0 is $\beta_0(G)$.

From (9.1), we also have that L is a positivie semi-definite matrix and hence all its eigenvalues are non-negative. Denote the eigenvalues by $\lambda_1 \leq \lambda_2 \ldots \leq \lambda_n$. Summarizing our above conclusions, here is the theorem.

Theorem 9.1.9. *Let $l = \beta_0(G)$. We have that $0 = \lambda_1 = \ldots = \lambda_l < \lambda_{l+1} \leq \ldots \lambda_n$. Thus, G is connected iff $\lambda_2 > 0$.*

Observe that $\partial_1^t x = (x_i - x_j)_{(v_i,v_j)\in\overrightarrow{E}}$.

Lemma 9.1.10. *Let $C_1, \ldots, C_l$ be the components of G. Define $x^1, \ldots, x^l$ as vectors that are constant on each of the components respectively i.e., $x^i_k = \mathbf{1}[k \in C_i]$. Then, we have that $x^1, \ldots, x^l$ form a basis for $KerL$.*

Proof. From the observation before the lemma, we have that $\|\partial_1^t x\|^2 = \sum_{v_i \sim v_j}(x_i - x_j)^2$ and from (9.1), we know that if $Lx = 0$, $\partial_1^t x = 0$ and hence $x_i = x_j$ for all $v_i \sim v_j$. This implies that x is a constant on each component and hence $x^1, \ldots, x^l \in KerL$. Since $r(KerL) = l$, it suffices to show that $x^1, \ldots, x^l$ are linearly independent. This follow easily as the vectors are supported on disjoint sets of vertices. $\square$

9.1.2 Spanning trees and Laplacian

We recall the famed Cauchy-Binet formula. For an $n \times m$ matrix A and a $m \times n$ matrix B with $n \leq m$, we have that

$$det(AB) = \sum_{S\subset[m],|S|=n} det(A[[n]|S])det(B[S|[n]]),$$

where $A[T|S]$ refers to the matrix A restricted to rows in T and columns in S. Let $W = V - \{1\}$ and we can apply the Cauchy-Binet formula to $L[W|W]$. As $L = \partial_1\partial_1^t$, we have that $L[W|W] = \partial_1[W|\overrightarrow{E}]\partial_1^t[\overrightarrow{E}|W]$.

$$det(L[W|W]) = \sum_{S\subset\overrightarrow{E}:|S|=n-1} det(\partial_1[W|S])det(\partial_1^t[S|W]) = \sum_{S\subset E:|S|=n-1} det(\partial_1[W|S])^2,$$

where the last equality is due to the fact that $\partial_1[W|S] = \partial_1^t[S|W]$. Since the sum of rows in $\partial_1[V|S]$ is 0, $r(\partial_1[W|S]) = r(\partial_1[V|S])$. Hence, if G is not connected, then $r(\partial_1[W|\overrightarrow{E}]) < n-1$ and $det(L[W|W]) = 0$. Thus, if $|E| < n-1$, the LHS and the RHS in the above equation are zero. Assume that G is connected.

Fix $S \subset \overrightarrow{E} : |S| = n-1$. Then $det(\partial_1[W|S]) \neq 0$ iff $\partial_1[W|S]$ is non-singular iff $\partial_1[W|S]$ has full column rank iff $r(\partial_1[W|S]) = r(\partial_1[V|S]) = n-1$ iff S is acyclic iff S is a spanning tree in G. Thus, we have that

$$det(L[W|W]) = \sum_{S\subset\overrightarrow{E},\text{spanning tree}} det(\partial_1[W|S])^2.$$

Suppose we show that $det(\partial_1[W|S]) \in \{0, -1, +1\}$, then we have that

$$det(L[W|W]) = |\text{Spanning trees of } G|. \tag{9.2}$$

The following lemma completes the proof.

Lemma 9.1.11. *For $W \subset V, S \subset \overrightarrow{E}$ such that $|W| = |S|$, we have that $det(\partial_1[W|S]) \in \{0, -1, +1\}$.*

Proof. We will prove by induction on $k = |W| = |S|$. Let $B = \partial_1[W|S]$. The case of $k = 1$ is trivial as the entries are $0, -1, +1$. Assume that the theorem holds for all $l < k$ for $k \geq 2$.

If each column of B has both a $+1$ and a -1 entry, then the sum of rows is 0 and $detB = 0$. If there is a column of 0's in B, then $detB = 0$ again. Hence, there is a column with only a single non-zero entry in B. Suppose this is the (w, s)-th entry. Then, we have that $detB = B_{w,s}det(\partial_1[W - \{w\}|S - \{s\}])$. By induction, the latter is $0, -1, +1$ and since $B_{w,s} = +1, -1$, we obtain the desiered conclusion. □

Corollary 9.1.12. *Let $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ be the eigenvalues of L. Then the number of spanning trees of G is $\frac{\prod_{i=2}^n \lambda_i}{n}$.*

Proof. Let ST be the number of spanning trees of G.

$$n \times ST = \sum_{W \subset [n], |W| = n-1} det(L[W|W]) = \sum_{W \subset [n], |W| = n-1} \prod_{i \in W} \lambda_i,$$

as the sum of $k$the principal minors of a matrix is equal to the sum of products of k eigenvalues. □

Theorem 9.1.13. *Let $W \subset [n]$ and $|W| = n - k$. Then $det(L[W|W])$ is the number of spanning forests in G with k components and each of the elements of W^c being in a distinct component.*

Proof. Let $W^c = \{w_1, \ldots, w_k\}$. Again by Cauchy-Binet, we have that

$$det(L[W|W]) = \sum_{S \subset E : |S| = n-k} det(\partial_1[W|S])^2$$

From Lemma 9.1.11, we know that $det(\partial_1[W|S]) = 0, -1, +1$ and hence it is enough to show that $\partial_1[W|S]$ is non-singular iff the edges of S forms a forest with k components with each w_i in a distinct component.

Let the edges of S forms a forest with k components with each w_i in a distinct component, say D_i. Let the edges in the ith component be S_i. Set $W_i = D_i - w_i$. We have that $\cup W_i = W$ and $\cup S_i = S$. Thus $\partial_1[W|S]$ is a block matrix with blocks $\partial_1[W_i|S_i], i = 1, \ldots, k$. By the proof of Corollary 9.1.12, we have that $det(\partial_1[W_i|S_i]) \in \{-1, +1\}$ as (D_i, S_i) forms a trees and hence $det(\partial_1[W|S]) \in \{-1, +1\}$

If $\partial_1[W|S]$ is non-singular, then the columns of S in $\partial_1[V|S]$ are linearly independent and hence the edges in S form a forest and they have k components as $|S| = n - k$. **to be completed** □

9.2 More properties of Adjacency Matrix

A subgraph H of G is said to be *elementary* if every component is a cycle or an edge. Denote by $c(H)$ and $c_1(H)$ the number of cycle components and edge components in a subgraph H respectively.

Theorem 9.2.1.
$$detA = \sum_{H:H \text{ is a sp. el. subgraph}} (-1)^{n-c_1(H)-c(H)} 2^{c(H)}.$$

Proof. Denoting the symmetric group of permutations by S_n, we have that

$$detA = \sum_{\pi \in S_n} (-1)^{n-|\pi|} A_\pi ; A_\pi := \prod_{i=1}^{n} A_{i,\pi(i)},$$

and $|\pi|$ is the number of cycles in the cycle decomposition of π. Define the graph $H_\pi = \{i \sim j : \pi(i) = j\}$. $A_\pi \in \{0,1\}$. We shall sketch the proof steps here and refer to [Bapat 2010, Theorem 3.8] for the details.

Firstly, $A_\pi = 1$ iff $\pi(i) \neq i$ for any i iff H_π is a elementary subgraph with the components in H_π corresponding to the cycles in π. Secondly, for an elementary subgraph H

$$|\{\pi \in S_n : H_\pi = H\}| = 2^{c(H)},$$

as reversing the orientation in a cycle of π still yields that $H_\pi = H$. Finally, since each cycle in π corresponds to a component in H_π, we have that $|\pi| = c(H) + c_1(H)$. Thus, the proof is complete by the determinantal formula above. □

Recall that the characteristic polynomial is $\phi_\lambda(A) = det(\lambda I - A) = \sum_{i=0}^{n} c_i \lambda^{n-i}$. Note that $c_0 = 1$. Further, we have that $c_k = (-1)^k \sum_{W \subset [n], |W| = k} det(A[W|W])$. Observe that $A[W|W] = A(H_W)$ where H_W is the induced subgraph on W. Thus by Theorem 9.2.1, we have that

$$det(A(H_W)) = \sum_{H \subset H_W : H \text{ el. sp.}} (-1)^{k-c_1(H)-c(H)} 2^{c(H)}$$

and so

$$c_k = (-1)^k \sum_{H:H \text{ el. subgraph, } |V(H)|=k} (-1)^{c_1(H)+c(H)} 2^{c(H)}.$$

Trivially, $c_1 = 0$.

Suppose that there is $c_3 = \ldots = c_{2k-1} = 0$. for some $k \geq 1$.

If there is a cycle of length 3, then $c_3 \neq 0$ as every elementary subgraph on 3 vertices is a cycle. Since $c_3 \neq 0$, there is no cycle of length 3. If there is a cyle of length 5, then there is an induced cycle of length 3 or 5 but since there is no cycle of length 3, we have only induced cycle of length 5. Again, since $c_3 = 0$, all elementary subgraphs on 5 vertices are induced 5-cycle and so $c_5 \neq 0$, a contradiction. Thus there are no induced 5-cycles.

Similarly, if there is a cycle of odd-length l, there is an induced cycle of odd-length at most l. But recursively applying the above argument, there are no induced odd-length cycles of length strictly smaller than l and hence the induced cycle is of length l. But this yields that $c_l \neq 0$, a contradiction for $l \leq 2k-1$. Thus if $c_3 = \ldots = c_{2k-1} = 0$, there are no odd-cycles of length at most $2k-1$.

Now, if H is a elementary subgraph on $2k+1$, then there exists an odd-cycle of length at most $(2k+1)$. By the previous paragraph, we have that there are no odd-cycles of length at most $2k-1$. Hence the odd-cyle is of length $2k+1$ and hence the number of induced $(2k+1)$-cycles is $-\frac{1}{2}c_{2k+1}$. Thus, we get the following theorem characterizing bi-partitness.

Theorem 9.2.2. *TFAE :*

1. *G is bipartite.*
2. *$c_{2k+1} = 0, k = 0, 1, \ldots$*
3. *A has symmetric spectrum i.e., if λ is an e.v. of A with multiplicity k, so is $-\lambda$.*

Proof. The obervations before the theorem prove that (i) is equivalent to (ii). To show (ii) is equivalent to (iii), observe that (ii) implies that $\phi_A(\lambda) = \prod_{i=1}^{m}(\lambda^2 - a_i)$ as $\phi_A(\lambda)$ has no odd coefficients. So, the eigenvalues are $\overset{+}{-}\sqrt{a_i}$ and since the eigenvalues are all real due to symmetry of A, we have that the spectrum of A is symmetric.

Conversely, if the spectrum of A is symmetric then $\phi_A(\lambda) = \prod_{i=1}^{m}(\lambda^2 - a_i^2)$ for some $a_i \in \mathbb{R}$ and thus $\phi_A(\lambda)$ has no odd coefficients. $\square$

9.3 Exercises

1. Let G be a graph with incidence matrix ∂_1 and let $B = (b_{ij})_{i,j \leq k}$ be a $(k \times k)$-submatrix of ∂_1 which is nonsingular. Show that there is precisely one permutation σ of $1, \ldots, k$ such that the product $\prod_{i=1}^{k} b_{i\sigma_i}$ is non-zero.

2. Let $\delta_i = \partial_i^T$ be the transpose of the incidence matrices ∂_i for $i = 0, 1$. Show that δ_i defines a linear map from C_{i-1} to C_i where $C_{-1} = \mathbb{F}$, the underlying field. Further show that $\delta_1 \circ \delta_0 = 0$. Describe the spaces $Im(\delta_i), Ker(\delta_i)$ for $i = 0, 1$. Can you express the number of connected components in terms of the ranks of these spaces ?

3. Let G be a graph with k components. Suppose B is a $(l \times l)$- submatrix of the incidence matrix ∂_1. with indexed by $\{e_1, \ldots, e_l\}$ and rows by $\{v_1, \ldots, v_l\}$. Further, Show that B is a non-singular matrix only if $l \leq r(\partial_1)$,$\{e_1, \ldots, e_l\}$ form a forest. What is the subgraph $\{e_1, \ldots, e_l\}$ when $l = r(\partial_1)$?

4. Compute the eigenvalues of the Laplacian and Adjacency matrices of the the cycle graph and the path graph.

5. Let G be a graph with n vertices, m edges and let $\lambda_1 \geq \ldots \lambda_n$ be the eigenvalues of the adjacency matrix of G. Show the following bounds for the eigenvalues

   (a) $\lambda_1 \leq \sqrt{\frac{2m(n-1)}{n}}$

(b) $\delta(G) \leq \lambda_1 \leq \Delta(G)$.

(c) If H is an induced subgraph on p vertices then

$$\lambda_n(G) \leq \lambda_p(H) \leq \lambda_1(H) \leq \lambda_1(G).$$

6. Let G be a connected graph on n vertices and m edges. Let ∂_1 be its incidence matrix under a fixed orientation of edges. Let $y = (y_1, \ldots, y_n)^T$ be a $(n \times 1)$-column vector such that for some $i \neq j \in [n]$, $y_i = +1, y_j = -1$ and $y_l = 0$ for $l \in [n] \setminus \{i, j\}$. Show that there exists a $(m \times 1)$-column vector x such that $\partial_1 x = y$. Give a graph-theoretic interpretation of the same.

7. Let $\lambda_1 \leq \ldots \leq \lambda_n$ be the eigenvalues of L. Compute the eigenvalues of $L + aJ$ for $a > 0$ where J is the all 1 matrix.

8. Compute the eigenvalues of the laplacian matrix L and adjacency matrix A of the Petersen graph. Calculate the number of spanning trees of the Petersen graph. (Hint : Show that $A^2 + A - 2I = J$)

9. Compute the eigenvalues of $L(G \times H)$ in terms of $L(G)$ and $L(H)$ where $G \times H$ is the cartesian product of G and H.

9.4 Further reading

The books of [Bapat 2010] or [Godsil and Royle 2013] are excellent sources for more details on graphs and matrices. Spectrum of the graph also plays an important role in studying "zeta function" on graphs (see [Terras 2010]) and uses what is known as the non-backtracking matrix of a graph. As mentioned before, higher-dimensional topological analogues of incidence matrix can be found in the book of [Edelsbrunner 2010]. For a more analytical connection of the Laplacian, see [Grigoryan 2018]. The Laplacian also plays a crucial role in the certain games (Dollar game or abelian sandpile model) defined on graphs and using this one can formulate and prove Riemann-Roch theorem for graphs; see [Corry and Perkinson 2018]. Abelian Sandpile models were originally introduced by physicists to model certain phenomena of 'particles organizing themselves' and still many interesting questions remain mathematically unproven about this model. See [Perkinson 2011] for connections between "algebraic geometry" and "sandpiles". See [Bond and Levine 2013] for more general models than abelian sandpiles known as "Abelian networks".

Chapter 10

Planar graphs

Can we draw graphs on a paper without edges crossing each other ? We can draw K_4 and so any graph on at most 4 vertices can be drawn. What about K_5 ? First, we shall clarify what we mean by 'drawing'.

Definition 10.0.1 (Planar embedding)**.** *An embedding of the graph $G = (V, E)$ in the plane is a pair of functionf $f, f_e, e \in E$ such that*

- $f : V \to \mathbb{R}^2$ *is an injection.*
- $\forall e = (u, v) \in E$, $f_e : [0, 1] \to \mathbb{R}^2$ *is a continuous simple path such that* $f_e(0) = f(u), f_e(1) = f(v), f_e([0, 1]) \cap f(V) = \{f(u), f(v)\}$.
- *for* $e \neq e'$, *we have that* $f_e([0, 1]) \cap f_{e'}([0, 1]) = f(e \cap e')$ *or equivalently* $f_e((0, 1) \cap f_{e'}((0, 1)) = \emptyset$.

Since we are concerned with finite graphs, we shall assume that all f_e's are polygonal line segments i.e., union of finitely many straight lines.

We shall never specify f_e directly but more via drawings. A *plane graph* is a graph G with its embedding $f, f_e, e \in E$. We shall view plane graphs as a subset of $\mathbb{R}^2$ by identifying G with $f(V) \cup \cup_{e \in E} f_e([0, 1]) \subset \mathbb{R}^2$. We shall denote $f_e([0, 1])$ by e and $f_e((0, 1))$ by $\mathring{e}$. A graph isomorphic to a plane graph is called a *planar graph*. We refer to [Diestel 2000, Chapter 4] for various topological and geometric facts that shall be used in some of the proofs as well as more details. The most important of these is the Jordan curve theorem.

Definition 10.0.2 (Faces)**.** *Let G be a plane graph. $\mathbb{R}^2 - G$ is an open set and consists of finitely many connected components. Each connected component is called a* face.

Here we list some properties of faces.

Proposition 10.0.3. *1. For any face F and an edge e, $\mathring{e} \cap \partial F = \emptyset$ or $e \subset \partial F$. Thus we have that* $\partial F = \cup_{e : \mathring{e} \cap \partial F \neq \emptyset} e$.

2. *If $H \subset G$ as a plane graph and F is a face of G, then there is a face F' of H such that $F \subset F'$.*

3. *Assuming as above, if $\partial F \subset H$ then $\partial F = \partial F'$.*

4. *An edge e belongs to a cycle iff $e \in \partial F_1 \cap \partial F_2$ for two distinct faces F_1, F_2.*

5. *An edge e is a cut-edge iff there exists a unique face F such that $e \in \partial F$.*

A corollary of the above proposition is that a forest has only one face. Further, we define the length of a face F as

$$l(F) := \sum_{e \in \partial F} (\mathbf{1}[e \text{ is in a cycle}] + 2\mathbf{1}[e \text{ is a cut-edge}]).$$

One can define a notion of 'traversal' (i.e., a closed walk) of a face such that the length of the closed walk is the length of the face.

Exercise 10.0.4. $\sum_F l(F) = 2|E|$.

10.1 Euler's formula and Kuratowski's theorem

Theorem 10.1.1 (Euler's formula)**.** *Let G be a plane graph with v vertices, e edges and f faces. Then $v - e + f = 1 + \beta_0(G)$ where recall that $\beta_0(G)$ is the number of connected components.*

Proof. The proof proceeds by showing the formula for forests and then inductively verifying it for connected graphs. For forest, observe that there is only one face because of the definition of the length of a face, that every edge is a cut-edge, $v - e = \beta_0(G)$ and Exercise 10.0.4.

If the formula holds for trees, then one can show that if G, H are two disjoint plane graphs then there exists $u \in G, v \in H$ such that $G' = G \cup H \cup (u, v)$ is also a plane graph with the number of faces being $f'(G) = f(G) + f(H) - 1$. □

Lemma 10.1.2. *Suppose G is a connected plane graph with $v \geq 3$. Then $e \leq 3v - 6$. Further if G has no triangles then $e \leq 2v - 4$.*

Proof. Note that if ∂F corresponds to a cycle, $l(F) \geq 3$ and if not $l(F) \geq 4$ as $v \geq 3$ and G is connected. Thus, we have by Euler's formula that

$$3(2 - v + e) = 3f \leq \sum_F l(F) = 2e$$

and so $e \leq 3v - 6$. Now, if G has no triangles, then $l(F) \geq 4$ for all F and the above inequality yields that $e \leq 2v - 4$. □

Using the above lemma, we can show that $K_{3,3}$ and K_5 are not planar. A direct constructive proof using some geometric arguments are also possible for non-planarity of $K_{3,3}$ and K_5.

We say H is *a minor* of a graph G if H is obtained from G by a sequence of the following operations : (i) Edge-deletion. (ii) Vertex-deletion or (iii) edge-contraction. The three operations preserve planarity (see Section 10.4.

Now, we shall state without proof one of the important theorems.

Theorem 10.1.3 (Kuratowski's theorem)**.** *A graph if planar iff it has no $K_{3,3}$ or K_5 minor.*

A graph is a triangulation if it is connected and $l(F) = 3$ for every face F. A planar graph is said to be *maximal* if addition of any edge makes it non-planar.

Lemma 10.1.4. *G is a plane graph with $v \geq 3$. TFAE*

1. *$e = 3v - 6$ and G is connected*

2. *G is a triangulation.*

3. *G is a maximal planar graph.*

Proof. The equivalence of (i) and (ii) is by noting that the inequality in the proof of Lemma 10.1.2 is an equality in this case. □

Exercise 10.1.5. *Show that (ii) and (iii) equivalent in the above lemma.*

10.2 Coloring of Planar graphs

Lemma 10.2.1. *Every planar graph has a vertex of degree at most* 5*.*

Proof. If every vertex has degree at least 6, then $2e \geq 6v$ which contradicts Lemma 10.1.2. □

Using the above lemma and induction, one can prove that a planar graph is 5-colorable. But we can do better with a little more arguments.

Theorem 10.2.2. *Every planar graph is* 5*-colorable.*

Proof. The theorem is trivially true for graphs with at most 5 vertices. By induction, we assume that the theorem holds for all graphs with at most n vertices.

Let G be a planar graph with $n + 1$ vertices. By Lemma 10.2.1, there is a vertex v of degree at most 5. Choose the same.

CASE 1 : If $deg(v) < 5$, then by induction $G - v$ is 5-colorable and we can assign v a color not assigned to any of its (at most 4) neighbours.

CASE 2 : Suppose $deg(v) = 5$. Let $N(v) = \{v_1, \ldots, v_5\}$. Since K_5 is not planar, $v_1, \ldots, v_5$ cannot form a complete graph and hence WLOG, assume that $v_1 \nsim v_2$. Construct a new planar graph G' by contracting the edge (v_1, v) and denoting the new vertex by v'. Further construct a planar graph G''

by contracting the edge (v', v_2) and denoting the new vertex v''. Since G'' has $n-1$ vertices and is planar, it is 5-colorable.

To construct a 5-coloring on G : Assign the same colors as in G'' to all vertices in G except v, v_1, v_2. Further assign the color of v'' to v_1 and v_2. Now, $N(v)$ has been colored using only 4 colors and so assign the 5th color to v. $\square$

10.3 ***Graph dual****

Let G be a plane graph. Construct a new (multi-) plane graph G^* called *the dual* as follows : Choose $x_i \in \mathring{F}_i$ for all faces F_i's. For all $e \in \partial F_i \cap \partial F_j$ for $i \neq j$, draw an edge between x_i and x_j. If $e \in \partial F_i$ is a cut-edge, we draw a loop at x_i. We can draw G^* as a plane graph as follows : Connect x_i to the mid-point of all edges e such that $e \in \partial F_i$ such that paths to no two edges intersect. Thus, if $e \in \partial F_i \cap \partial F_j$ for $i \neq j$ then x_i and x_j are connected to mid-points of e. If e is a cut-edge, draw two non-intersecting edges to mid-point of e.

For example, see the following figure (10.3) for a graph (blue) and its dual (in red). Two isomorphic

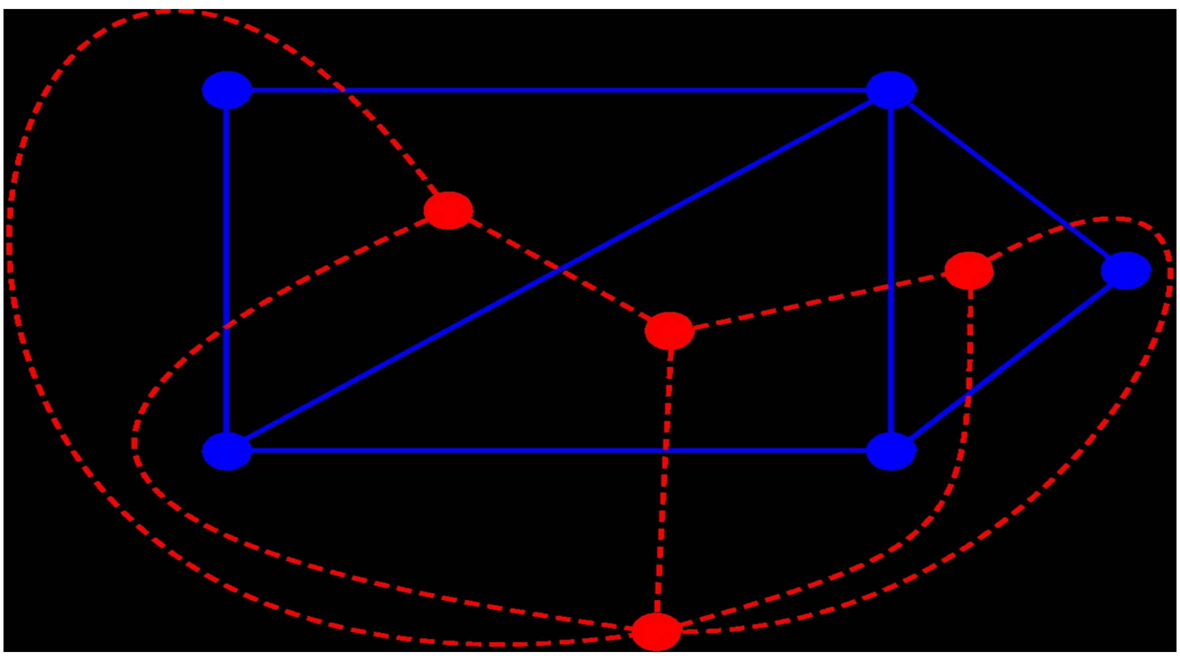


Figure 10.1: A graph and its dual

graphs can have different duals. See Figure 10.3

Exercise 10.3.1. *Show that the following are equivalent for a plane graph G.*

1. *G is bi-partite*

2. *$l(F)$ is even for all F.*

3. *The dual graph G^* is Eulerian.*

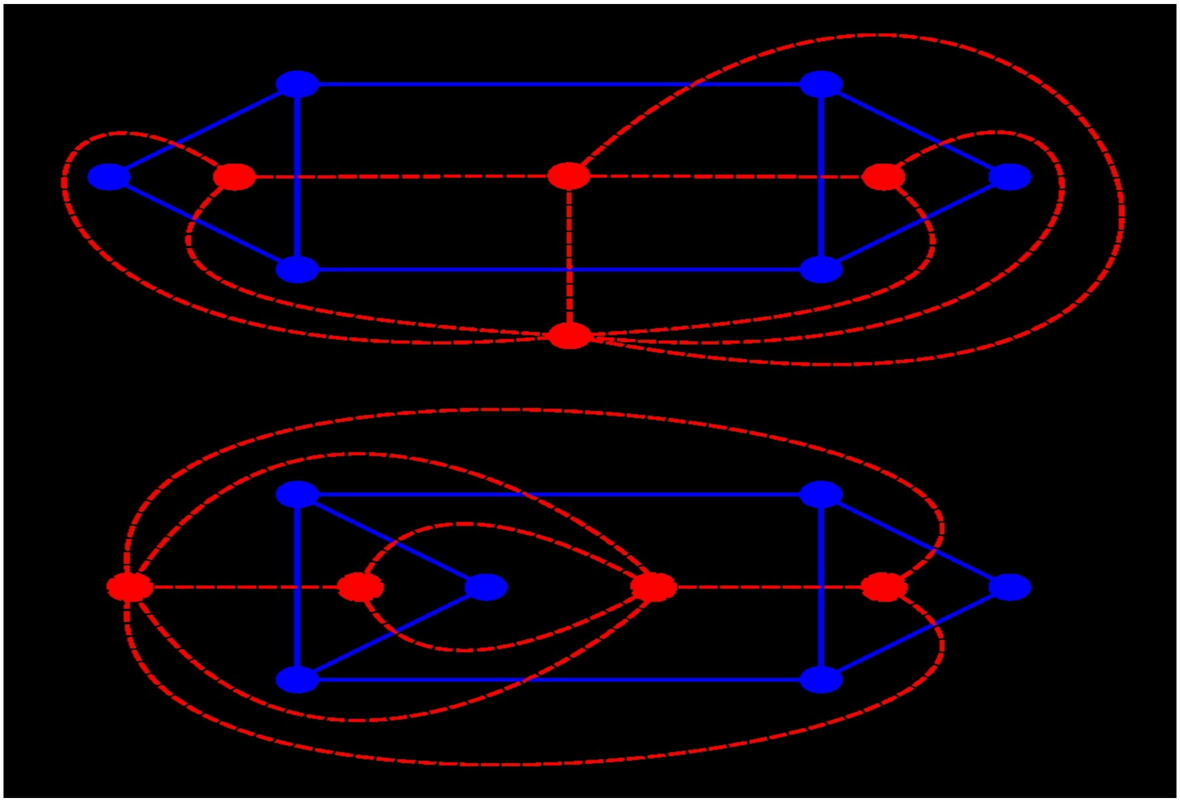


Figure 10.2: Two isomorphic graphs with same dual

Hint : Show that $d_{G^*}(x_i) = l(F_i)$.

10.4 Exercises

1. Merging two adjacent vertices of a planar graph yields another planar graph.
2. Any embedding of a planar graph will have the same number of faces.
3. Show that any connected triangle-free planar graph has at least one vertex of degree three or less. Prove by induction on the number of vertices that any connected triangle-free planar graph is 4-colorable.
4. Show that every planar graph with at least 4 vertices has at least 4 vertices of degree less than or equal to 5. (Hint : Consider a maximal planar graph.)

Additional Exercises : (not in syllabus or Tutorials)

1. Let G be a plane graph and G^* be its dual. Prove the following :
   - G^* is connected.
   - If G is connected, then each face of G^* contains exactly one vertex of G.

- $(G^*)^* = G$ iff G is connected.

2. Prove that a set of edges $T \subset E(G)$ in a connected plane graph form a spanning tree iff the duals of the edges $E(G) - T$ form a spanning tree of G^*.

3. Prove that every n-vertex self-dual plane graph has $2n - 2$ edges. For all $n \geq 4$, construct a simple n-vertex self-dual plane graph.

10.5 Further Reading

For more on planar graphs, see [Van Lint and Wilson, Chapter 33] and for graph embeddings on surfaces, see [Van Lint and Wilson, Chapter 35].

Bibliography

[Addario-Berry 2015] Addario-Berry L. "Partition functions of discrete coalescents: from Cayley's formula to Frieze's $\zeta(3)$ limit theorem." In XI Symposium on Probability and Stochastic Processes. (2015) (pp. 1-45). Birkhäuser, Cham.

[Aigner et al. 2010] Aigner, M., Ziegler, G. M., Hofmann, K. H., & Erdos, P. (2010). Proofs from the Book (Vol. 274). Berlin: Springer.

[Alon and Spencer 2004] Alon, N., and Spencer, J. H. (2004). The probabilistic method. John Wiley and Sons.

[Babai 1992] Babai, László, and Péter Frankl. Linear Algebra Methods in Combinatorics: With Applications to Geometry and Computer Science. Department of Computer Science, univ. of Chicag, 1992.

[Baker 2015] Baker, M. (2015). Hodge theory in combinatorics. `Hodgetheoryincombinatorics.`

[Baker 2018] Baker, M. (2018). Hodge theory in combinatorics. Bulletin of the American Mathematical Society, 55(1), 57-80.

[Bapat 2010] Bapat, R. B. (2010). Graphs and matrices (Vol. 27). London: Springer.

[Bauerschmidt et al. 2012] Bauerschmidt, Roland, Hugo Duminil-Copin, Jesse Goodman, and Gordon Slade. "Lectures on self-avoiding walks." *Probability and Statistical Physics in Two and More Dimensions* (D. Ellwood, CM Newman, V. Sidoravicius, and W. Werner, eds.), Clay Mathematics Institute Proceedings 15 (2012): 395-476.

[Bertsimas 1990] Bertsimas, Dimitris J., and Garrett Van Ryzin. "An asymptotic determination of the minimum spanning tree and minimum matching constants in geometrical probability." Operations Research Letters 9, no. 4 (1990): 223-231.

[Bollobas 2013] B. Bollobás, *Modern graph theory*, volume 184, (2013), Springer Science & Business Media.

[Bond and Levine 2013] B. Bond and L. Levine. (2013) Abelian Networks : Foundations and Examples, *arXiv:1309.3445v1.*

[Clay and Margalit 2017] Clay, Matt, and Dan Margalit, eds. Office Hours with a Geometric Group Theorist. Princeton University Press, 2017.

[Choudum 1986] Choudum, S. A. "A simple proof of the Erdos-Gallai theorem on graph sequences." Bulletin of the Australian Mathematical Society 33, no. 1 (1986): 67-70.

[Corry and Perkinson 2018] S. Corry and D. Perkinson. *Divisors and Sandpiles: An Introduction to Chip-Firing*, Volume 114, (2018), American Mathematical Soc.

[Diestel 2000] R. Diestel, *Graph theory*, (2000), Springer-Verlag Berlin and Heidelberg GmbH.

[Edelsbrunner 2010] Edelsbrunner, H., & Harer, J. (2010). Computational topology: an introduction. American Mathematical Soc..

[Frieze 1985] Frieze, Alan M. "On the value of a random minimum spanning tree problem." Discrete Applied Mathematics 10, no. 1 (1985): 47-56.

[Frieze and Pegden 2017] Frieze, Alan, and Wesley Pegden. "Separating subadditive Euclidean functionals." Random Structures & Algorithms 51, no. 3 (2017): 375-403.

[Gale and Shapley 1962] Gale, David, and Lloyd S. Shapley. "College admissions and the stability of marriage." The American Mathematical Monthly 69, no. 1 (1962): 9-15.

[Godsil and Royle 2013] Godsil, C., and Royle, G. F. (2013). Algebraic graph theory (Vol. 207). Springer Science & Business Media.

[Grigoryan 2018] Grigoryan, A. (2018). Grigor'yan, A. (2018). Introduction to Analysis on Graphs (Vol. 71). American Mathematical Soc..

[Grochow 2019] Grochow, Joshua. "New applications of the polynomial method: The cap set conjecture and beyond." Bulletin of the American Mathematical Society 56, no. 1 (2019): 29-64.

[Guth 2016] Guth, Larry. Polynomial methods in combinatorics. Vol. 64. American Mathematical Soc., 2016.

[Jukna 2011] Jukna, Stasys. Extremal combinatorics: with applications in computer science. Springer Science & Business Media, 2011.

[Krishnapur 2019] Krishnapur, M. Topics in Analysis : Lecture Notes. `http://math.iisc.ac.in/~manju/TA2019/Topicsinanalysis2019.pdf`, 2019.

[Lovasz 2011] Lovász, László. "Graph Theory Over 45 Years." In An Invitation to Mathematics, pp. 85-95. Springer, Berlin, Heidelberg, 2011.

[Pitman 1999] Pitman, Jim. "Coalescent random forests." Journal of Combinatorial Theory, Series A 85, no. 2 (1999): 165-193.

[Munkres 2018] Munkres, J. R. (2018). Elements of algebraic topology. CRC Press.

[Perkinson 2011] D. Perkinson, J. Perelman and John Wilmes. Primer for the Algebraic Geometry of Sandpiles, *arXiv:1112.6163*.

[Rhee 1992] Rhee, Wansoo T. "On the travelling salesperson problem in many dimensions." Random Structures & Algorithms 3, no. 3 (1992): 227-233.

[Smirnov 2011] Smirnov, Stanislav. "How do research problems compare with IMO problems?." In An Invitation to Mathematics, pp. 71-83. Springer, Berlin, Heidelberg, 2011.

[Spencer 1994] Spencer, J. (1994). Ten lectures on the probabilistic method (Vol. 64). SIAM.

[Stanley 2013] Stanley, Richard P. "Algebraic combinatorics." Springer 20 (2013): 22.

[Steele 1997] Steele, J Michael. "Probability theory and combinatorial optimization." vol. 69, (1997) : SIAM.

[Terras 2010] Terras, A. (2010). Zeta functions of graphs: a stroll through the garden (Vol. 128). Cambridge University Press.

[Tripathi 2010] Tripathi, Amitabha, Sushmita Venugopalan, and Douglas B. West. "A short constructive proof of the Erdős–Gallai characterization of graphic lists." Discrete Mathematics 310, no. 4 (2010): 843-844.

[Van Lint and Wilson] van Lint, Jacobus Hendricus and Wilson, Richard Michael *A course in combinatorics*, 2001, Cambridge university press.

[West 2001] D. B. West, *Introduction to graph theory*, Volume 2 (2001), Prentice hall Upper Saddle River.