Extremal Combinatorics

and

Graph Theory

A short course at UCSD

Table of contents

CHAPTER 1

Introduction

A ***graph*** G is a pair (V, E) where V is a set and E is a set of unordered pairs[1] of elements of V. The elements of V are called ***vertices*** and V is called the ***vertex set*** of the graph, and the elements of E are called ***edges***, and E is called the ***edge set*** of the graph. If G is a graph, we let $V(G)$ denote its vertex set and $E(G)$ its edge set. We note by $e(G)$ the number of edges in G. If u and v are two vertices of a graph $G = (V, E)$, then we say u and v are ***adjacent*** if $\{u, v\} \in E$ – in other words $\{u, v\}$ is an edge of G – and we say that vertex v is ***incident*** with edge e if $v \in e$.

We sometimes consider the following generalizations of graphs: a ***multigraph*** is a pair (V, E) where V is a set and E is a ***multiset*** of unordered pairs from V. In other words, we allow more than one edge between two vertices. A ***pseudograph*** is a pair (V, E) where V is a set and E is a ***multiset*** of unordered multisets of size two from V. A pseudograph allows ***loops***, namely edges of the form $\{a, a\}$ for $a \in V$. A ***digraph*** is a pair (V, E) where V is a set and E is a ***multiset*** of ordered pairs from V. In other words, the edges now have a direction: the edge (a, b) and edge (b, a) are different, and denoted in a digraph by putting an arrow from a to b or from b to a, respectively. A frequent source of digraphs comes from ***orientation*** of a graph: for each edge $\{a, b\}$ in the graph, replace it with either (a, b) or (b, a).

1.1. Degrees and neighborhoods

The ***neighborhood*** of a vertex v in a graph $G = (V, E)$, denoted $N_G(v)$, is the set of vertices of G which are adjacent to v. The ***degree*** of a vertex v in a graph G, denoted $d_G(v)$, is $|N_G(v)|$. When it is clear which graph G we are referring to, we write $d(v)$ and $N(v)$ instead of $d_G(v)$ and $N_G(v)$. The ***degree sequence*** of a graph G is the sequence of degrees of vertices of G in non-increasing order. A vertex of degree zero is called an ***isolated vertex***.

We write $\delta(G) = \min\{d_G(v) : v \in V\}$ and $\triangle(G) = \max\{d_G(v) : v \in V\}$ for the ***minimum degree*** and ***maximum degree*** of G, respectively. If all vertices in a graph have the same degree r, then the graph is said to be ***r-regular*** and if $r = 3$, they are referred to as ***cubic*** graphs. An important fact involving the degrees of a graph G, which we will use on numerous occasions, is the ***handshaking lemma***:

[1] We denote sets using braces, for instance $\{1, 2, 3\}$ is the set whose elements are $1, 2$ and 3, and we write $1 \in \{1, 2, 3\}$ to say "1 is an element of the set $\{1, 2, 3\}$." Note that a set precludes "repeated elements".

LEMMA 1. (HANDSHAKING LEMMA)
For any graph $G = (V, E)$,

$$\sum_{v \in V} d_G(v) = 2|E|. \tag{1.1}$$

PROOF. When we add up the degrees of vertices of G, every edge of G is counted twice, so the sum of the degrees is twice the number of edges. $\square$

The handshaking lemma gives an easy way to count the number of edges in a graph: it is just half the sum of the degrees of the vertices. Note if G is r-regular and has n-vertices, then the number of edges in G is $nr/2$, by the handshaking lemma (check this for the cube graph Q_4 in the figure below). A consequence of the handshaking lemma is that the number of vertices of odd degree in any graph must be even – otherwise the sum on the left above would be odd whereas the right side is even:

LEMMA 2. *For any graph $G = (V, E)$, the number of vertices of odd degree is even.*

EXAMPLE 1.1. The ***n-cube***, denoted Q_n, is the graph whose vertex set is the set of binary strings of length n, and whose edge set consists of all pairs of strings differing in one position. Let us see how many edges Q_n has as a formula in n. Since there are 2^n binary strings of length n, there are 2^n vertices in Q_n. Now each vertex v is adjacent to n other vertices – flip one position in v to get each string adjacent to v, and there are n possible positions to flip. So every vertex of the n-cube has degree n (in other words, it is n-regular), and so the number of edges in Q_n is

$$\frac{1}{2}\sum_{v \in V} d_{Q_n}(v) = \frac{1}{2} \cdot 2^n \cdot n = n2^{n-1}.$$

A manual count of the edges confirms this for Q_4, which is drawn below:

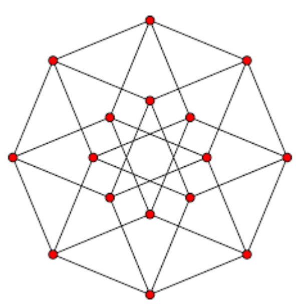


FIGURE 1.1. The 4-cube Q_4

The definition of neighborhood is extended to sets as follows: the ***neighborhood of a set*** X of vertices in a graph G, denoted $N_G(X)$, is set of vertices $x \notin X$ which are adjacent to at least one vertex in X:

$$N_G(X) = \{x \in V(G)\backslash X : \exists y \in X, \{x, y\} \in E(G)\}.$$

When $X = \{v\}$, this is precisely the usual definition of neighborhood.

1.2. Subgraphs

If H and G are graphs and $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$, then H is called a ***subgraph*** of G. To denote that H is a subgraph of G, we write $H \subseteq G$. If in addition $V(H) = V(G)$ then H is called a ***spanning subgraph*** of G.

We now define removal of edges and vertices from a graph G. For $X \subset V(G)$, we denote by $G - X$ the graph with vertex set $V(G)\backslash X$ and edge set $E = \{e \in E(G) : e \cap X = \emptyset\}$. For $L \subseteq E(G)$, we denote by $G - L$ the graph with vertex set $V(G)$ and edge set $E(G)\backslash L$. For sets $X.Y \subseteq V(G)$, let $e(X, Y)$ denote the number of edges of a graph with one end in X and one end in Y. The subgraph of G ***induced*** by a set $X \subseteq V(G)$, denoted $G[X]$, is precisely $G-(V\backslash X)$. A subgraph H of G is an ***induced subgraph*** if for some $X \subseteq V(G)$, $H = G[X]$. We write $e(X)$ for $e(G[X])$. If L is a set of edges of G, then the subgraph of G ***spanned by*** L is the graph with edge set L and vertex set $\bigcup_{e\in L} e$.

1.3. Basic classes of graphs

A ***path*** in a graph G is a subgraph P of G with vertex set of the form $\{v_1, v_2, \dots, v_k\}$ and edge set of the form $\{\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_{k-1}, v_k\}\}$. A graph is ***connected*** if any pair of vertices in the graph are the ends of at least one path. If a graph is not connected, we say it is ***disconnected***. The ***components*** of a graph $G = (V, E)$ are the maximal connected subgraphs of G – that is, the connected subgraphs such that no edge of G not already in the subgraph can be added while still preserving connectivity. A graph is k***-connected*** if for every set X of less than k vertices of G, $G - X$ is connected, and k***-edge-connected*** if for every set L of less than k edges of G, $G - L$ is connected. A ***bridge*** of a graph G is an edge e such that $G - e$ has more components than G. A ***cut vertex*** is a vertex $v \in V(G)$ such that $G - \{v\}$ is disconnected.

We can extend these notions to digraphs: a digraph is called ***strongly connected*** if any for any ordered pair of vertices (u, v) in the digraph, there exists a path with vertices $v_1, v_2, \dots, v_k$ such that $v_1 = u$ and $v_k = v$ and (v_i, v_{i+1}) is an edge of the digraph for $1 \leq i < k$ – this is called a ***directed path***. Similarly, a ***directed cycle*** is a digraph consisiting of edges $(v_1, v_2), (v_2, v_3), \dots, (v_{k-1}, v_k), (v_k, v_1)$. We next describe some graphs which we shall encounter frequently.

Complete graphs. The ***complete graph*** on n vertices, denoted K_n is the graph consisting of all possible edges on n vertices (in other words, every pair of vertices is adjacent). The ***empty graph*** on n vertices has no edges. In Figure 1.2, drawings of K_n for $2 \leq n \leq 6$ are given:

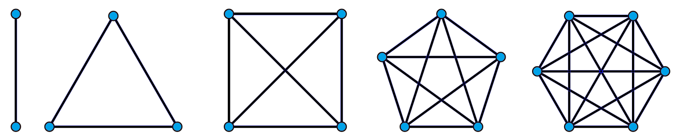


FIGURE 1.2. The complete graphs K_2 through K_6

The number of edges in K_n is $\binom{n}{2}$.

Bipartite graphs. Recall a ***partition*** of a set V consists of pairwise disjoint non-empty subsets whose union is V. A ***bipartite graph*** is a graph $G = (V, E)$ such that for some partition of V into two sets A and B such that every edge of G has the form $\{a, b\}$ with $a \in A$ and $b \in B$ (or in other words, no two vertices in A are adjacent, and no two vertices in B are adjacent). We call A and B the ***parts*** of G and refer to (A, B) as the ***bipartition of*** G. When $|A| = r$ and $|B| = s$ and all possible edges $\{a, b\}$ with $a \in A$ and $b \in B$ are included, then G is called the ***complete bipartite graph***, and denoted $K_{r,s}$. Note that the number of edges in a complete bipartite graph $K_{r,s}$ is exactly rs. In Figure 1.3, we draw the graphs $K_{2,3}$ and $K_{2,5}$.

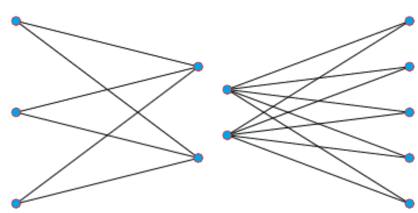


FIGURE 1.3. Complete bipartite graphs $K_{2,3}$ and $K_{2,5}$

Paths and Cycles. For $k \geq 3$, a ***k-cycle*** is the graph C_k with edge set

$$\{\{1, 2\}, \{2, 3\}, \{3, 4\}, \ldots, \{k - 1, k\}, \{k, 1\}\}.$$

For $k \geq 1$, a ***k-path*** is the graph P_k with edge set

$$\{\{1, 2\}, \{2, 3\}, \{3, 4\}, \ldots, \{k - 1, k\}, \{k, k + 1\}\}.$$

Note that a k-cycle has k edges and a k-path has k edges, and we often refer to the number k as the ***length*** of the cycle or path. It is convenient to represent a path as a sequence of vertices, for instance $v_1v_2 \ldots v_k$ represents a path with k vertices consisting of the edges v_iv_{i+1} for $i < k$. Similarly, a cycle can be represented as a circular sequence $v_1v_2 \ldots v_kv_1$. In Figure 1.4, we draw cycles C_3 and C_6.

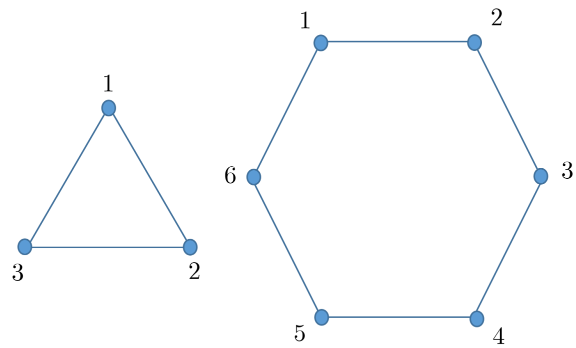


FIGURE 1.4. Cycles C_3 and C_6

LEMMA 3. *Let $k \geq 2$, and let G be a graph of minimum degree at least k. Then G contains a cycle of length at least $k+1$.*

PROOF. Let P be a longest path in G, say $v_1 v_2 \dots v_r$. Then $N(v_r) \subseteq V(P)$. Since $|N(v_r)| \geq \delta(G) \geq k$, $r \geq k+1$ and v_r has a neighbor v_i for some $i \leq r-k$. Now the cycle $v_i v_{i+1} \dots v_r v_i$ has length at least $k+1$, as required. $\square$

For $k \geq 2$, the complete graph K_{k+1} has minimum degree k and no cycle of length more than $k+1$, so the lemma is best possible. When $(k+1)|n$, a disjoint union of $n/(k+1)$ complete graphs K_{k+1} gives an n-vertex graph of minimum degree k with no cycle of length more than $k+1$.

1.4. Trees

A ***tree*** is a connected graph without cycles – a connected ***acyclic*** graph. An acyclic graph is called a ***forest*** – thus all components of a forest are trees. The following can be proved by induction on the number of vertices:

LEMMA 4. *A n-vertex connected graph is a tree if and only if it has $n-1$ edges.*

In any connected graph G, while there is a cycle, pick an edge of the cycle and remove it. Since this does not disconnect the graph, we repeat this procedure we eventually obtain a spanning subgraph of G which is acyclic and connected – a tree. We call this a ***spanning*** tree of the graph:

LEMMA 5. *Any connected graph contains a spanning tree.*

A spanning tree of the cube graph is given below in bold edges:

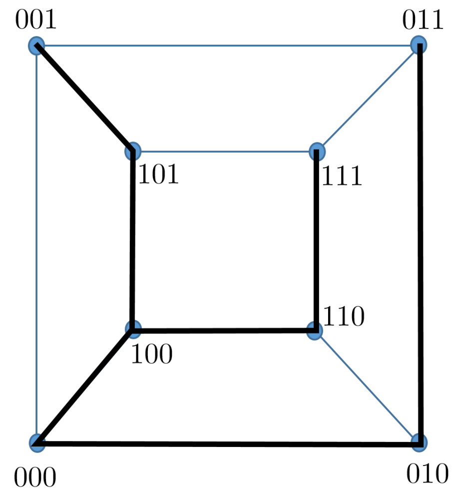


FIGURE 1.5. Spanning tree

1.5. Distance and search trees

The ***distance*** between vertices u and v in a connected graph G, denoted $d_G(u, v)$, is the length of a shortest uv-path. The set V of vertices of G together with the distance d_G form a ***metric space*** with d_G as the ***metric***. The most well-known metric space have the metric equal to the Euclidean distance between two points in a set of points in space. The definition of a metric d on a set X in general is a function $d : X \times X \to [0, \infty)$ such that for all $x, y, z \in X \times X$,

1. $d(x, y) = 0$ iff $x = y$
2. $d(x, y) = d(y, x)$
3. $d(x, z) \leq d(x, y) + d(y, z)$.

The third statement is the ***triangle inequality***. A ***geodesic*** in a graph G is a uv-path of length $d_G(u, v)$. A well-known algorithm for finding all shortest paths from a given vertex in a graph is known as ***Dijkstra's Shortest Path Algorithm***.

The maximum distance between any two vertices in a connected graph is called the ***diameter*** of G. The minimum r such that every vertex of G is at distance at most r from some vertex of G is called the ***radius*** of G. The triangle inequality implies that for every graph G of radius r and diameter d, $r \leq d \leq 2r$.

If v is a vertex in a connected graph G, we let $N_i(v)$ denote the set of vertices at distance exactly i from v, so that $N_1(v)$ is exactly the neighborhood of v and $N_0(v) = \{v\}$. We start with an arbitrary ordering of the vertices of G with some vertex v_0 as the first vertex. We build a spanning tree called a ***breadth-first search tree*** or ***BFS tree*** T in G as follows. First we add v to T – the ***root*** of the tree. At any stage of the construction, we have a tree T with vertex set $\{v_0, v_1, v_2, \ldots, v_k\}$. If $V(T) = V(G)$, stop. Otherwise, since G is connected, there exists a smallest integer

i such that $N(v_i)\backslash V(T) \neq \emptyset$. Choose v_{k+1} to be the smallest neighbor of v_i not in T in the ordering of the vertices of G, and add the edge $\{v_i, v_{k+1}\}$ to T.

LEMMA 6. *Let T be a BFS-tree rooted at a vertex v in a connected graph G. Then T is a spanning tree of G with $d_T(v, w) = d_G(v, w)$ for all $w \in V(G)$.*

The last statement in this lemma says that T preserves distances from v to all other vertices. The tree T is called a ***breadth-first search tree rooted at*** v. The sets $N_i(v)$ are sometimes called the ***layers*** of T, and the ***height*** of T is the maximum distance of any vertex from v.

EXAMPLE 1.2. The famous ***Petersen graph*** is drawn below, with vertices labelled 1 through 10. Let us apply the breadth-first search algorithm to find a spanning tree in G rooted at vertex 1. Of course, we start by adding 1 to the tree.

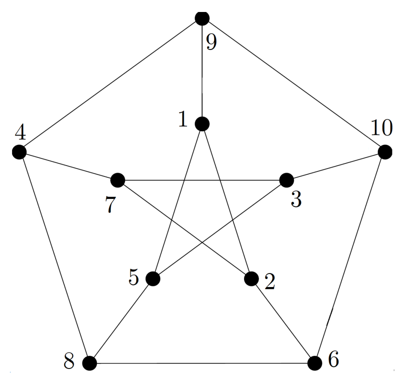


FIGURE 1.6. Petersen graph

We then add the neighbors of 1 in increasing order, namely, 2, 5 and then 9. So far our tree has edges $\{1, 2\}$, $\{1, 5\}$ and $\{1, 9\}$. Now we move on to the first vertex added in $N_1(v)$, namely 2. We add first the vertex 6 and then the vertex 7, with edges $\{2, 6\}$ and $\{2, 7\}$. Then move to the next added vertex in $N_1(v)$, namely 5. We add 3 and 8 and the edges $\{5, 3\}$ and $\{5, 8\}$. Finally, we move to the vertex 9, and add the vertices 4 and 10 and the edges $\{9, 4\}$ and $\{9, 10\}$. Then we stop since there are no vertices left to add. The tree has edges $\{1, 2\}, \{1, 5\}, \{1, 9\}, \{2, 6\}, \{2, 7\}, \{5, 3\}, \{5, 8\}, \{9, 4\}$ and $\{9, 10\}$, and the order in which vertices were added is $(1, 2, 5, 9, 6, 7, 3, 8, 4, 10)$. The layers of the tree are $N_0(v) = \{1\}$, $N_1(v) = \{2, 5, 9\}$, and $N_2(v) = \{6, 7, 3, 8, 4, 10\}$. The reader can check that the Petersen graph has diameter and radius equal to two.

Another way to generate a spanning tree in a graph G is the ***depth-first search*** algorithm. Let the vertices of G be ordered, and identify a vertex $v \in V(G)$ which will be the ***root*** of the ***depth-first search tree***. At any stage, pick a vertex x of the tree so far that is as far from v as possible in the tree, and has a neighbor w not in the tree. Select the first such neighbor w in the ordering of the vertices of G, and add the edge $\{x, w\}$ to the tree.

EXAMPLE 1.3. Consider the Petersen graph in Figure 1.6, with 1 being the root vertex. Then the depth first search algorithm gives the ordering $(1, 2, 6, 8, 4, 7, 3, 5, 10, 9)$. Note that after we add the edge $\{3, 5\}$ to the tree, the furthest vertex from the root and that has a neighbor not in the tree is 3, and we add $\{3, 10\}$. Then the next furthest vertex from the root which has a neighbor outside of the tree so far is 4 and we add the edge $\{4, 10\}$ to complete the tree.

There are many other algorithms for generating spanning trees of graph, and in particular, ***Prim's Algorithm*** and ***Kruskal's Algorithm*** for generating a minimum cost spanning tree in a graph where each edge has an associated cost.

1.6. Bipartite Graphs

One of the simplest things to check is whether a connected graph is bipartite. Namely, pick any vertex of the graph and place it in A. Then, all the neighbors of that vertex are forced to be in B. Then all their neighbors must be in A, and so on. We repeat this procedure on all components of the graph until all the vertices of the graph have been placed in A and in B, or the graph is not bipartite. This can be done via breadth-first search to characterize bipartite graphs:

LEMMA 7. *A graph G is bipartite if and only if it does not contain any odd cycles.*

PROOF. Since an odd cycle is not bipartite, bipartite graphs cannot contain odd cycles. Conversely, if a graph has no odd cycles, let T be a breadth-first search tree in G, rooted at some vertex v. We claim that $A = N_0(v) \cup N_2(v) \cup \dots$ and $B = N_1(v) \cup N_3(v) \cup \dots$ do not contain any edges of G, and therefore they are the parts in a bipartition of G. Suppose there exists an edge $\{x, y\}$ in A. Since edges of T connect consecutive layers, $\{x, y\}$ is not in T. Suppose $x, y \in N_{2i}(v)$. Let P be an xy-path in T. Then P together with $\{x, y\}$ forms a cycle C. On the other hand, P must have even length, since if $N_h(v)$ is the lowest layer that P intersects (see Figure 1.7), then P has length $(2i - h) + (2i - h) = 4i - 2h$. But then C has length $4i - 2h + 1$, so C is an odd cycle, which is a contradiction. Similarly, B does not contain any edges of G, so A and B are the parts of G and G is bipartite. $\square$

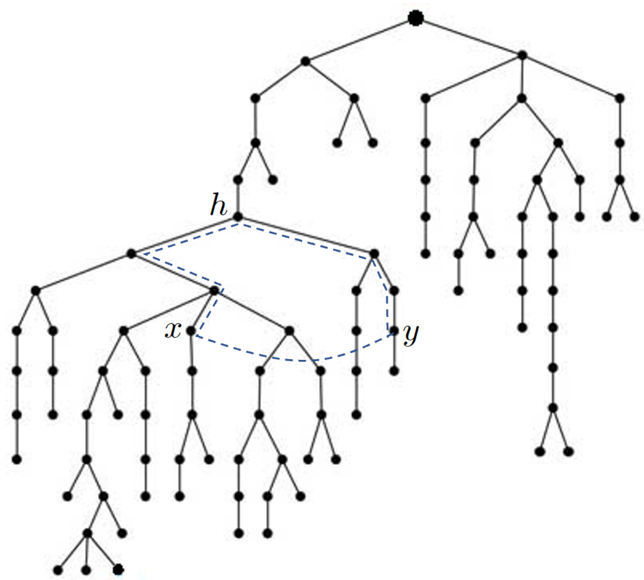


FIGURE 1.7. Breadth first search and bipartite graphs

1.7. Exercises

Question 1.7.1° For $n \geq 2$, let G_n be the ***grid graph***, whose vertex set is

$$V = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : 0 \leq x < n, 0 \leq y < n\}$$

and whose edge set is

$$E = \{\{(x, y), (x', y')\} : (x - x')^2 + (y - y')^2 = 1\}.$$

Determine the number of vertices and number of edges in G_n for each $n \geq 2$.

Question 1.7.2° Let $K_{n:r}$ denote the ***Kneser graph***, whose vertex set is the set of r-element subsets of an n-element sets, and where two vertices form an edge if the corresponding sets are disjoint.

(a) Describe $K_{n:1}$ for $n \geq 1$.
(b) Draw $K_{4:2}$ and $K_{5:2}$.
(c) Determine $|E(K_{n:r})|$ for $n \geq 2r \geq 1$.

Question 1.7.3° Prove that the vertices of an n-vertex connected graph can be ordered $(v_1, v_2, \ldots, v_n)$ so that for $i > 1$, v_i has at least one neighbor v_j with $j < i$.

Question 1.7.4° Let G be a digraph such that every vertex has positive in-degree. Prove that G contains a directed cycle.

Question 1.7.5° Find breadth-first search and depth-first search trees in the ***dodecahedron graph*** shown below.

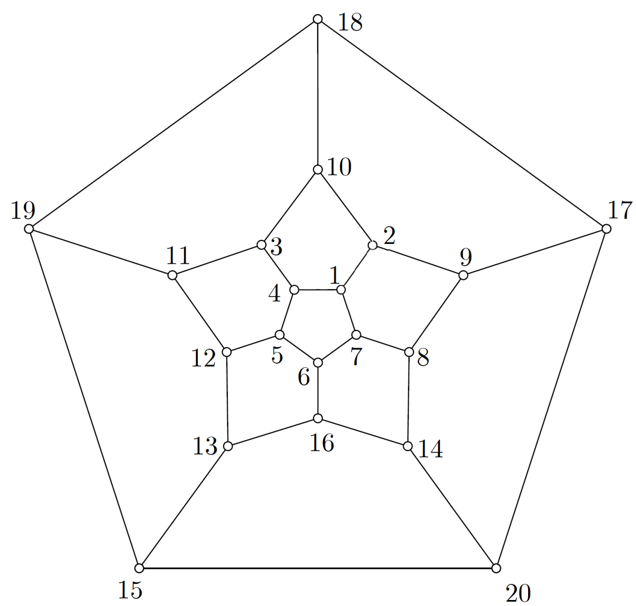


FIGURE 1.8. Dodecahedron graph

Question 1.7.6° Prove that any n-vertex graph with m edges has at least $m-n+1$ cycles.

Question 1.7.7° Let $G=(V,E)$ be a connected graph. Prove that for $1 \leq k \leq |V(G)|$, G has a connected subgraph with exactly k vertices.

Question 1.7.8° Let $G=(V,E)$ be a graph and $d(x,y)$ the distance between two vertices. Prove that (V,d) is a ***metric space***.

Question 1.7.9° Give a description of all graphs of radius 1 and diameter 2.

◇

Question 1.7.10. The ***line graph*** of a graph $G=(V,E)$ is the graph $L(G)=(E,F)$ whose vertex set is E and whose edge set is

$$F = \{\{e,f\} \subset E : e \cap f \neq \emptyset\}.$$

Determine $|E(L(G))|$ in terms of the degree sequence $(d_1, d_2, \ldots, d_n)$ of G.

Question 1.7.11.

(a) Prove that every graph with at least two vertices contains two vertices with the same degree.
(b) Is (a) true for multigraphs?
(c) For each $n \geq 2$ give an example of a graph with n vertices which does not have three vertices of the same degree.

Question 1.7.12. Let G be an n-vertex graph with $n \geq 2$ and $\delta(G) \geq (n-1)/2$. Prove that G is connected and that the diameter of G is at most two.

Question 1.7.13. Prove that for every graph G with radius r and diameter d, $r \leq d \leq 2r$. For each pair of positive integers r and d with $r \leq d \leq 2r$, give an example of a graph with radius r and diameter d.

Question 1.7.14.

(a) Let P and Q be longest paths in a connected graph G. Prove that

$$V(P) \cap V(Q) \neq \emptyset.$$

(b)* Let $P_1, P_2, \ldots, P_k$ be longest paths in a tree T. Prove that

$$V(P_1) \cap V(P_2) \cap \cdots \cap V(P_k) \neq \emptyset.$$

(c)* Prove that there is a connected graph such that the intersection of the vertex sets of all longest paths in G is empty.

Question 1.7.15. Let G be a graph whose vertex set is a set $V=\{p_1, p_2, \ldots, p_6\}$ of six people. Prove that there exist three people who are all friends with each other, or three people none of whom are friends with each other.

◇

Question 1.7.16* Let G be a connected graph with an even number of vertices. Prove that there is a spanning subgraph $H \subseteq G$ such that all vertices of H have odd degree.

Question 1.7.17* Let G be an n-vertex digraph such that

$$|N^+(v)| > \frac{1}{2}(3 - \sqrt{5})n$$

for every $v \in V(G)$. Prove that G contains a directed cycle of length two or three.

Question 1.7.18* Consider n people possessing unique items $u_1, u_2, \ldots, u_n$ of information that they wish to share with each other. Two people can call each other and share all the items of information they currently have.

(a) For $n \leq 4$, determine the minimum number of calls that can be made so that all information is shared amongst all n people.

(b) Prove that for $n \geq 5$ the minimum number of calls so that all n people have all items of information is $2n - 4$.

CHAPTER 2

Hamiltonian and Eulerian Graphs

A spanning cycle in a graph is called a ***hamiltonian cycle*** and a spanning path in a graph is called a ***hamiltonian path***. A graph is ***hamiltonian*** if it contains a hamiltonian cycle and ***traceable*** if it has a hamiltonian path. A graph is ***eulerian*** if all of its vertices have even degree. An ***eulerian tour*** of a graph with m edges is a sequence $(v_1, v_2, \ldots, v_m, v_1)$ of vertices such that every edge of the graph appears exactly once as $\{v_i, v_{i+1}\}$ or $\{v_{i+1}, v_i\}$). In this section, we discuss a necessary and sufficient condition for a graph to have an eulerian tour, and sufficient conditions for hamiltonian cycles.

2.1. Eulerian graphs

A ***digraph*** is a pair $G = (V, E)$ where V is a set and E is a multiset of ordered pairs of elements of V. Note that two vertices can be joined by many edges in either direction, and we allow ***loops***: a vertex may have an edge to itself. In a digraph G, let $N^+(v)$ and $N^-(v)$ denote the sets of vertices adjacent ***from*** v and ***to*** v, respectively. These are the ***out-neighborhood*** of v and the ***in-neighborhood*** of v respectively. Thus

$$N^+(v) = \{u : (v, u) \in E\} \qquad N^-(v) = \{u : (u, v) \in E\}.$$

For example, in the digraph drawn below, we have $N^+(x) = \{u, v, w\}$, $N^-(x) = \{v\}$.

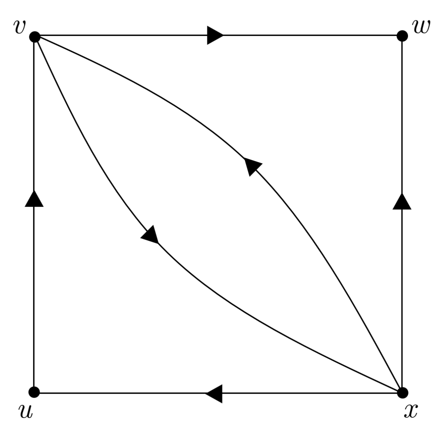


FIGURE 2.1. A digraph

The ***in-degree*** of a vertex v is $d^-(v) = |N^-(v)|$ and the ***out-degree*** is $d^+(v) = |N^+(v)|$. Digraphs arise frequently as ***orientations of graphs***: for a graph G and ach edge $\{u, v\} \in E(G)$, we replace $\{u, v\}$ either with (u, v) or (v, u). If G is a digraph, then the ***underlying multigraph*** is the multigraph such that $\{u, v\}$ is an edge whenever (u, v) is an edge of the digraph.

An ***eulerian tour*** of a digraph with m edges is a sequence $(v_1, v_2, \ldots, v_m, v_1)$ of vertices such that every edge of the digraph appears exactly once as a pair of consecutive vertices (v_i, v_{i+1}) in the sequence. In words, we can draw the digraph without retracing any edges. A digraph is ***eulerian*** if for every vertex v, the in-degree and out-degree are equal. The following is a necessary and sufficient condition for an eulerian tour in a graph or digraph.

THEOREM 1. *A graph has an eulerian tour if and only if it is eulerian and connected. A digraph has an eulerian tour if and only if it is eulerian and the underlying graph is connected.*

PROOF. We prove the first statement and leave the second as an exercise. Since an eulerian tour of a graph using an edge $\{v_{i-1}, v_i\}$ into vertex v_i must use the edge $\{v_i, v_{i+1}\}$ to leave, all vertices have even degree. Now suppose all vertices of G have even degree. Let $\tau = (v_1, v_2, \ldots, v_k)$ be the longest possible ***trail*** in G: each edge of G appears at most once as a pair of consecutive vertices in τ. If $v_k \neq v_1$, then an odd number of edges of τ contain each of v_1 and v_k, so there is an edge $\{v_k, v_{k+1}\}$ of G that is not traversed by τ. Now $(v_1, v_2, \ldots, v_k, v_{k+1})$ is a longer trail than τ, a contradiction. We conclude $v_k = v_1$ and τ is a tour in G. Since G is connected, there is an edge e not in the trail τ, say $\{v_i, v\} \in E(G)$. If v is not a vertex of the trail, then

$$(v_i, v_{i+1}, \ldots, v_k, v_1, v_2, \ldots, v_{i-1}, v_i)$$

is a tour of the same length as τ in G. If we add the edge $\{v_i, v\}$, we get the trail

$$(v_i, v_{i+1}, \ldots, v_k, v_1, v_2, \ldots, v_{i-1}, v_i, v)$$

which is longer than τ. If v is a vertex on the trail, say $v = v_j$ where $j < i$, then consider the trail $(v_i, v_{i+1}, \ldots, v_k, v_1, \ldots, v_{j-1}, v_j, v_i, v_{i-1}, \ldots, v_{j+1}, v_j)$ is a trail using the edge e and is one longer than τ. This contradiction completes the proof. $\square$

The proof of Theorem 1 gives an algorithm for finding an eulerian tour, known as ***Hierholzer's Algorithm***. The problem of finding eulerian tours is a special case of the ***Route Inspection Problem***.

2.2. Hamiltonian graphs

While Theorem 1 gives a simple necessary and sufficient condition for a graph to have an eulerian tour, no such simple condition is available for a graph to be hamiltonian. The problem of finding hamiltonian cycles is a special case of the ***Travelling Salesman Problem***. In this section, we consider a simple sufficient condition for a graph to be hamiltonian due to Dirac [**7**]:

THEOREM 2. (DIRAC'S THEOREM)
Let $n \geq 3$, and let G be an n-vertex graph of minimum degree at least $n/2$. Then G is hamiltonian.

PROOF. Suppose, for a contradiction, that there is a non-hamiltonian n-vertex graph of minimum degree at least $n/2$. Amongst all such graphs, let G be one with a maximum number of edges. If we add an edge $e = \{v_1, v_n\}$ between non-adjacent vertices of G, then we have a graph with a hamiltonian cycle C, and so $P = C - e$ is a hamiltonian uv-path in G, say $v_1 v_2 \dots v_n$. Let $N(v_1)^+ = \{v_{i+1} : v_i \in N(u)\}$ – this is the set of vertices which are immediately after neighbors of u on the path P. Then $N(v_1)^+ \cup N(v_n) \subseteq V(P) \backslash \{v_1\}$ as $\{v_1, v_n\} \notin E(G)$, so

$$|N(v_1)^+ \cup N(v_n)| \leq n - 1.$$

On the other hand, $|N(v_1)^+| + |N(v_n)| \geq n$, since G has minimum degree at least $n/2$. Therefore

$$|N(v_1)^+ \cap N(v_n)| = |N(v_1)^+| + |N(v_n)| - |N(v_1)^+ \cup N(v_n)| > 0.$$

Let $v_i \in N(v_1)^+ \cap N(v_n)$. Then $v_1 v_2 \dots v_{i-1} v_n v_{n-1} \dots v_i v_1$ is a hamiltonian cycle in G, as shown in Figure 2.2, a contradiction. So every n-vertex graph of minimum degree at least $n/2$ is hamiltonian. □

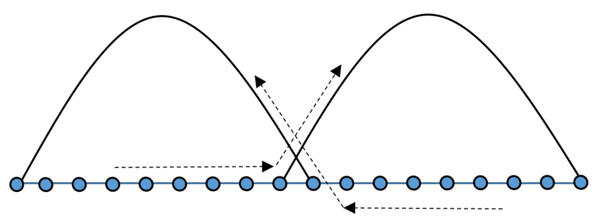


FIGURE 2.2. Finding a hamiltonian cycle

Let $k = \lfloor (n-1)/2 \rfloor$ – round $(n-1)/2$ down to the nearest integer. Then $G = K_{k,n-k}$ is non-hamiltonian, an has n vertices and minimum degree $k < n/2$. If n is odd, then the graph consisting of two copies of K_{k+1} joined at one vertex is also non-hamiltonian and has minimum degree k. These examples show that Theorem 2 is best possible – the condition on the minimum degree in the theorem cannot be lowered. The reader can check as an exercise that for $n \geq 2$, every n-vertex graph of minimum degree at least $n/2 - 1$ is traceable.

2.3. Pósa's Rotation Lemma

If P is path in a graph G, whose vertices are ordered from one end vertex u to the other end vertex v, then for a vertex $w \in V(P)$ we let w^+ be the vertex after w and w^- the vertex before w on the path. A key fact if P is a longest path in a graph G is that $N_G(u) \cup N_G(v) \subseteq V(P)$. Furthermore, for each $w \in N_G(v)$, the path

$Q = P - \{w, w^+\} + \{v, w\}$ has the same length as P and has $w^+ \in N(v)^+$ as an end vertex. The paths P and Q are obtained from one another by ***rotation at*** v, and both start at u. The idea of rotation was in fact behind the proof of Theorem 2. A key result is the ***rotation lemma*** due to Pósa [**20**]:

LEMMA 1. (PÓSA'S ROTATION LEMMA)
Let $t \geq 1$ and let G be a graph such that $|N(S)| > 2|S|$ for every $S \subseteq V(G)$ of size at most t. Then G contains a path of length at least $3t$.

PROOF. Let P be a longest path in G, ordered from one end u to the other end v. Let S be the set of vertices of P obtained by all possible rotations. The key observation is $N(S) \subseteq S^- \cup S^+$, which implies $|N(S)| \leq |S^-| + |S^+| \leq 2|S|$. It follows that $|S| \geq t + 1$. Let $T \subseteq S$ have size t. Then $|V(P)| \geq |T| + |N(T)| \geq 3t + 1$ so P has length at least $3t$. □

THEOREM 3. *Let G be a graph such that for every non-empty set $S \subseteq V(G)$,*

$$|N(S)| \geq \min\{2|S| + 1, |V(G) \backslash S|\}.$$

Then G is traceable.

PROOF. Let S be the set of vertices of a longest path P obtained by all possible rotations. Since $|N(S)| \leq |S^-| + |S^+| 2|S|$, we conclude $|N(S)| \geq |V(G) \backslash S|$. Since $S \cup N(S) \subseteq V(P)$, we obtain

$$|V(P)| \geq |S| + |N(S)| \geq |S| + |V(G) \backslash S| = |V(G)|.$$

Therefore P is a hamiltonian path in G, so G is traceable. □

2.4. Uniquely Hamiltonian Graphs

A graph is ***uniquely hamiltonian*** if it has exactly one hamiltonian cycle. The following theorem due to Smith for cubic graphs and Thomason in general gives an answer to this question when all the degrees of the vertices are odd: there are no uniquely hamiltonian graphs in which all the vertices have odd degree [**26**].

THEOREM 4. *Let G be a graph all of whose vertices have odd degree. Then there exist an even number of hamiltonian cycles containing any edge $e \in E(G)$. In particular, G is not uniquely hamiltonian.*

PROOF. Let $e = \{u, v\}$. If there are no hamiltonian cycles containing e, we are done. Suppose there is a hamiltonian cycle C containing e, and let $N_C(u) = \{v, w\}$ and $f = \{u, w\}$. Consider the hamiltonian uw-path $P = C - f$, ordered from u to w. Form a new graph H whose vertices are the hamiltonian paths of $G - f$ starting with the edge $\{u, v\}$, where two hamiltonian paths in G form an edge of H if they are obtained from one another by rotation (in Figure 2.3 below, the path P is shown in bold black edges). If Q is a hamiltonian path in H, starting with $\{u, v\}$ and ending at a vertex x, then there are exactly $d(x) - 1$ possible rotations, one for each edge containing x and not already used by Q, unless $\{u, x\}$ is an edge, in which case there are $d(x) - 2$ rotations (see Figure 2.3). In the latter case, Q together with e forms a hamiltonian cycle in G containing e. Since $d(x)$ is odd, $d(x) - 2$ is also odd. The number of vertices of H of odd degree is even, by Lemma 2, so there must be an even number of hamiltonian paths Q in $G - f$ which end at a neighbor of u (in Figure 2.3, H has three vertices, one corresponding to the path in bold black edges, one in dashed black edges, and one in dashed red edges, and H is a path of length two). Therefore G contains an even number of hamiltonian cycles containing e. $\square$

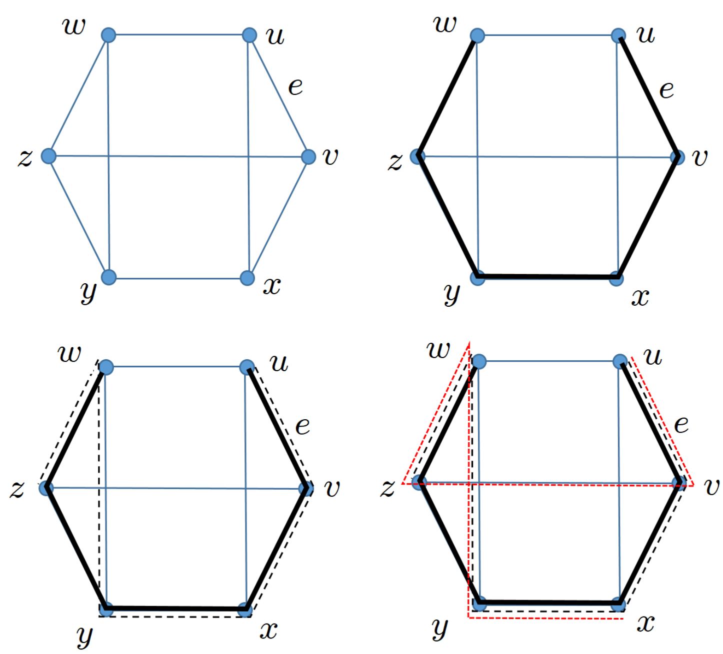


FIGURE 2.3. Finding a second hamiltonian cycle

2.5. Exercises

Question 2.5.1° Prove that if G is a connected graph with m edges such that two vertices $u, v \in V(G)$ have odd degree and all other vertices have even degree, then there exists a sequence $(v_0, v_1, v_2, \ldots, v_m)$ of vertices of G with $u = v_0$ and $v = v_m$ such that every edge of G appears exactly once as a pair $\{v_i, v_{i+1}\}$ (an ***eulerian trail***).

Question 2.5.2° A ***tournament*** is an orientation of a complete graph. Prove that every tournament contains a directed path containing all of its vertices.

Question 2.5.3° Let P be a longest path in a connected graph G, and suppose there exists a cycle C such that $P \subseteq C \subseteq G$. Prove that G is hamiltonian.

Question 2.5.4° For the ***Heawood graph*** shown below, draw the graph H from the proof of Theorem 4 where P is the hamiltonian path $(1, 2, 3, \ldots, 14)$. Then find a hamiltonian cycle different from $(1, 2, 3, \ldots, 14, 1)$.

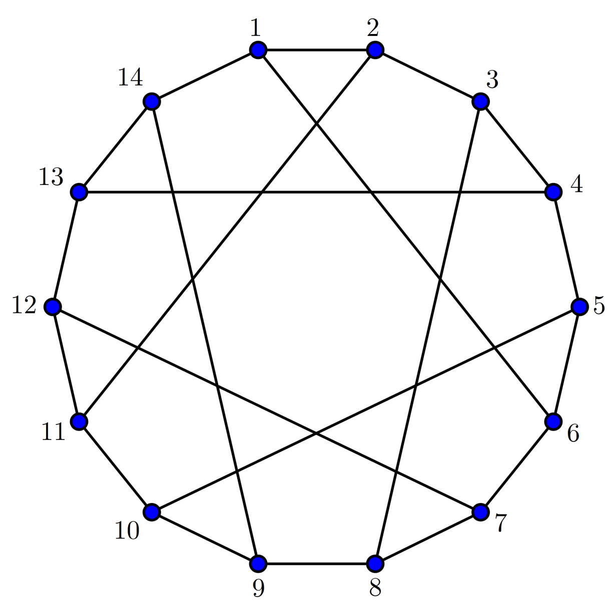


FIGURE 2.4. Finding a second hamiltonian cycle

Question 2.5.5° Let G be a connected cubic graph and $L(G)$ the ***line graph***[1] of G. Prove that $L(G)$ has an eulerian tour.

◇

[1]See Question 1.7.10.

Question 2.5.6. Prove that every graph has an orientation such that the difference between in and out degrees at each vertex is at most 1.

Question 2.5.7. Prove that a graph of minimum degree at least $k \geq 2$ containing no triangles contains a cycle of length at least $2k$.

Question 2.5.8. Let $n \geq 2k \geq 2$, and let G be an n-vertex graph such that for any two disjoint sets of k vertices, there is a perfect matching between the two sets. Prove that G is traceable.

Question 2.5.9. The ***closure*** of an n-vertex graph G, denoted $C(G)$, consists in adding edges between any two non-adjacent vertices u and v such that $d_G(u)+d_G(v) \geq n$. Prove that a graph G is hamiltonian if and only if $C(G)$ is hamiltonian.

Question 2.5.10. Let $n \geq 2$. Prove that an n-vertex graph with at least $\binom{n-1}{2} + 1$ edges is traceable. Give an example of an n-vertex graph with $\binom{n-1}{2}$ edges that is not traceable.

Question 2.5.11. Let G be a hamiltonian bipartite graph of minimum degree at least three. Prove that G contains at least two hamiltonian cycles.

◇

Question 2.5.12* Let C be a hamiltonian cycle in an eulerian graph G, and suppose every component of $G - E(C)$ is has an even number of vertices. Prove that there exists a hamiltonian cycle C' in G such that $C' \neq C$[2].

Question 2.5.13* Let G be an r-regular graph with $2r+1$ vertices, where $r \geq 2$. Prove that G is hamiltonian. Then give an example for each $r \geq 2$ of an r-regular graph with $2r+2$ vertices that is not hamiltonian.

Question 2.5.14* Let $k \geq 1$ and let G be a d-regular graph such that for any non-empty sets $S, T \subseteq V(G)$,

$$e(S,T) < \frac{d}{n}|S||T| + \frac{d}{3k}(|S|+|T|).$$

Prove that G contains a path of length at least $n(1 - 1/d - 1/k)$.

Question 2.5.15* Prove that if G is an eulerian digraph with m edges, then G contains a directed path of length at least $\sqrt{m/n}$.

[2]Question 1.7.16 may be helpful.

CHAPTER 3

Matching Theory

A ***matching*** in a graph is a set of pairwise vertex-disjoint edges of the graph. Any vertex contained in an edge of the matching is called ***saturated***, and any other vertex is called ***exposed***. In this section we are interested in determining the size of a ***maximum matching*** in a given graph and when a graph has a ***perfect matching*** or ***1-factor*** – that is, a matching covering all its vertices.

3.1. Hall's Theorem

Let X be a set of vertices in a multigraph G. We define $N_G(X)$ to be the ***neighborhood of*** X, namely

$$N_G(X) = \{y \in V(G)\backslash X : \{x,y\} \in E(G) \text{ for some } x \in X\}.$$

In other words, it is the set of vertices not in X adjacent to some vertex in X. Hall's Theorem gives a necessary and sufficient condition for a bipartite graph to have a perfect matching – and in fact a matching covering all vertices of one part. The condition (3.1) is known as ***Hall's Condition***.

Theorem 1. (Hall's Theorem)
Let G be a bipartite multigraph with parts A and B. Then G has a matching saturating A if and only if for every set $X \subset A$,

$$|N_G(X)| \geq |X| \tag{3.1}$$

In particular, if $|A| = |B|$ then G has a perfect matching.

Proof. If G has a matching M saturating A, then for every $X \subseteq A$ and there are $|X|$ neighbors of X in B via M. This gives Hall's condition. Now we suppose Hall's Condition is true, and show how to get a perfect matching in G. We proceed induction on $|A|$. If $|A| = 1$, then the statement is true. Suppose $|A| \geq 2$. We consider two cases.

Case 1. $|N_G(X)| > |X|$ *for all* $X \subset A$ *with* $X \neq A$.
In this case, pick any edge $\{a,b\}$ of G, and consider $H = G - \{a\} - \{b\}$. Then H has parts $A' = A\backslash\{a\}$ and $B\backslash\{b\}$. Furthermore, for any $X \subseteq A'$, $|N_H(X)| \geq |N_G(X)| - 1 \geq |X|$. Therefore Hall's Condition holds in H, so by induction, H has a matching M' saturating A'. Now add the edge $\{a,b\}$ to this matching to get a matching saturating A.

Case 2. $|N_G(X)| = |X|$ for some $X \subset A$ with $X \neq A$.
Let $Y = N_G(X)$. Let $I = G - (X \cup Y)$ and $H = G[X \cup Y]$. Then for any $Z \subseteq X$, by

(3.1) in G,

$$|N_H(Z)| = |N_G(Z)| \geq |Z|$$

so Hall's Condition holds in H. Therefore there is a matching M_H saturating X in H, by induction. For any $Z \subseteq A \backslash X$, by (3.1) in G,

$$|N_I(Z)| + |N_G(X)| = |N_G(Z \cup X)| \geq |Z \cup X| = |Z| + |X|.$$

However, $|N_G(X)| = |X|$ so we get

$$|N_I(Z)| \geq |Z|.$$

Therefore Hall's Condition holds in the graph I, so I has a matching M_I saturating $A \backslash X$. Finally, $M_H \cup M_I$ is a matching in G saturating A. $\square$

3.2. König-Ore Theorem

Hall's Theorem gives a formula for finding the size $\mu(G)$ of a maximum matching in a bipartite multigraph. For a bipartite graph $G(A, B)$, define $\mathrm{ex}(G, A) = |A| - \mu(G)$: this is the number of vertices of A exposed by a maximum matching. Hall's Theorem gives a formula for $\mathrm{ex}(G, A)$:

Theorem 2. (König-Ore Formula)
Let G be a bipartite multigraph with parts A and B. Then

$$\mathrm{ex}(G, A) = \max_{S \subset A} \{|S| - |N(S)|\}.$$

Proof. Let d be the right hand side of the identity above. Add d vertices to B, all adjacent to all vertices of A. Then Hall's Condition – namely $|N(X)| \geq |X|$ for all $X \subset A$ – is satisfied in this new graph, so it has a matching covering all vertices of A, by Hall's Theorem. It follows that G has a matching of size at least $|A| - d$. Therefore $\mathrm{ex}(G, A) \leq d$. Conversely, if M is a matching of size $|A| - \mathrm{ex}(G, A)$, then each set $S \subset A$ has at least $|S| - \mathrm{ex}(G, A)$ neighbours in B. In other words, $|N(S)| \geq |S| - \mathrm{ex}(G, A)$ for all S so $d = \max_{S \subset A}\{|S| - |N(S)|\} \leq \mathrm{ex}(G, A)$. $\square$

A ***1-factorization*** of a graph G is a collection of pairwise edge-disjoint 1-factors $M_1, M_2, \dots, M_r$ such that $G = M_1 \cup M_2 \cup \dots \cup M_r$. For example, for even values of n, the complete graph K_n has a 1-factorization.

Corollary 1. *Let $k \geq 1$ and let G be a k-regular bipartite multigraph. Then G has a 1-factorization.*

Proof. Let A and B be the parts of G. It suffices to prove that G has a perfect matching. To see this, we apply Hall's Theorem. For a set $X \subset A$ or $X \subset B$, there are $k|X|$ edges of G incident with exactly one vertex of X. There are also $k|N(X)|$ edges incident with exactly one vertex of $N(X)$. This set of edges contains all edges incident with X, so $k|N(X)| \geq k|X|$ and Hall's Condition is satisfied, and G has a perfect matching. $\square$

3.3. Tutte's 1-Factor Theorem

There is a natural condition for a (not necessarily bipartite) graph G to have a perfect matching: if S is a set of vertices of G and $H_1, H_2, \ldots, H_r$ are the ***odd components*** of $G - S$ – that is the components with an odd number of vertices – then none of the H_i can have a perfect matching, so each sends at least one edge of a perfect matching to S. In particular $|S| \geq r$, so we have for all $S \subset V(G)$, denoting by odd(G - S) the odd components of $G - S$,

$$|S| \geq \mathrm{odd}(G - S).$$

This is known as ***Tutte's Condition***. Note that if $S = \emptyset$, this asserts that G has an even number of vertices. Tutte's Theorem **[28]** shows, remarkably, that this is also a sufficient condition:

Theorem 3. (Tutte's 1-Factor Theorem)
Let G be a multigraph. Then G has a perfect matching if and only if for every set $S \subset V(G)$,

$$|S| \geq \mathrm{odd}(G - S). \tag{3.2}$$

Proof. If G has a perfect matching M, then for any $S \subseteq V(G)$, every odd component F of $G - S$, there is at least one exposed vertex of F for the matching $M \cap E(F)$. Each exposed vertex is adjacent in M to a vertex of S, so $|S| \geq \mathrm{odd}(G - S)$, which is Tutte's Condition. Now suppose G satisfies Tutte's Condition; we show how to find a perfect matching in G.

The proof we give is by induction on $|V(G)|$, the case $|V(G)| = 2$ holds since $G = K_2$ in that case. Suppose $|V(G)| > 2$, and let S be the largest subset of G such that equality holds in Tutte's Condition. Such a set S exists, because $|V(G)|$ is even, and so $G - \{s\}$ has at least one odd component for each $s \in V(G)$. Let F and H denote generic odd and even components of $G - S$.

Claim 1. *The graph H has a 1-factor.*

For any $R \subset V(H)$, we note

$$\mathrm{odd}(G - (R \cup S)) = \mathrm{odd}(H - R) + \mathrm{odd}(G - S)$$

since every odd component of $G - S$ is an odd component of $G - (R \cup S)$. By Tutte's Condition, $\mathrm{odd}(G - R \cup S) \leq |R| + |S|$. Since $\mathrm{odd}(G - S) = |S|$, we conclude $\mathrm{odd}(H - R) \leq |R|$ for all $R \subset V(H)$. By induction H has a 1-factor.

Claim 2. *The graph $F' = F - \{v\}$ has a 1-factor for any $v \in V(F)$.*

By induction, if this is false, then there exists a set $Q \subset V(F')$ such that $\mathrm{odd}(F' - Q) > |Q|$. Now for any set $R \subset V(F)$,

$$\mathrm{odd}(F - R) + |R| \equiv |V(F)| \equiv 1 \mod 2$$

since F has an odd number of vertices (this step is really key to the proof). Therefore $\mathrm{odd}(F' - Q) \geq |Q| + 2$. We also observe

$$\mathrm{odd}(G - S \cup \{v\} \cup Q) = \mathrm{odd}(G - S) - 1 + \mathrm{odd}(F' - Q)$$

since F is an odd component of $G-S$ but not of $G-S\cup\{v\}\cup Q$. If $T=S\cup\{v\}\cup Q$, then by Tutte's Condition, we get

$$\begin{aligned}|T| &\geq \operatorname{odd}(G-T)\\ &= \operatorname{odd}(G-S)-1+\operatorname{odd}(F'-Q) \quad \geq |S|+|Q|+1.\end{aligned}$$

This shows $\operatorname{odd}(G-T)=|T|$, contradicting the maximality of S, and the claim is proved.

Claim 3. *Let $G(S,C)$ be the bipartite graph formed from G by contracting each odd component of $G-S$ to a single vertex, and taking all edges with one end in S and one end in the set C of contracted vertices. Then $G(S,C)$ has a perfect matching.*

To prove this, we use Hall's Theorem: for every set $X\subset C$,

$$|X|=\operatorname{odd}(G-N(X))\leq |N(X)|$$

as required. Since $|S|=|C|=\operatorname{odd}(G-S)$, there is a 1-factor in $G(S,C)$.

To complete the proof of Tutte's 1-Factor Theorem, put together all the 1-factors that we found in Claims 1–3. Let $M_1,M_2,\ldots,M_r$ be 1-factors in the even components of G. Now let M be a 1-factor in $G(S,C)$. Then the edges of M form a matching in G, and for each odd component H_i of $G-S$, for $i\in\{1,2,\ldots,s\}$ where $s=\operatorname{odd}(G-S)$, there is exactly one vertex of H, say v_i, incident with an edge of M. Now Claim 2 gives a 1-factor N_i in $H-v_i$. Then

$$M\cup M_1\cup\cdots\cup M_r\cup N_1\cup N_2\cup\cdots\cup N_s$$

is a perfect matching of G. □

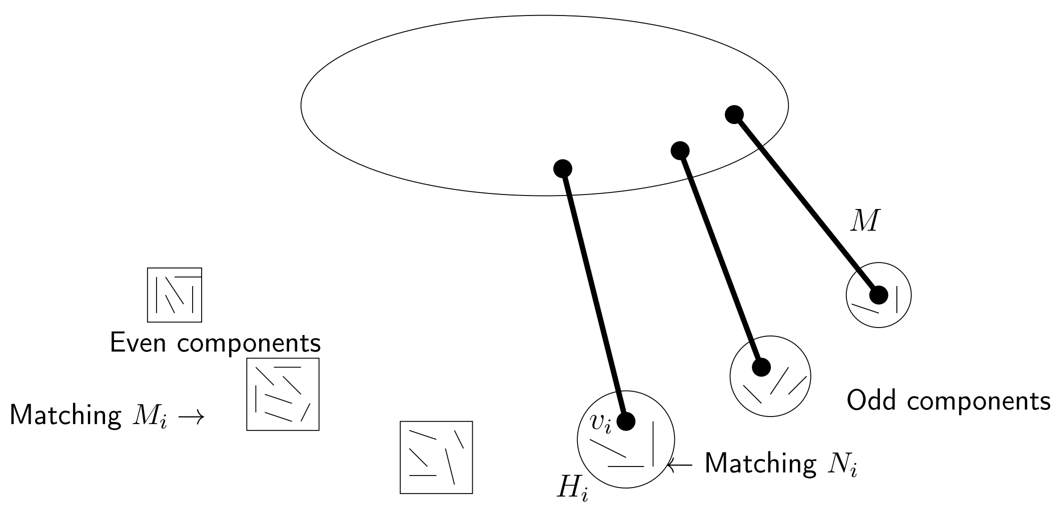


FIGURE 3.1. The proof of Tutte's Theorem

From Tutte's 1-Factor Theorem, we obtain the following condition for a ***cubic graph*** (3-regular graph) to have a perfect matching. A graph G is ***bridgeless*** if $G-e$ is connected for every edge $e\in E(G)$.

THEOREM 4. (PETERSEN'S THEOREM)
Any cubic bridgeless multigraph has a 1-factor.

PROOF. We have to check Tutte's Condition. Pick a set $S \subset V(G)$. If $S = \emptyset$, then Tutte's Condition holds since G has an even number of vertices and is connected. If $S \neq \emptyset$, then since G is bridgeless, there are at least two edges from S to each odd component of $G - S$. If H is an odd component of $G - S$, then it contains an even number of vertices of degree three, so it must send to S an odd number of edges – at least three. If $r = \mathrm{odd}(G-S)$, then we have $3r$ edges sent to S from odd components. On the other hand, G is cubic so S cannot accept more than $3|S|$ edges. So $3|S| \geq 3r$ and $|S| \geq r = \mathrm{odd}(G - S)$, which is Tutte's Condition. □

3.4. Tutte-Berge Formula

The ***Tutte-Berge Formula*** [**3**, **28**] generalizes Tutte's 1-Factor Theorem and the König-Ore Formula to finding the size of a ***maximum matching*** in a graph. We define $\mathrm{ex}(G)$ to be the minimum number of vertices of G exposed by a matching of G – thus $\mathrm{ex}(G) = |V(G)| - 2m$ where m is the number of edges in a maximum matching of G.

THEOREM 5. (TUTTE-BERGE FORMULA)
For any multigraph G,

$$\mathrm{ex}(G) = \max_{S \subset V(G)} \{\mathrm{odd}(G - S) - |S|\}. \tag{3.3}$$

A natural question is the size $\mu(G)$ of a largest matching in a connected cubic graph G. The following theorem answers the question, and may be generalized to r-regular graphs.

THEOREM 6. *Let G be a cubic multigraph on n vertices. Then G has a matching of size at least $2n/5$.*

Theorem 6 is best possible: the multigraph shown in Figure 3.2 is cubic with $n = 10$ vertices with no matching of size more than $4 = 2n/5$.

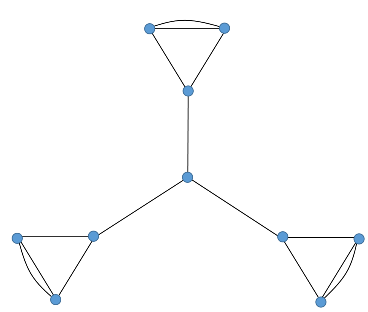


FIGURE 3.2. A cubic multigraph with no perfect matching

3.5. Tutte's f-Factor Theorem

Let G be a graph and $f : V(G) \to \{0, 1, 2, \dots\}$ a function. An ***f-factor*** of G is a spanning subgraph H of G such that $d_H(v) = f(v)$ for all $v \in V(G)$. Hall's Theorem can be converted to a necessary and sufficient condition for a bipartite graph to have an f-factor, as follows [**29**]:

THEOREM 7. (f-FACTOR THEOREM) *Let G be a bipartite graph with parts A and B. Then G has an f-factor if and only if for any two sets $X \subseteq A$ and $Y \subseteq B$,*

1. $$\sum_{u \in A} f(u) = \sum_{v \in B} f(b)$$
2. $$\sum_{x \in X} f(x) \leq e(X, Y) + \sum_{y \in B \setminus Y} f(y)$$

A similar extension of Tutte's 1-Factor Theorem to an f-Factor Theorem can be achieved, however the statement is more complicated.

3.6. Matching Algorithms

An ***alternating path*** in a graph G with a matching M is a path whose every alternate edge is in M – we call this ***M-alternating***. An ***augmenting path*** for a matching M is an alternating path whose first and last edges are not in M – we call this ***M-augmenting***. In the figure below, with a matching M shown in bold edges, $(1, 7, 6, 8)$ is an M-alternating path, whereas $(9, 4, 12, 6, 7, 2)$ is an M-augmenting path.

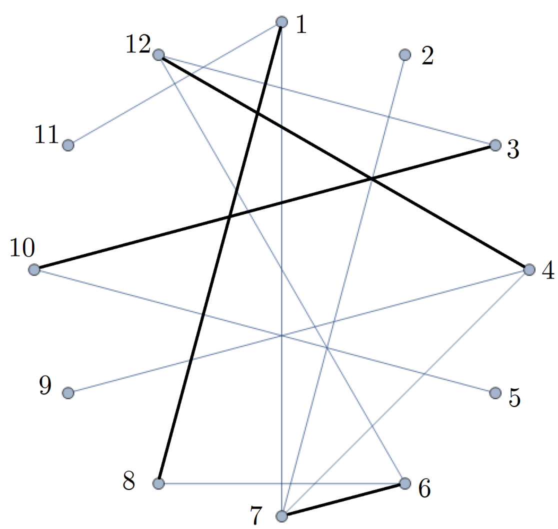


FIGURE 3.3. Alternating and augmenting paths

THEOREM 8. (BERGE)
A matching M in a graph G is a maximum matching if and only if M does not admit any augmenting paths.

This theorem shows that if we want to find a maximum matching in a graph, we should seek an augmenting path with respect to a given matching M. If no augmenting path exists, then M is a maximum matching, otherwise we augment M by swapping the edges of an augmenting path in M with the edges not in M. In particular, if U is the set of exposed vertices for M, any augmenting path has both ends in U, and can be constructed via a breadth-first search type algorithm. The algorithm for bipartite graphs is particularly simple, and known as the ***Hungarian Algorithm*** or ***Kuhn-Munkres Algorithm***, whereas ***Edmonds' Matching Algorithm*** handles the case of graphs that are not necessarily bipartite.

3.7. Edge coloring

A ***proper k-edge-coloring*** of a graph G is a function $\chi : E(G) \to \{1, 2, \dots, k\}$ such that if $e, f \in E(G)$ intersect, then $\chi(e) \neq \chi(f)$. In other words, any two edges which share a vertex must receive different colors. The minimum k for which G has a proper k-edge-coloring is the ***edge-chromatic number of G***, denoted $\chi'(G)$. A graph G is ***k-edge colorable*** if $\chi'(G) \leq k$, and ***k-edge-chromatic*** if $\chi'(G) = k$.

Thus $\chi'(G)$ is the minimum number of matchings whose union is $E(G)$, since the edges of each color form a matching. For instance, Corollary 1 gives $\chi'(G) = k$ when G is a k-regular bipartite graph. A more difficult problem is to show $\chi'(K_n) = n$ if n is odd and $\chi'(K_n) = n - 1$ if n is even. It is clear in general that $\chi'(G) \geq \triangle(G)$. ***König's Theorem*** states that $\chi'(G) = \triangle$ for any bipartite graph G – thus determining $\chi'(G)$ in bipartite graphs is easy:

THEOREM 9. (KÖNIG'S THEOREM)
For any bipartite graph G of maximum degree $\triangle$, $\chi'(G) = \triangle$.

PROOF. By induction on $|E(G)|$. If $|E(G)| = 0$ then the theorem is clear. Suppose $|E(G)| > 0$ and let $e = \{x, y\} \in E(G)$. By induction, the graph $G - e$ is $\triangle$-edge-colorable. If there is a color i which is not used on any edges incident with x or y, then we can assign color i to $\{x, y\}$ to get a $\triangle$-edge-coloring of G. So we may assume that the colors at x are $1, 2, \dots, \triangle - 1$ and the colors at y are $2, 3, \dots, \triangle$. Let H be the subgraph of G spanned by edges of colors 1 and $\triangle$. Then the component of H containing x is a path or a cycle. It cannot be a cycle, otherwise x would be incident with an edge of color 1 and color $\triangle$ in the cycle, contradicting that $\triangle$ is missing at x. So the component of H containing x is a path, P. If P ends at y, then since P has odd length we would have an edge of color 1 at y, a contradiction. So P ends at a vertex $z \neq y$. Now z is not incident with any edge of color 1 or $\triangle$ in $G - E(P)$, otherwise we could extend the path or the edge is incident with a vertex w of the path, but then the coloring would not be a proper edge-coloring. Now interchange colors 1 and $\triangle$ along the path P, to obtain a proper coloring of $G - e$ where the color 1 does not appear at x. Finally, assign e color 1 to get a proper coloring of G. $\square$

The next remarkable theorem tells us that for every graph G, $\chi'(G)$ is either $\triangle(G)$ or $\triangle(G) + 1$:

Theorem 10. (Vizing's Theorem)
For every graph G of maximum degree $\triangle$, $\chi'(G) = \triangle$ or $\chi'(G) = \triangle + 1$.

Graphs G with maximum degree $\triangle$ and $\chi'(G) = \triangle$ are referred to as ***class 1*** graphs, whereas those with $\chi'(G) = \triangle + 1$ are ***class 2*** graphs. A good characterization of these classes of graphs is not available – in fact classifying a graph as class 1 or class 2 is known to be difficult from the standpoint of computational complexity.

3.8. Exercises

Question 3.8.1° A school with 20 professors forms 10 committees, each containing 6 professors, such that every professor is on exactly 3 committees. Prove that it is possible to s elect a distinct representative from each committee.

Question 3.8.2° A ***tiling*** of an $m \times n$ chess board is a set of dominoes which cover all the squares on the chess board exactly once (each domino covers two adjacent squares).

(a) For which $m \geq 1$ and $n \geq 1$ does an $m \times n$ chess board having a tiling?
(b) If we remove two squares from an $m \times n$ chessboard, when do the remaining squares have a tiling?

Question 3.8.3° Let e be an edge of a connected cubic graph such that $G - e$ is disconnected. Prove that every perfect matching of G contains e.

Question 3.8.4° Is it possible for a cubic hamiltonian graph to have exactly one 3-edge-coloring?

Question 3.8.5° Determine $\chi'(G)$ when G is the Petersen graph.

Question 3.8.6° Classify as class 1 or class 2 all connected graphs with at most five vertices.

◇

Question 3.8.7.

(a) Prove that a tree has at most one perfect matching.
(b) Show that a tree has a perfect matching if and only if $\text{odd}(T-x) = 1$ for every $x \in V(T)$.

Question 3.8.8.

(a) Let G be an n by n bipartite graph of minimum degree more than $n/2$. Prove that G has a perfect matching.
(b) Let G be a $2n$-vertex graph of minimum degree at least n. Prove that G has a perfect matching.

Question 3.8.9. Let A_k be the set of subsets of $\{1, 2, \ldots, n\}$ of size k. Prove that for $k < n/2$, there is an injective function $f : A_k \to A_{k+1}$ such that $a \subseteq f(a)$ for all $a \in A_k$. For instance, if $k = 1$ and $n = 3$ then the function

$$f(\{1\}) = \{1, 2\} \quad f(\{2\}) = \{2, 3\} \quad f(\{3\}) = \{1, 3\}$$

is an example of such a function $f : A_1 \to A_2$.

Question 3.8.10. A ***latin square*** is a square array of symbols such that every symbol appears exactly once in every row and exactly once in every column.

(a) Give an example of an $n \times n$ array with n cells filled in with $1, 2, \ldots, n$ that cannot be completed to a latin square.
(b) Prove that if we fill in the first $r \leq n$ rows of an $n \times n$ array with symbols from $\{1, 2, \ldots, n\}$ so that no symbol appears more than once in every row or column, then the array can be completed to a latin square.

Question 3.8.11. Let A be an n by n matrix of zeros and ones. Suppose every row and every column of A has exactly k ones. Prove that we can pick n ones from A, no two in the same row or column.

Question 3.8.12. Prove that if $k \geq 1$ is odd and G is a k-regular $(k-1)$-edge-connected graph, then G has a perfect matching.

Question 3.8.13. Determine $\chi'(K_n)$ for all $n \geq 2$.

Question 3.8.14. Prove that every n-vertex bipartite multigraph G of maximum degree d is contained in a d-regular bipartite multigraph with $2n$ vertices. Deduce $\chi'(G) = d$.

Question 3.8.15. Let G be an n-vertex 4-regular multigraph. Prove that G has a matching with at least $n/3$ edges, and when n is a multiple of 3, describe 4-regular graphs with no larger matchings.

Question 3.8.16. Let $S_1, S_2, \ldots, S_n$ be sets such that for any $I \subseteq \{1, 2, \ldots, n\}$,

$$\left|\bigcup_{i \in I} S_i\right| \geq |I|.$$

Prove that it is possible to pick a ***system of distinct representatives*** for the sets S_i, namely distinct elements $s_1 \in S_1$, $s_2 \in S_2, \ldots, s_n \in S_n$.

Question 3.8.17. An ***independent set*** in a graph G is a set $X \subseteq V(G)$ such that $e(X) = 0$, and the ***independence number*** $\alpha(G)$ is the largest size of an independent set in G. A ***vertex cover*** of G is a set of vertices $X \subset V(G)$ such that $e \cap X \neq \emptyset$ for every edge $e \in E(G)$. The minimum size of a vertex cover of G, the ***vertex cover number***, is denoted $\beta(G)$.
(a) Prove that for any graph G, $\alpha(G) + \beta(G) = |V(G)|$.
(b) Prove that $\mu(G) \leq \beta(G) \leq 2\mu(G)$.

Question 3.8.18. Prove that a matching M in a graph is a maximum matching if and only if there is no augmenting path for M.

◇

Question 3.8.19* Suppose we fill in fewer than $n/2$ cells in an $n \times n$ array with symbols from $\{1, 2, \ldots, n\}$ so that no symbol appears more than once in every row or column. Prove that the array can be completed to a latin square.

Question 3.8.20* Prove that a bipartite graph with minimum degree at least d containing a perfect matching contains at least $d!$ perfect matchings. Is this best possible?

Question 3.8.21* Prove that a cubic $n \times n$ bipartite graph contains at least $(4/3)^n$ perfect matchings.

Question 3.8.22* Let $n = 2k+1$ and let A_k be the family of subsets of $\{1, 2, \ldots, n\}$ of size k. Define an injective function $f : A_k \to A_{k+1}$ such that $f(a) \subseteq a$ for all $a \in A_k$.

Question 3.8.23* An ***edge cover*** of G is a set E of edges of G such that for every vertex $v \in V(G)$, there is an edge of E containing v. The minimum size of an edge-cover, the ***edge cover number***, is denoted $\rho(G)$. Let $i(G)$ be the number of vertices of degree zero – ***isolated vertices*** – of G. Prove that $\mu(G) + \rho(G) + i(G) = |V(G)|$.

CHAPTER 4

Planar Graphs and Coloring

An ***embedding*** of a graph $G = (V, E)$ is a function $f : V \cup E \to \mathbb{R}^2 \cup \mathcal{C}$, where $\mathcal{C}$ is the set of continuous curves in $\mathbb{R}^2$, such that f is one-to-one, $f(v)$ is a point in $\mathbb{R}^2$ for each $v \in V$, and $f(\{u, v\})$ is a continuous curve in $\mathbb{R}^2$ with ends u and v when $\{u, v\} \in E$. The graph G is ***planar*** if we can choose f so that the curves $f(e) : e \in E$ meet only at their ends – that is no curve meets itself and any point in the intersection of two distinct curves is an endpoint of both of the curves. A drawing of G without crossings is called a ***plane embedding*** of G, or a ***plane graph***. Thus a graph is planar if and only if it has a plane embedding. The main theorem classifying planar graphs is ***Kuratowski's Theorem***. A ***subdivision*** of a graph H is a graph obtained by replacing each edge $e = \{u, v\}$ with a path P_e with ends u and v such that $V(P_e) \cap V(P_f) = e \cap f$ for all distinct $e, f \in E(G)$.

THEOREM 1. (KURATOWSKI'S THEOREM)
A graph is planar if and only if it does not contain a subdivision of K_5 or $K_{3,3}$.

We do not prove this theorem here, as it is beyond the scope of the course.

4.1. Faces and boundaries

If G is a plane graph, then $\mathbb{R}^2 \backslash G$ consists of a union of disjoint connected plane regions, which are called ***faces*** of G. We denote by $F(G)$ the set of faces of a plane graph G. Each finite plane graph has a unique face which has infinite area, which we refer to as the ***infinite face***. The ***boundary*** ∂F of a face F of G is the set of points in the topological closure of F which are not in the interior of F. Suppose the connected components of ∂F are $\gamma_1, \gamma_2, \ldots, \gamma_k$. A ***boundary walk*** of F is a shortest closed walk $(v_0, v_1, \ldots, v_d, v_0)$ of vertices of γ_i containing all the vertices of γ_i. The ***degree*** $d(F)$ of a face $F \in F(G)$ is the sum of the lengths of boundary walks of F.

EXAMPLE 4.1. The graph on the left in Figure 4.1 has six faces, all boundary walks of which are cycles of length four – so every face has degree four. The tree in the centre has only one face – the infinite face – and since a tree on n vertices has $n - 1$ edges and the boundary walk goes through each edge twice, the degree of the infinite face is $2(n - 1)$. In the graph on the right, there are two faces, one of degree six and one of degree ten.

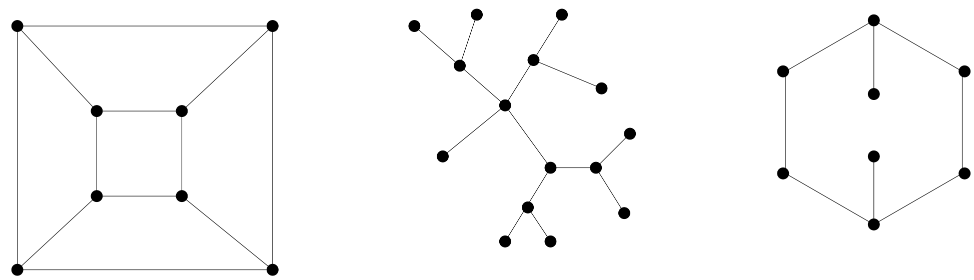


FIGURE 4.1. Faces of a plane graph

The degrees of the faces in a plane graph depend very much on the way the graph is drawn in the plane. There is a very useful analog of the handshaking lemma for face degrees in a plane graph. If we add up the degree of every face $F \in F(G)$, we observe that every edge of the graph is counted exactly twice:

LEMMA 1. *Let G be a plane graph. Then*

$$\sum_{F \in F(G)} d(F) = 2|E(G)|.$$

Euler's Formula relates $|F(G)|$, $|E(G)|$ and $|V(G)|$ as follows:

THEOREM 2. (EULER'S FORMULA)
Let G be a plane graph with k components. Then

$$|V(G)| - |E(G)| + |F(G)| = k + 1.$$

PROOF. We add $k-1$ edges to G to get a plane graph H. Suppose H has n vertices, m edges and f faces. We show $n - m + f = 2$ by induction on m. Since H is connected, $m \geq n - 1$ by Lemma 4. If $m = n - 1$, then H is a tree and $f = 1$, so $n - m + f = 2$. Now suppose $m \geq n$, in which case H contains a cycle C. Let e be an edge of C, and note $H - e$ is a connected plane graph with $m - 1$ edges and $f - 1$ faces. By induction, $n - (m - 1) + (f - 1) = 2$. which gives $n - m + f = 2$, and the proof for H is complete. Now $|V(G)| = n$, $|E(G)| = m - k + 1$ and $|F(G)| = f$, so $n - m + f = k + 1$. □

4.2. Non-planar graphs

A useful application of Euler's Formula is to give a simple necessary condition on the number of edges for planarity.

THEOREM 3. *Let G be a planar graph containing a cycle. Then*

$$|E(G)| \leq \frac{g}{g-2}(|V(G)| - 2), \tag{4.1}$$

where g is the length of a shortest cycle in G. In particular, for any planar graph G,

$$|E(G)| \leq 3|V(G)| - 6. \tag{4.2}$$

PROOF. Since G has a cycle, it has at least two faces. So for any face f, there is an edge e on the boundary of f and some other face of G. This means e is not a bridge in the graph consisting of vertices and edges in a closed walk containing e on the boundary of f, so this walk contains a cycle containing e. This shows $d(f) \geq g$ and so every face has degree at least g. Lemma 1 gives $g|F(G)| \leq 2|E(G)|$. Putting this in Euler's Formula, we get

$$|V(G)| - |E(G)| + \frac{2}{g}|E(G)| \geq 2$$

which, rearranged, gives the required bound on $|E(G)|$. The right side of the formula is maximized when $g = 3$, in which case we get $|E(G)| \leq 3|V(G)| - 6$. □

In particular, K_5 is not planar since $|E(K_5)| = 10$ and $g = 3$, and $K_{3,3}$ is not planar since $|E(K_{3,3})| = 9$ and $g = 4$. A ***maximal planar*** graph is a graph that is planar but the addition of any edge results in a non-planar graph. A ***maximal plane*** graph is a plane drawing of a maximal planar graph. Evidently, every face in a maximal plane graph with at least three vertices is a triangle (i.e. has degree three). Using Lemma 1 one can show that a maximal planar graph with $n \geq 3$ vertices has exactly $3n - 6$ edges and $2n - 4$ faces.

THEOREM 4. *Every planar graph has a vertex v of degree at most five.*

PROOF. This is clearly true for graphs G with at most six vertices. By Theorem 3, if G is planar with more than six vertices, then $|E(G)| \leq 3|V(G)| - 6$. By the handshaking lemma,

$$\sum_{v \in V(G)} d_G(v) \leq 6|V(G)| - 12 < 6|V(G)|.$$

This means some vertex v has $d_G(v) \leq 5$. □

4.3. Vertex Coloring

A ***proper k-coloring*** of a graph G is a function $\chi : V(G) \to \{1, 2, \ldots, k\}$ such that if $u, v \in V(G)$ are adjacent, then $\chi(u) \neq \chi(v)$. So we color the vertices with k colors in such a way that no two adjacent vertices have the same color. The ***chromatic number*** of G is denoted $\chi(G)$, and is the minimum k for which G has a proper k-coloring. For example, $\chi(K_n) = n$, and a graph G is bipartite if and only if $\chi(G) \leq 2$. We say that a graph is ***k-colorable*** if $\chi(G) \leq k$ and ***k-chromatic*** if $\chi(G) = k$.

EXAMPLE 4.2. Consider the ***Grötsch graph*** G below.

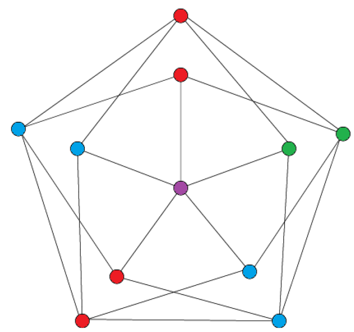


FIGURE 4.2. Proper coloring of the Grötsch graph.

We prove that $\chi(G) = 4$. A proper 4-coloring is shown, so $\chi(G) \leq 4$. To show that 4 colors are needed, we proceed as follows. Consider the "outer" cycle of length five. We know that 3 colors are needed to color this cycle, and we may assume that the colors around the cycle are red, blue, red, blue, green. If we are only allowed three colors, then the color of each vertex adjacent to the central vertex must be the same as its partner on the outer cycle. However, that means we used three colors in the neighborhood of the central vertex, so the central vertex must have a fourth color (purple in the picture).

Suppose that the vertices of a graph G can be ordered $(v_1, v_2, \ldots, v_n)$ so that v_i has at most d neighbors v_j with $j < i$. In particular, such an ordering exists if and only if the graph has no subgraph of minimum degree at least $d+1$, and we refer to such graphs as ***d-degenerate***. For example, Theorem 4 says planar graphs are 5-degenerate.

LEMMA 2. *Let G be a d-degenerate graph. Then $\chi(G) \leq d+1$.*

PROOF. Order the vertices $(v_1, v_2, \ldots, v_n)$ so that v_i has at most d neighbors v_j with $j < i$. Then assign v_1 color 1, and in general assign v_i the first color from $\{1, 2, \ldots, d+1\}$ that has not appeared on a neighbor v_j of v_i with $j < i$. Then this is a proper $(d+1)$-coloring of G, so $\chi(G) \leq d+1$. $\square$

This lemma together with Theorem 4 shows that planar graphs are 6-colorable. The famous 4-color theorem asserts more: every planar graph is 4-colorable. For general graphs, a central theorem on graph coloring is ***Brooks' Theorem***:

THEOREM 5. (BROOKS' THEOREM)
Let G be a connected graph of maximum degree $\triangle$. Then $\chi(G) \leq \triangle$ unless $G = K_{\triangle+1}$ or G is an odd cycle.

PROOF. We sketch the proof by induction on $|V(G)|$. If $\triangle \leq 2$, then G is a cycle or a path, and the theorem is evident. Suppose $\triangle \geq 3$. We may assume $G - \{v\}$ is connected for every $v \in V(G)$, otherwise $G = G_1 \cup G_2$ with $V(G_1) \cap V(G_2) = \{v\}$, and an appropriate inductive $\triangle$-coloring of G_1 and G_2 gives a $\triangle$-coloring of G. If $G \neq K_{\triangle+1}$, then there exist vertices v_1, v_{n-1} and v_n of G such that $\{v_1, v_{n-1}\}, \{v_1, v_n\} \in E(G)$ while $\{v_{n-1}, v_n\} \notin E(G)$. If $H = G - \{v_{n-1}\} - \{v_n\}$ is connected, then we may order the vertices $(v_1, v_2, \ldots, v_{n-2})$ of H so that for each $i \geq 2$, v_i has a neighbor v_j for some $j < i$. Now color v_n and v_{n-1} with color 1, and then color v_i with the first available color from $\{1, 2, \ldots, \triangle\}$ for $i \geq 2$. Since at most $\triangle - 1$ neighbors of v_i have been colored, this is possible. Finally, v_1 can be colored since color 1 is used on neighbors v_n and v_{n-1} of v_1, so only $\triangle - 1$ colors have been used in $N_G(v_1)$. Therefore $\chi(G) \leq \triangle$.

If H is not connected, then $G = G_1 \cup G_2$ where $V(G_1) \cap V(G_2) = \{v_{n-1}, v_n\}$. Furthermore, since $G - \{v_1\}$ and $G - \{v_2\}$ are connected, $G_1 + \{v_{n-1}, v_n\}$ and $G_2 + \{v_{n-1}, v_n\}$ have maximum degree at most $\triangle$. We note neither of G_1 nor G_2 is $K_{\triangle+1}$, so by induction, $\chi(G_1) \leq \triangle$ and $\chi(G_2) \leq \triangle$, in which case we combine colorings of G_1 and G_2 to get a coloring of G. $\square$

4.4. The Art Gallery Theorem

Let R be a closed connected region in the plane bounded by an n-sided polygon. Two points of R are ***mutually visible*** if there exists a straight line segment between the two points that is entirely contained in R. The ***art gallery problem*** is to determine the minimum size $f(R)$ of a set S of points in R such that for any point $x \in R$, there is a point $y \in S$ such that x and y are mutually visible. Evidently, if $f(R) \leq n$ for any region R bounded by an n-sided polygon. The art gallery theorem of Chvátal [**6**] gives the optimal value of $f(R)$:

THEOREM 6. (ART GALLERY THEOREM)
Let $n \geq 3$. For every n-sided polygonal region R, $f(R) \leq \lfloor n/3 \rfloor$. Furthermore, there exists an n-sided polygonal region R such that $f(R) = \lfloor n/3 \rfloor$.

PROOF. Let us ***triangulate*** the region R. In other words, we add straight lines between vertices of the boundary to obtain a plane graph G where every face other than the infinite face is a triangle inside R. We prove by induction on n that $\chi(G) = 3$. For $n = 3$, this is clear, since $G = K_3$. Suppose $n \geq 4$. Then some edge $e = \{u, v\}$ of G is not on the boundary of the infinite face of G. Then $\{u, v\}$ is a 2-vertex cut of G, so $G = G_1 \cup G_2$ where G_1 is a triangulation of a region R_1 and G_2 is a triangulation of a region R_2 such that $R_1 \cup R_2 = R$. Since both G_1 and G_2 have at least three vertices, $\chi(G_1) = 3$ and $\chi(G_2) = 3$. If $c_i : V(G_i) \to \{1, 2, 3\}$ is a proper 3-coloring of G_i, we can ensure $c_i(u) = 1$ and $c_i(v) = 2$ for $i \in \{1, 2\}$. Now let $c(x) = c_i(x)$ if $x \in V(G_i)$. Then $c : V(G) \to \{1, 2, 3\}$ is a proper 3-coloring of G, so $\chi(G) \leq 3$. Since G contains a triangle, we conclude $\chi(G) = 3$.

If $c : V(G) \to \{1, 2, 3\}$ is a proper 3-coloring of G, then for some $i \in \{1, 2, 3\}$, there are at most $\lfloor n/3 \rfloor$ vertices of color i. If $S = \{v \in V(G) : c(v) = i\}$, then $|S| \leq \lfloor n/3 \rfloor$ and every triangle in G contains exactly one vertex of S, since each triangle uses all three colors on its vertices. Now in a triangular region, any two points are mutually visible, so since R is a union of triangular regions, every point in R is visible from S, as required. To find R such that $f(R) = \lfloor n/3 \rfloor$, consider Figure 4.3. $\square$

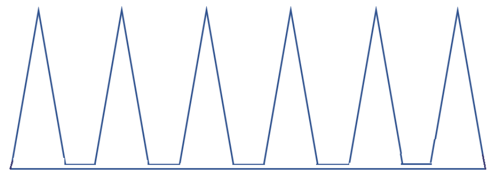


FIGURE 4.3. Art gallery R with $3k+2$ sides and $f(R) = k$.

The related problem of rectilinear art galleries – whose sides are parallel to the x-axis or y-axis – is more challenging. In this case, it is known that $n/4$ guards suffice.

4.5. The Five-Color Theorem

By Theorem 4 and Lemma 2, every planar graph is 6-colorable. In this section, we show that planar graphs are 5-colorable. To prove this, we need the notation of ***contraction*** of an edge $\{a, b\}$ in a graph G: this is the graph $G/\{a, b\}$ obtained from $G - \{a\} - \{b\}$ by adding a vertex x joined to all vertices in $N_G(a) \cup N_G(b)$.

THEOREM 7. (THE FIVE COLOR THEOREM)
Every planar graph is 5-colorable.

PROOF. Proceed by induction on $|V(G)|$. If $|V(G)| \leq 5$, then the theorem is true: assign all vertices different colors. Now suppose $|V(G)| > 5$. If there exists $v \in V(G)$ with $d_G(v) \leq 4$, then $G - \{v\}$ is 5-colorable by induction, and we can assign v a color not used on any neighbors of v. So we assume $\delta(G) \geq 5$. By Lemma 2, some vertex $v \in V(G)$ has $d_G(v) = 5$.

Since G is planar and K_5 is not planar, some pair $\{a, b\}$ of vertices of $N_G(v)$ are not adjacent. Let $F = G/\{a, v\}$ and let w be the vertex adjacent to all vertices in $N_G(a) \cup N_G(v)$. Note that F is planar: if G is drawn in the plane, then w is placed in the face of $G - \{a\} - \{v\}$ that contains the point where v was drawn in G. Let $H = F/\{w, b\}$, and let x be the vertex of H adjacent to all vertices in $N_F(b) \cup N_F(w)$. Note H is planar for the same reason that F – see Figure 4.4.

By induction, H has a proper coloring $c : V(H) \to \{1, 2, 3, 4, 5\}$. We now define a proper coloring $c' : V(G) \to \{1, 2, 3, 4, 5\}$. For each $u \in V(G)\backslash\{a, b, v\}$, let $c'(u) = c(u)$. Let $c'(a) = c'(b) = c(x)$ – we can do that since a and b are not adjacent. Finally, the number of colors used by neighbors of v in G in the coloring c' is at most four, since a and b got the same color. So there is a color i not used by any neighbor of v, and we let $c'(v) = i$. Then c' is a proper coloring of G. □

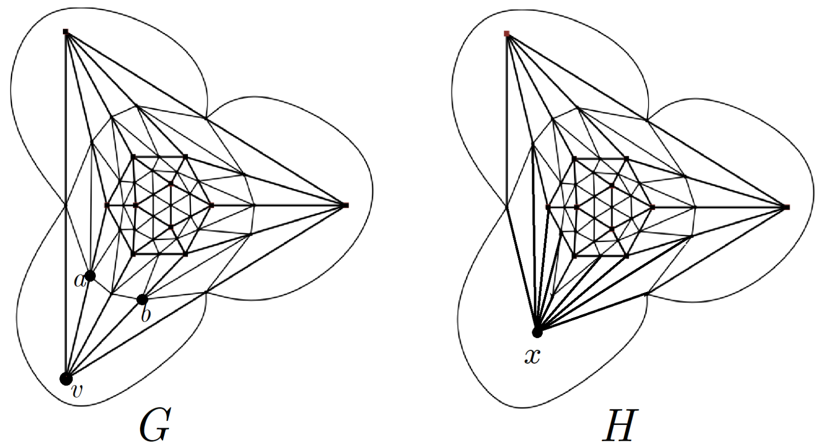


FIGURE 4.4. Proof of the 5-Color Theorem

An alternative proof of the 5-color theorem involves ***Kempe chains***, similar to the proof of Vizing's Theorem.

4.6. The Four-Color Theorem and Duality

Perhaps the most famous theorem in graph theory is the 4-color theorem, proved by Appel and Haken [**2**]: every planar graph is 4-colorable. The shortest proof is by Robertson and Seymour [**21**].

THEOREM 8. (THE FOUR-COLOR THEOREM)
Every planar graph is 4-colorable.

An interesting approach to the Four-Color Theorem is via ***duality***. Let G be a plane multigraph all of whose face boundaries are cycles. The ***combinatorial dual*** of G is the plane multigraph G^* obtained by placing a vertex v_f in the interior of each face $f \in F(G)$ and joining v_f to v_g by an edge for each edge in common to the boundary of faces f and g of G. Examples of duals are shown in Figure 4.5:

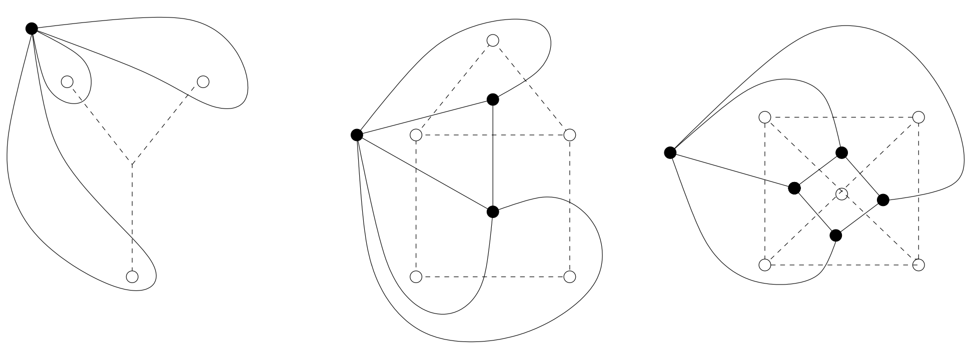


FIGURE 4.5. Duality

The ***map coloring problem*** is the problem of color the faces of a plane graph in such a way that whenever two faces share an edge, they have different colors. If k colors are used, we say the plane graph is ***k-face-colorable***. This is equivalent to finding a proper vertex-coloring of the dual of the plane graph. In particular, an approach to the Four-Color Theorem is to take the dual of any ***maximal plane*** graph with at least four vertices, and to try to color the faces with four colors. Since all the faces of a maximal plane graph are triangles, the dual is a cubic graph with no multiple edges.

THEOREM 9. *Every planar graph is 4-colorable if and only if every cubic planar graph is 3-edge-colorable.*

PROOF. Let G be a planar graph and let G_0 be a plane embedding of G. Then G_0 is contained in a maximal plane graph G_1. If every planar graph is 4-colorable, then G_1 is 4-colorable which means that the map G_1^* is 4-face-colorable and cubic. Since G_1 is 3-connected, no edge of G_1^* is a bridge so every edge of G_1^* is on the boundary of exactly two faces. Now assign edge-color 1 to those edges of G_1^* on the boundary of faces of color 1 and 2, or color 3 and 4, assign edge-color 2 to those edges of G_1^* on the boundary of faces of colors 1 and 3, or colors 2 and 4, and assign edge-color 3 to all remaining edges of G. One checks that this is a proper 3-edge-coloring of G^*, as required.

Define G, G_0, G_1, G_1^* as in the first part of the proof. If every cubic planar graph is 3-edge-colorable, then G_1^* has a proper 3-edge-coloring, with colors 1, 2 and 3. That is to say that $G_1^* = M_1 \cup M_2 \cup M_3$ where M_i is the perfect matching consisting of edges of color i. Then $H_1 = M_1 \cup M_2$ is a plane graph and $H_2 = M_1 \cup M_3$ is a plane graph. Colour the faces of H_1 with colors 1 and 2, and color the faces of H_2 with colors 1' and 2'. To get a 4-face-coloring of G_1^*, and hence a coloring of G, color a face F with color (i, j') if it is contained in a region of color i in H_1 and a region of color j' in H_2. Then the number of colors we used is four, and one checks that this a proper coloring of the faces of G_1^*. $\square$

The dual of a maximal plane graph with at least four vertices is cubic and does not contain vertex x and y such that $G - \{x\} - \{y\}$ is disconnected. Tait [**25**] conjectured that all these graphs are hamiltonian, and this would imply the Four Color Theorem, since then all those cubic graphs have edge-chromatic number three. Unfortunately, they are not all hamiltonian, as a counterexample of Tutte on forty-six vertices showed (Figure 4.6). Tutte's counterexample is shown below.

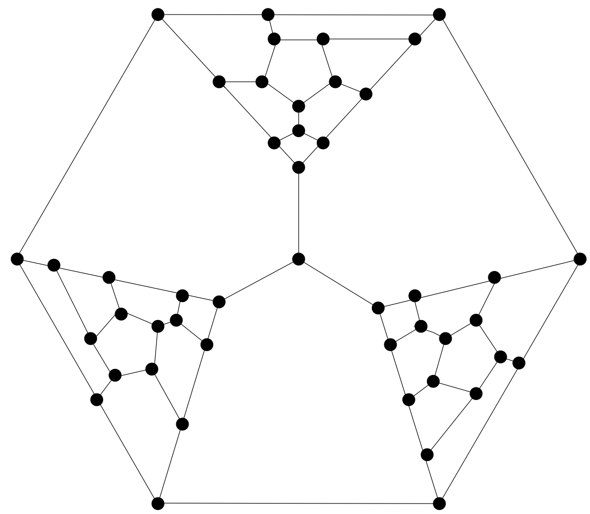


FIGURE 4.6. Tutte's Graph

4.7. Exercises

Question 4.7.1° A factory wishes to store the following chemicals in storage containers: hydrogen, helium, oxygen, chlorine, sulfur, and iron. The chemicals must be stored in separate containers if they are liable to react with one another. Determine the minimum number of containers to store all these chemicals.

Question 4.7.2° Determine $\chi'(G)$ and $\chi(G)$ for each of the graphs shown below.

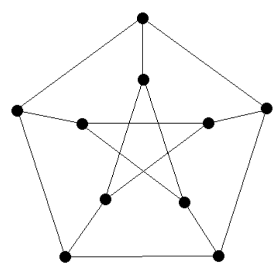


FIGURE 4.7. The Petersen graph

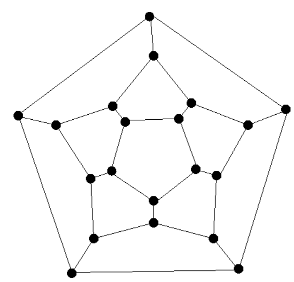


FIGURE 4.8. The dodecahedron graph

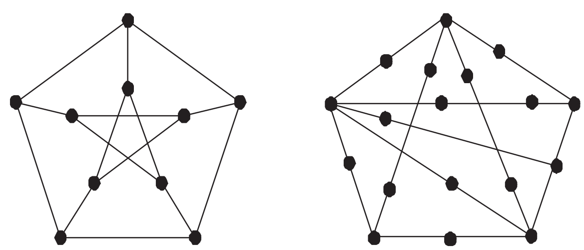
Question 4.7.3° Prove or disprove the existence of a plane multigraph with

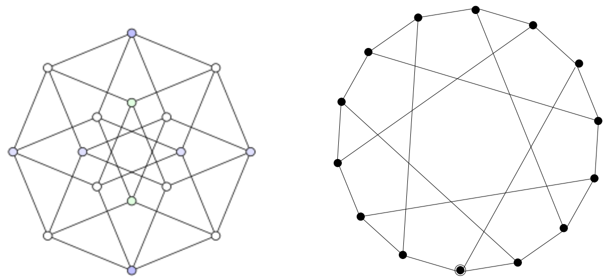
(a) minimum degree at least two whose faces all have different degrees
(b) minimum degree at least three whose faces all have different degrees.

Question 4.7.4° Let G be a union of graphs $G_1, G_2, \ldots, G_k$ such that $V(G_i) \cap V(G_j) = X$ and $G[X]$ is a complete graph. Prove that

$$\chi(G) \leq \max\{\chi(G_i) : 1 \leq i \leq k\}.$$

Question 4.7.5° Determine which of the graphs in the figure below is planar? Justify your answers.





◇

Question 4.7.6. Let G and H be graphs with $V(G) = V(H)$.

(a) Prove that $\chi(G \cup H) \leq \chi(G)\chi(H)$.
(b) Prove that K_{2^n+1} is not a union of n bipartite graphs.

Question 4.7.7. A maximal plane graph is a plane graph $G = (V, E)$ with $n \geq 3$ vertices such that if we join any two non-adjacent vertices in G, we obtain a non-plane graph.

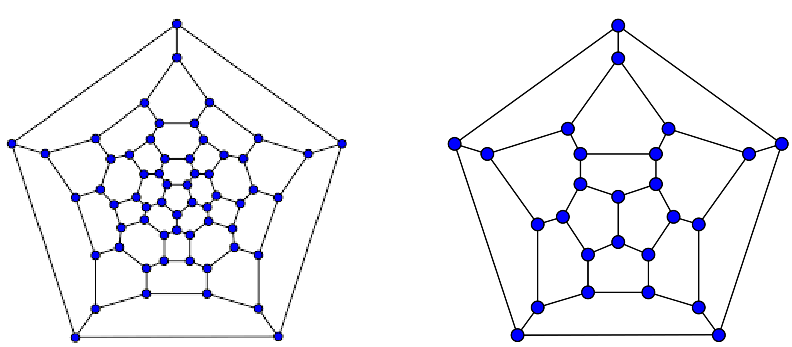
(a) Draw a maximal plane graph on six vertices.
(b) Show that a maximal plane graph on n points has $3n - 6$ edges and $2n - 4$ faces.
(c) A ***triangulation*** of an n-gon is a plane graph whose vertex set is the vertex set of a convex n-gon in the plane, whose infinite face boundary is a convex n-gon and all of whose other faces are triangles. How many edges does a triangulation of an n-gon have?

Question 4.7.8. Show that every triangle-free planar graph is 4-colorable.

Question 4.7.9.

(a) Give an example of a connected cubic planar graph that is not hamiltonian.
(b) Give an example of a cubic planar multigraph with no 3-edge-coloring.
(c)* Prove that the graph in Figure 4.6 is not hamiltonian.

Question 4.7.10. For which r is there a 3-regular plane graph with r faces of degree five and all other faces of degree six? Examples are drawn below.



Question 4.7.11. Suppose a person is standing in a room which has a painting on each of its walls. Prove that if the room has at most five walls, then the person can find a place to stand so as to see all the paintings at once. Prove that if the room has six walls or more, then it is possible that the person cannot find a place to stand so as to see all the paintings at once.

Question 4.7.12. Determine a 3-edge-coloring and hence a 4-face-coloring of the plane graph below.

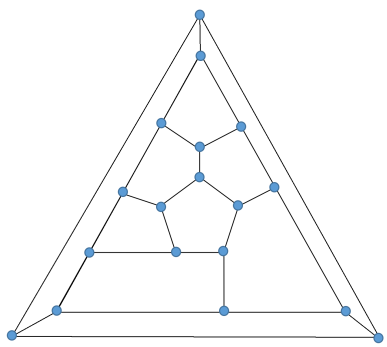


FIGURE 4.9. Edge-coloring and face-coloring

Question 4.7.13. Let G be a maximal plane graph with at least four vertices and $v \in V(G)$. Prove that there exists a plane cycle of length $d(v)$ in $G[N(v)]$. Using Question 4.7.11, prove that any planar graph can be drawn without crossings so that the edges are straight lines.

Question 4.7.14. Let $\omega(G)$ – the ***clique number of*** G – be the maximum number of vertices in a complete subgraph of a graph G.

(a) Prove that for every graph G, $\chi(G) \geq \omega(G)$.
(b) Prove that for every graph G, $\chi(G) \geq |V(G)|/\alpha(G)$.
(c) For each $k \geq 2$, find a graph G such that $\chi(G) = k+1$ and $\omega(G) = k$.

Question 4.7.15. Let $k \geq 2$ and let G be a graph of chromatic number k such that $\chi(G - \{v\}) < k$ for every $v \in V(G)$ (these are called k***-critical graphs***).

(a) If $k = 2, 3$, describe the graph G.
(b) Prove that $\delta(G) \geq k - 1$.
(c) Show that $G - \{x\}$ is connected for every $x \in V(G)$.

Question 4.7.16. Show that the maximum number of edges in an n-vertex graph of chromatic number k is at most $(k-1)n^2/2k$.

◇

Question 4.7.17* Let G be a planar graph with degree sequence $(d_1, d_2, \ldots, d_n)$ where $n \geq 3$. Prove that

$$\sum_{i=1}^{n}(6 - d_i) \geq 12.$$

Prove that a planar graph of minimum degree five has at least twelve vertices of degree five.

Question 4.7.18* Let G be a planar graph with degree sequence $(d_1, d_2, \ldots, d_n)$ where $n \geq 3$. Prove that for $3 \leq m \leq n$.

$$\sum_{i=1}^{m} d_i \leq 2n + 6m - 16.$$

Question 4.7.19* Prove that $n/4$ guards suffice to guard a rectilinear art gallery with n sides.

Question 4.7.20* The ***crossing number*** $\nu(G)$ of a graph G is the minimum number of pairs of crossing edges in a drawing of the graph in the plane. Prove that $\nu(K_5) = 1$ and prove that $\nu(G) = 2$ when G is the Petersen graph (see Figure 4.10).

Question 4.7.21* Prove that any triangle-free planar graph G has a partition (A, B) of $V(G)$ such that $G[A]$ and $G[B]$ are acyclic.

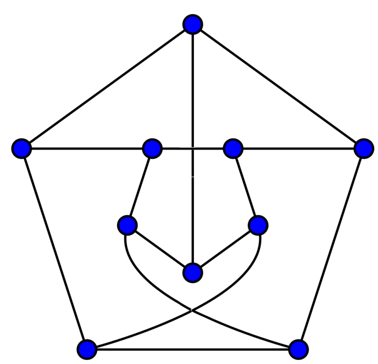


FIGURE 4.10. Crossings and the Petersen graph

Question 4.7.22* Let G be a graph. An ***orientation*** of G is a digraph $\vec{G}$ obtained by replacing each edge $\{a, b\} \in E(G)$ with either the arc (a, b) or the arc (b, a). Prove that if $\chi(G) \geq k$, then every orientation $\vec{G}$ of G contains a directed path of length at least k.

CHAPTER 5

Extremal Graph Theory

Let G be a graph and let F be a graph with at least one edge. We say that G is ***F-free*** if G does not contain a subgraph isomorphic to F. The basic problem of extremal graph theory is to determine the maximum number of edges in an n-vertex F-free graph. These maxima are collectively referred to as the ***extremal numbers*** or ***Turán numbers*** for F, and denoted $\mathrm{ex}(n, F)$. An F-free n-vertex graph with $\mathrm{ex}(n, F)$ edges is called an ***extremal graph*** for F. Ideally, for each graph F, we would like to determine not only $\mathrm{ex}(n, F)$ as well as the extremal graphs.

5.1. Cores and Maximum Cuts

A ***maximum cut*** of a graph is a bipartite subgraph with as many edges as possible. The following lemma attributed to Erdős states that every graph has a maximum cut with at least half the edges of the graph:

LEMMA 1. *Let G be a graph with m edges. Then G has a bipartite subgraph with at least $m/2$ edges.*

PROOF. Flip a fair coin at each vertex, and let A be the set of vertices where the coin turned up heads, and let B be the set of vertices where the coin turned up tails. Each edge of G has probability $1/2$ of having one end in A and one end in B. Therefore the average number of edges with one end in A and one end in B is $m/2$. It follows there must exist an A and a B such that at least $m/2$ edges of G have one end in A and one end in B. □

In the context of extremal graph theory, problems sometimes become easier in the context of bipartite graphs. For instance, consider the problem of showing that an n-vertex graph G with at least $2n$ edges contains an even cycle. By Lemma 1, there is a bipartite subgraph of G with at least n edges, which must contain a cycle. By Lemma 7, this cycle has even length.

Let G be a graph, and let H be obtained by repeatedly removing a vertex of degree less than k in the current graph. The (possibly empty) graph obtained after this procedure is called the ***k-core*** of the graph, and determines the largest subgraph of minimum degree at least k. A sufficient condition for the k-core to be non-empty is given by the following lemma:

LEMMA 2. *Let $k \geq 1$ and let G be an n-vertex graph with more than $(k-1)n - \binom{k}{2}$ edges. Then G has a subgraph of minimum degree at least k.*

PROOF. While there remains a vertex of degree at most $k-1$, we remove such a vertex. After $n-k$ steps, the number of edges is

$$|E(G)| - (k-1)(n-k) > (k-1)n - \binom{k}{2} - (k-1)(n-k) = \binom{k}{2}.$$

However there cannot be more than $\binom{k}{2}$ edges on the remaining k vertices, so the process must have terminated with a subgraph of minimum degree at least k. □

Certain extremal problems become simpler for graph of large minimum degree, in which case the k-core is relevant. For example, let $k \geq 1$ and let G be a graph of minimum degree at least k. If P is a longest path in G, with ***end vertices*** u and v, then $N(u) \cup N(v) \subseteq V(P)$, which implies P has length at least k. In fact, this argument works for trees too:

LEMMA 3. *Let $k \geq 1$ and let G be a graph of minimum degree at least k and T any tree with $k+1$ edges. If G does not contain T, then every component of G is a clique of order $k+1$.*

Using Lemma 2 together with Lemma 3, we conclude $\mathrm{ex}(n,T) \leq kn - \binom{k+1}{2}$ for every tree T with $k+1$ edges. Furthermore, if we consider an edge $\{w,v\}$ such that w is as close to u on the longest path P as possible, then the segment of P from w to v plus the edge $\{w,v\}$ forms a cycle C of length at least $k+1$. We remark that if G is bipartite, this cycle has length at least $2k$. Since $N(v) \subseteq V(P)$, all neighbors of v are vertices of $V(C)$. This discussion gives the following lemma:

LEMMA 4. *If G has minimum degree at least $k \geq 2$, then G contains a cycle C of length at least $k+1$, and if $k \geq 3$, then there is an edge e of G joining two non-adjacent vertices of C. If G is bipartite, then G contains a cycle C of length at at least $2k$, and if $k \geq 3$, then there is an edge e of G joining two non-adjacent vertices of C.*

The famous ***Erdős-Sós Conjecture*** [**11**] is as follows:

CONJECTURE 1. (ERDŐS-SÓS CONJECTURE)
If T is any tree with $k+1$ edges, then $\mathrm{ex}(n,T) \leq \frac{1}{2}kn$.

A proof of this conjecture was claimed by Ajtai, Komlós, Simonovits and Szemerédi [**1**]. The special case where the tree is a path is the ***Erdős-Gallai Theorem***, which we study next.

5.2. The Erdős-Gallai Theorem

Erdős and Gallai [**12**] proved the following theorem, which verifies the Erdős-Sós Conjecture for paths:

THEOREM 1. (ERDŐS-GALLAI THEOREM)
Let $k \geq 1$ and let G be an n-vertex P_k-free graph. Then $e(G) \leq (k-1)n/2$, with equality if and only if $k|n$ and every component of G is K_k.

PROOF. Let G be an n-vertex graph with at least $(k-1)n/2$ edges. We prove by induction on $n+k$ that G contains P_k unless every component of G is K_k. If G is disconnected, then some component of G contains a path of length k or equals K_k, by induction. If some component is K_k, we remove it and get a graph with $n-k$ vertices and $(k-1)(n-k)/2$ edges. By induction, that graph is a union of K_k or contains a path of length k, and we are done. Therefore we may assume G is connected. The theorem is clear for $k=1$, since a single edge forms a path P_1, and if there are no edges then every component of G is K_1. Suppose $k \geq 2$. If G contains a vertex v of degree less than $k/2$, then $G-\{v\}$ has at least $(k-1)(n-1)/2$ edges. By induction, $G-\{v\}$ contains a path of length k, unless $k|n-1$ and $G-\{v\}$ is a union of cliques K_k. Since G is connected, v has a neighbor in one of these cliques, and this gives a path of length k ending with v. So we may assume every vertex of G has degree at least $k/2$. By induction, G contains a path P of length $k-1$. The ends u and w of that P have all their neighbors on the path. As in Dirac's Theorem, there exists a neighbor of u that comes after a neighbor of w on the path from u to w, and this gives a cycle C of length k containing P. If there is a vertex x not in C, then since G is connected, there is a path from x to C, and in particular by adding an edge $\{w,x\}$ with $w \in V(C)$, we get a path of length k ending with x. So $V(C)=V(G)$, and so $|V(G)|=k$ and $G=K_k$, as required. $\square$

5.3. Mantel's and Turán's Theorems

The first theorem of extremal graph theory is ***Mantel's Theorem*** [**18**], which shows that the extremal triangle-free graphs are balanced complete bipartite graphs. In particular, for all $n \geq 1$, one deduces the extremal numbers for triangles:

$$\mathrm{ex}(n,K_3)=\Big\lfloor\frac{n^2}{4}\Big\rfloor. \tag{5.1}$$

Turán's Theorem [**27**] extends Mantel's Theorem to all complete graphs, by showing that for $r \geq 2$, the extremal K_{r+1}-free graphs are complete r-partite graphs. More precisely, let $(V_1,V_2,\dots,V_r)$ be a partition of an n-element set such that $|V_1| \leq |V_2| \leq \dots \leq |V_r| \leq |V_1|+1$. Then the ***Turán graph*** $T_r(n)$ is the complete r-partite graph with parts $V_1,V_2,\dots,V_r$. We observe $T_2(n)$ is the extremal triangle-free graph – a complete bipartite graph. The Turán graphs $T_r(n)$ also have the highest minimum degree amongst all graphs with the same number of vertices and edges, and any n-vertex r-partite graph with $e(T_r(n))$ edges must be isomorphic to $T_r(n)$.

THEOREM 2. (TURÁN'S THEOREM)
Let G be a K_{r+1}-free n-vertex graph. Then $e(G) \leq e(T_r(n))$ with equality if and only if $G=T_r(n)$. In particular,

$$\mathrm{ex}(n,K_{r+1})=e(T_r(n)). \tag{5.2}$$

PROOF. We prove by induction on n that if G is a K_{r+1}-free n-vertex graph with at least $e(T_r(n))$ edges, then $G=T_r(n)$. The cases $n \leq r$ follow from the fact that $T_r(n)=K_n$ for $n \leq r$. For $n \geq r+1$, let G be an n-vertex K_{r+1}-free graph with

$e(G) \geq e(T_r(n))$. Let H be a subgraph of G with $e(T_r(n))$ edges. As observed above, $\delta(H) \leq \delta(T_r(n))$. If v is a vertex of minimum degree in H, then

$$e(H - \{v\}) \geq e(T_r(n)) - \delta(T_r(n)) = e(T_r(n-1)). \tag{5.3}$$

By induction, $H - \{v\} = T_r(n-1)$. Let $W_1, W_2, \dots, W_r$ be the parts of $H - \{v\}$. If $N_H(v) \cap W_i \neq \emptyset$ for $1 \leq i \leq r$, then select $v_i \in N(v) \cap W_i$ for $1 \leq i \leq r$, and $\{v, v_1, v_2, \dots, v_r\}$ induces a K_{r+1} in H, a contradiction. Therefore $N(v) \cap W_i = \emptyset$ for some $i : 1 \leq i \leq r$. Let $V_j = W_j$ if $j \neq i$, and $V_i = W_i \cup \{v\}$. Then H is r-partite with parts $V_1, V_2, \dots, V_r$. Since H has $e(T_r(n))$ edges, we conclude $H = T_r(n)$. Since $T_r(n)$ is a maximal K_{r+1}-free graph, $G = H = T_r(n)$. $\square$

5.4. The Kövari-Sós-Turán Theorem

The notorious problem of determining the order of magnitude of Turán numbers for bipartite graphs is called the ***bipartite Turán problem***, and such problems are collectively referred to as ***degenerate extremal problems*** – see Füredi and Simonovits [**15**] for a comprehensive survey. In this section, we prove the following theorem, known as the ***Kövari-Sós-Turán Theorem*** [**17**].

Theorem 3. (Kövari-Sós-Turán Theorem)
Let $s, t \geq 1$ and $n \geq 1$. Then

$$\mathrm{ex}(n, K_{s,t}) \leq \tfrac{1}{2}((t-1)^{\frac{1}{s}} n^{2-\frac{1}{s}} + (s-1)n).$$

Proof. Let G be an n-vertex $K_{s,t}$-free graph. Suppose G has average degree $d > s-1$. The key observation is that since no set of s vertices in G has more than $t-1$ common neighbors,

$$\sum_{v \in V(G)} \binom{d(v)}{s} \leq (t-1)\binom{n}{s} \leq \frac{(t-1)n^s}{s!}. \tag{5.4}$$

We now use the simple fact that for $s \geq 1$, the function $f_s : \mathbb{R} \to \mathbb{R}$ defined by $f_s(x) = \binom{x}{s}$ for $x > s-1$ and $f_s(x) = 0$ for $x \leq s-1$ is convex. Recalling $d > s-1$, ***Jensen's inequality*** gives:

$$\sum_{v \in V(G)} \binom{d(v)}{s} \geq n\binom{d}{s} \geq \frac{n(d-s+1)^s}{s!}. \tag{5.5}$$

Together with (5.4), we obtain $(d-s+1)^s \leq (t-1)n^{s-1}$, and this gives the bound in the theorem. $\square$

The salient conjecture in the area is due to Zarankiewicz [**31**]:

Conjecture 2. (Zarankiewicz Conjecture)
For all $t \geq s \geq 2$, $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$.

5.5. Quadrilaterals and Sidon Sets

Theorem 3 shows that a graph G containing no cycle of length four satisfies

$$\sum_{v \in V(G)} \binom{d(v)}{2} \leq \binom{n}{2}.$$

By Jensen's Inequality, if G has average degree d, this implies

$$d(d-1) \leq n-1 \tag{5.6}$$

A challenging question is to determine extremal graphs without cycles of length four, in particular, to constuct a quadrilateral-free n-vertex graph with average degree d and $d(d-1) = n-1$. We make a diversion into combinatorial number theory before answering this question.

A ***Sidon set*** in an abelian group Γ is a set $A \subseteq \Gamma$ such that if $a - b = a' - b'$ with $a, b, a', b' \in A$, then $a = b$ or $a = a'$. In other words, every non-zero element of Γ can be represented in at most one way as a difference of elements in A. The problem of determining the maximum size of a Sidon set in an abelian group Γ of order n is notoriously difficult, and bears considerable similarity with the extremal problem for quadrilaterals.

LEMMA 5. *Let Γ be an abelian group with n elements and $A \subset \Gamma$ a Sidon set. Then*

$$|A|(|A|-1) \leq |\Gamma| - 1. \tag{5.7}$$

PROOF. For each $(a, b) \in A \times A$ with $a \neq b$, there is exactly one non-zero element of Γ of the form $a - b$. There are $|A|(|A|-1)$ pairs (a, b), and $|\Gamma| - 1$ non-zero elements in Γ, and (5.7) follows. $\square$

The above lemma implies Sidon sets cannot have size more than $\sqrt{n} + 1$ elements in an abelian group with n elements. There are a number of interesting examples of Sidon sets with close to this size – see O'Bryant for a survey [**19**]. We give a standard example of a Sidon set with q elements in $\mathbb{Z}_q \times \mathbb{Z}_q$ when q is an odd prime.

EXAMPLE 5.1. Let q be an odd prime, and let $\mathbb{Z}_q$ denote the ***cyclic group*** of integers modulo q. Let $\Gamma_q = \mathbb{Z}_q \times \mathbb{Z}_q$ with addition as the group operation. Let $n = |\Gamma_q| = q^2$. Then let

$$A = \{(x, x^2) \in \Gamma_q : x \in \mathbb{Z}_q\}.$$

Then $|A| = q = \sqrt{n}$, and A is a Sidon set: if in $\mathbb{Z}_q$ we have

$$x - y = z - w \quad \text{ and } \quad x^2 - y^2 = z^2 - w^2$$

then when $x \neq y$, we may divide the second equation by the first to get $x + y = z + w$. Adding the equations, we get $2x = 2z$, which implies $x = z$ in $\mathbb{Z}_q$ with q odd. Thus A is a Sidon set.

To connect Sidon sets to graph theory, we create from any set A in an abelian group Γ the ***Cayley sum graph*** $G(A)$: the vertex set of $G(A)$ is Γ and $\{x, y\} \in E(G(A))$ whenever $x + y \in A$. Technically speaking, $G(A)$ is a pseudograph: it may have loops, corresponding to pairs $\{x, x\}$ with $x + x \in A$. Then $G(A)$ has $|\Gamma|$ vertices and every vertex of $G(A)$ has degree exactly $|A|$. If $G(A)$ contains a quadrilateral (x, y, z, w, x), then there exist elements $a, b, c, d \in A$ such that

$$x + y = a \quad y + z = b \quad z + w = c \quad w + x = d.$$

This implies

$$a - b = d - c$$

and if A is a Sidon set, we must have $a = b$ or $a = d$. However, $a = b$ implies $x = z$ and $a = d$ implies $y = w$, contradicting that (x, y, z, w, x) is a quadrilateral. We conclude that if A is a Sidon set, then the Cayley sum graph $G(A)$ has no quadrilaterals.

Using the above example, $G(A)$ has q^2 vertices and $q^3/2$ edges and no quadrilaterals. However, $G(A)$ has loops corresponding to solutions to $2(x, y) = (a, a^2)$, so the number of loops is exactly q (one for each choice of $a \in \mathbb{Z}_q$). We conclude after removing loops that for each odd prime q, there is a graph with $q^3/2 - q$ edges on q^2 vertices with no quadrilaterals. Singer **[23]** showed that for infinitely many q, $\mathbb{Z}_{q^2+q+1}$ contains a Sidon set with $q + 1$ elements. The corresponding Cayley sum graph has $q^2 + q + 1$ vertices and exactly q^2 loops, so after removing loops we get a quadrilateral-free graph G_q with

$$|E(G_q)| = \frac{1}{2}(q + 1)(q^2 + q + 1) - q^2 = \frac{1}{2}q(q + 1)^2. \tag{5.8}$$

Such graphs were first described by Erdős, Rényi and Sós. Füredi **[15]** showed that these are in fact extremal graphs:

THEOREM 4. *Let $n = q^2 + q + 1$ where $q > 13$ is an odd prime power. Then*

$$\mathrm{ex}(q^2 + q + 1, C_4) = \frac{1}{2}(q + 1)(q^2 + q + 1). \tag{5.9}$$

These constructions may also be described geometrically, via finite projective planes. An example when $q = 3$ is shown below:

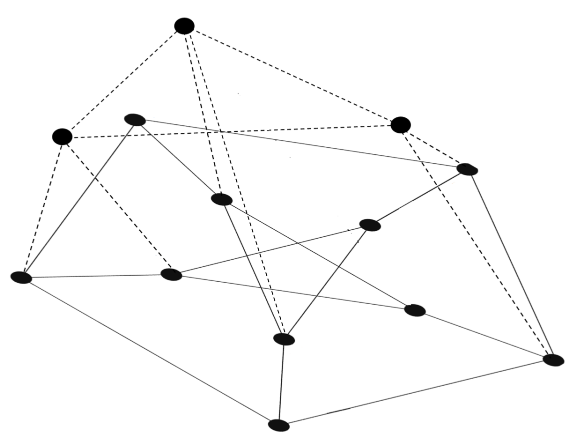


FIGURE 5.1. Extremal graph

5.6. The Moore Bound

The ***girth*** of a graph containing a cycle is the length of a shortest cycle in the graph. The ***distance*** between vertices u and v in a connected graph G, denoted $d_G(u,v)$, is the length of a shortest uv-path. The maximum distance between any two vertices in a connected graph is called the ***diameter*** of G. The minimum r such that every vertex of G is at distance at most r from some vertex of G is called the ***radius*** of G. If G is a d-regular graph of ***girth*** at least $2g+1$, then $|V(G)| \geq 1 + d\sum_{r=0}^{g-1}(d-1)^r$ by counting vertices at distance at most g from v. A similar bound holds for graphs of girth $2g+2$:

THEOREM 5. (MOORE BOUND)
Let G be an n-vertex d-regular graph of girth at least g, and let $k = \lfloor g/2 \rfloor$. Then

$$n \geq \begin{cases} 1 + d \cdot \sum_{i=0}^{k-1}(d-1)^i & \text{if } g \text{ is odd} \\ 2 \cdot \sum_{i=0}^{k-1}(d-1)^i & \text{if } g \text{ is even} \end{cases}$$

Equality holds only if G is a d-regular graph of diameter $\lfloor \frac{g}{2} \rfloor + 1$, and G is bipartite if g is even.

PROOF. A ***non-backtracking walk of length k*** in a graph is a sequence of vertices $(v_1, v_2, \dots, v_{k+1})$ such that $\{v_i, v_{i+1}\}$ is an edge of the graph for all $i \leq k$ and $\{v_i, v_{i+1}\} \neq \{v_{i+1}, v_{i+2}\}$ for all $i < k$. Suppose $g = 2k+1$. There are at least $d(d-1)^{k-1}$ non-backtracking walks of length i in G starting at a vertex v, and all of the walks of length at least 1 are paths with distinct ends in $V(G)\backslash\{v\}$, since the girth is at least g. Therefore

$$\sum_{i=0}^{k} d(d-1)^i \leq n-1, \tag{5.10}$$

and this gives the required bound. Now suppose $g = 2k$. For any edge $\{u,v\} \in E(G)$, there exist $2\sum_{i=1}^{k-1}(d-1)^i$ non-backtracking walks of length at most k starting with u or v in $G - \{u,v\}$. All of these walks have distinct ends and therefore

$$2\sum_{i=1}^{k-1}(d-1)^i \leq n-2. \tag{5.11}$$

This completes the proof of the lower bounds on n. We leave the case of characterization of equality as an exercise. $\square$

The graphs which achieve equality in the Moore Bound are called ***Moore graphs***. Moore graphs of girth $g \in \{3,4\}$ are clearly the complete graphs and complete bipartite graphs. Also, all cycles are Moore graphs. The interesting cases are the existence of d-regular Moore graphs of girth g when $d \geq 3$ and $g \geq 5$. The ***Petersen graph*** is a cubic Moore graph with girth $g = 5$, and a 7-regular Moore graph with girth $g = 5$ is the ***Hoffman-Singleton graph***, shown below:

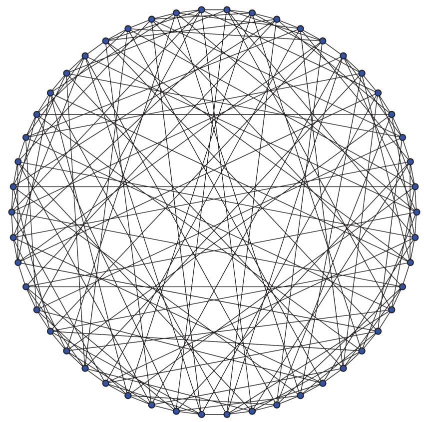


FIGURE 5.2. Hoffman-Singleton graph

Using linear algebra, it is possible to show that the only d-regular Moore graphs of girth five are when $d \in \{3, 7, 57\}$, but no construction of a 57-regular Moore graph of girth five is known.

5.7. Even Cycle Theorem

We saw that $\mathrm{ex}(q^2+q+1, C_4) = \frac{1}{2}q(q+1)^2$ when q is a large enough power of an odd prime, in Theorem 4. Furthermore, it can be shown from (5.6) that for all $n \geq 4$,

$$\mathrm{ex}(n, C_4) \leq \frac{n}{4}(1 + \sqrt{4n-3}).$$

In this section, we consider the extremal function for longer even cycles. The ***Even Cycle Theorem*** of Bondy and Simonovits [**5**] is as follows:

THEOREM 6. (EVEN CYCLE THEOREM)
Let $n, k \geq 2$. Then there exists a constant c_k such that

$$\mathrm{ex}(n, C_{2k}) \leq c_k n^{1+\frac{1}{k}}. \tag{5.12}$$

The proof involves a number of different ingredients, including breadth-first search trees (see Section 1.5), long cycles in graphs of large minimum degree (see Lemma 2 and Lemma 4), and the following ***coloring lemma*** [**30**]:

LEMMA 5.2. *Let F be a graph comprising a cycle C plus an edge, and suppose that (A, B) is a non-trivial partition of $V(F)$. Then for every $\ell < |V(F)|$ there exists a path $P \subset F$ of length ℓ such that one end of P is in A and the other end of P is in B, unless F is bipartite with parts A and B.*

The proof is left as an exercise. The key part of the proof is the next lemma, using breadth-first search trees. For a given breadth-first search tree T rooted at a vertex v in a graph G, let H_i denote the bipartite graph consisting of all edges of G between $N_i(v)$ and $N_{i+1}(v)$.

LEMMA 5.3. *Let $k \geq 2$, and let T be a breadth first search tree in a graph G rooted at a vertex v, and suppose $e(H_i) \geq k|V(H_i)|$ for some i. Then for some $r \leq i$, G contains cycles $C_{2r+2}, C_{2r+4}, \ldots, C_{2r+2k}$.*

PROOF. Let $H = H_i$ and $L = N_i(v)$. Since $e(H) \geq k|V(H)|$, by Lemma 2, H has a subgraph of minimum degree at least $k+1$. By Lemma 4, H contains a cycle C of length at least $2k+2$, and since $k \geq 2$, there exists an edge e of H between two non-adjacent vertices of C. Let $F = C + e$, and note that one of the parts of F is $V(F) \cap L$. Let U be a minimal subtree of T whose set of leaves is $V(C) \cap L$, so that U branches at its root u. Let $r \leq i$ be the height of U. Let A be the set of vertices of L in one branch, and $B = V(F)\backslash L$ – see Figure 5.3, where C has length six and $u = v$. Since $A \neq V(F) \cap L$, A and B are not the parts of F, so by Lemma 5.2, there exists a path P of length 2ℓ with one end in A and one end in B for $1 \leq \ell \leq k$. Then in U there exists a path Q of length $2r$ with the same ends as P, so $P_\ell \cup Q$ is a cycle of length $2r + 2\ell$. This works for $1 \leq \ell \leq k$, so the proof is complete. $\square$

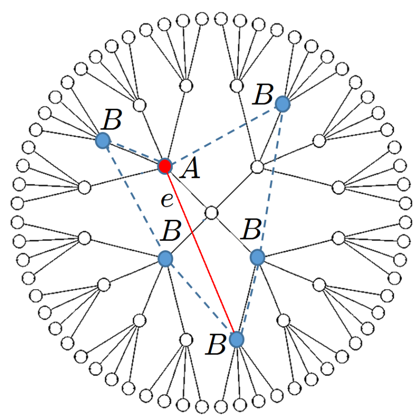


FIGURE 5.3. The sets A and B and edge e

PROOF OF EVEN CYCLE THEOREM. We show specifically that

$$\mathrm{ex}(n, C_{2k}) \leq (k-1)n^{1+\frac{1}{k}} + 4(k-1).$$

Let G be an n-vertex C_{2k}-free graph with at least $(k-1)n^{1+\frac{1}{k}} + 4(k-1)n$ edges. Then by Lemma 1, G has a bipartite subgraph J with at least $\frac{1}{2}e(G)$ edges. By Lemma 2, J has a connected subgraph H of minimum degree at least

$$d > (k-1)n^{\frac{1}{k}} + 4(k-1).$$

Let T be a breadth-first search tree of H rooted at a vertex $v \in V(H)$ with $L_i = N_i(v)$. We claim that for all $i < k$,

$$e(L_i) \leq k(|L_i| + |L_{i+1}|).$$

Otherwise, by Lemma 5.3, H contains cycles $C_{2r+2}, C_{2r+4}, \ldots, C_{2r+2k}$ for some $r \leq i < k$. However one of the cycles then has length $2k$, a contradiction. This proves the claim. We now claim that $|L_i| \geq n^{\frac{1}{k}}|L_{i-1}|$ for all $i \in \{1, 2, \dots, k\}$. The claim is true for $i = 1$ since H has minimum degree at least d. Having proved the claim for all $j < i$, we have

$$e(L_{i-1}, L_i) \geq d|L_{i-1}| - e(L_{i-2}, L_{i-1}) \geq d|L_{i-1}| - 4(k-1)|L_{i-1}|.$$

Since $e(L_{i-1}, L_i) \leq k(|L_i| + |L_{i-1}|)$, the definition of d shows $|L_i| > n^{\frac{1}{k}}|L_{i-1}|$, as required. Since this is valid for all $i \leq k$, we find $|L_k| > n^{1/k}|L_{k-1}| > \cdots > n^{k/k}|L_0|$ which gives $|L_k| > n$, a contradiction. □

The even cycle theorem is more difficult that the Moore Bound, due to the existence of short cycles in a C_{2k}-free graph. For example, the following extremal C_6-free graph in fact has many cycles of length four.

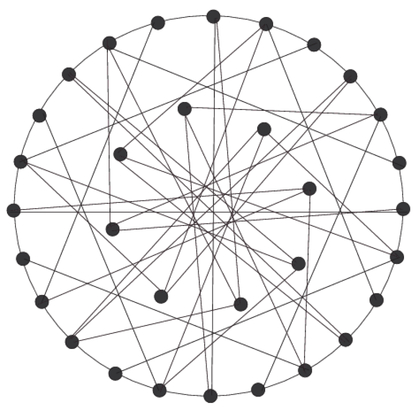


FIGURE 5.4. An extremal hexagon-free graph

The major conjecture in the area is due to Erdős and Simonovits [**13**]:

CONJECTURE 3. *For all $k \geq 2$, there exists a constant $c_k > 0$ such that for all $n \geq 2k$,*

$$\mathrm{ex}(n, C_{2k}) \geq c_k n^{1+\frac{1}{k}}. \tag{5.13}$$

5.8. Random graphs

The problem of constructing extremal F-free graphs is in general very difficult. When F has chromatic number $r + 1 \geq 3$, the Turán graph $T_r(n)$ is F-free, and therefore $\mathrm{ex}(n, F) \geq \mathrm{ex}(n, K_r)$ by Turán's Theorem. The Kövari-Sós-Turán Theorem addresses the case $F = K_{r,s}$, and shows that for any bipartite graph F, there exists a constant $c > 0$ such that $\mathrm{ex}(n, F) \leq n^{2-c}$, since F is certain contained in $K_{s,s}$ with $s = |V(F)|$. In this section, we use ***random graphs*** to give a lower bound for $\mathrm{ex}(n, F)$.

Let $n \geq 1$ and $p \in [0, 1]$. We consider the sample space Ω of all n-vertex graphs, say with vertex set $\{1, 2, \dots, n\}$, where a particular n-vertex graph G has probability measure

$$\mathbb{P}(G) = p^{e(G)}(1-p)^{\binom{n}{2}-e(G)}. \tag{5.14}$$

This probability space is realized by considering the $\binom{n}{2}$ edges of K_n independently, where each edge has probability p. If $p = 1$ we obtain K_n with probability 1. The graph obtained is denoted $G_{n,p}$ and is called the ***Erdős-Rényi Random Graph***.

EXAMPLE 5.4. We determine the probability that $(1, 2, \ldots, n)$ is a path in $G_{n,p}$. Each edge on the path has probability p, so the probability is p^{n-1}. This is not the same as the probability that $G_{n,p} = P_{n-1}$, which by (5.14) is

$$p^{n-1}(1-p)^{\binom{n-1}{2}}.$$

Finally, we consider the expected number of hamiltonian paths in $G_{n,p}$. There are $\frac{1}{2}n!$ hamiltonian paths in K_n, and each has probability p^{n-1} of being a path in $G_{n,p}$, as above. So the expected number of paths is

$$\sum_{P \subset K_n} \mathbb{P}(P \subset G_{n,p}) = \frac{1}{2}n! \cdot p^{n-1}.$$

We recall in general the ***expected value*** or ***mean*** of a discrete random variable X is

$$\mathbb{E}(X) = \sum_x x \cdot \mathbb{P}(X = x). \tag{5.15}$$

Since expectation is defined by a sum, it is a linear operator: for any $a, b \in \mathbb{R}$ and any random variables X and Y,

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

This property is referred to as ***linearity of expectation***. For example, if Y is the number of edges in $G_{n,p}$, then

$$\mathbb{E}(Y) = \sum_{e \in E(K_n)} \mathbb{P}(e \in G_{n,p}) = \sum_{e \in E(K_n)} p = p\binom{n}{2}.$$

One of the most fundamental inequalities in probability is ***Markov's Inequality***:

LEMMA 6. (MARKOV'S INEQUALITY)
Let X be a non-negative real-valued random variable and $\lambda > 0$. Then

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}(X)}{\lambda}. \tag{5.16}$$

When $p = 1/n$ and Y is the number of edges in $G_{n,p}$, $\mathbb{E}(Y) = p\binom{n}{2} = \frac{n-1}{2}$. Markov's Inequality shows that the probability that $G_{n,p}$ has at least $n-1$ edges is at most $\frac{1}{2}$, by taking $\lambda = n - 1$.

EXAMPLE 5.5. We examine the diameter of $G_{n,p}$. For two vertices u and v in $G_{n,p}$, the probability that there is no path of length one or two with ends u and v is exactly $(1-p)(1-p^2)^{2n-4}$, since there is one potential edge between u and v, and $n-2$ potential paths of length two, all of which are edge-disjoint and therefore occur independently in $G_{n,p}$. In particular, we conclude that the probability that u and v

are at distance at least three in $G_{n,p}$ is $(1-p)(1-p^2)^{2n-4}$. The expected number of pairs of vertices at distance at least three is exactly

$$\binom{n}{2}(1-p)(1-p^2)^{2n-4}.$$

For $p \in [0,1]$, we recall the inequality $1-p \leq e^{-p}$. If

$$p = \frac{2\log n}{\sqrt{2n-4}} \leq 1$$

then the number of pairs at distance at least three is at most $\binom{n}{2}e^{-2\log n} < \frac{1}{2}$. We conclude that with probability more than $\frac{1}{2}$, $G_{n,p}$ has diameter two.

5.9. Exercises

Question 5.9.1° Prove that an n-vertex graph G with at least $3n/2$ edges contains a cycle of length at least four.

Question 5.9.2° Let $k \geq 1$. Prove that an n-vertex bipartite graph containing no matching of size k has at most $(k-1)(n-k+1)$ edges for $n \geq 2k$.

Question 5.9.3° Determine for all $n \geq 1$ the value of $\mathrm{ex}(n, P_3)$.

Question 5.9.4° Determine for all $n \geq 1$ the value of $\mathrm{ex}(n, M)$ where M is a matching with two edges.

◇

Question 5.9.5. Let $n \geq k \geq 1$. Determine all n-vertex k-degenerate graphs with exactly $kn - \binom{k+1}{2}$ edges.

Question 5.9.6. Let $k \geq 2$ and let G be a graph with more than $\binom{k}{2}$ edges. Prove that G has a subgraph H with $|V(H)| = k$ and $\delta(H) \geq 1$.

Question 5.9.7. Let q be an odd prime number, and let G be the pseudograph with vertex set $\mathbb{Z}_q \times \mathbb{Z}_q$ such that (x_1, x_2) is adjacent to (y_1, y_2) whenever

$$x_2 + y_2 = x_1 y_1.$$

Prove that G does not contain any quadrilaterals. Suppose the equation $x^2 = 2$ has no solution. Determine the number of edges of G after loops are removed.

Question 5.9.8. Let G be a graph. Prove that there exists a partition (A, B) of $V(G)$ such that $e(A, B) \geq \frac{1}{2}e(G)$ and $|A| \leq |B| \leq |A| + 1$.

Question 5.9.9.

(a) Prove that for every $p \in [0, 1]$, there exists an n-vertex graph of girth at least five with at least m edges, where

$$m = p\binom{n}{2} - p^3\binom{n}{3} - 3p^4\binom{n}{4}.$$

(b) Prove that for large enough n, there exists an n-vertex graph of girth at least five with at least $\frac{1}{4}n^{4/3}$ edges.

Question 5.9.10. A ***bowtie*** is a graph B consisting of two triangles sharing exactly one vertex. Determine $\mathrm{ex}(n, B)$ for all $n \geq 1$.

Question 5.9.11. Let r be a positive integer, and let $f : \mathbb{R} \to [0, \infty)$ be defined by $f(x) = 0$ if $x < r - 1$ and $f(x) = x(x-1)\dots(x-r+1)$ if $x \geq r - 1$. Prove that f is convex on $\mathbb{R}$ and then show if $a = \frac{1}{n}(a_1 + a_2 + \cdots + a_n) \geq r - 1$ for positive integers $a_1, a_2, \ldots, a_n$, then

$$\frac{1}{n}\sum_{i=1}^{n}\binom{a_i}{r} \geq \binom{a}{r}.$$

Question 5.9.12. Let G be a bipartite graph with parts of sizes m and n, not containing a 4-cycle. Prove that

$$|E(G)| \leq m\sqrt{n} + m + n.$$

Question 5.9.13. Let G be an n-vertex graph not containing a 4-cycle. Prove that

$$|E(G)| \leq \frac{n}{2}(1 + \sqrt{4n-3}).$$

Question 5.9.14. Prove the ***coloring lemma***.

Question 5.9.15. Let $n \geq 1$. Show that the expected number of cycles in $G_{n,\frac{1}{n}}$ is at most $\frac{1}{2}\log n$.

$\diamond$

Question 5.9.16* Draw $n+1$ line segments between a set of n points in the plane. Prove that two of the line segments do not intersect. Is this best possible?

Question 5.9.17* Prove that for every graph G with m edges and n vertices, there exists a partition (A, B) of $V(G)$ such that $|A| = \lfloor n/2 \rfloor$ and $|B| = \lceil n/2 \rceil$ and

$$e(A,B) \geq \frac{m}{2} + \frac{m}{2(n-1)}.$$

Is this best possible?

Question 5.9.18* A ***pentagon*** is a cycle C of length five. Prove that the extremal C-free graphs are either complete bipartite graphs with $\lfloor n^2/4 \rfloor$ edges or $n \leq 7$ and the extremal C-free graphs consist of a clique of size $\min\{4, n\}$ and a clique of size $n - \min\{4, n\} + 1$ sharing exactly one vertex.

Question 5.9.19* Let G be an n-vertex d-regular graph. Prove that the number of triangles in G is at least

$$\frac{1}{3}d^2 n - \frac{1}{6}dn^2.$$

Question 5.9.20* Let G be a n-vertex d-regular graph. Prove that the number of quadrilaterals in G is at least

$$\frac{1}{2}\binom{n}{2}\binom{d(d-1)/(n-1)}{2}.$$

Question 5.9.21* Prove that for some constant $c > 0$, if G is any n-vertex graph of minimum degree at least three with m edges, then the number of paths of length three in G is at least cm^3/n^2.

Question 5.9.22* Prove that if G is any n-vertex graph with m edges not containing a 6-cycle, then

$$m \leq 2n^{4/3}.$$

Question 5.9.23* Let $n \geq r \geq 1$ and let G be an n-vertex graph with $\frac{1}{r}\binom{n}{2}$ edges. Prove that G contains a subgraph with $m \geq \sqrt{n/r}$ vertices and minimum degree at least $\frac{1}{r}(m-1)$.

CHAPTER 6

Ramsey Theory

Let $s \geq 2$ and $t \geq 2$ be integers. The ***Ramsey number*** $r(s,t)$ is the minimum n such that for every coloring of $E(K_n)$ with colors red and blue, there exists a $K_s \subset K_n$ all of whose edges are colored red or a $K_t \subset K_n$ all of whose edges are colored blue. We refer to these subgraphs as ***monochromatic***. Equivalently, $r(s,t)$ is the minimum n such that every n-vertex K_s-free graph has an ***independent set*** of t vertices – the graph is precisely the set of red edges in the afore-mentioned coloring of $E(K_n)$. A K_s-free graph with $r(s,t)-1$ vertices and no independent set of size t is called a ***Ramsey graph***, and a coloring of K_n with $n = r(s,t)-1$ vertices and no monochromatic K_s or K_t is called a ***Ramsey coloring***.

One observes that $r(2,t) = t$ for all $t \geq 2$, since a coloring of K_t either contains a red edge, or the entire K_t is blue. The existence of $r(s,t)$ for all $s,t \geq 2$ was established by Erdős and Szekeres [**14**]:

THEOREM 1. (ERDŐS-SZEKERES THEOREM)
For all $s,t \geq 2$,

$$r(s,t) \leq \binom{s+t-2}{s-1}. \tag{6.1}$$

PROOF. We observed $r(2,t) = t$ and we prove $r(s,t) \leq r(s,t-1) + r(s-1,t)$ for $s,t \geq 3$. Let $n = r(s,t)$ and let K_n be edge-colored red and blue. For a vertex $v \in V(K_n)$, if v has more than $r(s,t-1)$ blue neighbors, then amongst those neighbors we have a red K_s or a blue K_{t-1}. Together with v, the blue K_{t-1} gives a blue K_t, and we are done. If v has more than $r(s-1,t)$ red neighbors, then amongst those neighbors we have a red K_{s-1} or a blue K_t, and together with v, the red K_{s-1} gives a red K_s, and we are done. Therefore $r(s,t) \leq r(s,t-1) + r(s-1,t)$. Now we show by induction that $r(s,t) \leq \binom{s+t-2}{s-1}$. When $s = 2$, we already observed $r(2,t) = t$, as required. So we may assume $s \geq 3$, and similarly $t \geq 3$. By induction, and the first part of the proof,

$$r(s,t) \leq r(s,t-1) + r(s-1,t) \leq \binom{s+t-3}{s-1} + \binom{s+t-3}{s-2}.$$

Using ***Pascal's triangle identity***, the right hand side is exactly the upper bound in the theorem. $\square$

The theorem above shows $r(3,3) \leq 6$, and in fact $r(3,3) = 6$ since a pentagon is a Ramsey graph – it is triangle-free and has no independent set of size three. A corollary to the above theorem from (6.1) is

$$r(t,t) \leq 4^{t-1} \tag{6.2}$$

which follows from the fact that $\binom{2t-2}{t-1} \leq 2^{2t-2}$. A central difficulty in Ramsey Theory ◁◁ the construction of ***Ramsey graphs***. A natural construction is to let G be a union of $t-1$ complete graphs of size $s-1$, which shows $r(s,t) \geq (s-1)(t-1)+1$. It turns out that this construction gives a very poor lower bound on $r(s,t)$, and a much better lower bound comes from ***random graphs***, due to Erdős [**10**]:

THEOREM 2. (ERDŐS)
For all $t \geq 3$,

$$r(t,t) \geq \sqrt{2}^t. \tag{6.3}$$

PROOF. Let $n = \lfloor \sqrt{2}^t \rfloor$. Color the edges of K_n randomly, where red and blue each have probability $\frac{1}{2}$. The expected number of red K_t is

$$\binom{n}{t}\left(\frac{1}{2}\right)^{\binom{t}{2}}.$$

We recall

$$\binom{n}{t} = \frac{n(n-1)(n-2)\ldots(n-t+1)}{t!} \leq \frac{n^t}{t!} \leq \frac{\sqrt{2}^{t^2}}{t!}$$

and therefore the expected number of red K_t is at most

$$\frac{\sqrt{2}^{t^2}}{t!} \cdot \left(\frac{1}{2}\right)^{\binom{t}{2}} = \frac{\sqrt{2}^t}{t!} \leq \frac{\sqrt{8}}{6} < \frac{1}{2}.$$

Here we used $t \geq 3$. Similarly, the expected number of blue K_t is less than $\frac{1}{2}$. We conclude that the average coloring of K_n has less than 1 blue or red K_t, and therefore there exists a coloring of $E(K_n)$ with no monochromatic K_t. It follows that $r(t,t) > n$. □

The values of $r(t,t)$ are known only for $t \leq 4$ and the values $r(s,t)$ for $t \geq s \geq 3$ are know only for $s = 3$ and $t \leq 9$ and $s = 4$ and $t \leq 5$. The big conjecture in the area, due to Erdős, is as follows:

CONJECTURE 1. (ERDŐS) *There exists c such that*

$$\lim_{t\to\infty} r(t,t)^{1/t} = c. \tag{6.4}$$

The value of c, if it exists, is between $\sqrt{2}$ and 4 according to (6.2) and (6.3).

One of the special cases is $r(3,4)$: from the Erdős-Szekeres Theorem we find $r(3,4) \leq r(3,3) + r(2,4) = 6 + 4 = 10$. However it turns out that $r(3,4) = 9$. To see this, pick a vertex v in a red-blue colored K_9. If v has at least six blue neighbors, then amongst those neighbors we find a red triangle or a blue triangle, since $r(3,3) = 6$. This gives a red triangle or blue K_4, and we are done. So we can assume no vertex has at least six blue neighbors. Some vertex v must have fewer than five blue neighbors, otherwise the blue graph is 5-regular on nine vertices, which is not possible. That vertex v must have at least four red neighbors. Then none of them are joined by a red edge, else we get a red triangle, so they must contain a blue K_4, and we are done showing $r(3,4) \leq 9$. So see $r(3,4) = 9$, consider the colored K_8 shown below:

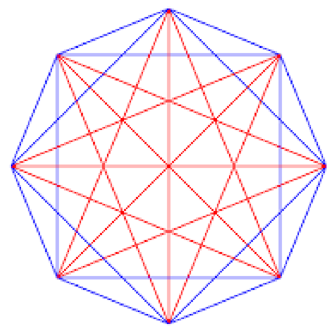


FIGURE 6.1. Ramsey coloring showing $r(3,4) \geq 9$

Using $r(3,4) = 9$, the Erdős-Szekeres Theorem says $r(4,4) \leq r(3,4) + r(4,3) = 18$. It turns out that $r(4,4) = 18$, due to the construction of the following Ramsey graph with 17 vertices. The vertex set of this graph G is $V(G) = \{0, 1, 2, \ldots, 16\}$, and $\{i, j\}$ is an edge if and only if $|i - j| \in \{1, 2, 4, 8\}$ modulo 17.

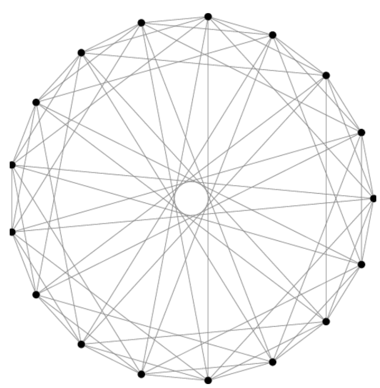


FIGURE 6.2. Ramsey graph showing $r(4,4) \geq 18$

6.1. Graph Ramsey Numbers

For graphs F and G, let $r(F, G)$ be the minimum n such that for every coloring of $E(K_n)$ with colors red and blue, there exists a copy of $F \subset K_n$ all of whose edges are colored red or a copy of $G \subset K_n$ all of whose edges are colored blue. We refer to these subgraphs as ***monochromatic***. In the case $F = K_s$ and $G = K_t$, we have $r(F, G) = r(s, t)$.

Example 6.1. Let $F = K_{1,2}$ and $G = K_t$. Then $r(F, G) = 2t - 1$. To see this, note that the red edges in a coloring of K_n with no red F must form a matching. If $n = 2t-1$, then we can find t vertices with no red edge between them, and they form a blue K_t. Therefore $r(F, G) \le 2t - 1$. Now color K_{2t-2} as follows: select any perfect matching and color it red, and color all remaining edges blue. This has no red F and no blue K_t, since each edge of the perfect matching in a complete blue subgraph has at most one vertex in that blue subgraph.

Upper bounds for $r(F, G)$ sometimes come from extremal graph theory: if we color $E(K_n)$ with two colors, and there is no red F or blue G, then the number of red edges is at most $\mathrm{ex}(n, F)$ and the number of blue edges is at most $\mathrm{ex}(n, G)$, and so $\mathrm{ex}(n, F) + \mathrm{ex}(n, G) \le \binom{n}{2}$.

Example 6.2. For instance, consider $r(C_4, K_t)$. By Turán's Theorem, for $n = r(C_4, t) \ge t$ we have

$$\mathrm{ex}(n, K_t) \le \binom{t-1}{2}\left(\frac{n}{t-1}\right)^2 = \frac{t-2}{t-1} \cdot \frac{n^2}{2}$$

If we color K_n with colors red and blue, then either we get a blue K_t or for $n \ge t$ the number of red edges is at least

$$\binom{n}{2} - \frac{t-2}{t-1} \cdot \frac{n^2}{2} = \frac{n^2}{2(t-1)} - \frac{n}{2}.$$

This must be at most $\mathrm{ex}(n, C_4)$, and by the Kövari-Sós-Turán Theorem,

$$\mathrm{ex}(n, C_4) \le \frac{n}{2}(1 + \sqrt{4n-3}).$$

Therefore for $n \ge t$,

$$\frac{n^2}{2(t-1)} - \frac{n}{2} \le \frac{n}{2}(1 + \sqrt{4n-3})$$

Solving for n gives

$$r(C_4, t) \le \frac{1}{2}(t^2 + t - 2 + (t-1)\sqrt{t^2 + 4t - 8}).$$

It can be shown that the right hand side is less than $(t+1)^2$, so $r(C_4, t) < (t+1)^2$.

6.2. Schur's Theorem

Let $r_k(t)$ denote the minimum n such that for every coloring of $E(K_n)$ with k colors, there exists a copy of $K_t \subset K_n$ all of whose edges are the same color.

Theorem 3. *For all $k \ge 2$,*

$$r_k(3) \le \lfloor ek! \rfloor + 1. \tag{6.5}$$

Proof. Proceed by induction on k, noting that $r_1(3) = 3 = \lfloor e \rfloor + 1$. Let $c : E(K_n) \to \{1, 2, \ldots, k\}$ be a coloring where $n = \lfloor ek! \rfloor + 1$. We have to show that there exists a monochromatic triangle. Suppose not. For $v \in V(K_n)$, some color i is on at least $(n-1)/k$ edges containing v, say $\{v, w_1\}, \{v, w_2\}, \ldots, \{v, w_s\}$. If any two of $w_1, w_2, \ldots, w_s$ are joined by an edge of color i, we have a triangle in color i, a contradiction. So color i is missing from the edges $\{w_i, w_j\}$, $1 \leq i < j \leq s$. Now since there is no monochromatic triangle, $s < r_{k-1}(3)$. But $s \geq (n-1)/k \geq \lfloor e(k-1)! \rfloor + 1 \geq r_{k-1}(3)$ by induction. This contradiction proves the theorem. $\square$

The above theorem gives a Ramsey theorem for the integers. Suppose we color each positive integer with one of a set of k colors. One may then ask for arithmetic structure which is guaranteed to appear in some color class. An early theorem in this vein was proved by Schur [**22**]:

THEOREM 4. (SCHUR'S THEOREM) *Suppose the set $S = \{1, 2, \dots, \lfloor ek! \rfloor + 1\}$ is colored with k colors. Then there exist $x, y, z \in S$ such that $x + y = z$ and x, y, z all have the same color.*

PROOF. Let $n = |S|$ and let K_n denote the complete graph on $[n]$ such that an edge $\{i, j\}$ is colored with the color of $|i - j|$. Then $E(K_n)$ is colored with k colors and has $r_k(3)$ vertices, so there must be a triangle with vertex set $\{h, i, j\}$ such that $|h - i|$ and $|i - j|$ and $|j - h|$ all have the same color. Then we may let $\{x, y, z\} = \{|h - i|, |i - j|, |j - h|\}$ so that $x + y = z$ as required. $\square$

Hilbert [**16**] showed in 1892 that if we color the positive integers with finitely many colors, then there exist arbitrarily large sets S such that translates of the set

$$\{\sum_{s \in S} s\epsilon_s : \epsilon_s \in \{0, 1\}\}$$

all have the same color. The sets are sometimes called ***affine cubes***, and the result of Hilbert is called ***Hilbert's cube lemma***. There have been many generalizations of this result in the literature. The question of which arithmetic substructures are guaranteed to occur in any coloring of the positive integers with finitely many colors is a difficult one, for instance it is not known whether one can guarantee that x, $x + y$ and xy all have the same color for some x and y.

6.3. Exercises

Question 6.3.1° For $n \geq 1$, let $f(n)$ denote the minimum positive integer N such that amongst any set of N points in the plane, there are n points forming the vertices of a convex polygon. Determine $f(n)$ for $n \leq 4$.

Question 6.3.2° Let F and G be graphs such that $\chi(F) = k \geq 2$ and G is connected with at least one edge. Prove that $r(F, G) \geq (\chi(F) - 1)(|V(G)| - 1)$.

Question 6.3.3° Prove that every red-blue coloring of $E(K_n)$ gives a monochromatic spanning tree. Is the same true for red-blue-green colorings?

Question 6.3.4° Prove that $r(4, 5) \leq 32$.

Question 6.3.5° Determine the smallest n such that if $[n]$ is colored with two colors, then there exist $x, y, z \in [n]$ all of the same color such that $x + y = z$.

◇

Question 6.3.6. Prove that in any red-blue coloring of $\mathbb{R}^2$, there exists an equilateral triangle whose vertices all have the same color.

Question 6.3.7. Let M_k denote a matching with k edges. Prove for $k, t \geq 1$ that $r(M_k, K_t) = 2k + t - 2$.

Question 6.3.8. Prove for $k \geq 1$ that $\lceil \frac{3}{2}k \rceil \leq r(P_k, P_k) \leq 2k$.

Question 6.3.9. Prove that for $n > 2^k$, every k-coloring of $E(K_n)$ gives a monochromatic odd cycle.

Question 6.3.10. Prove that if G is an n-vertex graph with more than $2n^2/5$ edges, then for every coloring of $E(G)$ with red and blue, there is a monochromatic triangle.

Question 6.3.11. Let $p \in [0, 1]$ and $1 \leq t \leq n$. Prove that the expected number of independent sets of size t in $G_{n,p}$ is $\binom{n}{t}(1-p)^{\binom{t}{2}}$.

Question 6.3.12. Let $k \geq 1$, and let G be a graph containing no independent set of size at least $|V(G)|/(k+1)$. Prove that there exists a subgraph H of G such that for all independent sets $I \subseteq V(H)$, $|N_H(I)| > k|I|$.

Question 6.3.13. Prove that $r(3, 5) = 14$.

Question 6.3.14. Prove that $r(C_4, C_4) = 6$.

Question 6.3.15. Prove that in any 2-coloring of the edges of K_n, there exist two monochromatic paths P and Q such that $P \cup Q$ is a hamiltonian cycle of K_n.

◇

Question 6.3.16* Prove that for all $n \geq 1$,

$$r(t,t) \geq n - \binom{n}{t} 2^{1-\binom{t}{2}}.$$

Then show that for $t \geq 3$,

$$r(t,t) \geq \frac{t}{e}\sqrt{2}^t.$$

Question 6.3.17* Prove for $k \geq 1$ that $r(P_k, P_k) = \lceil \frac{3}{2}k \rceil$.

Question 6.3.18* For $n \geq 1$, let $f(n)$ denote the minimum positive integer N such that amongst any set of N points in the plane, there are n points in convex position. Prove that $f(5) = 9$.

Question 6.3.19* Let the integers in $[n]$ be colored with two colors where n is large enough. Prove that there exist distinct $x, y, z, w \in [n]$ such that $x + y = z + w$ and x, y, z, w all have the same color.

Question 6.3.20* Let $n \geq 1$. Prove that if p is a large enough prime number, then there exist non-zero x, y and z such that $x^n + y^n - z^n$ is divisible by p.

Question 6.3.21* Let $k \geq 1$ and let σ be a sequence of $k^2 + 1$ distinct integers. Prove that σ contains an increasing sequence of length k or a decreasing sequence of length k.

Question 6.3.22* Prove that in any coloring of $\mathbb{R}^2$ with three colors, there are two points at distance one with the same color.

Question 6.3.23* Let $t \geq k \geq 2$ and $s \geq 2$. Prove that if $n^t < t!s^{\binom{t}{k}-1}$, then there exists an s-coloring of the subsets of $[n]$ of size k such that no subset of $[n]$ size t has all its subsets of size k the same color.

CHAPTER 7

Flows and Cuts

Let $G = (V, E)$ be a digraph and $s, t \in V$. We shall refer to s as the ***source*** vertex and t as the ***sink*** vertex in what follows. Let $\mathbb{R}_{\geq 0}$ denote the non-negative real numbers. We want to define what it means for the network to have a flow from s to t. A ***static st-flow*** is a function $f : E \to \mathbb{R}_{\geq 0}$ such that for all $x \in V \backslash \{s, t\}$:

$$\sum_{y \in N^+(x)} f(x, y) = \sum_{y \in N^-(x)} f(y, x)$$

The flow in edge (u, v) is denoted $f(u, v)$. This last requirement is known as ***Kirchoff's Law***, in words it states that the flow into a vertex must be equal to the flow out of a vertex, and that this must be true for all vertices apart from s and t. The flow in an edge $e \in E$ is denoted $f(e)$. The ***value of a flow*** f from s to t is defined by

$$v(f) = \sum_{y \in N^+(s)} f(s, y) - \sum_{y \in N^-(s)} f(y, s).$$

This is the net amount of flow leaving the source s.

The ***capacity*** of an edge $e \in E$ is a non-negative real number and denoted $c(e)$, and is the maximum possible amount of flow in the edge. An example of a network with capacities is shown below. Here $c(s, u) = c(v, w) = c(x, w) = c(v, u) = 1$ and $c(s, v) = c(u, w) = 4$ and $c(w, t) = 3$ and $c(u, x) = c(x, u) = c(x, t) = 2$.

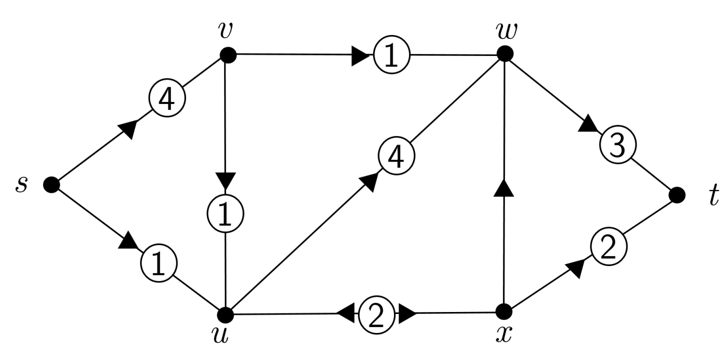


FIGURE 7.1. Capacities

We require that if f is a flow in the network, then $f(e) \leq c(e)$ for every edge $e \in E$. A ***maximum flow*** in a network with capacities is a flow f^* such that for every flow f, $v(f^*) \geq v(f)$ – it is a flow with largest value.

If $S \subset V$ and $s \in S$ and $t \notin S$, then the ***cut induced by S*** is the set of edges from S to $V \backslash S$ – this is the set of edges leaving S. This set of edges is denoted $(S, \overline{S})$, and is called an ***st-cut*** or simply a cut. The ***capacity*** of a cut $(S, \overline{S})$ is defined by

$$c(S, \overline{S}) = \sum_{e \in (S,\overline{S})} c(e).$$

It is clear that if f is a flow in the network, then $v(f) \leq c(S, \overline{S})$ for any cut $(S, \overline{S})$. A cut $(S, \overline{S})$ ***minimum cut*** is a cut $(S, \overline{S})$ such that $c(T, \overline{T}) \geq c(S, \overline{S})$ for every cut $(T, \overline{T})$ – so it is a cut with a minimum value of $c(S, \overline{S})$. Our main theorem says that the minimum capacity of a cut equals the maximum value of a flow. First we need a lemma:

LEMMA 1. *For any flow f and any cut $(S, \overline{S})$,*

$$v(f) = \sum_{x \in S} \sum_{y \notin S} f(x, y) - \sum_{x \in S} \sum_{y \notin S} f(y, x).$$

PROOF. By Kirchoff's Law, for all $x \notin \{s, t\}$,

$$\sum_{y \in N^+(x)} f(x, y) \quad = \quad \sum_{y \in N^-(x)} f(y, x)$$

Summing over $x \in S \backslash \{s\}$ we get

$$\sum_{x \in S \backslash \{s\}} \sum_{y \in N^+(x)} f(x, y) \quad = \quad \sum_{x \in S \backslash \{s\}} \sum_{y \in N^-(x)} f(y, x).$$

All those $f(x, y)$ with $x, y \in S$ cancel out: they are counted once in the left sum, and again once in the right sum. So we get

$$\sum_{x \in S \backslash \{s\}} \sum_{y \notin S} f(x, y) \quad = \quad \sum_{x \in S \backslash \{s\}} \sum_{y \notin S} f(y, x).$$

By definition of $v(f)$, this means

$$\sum_{x \in S} \sum_{y \notin S} f(x, y) - \sum_{x \in S} \sum_{y \notin S} f(y, x) - v(f) = 0$$

and this is the required result. $\square$

The definition of the value of the flow is for the specific cut $(S, \overline{S})$ with $S = \{s\}$, but this lemma says we can measure the value of a flow in any network by just looking at the net flow across a cut in the network.

7.1. The Max-Flow Min-Cut Theorem

The main theorem we prove is the following:

THEOREM 1. (Max-Flow Min-Cut Theorem)
In any network with capacities, the maximum value of a flow equals the minimum value of a cut.

PROOF. Let f be a maximum flow[1]. Then $v(f) \leq c(S, \overline{S})$ for every cut $(S, \overline{S})$ so $v(f) \leq \min c(S, \overline{S})$. To prove the theorem, we define a set $S \subset V$ with $c(S, \overline{S}) = v(f)$. First put s, the source, into S. Then for every edge (x, y) such that $x \in S$ and $c(x, y) > f(x, y)$, put $y \in S$, and for every edge (y, x) with $x \in S$ and $f(y, x) > 0$, put $y \in S$. We claim $t \notin S$ and $c(S, \overline{S}) = v(f)$. Suppose that $t \in S$. Then there exists a path $x_1 x_2 \dots x_r$ where $x_i \in S$ for all i and $x_1 = s$ and $x_r = t$ and, by definition of S,

$$c(x_i, x_{i+1}) - f(x_i, x_{i+1}) > 0 \qquad \text{or} \qquad f(x_{i+1}, x_i) > 0$$

for each i. Let ε be the smallest of all these positive numbers. Define a new flow g by taking $g(x_i, x_{i+1}) = f(x_i, x_{i+1}) + \varepsilon$ if $c(x_i, x_{i+1}) - f(x_i, x_{i+1}) > 0$, and taking $g(x_{i+1}, x_i) = f(x_{i+1}, x_i) - \varepsilon$ if $f(x_{i+1}, x_i) > 0$. Then $v(g) = v(f) + \varepsilon$, contradicting the maximality of f. We conclude $t \notin S$. Finally, since $f(y, x) = 0$ for every $x \in S$ and $y \notin S$, by Lemma 1 we have

$$v(f) = \sum_{x \in S} \sum_{y \notin S} f(x, y) - \sum_{x \in S} \sum_{y \notin S} f(y, x) = \sum_{(x,y) \in (S, \overline{S})} c(x, y) \quad = \quad c(S, \overline{S}).$$

This completes the proof. $\square$

The proof of the Max-Flow Min-Cut Theorem gives an algorithm for finding a maximum flow as well as a minimum cut. To construct a maximum flow f^* and a minimum cut $(S^*, \overline{S}^*)$, proceed as follows: start by letting f be the zero flow and $S = \{s\}$ where s is the source. Construct a set S as in the theorem: whenever there is an arc (x, y) such that $f(x, y) < c(x, y)$ and $x \in S$ and $y \notin S$, or an arc (y, x) such that $f(y, x) > 0$ and $x \in S$ and $y \notin S$, add y to S. If at the end of this procedure, $t \notin S$, then let $S^* = S$ to get a minimum cut and the current flow is a maximum flow. If at the end of this procedure $t \in S$, then there must be a path $x_0 x_1 x_2, \dots x_r$ where $s = x_0$ and $t = x_r$, along which f can be augmented by some value $\varepsilon > 0$. The value of ε is given in the proof above: it is

$$\min\{c(x_i, x_{i+1}) - f(x_i, x_{i+1}), f(x_{i+1}, x_i) | 0 \leq i < r\}.$$

Now restart with the augmented flow which is $f(x_i, x_{i+1}) + \varepsilon$ and $c(x_i, x_{i+1}) > f(x_i, x_{i+1})$ and $f(x_{i+1}, x_i) - \varepsilon$ if $f(x_{i+1}, x_i) > 0$, for each $i : 0 \leq i < r$. Now we start again with $S = \{s\}$ and the new flow a input.

EXAMPLE 7.1. Consider the network with capacities in Figure 7.1. According to the algorithm, start by letting f be the zero flow and $S = \{s\}$:

[1]Why does a maximum flow even exist? Prove that a maximum flow exists.

arc	flow	capacity
(s,u)	0	1
(s,v)	0	4
(v,w)	0	1
(x,u)	0	2
(u,x)	0	2
(u,w)	0	4
(x,t)	0	2
(w,t)	0	3

Since $f(s,v) = 0 < c(s,v) = 4$, we put v into S. Since $c(v,w) > f(v,w)$ we put w into S. Then put t into S since $c(w,t) > f(w,t)$. We stop since we have placed t in S. By the algorithm, there is a way to augment f: we consider the path $svwt$. We have the smallest difference between capacities and flows in the arcs of this path equal to 1. So we augment f to $f(s,v) = f(v,w) = f(w,t) = 1$. Now we start again with $S = \{s\}$ and the new flow.

arc	flow	capacity
(s,u)	0	1
(s,v)	1	4
(v,w)	1	1
(v,u)	0	1
(x,u)	0	2
(u,x)	0	2
(u,w)	0	4
(x,t)	0	2
(w,t)	1	3

Since $c(s,u) = 1$ and $f(s,u) = 0$, we add $u \in S$. Since $c(u,x) = 2$ and $f(u,x) = 0$, add $x \in S$. Since $c(x,t) = 2$ and $f(x,t) = 0$, add $t \in S$. So $S = \{s,u,x,t\}$ and since $t \in S$, we stop and we augment f by $\min\{1,2,2\} = 1$ along the path $suxt$ to get

arc	flow	capacity
(s,u)	1	1
(s,v)	1	4
(v,w)	1	1
(v,u)	0	1
(x,u)	0	2
(u,x)	1	2
(u,w)	0	4
(x,t)	1	2
(w,t)	1	3

Let $S = \{s\}$. Since $c(s,v) = 4$ and $f(s,v) = 1$, we can put $v \in S$. We cannot put $w \in S$ since $c(v,w) = 1 = f(v,w)$. But we can put $u \in S$ since $c(v,u) = 1$ and $f(v,u) = 0$. Then we can put $x \in S$ since $c(u,x) = 2$ and $f(u,x) = 1$. Finally we put $t \in S$ since $f(x,t) = 1$ and $c(x,t) = 2$. Since $t \in S$, we stop and augment f by $\min\{3,1,1,1\} = 1$ along the path $svuxt$ to get

arc	flow	capacity
(s,u)	1	1
(s,v)	2	4
(v,w)	1	1
(v,u)	1	1
(x,u)	0	2
(u,x)	2	2
(u,w)	0	4
(x,t)	2	2
(w,t)	1	3

Let $S = \{s\}$. Since $c(s,v) = 4$ and $f(s,v) = 2$, we put $v \in S$. But now $c(v,w) = f(v,w) = 1$, $c(v,u) = f(v,u) = 1$ and $c(s,u) = f(s,u) = 1$. So $S = \{s,v\}$ and this induces a minimum cut $(S,\overline{S})$. The flow we have just defined is a maximum flow, with value three, and notice $c(S,\overline{S}) = 3$, as expected.

7.2. Proof of Hall's Theorem

The Max-Flow Min-Cut Theorem gives an alternative proof of Hall's Theorem, as follows. We have a bipartite graph G with parts A and B satisfying Hall's Condition, $|N(X)| \geq |X|$ for all $X \subseteq A$ and all $X \subseteq B$. Orient all edges of G from A to B, add a vertex a joined to all vertices in A and a vertex b joined from all vertices in B, and assign all edges capacity 1. Here a and b play the rôle of source and sink. Let $S \subset V(G) \cup \{a,b\}$ contain a but not b. We claim $c(S,\overline{S}) = |A|$. Let $X = A \backslash S$ and $Y = B \backslash S$. Then $|N(A\backslash S)| \geq |A| - |S \cap A|$ and $|N(B\backslash S)| \geq |B| - |S \cap B|$ by Hall's Condition. It follows that

$$c(S,\overline{S}) \geq |A| - |S \cap A| + |B| - |S \cap B| \geq |A| + |B| - |S| \geq |A|.$$

A minimum cut therefore has $S = \{a\}$, and by max-flow min-cut, there is a flow of value $|A|$. The edges of G with unit flow form a perfect matching of G. $\square$ ◁◁

The max flow min cut theorem gives an efficient algorithm for finding a ***maximum matching*** in a bipartite graph G with parts A and B. If H is the digraph formed in the proof above by adding source a and sink b, and f is a maximum flow in H, then the set of edges $e \in E(G)$ such that $f(e) = 1$ is the edge-set of a maximum matching in G. ◁◁

7.3. Menger's Theorems

An ***edge-cut*** of a graph G is a set L of edges such that $G - L$ is disconnected. A graph G is ***k-edge-connected*** if the minimum size of an edge-cut in G is at least k. For vertices $u, v \in V(G)$, a ***uv-separator*** is a set L of edges such that u and v are in different components of $G - L$. The following is one of the cornerstones of connectivity theory in graphs:

THEOREM 2. (MENGER'S THEOREM - EDGE FORM) *The minimum size of a uv-separator in a graph G equals the maximum size of a set of pairwise edge-disjoint*

uv-paths in G. In particular, a graph is k-edge-connected if and only if each pair of its vertices is connected by k pairwise edge-disjoint paths.

Proof. Form a digraph D from G by placing an edge in both directions between each pair of adjacent vertices in G. Add capacity 1 to every edge, and let u be the source and v be the sink. If S is a uv-cut, then the capacity of S is equal to the number of edges with one end in S and one end in $\overline{S}$. That set L of edges is a uv-separator, and so the minimum capacity of a uv-cut is equal to the maximum value of a uv-flow, by the Max-Flow Min-Cut Theorem. Let f be a uv-flow, say with value k. Then for every edge $\{x, y\}$ of G, $f(x, y) \in \{0, 1\}$ and $f(y, x) \in \{0, 1\}$. If $f(x, y) = 1 = f(y, x)$, then change the flow to have $f(x, y) = 0 = f(y, x)$ – this does not change the value of the flow. Starting with the zero flow, we reached f by increasing the flow from u to v by one unit along k paths $P_1, P_2, \ldots, P_k$ from u to v. Since the capacity of every edge of the digraph is 1, and $f(x, y) = 1$ or $f(y, x) = 1$ but not both, the paths $P_1, P_2, \ldots, P_k$ are edge-disjoint. This proves the theorem. $\square$

In fact one can get a version of this theorem for digraphs directly: an edge-cut in a digraph G is a set L of edges of the digraph such that $G - L$ is not strongly connected. A digraph is ***strongly k-edge-connected*** if any edge-cut has size at least L. Then the above proof shows that a digraph is strongly k-edge-connected if and only if for every ordered pair (u, v) of vertices, there exist k edge-disjoint directed paths from u to v.

A ***vertex-cut*** of a graph G is a set of X of vertices such that $G - X$ is disconnected. A graph is ***k-connected*** if the minimum size of a vertex-cut is at least k. Let $u, v \in V(G)$. A set U of vertices is called a ***uv-separator*** if u and v are in different components of $G - U$. Paths P and Q are ***internally disjoint*** if $V(P) \cap V(Q)$ contains only ends of P and Q. The following is Menger's Theorem for vertex-cuts:

Theorem 3. (Menger's Theorem - Vertex Form) *Let $u, v \in V(G)$ be non-adjacent vertices. Then the minimum size of a uv-separator of vertices in a graph G equals the maximum size of a set of pairwise internally disjoint uv-paths in G. In particular, a graph is k-connected if and only if each pair of its vertices is connected by k pairwise edge-disjoint paths.*

This theorem can also be proved via the Max-Flow Min-Cut Theorem.

7.4. Exercises

Question 7.4.1° Find a maximum st-flow and minimum st-cut in the network shown below using the proof of the Max-Flow Min-Cut Theorem. Show all working. The current flow in the network is zero, and the capacities are shown as numbers next to the arcs.

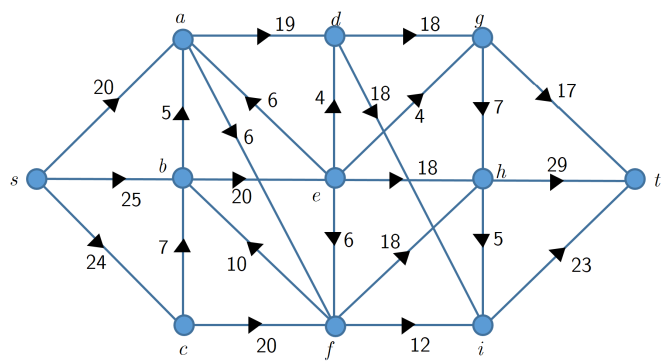


FIGURE 7.2. Network

Question 7.4.2° Find a maximum st-flow in the network shown in Figure 7.3, starting with the given flow f consisting of unit flow in the st-path of length four at the top of the diagram. Also find a minimum cut in the network. The capacities of the arcs are denoted by numbers next to each arc.

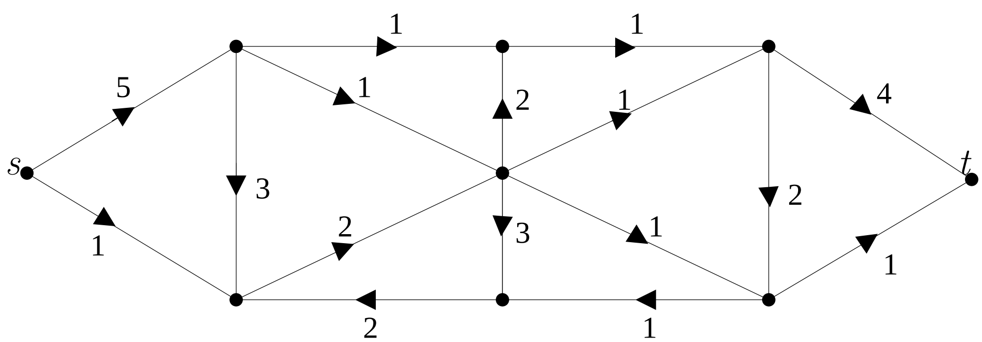


FIGURE 7.3. Network

Question 7.4.3° Find a maximum st-flow in the network shown in Figure 7.5, starting with the zero flow. Also find a minimum cut in the network. The capacities of the arcs are shown as numbers next to each arc.

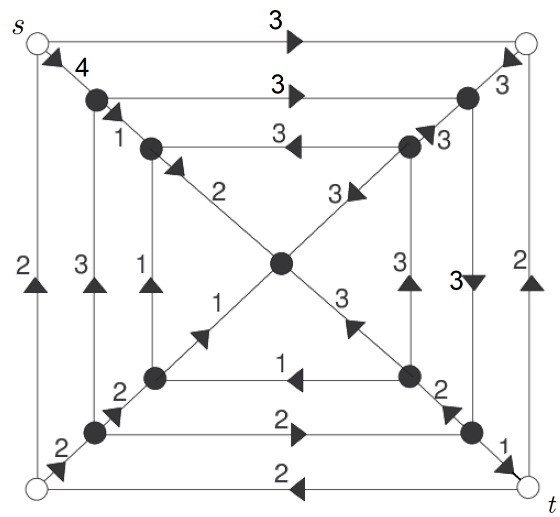


FIGURE 7.4. Network

Question 7.4.4° Find a maximum st-flow and minimum st-cut in the network below with source s and sink t by applying the Max-Flow Min-Cut Algorithm. The capacities of each arc are shown alongside the arcs as numbers below and the current flow is zero.

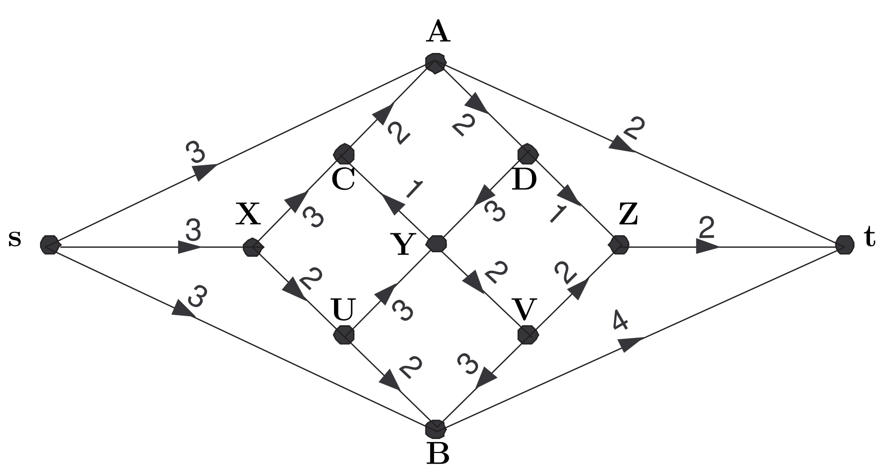


FIGURE 7.5. Network

◇

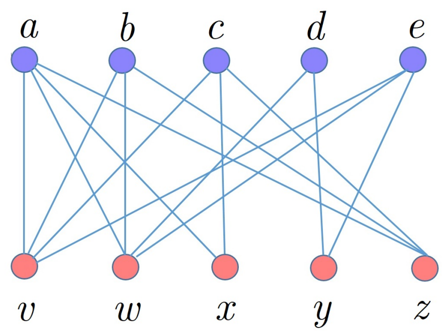
Question 7.4.5. In a network with a set Σ of sources and a set T of sinks, explain how you would find a maximum flow and minimum cut from Σ to T.

Question 7.4.6. Explain how to use the Max-Flow Min-Cut Theorem to find a maximum matching in a bipartite graph $G = (A \cup B, E)$.

Question 7.4.7. Let G be a digraph with source s and sink t and integer capacities. Prove that if f is a maximum flow, then $f(e)$ is an integer for each arc e of G.

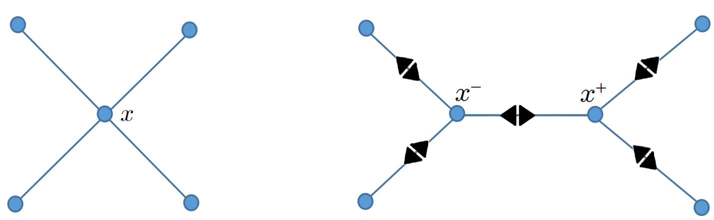
Question 7.4.8. Let G be a 3-connected graph. Prove that G contains an even cycle.

Question 7.4.9. Find a maximum matching in the bipartite graph below.



◇

Question 7.4.10* Prove the vertex form of Menger's Theorem for a graph G using the Max-Flow Min-Cut Theorem as follows: replace each vertex x of the graph with vertices x^- and x^+ joined by edges (x^-, x^+) and (x^+, x^-) of capacity 1, and add (x^+, y^-) and (y^-, x^+) of capacity 1 whenever $\{x, y\} \in E(G)$.



CHAPTER 8

Extremal Set Theory

Let $[n] := \{1, 2, \ldots, n\}$. A ***set system*** or ***hypergraph*** is a set $\mathcal{S}$ of subsets of $[n]$. The main goal of extremal set theory is to maximize or minimize the size of a set system $\mathcal{S}$ on $[n]$ subject to constraints on the sizes of the sets in $\mathcal{S}$ and binary relations between sets in $\mathcal{S}$, such as sizes of intersections or unions.

EXAMPLE 8.1. A ***laminar family*** is a set system $\mathcal{S}$ such that for any sets $S, T \in \mathcal{S}$, $S \cap T = \emptyset$ or $S \subseteq T$ or $T \subseteq S$. A laminar family

$$\{\{1\}, \{2\}, \{3\}, \{1,2,3\}, \{5\}, \{4,5,6\}, \{1,2,3,4,5,6\}, \{7,8,9,10,11,12\}, \{7,9,11,12\}, \{8,10\}\}$$

is shown below.

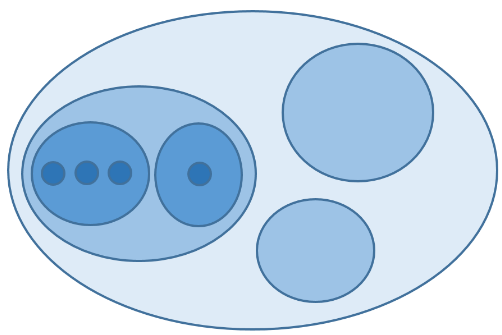


FIGURE 8.1. A laminar family

How many sets can a laminar family $\mathcal{S}$ of subsets of $[n]$ have? Clearly $\mathcal{S}_0 = \{\{i\} : 2 \leq i \leq n\}$ is a laminar family and if we add $\mathcal{C} = \{\emptyset, \{1\}, \{1,2\}, \{1,2,3\}, \ldots, \{1,2,\ldots,n\}\}$ we get a laminar family $\mathcal{S}_0 \cup \mathcal{C}$ of size $2n$. We leave it as an exercise to prove by induction that any laminar family $\mathcal{S}$ of subsets of $[n]$ has size at most $2n$, with equality if and only if the family if $\mathcal{S}_0 \cup \mathcal{C}$.

A set system $\mathcal{S}$ is ***r-uniform*** if every set in $\mathcal{S}$ has size r. It is convenient to let $\binom{[n]}{r}$ denote the set system of all subsets of $[n]$ of size r. The edge-set of a graph, for instance, is a 2-uniform set system.

8.1. Antichains – Sperner's Theorem

A ***chain*** is a sequence $(A_0, A_1, \ldots, A_n)$ of sets such that $A_0 \subset A_1 \subset \cdots \subset A_n$. A ***maximal chain*** has $|A_i| = i$ for all $i \leq n$. An ***antichain*** is a set system $\mathcal{S}$ such that whenever $e, f \in \mathcal{S}$ and $e \subseteq f$, $e = f$. In other words, no two sets in $\mathcal{S}$ are contained in one another. A simple example of an antichain is the set system $\binom{[n]}{r}$ for $0 \leq r \leq n$. ***Sperner's Theorem*** **[24]** answers the question of determining the size of a largest antichain of subsets of $[n]$.

Theorem 1. (Sperner's Theorem)
Let $\mathcal{S}$ be an antichain of subsets of $[n]$ and $m = \lfloor n/2 \rfloor$. Then

$$|\mathcal{S}| \leq \binom{n}{\lfloor n/2 \rfloor} \tag{8.1}$$

with equality for n even if and only if $\mathcal{S} = \binom{[n]}{m}$.

Proof. There are $n!$ maximal chains, since for each ordering $(i_1, i_2, \ldots, i_n)$ of $[n]$ we can form the maximal chain $(\emptyset, \{i_1\}, \{i_1, i_2\}, \ldots, \{i_1, i_2, \ldots, i_n\})$. We count pairs (e, C) where $e \in \mathcal{S}$ and C is a maximal chain containing e. For each set $e \in \mathcal{S}$, there are exactly $|e|!(n - |e|)!$ maximal chains containing e. For each maximal chain C, there exists at most one $e \in \mathcal{S}$ that C contains, since $\mathcal{S}$ is an antichain. Therefore

$$\sum_{e \in \mathcal{S}} |e|!(n - |e|)! \leq n!. \tag{8.2}$$

Next we note that if a and $n-a$ are further apart than b and $n-b$, then $a!(n-a)! > b!(n-b)!$. It follows that $|e|!(n-|e|)! \geq m!(n-m)!$ with equality if and only if $|e| = m$ or $|e| = m + 1$. In particular, from (8.2),

$$|\mathcal{S}| \cdot m!(n-m)! \leq n!$$

and this implies $|\mathcal{S}| \leq \binom{n}{m}$ with equality if and only if every set in $\mathcal{S}$ has size m if n is even. This completes the proof. $\square$

The proof above actually gives quite a bit more than the bound (8.1): we notice that if $H_r = \{A \in H : |A| = r\}$, then (8.2) gives

$$\sum_{r=0}^{n} |H_r| r!(n-r)! \leq n!.$$

Dividing by $n!$, we obtain the Lubell-Yamamoto-Melshalkin-Bollobás Inequality **[4]**, or ***LYMB*** inequality:

Theorem 2. *Let $\mathcal{S}$ be an antichain of subsets of $[n]$ and $H_r = \{A \in H : |A| = r\}$. Then*

$$\sum_{r=0}^{n} \frac{|H_r|}{\binom{n}{r}} \leq 1. \tag{8.3}$$

8.2. Delta Systems – Erdős-Rado Theorem

A ***delta-system*** or ***sunflower*** is a set system $\mathcal{S}$ such that for some set C and any distinct edges $e, f \in \mathcal{S}$, $e \cap f = C$. The set C is called the ***core*** of the delta system. For example, the 3-uniform delta systems consisting of three sets are $\{\{1,2,3\},\{4,5,6\},\{7,8,9\}\}$ (this has $\emptyset$ as the core), $\{\{1,2,3\},\{1,4,5\},\{1,6,7\}\}$ (this has $\{1\}$ as the core) and $\{\{1,2,3\},\{1,2,4\},\{1,2,5\}\}$ (this has $\{1,2\}$ as the core). The following theorem is due to Erdős and Rado [**9**]:

Theorem 3. (Erdős-Rado Theorem) *Let $r, s \geq 1$, and let $\mathcal{S}$ be an r-uniform set system. If $|\mathcal{S}| > r!(s-1)^r$, then $\mathcal{S}$ contains a delta system of size s.*

Proof. Let $\mathcal{S}$ be a family of r-sets without a sunflower of size s. Let $e_1, e_2, \dots, e_t$ be a maximum subfamily of pairwise disjoint sets in $\mathcal{S}$. Since a family of pairwise disjoint sets is a sunflower, we must have $t < s$. Now let $e = \bigcup_{i=1}^t e_i$. For every $a \in e$ consider the family

$$\mathcal{S}_a = \{f \backslash \{a\} : f \in \mathcal{S}, a \in f\}.$$

Now, the size of e is at most $(s-1)r$ and the size of each $\mathcal{S}_a$ is at most $(r-1)!(s-1)^{r-1}$ since a delta system in $\mathcal{S}_a$ gives a delta system in $\mathcal{S}$. Therefore $|\mathcal{S}| \leq (s-1)r(r-1)!(s-1)^{r-1} = r!(s-1)^r$. □

The most well-known conjecture in the area, due to Erdős, is as follows:

Conjecture 1. (Erdős) *There exists $c > 0$ such that for all $r \geq 3$, if $\mathcal{S}$ is an r-uniform set system with no sunflower of size three, then $|\mathcal{S}| \leq c^r$.*

The above proof gives $|\mathcal{S} \leq r!2^r$. To get an example of an r-uniform set system with no delta system of size three and 2^r sets, consider all sets $\{i_1, i_2, \dots, i_r\} \subset [2r]$ such that $i_j \in \{2j-1, 2j\}$ for $j \in [r]$. One can do slightly better and get a^r sets for some $a > 2$, however, the best and very recent advance on the above conjecture due to Alweiss, Lovett, Wu and Zhang (https://arxiv.org/abs/1908.08483) shows that if $\mathcal{S}$ is an r-uniform set system on $[n]$ containing no delta system of size three, then $|\mathcal{S}|$ is roughly at most $(\log n)^r$.

8.3. Intersecting Families – Erdős-Ko-Rado Theorem

A set system $\mathcal{S}$ is called ***intersecting*** if for all $e, f \in \mathcal{S}$, $e \cap f \neq \emptyset$. An easy example of an intersecting set system consists of all sets $e \subseteq [n]$ containing 1; this set system has size 2^{n-1}. This set system is an extremal set system:

Theorem 4. *Let $n \geq 1$ and let $\mathcal{S}$ be an intersecting set system of subsets of $[n]$. Then $|\mathcal{S}| \leq 2^{n-1}$.*

Proof. For each set $e \subseteq [n]$, at most one of e and $e^c = [n] \backslash e$ is in $\mathcal{S}$. The number of pairs $\{e, e^c\}$ is exactly 2^{n-1}, since $\{e, e^c\} = \{f, f^c\}$ when $f = e^c$. Therefore $|\mathcal{S}| \leq 2^{n-1}$. □

The proof gives many extremal intersecting set systems: for each pair $\{e, e^c\}$ place either e in $\mathcal{S}$ or e^c in $\mathcal{S}$. The extremal problem for intersecting set systems becomes trickier if we insist the set system is r-uniform. Let $\mathcal{S}_n^*$ denote the r-uniform set system consisting of all sets containing a fixed element of $[n]$. This problem was answered by Erdős, Ko and Rado [**8**]:

THEOREM 5. *Let $\mathcal{S}$ be an r-uniform intersecting set system of subsets of $[n]$ where $n \geq 2r+1$. Then*

$$|\mathcal{S}| \leq \binom{n-1}{r-1}$$

with equality if and only if $\mathcal{S} = \mathcal{S}_n^*$.

PROOF. Count pairs (σ, e) where σ is a cyclic permutation of $[n]$ and $\sigma(e)$ is a cyclic interval. There are $(n-1)!$ choices for σ. A key fact is that for $n \geq 2r+1$ the largest ◁◁ intersecting family of cyclic intervals in $[n]$ consists of all r intervals containing a fixed point in $[n]$, so having picked σ, there are at most r choices of an edge $e \in \mathcal{S}$ such that $\sigma(e)$ is a cyclic interval. On the other hand, there are $|\mathcal{S}|$ choices of e, and then $r!(n-r)!$ choices of σ such that $\sigma(e)$ is a cyclic interval. Therefore

$$|\mathcal{S}|r!(n-r)! \leq r(n-1)!$$

which gives the required bound. We leave the case of equality as an exercise. □ ◁◁

8.4. Linear algebra – L-intersecting families

Let L be a set of non-negative integers and $\mathcal{S}$ be a set system of subsets of $[n]$. Then $\mathcal{S}$ is ***L-intersecting*** if $|S \cap T| \in L$ for all distinct $S, T \in \mathcal{S}$. If L is the set of positive integers, then $\mathcal{S}$ is an intersecting set system, as in the last section. In this section, we address the extremal question of the maximum size of certain L-intersecting set systems of subsets of $[n]$. This illustrates and application of the ***linear algebra method***. For a complete description of this method and its applications, we refer the reader to the monograph of Babai and Frankl entitled linear algebra methods in combinatorics. For a set M of non-negative integers, we call a set system M-uniform if each of its sets has size in M.

THEOREM 6. *Let p be a prime and let L be a set multiples of p and M a set of non-multiples of p. If $\mathcal{S}$ is an L-intersecting M-uniform set system of subsets of $[n]$, or an M-intersecting L-uniform set system of subsets of $[n]$, then $|\mathcal{S}| \leq n$.*

PROOF. For each $S \in \mathcal{S}$, let $\chi(S)$ be the vector in $\{0,1\}^n$ such that $\chi(S)_i = 1$ if $i \in S$ and $\chi(S)_i = 0$ otherwise. We show that the vectors $\chi(S)$ are linearly independent over the finite field $\mathbb{F}_p$. First consider the case that $\mathcal{S}$ is L-intersecting and M-uniform. Since $S \cap T \in L$ for all distinct $S, T \in \mathcal{S}$, $\chi(S) \cdot \chi(T) \in L$ for all distinct $S, T \in \mathcal{S}$. Mod p, we have $\chi(S) \cdot \chi(T) = 0$. Furthermore, $\chi(S) \cdot \chi(S) = |S| \in M$. Suppose we can find linearly dependent vectors $x_1 = \chi(S_1), x_2 = \chi(S_2), \ldots x_k = \chi(S_k)$ where $S_1, S_2, \ldots, S_k \in \mathcal{S}$ are distinct sets in $\mathcal{S}$. Then for some elements $c_1, c_2, \ldots, c_k \in \mathbb{F}_p$, not all zero,

$$c_1x_1 + c_2x_2 + \cdots + c_kx_k = 0 \mod p.$$

If $c_i \neq 0$, then taking the dot product of this equation with x_i, we get

$$c_ix_i \cdot x_i = 0 \mod p.$$

However, $x_i \cdot x_i = |S_i| \in M$ which is not a multiple of p, so this is a contradiction. Therefore the $\chi(S)$ for $S \in \mathcal{S}$ are linearly independent over $\mathbb{F}_p$, so there cannot be more than n of them and $|\mathcal{S}| \leq n$. The proof if $\mathcal{S}$ is M-uniform and L-intersecting is similar. $\square$

The ***odd town problem*** is to determine the maximum number of clubs in a town with n residents such that every club has an even number of members and any two different clubs intersect in one member. Let $\mathcal{S}$ be the set system on $[n]$ whose sets are the set of members in each club. Then $|S| = 0 \bmod 2$ for every $S \in \mathcal{S}$, and $|S \cap T| = 1 \bmod 2$ for every distinct $S, T \in \mathcal{S}$. By the theorem, with $M = \{0\}$ and $L = \{1\}$, it follows that $|\mathcal{S}| \leq n$ – the town has at most n clubs.

8.5. Exercises

Question 8.5.1. Describe all r-uniform set systems $\mathcal{S}$ such that $|e \cap f| = r - 1$ for all distinct $e, f \in \mathcal{S}$.

Question 8.5.2. Let $\mathcal{S} \subseteq \binom{[n]}{r}$. Suppose every $i \in [n]$ is contained in exactly d sets in $\mathcal{S}$. Determine $|\mathcal{S}|$ in terms of n, d and r.

Question 8.5.3. Let $r \geq 2$ and let $\mathcal{S}$ be an r-uniform $\{1\}$-intersecting family on $[n]$. Prove that if n is large enough, then $\bigcap \mathcal{S} \neq \emptyset$.

Question 8.5.4. Prove that the set system consisting of all sets $\{i_1, i_2, \dots, i_r\} \subset [2r]$ such that $i_j \in \{2j-1, 2j\}$ for $j \in [r]$ does not contain a delta system of size three.

Question 8.5.5. A triangle consists of sets e, f, g such that $e \cap f \neq \emptyset$, $f \cap g \neq \emptyset$ and $g \cap e \neq \emptyset$ and $e \cap f \cap g = \emptyset$. Prove that if $\mathcal{S} \subseteq [2r]^{(r)}$ does not contain a triangle, then $|\mathcal{S}| \leq \binom{2r-1}{r-1}$.

Question 8.5.6. Let $\mathcal{S}$ be a triple system on $[n]$ such that $|S \cap T| \in \{0, 2\}$ for all distinct $S, T \in \mathcal{S}$. Prove that $|\mathcal{S}| \leq n$ and for each n a multiple of four, give an example to show that equality is possible.

Question 8.5.7. Let A be a set of real numbers such that $|a| \geq 1$ for each $a \in A$. Prove that there are at most $\binom{|A|}{\lfloor |A|/2 \rfloor}$ sets $S \subset A$ such that $0 < \sum_{a \in S} a < 1$.

$\diamond$

Question 8.5.8* Let $k \geq 0$ and let $\mathcal{S}$ be a set system on $[n]$ such that $|S \cap T| = k$ for all distinct $S, T \in \mathcal{S}$. Prove that $|\mathcal{S}| \leq n$.

Question 8.5.9* A quadrilateral consists of sets e, f, g, h of size $2r$ such that $|e \cap f| = |f \cap g| = |g \cap h| = |h \cap e| = r$ and $e \cap g = \emptyset$ and $f \cap h = \emptyset$. Prove that if $\mathcal{S} \subseteq \binom{[n]}{2r}$ contains no quadrilateral and $n \geq 2r$, then $|\mathcal{S}| \leq \binom{n}{r}^{3/2}$.

Question 8.5.10* Suppose a town has clubs $A_1, A_2, \dots, A_k$ such that $|A_i \cap A_j|$ is even for all $\{i, j\} \subset [k]$. Prove that $k \leq 2^{n/2}$ and show that equality is possible for even n.

Question 8.5.11* Prove for some $c > 2$ that there exists a set system $\mathcal{S}$ on $[n]$ of size at least c^n containing no delta system of size three.

Question 8.5.12* Suppose each set of integers is colored red or blue. Does there exist an infinite set all of whose subsets have the same color? For $k \geq 1$, does there exist an infinite set all of whose subsets of size k have the same color?

Appendix

A. Sets and Sequences

A ***set*** is an unordered collection of distinct objects. The objects are called ***elements*** of the set. We use braces to denote a set, for example, the set with elements $1, 2$ and 3 is denoted $\{1, 2, 3\}$. Since the elements are not ordered, we can rearrange the elements in the representation to get the same set, so $\{1, 2, 3\}$ and $\{3, 2, 1\}$ are the same set. The set with no elements is denoted $\{\}$ or $\emptyset$, and is called the ***empty set***. A set A is a ***subset*** of a set B, denoted $A \subset B$, if every element of A is also an element of B. We write $a \in A$ to denote that a is an element of set A. If A is a set with finitely many elements, we write $|\mathrm{A}|$ for the number of elements of the set A. Some standard infinite sets include $\mathbb{Z}$, the set of integers, $\mathbb{Z}_{\geq 0}$ the set of non-negative integers, and $\mathbb{R}$, the set of real numbers.

Recall that if A and B are sets, then $A \cap B = \{a : a \in A \text{ and } a \in B\}$, and $A \cup B = \{a : a \in A \text{ or } a \in B\}$. These are the ***intersection*** and ***union*** of the sets A and B respectively. Two sets A and B are ***disjoint*** if $A \cap B = \emptyset$. Sets $A_i : i \in S$ are ***pairwise disjoint*** if $A_i \cap A_j = \emptyset$ for all $i, j \in S$ with $i \neq j$. We write $\bigcup_{i \in S} A_i$ to denote the union of all sets A_i such that $i \in S$. If A and B are sets, then

$$A \times B = \{(a, b) : a \in A, b \in B\}$$

is the ***Cartesian product*** of A and B and $A^k = A \times A \times \cdots \times A$.

A ***sequence*** is an ordered collection of (not necessarily distinct) objects. The objects are called entries of the sequence. We use brackets to denote a sequence, for example $(1, 1, 2)$ denotes the sequence with entries $1, 1$ and 2. Since the entries are ordered, we can rearrange the elements in the representation to get a new sequence, so $(1, 1, 2)$ and $(1, 2, 1)$ are different sequences. When the entries are required to be distinct, the sequence is called a ***permutation*** of the set of its entries. For example, $(1, 2, 3)$ and $(2, 3, 1)$ are permutations of $\{1, 2, 3\}$. If a and b are sequences, then a is a subsequence of b if we can delete entries of b to get a. For example, $(1, 2, 3, 4)$ is a subsequence of $(1, 1, 2, 1, 3, 1, 4)$ obtained by deleting 1s. The ***length*** of a sequence with finitely many entries is the number of entries in the sequence.

Here is some notation involving products and sums of elements of sets and sequences: when we want to sum up the values of a function $f(i)$ for $i \in S$, where S is a set or a sequence, we write $\sum_{i \in S} f(i)$. The symbol we use for products is Π, so the product of $f(i)$ over $i \in S$ is denoted $\prod_{i \in S} f(i)$. We will be making extensive use of this notation.

B. Counting Sets and Sequences

Basic combinatorial questions involve counting sequences of finite length and sets of finite size. The following theorem tells us the total number of subsets of an n-element set:

THEOREM 1. *The number of subsets of an n-element set is 2^n.*

The next natural question is how many sequences of length n can be formed from a k-element set? For example, from the set $\{a, b\}$, we can form the sequences $(a, a), (a, b), (b, a)$ and (b, b) of length two. The answer is as follows:

THEOREM 2. *The number of sequences of length n from a k-element set is k^n.*

It should already be plain why this theorem is true: there are k choices for each entry of the sequence, and so k^n choices to fill up the sequence. By this logic, counting permutations is just as easy:

THEOREM 3. *The number of permutations of a set of size n is $n! := n(n-1)(n-2)\dots 1$. There are $(n)_k := n(n-1)\dots(n-k+1)$ sequences of k distinct elements in a set of size n.*

The notation $n!$ is read ***n factorial***, and denotes the product of all integers from 1 to n. Again, there are n choices for the first entry of a permutation, but then only $n-1$ for the next, $n-2$ for the next, and so on until the last entry, since all the entries are distinct. The last thing to count is the number of subsets of size k in an n-element set. The answer in general is given by the following theorem

THEOREM 4. *The number of sets of size k in an n-element set is*

$$\binom{n}{k} := \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}.$$

The numbers $\binom{n}{k}$ defined in this theorem are called ***binomial coefficients***, for reasons which we shall see shortly. One of the important identities is ***Pascal's Triangle Identity***: for $n \geq k \geq 1$,

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

C. Multiplication and Summation principles

All of the basic theorems in the last section have the same organizing principle, known as the ***multiplication principle***. Informally, the multiplication principle says that if we want to know how many sequences $(x_1, x_2, \dots, x_k)$ there are given that the number of choices for x_i is known, all we have to do is multiply together the number of choices (or decisions) for each x_i when $x_1, x_2, \dots, x_{i-1}$ have already been chosen.

PRINCIPLE A. (THE MULTIPLICATION PRINCIPLE) *Let X be a set. The number of sequences $(x_1, x_2, \dots, x_k) \subseteq X^k$ with a_i choices for x_i after having chosen $x_1, x_2, \dots, x_{i-1}$ for each $i = 1, 2, \dots, n$ is exactly $a_1 a_2 \dots a_n$.*

The proofs of Theorems 1, 2 and 3 come from this principle. Our argument for proving Theorem 1 uses the multiplication principle with two choices for each x_i, namely $x_i \in \{0,1\}$ for all i, in which case the number of choices is 2^n. The assignment $x_i = 1$ means i is placed in the set, and $x_i = 0$ means i is not placed in the set. So we have represented sets by binary sequences and there are 2^n binary sequences.

A second principle we use often is to break down a counting problem into a number of disjoint parts which are easier to deal with. We will refer to this as the summation principle:

Principle B. (The Summation Principle) *Let $A_1, A_2, \ldots, A_n$ be pairwise disjoint finite sets. Then*

$$\Big|\bigcup_{i=1}^{n} A_i\Big| = \sum_{i=1}^{n} |A_i|.$$

D. Inclusion-Exclusion Principle

A basic course in mathematics confirms $|A \cup B| = |A| + |B| - |A \cap B|$. This is a special instance of the ***inclusion-exclusion formula***, or ***combinatorial sieve***. Let $[n] = \{1, 2, \ldots, n\}$.

Principle C. (Inclusion-Exclusion) *Let $A_1, A_2, \ldots, A_n$ be sets of finite size. Then*

$$\Big|\bigcup_{i=1}^{n} A_i\Big| \quad = \quad \sum_{\emptyset \neq S \subseteq [n]} (-1)^{|S|+1} \ \Big|\bigcap_{i \in S} A_i\Big|.$$

Note that the sum is over all non-empty subsets of $[n]$, and when $n = 2$ it reduces to $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$. When the sets A_i are pairwise disjoint (they share no elements – meaning $A_i \cap A_j = \emptyset$ for all i, j), we get the summation principle.

E. Bijections

Let A and B be sets. A function $f : A \to B$ is called an ***injection*** (or ***one-to-one***) if whenever $x, y \in A$ are distinct, then $f(x) \neq f(y)$. The function f is a ***surjection*** (or ***onto*** B) if for every $b \in B$ there exists $x \in A$ such that $f(x) = b$. Finally, $f : A \to B$ is a ***bijection*** if f is an injection and a surjection.

Principle D. *Sets A and B have the same* ***cardinality*** *if and only if there is a bijection $f : A \to B$.*

We let $|A|$ denote the number of elements in a finite set A.

F. Mathematical Induction

Let $P(n)$ denote a logical statement for each positive integer n. Thus for each integer n, we can determine whether $P(n)$ is true or $P(n)$ is false. For example, $P(n)$ might be the statement that there is a prime larger than n, and so on. In the most basic form, the principle of mathematical induction can be stated as follows:

Principle E. (Mathematical Induction) *Let $P(n)$ be a statement for each positive integer n, and suppose that $P(1)$ is true, and $P(n) \to P(n+1)$ for each positive integer n. Then $P(n)$ is true for every positive integer.*

There are two steps in any induction. First one establishes the base case (in the terms above, one proves $P(0)$). Then, under the assumption that $P(n)$ is true, one attempts to prove that $P(n+1)$ must also be true (it is very important to get the order correct here – we are to show $P(n) \to P(n+1)$ and not $P(n+1) \to P(n)$).

In some instances, the following stronger form of induction is necessary:

Principle F. (Strong Induction) *Let $P(n)$ be a statement for each positive integer n, and suppose that $P(1)$ is true, and $P(n) \wedge P(n-1) \wedge \cdots \wedge P(1) \to P(n+1)$ for each positive integer n. Then $P(n)$ is true for every positive integer.*

G. The Pigeonhole Principle

One of the most fundamental principles in combinatorics is the ***pigeonhole principle***. It is entirely straightforward:

Principle G. (Pigeonhole Principle) *Let A and B be finite sets and $f : A \to B$ an injection. Then $|B| \geq |A|$.*

If we have a set of $n+1$ objects of n different types, then the pigeonhole principle says there exist two objects of the same type. We simply note that if A is the set of objects and $f(a)$ is the type of object $a \in A$, then $f : A \to B$ where B is the set of types and A is the set of objects. Since $|A| = n+1$ and $|B| = n$, f is not an injection, which means $f(a) = f(a')$ for some $a, a' \in A$. It is often in this form that the pigeonhole principle is used.

Notation

Bibliography

[1] Ajtai, M., Komlos, J., Simonovits, M., and Szemeredi, E.

[2] Appel, K., and Haken, W. A proof of the four color theorem. *Discrete Math. 16*, 2 (1976), 179–180.

[3] Berge, C. Sur le couplage maximum d'un graphe. *C. R. Acad. Sci. Paris 247* (1958), 258–259.

[4] Bollobás, B. *Combinatorics*. Cambridge University Press, Cambridge, 1986. Set systems, hypergraphs, families of vectors and combinatorial probability.

[5] Bondy, J. A., and Simonovits, M. Cycles of even length in graphs. *J. Combinatorial Theory Ser. B 16* (1974), 97–105.

[6] Chvátal, V. A combinatorial theorem in plane geometry. *J. Combinatorial Theory Ser. B 18* (1975), 39–41.

[7] Dirac, G. A. Some theorems on abstract graphs. *Proc. London Math. Soc. (3) 2* (1952), 69–81.

[8] Erdős, P., Ko, C., and Rado, R. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2) 12* (1961), 313–320.

[9] Erdős, P., and Rado, R. Intersection theorems for systems of sets. *J. London Math. Soc. 35* (1960), 85–90.

[10] Erdös, P. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc. 53* (1947), 292–294.

[11] Erdős, P. On some extremal problems in graph theory. *Israel J. Math. 3* (1965), 113–116.

[12] Erdős, P., and Gallai, T. On maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hungar 10* (1959), 337–356.

[13] Erdős, P., and Simonovits, M. Compactness results in extremal graph theory. *Combinatorica 2*, 3 (1982), 275–288.

[14] Erdös, P., and Szekeres, G. A combinatorial problem in geometry. *Compositio Math. 2* (1935), 463–470.

[15] Füredi, Z., and Simonovits, M. The history of degenerate (bipartite) extremal graph problems. In *Erdős centennial*, vol. 25 of *Bolyai Soc. Math. Stud.* János Bolyai Math. Soc., Budapest, 2013, pp. 169–264.

[16] Hilbert, D. Ueber die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten. *J. Reine Angew. Math. 110* (1892), 104–129.

[17] Kővari, T., Sós, V. T., and Turán, P. On a problem of K. Zarankiewicz. *Colloquium Math. 3* (1954), 50–57.

[18] Mantel, W. Problem 28. *Wiskundige Opgaven 10* (1907), 60–61.

[19] O'Bryant, K. A complete annotated bibliography of work related to sidon sequences. *Electron. J. Combin. 11*, 39 (2004).

[20] Pósa, L. Hamiltonian circuits in random graphs. *Discrete Math. 14*, 4 (1976), 359–364.

[21] Robertson, N., Sanders, D. P., Seymour, P., and Thomas, R. A new proof of the four-colour theorem. *Electron. Res. Announc. Amer. Math. Soc. 2*, 1 (1996), 17–25.

[22] Schur, I. Zur Theorie der linearen homogenen Integralgleichungen. *Math. Ann. 67*, 3 (1909), 306–339.

[23] Singer, J. A theorem in finite projective geometry and some applications to number theory. *Trans. Amer. Math. Soc. 43*, 3 (1938), 377–385.

[24] Sperner, E. Note zu der Arbeit von Herrn B. L. van der Waerden: "Ein Satz über Klasseneinteilungen von endlichen Mengen". *Abh. Math. Sem. Univ. Hamburg 5*, 1 (1927), 232.

[25] Tait, P. G. Listing's topologie. *Philosophical Magazine, 5th Series 17* (1884), 30–46.
[26] Thomason, A. G. Hamiltonian cycles and uniquely edge colourable graphs. *Ann. Discrete Math. 3* (1978), 259–268. Advances in graph theory (Cambridge Combinatorial Conf., Trinity College, Cambridge, 1977).
[27] Turán, P. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok 48* (1941), 436–452.
[28] Tutte, W. T. The factorization of linear graphs. *J. London Math. Soc. 22* (1947), 107–111.
[29] Tutte, W. T. The factors of graphs. *Canadian J. Math. 4* (1952), 314–328.
[30] Verstraëte, J. On arithmetic progressions of cycle lengths in graphs. *Combin. Probab. Comput. 9*, 4 (2000), 369–373.
[31] Zarankiewicz, K. Problem P 101. *Colloq. Math. 2*, 301 (1954), 19–30.

Index