CHAPTER 6

Möbius Inversion and Graph Colouring

6.1. Posets and Möbius Functions

August Ferdinand Möbius (1790-1868) introduced the function which bears his name in 1831 and proved the well-known inversion formula. He was an assistant to Carl Friedrich Gauss (1777-1855) and made important contributions in geometry and topology. The Möbius function is very important tool not only in combinatorics, but also in algebra and number theory.

A **poset** is a pair $(P, \leq)$ with P a set and $\leq$ a relation on P (that is, a subset of $P \times P$) satisfying

(1) $x \leq x$ for all $x \in P$ (reflexive property);
(2) $x \leq y$ and $y \leq x$ implies $x = y$ (antisymmetric property);
(3) $x \leq y$, $y \leq z$ implies $x \leq z$ (transitive property).

We call $\leq$ a **partial order** on P. If $x \leq y$ and $x \neq y$, we sometimes write $x < y$. An **interval** $[x, z]$ consists of elements of $y \in P$ satisfying $x \leq y \leq z$. A poset P is called **locally finite** if every interval is finite. We say y **covers** x if $x \leq y$ and the interval $[x, y]$ consists of only two elements, namely, x and y. The **Hasse diagram** of $(P, \leq)$ is given by representing elements of P as points in the Euclidean plane, joining x and y by a line whenever y covers x and putting y "higher" than x on the plane.

Here are some examples of posets.

(1) If S is a finite set and we consider the collection $P(S)$ of all subsets of S, partially ordered by set inclusion, is a locally finite poset.
(2) If $\mathbb{N}$ is the set of natural numbers, we can definite a partial order by divisibility. Thus, $a \leq b$ if and only if $a|b$. It is easily verified that this is a partial order.
(3) If S is a finite set we consider $\Pi(S)$, the collection of partitions of S. Given two partitions α and β we say $\alpha \leq \beta$, if every block of α is contained in a block of β. We sometimes refer to α as a **refinement** of β.

S. M. Cioabă and M. R. Murty, *A First Course in Graph Theory and Combinatorics*

(4) If $V(n,q)$ is the n-dimensional vector space over the finite field of order q, we can consider the poset of its subspaces partially ordered by inclusion.
(5) We can define a partial order on the elements of the symmetric groups S_n as follows. Let $\sigma \in S_n$. A permutation $\tau \in S_n$ is said to be a **reduction** of σ if $\tau(k) = \sigma(k)$ for all k except for $k = i, j$ where we have $\sigma(i) > \sigma(j)$ with $i < j$. We will write $\eta \leq \sigma$ if we can obtain η by a sequence of reductions from σ. This is called the **Bruhat order** on the symmetric group which makes S_n into a poset.

Given two posets $(P_1, \leq_1)$ and $(P_2, \leq_2)$, we can define their **direct product** as $(P_1 \times P_2, \leq)$, with partial order

$$(x_1, y_1) \leq (x_2, y_2) \quad \text{if} \quad x_1 \leq_1 x_2, \quad \text{and} \quad y_1 \leq_2 y_2.$$

If x and y are not comparable in P, we sometimes write $x \not\leq y$. Let F be a field and denote by $I(P)$ the set of intervals of P. The **incidence algebra** $I(P,F)$ is the F-algebra of functions

$$f : I(P) \to F$$

where we define multiplication (or convolution) by

$$(fg)(x,y) = \sum_{x \leq z \leq y} f(x,z)g(z,y).$$

Here we are writing $f(x,z)$ for $f([x,z])$. Given a locally finite poset P, its **Möbius function** μ is a map

$$\mu : P \times P \to \mathbf{Z}$$

defined recursively as follows. We set $\mu(x,y) = 0$ if $x \not\leq y$. Otherwise, we define it by the recursion

$$\sum_{x \leq z \leq y} \mu(x,z) = \delta(x,y)$$

where $\delta(x,y) = 1$ if $x = y$ and 0 otherwise. Observe that this equation can be written in "matrix form" as follows.

Define the **zeta function** of P by $\zeta(x,y) = 1$ if $x \leq y$ and zero otherwise. If for the moment, we assume P is finite, and we list our elements in some sequence $z_1, ..., z_n$ say. The matrix Z whose (i,j)-th entry is $\zeta(z_i, z_j)$ and the matrix M whose (i,j)-th entry is $\mu(z_i, z_j)$ satisfy $MZ = I$. This follows from the above recursion for μ. Thus, M is the inverse of the matrix Z. Since the inverse is both a left inverse as well as a right inverse, we deduce that $ZM = I$ which means

$$\sum_{x \leq z \leq y} \mu(z, y) = \delta(x, y).$$

THEOREM 6.1.1 (Möbius Inversion for Posets, Version 1). *Let* $(P, \leq)$ *be a locally finite poset and suppose that* $f : P \to \mathbf{R}$ *is given by*

$$f(x) = \sum_{y \leq x} g(y).$$

Then

$$g(x) = \sum_{y \leq x} \mu(y, x) f(y),$$

and conversely.

PROOF. We have that

$$\sum_{y \leq x} \mu(y, x) \sum_{z \leq y} g(z) = \sum_{z \leq x} g(z) \sum_{z \leq y \leq x} \mu(y, x) = g(x)$$

as required. The converse is left as an exercise. ∎

THEOREM 6.1.2 (Möbius Inversion for Posets, Version 2). *Let* $(P, \leq)$ *be a locally finite poset and suppose that*

$$f(x) = \sum_{y \geq x} g(y).$$

Then,

$$g(x) = \sum_{y \geq x} \mu(x, y) f(y),$$

and conversely.

PROOF. As before,

$$\sum_{y \geq x} \mu(x, y) \sum_{z \geq y} g(z) = \sum_{z \geq x} g(z) \sum_{x \leq y \leq z} \mu(x, y) = g(x),$$

as required. The converse is left as an exercise. ∎

6.2. Lattices

Given a poset $(P, \leq)$, we say z is a **lower bound** of x and y if $z \leq x$ and $z \leq y$. Any maximal element of the set of lower bounds for x and y is called a **greatest lower bound**. Such elements need not be unique as simple examples can show. The notions of **upper bound** and **least upper bound** are similarly defined. A **lattice** L is a pair $(L, \leq)$ such that $(L, \leq)$ is a poset and the greatest lower bound and least upper bound exist for any pair of elements x and y. We denote the greatest lower bound of x and y by $x \wedge y$ and least upper bound by $x \vee y$. For example, in the poset of the reals with the usual ordering, $x \wedge y$ is $\min(x, y)$ and $x \vee y$ is $\max(x, y)$. In the poset of the natural numbers partially ordered by divisibility, $x \wedge y$ is $\gcd(x, y)$, the greatest common divisor of x and y and $x \vee y$ is $\text{lcm}(x, y)$, the least common multiple of x and y. In the poset of subsets of a set S partially ordered by set inclusion, $x \wedge y$ is $x \cap y$ and $x \vee y$ is $x \cup y$.

Two posets $(P_1, \leq_1)$ and $(P_2, \leq_2)$ are said to be **isomorphic** if there is a one-to-one and onto map $f : P_1 \to P_2$ such that $x \leq_1 y$ if and only if $f(x) \leq_2 f(y)$.

Let S be a set of n elements and consider the poset $P(S)$ of subsets of S. Let $I = \{0, 1\}$ be the two element poset defined by $0 < 1$. One can show easily that $P(S)$ and I^n are isomorphic. For each subset A of S we define $f(A)$ to be the characteristic vector of A. It is then easily verified that this is the required isomorphism.

This observation allows us to compute the Möbius function of $P(S)$ very easily. Indeed, it is not hard to verify that if $(P_1, \leq_1)$ and $(P_2, \leq_2)$ are two locally finite posets, then the Möbius function of $P_1 \times P_2$ is given by

$$\mu((x_1, x_2), (y_1, y_2)) = \mu(x_1, y_1)\mu(x_2, y_2).$$

Now, the Möbius function for I is easily seen to be given by

$$\mu(x, y) = (-1)^{y-x}.$$

Thus, the Möbius function for I^n is given by

$$\mu((x_1, .., x_n), (y_1, ..., y_n)) = (-1)^{\sum_i (y_i - x_i)},$$

and using the isomorphism between $P(S)$ and I^n given above, we deduce that

$$\mu(A, B) = (-1)^{|B|-|A|}.$$

The Möbius inversion formula for sets now reads as:

THEOREM 6.2.1. *If*

$$F(A) = \sum_{A \subseteq B} G(B),$$

then

$$G(A) = \sum_{A \subseteq B} (-1)^{|B|-|A|} F(B),$$

and conversely.

We can specialize this to deduce the inclusion-exclusion principle. Indeed, suppose we have a set A with subsets A_i with $i \in I$. We would like to derive a formula for the size of

$$A \setminus \cup_i A_i.$$

For each subset J of I, we let $F(J)$ be the number of elements of A which belong to every A_i for $i \in J$. Let $G(J)$ be those elements which belong to every A_i for $i \in J$ and to no other A_i for $i \notin J$. Clearly,

$$F(K) = \sum_{J \supseteq K} G(J).$$

By Möbius inversion, we obtain

$$G(K) = \sum_{J \supseteq K} \mu(K, J) F(J).$$

What we seek is $G(\emptyset)$. Because $F(J) = |\cap_{i \in J} A_i|$, we retrieve the principle of inclusion and exclusion. Thus, the Möbius inversion formula is a vast generalization of this important principle.

6.3. The Classical Möbius Function

Let us consider the lattice $D(n)$ of divisors of a natural number n. By the unique factorization theorem, we see that if

$$n = p_1^{\alpha_1} \cdots p_k^{\alpha_k},$$

with p_i's being distinct primes, then

$$D(n) \simeq D(p_1^{\alpha_1}) \cdots D(p_k^{\alpha_k}).$$

Thus, in order to determine the Möbius function of $D(n)$, it suffices to determine for $D(p^a)$ for primes p. We observe that $\mu(p^i, p^j)$ is 1 if $i = j$, is -1 if $i = j - 1$ and 0 otherwise. By the product theorem, the Möbius function for $D(n)$ is easily computed: $\mu(a, b) = 0$ unless $a|b$ in which case it is the classical Möbius function $\mu(b/a)$ defined as follows: $\mu(n)$ is zero unless n is square-free in which case it is $(-1)^k$, where k is the number of prime factors of n. The Möbius inversion formula for the lattice of natural numbers partially ordered by divisibility is now seen as an immediate consequence of the general inversion formula.

One immediate application of the Möbius inversion formula is to count the number $\phi(n)$ of natural numbers less than n which are coprime to n. We see immediately that

$$\phi(n) = \sum_{d|n} \mu(d) n/d.$$

There are many applications of the inversion formula in counting problems. For instance, let us look at the following celebrated example. If we have an infinite supply of beads of λ colours, in how many ways can we make a necklace of n beads? Clearly, any necklace can be thought of as a sequence $(a_1, ..., a_n)$, where we identify any cyclic permutation of the sequence as giving rise to the same necklace. We will say that a necklace is **primitive** of length n if for no divisor $d < n$ it is not obtained by repeating n/d times a necklace of length d. We say a necklace has **period** d if it is obtained by repeating $\frac{n}{d}$ times a primitive necklace of length d. With these notions, we can count first the number of sequences to be λ^n. On the other hand, each sequence corresponds to some primitive necklace of period d which must necessarily divide n. If we let $M(d)$ be the number of primitive necklaces of length d, we have d places from which to start the sequence and so we obtain

$$\lambda^n = \sum_{d|n} dM(d).$$

By Möbius inversion, we get

$$M(n) = \frac{1}{n}\sum_{d|n} \mu(d)\lambda^{n/d}.$$

Now the total number of necklaces is

$$\sum_{d|n} M(d).$$

This can be simplified further. We have

$$\sum_{de=n}\sum_{ab=d} \frac{1}{d}\mu(a)\lambda^b = \sum_{abe=n} \frac{\lambda^b}{b}\frac{\mu(a)}{a} = \sum_{b|n}\frac{\lambda^b}{b}\sum_{ae=n/b}\mu(a)/a.$$

The inner sum is easily seen to be $\phi(n/b)/(n/b)$. Thus, the final formula is

$$\frac{1}{n}\sum_{b|n}\phi(n/b)\lambda^b.$$

6.4. The Lattice of Partitions

Let S be a finite set and $\Pi(S)$ the collection of its partitions. We make $\Pi(S)$ into a poset as follows. Recall that the components of a partition are called blocks (or equivalence classes). We say $\alpha \leq \beta$ if β refines the α. That is, each block of α is a union of blocks of β. For example,

$$\alpha = \{1,2\}\{3,4,5\} \leq \{1\}\{2\}\{3,5\}\{4\} = \beta.$$

It is easy to verify that this poset is a lattice with minimal element 0 given by the partition consisting of one block containing all the elements of S. The maximal element 1 is given by the partition consisting of singleton sets. Thus, the "greater" the partition, the larger the number of blocks.

We would like to determine the Möbius function of this lattice. To this end, let us define $b(\alpha)$ to be the number of blocks of the partition α. Let us fix a partition β with m blocks. If $\alpha \leq \beta$, then every block of α is a union of blocks of β and it is then clear that if we view β as a set of its blocks, then

$$[0, \beta] \simeq \Pi(\beta),$$

which will be useful in the computation of the Möbius function.

Let x be an indeterminate. For each partition α define $g(\alpha)$ to be the polynomial $(x)_{b(\alpha)}$. Then,

$$\sum_{\alpha \leq \beta} g(\alpha) = \sum_{\alpha \leq \beta} (x)_{b(\alpha)} = \sum_{k=1}^{m} S(m, k)(x)_k = x^m = x^{b(\beta)}$$

by a calculation from Chapter 2. By Möbius inversion,

$$g(\beta) = (x)_m = \sum_{\alpha \leq \beta} \mu(\alpha, \beta) x^{b(\alpha)} = \sum_{k=1}^{m} s(m, k) x^k,$$

by a calculation done (again) in the Chapter 2. Identifying the coefficients of x^k of both sides of the identity above gives

$$s(m, k) = \sum_{\alpha \leq \beta, b(\alpha) = k} \mu(\alpha, \beta).$$

Taking $k = 1$ gives

$$s(m, 1) = \mu(0, \beta).$$

Thus, the value of the Möbius function $\mu(0, \beta)$ depends only the number of blocks in β, namely $b(\beta)$. But recall that $(-1)^{m-1} s(m, 1)$ is the number of permutations of S_m with exactly one cycle in their disjoint cycle decomposition. The number of such permutations is $(m - 1)!$. Thus, we have proved that:

THEOREM 6.4.1. *For the lattice of partitions $\Pi(S)$ of an n element set, we have*

$$\mu(0, 1) = (-1)^{n-1}(n - 1)!.$$

We will now count the number of connected labeled graphs on n vertices. To this end, let us observe that any graph induces partition on the vertices given by its connected components. For each partition β of the n vertices, let $g(\beta)$ be the number of graphs whose partition of connected components is finer than β. Let $f(\beta)$ be the number of graphs whose partition of connected components is equal to β. Clearly,

$$g(\beta) = \sum_{\alpha \geq \beta} f(\alpha).$$

By Möbius inversion, we get

$$f(\beta) = \sum_{\alpha \geq \beta} \mu(\beta, \alpha) g(\alpha).$$

What we want to determine is $f(0)$. But this is

$$f(0) = \sum_{\alpha} \mu(0, \alpha) g(\alpha).$$

If $\alpha = \{B_1, ..., B_k\}$, then clearly

$$g(\alpha) = \sum_{\alpha} (-1)^{b(\alpha)-1} (b(\alpha) - 1)! 2^{\binom{|B_1|}{2}} \cdots 2^{\binom{|B_k|}{2}}.$$

6.5. Colouring Graphs

Graph colouring is one of the main topics in graph theory. We describe here some connections between this subject and Möbius inversion. More details regarding graph colouring will appear in Chapter 10 and Chapter 11.

Given a map M in the plane, let $p_M(\lambda)$ be the number of ways of colouring M properly using λ colours. We say that a colouring is **proper** if no two adjacent regions receive the same colouring. If $r(M)$ is the number of regions of the map, then the number of arbitrary colourings using λ colours is clearly $\lambda^{r(M)}$. Given any such colouring, we may "refine" it to get a proper colouring of a unique "submap" obtained by deleting the common boundary between the regions with the same colour. It is also clear that we may define a partial ordering on the set of "submaps" of M in the obvious way. Thus, we obtain

$$\lambda^{r(M)} = \sum_{A \subseteq M} p_A(\lambda).$$

By applying Möbius inversion on this poset of submaps, we obtain

$$p_M(\lambda) = \sum_{A \subseteq M} \mu(A, M) \lambda^{r(A)}.$$

This remarkable formula also shows that the number of ways of colouring the map M using only λ colours is given by a polynomial in λ of degree equal to the number of regions. This is not at all an obvious fact and yet by the theory of Möbius inversion, we were able to deduce it immediately.

The same result can be derived for colouring graphs. If X is a graph and $p_X(\lambda)$ is the number of properly colouring the vertices of X using λ colours, then we may derive a similar formula as follows. If X has $n(X)$ vertices, the number of arbitrary colourings of X using λ colours is $\lambda^{n(X)}$. Any such colouring can be refined to give a proper colouring of a subgraph obtained by contracting any two adjacent vertices that received the same colouring. The collection of subgraphs is a poset in the obvious way and thus, by Möbius inversion we see that

$$p_X(\lambda) = \sum_{A \subseteq X} \mu(A, X) \lambda^{n(A)},$$

which is again a polynomial in λ of degree equal to the number of vertices of the graph.

The **scheduling problem** is really a colouring problem. Suppose in a university we are to schedule exams so that no student has a time conflict. We construct a graph whose vertices are the courses for which we must schedule an exam. We join two vertices if the corresponding courses have a common student. The colours correspond to time slots and a proper colouring of the graph means that we assign time slots so that no student has a conflict.

Recall that given a graph X and an edge e by X/e we mean the contraction of X by e which means we create a new graph where the two vertices of e are identified.

THEOREM 6.5.1. *Let X be a simple graph and let $p_X(\lambda)$ be the number of ways of properly colouring X using λ colours. If e is an edge, then*

$$p_X(\lambda) = p_{X-e}(\lambda) - p_{X/e}(\lambda).$$

PROOF. Clearly, any proper colouring of X is also a proper colouring of $X - e$. Thus, we look at all proper colourings of $X - e$ and remove from this number those which are not proper colourings of X. This latter number corresponds to the situation where the two vertices of e get the same colour in $X - e$. But this corresponds to a proper colouring of X/e. ■

Since $X - e$ and X/e have at least one less edge than X, we see immediately by induction that $p_X(\lambda)$ is a polynomial in λ. However, a more precise theorem can be derived.

THEOREM 6.5.2. *The polynomial $p_X(\lambda)$ has degree $n = |V(X)|$ and integer coefficients alternating in sign and beginning as*

$$p_X(\lambda) = \lambda^n - |E(X)|\lambda^{n-1} + \cdots,$$

where $|E(X)|$ is the cardinality of the edge set.

PROOF. We prove this by induction on the number of edges of X. The claim holds trivially if $|E(X)| = 0$ for then $p_X(\lambda) = \lambda^n$. By induction, we may write

$$p_{X-e}(\lambda) = \lambda^n - (|E(X)| - 1)\lambda^{n-1} + a_2\lambda^{n-2} - \cdots$$

and

$$-p_{X/e}(\lambda) = \quad -\lambda^{n-1} + b_1\lambda^{n-2} - \cdots$$

where $a_2, \ldots$ and $b_1, \ldots$ are non-negative integers by the induction hypothesis. Adding these two equations gives

$$p_X(\lambda) = \lambda^n - |E(X)|\lambda^{n-1} + (a_2 + b_1)\lambda^{n-2} + \cdots$$

and the theorem is proved. ∎

Based on this result, we call $p_X(\lambda)$ the **chromatic polynomial** of X. This polynomial was introduced by George David Birkhoff (1884-1944) in 1912 as an attempt to attack the four-colour conjecture (now theorem). Showing that $p_X(4) > 0$ for any planar graph X is equivalent to the four-colour theorem.

For the complete graph K_n, the chromatic polynomial is easily seen to be

$$\lambda(\lambda - 1)(\lambda - 2) \cdots (\lambda - (n - 1)).$$

When we expand this as a polynomial in λ, we obtain

$$\sum_{k=0}^{n} s(n, k)\lambda^k$$

and the numbers $s(n, k)$ are the Stirling numbers of the first kind. From our theorem, we see that the $s(n, k)$ alternate in sign. Recall that $|s(n, k)|$ is the number of permutations of the symmetric group S_n with exactly k cycles in its unique factorization as a product of disjoint cycles.

The chromatic number $\chi(X)$ of the graph X is the smallest positive integer m so that $p_X(m) > 0$. The chromatic number of the complete graph K_n is clearly n. For the cycle graph C_n, the chromatic number is 2 or 3 according as n is even or odd. The four colour theorem is the assertion that the chromatic number of any graph obtained from a planar map is 4.

One can get a trivial bound for the chromatic number which is easily seen to be sharp in the cases of the complete graph and the odd cycles.

THEOREM 6.5.3. *Let $\Delta(X)$ denote the maximum degree of any vertex in a simple graph X. Then*

$$\chi(X) \leq 1 + \Delta(X).$$

PROOF. We use a *greedy* colouring by colouring the vertices in the order $1, 2, \ldots, n$ assigning to i the smallest-indexed colour not already used by its neighbours $j < i$.

Each vertex i will have at most Δ neighbours $j < i$ so this colouring will not use more than $\Delta + 1$ colours. ∎

As our remarks indicate, this theorem is sharp. However, a famous theorem of Brooks, proved in 1941, states that these are the only two counterexamples and if we exclude them, we have a sharper bound.

THEOREM 6.5.4 (Brooks, 1941). *If X is connected and not a complete graph or an odd cycle, then*

$$\chi(X) \leq \Delta(X).$$

The proof of this theorem is rather complicated and we will skip it here.

6.6. Colouring Trees and Cycles

Theorem 6.5.1 can be used to determine the chromatic polynomial of trees. In fact, any tree T has a leaf v (say). Let e be the unique edge containing vertex v. We have that

$$p_T(\lambda) = p_{T-e}(\lambda) - p_{T/e}(\lambda).$$

Since $T - e$ has two connected components, namely an isolated vertex and a tree with one less edge than T, we see that an inductive argument easily shows:

THEOREM 6.6.1. *Let T be a tree with n vertices. Then*

$$p_T(\lambda) = \lambda(\lambda - 1)^{n-1}.$$

PROOF. We apply induction on n and note that in the remark preceding the statement of the theorem, T/e is a tree on $n - 1$ vertices. Thus, induction gives

$$p_X(\lambda) = \lambda\{\lambda(\lambda - 1)^{n-2}\} - \lambda(\lambda - 1)^{n-2} = \lambda(\lambda - 1)^{n-1}.$$

∎

COROLLARY 6.6.2. *The chromatic number of a tree is 2.*

Theorems 6.5.1 and 6.6.1 can be used to determine the chromatic polynomial of the cycle C_n on n vertices. Deleting an edge from the cycle gives a tree on n vertices and contracting an edge gives a cycle on $n-1$ vertices. Thus, by an inductive argument we deduce:

THEOREM 6.6.3. *The chromatic polynomial of the cycle C_n is*

$$(\lambda-1)^n+(-1)^n(\lambda-1).$$

In particular, the chromatic number of C_n is 2 or 3 according as n is even or odd.

PROOF. For $n=3$, we verify the theorem directly:

$$\lambda(\lambda-1)(\lambda-2)=(\lambda-1)^3-(\lambda-1).$$

For the general case, by the remark preceding the theorem and the induction hypothesis, we get

$$\lambda(\lambda-1)^{n-1}-\{(\lambda-1)^{n-1}+(-1)^{n-1}(\lambda-1)\}$$

which is easily seen to be the stated expression. ∎

It is rather remarkable that the converse of Theorem 6.6.1 also holds. That is, if X is a graph with chromatic polynomial $p_X(\lambda)=\lambda(\lambda-1)^{n-1}$, then X is a tree. To see this, first note that if X consists of connected components $X_1, X_2, \ldots$ then the chromatic polynomial of X is just the product of the chromatic polynomials of the connected components. Secondly, any chromatic polynomial has $\lambda=0$ as a root. This can be seen in several ways. An immediate way to see it is to say that the number of ways of colouring a map using zero colours is zero. Another way is to see it is via an inductive argument from the contraction deletion Theorem 6.5.1. Thus, the order of the zero at $\lambda=0$ is at least equal to the number of connected components. Since the zero is of order 1 in our case, the graph is connected. In addition, the number of edges is $n-1$ which can be seen from computing the coefficient of the second term. Thus, X is connected and has exactly $n-1$ edges and so by Theorem 5.1.2, X is a tree. This proves:

THEOREM 6.6.4. *If X has chromatic polynomial $\lambda(\lambda-1)^{n-1}$, then X is a tree on n vertices.*

There are other classes of graphs except trees that are not isomorphic but share the same chromatic polynomial. An easy way to construct such graphs is by using the following theorem.

THEOREM 6.6.5. *Let X and Y be two graphs whose intersection is a complete graph K_r. Then*

$$p_{X \cup Y}(\lambda) = \frac{p_X(\lambda) \cdot p_Y(\lambda)}{\lambda(\lambda-1)\ldots(\lambda-r+1)}.$$

We leave the proof of this theorem as an exercise.

6.7. Sharper Bounds for the Chromatic Number

We will now connect eigenvalues of the adjacency matrix of a graph with its chromatic number. As preparation to this end, we will review the notion of **Rayleigh-Ritz quotient** or **ratio** from linear algebra.

Let A be a real symmetric matrix. If $x = (x_1, \ldots, x_n)^t$ and $y = (y_1, \ldots, y_n)^t$ are two n by 1 column vectors, then the inner product (x, y) is defined as $x_1y_1 + \cdots + x_ny_n$. For any non-zero column vector v, we call $(Av, v)/(v, v)$ the **Rayleigh-Ritz quotient** of v and denote it by $R(A, v)$. Denote by λ_{max} and λ_{min} the largest and smallest eigenvalues of A respectively. Then

$$\lambda_{\text{max}} = \max_{v \neq 0} \frac{(Av, v)}{(v, v)}$$

and

$$\lambda_{\text{min}} = \min_{v \neq 0} \frac{(Av, v)}{(v, v)}.$$

To see this, observe that if U denotes the matrix whose columns form an orthonormal basis of eigenvectors of A, then we may write

$$A = UDU^t,$$

where D is a diagonal matrix whose diagonal entries are the eigenvalues of A. Thus,

$$(Av, v) = v^tAv = v^tUDU^tv = \sum_i \lambda_i |(U^tv)_i|^2.$$

As each of the terms $|(U^tv)_i|^2$ is non-negative,

$$\lambda_{\text{min}} \sum_i |(U^tv)_i|^2 \leq v^tAv \leq \lambda_{\text{max}} \sum_i |(U^tv)_i|^2.$$

Since U is an orthogonal matrix, we have

$$\sum_i |(U^tv)_i|^2 = \sum_i |v_i|^2 = v^tv.$$

Thus, if $v \neq 0$,

$$\lambda_{\text{min}} \leq \frac{(Av, v)}{(v, v)} \leq \lambda_{\text{max}}.$$

The inequalities are easily seen to be sharp by considering the eigenvectors corresponding to $\lambda_{\max}$ and $\lambda_{\min}$ respectively, which proves our assertion. This result is usually referred to as the Rayleigh-Ritz theorem in the literature.

If X is a graph, let us denote $\lambda_{\max}(X)$ and $\lambda_{\min}(X)$ to be the largest and smallest eigenvalues of the adjacency matrix A_X of X. We also say that X' is a **subgraph** of X if $V(X') \subseteq V(X)$ and $E(X') \subseteq E(X)$. We begin by proving:

THEOREM 6.7.1. *If X' is a subgraph of X, then*

$$\lambda_{\max}(X') \leq \lambda_{\max}(X); \quad \lambda_{\min}(X') \geq \lambda_{\min}(X).$$

If $\Delta(X)$ and $\delta(X)$ denotes the maximal and minimal degrees of X, then

$$\delta(X) \leq \lambda_{\max}(X) \leq \Delta(X).$$

PROOF. The first part of the theorem is proved as follows. By relabeling the vertices, we may assume that the adjacency matrix A of X has a leading principal submatrix A_0 which is the adjacency matrix of X'. Let z_0 be chosen so that $A_0z_0 = \lambda_{\max}(A_0)z_0$ and $(z_0, z_0) = 1$. Let z be the column vector with $|V(X)|$ rows formed by adjoining zero to entries of z_0. Then,

$$\lambda_{\max}(A_0) = R(A_0, z_0) = R(A, z) \leq \lambda_{\max}(A).$$

Thus, $\lambda_{\max}(A_0) \leq \lambda_{\max}(A)$. The other inequality is proved in a similar way. For the second part, let u be a column vector each of whose entries is 1. Then, if $n = |V(X)|$ and d_i is the degree of vertex v_i, we have

$$R(A, u) = \frac{1}{n}\sum_{i,j} a_{ij} = \frac{1}{n}\sum_{i} d_i \geq \delta(X).$$

But the Rayleigh quotient $R(A, u)$ is at most $\lambda_{\max}(A)$ and so

$$\lambda_{\max}(X) \geq \delta(X).$$

For the other inequality, let v be an eigenvector corresponding to the eigenvalue $\lambda_0 = \lambda_{\max}(X)$. Let x_j be the largest positive entry of v. Then,

$$\lambda_0 x_j = (\lambda_0 v)_j = \sum_{i}^{*} x_i \leq \Delta(X)x_j$$

where the $*$ on the summation means we sum over the vertices adjacent to v_j. This proves the theorem. ∎

We will now relate the chromatic number to the largest eigenvalue of the adjacency matrix of X. To this end, we say a graph is ***t*-critical** if $\chi(X) = t$ and for all proper vertex subgraphs U of X, we have $\chi(U) < t$.

LEMMA 6.7.2. *Suppose X has chromatic number $t \geq 2$. Then X has a t-critical subgraph U such that every vertex of U has degree at least $t-1$ in U.*

PROOF. The set of all vertex subgraphs of X is non-empty and contains some graphs (for instance, X itself) that have chromatic number t. Let U be a vertex subgraph of X whose chromatic number is t which is minimal with respect to the number of vertices. Clearly, U is t-critical. Moreover, if $v \in V(U)$, then the vertex subgraph whose vertex set is $V(U)\backslash v$ is a vertex subgraph of U and has a vertex colouring with $t-1$ colours. If the valency of v in U were less than $t-1$, then, we could have extended this vertex colouring to U contradicting $\chi(U) = t$. ∎

The previous lemma has the following important consequences.

THEOREM 6.7.3 (Szekeres-Wilf 1968). *If X is a graph, then*

$$\chi(X) \leq 1 + \max_{Y \subseteq X} \delta(Y).$$

PROOF. By Lemma 6.7.2, there is a vertex subgraph U of X whose chromatic number is $\chi(X)$ and $\delta(U) \geq \chi(X) - 1$. Thus, we have

$$\chi(X) \leq 1 + \delta(U) \leq 1 + \max_{Y \subseteq X} \delta(Y).$$

∎

By a slight modification of the previous proof, we also get the following result.

THEOREM 6.7.4 (Wilf, 1967). *For any graph X, we have*

$$\chi(X) \leq 1 + \lambda_{\max}(X).$$

PROOF. As before, there is a vertex subgraph U of X whose chromatic number is $\chi(X)$ and $\delta(U) \geq \chi(X) - 1$. Thus, by Theorem 6.7.1, we have

$$\chi(X) \leq 1 + \delta(U) \leq 1 + \lambda_{\max}(U) \leq 1 + \lambda_{\max}(X),$$

as desired. ∎

6.8. Sudoku Puzzles and Chromatic Polynomials

The Sudoku puzzle has become a very popular puzzle that many newspapers carry as a daily feature. The puzzle consists of a 9×9 square grid in which some of the entries of the grid have a number from 1 to 9. One is then required to complete the grid in such a way that every row, every column, and every one of the nine 3×3 sub-grids contain the digits from 1 to 9 exactly once. The sub-grids are shown below.

For anyone trying to solve a Sudoku puzzle, several questions arise naturally. For a given puzzle, does a solution exist ? If the solution exists, is it unique ? If it is not unique, how many solutions are there ? Moreover, is there a systematic way of determining all the solutions ? How many puzzles are there with a unique solution ? What is the minimum number of entries that can be specified in a single puzzle to ensure a unique solution ? For instance, the next figure shows that the minimum is at most 17. We leave it to the reader to show that the puzzle below has a unique solution. It is unknown if a puzzle with 16 specified entries exists that yields a unique solution.

							1	
4								
	2							
				5		4		7
		8				3		
		1		9				
3			4			2		
	5		1					
			8		6			

We reinterpret the Sudoku puzzle as a vertex colouring problem in graph theory. We associate a graph with the 9×9 Sudoku grid as follows. The graph will have 81 vertices with each vertex corresponding to a cell in a grid. Two distinct vertices will be adjacent if and only if the corresponding cells in the grid are either in the same row, or same column, or the same sub-grid. Each completed Sudoku square then corresponds to a proper colouring of this graph. We put this problem in a more general and formal context. Consider an $n^2 \times n^2$ grid. To each cell in a grid, we associate a vertex labeled (i, j) with $0 \leq i, j \leq n^2 - 1$. We will say that (i, j) and (i', j') are adjacent if $i = i'$ or $j = j'$ or

$\lfloor \frac{i}{n} \rfloor = \lfloor \frac{i'}{n} \rfloor$ and $\lfloor \frac{j}{n} \rfloor = \lfloor \frac{j'}{n} \rfloor$. Recall that $\lfloor x \rfloor$ is the largest integer less than or equal to a. We will denote this graph by X_n and call it the Sudoku graph of rank n. An easy computation shows that X_n is a regular graph having degree $3n^2 - 2n - 1$. In the case $n = 3$, X_3 is 20-regular and in case $n = 2$, X_2 is 7-regular.

A Sudoku square of rank n will be a proper coloring of this graph using n^2 colours.

THEOREM 6.8.1. *For every natural number n, the chromatic number of the Sudoku square X_n is n^2.*

PROOF. It is easy to see that we need at least n^2 colours because the n^2 vertices of the same row or column create a complete subgraph of order n^2. For $0 \le i \le n^2 - 1$, write $i = t_i n + d_i$, where $0 \le t_i \le n - 1$ and $0 \le d_i \le n - 1$. Colour the vertex corresponding to the cell (i, j) of the Sudoku square by the colour $d_i n + t_i + n t_j + d_j$, reduced modulo n^2. We leave it as an exercise for the reader to show that this is a proper colouring of X_n with n^2 colours. ∎

A Sudoku puzzle corresponds to a partial colouring of X_n and the question is whether this partial colouring can be completed to a total proper colouring of the Sudoku graph X_n with n^2 colours. Given a partial proper colouring C of a graph G), one can show that the number of ways of completing this colouring to obtain a proper colouring with λ colours, is a polynomial in λ, provided that λ is greater than or equal to the number of colours used in C. We leave this as an exercise.

It is not obvious at the outset if a given puzzle has a solution. Also, it is always clear whether or not a puzzle has a unique solution. An obvious necessary condition to have a unique solution is that the partial Sudoku square must contain at least 8 distinct numbers from $\{1, \dots, 9\}$. This is not sufficient as the square below has exactly two solutions. The proof of this fact is left as an exercise.

9		6		7		4		3
			4			2		
	7			2	3		1	
5						1		
	4		2		8		6	
		3						5
	3		7				5	
		7			5			
4		5		1		7		8

6.9. Exercises

EXERCISE 6.9.1. Show that the five examples from the first are actually posets.

EXERCISE 6.9.2. Draw the Hasse diagram for S_3 with the Bruhat order and determine completely the Möbius function of this poset.

EXERCISE 6.9.3. If

$$G(x) = \sum_{n \le x} F(x/n)$$

prove that

$$F(x) = \sum_{n \le x} \mu(n) G(x/n).$$

EXERCISE 6.9.4. Show that

$$\sum_{n \le x} \mu(n)[x/n] = 1$$

where $[x]$ denotes the greatest integer less than x.

EXERCISE 6.9.5. Let $(P_1, \le_1)$ and $(P_2, \le_2)$ be two locally finite posets. Show that

$$\mu((x_1, y_1), (x_2, y_2)) = \mu(x_1, y_1)\mu(x_2, y_2).$$

EXERCISE 6.9.6. Let $(P, \le)$ be a finite poset. For $a \in P$, we will denote by $\downarrow a$ the set $\{x \in P : x \le a\}$ and $\uparrow a$ the set $\{x \in P : a \le x\}$. We say that P is **linearly ordered** if any two elements of P are comparable. Show that any partial ordering of P can be extended to a linear ordering as follows. View the poset $(P, \le)$ as a subset R of $P \times P$ satisfying the axioms: (1) $(a, a) \in R$, (2) $(a, b) \in R$ and $(b, a) \in R$ implies $a = b$ and (3) $(a, b) \in R$, $(b, c) \in R$ implies $(a, c) \in R$. A linear order can be regarded as a subset R' of $P \times P$ which has the additional property that for any $a, b \in P$ either $(a, b) \in R'$ or $(b, a) \in R'$. Let now a, b be incomparable in $(P, \le)$. Put $R' = R \cup (\downarrow a \times \uparrow b)$. Verify that R' is a partial order of P in which $(a, b) \in R'$. Deduce that any partial ordering of P can be extended to a linear ordering.

EXERCISE 6.9.7. Six different television stations are applying for channel frequencies and no two stations can use the same frequency if they are within 150 miles of each other. If the distances between the stations A, B, C, D, E and F are given by the matrix below, find the minimal number of frequencies needed.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{pmatrix} - & 85 & 175 & 200 & 50 & 100 \\ 85 & - & 125 & 175 & 100 & 160 \\ 175 & 125 & - & 100 & 200 & 250 \\ 200 & 175 & 100 & - & 210 & 220 \\ 50 & 100 & 200 & 210 & - & 100 \\ 100 & 160 & 250 & 220 & 100 & - \end{pmatrix} \end{array}.$$

EXERCISE 6.9.8. Prove that the sum of the coefficients of the chromatic polynomial of a graph X is zero unless X has no edges. Show that the coefficients of $p_X(\lambda)$ alternate in sign.

EXERCISE 6.9.9. If $X_1, \ldots, X_t$ are the components of X, then

$$p_X(\lambda) = \prod_{i=1}^{t} p_{X_i}(\lambda).$$

If $p_X(\lambda)$ is the chromatic polynomial of a graph X, show that we can write it as $\lambda^c f(\lambda)$ where $f(0) \neq 0$ and c is the number of connected components of X.

EXERCISE 6.9.10. The **join** of two graphs X and Y is defined as the graph obtained by joining every vertex of X to every vertex of Y. We denote this graph by $X \vee Y$. Show that $\chi(X \vee Y) = \chi(X) + \chi(Y)$.

EXERCISE 6.9.11. The wheel graph is $K_1 \vee C_n$. That is, the wheel graph is the cycle graph together with a vertex at the 'center' which is connected to all the vertices of C_n. Determine the chromatic polynomial of the wheel graph. Show also that

$$p_{K_{2,n}}(\lambda) = \lambda(\lambda - 1)(\lambda - 2)^n + \lambda(\lambda - 1)^n.$$

EXERCISE 6.9.12. Let $p_X(\lambda)$ be the chromatic polynomial of a connected graph X of order n. Show that

$$|p_X(\lambda)| \leq \lambda(\lambda - 1)^{n-1}$$

if $n \geq 3$.

EXERCISE 6.9.13. Compute the chromatic polynomial of the graph in Figure 6.1.

EXERCISE 6.9.14. Prove Theorem 6.6.5.

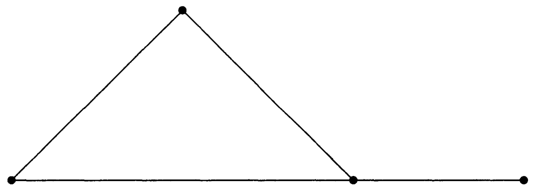

FIGURE 6.1

EXERCISE 6.9.15. Let X be a graph with n vertices, e edges and maximum degree Δ. Show that

$$\lambda_{max} \geq \max\left(\frac{2e}{n}, \sqrt{\Delta}\right).$$

When does equality occur ?

EXERCISE 6.9.16. Let G be a graph with n vertices and let C be a partial proper colouring of t vertices of G using k colours. If $p_{G,C}(\lambda)$ denotes the number of ways of completing this colouring using λ colours to a proper colouring of G, then prove that $p_{G,C}(\lambda)$ is a polynomial in λ with integer coefficients of degree $n - t$ for $\lambda \geq k$.

EXERCISE 6.9.17. Show that the chromatic number of a graph X satisfies

$$\chi(X) \leq \frac{1 + \sqrt{8e+1}}{2}.$$

EXERCISE 6.9.18. Let G_n be the graph whose vertex set is $[2n] = \{1, 2, \ldots, 2n\}$ and where (i, j) is an edge if and only if i and j have a common prime divisor. Show that the chromatic number of X_n is at least n.

EXERCISE 6.9.19. The **Kneser graph** $K(n, k)$ is the graph whose vertices are all the k-element subsets of $[n]$. Two k-subsets are adjacent in $K(n, k)$ if and only if they are disjoint. Show that the Petersen graph (Figure 10.2) is isomorphic to $K(5, 2)$ and that $\chi(K(n, k)) \leq n - 2k + 2$. The chromatic number of $K(n, k)$ actually equals $n - 2k + 2$ as proved by László Lovász in 1978, but this is a more difficult result.

EXERCISE 6.9.20. Let $c(X)$ denote the number of components of the graph X and for $F \subseteq E(X)$, denote by $X[F]$ the spanning subgraph of X with edge set F. Show that

$$p_X(\lambda) = \sum_{F \subseteq E(X)} (-1)^{|F|} \lambda^{c(X[F])}.$$