CHRIS GODSIL

“SEMINARS”

Version: 29/05/2023

Preface

Before covid this would have been a set of notes to serve as basis for seminar presentations, now they are just a set of notes. A very rough and incomplete set of notes.

Contents

To Do

1. Add bibliography, citations.
2. Check Theorem 5.5.1.
3. ch10: finish construction of Stockmeyer's tournments
4. ch1: Scattering: examples, eigenvectors; look into work by chemists (Fowler)
5. ch3.4: complete infinite paths
6. ch12: circulants of prime order determined by spectrum; examples of cospectral dihedral Cayley graphs
7. Prove Schwenk's limb replacement (currently Chapter 40)
8. ch26.4: Finish singular values, give Knuth's application to generalized tensor product
9. tidy up quadratic rank and CdV (eliminate repetition)
10. finish spectral centre of a tree

Now some tentative plans:

1. 600-cell: polytopes, quaternions, colouring.
2. Hales and Straus: 3-colouring projective planes, Kochen-Specker, Gleason.
3. Uniqueness of regular two-graph on 276 vertices.
4. Add proof of Gabriel and Henry that there is no pst on Laplacians. Or no fractional revival?
5. Wang's work on characterization of graphs by spectrum.
6. Quantum error correction—paper by Kribs et al.
7. Algebraic number theory: discriminants, integer equivalence, similarity mod p,...

8. Unitary geometry over finite fields.
9. Ku-Chen version of Gallai’s lemma.
10. add Haar wavelets to chapter on unitary weightings.

Recent Changes

1. Started chapter on spectral centre of a tree, (currently Chapter 46)
2. Added chapter on graph algebras, (currently Chapter 31)

27-7-21 Added Chapter 12 on walk-equivalent vertices.

04-08-21 Revised chapter on oriented graphs (currently Chapter 6).

28-09-21 Updated “Walking on the edges”.

30-09-21 Added Section 16.5 on unimodular tournaments.

Part I

Eigenvalues

1
Eigenvalues, Really

An algebraic integer is *totally real* if it and all its algebraic conjugates are real. The most relevant example is any graph eigenvalue.

Any algebraic integer α has a unique minimal polynomial, the monic polynomial ψ of least degree with α as a root. The zeros of ψ are precisely the algebraic conjugates of α, and so α is totally real if and only if the zeros of ψ are real. Thus $\sqrt{2}$ is totally real but $\sqrt[3]{2}$ is not.

An algebraic integer is *totally positive* if it and all its conjugates are real and positive. An algebraic integer is totally positive if and only if it is the square of a totally real algebraic integer.

Alan Hoffman (1973) asked whether every totally real algebraic integer is the eigenvalue of some graph. Bass, Estes and Guralnick (1994) used non-trivial number theory to show that this was the case. Here we will present a proof from 2013 due to Justin Salez,[1], showing that any totally real algebraic integer is an eigenvalue of a tree.(Hmm, the next significant work on this problem may not appear until 2033.)

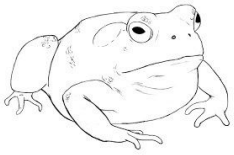

Figure 1.1: A toad, really, I'm positive

[1] "Every totally real algebraic integer is a tree eigenvalue", https://arxiv.org/abs/1302.4423

1.1 Totally Real

For our purposes we can define an algebraic integer to be an eigenvalue of an integer matrix.

We view $\mathbb{C}$ as a vector space over $\mathbb{Q}$. If $\alpha \in \mathbb{C}$, let M_α denote the linear map given by multiplication by α, thus if $x \in \mathbb{C}$, then

$$M_\alpha(x) := \alpha x.$$

Consider the vector space spanned by the non-negative integer powers of α. This vector space has finite dimension if and only if α satisfies a polynomial with rational coefficients or, equivalently, a polynomial with integer coefficients. Thus the vector space is finite dimensional if and only if α is an *algebraic number*. (It is an algebraic integer if the $\mathbb{Z}$**-module**) generated by the powers of α has finite dimension.)

If α is an algebraic integer, its algebraic conjugates are the eigenvalues of M_α and so α is totally real if and only if all eigenvalues of M_α are real.

If α is an algebraic integer of degree k, then the matrix representing M_α relative to the basis

$$1,\ \alpha,\ldots,\alpha^{k-1}$$

is an integer matrix of the form

$$\begin{pmatrix} 0 & b \\ I_{d-1} & a \end{pmatrix},$$

such matrices are known as *companion matrices*. (This shows that every algebraic integer, defined in the traditional way, is an eigenvalue of an integer matrix.)

1.1.1 Lemma. *The algebraic integers form a ring; the eigenvalues of graphs form a subring.*

Proof. The eigenvalues of

$$A \otimes I + I \otimes B$$

are the sum of the eigenvalues of A and B, the eigenvalues of

$$A \otimes B$$

are the products of the eigenvalues of A and B, and the eigenvalues of

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes A$$

are $\pm\theta$, where θ runs over the eigenvalues of A. This all shows that the set of algebraic integers is closed under sum, product and multiplication by -1.

Since these three operations take adjacency matrices of graphs to adjacency matrices of graphs, the second claim follows. □

1.1.2 Lemma. *If θ is an eigenvalue of a symmetric integer matrix, it is an eigenvalue of a graph.*

Proof. If B is not non-negative, replace each entry $B_{i,j}$ by 2×2 matrices

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_{i,j}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad -B_{i,j}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

according as $B_{i,j}$ is zero, positive or negative. This gives a non-negative matrix $\tilde{B}$ and $\phi(B,t)$ divides $\phi(\tilde{B},t)$. So we may assume B is non-negative.

Assume m is a non-negative integer matrix and that $|B_{i,j}| \le m$ for some positive integer m. We replace each entry of B by a circulant matrix of order $(m+1)\times(m+1)$: if $B_{i,j} = 0$ we use the zero matrix, if $B_{i,j} = m$ we use a 01-circulant matrix of order $(m+1)\times(m+1)$ with row sum m and diagonal entries zero. This provides us with a 01-matrix whose characteristic polynomial is divisible by $\phi(B,t)$.

If B is a 01-matrix with some diagonal entries equal to 1, it is the adjacency matrix of a graph X with loops, and now $X \times K_2$ is a simple graph with its characteristic polynomial divisible by $\phi(X,t)$. □

The algebraic numbers form a field. We see that β is algebraic if and only if it is an eigenvalue of a rational matrix, and thus if β is algebraic, there is an integer m such that $m\beta$ is an algebraic integer.

1.2 Tree Eigenvalues

We present Salez's result. First, a standard recurrence for the characteristic polynomial of a tree.

1.2.1 Theorem. *Assume T is a tree with a vertex a, let $b_1, \dots, b_k$ be the neighbours of a and let $T_1, \dots, T_k$ be the connected components of $T \setminus a$ (ordered so that $b_i \in V(T_i)$). Then*

$$\frac{\phi(T,t)}{\phi(T\setminus a,t)} = t - \sum_{b_i \sim a} \frac{\phi(T_i \setminus b_i, t)}{\phi(T_i,t)}.$$

□

Let $C_a(X,t)$ denote the generating function for the closed walks in X that start at a and let $C_a^{(1)}(X,t)$ be the generating function for the closed walks in X that start at a and return to a exactly once. (So the constant and linear terms of $C_a^1(X,t)$ are zero, and the coefficient of the quadratic term is the valency of a.)

1.2.2 Lemma. *If $a \in V(X)$ then*

(a) $C_a^{(1)}(X,t) = 1 - C_a(X,t)^{-1}$ *and* $C_a(X,t) = (1 - C_a^{(1)}(X,t))^{-1}$.

(b) We have

$$C_a(X,t) = \frac{t^{-1}\phi(X\setminus a, t^{-1})}{\phi(X,t^{-1})}.$$

(c) $C_a^{(1)}(X,\lambda^{-1}) = 1$ *if and only if* $\phi(X,\lambda) = 0$.

1.2.3 Theorem. *Let $\mathscr{C}(t)$ denote the set of all rational functions of the form $C_a^{(1)}(T,t)$, for rooted trees (T,a). Then $\mathscr{C}(t)$ is the smallest set of rational functions that*

(a) contains 0,

(b) is closed under addition,

(c) is closed under the map

$$f \mapsto \frac{t^2}{1-f}.$$

Proof. If $T = K_1$, then $C_a(t) = 1$ and $C_a^{(1)}(T,t) = 0$, so) $\in \mathscr{C}(t)$.

Let $\mathscr{C}(t)$ be as defined. If (T,a) is the 1-sum of (S_1,a) and (S_2,a), then

$$C_a^{(1)}(T,t) = C_a^{(1)}(S_1,t) + C_a^{(1)}(S_2,t).$$

Working carefully, we can translate Theorem 1.2.1 into

$$1 - \frac{\phi(T,t^{-1})}{t^{-1}\phi(T\setminus a,t^{-1})} = t^2 \sum_{b_i\sim a} \frac{t^{-1}\phi(T_i\setminus b_i, t^{-1})}{\phi(T_i,t^{-1})}.$$

and then

$$C_a^{(1)}(T,t) = \sum_{b_i \sim a} \frac{t^2}{1 - C_{b_i}^{(1)}(T_i,t)}.$$

This shows that $\mathscr{C}$ has the closure properties we claim.

A routine induction argument[2] on the height of a rooted tree yields that $\mathscr{C}(t)$ contains $C_a^{(1)}(T,t)$ for each rooted tree (T,a). □

[2] We follow Salez and skimp on details

As defined, $\mathscr{C}(t)$ is a set of rational functions, one for each rooted tree (T,a). We have

$$C_a^{(1)}(X,\lambda^{-1}) = 1 - \frac{\phi(X,\lambda)}{\lambda\phi(X\setminus a,\lambda)}$$

and so $\mathscr{C}(\lambda^{-1})$ is the set of values taken by this collection of rational functions on λ^{-1}. This set contains 0, is closed under addition, and closed under the map:

$$\alpha \mapsto \frac{1}{\lambda^2(1-\alpha)}.$$

Note that λ is an eigenvalue of some tree X if and only if 1 lies in $\mathscr{C}(\lambda^{-1})$. We now assume λ is a nonzero totally real algebraic integer, and prove that $1 \in \mathscr{C}(\lambda^{-1})$. The proof comes in three parts.

1.2.1 $\mathscr{C}(\lambda^{-1})$ contains a positive integer

We start with a preliminary technical lemma.

1.2.4 Lemma. *Let ζ be an algebraic integer of degree n, let ψ be the minimal polynomial of ζ and let $q_1 < \cdots < q_n$ be rational. Then there are integers $w_1,\ldots,w_n$ such that*

(a) $\sum_r \frac{w_r}{\zeta - q_r}$ is a positive integer.

(b) w_r and $(-1)^r\psi(q_r)$ have the same sign for all r.

Proof. Define polynomials $\ell_r(t)$ by

$$\ell_r(t) := \prod_{i:i\neq r}(t-q_i)$$

and define

$$L(t) := \prod_{r=1}^{n}(t-q_r)$$

Since $\ell_r(q_s) \neq 0$ if and only if $r = s$, the polynomials $\ell_r(t)$ form a basis for the vector space of rational polynomials of degree less than n.

As $L(t)$ and $\psi(t)$ are monic of degree n, we see that $\deg(L(t) - \psi(t)) < n$ and therefore there are rationals $w_1,\ldots,w_n$ such that

$$L(t) - \psi(t) = \sum_r w_r \ell_r(t). \tag{1.2.1}$$

From this it follows that

$$1 = \sum_r \frac{w_r}{\zeta - q_r}$$

The first claim in the statement of the lemma follows from this.

For the second claim, if we substitute $t = q_r$ in Equation (1.2.1), we get

$$-\psi(q_r) = w_r \ell_r(p_r).$$

The sign of $\ell_r(q_r)$ is $(-1)^{r-1}$ and therefore the second claim holds. □

Define a map μ on $\mathscr{C}(\lambda^{-1})$ by

$$\mu(f) := \frac{1}{\lambda^2(1-f)}$$

We apply Lemma 1.2.4 with $\zeta = \lambda^2$ and $q_i = i$. We claim that $(\lambda^2 - r)^{-1} \in \mathscr{C}(\lambda^{-1})$ for any positive integer k. First, $\mu(0) = \lambda^{-2}$, and so $k/\lambda^2 \in \mathscr{C}(\lambda^{-1})$ for any positive integer k. Finally

$$\mu(k/\lambda^2) = \frac{1}{\lambda^2\left(1 - \frac{k}{\lambda^2}\right)}$$

It follows that

$$\sum_r \frac{w_r}{\lambda^2 - r}$$

is the difference of two elements of $\mathscr{C}(\lambda^{-1})$. This implies that we have $\alpha \in \mathscr{C}(\lambda^{-1})$ and a positive integer d such that $\alpha - d \in \mathscr{C}(\lambda^{-1})$.

We claim that there β in $\mathscr{C}(\lambda^{-1})$ such that $\eta = \lambda^2(1-\alpha-\beta)$ is a totally positive algebraic number. If $\beta \in \mathscr{C}(\lambda^{-1})$, then

$$(\alpha+\beta), \quad (\alpha+\beta) - d$$

both lie in $\mathscr{C}(\lambda^{-1})$. If the integer m is large enough, $-1(\lambda^2 - m)$ is totally positive and so if the integer k is large enough, then $1-\alpha-\beta$ is totally positive and therefore $\lambda^2(1-\alpha-\beta)$ is totally positive. Reset $\alpha+\beta$ to α. Then

$$j(d-1)\alpha + j(\alpha - d) + i\lambda^{-2} \in \mathscr{C}(\lambda^{-1})$$

and

$$\begin{aligned}\mu((d-1)\alpha + j(\alpha-d) + i\lambda^{-2}) &= \frac{1}{\lambda^2((jd+1)(1-\alpha) - i\lambda^{-2})} \\ &= \frac{1}{(jd+1)\lambda^2(1-\alpha) - i},\end{aligned}$$

implying that if i, j are positive integers, then

$$\frac{1}{\lambda^2(1-\alpha) - \frac{i}{jd+1}} \in \mathscr{C}(\lambda^{-1}).$$

We apply Lemma 1.2.4 for a second time. Let $\psi(t)$ be the minimal polynomial of $\lambda^2(1-\alpha)$ and choose rationals $q_1, \ldots, q_n$ interleaving the zeros of ψ such that $\psi(q_r) = (-1)^r$. By Lemma 1.2.4, there are non-negative integers $m_1, \ldots, m_n$ such that

$$\sum_r \frac{m_r}{t - q_r}$$

is a positive integer. This proves our claim.

1.2.2 $\mathscr{C}(\lambda^{-1}) = -\mathscr{C}(\lambda^{-1})$

Set $\mathscr{C} = \mathscr{C}(\lambda^{-1})$. Then $\mathscr{C}$ contains zero and is closed under addition, hence so is $\mathscr{C} \cap (-\mathscr{C})$. We show that this intersection is also closed under μ, which implies that it equals $\mathscr{C}$. Let n be a positive integer in $\mathscr{C}$. If $\alpha \in \mathscr{C} \cap (-\mathscr{C})$, then $-\alpha \in \mathscr{C}$, hence

$$-(n-1)\alpha + n \in \mathscr{C}.$$

Therefore

$$\mu(-(n-1)\alpha + n) = \frac{-1}{\lambda^2(1-\alpha)(n-1)}$$

lies in $\mathscr{C}$, and consequently $-1/\lambda^2(1-\alpha) \in \mathscr{C}$.[3] This implies that $\mathscr{C} \cap (-\mathscr{C})$ is μ-closed.

[3] if $n = 1$, then λ is a tree eigenvalue (and we're done), or just use $2n$ in place of n

1.2.3 $\mathscr{C}(\lambda^{-1})$ *is a ring*

We start with the set $\mathscr{R}(\lambda^2)$ of numbers $p(\lambda^2)/q(\lambda^2)$ where

(a) $p, q \in \mathbb{Z}[t]$,

(b) q is monic,

(c) $\deg(p) < \deg(q)$.

It is immediate that $\mathscr{R}(\lambda^2)$ is a ring.

We claim that $\mathscr{R}(\lambda^2) = \mathscr{C}(\lambda^{-1})$. Since $\mathscr{R}(\lambda^2)$ contains zero, is closed under addition and is μ-closed, we see that $\mathscr{C}(\lambda^{-1}) \subseteq \mathscr{R}(\lambda^2)$.

Assume by induction on $\deg(q)$ that if $p(\lambda^2)/q(\lambda^2) \in \mathscr{R}(\lambda^2)$ and $\deg(q) \le n$, then $p(\lambda^2)/q(\lambda^2) \in \mathscr{C}(\lambda^{-1})$. Assume

$$q(t) = t^{n+1} + q_1 t^n + \cdots + q_{n+1}; \quad p(t) = p_0 t^n + \cdots + p_n.$$

We need to show that $p(\lambda^2)/q(\lambda^2) \in \mathscr{C}(\lambda^{-1})$. We treat two special cases first.

Suppose $p(t) = t^n$. By induction $(1+\lambda^{2m})^{-1} \in \mathscr{C}(\lambda^{-1})$. We have

$$\frac{1}{\lambda^{2n+2}} = \frac{1}{1 - \frac{1}{1+\lambda^{2n}}} - \frac{1}{\lambda^2}$$

and therefore $\lambda^{-(2n+2)} \in \mathscr{C}(\lambda^{-1})$. By induction, $\lambda^{-2i} \in \mathscr{C}(\lambda^{-1})$ for $i = 1, \ldots, n$, from which it follows that

$$\frac{\lambda^{2n}}{q(\lambda^2)} = \frac{1}{\lambda^2\left(1 + \frac{q_1}{\lambda^2} + \cdots + \frac{q_{n+1}}{\lambda^{2n+2}}\right)}$$

lies in $\mathscr{C}(\lambda^{-1})$.

Suppose $p(t)$ is monic of degree n and $p(0) = 1$. Then

$$r(t) := p(t) - \frac{q(t) - q(0)p(t)}{t}$$

lies in $\mathbb{Z}[t]$ and has degree at most $n-1$. Therefore

$$\frac{p(\lambda^2)}{q(\lambda^2)} = \frac{1}{\lambda^2\left(1 - \frac{r(\lambda^2)}{p(\lambda^2)} + \frac{q(0)}{\lambda^2}\right)}$$

lies in $\mathscr{C}(\lambda^{-1})$. It follows that $(\lambda^{2n}+1)/q(t) \in \mathscr{C}(\alpha^{-1})$ and consequently $1/q(t) \in \mathscr{C}(\lambda^{-1})$. Given this it is not hard to see that if $p(t)/q(t) \in \mathscr{R}(t)$ and $\deg(q) = n+1$, then $p(\lambda^2)/q(\lambda^2) \in \mathscr{C}(\lambda^{-1})$. Hence, by induction, $\mathscr{R}(\lambda^2) = \mathscr{C}(\lambda^{-1})$.

We can complete the proof that all totally real algebraic integers are tree eigenvalues by showing that $1 \in \mathscr{C}(\lambda^{-1})$. Since λ^2 is an algebraic integer it has a minimal polynomial $\alpha(t)$, which is monic with non-zero constant term. Hence we we may write

$$\alpha(t) = t\beta(t) + \alpha(0),$$

(where β is monic) and consequently

$$\beta(\lambda^2) = -\frac{\alpha(0)}{\lambda^2}$$

Therefore $\beta(\lambda^2) \in \mathscr{C}(\lambda^{-1})$. If $\deg(\beta) = k > 0$, then $\lambda^{-2}(\beta(\lambda^2) - \beta(0))$ is a monic polynomial of degree $k-1$, and so by induction we find that $1 \in \mathscr{C}(\lambda^{-1})$.

1.2.4 Comments

Well, heaven knows where this proof comes from. (I have followed Salez's argument closely, and, and apart from pointing out the role played by walk generating functions, have not made any significant alterations.)

Recall that

$$C_a^{(1)}(X, \lambda^{-1}) = 1 - \frac{\phi(X,\lambda)}{\lambda\phi(X\setminus a, \lambda)}.$$

If X is bipartite, the right side is an even function of λ, which is why we find λ^2 creeping in.

Salez states that $\mathscr{C}(\lambda^{-1})$ is the field $\mathbb{Q}(\lambda^2)$. He may well be right. The problem is that $\mathbb{Q}(t)$ is the quotient field of $\mathbb{Z}[t]$, but the rational functions p/q with $p, q \in \mathbb{Z}[t]$ and $q(t)$ monic are a subring of the quotient field. To be more precise, we are actually dealing with the rational functions p/q with $p, q \in \mathbb{Z}[t]$ such that is $q(t)$ monic and $q(\lambda^2) \neq 0$.

1.3 Laplacian Eigenvalues

Eigenvalues of Laplacians are totally real algebraic integers. Since they are all non-negative, it follows that they are totally positive. The Laplacian eigenvalues of $K_{1,n}$ are 0, 1 and $n-1$, with respective multiplicities 1, $n-1$ and 1, and therefore any non-negative integer is a Laplacian eigenvalue.

If L_1 and L_2 are the Laplacians of graphs X_1 and X_2, then

$$L(X_1 \square X_2) = (L(X_1) \otimes I) + (I \otimes L(X_2)).$$

This yields that the set of eigenvalues of Laplacians is closed under addition.

If $n = |V(X)|$, then

$$L(\overline{X}) = (n-1)I - \Delta - (J - I - A) = nI - J - L(X) = L(K_n) - L(X).$$

Hence if θ is an eigenvalue of $L(X)$ and $\theta \neq 0$, we see that $n - 1 - \theta$ is an eigenvalue of $L(\overline{X})$.

If X is k-regular and θ is an eigenvalue of X, then $k - \theta$ is an eigenvalue of $L(X)$.

1.3.1 Theorem. *Each Laplacian eigenvalue is a graph eigenvalue. If θ is a graph eigenvalue, then there is an integer k such that $k + \theta$ is a Laplacian eigenvalue.* □

Proof. Since the Laplacian is a symmetric integer matrix, it follows from Lemma 1.1.2 that any Laplacian eigenvalue is a graph eigenvalue.

For the second claim, we proved elsewhere[4] that every graph eigenvalue is an eigenvalue of a regular graph. (We may even take the regular graph to be a Cayley graph.) If θ is an eigenvalue of the k-regular graph X, then $-\theta$ is an eigenvalue of the k-regular graph $X \times K_2$ and so $k + \theta$ is a Laplacian eigenvalue. □

[4] Theorem 2.1 in C. D. Godsil, "Eigenvalues of graphs and digraphs". *Linear Algebra Appl.*, **46** (1982), 43–50.

If θ is an algebraic integer with minimal polynomial

$$t^d + a_1 t^{d-1} + \cdots + a_d,$$

its trace is $-a_1$ and its its *reduced trace* is $-a_1/d$. There is a conjecture that there are only finitely many totally positive algebraic integers with reduced trace less than two.

1.3.2 Lemma. *Let L be the Laplacian of a tree. Then either L has a non-zero eigenvalue with reduced trace less than two, or all eigenvalues have reduced trace equal to two.*

Proof. Let T be a tree on n vertices with characteristic polynomial $\phi(t)$ and assume that the factorization of ϕ into irreducibles is

$$\phi(t) = \prod_{i=1}^{k} \phi_i^{m_i}(t).$$

Let τ_i be the trace of a zero of ϕ_i and let d_i be its degree. Note that $\sum_i m_i d_i = n$. The coefficient of t^{n-1} in ϕ is $2n-2$, and therefore

$$2n - 2 = \sum_{i=1}^{k} m_i \tau_i.$$

If $\tau_i \geq 2d_i$ when $i > 1$, we find that

$$\sum_{i=2}^{k} m_i \tau_i \geq \sum_{i=2}^{k} 2 d_i m_i = 2n - 2.$$

This implies that all non-zero eigenvalues have reduced trace equal to two. □

Question: Is every totally positive algebraic integer a Laplacian eigenvalue?

1.4 Directed Graphs

Work on symbolic dynamics is concerned with the spectral radii of non-negative integer matrices. A *Perron number* is a real algebraic integer λ greater than one such that for any algebraic conjugate μ of λ, we have $\lambda > |\mu|$.

The follwing theorem is due to Lind. [5]

[5] See Section 11.1 in Marcus and Lind: "Symbolic Dynamics and Coding" CUP.

1.4.1 Theorem. *Any Perron number is the spectral radius of a non-negative integral matrix.* □

1.5 Interlacing Polynomials

We provide some background on interlacing, and use it to prove that if α is a totally real algebraic integer, some integer multiple of α is a tree eigenvalue.[6]

[6] this is, of course, a consequence of Salez's result

Let f and g be real polynomials of degree n and $n-1$ respectively, each with only real zeros. Assume that $\theta_1, \ldots, \theta_n$ are the zeros of f in non-decreasing order. We say that g *interlaces* f if each interval $[\theta_i, \theta_{i+1}]$ contains a zero of g. Hence if a zero of f has multiplicity k, then its multiplicity as a zero of g is at least $k-1$ and at most $k+1$. You may show that g interlaces f if and only if the rational function

$$\frac{g}{f}$$

has only simple poles, and the residue at each pole is positive (where we use l'Hospital's rule if the multiplicity of θ is greater than one). From this we see that the polynomials that interlace f form a convex cone. You might show that the residue at the zero θ of f is equal to

$$\frac{g(\theta)}{f'(\theta)}$$

If f is a polynomial with all zeros real then its derivative f' interlaces f (and the residue at a pole is the multiplicity of the corresponding zero).

1.5.1 Theorem. *Let f and g be real monic polynomials. The following statements are equivalent:*

(a) *g interlaces f.*

(b) *The rational function g/f has only simple poles, and the residues at its poles are positive.*

(c) *If t is not a zero of f then $(g(t)/f(t))' < 0$.*

(d) *If t is not a zero of g then $(f(t)/g(t))' > 0$.*

Proof. We show that (a) implies (c). If a zero θ of f has multiplicity k then it has multiplicity $k-1$, k or $k+1$ as a zero of g. In the first case θ is a pole of g/f, in the second case g/f is continuous at θ, and in the third it is a (simple) zero of g/f. So g/f is continuous on the interval between two consecutive poles and has a simple zero there. Since each pole is simple, the value of $g(t)/f(t)$ must change sign as we pass from points just to the the left of the pole to points just to the its right. Now (c) follows. The rest is left as exercises. □

1.5.2 Lemma. *Suppose f and g are monic and coprime and* $\deg(g) = \deg(f)-1$. *Then there are real numbers a and b and a monic polynomial h such that*

$$f(t) = (t-a)g(t) - bh(t).$$

If g interlaces f then $b>0$ and h interlaces g.

Proof. We have

$$\frac{f}{g} = t - a - b\frac{h}{g}.$$

Assume f has degree n. If g interlaces f, then f/g must have exactly $n-1$ poles (all simple) and n zeros (all simple too). Therefore h/g has exactly $n-1$ simple poles.

The are positive reals c_ψ such that

$$\frac{f(t)}{g(t)} = \sum_\psi \frac{c_\psi}{t-\psi}$$

and thus

$$\left(\frac{f(t)}{g(t)}\right)' = -\sum_\psi \frac{c_\psi}{(t-\psi)^2};$$

from this we see that $(f/g)'$ is not only positive at any point that is not a pole, but takes arbitrarily large values in the neighborhood of each pole. This implies that the residues of h/g at its poles must have the same sign as b. Since f, g and h are monic we see that $b>0$. □

An $n\times n$ matrix B is *tridiagonal* if $B_{i,j}=0$ whenever $|i-j|>1$ and all off-diagonal entries are non-negative. It is *irreducible* if $B_{i-1,i}B_{i,i-1}\neq 0$ for any i.

1.5.3 Lemma. *If f is a polynomial with distinct real zeros then it is the characteristic polynomial of an irreducible tridiagonal matrix whose entries lie in the extension of the rationals generated by the coefficients of f.*

Proof. Assume $n=\deg(f)$. If f is as stated then it is interlaced by f'. Set $f_0=f$ and $f_1=(1/n)f'$. Recursively define monic polynomials f_i such that

$$f_i = (t-a_i)f_{i+1} - b_i f_{i+2};$$

by the previous lemma $b_i > 0$ and f_{i+2} interlaces f_{i+1}. Let B be the $n \times n$ tridiagonal matrix with

$$B_{i,i} = a_{i-1}, \quad B_{i,i-1} = 1, \quad B_{i-1,i} = b_{i-1}, \qquad (i = 1, \ldots, n).$$

Then by induction we see that $f(t) = \det(tI - B)$. □

Suppose $f(t)$ is even and has distinct real zeroes and $g(t)$ is odd and interlaces f. If

$$f(t) = (t-a)g(t) - bh(t)$$

then

$$f(t) = f(-t) = (-t-a)g(-t) - bh(-t) = (t+a)g(t) - bh(-t)$$

and consequently

$$0 = 2ag(t) - b(h(-t) - h(t)).$$

If $\deg(f) = n$ then g has degree $n-1$ and h has degree $n-2$, so this implies both that $a = 0$ and h is even.

1.5.4 Corollary. *If f is even and its zeros are distinct, then it is the characteristic polynomial of a tridiagonal matrix where all diagonal entries are zero.* □

1.5.5 Theorem. *If θ is a totally real algebraic integer, then some integer multiple of θ is an eigenvalue of a tree.*

Proof. Let f be the minimal polynomial of θ over $\mathbb{Q}$. If $-\theta$ is a zero of f, then for each zero σ of f it follows that $-\sigma$ is also a zero. (Galois.) If $-\theta$ is not a zero of f then $f(-t)$ and $f(t)$ are coprime. Hence either f is even and its zeros are distinct, or $f(t)f(-t)$ is even and its roots are distinct. By the previous lemma we see that θ is an eigenvalue of an irreducible rational tridiagonal matrix T with zero diagonal.

Therefore there is a least positive integer m such that mT is an integer matrix with $m\theta$ as an eigenvalue. If D is the diagonal matrix with $D_{1,1} = 1$ and $D_{i,i} = m^{-1}$ when $i > 1$, then $S = mDTD^{-1}$ is an integer tridiagonal matrix with $S_{i,i-1} = 1$.

Assume that S is $n \times n$ and $\beta_{i-1} = S_{i,i+1}$. Construct a tree as follows. Start with one vertex, at level 0. Join this to β_0 new vertices, each at level 1. Now, recursively, join each vertex at level i to b_i new vertices, deemed to be at level $i+1$. The partition of the vertices of this tree by levels is equitable and its quotient is S. □

One consequence of this is that the set of graph eigenvalues that lie in a given number field K forms an order in $\mathcal{O}_K$.

2

Hagos: Reconstructing $\phi(X, t)$

In https://www.combinatorics.org/ojs/index.php/eljc/article/view/v7i1r12, Hagos proves that that the characteristic polynomial of a graph X is determined by the characteristic polynomials of its vertex-deleted subgraphs along with the characteristic polynomials of their complements. Here we present a version of his proof of this result.

2.1 Reconstruction

Let X be a graph on n vertices. The *deck* of X is a function ρ on the isomorphism classes of graphs on $n-1$ vertices; if Y is a graph on $n-1$ vertices, then $\rho(Y)$ is the number of vertices i in X such that $X \setminus i \cong Y$. Ulam's notorious reconstruction conjecture asserts that if $|V(X)| > 2$, then X is determined by its deck. (Note that K_2 and $2K_1$ have the same deck.) The reconstruction conjecture is false for directed graphs—Stockmeyer contructed pairs of non-isomorphic tournaments with the same deck.[1]

[1] Stockmeyer, P. K., The falsity of the reconstruction conjecture for tournaments, J. Graph Theory 1 (1977), 19–25

Despite very considerable effort, the reconstruction conjecture is still open. It is a fairly easy exercise to prove that a regular graph on more than two vertices is reconstructible. It is known that trees are reconstructible (Kelly), but this is not trivial.

A parameter of a graph is reconstructible if it is determined by its deck. The most trivial example would be the number of vertices. The number of edges is reconstructible, as is the degree sequence.

Since

$$\phi'(X, t) = \sum_{i \in V(X)} \phi(X \setminus i, t)$$

we can reconstruct the derivative $\phi'(X, t)$ of the characteristic polynomial of a graph.

2.1.1 Lemma. *If $|V(X)| = n$, and $k < n$, the number of closed walks in X of length k is reconstructible.*

Proof. We count pairs consisting of closed walks of length k and vertex not in the walk. Let $c_k(Y)$ denote the number of closed walks of length k

in a graph Y. If $n = |V(X)|$, then X gives us exactly $c_k(X)(n-k)$ pairs. On the other hand, there are exactly $c_k(X \setminus i)$ pairs with second coordinate i, implying that

$$c_k(X)(n-k) = \sum_{i \in V(X)} c_k(X \setminus i).$$

We conclude that $c_k(X)$ is reconstructible when $k < n$. □

The type of argument just used is standard.[2]

Tutte proved that the number of Hamiltonian cycles of X is reconstructible; it follows that the numbers

$$c_0(X), \ldots, c_n(X)$$

are reconstructible, and hence we can determine the eigenvalues of X. [3]

[2] Exercise: prove that the number of components of X is reconstructible

[3] Prove the "hence"—you will need invertibility of a Vandermonde matrix

2.2 Reconstructing the Walk Generating Function

Hagos shows that the generating function for all walks in X is determined by the characteristic polynomials of the vertex-deleted subgraphs of X and its complement. We present a version of his proof.

Let $C_i(Z,t)$ be the generating function for the closed walks in Z that start at i. Recall that

$$t^{-1} C_i(X, t^{-1}) = \frac{\phi(X \setminus i, t)}{\phi(X,t)}.$$

2.2.1 Lemma. *Let X and Y be graphs on vertex set $\{1, \ldots, n\}$ such that for each vertex i, we have $\phi(X \setminus i, t) = \phi(Y \setminus i, t)$. Then*

$$C_i(X,t) C_j(Y,t) = C_j(X,t) C_i(Y,t).$$

Proof. We have

$$\begin{aligned} t^{-2} C_i(X, t^{-1})\, C_j(Y, t^{-1}) &= \frac{\phi(X \setminus i, t)\, \phi(Y \setminus j, t)}{\phi(X,t)\, \phi(Y,t)} \\ &= \frac{\phi(Y \setminus i, t)\, \phi(X \setminus j, t)}{\phi(X,t)\, \phi(Y,t)} \\ &= t^{-2} C_j(X, t^{-1})\, C_i(Y, t^{-1}) \end{aligned}$$

□

In the proof, Hagos views walk generating functions as real functions of t. We have

$$W(X,t) = \mathbf{1}^T (I - tA)^{-1} \mathbf{1} = \sum_r \frac{\mathbf{1}^T E_r \mathbf{1}}{1 - t\theta_r}$$

and so if ρ is the spectral radius of X, we see that $W(X,t) \geq 0$ in the interval $[0, \rho^{-1}]$. Similarly $C(X,t) \geq 0$ on the same interval.

2.2.2 Theorem. *If X is not regular, the walk generating function $W(X,t)$ is determined by the characteristic polynomials of the vertex-deleted subgraphs of X and $\overline{X}$.*

Proof. We separate out two lemmas.

2.2.3 Lemma. *We have*

$$\begin{aligned}\big(W(X\setminus i,t)-W(X\setminus j,t)\big)\big(W(X,t)-W(Y,t)\big)\\ =\big(W_i(X,t)W_j(Y,t)-W_i(Y,t)W_j(X,t)\big)\times\\ \big(W_i(X,t)W_j(Y,t)+W_i(Y,t)W_j(X,t)\big)\,(C_i(Y,t)C_j(X,t))^{-1}.\end{aligned}$$

Proof. To see this, we first recall that (for any X)

$$W(X,t)=W(X\setminus i,t)+\frac{W_i(X,t)^2}{C_i(X,t)}.$$

Now

$$\begin{aligned}(W(X\setminus i,t)-W(X\setminus j,t))(W(X,t)-W(Y,t))\\ =(W(X,t)-W(X\setminus i,t))(W(Y,t)-W(Y\setminus j,t))\\ -(W(X,t)-W(X\setminus j,t))(W(Y,t)-W(Y\setminus i,t))\end{aligned}$$

and, since

$$W(Z,t)=W(Z\setminus i,t)+\frac{W_i(Z,t)^2}{C_i(Z,t)},$$

if follows that

$$\begin{aligned}\big(W(X\setminus i,t)-W(X\setminus j,t)\big)\big(W(X,t)-W(Y,t)\big)\\ =\frac{W_i(X,t)^2}{C_i(X,t)}\frac{W_j(Y,t)^2}{C_j(Y,t)}-\frac{W_j(X,t)^2}{C_j(X,t)}\frac{W_i(Y,t)^2}{C_i(Y,t)}\end{aligned}$$

By the lemma above, the denominators here are equal and our claim follows. □

Note that if $t\in[0,\rho^{-1}]$, then

$$\frac{W_i(X,t)W_j(Y,t)+W_j(X,t)W_i(Y,t)}{C_i(X,t)C_j(Y,t)}\geq 0.$$

This implies that

$$(d_X(i)-d_X(j))\big(W(X\setminus i,t)-W(X\setminus j,t)\big)\big(W(X,t)-W(Y,t)\big)$$

and

$$(d_X(i)-d_X(j))\big(W_i(X,t)W_j(Y,t)-W_i(Y,t)W_j(X,t)\big)$$

have the same sign on $[0,\rho^{-1}]$.

2.2.4 Lemma.

$$\begin{aligned}\sum_{i,j}(d_X(i)-d_X(j))\big(W_i(X,t)W_j(Y,t)-W_i(Y,t)W_j(X,t)\big)\\ =2n(W(X,t)-W(Y,t))\end{aligned}$$

Proof. Since corresponding vertices in X and Y have the same valency,

$$\begin{aligned}(d_X(i)-d_X(j))\big(W_i(X,t)W_j(Y,t)-W_i(Y,t)W_j(X,t)\big)\\ = d_X(i)W_i(X,t)W_j(Y,t)+d_X(j)W_j(X,t)W_i(Y,t)\\ -d_Y(i)W_i(Y,t)W_j(X,t)-d_Y(j)W_j(Y,t)W_i(X,t).\end{aligned}$$

If we sum this over all i and j, the result is twice

$$W(Y,t)\sum_i d_X(i)W_i(X,t)-W(X,t)\sum_j d_Y(j)W_j(Y,t). \tag{2.2.1}$$

We note that

$$(I-tA)^{-1}=I+A(I-tA)^{-1}$$

whence

$$W(X,t)=\mathbf{1}^T(I-tA)^{-1}\mathbf{1}=n+t\sum_i d_X(i)W_i(X,t).$$

Therefore the expression in (2.2.1) is equal to

$$W(Y,t)(W(X,t)-n)-W(X,t)(W(Y,t)-n)=n(W_X(t)-W_Y(t)). \qquad \square$$

It follows that

$$(d_X(i)-d_X(j))\big(W(X\setminus i,t)-W(X\setminus j,t)\big)\big(W(X,t)-W(Y,t)\big)$$

and

$$W_X(t)-W_Y(t)$$

have the same sign on $[0,\rho^{-1}]$. However

$$W(X\setminus i,t)-W(X\setminus j,t)=(d_X(j)-d_X(i))\,t+o(t^2)$$

and therefore

$$(d_X(i)-d_X(j))\big(W(X\setminus i,t)-W(X\setminus j,t)\big)\le 0$$

for t in some interval $[0,\epsilon]$ with $0<\epsilon\le\rho^{-1}$. Hence either $W(X,t)=W(Y,t)$ (as desired) or X is regular. $\square$

We can now show that $\phi(X,t)$ is determined by the characteristic polynomials of the vertex-deleted subgraphs of X and $\overline{X}$.

The key observation is that $\phi'(X,t)$ and a zero of $\phi(X,t)$ together determine $\phi(X,t)$. Hence if X is regular with valency k (which we can read off from the polynomials $\phi(X\setminus i,t)$), then $\phi(X,k)=0$ and $\phi'(X,t)=\sum_i\phi(X\setminus i,t)$ and so $\phi(X,t)$ is determined. If X is not regular, we have $W(X,t)$ and the poles of $W(X,t)$ are eigenvalues of X. So we have $\phi(X,t)$.

Remark: If X and Y are strongly regular with the same parameters, then the subgraphs $X\setminus i$ are cospectral to each other, with cospectral complements and they are cospectral to the vertex deleted subgraphs of Y and $\overline{Y}$.

2.3 Reconstructing Eigenvectors

If $n = |V(X)|$, we have the identity (see G&M "Spectral conditions…")

$$t^{-1}W_i(X,t^{-1}) = \left[(-1)^n \frac{\phi(X\setminus i,t)}{\phi(X,t)}\left\{\frac{\phi(\overline{X},-t-1)}{\phi(X,t)} - \frac{\phi(\overline{X\setminus i},-t-1)}{\phi(X\setminus i,t)}\right\}\right]^{1/2}$$

from which it follows that $W_i(X,t)$ is determined by the polynomial deck of X. We have

$$W_i(X,t) = e_i^T(I-tA)^{-1}\mathbf{1} = \sum_r \frac{e_i E_r \mathbf{1}}{1-t\theta_r}.$$

If $E_r\mathbf{1} \neq 0$, then it is an eigenvector for X and, given the walk generating functions $W_i(X,t)$, we can construct the eigenvector. [4]

If X is controllable, all eigenvectors are main, and so X itself can be reconstructed from the polynomial deck alone.

[4] If $E_r\mathbf{1} \neq 0$, then θ_r is a *main eigenvalue* and $E_r\mathbf{1}$ is a *main eigenvector*

2.4 Problem

Cvetković and Gutman started all this by raising the following question:

Is $\phi(X,t)$ determined by the characteristic polynomials of its vertex-deleted subgraphs?

Schwenk doubted this was true. (Citations are in Hagos.)

3
Cayley Graphs for Cyclic and Dihedral Groups

We consider spectral properties of Cayley graphs for some cyclic and dihedral groups.

3.1 Circulants

We prove a result due to Elspas and Turner[1]. Our circulants may be directed. See also Turner.[2]

3.1.1 Theorem. *Two circulants of the same prime order are isomorphic if and only if they have the same minimal polynomial.*

Proof. [words missing]. □

The circulants on 20 vertices with connection sets

$$\{\pm 2, \pm 3, \pm 4, \pm 7\}, \quad \{\pm 3, \pm 6, \pm 7, \pm 8\}$$

are cospectral but not isomorphic. No undirected circulants with fewer vertices are cospectral but not isomorphic.[3]

[1] Bernard Elspas and James Turner "Graphs with circulant adjacency matrices", *Journal of Combinatorial Theory*, **9** (1970), 297–307

[2] J. Turner "Point-symmetric graphs with a prime number of points", *J. Comb. Theory*, **3)** (1967), 136–145

[3] C. Godsil, D. A. Holton, B. McKay "The spectrum of a graph", prehistoric

3.2 Cayley Graphs for the Dihedral Group

We consider dihedral groups of order $2m$, where m is odd. Such a group has a cyclic normal subgroup of index two, and m elements of order two. Let D be dihedral of order $2m$ and let B be the cyclic subgroup of order m in D. Let b be a generator for B and let a be an involution. Then $b^m = 1$ and $b^a = b^{-1}$. We note that

$$(ab^k)^2 = (ab^k a)b^k = b^{-k}b^k = 1$$

and accordingly the products ab^k for $k = 0, \ldots, m-$ are distinct involutions. If we view D as a regular permutation group on $2m$ elements, then B is the kernel of the sign map.

If $\mathscr{C}$ is the connection set for a Cayley graph of D, we define

$$\mathscr{C}_0 = \mathscr{C} \cap B, \quad \mathscr{C}_1 = \mathscr{C} \setminus B.$$

It follows that $X(D,\mathscr{C})$ is the edge-disjoint union of a Cayley graph $X(D,\mathscr{C}_0)$ for the cyclic group of order m and a Cayley graph $X(D,\mathscr{C}_1)$. Since $\mathscr{C}_0$ does not generate D, the first Cayley graph is not connected, it consists of two copies of a Cayley graph for $\mathbb{Z}_m$. The Cayley graph $X(D,\mathscr{C}_1)$ is bipartite.

3.2.1 Lemma. *The adjacency matrices $A(X(D,\mathscr{C}_0))$ and $A(X(D,\mathscr{C}_1))$ commute. There are circulant matrices M and N of order m such that $M = M^T$ and*

$$A(X) = \begin{pmatrix} M & N \\ N^T & M \end{pmatrix}. \qquad \square$$

Proof. Let P be the permutation matrix representing b and let S represent a. Since $ab = b^{-1}a$, we have

$$S(P^k + P^{-k}) = (P^{-k} + P^k)S.$$

We leave the rest as an exercise. $\square$

One consequence of this is that each eigenvalue of $X(D,\mathscr{C})$ is the sum of an eigenvalue of $X(D,\mathscr{C}_0)$ and an eigenvalue of $X(D,\mathscr{C}_1)$. [4]

[4] The trick is to determine the pairing

3.3 Irreducible Modules

Suppose S and T represent involutions and let z be an eigenvector for ST, with eigenvalue λ. Then S swaps the vectors z and Sz and

$$Tz = S(ST)z = \lambda z, \quad TSz = (ST)^{-1}z = \lambda^{-1}z$$

which shows us that the span of z and Sz is invariant under S and T. Hence it is a module for D with dimension at most two. Since the algebra generated by S and T is closed under transpose, the orthogonal complement to the span of $\{z, Sz\}$ is also a D-module.

Therefore $\mathbb{R}^{2m}$ decomposes into irreducible D-modules of dimension one and two. Since m is odd, the commutator subgroup of D is B and therefore D has exactly two linear characters (trivial and sign). As the sum of the squares of the dimensions of the irreducible modules is equal to $2m$, it follows that D has exactly $(m-1)/2$ irreducible modules of dimension two.

Each irreducible module is contained in an eigenspace of $X(D,\mathscr{C})$.

Turning to details, assume S and T are generators for D, equivalently $|ST| = m$. Then λ is an m-th of unity.[5] The irreducible modules of dimension two are thus indexed by the non-trivial m-th roots of unity.

[5] $(ST)^m = 1$

The involutions in D are represented by the matrices $S(ST)^k$.

The matrix representing the action S on $\mathrm{span}(z, Sz)$ is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

As

$$STSz = S(ST)^{-1}z = \lambda^{-1}Sz$$

we see that ST is represented by

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}.$$

Therefore $(ST)^k$ is represented by

$$\begin{pmatrix} \lambda^k & 0 \\ 0 & \lambda^{-k} \end{pmatrix}$$

and $S(ST)^k$ is represented by

$$\begin{pmatrix} 0 & \lambda^{-k} \\ \lambda^k & 0 \end{pmatrix}.$$

If $\Omega \subseteq \{0, \ldots, m-1\}$, we conclude that

$$\sum_{k\in\Omega} S(ST)^k$$

is represented on $\mathrm{span}(z, Sz)$ by

$$\begin{pmatrix} 0 & \sum_{k\in\Omega} \lambda^{-k} \\ \sum_{k\in\Omega} \lambda^k & 0 \end{pmatrix};$$

the eigenvalues of this matrix are

$$\pm\left(\sum_{k\in\Omega} \lambda^{-k} \sum_{k\in\Omega} \lambda^k\right)^{1/2}.$$

4
Directed Graphs

If an operator is not normal, the relation between its eigenvalues and its eigenvalues can be quite weak. And even if a matrix is normal, its principal submatrices need not be normal.

4.1 Matrices and Algebras

Let X be a directed graph with adjacency matrix A. Automorphisms of X are given by permutation matrices that commute with A, but if $PA = AP$ then $P^T A^T = A^T P^P$ and therefore $PA^T = A^T P$. So the automorphisms of X lie in the commutant of the algebra $\langle A, A^T \rangle$. This is much stronger than requiring them to lie in the commutant of the adjacency algebra $\langle A \rangle$. I expect that, in general, $\langle A, A^T \rangle$ is the full matrix algebra, and so $\langle A, A^T \rangle$ is not a particularly useful invariant.

The directed graph with adjacency matrix A^T is known as the *converse* of X. We will dnoete it by X^T.

Two symmetric matrices of the same order are similar if and only if their characteristic polynomials are equal. However if A is an $n \times n$ upper triangular matrix with diagonal entries zero, then $\phi(A, t) = t^n$. Thus all strictly upper triangular matrices of the same order have the same characteristic polynomial. We do have the following:

4.1.1 Lemma. *Two square matrices A and B over the field $\mathbb{F}$ are similar if and only if $tI - A$ and $tI - B$ have the same elementary divisors.* □

Another characterization is that A and B are similar if the have the same rational normal form (aka Frobenius normal form). One consequence of these characterizations is that A^T and A are similar.

4.2 Resolvents

Let A be an $n \times n$ matrix. The *resolvent* $R(z)$ of A is the matrix $(zI - A)^{-1}$. To add combinatorial interest, if A is the adjacency matrix of a directed graph, the resolvent is a modified form of the matrix walk generating function.

If A is a Hermitian matrix, we have

$$(tI - A)^{-1} = \sum_r (t - \theta_r)^{-1} E_r.$$

We can view this as a partial fraction decomposition of $(tI - A)^{-1}$. We derive a version of this for square matrices over any field that contains all the eigenvalues of A.

As

$$R(z) = \frac{1}{\det(zI - A)} \operatorname{adj}(zI - A),$$

each entry of $R(z)$ is a rational function. Let θ be an eigenvalue of A with multiplicity m. Then there are matrices A_i such that

$$R(z) = \sum_{r=-m}^{\infty} A_r (z - \theta)^r;$$

we will determine these matrices. The key is the following simple identity.

4.2.1 Theorem. *If $R(z)$ is the resolvent of some matrix then*

$$R(z) - R(w) = -(z - w) R(z) R(w).$$

Proof. Let $R(z)$ be the resolvent of A. Then

$$(zI - A)(R(z) - R(w))(wI - A) = (wI - A) - (zI - A) = (w - z)I,$$

whence the result follows. □

We note one consequence of this.

4.2.2 Lemma. *The eigenvalues of a symmetric real matrix are real.*

Proof. Let B be the matrix got from A by deleting its first row and column. Considering the (1,1)-entry of the previous identity, we have

$$\frac{\phi(B, z)}{\phi(A, z)} - \frac{\phi(B, w)}{\phi(A, w)} = -(z - w) \sum_k R(z)_{1,k} R(w)_{k,1}.$$

When A is symmetric this yields

$$\phi(B, z)\phi(A, w) - \phi(B, w)\phi(A, z) = -(z - w)\phi(A, z)\phi(A, w) \sum_k R(z)_{1,k} R(w)_{1,k}.$$

Assume by way of contradiction that A is a symmetric matrix of smallest order with a eigenvalue θ that is not real. Then the left side here is zero while if we define $f_k(z) = \phi(A, z) R_{1,k}(z)$, the right side is equal to

$$(\bar{\theta} - \theta) \sum_k f_r(\theta) f_r(\bar{\theta}).$$

Since f_r is a polynomial with real coefficients, all summands are non-negative and since $f_1(z) = \phi(B, z)$, neither $f_1(\theta)$ nor $f_1(\bar{\theta})$ are zero. Therefore the right side is positive, which is impossible. □

4.2.3 Lemma. *Suppose that $R(z)$ is the resolvent of A and that θ is an eigenvalue of A with multiplicity m. If $R(z) = \sum_{-m}^{\infty} A_r (z-\theta)^r$ then*

$$A_r A_s = \begin{cases} -A_{r+s+1}, & r, s \geq 0; \\ A_{r+s+1}, & r, s \leq -1; \\ 0, & \text{otherwise.} \end{cases}$$

Proof. We assume that 0 is an eigenvalue of A, and seek to determine the coefficients A_r in the expansion $R(z) = \sum_{r \geq -m} A_r z^r$. From 4.2.1 we have

$$-\sum_{r,s \geq -m} A_r A_s z^r w^s = -R(z)R(w) = \frac{R(z) - R(w)}{z - w} = \sum_{r \geq -m} A_r \frac{z^r - w^r}{z - w}.$$

The lemma follows for $\theta = 0$ by comparing coefficients of $z^i w^j$ in the two series above, and the general result is an easy consequence of this. □

From this result we see that the matrices A_i, $i = -m, -m+1, \ldots$ commute. We also find that:

$$\begin{aligned} A_r &= (-1)^r A_0^{r+1}, \qquad \text{if } r \geq 0, \\ A_{-r} &= (A_{-2})^{r-1}, \qquad \text{if } r \geq 2, \\ A_{-1} A_{-r} &= A_{-r}, \qquad \text{if } r \geq 0. \end{aligned}$$

Therefore the coefficients in our Laurent series for $R(z)$ are determined by A_0, A_{-1} and A_{-2}, where $(A_{-1})^2 = A_{-1}$ and $(A_2)^m = 0$. Thus A_{-1} is idempotent and A_{-2} is nilpotent, let us denote them respectively by E_θ and N_θ. Now note that

$$(tI - A)R(z) = ((t-z)I + zI - A)R(z) = (t-z)R(z) + I;$$

Putting $t = \theta$ in this yields

$$\begin{aligned} (\theta I - A)A_r &= A_{r-1}, \quad r \neq 0, \\ (\theta I - A)A_0 &= A_{-1} - I, \\ (\theta I - A)A_{-m} &= 0. \end{aligned}$$

Hence

$$N_\theta = (\theta I - A)E_\theta.$$

Define the *principal part* $P_\theta(z)$ of $R(z)$ by

$$P_\theta(z) := \sum_{r=1}^{m} A_{-r}(z-\theta)^{-r}.$$

Thus

$$\begin{aligned} P_\theta(z) &= (z-\theta)^{-1} E_\theta + \sum_{r=1}^{m-1} N_\theta^r (z-\theta)^{-r} \\ &= (z-\theta)^{-1} \sum_{r=0}^{m-1} (\theta I - A)^r E_\theta (z-\theta)^{-r} \end{aligned}$$

The following result provides a partial fraction decomposition of the resolvent.

4.2.4 Theorem. *Let $R(z)$ be the resolvent of A and let $P_\theta(z)$ be the principal part of $R(z)$ at θ. Then $R(z) = \sum_\theta P_\theta(z)$.*

Proof. A rational function in z is called *proper* if the degree of its numerator is less than the degree of its denominator. A proper rational function with no poles is constant. The set of proper rational functions is a vector space.

We note that the entries of $R(z)$ and the entries of $P_\theta(z)$ are proper rational functions. Hence each entry of the difference

$$R(z) - \sum_\theta P_\theta(z);$$

is a proper rational function. By the construction of $P_\theta(z)$, these rational functions have no poles. As both $R(z)$ and $P_\theta(z)$ converge to zero as $z \to \infty$, our theorem follows. □

We know that, if m is the multiplicity of θ as an eigenvalue of A then $A_r = 0$ when $r < -m$, equivalently $(\theta I - A)^m E_\theta = 0$. This implies that the order of the pole of $R(z)$ at θ is at most m.

4.2.5 Theorem. *The order of the pole of $R(z)$ at θ is equal to the multiplicity of θ as a zero of the minimal polynomial of A.*

Proof. Let $\psi(z)$ denote the minimal polynomial of A, let $\nu(\theta)$ be the multiplicity of θ as a zero of $\psi(z)$ and suppose

$$\psi_\theta(z) = \frac{\psi(z)}{(z-\theta)^{\nu(\theta)}}.$$

Let $\mathscr{A}_\theta$ denote the space spanned by the matrices $(\theta I - A)^i E_\theta$ and let $d(\theta)$ be its dimension. Thus $d(\theta)$ is the greatest integer such that $(\theta I - A)^{d(\theta)-1} E_\theta \neq 0$.

As $(\theta I - A)^{d(\theta)} P_\theta(z) = 0$, it follows that

$$\prod_\theta (\theta I - A)^{d(\theta)} R(z) = 0.$$

Since $R(z)$ is invertible, this implies that

$$\prod_\theta (\theta I - A)^{d(\theta)} = 0.$$

From the definition of the minimal polynomial we then deduce that $\nu(\theta) \leq d(\theta)$, for all eigenvalues θ of A. We show next that $\nu(\theta) = d(\theta)$.

The matrices $(\theta I - A)^i E_\theta$ for $i = 0, 1, \ldots, d(\theta) - 1$ form a basis for $\mathscr{A}_\theta$. As $\psi_\theta(\theta) \neq 0$, it follows that the matrix representing the action of $\psi_\theta(A)$ relative to this basis is triangular, with non-zero diagonal entries. In particular, it is invertible. On the other hand, if $M \in \mathscr{A}_\theta$, then

$$0 = (\theta I - A)^{\nu(\theta)} \psi_\theta(A) M = \psi_\theta(A)(\theta I - A)^{\nu} \nu(\theta) M,$$

and this implies that $(\theta I - A)^\nu$ acts as the zero operator on $\mathscr{A}_\theta$. It follows that $\nu(\theta) \geq d(\theta)$. □

4.2.6 Corollary. *For each eigenvalue θ, the matrix E_θ is a polynomial in A.*

Proof. Since $zR(z) \to I$ and $zP_\theta(z) \to E_\theta$ as $z \to \infty$, 4.2.4 implies that

$$I = \sum_\theta E_\theta. \tag{4.2.1}$$

It follows from the proof of 4.2.5 that $\psi_\theta(A)E_\tau = 0$ if $\tau \neq \theta$, whence (4.2.1) yields that

$$(\theta I - A)^i \psi_\theta(A) = (\theta I - A)^i \psi_\theta(A) E_\theta.$$

Referring to the proof of 4.2.5 again, we see that $\psi_\theta(A)E_\theta$ lies in $\mathscr{A}_\theta$. It is not hard to show that the matrices $(\theta I - A)^i \psi_\theta(A)E_\theta$ for $i = 0, 1, \ldots, \nu(\theta)$ form a basis for $\mathscr{A}_\theta$, and accordingly each matrix in $\mathscr{A}_\theta$ must be a polynomial in A. □

4.2.7 Corollary. *Any square matrix A is the sum of a diagonalizable and a nilpotent matrix, each of which is a polynomial in A.*

Proof. As $E_\theta E_\tau = 0$ when $\theta \neq \tau$ and $E_\theta^2 = E_\theta$, the column space of E_θ is an eigenspace for all the idempotents E_τ. Given this, (4.2.1) implies that $\mathbb{F}^n$ is the direct sum of eigenspaces of E_θ. Hence E_θ is diagonalizable; more generally any linear combination of the matrices E_θ is diagonalizable. It is also a polynomial in A.

As $AE_\theta = E_\theta + N_\theta$, it also follows from (4.2.1) that

$$A = \sum_\theta (\theta E_\theta + N_\theta) = \sum_\theta \theta E_\theta + \sum_\theta N_\theta.$$

Since $N_\theta N_\tau = 0$ when $\theta \neq \tau$, it follows that $\sum_\theta N_\theta$ is nilpotent. Since $N_\theta = (\theta I - A)E_\theta$, we see that N_θ is a polynomial in A and, therefore, $\sum_\theta N_\theta$ is too. □

The last result implies that symmetric matrices are diagonalizable—if A is symmetric, so is any polynomial in A, but the only symmetric nilpotent matrix is the zero matrix. It is slightly more difficult to see that the only normal nilpotent matrix is the zero matrix; from this it follows that normal matrices are diagonalizable.

4.3 Graphs and Minimal Polynomials

We first show that the number of connected graphs with a given set of eigenvalues is finite. For this we need some eigenvalue bounds.

4.3.1 Lemma. *Let $\hat{d}$ be the average valency of the connected graph X and let $d_{\max}$ be its maximum valency. If ρ is the spectral radius of X, then*

$$\hat{d} \leq \rho \leq d_{\max}.$$

Further if one equality is tight then both are (and X is regular). □

If Y is a proper subgraph of X, and let ρ_Y and ρ_X denote their respective spectral radii. Then $\rho_X \leq \rho_Y$ and if X is connected, this inequality is strict. If X has a vertex of degree m then $K_{1,m}$ is a subgraph of X with spectral radius $\sqrt{m}$, whence $\sqrt{d_{\max}}$ is another lower bound on the spectral radius of X.

4.3.2 Lemma. *If X has diameter D, then the number of distinct eigenvalues of X is at least $D+1$.* □

We see now that if X is a connected graph with at most s distinct eigenvalues then the diameter of X is at most $s-1$. If the largest eigenvalue is ρ then the maximum valency of X is at most ρ^2. So we have bounded the maximum valency of X and its diameter, and therefore the number of vertices of X is bounded. John Stembridge has extended this result to directed graphs.

Note that if a graph has exactly s distinct eigenvalues, then the minimal polynomial of its adjacency matrix has degree s. Our directed graphs may have multiple loops and arcs. We state and prove Stembridge's result:

4.3.3 Theorem. *For a given polynomial p, there are only finitely many strongly connected directed graphs X such that $p(A(X)) = 0$.*

Proof. Suppose $\deg(p) = r$ and let B be the matrix

$$I + \cdots + A^{r-1}.$$

Since the space of polynomials in A is spanned by $I, \ldots, A^{r-1}$ and since X is strongly connected, we see that all entries of B are positive integers. Therefore if $v = |V(X)|$, then $B\mathbf{1} \geq v\mathbf{1}$.

Let ρ be the spectral radius of A; the spectral radius of B is then

$$R = \frac{\rho^r - 1}{\rho - 1}.$$

Let w be a left eigenvector of A with eigenvalue ρ, we may assume it is positive. Then $w^T B = R w^T$ and

$$\rho w^T \mathbf{1} = w^T B \mathbf{1} \geq v w^T \mathbf{1}$$

and since $w > 0$, it follows that $v \leq R$.

We have bounded the number of vertices of X, we now bound the multiplicity of an arc. The spectral radius of X is an upper bound on the spectral radius of any subgraph.

As X is strongly connected and $p(A) = 0$, each arc of X lies in a directed cycle of length at most r. Define the weight of a cycle to be the product of the entries $A_{i,j}$, where (i, j) runs over the arcs in the cycle. The characteristic polynomial of the adjacency matrix of a directed cycle with length s and weight ω is $x^s - w$. We conclude that if X contains a directed cycle of length s and weight ω, then

$$\omega^{1/s} \leq \rho$$

Therefore ρ^r is an upper bound on the multiplicity of an arc in X. □

4.4 Oriented Graphs

An *oriented graph* is a directed graph with no loops and at most one arc between any two vertices. In matrix terms, $A \circ A^T = 0$. An oriented complete graph is a *tournament*. The signed adjacency matrix of an oriented graph is $A - A^T$, which is a skew-symmetric matrix. Any principal submatrix of $A - A^T$ is the signed adjacency matrix of an induced subgraph. In this sense then, oriented graphs are better behaved than the class of directed graphs with normal adjacency matrix.

Suppose X is bipartite with adjacency matrix

$$A = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}.$$

Then

$$\begin{pmatrix} iI & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}\begin{pmatrix} -iI & 0 \\ 0 & I \end{pmatrix} = i\begin{pmatrix} 0 & B \\ -B^T & 0 \end{pmatrix}$$

and so iA is similar to a skew symmetric matrix. Many results about oriented graphs may this yield information about bipartite graphs.

If M is skew symmetric, then iM is Hermitian and its eigenvalues are all real. Hence the eigenvalues of M are purely imaginary. Further, since

$$-M = M^T$$

and since M and M^T are similar, the eigenvalues of M are symmetric about the x-axis in the complex plane.[1] A consequence is that a skew symmetric matrix of odd order cannot be invertible.

[1] Is there a more elementary proof of this?

If P is a permutation matrix and S is skew symmetric, then $P^T SP$ is skew symmetric and we can identify the automorphisms of an oriented graph with the permutation matrices that commute with $S = A - A^T$. This is perhaps not the right group. Define a signed permutation matrix to be the product PD of a permutation matrix P with a diagonal matrix D having diagonal entries equal to ± 1. Signed permutation matrices are monomial matrices. The *signed automorphism group* of an oriented graph X is the set of signed permutation matrices Q that commute with S. We see that $-I$ is a signed automorphism which acts trivially. We could quotient it out, but choose not to bother.

An oriented graph Y is a *switching* of the oriented graph X if we get Y from X by partitioning $V(X)$ into two cells and then reversing each arc that joins vertices in different cells. If Y is a switching of X, there is a diagonal matrix D with $D^2 = I$ such that

$$DS(X)D = S(Y).$$

So a signed automorphism of an oriented graph is the composition of a switching and a permutation.

4.5 Unimodular Tournaments

[This material is based on "On unimodular tournaments" by Belkouche et al.[2]] A tournament is *unimodular* if the determinant of its skew adjacency matrix is 1. (The determinant of a skew symmetric matrix is always a perfect square.)

[2] arXiv:2109.11809

4.5.1 Lemma. *If S is skew symmetric of order $n \times n$, then* $\det(S)$ *is zero if n is odd, and is the square of an odd integer if n is even.*

Proof. For the square part, look up Pfaffians. As $S^T = -S$, we have

$$\det(S) = \det(S^T) = (-1)^n \det(S)$$

and therefore $\det(S) = 0$ if n is odd. If n is even, we note that, modulo 2,

$$S \cong J - I$$

and $\det(J - I) = (n-1)(-1)^{n-1}$. Therefore $\det(S)$ is odd when n is even. □

Define the signed permutation operator P on $\mathbb{R}^n$ by

$$Pe_r = \begin{cases} e_{r+1}, & r < n; \\ -e_1, & r = n. \end{cases}$$

The skew adjacency matrix S of the transitive tournament on n vertices is given by

$$S_{i,j} = \begin{cases} -1, & i > j; \\ 0, & i = j; \\ 1, & i < j. \end{cases}$$

Then $P^n = -I$ and $S = P + \cdots P^{n-1}$.

A tournament with skew adjacency matrix S is *invertible* if S^{-1} is the skew adjacency matrix of a tournament. The transitive determinant on an even number of vertices is invertible. (It is switching equivalent to its inverse.)

4.5.2 Lemma (Belkouche et al.)**.** *Let S be the skew adjacency matrix of a tournament T on $n = 2m$ vertices. Then:*

(a) The diagonal entries of S^{-1} are zero.

(b) The off-diagonal entries of S^{-1} are odd.

Proof. The diagonal entries of S^{-1} are determinants of skew-adjacency matrices of tournaments on an odd number of vertices, so $S^{-1} \circ I = 0$.

If $D \subseteq V(T)$, then (thanks to Jacobi)

$$\det((tI - S)^{-1})_{D,D} = \frac{\phi(S(T \setminus D), t)}{\phi(S, t)}$$

and consequently

$$\det((S^{-1})_{D,D}) = (-1)^{|D|}\det\big((-S^{-1})_{D,D}\big) = \frac{\det(T\setminus D)}{\det(T)} = \det(T\setminus D).$$

Assume $D = \{i, j\}$. Then

$$(S^{-1})^2_{i,j} = \det((tI - S)^{-1})_{D,D} = \det(T\setminus\{i, j\}) \tag{4.5.1}$$

which implies $(S^{-1})_{i,j}$ is odd. □

The next result is an easy consequence of (4.5.1).

4.5.3 Theorem (Belkouche et al)**.** *Let T be a unimodular tournament with skew-adjacency matrix S. The following are equivalent:*

(a) *S^{-1} is the skew-adjacency matrix of a tournament.*

(b) *Each subtournament $T\setminus\{i, j\}$ is unimodular.*

(c) *The coefficient of t^2 in the characteristic polynomial of T is $\binom{n}{2}$.* □

When are T and T^{-1} cospectral? Switching equivalent?

5
Graphs on Eigenvalues

Let D be the initial state of a continuous quantum walk on X. Then the state $D(T)$ at time t is

$$D(t) = \sum_{r,s} e^{\mathrm{i}t(\theta_r - \theta_s)} E_r D E_s$$

We define the *eigenvalue graph* $\mathscr{E}(X,D)$ induced by D to be the graph with the distinct eigenvalues of X as its vertices, and with an edge joining θ_r and θ_s if $E_r D E_s \neq 0$. This graph will have loops.

If $D(t)$ is real, then $e^{\mathrm{i}t(\theta_r-\theta_s)}$ is real whenever $E_r D E_s \neq 0$, and if $e^{\mathrm{i}t(\theta_r-\theta_s)}$ is real, then $t(\theta_r - \theta_s)$ is an integer multiple of π.

5.1 Edges and Loops

Since we have

$$D = IDI = \left(\sum_r E_r\right) D \left(\sum_r E_r\right) = \sum_{r,s} E_r D E_s,$$

we see that our graph has edges (or, at least, loops).

5.1.1 Lemma. *If $\{\theta_r, \theta_s\}$ is an edge in $\mathscr{E}(X,D)$, there are loops on θ_r and θ_s.*

Proof. We have

$$E_r D E_s = (E_r D^{1/2})(E_s D^{1/2})^T$$

and therefore if $E_r D E_s \neq 0$, then neither $E_r D^{1/2}$ nor $E_s D^{1/2}$ is zero. Consequently

$$E_r D E_r = (E_r D^{1/2})(E_r D^{1/2})^T \neq 0$$

and similarly $E_s D E_s \neq 0$.

5.1.2 Lemma. *If D is a pure state, $\mathscr{E}(X,D)$ has just one connected component, a clique. In general, the complement of a component of $\mathscr{E}(X,D)$ has chromatic number at most* $\mathrm{rk}(D)$.

Proof. Suppose $D = zz^*$. Then

$$E_r D E_s = (E_r z)(E_s z)^*$$

and therefore $E_r D E_s \neq 0$ if and only if $E_r z$ and $E_s z$ are not zero.

Now assume $\mathrm{rk}(D) = m$. Then D can be written as a convex combination of at most m pure states $D_1, \ldots, D_m$. It follows that $E_r D E_s \neq 0$ if and only if $E_r D_i E_s \neq 0$ for some i. Since

$$E_r D E_s = \sum_k E_r D_k E_s$$

it follows that $\mathscr{E}(X, D)$ is a spanning subgraph of the union of the graphs $\mathscr{E}(X, D_k)$, proving the second claim. □

For the case where D is a subset state, we will have more to say in the next section.

5.1.3 Lemma. *There is a loop but no edge on θ_r if and only if D and E_r commute.*

Proof. If $E_r D E_s = 0$ when $s \neq r$, then

$$E_r D = E_r D I = \sum_s E_r D E_s = E_r D E_r.$$

Therefore $E_r D$ is symmetric, and therefore E_r and D commute. □

5.1.4 Corollary. *The eigenvalue graph $\mathscr{E}(X, D)$ has no edges if and only if D commutes with A.*

Proof. Recall that the projection of a matrix D onto the commutant of A is $\sum_r E_r D E_r$. We have

$$D = I D I = \sum_{r,s} E_r D E_s$$

and, as the matrices $E_r D E_s$ are pairwise orthogonal, $E_r D E_s = 0$ when $r \neq s$ if and only if D commutes with A. □

If D commutes, the quantum walk with initial state D is not very interesting.[1]

[1] it's constant!

5.1.5 Lemma. *Let Γ be the Galois group of the splitting field of the characteristic polynomial of X. Each element of Γ induces an automorphism of the eigenvalue graph.*

5.2 Subset States

If $S \subseteq V(X)$, then D_S is the diagonal 01-matrix with $(D_S)_{a,a} = 1$ if and only if $a \in S$. We refere to D_S as a *subset state* although strictly speaking the subset state is $|S|^{-1} D_S$. The matrix D_S represents orthogonal projection onto the subspace

$$\mathrm{span}\{e_a : a \in S\},$$

and therefore a matrix M commutes with the projection D_S if and only if $\mathrm{span}\{e_a : a \in S\}$ is M-invariant.

If $a \in V(X)$, we write D_a instead of $D_{\{a\}}$.[2]

[2] this bothers some referees

We note that

$$E_r D_S E_r = (E_r D_S)(D_S E_r)$$

and so the matrices

$$E_r D_S E_r, \quad D_S E_r^2 D_S = D_S E_r D_S$$

have the same nozero eigenvalues, with the same multiplicities; in particular they have the same rank. Note that $D_S E_r D_S$ is the submatrix of E_r with rows and columns indexed by the vertices in S.

5.2.1 Lemma. *Assume $S \subseteq V(X)$. Then the set of loops in $\mathscr{E}(X, D_S)$ is the union of the set of loops in the graphs $\mathscr{E}(X, D_a)$ for a in S.*

Proof. We have

$$E_r D_S E_r = \sum_{a \in S} E_r D_a E_r.$$

As the matrices $E_r D_a E_r$ are positive semidefinite, their sum is zero if and only if each matrix is zero. □

5.2.2 Lemma. *Assume $S \subseteq V(X)$ and neither $E_r D_S E_r$ nor $E_s D_S E_s$ is zero, but $E_r D_S E_s = 0$. Then:*

(a) If $|S| = 2$, then $\mathrm{rk}(E_r D_S E_r) = \mathrm{rk}(E_s D_S E_s) = 1$.

(b) If $|S| = 3$, either $\mathrm{rk}(E_r D_S E_r) = 1$ or $\mathrm{rk}(E_r D_S E_r) = 1$, and neither has rank three.

Proof. If $E_r D_S E_s = 0$, then

$$\sum_{a \in S} E_r D_a E_s = 0.$$

Suppose $S = \{a, b\}$. Then

$$E_r D_a E_s = -E_r D_b E_s$$

and $E_r e_a$ and $E_r e_b$ are parallel; similarly $E_s e_a$ and $E_s e_b$ are parallel.

Now we assume $S = \{a, b, c\}$. If $M_i = x_i y_i^*$ for $i = 1, 2, 3$ and

$$M_1 + M_2 + M_3 = 0,$$

then

$$\mathrm{rk}(M_1 + M_2) = rk(-M_3) = 1.$$

If x_1 and x_2 are linearly independent, there is a vector z_1 such that

$$z_1^* x_1 = 1, \; z^* x_2 = 0$$

whence

$$z_1^*(M_1 + M_2) = z^* M_1 = y_1^*$$

Similarly there is z_2 such that $z_2^*(M_1 + M_2) = y_2$, and therefore the row space of $M_1 + M_2$ contains y_1 and y_2. As $\mathrm{rk}(M_1 + M_2) = 1$, we infer that y_1 and y_2 are parallel. Consequently one of the spaces

$$\mathrm{span}\{x_1, x_2, x_3\}, \quad \mathrm{span}\{y_1, y_2, y_3\}$$

is 1-dimensional. If

$$x y_1^* + x y_2^* + x y_3^* = 0,$$

then $y_1 + y_2 + y_3 = 0$. □

6
The Matching Polynomial

We use $p(X,k)$ to denote the number of k-matchings in the graph X. The *matching polynomial* of X on n vertices is

$$\mu(X,t) := \sum_k (-1)^k p(X,k)\, t^{n-2k}.$$

6.1 Background

We summarize some results on the matching polynomial (not all of which are currently in use here). Aside from the last theorem (due to Ku and Chen), they all appear in AC).

6.1.1 Lemma. *If $e = ab$ is an edge in X,*

$$\mu(X,t) = \mu(X\setminus e,t) - \mu(X\setminus ab,t). \qquad \square$$

If u is a vertex in X, the *path tree* $T(X,u)$ has the paths in X that start at u as its vertices, with two paths adjacent if one is a maximal proper subgraph of the other. It has the pleasant feature that

$$\frac{\mu(X\setminus u,t)}{\mu(X,t)} = \frac{\phi(T(X,u)\setminus u,t)}{\phi(T(X,u),t)}.$$

and, if X is connected, $\mu(X,t)$ divides $\phi(T(X,u),t)$. Since the eigenvalues of $T(X,u)\setminus u$ interlace the zeros of $T(X,u)$, we see that zeros of $\mu(X\setminus u,t)$ interlace the zeros of $\mu(X,t)$. We also deduce that if X is connected, the largest zero of $\mu(X,t)$ is simple.

6.1.2 Theorem (Heilmann and Lieb)**.** *Let X be a graph with maximum valency Δ. If $\Delta \geq 2$, then the matching zeros of X lie in the open interval $(-2\sqrt{\Delta-1}, 2\sqrt{\Delta-1})$.*

Proof. Let u be a vertex of X with valency Δ. Then the path-tree of X based at u has maximum valency Δ, and the spectral radius of a tree is less than $2\sqrt{\Delta-1}$. $\square$

6.1.3 Lemma. *The matching zeros of $X\setminus u$ interlace those of X.* $\square$

Equivalently the poles of $\mu(X\setminus u,t)/\mu(X,t)$ are simple and their residues are positive.

6.1.4 Lemma. *If P is a path in X, then $\mu(X\setminus P,t)/\mu(X,t)$ has only simple poles.*

Both claims in the next theorem are consequences of this lemma.

6.1.5 Theorem. *The number of distinct zeroes of $\mu(X,t)$ is at least the number of vertices in the longest path of X. The multiplicity of a zero is at most equal to the mininum number of vertex-disjoint paths needed to cover $V(X)$.*

Let $\mathrm{mult}(\theta,X)$ denote the multiplicity of θ as a zero of $\mu(X,t)$.

6.1.6 Theorem (Ku and Chen)**.** *If X is connected and $\mu(\theta,X\setminus a)<\nu(\theta,X)$ for each vertex a of X, then θ is a simple zero of $\mu(X,t)$.* □

When $\theta=0$, this is Gallai's lemma. It implies that if X is vertex transitive, all its matching zeros are simple.

The matching polynomial of $X^{(S)}$ is determined by the sum

$$\sum_{a\in S}\mu(X\setminus a,t)$$

and therefore, if

$$\sum_{a\in S}\mu(X\setminus a,t)=\sum_{a\in T}\mu(X\setminus a,t),$$

then $X^{(S)}$ and $X^{(T)}$ are comatching. Hence[1], if X is vertex transitive, all cones over k-vertex subsets of $V(X)$ are comatching.

[1] as observed by Xiaojing Wang in her Ph.D. thesis

6.2 Multiplicities and Path Sums

We prove some of the lemmas from the previous section.

If $a,b\in V(X)$, let $\mathscr{P}(X)_{a,b}$ (or simply$\mathscr{P}_{a,b}$) denote the set of all paths in X from a to b. Similarly $\mathscr{P}_a$ denote the set of paths that start at a. We use $X\setminus P$ to denote the graph we get by deleting the vertices in P. The key result is the following theorem due to Heilmann and Lieb:

6.2.1 Theorem. *If $a,b\in V(X)$, then*

$$\mu(X\setminus a,t)\mu(X\setminus b,t)-\phi(X,t)\phi(X\setminus ab,t)=\sum_{P\in\mathscr{P}_{a,b}}\mu(X\setminus P,t)^2. \tag{6.2.1}$$

Proof. We first ignore the cases where a and b lie in distinct components, or where a and b lie on a cut-edge of X.

So we assume that there is at least one path in $\mathscr{P}_{a,b}$ with length at least two, and proceed by induction on $|E(G)|$. Assume that $e=au$ is an edge that lies on an ab-path and set $Y=X\setminus e$. By induction,

$$\mu(Y\setminus a,t)\mu(Y\setminus b,t)-\phi(Y,t)\phi(Y\setminus ab,t)=\sum_{P\in\mathscr{P}_{a,b}}\mu(Y\setminus P,t)^2. \tag{6.2.2}$$

Now $X \setminus a = Y \setminus a$ and $X \setminus au = Y \setminus au$ and therefore the difference between the left sides of Equations (6.2.1) and(6.2.2) is

$$\mu(X \setminus a, t)(\mu(X \setminus b, t) - \mu(Y \setminus b, t)) - \phi(X \setminus ab, t)(\phi(X, t) - \phi(Y, t)).$$

Applying the edge-deletion recurrence (twice), we see this difference is equal to

$$-\mu(X \setminus a, t)\mu(X \setminus aub, t) + \phi(X \setminus ab, t)\phi(X \setminus au, t)$$

and by our induction hypothesis (applied to $X \setminus a$), we conclude that this difference equals

$$\sum_{P \in \mathscr{P}(X \setminus a)_{u,b}} \mu((X \setminus a) \setminus P, t)^2.$$

We note that if $P \in \mathscr{P}(X)_{a,b}$, then $X \setminus P = Y \setminus P$ and so the difference between the rides sides of Equations (6.2.1) and(6.2.2) is the sum of terms $\mu(X \setminus P, t)^2$, where P runs over the paths in $\mathscr{P}(X)_{a,b}$ that use the edge au. This completes the proof. □

If we adopt the convention that $\mu(X \setminus ab, t) = 0$ when $a = b$, then the theorem holds in this case too.

This theorem has a number of consequences. Recall that

$$\sum \mu(X \setminus b, t) = \mu'(X, t).$$

Hence if we fix a and sum the identity in Theorem 6.2.1 over b in $V(X)$, we get

$$\mu(X \setminus a, t)\mu'(X, t) - \mu(X, t)\mu'(X \setminus a, t) = \sum_{P \in \mathscr{P}_a} \mu(X \setminus P, t)^2.$$

Next[2], sum this identity over a in $V(X)$ to produce

$$\mu'(X, t)^2 - \mu(X, t)\mu''(X, t) = \sum_{P \in \mathscr{P}(X)} \mu(X \setminus P, t)^2$$

(where we sum over all paths in X).[3]

If θ is a matching zero of multiplicity ν, then by interlacing it is a zero of

$$\mu(X \setminus a, t)\mu(X \setminus b, t) - \phi(X, t)\phi(X \setminus ab, t)$$

with multiplicity at least $2\nu - 2$. Hence it must be zero of $\mu(X \setminus P, t)$ with multiplicity at least $\nu - 1$, for each path in X. This proves Lemma 6.1.4.

We prove that if a P is a path in X, then the number of distinct zeros of X is at least $|V(P)|$. Let $\mathrm{mult}(\theta, Y)$ denote the multiplicity of θ as a matching zero of Y. Then if P is path in X, we have $\mathrm{mult}(\theta, Y) \leq |V(X \setminus P)|$ and therefore

$$\sum_{\theta} (\mathrm{mult}(\theta, X) - 1) \leq |V(X \setminus P)|,$$

yielding our claim.

[2] since summing over b that was fun

[3] You might amuse yourself by verifying that this is correct for the empty graph on n vertices

7
Matching Integral Graphs

Which graphs have the property that all zeros of their matching polynomial are integers.

7.1 Graphs with all Matching Zeros Simple and Integral

In this section and the next, we present results from a paper by Akbari et al.[1]

Lemma 6.1.1 provides an easy proof that the graphs in the margin have the same matching polynomial.

[1] S. Akbari, P. Csikvári, A. Ghafaria, S. Khalashi Ghezelahmad, M. Nahvi. "Graphs with integer matching polynomial zeros". https://arxiv.org/abs/1608.00782v3

Figure 7.1: G_1, G_2: both with matching polynomial $t(t^2-1)(t^2-4)$

7.1.1 Theorem. *If all matching zeros of x are simple and integral, then it is one of K_1, K_2, G_1, G_2, $\overline{C_3 \cup C_4}$.*

Proof. Assume X has n vertices and e edges.

First we treat the case where n is even, say $n = 2k$. If 0 is a matching zero, its multiplicity must be even, so it is not a matching zero. Assume that the positive zeros of $\mu(G,x)$ are $\theta_1,\ldots,\theta_k$.

$$\mu(X,t) = \prod_{i=1}^{k}(t^2-\theta_i^2).$$

Then the coefficient of t^{n-2} in $\mu(X,t)$ is

$$\sum_{i=1}^{k}\theta_i^2 \geq \sum_{i=1}^{k} i^2 = \frac{1}{6}k(k+1)(2k+1).$$

On the other hand, the coefficient of t^{n-2} is the number of 1-matchings aka edges of X. Therefore

$$\frac{1}{6}k(k+1)(2k+1) = |E(X)| \leq \binom{2k}{2},$$

which implies $n \leq 6$. We can deal with these cases in sage.

Now assume $n = 2k+1$ and let $0,\theta_1,\ldots,\theta_k$ be the non-negative matching zeros. Then

$$\binom{2k+1}{2} \geq |E(X)| = \sum_{i=1}^{k}\theta_i^2 \geq \sum_{i=1}^{l} i^2 = \frac{1}{6}k(k+1)(2k+1),$$

which implies that $n \leq 11$.

Suppose $n = 11$. Then $\Delta \leq 10$ and the matching zeros lie in $[-5,5]$. Hence the non-negative zeros are $0,1,\ldots,5$ and so

$$\begin{aligned}\mu(X,t) &= t(t^2-1)(t^2-4)(t^2-9)(t^2-16)(t^2-25)\\ &= t^{11}-55t^9+1023t^7-7645t^5+21076t^3-14400t.\end{aligned}$$

Hence $e = p(X,1) = 55$ and consequently $X = K_{11}$, but

$$\mu(K_{11},t) = t^{11}-55t^9+990t^7-6930t^5+17325t^3-10395t.$$

So now suppose $n = 9$. Then $\Delta \leq 8$ and, again, our zeros lie in $[-5,5]$. Since

$$p(X,1) = \theta_1^2+\theta_2^2+\theta_3^2+\theta_4^2 \leq 36,$$

the positive zeros are 1, 2, 3, 4 and

$$\mu(X,t) = (t^2-1)(t^2-4)(t^2-9)(t^2-16) = t^9-30t^7+273t^5-820t^3+576t.$$

Let d_i be the valency of the i-th vertex of X. Then

$$p(X,2) = \binom{30}{2} - \sum_{i=1}^{9}\binom{d_1}{2},$$

implying that

$$\sum_{i=1}^{9}\binom{d_i}{2} = 435-273 = 162.$$

The average valency of X is $\overline{d} = 20/3$, and hence Jensen's inequality[2] yields that

$$18 = \frac{1}{9}\sum_{i=1}^{9}\binom{d_i}{2} \geq \binom{\overline{d}}{2} = 18.89.$$

[2] $E(f(X)) \geq f(E(X))$ if f is convex

So n even or odd, if the matching zeros of X are simple and integral, then $n \leq 7$. It is now easy to complete the proof of the theorem by direct computation. Or you can amuse youself by finding clever tricks, as Akbari et al. do. □

It may be surprising that an analog of the previous theorem holds for the characteristic polynomial:

7.1.2 Theorem. *If all eigenvalues of X are simple and integral, then X has at most 10 vertices.*

Proof. Lemma 7.10.1 in GS&QW states that if X is a graph on n vertices and σ is the minimum distance between two of the n eigenvalues of X, then $\sigma^2 < 12/(n+1)$. This implies that if the eigenvalues of X are distinct integers, $|V(X)| \leq 10$. □

7.2 Regular Graphs

Akbaria et al. also prove that the only connected regular graph with integer matching zeros is $\overline{C_3 \cup C_4}$. They use the following graph theoretic result (due to Dirac):

7.2.1 Lemma. *If $k \geq 3$ and X is k-regular, it contains a path of length at least* $\min\{2k+1, n\}$. □

7.2.2 Theorem. *If X is regular and connected and its matching zeros are integers, then X is K_2 or $\overline{C_3 \cup C_4}$.*

Proof. Assume $k \geq 3$. Then all matching zeros of S lie in the interval $(2\sqrt{k-1}, 2\sqrt{k-1})$ and therefore therefore the number of distinct matching zeros is at most

$$2\lfloor 2\sqrt{k-1} \rfloor + 1.$$

If X contains a path of length $2k+1$, we then have the inequality

$$2k+1 \leq 2\lfloor 2\sqrt{k-1} \rfloor + 1,$$

implying that $k \leq 2$ (we'll come back to this). If $2k+1 \geq n$, then the lemma tells us that X has a Hamilton path, whence its matching zeros are simple, and the result follows from Theorem 7.1.1.

If $k = 2$, then X is a cycle. The matching zeros of a cycle are simple and lie in $(-2, 2)$, hence if $n \geq 6$, the zeros cannot all be integers. Finish with a computation. □

If the eigenvalues of a graph X with maximum valency k are integers, they lie between $-k$ and k and therefore X has at most $2k+1$ distinct eigenvalues. This implies that the diameter of X is at most $2k$, and so for each positive integer k, there are only finitely many connected graphs with only integer eigenvalues and maximum valency k.

7.3 Other Graphs

We note that random graphs have Hamilton paths,[3] and so their matching polynomials have only simple zeros.

[3] we're being sloppy, google

Akbari et al. prove finiteness results for claw-free graphs and graphs with a perfect matching. We present the latter.

7.3.1 Theorem. *If X has a perfect matching, then $\mu(X, t)$ has a zero in $(0, 1]$. If X has a perfect matching and does not have a zero in $(0, 1)$, it is mK_2 (for some m).*

Proof. Assume X is a graph on $2n$ vertices. If X has a perfect matching, 0 is not a matching zero and so we have

$$\mu(X, t) = \prod_{i=1}^{n} (t^2 - \theta_i^2)$$

with $\theta_i^2 > 0$ for all i.

We note that

$$\frac{p(X, n-1)}{p(X, n)} = \sum_{i=1}^{n} \frac{1}{\theta_i^2}.$$

Since every $(n-1)$-matching extends to a perfect matching in at most one way, and since each perfect matching contains n $(n-1)$-matchings, $p(X, n-1)/p(X, n) \geq n$. So the average value of θ_i^{-2} is at least one. Hence either there is i such that $\theta_i^2 < 1$, or $\theta_i^2 = 1$ for all i. In the latter case, $\sum_i \theta_i^2 = n$, which tells us that $|E(X)| = n$, i.e., all edges of X belong to the perfect matching. □

Any claw-free graph has a matching that misses at most one vertex[4]. (Proving this would be a reasonable exercise in a course that treat's Tutte's theorem.) Hence a connected claw-free graph on an even number of vertices with only integer matching zeros is K_2. In the work of Akbari et al. on claw-free graphs, the claw-free graphs on an odd number vertices form the hard part of of their proof. Still we ask the more general question: are there are only finitely many graphs with integer matching zeros and a near-perfect matching (i.e., a matching that misses just one vertex).

[4] Sumner

Akbaria et al. pose the following:

Question: Are there only finitely many 2-connected graphs with only integer matching zeros?

[The answer is no, Ghorbani 2021 arxiv]

8
Integral Trees

A tree is *integral* if all its eigenvalues are integers. They are scarce and, for many years, it was an open question whether there were integral trees of arbitrarily large diameter. The question was resolved in ???? by Peter Csikvári, who proved that every set of positive integers was the set of eigenvalues of some tree.

8.1 Csikvári's Trees

We choose positive integers $s_1, \ldots, s_m$ and define rooted trees $Cs(s_1, ..., s_m)$ recursively.

Assume T is a tree where all vertices of degree one lie in the same colour class, and define $T(s)$ to be the tree we get by rooting a star $K_{1,s}$ at each vertex in the colour class of the end-vertices. The vertices of degree one in $T(s)$ all lie in the same colour class, and so we can repeat the process as many times as we care to. The result of this process is a tree determined by the initial tree T and a sequence $(s_1, \ldots, s_m)$ of positive integers. If $T = K_{1,s_1}$, we denote the final tree by $\mathrm{Cs}(s_1, \ldots, s_m)$.

Our goal now is to determine the eigenvalues of $T(s)$.

8.1.1 Lemma. *If the adjacency matrix A of T is equal to*

$$\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}$$

and B is $m \times n$, then

$$\phi(T(s), s) = t^m \det((t^2 - s)I - B^T B).$$

Proof. Since $T(s)$ is rooted product, we can apply the machinery from Section **??**. Assume that A is as given and that B is $m \times n$. Assume further that the first m vertices of T are black and the remaining n vertices are white. Assume the end-vertices are white.

Let R be a graph (e.g., a star) with root a and set

$$\psi(t) = \frac{\phi(R, t)}{\phi(R \setminus a, t)}.$$

Then Theorem **??** yields that

$$\phi(T(s),t) = \phi(R\setminus a,t)^n \det\begin{pmatrix} tI & -B \\ -B^T & \psi(t)I \end{pmatrix}$$

and, as

$$\begin{pmatrix} I & 0 \\ t^{-1}B^T & I \end{pmatrix}\begin{pmatrix} tI & -B \\ -B^T & \psi(t)I \end{pmatrix} = \begin{pmatrix} tI & -B \\ 0 & \psi(t)I - t^{-1}B^TB \end{pmatrix},$$

we conclude that

$$\phi(T(s),t) = t^{m-n}\phi(R\setminus a,t)^n \det(t\psi(t)I - B^TB).$$

If $R = K_{1,s}$ (rooted at its central vertex), then

$$\psi(T(s),t)) = \frac{t^2-s}{t}$$

and then

$$\phi(T(s),t) = t^{m+(s-1)n}\det((t^2-s)I - B^TB). \qquad \square$$

If

$$A = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix},$$

then

$$A^2 = \begin{pmatrix} BB^T & 0 \\ 0 & B^TB \end{pmatrix}.$$

Since BB^T and B^TB have the same nonzero eigenvalues with the same multiplicities, we see that the eigenvalues of BB^T are the squares of the eigenvalues of T. Hence if $\theta_1,\dots,\theta_m$ are the positive eigenvalues of T, the eigenvalues of $T(s)$ are

$$\sqrt{s+\theta_i^2}, \qquad (i = 1,\dots,m).$$

8.1.2 Theorem. *The positive eigenvalues of* $\mathrm{Cs}(s_1,\dots,s_m)$ *are*

$$\sqrt{s_m},\ \sqrt{s_m+s_{m-1}},\dots,\sqrt{s_m+\cdots+s_1}. \qquad \square$$

If we have positive integers $r_1,\dots,r_m$ *and set* r_0 *and define*

$$s_i = r_i^2 - r_{i-1}^2, \qquad (i = m,\dots,1),$$

then $\mathrm{Cs}(s_1,\dots,s_m)$ *is a tree of diameter* $2m$ *with all eigenvalues integers.* $\square$

If we have positive integers $r_1,\dots,r_m$ and define, for $i = m, m-1,\dots,1$,

$$s_i = r_i^2 - r_{i-1}^2$$

E. Ghorbani, A. Mohammadian B. and Tayfeh-Rezaie have constructed integral trees of arbitrarily large odd diameter[1]. They have also shown that for each integer k there are only finitely many integral trees with nullity k.[2]

[1] https://arxiv.org/abs/1011.4666

[2] https://arxiv.org/abs/1207.1802v2

9
Algebraic Matching Theory

The theory of matchings is a larger and very well-developed part of graph theory. If $\theta \in \mathbb{R}$ and X is a graph, them $\text{mult}(\theta, X)$ denotes the multiplicity of θ as a zero of $\mu(X, t)$. We observe that the maximum number of vertices in a matching in X is $n - \text{mult}(0, X)$. This leads to a project:[1] rewrite a result about matchings in terms of $\text{mult}(0, X)$, then ask "What happens if we replace 0 by θ?"

[1] another can of worms

To be more concrete, Gallai's lemma states that if X is connected and each vertex is not covered by some matching of maximum size, then the maximum matchings of X each miss exactly one vertex. (This lemma plays an important role in matching theory. It implies, for example that a maximum matching in a connected vertex transitive graph misses at most one vertex.) Our translation of Gallai's lemma is that, if X is connected and for each vertex a in X,

$$\text{mult}(0, X \setminus a) < \text{mult}(0, X),$$

then $\text{mult}(0, X) = 1$.

9.1 Factor-Critical Graphs

It amuses us to give some basic matching theory. We use $\nu(X)$ to denote the maximum number of edges in a matching in X.

Let M and N be two matchings in the graph X. The graph formed by the edges in the symmetric difference $M \oplus N$ has a very simple structure, but the details of this structure are nonetheless very important. We first observe the each vertex is covered by at most two edges from $M \cup N$. Hence $M \oplus N$ has maximum valency two, and so each connected component is either a path of a cycle. If C is a cycle in $M \cup N$ then each vertex in C lies on one edge from M and one from N. It follows that C must contain an even number of edges (and so $M \cup N$ is bipartite). You might also show that if M and N are matchings of maximal size, the paths in $M \oplus N$ all have even length.

An *alternating path* with respect to a matching M is a path where every second edge is in M; alternating cycles are defined similarly. Thus each component of $M \oplus N$ is an alternating path or an alternating cycle.

The following lemma, die to de Caen. is central.

9.1.1 Lemma. *Suppose u and v are vertices in the graph X such that any maximum matching covers at least one of them. If M_u is a maximum matching in X that misses u and M_v is a maximum matching which misses v, there is a path of even length in $M_u \oplus M_v$ with u and v as its end-vertices.*

Proof. Let H be the subgraph formed by the edges in $M_u \oplus M_v$. The components of H are even cycles and paths, as both M_u and M_v have maximum size, each path must have even length.

Our hypothesis implies that u is covered by M_v and v is covered by M_u. Hence u and v have valency one in H and therefore they are end-vertices of paths.

Suppose first that they lie on distinct paths and let P be the path containing u. Then P is an M_v-alternating path with even length and $M_v \oplus E(P)$ is a maximum matching which misses both u and v.

But no maximum matching misses both u and v, therefore u and v must be end-vertices of the same path from H. As all paths in H have even length, we are finished. □

A vertex u in X is missed by a maximum matching if and only if $\nu(G \setminus u) = \nu(G)$. A vertex that is missed by a maximum matching will be called *avoidable*.

9.1.2 Lemma. *If G is bipartite, its avoidable vertices form a coclique.*

Proof. Observe that if uv is an edge in X then no maximum matching can miss both u and v. Suppose that neither u nor v lies in every maximum matching. Then there are maximum matchings M_u and M_v that miss u and v respectively. By the previous lemma there is a uv-path of even length in $M_u \oplus M_v$. Together with the edge e, this path forms an odd cycle. As X is bipartite this is impossible; we conclude that two avoidable vertices cannot be adjacent. □

This lemma implies that in a connected bipartite graph, there is always a vertex that lies in every matching of maximum size.

A graph G is *factor-critical* if it is connected and each vertex-deleted subgraph has a perfect matching. By Lemma 9.1.2, the only factor-critical bipartite graph is K_1. The odd circuits provide a less trivial class of examples; more generally any conncted vertex-transitive graph with an odd number of vertices is factor-critical (as we shall see).

9.1.3 Lemma. *Let u, v and x be distinct avoidable vertices in G. If no maximal matching in G misses both u and x, and no maximal matching misses both x and v, then no maximal matching misses both u and v.*

Proof. Let M_x be a maximum matching that misses x and suppose that N is a maximum matching that misses both u and v. By Lemma 9.1.1 applied to u and x, there is an alternating path in $M_x \cup N$ joining u to x. By the same argument applied to x and v, there is an alternating path from x to v in $M_x \cup N$. Since there can be only one alternating path in $M_x \cup N$ starting at x, this implies that $u = v$, a contradiction. Therefore there is no maximum matching in G that misses u and v. □

This lemma implies that the relation "not both missed by a maximum matching" is an equivalence relation on the set of avoidable vertices of G.

Next we prove Gallai's lemma.

9.1.4 Lemma. *If X is connected and each vertex is missed by a maximum matching, then for each vertex v the graph $X \setminus v$ has a perfect matching.*

Proof. As X is connected and all vertices are avoidable, from the previous lemma, we see that no two vertices of G are missed by a maximum matching. Therefore each maximum matching of G misses exactly one vertex. □

9.1.5 Corollary. *A maximum matching in a connected vertex-transitive graph misses at most one vertex.*

Proof. Suppose G is connected and vertex transitive. If some vertex v is covered by every maximum matching then, since G is vertex transitive, each vertex in G is covered by a maximum matching. Therefore G has a perfect matching.

Otherwise each vertex of G is avoidable and G is factor critical. In this case the result follows from Gallai's lemma. □

9.2

10
Characteristic Matrices of Graphs

10.1 The Characteristic Polynomial

Let X be a graph on n vertices. The *characteristic matrix* $P(X)$ of X is the $n \times n$ matrix whose i-th row consists of the coefficients of $\phi(X \setminus i, t)$. (The columns are ordered with higher degree terms first, although we may choose to vary this).[1] The sum of the rows of P gives the coefficients of the derivative of the characteristic polynomial of X.

[1] with warning, if you're lucky

Our main interest is in the rank of P. We start with some simple observations. We note that $P(X)e_2 = 0$, because our graphs do not have loops, and so $\mathrm{rk}(P(X)) \leq n-1$. If X is bipartite, every second column of $P(X)$ is zero and so $\mathrm{rk}(P(X)) \leq \lfloor (n+1)/2 \rfloor$. If X is regular, the first and third column are parallel and $\mathrm{rk}(P) \leq n-2$.

If θ is an eigenvalue of X with multiplicity $m > 1$, then it is an eigenvalue of $X \setminus i$ with multiplicity at least $m-1$. Hence the vector

$$\begin{pmatrix} \theta^{n-1} \\ \vdots \\ 1 \end{pmatrix}$$

lies in $\ker(P)$. If $m \geq 3$, then θ is a zero of the derivative $\phi(X \setminus i, t)'$ and so the vector

$$\begin{pmatrix} (n-1)\theta^{n-2} \\ (n-2)\theta^{n-3} \\ \vdots \\ 0 \end{pmatrix}$$

also lies in $\ker(P)$. It follows the $\dim(ker(P)) \geq m-1$.

Clearly $\mathrm{rk}(P) = 1$ if P is walk regular. The number of orbits of $\mathrm{Aut}\, X$ is an upper bound on $\mathrm{rk}(P)$.

We have

$$\frac{\phi(X \setminus i, t)}{\phi(X, t)} = \sum_r \frac{(E_r)_{i,i}}{t - \theta_r}.$$

We define the function on the eigenvalues of X with r-th entry equal to $(E_r)_{i,i}$ to be the *spectral density* of the vertex i This function is non-

negative and sums to one, so it is a probability density on the eigenvalues of X. (The support of the spectral density is the eigenvalue support of i.)

10.1.1 Lemma. *The rank of the characteristic matrix of X is at most the number of distinct eigenvalues of X.*

Proof. The distinct rational functions in the spectral decomposition of $(tI - A)^{-1}$ are linearly independent; the poles of these rational functions are a subset of the distinct eigenvalues of X. □

10.2 The Average Mixing Matrix

One reason the characteristic matrix is interesting is because of its connections to the average mixing matrix. Recall that the *average mixing matrix* $\widehat{M}$ is given by

$$\widehat{M} = \sum_r E_r \circ E_r.$$

We know that if $\mathrm{rk}(\widehat{M}) = 1$ then X is K_1 or K_2. We would like to know the graphs for which $\mathrm{rk}(\widehat{M}(X)) = 2$, but have made no real progress.[2]

[2] despite significant effort :-(

If the distinct eigenvalues of X are $\theta_0, \ldots, \theta_d$, then

$$(E_r)_{a,a} = \frac{\phi(X \setminus i, \theta_r)}{\phi'(X, \theta_r)},$$

with the understanding that we need to call on L'Hospital's rule if the multiplicity of the eigenvalue is greater than one. We define

$$\phi_{a,b}(X, t) := \sqrt{\phi(X \setminus a, t)\phi(X \setminus b, t) - \phi(X \setminus \{a, b\}, t)\phi(X, t)}$$

and 'recall' that

$$(E_r)^2_{a,b} = \frac{\phi(X \setminus a, \theta_r)\phi(X \setminus b, \theta_r)}{\phi'(X, \theta_r)^2}$$

(with a similar warning when the multiplicity of θ_r is greater than one).

10.2.1 Theorem. *Let P be the characteristic matrix of X. If the eigenvalues of X are simple, then* $\mathrm{rk}(\widehat{M}) = \mathrm{rk}(P)$.

Proof. We have

$$(E_r)_{u,v} = \lim_{x \to \theta_r} \frac{(x - \theta_r)\big(\phi(X \setminus u, x)\phi(X \setminus v, x) - \phi(X \setminus uv, x)\phi(X, x)\big)^{1/2}}{\phi(X, x)}$$

and since

$$\lim_{x \to \theta_r} \frac{\phi(X, x)}{x - \theta_r} = \phi'(X, \theta_r)$$

we conclude that if θ_r is simple

$$\big((E_r)_{u,v}\big)^2 = \frac{\phi(X \setminus u, \theta_r)\,\phi(X \setminus v, \theta_r)}{\phi'(X, \theta_r)^2}.$$

If B is the $n \times n$ matrix with ur-entry $\phi(X \setminus u, \theta_r)$ and Δ is the $n \times n$ diagonal matrix with r-th diagonal entry $\phi'(X, \theta_r)$, it follows that

$$\widehat{M} = B\Delta^{-2}B^T.$$

Assume $n = |V(X)|$ and let $\theta_1, \ldots, \theta_n$ be the eigenvalues of X. Let V be the $n \times n$ Vandermonde matrix with ij-entry θ_j^{i-1}. Let ϕ be the characteristic polynomial of X.

Let P be the characteristic matrix of X, with columns in increasing order of degree. Then $PV = B$ and

$$\widehat{M} = PV\Delta^{-2}V^TP^T.$$

This proves the theorem. □

10.3 The Matching Polynomial

We use $\mu(X, t)$ to denote the matching polynomial of X. It shares many properties with the characteristic polynomial, in particular it is equal to the characteristic polynomial if and only if X is a forest. Also

$$\sum_{a \in V(X)} \mu(X, t) = \mu(X, t)',$$

in analogy with the characteristic polynomial.

We define the characteristic matrix by using the polynomials $\mu(X \setminus i, t)$ to provide the rows. Since the matching polynomial is an even function if n is even and is odd if n is, by considering the possioble number of nonzero columns of P, we have $rk(P) \leq \lfloor (n+1)/2 \rfloor$.

We note that the characteristic and matching polynomials of a graph are equal modulo two, so the corresponding characteristic matrices are equal modulo two.[3]

[3] and now we start to wonder about the rank mod two—there's no end to the questions…

We say Y is an *extension* if it is obtained by adding one new vertex and joining it to the vertices in some subset of $V(X)$. We will usually assume that $V(X) = \{1, \ldots, n\}$, and ∞ to denote the new vertex. The extension is determined by a 01-vector, the characteristic vector of the set of neighbours of the new vertex. It is comparatively simple to compute the matching polynomial of the extension:

$$\mu(Y, t) = t\mu(X, t) - \sum_{i : i \sim 0} \mu(X \setminus i, t).$$

Therefore if b is the characteristic vector of the neighbours of 0 and P is the characteristic matrix of X, then the matching polynomial of Y is determined by the vector b^TP.

Graphs Y and Z are *comatching* if $\mu(Y, t) = \mu(Z, t)$.

10.3.1 Lemma. *Let b_1 and b_2 be the characteristic vectors of subset of $V(X)$. The extensions corresponding to these vectors are comatching if and only if $(b_1 - b_2)^TP = 0$.* □

Hence we can look for comatching extensions of X by looking for $(0,\pm 1)$ vectors in $\ker(P^T)$. Caelan Wang worked with this in her Ph.D. thesis.

10.4 Laplacians

Here we work with the matrix whose rows are characteristic polynomials of the matrices $L(X\setminus e)$, where e runs over the **edges** of X. Since the constant term of these polynomials is zero, we only use the first $n-1$ coefficients of each polynomials, so P is a matrix of order $|E(X)|\times(|V(X)-1|)$.

First question: what is $\sum_{e\in E(X)}\phi(L(X\setminus e),t)$?

The coefficient of t^{n-m} is the number of rooted forests with exactly m edges (and by a rooted forest, we mean a forest where each component tree is rooted). Thus the constant term is zero and the coefficient of t is n times the number of spanning trees. It follows that

$$\sum_{e\in E(X)}\phi(L(X\setminus e),t)=\frac{d}{dt}\left(t^{|E(X)-|V(X)}\phi(L(X),t)\right).$$

Assume $ab\in E(X)$ and

$$H:=(e_a-e_b)(e_a-e_b)^T.$$

We point out that $F_1=\frac{1}{n}J$, whence $(e_a-e_b)^TF_1(e_a-e_b)=0$ If $L=L(X)$, then

$$\begin{aligned}\frac{\det(tI-L+H)}{\det(tI-A)}&=\det(I-(tI-L)^{-1}(e_a-e_b)(e_a-e_b)^T)\\&=1-(e_a-e_b)T(tI-L)^{-1}(e_a-e_b)\\&=1-\sum_{r\geq 2}\frac{(e_a-e_b)^TF_r(e_a-e_b)}{t-\lambda_r}\end{aligned}$$

We find that

$$(e_a-e_b)^TF_r(e_a-e_b)=(F_r)_{a,a}+(F_r)_{b,b}-2(F_r)_{a,b}$$

Since $F_r\succcurlyeq 0$, the right side is non-negative and the sum over r is two. So we may view

$$\frac{1}{2}(e_a-e_b)^TF_r(e_a-e_b),\qquad (r=0,1,\ldots,d)$$

as the spectral density of an edge.

10.4.1 Lemma. *The rank of the characteristic matrix of $L(X)$ is at most the number of distinct eigenvalues of $L(X)$.* □

Obviously we do not have much to say about the characteristic matrix of Laplacians.

11
McKay's Limbs

We introduce a remarkable tree on 16 vertices found by Brendan McKay.[1]

[1] On the spectral characterisation of trees

11.1 A Remarkable Tree

The tree is in the margin.

Here is the list of its edges.

$$[(0,1),(1,2),(2,3),(2,10),(3,4),(4,5),(5,6),(5,12),$$
$$(6,7),(7,8),(8,9),(8,15),(10,11),(12,13),(13,14)]$$

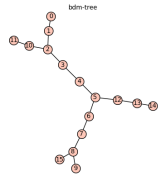

Figure 11.1: A tree on 16 vertices

The vertices of interest are 3 and 6. we denote the tree by B. We use $D(X)$ to denote the distance matrix of X, while $\mathrm{Line}(X)$ is the line graph of X. (Note that for bipartite graphs, the spectrum of the Laplacian and the line graph provide the same information.)

The following list of properties indicates why this tree is remarkable.

(a) $B\setminus 3$ and $B\setminus 6$ are cospectral, with cospectral complements.

(b) $\phi(L(B\setminus 3)) \neq \phi(L(B\setminus 6))$.

(c) Let C_3 and C_6 be the trees produced by adding a vertex of valency one at 3 and 6 respectively. Then:

  (i) $\phi(C_3,t) = \phi(C_6,t)$.

  (ii) $\phi(\overline{C_3},t) = \phi(\overline{C_6},t)$.

  (iii) $\phi(L(C_3),t) = \phi(L(C_6),t)$.

  (iv) $\phi(L(\overline{C_3}),t) = \phi(L(\overline{C_6}),t)$.

  (v) $\phi(D(C_3),t) = \phi(D(C_6),t)$.

  (vi) $\phi(\overline{D(C_3)},t) = \phi(\overline{D(C_6)},t)$.

  (vii) $\phi(\mathrm{Line}(\overline{C_3}),t) = \phi(\mathrm{Line}(\overline{C_6}),t)$.

  (viii) $\phi(\overline{\mathrm{Line}(C_3)},t) = \phi(\overline{\mathrm{Line}(C_6)},t)$.

We stated results for the 1-sum with K_2; we can replace K_2 by any rooted tree. McKay provides what he calls a proof by brute force.

11.2 Isomorphisms for Trees

Let $\Delta(X)$ denote the diagonal matrix of valencies of X. McKay also offers the following (with another proof by brute force).

11.2.1 Theorem. *Let S and T be trees. Then S and T are isomorphic if and only if for each monomial $m(x,y)$ in two non-commuting variables we have*

$$\mathrm{tr}(m(A(S),\Delta(S))) = tr(m(A(T),\Delta(T))). \qquad \square$$

If $\mathrm{tr}(m(A(S),\Delta(S))) = tr(m(A(T),\Delta(T)))$ for all monomials m, then for all non-negative integers k

$$\mathrm{tr}(m(A(S),\Delta(S))^k) = tr(m(A(T),\Delta(T))^k),$$

and therefore $m(A(S),\Delta(S))$ and $m(A(T),\Delta(T))$ have the same characteristic polynomial.[2]. However if there is an orthogonal matrix L such that

[2] These matrices are not symmetric, so they might not be similar

$$L^T A(T)L = A(S), \qquad L^T \Lambda(T)L = L^T \Lambda L,$$

then the condition of the theorem holds for all monomials, and we infer that the trees S and T are isomorphic.

11.3 Distance Matrices of Trees

If X is a graph on n vertices, its *distance matrix* is the $n \times n$ matrix $D = D(X)$ such that

$$D_{u,v} := \mathrm{dist}_X(u,v).$$

This a symmetric matrix with zero diagonal.

Distance matrices of trees were first studied by Graham and Pollak [3], in work that was motivated by problems about embeddings of graphs into hypercubes. (And these problems arose from a practical problem in telecommunication.)

[3]

11.3.1 Theorem. *If T is a tree on n vertices, then* $\det D(T) = (n-1)(-1)^{n-1}2^{n-2}$.

Proof. The proof we present is due to Yan and Yeh [4]. It uses induction on n and a determinental identity due to Jacobi which we state now. If A is an $n \times n$ matrix and $S, T \subseteq \{1,\ldots,n\}$, let $A(S|T)$ denote the submatrix of A we get by deleting from A the rows i such that $i \in S$ and the columns j such that $j \in T$. Then:

[4]

11.3.2 Lemma. *For any square matrix D, we have*

$$\det(D)\det(D(1n|1n)) = \det(D(1|1))\det(D(n|n)) - \det(D(1|n))\det(D(n|1)). \qquad \square$$

Now we prove Theorem 11.3.1 for a tree T on n vertices. The result is immediate if $n = 2$, so assume $n \geq 3$. Our tree has at least two vertices of valency one, assume these are the vertices 1 and n. Suppose that vertices i and j respectively are the unique neighbors of 1 and n in T. If D is the distance matrix of T, then

$$De_1 - De_i = \mathbf{1} - 2e_1, \qquad De_n - De_j = \mathbf{1} - 2e_n$$

whence

$$De_1 - De_i + De_j - De_n = -2e_1 + 2e_n.$$

From this we deduce that

$$\det(D) = -2\det(D(1|1)) + 2(-1)^{n-1}\det(D(n|1)). \tag{11.3.1}$$

We note that, since vertices 1 and n have valency one, the matrices $D(1|1)$, $D(n|n)$ and $D(1n|1n)$ are respectively the distance matrices of the trees $T\setminus 1$, $T\setminus 2$ and $T\setminus\{1,2\}$. By induction, the equation above implies that

$$\det(D) = -2[-(n-2)(-2)^{n-3}] + 2(-1)^{n-1}\det(D(n|1))$$

and using Lemma 11.3.2 we find that

$$\det(D)[-(n-3)(-2)^{n-4})] = [-(n-2)(-2)^{n-3}]^2 - \det(D(1|n))^2.$$

The theorem follows from these two equations.

11.4 Partial Orders

There is natural construction of a partial order from a rooted tree. Assume we have a tree T with a root vertex x. If $u, v \in V(T)$, we declare that $\leq v$ if u lies on the unique path from x to v. We define the incidence matrix Z of this partial order to be the $|V(T)| \times |V(T)|$ matrix given by

$$(Z)_{u,v} = \begin{cases} 1, & \text{if } u \leq v; \\ 0, & \text{otherwise.} \end{cases}$$

We may assume without loss that Z is upper triangular. Since $(Z)_{u,u} = 1$ for each vertex u of T, it follows that $\det(Z) = 1$ and Z is invertible.

Define the square matrix H by

$$H := \mathbf{1}e_1^T + e_1\mathbf{1}^T - 2I;$$

thus

$$H = \begin{pmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & -2I \end{pmatrix}.$$

Observe that

$$H\begin{pmatrix} 1 & 0^T \\ -\frac{1}{2}\mathbf{1} & I \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(n-1) & \mathbf{1}^T \\ 0 & -2I \end{pmatrix}.$$

and so

$$\det(H) = \frac{1}{2}(n-1)(-2)^{n-1} \tag{11.4.1}$$

Graham and Lovász [5] proved the following result. Since $\det(Z) = 1$, the theorem of Graham and Pollak is an immediate consequence. [5]

11.4.1 Theorem. *If D is the the distance matrix of the tree T, then*

$$D(T) = Z^T H Z.$$

Proof. Assume that we have arranged for Z to be upper triangular. If we denote $Z^T\mathbf{1}$ by d, then

$$Z^T\mathbf{1}e_1^T Z + Z^T e_1 \mathbf{1}^T Z = d\mathbf{1}^T + \mathbf{1}d^T$$

and therefore

$$Z^T H Z = d\mathbf{1}^T + \mathbf{1}d^T - 2Z^T Z.$$

If 0 is the root of T, the ij-entry of the right side here is equal to

$$\begin{aligned}(1+\mathrm{dist}(0,i)) + (1+\mathrm{dist}(0,j)) - 2(1+\mathrm{dist}(0,i\wedge j))&\\ = \mathrm{dist}(0,i)) + \mathrm{dist}(0,j) - 2\,\mathrm{dist}(0,i\wedge j)&\\ = \mathrm{dist}(i,j).&\end{aligned}$$

□

11.5 Inverses

From the observation that $D(T) = Z^T H Z$ we deduce that

$$D(T)^{-1} = Z^{-1} H^{-1} Z^{-T}.$$

Since Z is upper triangular and $(Z)_{i,i} = 1$ for all i, we see that Z^{-1} is an integer matrix and is also upper triangular. In fact we can be much more precise.

11.5.1 Lemma. *Let T be a rooted tree and let Z be the incidence matrix of the partial order it determines. Then*

(a) $(Z^{-1})_{u,u} = 1$ *for all vertices u;*

(b) *If u is adjacent to v and lies on the unique path from the root to v, then* $(Z^{-1})_{u,v} = -1$.

(c) *Otherwise* $(Z^{-1})_{u,v} = 0$.

Proof. It is straightforward that the matrix defined by (a), (b) and (c) is Z^{-1}. Alternatively we note that each interval of the poset determined by T is a chain, and then use standard facts about the Möbius function. □

We note an interesting and useful consequence of this lemma. The columns of Z^{-1} are indexed by the vertices of T. The column corresponding to the root is the characteristic vector of the root, viewed as a subset of $V(T)$. The remaining columns are the signed characteristic vectors of the edges of T. Hence

$$Z^{-1} = \begin{pmatrix} e_1 & \tilde{B} \end{pmatrix}$$

where $\tilde{B}$ is the incidence matrix of an orientation of T. (Depending on assumptions, it may be taken to be the orientation where each edge is directed towards the root.)

11.5.2 Theorem. *Let T be a tree with valency matrix Δ and let β denote the vector $(2I-\Delta)\mathbf{1}$. Then*

$$D^{-1} = \frac{1}{2n-2}\beta\beta^T - \frac{1}{2}(\Delta - A).$$

Proof. By multiplication we can verify that

$$H^{-1} = \frac{1}{n-1}\begin{pmatrix} 2 & \mathbf{1}^T \\ \mathbf{1} & \frac{1}{2}(J-(n-1)I) \end{pmatrix}.$$

Denote $\tilde{B}\mathbf{1}$ by γ. We have

$$\begin{pmatrix} e_1 & \tilde{B} \end{pmatrix}\begin{pmatrix} 2 & \mathbf{1}^T \\ \mathbf{1} & \frac{1}{2}(J-(n-1)I) \end{pmatrix}\begin{pmatrix} e_1^T \\ \tilde{B}^T \end{pmatrix}$$
$$= 2e_1e_1^T + e_1\gamma^T + \gamma e_1^T + \frac{1}{2}(\gamma\gamma^T - (n-1)\tilde{B}\tilde{B}^T).$$

Now

$$4e_1e_1^T + 2e_1\gamma^T + 2\gamma e_1^T + \gamma\gamma^T = (2e_1+\gamma)(2e_1+\gamma)^T$$

and since $2e_1+\gamma = \beta$ and $\tilde{B}\tilde{B}^T = \Delta - A$, the result follows. □

Our next lemma is described by Merris [6] as an unpublished result of William Watkins.

11.5.3 Lemma. *If D is the distance matrix of a tree, then $\tilde{B}^T D\tilde{B} = -2I$.*

Proof. We have

$$Z^{-T}DZ^{-1} = \begin{pmatrix} e_1^T \\ \tilde{B}^T \end{pmatrix} D \begin{pmatrix} e_1 & \tilde{B} \end{pmatrix}$$
$$= \begin{pmatrix} e_1^T De_1 & e_1^T D\tilde{B} \\ \tilde{B}^T De_1 & \tilde{B}^T D\tilde{B} \end{pmatrix}$$

and since $Z^{-T}DZ = H$, the result follows. □

11.5.4 Corollary. *If D is the distance matrix of a tree, then*

$$(\Delta - A)D(\Delta - A) = -2(\Delta - A).$$

Proof. From the lemma we have

$$-2(\Delta - A) = -2\tilde{B}\tilde{B}^T = \tilde{B}(\tilde{B}^T D\tilde{B})\tilde{B}^T = (\Delta - A)D(\Delta - A).$$ □

This corollary implies that $\Delta - A$ is a generalised inverse of D.

11.6 Reduced Walks

A walk in a graph X is *reduced* if it does not contain a subsequence of the form uvu. The second and penultimate vertices in a reduced closed walk of length greater than two might be the same. If $|V(X)| = n$, then the matrix generating series $\Phi(X,t)$ is defined by declaring that $(\Phi(X,t))_{u,v}$ is the generating series for the reduced walks in X from u to v, for all vertices u and v of X. We see that if X is a tree, there is exactly one reduced walk between a given pair of vertices, and the length of the walk is the distance between the vertices. Hence if T is a tree, the entries of $\Phi(T,t)$ are polynomials of degree at most the diameter of T. Equivalently we can write

$$\Phi(T,t) = \sum_{r=0} t^r D_r,$$

where $(D_r)_{u,v} = 1$ if $\mathrm{dist}(u,v) = r$ and is otherwise zero. If T is a tree, then $\Phi'(T,1) = D(T)$.

If $A = A(X)$, define $p_r(A)$ to be the matrix (of the same order as A) such that $(p_r(A)_{u,v})$ is the number of reduced walks in X from u to v. Thus

$$\Phi(X,t) = \sum_r t^r p_r(A).$$

Observe that

$$p_0(A) = I, \quad p_1(A) = A, \quad p_2(A) = A^2 - \Delta,$$

where Δ is the diagonal matrix of valencies of X. If $r \geq 3$ we have the recurrence

$$Ap_r(A) = p_{r+1}(A) + \Delta_1 p_{r-1}(A).$$

These calculations were first carried out by Biggs, who observed the implication that $p_r(A)$ is a polynomial in A and Δ, of degree r in A.

We define Δ_1 to be $\Delta - I$. Our next theorem combines two results from Chan and Godsil [7].

[7] Ada Chan and Chris Godsil. Type-ii matrices and combinatorial structures. *Combinatorica*, 30(1):1–24, 2010

11.6.1 Theorem. *For any graph X on at least two vertices,*

$$\Phi(X,t)(I - tA + t^2\Delta_1) = (1 - t^2)I.$$

Furthermore, $\det(I - tA + t^2\Delta_1) = 1 - t^2$ *if and only if T is a tree.*

12
Degrees of Cospectrality

Two graphs X_1 and X_2 with respective adjacency matrices A_1 and A_2 and degree matrices D_1 and D_2 are *degree cospectral* if there is an invertible matrix M such that

$$M^{-1}A_1M = A_2, \quad M^{-1}D_1M = D_2.$$

Clearly cospectral regular graphs are degree cospectral. There is more to be said, and we say it here.

Wanting's construction

12.1 Cospectral Complements

We start by considering an analogous concept. It is an experimental fact that when small[1] graphs are cospectral, their complements are often cospectral. If X_1 and X_2 are cospectral with cospectral complements, we say that they are *cocospectral*[2] Clearly cospectral regular graphs are cocospectral, as are pairs of cospectral graphs created by local switching.

We present two interesting related results.

We use $W(X,t)$ to denote the generating function $\sum_{n\geq 0}\mathbf{1}^TA^n\mathbf{1}$ for all walks in X.[3]

[1] up to nine vertices

[2] Is there a better term?

[3] Compute this generating function when X is k-regular.

12.1.1 Lemma. *Assume X_1 and X_2 are cospectral graphs. Then their complements are cospectral if and only if $W(X_1,t) = W(X_2,t)$.* □

The following is due to Johnson and Newman:

12.1.2 Theorem. *If X_1 and X_2 are cocospectral, there is an orthogonal matrix M such that*

$$M^{-1}A_1M = A_2, \quad M^{-1}\overline{A_1}M = \overline{A_2}.$$ □

Equivalently, there is an orthogonal matrix M such that

$$M^{-1}A_1M = A_2, \quad MJ = JM.$$

12.2 Degree Cospectral Graphs

We note that if two graphs are degree cospectral, then their Laplacian and signless Laplacian matrices are also cospectral. Since the matrix $D^{1/2}$ is a polynomial in the degree matrix D, the normalized Laplacians are also cospectral.

12.2.1 Lemma. *A graph X is connected if and only if $J \in \langle A, D\rangle$.*

Proof. If X is not connected, then clearly J does not lie in $\langle A, D\rangle$. On the other hand, the Laplacian $D - A$ of X lies in $\langle A, D\rangle$ and J is a polynomial in $D - A$ if and only if X is connected. □

12.2.2 Lemma. *If X_1 and X_2 are connected and degree cospectral, they are cocospectral.*

Proof. Assume X_1 and X_2 are connected and degree cospectral. If M is the similarity matrix, then

$$M^{-1}(D_1 - A_1)M = D_2 - A_2.$$

There is a polynomial $p(t)$ such that $p(D_1 - A_1) = J$ and $p(D_2 - A_2) = J$. Therefore $M^1 JM = J$, implying that X_1 and X_2 are cocospectral. □

If X_1 and X_2 are degree cospectral, then

$$M^{-1}A_1M = A_2, \quad M^{-1}D_1M = D_2.$$

Since D_1 and D_2 are diagonal, it follows that there is a permutation matrix P such that $P^T D_2 P = D_1$, and consequently

$$P^T M^{-1}A_1MP = P^T A_2P, \quad P^T M^{-1}D_1MP = D_1.$$

Hence if X_1 and X_2 are degree cospectral, we may assume $D_1 = D_2$ and that L commutes with D_1. This implies that M is block-diagonal, with blocks indexed by vertices of given degree.

12.2.3 Theorem. *Let X_1 and X_2 be degree cospectral, let $\mathscr{D}$ be subset of the set of degrees of X_1, and let $X_i(\mathscr{D})$ (for $i = 1, 2$) be the graph we get from X_i by deleting each vertex with degree in $\mathscr{D}$.*

Our next result is a reformulation of a result due to McKay [4]. [4]

12.2.4 Theorem. *Two trees are degree cospectral if and only if they are isomorphic.*

Proof. Let $\mathscr{P}$ denote the set of all monomials in two non-commuting variables. Then McKay proved that two trees T_1 and T_2 are isomorphic if, for each polynomial p in $\mathscr{P}$, we have $\mathrm{tr}(p(A_1, D_1)) = \mathrm{tr}(p(A_2, D_2))$. Since similar matrices have the same trace, our claim follows. □

12.3 Non-Backtracking Walks

We use $\Phi(X,t)$ to denote the matrix generating function for non-backtracking walks in X. Recall (from Theorem 11.6.1) that

$$\Phi(X,t)(I - tA + t^2(D-I)) = (1-t^2)I$$

We see at once that if X_1 and X_2 are degree cospectral, so are $\Phi(X_1,t)$ and $\Phi(X_2,t)$.

12.4 Polynomials and Similarity

Following Wang et al.[5], we define polynomials

$$\psi(X,s,t,u) := \det(sI - tA - uD).$$

We see that

$$\psi(X,s,1,0),\ \psi(X,s,1,-1),\ \psi(X,s,1,-1)$$

are respectively the characteristic polynomials of the adjacency, Laplacian and signless Laplacian of X. Also, if X has no isolated vertices,

$$\det(uI - D^{-1/2}AD^{-1/2}) = \det(D^{-1})\det(uD - A) = \det(D^{-1})\psi(X,0,1,-u)$$

and therefore $\psi(X,s,t,u)$ also determines the normalized Laplacian.

If X and Y are degree-cospectral, then it follows at once that $\psi(X,s,t,u) = \psi(Y,s,t,u)$. The question is, to what extent is the converse true?

12.4.1 Theorem. *Two matrices A and B over a field are similar if and only if the matrices $tI - A$ and $tI - B$ have the same Smith normal form.* □

Note that if A and B are matrices over the field $\mathbb{F}$, then $tI - A$ and $tI - B$ are matrices over the ring $\mathbb{F}[t]$, which is a principal ideal domain.

Our problem is that $\psi(X,s,t,u)$ is the characteristic polynomial of the matrix $tA + uD$, which is a matrix over the ring $\mathbb{R}[s,t]$.[6] As

$$\det(sI - tA - uD) = u^n \det\left(\frac{s}{t}I - A - \frac{u}{t}\right) = u^n\psi\left(X, s/t, 1, u/t\right),$$

we are free to assume $t = 1$. In this case $A + uD$ is a matrix over $\mathbb{R}[u]$. If we wish to consider similarity, we want our matrices to have inverses and this means we need to work over a field. Thus we want fields that contain $\mathbb{R}[u]$. One choice is the field of rational functions $\mathbb{R}(u)$. A second is the field of Laurent series, $\mathbb{R}((u))$.[7] We may view $\mathbb{R}(u)$ as a subfield of $\mathbb{R}((u))$.

We arrive at three questions: Given matrices $A_i + uD_i$ (for $i = 1,2$),

(a) Are they similar over $\mathbb{R}((u))$?

(b) Are they similar over $\mathbb{R}(u)$?

(c) Are they similar over $\mathbb{R}$.

[5] Wei Wang, Feng Li, Hongliang Lu, and Zongben Xu. Graphs determined by their generalized characteristic polynomials. *Linear algebra and its applications*, 434(5): 1378–1387, 2011

[6] Which is not a principal ideal domain.

[7] which may be referred to as formal Laurent series

In the first two cases we can (at least in principal) decide similarity by computing the Smith normal forms. This does not work in case (c), which is unfortunate because similarity over $\mathbb{R}$ is degree-cospectrality.

Eigenvalues are not obviously useful—the eigenvalues of a matrix over $\mathbb{F}$ lie in an algebraic closure of $\mathbb{F}$, and there does not appear to be any reasonable way of computing these when $\mathbb{F}$ is $\mathbb{R}(u)$ or $\mathbb{R}((u))$. The algebraic closure of the Laurent series is the field of Puiseaux series, so it has a name at least.

13
Random Walks

13.1 Random Variables

In classical physics, each time we repeat the same experimental process we obtain the same result. Indeed, it has been said that repeating the same process repeatedly in the hope of a different outcome is the essence of madness. From this viewpoint quantum physics is mad: if we repeat the same process then the outcome can vary.

The outcome of a classical experiment is determined. The outcome of an experiment in quantum physics is a *random variable*. A standard example of a random variable is the outcome of a toss of a fair coin. Here we expect that if we toss the coin 1000 times, we will observe heads close to 500 times. If the frequency of heads is markedly different from this we deduce that the coin is not fair—we are dealing with a different random variable.

The mathematical machinery for working with random variables is simple, but not particularly intuitive. We start with an event set Ω, the set of possible outcomes of our random process. Thus if the value of random variable is the result of tossing a dice, then $\Omega = \{1, \ldots, 6\}$. To complete the description of the random variable we must specify a probability measure μ; this associates a probability to each measurable subset of Ω and satisfies $\mu(\Omega) = 1$. (If Ω is finite we may drop the word "measurable".) In the cases of most interest to us, Ω will be a countable set and μ will be given by a non-negative real function on Ω whose values sum to 1; this function is a *probability density*. The pair (Ω, μ) may be called an event space. A *random variable* is a function on Ω.

By way of example, we consider tossing an unbiased coin. Here there are two outcomes, heads and tails, which we denote by H and T. Since the coin is unbiased, the probability of heads and the probability of tails are both equal to 0.5. If we were betting on the result of a toss of this coin, winning one dollar when it came up heads and losing one dollar otherwise, our winnings on a toss is a random variable. If we denoted this variable by Y, then Y^2 is also a random variable, but not very interesting since its value

is always one.

We may choose to toss the coin n times. In this case there are 2^n possible outcomes, each with probability 2^{-n}. Each outcome is a sequence of length n over the alphabet $\{H, T\}$. We can define Y_i to be our winnings on the i-th toss. Then the random variable

$$\sum_{i=1}^{n} Y_i$$

is a random variable whose value is our net gain. Information about this sum is naturally interesting to us.

We might also consider tossing this coin until our net winnings reach a fixed value, a say, for the first time. Then our random variable is the number of tosses required. Here the event space Ω is the set of non-negative integers. However some care is needed here since, although true, it is not obvious that $\mu(\Omega) = 1$.

13.2 First Returns

We determine the average number of steps before a random walk on a graph X returns to its starting vertex. We assume throughout that X is connected. Define

$$W(X, t) := (I - t\Delta^{-1}A)^{-1};$$

its ab-entry is the probability generating function for the probability that a walk starting at a is on vertex b after n steps. Let $R_u(X, t)$ be the generating function for the probability that a walk starting at u returns to u for the first time after n steps. As in Section **??** we have

$$W_{u,u}(t) = \frac{1}{1 - R_u(t)}$$

and consequently

$$R_u(t) = 1 - \frac{1}{W_{u,u}(t)}.$$

From this it follows that if the distribution of return times to vertices a and b are equal if and only these vertices are cospectral relative to $\hat{A}$.

Set $B = \Delta^{-1}A$ (for convenience); then

$$B = \Delta^{-1/2}\hat{A}\Delta^{1/2}$$

and therefore using the spectral decomposition of $\hat{A}$ we get

$$B = \sum_r \theta_r \Delta^{-1/2} E_r \Delta^{1/2}.$$

We set $F_r = \Delta^{-1/2}E_r\Delta^{1/2}$ and view this as a spectral decomposition for B (which is not symmetric in general). Note that the matrices F_r are idempotents, $F_rF_s = 0$ if $r \neq s$ and $\sum F_r = I$. Therefore

$$W(X, t) = \sum_r \frac{1}{1 - t\theta_r} F_r.$$

From the calculations in the previous section, assuming $\theta_1 = 1$, we have

$$F_1 = \frac{1}{2e} J\Delta.$$

The expected time for the first return to the vertex u is $R'_u(1)$, which we compute. From our expression for $R_u(t)$ we have

$$R'_u(t) = \frac{W'_{u,u}(t)}{W_{u,u}(t)^2}.$$

Since

$$W_{u,u}(t) = \sum_r \frac{(F_r)_{u,u}}{1 - t\theta_r}$$

where $\theta_1 = 1$ (say) and $\theta_r < 1$ when $r > 1$. Hence we find that

$$R'_u(1) = \frac{(F_r)_{u,u}}{(F_r)^2_{u,u}} = \frac{1}{(F_r)_{u,u}} = \frac{2e}{d_u},$$

where d_u denotes the valency of a.

There are easier and more elegant ways to compute the first return time.

13.3 Partial Fractions

We compute a partial fraction decomposition for $R_u(t)$. If M is a square matrix we use $M_{(i)}$ to denote the matrix we get by deleting the i-th row and column from M.

13.3.1 Lemma. *If M is symmetric and*

$$M = \begin{pmatrix} 0 & b^T \\ b & M_{(1)} \end{pmatrix},$$

then

$$\frac{\phi(M,t)}{\phi(M_{(1)},t)} = t - b^T(tI - M_{(1)})^{-1}b.$$

Proof. We have

$$\begin{pmatrix} t & -b^T \\ -b & tI - M_{(1)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & tI - M_{(1)} \end{pmatrix} \begin{pmatrix} t & b^T \\ (tI - M_{(1)})^{-1}b & I \end{pmatrix}$$

from which it follows that

$$\det(tI - M) = \det(tI - M_{(1)})(t - b^T(tI - M_{(1)})^{-1}b). \qquad \square$$

If $f(t)$ is a polynomial of degree n, we use $\tilde{f}(t)$ to denote the *reversed polynomial* $t^m f(t^{-1})$. With this notation

$$W_{u,u}(t) = \frac{\tilde{\phi}(\hat{A}_{(u)},t)}{\tilde{\phi}(\hat{A},t)}$$

and hence we obtain:

13.3.2 Lemma. *If $\hat{A}_{(u)}$ has the spectral decomposition $\hat{A}_{(u)} = \sum \varphi_r E_r$, then*

$$R_u(t) = t^2 \sum_r \frac{b^T E_r b}{1 - t\varphi_r}$$

As a corollary

$$R'_u(1) = 2 + \sum_r \frac{\varphi_r b^T E_r b}{(1-\varphi_r)^2}.$$

We observe that

$$R_u(1) = b^T (I - \hat{A}_{(u)})^{-1} b, \qquad R'_u(1) = 2 + b^T \hat{A}_{(u)} (I - \hat{A}_{(u)})^{-1} b.$$

It is not immediately obvious that the above expression for $R_u(1)$ evaluates to 1, but it follows at once if we set $t = 1$ in Lemma 13.3.1.

13.4 Hitting Times

If a and b are distinct vertices in the graph X, then we can consider the probability that a walk starting at a arrives at b for the first time at time n. We denote the corresponding probability generating function by $H_{a,b}(X,t)$. We have

$$H_{a,b}(t) W_{b,b}(t) = W_{a,b}(t)$$

and therefore

$$H_{a,b}(X,t) = \frac{W_{a,b}(X,t)}{W_{b,b}(X,t)}.$$

It will be consistent and useful to set $H_{a,a}(t) = 1$; then if $D(X,t)$ is the diagonal matrix with $D_{a,a}(X,t) := W_{a,a}(X,t)$ we can define

$$H(X,t) = W(X,t) D(X,t)^{-1}.$$

We denote the expected time that a walk starting at a hits b by $h_{a,b}$.

13.4.1 Lemma. *If a and b are vertices in X, then*

$$h_{a,b} = 2|E(X)| \sum_{r>1} \frac{1}{1-\theta_r} \left(\frac{(E_r)_{b,b}}{d_b} - \frac{(E_r)_{a,b}}{\sqrt{d_a d_b}} \right).$$

Proof. We have the partial fraction decomposition

$$W(X,t) = \sum_r \frac{1}{1 - t\theta_r} F_r$$

from which we have

$$H_{a,b}(t) = \frac{\sum_r \frac{1}{1-t\theta_r} (F_r)_{a,b}}{\sum_r \frac{1}{1-t\theta_r} (F_r)_{b,b}}$$

Assuming $\theta_1 = 1$ and noting that $(F_1)_{a,b} = (F_1)_{b,b}$, we can write the above fraction as

$$\frac{\alpha + \frac{c}{1-t}}{\beta + \frac{c}{1-t}}$$

where

$$\alpha = \sum_{r>1} \frac{1}{1-t\theta_r}(F_r)_{a,b}, \qquad \beta = \sum_{r>1} \frac{1}{1-t\theta_r}(F_r)_{b,b}.$$

We find that

$$H'_{a,b}(1) = \frac{\beta(1)-\alpha(1)}{c} = \frac{1}{c}\sum_{r>1} \frac{(F_r)_{b,b}-(F_r)_{a,b}}{1-\theta_r}$$

and $c = 2|E(X)|/d_b$. Since

$$(F_r)_{a,b} = \sqrt{\frac{d_a}{d_b}}(E_r)_{a,b}$$

our proof is complete. □

13.5 Symmetry of Hitting Times

We have computed the probability generating functions $H(t)_{a,b}$ for the hitting time, along with the expected hitting times. In this section we consider some applications of these expressions. In particular we consider the relation between the hitting time for walks starting at a and ending on b, and the hitting time for walks starting at b and ending on a.

13.5.1 Lemma. *If a and b are vertices in X, then $H(t)_{a,b} = H(t)_{b,a}$ if and only if a and b are cospectral relative to $\hat{A}$.*

Proof. We have

$$W(X,t) = \Delta^{-1/2}(I - t\hat{A})^{-1}\Delta^{1/2}$$

whence

$$W(X,t)_{a,b} = \frac{d_b}{d_a}W(X,t)_{b,a}.$$

Since $H(t)_{a,b} = W(t)_{a,b}/W(t)_{b,b}$, it follows that $H(t)_{a,b} = H(t)_{b,a}$ if and only if

$$\frac{d_b}{d_a}\frac{W_{b,a}(X,t)}{W(X,t)_{b,b}} = \frac{W_{b,a}(X,t)}{W(X,t)_{a,a}},$$

that is, if and only if

$$\frac{1}{d_a}W(X,t)_{a,a} = \frac{1}{d_b}W(X,t)_{b,b}.$$

Since the constant term of $W(X,t)_{u,u}$ is equal to 1, if we have equality here then $d_a = d_b$. As $W(X,t)_{u,u} = (I - t\hat{A})^{-1}_{u,u}$, we are done. □

We consider what it means when $h_{a,b} = h_{b,a}$. As a first step we observe that

$$h_{a,b} - h_{b,a} = 2|E(X)|\sum_{r>1}\frac{1}{1-\theta_r}\left(\frac{(E_r)_{b,b}}{d_b} - \frac{(E_r)_{a,a}}{d_a}\right),$$

which suggests we should define a 'potential function' υ_a by

$$\upsilon_a = \frac{1}{d_a}\sum_{r>1}\frac{(E_r)_{a,a}}{1-\theta_r}.$$

Then if $M^{\#}$ denotes the pseudoinverse of M (so $MM^{\#}$ is the projection onto the column space of M), we have

$$v_a = \frac{1}{d_a}((I - \hat{A})^{\#})_{a,a}$$

and also

$$h_{a,b} - h_{b,a} = v_b - v_a.$$

[*** I have to confirm the expression for v_a, I'd have expected to multiply, not divide by d_a.***]

13.5.2 Corollary. *If a, b and c are distinct vertices in X, then*

$$h_{a,b} + h_{b,c} + h_{c,a} = h_{a,c} + h_{c,b} + h_{b,a}. \qquad \square$$

Of course we may replace the 3-cycle here by a cycle of any length.

13.6 Generating Functions

The *return time* of a random walk on X starting at vertex i is the expected number of steps before the particle returns to the vertex i, if X is infinite then the return time could be infinite. If u and v are vertices of X, the *hitting time* $H(u,v)$ is the expected number of steps before a random walk starting at u arrives at v. (Note that $H(u,u) = 0$.)

We derive information about return and hitting times using generating functions. Since

$$(I - tM)^{-1} = \sum_{r \geq 0} t^r M^r,$$

then ij-entry of $(I - tM)^{-1}$ is the generating function for the probability that a random walk in X from i is on j after exactly n steps. We use $P(t)$ as an abbreviation for $(I - tM)^{-1}$ and $P_{i,j}(t)$ for its ij entry.

13.6.1 Lemma. *If v_i is the valency of the i-th vertex in X, then*

$$P_{j,i}(t) = \frac{v_i}{v_j} P_{i,j}(t).$$

Proof. We have

$$P_{i,j}(t) = ((I - tM)^{-1})_{i,j} = (\Delta^{-1/2}(I - tN)^{-1}\Delta^{1/2})_{i,j} = \sqrt{\frac{v_j}{v_i}}((I - tN)^{-1})_{i,j}$$

and so, since N is symmetric,

$$\begin{aligned} P_{j,i}(t) &= \sqrt{\frac{v_i}{v_j}}((I - tN)^{-1})_{j,i} \\ &= \sqrt{\frac{v_i}{v_j}}((I - tN)^{-1})_{i,j} \\ &= \frac{v_i}{v_j} P_{i,j}(t). \end{aligned}$$

$\square$

13.7 The n-Cube

Let Q_n denote the n-cube, with the 01-vectors of length n as its vertices, and let A be its adjacency matrix. We want to compute its hitting times.

13.7.1 Lemma. *Suppose A is a real symmetric matrix with an eigenvalue θ and $z_1, \dots, z_m$ is an orthonormal basis for the eigenspace belonging to θ. Then*

$$E_\theta = \sum_i z_i z_i^T$$

is the idempotent belonging to θ in the spectral decomposition of A. □

For the d-cube, if $a \in \mathbb{Z}_2^d$ then the map ψ_a on $\mathbb{Z}_2^d$ given by

$$\psi_a(x) = (-1)^{a^T x}$$

is an eigenvector for A with eigenvalue $d - 2\,\mathrm{wt}(a)$. This a ± 1 vector and so its norm is 2^d. If $0 \le j \le d$, then the idempotent associated with the eigenvalue $d - 2j$ is

$$E_j = 2^{-d} \sum_{a:\mathrm{wt}(a)=j} \psi_a \psi_a^T$$

and

$$(E_j)_{x,y} = 2^{-d} \sum_{a:\mathrm{wt}(a)=j} (-1)^{a^T(x+y)}.$$

If $\ell = \mathrm{wt}(x+y)$, then

$$(E_j)_{x,y} = 2^{-d} \sum_{i=0}^{j \wedge \ell} \binom{\ell}{i} \binom{d-\ell}{j-i} (-1)^i.$$

If we take $x = \mathbf{0}$ and $y = \mathbf{1}$, then $\ell = d$ and

$$(E_j)_{\mathbf{0},\mathbf{1}} = 2^{-d} \binom{d}{j} (-1)^j, \qquad (E_j)_{\mathbf{1},\mathbf{1}} = 2^{-d} \binom{d}{j}$$

By (**??**),

$$\begin{aligned} H_{\mathbf{0},\mathbf{1}} &= \sum_{j=1}^{d} \frac{1}{1 - \frac{d-2j}{d}} \binom{d}{j} (1 - (-1)^j) \\ &= \sum_{j=1}^{d} \frac{d}{2j} \binom{d}{j} (1 - (-1)^j) \\ &= \sum_{j \text{ odd}} \frac{d}{j} \binom{d}{j}. \end{aligned}$$

Observe that

$$\begin{aligned} \sum_{j \text{ odd}} \frac{d}{j} \binom{d}{j} &\ge \sum_{j \text{ odd}} \frac{d+1}{j+1} \binom{d}{j} \\ &= \sum_{j \text{ odd}} \binom{d+1}{j+1} \\ &= 2^d - 1. \end{aligned}$$

It can be shown that this sum is bounded above by 2^{d+1}, and that it is asymptotic to 2^d.

13.8 Commute Times

The *commute time* $K_{i,j}$ is the average time taken for a walk that visits j, having started at i. We have

$$K_{i,j} = H_{i,j} + H_{j,i}.$$

From (**??**),

$$K_{i,j} = 2e \sum_{s=2}^{r} \frac{1}{1-\lambda_s} \left(\frac{(E_s)_{j,j}}{v_j} + \frac{(E_s)_{i,i}}{v_i} - 2\frac{(E_s)_{i,j}}{\sqrt{v_i v_j}} \right)$$

If X has v vertices and is k-regular, this reduces to

$$K_{i,j} = v \sum_{s=2}^{r} \frac{1}{1-\lambda_s} ((E_s)_{j,j} + (E_s)_{i,i} - 2(E_s)_{i,j})$$

The key observation here is that, for all s

$$(E_s)_{j,j} + (E_s)_{i,i} - 2(E_s)_{i,j} \geq 0,$$

however we leave the proof as an exercise. Since $\sum_s E_s = I$, we also have

$$\sum_{s=2}^{r} E_s = I - E_1 = I - \frac{1}{2e} \Delta^{1/2} J \Delta^{1/2} = I - \frac{1}{v} J$$

and hence

$$\sum_{s=2}^{r} ((E_s)_{j,j} + (E_s)_{i,i} - 2(E_s)_{i,j}) = 2\frac{v-1}{v} - 2\frac{1}{v} = \frac{2v-4}{v}.$$

Since $-1 \leq \lambda_i \leq 1$, it follows that

$$v - 2 \leq K_{i,j} \leq \frac{1}{1-\lambda_2}(2v-4).$$

13.9 Mixing

We have seen that if M is the transition matrix of an aperiodic irreducible Markov chain then M^r converges to a positive matrix of the form $\mathbf{1}z^T$. Although this is very useful to know, in many cases we may need information about how quickly this sequence of matrices converges. Thus, if we denote the limit by M^∞, we want information about

$$\|M^r - M^\infty\|$$

as a function of r.

We consider random walks on regular directed graphs. (By regular we mean that the out- and in-valency of each vertex is equal.) Assume each vertex in X has out-valency k. Then

$$M = \frac{1}{k} A$$

is the transition matrix of a random walk on X. Our proofs of the existence of a limiting probability distribution still apply, but any argument that depends on spectral decomposition for symmetric matrices does not hold.

13.9.1 Theorem. *Let A be the adjacency matrix of a regular directed graph with valency k and define*

$$\sigma_2 = \max_{x \in \mathbf{1}^\perp} \frac{\|Ax\|}{\|x\|}.$$

Let x_S and x_T be the characteristic vectors of subsets S and T of $V(X)$. Then

$$\left| \langle x_S, A^r x_T \rangle - \frac{|S||T|}{v} k^r \right| \leq \sigma_2^r \sqrt{\left(|S| - \frac{|S|^2}{v} \right) \left(|T| - \frac{|T|^2}{v} \right)}.$$

Proof. Assume $v = |V(X)|$ and define

$$u_S = x_S - \frac{|S|}{v}\mathbf{1}, \quad u_T = x_T - \frac{|T|}{v}\mathbf{1}.$$

Note that u_S and u_T are the orthogonal projections of x_S and x_T onto $\mathbf{1}^\perp$, and therefore they are orthogonal to $\mathbf{1}$.

Since $A^T\mathbf{1} = k\mathbf{1}$, both A and A^T fix the space of constant vectors and therefore both A and A^T fix $\mathbf{1}^\perp$. Hence if $x \in \mathbf{1}^\perp$ then $Ax \in \mathbf{1}^\perp$ and it follows that if $u \in \mathbf{1}^\perp$, then

$$\|A^r u\| \leq \sigma_2 \|u\|.$$

Now

$$x_S = \frac{|S|}{v}\mathbf{1} + u_S, \quad x_T = k^r \frac{|T|}{v}\mathbf{1} + A^r u_T$$

and therefore

$$\langle x_S, A^r x_T \rangle = \frac{|S||T|}{v} k^r + \langle u_S, A^r u_T \rangle.$$

By Cauchy-Schwarz,

$$|\langle u_S, A^r u_T \rangle| \leq \|u_S\| \, \|A^r u_T\| \leq \sigma_2^r \|u_S\| \, \|u_T\|$$

Since u_S and $\mathbf{1}$ are orthogonal,

$$\|u_S\|^2 = \|x_S^2\| - \frac{|S|^2}{v^2}\|\mathbf{1}\| = |S| - \frac{|S|^2}{v} = |S|\Big(1 - \frac{|S|}{v}\Big).$$

A similar statement holds for $\|u_T\|$, and the theorem follows. □

Note that $\langle x_S, A^r x_T \rangle$ is the number of walks in X from a vertex in S to a vertex in T. If we take S and T to be distinct vertices of X, this theorem implies that if $M = \frac{1}{k}A$, then

$$\left| (M^r)_{i,j} - \frac{1}{v} \right| \leq \left(\frac{\sigma_2}{k} \right)^r.$$

13.10 Continuous Markov Chains

We consider continuous walks. Here we have walks on a graph where in a given small time interval δ_t we move from the current vertex i with probability proportional to δ_t/d_i and, if we do move, all neighbors are equally likely. We can approximate this by a discrete walk with transition matrix

$$M_\epsilon := I - \epsilon\Delta + \epsilon A = I + \epsilon(A - \Delta)$$

We note that

$$M_\epsilon^{k/\epsilon} \approx \exp(k(A - \Delta))$$

In general a continuous Markov chain is described by a matrix

$$\exp(tB)$$

which is required to be row stochastic for all t. It will be non-negative for all t if $I + \epsilon B$ is non-negative whenever ϵ is positive and small enough, which means that all off-diagonal entries of B should be non-negative. The row sums of $\exp(tB)$ equal 1 if and only if $B\mathbf{1} = 0$, that is, if and only if the rows of B sum to 0.

The continuous random walk on a graph is given by

$$\exp(t(A - \Delta));$$

we saw above that it can be viewed as a limit of a sequence of discrete random walks on X.

There is a connection between continuous quantum walks and classical continuous walks on a graph. Consider $U(t)$ for values of t such that t^3 is negligible. Then

$$U(t) \approx I + i\,t A - \frac{1}{2}t^2 A^2$$

and therefore

$$U(t) \circ U(-t) \approx I + t^2(A \circ A) - t^2(A^2 \circ I).$$

If D is the diagonal matrix of valencies of X, we can express the right side here as $I + t^2(A - D)$. So a continuous quantum walk at time t looks like a classical continuous walk at time t^2 (when terms of order t^3 or higher are negligible).

As an exercise you might show that if E is diagonal and

$$U(t) := \exp(i\,t(A + E)),$$

then $U(t) \circ U(-t) \approx I + t^2(A - D)$.

Line Digraphs, Non-backtracking Walks

13.11 Line Digraphs

We start with line graphs of directed graphs.

The *line digraph* $LD(X)$ of a directed graph X has the arcs of the directed graph as vertices, with arcs ab and cd adjacent if $b = c$. The *strict line digraph* has the same vertex set, but the arcs ab and cd are adjacent if and only if $b = c$ and $a \neq d$. (We will assume our directed graphs do not have loops.)

The *head* of the arc ab is the vertex b, the *tail* is the vertex a. Define the hear-arc incidence matrix D_h to be the 01-matrix with rows indexed by vertices, columns by arcs and with $(D_h)_{u,\alpha} = 1$ if the vertex u is the head of the arc α. Define the tail-arc matrix D_t analogously. The matrix $D_h^T D_h$ is diagonal, and its i-th diagonal entry is the in-valency of the i-th vertex. Similarly $D_t^T D_t$ is diagonal and records the out-valencies of the vertices. We recall that a directed graph is *Eulerian* if each vertex has the same in- and out-valency. If X is an Eulerian directed graph, $D_h^T D_h = D_t^T D_t$.

13.11.1 Lemma. *If X is a directed graph then $D_t D_h^T = A(X)$ and $D_h^T D_t = A(LD(X))$.* □

We observe that $D_t^T D_h = (A(LD(X)))^T$. Also

$$\det(I - sA(X)) = \det(I - sD_t D_h^T) = \det(I - sD_h^T D_t) = \det(I - sA(LD(X))).$$

That's it for line digraphs of directed graphs. We now restrict ourselves to line digraphs of graphs. Here we define an *arc* to be an ordered pair of adjacent vertices, so each edge provides two arcs. (Note that, viewed as directed graphs, all graphs are Eulerian.)

One more matrix. We use R to denote the *arc-reversal matrix*. It is the permutation matrix with rows and columns indexed by the arcs of X, and $R_{\alpha,\beta} = 1$ if and only if β is the reverse of β. Clearly $R^2 = I$ and $D_t R = D_h$, hence

$$RD_h^T D_t R = D_t^T D_h = (A(LD(X)))^T.$$

The adjacency matrix of the strict line digraph is $A(LD(X)) - R$.

13.12 Non-backtracking Walks

A walk in a graph is *non-backtracking* if it does not contain a subsequence of the form uvu. To put this a little differently, a walk in a graph is specified by a sequence of arcs, and the walk is non-backtracking if it does not use an arc and its reverse in consecutive steps. Thus each non-backtracking walk on the graph X corresponds to a walk on the strict line digraph.

We define $\Phi(X,t)$ to be the $|V(X)| \times |V(X)|$ matrix with uv-entry equal the generating function for the non-backtracking walks from u to v. If $A = A(X)$, define $p_r(A)$ to be the matrix (of the same order as A) such that $(p_r(A)_{u,v})$ is the number of reduced walks in X from u to v. Thus

$$\Phi(X,t) = \sum_r t^r p_r(A).$$

Observe that

$$p_0(A) = I, \quad p_1(A) = A, \quad p_2(A) = A^2 - \Delta,$$

where Δ is the diagonal matrix of valencies of X. If $r \geq 2$ we have the recurrence

$$Ap_r(A) = p_{r+1}(A) + \Delta_1 p_{r-1}(A). \tag{13.12.1}$$

These calculations were first carried out by Biggs, who observed the implication that $p_r(A)$ is a polynomial in A and Δ, of degree r in A. Note that if $A_r(X)$ denotes the r-th distance matrix of X and T is a tree, then

$$\Phi(X,t) = \sum_r t^r A_r(T).$$

We define Δ_1 to be $\Delta - I$. Our next theorem combines two results from Chan and Godsil [1].

[1]

13.12.1 Theorem. *For any graph X on at least two vertices,*

$$\Phi(X,t)(I - tA + t^2\Delta_1) = (1-t^2)I.$$

Furthermore, $\det(I - tA + t^2\Delta_1) = 1 - t^2$ *if and only if T is a tree.* □

If X is weighted, the weight of a walk is the product of the weights of the arcs its uses.

From the recurrence in Equation (13.12.1), we see that if $r \geq 2$, then

$$\begin{pmatrix} p_{r+1}(A) \\ p_r(A) \end{pmatrix} = \begin{pmatrix} A & -\Delta_1 \\ I & 0 \end{pmatrix}^{r-1} \begin{pmatrix} A^2 - \Delta \\ A \end{pmatrix}.$$

13.13 Strict Line Digraphs

We want spectral information on the strict line digraph. This matrix lies in the algebra generated by $A(LD(X))$ and R, and it is best to work at the level of this algebra.

The matrices D_h^T and D_t^T are characteristic matrices of partitions of the arcs of X. If $\Delta := D_h D_h^T$, then the matrices

$$D_h^T \Delta^{-1/2}, \quad D_t^T \Delta^{-1/2}$$

are normalized characteristic matrices and therefore the matrices

$$P_h = D_h^T \Delta^{-1} D_h, \quad P_t = D_t^T \Delta^{-1} D_t$$

are projections from the space of functions on the arcs of X to the functions on arcs that are constant on arcs with the same head, respectively the same tail. We observe that

$$P_h \mathbf{1} = D_h^T \Delta^{-1} D_h \mathbf{1} = D_h^T \Delta^{-1} \Delta \mathbf{1} = \mathbf{1}.$$

We have

$$\Delta^{-1/2} D_t D_h^T \Delta^{-1/2} = \Delta^{-1/2} A \Delta^{-1/2}$$

and thus $\Delta^{-1/2} D_t D_h^T \Delta^{-1/2}$ is the *normalized Laplacian* of X. This has the same non-zero eigenvalues with the same multiplicities as

$$D_h^T \Delta^{-1/2} \Delta^{-1/2} D_t = D_h^T \Delta^{-1} D_t = D_h^T \Delta^{-1} D_h R = P_h R.$$

As noted above, each row of $P_h R$ sums to one and therefore $P_h R$ is the transition matrix of the usual random walk on $LD(X)$.

The matrices $2P_h - I$ and R are orthogonal, hence so is the product $(2P_h - I)R$. This is the transition matrix for the Grover walk on X.

14
Spectral Characterization of Controllable Graphs

This an exposition of work from a series of papers by Wang (sometimes with coauthors). It is joint work with Chen Xie.

14.1 Controllable Graphs

The *walk matrix* of a graph X on n vertices with adjacency matrix A is

$$W_X = \begin{pmatrix} \mathbf{1} & A\mathbf{1} & \cdots & A^{n-1}\mathbf{1} \end{pmatrix}$$

We say that X is *controllable* if W_X is invertible; equivalently, if the $\mathbb{R}[A]$-submodule of $\mathbb{R}^n$ generated by $\mathbf{1}$ is equal to $\mathbb{R}^n$. We have discussed a more general version of this concept elsewhere, so our treatment will be somewhat sketchy.

If two rows of W_X are equal, then $\mathrm{rk}(W_X) < n$ and X is not controllable. Hence if X is controllable, its automorphism group must be trivial. Also if A has an eigenvalue θ with multiplicity greater than one, then this eigenspace must contain an eigenvector z orthogonal to $\mathbf{1}$. In this case $z^T A^k \mathbf{1} = 0$ for all k, whence $z^T W = 0$ and again X is not controllable. A variant of this argument shows that if X and $\overline{X}$ have a common eigenvector, then X is not controllable.

Assume X is controllable and that

$$A = \sum_{j=1}^{n} \theta_j E_j$$

is the spectral decomposition of A. Let V be the Vandermonde matrix based on thre eigenvalues of A, thus $V_{i,j} = \theta_{j-1}^{i-1}$. Then

$$A^r \mathbf{1} = \sum_j \theta_j^r E_j \mathbf{1}$$

and therefore

$$W_X = \begin{pmatrix} E_1\mathbf{1} & \cdots & E_n\mathbf{1} \end{pmatrix} V^T.$$

Assume D is the $n \times n$ diagonal matrix with $D_{j,j} = \mathbf{1}^T E_j \mathbf{1}$. As the vectors $E_j\mathbf{1}$ and $E_k\mathbf{1}$ are orthogonal if $j \neq k$ and as

$$(E_j\mathbf{1})^T E_j \mathbf{1} = \mathbf{1}^T E_j^2 \mathbf{1} = \mathbf{1}^T E_j \mathbf{1},$$

we see that

$$W_X^T W_X = VDV^T.$$

One consequence of this is that

$$\det(W_X)^2 = \det(V)^2 \prod_{j=1}^{n} \mathbf{1}^T E_j \mathbf{1};$$

we point out that $\det(V)^2$ is the discriminant of the characteristic polynomial of X. So X is controllable if and only none of the vectors $E_j\mathbf{1}$ is zero.

14.1.1 Lemma. *If X is a controllable graph on n vertices with spectral idempotents $E_1, \ldots, E_n$, then the matrices $E_i J E_j$ for $1 \le i, j \le n$ form an orthogonal basis for the set of all $n \times n$ matrices.*

Proof. Using the trace inner product on matrices, we have

$$\langle E_i J E_j, E_k J E_\ell \rangle = \operatorname{tr}(E_j J E_i E_k J E_\ell) = \operatorname{tr}(E_\ell E_j J E_i E_k J)$$

and the last term is zero unless $j = \ell$ and $i = k$. In the latter case, the trace reduces to

$$\mathbf{1}^T E_i \mathbf{1}\mathbf{1}^T E_j \mathbf{1}. \qquad \square$$

A corollary of this is that if X is controllable if and only if A and J generate the algebra of all $n \times n$ matrices.

14.2 Cocospectral Controllable Graphs

Graphs X and Y are *cocospectral* if they are cospectral and their complements are cospectral. Haemers has conjectured that almost all graphs are determined by their spectrum. We are concerned here with a slightly weaker conjecture: almost all graphs are determined by their characteristic polynomials and the characteristic polynomials of their complements. Further, O'Rourke and Touri [1] have proved that almost all graphs are controllable. [1]

We say a graph X is determined by its *generalized spectrum* if it is determined by $\phi(X, t)$ and $\phi(\overline{X}, t)$, i.e., any graph cocospectral to X is isomorphic to it.

14.2.1 Lemma. *Suppose X and Y are cospectral graphs. Then their complements are cospectral if and only if, for each non-negative integer k, the number of walks in X of length k is equal to the number of walks of length k in Y.*

Proof. If $\mathscr{F}(X, t)$ is the generating function for walks in X and $n = |V(X)|$, then

$$t^{-1}\mathscr{F}(X, t^{-1}) = (-1)^n \left(\frac{\phi(\overline{X}, -t-1)}{\phi(X, t)} - 1 \right).$$

(For details see, e.g., the exercises to Chapter 4 in [2]). $\square$ [2]

This result is relevant because

$$(W_X^T W_X)_{i,j} = \mathbf{1}^T A^{i+j-2}\mathbf{1},$$

whence if graphs X and Y are cocospectral,

$$W_X^T W_X = W_Y^T W_Y.$$

We discuss the converse below.

Suppse X and Y are cospectral graphs. Then there is an orthogonal matrix Q such that $Q^T A_X Q = A_Y$ and, if X and Y are cocospectral, we may assume that $Q\mathbf{1} = \mathbf{1}$. As

$$Q^T A_X^k \mathbf{1} = A_Y^k Q^T \mathbf{1} = A_Y^k \mathbf{1}$$

it follows that

$$Q^T W_X = W_Y.$$

This leads to the following important result of Wang [3]: [3]

14.2.2 Theorem. *Assume X and Y are controllable graphs. If X and Y are cocospectral, there is a rational orthogonal matrix Q such that $Q\mathbf{1} = \mathbf{1}$ and $Q^T A_X Q = A_Y$.*

Proof. If X and Y are controllable and $Q^T W_X = W_Y$, then

$$Q^T = W_Y W_X^{-1}$$

and W_X^{-1} is rational. □

14.2.3 Lemma. *Suppose X is a controllable graph. Then X is cocospectral to Y if and only if $W_X^T W_X = W_Y^T W_Y$.*

Proof. By Lemma 14.2.1, we see that if X and Y are cocospectral, then their walk generating functions are equal. Consequently $W_X^T W_X = W_Y^T W_Y$.

For the converse, assume X is controllable and $W_X^T W_X = W_Y^T W_Y$. Then Y is controllable and $P = W_Y W_X^{-1}$ is orthogonal. As $W_X e_1 = \mathbf{1}$,

$$P\mathbf{1} = W_Y W_X^{-1}\mathbf{1} = W_Y e_1 = \mathbf{1},$$

and therefore X and Y are cocospectral. □

14.3 Irreducible Characteristic Polynomials

[This material is due to CG] If X is controllable and we lexicographically order the rows of W_X, the resulting matrix serves as a canonical label for X. If $a \in V(X)$, let $\mathscr{F}_a$ be the generating function

$$\sum_{n\geq 0} t^n e_a A^n \mathbf{1},$$

thus it is the generating function for the walks in X that start at a. It follows that the set of walk generating functions $\mathscr{F}_a$ for a in $V(X)$ determines X. From the exercises to Chapter 4 in [4]), the series $\mathscr{F}_a$ is determined by the four polynomials

$$\phi(X,t),\ \phi(\overline{X},t),\ \phi(X\setminus a,t),\ \phi(\overline{X}\setminus a,t).$$

Hence a controllable graph is determined by a set of $2n+2$ polynomials.

The next result is from [cg: control].

14.3.1 Lemma. *If the characteristic polynomial of X is irreducible, then X is controllable.*

Proof. We have $A = A_X$ acting on $\mathbb{R}^n$ and we observe that if there is a proper non-zero subspace U of $\mathbb{R}^n$ fixed by A, the characteristic polynomial of A restricted to U is a proper non-constant divisor of $\phi(X,t)$.

If X is not controllable, there is a non-zero vector a such that

$$\begin{pmatrix}\mathbf{1} & A\mathbf{1} & \cdots & A^{n-1}\mathbf{1}\end{pmatrix} a = 0$$

ans therefore there is a polynomial $\alpha(t)$ with degree at most $n-1$ such that $\alpha(A)\mathbf{1} = 0$. Now $\alpha(A) \neq 0$ (because $\phi(X,t)$ is the minimal polynomial of A and it is irreducible) and $\mathrm{rk}(\alpha(A)) \leq n-1$ (because it lies in $\mathbf{1}^{\perp}$). Hence the column space of $\alpha(A)$ is a non-zero proper A-invariant subspace of $\mathbb{R}^n$, and so A is not irreducible. □

A weak form of a converse to this is true: if X is controllable then there choices of z such that the characteristic polynomial of $A + zJ$ is irreducible. [Hilbert's irreducibility theorem.]

We use $\phi_{a,b}(X,t)$ to denote the (a,b)-entry of the adjugate of $tI - A$. From Section 4.1 of [5]), we have

$$\phi_{a,b}(X,t)^2 = \phi(X\setminus a,t)\phi(X\setminus b,t) - \phi(X,t)\phi(X\setminus\{a,b\},t)).$$

We use this to show that if $\phi(X,t)$ is irreducible, then given $\phi(X\setminus a,t)$ and $\phi(X\setminus b,t)$, we can determine if a and b are adjacent. Hence if $\phi(X,t)$ is irreducible, X is determined by the polynomials $\phi(X\setminus a)$ for $a \in V(X)$.

To prove the claim, suppose that there are polynomials f and g with degree at most $n-2$ such that

$$f(t)^2 = \phi(X\setminus a,t)\phi(X\setminus b,t) - \phi(X,t)g(t)$$

Subtracting this from the earlier equation yields that

$$\phi_{a,b}(X,t)^2 - f^2 = \phi(X,t)(g(t) - \phi(X\setminus\{a,b\},t)).$$

The left side is the product of two polynomials of degree at most $n-2$, and it is divisible by the irreducible polynomial $\phi(X,t)$ of degree n. Therefore the left side is zero and so $g = \phi(X\setminus\{a,b\},t)$. Since we can read off the

number of edges of a graph from its characteristic polynomial, our claim is proved. (The argument we've just used is essentially the argument Tutte use to prove that a graph whose characteristic polynomial is irreducible is vertex reconstructible.)

14.4 Elementary Divisors

Since W_X in an integer matrix, there are unimodular integer matrices L and R and a diagonal integer matrix D such that $W = LDR$; further $D_{i-1,i-1}$ divides $D_{i,i}$ for $i = 2,\dots n$. (So D is the Smith normal form of W_X, and the diagonal entries of D are its elementary divisors.) We note that for a prime p, we have $\mathrm{rk}_p(W_X) = \mathrm{rk}_p(D)$.

We have the following pretty result due to Wang.

14.4.1 Lemma. *Let X be a graph on n vertices with walk matrix W. If n is even, then all entries of W^TW are even; if n is odd, then $(W^TW)_{1,1}$ is odd and all other entries are even.*

Proof. First,

$$(W^TW)_{i,j} = \mathbf{1}^TA^{i+j-2}\mathbf{1},$$

thus this entry is the number of walks in X of length $i+j-2$. If u and v are distinct vertices in X, the number of walks of length k for u to v is equal to the number of walks of length k from v to u.

Hence if $k > 0$, the number of walks of length k in X has the same parity as the number of closed walks of length k, which is equal to $\mathrm{tr}(A^k)$. Now

$$\mathrm{tr}(A^k) = \mathrm{tr}(AA^{k-1}) = \mathrm{sum}(A \circ A^{k-1})$$

and we can view $A \circ A^{k-1}$ as a weighted graph, whence $\mathrm{sum}(A \circ A^{k-1})$ is the sum of the valencies of the vertices of the weighted graph, and is therefore even. □

14.4.2 Corollary. *If X is a graph on n vertices with walk matrix W, then $\mathrm{rk}_2(W)$ is at most $\lfloor \frac{n+1}{2} \rfloor$.*

In terms of the elementary divisors we may express this result by stating that if $2j \le n+1$, then d_j is odd.

If we look at the elementary divisors of graphs, we find that the Smith normal form is in general quite restricted. If $j = \lfloor (n+1)/2 \rfloor$, then very often

(a) $d_1 = \cdots = d_j = 1$.

(b) $d_{j+1} = \cdots = d_{n-1} = 2$

(c) $d_n/2$ is odd.

Thus if p is an odd prime, $\mathrm{rk}_p(W) \geq n-1$ and 2^j is the larget power of two that divides $\det(W)$. (It follows from the last claim that $\mathrm{rk}_2(W)$ meets its lower bound).

We struss that above results are not theorems (or even lemmas), they simply summarize what we observe on graphs up to 8 vertices, and on random graphs with larger sets of vertices.

14.5 Level Two

For graphs X and Y, we define the set of orthogonal matrices $\mathcal{Q} = \mathcal{Q}_{X,Y}$ by

$$\mathcal{Q} = \{Q \in O(n,\mathbb{Q}) : Q\mathbf{1} = \mathbf{1},\ Q^T A_X Q = A_Y\}.$$

If Q is a rational orthogonal matrix, its *level* is the least positive integer ℓ such that ℓQ is an integer matrix. A rational orthogonal matrix of level one is a monomial matrix.

As we saw, if X is controllable and cocospectral with Y, then $Q = A_Y A_X^{-1}$ is orthogonal and rational and $Q\mathbf{1} = \mathbf{1}$. It follows that if the only rational orthogonal matrices Q such that $Q^T A_X Q$ is integral are those of level one, then X is determined by its spectrum and the spectrum of $\overline{X}$.

14.5.1 Lemma. *Assume X is controllable and let $d_1, \ldots, d_n$ be the elementary divisors of W_X. Then the level of any matrix in $\mathcal{Q}_{X,Y}$ divides d_n.*

Proof. Assume Q has level ℓ and let D be the Smith normal form of W_X. There are unimodular matrices L and R such that $W_X = LDR$, from which we see that $d_n W_X^{-1}$ is an integer matrix.

If $Q^T A_X Q = A_Y$, then $Q^T = W_Y W_X^{-1}$ and therefore ℓQ^T is integral. □

We can use the next result to show that in many cases all matrices in $\mathcal{Q}$ havel at most two.

14.5.2 Theorem. *Assume X is controllable, Q is rational and orthogonal with level ℓ, and $Q^T A_X Q = A_Y$. If p is an odd prime and p divides ℓ, then there is vector z in $GF(p)^n$ such that $z^T W_X \equiv 0$ (mod p) and $z^T z \equiv 0$ (mod p).*

Proof. From the definition of level, there is an index r such that $x^T = \ell e_r^T Q$ is not zero mod p. Then $x^T x = \ell^2$, which is zero mod p.

As QW_X is an integral matrix (it is equal to W_Y), it follows that $e_r^T QW_X$ is integral and accordingly

$$x^T W_X = \ell e_r^T QW_X \equiv 0 \mod p.$$ □

As we noted earlier, for many graphs the elementary divisors $d_1, \ldots, d_{n-1}$ are even. Hence if p is an odd prime that divides d_n, then $\mathrm{rk}_p(W_X) = n-1$ and therefore the kernel of W_X^T over $GF(p)$ is 1-dimensional. If a vector x spans this kernel and $x^T x \not\equiv 0 \bmod p$, then we deduce that p does not divide the level of Q.

THe results presented up to this point come from [wang,xu:suficient].

14.6 Rational Orthogonal Matrices with Level Two

Suppose Q is rational and orthogonal with level two. The $2Q$ is an integer matrix, and it follows that each row has either one non-zero entry, necessarily ± 1, or it has exaxctly four non-zero entries, each equal to $\pm 1/2$.

We focus on indecoposable matrices of index two. So each row will have four non-zero entries, and two rows overlap in zero, two or four positions.

In [wang-xu: excluding], the orthogonal matrices of index two are completely determined. There are two 'sporadic' examples, of order 4×4 and 8×8 and, depending on what is considered different, one or two infinite families.

14.7 Generalized, Sufficient

We turn to two papers by Wang [wang: generalized, sufficient]. The result of interest to us is as follows: if $\mathscr{G}_n$ denotes the set of controllable graphs on n vertices such that $2^{-\lfloor n/2 \rfloor} \det(W_X)$ is square-free and odd, then each graph in $\mathscr{G}_n$ is determined by its generalized spectrum.

In this section we work on the following theorem, from [wang: generalized]

14.7.1 Theorem. *Suppose X is a controllable graph on n vertices such that*

$$2^{-\lfloor n/2 \rfloor} \det(W_X)$$

is square-free and odd. If p is an odd prime such that p divides $\det(W_X)$ *but p^2 does not, then p does not divide the level of any matrix in $\mathscr{Q}$.*

Equivalently we prove that any matrix in $\mathscr{Q}$ has level one or two.

14.7.2 Lemma. *Let A be the adjacency matrix of a controllable graph on n vertices. If p is a prime such that* $\mathrm{rk}_p(W) = n-1$*, then for any scalar λ, we have* $\mathrm{rk}_p(A - \lambda I) \geq n-2$.

Proof. Let C_ϕ be the companion matrix of the characteristic polynomial of A. If we delete the first row and last column from C_ϕ, the resulting matrix is I_{n-1}. Also $(C_\phi)1, n = \pm \det(A)$. Hence

$$\mathrm{rk}_p(C_\phi - \lambda I) = \begin{cases} n-1, & \text{if } p \mid \det(A); \\ n, & \text{if } p \nmid \det(A). \end{cases}$$

The companion matrix is relevant because

$$(A - \lambda I) W_X = W_X (C_\phi - I).$$

Over $GF(p)$, the rank of the right side is at least $n-2$ and therefore the dimension of the left kernel of $A - \lambda I$ is at most 2. □

14.7.3 Lemma. *If* $\mathrm{rk}_p(W_X) = n-1$, *then* $\ker(W_X^T)$ *is an eigenspace for* A.

Proof. Let $W^\#$ denote the adjugate of W. Then $W^\# W = \det(W) I$ and if $p \mid \det(W)$, we have $W^\# W = 0$. If $\mathrm{rk}_p(W) = n-1$ then, over $GF(p)$, some minor of W order $(n-1)\times(n-1)$ is. not zero, and hence $W^\#$ is not zero over $GF(p)$. Since $\mathrm{rk}(W^\#) = 1$, there are non-zero vectors y and z such that

$$W^\# = yz^T.$$

It follows that $z^T W = 0$. Now

$$0 = z^T W C_\phi = z^T A W$$

which implies that $z^T A$ is a scalar multiple of z^T. □

If $u \in \mathbb{R}^n$, we define $W_X(u)$ to be the matrix

$$\begin{pmatrix} u & Au & \cdots & A^{n-1}u \end{pmatrix};$$

thus $W_X = W_X(\mathbf{1})$.

We assume henceforth that over $GF(p)$ we have $z^T W = 0$ and $Az = \lambda_0 z$. We are going to prove that if p divides the level ℓ and $\mathrm{rk}_p(W) = n-1$, then $p^2 \mid d_n$. As we saw. $\mathrm{rk}_p(A - \lambda_0 I)$ is $n-1$ or $n-2$. We treat these two cases separately.

Case 1: $\mathrm{rk}_p(A - \lambda_0 I) = n-1$

Claim: *(a)* $\mathbf{1} \in \mathrm{col}(A - \lambda_0 I)$
Since $z^T W = 0$ we have $z^T \mathbf{1} = 0$, whence $\mathbf{1} \in z^\perp$. But $z^\perp$ is equal to the row space of $A - \lambda_0 I$ and, since A is symmetric, the claim follows.

Claim:

Claim: *(c)* $z \in \mathrm{col}(A - \lambda_0 I)$
Since $z^T z = 0$ we see that $z \in z^\perp$ and, from the proof of Claim (a), $z^\perp = \mathrm{col}(A - \lambda_0 I)$.

14.8 Isomorphism of Controllable Graphs

Let X be a graph on n vertices with adjacency matrix A. Let σ be the characteristic vector of a non-empty subset S of $V(G)$. We say that the pair (X, β) is *controllable* if the matrix

$$W_X(\sigma) = \begin{pmatrix} \sigma & A\sigma & \cdots & A^{n-1}\sigma \end{pmatrix}$$

is invertible. The caes of most interest to us are when $S = V(X)$ (and σ is the all-ones vertor $\mathbf{1}$), and when S is a vertex. If $\sigma = \mathbf{1}$ then we abbreviate $W_X(\mathbf{1})$ to W_X, and call it the *walk matrix* of X. In this case we say simply that X is controllable.

We oberve that the column space of $W_X(\sigma)$ is the smallest subspace of $\mathbb{R}^n$ that contains σ and is invariant under the matrices A and $\sigma\sigma^T$. If $\sigma = \mathbf{1}$, then $\sigma\sigma^T = J$ and $\mathrm{col}(W_X)$ is invariant under A and J. A subspace is invariant under A and J if and only if it is invariant under $J - I - A$ and J, whence we see that a graph is controllable if and only if its complement $\overline{X}$ is controllable.

We see that the ui-entry of $W_X(\sigma)$ is the number of walks of length i in X that start at a vertex in S and end at u.

Clearly $W_X(\sigma)$ is not invertible if two of its rows are equal. Hence if there is a non-identity automorphism of X that fixes S, then (X,σ) is not controllable.

If X is a controllable graph, we say that W_X is *normalized* if its rows form a lexicopgraphically increasing sequence of vectors in $\mathbb{Z}^n$. (Actually you may choose any total ordering of elements of $\mathbb{Z}^n$ you like.) Then if X and Y are two controllable graphs on n vertices, then X and Y are isomorphic if and only if the normalized walk matrices of X and Y are equal.

O'Rourke and Touri [6] proved my conjecture that a random graph is controllable. Hence we have a Las Vegas algorithm for graph isomorphism: compute $\det(W_X)$, if this is not zero then the normalized walk matrix is a canonical label; if it is zero, report failure.

[6]

14.9 Resultants

We use $\mathrm{Res}(f,g)$ to denote the resultant of two polymomials f and g.

14.9.1 Theorem. *If $|V(X)| = n$ and W is the walk matrix of X, then*

$$(-1)^n \mathrm{Res}(\phi(\overline{X}, -t-1), \phi(X,t)) = \det(W)^2$$

Proof. Suppose $A = A(X)$ has spectral decomposition

$$A = \sum_{r=1}^{m} \theta_r E_r$$

and let V be the Vandermonde matrix of the eigenvalues of A, with i-th row equal to

$$\begin{pmatrix} \theta_1^{i-1} & \cdots & \theta_n^{i-1} \end{pmatrix}.$$

Then

$$A^k\mathbf{1} = \sum_r \theta_r^k E_r \mathbf{1}$$

and so if M is the matrix

$$M = \begin{pmatrix} E_1\mathbf{1} & \cdots & E_m\mathbf{1} \end{pmatrix},$$

then

$$W_X = MV^T.$$

One consequence of this is that if X is controllable, then the eigenvalues of A are all simple. We also observe that the columns of M are pairwise orthogonal and so if Δ is the diagonal matrix with ii-entry $\mathbf{1}^T E_i \mathbf{1}$, then $M^T M = \Delta$ and

$$W_X^T W_X = V M^T M V^T = V \Delta V^T.$$

For later use, note that this implies that

$$\det(W_X)^2 = \det(V)^2 \det(\Delta). \tag{14.9.1}$$

Next,

$$\begin{aligned}\phi(\overline{X}, t) = \det(tI - (J - I - A)) &= \det((t+1)I + A)\det(I - ((t+1)I + A))^{-1} J) \\ &= \det((t+1)I + A)\,(1 - \mathbf{1}^T((t+1)I + A))^{-1}\mathbf{1})\end{aligned}$$

and consequently

$$\begin{aligned}(-1)^n \frac{\phi(\overline{X}, -t-1)}{\phi(tI - A)} &= (1 - \mathbf{1}^T((-tI + A))^{-1}\mathbf{1}) \\ &= 1 + \sum_r \frac{\mathbf{1}^T E_r \mathbf{1}}{t - \theta_r}.\end{aligned}$$

If we multiply both sides of this identity by $t - \theta_s$ and then let $t \to \theta_s$, we deduce that

$$\phi(\overline{X}, -\theta_s - 1) = (-1)^n (\mathbf{1}^T E_r \mathbf{1})\, \phi'(X, \theta_s I)$$

and therefore

$$\mathrm{Res}(\phi(\overline{X}, -t-1), \phi(X, t)) = (-1)^{n^2} \det(\Delta) \prod_s \phi'(X, \theta_s).$$

Here the final product is the discriminant of $\phi(X, t)$, which is equal (up to sign), to $\det(V)^2$. Using Equation (14.9.1), we find that

$$\mathrm{Res}(\phi(\overline{X}, -t-1), \phi(X, t)) = (-1)^n \det(W)^2. \qquad \square$$

14.9.2 Corollary. *The graph X is controllable if and only if $\phi(X, t)$ and $\phi(\overline{X}, -t-1)$ are coprime.* $\square$

Anothe proof of this is given in [7].

We point out that $(W^T W)_{i-1, j-1}$ counts the number of walks in X of length $i + j - 2$.

14.10 Controllable Vertices

Suppose a is a vertex in X with characteristic vector e_a and let W_a denote the walk matrix relative to a. The vertex u is controllable if W_a is invertible.

14.10.1 Theorem. *If W_a is the walk matrix of the vertex a in X, then $\det(W)^2$ is equal (up to sign) to the resultant of $\phi(X, t)$ and $\phi(X \setminus a, t)$.*

Proof. Let V be the Vandermonde matrix of the eigenvalues of A and let M be the matrix with the vectors $E_r e_a$ as its columns. Let Δ be the diagonal matrix with i-th entry equal to $e_a^T E_r e_a = (E_r)_{a,a}$. Then

$$W_a = MV^T$$

and

$$W_a^T W_a = V\Delta V^T.$$

As before, $\det(V)^2$ is the discriminant of $\phi(X,t)$. We also see that $\det(W_a) = 0$ if the eigenvalues of A are not all simple.

We have

$$(E_r)_{a,a} = \frac{\phi(X\setminus a, \theta_r)}{\phi'(X,\theta_r)}$$

and accordingly

$$\det(W_a)^2 = \mathrm{Res}(\phi(X\setminus a,t)), \phi(X,t)).$$ □

This result implies that a is controllable in X if and only if $\phi(X\setminus a,t)$ and $\phi(X,t)$ are coprime.

As an exercise, the reader may like to prove that if G is the $n\times n$ matrix with ij-entry $\mathrm{tr}(A^{i+j-2})$, then $\det(G)$ is equal to the discriminant of $\phi(X,t)$. It may help to note that G is the Gram matrix of the set of powers

$$\{I, A, \ldots, A^{n-1}\}$$

relative to the trace inner product.

15
Computations

We discuss some computational techniques.

15.1 Rules of Numerical Linear Algebra

(i) Never write your own code to solve standard problems (linear equations, eigenvalues and singular values, least squares…).

(ii) Never compute the inverse of a matrix (unless it is unitary or orthogonal, perhaps triangular).

(iii) Never multiply two matrices together.

(iv) Don't compute the rank of a real or complex matrix. (Use singular values.)

15.2 Spectral Decomposition of Normal Matrices

In working with discrete quantum walks, we sometimes need to compute the spectral decomposition of a unitary matrix. We have good routines for computing the eigenvalues and eigenvectors of Hermitian matrices; our goal is to reduce to this case.

A matrix M is normal if M and M^* commute. In this case the algebra generated by M and M^* is commutative and closed under conjugate-transpose, therefore it is semisimple and there is an orthogonal change of basis that diagonalizes it.

We define the real and imaginary parts of the matrix M to be the matrices

$$A = \frac{1}{2}(M + M^*), \qquad B = \frac{1}{2i}(M - M^*).$$

We have $M = A + iC$, as it should. We also see that A and B are Hermitian[1] and, since M is normal, A and B commute. Since A and B commute, each eigenspace of A is B-invariant, and accordingly each eigenspace W is a direct of eigenspaces of the restriction of B to W. (We will be lazy and refer to them as eigenspaces of B.)

[1] $M - M^*$ is skew Hermitian

Since A is Hermitian, standard code will return a list of eigenvalues and an orthonormal basis of eigenvectors. We then write code that takes this list and returns a list of simple eigenvalues along with an orthonormal basis for each eigenspace of A.

Our problem now is to split each eigenspace of A into eigenspaces for $M - M^*$.[2]. Assume that U is a matrix whose columns form an orthonormal basis for the λ-eigenspace. The matrix that represents the restriction of $M - M^*$ to the eigenspace is $UU^*(M - M^*)UU^*$. So the procedure is:

[2] i.e., eigenspaces for the restriction of $M - M^*$

(a) Compute $U^*(M - M^*)U$.

(b) Compute the spectral decomposition of $U^*(M - M^*)U$, getting idempotents $F_1, \ldots, F_m$.

(c) Return the idempotents UF_rU^* for $r = 1, \ldots, m$.

We note that this breaks Rule (iii) multiple times.[3]

[3] It's an imperfect world

We remark that the python code that computes the spectral idempotents from the bases for the eigenspaces is about 10 times slower than the library routine that computes the eigenvalues and eigenvectors.

15.3 Spectral Decomposition of Unitary Matrices

If our matrices are unitary, not just normal, we can remove some warts. We suppose M is unitary $n \times n$ with eigenvalues

$$e^{i\theta_1}, \ldots, e^{i\theta_n}.$$

The eigenvalues of $\frac{1}{2}(M + M^*)$ are then

$$\cos(\theta_r), \quad (r = 1, \ldots, n)$$

and the eigenvalues of $\frac{1}{2i}(M - M^*)$ are

$$\sin(\theta_r), \quad (r = 1, \ldots, n).$$

So if the columns of U are an orthonormal basis for the λ-eigenspace of $M + M^*$, the eigenvalues of $\frac{1}{2i}(M - M^*)$ are $\pm\sqrt{1 - \lambda^2}$.

Let B denote $\frac{1}{2i}(M - M^*)$ and assume that the columns of U are an orthonormal basis for the eigenspace of $\frac{1}{2}(M + M^*)$ with eigenvalue $\cos(\theta_r)$. Then $B^2 = \sin^2(\theta_r)I$ on $\mathrm{col}(U)$ and therefore

$$(B - \sin(\theta_r)I)(B + \sin(\theta_r)) = 0.$$

So the matrices

$$\frac{1}{\sin(\theta_r)}(B - \sin(\theta_r)I), \quad \frac{1}{\sin(\theta_r)}(B + \sin(\theta_r)I)$$

are pairwise orthogonal idempotents. If we denote them by $B_r(+)$ and $B_r(-)$ respectively and let F_r be the spectral idempotent of $M + M^*$ associated to the eigenvalue θ_r. Then the matrices $F_rB_r(+)F_r$ and $F_rB_r(-)F_r$ are

spectral idempotents for M. (One of these matrices might be zero, in which case we ignore it.)

To summarize, for unitary matrices, we can drop step (b) from our procedure for normal matrices.

The normal matrices that we want to diagonalize are those coming from dicrete quantum walks since, for discrete walks, we need only diagonalize the Hamiltonian. In this case unitary matrices we deal with are indexed by the arcs of a graph X, hence they are large. It it quite possible that, for interesting classes of discrete walks,[4] there is an algorithm that works with matrices of order $|V(X)| \times |V(X)|$.[5]

[4] two reflections? bipartite?

[5] Exercise!

15.4 Singular Values

We will be peasants and define the *singular values* of a complex matrix M to be the eigenvalues of MM^* (or M^*M, whichever is more convenient). The problems with this definition are that it does not indicate why singular values might be useful, and the implied algorithm requires us to break Rule (iii).

We restrict ourselves to real matrices, and start by considering a simply stated problem: given an $m \times n$ matrix M, find the rank-1 matrix xy^T such that $\|M - xy^T\|$ is minimal. (Note that if we can solve this problem. the obvious next step is to determine the best rank-1 approximation to $M - xy^T$ and leads us to a more general problem: find the best rank-k approximation to M.)

Let A be an $m \times n$ matrix over $\mathbb{C}$. We start by looking for the rank-1 matrix M such that $\|A - M\|$ is minimal. It is convenient to express M as λxy^*, where x and y are unit vectors. Now

$$\begin{aligned}\langle A - \lambda xy^*, A - \lambda xy^*\rangle &= \langle A, A\rangle - \overline{\lambda}\langle xy^*, A\rangle - \lambda\langle A, xy^*\rangle + \lambda\overline{\lambda}\langle xy^*, xy^*\rangle \\ &= \langle A, A\rangle - \overline{\lambda}x^*Ay - \lambda y^*A^*x + \lambda\overline{\lambda} \\ &= \langle A, A\rangle - x^*Ay\,y^*A^*x + (\lambda - x^*Ay)(\overline{\lambda} - y^*Ax) \\ &= \langle A, A\rangle - |x^*Ay|^2 + |\lambda - x^*Ay|^2,\end{aligned}$$

from which we see that, given x and y, we should take $\lambda = x^*Ay$. By Cauchy-Schwarz

$$|x^*Ay|^2 \le \langle x, x\rangle\langle Ay, Ay\rangle = \langle Ay, Ay\rangle$$

and equality holds if and only if Ay is a scalar multiple of x. In other words, we want

$$x = \frac{1}{\|Ay\|}Ay$$

and then

$$\lambda = x^*Ay = \frac{1}{\|Ay\|}y^*A^*Ay = \|Ay\|.$$

If θ is the largest eigenvalue of A^*A, it follows that we get an optimal rank-1 approximation to A by choosing an eigenvector y for A^*A with eigenvalue θ and norm 1 and taking the approximation to be Ayy^*.

Observe that

$$AA^*Ay = \theta Ay$$

and therefore Ay is an eigenvector for AA^* with eigenvalue θ.

Consider the difference $A - Ayy^*$. Since $Ay \neq 0$ and $(A - Ayy^*)y = 0$ and since the $\mathrm{col}(A - Ayy^*)$ is a subspace of $\mathrm{col}(A)$, we see that $\mathrm{rk}(A - Ayy^*) = \mathrm{rk}(A) - 1$.

16
Spectra of Infinite Graphs

We consider spectral questions on infinite graphs with bounded valency.

16.1 Some Sources

Books:

1. Peter Lax. "Functional Analysis".
2. Gerhard Teschl. "Mathematical Methods in Quantum Mechanics".
3. Brian C. Hall. "Quantum Theory for Mathematicians".
4. E. B. Davies. "Spectral Theory and Differential Operators".
5. E. B. Davies. "Linear Operators and their Spectra". [1]
6. David Borthwick. "Spectral Theory".
7. Angus E. Taylor. "Introduction to Functional Analysis".[2]

[1] https://www.maths.ed.ac.uk/~v1ranick/papers/davies.pdf.

[2] Can be found on line.

Papers:

1. Godsil and Mohar. Walk generating functions and spectral measures of infinite graphs. LAA **107** (1988), 191–206.
2. B. Mohar. The spectrum of an infinite graph. LAA **48** (1982), 245–256.

Neither of these lists is exhaustive; these are just sources I've looked at in compiling these notes.

My main aim is to provide a self-contained treatment of the material from the two papers list above.

16.2 Banach Spaces

A *Banach space* is a complete normed vector space, real or complex. Typical examples are finite-dimensional vector spaces, spaces of sequences,

and spaces of continuous or measurable functions. If the norm comes from an inner product, we have a *Hilbert space*.

Our concern is with linear operators on Banach spaces. Our standard example will be the adjacency matrix A of a graph X acting on the space of real functions on $V(X)$. (If $f \in \mathbb{R}^{V(X)}$, then the value $(Af)(u)$ of Af at u is $\sum_{v\sim u} f(v)$.) The bounded functions on $V(X)$ form a Banach space, as do the square-summable functions. We denote these spaces by $\ell_\infty(X)$ and $\ell_2(X)$.

If A is a linear operator on a Banach space $\mathscr{B}$, the *operator norm* of A is

$$\sup_{\|u\|=1} \|Au\|.$$

If this is finite, we say that A is *bounded*. The adjacency operator is bounded on $\ell_\infty(X)$ if and only if X is bounded, that is, there is an upper bound on the valency of a vertex.

16.2.1 Theorem. *A linear operator on a Banach space is bounded if and only if it is continuous.* □

16.3 Unbounded Operators

For the purposes of quantum physics, we need to adjust our set-up. Assume $\mathscr{H}$ is a Hilbert space and let $\mathfrak{D}$ be a dense subspace of $\mathscr{H}$. Then a *linear operator* is a linear mapping from $\mathfrak{D}$ to $\mathscr{H}$. The subspace $\mathfrak{D}$ is the *domain* of A, and is usually denoted by $\mathfrak{D}(A)$.

In this context, a linear operator A is *bounded* if

$$\sup_{v\in\mathfrak{D}(A)\setminus 0} \frac{\|Av\|}{\|v\|} < \infty.$$

otherewise it is *unbounded*. If A is bounded, there is a unique extension to a bounded operator on $\mathscr{H}$.[3] The operators arising in quantum physics are commonly unbounded.

[3] see, e.g., Teschl, Theorem 0.29

We turn to a combinatorially interesting example. The *adjacency operator* $A(X)$ of a graph X acts on the space of real functions on $V(X)$. If $f \in \mathbb{R}^{V(X)}$ and $u \in V(X)$,

$$(Af)(u) := \sum_{v\sim u} f(v).$$

We can form three Banach spaces from $\mathbb{R}^{V(X)}$, using the norms $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$. These are, respectively, the real sequences that are summable, square summable, and bounded. The square-summable sequences form a Hilbert space, which will be our preferred situation.

A graph is *locally finite* if each vertex has finite valency. (All the graphs we use will be locally finite.) A graph is *bounded* if there is an integer K such that every vertex has valency at most K. If $\mathbf{1}$ is the function taking value 1 on each vertex of X, then $A\mathbf{1}$ is bounded if and only if X is bounded.

The space $c\S$ of real sequences with finite support is a dense subspace of the Hilbert space $\ell_2(\mathbb{R})$. If A is the adjacency operator of a locally finite graph and $f \in \mathscr{S}$, then $Af \in \mathscr{S}$. Hence A is an example of an unbounded operator.

It is not hard to show that if f is summable and X is bounded, then Af is summable. Thus the adjacency operator of a bounded graph is bounded under both the $\|\cdot\|_1$ and $\|\cdot\|_\infty$. But what about $\|A\|_2$? It is easy to show that if $\|A\|_2 < \infty$, then X is bounded. The converse requires more effort.

It is a theorem [4] that if $A \in \ell_1(\mathbb{R}) \cap \ell_\infty(\mathbb{R})$, then $A \in \ell_2(\mathbb{R})$ and

[4] Exercise 1 on page 224 of Taylor

$$\|A\|_2^2 \leq \|A\|_1 \|A\|_\infty.$$

This implies that the adjacency operator of a bounded graph has finite norm on $\ell_2(\mathbb{R})$. We offer a proof of a slightly less precise result due to Schur.

16.3.1 Theorem. *The adjacency operator of a graph is a bounded operator on $\ell_2(\mathbb{R})$ if and only if X is bounded.*

Proof. It is, as noted early, easy to see that if $\|A(X)\| < \infty$, then X is bounded. We turn to the converse.[5] We assume that M is an upper bound on the valency of a vertex in X and show that $\|A\|_2 \leq M$.

[5] following Taylor, Theorem 6.12-A

Part II

Other Fields

17
Adjacency Matrices over $GF(p)$

17.1 Similarity

Two matrices A and B over a ring R as *equivalent* if there are matrices P and Q over R with determinant one such that

$$PAQ = B.$$

Note that A and B must have the same order, but need not be square. We are taught that two matrices over $\mathbb{Q}$ are equivalent if and only if they have the same reduced row-echelon form.

17.1.1 Theorem. *Matrices A and B over $\mathbb{F}$ are similar if and only if $tI - A$ and $tI - B$ are equivalent over $\mathbb{F}[t]$.* □

It might seem that this theorem reduces a hard problem to a harder problem. However:

17.1.2 Theorem. *Matrices A and B over a principal ideal domain are equivalent if and only if they have the same Smith normal form.* □

The principal ideal domains we will deal with are $\mathbb{Z}$ and $\mathbb{F}[t]$.

Let A be an $m \times n$ matrix over the principal ideal domain R. Define $\psi_r(A)$ to be the gcd of the $r \times r$ submatrices of A. If $r < \min\{m, n\}$, then ψ_r divides ψ_{r+1}. The ratios ψ_{r+1}/ψ_r are the *elementary divisors* of A.

17.1.3 Theorem. *If R is a principal ideal domain, then two matrices are equivalent over R if and only if they have the same elementary divisors.* □

The elementary divisors determine, and are determined, by the Smith normal form, so there is little difference between the previous two thorems.

If R is a field and A is $n \times n$, then $\psi_n(A)$ is the characteristic polynomial of A and $\psi_{n-1}(A)$ is its minimal polynomial.

If M and N are equivalent matrices over $\mathbb{Z}$, then for any prime p,

$$\mathrm{rk}_p(M) = \mathrm{rk}_p(N).$$

17.1.4 Theorem. *Let A and B be square matrices over $\mathbb{F}$. Then A and B are similar if and only if they have same rational normal form.*[1]

[1] aka Frobenius normal form

Part III

Association Schemes

18
A Tensor Identity

We use $A \otimes B$ to denote the Kronecker product of two matrices A and B.

18.1 Seidel's Identity

We offer an exalted version of Seidel's identity, due to Koppinen.

18.1.1 Theorem. *Let $\mathscr{A}$ be an association scheme with d classes. Then*

$$\sum_{i=0}^{d} \frac{1}{v v_i} A_i \otimes A_i^T = \sum_{i=0}^{d} \frac{1}{m_i} E_i \otimes E_i.$$

Proof. Suppose that V is an inner product space and $u_1, \ldots, u_k$ and $v_1, \ldots, v_k$ are two orthogonal bases for a subspace U of V. If

$$R = \sum_{i=1}^{k} \frac{1}{\langle u_i, u_i \rangle} u_i u_i^*$$

and

$$S = \sum_{i=1}^{k} \frac{1}{\langle v_i, v_i \rangle} v_i v_i^*,$$

and $x \in V$, then Rx and Sx are both the orthogonal projection of x onto U. So $Rx = Sx$ for all x and therefore $R = S$. Since

$$x y^* = x \otimes y^*,$$

we thus have

$$\sum_{i=1}^{k} \frac{1}{\langle u_i, u_i \rangle} u_i \otimes u_i^* = \sum_{i=1}^{k} \frac{1}{\langle v_i, v_i \rangle} v_i \otimes v_i^*. \tag{18.1.1}$$

Now let $\mathrm{vec} : \mathrm{Mat}_{m \times n}(\mathbb{C}) \to \mathbb{C}^{mn}$ be the linear map given by

$$\mathrm{vec}(A) = \begin{pmatrix} Ae_1 \\ \vdots \\ Ae_n \end{pmatrix}.$$

If $M \in \mathrm{Mat}_{n \times n}(\mathbb{C})$, let $M^{\#}$ denote the linear map from $\mathrm{Mat}_{n \times n}(\mathbb{C})$ to $\mathbb{C}$ given by

$$M^{\#}(X) := \mathrm{tr}(M^* X).$$

Note that

$$M^{\#}(X) = \mathrm{vec}(M)^* \mathrm{vec}(X).$$

Then (18.1.1) yields that

$$\sum_{i=0}^{d} \frac{1}{v v_i} A_i \otimes A_i^{\#} = \sum_{i=0}^{d} \frac{1}{m_i} E_i \otimes E_i^{\#}.$$

Consequently

$$\sum_{i=0}^{d} \frac{1}{v v_i} A_i \otimes \mathrm{vec}(A_i)^T = \sum_{i=0}^{d} \frac{1}{m_i} E_i \otimes \mathrm{vec}(\overline{E}_i)^T$$

and therefore

$$\sum_{i=0}^{d} \frac{1}{v v_i} A_i \otimes A_i = \sum_{i=0}^{d} \frac{1}{m_i} E_i \otimes \overline{E}_i.$$

Let I denote the identity map on $\mathrm{Mat}_{v\times v}(\mathbb{C})$ and τ the transpose map. If we apply $I \otimes \tau$ to both sides of this identity, the result follows. □

We let $\mathscr{K}$ denote either of the two sums in the statement of 18.1.1. Since $E_j \otimes E_j$ is self-adjoint, we have $\mathscr{K}^* = \mathscr{K}$ and therefore we also have

$$\mathscr{K} = \sum_{i=0}^{d} \frac{1}{v v_i} A_i^T \otimes A_i.$$

18.2 Applications

We present three applications of our tensor identity.

First, suppose $X \in \mathrm{Mat}_{v\times v}(\mathbb{C})$ and $T : \mathbb{C}[\mathscr{A}] \otimes \mathbb{C}[\mathscr{A}] \to \mathbb{C}[\mathscr{A}]$ is the linear mappping given by

$$T(C \otimes D) = \mathrm{tr}(DX)C.$$

Therefore

$$T(\mathscr{K}) = \sum_{i=0}^{d} \frac{1}{v v_i} \mathrm{tr}(A_i^T X) A_i = \sum_{i=0}^{d} \frac{1}{m_i} \mathrm{tr}(E_i X) E_i.$$

This shows that **??** is a consequence of 18.1.1.

An association scheme $\mathscr{A}$ with d classes is *pseudocyclic* if its valencies $v_1, \ldots, v_d$ are all equal and its multiplicities m_i are all equal. If we denote the common value of these parameters by m, then $v = dm + 1$. Koppinen's identity yields that

$$\mathscr{K} = \frac{1}{v} I + \frac{1}{vm} \sum_{i=1}^{d} A_i^{\otimes 2} = E_0 + \frac{1}{m} \sum_{i=1}^{d} E_i^{\otimes 2}.$$

Here

$$\sum_{i=1}^{d} A_i^{\otimes 2}$$

is the adjacency matrix of a regular graph. The previous equality shows that it has exactly three eigenvalues ($vm - m$, $v - m$ and $-m$), and therefore it is the adjacency matrix of a strongly regular graph.

The simplest example of a pseudocyclic scheme is the scheme with d classes associated to the odd cycle C_{2d+1}. (In this case the strongly regular graph is $L(K_{2d+1,2d+1})$.)

We offer another proof of the inequality (**??**).

18.2.1 Theorem. *Let $\mathscr{A}$ be an association scheme with d classes on v vertices and let R be a subset of $\{1,\dots,d\}$. If C is an R-clique and D is an R-coclique, then $|C||D| \le v$.*

Proof. Let C be an R-clique and D an R-coclique, with characteristic vectors y and z respectively. Let S be the subset $C \times D$ of $V \times V$, with characteristic vector x. Then $x = y \otimes z$ and

$$x^T(A_i \otimes A_i)x = y^T A_i y\, z^T A_i z = 0$$

if $i \neq 0$. So

$$x^T x = x^T \left(\sum_i \frac{1}{v v_i} A_i \otimes A_i \right) x = \sum_{j=0}^{d} \frac{1}{m_j} x^T (E_j \otimes \overline{E_j}) x.$$

The matrices E_i are positive-semidefinite, and therefore so are the matrices $E_i \otimes \overline{E_i}$. Consequently each term in the last sum is non-negative, and thus

$$|S| = xx^T \ge x^T (E_0 \otimes E_0) x = \frac{|S|^2}{v^2}.$$

Therefore $|S| \le v$. □

19
Galois Theory

We are going to use Galois theory to establish a correspondence between certain subfields of L and subschemes of $\mathscr{A}$. This may be viewed as an extension of work of Bridges and Mena [bm2] and of Hou [xdh].

19.1 Bose-Mesner Automorphisms

Let $\mathscr{A}$ be an association scheme with Bose-Mesner algebra $\mathbb{C}[\mathscr{A}]$. A linear map $M \mapsto M^\psi$ on $\mathbb{C}[\mathscr{A}]$ is an *algebra automorphism* if for all M and N in $\mathbb{C}[\mathscr{A}]$:

(a) $(MN)^\psi = M^\psi N^\psi$.

(b) $(M \circ N)^\psi = M^\psi \circ N^\psi$.

It follows immediately that ψ maps Schur idempotents to Schur idempotents and matrix idempotents to matrix idempotents. Using this, we will prove:

(c) ψ is invertible.

We have $J \circ J = J$ and therefore

$$J^\psi \circ J^\psi = J^\psi.$$

Hence J^ψ is a 01-matrix. We also have $J^2 = vJ$ and so

$$(J^\psi)^2 = J^\psi;$$

it follows that $J^\psi = J$. Consequently

$$J = J^\psi = \sum_i A_i^\psi,$$

from which we see that ψ permutes the Schur idempotents. Therefore it maps a basis of $\mathbb{C}[\mathscr{A}]$ to a basis of $\mathbb{C}[\mathscr{A}]$, and therefore it is invertible. We also see that ψ must permute the set of matrix idempotents.

Since

$$\mathrm{sum}((A_iA_j)\circ I) = \mathrm{tr}((A_iA_j)I) = \mathrm{tr}(A_iA_j) = \langle A_i^T, A_j\rangle,$$

we find that $(A_iA_j)\circ I \neq 0$ if and only if $A_j = A_i^*$. Hence

(d) $(M^*)^\psi = (M^\psi)^*$.

This completes our list of properties of an algebra automorphism.

The transpose map is an algebra automorphism, which is non-trivial if $\mathscr{A}$ is not symmetric.

We are going to use algebra automorphisms to construct subschemes. If ψ is an algebra automorphism, the *fixed-point space* of ψ is the set of matrices in $\mathbb{C}[\mathscr{A}]$ that are fixed by ψ. This is evidently a subspace of $\mathbb{C}[\mathscr{A}]$, as the name implies.

19.1.1 Lemma. *The fixed-point space of an algebra automorphism of an association scheme is the Bose-Mesner algebra of a subscheme.*

Proof. The fixed-point space is closed under multiplication, Schur multiplication and contains I and J. □

By way of example, consider the transpose map acting on $\mathscr{A}$. Its fixed-point space is spanned by those Schur idempotents that are symmetric, together with the matrices

$$A_i + A_i^T,$$

where A_i is not symmetric. By the lemma, these matrices are the Schur idempotents of a symmetric subscheme of $\mathscr{A}$.

19.2 Galois

Let $\mathscr{A}$ be an association scheme. The *splitting field* of $\mathscr{A}$ is the extension $\mathbb{F}$ of the rationals generated by the eigenvalues of the scheme. The *Krein field* is the extension of the rationals generated by the Krein parameters. From the relation between the dual eigenvalues and the eigenvalues we see that the splitting field is also generated by the dual eigenvalues. From our expression for the Krein parameters in terms of the eigenvalues, the Krein field is a subfield of $\mathbb{F}$.

Let $\mathscr{A}$ be an association scheme with splitting field L, and Krein field K. Let Γ be the Galois group of $L/\mathbb{Q}$ and let H be the Galois group of L/K. (So H is a subgroup of Γ.)

If $\sigma \in \Gamma$ and $M \in L[\mathscr{A}]$, define M^σ to be matrix obtained by applying σ to each entry of M. This gives the *entry-wise* action of Γ. This **is not** an L-linear map.

We define a second action of Γ on $L[\mathscr{A}]$. Suppose $\tau \in \Gamma$ and $M \in L[\mathscr{A}]$. Then $M = \sum_j a_jE_j$ and we define $M^{\hat{\tau}}$ by

$$M^{\hat{\tau}} = \sum_j a_jE_j^\tau.$$

This **is** an L-linear map.

19.2.1 Theorem. *Let $\mathscr{A}$ be an association scheme with splitting field L and Krein field K. If τ is an element of the Galois group of $L/\mathbb{Q}$, then $\hat{\tau}$ is an algebra automorphism if and only if τ fixes each element of K.*

Proof. There are a number of steps to the argument.

If $M \in L[\mathscr{A}]$ and $M = \sum_j a_j E_j$ then, since $E_j^* = E_j$, we have

$$(M^*)^\tau = \sum_j a_j^* E_j^\tau = (M^{\hat{\tau}})^*.$$

Next, if M and N belong to $L[\mathscr{A}]$ and $\sigma \in \Gamma$, then

$$(MN)^\sigma = M^\tau N^\sigma, \qquad (M \circ N)^\sigma = M^\sigma \circ N^\sigma.$$

It follows from this that, if $A_i \in \mathscr{A}$, then $A_i^\sigma \in \mathscr{A}$ and similarly E_j^σ is a principal idempotent for each j. (Note, however that this entry wise action is linear over $\mathbb{Q}$, but not over L.)

Since $(E_i)^\tau (E_j)^\tau = (E_i E_j)^\tau$, we have

$$(MN)^{\hat{\tau}} = M^{\hat{\tau}} N^{\hat{\tau}}.$$

We show that $\hat{\tau}$ commutes with Schur multiplication if and only if $\tau \in H$. On the one hand,

$$(E_i \circ E_j)^{\hat{\tau}} = \frac{1}{v} \sum_r q_{i,j}(r) E_r^{\hat{\tau}} = \frac{1}{v} \sum_r q_{i,j}(r) E_r^\tau$$

while, on the other

$$E_i^{\hat{\tau}} \circ E_j^{\hat{\tau}} = E_i^\tau \circ E_j^\tau = (E_i \circ E_j)^\tau = \frac{1}{v} \sum_r q_{i,j}(r)^\tau E_r^\tau.$$

Comparing these two equations yields that

$$(E_i \circ E_j)^{\hat{\tau}} = E_i^{\hat{\tau}} \circ E_j^{\hat{\tau}}$$

for all i and j, if and only if τ fixes each Krein parameter.

From this we see that $\hat{\tau}$ is an algebra automorphism of $\mathscr{A}$ if and only if τ fixes each element of K. □

Using related, but distinct, actions of the Galois group of L/K, Munemasa [mune] proved that H lies in the centre of Γ. (Similar results appear in [dbg, coga].) Since the argument is short, we present a version of it here. If $\sigma \in \Gamma$ then E_j^σ is a principal idempotent. Therefore

$$E_j^{\sigma\hat{\tau}} = E^{\sigma\tau} = \frac{1}{v} \sum_i q_j(i)^{\sigma\tau} A_i$$

and similarly,

$$E_j^{\hat{\tau}\sigma} = E^{\tau\sigma} = \frac{1}{v} \sum_i q_j(i)^{\tau\sigma} A_i.$$

Noting that $A_i^\sigma = A_i$ and that $\hat{\tau}$ is linear, we also have

$$\left(\sum_i q_j(i)A_i\right)^{\sigma\hat{\tau}} = \sum_i q_j(i)^\sigma A_i^{\hat{\tau}} = \left(\sum_i q_j(i)A_i\right)^{\hat{\tau}\sigma}.$$

As the first term here equals $E_j^{\sigma\hat{\tau}}$ and the second equals $E_j^{\hat{\tau}\sigma}$, we conclude that

$$q_j(i)^{\sigma\tau} = q_j(i)^{\tau\sigma}.$$

Since the dual eigenvalues generate L, this implies that σ and τ commute, for all σ in Γ and all τ in H. Therefore H lies in the centre of Γ.

19.2.2 Theorem. *Let $\mathscr{A}$ be an association scheme with splitting field L and Krein field K and let H be the Galois group of L/K. Let F be a subfield of L that contains K and let H_F be the corresponding subgroup of H. Then the matrices in $L[\mathscr{A}]$ with eigenvalues and entries in F are the Bose-Mesner algebra over F of the subscheme fixed by the elements $\hat{\tau}$, for τ in H_F.*

Proof. Let $\widehat{H}_F$ denote the group formed by the mappings $\hat{\tau}$, for τ in H_F. Let $\mathscr{F}$ denote the set of matrices in $L[\mathscr{A}]$ with eigenvalues and entries in F. If $M \in L[\mathscr{A}]$ and $M = \sum_i a_i E_i$ then

$$M^{\hat{\tau}\tau^{-1}} = \sum_i a_i^{\tau^{-1}} E_i.$$

This shows that a 01-matrix in $L[\mathscr{A}]$ is fixed by $\hat{\tau}$ if and only if its eigenvalues are fixed by τ; thus a 01-matrix lies in $\mathscr{F}$ if and only if it is fixed by $\widehat{H}_F$.

Clearly $\mathscr{F}$ is a transpose-closed algebra. Suppose M and N belong to $\mathscr{F}$ and

$$M = \sum_i a_i E_i, \qquad N = \sum_i b_i E_i.$$

Then

$$M \circ N = \sum_{i,j} a_i b_j E_i \circ E_j$$

and, as the eigenvalues of $E_i \circ E_j$ lie in F, it follows that the eigenvalues of $M \circ N$, along with its entries, lie in F. Therefore $\mathscr{F}$ is Schur-closed. This implies that $\mathscr{F}$ is spanned by 01-matrices.

Consequently $\mathscr{F}$ is the span over F of the 01-matrices in $L[\mathscr{A}]$ with eigenvalues in F. This completes the proof. □

If F is a subfield of L that contains K, we use $\mathscr{A}/F$ to denote the subscheme of $\mathscr{A}$ corresponding to F.

19.3 Applications

An association scheme $\mathscr{A}$ is *metric* if its elements $A_0, \ldots, A_d$ can be ordered so that A_i is polynomial of degree i in A_1, for $i = 0, 1, \ldots, d$.

19.3.1 Lemma. *Let $\mathscr{A}$ be a symmetric association scheme with splitting field L and Krein field K. If $\mathscr{A}$ is metric then $[L:K] \leq 2$.*

Proof. Suppose that $A_0, \ldots, A_d$ are the minimal Schur idempotents of $\mathscr{A}$, and that $\mathscr{A}$ is metric relative to A_1. Let τ be an element of the Galois group H of L/K. Then $A_1^{\hat{\tau}}$ is a minimal Schur idempotent for $\mathscr{A}$, and it follows that $\mathscr{A}$ is metric relative to $A_1^{\hat{\tau}}$. By [bcn: Theorem 4.2.12] we know that $\mathscr{A}$ is metric with respect to at most two of its classes. As each A_i is a rational polynomial in A_1, any element of H which fixes A_1 must fix each A_i and therefore $|H| \leq 2$. □

If $\mathscr{A}$ has the property that the valencies v_i are all distinct then each minimal Schur idempotent must be fixed under the eigenvalue action of an element of L/K; hence for schemes with this property L and K must coincide.

Let G be a finite group of order v. We may view the complex group algebra $\mathbb{C}[G]$ as an algebra of $v \times v$ matrices, with permutation matrices representing the elements of G. Then centre of $\mathbb{C}[G]$ is then an association scheme. The matrices in this scheme correspond to the conjugacy classes of G and the principal idempotent to the irreducible characters of G. For these schemes the Krein parameters are known to be rationals. If G has exponent m then the splitting field L is the extension of $\mathbb{Q}$ by a primitive m-th root of unity. Each subfield of L thus determines a subscheme. In particular, if some character of G is not rational valued then the rational matrices with rational eigenvalues are the Bose-Mesner algebra over $\mathbb{Q}$ of a proper subscheme.

When G is abelian we can say more. If we view the elements of G as $v \times v$ permutation matrices then G itself is an association scheme. Bridges and Mena [bm2] proved that, if $\mathscr{A} = G$ then $\mathscr{A}/\mathbb{Q}$ has dimension equal to the number of cyclic subgroups of G. They also determined the minimal Schur idempotents of $\mathscr{A}/\mathbb{Q}$: if $g \in G$, let $[g]$ denote the set of elements h of G; the corresponding sum in the Bose-Mesner algebra $\mathbb{C}[G]$ is a 01-matrix and can be shown to have rational eigenvalues.

We present one application, proved independently by R. A. Liebler (private communication).

19.3.2 Lemma. *A regular abelian group of automorphisms of the n-cube has exponent dividing* 4.

Proof. Let G be an abelian group acting regularly on the n-cube Q_n, and suppose that G has exponent 2^m, where $m \geq 3$. Let g be an element of G with order 2^m. Then $[g]$ consists of all powers g^i, where i is odd and less than 2^m. This implies that $[g]$ is the adjacency matrix of the graph formed by 2^{n-m} vertex disjoint copies of $K_{2^{m-1},2^{m-1}}$.

Let $\mathscr{A}$ be the association scheme formed by the elements of G. As the eigenvalues of Q_n are integers, its adjacency matrix belongs to $\mathscr{A}/\mathbb{Q}$. There-

fore it is a sum of matrices $[g]$, where g ranges over a generating set for G. At least one element of this generating set must have order 2^m and, consequently Q_n must contain an induced subgraph isomorphic to $K_{4,4}$.

We complete our argument by showing that $K_{3,3}$ cannot be an induced subgraph of Q_n. This proof is by induction on n. The crucial property of $K_{3,3}$ is that we cannot disconnect it by deleting the edges of a matching. For all matchings in $K_{3,3}$ lie in a matching of size three, all 3-matchings in $K_{3,3}$ are equivalent under its automorphism group and $K_{3,3}$ with a 3-matching deleted is C_6, the cycle on six vertices. The n-cube on the other hand is the Cartesian product of K_2 with Q_{n-1}, hence we may delete a perfect matching from Q_n, obtaining two disjoint copies of Q_{n-1} as a result. So any induced $K_{3,3}$ in Q_n must be contained in one of these copies of Q_{n-1}, and hence our claim follows. □

The abelian group $\mathbb{Z}_4^n$ acts regularly on Q_{2n}, since Q_{2n} is isomorphic to the Cartesian product of n copies of C_4. Thus the hypothesis of the lemma cannot be weakened.

We explain briefly why the last result is of interest. Any abelian group of exponent dividing four and order 4^n acts a regular group of automorphisms of the Hamming scheme $H(2n,2)$. Hence we can identify its vertices with the elements of the group $\mathbb{Z}_4^n$, or with the elements of $\mathbb{Z}_2^{2n}$. An additive code over $\mathbb{Z}_4$ is a subset which forms a subgroup of $\mathbb{Z}_4^n$, a linear binary code is a subset which is a subgroup of $\mathbb{Z}_2^{2n}$. A code can be additive over $\mathbb{Z}_4$ but not over $\mathbb{Z}_2$. In [hkcss] it is shown that the Kerdock codes, which are non-linear binary codes, are additive codes over $\mathbb{Z}_4$. Thus the above result indicates one obstacle to extending the results in [hkcss] to codes over $\mathbb{Z}_{2^m}$ when $m \geq 3$.

20
Gauss and Jacobi Sums

20.1 From Sets to Gauss Sums

Let C be a subset of the abelian group G. We have seen that the eigenvalues of $X(G,C)$ are given by the sums $\varphi(C)$, where φ runs over the characters of G. This leads to another interpretation of the eigenvalues of X. We allow X to be directed and will tolerate loops.

If x_C is the characteristic function of C as a subset of G and we view the characters of G as an orthogonal basis for $\mathbb{C}^v$, then $\varphi(C)$ is the coefficient of x_C relative to this basis; equivalently

$$\varphi(C) = \langle x_C, \varphi \rangle.$$

Since $\langle \varphi, \varphi \rangle = v$, we thus have

$$x_C = \frac{1}{\sqrt{v}} \sum_{\varphi} \varphi(C)\varphi.$$

From this viewpoint, the eigenvalues of $X(G,C)$ are the Fourier coefficients of the characteristic function x_C.

We turn to the cases where G is the additive group of a field with order q. In this case we have a second abelian group at hand—the multiplicative group $\mathbb{F}^*$ of $\mathbb{F}$ and our connection set is normally a subset of $\mathbb{F}^*$. Characters on $\mathbb{F}^*$ are called multiplicative characters of $\mathbb{F}$, to distinguish them from the additive characters we have considered up till now. The multiplicative characters form an orthogonal basis for the space of complex functions on $\mathbb{F}^*$. We constructed Paley graphs (and tournaments) by choosing the connection set C to be a subgroup of index two in $\mathbb{F}^*$. However we can also specify this connection set in terms of multiplicative characters.

The *order* of a character ψ is the least integer k such that ψ^k is the trivial character. The order is well-defined because the characters form a group. Considering $\mathbb{F}^*$, if $r|(q-1)$ then there is a unique homomorphism from $\mathbb{F}^*$ onto $\mathbb{Z}_r$, whose kernel is the unique subgroup of $\mathbb{F}^*$ with index r. The composition of a character on $\mathbb{Z}_r$ with this homomorphism is a character of $\mathbb{F}^*$ with order dividing r. In particular $\mathbb{F}^*$ has characters of order r, and

they all have the same kernel. A *quadratic character* is a character of order two. (We do not consider the order of additive characters on $\mathbb{F}$, because the order of any non-trivial additive character is equal to the characteristic of $\mathbb{F}$.)

Now suppose the order q of $\mathbb{F}$ is odd. Then $\mathbb{F}$ has a quadratic character, which we denote by σ. If $x \in \mathbb{F}^*$, then $\sigma(x) \pm 1$ and $\sigma(x^2) = 1$. Thus $\ker(\sigma)$ is the set S of squares in $\mathbb{F}^*$. If β is an additive character on $\mathbb{F}$, then the eigenvalue of the Paley graph belonging to β is $\beta(S)$. We have

$$\beta(S) = \langle \beta, x_S \rangle.$$

Since x_S is a function on $\mathbb{F}^*$ we can write it as a linear combination of multiplicative characters of $\mathbb{F}^*$, in fact

$$x_S = \frac{1}{2}(1 + \sigma).$$

Therefore

$$\beta(S) = \frac{1}{2} \sum_{x \in \mathbb{F}^*} \beta(x) + \frac{1}{2} \sum_{x \in \mathbb{F}^*} \beta(x)\sigma(x).$$

Since $\beta(0) = 1$, the first sum is -1 and the problem is the second sum.

There is a notational point here. The second sum is the inner product of two functions on $\mathbb{F}^*$, namely σ and the restriction of β to $\mathbb{F}^*$. It is equal to the inner product (on $\mathbb{F}$) of β and the function we get from σ by declaring $\sigma(0) = 0$. This second viewpoint is the universal standard and we will follow it from now on. Thus we will write the second sum as

$$\sum_{x \in \mathbb{F}} \beta(x)\sigma(x)$$

and we extend each multiplicative character ψ to a function on $\mathbb{F}$ by declaring that $\psi(0) = 0$. We do this even when ψ is the trivial character on $\mathbb{F}^*$!

To summarise, we have shown that the problem of computing the eigenvalues of the Paley graph can be "reduced" to the problem of evaluating a sum

$$\sum_{x \in \mathbb{F}} \beta(x)\sigma(x),$$

where β is an additive character on $\mathbb{F}$ and σ is a multiplicative character. Sums of this form are known as *Gauss sums*, and play a very important role in number theory.

20.2 Evaluating Gauss Sums

Fix a field $\mathbb{F}$ of order q. If β is an additive and σ a multiplicative character of $\mathbb{F}$, define

$$G(\beta, \sigma) := \sum_{x \in \mathbb{F}} \beta(x)\sigma(x).$$

If $c \in \mathbb{F}$, we define

$$\beta_c : x \to \beta(cx);$$

this is again an additive character on $\mathbb{F}$.

20.2.1 Theorem. *If ξ is an additive character and ψ a multiplicative character on $\mathbb{F}$, then:*

(a) If $c \in \mathbb{F}$, then $G(\xi_c, \sigma) = \sigma(c^{-1})G(\xi, \sigma)$.

(b) $G(\xi, \overline{\psi}) = \psi(-1)\overline{G(\xi, \psi)}$.

(c) If $\psi \neq 1$, then $G(\xi, \psi)\overline{G(\xi, \psi)} = q$.

Proof. We have

$$\begin{aligned} G(\beta_c, \sigma) &= \sum_{x \in \mathbb{F}} \beta(cx)\sigma(x) \\ &= \sum_{y \in \mathbb{F}} \beta(y)\sigma(c^{-1}y) \\ &= \sigma(c^{-1}) \sum_{y \in \mathbb{F}} \beta(y)\sigma(y) \\ &= \sigma(c^{-1})G(\beta, \sigma). \end{aligned}$$

For (b), we have

$$\begin{aligned} g(\overline{\psi}) &= \sum_x \xi(x)\overline{\psi(x)} \\ &= \sum_x \overline{\overline{\xi(x)}\psi(x)} \\ &= \sum_x \overline{\xi(-x)\psi(x)} \\ &= \sum_x \overline{\xi(x)\psi(-x)} \\ &= \overline{\psi(-1)} \sum_x \overline{\xi(x)\psi(x)} \\ &= \overline{\psi(-1)}\,\overline{g(\psi)} \end{aligned}$$

Since $\psi(-1) = \pm 1$, we have (b). For (c),

$$\begin{aligned} g(\psi)\,\overline{g(\psi)} &= \sum_{x,y \neq 0} \xi(x)\psi(x)\overline{\xi(y)}\,\overline{\psi(y)} \\ &= \sum_{x,y \neq 0} \xi(x-y)\psi(xy^{-1}) \\ &= \sum_{z,y \neq 0} \xi(zy-y)\psi(z) \\ &= \sum_{z=1, y \neq 0} \xi(zy-y)\psi(z) + \sum_{z \neq 0,1} \sum_{y \neq 0} \xi(zy-y)\psi(z). \end{aligned}$$

Here the first sum is

$$\sum_{y \neq 0} \xi(0)\psi(1) = \sum_{y \neq 0} 1 = (q-1)$$

while if $z - 1 \neq 0$ then

$$\sum_{y \neq 0} \xi((z-1)y) = -1$$

and so the second sum is

$$\sum_{z\neq 0,1}(-1)\psi(z)=1.$$

Thus (c) follows. □

20.3 *Paley for Number Theorists*

We apply the previous results to the task of evaluating the eigenvalues of the Paley graph.

Our initial task is to compute $g(\sigma)$, where σ is the quadratic character on $\mathbb{F}$. Since the values of σ are real, $\overline{\sigma}=\sigma$. Hence by Theorem 20.2.1 we have that

$$\sigma(-1)g(\psi)^2=q$$

and hence

$$g(\psi)=\pm\sqrt{\sigma(-1)q}.$$

Note that $\sigma(-1)$ is 1 or -1 according as -1 is a square or a non-square in $\mathbb{F}$.

Hence if -1 is a square, then $g(\sigma)=\pm\sqrt{q}$. So $G(\xi,\sigma)=\pm\sqrt{q}$ and so if $c\neq 0$, then

$$G(\xi\circ c,\sigma)=\sigma(c^{-1})G(\xi,\sigma).$$

Here we use $\xi\circ c$ to denote the composition of ξ with multiplication by c. Note that $\sigma(c^{-1})=\sigma(c)$. We conclude that the eigenvectors ξ and $\xi\circ c$ have the same eigenvalue if and only if c is a square. It follows that the Paley graph has $(q-1)/2$ eigenvalues equal to $\frac{1}{2}(-1+\sqrt{q})$ and $(q-1)/2$ equal to $\frac{1}{2}(-1-\sqrt{q})$.

20.4 *Cyclotomic Schemes*

Suppose $\mathbb{F}=GF(q)$ is a field of characteristic p and $q=rm+1$. Let a be a generator of $\mathbb{F}^*$, let D denote the subgroup of r-th powers in $\mathbb{F}^*$ and let D_i denote the coset a^iD of D. (So $D_0=D$.)

Let θ be fixed non-trivial p-th root of 1 in $\mathbb{C}$ and let ψ_a be the additive character on $\mathbb{F}$ given by

$$\psi_a(x)=\theta^{\mathrm{tr}(ax)}.$$

Let ξ be a multiplicative character of order r on $\mathbb{F}^*$; we work with Gauss sums

$$G(\psi,\xi)=\sum_{x\in\mathbb{F}}\psi(x)\xi(x).$$

Now

$$\xi(a^ix^r)=\xi(a^i)\xi(x^r)=\xi(a^i)$$

and thus ξ is constant on the cosets of D in $\mathbb{F}^*$. We have

$$\sum_{j=0}^{r-1}\xi^j(x)=\begin{cases}r, & x\in\ker(\xi);\\ 0, & \text{otherwise.}\end{cases}$$

Consequently

$$\frac{1}{r}\sum_{j=0}^{r-1} G(\psi, \xi^j) = \sum_{y \in D} \psi(y).$$

21
Pseudocyclic Schemes

An association scheme on d classes is *pseudocyclic* if its non-trivial multiplicities are equal. As we will see, this condition implies that the non-trivial valencies are equal. The Paley graphs provide one class of examples.

21.1 Cyclotomic Schemes

Let $\mathbb{F}$ be a finite field of order q and suppose d is a positive integer that divides $q-1$. Let D be the subset of d-th powers in $\mathbb{F}$, let g be a primitive element in $\mathbb{F}$ and for $r = 1, \ldots, d$, define

$$D_r := g^{r-1}D.$$

These sets D_i form a complete set of cosets for D in $\mathbb{F}^*$. Define X_i to be the Cayley graph $X(\mathbb{F}, D_i)$ and let A_i be its adjacency matrix. The graphs X_i form a pseudocyclic association scheme, but some effort is required to prove this. Multiplication by g induces a permutation of the elements of $\mathbb{F}$, and this permutation induces isomorphisms

$$X_1 \to X_2 \to \cdots \to X_d \to X_1.$$

In particular the graphs X_i are pairwise isomorphic.

We offer examples. If $\mathbb{F} = GF(p)$ (where p is odd) then

$$\{x^{(p-1)/2} : x \in \mathbb{F}^*\} = \{1, -1\}.$$

Here each graph is isomorphic to the cycle C_p. The squares in $\mathbb{F}$ form a subgroup of $\mathbb{F}^*$ of index two. If $q \equiv 1$ modulo four, then X_1 and X_2 are isomorphic to the Paley graph; if $q \not\equiv 1$ modulo four, we have Paley tournaments.

To prove that this construction gives an association scheme, we need only show that $A_i A) j$ lies in the span of the set $\{A_0, \ldots, A_d\}$. For this we use an eigenvalue argument. If ξ is an additive character of $\mathbb{F}$, then it is an eigenvector for the scheme and the eigenvalue of A_i is $\xi(D_i)$. If $a \in \mathbb{F}$

then ξ_a is the character whose value on x is $\xi(ax)$; by varying a we get all additive characters of $\mathbb{F}$. We see that

$$\xi_a(D_i) = \xi(ag^i D).$$

If $a \neq 0$ then $ag^i D = g^j D$ for some j and hence we see that each graph in the scheme has at most $d+1$ distinct eigenvalues.

21.1.1 Lemma. *The graphs X_i form an association scheme with d classes. The intersection number $p_{i,j}(r)$ is the number of ways an element z in D_r can be written as a sum $x+y$, where $x \in D_i$ and $y \in D_j$.* □

21.2 Pseudocyclic Schemes

Let $\mathscr{A}$ be an association scheme on v vertices with d classes, valencies $v_0, \ldots, v_d$ and multiplicities $m_0, \ldots, m_d$. (So $v_0 = m_0 = 1$.) Let Δ_v and Δ_m denote the diagonal matrices of valencies and multiplicities respectively, and let P and Q be its matrices of eigenvalues and dual eigenvalues.

21.2.1 Theorem. *An association scheme on v vertices with d classes is pseudocyclic if and only if*

(a) $v_1 = \cdots = v_d$ *and*

(b) $\sum_{i=0}^{d} p_{r,i}(i) = \frac{v-1}{d} - 1$ *for* $r = 1, \ldots, d$

Proof. Let $B_0, \ldots, B_d$ be the $(d+1) \times (d+1)$ matrices representing multiplication by $A_0, \ldots, A_d$ on the Bose-Mesner algebra of $\mathscr{A}$. Thus

$$(B_r)_{i,j} = p_{r,j}(i).$$

As the eigenvalues of B_r are the eigenvalues of A_r, we have

$$\sum_{i=0}^{d} p_r(i) = \mathrm{tr}(B_r) = \sum_{j=0}^{d} p_{r,j}(j).$$

Now assume that (a) and (b) hold set t equal to $(v-1)/d$. Since

$$p_r(0) = v_r = \frac{v-1}{d} = t$$

and

$$\sum_{i=0}^{d} p_r(i) = t - 1,$$

when $r > 0$ we have

$$v_r + t \sum_{i=1}^{d} = t + t(-1) = 0.$$

So

$$\mathbf{1}^T \Delta_v P = v \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix}$$

and therefore

$$v\begin{pmatrix}1 & 0 & \dots & 0\end{pmatrix}Q = \mathbf{1}^T\Delta_v PQ = v\mathbf{1}^T\Delta_v$$

This shows that the first row of Q is $\mathbf{1}^T\Delta_v$. As the first row of Q is equal to $\mathbf{1}^T\Delta_m$, it follows that $\Delta_m = \Delta_v$. Hence we have shown that if (a) and (b) hold, the non-trivial multiplicities of $\mathscr{A}$ are equal.

Now suppose that the non-trivial multiplicities of $\mathscr{A}$ are equal to t. (So $t = (v-1)/d$.) We have

$$\Delta_m P = Q^T\Delta_v$$

whence

$$P^T\Delta_m P = v\Delta_v$$

and so

$$\det(P)^2\prod_{i=1}^{d} m_i = v^{d+1}\prod_{j=1}^{d} v_j.$$

Since the eigenvalues of the scheme are algebraic integers, this implies that $\det(P)^2$ is an integer and that t^d divides the right side. Since t and v are coprime, it follows that t^d divides $\prod_{j=1}^{d} v_j$.

By the arithmetic-geometric mean inequality

$$\prod_{j=1}^{d} v_j \le \left(\frac{v_1+\cdots+v_d}{d}\right)^d \le t^d$$

from which it follows now that the valencies $v_1,\dots,v_d$ are equal.

It remains to show that (b) holds. Since

$$m_1 = \cdots = m_d = t,$$

and since $\operatorname{tr}(A_r) = 0$,

$$0 = v_r + t\sum_{i=1}^{d} p_r(i).$$

Consequently

$$\sum_{i=0}^{d} p_r(i) = v_r - \frac{v_r}{t} = t-1,$$

as required. □

21.3 A Scheme that is not Pseudocyclic

The scheme belonging to $L(K_7)$ has matrix of eigenvalues

$$\begin{pmatrix}1 & 10 & 10\\ 1 & 3 & -2\\ 1 & -2 & 1\end{pmatrix},$$

and thus it an example of a scheme that has all non-trivial valencies equal but is not pseudocyclic. We compute B_1 and B_2.

The parameters of $L(K_7)$ are $(21,10;5,4)$. Therefore

$$p_{1,1}(1) = 5, \quad p_{1,1}(2) = 4.$$

The parameters of $\overline{L(K_7)}$ are $(21,10;3,6)$ and consequently

$$B_1 = \begin{pmatrix} 0 & 10 & 0 \\ 1 & 5 & 4 \\ 0 & 4 & 6 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} 0 & 0 & 10 \\ 0 & 4 & 6 \\ 1 & 6 & 3 \end{pmatrix}.$$

So

$$\mathrm{tr}(B_1) = 11, \quad \mathrm{tr}(B_2) = 7$$

where both traces would equal 9 for a pseudocyclic scheme.

21.4 Three-Class Schemes: Intersection Matrices

We study pseudocyclic schemes with three classes. The theory presented is due to Rudi Mathon. Let $\mathscr{A}$ be such a scheme on $v = 3m+1$ vertices. So all non-trivial valencies and multiplicities are equal to m. There are integers a, b and c such that

$$A_1^2 = mI + aA_1 + bA_2 + cA_3. \tag{21.4.1}$$

We aim to determine the matrices B_1, B_2 and B_3, starting with B_1.

We note that

$$(A_iA_j) \circ A_k = p_{i,j}(k)A_k$$

and consequently

$$vv_k p_{i,j}(k) = \mathrm{tr}(A_iA_jA_k).$$

If $j,k \neq 0$ then $v_j = v_k$ and so, in a pseudocyclic scheme,

$$p_{i,j}(k) = p_{i,k}(j).$$

Thus we may write

$$B_1 = \begin{pmatrix} 0 & m & 0 & 0 \\ 1 & a & b & c \\ 0 & b & x & y \\ 0 & c & y & z \end{pmatrix}$$

where x, y and z are to be determined. Each row of B_1 sums to m and thus

$$a+b+c = m-1.$$

Also

$$m-1 = \mathrm{tr}(B_1) = a+x+z$$

and thus we have three equations for x, y and z. The solution is

$$x = c, \quad y = a+1, \quad z = b$$

and thus

$$B_1 = \begin{pmatrix} 0 & m & 0 & 0 \\ 1 & a & b & c \\ 0 & b & c & a+1 \\ 0 & c & a+1 & b \end{pmatrix}.$$

Now we deal with B_2, which we may write as

$$B_2 = \begin{pmatrix} 0 & 0 & m & 0 \\ 0 & b & c & a+1 \\ 1 & c & x & y \\ 0 & a+1 & y & z \end{pmatrix}.$$

(Note that the entries in the second row of B_2 have been read off from B_1.) Arguing in a similar way, we find that

$$B_2 = \begin{pmatrix} 0 & 0 & m & 0 \\ 0 & b & c & a+1 \\ 1 & c & a & b \\ 0 & a+1 & b & c \end{pmatrix}.$$

At this point the only intersection number we have not written down is $p_{3,3}(3)$. Since this is determined by the row sum, we have

$$B_3 = \begin{pmatrix} 0 & 0 & 0 & m \\ 0 & c & a+1 & b \\ 0 & a+1 & b & c \\ 1 & b & c & a \end{pmatrix}.$$

If P is the 4×4 permutation matrix corresponding to the permutation $(0)(123)$ of $\{0,1,2,3\}$, then

$$P^T B_1 P = B_2, \quad P^T B_2 P = B_3, \quad P^T B_3 P = B_1$$

and therefore the map $M \mapsto P^T M P$ is an automorphism of the matrix algebra generated by the B_i's. It follows that the linear map on the Bose-Mesner algebra that fixes I and sends

$$A_1 \to A_2 \to A_3 \to A_1$$

is an automorphism of the Bose-Mesner algebra. (It is regrettable that we need so much computation to derive this symmetry.)

21.5 Three-Class Schemes: Eigenvalues

We can express the eigenvalues of a cyclotomic 3-class scheme in terms of v and $p_{1,1}(1)$.

21.5.1 Theorem. *If $\mathscr{A}$ is a cyclotomic 3-class scheme on v vertices, then the eigenvalues of A_1 are the zeros of*

$$t^3 + t^2 - mt - \beta$$

where

$$m = \frac{1}{3}(v-1), \quad \beta = \frac{1}{3}(v(a+1) - m^2), \quad a = p_{1,1}(1).$$

Proof. In addition to the identity

$$A_1^2 = mI + aA_1 + bA_2 + cA_3$$

we have the following

$$\begin{aligned} A_1A_2 &= bA_1 + cA_2 + (a+1)A_3 \\ A_2A_3 &= (a+1)A_1 + bA_2 + cA_3 \\ A_1A_3 &= cA_1 + (a+1)A_2 + bA_3. \end{aligned}$$

(Here the first equation follows from the entries of B_1 and other two from symmetry.) It follows that

$$A_1^2 - A_2A_3 = mI - A_1$$

and consequently

$$A_1^3 + A_1^2 - mA_1 = A_1A_2A_3.$$

The product $A_1A_2A_3$ is invariant under cyclic permutations of its indices, and so there are scalars β and γ such that

$$A_1A_2A_3 = \beta I + \gamma J$$

and then

$$A_1^3 + A_1^2 - mA_1 - \beta I = \gamma J.$$

Since $A_1\mathbf{1} = m\mathbf{1}$,

$$(A_1^3 + A_1^2 - mA_1)\mathbf{1} = m^3\mathbf{1} + m^2\mathbf{1} - m^2\mathbf{1} = m^3\mathbf{1}$$

and therefore

$$m^3 = \beta + v\gamma.$$

Also

$$\operatorname{tr}(A_1A_2A_3) = \operatorname{sum}((A_1A_2) \circ A_3) = p_{1,2}(3)\,v\,v_3 = (a+1)\,vm$$

whence

$$(a+1)\,vm = v\beta + v\gamma.$$

It follows that

$$\beta = \frac{1}{3}((a+1)v - m^2), \quad \gamma = \frac{1}{3}(m^2 - (a+1)).$$ □

There is another way to compute β. The eigenvalues of A_1 are m and the three roots of

$$t^3 + t^2 - mt - \beta.$$

Therefore

$$m\beta = \det(B_1) = -m(bc - (a+1)^2)$$

and so

$$\beta = (a+1)^2 - bc.$$

It follows that given m and a we can compute bc and $b+c$

The coefficients of A_1 and A_2 in $A_1A_2A_3$ are respectively $a^2+b^2+c^2+a$ and $ab+bc+ac+b+c$, so

$$a^2 + b^2 + c^2 + a = ab + bc + ac + b + c.$$

The cyclotomic schemes are, or course, examples of pseudocyclic schemes and so we can construct 3-class examples on q points whenever $q \equiv 1 \bmod 3$. (As -1 is always a cube, we get symmetric schemes.) Suppose q is even and C is a conic in $PG(2,q)$. Then the orthogonal group acts generously transitively on the passants to C, and the resulting association scheme is pseudocyclic.

22
Graph Algebras

We consider some of the algebras we can attach to a graph.

22.1 The adjacency algebra

If R is a ring and $A = A(X)$, the ring $R[A]$ of polynomials in A is an algebra (in the moxt general sense). We restrict ourselves to the case where R is a field. If X has diameter d, the dimension of $\mathbb{F}[A]$ is at least $d+1$. If $\mathbb{F} = \mathbb{R}$ then $\dim(\mathbb{F}[A])$ is equal to the degree of the minimal polynomial of A; if $\mathbb{F} = \mathbb{R}$ and $A = A^T$, this is the number of distinct eigenvalues of A.

We note that $\mathbb{R}[A_1]$ and $\mathbb{R}[A_2]$ are isomorphic if and only if X_1 and X_2 are cospectral.

The automorphism group of X lies in the commutant of $\mathbb{F}[A]$.

One question is what we can say about when two adjacency algebras of $\mathbb{F}$ are isomorphic? (The characteristic polynomial is not the right invariant). Note that $\mathbb{F}[A]$ is not semisimple in general.

22.2 Extended adjacency algebras: A and J

We can extend an algebra by adding new elements. We focus on extensions of the adjacency algebra $\mathbb{R}[A]$ of a graph. The key point is that if we add a symmetric matrix M to this algebra, the extended algebra is still closed under transpose and therefore is semisimple.

The first particular case is the algebra $\langle A, J\rangle$. Note that $\langle A, J\rangle = \langle A(\overline{X}), J\rangle$. The automorphism group of X lies in the commutant of $\mathbb{F}[A]$.

If X is connected and regular, then J is a polynomial in A and our "extension" has not changed anything. On the other hand:

22.2.1 Theorem. *Let X be a graph on n vertices. Then $\langle A, J\rangle = \mathrm{Mat}_{n\times n}(\mathbb{R})$ if and only if no eigenvector of X is orthogonal to $\mathbf{1}$.* □

If the stated eigenvector condition holds we say X is *controllable*. If X is controllable, the matrices

$$A^i J A^j, \quad 0 \le i, j \le n-1$$

are a basis for $\mathrm{Mat}_{n\times n}(\mathbb{R})$.

An eigenvalue is a *main eigenvalue* if its eigenspace contains an eigenvector not orthogonal to $\mathbf{1}$. If m is the multiplicity of the eigenvalue θ, then either the eigenspace of θ lies in $\mathbf{1}^\perp$ or the subspace of eigenvectors orthogonal to $\mathbf{1}$ has codimension one, and θ is a main eigenvalue. Define the reduced multiplicity of $\nu(\theta)$ to be dimension the intersection of the θ-eigenspace with $\mathbf{1}^\perp$.

22.2.2 Lemma. *Let X be a graph with ℓ distinct eigenvalues, and let μ be the number of main eigenvalues of X. Then*

$$\dim(\langle A, J\rangle) = \mu^2 + \ell - \mu. \qquad \square$$

We sketch a proof. The $\langle A, J\rangle$-module generated by $\mathbf{1}$ is irreducible [1]; we denote it by M. The direct sum decomposition

[1] your problem

$$\mathbb{R}^n = M \oplus M^\perp$$

is invariant under $\langle A, J\rangle$. Note that $M^\perp$ is spanned by the eigenvectors of A in $\mathbf{1}^\perp$. The restriction of $\langle A, J\rangle$ to M is isomorphic to $\mathrm{Mat}_{\mu\times\mu}(\mathbb{R})$; its restriction to $M^\perp$ is commutative and has $\ell - \mu$ distinct eigenvalues.

22.2.3 Theorem (Johnson & Newman)**.** *Let X_1 and X_2 be graphs. Then X_1 and X_2 are cospectral with cospectral complements if and only if there is an orthogonal matrix L such that*

$$L^{-1}A_1L = A_2, \quad L^{-1}JL = J. \qquad \square$$

So if X_1 and X_2 are controllable graphs on n vertices, then $\langle A_1, J\rangle$ and $\langle A_2, J\rangle$ are isomorphic, but there is an algebra isomorphism L such that $L^{-1}A_1L = A_2$ and $L^{-1}JL = J$ if and only if X_1 and X_2 are cocospectral.[2]

[2] cospectral with cospectral complements

If X_1 and X_2 are controllable, there are cocospectral if and only if

$$\mathbf{1}^T A_1^k \mathbf{1}^T = \mathbf{1}^T A_2^k \mathbf{1}^T$$

for all nonnegartive integers k.

22.3 Extended adjacency algebras: A and Δ

Let Δ denote the diagonal matrix of valencies of X. The algebra $\langle A, \Delta\rangle$ is semisimple and the automorphism group of X lies in its commutant.

If X is connected, then J is a polynomial in A and Δ.[3]. We assume that X is connected, and then $\langle A, J\rangle \le \langle A, \Delta\rangle$.

[3] Exercise!

We have very little to say about the algebra $\langle A, \Delta\rangle$, but there is one result of interest for trees.

22.3.1 Theorem (B. D. McKay)**.** *Let T_1 and T_2 be trees with respective adjacency matrices A_1 and A_2, and valency matrices Δ_1 and Δ_2. If T_1 and T_2 are not isomorphic, there is a polynomial $p(x, y)$ such that*

$$\phi(p(A_1, \Delta_1), t) \neq \phi(p(A_2, \Delta_2), t). \qquad \square$$

We point out that, in general, $p(A, \Delta)$ is not symmetric.

If T_1 and T_2 are isomorphic, there is a permutation matrix P such that

$$P^T A_1 P = A_2, \quad P^T \Delta_1 P = \Delta_2.$$

If T_1 and T_2 are not isomorphic, there is no invertible linear map L such that

$$L^{-1} A_1 L = A_2, \quad L^{-1} \Delta_1 L = \Delta_2.$$

If T_1 and T_2 are controllable, $\langle A_1, \Delta_1 \rangle$ and $\langle A_2, \Delta_2 \rangle$ are isomorphic, because they are both equal to the full matrix algebra.

If $L^{-1} \Delta_1 L = \Delta_2$, then $L^{-1} p(\Delta_1) L = p(\Delta_2)$ for any polynomial p.S This implies that there is a permutation matrix P such that $P^T \Delta_2 P = \Delta_1$; consequently $P^T L^{-1} \Delta_1 L P = \Delta_1$ and therefore LP is block diagonal.

23
Coherent Things

A *coherent algebra* $\mathscr{C}$ is a (finite-dimensional) matrix algebra over a subfield of $\mathbb{C}$ that is

(a) Contains J and is closed under Schur product.

(b) Is closed under transpose and complex conjugation.

From (a) we see that $\mathscr{C}$ is a commutative algebra relative to Schur multiplication. (By default our rings and algebras must have an identity element.) Condition (b) implies that $\mathscr{C}$ is a semisimple matrix algebra.

The Bose-Mesner algebra of an association scheme is a coherent algebra.

Coherent algebras are also known as cellular algebras.

Weisfeiler and Leman introduced *cellular algebras* in 1968 and, in 1970, a paper of Donald Higman introduced *coherent algebras*[1] In 1980?, Higman wrote about coherent algebras. In ??, Graham and Lehrer introduced what they called cellular algebras (in the context of representation theory). Because of this usage, my feeling is that coherent algebras is the better term.

[1] definitions to come

23.1 Coherent Configurations

We start with a simple but very useful result. We say that a Schur idempotent is *primitive* if it is not zero and cannot be expressed as a sum of two non-zero Schur idempotents. The Schur product of two distinct primitive idempotents is zero.

23.1.1 Theorem. *If the vector space of matrices $\mathscr{M}$ is closed under Schur product, it has a unique basis of consisting of primitive Schur idempotents.*

Proof. If p is a polynomial

$$p(t) = p_0 t^k + \cdots + p_k$$

and A is a matrix, we define the *Schur polynomial* $p \circ A$ to be

$$p_0 A^{\circ k} + \cdots + p_k J.$$

If λ is an entry of the matrix A, let p_λ be the polynomial that takes the value 1 on λ and 0 on all other entries of A. Then $p \circ A$ is a 01-matrix that lies in $\mathscr{M}$, and it follows that $\mathscr{M}$ is spanned by 01-matrices and consequently by a set of primitive idempotents. Therefore $\mathscr{M}$ has a basis of 01-matrices, which we must show is unique. The primitive Schur idempotents span $\mathscr{M}$. The Schur product of two distinct primitive Schur idempotents is zero and, given this, it is easy to show that the primitive Schur idempotents form a basis for $\mathscr{M}$. □

Each Schur idempotent is the adjacency matrix of a directed graph, possibly with loops. A set of directed graphs is a *coherent configuration* if it is the set of primitive Schur idempotents of a coherent algebra.

The basis of prmitive Schur idempotents is closed under transpose.[2]

[2] Prove it!

As a coherent algebra contains I, we see that I is a sum of primitive Schur idempotents and these idempotnents are necessarily diagonal matrices, and determine a partition of the vertices of the coherent configuration. The cells of this partition are known as *fibres*. A coherent algebra (or configuration) is *homogeneous* if I is a primitive Schur idempotent. This leads to the following exercise.

23.1.2 Theorem. *A commutative coherent algebra is homogeneous.* □

Suppose $\mathscr{C}_1$ and $\mathscr{C}_2$ are coherent algebras of $n \times n$ matrices and and L is an invertible matrix such that $M \to L^{-1}ML$ is an algebra isomorphism. Then L must map the diagonal elements of the Schur basis of $\mathscr{C}_1$ to diagonal elements of $\mathscr{C}_2$.[3]. It follows that there is a permutation matrix P such that LP is block diagonal, with blocks corresponding to the fibres of $\mathscr{C}_1$. (The fibre partition is a refinement of the valency partition.)

[3] work out why

23.2 Permutation Groups and Coherent Algebras

Coherent algebras play a significant role in the study of automorphism groups of graphs, and this is the original motivation for the concept.

You should verify the following:

23.2.1 Lemma. *Let A and B by $m \times n$ matrices. If P is an $m \times m$ permutation matrix, then $P(A \circ B) = (PA) \circ (PB)$.* □

23.2.2 Corollary. *If $\mathscr{P}$ is a set of $n \times n$ permutation matrices, the commutant of $\mathscr{P}$ is a coherent algebra. The fibres of this coherent algebra are the orbits of the permutation group generated by $\mathscr{P}$.*

Proof. If A and B commute with a permutation matrix P, then

$$P(A \circ B) = (PA) \circ (PB) = (AP) \circ (BP) = (A \circ B)P.$$

Since the commutant of a set of matrices is a matrix algebra, the commutant of $\mathscr{P}$ is a coherent algebra.

We leave the statement about fibres as an exercise. □

The coherent algebra generated by a set $\mathscr{M}$ of matrices is the smallest coeherent algebra that contains $\mathscr{M}$.

23.2.3 Lemma. *If A is the adjacency matrix of the graph X, then* $\mathrm{Aut}(X)$ *lies in the commutant of the coherent algebra generated by A.* □

Thus if the coherent algebra generated by A is the full matrix algebra (for example, if X is controllable) then $\mathrm{Aut}(X)$ is trivial. (The coherent algebra generated by A can be computed in polynomial time.)

We given an example of a coherent algebra that is not commutative and is not the commutant of a permutation group. Let N be the vertex-block incidence matrix of a 2-(v,k,λ)-design with b blocks. Then the matrices

$$\begin{pmatrix} I_v 0 & \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & I_b \end{pmatrix}, \begin{pmatrix} J - I_v 0 & \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & J - I_b \end{pmatrix}, \begin{pmatrix} 0 & N \\ N^T & 0 \end{pmatrix}, \begin{pmatrix} 0 & J - N \\ J - N^T & 0 \end{pmatrix}$$

are Schur idempotents that sum to J. You may prove that they form a coherent configuration if and only if the design is quasi-symmetric, that is, there are constant a and b such that any two distinct blocks intersect in a or b points.

A quantum permutation of index d is an $n \times n$ matrix whose entries are $d \times d$ projections, such that the entries in each row and each column sum to I_d. The commutant of a set of quantum permutations is a Schur-closed, and hence forms a coherent algebra.

23.3 Isomorphism

Now we get to the messy part. A homomorphism of algebras is a ring homomorphism that commutes with scalar multiplication, and a invertible homomorphism is an isomorphism. If $\mathscr{C}_1$ and $\mathscr{C}_2$ are coherent algebras, it is obvious that an algebra homomorphism from $\mathscr{C}_1$ to $\mathscr{C}_2$ need not preserve the Schur product; if it does we will call it a *coherent homomorphism*.

23.3.1 Theorem. *Let $\mathscr{M}$ and $\mathscr{N}$ be coherent algebras and let Ψ be an algebra homomorphism from $\mathscr{M}$ to $\mathscr{N}$. If Ψ commutes with Schur product and $\Psi(J) \neq 0$, then Ψ is injective.*

Proof. Let $\Psi : \mathscr{M} \to \mathscr{N}$ be an algebra homomorphism and assume $K = \ker(\Psi)$. The K is an ideal of $\mathscr{M}$. If $R \in \mathscr{M}$ and $S \in K$, then

$$\Psi(R \circ S) = \Psi(R) \circ \Psi(S) = 0.$$

This shows that $R \circ S \in K$ if $S \in K$. It follows that K has a basis of Schur idempotents. Suppose S a non-zero Schur idempotent in K. Then $JSJ \in K$, but

$$JSJ = \mathbf{1}(\mathbf{1}^T S \mathbf{1})\mathbf{1}^T = \mathbf{1}^T S \mathbf{1} J.$$

As $\mathbf{1}^T S \mathbf{1} > 0$ this implies that $J \in K$. □

23.3.2 Theorem. *An algebra homomorphism from $\mathscr{C}_1$ to $\mathscr{C}_2$ commutes with Schur product if and only if it maps the primitive Schur idempotents of $\mathscr{C}_1$ to the primitive Schur idempotents of $\mathscr{C}_2$.* □

An automorphism of a matrix algebra is *inner* if it is given by a map $M \mapsto A^{-1}MA$. It is true that any automorphism of $\mathrm{Mat}_{n\times n}(\mathbb{C})$ is inner,[4] but this is not true for

[4] Noether-Skolem

$$\mathrm{Mat}_{2\times 2}(\mathbb{C}) \oplus \mathrm{Mat}_{2\times 2}(\mathbb{C}).$$

To see this, note that the permutation

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

gives an automorphism that is not inner. I believe it is true that if $\mathscr{C}$ is a semisimple matrix algebra and the commutant of $\mathscr{C}$ is commutative, all automorphisms are inner.

An isomorphism between matrix algebras is *combinatorial* if it is given by a map $M \mapsto P^T MP$ for some permutation matrix P. Combinatorial isomorphisms are necessarily coherent.

23.4 Subalgebras

If $\mathscr{C}$ is a matrix algebra and E is an idempotent in $\mathscr{C}$, then $E\mathscr{C}E$ is subspace of $\mathscr{C}$ that is closed under matrix multiplication. It is not a subalgebra, because it does not contain the identity of $\mathscr{C}$. However the sum

$$E\mathscr{C}E + (I-E)\mathscr{C}(I-E)$$

is a subalgebra. If E is diagonal and 01 and $\mathscr{C}$ is coherent, then $E\mathscr{C}E$ is Schur closed and the subalgebra just given is coherent.

Let $\mathscr{C}$ be the coherent algebra generated by $A(X)$. If $\gamma \in \mathrm{Aut}(X)$, then the map

$$(u,v) \mapsto (u\gamma, v\gamma)$$

is a permutation of $V(X) \times V(X)$, lies in the commutant of $\mathscr{C} \otimes \mathscr{C}$. The subsets

$$\{(u,u) : u \in V\}, \quad \{(u,v) : u \sim v\}, \quad (u,v) : u \neq v, u \not\sim v$$

are unions of orbits of this action of $\mathrm{Aut}(X)$. (If X is complete, the third orbit is empty.)

Define an $n^2 \times n^2$ idempotent F_0 by

$$(F_0)_{(u,v),(u,v)} = \begin{cases} 1, & \text{if } u = v; \\ 0, & \text{otherwise.} \end{cases}$$

Note that F_0 is diagonal. Define F_1 to be the diagonal 01-idempotent with

$$(F_1)_{(u,v),(u,v)} = 1$$

whenever $u \sim v$. Finally set $F_2 = I - F_0 - F_1$. (We point out that we are not using the automorphism group to define these idempotents.)

The coherent algebra generated by $\mathscr{C} \otimes \mathscr{C}$ and F_0 is the *2-extension* of $\mathscr{C}$.

23.5 Jaeger Algebras

Let $\mathscr{M}$ denote $\mathrm{Mat}_{n\times n}(\mathbb{C})$. If $A \in \mathrm{End}(V)$, define operators X_A, Δ_A and Y_A on $V\otimes$ by

$$X_A M = AM, \quad \Delta_A(M) = A \circ M, \quad Y_A(M) = MA^T.$$

If $\mathscr{A}$ is a subalgebra of $\mathscr{M}$, define $\mathscr{J}_3(\mathscr{A})$ to be the algebra generated by the operators X_A and Δ_A for A in $\mathscr{M}$. Define $\mathscr{J}_4(\mathscr{A})$ to be the algebra generated by $\mathscr{J}_3(\mathscr{A})$ and the operators Y_A for A in $\mathscr{A}$. We say that $\mathscr{J}_3$ and $\mathscr{J}_4$ are *Jaeger algebras*.

We need to explain the transpose in the definition of Y_A and the indexing. If A_i and B_i are $n \times n$ matrices (with $i = 0, 1$), the map

$$M \mapsto AMB^T$$

is an endomorphism of M and all endomorphisms of M are linear combinations of endomorphisms of this form. Thus we have a map from $\mathrm{End}(V) \otimes \mathrm{End}(V)$ into $\mathrm{End}(\mathscr{M})$. Further

$$A_1 A_2 M B_2^T B_1 = (A_1 A_2) M (B_1 B_2)^T$$

and therefore this map is a homomorphism. Consequently $\mathscr{M}$ is a module over $\mathrm{End}(V) \otimes \mathrm{End}(V)$.

Indexing. Let V a vector space. We define some operators on $V^{\otimes r}$. Assume $X_A(i) \in \mathrm{End}(V)$ and define to be the product operator acting as A on the i-th component and as I on the remaining components. If $1 \le i \le r-1$, let $\Delta_A(i)$ act as Δ_A on the i-th and $(i+1)$-th components[5] and as the identity on the remaining components. Then the algebra generated by the operators $X_A(1)$ for A in $\mathscr{A}$ is $\mathscr{J}_2(\mathscr{A})$ and the algebra generated by

$$\{X_A(1), \Delta_A(1) : A \in \mathscr{A}\}$$

is $\mathscr{J}_3(\mathscr{A})$.[6]

[5] identify $V \otimes V$ with $\mathrm{End}(V)$

[6] the indexing is coming from the theory of braid groups

23.6 Modules for Jaeger algebras

Assume $\mathscr{A}$ is the Bose-Mesner algebra of some association scheme on n vertices. Let $e_1, \ldots, e_n$ denote the standard basis. It is easy to see that the matrices

$$\{ue_r^T : u \in \mathbb{R}^n\}$$

form a $\mathscr{J}_3$-submodule of $\mathrm{Mat}_{n\times n}(\mathbb{R})$, and that $\mathrm{Mat}_{n\times n}(\mathbb{R})$ is the direct sum of such submodules.

Part IV

Geometry

24
Isoclinic Subspaces, Covers and Codes

24.1 Isoclinic Subspaces

Let U and V be two k-dimensional subspace of an inner product space W, and let P and Q be the corresponding orthogonal projections. Then P maps the unit sphere in V to an ellipsoid in U. The shape of this ellipsoid is determined by the extreme points of the function

$$\|Pv\|^2 = v^* P^* P v = v^* P v,$$

where v runs over the unit vectors in V. We say that V is *isoclinic* to U is there is a constant λ such that

$$v^* P v = \lambda v^* v.$$

If V is isoclinic to U with parameter λ, then

$$x^* Q^* P Q x = \lambda x^* Q^* Q X = \lambda x^* Q x$$

for all x in w. Hence we see see that U and V are isoclinic with parameter λ if and only if

$$QPQ = \lambda Q.$$

Thus we have translated a geometric condition into a linear algebraic one. Our next result shows that is a symmetric relation.

24.1.1 Lemma. *The subspace U is isoclinic to V if and only if V is isoclinic to U.*

Proof. Let R be a matrix whose columns form an orthonormal basis for U, and let S be a matrix whose columns form an orthonormal basis for V. Then

$$RR^* = P, \quad SS^* = Q$$

and

$$QPQ = SS^* RR^* SS^* = S(S^* RR^* S)S^*.$$

If $QPQ = \lambda Q$, then it follows that

$$\lambda SS^* = S(S^*RR^*S)S^*$$

and therefore

$$\lambda I = S^*S(S^*RR^*S)S^*S = S^*RR^*S.$$

Hence $R^*SS^*R = \lambda I$ and so

$$\lambda P = \lambda RR^* = R(R^*SS^*R)R^* = PQP. \qquad \square$$

Note that $\mathrm{tr}(PQP) = \mathrm{tr}(QPQ)$, and so if $\mathrm{rk}(P) = \mathrm{rk}(Q)$ and $QPQ = \lambda P$, then $PQP = \lambda P$. A consequence of the proof is that U and V are isoclinic if and only the matrix $\lambda^{-1/2}R^*S$ is orthogonal.

As exercises, prove that if P and Q are projections then $(P-Q)^2$ commutes with P and Q. Also if U and V are isoclinic with parameter λ, then

$$(P-Q)^3 = (1-\lambda)(P-Q).$$

This implies that the eigenvalues of $P-Q$ are

$$0,\ \pm\sqrt{1-\lambda};$$

since $\mathrm{tr}(P-Q) = 0$, the non-zero eigenvalues have equal multiplicity.

24.2 Matrices

We investigate sets of pairwise isoclinic k-subspaces in $\mathbb{R}^n$. Let U be the column space of the matrix

$$R = \begin{pmatrix} I_k \\ 0 \end{pmatrix}.$$

Suppose S is the $n \times k$ matrix

$$S = \begin{pmatrix} Y \\ Z \end{pmatrix}$$

where $S^*S = I_k$. Then the column spaces of R and S are λ-isoclinic if and only if

$$\lambda I = S^*RR^*S = Y^*Y.$$

Since

$$I = S^*S = Y^*Y + Z^*Z$$

we then have $Z^*Z = (1-\lambda)I$. If

$$T = \begin{pmatrix} \lambda^{1/2} I \\ \lambda^{-1/2} ZY^* \end{pmatrix}$$

then $T = \lambda^{-1/2}SY^*$, so $\mathrm{col}(T) = \mathrm{col}(S)$ and $T^*T = I$.

24.2.1 Lemma. *If V is λ-isoclinic to the column space of*

$$\begin{pmatrix} I_k \\ 0 \end{pmatrix}$$

then V is the column space of a matrix

$$\begin{pmatrix} \lambda^{1/2} I_k \\ \lambda^{-1/2} Z \end{pmatrix}$$

where $Z^* Z = (1-\lambda) I$. □

Assume $Y^* Y = Z^* Z = (1-\lambda) I$. Then the column spaces of the matrices

$$\begin{pmatrix} \lambda^{1/2} I_k \\ \lambda^{-1/2} Y \end{pmatrix}, \quad \begin{pmatrix} \lambda^{1/2} I_k \\ \lambda^{-1/2} Z \end{pmatrix}$$

are ν-isoclinic if and only if the matrix

$$\nu^{-1/2}(\lambda I + \lambda^{-1} Y^* Z)$$

is orthogonal.

24.3 Equiangular Subspaces

Suppose that $P_1, \ldots, P_m$ are projections onto e-dimensional subspaces of d-dimensional vector space. We say that they are *equiangular* if there is a scalar α^2 such that

$$\operatorname{tr}(P_i P_j) = \alpha^2$$

whenever $i \neq j$. We note that

$$\operatorname{tr}(P-Q)^2 = 2e - 2\operatorname{tr}(PQ)$$

where $\operatorname{tr}(P-Q)^2$ is the Euclidean distance between the matrices P and Q. So we could have used "equidistant" in place of "equiangular".

24.3.1 Lemma. *An equiangular set of projections is linearly independent.*

Proof. Suppose we have scalars $c_1, \ldots, c_m$ such that

$$0 = \sum_i c_i P_i.$$

Then

$$0 = \sum_i \operatorname{tr}(P_r P_i) = c_r e + \alpha^2 \sum_{i \neq r} c_i = e(c_r - \alpha^2) + \alpha_2 \sum_i c_i.$$

From this we deduce that c_r is independent of r and hence that $c_r = 0$ for all r. □

The projections P_i are Hermitian and so, if we work over $\mathbb{C}$, they lie in a real vector space of dimension d^2. Over $\mathbb{R}$ they lie in a space of dimension $d(d+1)/2$. These upper bounds are known as the *absolute bounds*. The bound supplied by the following theorem is the *relative bound*.

24.3.2 Theorem. *If the projections $P_1, \dots, P_m$ are equiangular with angle α^2 and $d\alpha^2 \le e$, then*

$$m \le \frac{d(e-\alpha^2)}{e^2 - d\alpha^2},$$

equality holds if and only if

$$\sum_i P_i = \frac{me}{d} I.$$

Proof. We set

$$S := \sum_i \Big(P_i - \frac{e}{d} I \Big)$$

Then $S = S^*$ and therefore $\operatorname{tr}(S^2) \ge 0$, which yields

$$\begin{aligned} 0 &\le \sum_i \operatorname{tr}\Big(P_i - \frac{e}{d} I \Big)^2 + \sum_{i \ne j} \operatorname{tr}\Big[\Big(P_i - \frac{e}{d} I \Big) \Big(P_j - \frac{e}{d} I \Big) \Big] \\ &= m\Big(e - \frac{e^2}{d} \Big) + m(m-1)\Big(\alpha^2 - \frac{e^2}{d} \Big). \end{aligned}$$

Our bound follows from this. If equality holds that $\operatorname{tr}(S^2) = 0$ and therefore $S = 0$. □

If P and Q are projections onto isoclinic spaces with parameter λ, then

$$\lambda e = \operatorname{tr}(\lambda P) = \operatorname{tr}(PQP) = \operatorname{tr}(PQ) = \alpha^2.$$

Thus $\lambda = \alpha^2/e$ and our expression for m becomes

$$m = \frac{d(1-\lambda)}{e - d\lambda}.$$

This bound (for equi-isoclinic subspaces) is due to Lemmens and Seidel. They also note that the absolute bound cannot be tight if $e > 1$, because the projections P_i lie in the subspace of mappings Q such that $P_1 Q P_1$ is a scalar multiple of P_1 and this has codimension $e(e+1)/2$.

A set $P_1, \dots, P_m$ of projections with rank e such that

$$\sum P_i = \frac{me}{d}$$

is known as a *tight fusion frame*. If $e = 1$, it is a *tight frame*.

If R_i is a matrix whose columns form an orthonormal basis for the column space of P_i, then

$$P_i = R_i R_i^*.$$

So if $\sum_i P_i = (me/d) I$, then

$$\frac{me}{d} I = \sum_i R_i R_i^*.$$

If $\mathscr{R}$ denotes the $d \times me$ matrix

$$\begin{pmatrix} R_1 & \dots & R_m \end{pmatrix}$$

then

$$\mathscr{R}\mathscr{R}^* = \sum_i R_i R_i^* = \frac{me}{d} I$$

and accordingly $\mathscr{R}^*\mathscr{R}$ is a scalar multiple of a projection of order $me \times me$. (It has a block decomposition where the ij-block is $R_i^* R_j$; this block is a scalar multiple of an orthogonal matrix.

24.4 Error Correction

Let $\mathscr{C}$ be an e-dimensional subspace of $\mathbb{C}^d$. A matrix A in $U(d)$ is *detectable* if for any two vectors x and y in $\mathscr{C}$, we have $\langle x, y\rangle = 0$ if and only if $\langle x, Ay\rangle = 0$. We note that $A^{-1} = A^*$ is detectable if and only if A is.

If $x, y \in \mathscr{C}$, then

$$\langle x, Ay\rangle = \langle Px, APy\rangle = \langle x, PAPy\rangle$$

and here PAP maps $\mathscr{C}$ to itself. Therefore A is detectable if and only if PAP maps $x^\perp \cap \mathscr{C}$ into itself, for each x in $\mathscr{C}$.

24.4.1 Theorem. *Let $\mathscr{C}$ be an e-dimensional subspace of $\mathbb{C}^d$, where $e \geq 3$. A matrix A is detectable if and only if U and AU are isoclinic.*

Proof. Let P represent orthogonal projection onto U and assume A is detectable. Then PAP fixes $\mathscr{C}$ and fixes the subspace $x^\perp \cap \mathscr{C}$ for each x in $\mathscr{C}$. Thus it fixes each hyperplane in $\mathscr{C}$, and therefore it must be a scalar matrix. If the columns of R are an orthonormal basis for $\mathscr{C}$, we have

$$\alpha I = PAP = RR^* ARR^*$$

and hence $R^* AR = \lambda I$. Now

$$PA^* PAP = RR^* A^* RR^* ARR^* = \alpha\bar{\alpha} P$$

and we conclude that $\mathscr{C}$ and $A\mathscr{C}$ are isoclinic.

We turn to the converse. Assume $A \in U(d)$ and $\mathscr{C}$ and $A\mathscr{C}$ are isoclinic. Assume further that R is a $d \times e$ matrix whose columns form an orthonormal basis for $\mathscr{C}$. Then RR^* represents projection onto $\mathscr{C}$ and $A^* RR^* A$ represents projection onto $A\mathscr{C}$. Therefore

$$\lambda RR^* = RR^* A^* RR^* ARR^*$$

and accordingly

$$\lambda I = R^* A^* RR^* AR.$$

This implies that $\lambda^{-1/2} R^* AR$ is unitary. Assume $x = Rw$ and $y = Rz$. Then $x, y \in \mathscr{C}$ and $x \perp y$, then $x^* Ay = 0$ (because A is unitary). We conclude that A is detectable. □

24.5 Isoclinic Subspaces from Covers

Two subsets of $\mathbb{C}^d$ are *congruent* if there is a unitary mapping that takes the first subset to the second. If the subsets are finite, and are given as the columns of matrices M and N, then they are congruent if and only if there is a unitary matrix A and a permutation matrix P such that $AMP = N$.

24.5.1 Lemma. *Two spanning sets of vectors $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$ are congruent if and only if their Gram matrices are permutation equivalent.* □

Proof. Let U_1 and U_2 be the matrices with the vectors $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$ repectively as columns. Reordering the columns of U_1 as needed, the two sets of vectors are congruent if and only if there is an orthogonal matrix Q such that $QU_1 = U_2$. If such Q exists,

$$U_2^T U_2 = U_1^T Q^T Q U_1 = U_1^T U_1$$

and the Gram matrices are equal.

So now we assume that $U_1^T U_1 = U_2^T U_2$. Since our vectors span, the rows of U_1 are linearly independent and hence U_1 has a right inverse R. Then

$$I = R^T U_1^T U_1 R = R^T U_2^T U_2 R$$

and therefore $Q = U_2 R$ is orthogonal.

Next, since $U_1 R = I$, the matrix RU_1 is idempotent and, as $U_1 R U_1 = U_1$, it acts as the identity on the row space of U_1.

We now note that, since $U_1^T U_1 = U_2^T U_2$, the row spaces of U_1 and U_2 are equal. Therefore

$$QU_1 = U_2 R U_1 = U_2.$$ □

A set of vectors $x_0, \ldots, x_r$ in $\mathbb{R}^d$ forms a *regular r-simplex* if its Gram matrix is a non-zero scalar multiple of $rI_r - J_r$. The vectors $x_0, \ldots, x_r$ are the *vertices* of the simplex. The span of a regular r-simplex has dimension $r-1$. Any two regular r-simplices are congruent, in fact any bijection from the vertices of one simplex to the vertices of the other extends to an orthogonal mapping (by the previous lemma).

Suppose $\mathscr{C}$ and $\mathscr{D}$ are subspaces with dimension e, with associated projections P and Q respectively. Then $\mathscr{C}$ and $\mathscr{D}$ are isoclinic if the restriction of P to $\mathscr{D}$ is a scalar multiple of an orthogonal operator.

We can construct isoclinic subspaces from antipodal distance-regular graphs. Suppose X is distance-regular on n vertices. Assume θ is an eigenvalue of X with multiplicity d and corresponding spectral idempotent E. If $u \in V(X)$, then the map $u \mapsto Ee_u$ assigns a vector in $\mathbb{R}^m$ to each vertex of X—we call it a representation of X on the θ-eigenspace of X. Since X is distance regular, $E_{u,v}$ is determined by the distance between x and y in X, in particular the vectors Ee_u all have same length (namely $\sqrt{d/n}$).

24.5.2 Theorem. *Let X an antipodal distance-regular graph with fibres of size r and let θ be an eigenvalue of X that is not an eigenvalue of the quotient. Then the images of the fibres under the representation on the θ-eigenspace are pairwise isoclinic subspaces of dimension $r-1$. The parameter of isoclinism is determined by the distance between the fibres in X.*

Proof. Let F be a fibre with vertices $1,\dots,r$, set $E = E_\theta$ and $y_i = Ee_i$ for $i \in F$. Let $\mathscr{F}$ denote the span of the vectors Ea_i.

As

$$\sum_i y_i = (A_d + I)y_1$$

we see that

$$\sum_i y_i = E(A_d + I)y_1.$$

Since $r^{-1}(A_d + I)$ is an idempotent, representing projection onto the space of vectors constant on the fibres of X, it follows that $E(I + Y_d) = 0$. We conclude that vectors y_i sum to zero and, since the vertices in F are pairwise equidistant, their image is a regular simplex.

Suppose b is a vertex in X at distance i from F and that $2i < D$. Set $x = Ee_b$. Assume b is at distance i from a_1; then it is at distance $D-i$ from each of $a_2,\dots,a_r$. Accordingly

$$0 = \langle x, \sum_{i=1}^{r} y_i\rangle = \langle x, y_1\rangle + (r-1)\langle x, y_2\rangle$$

and similarly

$$0 = \langle y_1, \sum_{i=1}^{r} y_i\rangle = \langle y_1, y_1\rangle + (r-1)\langle y_1, y_2\rangle$$

Now we calculate that

$$x - \frac{\langle y_1, x\rangle}{\langle y_1, y_1\rangle} y_1$$

is orthogonal to the vectors $y_1 - y_i$ for $i = 2,\dots,r$, and therefore the vector

$$\frac{\langle y_1, x\rangle}{\langle y_1, y_1\rangle} y_1$$

is the projection of x onto $\mathscr{F}$.

Since each vertex in the fibre of x is at distance i from a vertex in F, we deduce that orthogonal projection P onto $\mathscr{F}$ maps the regular simplex spanned by the fibre of X onto $\alpha = \langle y_1, x\rangle / \langle y_1, y_1\rangle$ times the image of F. Therefore the restriction of $\alpha^{-1}P$ to the span of the fibre of x is an orthogonal mapping, and so the spans of two fibres at distance i are isoclinic.

If $2i = d$, then x is at the same distance from each vertex in $\mathscr{F}$, whence $\langle x, y_i\rangle = 0$ and therefore the images of distinct fibres are orthogonal subspaces—still isoclinic. □

If X is a distance-regular antipodal r-fold cover of Y, then fibres in the preimage of a clique in Y give rise to a set of equi-isoclinic subspaces of dimension $r-1$.

A distance-regular antipodal r-fold cover of $K_{n,n}$ has diameter four. It follows that that the images of the fibres corresponding to vertices in one of the colour classes are pairwise orthogonal. The eigenvalues of this cover are the eigenvalues of $K_{n,n}$ and $\pm\sqrt{n}$, each with multiplicity $(r-1)n$. Hence the images of the fibres in a given colour class form an orthogonal decomposition of $\mathbb{R}^{(r-1)n}$ into n subspaces of dimension $(r-1)$.

24.6 Equi-isoclinic Subspaces and Unitary Covers

Let $\mathscr{C}_1,\dots,\mathscr{C}_m$ be a set of pairwise λ-isoclinic e-dimensional subspaces of $\mathbb{C}^d$, and let $R_1,\dots,R_m$ be $d\times e$ matrices such that $R_i^*R_i = I_e$ and $P_i = R_iR_i^*$ is the projection onto $\mathscr{C}_i$. Let G be the $me\times me$ block matrix with ij-block equal to $R_i^*R_j$; we might privately think of G as a kind of Gram matrix.

We see that $G^* = G$. The projections $P_1,\dots,P_m$ form a tight fusion frame if and only if G is a scalar times an idempotent. The subspaces $\mathscr{C}_1,\dots,\mathscr{C}_m$ are pairwise isoclinic if and only if each block is a scalar times a unitary matrix. If the subspaces are pairwise λ-isoclinic, then each off-diagonal block of $\lambda^{-1/2}(G-I)$ is unitary. Thus a set of m pairwise equi-isoclinic d-dimensional subspaces determines a map d from the arcs of the complete graph K_m into the unitary group into the unitary group, such that $f(i,j)f(j,i) = 1$ for each arc ij. We call it a *unitary arc function* on K_m. We extend f to a function on the walks in K_m: if $w = v_0\cdots v_n$ is a walk, then

$$f(w) = f(v_0v_1)\cdots f(v_{n-1}v_n).$$

If $A_1,\dots,A_m$ are matrices from $U(d)$, then the function

$$(i,j)\mapsto A_i^*f(i,j)A_i$$

is a function on the arcs that takes the same value on closed walks that f does. The corresponding block matrix G is similar to G. It follows that we may assume that f takes the value I on the arcs from a spanning tree, in which case we say the function is *normalized*. In particular we may assume that

$$f(1,i) = I = f(i,1)$$

for all $i\neq 1$.

The reduced closed walks at a given graph form the fundamental group of the graph; a normalized unitary arc function determines a homomorphism from the fundamental group into the unitary group. Hence it gives a unitary representation of the fundamental group.

24.7 Lines from Subspaces

If x and y are unit vectors and $\langle x,y\rangle\langle y,x\rangle = \lambda$

$$xx^* yy^* xx^* = x\langle x,y\rangle\langle y,x\rangle x^* = \lambda xx^*$$

and so the spans of x and y are 1-dimensional λ-isoclinic subspaces.

24.7.1 Lemma. *Let $\mathscr{C}$ and $\mathscr{D}$ be a pair of λ-isoclinic subspaces and let x and y be unit vectors such that $x \in \mathscr{C}$ and $y \in \mathscr{D}$ and $|\langle x,y\rangle|^2 = \lambda$. If P represents orthogonal projection onto $\mathscr{C}$, then $Py = \langle x,y\rangle x$.*

Proof. Set $\gamma = \langle x,y\rangle$. We have

$$\langle Py - \gamma x, Py - \gamma x\rangle = \langle Py, Py\rangle - \gamma\langle Py, x\rangle - \bar{\gamma}\langle x, Py\rangle + \gamma\bar{\gamma}\langle x,x\rangle.$$

Here, because $\mathscr{C}$ and $\mathscr{D}$ are λ-isoclinic,

$$\langle Py, Py\rangle = \lambda\langle y,y\rangle = \lambda,$$

and

$$\gamma\langle Py, x\rangle = \gamma\langle y, Px\rangle = \gamma\langle y,x\rangle = \lambda,$$

similarly $\bar{\gamma}\langle x, Py\rangle = \bar{\gamma}\langle x,y\rangle = \lambda$. Hence $\langle Py - \gamma x, Py - \gamma x\rangle = 0$. □

The following result is an extension of a result from Lemmens and Seidel. It gives a necessary condition for a set of equi-isoclinic subsapces to contain a set of equiangular lines.

24.7.2 Theorem. *Let $\mathscr{C}_1,\dots,\mathscr{C}_m$ be a set of pairwise λ-isoclinic subspaces in $\mathbb{C}^d$, with associated projections $P_1,\dots,P_m$. Let $R_1,\dots,R_m$ be matrices with orthonormal columns such that $P_i = R_i R_i^*$. Let f denote the corresponding unitary arc function. If $z_1,\dots,z_m$ are unit vectors such that $z_i \in \mathscr{C}_i$ and*

$$\langle z_i, z_j\rangle\langle z_j, z_i\rangle = \lambda, \quad i \neq j,$$

then for any closed walk w on K_m starting at vertex 1, the vector $R_1^* z_1$ is an eigenvector for $f(w)$.

Proof. We have $f(i,j) = R_i^* R_j$ for each arc (i,j). Now

$$P_1 P_{i_1}\cdots P_{i_k} P_1 z_1 = z_1 z_1^* z_{i_1} z_{i_1}^* \cdots z_{i_k} z_{i_k}^* z_1 = \gamma z_1$$

and

$$P_1 P_{i_1}\cdots P_{i_k} P_1 z_1 = R_1 R_1^* (P_{i_1}\cdots P_{i_k}) R_1 R_1^* z_1,$$

it follows that $R_1^* z_1$ is an eigenvector for the product

$$f(1,i_1)\cdots f(i_k,1).$$ □

This result tells us that if a set of equi-isoclinic subspaces contains a set of equiangular lines, then the group generated by the arc function on closed walks has a 1-dimensional invariant subspace. Equivalently it has a non-trivial linear representation.

25
Unitals

Let V be a vector space over $\mathbb{F}$ and let σ be an automorphism of $\mathbb{F}$ with order two. A form $\langle x, y\rangle$ on V is *unitary* relative to σ if it is linear in the second variable and

$$\langle y, x\rangle = \langle x, y\rangle^\sigma.$$

Note that the form is semilinear in the first variable:

$$\langle cx, y\rangle = c^\sigma \langle x, y\rangle.$$

If $x \in V$ then $x^\perp$ is defined by

$$x^\perp := \{y : \langle x, y\rangle = 0\}.$$

Similarly we define $U^\perp$ for a subspace U of V. If $a^\perp = V$ if and only if $a = 0$, we say that the form is non-degenerate. The happens if and only if A is invertible. If the form is non-degenerate, then $(U^\perp)^\perp = U$ and $\dim(U^\perp) = \dim(V) - \dim(U)$.

We give a construction. An $n \times n$ matrix A over $\mathbb{F}$ is σ-Hermitian if

$$(A^\sigma)^T = A.$$

Thus the identity matrix is σ-Hermitian. We define a form on the vector space $\mathbb{F}^n$ by

$$\langle x, y\rangle := (x^\sigma)^T A y.$$

Then

$$\langle x, y\rangle^\sigma = x^T A^\sigma y^\sigma = (y^\sigma)^T (A^\sigma)^T x$$

and therefore if A is σ-Hermitian, then

$$\langle x, y\rangle^\sigma = \langle y, x\rangle.$$

This form is non-degenerate if and only if A is invertible. It is semilinear in the first variable and linear in the second. If A is σ-Hermitian and invertible, we call $\langle x, y\rangle$ a σ-Hermitian form on V. (And before long we will drop the reference to σ.)

The set of vectors x such that $\langle x, x\rangle = 0$ is called a *Hermitian variety*.

We are only concerned with finite fields, and in this case any field automorphism of order two arises as the q-th power map on a field of order q^2. We may take $A = I$, and then

$$\langle x, y\rangle = \sum_i x_i^q y_i.$$

If $\mathbb{F}$ is our field and $\mathbb{F}_0$ is its subfield of order q, then the map $x \mapsto x^{q+1}$ is the norm relative to $\mathbb{F}_0$; it is a homomorphism from $\mathbb{F}^*$ onto $\mathbb{F}_0^*$. We denote the norm of x by $N(x)$ or, if more precision is needed, by $N_{\mathbb{F}/\mathbb{F}_0}(x)$.

25.1 Uniqueness

We show that unitary forms are unique, up to a change of basis.

25.1.1 Lemma. *Suppose γ is a non-degenerate unitary form on V. If U is a subspace of V and $U \cap U^\perp = 0$, then the restriction of γ to U is non-degenerate.*

Proof. Let γ_0 and γ_1 denote the restriction of γ to U and $U^\perp$ respectively. Suppose $x \in U$ and $\langle x, u\rangle = 0$ for all u in U. If $v \in U^\perp$, then

$$\gamma(x, au + bv) = a\gamma(x, u) + b\gamma(x, v) = 0.$$

As $V = U + U^\perp$, it follows that $\gamma(x, y) = 0$ for all y in V, and therefore $x = 0$. Thus γ_1 is non-degenerate, and the second claim follows similarly. □

A set of vectors S in V is orthogonal relative to the unitary form γ if $\gamma(x, y) = 0$ for pair of distinct vectors x and y from S. An orthogonal set of vectors S is *orthonormal* if $\gamma(x, x) = 1$ for each element x of S. A vector x is isotropic if $\gamma(x, x) = 0$.

25.1.2 Theorem. *If γ is a non-degenerate unitary form on a vector space V over a finite field, then V has an orthonormal basis.*

Proof. We first show that there is a non-isotropic vector x_1. Assume by way of contradiction that $\gamma(x, x) = 0$ for all x in V. Then for all x and y

$$0 = \gamma(x + y, x + y) = \gamma(x, x) + \gamma(y, y) + \gamma(x, y) + \gamma(y, x) = \gamma(x, y) + \gamma(y, x).$$

If $a \in \mathbb{F}$, then

$$0 = \gamma(x, ay) + \gamma(ay, x) = a\gamma(x, y) + a^\sigma\gamma(y, x)$$

and hence

$$(a^\sigma - a)\gamma(x, y) = 0$$

for all a in $\mathbb{F}$. Therefore $\gamma(x, y) = 0$ and γ is the zero form.

Thus there is non-isotropic vector x_1 in V. Let U denote the span of x_1. Then the restriction of γ to U is non-degenerate, as is its restriction to $U^\perp$.

By induction on its dimension, $U^\perp$ has an orthogonal basis, and the union of this basis with x_1 is orthogonal.

We convert our orthogonal basis to an orthonormal basis. We have

$$\gamma(ax, ax) = aa^\sigma\gamma(x, x) = a^{q+1}\gamma(x, x).$$

Since $\gamma(x, x) \in \mathbb{F}_0$ and since the norm map is onto, if $x \neq 0$ we can choose a so $\gamma(x, x) = 1$. Therefore V has an orthonormal basis. □

It follows that if $\mathbb{F}$ is finite, then we are free to assume that our unitary form is given by

$$\langle x, y\rangle = \sum_i x_i^{1+q} y_i.$$

25.2 Hermitian Geometry in the Plane

We are interested in the geometry of the isotropic points of a unitary polarity on a vector space over a finite field.

Although projective lines are simple, we need to determine what happens with them. Suppose $\dim(V) = 2$ and γ is a unitary polarity on V. Assume $\{x, y\}$ is an orthonormal basis for V. Then

$$\gamma(x + ty, x + ty) = 1 + t^{q+1}$$

and hence there are $q + 1$ elements t (in our field of order q^2) such that $x + ty$ is isotropic. Suppose there is no orthonormal basis. Then γ is degenerate and there is a point y such that $\gamma(y, x) = 0$ for all x. If γ is not the zero form, then there is a non-isotropic point x such that $\gamma(x, y) = 0$ and consequently

$$\gamma(tx + y, tx + y) = t^{q+1}\gamma(x, x) + t^q\gamma(x, y) + t\gamma(y, x) = t^{q+1}\gamma(x, x) \neq 0.$$

Hence y is the only isotropic point on ℓ (and $\ell \subseteq y^\perp$). Thus we have:

25.2.1 Lemma. *If ℓ is a line in a projective space over a field of order q^2 with a unitary polarity, then the number of isotropic points on ℓ is 1, $q + 1$ or $q^2 + 1$.*

25.2.2 Lemma. *Let γ be a vector space of dimension d with a non-degenerate unitary polarity. If U is a singular subspace of V, then* $2\dim(U) \leq \dim(V)$.

Proof.

A line with exactly $q + 1$ isotropic points on it called a *hyperbolic line*. A line containing exactly one isotropic point is an absolute line. For suppose ℓ is absolute and x is an isotropic point on ℓ. If $x = \ell^\perp$ then x is the unique isotropic point on ℓ. If $x \neq \ell^\perp$ then x and $\ell^\perp$ are distinct orthogonal isotropic points on ℓ, and therefore all points on ℓ are isotropic.

We turn next to projective planes.

25.2.3 Theorem. *If γ is a non-degenerate unitary polarity on the projective plane over a field of order q^2, then it has exactly q^3+1 isotropic points.*

Proof. The first step is to show that there are at least two isotropic points. By Lemma 25.2.1 there is an isotropic point x. Suppose ℓ is a line on x not equal to $x^\perp$. Then ℓ is not absolute and so contains exactly $q+1$ absolute points. Therefore the total number of absolute points is $1+q^3$. □

25.3 A Distance-Regular Graph

We return to the graph of Section **??**, and show that in the case of a unitary polarity it is distance regular. Denote it by X and assume our field has order q^2 with $q>2$.

Suppose x and y are distinct non-absolute points. Then all common neighbours of x and y lie on both $x^\perp$ and $y^\perp$. If x and y are distinct, these lines are distinct and so $x^\perp \cap y^\perp$ is a point, z say. If z is not absolute then x and y have exactly one common neighbor, namely z.

Suppose $z = x^\perp \cap y^\perp$ is absolute, let H denote the set of all absolute points and let V denote the set of all non-absolute points. There are q^2-q non-absolute points on $y^\perp$. There $q+1$ absolute lines on x, one of which contains z. If $w \in y^\perp$ and $p = x^\perp \cap w^\perp$ is absolute, then $x \vee w$ is an absolute line on x. So there are at least q^2-2q points w on $y^\perp$ such that $x \vee w$ is not absolute, and therefore x and y are distance three. Accordingly the diameter of our graph is three.

Since $a_1 = c_2 = 1$,

$$A_1^2 = kI + A_1 + A_2$$

and thus

$$A_2 = A_1^2 - A_1 - kI.$$

Further

$$A_3 = J - I - A_1 - A_2$$

and since J is a cubic polynomial in A_1, we see that A_3 is also a cubic polynomial in A_1. Consequently A_1 is a distance-regular graph with diameter three. Its matrix of eigenvalues is

$$\begin{pmatrix} 1 & q^2-q & (q^2-q)(q^2-q-2) & (q+1)(q^2-1) \\ 1 & q & 0 & -1-q \\ 1 & -1 & q-q^2 & q^2-q \\ 1 & -q & 2q & -1-q \end{pmatrix}.$$

26
Semifields and Relative Difference Sets

We show how to construct relative difference sets from a class of translation planes.

26.1 Affine Planes

Let $\mathbb{F}$ be a field of order q. (Everything in this section works over any field, but assuming that $\mathbb{F}$ is finite will allow us to be more concrete.) The points of the affine plane over $\mathbb{F}$ are the q^2 elements of $\mathbb{F} \times \mathbb{F}$. We define the lines as follows. If $m, c \in \mathbb{F}$, the set of points

$$\{(x, mx + c) : x \in \mathbb{F}\}$$

is the line with *slope* m. If $a \in \mathbb{F}$, the set of points

$$\{(a, x) : x \in \mathbb{F}\}$$

forms a line with infinite slope. The set of lines with given slope forms a parallel class. We have q lines in each parallel class and $q + 1$ parallel classes.

We can construct a bipartite graph X as follows. The two classes of vertices are the q^2 affine points and the q^2 lines with finite slope. A point is incident with the lines that contain it; hence the graph is regular with valency q. It is not hard to verify that X is a q-fold cover of $K_{q,q}$. The antipodal classes are the point sets of the lines with infinite slope and the parallel classes of lines.

We define automorphisms $h_{a,b}$ of the affine plane by

$$h_{a,b} : (x, y) \mapsto (x + a, y + ax + b).$$

It can be shown that these maps form an abelian group and gives rise to a relative difference set, but we do not do so now because we will derive a more general result later.

26.2 Translation Planes

We introduce a more general class of affine planes. Let $\mathbb{F}$ be a field of order q, and let V be a vector space over $\mathbb{F}$ with dimension k. Let Σ be a set of $q^k - 1$ non-zero endomorphisms of V. We define an incidence structure as follows. The points are the elements of $V \times V$. If $\sigma \in \Sigma \cup 0$ and $c \in V$, then the set

$$\{(x, \sigma x + c) : x \in V\}$$

is a line with finite slope. We denote this line by $[\sigma, c]$. If $a \in V$, the set

$$\{(a, x) : x \in \mathbb{F}\}$$

is a line with infinite slope. Since $|V| = q^k$, our incidence structure has q^{2k} points. There are q^{2k} lines with finite slope and q^k lines of infinite slope.

What conditions on the elements of Σ guarantee that this incidence structure is an affine plane or order q^k? Suppose $\rho, \sigma \in \Sigma$ and $c, d \in \mathbb{F}$. Then

$$\rho x + c = \sigma x + d$$

if and only if

$$(\rho - \sigma)x = d - c.$$

Hence two lines of different finite slope will have exactly one point in common if and only if the difference of any two elements of $\Sigma \cup 0$ is invertible. This condition also implies that two lines in the same parallel class have no points in common. More generally any two distinct lines in different parallel classes have exactly one point in common, but we leave the proof of this to the reader.

If (x, c) and (y, d) are two distinct points and $x = y$, then they lie on a line of infinite slope. If $x \neq y$, then they line on the line $[\sigma, a]$ if and only if

$$\sigma x + a = c, \quad \sigma y + a = d.$$

These two equations hold if and only if

$$\sigma(x - y) = c - d.$$

If the difference of two distinct elements of $\Sigma \cap 0$ is invertible and x is a non-zero element of V, then $\rho x \neq \sigma x$ for any two distinct elements of Σ. Consequently there is a unique element σ of Σ such that $\sigma(x - y) = c - d$, and so the two given points lie on the line $[\sigma, c - \sigma x]$.

We conclude that the incidence structure derived from Σ is an affine plane if and only if the difference of any two elements of $\Sigma \cup 0$ is invertible.

We define a class of automorphisms. If $a, b \in V$, let $\tau_{a,b}$ be the map

$$\tau_{a,b}(x, y) = (x + a, y + b).$$

This gives q^{2k} automorphisms, which form an elementary abelian group of order q^{2k} which acts regularly on the points of the plane. This is known as the group of *translations* of the plane. Since

$$\sigma x + c + b = \sigma(x + a) + c + b - \sigma a,$$

it follows that a translation maps each line to a line in the same parallel class.

As in the previous section we can construct a q^k-fold cover of K_{q^k,a^k} with colour classes (x, y) (where $x, y \in V$) and (σ, a) (where $\sigma \in \Sigma \cup 0$ and $a \in V$). Note that since the translation group does not act transitively on the set of lines of finite slope, we cannot use it to construct a relative difference set.

26.3 Relative Difference Sets

Let V be a vector space of dimension k over the field of order q and let Σ be a set of $q^k - 1$ endomorphisms of V such that the difference of any two distinct elements of $\Sigma \cup 0$ is invertible. If $\sigma \in \Sigma \cup 0$ and $a \in V$, we define a map $h_{\rho,a,b}$ on $V \times V$ by

$$h_{\rho,a,b}(x, y) := (x + a, y + \rho x + b).$$

Let u be a fixed non-zero element of V and define maps $H_{\rho,b}$ by

$$H_{\rho,b} := h_{\rho,\rho u,b}.$$

We say that Σ is *commutative* if $\rho\sigma = \sigma\rho$ for all elements ρ and σ of Σ.

26.3.1 Lemma. *If $\Sigma \cup 0$ is closed under addition, then the maps $h_{\rho,a,b}$, where $\rho \in \Sigma \cup 0$ and $a, b \in V$, form a group of collineations of the affine plane determined by Σ. This group has order q^{3k}; if q is odd its exponent is p, if q is even its exponent is four.*

Proof. We have

$$\begin{aligned} h_{\sigma,c,d}\, h_{\rho,a,b}(x, y) &= h_{\sigma,c,d}(x + a, y + \rho x + b) \\ &= (x + a + c, y + \rho x + b + \sigma x + \sigma a + d) \\ &= h_{\sigma+\rho,a+c,b+\sigma a+d}(x, y), \end{aligned}$$

and therefore the maps $h_{\rho,a,b}$ are closed under composition. This identity also shows that the maps $H_{\rho,b}$ are closed under composition.

We see that $h_{\sigma,c,d}$ maps the point $(x, \tau x + f)$ to

$$(x + c, \tau x + f + \sigma x + d) = (x + c, (\tau + \sigma)(x + c) + f + d - (\tau + \sigma)c),$$

and therefore it maps lines to lines if and only if, for each element τ of $\Sigma \cup 0$,

$$\tau + \sigma \in \Sigma \cup 0.$$

It follows that the maps $h_{\rho,a,b}$ send lines to lines if and only if $\Sigma \cup 0$ is closed under addition. Therefore each element σ of Σ has an additive inverse which we denote by $-\sigma$. From this it follows that

$$h_{-\rho,-a,-b-\rho a} = h_{\rho,a,b}^{-1},$$

consequently the maps $h_{\rho,a,b}$ form a group and the maps $H_{\rho,b}$ form a subgroup.

You may verify that

$$h_{\rho,a,b}^{k}(x,y) = (x + ka, y + k\sigma x + kb + \binom{k}{2}\sigma a).$$

If the characteristic of $\mathbb{F}$ is p, then our group has exponent p if p is odd, and exponent four if $p = 2$. □

Note that the group formed by the maps $h_{\rho,a,b}$ contains the group of translations of the affine plane, which is an abelian group of order q^{2k}. The maps $H_{\rho,b}$ form a second abelian subgroup of the same order. (One consequence of this is that the larger group is is nilpotent of class two.)

26.3.2 Lemma. *Suppose Σ is commutative and $\Sigma \cup 0$ is closed under addition. Then the maps $H_{\rho,b}$, where $\rho \in \Sigma \cup 0$ and $b \in V$, form an abelian group of collineations of the affine plane determined by Σ which acts transitively on the points of the plane, and transitively on the set of lines of finite slope.*

We have

$$h_{\sigma,c,d}h_{\rho,a,b} = h_{\sigma+\rho,a+c,b+d+\sigma a}, \qquad h_{\rho,a,b}h_{\sigma,c,d} = h_{\sigma+\rho,a+c,b+d+\rho c},$$

and therefore $h_{\rho,a,b}$ and $h_{\sigma,c,d}$ commute if and only if

$$\rho c = \sigma a.$$

Then

$$H_{\sigma,d}H_{\rho,b} = H_{\sigma+\rho,b+\sigma\rho u}$$

Since $\sigma\rho = \rho\sigma$, we deduce that the maps $H_{\rho,b}$ and $H_{\sigma,d}$ form an abelian group of order q^{2k}. Since no non-identity element of this group fixes a point, it must act transitively on points. The image of the line $[\tau, f]$ under $H_{\sigma,d}$ is the line $[\tau+\sigma, f+d-(\tau+\sigma)\sigma u]$, and so unless $\sigma = 0$ and $d = 0$, it follows that $H_{\sigma,d}$ does not fix a line. Therefore our abelian group acts transitively on the lines of finite slope. □

To summarise, suppose V has dimension k over $\mathbb{F}$ and let Σ be a set of $q^k - 1$ endomorphisms of V, such that

(a) The difference of two elements of $\Sigma \cup 0$ is invertible.

(b) The set $\Sigma \cup 0$ is closed under addition.

(c) If $\rho, \sigma \in \Sigma$, then $\rho\sigma = \sigma\rho$.

Then the maps $H_{\rho,b}$ form an abelian group acting transitively on affine points and on lines of finite slope.

26.3.3 Corollary. *The set*

$$D := \{H_{\sigma,0}\}$$

is a relative difference set with parameters $(q^k, q^k, q^k, 1)$. □

If $k = 1$ then we may take Σ to be the non-zero elements of $\mathbb{F}$. In this case our affine plane is the usual affine plane of order q. Hence for each finite Pappian plane of order q, we have a relative difference set and hence a set of $q+1$ mutually unbiased bases in $\mathbb{C}^q$. (In this case the relative difference sets were first constructed by Hughes [1].)

[1]

27
Extraspecial 2-Groups

We discuss extraspecial p-groups, emphasizing the case $p = 2$.

27.1 The Frattini Subgroup

A p-group G is *extraspecial* if:

(a) G is not abelian.

(b) $Z(G) \cong \mathbb{Z}_p$.

(c) $G/Z(G)$ is an elementary abelian p-group.[1]

[1] "elementary abelian p-group" is how group theorists refer to vector spaces over $\mathbb{Z}_p$

The first two conditions are straightforward enough, but the third has some implications. First, $G/Z(G)$ is abelian if and only if $G' \leq G$ and, since G is not abelian we must have $G' = Z(G)$. Second, since $G/Z(G)$ has exponent p, we see that $Z(G)$ contains the p-th power of each element of G. The subgroup of G generated by the p-th powers of its elements is denoted ??.

There is another subgroup in play here: the *Frattini subgroup* $\Phi(G)$ of G is the intersection of the maximal subgroups of G. It is a normal, indeed characteristic, subgroup of G. We offer a second viewpoint. We call an element x of G a *non-generator* if whenever $\langle S \rangle = G$ and $x \in S$, then $\langle S \setminus x \rangle = G$. (For example, any element of order p in a cyclic group of order p^2.)

27.1.1 Lemma. *For any finite group G, the Frattini subgroup is the set of non-generators of G.* □

We write $H \trianglelefteq G$ to denote that H is

When G is a p-group, the only case we really care about, there is a second description of $\Phi(G)$. This follows from the fact that if Q is a maximal subgroup of the p-group P, then $Q \trianglelefteq P$ and $P/Q \cong \mathbb{Z}_p$. If $M_1, \ldots, M_k$ are normal subgroups of a group G, then there is an injection

$$\frac{G}{\cap_i M_i} \to \frac{G}{M_1} \times \cdots \times \frac{G}{M_k};$$

27.1.2 Lemma. *If G is a p-group, $G/\Phi(G)$ is elementary abelian.* □

Let $\Omega_p(G)$ denote the subgroup of G generated by the p-th powers of the elements of G.

27.1.3 Lemma. *If G is a p-group, $\Phi(G) = G'\Omega_p(G)$.* □

27.1.4 Theorem. *A Sylow p-subgroup of $\Phi(G)$ is a normal subgroup of G.*

Proof. Let P be a Sylow p-subgroup of $\Phi(G)$ and assume $g \in G$. Since $\Phi(G)$ is normal, $P^g \leq \Phi(G)$; since the Sylow p-subgroups are conjugate, there is an element x in $\Phi(G)$ such that $P^g = P^x$. Then $P^{gx^{-1}} = P$, and therefore $gx^{-1} \leq N_G(P)$. This shows that any element of G lies in $N_G(P)\Phi(G)$ and since $\Phi(G)$ consists of non-generators, we have $G = N_G(P)$. □

It follows that any Sylow p-subgroup of $\Phi(G)$ is normal, and hence $\Phi(G)$ is nilpotent. [2]

[2] Thus will have very little impact on us, because G will be a p-group and so it, along with each of its subgroups, is nilpotent

27.2 Commutator Calculus

If $x, y \in G$, then

$$[x, y] := x^{-1}y^{-1}xy$$

We say $[x, y]$ is the *commutator* of x and y. We note some basic properties of this operation.

(a) $[x, y]^{-1} = [y, x]$.

(b) $[x, yz] = [x, z][x, y]^z$.

(c) $[x, y^{-1}] = [y, x]^{y^{-1}}$.

This identities are not easy to work with in general. However the simplify considerably if $G' \leq Z(G)$; in particular we then have

$$[x, yz] = [x, y][x, z]$$

and thus the map $x \mapsto [a, x]$ is a homomorphism from G to G' and the kernel of this homomorphism contains $Z(G)$.

When G is an extraspecial p-group this commutator map is a bilinear map from the vector space $G/Z(G)$ to the field $Z(G)$. Since $[x, x] = 1$, this map is an alternating bilinear form on $G/Z(G)$.

Part V

Graphs

28
Non-Reconstructible Tournaments

Stockmeyer[1] constructed non-reconstructible pairs of tournaments on $2^n + 1$ and 2^{n+2}. There was an error in this paper (noted in an erratum by Stockmeyer). Subsequently Kocay wrote a paper fixing the error and providing a different proof.[2]

Prior to Stockmeyer's work, non-reconstructible pairs of tournaments on 5,6 and 8 vertices were known. We present his construction.

[1] Paul K. Stockmeyer "The falsity of the reconstruction conjecture for tournaments", *J. Graph Theory*, **1** (1977), 19–25

[2] William Kocay "On Stockmeyer's non-reconstructible tournaments", *J. Graph Theory*, **9** (1985), 473–476

28.1 A Big Tournament

Any rational number m can be written in the form

$$m = 2^k \frac{r}{s}$$

where r and s are odd. We refer to r/s as the *odd part* of m and denote it by $\text{odd}(m)$.

We define a directed graph $\mathscr{T}(\mathbb{Z})$ with vertex set $\mathbb{Z}$ by declaring (i, j) to be an arc if $\text{odd}(j - i) \cong 1$ modulo 4. Then $\mathscr{T}(\mathbb{Z})$ is a Cayley graph for $\mathbb{Z}$ (and so it is vertex transitive). It is easy to see that $(0, i)$ is an arc if and only if $(-i, 0)$ is an arc, and accordingly $\mathscr{T}(\mathbb{Z})$ is a tournament.

Define $\mathscr{C}_1$ to be the set of positive integers k such that $\text{odd}(k) \cong 1$ modulo 4 and define $\mathscr{C}_3$ to be the set of positive integers k such that $\text{odd}(k) \cong 4$ modulo 4 Then $\mathscr{C}_1$ is the connection set for our Cayley graph. We have

$$\mathscr{C}_1 \cap \mathscr{C}_3 = \emptyset, \quad \mathscr{C}_1 \cup \mathscr{C}_3 = \mathbb{N}$$

A directed graph has both a complement (with adjacency matrix $J - I - A$) and a *converse* (with adjacency matrix A^T). For a tournament, the complement and the converse are equal.

Multiplication by -1 maps $\mathscr{T}(\mathbb{Z})$ onto its converse.

There is a recursive structure to $\mathscr{T}(\mathbb{Z})$: the subtournament induced by the even vertices and the subtournament induced by the odd vertices are both isomorphic to $\mathscr{T}(\mathbb{Z})$. Multiplication by 2 is an injective endomorphism of $\mathscr{T}(\mathbb{Z})$, with the tournament induced by the even integers as its

image. Similarly the map $i \mapsto 2i-1$ is an injective endomorphism with the tournament induced by the odd integers as its image.

Further, if for $i = 0,1,2,3$ we use $\mathbb{Z}_i$ to denote the integers congruent to i modulo four, then the tournaments induced by these sets are isomorphic and each vertex in $\mathbb{Z}_i$ dominates each vertex in $\mathbb{Z}_{i+1}$.[3]

[3] with subscripts computed modulo four

A Cayley graph $X(G,\mathscr{C})$ is a *graphical regular representation* for G if $\mathrm{Aut}(X) \cong G$, equivalently if $\mathrm{Aut}(X)_1 = \langle 1\rangle$. The usual acronym is GRR. [4] Two vertices u and v in the graph X are *pseudosimilar* if $X\setminus u$ and $X\setminus v$ are isomorphic, but there is no automorphism of X that maps u to v.

[4] Obviously one is needed, much as I dislike them

28.1.1 Theorem. *Let X be a GRR for G. If g is a non-identity element of G with odd order, then the vertices g and g^{-1} are pseudosimilar in $X\setminus 1$.*

Proof. Since X is a GRR, the vertex-deleted subgraph $X\setminus 1$ is asymmetric.

Right multiplication by G maps the ordered pair of vertices $(g^{-1},1)$ to $(1,g)$, and therefore $X\setminus\{g^{-1},1\}$ and $X\setminus\{1,g\}$ are isomorphic. □

Thus if X is a GRR for a group of odd order, each vertex in $X\setminus 1$ has a pseudosimilar mate.[5] For any integer i, we have that $\mathscr{T}(\mathbb{Z})\setminus\{-i,0\}$ and $\mathscr{T}(\mathbb{Z})\setminus\{0,i\}$ are isomorphic. The only groups of odd order without TRRs are $\mathbb{Z}_3^2$ and $\mathbb{Z}_3^3$.

[5] I'm thinking that $\mathscr{T}(\mathbb{Z})$ is a TRR (yes TRR is a thing)

Question: Is the stabilizer of a vertex in $\mathrm{Aut}(\mathscr{T}(\mathbb{Z}))$ trivial?

28.2 Smaller Tournaments

Stockmeyer's tournaments are subtournaments of $\mathscr{T}(\mathbb{Z})$. If $\mathscr{T}(m)$ denotes the tournament induced by the vertices $1,\ldots,m$, then Stockmeyer uses the tournaments $\mathscr{T}(2^k)$.

We derive some basic properties of these tournaments.

28.2.1 Lemma. *We have:*

(a) $\mathscr{T}(2^k)$ is self-converse.

(b) The first 2^{n-1} vertices of $\mathscr{T}(2^k)$ have out-valency 2^{n-1}; the last 2^{n-1} vertices have out-valency $2^{n-1}-1$.

(c) The automorphism group of $\mathscr{T}(2^k)$ is trivial.

Proof. Multiplication by -1 maps $\mathscr{T}(\mathbb{Z})$ to its converse; the composition of this map with a right shift by 2^k+1 maps the set $\{1,\ldots,2^k\}$ to itself and maps $\mathscr{T}(2^k)$ to its converse.

[*** tbc ***]

28.2.2 Lemma. *If $i \in \{1,\ldots,2^k$, then $\mathscr{T}(2^k)\setminus i \cong \mathscr{T}(2^k)\setminus(2^k+1-i)$.*

28.3 Non-Reconstructible Tournaments

Let $B(2^k)$ be the tournament we get by adding a new vertex 0 that dominates the even vertices in $\mathscr{T}(2^k)$ and is dominated by the odd vertices. Let $C(2^k)$ be the tournament we get by adding a new vertex 0 that dominates the odd vertices in $\mathscr{T}(2^k)$ and is dominated by the even vertices.

28.3.1 Theorem. *The tournaments $B(2^k)$ and $C(2^k)$ are not isomorphic.*

28.3.2 Theorem. *The tournaments $B(2^k)\backslash 0$ and $C(2^k)\backslash 0$ are isomorphic. For $i = 1,\ldots,2^k$, the tournaments $B(2^k)\setminus i$ and $C(2^k)\setminus 2^k+1-i$ are isomorphic.*

For $k = 3,4,5,6$, the characteristic polynomials of the adjacency matrices of $B(2^k)$ and $C(2^k)$ differ by 1.[6] The skew symmetric adjacency matrices are related by switching on the vertex 0.

[6] I have no idea

29
Vector Colourings and Homomorphisms

The material here is based in part on my colouring notes, and on work of Roberson et al. [1]

[1]

We define a *representation* of a graph X to be map from $V(X)$ into an inner product space, often $\mathbb{R}^d$. A representation is *spherical* if all vectors in its image have the same length. The *Gram matrix* of a representation ψ is the matrix G such that $G_{i,j} = \langle \psi(i), \psi(j) \rangle$ for all vertices i and j. If ψ and ρ are representations of X, there is an orthogonal matrix mapping the image of ψ onto the iage of ρ if and only if the Gram matrices of the two representations are equal.

We do **not** assume that representations are injective.

29.1 The Vector Chromatic Number

In this section we study the *vector chromatic number* $\chi_v(X)$ of a graph X. This has two useful features—it interpolates the clique and chromatic number of X:

$$\omega(X) \leq \chi_v(X) \leq \chi(X)$$

and it can be computed in polynomial time. (The latter distinguishes it from the fractional chromatic number.)

To begin we define a family of infinite graphs. Let $S(n,a)$ denote the graph with the unit vectors in $\mathbb{R}^n$ as its vertices, where unit vectors x and y are adjacent if and only if $\langle x,y \rangle \leq a$. (We will only be concerned with the case where $a < 0$.) We say that a graph X is *vector β-colourable* if there is a homomorphism from X into $S(n,-(\beta-1)^{-1})$ for some n.

29.1.1 Lemma. *If there is a homomorphism from X to Y, then $\chi_v(X) \leq \chi_v(Y)$.* □

If $\beta < 2$, any vector β-colourable graph is empty. Since $S(n,-1)$ consists of disjoint copies of K_2, a graph is vector 2-colourable if and only if it is bipartite.

29.1.2 Lemma. *The complete graph K_n is vector n-colourable.*

Proof. Let $e_1, \ldots, e_n$ denote the standard basis vectors in $\mathbb{R}^n$. Let f_i be given by

$$f_i := e_i - \frac{1}{n}\mathbf{1}.$$

Then

$$\langle f_i, f_j \rangle = \begin{cases} 1 - \frac{1}{n}, & i = j; \\ -\frac{1}{n}, & i \neq j. \end{cases}$$

If

$$\hat{f}_i := \frac{1}{\|f_i\|} f_i,$$

then the map $i \mapsto \hat{f}_i$ is a vector n-colouring. □

29.1.3 Corollary. *If X is a graph, then $\chi_v(X) \leq \chi(X)$.* □

29.1.4 Lemma. *If X is a graph, then $\omega(X) \leq \chi_v(X)$.*

Proof. Assume

$$V(X) := \{1, \ldots, n\}$$

and that $i \mapsto x_i$ is a vector β-colouring of X. Suppose the subset C of $V(X)$ is a clique and that

$$x_C := \sum_{i \in C} x_i.$$

Then

$$\begin{aligned} 0 \leq \langle x_C, x_C \rangle &= |C| + \sum_{i \neq j} \langle x_i, x_j \rangle \\ &\leq |C| + (|C|^2 - |C|)\left(-\frac{1}{\beta - 1}\right) \\ &= |C|\left(1 - \frac{|C| - 1}{\beta - 1}\right) \end{aligned}$$

This implies that $\beta \geq |C|$. □

Given this result, we have proved that the vector chromatic number of K_n is n.

We introduce a variant of vector colouring, which is actually better known. Let $S^=(n, a)$ denote the graph with the unit vectors in $\mathbb{R}^n$ as its vertices, where unit vectors x and y are adjacent if and only if $\langle x, y \rangle = a$. If $\beta > 0$, we say that a graph X has a *strict vector β-colouring* if there is a homomorphism from X into $S^=(n, -(\beta - 1)^{-1})$ for some n. The strict vector chromatic number is the minimum possible value of β; we denote it by $\chi_{sv}(X)$. A strict vector β-colouring is a vector β-colouring; hence the strict vector chromatic number is an upper bound on the vector chromatic number.

The optimal vector colouring for K_n we constructed above is a strict colouring, whence the strict vector chromatic number is a lower bound on the chromatic number. Karger, Motwani and Sudan (arXiv:cs/9812008) proved that the strict vector chromatic number of a graph X is equal to $\theta(\overline{X})$, where θ denotes the function introduced by Lovász.

29.2 Semidefinite Optimization

Both the vector chromatic number and the strict vector chromatic number can be expressed as the value of semidefinite programs. We wil see that this provides some very powerful tools.

There is no difficulty in formulating $\chi_v(X)$ as the value of a semidefinite program:

29.2.1 Lemma. *The vector-chromatic number of X is equal to*

$$\begin{aligned} &\min\ t\\ M\circ I &= (t-1)I\\ M\circ A &\leq -A\\ M &\succcurlyeq 0. \end{aligned}$$

□

If the matrix M is a feasible solution to the above program, then $(t-1)^{-1}M$ is the Gram matrix of a vector t-colouring of X.

The dual to this program is given by:

29.2.2 Lemma. *The vector-chromatic number of X is equal to*

$$\begin{aligned} &\max \operatorname{sum}(B)\\ B\circ\overline{A} &\geq 0\\ \operatorname{tr}(B) &= 1\\ B &\succcurlyeq 0. \end{aligned}$$

□

If $e = |E(X)|$ and τ is the least eigenvalue of A, then

$$B = \frac{1}{n(-\tau)}(A-\tau I)$$

is dual feasible, with value

$$1-\frac{2e/n}{\tau}.$$

29.2.3 Lemma. *If M is primal feasible with value t and B is dual feasible with value s, then $t-s \geq \langle M,B\rangle \geq 0$.*

Proof. Since M and B are positive semidefinite, $\langle M,B\rangle \geq 0$. Further

$$\begin{aligned} \langle M,B\rangle = \operatorname{sum}(M\circ B) &= (t-1)\operatorname{sum}(B\circ I) + \operatorname{sum}((M-(t-1)I)\circ B)\\ &\leq t-1-(s-1)\\ &= t-s. \end{aligned}$$

□

An immediate consequence of this is that M and B are respectively primal and dual optimal, then $\langle M,B\rangle = 0$ and hence $MB = 0$. We will make use of this.

Given that $A-\tau I$ is $1/(-n\tau)$ times a dual feasible solution, we also deduce the following.

29.2.4 Corollary. *For any graph X we have*

$$\chi_v(X) \geq 1-\frac{2e/n}{\tau}.$$

□

29.3 Strict Vector Colouring

We can easily modify our semidefinite programs for χ_v to obtain programs for χ_{sv}. The primal is

$$\begin{aligned} &\min\ t\\ &M\circ I = (t-1)I\\ &M\circ A = -A\\ &M \succcurlyeq 0 \end{aligned}$$

and the dual is

$$\begin{aligned} &\max \operatorname{sum}(B)\\ &B\circ \overline{A} = 0\\ &\operatorname{tr}(B) = 1\\ &B \succcurlyeq 0. \end{aligned}$$

29.3.1 Lemma. *If M is primal feasible and B is dual feasible for the strict program, with values t and s respectively, then $t-s = \langle M, V\rangle$.*

Hence M and B are optimal for χ_{sv} if and only if they are feasible and $MB = 0$.

Proof.

$$\begin{aligned} \langle M, B\rangle = \operatorname{sum}(M\circ B) &= (t-1)\operatorname{sum}(B\circ I) + \operatorname{sum}((M-(t-1)I)\circ B)\\ &= t-1-(s-1)\\ &= t-s. \end{aligned}$$

□

Suppose M is primal feasible with value t and B is dual feasible. Then

$$\begin{aligned} (MB)_{i,i} &= (t-1)B_{i,i} + \sum_j M_{i,j}B_{i,j}\\ &= (t-1)B_{i,i} - \sum_{j\sim i} B_{i,j}\\ &= tB_{i,i} - (B\mathbf{1})_i. \end{aligned}$$

29.3.2 Lemma. *If B is an optimal solution for the dual spd for χ_{sv}, then $(B\mathbf{1})_i = B_{i,i}\chi_{sv}(X)$ for each vertex i.* □

This leads to a variant of the ratio bound provided by Lemma 29.2.4.

29.3.3 Lemma. *For any graph X we have*

$$\chi_{sv}(X) \geq 1 - \frac{2e/n}{\tau};$$

if equality holds then X is regular.

Proof. The inequality follows from Lemma 29.2.4 (since $\chi_{sv}(X) \geq \chi_v(X)$). Let M and B respectively be primal and dual optimal for χ_{sv}. Then $MB = 0$ and, by our calculation above,

$$\chi_{sv} B_{i,i} = (B\mathbf{1})_i$$

for each i. If equality holds in our bound

$$\frac{1}{n(-\tau)}(A - \tau I)$$

is dual optimal and therefore X is regular. □

29.4 1-Walk Regular Graphs

There is an interesting class of graphs where we can give an explicit formula for $\chi_v(X)$ (and $\chi_{sv}(X)$).

We define a graph to be *1-walk regular* if:

(a) There are constants c_k such that $A^k \circ I = c_k I$, and

(b) there are constants d_k such $A^k \circ A = d_k A$.

We note that graphs that satisfy the first these conditions are known as walk-regular, here it may be less confusing if we refer to them as 0-walk regular. A 0-walk regular graph is necessarily regular. Any graph that is vertex and edge transitive is 1-walk regular, as is any graph that is a single class in an association scheme. We characterize 1-walk regular graphs in terms of the spectral idempotents:

29.4.1 Lemma. *Let A be the adjacency matrix of X and let $E_1, \ldots, E_d$ be its spectral idempotents. Then X is 1-walk regular if and only if there are constants γ_k and δ_k such that*

$$E_i \circ I = \gamma_i I, \qquad E_i \circ A = \delta_i A.$$ □

Suppose A is 1-walk regular with valency k. If $\theta_1, \ldots, \theta_d$ are the eigenvalues of A and their respective multiplicities are $m_1, \ldots, m_d$, then

$$\gamma_i = \frac{m_i}{n}.$$

Further

$$nk\delta_i = \mathrm{sum}(\delta_i A) = \mathrm{sum}(E_i \circ A) = \mathrm{tr}(AE_i) = \theta_i m_i$$

and thus

$$\delta_i = \frac{m_i \theta_i}{nk}.$$

One consequence of these calculations is that

$$\frac{\delta_i}{\gamma_i} = \frac{\theta_i}{k}.$$

If X is 1-walk-regular and E is the projection onto its τ-eigenspace and $uv \in E(X)$, then

$$\frac{E_{u,v}}{E_{v,v}} = \frac{\tau}{k}.$$

It follows that E is the Gram matrix of a vector colouring, with value $1 - \frac{k}{\tau}$.

29.4.2 Theorem. *Suppose X is a k-regular graph with least eigenvalue τ. If X is 1-walk regular, then*

$$\chi_v(X) = \chi_{sv}(X) = 1 - \frac{k}{\tau}.$$

Proof. We have

$$\chi_{sv}(X) \geq \chi_v(X) \geq 1 - \frac{k}{\tau}.$$

We construct optimal solutions for the sdp for χ_{sv} with value $1 - \frac{k}{\tau}$. This is easy, if E denotes orthogonal projection onto the τ-eigenspace we take

$$M = \frac{nk}{m_\tau(-\tau)}E, \quad B = \frac{1}{n(-\tau)}(A - \tau I).$$

Clearly $MB = 0$ and it is easy to check that M and B are primal and dual feasible, with the required value. The theorem now follows from Lemma 29.3.1. □

29.5 Fractional Chromatic Number

We take the fractional chromatic number $\chi_f(X)$ of a graph X to be

$$\min\left\{\frac{k}{n} : X \to K_{n:k}\right\}$$

It is a theorem that this minimum always exists.

29.5.1 Lemma. *For the Kneser graphs,*

$$\chi_{sv}(K_{n:k}) = \chi_f(K_{n:k}) = \frac{n}{k}.$$

Proof. The first inequality holds because Kneser graphs are 1-walk regular, the second because $\alpha(K_{n:k}) = \binom{n-1}{k-1}$ by EKR (and for a vertex-transitive graph, $\chi_f(X) = |V(X)|/\alpha(X)$). □

If follows that:

29.5.2 Lemma. *For any graph X, we have $\chi_{sv}(X) \leq \chi_f(X)$.*

This gives a string of inequalities:

$$\omega(X) \leq \chi_v(X) \leq \chi_{sv}(X) \leq \chi_f(X) \leq \chi(X).$$

30
Quadratic Rank

The *quadratic map* q from $\mathbb{R}^m$ to $\mathbb{R}^{\binom{m+1}{2}}$ maps a vector $(u_1, \ldots, u_m)$ to the vector

$$(u_i u_j)_{i \le j}.$$

If U is an $n \times m$ matrix then $q(U)$ denotes the $n \times \binom{m+1}{2}$ matrix we get by applying q to each row of U. The *quadratic rank* of U is the rank of $q(U)$.

The quadratic rank of U is less than $\binom{m+1}{2}$ if and only if the columns of $q(U)$ are linearly independent. This happens if and only if there are scalars $b_{i,j}$, not all zero, such that for each row of B we have

$$\sum_{i,j:i \le j} b_{i,j} u_i u_j.$$

Equivalently, there is a non-zero $m \times m$ symmetric matrix B such that $u^T B u = 0$. (Note that, if $i \neq j$ then $B_{i,j} = b_{i,j}/2$.) In other terms, the quadratic rank of U is less than $\binom{m+1}{2}$ if and only if the rows of U lie on a homogeneous quadric. It follows that, if R is an invertible $m \times m$ matrix then U and UR have the same quadratic rank. Consequently the quadratic rank of U is a property of its column space, rather than of the matrix itself.

If $v_1^T, \ldots, v_n^T$ are the rows of U then the quadratic rank of U is the dimension of the space spanned by the matrices vv^T. In particular, the quadratic rank of U is $\binom{m+1}{2}$ if and only if the matrices $v_i v_i^T$ span the space of all $m \times m$ symmetric matrices.

It is often convenient to describe a representation ρ of X by an $n \times m$ matrix, U say, with rows indexed by $V(X)$. Then $\rho(v)$ is the v-row of U. Eigenspaces provide useful representations—simply choose U to be a matrix whose columns form an orthogonal basis for the given eigenspace.

30.1 Quadratic Edge-Rank

Suppose ρ is a representation of a graph X and let $\widetilde{U}$ be the matrix with rows

$$\rho(v),\ v \in V(X), \quad \rho(v) - \rho(w),\ uv \in E(X).$$

We call $\widetilde{U}$ the *edge extension* of U. Define the *quadratic edge-rank* of ρ to be the quadratic rank of $\widetilde{U}$. Note that if X has e vertices and e edges then any representation of X has quadratic edge-rank at most $v+e$.

If the quadratic edge-rank of ρ is less than $\binom{m+1}{2}$ then there is an $m \times m$ symmetric matrix B such that

$$\rho(u)B\rho(u)^T = 0, \quad \forall u \in V(X) \tag{30.1.1}$$

and

$$(\rho(u)-\rho(v))B(\rho(u)-\rho(v))^T = 0, \quad \forall uv \in E(X). \tag{30.1.2}$$

We can summarize our deliberations thus:

30.1.1 Lemma. *Let ρ be a representation of X in $\mathbb{R}^m$. The quadratic edge-rank of ρ is less than $\binom{m+1}{2}$ if and only if there is a non-zero homogeneous quadric which contains the image of each vertex of X, and contains all points on the lines that join the images of each pair of adjacent vertices.* □

30.1.2 Lemma. *Suppose θ is an eigenvalue of the symmetric matrix Q with multiplicity m, and let ρ denote the corresponding representation. The quadratic edge-rank of the θ-eigenspace of Q is equal to the dimension of the span of the matrices*

$$\rho(u)^T\rho(v) + \rho(v)^T\rho(u), \quad u \simeq v.$$ □

The rank of a matrix is equal to the largest integer k such that the determinant of some $k \times k$ submatrix is non-zero. Given this, it is not hard to see that a small perturbation of a matrix cannot increase its rank. Further, if the columns of a matrix are linearly independent then a small perturbation does not change its rank. It follows that if an $n \times m$ matrix U has quadratic rank $\binom{m+1}{2}$, then any small perturbation of U has quadratic rank $\binom{m+1}{2}$.

We note that the quadratic rank is defined for any subspace of $\mathbb{R}^n$, not just for eigenspaces. There is one simple but useful consequence of this.

30.1.3 Lemma. *Suppose W is an m-dimensional subspace of $\mathbb{R}^n$ with quadratic rank $\binom{m+1}{2}$. If W_1 is a subspace of W with dimension k, its quadratic rank is $\binom{k+1}{2}$.* □

30.2 Uniqueness

Assume X is 1-walk regular. We consider when the optimal solutions for the sdp for χ_{sv} are unique.

If $(A-\tau I)N = 0$, then the columns of N must lie in the τ-eigenspace of A. If N is symmetric and E is the projection on the τ-eigenspace, we conclude the $N = ENE$. Let U be a matrix whose columns form an orthonormal basis for the τ-eigenspace. Then $E = UU^T$ and so

$$N = ENE = U(U^TNU)U^T$$

30.2.1 Lemma. *Assume U is an $n \times m$ matrix whose columns form an orthonormal basis for the τ-eigenspace of A. If $N = N^T$ and $(A - \tau I)N = 0$, then $N = URU^T$ for some $m \times m$ symmetric matrix R.* □

Suppose M_1 and M_2 are primal optimal solutions to the sdp for the strict vector chromatic number of a 1-walk regular graph. Then

$$(A - \tau I)M_1 = (A - \tau I)M_2 = 0$$

and consequently

$$M_1 - M_2 = USU^T$$

for a symmetric matrix S. We note that

$$(M_1 - M_2) \circ (I + A) = 0,$$

and therefore if u_i denotes the i-th row of U and $i \simeq j$,

$$0 = (USU^T)_{ij} = u_i S u_j^T = \langle S, u_j^T u_i \rangle.$$

Further

$$\langle S, u_j^T u_i \rangle = \mathrm{tr}(S u_j^T u_i) = u_i S u_j^T = u_j S u_i^T = \langle S, u_i^T u_j \rangle,$$

which implies that

$$\langle S, u_j^T u_i \rangle = \frac{1}{2} \langle S, u_i \vee u_j \rangle.$$

Consequently $(USU^T)_{ij} = 0$ if and only if $\langle S, u_i \vee u_j \rangle$.

30.2.2 Theorem. *Let X be a 1-walk regular graph with least eigenvalue τ of multiplicity m. Then X has a unique strict vector t-coloring with $t = 1 - \frac{k}{\tau}$ if and only the quadrtic edge-rank of the τ-eigenspace of X is $\binom{m+1}{2}$.* □

Since the columns of U are eigenvectors for A, we have

$$\tau u_i = \sum_{j \sim i} u_j$$

and consequently

$$\sum_{j \sim i} u_i \vee u_j = \tau u_i \vee u_i.$$

Hence in computing the extended quadratic rank, we only need to consider the matrices of the form $u_i \vee u_j$ where $ij \in E(X)$ (and therefore the quadratic rank is at most $|E(X)|$).

30.3 Unique Vector Colourings and Cores

Suppose X is 1-walk regular and has a unique optimal strict vector t-colouring. Let Y denote the core of X, and let ψ be a homomorphism from X to Y. Then $\chi_{sv}(Y) = \chi_{sv}(X)$ and therefore there is a vector t-colouring

φ of Y. The composition $\varphi \circ \psi$ is then a strict vector t-colouring of X; by uniqueness the Gram matrix of this composition is the projection E onto the τ-eigenspace of X.

We say that a homomorphism is *locally injective* if it does not identify two vertices at distance two in X. (Thus any two distinct points in the same fibre of a locally injective homomorphism are distance at least three.) We recall a result of Nešetřil that any locally injective map from a graph to its core must be an isomorphism.

30.3.1 Lemma. *If every strict vector colouring of X is locally injective, then X is a core.* □

Proof. Suppose $\psi : X \to Y$ and $\chi_{sv}(X) = \chi_{sv}(Y)$. Let ρ and σ respectively be optimal strict vector colourings of X and Y. Then the composition $\sigma \circ \psi$ is an optimal vector coloring of X and, since it is locally injective, so is ψ. Now take Y to be the core of X and apply Nešetřil's result to deduce that ψ must be an isomorphism. □

30.3.2 Lemma. *If X is 2-walk regular and not bipartite, then its τ-representation is locally injective.* □

Proof. Let ρ denote the τ-representation of X. If vertices i and j are at distance two and $\rho(i) = \rho(j)$, then ρ maps any pair of vertices at distance two to the same point. But this implies that X is bipartite. □

30.3.3 Corollary. *Assume X is 2-walk regular and not bipartite, and let m be the multiplicity if τ. If the quadratic edge-rank of the τ-representation is $\binom{m+1}{2}$, then X is a core.* □

30.4 Cores of Distance-Regular Regular Graphs

We present a version of David Roberson's proof that the image any homomorphism from a strongly regular graph is either the graph itself or a clique—this the homomorphism is either an isomorphism or a colouring.

We work in greater generality than required. We assume that X and Y are distance-regular graphs with the same parameters and that $\phi : X \to Y$ is a homomorphism. The composition of ϕ with the optimal strict vector colouring of Y arising from its τ-eigenspace is an optimal strict vector colouring of X. Let F be the Gram matrix of the composite colouring and let $E = E_\tau$. Then

$$(A - \tau I)F = (A - \tau I)F = 0$$

and $(F - E) \circ (I + A) = 0$.

30.4.1 Theorem. *Assume X is a non-bipartite distance-regular graph with diameter d, and let ϕ be an endomorphism of X such that $\phi(X)$ has diameter e. If $N_e(u)$ is connected for each vertex u of X, then ϕ is an automorphism.*

Proof. Assume u and v are vertices of X at distance e and $\mathrm{dist}(\phi(u),\phi(v)) = e$. (We may take two vertices at maximum distance in $\phi(X)$.) Let w be a neighbour of u. Then $\mathrm{dist}(\phi(w),\phi(v)) \in \{e-1, e\}$. Let C_1 be the set of neighbours w of u such that $\mathrm{dist}(\phi(w),\phi(v)) = e-1$ and let C_0 denote the set of neighbours of u such that $\mathrm{dist}(\phi(w),\phi(v) = e$. Then C_0 and C_1 partition the neighbours of u. We see also that if $w \in C_0$, then

$$(F)_{\psi(u),\psi(v)} = E_{u,v}$$

We have

$$\begin{aligned} 0 = ((A-\tau I)(E-F))_{u,v} &= \sum_{w\in V(X)} (A-\tau I)_{u,w}(E-F)_{w,v} \\ &= -\tau(E-F)_{u,v} + \sum_{w\sim u} (E-F)_{w,v} \\ &= \sum_{w\in C_0} (E-F)_{w,v}. \end{aligned}$$

Let $\gamma_0,\dots,\gamma_d$ be the cosine sequence for the τ-eigenspace of X (and Y). Then $\gamma_0 = 1$ and the terms of the sequence alternate in sign (see cg-blue [1]). [1]
It follows that

$$\sum_{w\in C_0} (E-F)_{w,v} = (\gamma_e - \gamma_{e-1})|C_1|$$

and therefore $C_1 = \emptyset$.

At this point the algebra is over and we are back to pure graph theory. We have proved that if $w \sim u$ and $\mathrm{dist}(\phi(u),\phi(v)) = e$, then $\mathrm{dist}(\phi(w),\phi(v)) = e$. Hence ϕ maps $N_1(u)\cap N_e(v)$ into $N_1(\phi(u)\cap N_e(\phi(v)))$. Repeating this argument with u replaced w and w by a suitable neighbour of W, we deduce that ϕ maps each component of $N_e(v)$ into a component of $N_e(\phi(v))$. Since $N_e(v)$ is connected (by hypothesis), $\phi(N_e(v)) \subseteq N_e(\phi(v))$, and by an induction argument using the fact that X_e is connected, we deduce that ϕ must map pairs of vertices at distance e to pairs of vertices at distance e.

If ϕ is not an automorphism, there must be a pair of vertices at distance two in X mapped to the same vertex by ϕ. Suppose these vertices are u and x, and let v be a vertex in X in $N_e(u)\cap N_{e-2}(x)$. Then $\mathrm{dist}(\phi(u),\phi(v)) \le e-2$, contradicting the conclusion of the previous paragraph. □

We can also say something about the case where X is antipodal of diameter d. Suppose ϕ is an endomorphism of X and u and v are vertices such that $\mathrm{dist}(\phi(u),\phi(v)) = d$. Then all neighbours of u are at distance $d-1$ from v, but the argument above shows that $C_1 = \emptyset$. Hence the core of X has diameter at most $d-1$.

Suppose X is 1-walk regular and its core is a complete graph. We have

$$\omega(X) \leq 1 - \frac{k}{\tau} \leq \chi(X),$$

and so $\chi_f(X) = \chi(X)$ and

$$\alpha(X) = \frac{|V(X)|}{1 - \frac{k}{\tau}}.$$

One consequence is that k/τ must be an integer. The fibres of the colouring form an equitable partition.

31
The Colin de Verdière Number

This chapter provides an introduction to some properties of the Colin de Verdière number of a graph. They are heavily dependent on a survey by Van der Holst, Lovász and Schrijver. There are two possible novelties. We make use of matrix perturbation theory, and we use quadratic rank to describe the so-called strong Arnold hypothesis.

If $u, v \in V(X)$, we write $u \simeq v$ to denote that u and v are equal or adjacent.

31.1 Perturbation

The following discussion summarizes Theorems II.5.4 and II.6.8 of Kato [].

Let A and H be real symmetric $n \times n$ matrices. We concern ourselves with the eigenvalues of $A + tH$, for small values of t; we will see that it is possible to view these as perturbations of the eigenvalues of A. Assume θ is an eigenvalue of A with multiplicity m, and let P be the projection on to the associated eigenspace. Then there is a matrix-valued function $P(t)$ such that

(a) $P(0) = P$.

(b) $P(t)$ is a real analytic function of t, and a projection.

(c) The column space of $P(t)$ is invariant under $A + tH$.

Note that $\mathrm{rk}(P) = m$. As $\mathrm{rk}(P(t)) = \mathrm{tr}(P(t))$, it follows that $\mathrm{rk}(P(t))$ is a continuous integer-valued function. Therefore

(d) $\mathrm{rk}\, P(t) = m$.

From (c) it follows that the column space of $P(t)$ is a sum of eigenspaces of $A + tH$. These eigenspaces can be viewed as arising by splitting the θ-eigenspace of A.

The eigenvalues associated with these eigenspaces are analytic functions $\theta_1(t), \ldots, \theta_k(t)$ such that $\theta_i(0) = \theta$. If U is a matrix whose columns form an orthonormal basis for the columns of $P = P(0)$, then $P = UU^T$ and

the derivatives $\theta_1'(0), \ldots, \theta_k'(0)$ are the eigenvalues of $U^T HU$. The dimension of the $\theta_i(t)$ eigenspace equals the dimension of the $\theta_i'(0)$-eigenspace of $U^T HU$.

31.1.1 Lemma. *Let Q be a symmetric matrix and suppose that the columns of the matrix U for an orthonormal basis for* $\ker(Q)$*. If K is symmetric then the corank of $Q + tK$ equals the corank of Q for all sufficiently small values of t if and only if $U^T KU = 0$.*

Proof. The matrix UU^T is the orthogonal projection onto $\ker(Q)$. As $U^T U = I$, we see that $U^T KU = 0$ if and only if $PKP = 0$. □

31.2 The Strong Arnold Hypothesis

If A and B are two matrices of the same order, we use $A \circ B$ to denote their Schur product, which is defined by the condition

$$(A \circ B)_{i,j} = A_{i,j} B_{i,j}.$$

If X is a graph on n vertices, we define a *generalized Laplacian* for X to be a symmetric matrix Q such that $Q_{u,v} < 0$ if u and v are adjacent vertices in X and $Q_{u,v} = 0$ if u and v are distinct and not adjacent. (There are no constraints on the diagonal entries of Q.) Examples are the usual Laplacian and $-A$, where A is the adjacency matrix of X. Note that we have not assumed that the least eigenvalue of Q is simple, although this will hold if X is connected, by Perron-Frobenius.

We associate two spaces of symmetric matrices to each generalized Laplacian Q. Let $\mathcal{N}_Q$ denote the space of symmetric $n \times n$ matrices N such that

$$N \circ I = N \circ Q = 0$$

and let $\mathcal{K}_Q$ denote the space of symmetric $n \times n$ matrices K such that

$$QK = 0.$$

We say that Q satisfies the *Strong Arnold Hypothesis* if $\mathcal{N}_Q \cap \mathcal{K}_Q = 0$. This is often abbreviated to SAH. If θ is an eigenvalue of Q, we say that its associated eigenspace satisfies the SAH if $Q - \theta I$ does.

To give a very small example, suppose that $\ker Q$ has dimension one, and is spanned by a vector x. Any symmetric matrix H such $QH = 0$ must be a multiple of xx^T, but $(xx^T)_{i,i} = (x_i)^2$ and if $I \circ xx^T = 0$ then $x = 0$. Hence, if $\dim \ker Q = 1$, then Q satisfies the SAH.

We now describe a second version of the SAH. The space of symmetric $n \times n$ matrices is an inner product space, relative to the bilinear form

$$\langle A, B \rangle = \operatorname{tr}(AB).$$

We have the following:

31.2.1 Lemma. *The SAH holds for Q if and only if $\mathcal{N}_Q^\perp + \mathcal{K}_Q^\perp$ is the space of all symmetric $n \times n$ matrices.* □

Clearly $\mathcal{N}_Q$ consists of the symmetric matrices N such that $N_{u,v} = 0$ whenever $u \simeq v$. Hence $H \in \mathcal{N}_Q^\perp$ if and only if $Q + tH$ is a generalized Laplacian for all sufficiently small values of t.

To characterize $\mathcal{K}_Q^\perp$, we need a preliminary result.

31.2.2 Lemma. *Let Q be a symmetric matrix with corank m, and let U be a matrix whose columns form an orthonormal basis for* $\ker Q$*. Then a symmetric matrix K satisfies $QK = 0$ if and only if there is a symmetric matrix B such that $K = UBU^T$.*

Proof. The stated condition is sufficient, we prove that it is also necessary. If $QK = 0$ then the column space of K lies in $\ker Q$. As K is symmetric, there is a matrix U_1 whose columns lie in the column space of K and a symmetric matrix B_1 such that $K = U_1 B_1 U_1^T$. There is a matrix, R say, such that $U_1 = UR$ and therefore $K = U(RB_1R^T)U^T$, as required. □

It follows that $\mathcal{K}_Q^\perp$ consists of the matrices H such that $\langle UBU^T, H\rangle = 0$ for all symmetric matrices B. As

$$\langle UBU^T, H\rangle = \operatorname{tr}(UBU^T H) = \operatorname{tr}(U^T HUB) = \langle U^T HU, B\rangle,$$

we see that $H \in \mathcal{K}_Q^\perp$ if and only if $U^T HU$ is orthogonal to all symmetric matrices B. But this implies that $U^T HU = 0$ and therefore $\mathcal{K}_Q^\perp$ consists of the symmetric matrices H such that $U^T HU = 0$. Using Lemma 31.1.1, we conclude that $H \in \mathcal{K}_Q^\perp$ if and only if $Q + tH$ has the same corank as Q, for all sufficiently small values of t.

31.3 Quadratic Rank

We relate quadratic rank to the Strong Arnold hypothesis.

31.3.1 Theorem. *Let Q be symmetric and let θ be an eigenvalue of Q with multiplicity m. The θ-eigenspace of Q satisfies the Strong Arnold hypothesis if and only if the corresponding representation has quadratic edge-rank $\binom{m+1}{2}$.* □

If (30.1.1) holds, then (30.1.2) holds for the edge uv if and only if

$$\rho(v)B\rho(u)^T + \rho(u)B\rho(v)^T = 0;$$

equivalently if and only if

$$\langle B, \rho(u)^T\rho(v) + \rho(v)^T\rho(u)\rangle = 0.$$

If x, y are row vectors, we define

$$x \vee y := x^T y + y^T x.$$

We see that $x \vee y$ is a symmetric $m \times m$ matrix with rank at most two; the matrices $x \vee y$ span the space of symmetric $m \times m$ matrices.

If (30.1.1) holds, then (30.1.2) holds for the edge uv if and only if

$$\rho(v)B\rho(u)^T + \rho(u)B\rho(v)^T = 0;$$

equivalently if and only if

$$\langle B, \rho(u)^T\rho(v) + \rho(v)^T\rho(u)\rangle = 0.$$

31.4 A Minor-Monotone Parameter

Let X be a graph and let $\mathscr{Q}$ denote the set of all generalized Laplacians Q such that:

(a) $\lambda_1(Q)$ is simple.

(b) The λ_2-eigenspace of Q satisfies the SAH.

The *Colin de Verdière* number of X is the maximum multiplicity of λ_2, over all matrices in $\mathscr{Q}$. We denote it by $\mu(X)$.

31.4.1 Theorem. *If $e \in E(X)$ then $\mu(X \setminus e) \le \mu(X)$.*

Proof. Let Y denote $X \setminus e$ and let Q be a generalized Laplacian for Y such that λ_2 has multiplicity equal to $\mu(Y)$. Let Ξ be the adjacency matrix of e, viewed as a subgraph of X with $|V(X)|$ vertices. By Lemma 31.2.1, we can write Ξ as a sum $N + K$ where $N \in \mathscr{N}_Q^\perp$ and $K \in \mathscr{K}_Q^\perp$.

Therefore $Q + tK$ is a generalized Laplacian for X when $t \neq 0$. As $K \in \mathscr{K}_Q^\perp$, it follows from **??** that $\lambda_2(Q + tK) = \lambda_2(Q)$ has multiplicity equal to $\mu(Y)$ whenever t is small enough. By our remarks at the end of **??**, the representations of X associated with $\lambda_2(Q + tK)$ and $\lambda_2(Q)$ have the same quadratic rank, and so the SAH holds for $Q + tK$.

Since the multiplicity of $\lambda_2(Q + tK)$ is constant for small t, we also see that $\lambda_1(Q + tK)$ is simple. □

31.4.2 Theorem. *If $e \in E(X)$ then $\mu(X/e) \le \mu(X)$.*

Proof. Suppose $e = 12$ and

$$Q(X/e) = \begin{pmatrix} a & b^T \\ b & Q_1 \end{pmatrix}.$$

We assume that $\lambda_2(Q(X/e)) = 0$. Let Y be the graph $K_1 \cup (X/e)$ and let $Q = Q(Y)$ be a generalized Laplacian for Y. We may assume $Q(Y)$ has the form

$$\begin{pmatrix} \epsilon & 0 & 0 \\ 0 & a & b^T \\ 0 & b & Q_1 \end{pmatrix},$$

where $\epsilon > 0$, and will be restricted further shortly. Let Ξ be the matrix

$$\Xi = \begin{pmatrix} 0 & -1 & c^T \\ -1 & 0 & -c^T \\ -c & 0 & 0 \end{pmatrix}.$$

Here c is a non-positive vector such that $Q + \Xi$ is a generalized Laplacian for X. So, for example:

$$c_i = \begin{cases} 0, & \text{if } 1 \not\sim i \text{ and } 2 \not\sim i; \\ b_i, & \text{if } 1 \sim i \text{ and } 2 \not\sim i; \\ 0, & \text{if } 1 \not\sim i \text{ and } 2 \sim i; \\ b_i/2, & \text{otherwise.} \end{cases}$$

As before, $\Xi = N + K$, where $N \in \mathscr{N}_Q^\perp$ and $K \in \mathscr{K}_Q^\perp$. Consider the matrix pencil $Q + tK$. For small values of t we know that the rank of $Q + tK$ does not change, and the SAH holds for $\ker(Q + tK)$. Choose some positive value of t that works, and assume that ϵ was chosen so that $\epsilon < t^2/(1-t)$. If we multiply the first row and colum of $Q + tK$ by $(1-t)/t$, we get the matrix

$$Q' = \begin{pmatrix} \epsilon(1-t)^2/t^2 & t-1 & (1-t)c^T \\ t-1 & a & b^T - tc^T \\ (1-t)c & b - tc & Q_1 \end{pmatrix}$$

This operation does not change the rank, and it is not hard to see that SAH holds for $\ker(Q')$ and that $\lambda_1(Q')$ is simple.

Let Q'' be the matrix we get from Q' by subtracting its first row from its second, and the first column from the second. We observe that Q'' is a generalized Laplacian for X (at last), and that its rank is equal to the rank of Q'. We have to show that the SAH holds.

Let U be the $n \times m$ matrix whose columns form a basis for $\ker(Q')$, and let M be the elementary matrix we get by adding the first row of I to its second row. Thus $Q'' = MQ'M^T$ and the columns of $M^{-1}U$ are a basis for $\ker(Q'')$. Since the SAH holds for Q', the space of all $m \times m$ symmetric matrices is spanned by the matrices

$$u_i u_i^T, \quad (u_i - u_j)(u_i - u_j)^T, \quad i \in V(X),\ ij \in E(X)$$

where u_i is the i-th row of U. Let v_i denote the i-th row of $M^{-1}U$.

We have $v_1 = u_1$, $v_2 = u_1 - u_2$ and $v_1 - v_2 = u_2$. Hence the SAH holds. As Q'' and Q' are congruent, it follows from Sylvester's law of inertia that $\lambda_1(Q'')$ is simple. □

The previous two results combine to give the most important property of the Colin de Verdière number:

31.4.3 Corollary. *If Y is a minor of X then $\mu(Y) \le \mu(X)$.*

31.5 Properties

We derive some further relations between the Colin de Verdière number of a graph and its subgraphs. We begin with a technical result.

31.5.1 Lemma. *If an eigenspace of X contains two eigenvectors with disjoint supports, then the SAH hypothesis fails.*

Proof. Suppose x and y are vectors and U is the matrix

$$U = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}.$$

As the Schur product of the columns of U is zero, it follows that $\operatorname{qrk}(U) = 2$. By 30.1.3, we deduce that the SAH fails for any subspace that contains x and y. □

We now determine $\mu(K_n)$ and $\mu(\overline{K_n})$. It is easy to see that $\mu(K_1) = 0$. Suppose $n \geq 2$. Then $-J$ is a generalized Laplacian for K_n with $\lambda_2(Q) = 0$ having multiplicity $n-1$. Here $\mathcal{N}_Q^\perp$ is the space of all symmetric matrices, and so the SAH holds by Lemma 31.2.1. (Conversely, it is not too hard to show that $\mu(X) = |V(X)| - 1$ if and only if $X = K_n$.)

Next we consider $X = \overline{K_n}$. Here $\mathcal{N}_Q^\perp$ is the space of diagonal matrices. Suppose Q is a generalized Laplacian for X such that λ_2 is simple. We may assume without loss that the associated eigenvector is e_1, the first standard basis vector. Then

$$\mathcal{K}_Q^\perp = \{H : e_1^T H e_1 = 0\} = \{H : H_{1,1} = 0\}.$$

Thus $\mathcal{K}_Q^\perp$ has codimension 1 in the space of symmetric matrices, and so Lemma 31.2.1 again yields that the SAH holds. Hence $\mu(X) \geq 1$.

Suppose now that λ_2 has multiplicity at least two. We may assume that the eigenspace contains e_1 and e_2, whence 31.5.1 yields that the SAH fails.

31.5.2 Lemma. *If X has at least one edge, then $\mu(X)$ equals the maximum value of $\mu(Y)$, where Y ranges over the components of X.*

Proof. By 31.4.1, $\mu(X) \geq \mu(C)$, where C runs over the components of X.

Let Q be a generalized Laplacian for X that realises $\mu(X)$, with $\lambda_2(Q) = 0$. Exactly one component Y of X has least eigenvalue equal to λ_1; all other components must have least eigenvalue no less than zero. If two distinct components of X have eigenvalue zero then the kernel of Q contains two vectors with disjoint support, and the SAH fails.

If some component of X has least eigenvalue zero, then since a component is connected, its least eigenvalue is simple and $\mu(X) = 1$. As X contains an edge, some component C has $\mu(C) \geq 1$, and the lemma follows.

Otherwise 0 must be an eigenvalue of Y and $\mu(X) = \mu(Y)$. □

Suppose Q is a generalized Laplacian for X and ρ is the representation on some eigenspace of Q with eigenvalue θ and dimension m. Let u be a vertex of X and let W be the subspace of the eigenspace spanned by the eigenvectors that vanish on u. The restriction to $V(X)\setminus u$ of any of these eigenvectors is an eigenvector for $Q(X\setminus u)$, with eigenvalue θ. This shows that the multiplicity of θ as an eigenvalue of $Q(X\setminus u)$ is at least $m-1$. It follows from spectral decomposition that the dimension is $m-1$ if and only if $\rho(u)\neq 0$. (See Lemma 8.13.1 in Godsil & Royle.)

31.5.3 Lemma. *Let Q be a generalized Laplacian for X, and let ρ be the representation on the λ_2-eigenspace of Q. Assume $\lambda_2(Q)$ has multiplicity m and let u be a vertex in X such that*

(a) $\rho(u)\neq 0$.

(b) The images of u and its neighbours span $\mathbb{R}^m$.

Then if the SAH holds for $Q(X\setminus u)$, it holds for Q.

Proof. By hypothesis, the vectors in $\ker(Q-\lambda_2 I)$ that do not have u in their support span a space W of dimension $m-1$, with quadratic rank $\binom{m}{2}$. Let $x_1,\ldots,x_{m-1}$ be a basis for W and let z be an eigenvector for Q with eigenvalue λ_2 such that $z_u=1$.

Suppose the SAH fails for Q. Then there is a non-zero symmetric $m\times m$ matrix B such that

$$\rho(v)^T B\rho(w)=0$$

if $v=w$ or $v\sim w$. As $\rho(u)$ is the first standard basis vector in $\mathbb{R}^m$ and as the images of u and its neighbours span $\mathbb{R}^m$, it follows that $\rho(u)^T B=0$. Thus the first row and column of B are zero. If B is not zero the the SAH fails for $Q(X\setminus u)$. □

31.5.4 Lemma. *If $u\in V(X)$ then $\mu(X\setminus u)\geq \mu(X)-1$. If u is adjacent to each vertex in $V(X)\setminus u$ and $|V(X)|\geq 2$, then $\mu(X\setminus u)=\mu(X)-1$.*

Proof. Suppose that Q is a generalized Laplacian for X that realizes the Colin de Verdière number of X. Let u be a vertex of X, let Q_u be the matrix we get by deleting the u-row and u-column from Q and let W be the space consisting of the λ_2-eigenvectors x of Y such that $x_u=0$. It is easy to verify that deleting the u-coordinate from a vector in W gives an eigenvector for Q_u with eigenvalue $\lambda_2(Q)$. As the SAH holds for Q, it holds for W.

To complete the proof of the first claim, we show that if Q is a generalized Laplacian for X and W is a k-dimensional subspace of the λ_2-eigenspace which satisfies the SAH, then $\mu(X)\geq k$.

Suppose $\lambda_2(Q)$ has multiplicity m and let F be the projection onto the complement of W in the λ_2-eigenspace. Let U_1 be an $n\times k$ matrix whose columns are a basis for W. As the SAH holds for W, we have $F=N+K$, where $N\in\mathcal{N}_Q^{\perp}$ and $U_1^T KU=0$. Let U be a $n\times n$ matrix with a basis of the

λ_2-eigenspace as columns. Then $U^T NU = U^T(F-K)U$ has eigenvalue 0 with multiplicity k and 1 with multiplicity $m-k$. So, for small non-zero values of t, we find that $\lambda_2(Q+tN)$ has multiplicity k and satisfies the SAH.

Finally, suppose u is adjacent to all vertices in $X \setminus u$. By the previous result, we may assume that $X \setminus u$ is connected. Let Q' be a matrix realizing $\mu(X \setminus u)$ with $\lambda_2(Q') = 0$, and let z be an eigenvector of Q' with eigenvalue $\lambda_1 = \lambda_1(Q)$. We may choose z so that $z < 0$ and $\|z\| = 1$. Let Q be the generalized Laplacian given by

$$Q = \begin{pmatrix} \lambda_1^{-1} & z^T \\ z & Q' \end{pmatrix}.$$

Then $\ker Q$ contains $(\lambda_1, z)^T$ and all vectors of the form $(0, x)^T$, where $x \in \ker(Q')$. The SAH holds by the previous lemma.

The least eigenvalue of Q is simple, because X is connected. By interlacing, $\lambda_2(Q) = \lambda_2(Q')$. We conclude that $\mu(X) = \mu(X \setminus u) + 1$. □

31.5.5 Corollary. *Suppose C is a vertex cutset in X and let $Y_1, \ldots, Y_r$ be the components of $X \setminus C$. Then $\mu(X) \leq |C| + \max_i \mu(Y_i)$.*

32
Hom-idempotent Graphs

A graph X is *hom-idempotent* if there is homomorphism from $X \Box X$ to X. We present work of Larose, Laviolette and Tardif, showing that X is hom-idempotent if and only if it is homomorphically equivalent to a normal Cayley graph.

32.1 Examples

We first prove that K_n is hom-idempotent. View $K_n \Box K_n$ as a grid. A coclique in K_n contains at most one vertex in each row, and at most one in each column. Hence $\alpha(K_n \Box K_n) \leq n$ and we see that n-colourings of this graph correspond to $n \times n$ Latin squares. Of which there are many; it is traditional to choose the addition table for the $\mathbb{Z}_n$.

we use this to prove that

$$\chi(X \Box Y) = \max\{\chi(X), \chi(Y)\}.$$

Since X and Y are subgraphs of the product, the right side is a lower bound. On the other hand, if both X and Y admit m-colourings, then

$$X \Box Y \to K_m \Box K_m,$$

implying that $\chi(X \Box Y) \leq m$.

The Petersen graph is not hom-idempotent. To prove this we need two facts about fractional chromatic number. First, if $X \to Y$, then $\chi_f(X) \leq \chi_f(Y)$. Second, if X is vertex transitive, then $\chi_f(X) = |V(X)|/\alpha(X)$.

If P denotes the Petersen graph, $\chi_f(P) = 5/2$. We compute $\chi_f(P \Box P)$; since this product is vertex transitive this means we compute its coclique number.

Let $\alpha_r(X)$ denote the maximum size of an induced r-colourable subgraph of X. Recall that

$$\alpha_r(X) = \alpha(X \Box K_r)$$

and therefore $\alpha(P \Box K_2) = 7$. The subgraph of $P \Box P$ induced by two vertical copies of P is isomorphic to either $2P$ or $P \Box K_2$, in the latter case we say

that the copies are adjacent. If follows that if the coclique S contain four vertices in a vertical copy of P (the maximum possible), then it contains at most three vertices in each of the three copies of P adjacent to it. Hence $\alpha(P \square P) \leq 37$ and so

$$\chi_f(P \square P) \geq \frac{100}{37} > \frac{5}{2}.$$

32.2 Normal Cayley graphs are Hom-idempotent

A Cayley graph is *normal* if its connection set is a union of conjugacy classes. Any Cayley graph for an abelian group is normal.

32.2.1 Lemma. *If $X(G,\mathscr{C})$ is a normal Cayley graph, the map from $X \square X$ to X that sends (g,h) to gh is a homomorphism to X.*

Proof. Assume (g,h) and (g',h') are adjacent in $X \square X$. If $g = g'$, then $gh \neq g'h'$. Similarly if $h = h'$ the images of these two vertices are distinct.

Next

$$(g'h')(gh)^{-1} = g'h'h^{-1}g^{-1}$$

and so if $g = g'$, then $(g'h')(gh)^{-1}$ is conjugate to $h'h^{-1}$ and so lies in $\mathscr{C}$. Therefore if $h \sim h'$, then $g'h'$ and gh are adjacent. □

32.3 Hom-idempotent Graphs are Normal Cayley Graphs?

We show that hom-idempotent graphs are homomorphically equivalent to normal Cayley graphs.[1]

[1] the section title is long enough without "homomorphically equivalent", hence the question mark

A *shift* on a graph is a graph automorphism that maps each vertex to a neighbour. If σ is a shift, so is σ^{-1}. If σ is a shift and α is an automorphism, then $\alpha^{-1}\sigma\alpha$ is again a shift.

Suppose $X = X(G,\mathscr{C})$ is a normal Cayley graph and that x and y are adjacent vertices in X. Assume $g = yx^{-1}$, so $g \in \mathscr{C}$. If $z \in V(X)$, then since $\mathscr{C}$ is closed under conjugation, $zgz^{-1} \in \mathscr{C}$. Therefore $z \sim zg$ and it follows that right multiplication by g is an automorphism of X that takes each vertex to a neighour, i.e., it is a shunt.

32.3.1 Theorem. *A graph X is hom-idempotent if and only if is homomorphically equivalent to the Cayley graph for* Aut X *with the set of shifts of $X^\bullet$ as its connection set.*

Proof. We observe that X is hom-idempotent if and only if $X^\bullet$ is, so we assume that X is a core.[2]

[2] it will save a lot of bullets

Lemma 32.2.1 tells us that normal Cayley graphs are hom-idempotent.

Suppose X is a core and that there is a homomorphism ψ from $X \square K_2$ to X. Assume $V(K_2) = \{0,1\}$. The $(V(X),0)$ and $(V(X),1)$ induce copies of X and, since X is a core the restriction of ψ to one of these copies of X is an automorphism. Let ψ_i (for $i = 0,1$) denote the restriction of ψ to $(V(X),i)$.

By composing ψ with an automorphism of its image, we may assume that ψ_0 is the identity. So if $u \in V(X)$, then $\psi_0((u,0)) = u$ and $\psi_1(u,1)$ must be a neighbour of u (in X). Hence ψ_1 is a shunt on X.

Now suppose we have a homomorphism $\psi : X \Box Y \to X$. For each vertex y in Y, this gives us a homomorphism ψ_y from X to itself, and so $y \mapsto \psi_y$ maps Y to automorphisms of X. If $y \sim z$ then $\psi_y^{-1}\psi_z$ is a shunt, and so we have shown that if X is a core, a homomorphism from $X \Box Y$ determines a homomorphism from Y into the Cayley graph for $\mathrm{Aut}\, X$ with the shunts of X as its connection set. □

32.4 A Feasibility Condition for Automorphisms

We derive a result (due to G. Higman) which implies that the Petersen graph does not admit any shifts.

Suppose $A = A(X)$ and assume it has spectral decomposition

$$A = \sum_r \theta_r E_r.$$

Recall that the automorphism group of X can be identified with group of permutation matrices P such that $PA = AP$. If $P \in \mathrm{Aut}\, X$, then P must commute with each projection E_r and this implies in turn that P leaves invariant the eigenspace associated to θ_r.

If a matrix P leaves a subspace U of $\mathbb{R}^n$ invariant, then it determines a linear mapping from U to itself. We call the trace of this linear mapping the *trace of P restricted to U*, and denote it by $\mathrm{tr}_U(P)$. We note that tr_U is a sum of eigenvalues of P.

32.4.1 Lemma. *Let X be a graph, let P be a permutation matrix in* $\mathrm{Aut}\, X$ *and let U be an eigenspace of X. If E is the orthogonal projection on U then* $\mathrm{tr}(PE)$ *is an algebraic integer.*

Proof. As $\mathrm{Aut}\, X$ is finite, there is a least positive integer m such that $P^m = I$. Therefore the eigenvalues of P are zeros of the monic polynomial $x^m - 1$ and $\mathrm{tr}_U(P)$ is an algebraic integer.

Let $u_1, \ldots, u_n$ be an orthonormal basis for $\mathbb{R}^n$ such that $u_1, \ldots, u_m$ is an orthonormal basis for U. Then

$$\mathrm{tr}_U(P) := \sum_{i=1}^m \langle u_i, Pu_i\rangle = \sum_{i=1}^m \langle Eu_i, PEu_i\rangle = \sum_{i=1}^m \langle u_i, EPEu_i\rangle$$

Next, $Eu_j = 0$ if $j > m$ and so the last sum is equal to

$$\sum_{i=1}^n \langle u_i, EPEu_i\rangle.$$

This equals the trace of EPE. Since $E^2 = E$ we have

$$\mathrm{tr}(EPE) = \mathrm{tr}(PEE) = \mathrm{tr}(PE),$$

whence the lemma follows. □

Note that $\mathrm{tr}_U(P) = \mathrm{tr}(PE)$. In general it is not at all convenient to compute $\mathrm{tr}_U(P)$. For distance-regular graphs it is easier though.

32.4.2 Lemma. *Let X be a distance-regular graph with valency k on v vertices. If P is an automorphism that maps each vertex to a vertex at distance $r \neq 0$, then*

$$\frac{m_j p_r(j)}{v_r}$$

is an algebraic integer.

Proof. First

$$\mathrm{tr}(PA_i) = \mathrm{sum}(P \circ A_i)$$

and thus $\mathrm{tr}(PA_i) = 0$ if $i \neq r$ and $\mathrm{tr}(PA_r) = v$. We have

$$E_j = \frac{1}{v} \sum q_j(i) A_i.$$

and therefore

$$\mathrm{tr}(E_j P) = \frac{q_j(r)}{v} \mathrm{tr}(PA_r) = \frac{q_j(r)}{.}$$

Now $A_r E_j = p_r(j) E_j$ and so

$$p_r(j) m_j = \mathrm{tr}(A_r E_j) = \mathrm{sum}(A_r \circ E_j) = \frac{1}{v} q_j(r) v v_r,$$

hence

$$q_j(r) = \frac{p_r(j) m_j}{v_r}.$$

We apply this to the Petersen graph with $r = 1$ and $j = 1$. We have $p_1(1) = 1$, $m_1 = 5$ and $k = 3$. Since $5/3$ is not an algebraic integer, there are no shifts. (In particular, the Petersen graph is not a circulant.)[3]

[3] there is a much simpler proof of this: a Cayley graph for an abelian group with valency at least three has girth at most four

33
Counting Trees

33.1 Burnside and Pólya

We want to count isomorphism classes of trees on n vertices, and it happens that counting (isomorphism classes of) rooted trees is the way to go.

Let $R(x)$ be the generating series for rooted trees. Then

$$R(x) = x + x^2 + 2x^3 + \cdots .$$

Our aim is to derive an equation satisfied by $R(x)$:

$$R(x) = x \exp\left(\sum_{k \geq 1} \frac{1}{k} R(x^k) \right).$$

From this we can derive both a recurrence for the coefficients of $R(x)$, and precise asymptotic information about the growth of these coefficients with n.

We deem a *rooted forest* to be a forest where each component is rooted. Let $\mathscr{R}$ denote the set of all rooted trees. The symmetric group $\mathrm{Sym}(m)$ acts on $\mathscr{R}^m$, and the orbits of $\mathrm{Sym}(m)$ on $\mathscr{R}^m$ correspond to the rooted forests with exactly m components. We can compute this using Burnside's lemma: for each permutation π in $\mathrm{Sym}\, m$ we compute the generating series S_π for the elements of $\mathscr{R}^m$ that are fixed by π. The generating series F_m for rooted forests with m components is then

$$\frac{1}{m!} \sum_{\pi} S_\pi .$$

By way of example, suppose $m = 2$. Then

$$S_{(1)}(x) = R(x)^2, \qquad S_{(12)} = R(x^2)$$

and so

$$F_2(x) = \frac{1}{2}(R(x)^2 + R(x^2)).$$

We note that $xF_m(x)$ is the generating series for rooted trees where the root vertex has valency m, and consequently

$$R(x) = x \sum_{m\geq 0} F_m(x).$$

We work out the exponential generating series for cycles, where a cycle of length ℓ is mapped to the variable a_ℓ. Since there are $(n-1)!$ cycles of length n, the generating series is

$$\sum_{m\geq 1} \frac{a_m}{m}.$$

We view a cycle as a connected permutation, and then we see that the generating series for all permutations is

$$\exp\left(\sum_{m\geq 1} \frac{a_m}{m}\right).$$

If $\mu_\ell(\pi)$ denotes the number of cycles of length ℓ in the permutation π, then in this series a permutation π is mapped to the monomial

$$\prod_{\ell\geq 1} a_\ell^{\mu_\ell(\pi)}.$$

Thus

$$\exp\left(\sum_{m\geq 1} \frac{a_m}{m}\right) = \sum_{m\geq 1} \sum_{\pi\in \mathrm{Sym}\, m} \frac{\prod_{\ell\geq 1} a_\ell^{\mu_\ell(\pi)}}{m!}.$$

If we substitute $R(x^\ell)$ for a_ℓ in this, the left side becomes

$$\sum_{m\geq 0} F_m(x)$$

and so we have our equation for $R(x)$.

33.2 Counting Trees

We derive an expression for the generating series for trees in terms of the generating series for rooted trees.

An edge uv in a graph is *symmetric* if there is an automorphism that swaps its ends. In a tree there is at most one symmetric edge (which would be the center of the tree). An edge that is not symmetric is *aymmetric*. The following result is known as *Otter's lemma*.

33.2.1 Lemma. *Let T be a tree. Then the number of orbits of vertices of T minus the number of orbits of asymmetric edges is 1.*

Proof. We get to use Burnside again. Let G be the automorphism group of T and suppose $\gamma \in G$. If γ fixes two vertices u and v, then it must fix each vertex on the unique path in T that joins u to v. Therefore the fixed points of γ induce a subtree T_γ.

Now assume T does not have a symmetric edge.. Then any automorphism fixes each vertex in the center of T and so T_γ is not empty, for any automorphism γ. Then

$$|V(T_\gamma)| - |E(T_\gamma)| = 1$$

and the lemma follows.

Now assume T has symmetric edge $e = uv$. Let α be an automorphism that swaps u and v. Then we have the coset decomposition

$$G = G_u \cup G_u\alpha$$

If $\gamma \in G_u$, then it fixes one more vertex than its does edges, but $\gamma\alpha$ fixes no vertex and one edge. Hence the lemma holds in this case too. □

If T is a tree, let $v^*(T)$ and $e^*(T)$ denote the number of vertex orbits and the number of asymmetric edge orbits respectively and define

$$\beta(T) := v^*(T) - e^*(T).$$

33.2.2 Theorem. *If $R(x)$ is the generating series for rooted trees, then the generating series for trees is*

$$T(x) = R(x) - \frac{1}{2}R(x)^2 + \frac{1}{2}R(x^2).$$

Proof. Let R_n be the number of rooted trees on n vertices and let ER_n be the number of trees rooted at an asymmetric edge. Then $R_n - ER_n$ is the number of trees on n vertices (by Otter's lemma). The generating series for trees rooted at a symmetric edge is $R(x^2)$, and therefore the generating series $ER(x)$ for trees rooted at an asymmetric edge is given by

$$ER(x) = \frac{1}{2}(R(x)^2 - R(x^2)).$$

Hence the generating series for all trees is

$$T(x) = R(x) - \frac{1}{2}(R(x)^2 - R(x^2)).$$ □

33.3 A Recurrence for Rooted Trees

We note two identities that hold for any power series with constant term zero. The calculations are reasonably straightforward, so we omit them.

We use $[x^n, F(x)]$ to denote the coefficient of x^n in the series $F(x)$. The proof of the following is straightforward and omitted.

33.3.1 Lemma. *If $R(x) = \sum_{n\geq 1} R_n x^n$, then*

$$\sum_{n\geq 1} \frac{R(x^n)}{n} = \sum_{n\geq 1}\left(\sum_{d|n} \frac{dR_d}{n}\right)x^n = \log\left(\prod_{n\geq 1}(1-x^n)^{-R_n}\right).$$ □

33.3.2 Lemma. *If*

$$B(x) = \exp\left(\sum_{n\geq 1} \frac{a_n}{n} x^n\right)$$

and $B_n := [x^n, B(x)]$, *then* $B_0 = 1$ *and, if* $n \geq 1$,

$$B_n = \frac{1}{n} \sum_{i=1}^{n} B_{n-i} a_i.$$

Proof. We have

$$B'(x) = B(x) \sum_{n\geq 1} a_n x^n$$

and consequently

$$(n+1) B_{n+1} = \sum_{i=0}^{n} B_{n-i} a_{i+1}.$$

The lemma follows at once. □

33.3.3 Theorem. *If* R_n *is the number of rooted trees on* n *vertices, then*

$$R_{n+1} = \frac{1}{n} \sum_{i=1}^{n} \left(\sum_{d|i} d R_d\right) R_{n+1-i}.$$ □

34
The Spectral Centre of a Tree

Brouwer and Haemers [1] have introduced the concept of the *spectral centre* of a tree. We give a presentation of this.

[1]

34.1 The Centre of a Tree

The *eccentricity* of a subset S of vertices in a graph is the maximum distance of a vertex from S.[2]. The *radius* of X is the minimum value of the eccentricity of a vertex. A vertex u in X is central if its eccentricity is equal to the radius of X.

[2] I sometimes call it the *covering radius*

Here we are only concerned with the eccentricity of trees. We note that if $|V(T)| > 2$, then a vertex of valency one cannot be central; it follows that if $|V(T)| > 2$ and S is the tree we get from T by deleting all vertices of degree one then the centre of S is equal to the centre of T. Using this you may show that the centre of a tree is either a single vertex of an edge.

For a second exercise, you might prove that if P is a path of maximal length in the tree T, the centre of P is the centre of T. (There is always a vertex that lies on all paths of maximal length, as you might also prove.)

A tree also has a *centroid*. This is either a vertex u such that each component of $T \setminus u$ has fewer than $|V(T)|/2$ vertices, or an edge $\{u, v\}$ such that deleting the edge uv leave two components each with exactly $|V(T)|/2$ vertices. The centroid is unique.[3]

[3] Your move—prove this.

34.2 Supports of Eigenvectors

We start with a result valid for all graphs. If z is an eigenvector for X and $u \in V(X)$, we use $z(u)$ to denote the entry of z indexed by u.

34.2.1 Theorem. *Let u be a vertex in the graph X. Then θ is a common eigenvalue of X and $X \setminus 1$ if and only if there is an eigenvector z for X such that $z(1) = 0$.*

Proof. We take it as obvious that the existence of the eigenvector as described implies the existence of the common eigenvalue.

So assume θ is an eigenvalue of X and $X \setminus 1$. Set

$$A = \begin{pmatrix} 0 & b^T \\ b & A_1 \end{pmatrix}.$$

Then

$$\begin{pmatrix} 0 & b^T \\ b & A_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b^T y \\ xb + A_1 y \end{pmatrix}$$

and therefore if

$$\begin{pmatrix} 0 & b^T \\ b & A_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \theta \begin{pmatrix} x \\ y \end{pmatrix},$$

then

$$\theta x = b^T y, \quad (\theta I - A_1) y = xb.$$

Suppose w is an eigenvector for $X \setminus 1$ with eigenvalue θ. Then

$$w^T (\theta I - A_1) = 0$$

and consequently $x w^T b = 0$. Hence there are two possibilities: either $x = 0$, or $b^T w = 0$. In the first case

$$\begin{pmatrix} 0 \\ y \end{pmatrix}$$

is a θ-eigenvector for X that vanishes on 1, in the second case

$$\begin{pmatrix} 0 \\ z \end{pmatrix}$$

is the eigenvector we want. □

34.3 Sign-Changes

If z is an eigenvector for X and $uv \in E(X)$, we say there is a *sign-change* at uv if $z(u)z(v) < 0$.

Thus if X is connected and z is an eigenvector for the spectral radius of X, there are no sign-changes. However suppose ρ is the spectral radious of X and λ is an eigenvalue not equal to ρ. If z is an eigenvector for ρ and y is an eigenvector for λ, then $z^T y = 0$; since the entries of z all have the same sign, y must have both positive and negative entries and (if X is connected), w must change sign on some edge.

If $z \in \mathbb{R}^{V(X)}$, its *positive support* $\mathrm{supp}^+(z)$ is the set

$$\{u \in V(X) : z(u) > 0\}.$$

Similarly there is the *negative support* $\mathrm{supp}^-(x)$. We have

$$\mathrm{supp}^+(-z) = \mathrm{supp}^-(z), \qquad \mathrm{supp}^-(x) = \mathrm{supp}^+(x)$$

and so the actual signs do not matter much. We define a *sign component* of X to be a component of one of the subgraphs induced by $\mathrm{supp}^+(x)$ of

$\text{supp}^-(x)$. We observe that any edge that joins two sign-components must be a sign-change.

If $\lambda \geq 0$, a vector x is λ-subharmonic if $x \geq 0$ and $Ax \geq \lambda x$. To get an example, suppose x is an eigenvector with with eigenvalue θ and let $|y|$ denote the vector we get from y we get by replacing the entries of y by their absolute values. Then $Ax = \theta x$ and

$$|\theta||x| = |\theta x| \leq |Ax| \leq A|x|,$$

showing that $|x|$ is $|\theta|$-subharmonic.

34.3.1 Lemma. *If $\lambda > 0$ and z is a λ-eigenvector of X, then the restriction of z to a sign-component of z is λ-subharmonic, and the spectral radius of the subgraph induced by a sign-component is greater than λ.*

Proof. Let S be the subgraph of X induced by a sign-component of z. For each vertex i of X, we have

$$\lambda z(i) = \sum_{j \sim i} z(j).$$

Let y denote the restriction of z to $V(S)$ and let B be the submatrix of A with rows and columns indexed by $V(S)$. Assume $y \geq 0$ and $i \in V(S)$. If all neighbours of i lie in $V(S)$, then $(By)_i = \lambda y_i$; if some neighbour j of i is not in $V(S)$, then $z(j) < 0$ and then $(By)_i > \lambda y_i$. Therefore y is λ-subharmonic for S.

If we normalize z so that $\|y\| = 1$, then

$$y^T By > \lambda y^T y = \lambda$$

and therefore the spectral radius of S is greater than λ. □

34.4 Sign-changes on Trees

Our first result is simple.

34.4.1 Lemma. *If z is an eigenvector of a tree T and no entry is zero, the number of sign-changes is one less than the number of sign-components.*

Proof. The disjoint union of the sign-components is a spannning forest F of T. Since any two sign-components are joined by at most one edge, $|E(T)| - |E(F)|$ is equal to the number of components of F. As any two sign-components are joined by at most one edge and, as that edge is a sign change,the lemma follows. □

To prove the main result of this section, we need a formula for the entries of the eigenvectors of a tree.

34.4.2 Theorem. *Let T be a tree with vertex set $\{1, \ldots, n\}$, and let $P(i)$ denote the unique path in T from 1 to i. Let $z_T(\lambda)$ denote the vector with i-th entry equal to $\phi(T \setminus P(i), \lambda)$ for $i = 1, \ldots, n$. If λ is an eigenvalue of T, then $z_T(\lambda)$ is a eigenvector for T with eigenvalue λ.* □

For example, with P_3, we have $\lambda^3 - 2\lambda = \phi(P_3, \lambda)$ and

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda^2 - 1 \\ \lambda \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda \\ \lambda^2 \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} \lambda^2 - 1 \\ \lambda \\ 1 \end{pmatrix} - (\lambda^3 - 2\lambda) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

34.4.3 Theorem. *Let T be a tree with distinct eigenvalues $\lambda_1, \ldots, \lambda_d$, in decreasing order. Let z be an eigenvector for T with eigenvalues θ_i. If no entry of z is zero, it has exactly $i-1$ sign changes.* □

Proof. We start by proving that if λ is an eigenvalue of T, the number of sign-changes in the eigenvector $z_T(\lambda)$ of T is equal to the number of eigenvalues of $T \setminus 1$ that are greater than λ.

We proceed by induction on n. The case $n = 2$ is immediate, so we assume $n \geq 3$. For a graph X we use $\nu(X)$ to denote the number of eigenvalues of X greater than λ. Assume $2 \sim 1$, let U be the component of $T \setminus 1$ that contains 2 and let U' be the subgraph of T induced by the vertices not in U and not equal to 1. Let y be the restriction of $z_T(\lambda)$ to $V(U)$.

Given the obvious bijection between paths in T from 1 to j in U and paths in U from 2 to j, we have

$$y = \phi(U', \lambda) z_U(\lambda).$$

Hence, by induction, the number of sign changes of y is equal to $\nu(U \setminus 2)$.

Is there a sign change on the edge 12? Let z_i denote the value of z on the vertex i. We have

$$\phi(T \setminus 1, t) = \phi(U, t)\phi(U', t), \quad \phi(T \setminus P_{1,2}, t) = \phi(U \setminus 2, t)\phi(U', t).$$

and so the sign of $z_1 z_2$ is equal to the sign of

$$\frac{\phi(U \setminus 2, \lambda)}{\phi(U, \lambda)}$$

34.5 Supports of Eigenvectors of Trees

If e is an edge in a tree T, then $T \setminus e$ has two connected components, each of which we will call a *branch* of T. If u is a vertex in X, each component of $T \setminus u$ is a branch. We follow Brouwer and Haemers, who are following Fiedler.

34.5.1 Theorem. *If θ is an eigenvalue of the tree T and there is a θ-eigenvector of T that vanishes on some vertex of T, then there is a vertex u such that all θ-eigenvectors of X vanish on u.*

Proof. If θ is a simple eigenvector, we are done. Assume not.

If z is a θ-eigenvector and $z(u) = 0$, then $T \setminus u$ is a vertex-disjoint union of branches. Among all branches that arise in this way, choose a branch

S such that $V(S)$ is minimal and let z the θ-eigenvector used to get S. Let $e = \{u, v\}$ be the edge that joins S to the rest of T. Note that $v \in V(S)$ and that the restriction of z to $V(S)$ is a θ-eigenvector for S.

Let y be a second θ-eigenvector for T. If $y(u) \neq 0$, then y and z are linearly independent and so there is a non-zero linear combination of y and z that vanishes on a vertex of S. Since this contradicts our choice of S, we conclude that all θ-eigenvectors of T vanish on u. □

Let E_θ be the spectral idempotent associated to the eigenvalue θ of T. We have a reformulation of the theorem:

34.5.2 Corollary. *Let T be a tree with an eigenvalue θ. If some θ-eigenvector of T has an entry equal to zero, then E_θ has a row (and column) equal to zero.*

Proof. Under the hypotheses, all θ-eigenvectors vanish on some vertex u. Since we can write

$$E_\theta = \sum_{i=1}^{m} z_i z_i^T$$

where $z_1, \ldots, z_m$ is an orthonormal basis for $\ker(\theta I - A(T))$. From the theorem, there is a vertex u such that $z_i(u) = 0$ for $i = 1, ldots, m$ and accordingly $(E_\theta)_{u,u} = 0$. As E_θ is positive semidefinite, its u-row and u-column are both zero. □

Let $Z(\theta)$ denote the set of common zeros of the θ-eigenvectors of T and let m denote the multiplicity of θ. The components of $T \setminus Z(\theta)$ are the *eigenvalue components* of T for θ. You might show that the multiplicity of θ in an eigenvalue component is at most one, and that if the number of components of T is equal to c, then $m = c - |Z(\theta)|$.

The eigenvalue support of a vertex u in a graph X is equal to the set of eigenvalues of X if and only if $E_\theta e_u \neq 0$ for each eigenvalue θ. Hence if X is a tree and $u \in V(X)$, then either the eigenvalue support of u is the set of eigenvalues of X, or there is an eigenvalue θ such that $(E_\theta)_{u,u} = 0$.

34.6 The Spectral Centre of a Tree

34.6.1 Theorem. *Let T be a tree with at least two vertices and with second-largest eigenvalue λ of multiplicity m. Then there is a unique minimal subtree Y such that no eigenvalue of $T \setminus Y$ is greater than λ. If $m > 1$, then Y is a vertex; if $m = 1$ then Y is an edge.*

35
Covers

35.1 Directed Graphs

A *directed graph* consists of vertex set V, an arc set E and two maps f_0 and f_1 from E to V such that, if $\alpha \in E$ then $f_0(\alpha)$ is the tail of α and $f_1(\alpha)$ is its head. If $u, v \in V(X)$ and say that $\alpha = uv$ is an arc, we mean that $f_0(\alpha) = u$ and $f_1(\alpha) = v$. We say X is a *graph* if, for each arc α there is an arc α' such that

$$f_0(\alpha) = f_1(\alpha'), \quad f_1(\alpha) = f_0(\alpha).$$

If $X = (V_X, E_X, f_0, f_1)$ and $AY = (V_Y, E_Y, g_0, g_1)$ are directed graphs, a *homomorphism* from X to Y consists of maps $\varphi_0 : V_X \to V_Y$ and $\varphi_1 : E_X \to E_Y$ such that

$$\varphi(f_0(u)) = g_0(\varphi(u)), \quad \varphi(f_1(u)) = g_1(\varphi(u)).$$

More succinctly, $\varphi \circ f_i = g_i \circ \varphi$. In this case sat that α and α' form an *edge* of X.

A homomorphism φ is a *covering map* if, for each vertex u, if it induces a bijection from the arcs that start at u to the arcs that start at $\varphi(u)$.

35.2 Products, Coproducts

Directed graphs and their homomorphisms form a category. An object α in a category is *initial* if for each object β there is an arrow $\alpha \mapsto \beta$, unique up to isomorphism. (For example, 0 in the category of abelian groups.) The dual to an initial object is a *terminal object*. (For example, 0 in the category of abelian groups.)

Given an object α in a category $\mathscr{C}$, we form a new category whose objects are the pairs (β, f), where β is an object in $\mathscr{C}$ and $f : \alpha \to \beta$ is an arrow in $\mathscr{C}$. If (β_1, f_1) and (β_2, f_2) are pairs in our new category and g is an arrow from β_1 to β_2 such that

$$f_2 \circ g = f_1$$

then we take g to be an arrow in our new category. We denote this category by $\mathscr{C}/\alpha$ (and call it the *slice category over* α).

36
Fractional Isomorphism

36.1 Majorization

If $x, y \in \mathbb{R}^n$, we say that x *majorizes* y if there is a doubly stochastic matrix S such that $y = Sx$. We write $x \succ y$ to denote that x majorizes y. We see that

$$\mathbf{1}^T y = \mathbf{1}^T Sx = \mathbf{1}^T x.$$

As the product of doubly stochastic matrices is doubly stochastic, it follows that if $x \succ y$ and $y \succ z$, then $x \succ z$. Hence majorization is a preorder; it is not a partial order.

Let $\Pi(z)$ denote the convex hull of all vectors that arise by permuting the entries of z.[1]

[1] These vectors are all majorized by z, since permutation matrices are doubly stochastic.

36.1.1 Lemma. *Let R and S be doubly stochastic matrices. Assume that R is a convex combination of permutation matrices $P_1, \ldots, P_r$, say $R = \sum_i a_i R_i$; assume also that S is a convex combination of permutations $Q_1, \ldots, Q_s$, say $Q = \sum_i b_i Q_i$.*

(a) If x and y are vectors such that $x = Ry$ and $y = Sx$, then $x = P_i y$ and $y = Q_j x$ for all i and j.

(b) If $x = Ry$ and $y = Sx$ and S is indecomposable, then $x = y = c\mathbf{1}$ for some scalar c.

Proof. (a) If $x = Ry$, then $x = \sum_i a_i P_i y$ and hence x is a convex combination of the vectors $P_i y$. Therefore $x \in \Pi(y)$. Since $\Pi(y)$ is closed under multiplication by permutation matrices, it follows that $Px \in \Pi(y)$ for all permutation matrices P and therefore $\Pi(x) \subseteq \Pi(y)$. Similarly we find that $\Pi(y) \subseteq \Pi(x)$ and so $\Pi(x) = \Pi(y)$. Since x and y are extreme points, we conclude that $x = P_i y$ for all i and $y = Q_j x$ for all j.

(b) We have $x = PSx$ for a permutation matrix P and, since S is indecomposable, so is PS. By Perron-Frobenius, PS has a unique positive eigenvector, evidently $\mathbf{1}$. Hence $x = c\mathbf{1}$ for some c. As $y = Q_j x$, we have $y = x$. □

36.2 Fractional Isomorphism

Graphs X and Y with respective adjacency matrices A and B are *fractionally isomorphic* if there is a doubly stochastic matrix R such that $AR = RB$. A fractional automorphism of X is a fractional isomorphism from X to itself.

Since permutation matrices are doubly stochastic, isomorphic graphs are fractionally isomorphic. Any two k-regular graphs on n vertices are fractionally isomorphic, take $R = n^{-1}J$.

Let $\delta(X)$ denote the coarsest equitable partition of X; this is a refinement of the degree partition.

36.2.1 Theorem. *Let X and Y be graphs. The following claims are equivalent:*

(a) X and Y are fractionally isomorphic.

(b) $X/\delta(X)$ and $Y/\delta(Y)$ are isomorphic.

(c) X and Y have a common finite cover.

(d) X and Y have the same universal cover. □

The equivalence of the last three claims is due to Leighton [2]. The first claim was added to the list by Scheinerman et al. [3].

Our next result shows that if graphs X and Y are fractionally isomorphic, then their disjoint union $X \cup Y$ admits a fractional automorphism.

36.2.2 Lemma. *If S is doubly stochastic and $AS = SB$, then*

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} 0 & S \\ S^T & 0 \end{pmatrix} = \begin{pmatrix} 0 & S \\ S^T & 0 \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$ □

If X is a graph on n vertices with adjacency matrix A, its *walk matrix* is the matrix

$$W_X := \begin{pmatrix} \mathbf{1} & A\mathbf{1} & \cdots & A^{n-1}\mathbf{1} \end{pmatrix}.$$

Define $\Pi(X)$ to be the convex hull of the set of matrices PW_X, where P runs over all permutation matrices.

36.2.3 Theorem. *If X and Y are fractionally isomorphic, there is a permutation matrix P such that $W_Y = PW_X$.*

Proof. We use the idea of the proof of Lemma 36.1.1.

Let A and B respectively by the adjacency matrices of X and Y. Assume R is doubly stochastic and $AR = RB$. Then $A^kR = RB^k$ for all non-negative integers k and as $R\mathbf{1} = \mathbf{1}$, we have $W_X = RW_Y$. Since $R^TA = BR^T$, we also have $R^TW_Y = W_X$.

As R is doubly stochastic, we can write it as a convex combination

$$R = \sum_i a_i P_i$$

and therefore $W_X \in \Pi(Y)$. Similarly $W_Y \in \Pi(X)$, and so we conclude that there is a permutation matrix P such that $W_Y = PW_X$. □

Part VI

Rings

37
Rings

The rings $\mathbb{Z}$ and $\mathbb{F}[t]$ are principal ideal domains. Since these rings are of combinatorial interest we develop some of their basic properties. The rings we consider are (mostly) commutative.[1]

[1] Occasionally we slip up and admit matrix rings into the discussion

37.1 Ideals

A ring is *Noetherian* if there is no strictly increasing chain of ideal with infinite length. (E.g., $\mathbb{Z}$, $\mathbb{F}[t]$). A ring is *Artinian* if there is no strictly decreasing chain of ideal with infinite length. (Neither $\mathbb{Z}$ nor $\mathbb{F}[t]$ are not Artinian; finite-dimensional algebras are.)

37.1.1 Lemma. *A commutative ring R is Noetherian if and only if each ideal is finitely generated.* □

An *integral domain* is a commutative ring with no zero divisors. (We will often abbreviate "integral domain" to "domain".) An ideal I in a ring R is *prime* if $a, b \in R$ and $ab \in I$, then either a or b lies in I. The zero ideal in a domain is prime.

37.1.2 Lemma. *An ideal I in a commutative ring is prime if and only R/I is an integral domain; it is maximal if and only if R/I is a field.* □

Note that maximal ideals are prime.

Suppose the additive subgroup of a domain R has finite order n. Since R has no zero divisors, we see that the integer n is prime, and hence R is an algebra over the field $\mathbb{Z}_p$ for some prime p. If n is infinite, R contains a copy of $\mathbb{Z}$.

37.1.3 Lemma. *A finite integral domain is a field.*

Proof. Let R be a finite integral domain and assume a is a non-zero element of R. Since R is finite the sequence of powers of a has only finitely many distinct terms, and so there are integers m and n, with $m < n$ such that $a^m = a^n$. Therefore

$$0 = a^n - a^m = a^m(a^{n-m} - 1)$$

and, since F is a domain, we have $a^{n-m} = 1$. Therefore a^{n-m-1} is a multiplicative inverse to a, and we have shown that each non-zero element of R is invertible. Therefore R is a field. □

A subset S of a domain R is *multiplicatively closed* if it is closed under multiplication[2] and does not contain 0. Note that S necessarily contains the empty product 1. Examples:

[2] go figure

(a) $R \setminus 0$.

(b) More generally, the complement of a prime ideal.

(c) All powers of any element that is not a unit.

If S is multiplicatively closed, we can form the *ring of fractions* $S^{-1}R$ consisting of elements r/s (for r in R and s in S). If $S = R \setminus 0$, this gives us the *quotient field* of the domain R.

You might show that the ideals of $S^{-1}R$ are the ideals that are disjoint from S.

37.1.4 Theorem. *Let S be a multiplicatively closed subset of the domain R. If I is an ideal of R, maximal subject to $S \cap I = \emptyset$, then I is prime.*

Proof. Assume S is multiplicatively closed and I is an ideal, maximal subject to the condition $I \cap S = \emptyset$.

Assume by way of contradiction that I is not prime, and that a and b are elements of $R \setminus I$ with $ab \in I$. The ideal $I + \langle a \rangle$ contains element of S, we assume that for i in I and x in R, we have $i + ax \in S$. Similarly there if j in I and y in R such that $j + by \in S$. Now since $i, j, ab \in I$,

$$(i + ax)(j + by) = ij + iby + jax + abxy \in I$$

But this implies that S is not disjoint from I.

37.2 Principal Ideal Domains

A ring R is a *principal ideal domain*[3] if it is an integral domain and every ideal is principal. We note that if R is a pid and I is an ideal in R, the quotient R/I is again a pid.

[3] abbreviated "pid"

37.2.1 Lemma. *Any non-zero prime ideal in a principal ideal domain is maximal.*

Proof. Assume I is a non-zero prime ideal in R. Let p be a generator of I and suppose I is not maximal. Then I is properly contained in a maximal ideal M; assume that m generates M. Since $p \in M$, there is x in R such that $mx = p$. Since P is prime either $x \in P$ or $m \in P$. In the latter case, $P = M$, a contradiction. If $x \in P$, then $x = py$ (for some y in R) and therefore $myp = p$. This implies that $p(my - 1) = 0$. Since R is a domain we have $my = 1$, therefore m is invertible and $M = R$, a second contradiction. □

An *algebraic integer* is an eigenvalue of a square integer matrix. The set of all algebraic integers forms a subring of $\mathbb{C}$ that contains $\mathbb{Q}$.[4] If $\mathbb{F}$ is a finite extension of $\mathbb{Q}$, the algebraic integers in $\mathbb{F}$ form a ring, an integral domain. It may, or may not, be a principal ideal domain.[5]

We offer an example of a pid, not $\mathbb{Z}$ or $\mathbb{F}[t]$. Let θ be a complex number and let R be the set of rational functions that do not have θ as a pole. This is a commutative ring and any element of R that is not zero at θ has an inverse in R. Let I be the set of elements of R that vanish on θ, this is the ideal generated by $z-\theta$. So R has exactly one maximal ideal, which is principal. (Thus R is an example of a *local ring*.)

[4] consider the eigenvalues of $A \otimes B$ and $(A \otimes I) - (I \otimes B)$

[5] go bother a number theorist

38
Number Theory

38.1 Algebraic Integers

A complex number θ is an *algebraic integer* if it is a zero of a monic polynomial with integer coefficients. The *minimal polynomial* ψ_θ of an algebraic integer θ is the monic polynomial of least degree that is satisfied by θ. Recall that if f and g are polynomials over $\mathbb{Z}$ then there are polynomials a and b such that $af + bg$ is equal to the greatest common divisor of f and g. If f and g are monic then a and b are too. From this it follows that the minimal polynomial is unique, and that if f is a polynomial satisfied by θ, then ψ_θ divides f.

Gauss proved that if an integer polynomial factors over $\mathbb{Q}$, then it factors over $\mathbb{Z}$. Given this we deduce that the minimal polynomial of θ is irreducible over $\mathbb{Q}$.

The following lemma provides a very useful tool. Here and elsewhere I use $\leq$ to denote inclusion of modules, rings, fields....

38.1.1 Lemma. *Let M be a finitely generated $\mathbb{Z}$-submodule of $\mathbb{C}$. If $\alpha \in \mathbb{C}$ and $\alpha M \leq M$, then α is an algebraic integer.*

Proof. Suppose $b_1, \ldots, b_r$ is a generating set for M. Then the elements αb_i of M are integer linear combinations of $b_1, \ldots, b_r$, and therefore there is an integer matrix $A = (a_{i,j})$ such that

$$\alpha b_i = \sum_r a_{i,r} b_r.$$

If f is the characteristic polynomial of A then f is monic and its coefficients are integers. Since $f(\alpha) = 0$, we conclude that α is an algebraic integer. □

38.1.2 Theorem. *The set of all algebraic integers is a ring.*

Proof. Suppose α and β are algebraic integers of degree m and n respectively, and let M be the $\mathbb{Z}$-submodule of $\mathbb{C}$ generated by the mn numbers

$$\alpha^r \beta^s, \qquad 0 \leq r \leq m-1,\ 0 \leq s \leq n-1.$$

Then $(\alpha-\beta)M \le M$ and $\alpha\beta M \le M$, which tells us that $\alpha-\beta$ and $\alpha\beta$ are algebraic integers. □

Let R be a domain and let K be its field of fractions. We say that R is *integrally closed* if any element of K that satisfies a monic polynomial in $R[t]$ is an element of R. The canonical example is $\mathbb{Z}$.

38.1.3 Theorem. *Let K be a subfield of $\mathbb{C}$ and let R be the set of all algebraic integers in K. Then R is integrally closed.*

Proof. Suppose $\alpha \in K$ and that α satisfies a monic polynomial of degree n. Then the powers

$$1, \alpha, \ldots, \alpha^{n-1}$$

form a basis for $R[\alpha]$, and therefore $R[\alpha]$ is a finitely generated $\mathbb{Z}$-submodule of $\mathbb{C}$. Since it is invariant under multiplication by α, it follows that α is an algebraic integer. □

38.2 Number Fields

An *algebraic number field* K is a field K such that $\mathbb{Q} \le K \le \mathbb{C}$ and $[K:\mathbb{Q}] < \infty$. We define $\mathcal{O}_K$ to be the set of all algebraic integers that are contained in K; it is a subring of the ring of all algebraic integers. The ring $\mathbb{Z}[\theta]$ is a subring of $\mathcal{O}_K$, but it need be equal to it. (Consider $\mathbb{Q}(\sqrt{5})$.)

38.2.1 Theorem. *If K is an algebraic number field then $K = \mathbb{Q}(\theta)$ for some algebraic integer θ.*

Proof. Suppose α and β are algebraic numbers. We show that there is an algebraic number γ such that

$$\mathbb{Q}(\alpha,\beta) = \mathbb{Q}(\gamma).$$

Let f and g respectively be the minimal polynomials of α and β, and let $\alpha_1, \ldots, \alpha_k$ and $\beta_1, \ldots, \beta_\ell$ respectively be the distinct roots of these polynomials in $\mathbb{C}$. Assume $\alpha = \alpha_1$ and $\beta = \beta_1$ and choose a rational number e such that

$$\alpha + e\beta \neq \alpha_i + e\beta_j, \quad i, j \neq 1.$$

Set $\gamma = \alpha + e\beta$. Then both polynomials $f(\gamma - et)$ and $g(t)$ vanish at β, and therefore both polynomials are divisible by $t-\beta$. Now $f(\gamma - et) = 0$ if and only if $\gamma - et = \alpha_i$ for some i, and this is equivalent to requiring that

$$\alpha + e\beta - et = \alpha_i$$

and hence that

$$e = \frac{\alpha_i - \alpha}{\beta - t}.$$

But if $g(t) = 0$ and $t \neq \beta$ then $t = \beta_j$ (with $j \neq 1$), so given our choice of e, we see that β is the only common zero of $f(\gamma - et)$ and $g(t)$.

Now the greatest common divisor of $f(\gamma - et)$ and $g(t)$ belongs to $\mathbb{Q}(\gamma)[t]$, whence $t - \beta \in \mathbb{Q}(\gamma)[t]$ and thus $\beta \in \mathbb{Q}(\gamma)$. Consequently $\alpha \in \mathbb{Q}(\gamma)$ and therefore $\mathbb{Q}(\gamma) = \mathbb{Q}(\alpha, \beta)$. It follows by an easy induction that $K = \mathbb{Q}(\gamma)$ for some algebraic number γ; since some integer multiple of γ is guaranteed to be an algebraic integer, the theorem holds. □

If K is a number field and $a \in K$, then $L = \mathbb{Q}(a)$ is a subfield of K and therefore K is a vector space over L. Hence

$$[K : \mathbb{Q}] = [K : L][L : \mathbb{Q}],$$

which shows that the degree of L divides the degree of K.

38.2.2 Lemma. *If K is a number field and $a \in K$, there is an integer m such that ma is an algebraic integer.* □

Proof. Suppose that a is algebraic and

$$f_0 a^d + \cdots + f_d = 0$$

for suitable integers f_r. If we multiply this by f_0^{d-1} we find that

$$0 = (f_0 a)^d + f_0 f_1 (f_0 a)^{d-1} + \cdots + f_0^d f_d$$

and therefore $f_0 a_d$ is an algebraic integer. □

38.2.3 Corollary. *Any number field has a basis that consists of algebraic integers.* □

Equivalently, $\mathcal{O}_K$ contains a basis for K.

38.3 Galois

Two algebraic integers are *conjugate* if they have the same minimal polynomial. If θ is an algebraic integer and

$$\psi_\theta(t) = t^d - a_1 t^{d-1} - \cdots - a_d$$

then a_1 is the sum of all algebraic conjugates of θ (since it is the sum of the zeros of ψ_θ) and thus we see that this sum is an integer. Similarly a_d is, up to sign, the product of the zeros of ψ_θ and so the product of the conjugates of θ is an integer. More generally a_k is the sum of all products of k distinct conjugates, and we are lead to the conclusion that any symmetric function in the conjugates of θ is an integer.

If $\mathbb{E}$ is an extension of a field $\mathbb{F}$, then the *Galois group* of the extension is the group of automorphisms of $\mathbb{E}$ that leave each element of $\mathbb{F}$ fixed. Here we are only concerned with the case where $\mathbb{E}$ is contained in $\mathbb{C}$ and its dimension as a vector space over $\mathbb{F}$ is finite. If ψ is a polynomial with

coefficients in $\mathbb{F}$ and θ is an element of $\mathbb{E}$ that satisfies ψ, then any element of the Galois group must map θ to another zero of ψ. In particular, if ψ is the minimal polynomial of θ over $\mathbb{F}$ then any element of the Galois group must permute those zeroes of ψ that lie in $\mathbb{E}$. This if $\mathbb{F} = \mathbb{Q}$ and ψ is the minimal polynomial of an algebraic integer θ, then each element of the Galois group permutes the set of algebraic conjugates of θ that lie in $\mathbb{E}$.

Now suppose we choose $\mathbb{E}$ to be an extension of $\mathbb{Q}$ such that ψ factors into linear terms in $\mathbb{E}$ and, given this, the dimension of $\mathbb{E}$ over $\mathbb{Q}$ is minimal. (We call $\mathbb{E}$ a *splitting field* for ψ over $\mathbb{Q}$. They exist and they are all isomorphic, but proof is required.) Then

(a) The Galois group acts transitively on the zeros of ψ: in fact two elements of $\mathbb{E}$ are algebraically conjugate if and only if one is mapped to the other by an element of the Galois group.

(b) Any element of $\mathbb{E}$ that is left fixed by each element of the Galois group is an integer.

Here (a) follows from the fact that if ψ is irreducible over $\mathbb{E}$ and a and b are roots of ψ in $\mathbb{E}$, then $\mathbb{E}(a)$ and $\mathbb{E}(b)$ are isomorphic; the second is a consequence of the result that if $\mathbb{E}$ is a splitting field over $\mathbb{F}$, then the fixed field of the Galois group is $\mathbb{F}$.

38.4 Trace and Norm

If K is a finite extension of a field L and $a \in K$ then multiplication by a is an L-linear map from K to itself. The *trace* $\mathrm{Tr}_{K/L}(a)$ of a is the trace of this map. The *norm* $N_{K/L}(a)$ of a is the determinant of this map. You might verify the following relations hold when L is in turn a finite extension of a subfield M:

$$\mathrm{Tr}_{K/M}(a) = [K : L]\, Tr_{L/M}(a), \qquad N_{K/M}(a) = N_{L/M}(a)^{[K:L]}.$$

Both the norm and trace lie in the base field.

Once we decide to view the elements of K as linear operators on a vector space over L, we may choose to go further and view them as $n \times n$ matrices over L (where $n = [K : L]$). Now K can be viewed as a subalgebra of the $n \times n$ matrices over L, and clearly it is a commutative algebra. As an algebra it is semisimple—the only nilpotent ideal is the zero ideal. If $K \leq \mathbb{C}$ and $L = \mathbb{Q}$, it follows that we can simultaneously diagonalize the elements of K. In other word $\mathbb{C}^n$ is the direct sum of n 1-dimensional subspaces each of which is an eigenspace for each element of K. If we fix a common eigenvector, z say, then the map that sends an element of K to its z-eigenvalue is a homomorphism from K to $\mathbb{C}$. This map is known as a *complex embedding* of K.

38.4.1 Lemma. *If $\sigma_1, \ldots, \sigma_n$ are the distinct complex embeddings of the number field $K = \mathbb{Q}(a)$ then*

$$\mathrm{Tr}_{K/\mathbb{Q}}(a) = \sum_{i=1}^{n} \sigma_i(a), \qquad N_{K/\mathbb{Q}}(a) = \prod_{i=1}^{n} \sigma_i(a). \qquad \square$$

If α is an algebraic integer, then its trace and norm are integers.

The map

$$(x, y) \mapsto \mathrm{Tr}_{K/\mathbb{Q}}(xy)$$

is the *trace form* on K.

38.4.2 Theorem. *The trace form is non-degenerate bilinear form on K.*

Proof. Only degeneracy is a problem. If $\mathrm{Tr}(ax) = 0$ for all x, then $\mathrm{Tr}(a^k) = 0$ whenever $k > 1$. We can express the coefficients in a monic polynomial in terms of the powers of its roots, using these expressions we find that if $K = \mathbb{Q}(a)$, then the minimal polynomial of a has the form t^m, and hence $a = 0$. $Q(a)$ is a proper subfield, use transitivity. $\square$

The fact that the trace form is non-degenerate provides another way to prove that K is semisimple as an algebra over $\mathbb{Q}$. Number theorists would say that K is a *separable extension*.

38.5 Lattices

We define a *lattice* to be an additive subgroup of a real vector space V that is generated by a basis. Equivalently it is an $\mathbb{Z}$-submodule of V with rank equal to $\dim(V)$. We say L_1 is a *sublattice* of L_2 if L_1 and L_2 are lattices and $L_1 \leq L_2$.

A subset L of $\mathbb{R}^n$ is *discrete* if there is a positive real number ϵ such that the balls of radius ϵ about the elements of S are pairwise disjoint. Any lattice is discrete. For suppose $a_1, \ldots, a_n$ is a basis that generates L and let A be the matrix with $a_1, \ldots, a_n$ as its columns. Then $A^T A$ is positive semidefinite, with least eigenvalue λ say, and consequently

$$\|Az\|^2 = z^T A^T A z \geq \lambda z^T z.$$

If z is integral and not zero, then $z^T z \geq 1$.

38.5.1 Theorem. *Any discrete additive additive subgroup of $\mathbb{R}^n$ that contains a basis is a lattice.*

Proof. Suppose L is an additive subgroup of $\mathbb{R}^n$ that contains basis $y_1, \ldots, y_n$. We must show that there is a basis $x_1, \ldots, x_n$ such that each vector in L is an integral linear combination of $x_1, \ldots, x_n$. Let V_k denote the vector space spanned by $y_1, \ldots, y_k$.

Choose x_1 in L so that $x_1 = c_1 y_k$ where $c_1 > 0$ and is minimal. Now assume we have have $x_1, \ldots, x_k$ in hand. The set of vectors in L

$$c_{k+1} y_{k+1} + \sum_{i=1}^{k} c_{k+1,i} y_i$$

where all coefficients lie in $[0,1]$ is bounded. Since L is discrete, there is a vector in this set for which c_{k+1} is minimal and positive, and we take this to be x_{k+1}.

To complete the proof we must show that $x_1,\dots,x_n$ is an integral basis for L. Suppose it is not, and choose a vector x in L that is not an integer linear combination of $x_1,\dots,x_n$. Assume

$$x = \sum_{i=1}^{n} d_i x_i$$

where the coefficients d_i are not all integers. If d_k is the largest-indexed coefficient that is not an integer, then

$$w = \sum_{i=1}^{k} d_k x_k$$

is a vector in L that is not an integer linear combination of $x_1,\dots,x_n$. Replacing w by

$$w - \lfloor d_k \rfloor x_k$$

if needed, we may assume that $0 < d_k < 1$. But now if we express $x_1,\dots,x_k$ in terms of $y_1,\dots,y_k$, then in the resulting expression for w the coefficient of y_k is $d_k c_k < c_k$. This contradicts our choice of x_k. □

In the context of the above proof, let X and Y be the matrices with columns $x_1,\dots,x_n$ and $y_1,\dots,y_n$ respectively. Then the first part of proof shows that $X = YC$ where C is lower triangular and $0 < C_{i,i} < 1$ for all i. Since the columns of X form an integral basis and $Y = XC^{-1}$ it follows that C^{-1} is an integer matrix.

38.5.2 Lemma. *Let M be a finitely generated $\mathbb{Z}$-submodule of K. If $\alpha \in K$ and $\alpha M \le M$, then $\alpha \in \mathcal{O}_K$ and $|N(\alpha)| = [M : \alpha M]$.*

Proof. Since M is a lattices in its span over K, it has a free basis $\omega_1,\dots,\omega_r$. Let A be the matrix that represents multiplication by α on M. Then $|\det(A)| = [M : AM]$. Let $f(t)$ be the characteristic polynomial of A. Then f is irreducible and therefore f is the minimal polynomial of A. So $\det(A) = N(\alpha)$. □

38.6 Dual Lattices

If $a_1,\dots,a_n$ is a basis of V, its *dual basis* is the sequence $\beta_1,\dots,\beta_n$ from the dual space V^* of V, such that

$$\beta_j(a_i) = \delta_{i,j}.$$

A *character* on a lattice is an element, γ say, of the dual space V^* that takes only integer values on L. If $a_1,\dots,a_n$ is an integral basis for L then its

dual basis $\beta_1, \ldots, \beta_n$ is an integral basis for the $\mathbb{Z}$-module generated by the characters. So they generate a lattice L^*, called the *dual lattice* of L.

In practice we identify the dual space V^* with V by choosing an isomorphism. An isomorphism from V to V^* is a non-degenerate bilinear form. If ψ is such a form then a dual basis to $a_1, \ldots, a_n$ consists of vectors $b_1, \ldots, b_n$ such that

$$\psi(b_j, a_i) = \delta_{i,j}$$

and L^* is the lattice generated by $b_1, \ldots, b_n$. Of course the customary form is $x^T y$, in which case if $a_1, \ldots, a_n$ are the columns of A, then the columns of A^{-T} are the dual basis. Strictly speaking, to define a lattice we should always specify a bilinear form. We instead take the view that, if the form is not specified then we are using the usual Euclidean inner product $\langle , \rangle$. A lattice is *Euclidean* if the bilinear form is positive semidefinite.

By way of example consider the lattice L generated by the standard basis in $\mathbb{R}^n$. Then $\langle e_j, e_i \rangle = \delta_{i,j}$ and so $L = L^*$. More generally, if L has the property that $\langle a, b \rangle \in \mathbb{Z}$ for all a and b in L, then $L \leq L^*$. In this case we say that L is an *integral*) lattice. (This does not necessarily mean that the vectors in L are integer vectors. For example if L is generated by the columns of A, then L is integral if $A^T A$ is integral.)

38.7 Index and Orders

We use $\mathcal{O}_K$ to denote the set of all algebraic integers in K; this is a ring (since it is the intersection of two subrings of $\mathbb{C}$).

Suppose R is a subring of S; if we view S as an additive group (or $\mathbb{Z}$-module) then R is a subgroup. The *index* of R in S is the order of the quotient group S/R. An *order* of K is a subring of $\mathcal{O}_K$ that, viewed as an additive group, has finite index in $\mathcal{O}_K$. If $K = \mathbb{Q}(\theta)$ then $\mathbb{Z}[\theta]$ is an order. Ideals are subrings and we will use the next result to show that all ideals of $\mathcal{O}_K$ are orders.

38.7.1 Lemma. *If α is an algebraic integer then $N(\alpha)$ is an integer polynomial in α.*

Proof. If ψ is the minimal polynomial of α then we can write it in the form

$$\psi(t) = t\varphi(t) + N_{\mathbb{Q}(\alpha)/Q}(\alpha).$$

where the coefficients of φ are integers. Since $\psi(\alpha) = 0$, this implies that $N(\alpha)$ in an integer polynomial in α. □

38.7.2 Lemma. *Any ideal of an order is an order.*

Proof. Suppose R is an order in $\mathcal{O}_K$ and J is a ideal in R. If α is a non-zero element of J, then $m = N(\alpha)$ is an integer in J. Simce the ideal mR has finite index in R, so must J. □

38.7.3 Theorem. *Any order in $\mathcal{O}_K$ has an integral basis.*

Proof. Suppose R is an order in $\mathcal{O}_K$. Since it has finite index in $\mathcal{O}_K$, there is an integer m such that $[\mathcal{O}_K : R]\mathcal{O}_K \leq R$, and therefore R contains a basis. Hence it is a lattice and therefore it has an integral basis. □

For us the most important example of an order will be the ring $\mathbb{Z}[\theta]$.

We note the following, but leave the proof as an exercise.

38.7.4 Lemma. *A subring R of $\mathcal{O}_K$ is n order if and only if the field of fractions of R is equal to K.* □

38.8 Discriminants

Let K be a number field with degree n and let $\sigma_1, \ldots, \sigma_n$ be its complex embeddings. If $(\alpha_1, \ldots, \alpha_n)$ is an n-tuple of elements of K, we define the *discriminant* $\operatorname{disc}(\alpha)$ of α to be

$$\det(\sigma_i(a_j))^2.$$

If $a_1, \ldots, a_n$ is an integral basis for $\mathcal{O}_K$, its discriminant is called the discriminant of K and be denoted by d_K. And to add to the confusion, the discriminant of a polynomial f is the resultant of f and f', and is denoted by $\operatorname{disc}(f)$.

38.8.1 Lemma. *If $(\alpha_1, \ldots, \alpha_n)$ is an n-tuple of elements of K, its discriminant is equal to*

$$\det(\operatorname{Tr}(a_i a_j)).$$

Proof. Define the $n \times n$ matrix S by $S_{i,j} = \sigma_i(a_j)$. Then

$$(S^T S)_{i,j} = \sum_{r=1}^{n} \sigma_r(a_i)\sigma_r(a_j) = \sum_{r=1}^{n} \sigma_r(a_i a_j) = \operatorname{Tr}(a_i a_j).$$

Since $\operatorname{disc}(\alpha) = \det(S^T S)$, our result follows. □

Suppose θ is an algebraic integer with minimal polynomial f of degree n. Then

$$\operatorname{disc}(1, \theta, \ldots, \theta^{n-1}) = (-1)^{n(n-1)2} \operatorname{disc}(f) = N(f'(\theta)).$$

38.8.2 Lemma. *Let θ be an algebraic integer whose conjugates are all real. Then the trace form on $\mathbb{Q}(\theta)$ is positive definite.*

Proof. Let f be the minimal polynomial of θ. Our hypothesis implies that all roots of f are real. Let $K = \mathbb{Q}(\theta)$. All entries of the matrix $S = (\sigma_i(\theta^j))$ are real and therefore $S^T S$ is positive definite. □

Taussky has proved that if K has exactly r real embeddings and $2s$ complex embeddings, then the matrix $\operatorname{Tr}(a_i a_j)$ (for an integral basis $a_1, \ldots, a_n$) has $r + s$ positive eigenvalues and s negative eigenvalues.

If $m = [\mathcal{O}_K : \mathbb{Z}[\theta]]$, then

$$\operatorname{disc}(1, \theta, \ldots, \theta^{n-1}) = m^2 d_K$$

Stickleberger's criterion asserts d_K is always congruent to 0 or 1 mod 4. If d_K is square-free then $\mathcal{O}_K = \mathbb{Z}[\theta]$.

38.9 Modules

Let K be a number field. We are interested in finitely generated $\mathbb{Z}$-submodules of K. In this context we will refer to them simply as modules. If $[K : \mathbb{Q}] = n$ than any module has rank at most n, if equality holds we will call it a *lattice*. (This is an abuse of notation.) If M and N are modules then they have a well-defined product

$$MN = \{xy : x \in M, y \in N\}$$

which is again a module. Similarly their sum $M + N$ is a module. If M is a module and $\alpha \in K$ then αM is a module. Two modules M and N are *equivalent* if there is α in K such that $N = \alpha M$.

If $\alpha \in K$ and $\alpha M \leq M$ then, by Lemma 38.1.1, it follows that $\alpha \in \mathcal{O}_K$. The set

$$\{\alpha \in K : \alpha M \leq M\}$$

is a subring of $\mathcal{O}_K$. This subring is an invariant of the equivalence class of M, and M is a module over this ring.

Suppose R is a subring of $\mathcal{O}_K$ whose field of fractions coincides with K. Then R must contain a basis for K over $\mathbb{Q}$ and therefore R is a lattice.

38.9.1 Lemma. *If R is an order in the number field K, then the number of equivalence classes of R-modules is finite.*

Proof. Let $\omega_1, \ldots, \omega_n$ be an integral basis for R. If σ is a complex embedding of K, define

$$r_\sigma := \sum_{i=1}^{n} |\omega_i^\sigma|.$$

and set $\rho = \prod_\sigma r_\sigma$. Suppose

$$\delta = \sum_{i=1}^{n} h_i \omega_i$$

where $h_1, \ldots, h_n \in \mathbb{Z}$ (so $\delta \in R$) and set H equal to $\max\{|h_i|\}$. Then

$$|N(\delta)| = \prod_\sigma |h_1 \omega_1^\sigma + \cdots + h_n \omega_n^\sigma| \leq H^n \rho.$$

Let M be an R-module. Since K is the quotient field of R, there is an element α in K such that $\alpha M \leq R$, so we may suppose $M \leq R$. Choose the integer k so that

$$k^n \leq [R : M] < (k+1)^n.$$

Then the number of residue classes of R mod M is less than $(k+1)^n$, and therefore one of these residue classes must contain at least numbers β, γ of the form

$$\ell_1\omega_1 + \cdots + \ell_n\omega_n$$

where the ℓ_i are integers in the interval $[0, \ell]$. So $\alpha = \beta - \gamma \in M$.

From our initial calculation

$$|N(\alpha)| \leq \ell^n R \leq [R : M]\rho$$

and, as $\alpha \in M$, we see that $\alpha R \leq M$. Therefore

$$M \leq R \leq \alpha^{-1} M.$$

and

$$[\alpha^{-1}M : R][R : M] = [\alpha^{-1}M : M] = [M : \alpha M] = |N(\alpha)|.$$

Accordingly

$$[\alpha^{-1}M : R] = |N(\alpha)|[R : M]^{-1} \leq \rho$$

and this we have shown that M is equivalent to a module $N = \alpha^{-1}M$ such that $N \geq R$. If $a := [N : R]$ then $aN \leq R$ and hence

$$R \leq N \leq a^{-1}R.$$

Since $a \leq \rho$ there are only finitely many choices for a and, for each possible value of a, there are only finitely many submodules in $a^{-1}R/R$. We conclude that there are only finitely many equivalence classes of R-modules in K. □

Notes

By the time most of this chapter was written, my favorite sources for algebraic number theory were Koch "Number Theory" [1], Stein's "Algebraic Number Theory" (available on-line at http://wstein.org/books/ant/), Cassel's [2] "Local Fields" , Fröhlich and Taylor [3], and Lorenzini [4] "An Invitation to Arithmetic Geometry". The material on lattices is from Siegel [5] (which I like very much).

[1]
[2]
[3]
[4]
[5]

39
Integral Similarity of Matrices

39.1 Matrix Algebras

Two $n \times n$ integer matrices are *integrally similar* if there is a matrix L in $GL(n,\mathbb{Z})$ such that $L^{-1}AL = B$. (Note that $\det(L) = \pm 1$.) We would like to be able to decide when two matrices are integrally similar. We will prove a theorem due to Latimer and MacDuffee which reduces the problem to number theory, provided the characteristic polynomial of A is separable over $\mathbb{Q}$. (A polynomial f is *separable* if f and f' are coprime.) If f is separable then the algebra $\mathbb{Q}(t)/(f)$ is isomorphic to a sum of fields $\mathbb{Q}(t)/(f_i)$, where f_i runs over the irreducible factors of f.

39.1.1 Theorem. *Let f be a separable polynomial in $\mathbb{Z}[t]$, let K be the algebra $\mathbb{Q}[t]/(f)$ and let θ be the image of t in K. Then the integer equivalence classes of matrices with f as characteristic polynomial correspond to the ideal classes in $\mathbb{Z}[\theta]$.*

Proof. Suppose A is an integer matrix with f as its characteristic polynomial.

If M is the adjugate matrix of $\theta I - A$, then

$$(\theta I - A)M = \det(\theta I - A)I = f(\theta)I = 0$$

and so the columns of M are eigenvectors for A with eigenvalue θ. Let z be the first column of M.

If y is a vector with entries from R, let (y) denote the R-module generated by its entries.

Since $\theta M = AM$, we see that $\theta(z) \subseteq (z)$ and therefore (z) is an ideal of $\mathbb{Z}[\theta]$. If y is any eigenvector for A with entries from R and eigenvalue θ, then $y = cz$ for some c in K and, since K is quotient field of $\mathbb{Z}[\theta]$, the ideals (y) and (z) are equivalent. Thus A determines an equivalence class of ideals in $\mathbb{Z}[\theta]$. If N is an invertible integer matrix and $A = N^{-1}BN$ then Nz is an eigenvector for B with eigenvalue θ. As $(Nz) = (z)$, we see that B and A determine the same ideal class.

For the converse, suppose J and K are equivalent ideals in $\mathbb{Z}[\theta]$, say

$$\alpha J = \beta K$$

for α, β in R. If $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ are integral bases for J and K respectively and

$$\frac{\alpha}{\beta} a_i \in K$$

for all i. Hence if a and b denote the vectors with entries $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, there is an integer matrix C such that

$$\frac{\alpha}{\beta} a = Cb.$$

Similarly there is an integer matrix D such that

$$\frac{\beta}{\alpha} b = Da$$

and we see that $CD = I$. Now if M and N are integer matrices such that

$$\theta a = Ma, \qquad \theta b = Nb$$

and now

$$CNb = \theta Cb = \frac{\alpha}{\beta} \theta a = MCb.$$

Since C, M and N are integers this implies that $CNb^\sigma = MCb^\sigma$ for all σ in the Galois group of K, and consequently $CN = MC$. Thus M and N are integrally similar. □

39.2 Matrix Similarity over GF(p)

If integer matrices A and B are similar over $\mathbb{Z}$, then their images mod p are similar over $GF(p)$. It follows that if the adjacency matrices of two graphs are integrally similar, then the matrices must have the same rank mod p, for all primes. For strongly-regular graphs, the p-rank of the adjacency matrix is determined by the parameters expect possibly when p divides $(\theta - \tau)^2$.

However equivalence mod p for all primes does not imply integral equivalence, as we show now. Suppose

$$A = \begin{pmatrix} 1 & -5 \\ 3 & -1 \end{pmatrix}.$$

If

$$U = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{5}{3} \end{pmatrix}$$

then $UAU^{-1} = A^T$. Hence A and A^T are similar over $GF(p)$ if $p \neq 3$. If $p = 3$ and we take

$$U = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

then, over $GF(3)$, we again get $UAU^{-1} = A^T$. Therefore A and A^T are conjugate over $GF(p)$, for all primes p.

The characteristic polynomial of A is $t^2 + 14$. Set θ equal to $\sqrt{-14}$. Then the adjugate of $\theta I - A$ is

$$\begin{pmatrix} \theta + 1 & -5 \\ 3 & \theta - 1 \end{pmatrix}$$

The minimal polynomial of A over $\mathbb{Q}$, when reduced mod p, need not be the minimal polynomial of A over $\mathbb{Z}_p$ (although the latter must divide the former). For example the minimal polynomial of the following matrix over $GF(2)$ is $x^3 + x$ but the mod 2 reduction of the minimal polynomial over $\mathbb{Q}$ is $x^5 + x$.

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

39.3 An Example

If we want to decide if the matrix

$$A = \begin{pmatrix} 1 & -5 \\ 3 & -1 \end{pmatrix}$$

is similar (over some ring) to its transpose then we can first solve the equations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -5 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ -5 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

which are equivalent to

$$\begin{aligned} a + 3b &= a + 3c \\ -5a - b &= b + 3d \\ c + 3d &= -5a - c \\ -5c - d &= -5b - d. \end{aligned}$$

This implies that $b = c$ and

$$d = -\frac{1}{3}(5a + 2b),$$

and thus the general solution has the form

$$\begin{pmatrix} a & b \\ b & -\frac{1}{3}(5a + 2b) \end{pmatrix}.$$

The determinant of this matrix is

$$-\frac{1}{3}(5a^2+2ab-3b^2)=\frac{1}{3}(a+b)(5a-3b)$$

and so A and A^T are integrally similar if and only if we can choose integers a and b so that $3|(a+b)$ and $(a+b)(5a-3b)=\pm 3$. It is easy to verify that there are no solutions.

If the characteristic polynomial of a matrix A is irreducible and $A^T = L^{-1}AL$ then L is necessarily symmetric.

Notes

The theory of integral equivalence of matrices starts with Latimer and MacDuffee [1]. (It is also treated in Newman [2], but there is more information in the original, and it's no harder to read.) Our treatment follows Buccino [3].

1
2
3

Part VII

Problems and Projects

40
State Transfer

1. Show that perfect state transfer does not occur on trees with more than three vertices. There are many interesting special cases. For example, show that if we have perfect state transfer between adjacent vertices on a tree, the the tree is P_2. Or show that if we have perfect state transfer between vertices of valency one with a common neighbour, then we are on P_3.

2. Let u and v be vertices in X, let S and T denote the set of edges of X that contain u and v respectively (we view these as subsets of the vertices of the line graph $L(X)$) and let χ_u and χ_v be the characteristic vectors of these subsets. If, relative to the unsigned Laplacian, we have perfect state transfer from u to v, then there is perfect state transfer (relative to the adjacency matrix) from $\chi_u\chi_u^T$ to $\chi_v\chi_v^T$ in $L(X)$. (See Subsection **??**.) We note that S and T induce cliques in $L(X)$.

   Find an example of a graph X where we have state transfer from D_S to D_T on $L(X)$.

3. In all known examples of perfect state transfer, the phase factor is a root of unity. Is there an example where it is not? (It can be shown that if the sum of the eigenvalues in the eigenvalue support is zero, the phase factor is a root of unity. I don't know of examples of perfect state transfer where this sum is not zero.)

4. For which odd integers $n \geq 5$ is there a time $t > 0$ such that $U_{P_n}(t)_{1,n} = 0$?

5. Which cubelike graphs admit pst at time $\pi/8$? (This is really asking for a characterization of certain codes.)

6. Investigate graphs where $\frac{1}{n}J$ is periodic. What happens at half the period?

7. Let C_1 and C_2 be cliques of size k in a strongly regular graph X and let D_1 and D_2 respectively denote the corresponding diagonal density matrices. Investigate when we get transfer from D_1 to D_2 using $A(X)$ as the Hamiltonian. [*** Never ***]

8. Let X be a graph on k vertices. Identify $V(X)$ with a k-subset of $\{1,\dots,n\}$, giving us an n-vertex graph Y_1. Construct Y_2 from X using a different identification. Then the Laplacians of Y_1 and Y_2 gives us density matrices D_1 and D_2. Investigate when we can get transfer from D_1 to D_2 using the Laplacian of some graph with vertex set $\{1,\dots,n\}$.

41
Mixing

1. Which odd cycles admit uniform mixing? The only known example is K_3, the first open case is C_9. (Natalie proved that no cycle of prime order p with $p > 3$ admits uniform mixing.)

2. If the average mixing matrix of a connected graph X has rank one, then $|V(X)| \leq 2$. Are there infinitely many connected graphs whose average mixing matrix has rank two? (Ferdinand Ihringer and ??? have an example on 64 vertices.)

3. The only known non-regular graphs with uniform mixing are $K_{1,3}$ and its Cartesian powers. Find another example.

4. Investigate local uniform mixing (and pst) on oriented graphs. Tournaments seem like a good place to start.

42
None of the Above

1. Is every totally positive algebraic integer a Laplacian eigenvalue?
2. If X is switching self-complementary, does it follow that $|V(X)|$ is congruent to 0 or 1 modulo four? (See Section **??**.)
3. Is there a graph X which is not vertex-transitive, but all its 1-vertex deleted subgraphs have the same matching polynomial?
4. Are almost all graphs characterized by their matching polynomial?
5. Is the stabilizer of a vertex in $\mathrm{Aut}(\mathscr{T}(\mathbb{Z}))$ trivial? (It's possible that this is not difficult.)

Part VIII

I’m thinking...

43
Colouring Projective Spaces

Problem: can we colour the points of the real projective plane with three colours, so that all three colours are used, and no line contain all three colours?
Solution: choose a point p and a line L on p; colour p red, colour the points of $L \setminus p$ white, and colour the remaining points blue.
Problem: is there another solution?

43.1 Valuations

A *valuation* of a field $\mathbb{F}$ is a real function v (say) trying be a metric. It satisfies

(a) $v(x) = 0$ if and only if $x = 0$.

(b) $v(xy) = v(x)v(y)$.

(c) $v(x+y) \le v(x) + v(y)$.

Note that (b) implies that $v(1) = 1$. The immediate example of a valuation is the usual absolute value on $\mathbb{R}$; the norm on $\mathbb{C}$ is a second example. The valuations relevant to us satisfy a strengthening of (c):

(c+) $v(x+y) \le \max\{v(x), v(y)\}$.

If this holds, we have a *non-archimedean valuation*. Note that we will often denote a valuation by $|x|$ rather than $v(x)$.

Given a valuation, a sequence $(a_r)_{r\ge 0}$ converges to 0 if the sequence of real numbers $(v(a_r))_{r\ge 0}$ converges to 0.

For our first example of a non-archimedean valuation, take $\mathbb{F}$ to be the field of Laurent series in t over your favourite integral domain. If f is a Laurent series, there is a unique integer k and a series f_0 with nonzero constant term such that

$$f(t) = t^k f_0(t)$$

Define $v(f) := 2^{-k}$.[1]

[1] and check that it is a non-archimedean valuation

Our second example is a family of valuations on the rationals. Choose a prime p. If $x \in \mathbb{Q}$, there is an integer k and integers a and b coprime to p such that

$$x = p^k \frac{a}{b}.$$

Define $|x|_p$ to be 2^{-k}. This is the *p-adic valuation*.

If v is a non-archimedean valuation on $\mathbb{F}$, the set

$$\mathcal{O} := \{x : v(x) \leq 1\}$$

is a ring, called a *valuation ring*. The subset

$$\mathcal{M} := \{x : v(x) < 1\}$$

is a maximal ideal. (To see this note that if $v(x) = 1$, then $x^{-1} \in \mathcal{O}$ and consequently any element of $\mathcal{O} \setminus \mathcal{M}$ is invertible. Therefore any ideal of $\mathcal{O}$ is contained in $\mathcal{M}$.) The quotient $\mathcal{O}/\mathcal{M}$ is a field, the *residue class field*. If $|\cdot|_p$ is the p-adic valuation on $\mathbb{Q}$, then $\mathcal{O}$ is the ring of p-adic integers, a subring of $\mathbb{Z}$, and the residue class field is $\mathbb{Z}_p$.

What do we do with valuations? Given a valuation $|\cdot|$ on $\mathbb{F}$ we can define *Cauchy sequences* over $\mathbb{F}$. The set of all sequences $\mathbb{F}^{\mathbb{Z}}$ forms a ring under pointwise addition and multiplication. Declare two Cauchy sequences to be *equivalent* if their difference converges to zero. The *completion* of $\mathbb{F}$ relative to our evaluation is the ring of equivalence classes of Cauchy sequences. You recall this, because if $\mathbb{F} = \mathbb{R}$ and $|\cdot|$ is the familiar valuation on $\mathbb{Q}$, this is a standard construction of $\mathbb{R}$.

Using the strong version of the triangle inequality, we can prove that the sequence of partial sums $\sum_r a_r$ is a Cauchy sequence if and only if $(a_r)_{r\geq 0}$ converges to 0.[2]

[2] this is an improvement over what we meet in Calculus

The completion of $\mathbb{Q}$ relative to $|\cdot|_p$ is the field of *p-adic numbers*.

43.2 Valuation Rings

A subring D of $\mathbb{F}$ is a *valuation ring* if for each non-zero element x of $\mathbb{F}$, either $x \in D$ or $x^{-1} \in D$. As

$$1 = |1| = |xx^{-1}| = |x||x^{-1}|,$$

the ring $\mathcal{O}$ from the previous section is a valuation ring. The fraction field of D is $\mathbb{F}$.

We present three interesting properties of valuation rings.

43.2.1 Lemma. *Assume $\mathcal{O}$ is a valuation ring in $\mathbb{F}$. Then*

(a) The set $\mathcal{M}$ of non-units in $\mathcal{O}$ form an ideal.

(b) Any proper ideal of $\mathcal{O}$ is contained in $\mathcal{O}$.

(c) $\mathcal{O}$ is integrally closed in $\mathbb{F}$.[3]

[3] we will not need integral closure, but it is very important

Proof. If $a, b \in \mathscr{M}$ then either a/b or b/a lies in $\mathscr{M}$. If $a/b \in \mathscr{M}$, then $1 + a/b$ is not in $\mathscr{M}$ (or we would have $1 \in \mathscr{M}$). Therefore $1 + a/b$ is a unit. If

$$a + b = b\left(1 + \frac{a}{b}\right)$$

were a unit, this would imply that b is a unit. So $a + b \in \mathscr{M}$. If $a/b \notin \mathscr{M}$, then $b/a \in \mathscr{M}$ and

$$a + b = a\left(1 + \frac{b}{a}\right).$$

This again yields that $a + b \in \mathscr{M}$. If $x \in \mathscr{O}$ and $a \in \mathscr{M}$, then xa cannot be a unit, and therefore $xa \in \mathscr{M}$. This proves (a).

The elements of any ideal are non-units, so (b) holds.

For (c), first an explanation. A subring D of $\mathbb{F}$ is *integrally closed* whenever α in $\mathbb{F}$ is a root of a monic polynomial $p(t)$ from $D[t]$, we have $\alpha \in D$.[4] So suppose we have and equation

[4] think of $D = \mathbb{Z}$ in $\mathbb{F} = \mathbb{Q}$

$$\alpha^n + c_{n-1}\alpha^{n-1} + \cdots + c_1\alpha + c_0 = 0.$$

with $c_0, \ldots, c_{n-1}$ in $\mathscr{O}$. If $\alpha \notin \mathscr{O}$, then $\alpha^{-1} \in \mathscr{O}$ and if we multiply the above equation by α^{1-n}, we get

$$\alpha = -c_{n-1} - c_{n-2}\alpha^{-1} - \cdots - c_0\alpha^{1-n} \in \mathscr{O}. \qquad \square$$

It can be shown that if $\mathscr{O}$ is Noetherian, it is a principal ideal domain. A ring with a unique maximal right ideal is known as a local ring, so valuation rings are local.

If $\mathscr{O}$ is a valuation ring in $\mathbb{F}$ and $\mathscr{M}$ is a principal ideal with generator t, we can define the *order* $\operatorname{ord}(x)$ of x in $\mathscr{O}$ to be the least integer k such that $t^k x$ is a unit. If we now define

$$|x| := 2^{-\operatorname{ord}(x)}$$

then you may show that this gives us a non-archimedean valuation on $\mathbb{F}$.

43.3 From Valuations to Colourings

Let $\mathscr{P}$ be a projective plane over a field $\mathbb{F}$ (for us, this will mostly be a subfield of $\mathbb{C}$). A proper 3-colouring of $\mathscr{P}$ is a surjection onto some set of three colours, such that each line gets exactly two colours. There is a more graphic definition. The *incidence graph* of $\mathscr{P}$ is the bipartite graph with the points and lines as vertices, with a point adjacent to the lines it is incident with. (A bipartite graph is the incidence graph of a projective plane if and only if its diameter is three and its girth is six. To avoid degenerate cases, we may assume that the minimum valency is at least three.) A proper 3-colouring is a graph homomorphism from the incidence graph of $\mathscr{P}$ **onto** C_6. (We may view C_6 as the incidence graph for the projective plane of order one.)

We show how to construct 3-colourings from valuations. Define two vectors to be equivalent if they span the same line. We represent the points of $\mathscr{P}$ by homogeneous coordinate vectors, that is, equivalence classes of row vectors from $\mathbb{F}^3$. (The lines can be represented by equivalence classes of column vectors). We will use the same symbol to denote a vector and its equivalence class.

Now we associate a 3-colouring to a valuation $|\cdot|$ on $\mathscr{P}$. Suppose the point p is represented by the vector

$$\begin{pmatrix} x & y & z \end{pmatrix}$$

Our colouring scheme is:

(a) If $|x| < |z|$ and $|y| < |z|$, colour p red.

(b) If $|x| \geq |z|$ and $|y| < |z|$, colour p white.

(c) If $|x| \geq |z|$ and $|y| \geq |z|$, colour p blue.

The points p_1, p_2, p_3 represented by the vectors

$$\begin{pmatrix} x_1 & y_1 & z_1 \end{pmatrix}, \quad \begin{pmatrix} x_2 & y_2 & z_2 \end{pmatrix}, \quad \begin{pmatrix} x_3 & y_3 & z_3 \end{pmatrix}$$

are collinear if and only if

$$\det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} = 0.$$

If p_1, p_2, p_3 are coloured red, white and blue respectively, then then $|z_1 x_2 y_3| \geq 1$, but the other five monomials have value less then one and, since our valuation is non-archimedean, the determinant cannot be zero.

We give a more concrete example. The *trivial valuation* v has $v(x) = 1$ if and only if $x \neq 0$. Then the assignments:

- If $x \neq 0$, then (x, y, z) is coloured white.
- If $y \neq 0$, then $(0, y, z)$ is coloured blue.
- If $z \neq 0$, then $(0, 0, z)$ is coloured red.

form a proper 3-colouring. It is the example we offered in the preamble to this chapter.

43.4 From Colourings to Valuations

Now for the heavy lifting—we construct valuations from proper 3-colourings.[5]

[5] this is the direction we will need

We declare the point $(0, 0, 1)$ to be the origin of our plane and the points with third coordinate zero are the points at infinity. Assume the origin is coloured red, that $(1, 0, 0)$ is coloured white and $(0, 1, 0)$ is coloured blue.

44
Laplacians

44.1 Basics

Let Δ be the diagonal matrix of valencies of the graph X. The *Laplacian* $L(X)$ of X is the matrix $\Delta - A$. The *unsigned Laplacian* is $\Delta + A$. We use $\lambda_1, \ldots, \lambda_n$ to denote the eigenvalues of X in increasing order, but $\lambda_1 = 0$. We write $\lambda_i(X)$ to denote the i-th largest eigenvalue of $L(X)$.

The incidence matrix of an orientation of X is the $|V(X)| \times |E(X)|$ matrix D, where the columns are the signed characteristic vectors of the edges.[1] We have

$$L(X) = DD^T$$

from which we deduce that $L \succcurlyeq 0$ and that $\mathrm{rk}(L) = \mathrm{rk}(D)$. If X has n vertices and c connected components, then $\mathrm{rk}(D) = n - c$. Any vector that is constant on a component lies in $\ker(L)$.

If S is a diagonal matrix indexed by $E(X)$ and $S_{e,e} = \pm 1$ for each edge, then

$$DS(DS)^T = DD^T = L$$

and so our choice of orientation for X is irrelevant. If X is bipartite, there is a diagonal matrix R with diagonal entries ± 1 such that RD is non-negative. Hence RD is the usual vertex-edge incidence matrix of X and $RDD^TR = \Delta + A$. Hence when X is bipartite, the unsigned Laplacian and the Laplacian are similar. On the other hand, $\mathrm{rk}(RD)$ is n less the number of bipartite components of X and so, if X is connected, $\Delta + A$ and $\Delta - A$ are similar if and only if X is bipartite.

[1] There is disagreement about which is the head and which the tail of an arc. Just be consistent with yourself

44.2 Eigenvalue Bounds

If Y is obtained from X by adding an edge, then

$$L(Y) = L(X) + L(K_2)$$

and since $L(K_2) \succcurlyeq 0$. it follows that $\lambda_r(Y) \geq \lambda_r(X)$.

The matrix $D^T D$ is a signed adjacency matrix of the line graph of X and consequently $D^T D$ and L have the same non-zero eigenvalues, with the same multiplicities.

44.2.1 Lemma. *The spectral radius of the Laplacian of X is bounded above by the spectral radius of its line graph; equality holds for a connected graph X if and only if X is bipartite.* □

Proof. Let M be the adjacency matrix of the line graph of X and set $N = D^T D$. We use $|z|$ to denote the vector such that $|z|_i = |z_i|$ and define $|N|$ similarly. Then

$$|z^T N z| \le |z|^T |N||z| = |z|^T M |z|,$$

which proves the bound.

What of equality? In this case, our inequality is based on the triangle inequality, and the triangle inequality is tight if and only if all terms in the sum have the same sign. This means we must have

$$z_e z_f N_{e,f} \ge 0$$

for all edges e and f, and therefore $N_{e,f}$ and $z_e z_f$ must have the same sign for all e and f. If S is a diagonal (± 1)-matrix such that $Sz \ge 0$, then $SNS \ge 0$. Therefore $SNS = M$ and so X is bipartite. □

The spectral radius of a graph is bounded above by its maximum valency, and therefore the spectral radius of the line graph is at most

$$\max\{d_u + d_v : u, v \in V(X),\ u \sim v\}.$$

Consequently this is an upper bound on the spectral radius of L.

A set of columns of D is linearly independent if and only if the corresponding edges form an acyclic subgraph of X. If

$$\phi(L, t) = t^n + a_1 t^{n-1} \cdots + a_{n-1} t$$

then a_k is the number of spanning forests of X with exactly k rooted components. Therefore a_1 is $(-1)^{n-1} n$ times the number of spanning trees of X. Since $(-1)^{n-1} a_{n-1}$ is the product of the non-zero eigenvalues of X, for K_n we have $|a_{n-1}| = n^{n-1}$ and from this we deduce that the number of trees on the vertex set $\{1, \ldots, n\}$ is n^{n-2} when $n \ge 2$.

44.2.2 Lemma. *The vertex connectivity of X is bounded below by $\lambda_2(X)$.*

Proof. Suppose $S \subseteq V(X)$ and $X \setminus S$ is not connected. Then $\lambda_2(X \setminus S) = 0$ and by our remarks just above, it follows that $\lambda_2(X) \le k$. □

Following Fiedler, $\lambda_2(X)$ is known as the *algebraic connectivity* of X.

44.2.3 Lemma. *If T is a tree on at least three vertices, then*

$$\lambda_2(P_n) \le \lambda_2(T) \le 1.$$

If the lower bound is tight, $T = P_n$; if the upper bound is tight, $T = K_{1,n}$. □

The upper bound follows from the fact that a tree on at least three vertices has vertex connectivity equal to 1.

We note that

$$L(K_n) = nI - J$$

and its eigenvalues are 0 and n. Any non-zero vector that sums to zero on $V(X)$ is an eigenvector for n, and so n has multiplicity $n-1$.

We have

$$L(\overline{X}) = nI - J - L(X).$$

If $\mathbf{1}^T z = 0$ and $L(X)z = \lambda z$, then $L(\overline{X})z = (n-\lambda)z$. Therefore if n has multiplicity k as an eigenvalue of $L(X)$, then 0 is an eigenvalue of $L(\overline{X})$ with multiplicity $k+1$. We conclude that $\overline{X}$ has at least $k+1$ connected components. In particular, if $|V(X)|$ is an eigenvalue of the Laplacian of X, then X is a join.

The eigenvalues of a join are the eigenvalues of the complement of $\overline{X} \cup \overline{Y}$. If $m = |V(X)|$ and $n = |V(Y)|$, this has eigenvalues 0, $m - \lambda_i$, $n - \mu_j$ and the eigenvalues of its complement are 0, $n + \lambda_i$, $m + \mu_j$. Since the complement of a disconnected graph is connected, 0 is simple.

The eigenvalues of the cone over X are 0 and $n+1$ (both simple) and $\lambda_i + 1$, where λ_i runs over the $n-1$ largest eigevenvalues if X. If X is obtained from Y by deleting a vertex, then Y is a spanning subgraph of the cone $\hat{X}$ over X, and therefore if $2 \le r \le n$,

$$\lambda_r(Y) \le \lambda_r(\hat{X}) \le \lambda_r(X) + 1.$$

44.2.1 Cut-Edges

Suppose X and Y are vertex-disjoint graphs and we construct the graph Z by joining a vertex a in X to a vertex b in Y. Then

$$\phi(L(Z)) = \phi(L(X))\phi(L(Y)) - \phi(L(X))\phi_b(L(Y)) - \phi_a(L(X))\phi(L(Y)).$$

We consider a special case of the previous identity, when $Y = K_1$. Here

$$\phi(L(Z), t) = (t-1)\phi(L(X), t) - t\phi_1(L(X), t).$$

44.3 Trees

We consider Laplacians of trees. Since trees are bipartite, the signed and unsigned Laplacians are similar, and the eigenvalues of $L(T)$ are the eigenvalues of its line graph.

The only trees that are joins are the stars, so $|V(T)|$ is a Laplacian eigenvalue of T if and only if T is a star. The complement of the star $K_{1,n}$ is $K_1 \cup K_n$, with Laplacian eigenvalues 0 (multiplicity two) and n (multiplicity $n-1$). Hence the eigenvalues of $K_{1,n}$ are 0 and n (both simple) and 1 with multiplicity $n-1$.

44.3.1 Theorem. *Let T be a tree on n vertices and let λ be an integer eigenvalue of $L(T)$. If $\lambda > 1$, then:*

(a) λ divides n.

(b) No entry of an eigenvector for λ is zero.

(c) λ is simple.

Proof. The product of the positive eigenvalues of T is n, which implies the first claim,

Assume λ is an integer eigenvalue and z is an eigenvector for λ. Suppose that $z_1 = 0$. Let k be the valency of 1 and let $S_1, \ldots, S_k$ be the connected components of $T \setminus 1$. Then the restriction of z to S_i is an eigenvector for the submatrix M_i of $L(T)$ with rows and columns indexed by $V(M_i)$; the eigenvalue is λ.

It follows that λ divides $\det(M_i)$. Reordering vertices if needed, we see that

$$M_i - e_1 e_1^T$$

is the Laplacian of S_i. If $M_i(1|1)$ is the matrix we get by deleting the first row and column from M_i, then $\det(M_i(1|1))$ is equal to the number of spanning trees of S_i. Therefore $\det(M_i(1|1)) = 1$, implying that $\lambda = 1$.

If the multiplicity m of λ is at least two, then the intersection of the eigenspace with the orthogonal complement to e_1 has dimension at least $m-1$. Therefore the λ-eigenspace contains an eigenvector that is zero on 1. □

44.4 Periodic States

We are going to abuse notation and refer to positive semidefinite matrices as density matrices, whether or not their trace is 1. A *subset state* is a diagonal matrix with diagonal entries 0 and 1. If $S \subseteq V(X)$, then D_S is diagonal 01-matrix with $(D_S)_{a,a} = 1$ if and only if $a \in S$. A state D is periodic at time τ if D and $U(\tau)$ commute; if D_S is periodic, we may say that S itself is periodic. We want to think about periodic subset-states for Laplacian walks.

Observe that D_S represents orthogonal projection onto

$$\operatorname{span}\{e_a : a \in S\}$$

and therefore if a matrix M commutes with D_S, the subspace $\operatorname{span}\{e_a : a \in S\}$ is M-invariant. It follows that if the subset $S = \{a, b\}$ is periodic, we have fractional revival on $\{a, b\}$.

Assume L has spectral decomposition

$$L = \lambda_r F_r.$$

If D is the initial state for a walk, the state $D(t)$ at time t is

$$D(t) = \sum_{r,s} e^{it(\lambda_r - \lambda_s)} E_r D E_s.$$

The *eigenvalue support* of D is the set of pairs

$$\{(\lambda_r, \lambda_s) : E_r D E_s \neq 0\}$$

(If D is pure, i.e., $D = zz^*$ for some vector z, then $E_R z z^* E_s = 0$ implies that $E_r z = 0$ or $E_s z = 0$. In this case the eigenvalue support is determined by the eigenvalues λ_r such that $E_r z \neq 0$.) Note that if $D^2 = D$, then

$$E_r D E_r = E D_r (E D_r)^T$$

and so $E_r D E_r = 0$ if and only if $E_r D = 0$.

44.4.1 Lemma. *If D is a rational matrix, its eigenvalue support is closed under the action of the Galois group.* □

44.4.2 Theorem. *All elements of the eigenvalue support of a periodic subset state are integers.*

Proof. If D and $D(t)$ are real, then $e^{it(\lambda_r - \lambda_s)}$ is real whenever $E_r D E_s \neq 0$. Hence $\sin(t(\lambda_r - \lambda_s)) = 0$ on the eigenvalue support and therefore there are integers $m_{r,s}$ such that

$$t(\lambda_r - \lambda_s) = m_{r,s}\pi.$$

This implies that if $E_r D E_s \neq 0$ and $E_k D E_\ell \neq 0$ and $\lambda_k \neq \lambda_\ell$, then

$$\frac{\lambda_r - \lambda_s}{\lambda_k - \lambda_\ell} \in \mathbb{Q}.$$

If we fix r and s and take the product of all these ratios as $(\lambda_k, \lambda_\ell)$ runs over the pairs with $E_k D E_\ell \neq 0$, we find that for some integer m,

$$\frac{(\lambda_r - \lambda_s)^m}{\prod_{(k,\ell)} (\lambda_k - \lambda_\ell)} \in \mathbb{Q}$$

By the previous lemma, the denominator here is an integer, whence $(\lambda_r - \lambda_s)^m \in \mathbb{Q}$. Now $\lambda_r - \lambda_s$ is an algebraic integer, so $(\lambda_r - \lambda_s)^m$ is an algebraic integer and, since it is rational, it is an integer. As $\lambda_k - \lambda_\ell$ is real, $m \in \{1, 2\}$.

However 0 is an eigenvalue, with eigenspace spanned by $\mathbf{1}$ and $\mathbf{1}^T D \mathbf{1} = |S| > 0$. Therefore if X is connected, $E_1 D E_r = 0$ if and only if $D E_r = 0$, and thus the eigenvalue support of D contains all pairs $(0, s)$ such that $E_s D E_s \neq 0$.

We have

$$D(t) = \sum_{r,s} e^{it(\lambda_r - \lambda_s)} F_r D F_s$$

and so if D and $D(t)$ are real, $e^{it(\lambda_r - \lambda_s)}$ is real for all r and s. As $\lambda_1 = 0$, we have that $e^{it\lambda_r}$ is real, i.e., $\sin(t\lambda_r) = 0$. Accordingly λ_r / λ_s is rational (assuming $s \neq 1$).

Assume $S = \{\sum \sigma 1 k\}$ is the eigenvalue support of the pure state. Then $\sum_r \sigma_r$ is an integer, K say, and then

$$K = \sum_r \sigma_r = \sigma_1 \sum_r \frac{\sigma_r}{\sigma_n}$$

and so σ_1 is rational and so S consists of integers. □

44.4.3 Corollary. *If the subset S is periodic under the Laplacian walk on X, it is also periodic for the Laplacian walk on $\overline{X}$.*

44.5 No Laplacian PST on Trees

[2]

Suppose the Laplacian L of X has spectral decomposition

$$L = \sum_{r=1}^{m} \lambda_r F_r,$$

where we assume that

$$0 = \lambda_1 < \cdots < \lambda_m.$$

If we have perfect state transfer from vertex a to vertex b at time τ, we also have perfect state transfer from b to a at time τ, and it follows that the subset $\{a, b\}$ is periodic (at time 2τ).

If $u \in V(X)$, then $D_u = e_u e_u^T$ is a pure state and its eigenvalue support os the set of eigenvalues λ_r such that $F_r e_u e_u^T F_r \neq 0$ (equivalently such that $F_r e_u \neq 0$). The all-ones vector is an eigenvector with eigenvalue 0; the corresponding spectral idempotent is $n^{-1}J$ and so (if X is connected, which we will be assuming), then 0 lies in the eigenvalue support of each vertex.

If we have perfect state transfer from a to b at time τ, then

$$e_b e_b^T = U(\tau) e_a e_a^T U(-\tau)$$

and, since $U(t)F_r = e^{i\lambda_r} F_r$

$$F_r e_b e_b^T F_r = F_r e_a e_a^T F_r.$$

Therefore $F_r e_b = \pm F_r e_a$. Define ϵ_r to be the ± 1-valued function on the eigenvalue support of a such that $F_r e_b = \epsilon_r F_r e_a$. If $\epsilon_r = 1$ for all r, then

$$e_b = \sum_r F_r e_b = \sum_r \epsilon_r F_r e_a = \sum_r F_r e_a = e_a.$$

We say that $\{\theta_r : \epsilon_r = 1\}$ is the *positive support* of a and that $\{\theta_r : \epsilon_r = -1\}$ is the *negative support*, we denote them by Λ^+ and Λ^- respectively. By what we have just seen, both supports are non-empty.

44.5.1 Lemma. *Assume X is connected and suppose a and b are strongly cospectral vertices in X. Then Λ^+ and Λ^- are nonempty, and if $|\Lambda^+| = 1$ or $|\Lambda^-| = 1$ then $X \cong K_2$.*

[2] Coutinho and Liu. "No Laplacian perfect state transfer on trees". arXiv:1408.2935, Chan et al. "Laplacian fractional revival on graphs". arXiv

Proof. We have

$$e_a = \sum_r E_r e_a$$

and if a and are strongly cospectral,

$$e_b = \sum_r \epsilon_r E_r e_a.$$

Since a and b are distinct, $\Lambda^- \neq \emptyset$.

From the above two equations, we have

$$e_a + e_b = \sum_{r:\theta_r \in \Lambda^+} F_r e_a, \qquad e_a - e_b = \sum_{r:\theta_r \in \Lambda^-} F_r e_a$$

If $|La^+| = 1$, then $\Lambda^+ = \{0\}$, but this implies that $e_a + e_b = 2F_1 e_a$. Since $F_1 = n^{-1}J$, we find that $n = 2$. [Remark: this works for the adjacency matrix too, since all entries of the Perron vector are positive.]

Assume $|\Lambda^-| = 1$, say $\Lambda^- = \{\lambda_k\}$. Then

$$e_b = (I - 2F_k)e_a. \tag{44.5.1}$$

As $(I - 2F_k)^2 = I$, we also have $e_a = (I - 2F_k)e_b$ and thus $I - 2F_k$ is an orthogonal matrix that swaps e_a and e_b. Therefore if we set

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

then we may assume that $I - 2F_k$ has the form

$$\begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}.$$

Here the submatrix Q is orthogonal.

$$e_a - e_b = \sum_r E_r e_a - \sum_r \epsilon_r E_r e_a = 2E_s e_a.$$

This implies that $e_a - e_b$ is an eigenvector of L with eigenvalue λ_s. Since $e_a - e_b$ is an integer vector, λ_s is an integer. As $e_a - eb$ has entries equal to zero, $\lambda_s = 1$.

Assume a and b have valency d_a and d_b respectively. Then if $a \sim b$

$$e_a^T L(e_a - e_b) = d_a + 1, \quad e_b^T L(e_a - e_b) = -1 - d_b$$

and if $\not\sim b$, then

$$e_a^T L(e_a - e_b) = d_a, \quad e_b^T L(e_a - e_b) = -d_b$$

Since $\lambda_s = 1$ we see that a and b have valency one and are not adjacent. If $u \neq a, b$, we have

$$e_u^T L(e_a - e_b) \in \{0, \pm 1\}$$

and so if $e_a - e_b$ is an eigenvector and $u \neq a, b$, then $e_u^T L(e_a - e_b) = 0$. Hence a and b have same set of neighbours.

Let g be the gcd of the eigenvalue support of a.

44.5.2 Lemma. *We have* $\epsilon_r = (-1)^{\lambda_r/g}$.

44.5.3 Theorem. *Assume* X *is a tree and* a *and* b *are strongly cospectral vertices in* X. *If* λ_r *is a non-zero element of the eigenvalue support of* a *and* $\epsilon_r = -1$, *then* λ_r *is a power of two.*

45
The 600-Cell

The 600-cell is a regular convex polytope in $\mathbb{R}^4$ with 120 vertices and 600 facets.[1] We are concerned with a paper of Fisk [2], concerning the proper colourings of its 1-skeleton. There will be many detours. The material on quaternions is based on ??.

[1] Details to come

[2]

45.1 Regular Polytopes

A *polytope* is the convex hull of a finite set of points in $\mathbb{R}^d$. We are going to assume familiarity with the basics of convexity. This is a topic where the facts that are true are usually relatively easy to prove, unfortunately there are many facts which seem obviously true that are false.

Recall that y is an *affine linear combination* of $x_1, \ldots, x_m$ if there are real scalars $a_1, \ldots, a_m$ such that

$$\sum_r a_r = 1, \qquad y = \sum_r a_r x_1.$$

Equivalently the vector

$$\begin{pmatrix} y \\ 1 \end{pmatrix}$$

(in $\mathbb{R}^{d+1}$) is a linear combination of the vectors

$$\begin{pmatrix} x_r \\ 1 \end{pmatrix}, \quad (r = 1, \ldots, m).$$

An affine subspace is a subset of $\mathbb{R}^d$ that is closed under affine linear combinations. Affine subspaces are cosets of vector subspaces and the dimension of an affine subspace is the usual dimension of the associated vector subspace. The dimension of a polytope is the minimum dimension of an affine space that contains it. Points are 0-dimensional polytopes and line segements (of finite length) are 1-dimensional polytopes.

An affine hyperplane in $\mathbb{R}^d$ is given by a vector h in $\mathbb{R}^d$ and scalar c; the hyperplane is the set

$$\{x \in \mathbb{R}^d : h^T x = c\}$$

An affine hyperplace is a *supporting hyperplane* of a polytope $\mathscr{P}$ if it contains points from $\mathscr{P}$, but does not separate two points of $\mathscr{C}$. If H is a supporting hyperplane for $\mathscr{P}$, then $H \cap \mathscr{P}$ is a convex polytope. It is a *face* of $\mathscr{P}$. If $\dim(\mathscr{P}) = d$, the faces of dimension $d-1$ are the *facets* of $\mathscr{P}$. A face of dimension 0 is a vertex, a face of dimension 1 is an edge.[3]

[3] This is the origin of the graph theory terminology

The hyperplanes H such that $H \cap \mathscr{P}$ is a facet of $\mathscr{P}$ form the vertices of the *dual polytope* of $\mathscr{P}$.

An example. Let M be an $n \times d$ matrix with distinct rows. Its rows form a set of n points in $\mathbb{R}^d$, and hence form a convex polytope $\mathscr{P}$; the vertices of $\mathscr{P}$ will be a subset of the rows of M and $\dim(\mathscr{P})$ is the rank of M. Let h be a vector in $\mathbb{R}^d$ and let c be the maximum value of an entry of Mh. Then the afine hyperplane specified by (h, c) is a supporting hyperplane the rows $e_i^T M$ such that $e_i^T Mh = c$ form a face of $\mathscr{P}$. You might amuse yourself[4] by proving that, if the rows of M all have the same length, then each row is a vertex of $\mathscr{P}$. Note that the rows $e_i^T M$ such that $e_i^T Mh$ takes its minimum value form a second face of $\mathscr{P}$. We can put all this an other way: if z lies in the column space of M, the rows $e_r^T M$ such that z_r is maximal form a face of $\mathscr{P}$.

[4] :-)

The vertices and edges of a polytope form a graph, known as its *1-skeleton*. Every 3-connected planar graph arises as the 1-skeleton of a 3-dimensional convex polytope. (The faces of the embedding on the sphere are the 2-dimensional faces of the polytope.)

The set of all non-empty proper faces of a polytope $\mathscr{P}$ is partially ordered by inclusion. If $\dim(\mathscr{P}) = e$, then any flag has size at most e. A polytope is *regular* if, given any two maximal flags $\mathscr{F}_1$ and $\mathscr{F}_2$, there is an invertible affine map[5] which fixes $\mathscr{P}$ and sends $\mathscr{F}_1$ to $\mathscr{F}_2$. We could also view the set of all faces of $\mathscr{P}$ as a simplicial complex.

[5] affine maps send z to $Az + b$

It is not hard to see that each face of a regular polytope is itself a regular polytope.

There are three infinite families of regular polytopes:

(a) Simplices: the convex hull of the standard basis vectors in $\mathbb{R}^d$ (with dimension $d-1$).

(b) Hypercubes: the convex hull of the ± 1-vectors in $\mathbb{R}^d$, this is the set of vectors z such that $\|z\|_\infty \leq 1$.

(c) Cross-polytopes: these are the duals of the hypercubes, and can be realized as the set of vectors z such that $\|z\|_1 \leq 1$; in $\mathbb{R}^3$ we have the octahedron.

There are exactly four more regular-polytopes. The icosahedron and the dodecahedron in $\mathbb{R}^3$ will be familiar. In $\mathbb{R}^4$ we have the *600-cell* (with 120 vertices and 600 facets) and its dual, the *600-cell* (with 600 vertices and 120 facets).

In the following sections we will construct the 600-cell using the quaternions.

45.2 Quaternions: The Basics

The simplest way to define the quaternions $\mathbb{H}$ is to say they are the 4-dimensional real algebra with a basis

$$1,\ i,\ j,\ k$$

subject to the relations

$$i^2 = j^2 = k^2 = -1$$

and

$$ij = k,\ jk = i,\ ki = j.$$

The difficulty with defining an algebra by generators and relations is that we need to check that it is associative. In this case it is easy enough to check that $x(yz) = (xy)z$ for all choices of x, y, z from $\{i, j, k\}$.

The subspace spanned by 1 and i is isomorphic to $\mathbb{C}$.

Note that $\mathbb{H}$ is **not** an algebra over $\mathbb{C}$

If $a \in \mathbb{H}$, we define linear maps L_a and R_a from $\mathbb{H}$ to itself by

$$L_a(z) = az, \quad R_a(z) = za.$$

The associativity of the quaternions is equivalent to the fact that L_a and R_b commute for $a, b \in \mathbb{H}$. Relative to our canonical basis, L_i, L_j and L_k are respectively represented by the matrices

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

For R_i, R_j, R_k we have the matrices

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

We will be sloppy and use L_z and R_z to denote both the linear mapping and the matrix.

Exercise: Verify that

$$L_i R_i = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Now, if $z := a + bi + cj + dk$ then

$$L_z = \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}, \qquad R_z = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix}.$$

The rows of L_z are orthogonal and therefore

$$L_z L_z^T = (a^2 + b^2 + c^2 + d^2) I$$

You may check that for x, y, z in $\{i, j, k\}$, we have $L_{xy} = L_x L_y$; from this it follows that the map $z \mapsto L_z$ is an isomorphism from $\mathbb{H}$ into a 4-dimensional subspace of $\mathrm{Mat}_{4\times 4}(\mathbb{R})$. Each matrix in this subspace is invertible.[6]

[6] the map $z \mapsto R_z$ is **not** an isomorphism, but $z \mapsto R_z^T$ is

45.3 Conjugates, Norms and Traces

The *conjugate* $\bar{z}$ is given by

$$\bar{z} := a - bi - cj - dk.$$

The restriction of conjugation to $\mathbb{C}$ is the usual complex conjugation. We have

$$L_{\bar{z}} = L_z^T,$$

one consequence of which is that

$$\overline{wz} = \bar{z}\bar{w}.$$

A quaternion z is *real* if $\bar{z} = z$.

If $z \in \mathbb{H}$ we define its *trace* $\mathrm{Tr}(z)$ to be $z + \bar{z}$ and its *norm* $N(z)$ to be $z\bar{z}$. We see that Tr is an $\mathbb{R}$-linear map from $\mathbb{H}$ to $\mathbb{R}$. Since $\mathrm{Tr}(z) = \frac{1}{2}\,\mathrm{tr}(L_z)$ we have

$$\mathrm{Tr}(wz) = \mathrm{Tr}(zw).$$

If $z = a + bi + cj + dk$, then

$$\mathrm{Tr}(z) = 2a, \quad N(z) = a^2 + b^2 + c^2 + d^2.$$

You may verify that $N(wz) = N(w)N(z)$. and that $N(z)^2 = \det(L_z)$

45.3.1 Lemma. *If $z \in \mathbb{H}$, then $z^2 - \mathrm{Tr}(z)z + N(z) = 0$.*

Proof. This immediate:

$$z^2 - \mathrm{Tr}(z)z + N(z) = z^2 - (z + \bar{z})z + z\bar{z} = 0. \qquad \square$$

We learn from this that the minimal polynomial of L_z is quadratic. The discriminant of this quadratic is

$$4a^2 - 4(a^2 + b^2 + c^2 + d^2) = -4(b^2 + c^2 + d^2)$$

and therefore the eigenvalues of L_z are $a \pm i\sqrt{b^2 + c^2 + d^2}$.

A quaternion z is *pure* if $\mathrm{Tr}(z) = 0$; if z is pure and $N(z) = 1$, then $z^2 = -1$. So the algebra generated by 1 and z is isomorphic to $\mathbb{C}$.

46
Low Rank Average Mixing

Everything on mixing matrices in this chapter is joint work with a large number of people (Zhan, Ihringer, Guo, Chan,...).

If the average mixing matrix of a connected graph X has rank one, then $|V(X)| = 2$. Are there infinitely many connected graphs where the average mixing matrix has rank two?

46.1 Walk-Regular Matrices with only Simple Eigenvalues

We say a square matrix N is *walk regular* if, for all non-negative integers k, the diagonal of N^k is constant.

46.1.1 Lemma. *Let N be a normal matrix with only simple eigenvalues. Then N is walk regular if and only all eigenvectors of N are flat.*

Proof. Assume M is normal and all its eigenvalues are simple.

Since N is normal, there is a unitary matrix Q and a diagonal matrix D such that

$$N = QDQ^*$$

If $Q_r := Qe_r$ and $\lambda_r := D_{r,r}$, then

$$N = \sum_r \lambda^r Q_r Q_r^*$$

is the spectral decomposition of N and therefore $Q_r Q_r^*$ is a spectral idempotent of N. Hence it is a polynomial in N. Its diagonal is constant if and only if Q_r is flat. □

We have no applications for this result, so we restrict our scope.

46.1.2 Corollary. *Let H be an $n \times n$ Hermitian matrix with only simple eigenvalues and let Q be the matrix of eigenvectors of H. Then H is walk regular if and only if $\sqrt{n}Q$ is a real Hadamard matrix.*

If H is Hermitian with only simple eigenvalues and walk regular, its eigenvalues must be a signed sum of entries of H.

Theorem 7.3.2 from GSQW asserts that a walk-regular graph on at least three vertices has at most $|V(X)/2|$ simple eigenvalues. [1]

[1] So K_1 and K_2 are the only walk-regular graphs with all eigenvalues simple.

46.2 Sums

We recall that

$$\widehat{M} = \sum_r E_r^{\circ 2}.$$

We will make much use of the following.

46.2.1 Lemma. *If E and F are matrices, then*

$$\mathrm{rk}(E \circ F) \leq \mathrm{rk}(E)\,\mathrm{rk}(F)$$

and

$$rk(E^{\circ 2}) \leq \binom{\mathrm{rk}(E)+1}{2}.$$

Proof. For the first inequality we merely note that $E \circ F$ is a submatrix of $E \otimes F$.

For the second inequality, suppose $f_1, \ldots, f_k$ are rows of E that form a basis for $\mathrm{row}(E)$. Since

$$(x+y) \circ (x+y) = x \circ x + y \circ y + 2x \circ y,$$

we see that $\mathrm{row}(E)$ is spanned by the k vectors $f_1, \ldots, f_k$ along with the $\binom{k}{2}$ products $f_i \circ f_j$. □

It follows that $\mathrm{rk}(E^{\circ 2}) = 1$ if and only $\mathrm{rk}(E) = 1$. We point out that if $E = zz^T$, then

$$E \circ E = (z \circ z)(z \circ z)^T.$$

46.2.2 Lemma. *If E and F are positive semidefinite matrices of the same order*

$$\ker(E+F) = \ker(E) \cap \ker(F).$$

Proof. If E and F are positive semidefinite, so is $E+F$. If $(E+F)x = 0$, then

$$0 = x^T(E+F)x = x^T Ex + x^T Fx$$

and therefore $x^T Ex = x^T Fx = 0$. Since E and F are positive semidefinite, this implies that $Ex = Fx = 0$. □

We see that if $\mathrm{rk}(\widehat{M}) = 1$, then $\mathrm{rk}(E_r) = 1$ and $E_r^{\circ 2} = E_s^{\circ 2}$ for all r and s. If $|V(X)| = n$, this implies that X has n simple eigenvalues. If $E_r^{\circ 2} = E_s^{\circ 2}$, then

$$I \circ E_r^{\circ 2} = I \circ E_s^{\circ 2}$$

and accordingly $I \circ E_r = I \circ E_s$. Since $\sum E_r = I$, we see that

$$I \circ E_r = \frac{1}{n} I.$$

Therefore X is walk regular, and a walk-regular graph with only simple eigenvalues has at most two vertices.

46.3 Rank Two

46.3.1 Lemma. *If E is positive semidefinite and* $\mathrm{rk}(E) = 2$*, then* $\mathrm{rk}(E \circ E) \geq 2$*.*

Proof. We may assume that $E = xx^T + yy^T$ and so

$$\begin{aligned} E \circ E &= (xx^T) \circ (xx^T) + (yy^T) \circ (yy^T) + 2(xx^T) \circ (yy^T) \\ &= (x \circ x)(x \circ x)^T + (y \circ y)(y \circ y)^T + 2(x \circ y)(x \circ y)^T \\ &= \begin{pmatrix} x \circ x & y \circ y & \sqrt{2}x \circ y \end{pmatrix} \begin{pmatrix} x \circ x & y \circ y & \sqrt{2}x \circ y \end{pmatrix}^T. \end{aligned}$$

If we define

$$W = \begin{pmatrix} x \circ x & y \circ y & \sqrt{2}x \circ y \end{pmatrix}$$

then $\mathrm{rk}(E^{\circ 2}) = \mathrm{rk}(W)$.

Suppose $\mathrm{rk}(W) = 1$. Then for some scalars β and γ we have $y \circ y = \beta x \circ x$ and $x \circ y = \gamma x \circ x$. The second condition implies that x_i and y_i have the same sign and, given this, the first condition yields that x and y are parallel. Therefore $\mathrm{rk}(E) = 1$, a contradiction. □

Note that if H is an $n \times n$ Hadamard matrix, then $\mathrm{rk}(H) = n$ and $\mathrm{rk}(H^{\circ 2}) = 1$. So the condition that $E \succcurlyeq 0$ is needed.

Assume $\mathrm{rk}(W) = 2$. There are two possibilities: either

(a) All eigenvalues of X are simple, or

(b) tThere is an eigenvalue of multiplicity two.

Bibliography

Ada Chan and Chris Godsil. Type-ii matrices and combinatorial structures. *Combinatorica*, 30(1):1–24, 2010.

Wei Wang, Feng Li, Hongliang Lu, and Zongben Xu. Graphs determined by their generalized characteristic polynomials. *Linear algebra and its applications*, 434(5):1378–1387, 2011.

Index