Association Schemes[1]

Chris Godsil
Combinatorics & Optimization
University of Waterloo

[1] June 3, 2010

Preface

These notes provide an introduction to association schemes, along with some related algebra. Their form and content has benefited from discussions with Bill Martin and Ada Chan.

Contents

Chapter 1

Schemes and Algebras

Our first three chapters provide an introduction to the basic theory of association schemes and to some of their applications. In this chapter we introduce association schemes and describe their structure.

1.1 Definitions and Examples

We try to motivate the definitions to come. Suppose X is a graph with vertex set V and diameter d. For $i = 1, \ldots, d$ we define X_i to be the graph with vertex set V, where two vertices are adjacent in X_i if they are at distance i in X. (So $X = X_1$.) Let A_i denote the adjacency matrix of X_i, set A_0 equal to I and consider the matrix algebra $\mathbb{C}[\mathscr{A}]$ over $\mathbb{C}$ generated by $A_1, \ldots, A_d$.

If we identify the automorphism group of X with the set of permutation matrices that commute with A_1, then each automorphism of X lies in the commutant of $\mathbb{C}[\mathscr{A}]$. Thus, for example, if $\mathbb{C}[\mathscr{A}] = \mathrm{Mat}_{n\times n}(\mathbb{C})$, then the automorphism group of X must be the identity group. Since the matrices $A_0, \ldots, A_d$ are linearly independent, $\mathbb{C}[\mathscr{A}]$ has dimension at least $d+1$. This suggests that the case where $\dim(\mathbb{C}[\mathscr{A}]) = d+1$ should be interesting. In fact the dimension of $\mathbb{C}[\mathscr{A}]$ is $d+1$ if and only if the matrices $A_0, \ldots, A_d$ form an association scheme.

An *association scheme* with d classes is a set $\mathscr{A} = \{A_0, \ldots, A_d\}$ of 01-matrices such that

(a) $A_0 = I$.

(b) $\sum_{i=0}^{d} A_i = J$.

(c) $A_i^T \in \mathscr{A}$ for each i.

(d) $A_i A_j = A_j A_i \in \mathrm{span}(\mathscr{A})$.

Note that (b) implies that the matrices $A_0, \ldots, A_d$ are linearly independent, and (d) that the algebra they generate has dimension $d+1$. Since J is the sum of the A_i, it commutes with each A_i, which implies that all rows and columns of A_i have the same sum.

An association scheme is *symmetric* if each matrix in it is symmetric. We view $A_1, \ldots, A_d$ as adjacency matrices of directed graphs $X_1, \ldots, X_d$, with common vertex set V. We say two vertices u and v are *i-related* if uv is an arc in X_i.

1.1.1 Example. The *Johnson scheme* $J(v,k)$. The vertex set of this scheme is the set of all k-subsets of a fixed set of v elements. Two vertices α and β are i-related if $|\alpha \cap \beta| = k - i$. This scheme has k classes.

1.1.2 Example. The *Grassman scheme* $J_q(v,k)$. The vertex set is the set of all subspaces of dimension k of the vector space of dimension n over $GF(q)$. Subspaces α and β are i-related if $\dim(\alpha \cap \beta) = k - i$. This scheme has k classes.

1.1.3 Example. The *Hamming scheme* $H(n,q)$. Let Q be an alphabet of q symbols. The vertex set of $H(n,q)$ is Q^n, the set of all words of length n over Q. Two words are i-related if they differ in exactly i coordinate positions. This scheme has n classes.

1.1.4 Example. The *bilinear forms scheme*. The vertices are the $m \times n$ matrices over the field of q elements. Two matrices A and B are i-related if $\mathrm{rk}(A-B) = i$. The number of classes in this scheme is the minimum of m and n.

1.1.5 Example. The conjugacy classes of a finite group Γ. Let the congugacy classes of Γ be $C_0, \ldots, C_d$, where $C_0 = \{1\}$. The vertex set of this scheme consists of the elements of Γ, and two group elements g and h are i-related if $hg^{-1} \in C_i$. This is our first example of a scheme that is not symmetric.

1.1.6 Example. Let Z be the complete graph on n^2 vertices. A *parallel class* in Z is a subgraph isomorphic to nK_n. Two parallel classes are orthogonal is they are edge-disjoint. A *partial spread* is a set of pairwise orthogonal parallel classes $C_1, \ldots, C_t$. Define A_i to be the adjacency matrix of the i-parallel class, set $A_0 = I$ as usual and define

$$A_{t+1} = J - \sum_{i=1}^{t} A_i.$$

Then $A_0, \ldots, A_{t+1}$ is a symmetric association scheme. (These schemes correspond to orthogonal arrays with index 1.)

1.2 Strongly Regular Graphs

The simplest association schemes are the schemes with one class. In this case we have $A_0 = I$ and $A_1 = J - I$; the directed graph X_1 is the complete graph itself. We cannot think of anything intelligent to say about this situation, so we turn to the next simplest case. These are the symmetric schemes with two classes, and are equivalent to strongly-regular graphs.

Rather than offer the necessary definitions, we consider a classical example. We consider graphs with diameter two and maximum degree k. If X is such a graph and $u \in V(X)$, then u has at most k neighbours, and at most $k(k-1)$ vertices lie at distance two from u. Therefore

$$|V(X)| \leq 1 + k + k^2 - k = k^2 + 1.$$

If equality holds then X is k-regular and its girth is at least five. This leads us to study k-regular graphs on k^2+1 vertices with diameter two. Suppose X is such a graph and let A be its adjacency matrix.

We claim that

$$A^2 + A - (k-1)I = J. \tag{1.2.1}$$

This is an easy consequence of the fact that the ij-entry of A^2 is the number of walks of length two from i to j in X. The number of walks of length two that start and finish at the same vertex is the valency of the vertex, and therefore since X is regular, $(A^2)_{i,i} = k$. The number of walks of length two that start at a given vertex i and end at the adjacent vertex j is the number of triangles in X that contain the edge ij. Therefore $(A^2)_{i,j} = 0$ in this case. Finally if i and j are distinct and not adjacent in X then, since there are no 4-cycles in X and since the diameter of X is two, $(A^2)_{i,j} = 1$. Equation (1.2.1) follows from these facts.

We explain the connection with association schemes. The adjacency matrix $\overline{A}$ of the complement $\overline{X}$ of X is $J - I - A$. From (1.2.1)

$$\overline{A} = J - I - A = A^2 - kI.$$

Since $\overline{A}$ is thus a polynomial in A, it commutes with A. We also see that A^2 is a linear combination of I, A and $\overline{A}$. Since $AJ = JA = kJ$ we can also show that $A\overline{A}$

and $\overline{A}^2$ are linear combinations of I and $\overline{A}$. We conclude that the matrices I, A and $\overline{A}$ form a symmetric association scheme with two classes.

We can use (1.2.1) to obtain more information about our graphs. They key is that we can compute the eigenvalues of A.

First note that the all-ones vector $\mathbf{1}$ is an eigenvector for A; in fact

$$A\mathbf{1} = k\mathbf{1}$$

and so the corresponding eigenvalue of the valency k. Suppose λ is an eigenvalue of A with eigenvector z. We may assume that z is orthogonal to $\mathbf{1}$, whence $Jz = 0$. Therefore

$$0 = Jz = (A^2 + A - (k-1)I)z = (\lambda^2 + \lambda - k + 1)z$$

and so λ is a zero of the quadratic polynomial

$$t^2 + t - k + 1.$$

Denote the roots of this by θ and τ. Since $\theta\tau = 1 - k$ we may assume that $\theta > 0 > \tau$. Let m_θ and m_τ denote respectively the multiplicity of θ and τ as an eigenvalue of A. Since X has $k^2 + 1$ vertices and k is an eigenvalue with multiplicity at least one, we have

$$1 + m_\theta + m_\tau = k^2 + 1. \tag{1.2.2}$$

Also $\mathrm{tr}(A) = 0$ and consequently

$$k + m_\theta\theta + m_\tau\tau = 0. \tag{1.2.3}$$

These two equations imply that

$$m_\tau = \frac{\theta k^2 + k}{\theta - \tau} \tag{1.2.4}$$

The existence of this expression for the multiplicity of an eigenvalue is a consequence of the fact that we are dealing with an association scheme. The fact that its right side must be an integer provides a very useful constraint. The ensuing calculations show how we may put it to work.

We distinguish two cases. First, suppose that θ and τ are irrational. We have

$$0 = k + (m_\theta - m_\tau)\theta + m_\tau(\theta + \tau) = k - m_\tau + (m_\theta - m_\tau)\theta$$

and since $k - m_\tau$ is an integer and θ is irrational, it follows that $m_\theta - m_\tau = 0$. Then (1.2.3) yields that $k = m_\theta = m_\tau$ and so (1.2.2) now yields that $k^2 - 2k = 0$. The only useful solution to this is $k = 2$, when we see that $X = C_5$.

Thus we may assume that θ and τ are rational, and hence they are integers. Since θ and τ are the roots of $t^2 + t - k + 1$, we have

$$(\theta - \tau)^2 = 1 + 4(k-1) = 4k - 3$$

and therefore $4k - 3$ must be a perfect square. Since $4k - 3$ is odd, we may assume

$$4k - 3 = (2s+1)^2$$

and therefore

$$k = s^2 + s + 1.$$

From this it follows that $\theta = s$ and $\tau = -s - 1$ and consequently

$$m_\tau = \frac{(s^2+s+1)(s(s^2+s+1)+1)}{2s+1}$$

Now

$$4s^2 + 4s + 4 = (2s+1)^2 + 3$$

and

$$8s^3 + 8s^2 + 8s + 8 = 2s(2s+1)^2 + 3(2s+1) + 5 = (4s^2+4s+3)(2s+1) + 5.$$

Hence there is a polynomial p with integer coefficients such that

$$32m_\tau = p(s) + \frac{15}{2s+1}.$$

We conclude that m_τ is an integer if and only if $2s+1$ divides 15. This implies that

$$s \in \{1, 2, 7\}$$

and so

$$k \in \{3, 7, 57\}.$$

To summarise, we have shown that if there is a k-regular graph of diameter two on $k^2 + 1$ vertices, then k is 2, 3, 7 or 57 (and v is 5, 10, 50 or 3250). The case $k = 2$ is realized by C_5. The case $k = 3$ is realized by the Petersen graph and the case $k = 7$ by the famous Hoffman-Singleton graph. We do not know if there is a graph with valency 57. This is an old and famous open question.

1.3 The Bose-Mesner Algebra

The *Bose-Mesner algebra* of an association scheme $\mathscr{A} = \{A_0, \ldots, A_d\}$ is the algebra generated by the matrices $A_0, \ldots, A_d$; equivalently it is the complex span of these matrices. There is a second multiplication on the Bose-Mesner algebra which will prove to be very important. We define the Schur product $A \circ B$ of two matrices of the same order by

$$(A \circ B)_{i,j} := A_{i,j} B_{i,j}.$$

This is a commutative and associative product with J as unit. Since the set $\mathscr{A} \cup 0$ spans the Bose-Mesner algebra, and since this set is closed under Schur multiplication, it follows that the Bose-Mesner algebra is closed under Schur multiplication. Hence it is an algebra with respect to Schur multiplication. The Bose-Mesner algebra is also closed under complex conjugation and the transpose map.

A *coherent algebra* is a matrix algebra over $\mathbb{C}$ that is Schur-closed, closed under transpose and complex conjugation, and contains I and J. Any Bose-Mesner algebra is a commutative coherent algebra. We will discuss coherent algebras at greater length in Chapter **??**, but we offer some simple observations now.

1.3.1 Lemma. *A commutative coherent algebra is the Bose-Mesner algebra of an association scheme.* □

Define the *commutant* of a set of matrices to be the set of all matrices that commute with each element of the set.

1.3.2 Lemma. *The commutant of a set of $v \times v$ permutation matrices is a coherent algebra.*

Proof. It suffices to show that the commutant of a single permutation matrix P is a coherent algebra. The key point is then to show that the commutant of P is Schur-closed.

Suppose M and N commute with P. Then

$$P(M \circ N) = (PM) \circ (PN) = (MP) \circ (NP) = (M \circ N)P$$

and therefore the commutant of P is Schur-closed. □

A permutation group Γ on a set V is *generously transitive* if, for each pair of points u and v in V, there is an element γ of Γ such that

$$u\gamma = v, \quad v\gamma = u.$$

Clearly a generously transitive permutation group is transitive.

1.3.3 Lemma. *The commutant of a permutation group is the Bose-Mesner algebra of a symmetric association scheme if and only if the group is generously transitive.*

Proof. Let Γ be a permutation group on V. The commutant of Γ is a coherent algebra, so we need only decide when it is commutative. We note Γ acts as a group of permutations of $V \times V$, and the orbits of Γ form a partition of this set. Each orbit is a directed graph, and the adjacency matrices of the orbits form a basis for the commutant of Γ.

The set

$$\{(v, v) : v \in V\},$$

known as the diagonal of $V \times V$, is a union of orbits of Γ, and is a single orbit if and only if Γ is transitive. Suppose u and v are distinct. Then uv and vu lie in the same orbit if and only if there is an element of Γ that swaps u and v.

Hence if Γ is transitive, then it is generously transitive if and only if all matrices in the commutant of Γ are symmetric. Since the product of two symmetric matrices A and B is symmetric if and only if $AB = BA$, the lemma follows. □

This lemma can be used to verify that the schemes $J(v, k)$, $J_q(v, k)$, $H(n, q)$ and $\mathrm{Mat}_{m\times n}(\mathbb{F})$ are symmetric, with the stated number of classes.

1.4 Idempotents

Let $\mathbb{C}[\mathscr{A}]$ be the Bose-Mesner algebra of the association scheme

$$\mathscr{A} = \{A_0, \ldots, A_d\}.$$

The matrices $A_0, \ldots, A_d$ form a basis, each element of which is a Schur idempotent. In this section we identify a second basis, consisting of matrix idempotents.

Two idempotents E and F are *orthogonal* if $EF = 0$. For example, if E is an idempotent, then E and $I - E$ are orthogonal idempotents. We define a partial

ordering on the idempotents of a commutative algebra $\mathbb{C}[\mathscr{A}]$. Suppose E and F are idempotents in $\mathbb{C}[\mathscr{A}]$. We write $E \leq F$ if $FE = E$. This relation is reflexive, antisymmetric and transitive; therefore it is a partial order. A *minimal idempotent* is a minimal element of the set of non-zero idempotents. If E and F are idempotents, then $EF \leq E, F$; it follows that if E and F are minimal, then they are orthogonal.

1.4.1 Theorem. *Let $\mathscr{B}$ be a commutative matrix algebra with identity over an algebraically closed field. Assume that if $N \in \mathscr{B}$ and $N^2 = 0$, then $N = 0$. Then $\mathscr{B}$ has a basis of pairwise orthogonal idempotents.*

Proof. As a first step, we show that each element of $\mathscr{B}$ is a linear combination of idempotents.

Suppose $A \in \mathscr{B}$. Let $\psi(t)$ be the minimal polynomial of A and assume that

$$\psi(t) = \prod_{i=1}^{k} (t - \theta_i)^{m_i}.$$

If

$$\psi_i(t) := \frac{\psi(t)}{(t-\theta_i)^{m_i}},$$

then the polynomials $\psi_1, \ldots, \psi_k$ are coprime, and therefore there are polynomials $f_1(t), \ldots, f_k(t)$ such that

$$1 = \sum_i f_i(t)\psi_i(t).$$

Therefore

$$I = \sum_i f_i(A)\psi_i(A). \tag{1.4.1}$$

If $i \neq j$, then $\psi_i(A)\psi_j(A) = 0$ because ψ divides $\psi_i\psi_j$. Hence if we multiply both sides of (1.4.1) by $f_i(A)\psi_i(A)$, we find that

$$f_i(A)\psi_i(A) = (f_i(A)\psi_i(A))^2.$$

Thus $f_i(A)\psi_i(A)$ is an idempotent, which we denote by E_i. We note that $E_iE_j = 0$ if $i \neq j$. Since ψ divides $(t-\theta_i)^{m_i}\psi_i(t)$, we have

$$(A - \theta_i I)^{m_i} E_i = 0.$$

Consequently

$$[(A - \theta_i I)E_i]^{m_i} = 0,$$

and, given our hypothesis, it follows that $(A-\theta_i I)E_i = 0$. We may rewrite (1.4.1) as

$$I = E_1 + \cdots + E_k$$

and so

$$A = AE_1 + \cdots + AE_k = \theta_1 E_1 + \cdots + \theta_k E_k.$$

This expresses A as a linear combination of idempotents.

We have shown that $\mathscr{B}$ is spanned by idempotents. The essential problem that remains is to show that minimal idempotents exist. Suppose E and F are distinct idempotents and $E \leq F$. Then

$$F(I-E) = F - E \neq 0$$

but $E(I-E) = 0$. Hence the column space of E must be a proper subspace of the column space of F. Therefore if $E_1, \ldots, E_m$ are distinct idempotents and

$$E_1 \leq \cdots \leq E_m$$

then $m \leq n+1$. We conclude that minimal idempotents exist.

Now we prove that each idempotent is a sum of minimal idempotents. Suppose F is an idempotent and E is a minimal idempotent. If $EF \neq 0$, then $EF \leq E$ and therefore $EF = E$. This also shows that distinct minimal idempotents are orthogonal. Let F_0 be the sum of the distinct minimal idempotents E such that $E \leq F$. Then F_0 is an idempotent. If $F_0 \neq F$ then $F - F_0$ is an idempotent and so there is a minimal idempotent below it, which contradicts our choice of F_0. We conclude that $\mathscr{B}$ is spanned by minimal idempotents. □

Suppose $\mathscr{B}$ is a Schur-closed algebra that contains J over some field. Then 1.4.1 implies that $\mathscr{B}$ has a basis of 01-matrices. Of course this can be proved more directly (and with less effort).

A matrix N is nilpotent if $N^k = 0$ for some k. Theorem 1.4.1 asserts that a commutative matrix algebra with identity has a basis of orthogonal idempotents if there are no non-zero nilpotent matrices in it. Since a non-zero linear combination of pairwise orthogonal idempotents cannot be nilpotent, this condition is necessary too. A commutative algebra is *semisimple* if it contains no non-zero nilpotent elements.

1.5 Idempotents for Association Schemes

We will apply the theory of the last section to Bose-Mesner algebras.

1.5.1 Theorem. *Suppose $\mathscr{B}$ is a commutative subalgebra of* $\mathrm{Mat}_{v\times v}(\mathbb{C})$ *that is closed under conjugate transpose and contains the identity. Then $\mathscr{B}$ has a basis of matrix idempotents $E_0,\dots,E_d$ such that*

(a) $E_iE_j = \delta_{i,j}E_i$.

(b) The columns of E_i are eigenvectors for each matrix in $\mathbb{C}[\mathscr{A}]$.

(c) $\sum_{i=0}^{d} E_i = I$.

(d) $E_i^* = E_i$. □

Proof. Suppose $N \in \mathbb{C}[\mathscr{A}]$ and $N^2 = 0$. Then

$$0 = (N^*)^2N^2 = (N^*N)^2$$

and hence

$$0 = \mathrm{tr}((N^*N)^2) = \mathrm{tr}((N^*N)^*(N^*N)).$$

If $H := N^*N$, then $\mathrm{tr}(H^*H) = 0$ if and only if $H = 0$, so we deduce that $N^*N = 0$. But then $\mathrm{tr}(N^*N) = 0$ and therefore $N = 0$. Hence $\mathbb{C}[\mathscr{A}]$ satisfies the hypotheses of Theorem 1.4.1, and so it has a basis of pairwise orthogonal idempotents, which we denote by $E_0,\dots,E_d$. Thus (a) is proved.

If $A \in \mathbb{C}[\mathscr{A}]$, then

$$A = \sum_i a_iE_i$$

for suitable scalars a_i. Since the idempotents E_i are orthogonal,

$$AE_r = a_rE_r.$$

This shows that the columns of E_r are eigenvectors for A, and the scalars a_i are eigenvalues of A. So (c) is proved.

Since $I \in \mathbb{C}[\mathscr{A}]$, it is a linear combination of $E_0,\dots,E_d$:

$$I = \sum_i a_iE_i.$$

Since the scalars a_i are eigenvalues for I, they must all equal 1. Hence (d) holds.

Finally we show that the idempotents E_i are Hermitian. Since $\mathbb{C}[\mathscr{A}]$ is closed under transpose and complex conjugation, $E_i^* \in \mathbb{C}[\mathscr{A}]$. Therefore there are scalars $a_0,\dots,a_d$ such that

$$E_i^* = \sum_j a_jE_j$$

and so

$$E_i^* E_i = f_i E_i.$$

Since $\mathrm{tr}(E_i^* E_i) > 0$ and $\mathrm{tr}(E_j) > 0$, it follows that $f_i \neq 0$. But E_i^* is a minimal idempotent, and therefore $f_j = 0$ if $j \neq i$. This implies that E_i^* is a scalar multiple of E_i, but $\mathrm{tr}(E_i) = \mathrm{tr}(E_i^*)$, and therefore $E_i^* = E_i$. □

This theorem applies immediately to the Bose-Mesner algebra of an association scheme. In this case $\frac{1}{v}J \in \mathscr{B}$; since this is an idempotent with rank one, it must be minimal and therefore it is equal to one of the idempotents E_i. It is conventional to assume it is E_0.

If A_i is Schur idempotent in $\mathscr{A}$, so is A_i^T. If E_j is a matrix idempotent, so is E_j^T (which is equal to $\overline{E}_j$). We adopt the useful convention that

$$A_{i'} := A_i^T$$

and

$$E_{j'} := E_j^T = \overline{E}_j.$$

Note that $v_{i'} = v_i$ and $m_{j'} = m_j$.

To give a better idea of the power of 1.4.1, we use it to derive one of the basic results in linear algebra. A complex matrix A is *normal* if $AA^* = A^*A$. We adopt the convention that the algebra generated by a set of matrices always contains the identity.

1.5.2 Theorem. *If A is normal, then A is unitarily similar to a diagonal matrix.*

Proof. The algebra generated by A and A^* is commutative and closed under conjugate-transpose. Hence it has a basis of orthogonal idempotents $F_1, \ldots, F_d$. Since each F_i is Hermitian, the condition $F_i F_j = 0$ implies the column spaces of F_i and F_i are orthogonal. If follows that there is an orthogonal basis of eigenvectors of A. □

Notes

There a number of useful references for association schemes. Bannai and Ito [**?**], is the oldest of these, but carries its age well. It views the subject from a group theoretic viewpoint. Bailey's book [**?**] is more recent and views association schemes from the viewpoint of design theory. Since this is the origin of the

subject, this is a very natural approach. We note that Bailey restricts herself to what we call symmetric association schemes; for design theory this is very natural. However it excludes the association schemes arising from the conjugacy classes of a finite group and as the only real cost in allowing non-symmetric schemes is the use of $\mathbb{C}$ rather than $\mathbb{R}$, and we have happily chosen to pay it.

Brouwer, Cohen and Neumaier's book on distance-regular graphs [**?**] offers a lot of information on association schemes. Zieschang [**?**] allows his association schemes to be infinite and/or non-commutative. For an algebraist this can be very interesting, but the resulting theory does not seem to have much contact with the combinatorial questions that we are interested in.

The classic source of information on association schemes (in the sense we use the term) is Delsarte's thesis [**?**]. A copy of this is available online at http://users.wpi.edu/~martin/RESEARCH/philips.pdf. One of Delsarte's main contributions was to demonstrate that the theory of association schemes provides an extremely useful framework for work in coding theory.

Chapter 2

Parameters

To each association scheme there are four associated families of parameters: the eigenvalues,the dual eigenvalues, the intersection numbers and the Krein parameters. We introduce these and present a few of their applications. We will see that the algebraic structure of an association scheme is entirely determined by its eigenvalues.

2.1 Eigenvalues

There are scalars $p_i(j)$ such that

$$A_i = \sum_{r=0}^{d} p_i(r)E_r, \qquad (i = 0,\ldots,d) \tag{2.1.1}$$

and scalars $q_i(j)$ such that

$$E_j = \frac{1}{v}\sum_{r=0}^{d} q_j(r)A_r. \qquad (j = 0,\ldots,d) \tag{2.1.2}$$

The scalars $p_i(j)$ are called the *eigenvalues* of the scheme. Since they are eigenvalues of the 01-matrices A_i, they are algebraic integers. Note that

$$A_i J = p_i(0)J$$

and therefore $p_i(0)$ is equal to the common value of the row sums of A_i. We define

$$v_i := p_i(0),$$

call $v_0, \ldots, v_d$ the *valencies* of the scheme. Because $I = \sum_i E_i$, we also have $p_0(i) = 1$ for each i.

The eigenvalues of A_i^T are the numbers $\overline{p_i(j)}$, for $i = 0, 1, \ldots, d$.

The scalars $q_i(j)$ are the *dual eigenvalues* of the scheme. Since

$$E_0 = \frac{1}{v} \sum_i A_i,$$

we have $q_0(i) = 1$. The columns of E_i are eigenvectors for each matrix in $\mathbb{C}[\mathscr{A}]$, and so its column space is an eigenspace for $\mathbb{C}[\mathscr{A}]$. The dimension of this eigenspace is the rank of E_i. Since E_i is an idempotent, all its eigenvalues are equal to 1 and

$$\mathrm{rk}(E_i) = \mathrm{tr}(E_i).$$

The quantities $\mathrm{tr}(E_i)$ are the *multiplicities* of the scheme. From refEA we have

$$\mathrm{tr}(E_i) = \frac{1}{v} \sum_i q_i(r) \, \mathrm{tr}(A_r).$$

Now $\mathrm{tr}(A_r) = 0$ if $r \neq 0$ and $\mathrm{tr}(A_0) = v$, so we find that

$$\mathrm{tr}(E_i) = q_i(0).$$

We use m_i to denote $\mathrm{tr}(E_i)$.

The *eigenmatrix* of $\mathbb{C}[\mathscr{A}]$ is the $(d+1) \times (d+1)$ matrix P given by

$$P_{i,j} = p_j(i).$$

The dual eigenmatrix Q is the $(d+1) \times (d+1)$ matrix Q given by

$$Q_{i,j} = q_j(i).$$

From 2.1.1 and 2.1.2, we have

$$PQ = vI.$$

One consequence of this is that the dual eigenvalues of $\mathbb{C}[\mathscr{A}]$ are determined by the eigenvalues. As we proceed we will see that much of the structure of an association scheme is determined by its eigenmatrix.

2.2 Strongly Regular Graphs

A graph X is *strongly regular* if it is neither complete nor empty and there are integers k, a and c such that:

(a) X is regular with valency k.

(b) Any two adjacent vertices have exactly a common neigbours.

(c) Any two distinct non-adjacent vertices have exactly c common neighbours.

If A is the adjacency matrix of X, these conditions are equivalent to the two matrix equations

$$AJ = kJ, \qquad A^2 = kI + aA + c(J - I - A).$$

It is usually better to write the second of these as

$$A^2 - (a-c)A - (k-c)I = cJ.$$

A strongly regular graph on v vertices with parameters k, a and c as above is called a $(v, k; a, c)$ strongly regular graph.

It is straightforward to use the above matrix equations to show that if A is the adjacency matrix of a strongly regular graph, then

$$I, A, J - I - A$$

form an association scheme with two classes. Conversely, any association scheme with two classes arises from a strongly regular graph.

Suppose A_1 is the adjacency matrix of a strongly regular graph X and $\mathscr{A}$ is the corresponding association scheme, with matrix idempotents E_0, E_1 and E_2. If X is k-regular, then

$$A_0 = E_0 + E_1 + E_2, \quad A_1 = kE_0 + \theta E_1 + \tau E_2.$$

This equations determine two columns of the eigenmatrix P. Since $A_2 = J - I - A_1$, we also have

$$A_2 = (v - 1 - k)E_0 - (\theta + 1)E_1 - (\tau + 1)E_2.$$

Therefore

$$P = \begin{pmatrix} 1 & k & v-1-k \\ 1 & \theta & -\theta - 1 \\ 1 & \tau & -\tau - 1 \end{pmatrix}$$

from which we compute that

$$Q = \frac{1}{\theta-\tau}\begin{pmatrix} \theta-\tau & -k-(v-1)\tau & k+(v-1)\theta \\ \theta-\tau & v-k+\tau & k-v-\theta \\ \theta-\tau & \tau-k & k-\theta \end{pmatrix}$$

The entries in the first row of Q give the multiplicities of the eigenvalues of the graph. One consequence of this is that the ratio

$$\frac{\theta(v-1)+k}{\theta-\tau}$$

must be an integer. Constraints of this form play a major role in the theory of distance-regular graphs.

2.3 Intersection Numbers

Suppose $\mathscr{A}$ is a scheme with d classes. Since $\mathbb{C}[\mathscr{A}]$ is closed under multiplication, there are constants $p_{i,j}(k)$ such that

$$A_i A_j = \sum_{k=0}^{d} p_{i,j}(k) A_k.$$

We call these the *intersection numbers* of the scheme. We see that

$$p_{i,j}(k) A_k = A_k \circ (A_i A_j),$$

from which it follows that the intersection numbers are non-negative integers. We see also that

$$p_{i,j}(k) = \frac{\mathrm{sum}(A_k \circ (A_i A_j))}{v v_k} = \frac{\mathrm{tr}(A_k^T A_i A_j)}{v v_k}. \tag{2.3.1}$$

We define the *intersection matrices* $B_0, \ldots, B_d$ by

$$(B_i)_{j,k} := p_{i,j}(k).$$

If π denotes the relation partition of $V(\mathscr{A})$ with respect to v, then $B_i = A/\pi$. Hence the matrices $B_0, \ldots, B_d$ generate a commutative algebra of $(d+1)\times(d+1)$ matrices which is isomorphic to $\mathbb{C}[\mathscr{A}]$ as an algebra. (However it is not Schur-closed in general.)

The intersection numbers are determined by the eigenvalues of the scheme. The eigenvalue of $A_k^T A_i A_j$ on the column space of E_ℓ is

$$p_i(\ell)p_j(\ell)\overline{p_k(\ell)}$$

whence 2.3.1 implies that

$$p_{i,j}(k) = \frac{1}{v v_k}\sum_{\ell=0}^{d} m_\ell p_i(\ell)p_j(\ell)\overline{p_k}(\ell).$$

Let $X_1,\ldots,X_d$ be the graphs of an association scheme. If X_i has diameter s then the matrices

$$A_i^0,\ldots,A_i^s$$

are linearly independent. (It might be easier to see that the first $s+1$ powers of A_i+I are linearly independent.) Therefore the diameter of X_i is bounded above by d, the number of classes of the scheme.

An association scheme with d classes is *metric* with respect to the i-th relation if the diameter of X_i is d. If the scheme is metric with respect to the i-th relation, then X_i is said to be a *distance-regular graph*. The Johnson scheme, the Grassman scheme, the Hamming scheme and the bilinear forms scheme are all metric with respect to their first relation. A primitive strongly regular graph is primitive with respect to each non-identity relation. An association scheme may be metric with respect to more than one relation. The standard example is the Johnson scheme $J(2k+1,k)$, which is metric with respect to A_1 and A_k.

If $\mathscr{A}$ is metric with respect to A_i and $s \le d$, then $(I+A_i)^s$ is a linear combination of exactly $s+1$ distinct Schur idempotents. It is customary to assume $i=1$, and to order the Schur idempotents so that $(I+A_1)^s$ is a linear combination of $A_0,\ldots,A_s$. With this convention, the intersection matrix B_1 is tridiagonal.

2.4 Krein Parameters

We consider the parameters dual to the intersection numbers. Let $\mathscr{A}$ be a scheme on v vertices with d classes. Then there are constants $q_{i,j}(k)$ such that

$$E_i \circ E_j = \frac{1}{v}\sum_{k=0}^{d} q_{i,j}(k)E_k. \tag{2.4.1}$$

We call these constants the *Krein parameters* of the scheme. We have

$$q_{i,j}(k)E_k = vE_k(E_i \circ E_j)$$

and therefore

$$q_{i,j}(k) = v\frac{\mathrm{tr}(E_k(E_i \circ E_j))}{m_k} = v\frac{\mathrm{sum}(\overline{E}_k \circ E_i \circ E_j)}{m_k}$$

Now

$$\overline{E}_k \circ E_i \circ E_j = \frac{1}{v^3}\sum_{\ell=0}^{d} q_i(\ell)q_j(\ell)\overline{q_k(\ell)}A_\ell$$

which yields

$$q_{i,j}(k) = \frac{1}{vm_k}\sum_{\ell=0}^{d} q_i(\ell)q_j(\ell)\overline{q_k(\ell)}v_\ell = \frac{m_i m_j}{v}\sum_{\ell=0}^{d}\frac{\overline{p_\ell(i)}\,\overline{p_\ell(j)}p_\ell(k)}{v_\ell^2}.$$

(Here the second equality is derived using 2.3.1). We see that the Krein parameters are determined by the eigenvalues of the scheme.

If M is a square matrix and $p(t)$ a polynomial, we define the *Schur polynomial* $p \circ M$ to be the matrix with

$$(p \circ M)_{i,j} = p(M_{i,j}).$$

We define the *Schur diameter* of a matrix M to be the least integer s such that there is a polynomial p with degree s and $p \circ M$ is invertible. (If A is the adjacency matrix of a directed graph, the diameter of the graph is the least integer s such that there is a polynomial p of degree s and $p \circ A$ is Schur invertible.)

2.4.1 Lemma. *If E is a square matrix with Schur diameter s, the Schur powers*

$$J, E, \ldots, E^{\circ s}$$

are linearly independent.

Proof. If $E^{\circ(r+1)}$ lies in the span U_r of the first r Schur powers of E, then U_r is invariant under Schur multiplication by E_r. Therefore U_r contains all Schur polynomials in E. If $r < s$, no Schur polynomial in E is invertible, which contradicts our hypothesis. It follows that spaces $U_0, \ldots, U_s$ form a strictly increasing sequence, and this implies the lemma. □

Let $\mathscr{A}$ be an association scheme with d classes. If E_i is a matrix idempotent of $\mathscr{A}$ with Schur diameter s, then $s \leq d$. We say $\mathscr{A}$ is *cometric* with respect to E_i if the Schur diameter of E_i is d. The Johnson scheme, the Grassman scheme, the Hamming scheme and the bilinear forms scheme are all cometric. A primitive strongly regular graph is primitive with respect to each non-identity idempotent. If $\mathscr{A}$ is cometric with respect to the idempotent E, then it is conventional to order the idempotents so that $E^{\circ r}$ is a linear combination of $E_0, \ldots, E_r$.

In the following we make use of the Kronecker product of matrices. What we need is summarised in Section 4.1.

Examples show that the Krein parameters need not be non-negative integers, or even rational. We do have the following.

2.4.2 Theorem. *The Krein parameters are non-negative real numbers.*

Proof. From (2.4.1), we see that the Krein parameters are the eigenvalues of the matrix $vE_i \circ E_j$. The matrices E_i and E_j are positive semidefinite, and therefore $E_i \otimes E_j$ is a positive semidefinite matrix. The matrix $E_i \circ E_j$ is a principal submatrix of this Kronecker product, and therefore it is positive semidefinite too. Hence its eigenvalues are non-negative real numbers. □

We offer a second proof that the Krein parameters are non-negative real numbers.

Let $\mathscr{A}$ be an association scheme on v vertices and let $e_1, \ldots, e_v$ denote the standard basis for $\mathbb{C}^v$. Define $\mathscr{T}$ by

$$\mathscr{T} = \sum_{i=1}^{v} e_i \otimes e_i \otimes e_i.$$

2.4.3 Lemma. *Let $\mathscr{A}$ be an association scheme. Then*

$$q_{i,j}(k) = \frac{v}{m_k} \| (E_i \otimes E_j \otimes E_{k'}) \mathscr{T} \|^2,$$

and $q_{i,j}(k) = 0$ if and only if $(E_i \otimes E_j \otimes E_{k'})\mathscr{T} = 0$.

Proof. We have

$$\mathrm{sum}(E_i \circ E_j \circ E_k) = \mathscr{T}^* (E_i \otimes E_j \otimes E_k) \mathscr{T}.$$

Since $E_i \otimes E_j \otimes E_k$ is idempotent and self-adjoint,

$$\mathrm{sum}(E_i \circ E_j \circ E_k) = \| (E_i \otimes E_j \otimes E_k) \mathscr{T} \|^2.$$

Both claims of the lemma follow. □

If $q_{i,j}(k) = 0$, then $E_k(E_i \circ E_j) = 0$ and therefore each column of $E_{k'} = E_k^T$ is orthogonal to each column of $E_i \circ E_j$. We will need the following strengthening of this result.

2.4.4 Lemma. *Let $\mathscr{A}$ be an association scheme on v vertices. If $q_{i,j}(k) = 0$ and x, y and z are three elements of $\mathbb{C}^v$, then $E_{k'}z$ is orthogonal to $E_i x \circ E_j y$.*

Proof. We have

$$\mathscr{T}^*(E_i \otimes E_j \otimes E_{k'})(x \otimes y \otimes z) = \mathbf{1}^*(E_i x \circ E_j y \circ E_{k'} z).$$

The right side is zero if and only if $E_{k'}z$ is orthogonal to $E_i x \circ E_j y$. The left side is zero if $q_{i,j}(k) = 0$. □

Suppose $\mathscr{A}$ is cometric with respect to E_1. A *harmonic polynomial* of degree i is defined to be an element of the column space of E_i. A *polynomial function* of degree i is a linear combination of harmonic polynomials with degree at most i. The previous result implies that if f is a polynomial with degree 1 and g is a polynomial with degree i, then $f \circ g$ has degree at most $i+1$. Note that $f \circ g$ is just the usual product of functions.

2.5 The Frame Quotient

Let $\mathscr{A}$ be an association scheme with d classes on the vertex set v. Let e_u be the characteristic vector of the vertex u and let H be the matrix

$$H := (A_0 e_u, A_1 e_u, \cdots, A_d e_u).$$

Then H is the characteristic matrix of the relation partition relative to the vertex u, and an easy computation shows that the column space of V is A-invariant. Hence there are $(d+1) \times (d+1)$ matrices $B_0, \ldots, B_d$ such that

$$A_r H = H B_r.$$

Since

$$A_r A_i e_u = \sum_{j=0}^{d} p_{r,i}(j) A_j e_u,$$

we find that

$$(B_r)_{i,j} = p_{r,j}(i).$$

The matrices B_r are the *intersection matrices* of the scheme. They form an algebra of $(d+1)\times(d+1)$ matrices isomorphic to the Bose-Mesner algebra of $\mathscr{A}$, because

$$B_r B_s = \sum_{i=0}^{d} p_{r,s}(i) B_i.$$

There are also matrices $F_0,\dots,F_d$ such that $E_r H = HF_r$ and

$$B_r = \sum_{i=0}^{d} p_r(i) F_i.$$

Since $E_r^2 = E_r$, we have $HF_r = HF_r^2$ and since the columns of H are linearly independent, it follows that F_r is an idempotent and

$$I = \sum_r F_r.$$

As $\mathrm{tr}(F_r)$ is a positive integer, this implies that $\mathrm{tr}(F_r) = 1$ for all r. Therefore

$$\mathrm{tr}(B_{r'} B_s) = \sum_i p_{r'}(i) p_s(i) = \sum_i \overline{p_r(i)} p_s(i) = (P^* P)_{r,s}.$$

One consequence of this is that the entries of P^*P are integers.

2.5.1 Theorem. *Let P be the eigenmatrix of the association scheme $\mathscr{A}$, let p be a prime and let $\mathbb{F}$ denote $GF(p)$. Then $\mathbb{F}[\mathscr{A}]$ contains a non-zero nilpotent matrix if and only if p divides*

$$v^{d+1} \prod_{i=0}^{d} \frac{v_i}{m_i}.$$

Proof. Let $G = P^*P$ viewed as a matrix over $\mathbb{F}$. Suppose $M \in \mathbb{F}[\mathscr{A}]$ and $MH = HN$. By changing u if needed, we may assume that $MH \neq 0$. If $M^2 = 0$ then $HN^2 = 0$ and therefore $N^2 = 0$. Hence $(B_r N)^2 = 0$ for each r and so

$$\mathrm{tr}(B_{r'} N) = 0$$

for all r. Since N is an $\mathbb{F}$-linear combination of $B_0,\dots,B_d$, this implies that the null space of G is not zero.

Suppose conversely that the null space of G is not zero. If $Gx = 0$ where $x \neq 0$ and

$$N := \sum_r x_r B_r,$$

then $\mathrm{tr}(B_{r'}N) = 0$ for all r. Therefore $\mathrm{tr}(N^k) = 0$ when $k > 0$, and so N is nilpotent.

We conclude that $\mathbb{F}[\mathscr{A}]$ contains a nilpotent element if and only if $\det(G) = 0 \mod p$. As we will see in Section 3.1,

$$P^* D_m P = v D_v$$

and therefore

$$\det(P^* P) = \det(v D_v) / \det(D_m).$$

The theorem follows immediately. □

The expression

$$v^{d+1} \prod_{i=0}^{d} \frac{v_i}{m_i}$$

is known as the *Frame quotient* of the scheme. It is known that for each k and any prime p,

$$|\{i : p^k | m_i\}| \leq |\{i : p^k | v v_i\}|.$$

One consequence of the previous theorem is that $\mathbb{F}[\mathscr{A}]$ is semisimple if and only if the Frame quotient is not divisible by p, the characteristic of $\mathbb{F}$.

The Frame quotient of the Petersen graph is

$$1000 \frac{18}{20} = 900.$$

It is not a surprise that this quotient is a perfect square, since the following simple observation holds.

2.5.2 Lemma. *If the eigenvalues of an association scheme are integers, the Frame quotient is the square of an integer.* □

Notes

There is little to say about this section; our approach is straightforward and fairly standard. We have not addressed the problem of determining the parameters of an association scheme. The actual approach taken will depend on how the scheme is presented. If the scheme is the centralizer of a multiplicity free permutation representation of a group, then it may be possible to use character theory. In general though the problem is usually difficult for association schemes with more than three classes.

Chapter 3

An Inner Product

Here we find that the Bose-Mesner algebra of an association scheme is an inner product space. The inner product can be computed in two different ways, and both the matrix and the Schur idempotents form orthogonal bases relative to it. This leads immediately to one of the most important application of association schemes, namely the linear programming method developed by Delsarte [**?**].

3.1 An Inner Product

There is one important property of Bose-Mesner algebras still to be discussed. If $M, N \in \mathrm{Mat}_{m\times n}(\mathbb{C})$, we define

$$\langle M, N\rangle := \mathrm{tr}(M^* N).$$

As is well known, this is a complex inner product on $\mathrm{Mat}_{m\times n}(\mathbb{C})$. Note that

$$\langle N, M\rangle = \overline{\langle M, N\rangle}.$$

If $\mathrm{sum}(M)$ denotes the sum of the entries of the matrix M, then

$$\mathrm{tr}(M^* N) = \mathrm{sum}(\overline{M} \circ N)$$

and therefore

$$\langle M, N\rangle = \mathrm{sum}(\overline{M} \circ N).$$

It follows that the Bose-Mesner algebra of an association scheme $\mathscr{A}$ is an inner product space. If

$$\mathscr{A} = \{A_0, \ldots, A_d\}$$

then $\mathscr{A}$ is an orthogonal set: if $i \neq j$, then $A_i \circ A_j = 0$ and therefore

$$\langle A_i, A_j \rangle = \mathrm{sum}(A_i \circ A_j) = 0.$$

Similarly if $i \neq j$, then $E_i E_j = 0$ and

$$\langle E_i, E_j \rangle = \mathrm{tr}(E_i^* E_j) = \mathrm{tr}(E_i E_j) = 0.$$

We have

$$\langle A_i, E_j \rangle = \mathrm{tr}(A_i^T E_j) = \mathrm{tr}(\overline{p_i(j)} E_j) = m_j \overline{p_i(j)}$$

and

$$\langle A_i, E_j \rangle = \mathrm{sum}(A_i \circ E_j) = q_j(i) v_i.$$

Hence we have the important relation:

$$\frac{q_j(i)}{m_j} = \frac{\overline{p_i(j)}}{v_i}. \tag{3.1.1}$$

We express this last identity in matrix terms. Let D_m be the $(d+1) \times (d+1)$ diagonal matrix with i-th diagonal entry equal to m_i and let D_v be the $(d+1) \times (d+1)$ diagonal matrix with i-th diagonal entry equal to v_i. Then 3.1.1 implies that

$$Q D_m^{-1} = D_v^{-1} P^*$$

or equivalently that

$$Q = D_v^{-1} P^* D_m$$

Since $PQ = vI$, it also follows that

$$v D_v = P^* D_m P.$$

3.2 Orthogonal Projection

Now suppose $M \in \mathrm{Mat}_{v \times v}(\mathbb{C})$ and let $\widehat{M}$ denote the orthogonal projection of M onto $\mathbb{C}[\mathscr{A}]$. Then

$$\widehat{M} = \sum_{i=0}^{d} \frac{\langle A_i, M \rangle}{\langle A_i, A_i \rangle} A_i,$$

since $A_0, \ldots, A_d$ is an orthogonal basis for $\mathbb{C}[\mathscr{A}]$. But we also have

$$\widehat{M} = \sum_{j=0}^{d} \frac{\langle E_j, M \rangle}{\langle E_j, E_j \rangle} E_j.$$

This yields a new proof of the following important result, which I am ascribing to J. J. Seidel.

3.2.1 Theorem. *If the matrices $A_0, \ldots, A_d$ form an association scheme on v vertices with idempotents $E_0, \ldots, E_d$ and $M \in \mathrm{Mat}_{v\times v}(\mathbb{C})$, then*

$$\widehat{M} = \sum_{i=0}^{d} \frac{\langle A_i, M\rangle}{v v_i} A_i = \sum_{j=0}^{d} \frac{\langle E_j, M\rangle}{m_j} E_j.$$

Proof. We note that

$$\langle A_i, A_i\rangle = \mathrm{sum}(A_i \circ A_i) = \mathrm{sum}(A_i) = v v_i$$

and

$$\langle E_j, E_j\rangle = \mathrm{tr}(E_j) = m_j. \qquad \square$$

The way to view this result is that the first expression for $\widehat{M}$ gives us its entries, while the second gives us its eigenvalues. The set

$$\{i : 1 \le i \le d,\ \langle M, A_i\rangle \neq 0\}$$

is called the *degree set* of M, and its size is the *degree* of M. The set

$$\{i : 1 \le i \le d,\ \langle M, E_i\rangle \neq 0\}$$

is called the *dual degree set* of M, and its size if the *dual degree* of M.

3.3 Linear Programming

Suppose $\mathscr{A}$ is an association scheme with vertex set V and d classes. If C is a subset of V, its *degree set* if the set of integers i such that some pair of distinct vertices in C is i-related. (This usage is consistent with the usage introduced at the end of the previous section.) The degree set of C is a subset of $\{1, \ldots, d\}$. If $R \subseteq \{1, \ldots, d\}$, we call a subset C of V an *R-clique* if the degree set of C is contained in R. If thedegree set of C is disjoint from R, we call it an *R-coclique*. A clique in X_i is an $\{i\}$-clique.

Suppose y is the characteristic vector of an R-clique in $\mathscr{A}$ and $M = yy^T$. Then the projection $\widehat{M}$ of M onto the Bose-Mesner algebra of $\mathscr{A}$ satisfies

(a) $\widehat{M} \succcurlyeq 0$.

(b) If $i \notin R \cup \{0\}$, then $\widehat{M} \circ A_i = 0$.

Since

$$\widehat{M} = \sum_{i=0}^{d} \frac{y^T A_i y}{v v_i} A_i = \sum_{i=0}^{d} \frac{y^T E_i y}{m_i} E_i$$

we have

$$\mathrm{tr}(\widehat{M}) = \frac{y^T y}{v} = \frac{|C|}{v}$$

and

$$\mathrm{sum}(\widehat{M}) = \frac{y^T J y}{v} = \frac{|C|^2}{v}.$$

Accordingly

$$|C| = \frac{\mathrm{sum}(\widehat{M})}{\mathrm{tr}(\widehat{M})}.$$

We summarise our conclusions.

3.3.1 Theorem. *Let $\mathscr{A}$ be an association scheme with d classes and let C be an R-clique in it. Then*

$$|C| \le \max_M \frac{\mathrm{sum}(M)}{\mathrm{tr}(M)},$$

where M runs over the positive semidefinite matrices in $\mathbb{C}[\mathscr{A}]$ such that $M \circ A_i = 0$ if $i \notin R \cup \{0\}$. □

We next derive a bound on the size of an R-coclique. Let N be a matrix in $\mathbb{C}[\mathscr{A}]$ such that

(a) $N \succcurlyeq 0$.

(b) If $i \notin R \cup \{0\}$, then $N \circ A_i \le 0$.

Assume

$$N = \sum_i a_i A_i = \sum_i b_i E_i$$

and let x be the characteristic vector of an R-coclique S. If $i \in R$ then $x^T A_i x = 0$ and, if $i \notin R \cup \{0\}$, then $a_i x^T A_i x \le 0$. Consequently

$$x^T N x = \sum_i a_i x^T A_i x \le a_0 x^T x = a_0 |S|$$

and

$$x^T N x = \sum_j b_j x^T E_j x \geq b_0 x^T E_0 x \geq \frac{b_0}{v}|S|^2.$$

Hence

$$|S| \leq v\frac{a_0}{b_0} = v\frac{\mathrm{tr}(N)}{\mathrm{sum}(N)}.$$

Thus we have the following. (Note that $\mathrm{tr}(N) = a_0 v$ and $\mathrm{sum}(N) = b_0 v$.)

3.3.2 Theorem. *Let $\mathscr{A}$ be an association scheme with d classes and let S be an R-coclique in it. Then*

$$|S| \leq \min_N v\frac{\mathrm{tr}(N)}{\mathrm{sum}(N)}$$

where N runs over the set of positive semidefinite matrices in $\mathbb{C}[\mathscr{A}]$ such that $N \circ A_i \leq 0$ if $i \notin R \cup \{0\}$.

From this theorem we also see that

$$\frac{v}{|S|} \geq \max_N \frac{\mathrm{sum}(N)}{\mathrm{tr}(N)}$$

where N runs over the set of positive semidefinite matrices in $\mathbb{C}[\mathscr{A}]$ such that $N \circ A_i \leq 0$. Hence the same inequality holds when N runs over the smaller set of positive semidefinite matrices in $\mathbb{C}[\mathscr{A}]$ such that $N \circ A_i = 0$ if $i \notin R \cup \{0\}$. It follows from Theorem 3.3.1 that if C is an R-clique, then

$$\frac{v}{|S|} \geq |C|.$$

Thus we have proved that if C is an R-clique and S is an R-coclique in $\mathscr{A}$, then

$$|C||S| \leq v. \tag{3.3.1}$$

This inequality is due to Delsarte. We offer an alternative derivation of it in Section 4.6.

If P is the matrix of eigenvalues of $\mathscr{A}$ and $a = (a_0, \ldots, a_d)$, then the eigenvalues of the matrix $M = \sum_i a_i A_i$ are the entries of the vector Pa. Since

$$\mathrm{tr}(M) = v a_0, \qquad \mathrm{sum}(M) = v e_0^T P a,$$

we see that $|S|$ is bounded above by the value of the following linear program

$$\begin{aligned} \max e_0^T P a \\ a_0 = 1, a_i = 0 \text{ if } i \in R \\ Pa \geq 0 \\ a \geq 0. \end{aligned}$$

Alternatively, suppose $b = (b_0, \ldots, b_d)$. Then the entries of the matrix $N = \sum_j b_j E_j$ are the entries of the vector $P^{-1}b$. Since $PQ = vI$ and

$$\text{sum}(N) = b_0, \qquad \text{tr}(N) = e_0^T Q b,$$

we see that $|S|$ is bounded above by the reciprocal of the value of the linear program

$$\begin{aligned} \min e_0^T Q b \\ b \geq 0 \\ b_0 = 1, e_i^T Q b \leq 0 \text{ if } i \in R. \end{aligned}$$

In working with these linear programs, it can be useful to recall that $Q = D_v^{-1} P^* D_m$. If in the last linear program we replace the constraints $e_i^T Q b \leq 0$ by $e_i^T Q b = 0$, the resulting linear program is dual to the first.

3.4 Cliques and Cocliques

We use the theory of the previous section to derive some specific bounds. Let $\mathscr{A}$ be an association scheme on v vertices with d classes.

Suppose first that C is a 1-clique, that is, a clique in the graph X_1 with adjacency matrix A_1. We seek to use Theorem 3.3.1 to obtain an upper bound on $|C|$. If $M \in \mathbb{C}[\mathscr{A}]$ and $M \circ A_i = 0$ if $i \neq 0, 1$, then

$$M = aI + bA_1.$$

Hence

$$\frac{\text{sum}(M)}{\text{tr}(M)} = \frac{av + bvv_1}{av} = 1 + v_1 \frac{b}{a}.$$

Here v_1 is the valency of X_1, and want to choose a and b to maximise the last term, subject to the condition that $M \succcurlyeq 0$. Since our objective function depends

only on the ratio b/a, we may assume $b = \pm 1$. If $b = -1$, then the least eigenvalue of $aI - A_1$ is $a - v_1$, and we maximise our objective function by taking $a = v_1$. The value of the objective function is 2. If $b = 1$ and the least eigenvalue of A_1 is τ, then the least eigenvalue of $aI + A_1$ is $a + \tau$ and we maximise our objective function by taking $a = -\tau$. This gives a bound

$$|C| \leq 1 - \frac{v_1}{\tau}.$$

This bound is never less than 2, and so it is the linear programming bound on a 1-clique.

3.4.1 Lemma. *If X is a graph in an association scheme with valency k and least eigenvalue τ, then*

$$\omega(X) \leq 1 - \frac{k}{\tau}.$$ □

By using Theorem 3.3.2, we can derive an upper bound on the size of a 1-coclique in a union of classes from an association scheme. Suppose A is the adjacency matrix of such a graph X with valency k, and that its least eigenvalue is τ. If $N := A - \tau I$, then $N \succcurlyeq 0$ and

$$\mathrm{tr}(N) = -v\tau, \qquad \mathrm{sum}(N) = vk - v\tau$$

By Theorem 3.3.2, this results in the bound

$$\alpha(X) \leq \frac{v}{1 - \frac{k}{\tau}}.$$

This bound actually holds for all regular graphs. Note that here we did not need to solve the linear program in Theorem 3.3.2, any matrix which satisfies the conditions provides an upper bound.

We give an application of the inequality (3.3.1). Let $\mathscr{A}$ be the Hamming scheme $H(n,q)$. Let B_e denote the ball of radius e about the zero word in the Hamming graph. Then

$$\beta_e := |B_e| = \sum_{i=0}^{e} \binom{n}{i} (q-1)^i.$$

Any two words in B_e are at distance at most $2e$. If $R := \{2e+1, \ldots, n\}$, then B_e is an R-coclique, while an R-clique is a code with minimum distance $2e+1$. So (3.3.1) yields that

$$|C| \leq \frac{q^n}{\beta_e};$$

in coding theory this is the *sphere-packing bound*.

Note that if C is an R-clique and D is an R-coclique, then $|C \cap D| \leq 1$. Hence if we could partition the vertex set V of $\mathscr{A}$ into disjoint copies of an R-coclique D, no code has more than one vertex in any cell of this partition and so we trivially get the bound $|C| \leq |V|/|D|$. Suppose q is a prime power and the vertices of the Hamming scheme $H(n,q)$ are taken to be the vectors in $V(n,q)$. If D is an R-coclique and a subspace, then the cosets of D partition the vertices of $H(n,q)$ into copies of D and therefore any S-clique contains at most $q^n/|D|$ vertices. The above result enables us to derive the same bound, given only a single copy of D. From a coding theorist's viewpoint, association schemes provide a tool which enables us to extend results about linear codes to the general case. This crucial fact is due to Delsarte.

3.5 Feasible Automorphisms

Let $\mathscr{A}$ be an association scheme with d class on v vertices. Let P be a $v \times v$ permutation matrix. Then P is an automorphism of $\mathscr{A}$ if it commutes with each Schur idempotent A_i or equivalently if it commutes with each matrix in $\mathbb{C}[\mathscr{A}]$.

We derive a necessary condition for P to be an automorphism, due to G. Higman.

Let σ denote the permutation associated with P. Define $v_i(\sigma)$ to be the number of vertices u in the scheme such that u is i-related to u. Then

$$v_i(\sigma) = \mathrm{sum}(P \circ A_i).$$

We compute the projection $\widehat{P}$ of P onto $\mathbb{C}[\mathscr{A}]$:

$$\widehat{P} = \sum_{i=0}^{d} \frac{v_i(\sigma)}{v v_i} A_i = \sum_{i=0}^{d} \frac{\langle P, E_i \rangle}{m_i} E_i.$$

Therefore

$$\widehat{P} E_j = \sum_{i=0}^{d} \frac{v_i(\sigma)}{v v_i} p_i(j) E_j = \frac{\langle P, E_j \rangle}{m_j} E_j$$

and consequently

$$\langle P, E_j \rangle = \frac{m_j}{v} \sum_{i=0}^{d} \frac{p_i(j)}{v_i} v_i(\sigma).$$

We claim that if P is an automorphism, then $\langle P, E_j \rangle$ must be an algebraic integer. For since E_j is idempotent and Hermitian, we may write it as

$$E_j = UU^*$$

where U is a $v \times m_j$ matrix such that $U^*U = I$. Hence

$$\langle P, E_j \rangle = \mathrm{tr}(E_j P) = \mathrm{tr}(UU^* P) = \mathrm{tr}(U^* PU).$$

If P commutes with E_j, then

$$PUU^* = UU^* P$$

and therefore

$$PU = U(U^* PU).$$

This implies that the characteristic polynomial of U^*PU divides the characteristic polynomial of P, and therefore $\mathrm{tr}(U^*PU)$ is a sum of eigenvalues of P. Hence it is an algebraic integer.

We apply this theory to the Petersen graph. Suppose σ is an automorphism of this graph which maps each vertex to an adjacent vertex. Thus

$$\nu_0(\sigma) = 0,\ \nu_1(\sigma) = 10,\ \nu_2(\sigma) = 0.$$

The eigenvalues of the Petersen graph are -2, 1 and 3 with respective multiplicities 4, 5 and 1. If E_1 is the matrix idempotent associated to the eigenvalue 1 and A_1 is the adjacency matrix of the Petersen graph, then

$$\langle P, E \rangle = \frac{5}{10} \times \frac{1}{3} \times 10 = \frac{4}{3}.$$

Since $4/3$ is not an algebraic integer, we conclude that no automorphism of the Petersen graph maps each vertex to an adjacent vertex.

Suppose H is a projection that commutes with $\mathbb{C}[\mathscr{A}]$. Then the above argument yields that

$$\langle H, E_j \rangle = \frac{m_j}{v} \sum_{i=0}^{d} \frac{p_i(j)}{v_i} \langle H, A_i \rangle$$

is a non-negative integer. (The value of this observation is unclear, but in principle it could be used to show that certain equitable partitions do not exist.)

Notes

The observation that the Bose-Mesner algebra of an association scheme is an inner product space is surprising useful, and allows a comparatively easy approach to the linear programming method. Nonetheless the results in this chapter are all standard. The linear programming method was developed in [**?**] by Delsarte. The method developed in Section 3.5 is an unpublished idea of G. Higman, and is used in [**?**] to show that a Moore graph of diameter two and valency 57 cannot be vertex transitive.

Chapter 4

Products and Tensors

We show how to use the Kronecker product of matrices, or equivalently the tensor product of algebras, to construct new association schemes from old.

4.1 Kronecker Products

If A and B are matrices and $A = (a_{i,j})$, we define their *Kronecker product* $A \otimes B$ to be the matrix we get by replacing $a_{i,j}$ with the matrix $a_{i,j}B$, for all i and j. (We have made use of this already in Section 2.4.) We summarise some of the basic properties of this operation.

First it is linear in each variable and, for any scalar c

$$cA \otimes B = A \otimes cB.$$

We have

$$\operatorname{tr}(A \otimes B) = \operatorname{tr}(A)\operatorname{tr}(B)$$

and

$$(A \otimes B)^T = A^T \otimes B^T.$$

4.1.1 Lemma. *If the matrix products AC and BD are defined, then*

$$(A \otimes B)(C \otimes D) = AC \otimes BD. \qquad \square$$

One consequence of this is that if x and y are eigenvectors of A and B respectively, then $x \otimes y$ is an eigenvector of $A \otimes B$. It follows that if A and B are positive semidefinite, so is $A \otimes B$. Note also that

$$A \otimes B = (A \times I)(I \otimes B) = (I \otimes B)(A \otimes I).$$

The Kronecker product also interacts nicely with the Schur product:

4.1.2 Lemma. *If A and C are matrices of the same order and B and C are matrices of the same order, then*

$$(A \otimes B) \circ (C \otimes D) = (A \circ C) \otimes (B \circ D). \qquad \square$$

4.1.3 Lemma. *There is a permutation matrix P such that $P^2 = I$ and*

$$P(A \otimes B)P = B \otimes A.$$

Proof. Assume A has m columns and B has n. Let $e_1, \ldots, e_m$ and $f_1, \ldots, f_n$ denote the standard bases for $\mathbb{F}^m$ and $\mathbb{F}^n$ respectively. Then

$$e_i \otimes f_j, \quad 1 \leq i \leq m,\ 1 \leq j \leq n$$

and

$$f_j \otimes e_i, \quad 1 \leq i \leq m,\ 1 \leq j \leq n$$

are two orderings of the standard basis for $\mathbb{F}^{mn}$. Define P to be the matrix that represents the linear mapping that takes $e_i \otimes f_j$ to $f_j \otimes e_i$, for all i and j. Then P is a permutation matrix and $P^2 = I$. Finally

$$(B \otimes A)P(e_i \otimes f_j) = Be_j \otimes Ae_i = P(Ae_i \otimes Be_j), = P(A \otimes B)(e_i \otimes f_j).$$

and as this holds for all i and j and as $P^T = P$, the result follows. $\square$

4.2 Tensor Products

The Kronecker product is a concrete realisation of the tensor product of vector spaces, which we introduce now.

Roughly speaking, tensor products are a tool we use to avoid discussing bilinear functions. We recall that if V_1, V_2 and W are vector spaces over a field, then a map β from $V_1 \times V_2$ is *bilinear* if it is linear in each variable. Note here that, although $V \times W$ is the underlying set of vectors for the vector space $V_1 \oplus V_2$, the bilinear map β is not a linear map from $V_1 \oplus V_2$. (A good example of a bilinear map is the determinant of a 2×2 real matrix, viewed as a function from the two columns of the matrix to $\mathbb{R}$. Thus here $V_1, V_2 \cong \mathbb{R}^2$ and $W = \mathbb{R}$.) Although bilinear maps are not linear maps, the set of bilinear maps from $V_1 \times V_2$ to W is a vector space.

The solution to our avoidance problem is to construct a new vector space, denoted $V_1 \otimes V_2$ and a canonical bilinear map

$$\beta : V_1 \times V_2 \to V_1 \otimes V_2$$

so that for each bilinear map γ from $V_1 \times V_2$ to W there is a linear map

$$g : V_1 \otimes V_2 \to W$$

such that $\gamma = g \circ \beta$. Since β is determined by the three vector spaces V_1, V_2 and W, we see that we can work with linear maps from $V_1 \otimes V_2$ to W in place of bilinear maps from $V_1 \times V_2$. (Thus we have simplified the maps we need to deal with by complicating our objects. This is not an uncommon practice in mathematics, and the trade-off is usually worth it.)

The construction of $V_1 \otimes V_2$ is a two-stage process—we derive it as a quotient of a larger vector space. Let U denote the vector space of all functions with finite support from $V_1 \times V_2$ to the underlying field. Less formally we may identify U as the set of (finite) linear combinations

$$\sum_i a_i (x_i, y_i)$$

where $(x_i, y_i) \in V_1 \times V_2$. Next we introduce the subspace U_0 of U spanned by all linear combinations of the following vectors:

$$\begin{aligned}
&(x_1 + x_2, y) - (x_1, y) - (x_2, y) \\
&(x, y_1 + y_2) - (x, y_1) - (x, y_2) \\
&(ax, y) - a(x, y) \\
&(x, ay) - a(x, y).
\end{aligned}$$

Finally we define $V_1 \otimes V_2$ to be the quotient space U/U_0; the canonical bilinear map β is defined by

$$\beta(x, y) := x \otimes y.$$

The problem left is to prove that each bilinear map from $V_1 \times V_2$ can be expressed uniquely as a composition of a linear map from $V_1 \otimes V_2$ to W. We leave this to the reader—there are lots of places where you can look it up! We note that the Kronecker product of $\mathbb{F}^m$ and $\mathbb{F}^m$ is isomorphic to $\mathbb{F}^m \otimes \mathbb{F}^n$ (and in turn this isomorphic to $\mathbb{F}^{mn}$). In particular if V and W have finite dimension, then

$$\dim(V \otimes W) = \dim(V) \dim(W).$$

Having introduced tensor products of vector spaces, we turn to algebras. Suppose V_1 and V_2 are vector spaces and $A_i \in \mathrm{End}(V_i)$. We define $A_1 \otimes A_2$ by decreeing that if $x_i \in V_i$, then

$$(A_1 \otimes A_2)(x_1 \otimes x_2) := A_1 x_1 \otimes A_2 x_2.$$

It remains to be checked that this product satisfies the rules we gave for the Kronecker product (you may do this) and hence that the span of the products $A_1 \otimes A_2$ is an algebra. It is isomorphic as a vector space to $\mathrm{End}(V_1) \otimes \mathrm{End}(V_2)$.

Tensor products may be generalised. We could attempt to work with vector space with infinite dimension or, in place of vector spaces, we could use modules. In both cases the subject becomes much more more subtle—for example the tensor product of the $\mathbb{Z}$-modules $\mathbb{Z}_2$ and $\mathbb{Z}_3$ is the zero module. But even the case of tensor products of finite dimensional vector spaces there can be surprises.

We offer some exercises. The complex numbers are an algebra over the reals, show that $\mathrm{Mat}_{d\times d}(\mathbb{R}) \otimes \mathbb{C} \cong \mathrm{Mat}_{d\times d}(\mathbb{C})$. Show that the tensor product of $\mathbb{F}[x]$ with itself is isomorphic to the ring of polynomials in two non-commuting variables $\mathbb{F}[x, y]$.

Suppose V is a finite dimensional vector space with dual space V^*. If

$$x \otimes f \in V \otimes V^*$$

then we can define an associated mapping from V to itself by

$$(x \otimes f)(v) := f(v)x, \qquad v \in V.$$

Here $f(v)$ is a scalar, and it is easy to verify that $x \otimes f \in \mathrm{End}(V)$. Do this, and also show that

$$V \otimes V^* \cong \mathrm{End}(V).$$

If $V = \mathbb{F}^n$, with elements viewed as column vectors and we identify V^* with $\mathbb{F}^n$ viewed as row vectors, then we may identify $x \otimes f$ with the matrix xy^T for suitable y. The isomorphism above is then equivalent to the observation that every matrix can be written as a sum of rank-one matrices.) If $V = \mathbb{F}^n$ we are prepared to identify V^* with V, then we can identify $V \otimes V$ with the space of $n \times n$ matrices.

4.3 Tensor Powers

We consider constructions of association schemes that make use of the tensor product.

4.3.1 Lemma. *If $A_o,\dots,A_d$ and $B_0,\dots,B_e$ are two association schemes with d and e classes respectively, then the matrices*

$$A_i \otimes B_j, \quad 0 \le i \le d,\ 0 \le j \le e$$

form an association scheme with $de+d+e$ classes, and that the Bose-Mesner algebra of this product is the tensor product of the Bose-Mesner algebras of its factors.

Proof. This is not hard to verify directly. Alternatively let the two schemes be denoted by $\mathscr{A}$ and $\mathscr{B}$ respectively. It follows from Lemma 4.1.1 and Lemma 4.1.2 that the tensor product

$$\mathbb{C}[\mathscr{A}] \otimes \mathbb{C}[\mathscr{B}]$$

is closed under matrix and Schur multiplication. Since it contains J and is transpose-closed, we deduce that it is the Bose-Mesner algebra of a scheme. The dimension of this algebra is $(d+1)(e+1)$ and hence this product scheme has the stated number of classes. □

Similarly we have a power construction:

4.3.2 Lemma. *If $\mathscr{A}$ is an association scheme with d classes, then $\mathbb{C}[\mathscr{A}]^{\otimes k}$ is the Bose-Mesner algebra of an association scheme with $(d+1)^k - 1$ classes.* □

It is not hard to construct new association schemes with a large number of classes, hence the previous two constructions are not as useful as we might hope. However there is an interesting construction based on the tensor power, which we develop now.

Suppose V is a vector space. We define an action of $\mathrm{Sym}(k)$ on $V^{\otimes k}$ by declaring that if

$$x_1 \otimes \cdots \otimes x_k$$

and $\sigma \in \mathrm{Sym}(k)$, then

$$\sigma : x_1 \otimes \cdots \otimes x_k \mapsto x_{1\sigma} \otimes \cdots \otimes x_{k\sigma}.$$

It follows that σ induces a linear map from $V^{\otimes k}$ to itself (which we will denote by σ. If $e_1, \ldots, e_d$ is a basis for V, then the products

$$e_{i_1} \otimes \cdots \otimes e_{i_k}$$

form a basis for $V^{\otimes k}$. Since σ permutes the elements of this basis, the matrix representing σ is a permutation matrix.

Note that some elements of $V^{\otimes k}$ are left fixed by the action of $\mathrm{Sym}(k)$. As examples we have the diagonal terms

$$e_i \otimes \cdots \otimes e_i$$

and, when $k = 2$, the sum

$$e_1 \otimes e_2 + e_2 \otimes e_1$$

is fixed by $\mathrm{Sym}(2)$. We define the *k-th symmetric power* of V to be the subspace of $V^{\otimes k}$ formed by the vectors that are fixed by each element of $\mathrm{Sym}(k)$. If $\dim(V) = d$, then its k-th symmetric power has dimension $\binom{d+k-1}{k}$.

4.3.3 Theorem. *If $\mathscr{A}$ is an association scheme with d classes, then the k-th symmetric power of $\mathbb{C}[\mathscr{A}]$ is an association scheme with $\binom{d+k}{k} - 1$ classes.*

Proof. The k-th symmetric power of $\mathbb{C}[\mathscr{A}]$ is the centralizer of a set of permutation matrices, and therefore it is Schur-closed by Lemma **??**. It is closed under matrix multiplication and transpose and contains I and J, and it is commutative since $\mathbb{C}[\mathscr{A}]$ is. Therefore it is the Bose-Mesner algebra of an association scheme. □

We call the scheme produced by this construction the *k-th symmetric power* of $\mathscr{A}$, and we denote it by $H(k, \mathscr{A})$.

We note the proof of the previous theorem theorem also yields that a symmetric power of a coherent algebra is again a coherent algebra, and this power is homogeneous if the input is.

4.4 Generalized Hamming Schemes

In this section we offer an alternative, more concrete, construction of the symmetric power and consider some examples.

Suppose $\mathscr{A}$ is an association scheme with Schur idempotents $A_0, \ldots, A_d$ and vertex set V. If u and v are two elements of V^n, let $h(u, v)$ be the vector of length

$d+1$ whose i-th entry $h_i(u,v)$ is the number of coordinates j such that u_j and v_j are i-related. The entries of $h(u,v)$ sum to n; conversely any non-negative vector of length n whose entries sum to n is equal to $h(u,v)$ for some u and v. If α is a non-negative vector of length $d+1$ and $\mathbf{1}^T\alpha = n$, define A_α to be the 01-matrix with rows and columns indexed by V^n and with $(A_\alpha)_{u,v} = 1$ if and only if $h(u,v) = \alpha$. This set of matrices forms the k-th symmetric power of $\mathscr{A}$. If $\mathscr{A}$ is the scheme with one class on q vertices, then $H(n,\mathscr{A})$ is the Hamming scheme $H(n,q)$.

By way of a more particular example, suppose I, A_1 and A_2 form an association scheme with two classes, i.e., the association scheme of a strongly regular graph. The Schur idempotents of $\mathscr{A}\otimes\mathscr{A}$ are the nine matrices

$$\begin{array}{ccc} I, & I\otimes A_1, & A_1\otimes I, \\ I\otimes A_2, & A_2\otimes I, & A_1\otimes A_2, \\ A_2\otimes A_1, & A_1\otimes A_1, & A_2\otimes A_2. \end{array}$$

The Schur idempotents of $H(2,\mathscr{A})$ are

$$\begin{array}{ccc} I, & I\otimes A_1 + A_1\otimes I, & I\otimes A_2 + A_2\otimes I, \\ A_1\otimes A_2 + A_2\otimes A_1, & A_1\otimes A_1, & A_2\otimes A_2. \end{array}$$

4.5 A Tensor Identity

We use $A\otimes B$ to denote the Kronecker product of two matrices A and B. We offer a more exalted version of Seidel's identity, due to Koppinen.

4.5.1 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

$$xy^* = x \otimes y^*,$$

we thus have

$$\sum_{i=1}^{k} \frac{1}{\langle u_i, u_i \rangle} u_i \otimes u_i^* = \sum_{i=1}^{k} \frac{1}{\langle v_i, v_i \rangle} v_i \otimes v_i^*. \tag{4.5.1}$$

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 (4.5.1) 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 $\mathcal{K}$ denote either of the two sums in the statement of Theorem 4.5.1. Since $E_j \otimes E_j$ is self-adjoint, we have $\mathcal{K}^* = \mathcal{K}$ and therefore we also have

$$\mathcal{K} = \sum_{i=0}^{d} \frac{1}{v v_i} A_i^T \otimes A_i.$$

4.6 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 Theorem 3.2.1 is a consequence of Theorem 4.5.1.

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 (3.3.1).

4.6.1 Theorem. *Let $\mathscr{A}$ be an association scheme with d classes on v vertices and let R be a subset of $\{1, \ldots, d\}$. 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

$$|S| = x x^T \geq x^T (E_0 \otimes E_0) x = \frac{|S|^2}{v^2}.$$

Therefore $|S| \leq v$. □

Notes

Bailey [**?**] also offers a detailed treatment of constructions based on tensor products. Delsarte [**?**, ???] introduced what we called the generalised Hamming schemes, calling them ???. We will see that this viewpoint leads to an elegant approach to the computation of the matrix of eigenvalues for the Hamming scheme. Koppinen's identity appears in [**?**]. Its applications in Section 4.6 are new, although the results themselves are not. (In particular the pseudocyclic schemes we present were first found by [**?**].)

Chapter 5

Subschemes and Partitions

If a subspace of the Bose-Mesner algebra of an association scheme $\mathscr{A}$ contains I and J and is closed under Schur and matrix multiplication, it must be the Bose-Mesner algebra of an association scheme, $\mathscr{B}$ say. We say that $\mathscr{B}$ is a *subscheme* of $\mathscr{A}$. In this chapter we study subschemes and some related matters.

5.1 Equitable Partitions

Let V be set of size v and let π be a partition of V with k cells. Then π is a set, each element of which is a subset of V. The *characteristic matrix* of π is the $v \times k$ matrix whose columns are the characteristic vectors of the cells of π. The column space of π is the space of functions on V that are constant on the cells of π; we denote this space by $F(\pi)$.

If $\mathscr{B}$ is an algebra of matrices with rows and columns indexed by V, we say that a partition π of V is *equitable* relative to $\mathscr{B}$ if $F(\pi)$ is $\mathscr{B}$-invariant. The algebras of interest to us will be generated by adjacency matrices of graphs. Suppose A is the adjacency matrix of a graph X and π is a partition of $V(X)$ with characteristic matrix H. Then $F(\pi)$ is A-invariant if and only if there is a $k \times k$ matrix B such that

$$AH = HB. \tag{5.1.1}$$

We call B the *quotient* of A relative to π, and denote it by A/π. If the cells of π are $C_1, \ldots, C_k$, then 5.1.1 holds if and only if for each i and j, the number of neighbours in C_i of vertex in C_j is determined by i and j (and is equal to $B_{i,j}$). Note that $H^T H$ is diagonal and invertible, whence

$$B = (H^T H)^{-1} H^T A H.$$

Hence B is determined by A and H.

We consider two classes of examples. Let A be the adjacency matrix of a graph X. A partition π of $V(X)$ is an equitable partition of X if it is A-invariant. A subspace that contains $\mathbf{1}$ is J-invariant, and so if π is A-invariant, it is invariant under the algebra generated by A and J. (The latter algebra is often the algebra of all matrices.) An *orbit partition* of a graph X is a partition whose cells are the orbits of some group of automorphisms of X. Any orbit partition of a graph X is equitable, but not all equitable partitions of a graph are orbit partitions. For example, X is regular if and only if the partition with $V(X)$ as its only cell is equitable.

Our second class of examples concerns association schemes. If $\mathscr{A}$ is an association scheme with vertex set V, we call a partition π of V an equitable partition of the scheme if it is equitable relative to the Bose-Mesner algebra $\mathbb{C}[\mathscr{A}]$. Suppose $v \in V$ and let C_i denote the set of vertices w such that v is i-related to w. Then $\pi_v = \{C_0, \ldots, C_d\}$ is a partition of $V(X)$ which is an equitable partition of the scheme. We call it the *relation partition* with respect to v. The quotient matrix A_i/π_v is independent of v.

A subspace of $\mathbb{F}^V$ is equal to $F(\pi)$ for some π if and only if it is closed under Schur multiplication and contains the vector $\mathbf{1}$. The following result is not too hard to verify.

5.1.1 Lemma. *If π and σ are partitions of V, then*

$$F(\pi) \cap F(\sigma) = F(\pi \vee \sigma). \qquad \square$$

(Here $\pi \vee \sigma$ denotes the join of π and σ in the lattice of partitions of V.) From this lemma, we see that if π and σ are equitable partitions relative to $\mathscr{B}$, then $\pi \vee \sigma$ is also equitable.

If $p(t)$ is a polynomial, then 5.1.1 implies that $p(A)H = Hp(B)$. Consequently we have a homomorphism from the algebra of polynomials in A to the algebra of polynomials in B. It follows that the minimal polynomial of B divides the minimal polynomial of A, but something a little stronger is true.

5.1.2 Lemma. *Suppose A, B and H are matrices such that $AH = HB$. If the columns of H are linearly independent, the characteristic polynomial of B divides the characteristic polynomial of A.*

Proof. Assume A is $v \times v$ and let $x_1, \ldots, x_v$ be a basis for $\mathbb{F}^v$ such that $x_1, \ldots, x_k$ are the columns of H in order. Then relative to this basis, the matrix representing

A has the form

$$\begin{pmatrix} B & * \\ 0 & A_1 \end{pmatrix}$$

where A_1 is square. Therefore

$$\det(tI - A) = \det(tI - B)\det(tI - A_1). \qquad \square$$

We give one of the standard applications of this result. Let X be a graph. A *perfect 1-code* in X is a subset C of $V(X)$ such that:

(a) Any two vertices in C are at distance at least three.

(b) Any vertex not in C is adjacent to exactly one vertex in C.

Suppose C is a perfect 1-code in X, and let π be the partition with cells $C_1 = C$ and $C_2 = V(X)\setminus C$. If X is regular with valency k, this is an equitable partition of X, with quotient matrix

$$B = \begin{pmatrix} 0 & k \\ 1 & k-1 \end{pmatrix}.$$

Both rows of this matrix sum to k, so k is an eigenvalue for B. Since $\mathrm{tr}(B) = k-1$ it follows that -1 is an eigenvalue for B. We conclude that if a regular graph X contains a perfect 1-code, then -1 must be an eigenvalue of A.

Note that if X is k-regular and has v vertices, then a perfect 1-codes has size $v/(k+1)$. Thus we have a second necessary condition: $k+1$ must divide v.

Let π be a partition with characteristic matrix H and suppose D is the non-negative diagonal matrix such that $D^2 = H^T H$. Then all columns of HD^{-1} have length 1, and so the columns of this matrix form an orthonormal set. We call HD^{-1} the *normalized characteristic matrix* of π, for some purposes it is more useful than H. If $G = HD^{-1}$ then GG^T represents orthogonal projection onto $F(\pi)$. We note that $F(\pi)$ is A-invariant if and only if A and G commute. (This is a special case of a general fact from linear algebra: if E is idempotent, then $\mathrm{col}(E)$ is A-invariant if and only if $EA = AE$.)

5.2 Subschemes and Partitions

Suppose $\mathscr{C}$ and $\mathscr{D}$ are association schemes on the same vertex set. We say that $\mathscr{C}$ is a *subscheme* of $\mathscr{D}$ is the Bose-Mesner algebra of $\mathscr{C}$ is a subspace of the Bose-Mesner algebra of $\mathscr{D}$. To give a very simple example, the association

scheme of the complete graph on v vertices is a subscheme of any scheme on v vertices. We will meet more examples soon.

We admit that what we call a subscheme is often (more often?) called a *fusion scheme* and, to make it worse, the term subscheme is used for another concept.

5.2.1 Lemma. *Let $\mathscr{A}_1$ and $\mathscr{A}_2$ be association schemes. If $\mathscr{A}_1$ is a subscheme of $\mathscr{A}_2$, then each Schur idempotent of $\mathscr{A}_1$ is a sum of Schur idempotents of $\mathscr{A}_2$.*

Proof. Suppose $\mathscr{A}_1 = \{A_0, \ldots, A_d\}$. If $\mathscr{A}_1$ is a subscheme of $\mathscr{A}_2$, then A_i must be a linear combination of Schur idempotents of $\mathscr{A}_2$; since A_i is a 01-matrix, the coefficients in this linear combination must be 0 or 1. □

Let $\mathscr{A}$ be an association scheme with d classes, and let π be a partition of $\{0, 1, \ldots, d\}$ with cells $\{C_0, \ldots, C_e\}$. Define M_i to be sum of the matrices in the set

$$\{A_r : r \in C_i\}.$$

The matrices M_i are Schur idempotents, but not minimal Schur idempotents. We call them the Schur idempotents *corresponding to the cells of π*. We wish to characterise the partitions π such that the matrices M_i are the minimal Schur idempotents of a subscheme of $\mathscr{A}$. We note three necessary conditions:

(a) Since some M_i must be the identity matrix, some cell of π must equal $\{0\}$.

(b) The set of matrices M_i is transpose-closed.

(c) The algebra generated by the matrices M_i must have dimension $|\pi|$.

It is not hard to see that together these three conditions are also sufficient. There is a direct way to determine the dimension of the algebra, which we discuss next.

5.2.2 Theorem. *Let $\mathscr{A}$ be an association scheme with d classes and eigenmatrix P. Let π be a partition of $\{0, 1, \ldots, d\}$ with cells $\{C_0, \ldots, C_e\}$ and with characteristic matrix S. Assume that the set of Schur idempotents corresponding to π is transpose-closed and contains I. Then the dimension of the algebra they generate is equal to the number of distinct rows of PS.*

Proof. Let $E_0, \ldots, E_d$ be the minimal matrix idempotents of $\mathscr{A}$. Let $M_0, \ldots, M_e$ be the Schur idempotents corresponding to the cells of π, and let $\mathscr{M}$ be the algebra they generate. Then $\mathscr{M}$ is commutative and closed under transposes

and complex conjugation. We apply the theory of Sections 1.4 and 1.5. From this it follows that $\mathscr{M}$ has a basis of pairwise orthogonal idempotents $F_1, \ldots, F_\ell$. Each of these idempotents lies in $\mathbb{C}[\mathscr{A}]$, and hence F_i is the sum of the E_r such that $E_r \leq F_i$.

Let $m_i(j)$ denote the eigenvalue of M_i associated to the idempotent E_j of $\mathscr{A}$. Then

$$m_j(i) = (PS)_{i,j}.$$

If $E_r \leq F_i$, then $M_j E_r = m_j(i) E_r$.

Suppose F_r and F_s are distinct idempotents of $\mathscr{M}$. If $m_{i,r} = m_{i,s}$ for all i, then $F_r + F_s$ together with the idempotents F_i with $i \notin \{r, s\}$ spans $\mathscr{M}$. Therefore for each r and s, there is a matrix M_i such that $m_i(r) \neq m_i(s)$.

We conclude that E_r and E_s are summands of the same idempotent F_i if and only if rows r and s of PS are equal. The theorem follows. □

The Schur idempotents corresponding to π are linearly independent, so we see that the dimension of the algebra they generate is at least $|\pi| = e + 1$.

5.2.3 Corollary. *Let $\mathscr{A}$ be an association scheme with d classes and eigenmatrix P. Let π be a partition of $\{0, 1, \ldots, d\}$ with cells $\{C_0, \ldots, C_e\}$ and with characteristic matrix S. Assume that the set $\{M_0, \ldots, M_e\}$ of Schur idempotents corresponding to π is transpose-closed and contains I. Then $\{M_0, \ldots, M_e\}$ are the minimal Schur idempotents of a subscheme of $\mathscr{A}$ if and only if PS has exactly $e + 1$ distinct rows.* □

Let P be an $m \times n$ matrix and let σ be a partition of its columns with characteristic matrix S. Define two rows of P to be equivalent if the corresponding rows of PS are equal, and let ρ be the partition of the rows of P with the equivalence classes of this relation as its cells. We say that ρ is the partition *induced* by σ.

5.2.4 Lemma. *Let P be an $m \times n$ matrix, let σ be a partition of its columns and let ρ be the partition of its rows induced by σ. If the columns of P are linearly independent, then $|\rho| \geq |\sigma|$.*

Proof. If the columns of P are linearly independent, then $PSx = 0$ if and only if $Sx = 0$. But the columns of S are linearly independent and so $Sx = 0$ if and only if $x = 0$. Hence the columns of PS are linearly independent. Therefore

$$|\sigma| = \mathrm{rk}(PS).$$

Since the number of distinct rows of PS is an upper bound on its rank, the lemma follows. □

We call a partition σ of the columns of a matrix *tight* if $|\sigma^*| = |\sigma|$.

Suppose σ is a partition of the columns of M, with induced partition ρ. Let S and R respectively denote the characteristic matrices of σ and π. Then each column of PS is a linear combination of the columns of R, and so there is a matrix M_1 such that

$$MS = RM_1.$$

Here M_1 has order $|\rho| \times |\sigma|$; if the columns of M are linearly independent, then so are the columns of M_1.

Let P be the eigenmatrix of an association scheme $\mathscr{A}$ with d classes. If σ is a tight partition of the columns of P, then

$$PS = RP_1$$

where R is the characteristic matrix of σ^*. In this case P is invertible, and so P_1 is invertible. If Q is the dual eigenmatrix of $\mathscr{A}$, then $QP = vI$ and we have

$$vSP_1^{-1} = QR.$$

This implies that σ^* is a tight partition of the columns of Q, with induced partition equal to σ.

If the Schur idempotents associated with the cells of σ form a subscheme of $\mathscr{A}$, then P_1 is the matrix of eigenvalues of the subscheme.

5.3 Primitivity

An association scheme with d classes is *primitive* if each of its graphs $X_1, \ldots, X_d$ is connected. (Although these graphs may directed, they are finite and all vertices have the same in- and out-valency, hence the strong weak components coincide.) An association scheme that is not primitive is *imprimitive*.

In the binary Hamming scheme $H(n,2)$ we see that X_n consists of 2^{n-1} disjoint edges. The graph X_1 is the n-cube, which is bipartite, and so the graphs X_{2i} are all disconnected.

The imprimitive strongly regular graphs are the graphs mK_n (where $m, n > 1$) and their complements. Product schemes are imprimitive in all non-trivial cases.

Suppose $\mathscr{A} = \{A_0, \ldots, A_d\}$ is an association scheme with vertex set V and let V_1 be a subset of V. The principal submatrices of $A_0, \ldots, A_d$ with rows and columns indexed by V_1 generate a matrix algebra with dimension at least $s+1$,

where s is the degree of V_i. If the dimension is $s+1$, it is the Bose-Mesner algebra of an association scheme with s classes.

5.3.1 Lemma. *Suppose X is the graph of a minimal Schur idempotent in the association scheme $\mathscr{A}$. If X is not connected, then the restriction of $\mathscr{A}$ to a component of X is an association scheme whose parameters are independent of the choice of components. The partition whose cells are the components of X_1 is equitable.*

Proof. Suppose that X_1 is not connected and let σ be the partition whose cells are the components of X_1. We prove that $SS^T \in \mathbb{C}[\mathscr{A}]$, which implies that σ is equitable.

If u and v are i-related vertices in the same component of X_1, they are joined by a walk, each arc of which lies in X_1. It follows that any pair of i-related vertices in $\mathscr{A}$ lie in the same component of X_1. Consequently the graph on $V(\mathscr{A})$ with two vertices adjacent if and only if they lie in the same component of X_1 is the edge-disjoint union of graphs from $\mathscr{A}$. Since $SS^T - I$ is the adjacency matrix of this graph, our claim is proved.

If $SS^T \in \mathbb{C}[\mathscr{A}]$, then SS^T commutes with J, and therefore the cells of σ all have the same size, which we denote by c. Since $SS^T \in \mathbb{C}[\mathscr{A}]$, Seidel's identity yields

$$SS^T = \sum_{i=0}^{d} \frac{\langle SS^T, A_i\rangle}{vv_i} A_i = \sum_{j=0}^{d} \frac{\langle SS^T, E_j\rangle}{m_j} E_j.$$

Since SS^T is a Schur idempotent, the non-zero coefficients in the first sum are all equal to 1. Hence $A_i \circ (SS^T)$ is either 0 or A_i. Since $\frac{1}{c}SS^T$ is a matrix idempotent, the non-zero coefficients in the second sum are all equal to c.

If

$$\mathscr{D} := \{i : A_i \circ (SS^T) \neq 0\},$$

then since

$$(SS^T)^2 = cSS^T,$$

it follows that if i and j belong to $\mathscr{D}$ and $p_{i,j}(k) \neq 0$, then $k \in \mathscr{D}$. Therefore the span of the set

$$\{A_i : i \in \mathscr{D}\}$$

is closed under multiplication. Each of these matrices is block diagonal, with blocks of size c. The matrices we get by taking the r-th block of A_i for each i in $\mathscr{D}$ form an association scheme, and the r association schemes we get in this way all have the same parameters. □

Suppose

$$\mathscr{E} := \{j : \langle S, S\rangle^T, E_j \neq 0\}.$$

Since SS^T is a Schur idempotent, it follows that if i and j belong to $\mathscr{E}$ and $q_{i,j}(k) \neq 0$, then $k \in \mathscr{E}$. So the span of the set

$$\{E_j : j \in \mathscr{D}\}$$

is closed under the Schur product.

If $j \in \mathscr{E}$ then

$$E_j = F_j \otimes \frac{c}{v} J_{v/c}.$$

Hence the F_j's are a set of pairwise orthogonal idempotents and

$$\sum_{j \in \mathscr{E}} F_j = I.$$

It follows that the matrices F_j (for $j \in \mathscr{E}$) form an association scheme whose vertices are the cells of σ.

5.4 Simple Subsets

Let $\mathscr{A} = \{A_0, \ldots, A_d\}$ be an association scheme with vertex set V and let C be a subset of V with characteristic vector x. We will work with the cyclic subspace generated by x, which is the subspace generated by the vectors

$$A_0 x, A_1 x, \ldots, A_d x.$$

We view this as a space of functions on V, and denote it by $\mathscr{D}(C)$.

The vectors

$$E_0 x, E_1 x, \ldots, E_d x$$

form a second basis for $\mathscr{D}$; since these vectors are orthogonal it follows that $\dim(\mathscr{D})$ is equal to the number of non-zero vectors in this set. If s^* denotes the dual degree of C, then we have

$$\dim(\mathscr{D}(C)) = 1 + s^*.$$

If $u \in V$, we define its *profile* relative to C to be the vector

$$(e_u^T A_0 x, \ldots, e_u^T A_d x)$$

The i-th entry of this is the number of vertices v in C such that u is i-related to v. We can partition the vertices of C by their profiles, and we call it the partition *induced* by C. Any element of $\mathscr{D}(C)$ is constant on the cells of this partition, and therefore if $\pi(C)$ is the partition induced by C,

$$\dim(\mathscr{D}(C)) \leq |\pi(C)|.$$

We call C a *simple* subset of V if equality holds.

Let $F(\pi)$ denote the space of functions on V that are constant on the cells of π. This space has dimension $|\pi|$ and contains $\mathscr{D}(\pi)$. This immediately yields the following.

5.4.1 Lemma. *Let C be a subset of the vertices of the association scheme $\mathscr{A}$ with induced partition π. Then C is simple if and only if $\mathscr{D}(C) = F(\pi)$.* □

Note also that $F(\pi)$ is $\mathbb{C}[\mathscr{A}]$-invariant if and only if π is equitable. Hence the partition induced by a simple subset is equitable. (The converse is false, although examples are not trivial to find.)

5.4.2 Example. If C is a linear code in $H(n,q)$, then the profile of u relative to C is the distance distribution of the coset $C + e_u$ of C.

5.4.3 Example. If C is just a single vertex u of $\mathscr{A}$, then

$$e_u^T E_j e_u > 0$$

for $j = 0, 1, \ldots, d$ and so the dual degree of $\{u\}$ is d. The partition of V induced by u is just its relation partition, which has $d+1$ cells. So $\{u\}$ is simple.

5.4.4 Example. Suppose $\mathscr{A}$ is imprimitive, and that C is the vertex set of a connected component of A_1.

5.5 Completely Regular Subsets

Suppose the association scheme $\mathscr{A}$ is metric with respect to A_1. If V is the vertex set of $\mathscr{A}$ and $C \subseteq V(\mathscr{A})$, we define C_i to be the subset of V consisting of the vertices that are at distance i (in X_1) from C. Thus $C_0 = C$. The sets C_i form a partition of V which we call the *distance partition*. The maximum distance of a vertex from C is the *covering radius* of C. The covering radius of C is t if and only if its distance partition has exactly $t+1$ cells. If the distance partition of C is equitable, we say C is *completely regular*. The canonical example is a perfect code in the Hamming scheme.

5.5.1 Lemma. *If $\mathscr{A}$ is a metric association scheme and C is a completely regular subset of $V(\mathscr{A})$, then C is simple.*

Proof.

5.5.2 Lemma. *If $\mathscr{A}$ is a metric association scheme and C is a completely regular subset of $V(\mathscr{A})$ with covering radius t and dual degree s, then $t \leq s$.*

Translation drgs correspond to cr codes in Hamming scheme.
$\delta \geq 2\delta - 1$ implies cr.
Lloyd's theorem.
Examples.
Do quotients schemes a là Godsil and Martin

Chapter 6

Translation Schemes

Suppose Γ is an abelian group of order v. The conjugacy class scheme on Γ is a scheme with $v-1$ classes, and each minimal Schur idempotent is a permutation matrix. Many interesting schemes arise as subschemes of these; they are known as translation schemes.

6.1 Characters

Let Γ be a finite abelian group. A *character* of Γ is a homomorphism from Γ into the multiplicative group formed by the non-zero complex numbers. The set of all characters of Γ is denoted by Γ, and is called the *character group* of Γ. If $\psi \in \Gamma^*$ and $g \in \Gamma$, then g^k for some integer k. Therefore

$$\psi(1) = \psi(g^k) = \psi(g)^k,$$

whence we see that $\psi(g)$ is a k-root of unity. It follows that

$$\psi(g^{-1}) = \overline{\psi(g)}.$$

The *trivial* character is the map that sends each element of Γ to 1. If φ and ψ are characters, we define the map $\varphi\psi$ by

$$\varphi\psi(g) := \varphi(g)\psi(g).$$

Using this definition it follows that Γ^* is an abelian group. If $\psi \in \Gamma^*$, then $\psi^{-1} = \overline{\psi}$.

To give an example, suppose $\Gamma = \mathbb{Z}_n$. Let θ be an n-th root of unity in $\mathbb{C}$ and let g be a generator for Γ. Then the map

$$g^k \mapsto \theta^k$$

is readily seen to be a character of Γ. Thus each n-th root of unity determines a character of Γ, and these characters form a subgroup of Γ^* with order n. For further progress, we need the following.

6.1.1 Lemma. *If ψ is a non-trivial character of the finite abelian group Γ, then*

$$\sum_{g\in\Gamma} \psi(g) = 0.$$

Proof. If $a \in G$ then

$$\sum_{g\in\Gamma} \psi(g) = \sum_{g\in\Gamma} \psi(ag) = \psi(a) \sum_{g\in\Gamma} \psi(g),$$

whence we see that if $\psi(a) \neq 1$, then $\sum_g \psi(g) = 0$. □

If $S \subseteq \Gamma$ and $\psi \in \Gamma^*$, we define

$$\psi(S) = \sum_{g\in S} \psi(g).$$

The previous result thus states that if ψ is not trivial, then $\psi(\Gamma) = 0$.

6.1.2 Corollary. *If φ and ψ are characters of Γ, then*

$$\sum_{g\in\Gamma} \varphi(g)\overline{\psi(g)} = \begin{cases} |\Gamma|, & \text{if } \varphi = \overline{\psi}; \\ 0, & \text{otherwise.} \end{cases}$$

Proof. Apply the lemma to the product $\varphi\overline{\psi}$. □

We define the sum in this corollary to be the *inner product* of φ and ψ; we see that distinct characters are orthogonal. It follows that the elements of Γ^* are linearly independent elements of the vector space $\mathbb{C}^G$ of complex-valued functions of Γ. Since this space has dimension $|\Gamma|$, we conclude that

$$|\Gamma^*| \leq |\Gamma|.$$

We can now show that Γ^* and Γ are isomorphic abelian groups. We saw above that $\mathbb{Z}_n^*$ contains a subgroup isomorphic to Γ, and therefore

$$\mathbb{Z}_n^* \cong \mathbb{Z}_n.$$

A finite abelian group is the direct product of cyclic groups. If A and B are finite abelian groups then we may assume inductively that

$$(A \times B)^* \cong A^* \times B^*,$$

and so our claim follows.

Let Γ be a finite abelian group of order n. A *character table* of Γ is the $n \times n$ matrix with ij-entry equal to the value of the i-character on the j-th element of Γ. By 6.1.2,

$$HH^* = nI.$$

Also

$$H \circ \overline{H} = J.$$

For example, the character table of $\mathbb{Z}_2^n$ may be taken to be the Kronecker product of n copies of

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

For another example, let Γ be $\mathbb{Z}_n$ and suppose η is a primitive n-th root of unity. The matrix P with rows and columns indexed by Γ and with

$$P_{i,j} = \eta^{ij}$$

is a character table for Γ. Since this is symmetric, any finite abelian group has a symmetric character table.

6.2 Translation Graphs

Let G be a group and suppose $C \subseteq G$. The *Cayley graph* $X(C)$ is the graph with vertex set G and arc set

$$\{(g,h) : hg^{-1} \in C\}.$$

Define

$$C^{-1} = \{c^{-1} : c \in C\}$$

and call C *inverse-closed* if $C = C^{-1}$. Then $X(C)$ is a directed graph if and only if C is not inverse-closed, and it will contain loops if $1 \in C$. We do not insist that Cayley graphs be undirected, but we do insist that they do not have loops.

If $a \in G$, let ρ_a be the map that sends x in G to xa. Then ρ_a is a permutation of G and an automorphism of $X(C)$. Hence G acts as a regular group of automorphisms of $X(C)$. Conversely, if G acts as a regular group of automorphisms of a graph X, we may choose a vertex v in X and define C to be the set of elements g of G such that (v, gv) is an arc in X. Then X is isomorphic to the Cayley graph $X(C)$.

We define a *translation graph* to be a Cayley graph for an abelian group. One advantage of translation graphs is that their eigenvalues and eigenvectors are more accessible, as we show now.

Suppose Γ is an abelian group of order v. Each character of G can be extended to a function on the subsets of Γ as follows. Suppose $\psi \in \Gamma^*$ and $S \subseteq \Gamma$. Then

$$\psi(S) := \sum_{g \in S} \psi(g).$$

6.2.1 Lemma. *Let X be a Cayley graph for the abelian group Γ, relative to the subset C. Each character ψ of Γ is an eigenvector for $A(X)$ with eigenvalue $\psi(C)$.*

Proof. A function ψ on $V(X)$ is an eigenvector if there is a complex number λ such that

$$\lambda\psi(g) = \sum_{h \sim g} \psi(h)$$

Since

$$\sum_{h \sim g} \psi(h) = \sum_{c \in C} \psi(cg) = \psi(g) \sum_{c \in C} \psi(c) = \psi(g)\psi(C),$$

we see that ψ is an eigenvector with eigenvalue $\psi(C)$.

6.3 Translation Schemes and their Duals

Let Γ be a finite abelian group of order v. Each element of Γ gives rise to a permutation of Γ—the permutation corresponding to a maps g in Γ to ga. Hence for each element g in Γ we have a permutation matrix $P(g)$; the map $g \mapsto P(g)$ is a group homomorphism. Therefore

$$P(g)P(h) = P(h)P(g), \qquad P(g^{-1}) = P(g)^T.$$

We have $P(1) = I$ and $\sum_g P(g) = J$. Hence the matrices $P(g)$ form an association scheme with $v-1$ classes. (This is in fact the conjugacy class scheme on Γ, but the description we have just presented may be more transparent.) We call it the *abelian group scheme* on Γ.

6.3.1 Lemma. *Let $\mathscr{A}$ be an association scheme with v vertices. Then $\mathscr{A}$ has $v-1$ classes if and only if it is the association scheme of an abelian group.*

Proof. Suppose $\mathscr{A}$ has v vertices and v classes $A_0, \ldots, A_{v-1}$. Since $\sum_i A_i = J$, we have $v_i = 1$ for each i. It follows that A_i is a permutation matrix, and that together they form an abelian group of order v. □

We define a *translation scheme* to be a subscheme of an abelian group scheme. The Hamming schemes and the bilinear forms schemes are translation schemes.

6.3.2 Example. Let $\mathbb{F}$ be a finite field of order q and suppose K is a subgroup of F^*, the multiplicative group of $\mathbb{F}$. The *cyclotomic scheme* has the elements of $\mathbb{F}$ as its vertices, and (u,v) is i-related if and only if $v-u$ lies in the i-th coset of K. Hence if $k = |K|$, this scheme has $(q-1)/k$ classes each of valency k. This scheme is symmetric if and only if $-1 \in K$. It is a translation scheme relative to the additive group of $\mathbb{F}$. If $q = p^n$ for some prime n, then the scheme is linear if and only if K contains $GF(p)^*$.

Let $\mathscr{A}$ be a subscheme of the scheme coming from the abelian group Γ. Then Γ acts by right multiplication as a group of permutations on itself, and thus Γ acts transitively on the vertices of $\mathscr{A}$. In particular, $\Gamma \le \mathrm{Aut}(X_i)$ for $i = 1, \ldots, d$ and therefore each X_i is a Cayley graph for Γ relative to a subset C_i. The sets C_i partition $\Gamma \setminus 1$ and are closed under inverses, that is, for each i we have $C_i^{-1} = C_j$ for some j.

The matrix of eigenvalues P of an abelian group scheme is the character table of the group. Thus the columns of P are indexed by the elements of the group, the rows by the characters and the ij-entry is the value of the i-th character on the j element. Assume π is a partition $C_0, \ldots, C_d$ of Γ such that $C_0 = \{1\}$ and the set of cells is inverse-closed. Let S be the characteristic matrix of π. Then by Theorem 5.2.2, the dimension of the algebra generated by the matrices $A_i = A(X(C_i))$ is equal to the number of distinct rows of PS. Further, by Corollary 5.2.3, is this dimension is $e+1$, then $e \ge d$ and equality holds if and only if $A_0, \ldots, A_d$ form an association scheme.

If equality holds then π determines a partition of Γ^* into $d+1$ cells, $D_0, \ldots, D_d$ say. It is not hard to show that one of these cells consists of the trivial charac-

ter, and that the set of cells in inverse-closed. Hence we obtain an association scheme on Γ^*. We call this scheme the *dual* of the scheme determined by π. Thus translation schemes come in dual pairs.

6.4 Linear Graphs

Let V be the vector space of dimension n over the field $\mathbb{F}$ of order q. Then V is an abelian group of order q^n and in this case the characters of V can be constructed from the characters of the additive group of the underlying field, as follows. If φ is a non-trivial character of the additive group of $\mathbb{F}$ and $a \in V$, define the map φ_a from V to $\mathbb{C}$ by

$$\varphi_a(x) = \varphi(a^T x).$$

Then φ_a is a character of V and, as a runs over the elements of V, we obtain all characters of V. The kernel of φ_a consists of the set of vectors x such that $a^T x = 0$, which we denote by $a^\perp$.

A Cayley graph $X(C)$ for the vector space V is *linear* if $0 \notin C$ and C is closed under multiplication by the non-zero elements of $\mathbb{F}$. (If C satisfies these conditions, we say it is *linear* subset of V.) If C is linear and $a \in V$, there is a simple expression for $\varphi_a(C)$ which we derive now.

Let $\mathbb{F}^*$ denote the set of non-zero elements of $\mathbb{F}$. If $a^T x \neq 0$, then the set

$$\{\lambda a^T x : \lambda \in \mathbb{F}^*\}$$

consists of the distinct non-zero elements of $\mathbb{F}$. Hence

$$\sum_{\lambda \in \mathbb{F}^*} \varphi_a(\lambda x) = \sum_{\lambda \in \mathbb{F}^*} \varphi(\lambda a^T x) = -1.$$

If $a^T x = 0$, this sum equals $q-1$, and so we conclude that

$$\varphi_a(C) = \frac{1}{q-1}(q|C \cap a^\perp| - |C|).$$

We note that if $\lambda \in \mathbb{F}^*$, then $(\lambda a)^\perp = a^\perp$, and so $\varphi_{\lambda a}(C) = \varphi_a(C)$ if $\lambda \in \mathbb{F}^*$.

6.4.1 Example. Let V be the vector space of dimension four over $\mathbb{Z}_2$ and suppose

$$C = \{e_1, e_2, e_3, e_4, e_1 + e_2 + e_3 + e_4\}.$$

Over $\mathbb{Z}_2$, any subset that does not contain 0 is linear. If $a = 0$ then $a^\perp = V$ and $\varphi_a(C) = 5$. If $a = e_1$ then

$$C \cap a^\perp = \{e_2, e_3, e_4\}$$

and so

$$\varphi_a(C) = 2 \times 3 - 5 = 1.$$

If $a = e_1 + e_2 + e_3 + e_4$ we find that

$$\varphi_a(C) = 2 \times 1 - 5 = -3.$$

You may check that these are all the eigenvalues of $X(C)$; hence it is strongly regular. (This is the Clebsch graph. It is an interesting exercise to show that the vertices at distance two from 0 induce a copy of the Petersen graph.)

6.5 Geometry, Codes and Graphs

Let V be the vector space of dimension d over the finite field $GF(q)$. The 1-dimensional subspaces of V are the points of the projective space $PG(d-1,q)$. Suppose $\Omega \subseteq PG(d-1,q)$. We can represent the set Ω by the columns of a $d \times |\Omega|$ matrix M whose columns are homogeneous coordinate vectors for the elements of Ω. We call the row space of M the *code* of Ω. The kernel of M is the *dual code* of Ω. We will usually denote the code of Ω by $C(\Omega)$, or by C. The dual code of C is $C^\perp$.

If $\Omega \subseteq PG(d-1,q)$, then $\langle\Omega\rangle$ denotes the smallest projective subspace that contains Ω. The *dimension* of Ω is the projective dimension of $\langle\Omega\rangle$; the *rank* $\mathrm{rk}(\Omega)$ is the dimension of the subspace of V corresponding to the projective subspace $\langle\Omega\rangle$. (The rank is one greater than the projective dimension.) We note that $\mathrm{rk}(\Omega)$ is equal to the dimension of its code.

Using the machinery we have just defined, we can translate geometric questions about Ω into questions about its code. However there is also a translation into graph theory. Suppose M is a matrix representing Ω. Let $X(\Omega)$ denote the Cayley graph for the additive group of V with the non-zero scalar multiples of the columns of M as its connection set. Thus X is a Cayley graph on q^d vertices, with valency $(q-1)|\Omega|$. It is connected if and only $\mathrm{rk}(M) = d$, and this holds if and only if $\mathrm{rk}(\Omega) = d$.

If C is a subspace of V, its *coset graph* is the graph with the cosets of C as its vertices, and the number of edges joining two cosets C_1 and C_2 is equal to the number of vectors in C_2 at Hamming distance one from a given vector in C_1. This definition allows a coset graph to have loops as well as multiple edges.

6.5.1 Lemma. *The coset graph of a code C is simple if and only if the minimum distance of C is at least three.* □

Note that the columns of M are distinct, and so the dual code of Ω has minimum distance at least three. (A code with minimum distance at least three is often called a *projective code*.)

6.5.2 Lemma. *If $\Omega \subseteq PG(d-1,q)$, then $X(\Omega)$ is the coset graph of the dual code of Ω.* □

There is also a direct geometric description of $X(\Omega)$. View $PG(d-1,q)$ as the hyperplane at infinity of the affine geometry $AG(d,q)$. The vertices of $AG(d,q)$ are the elements of V and its subspaces are the cosets of the linear subspaces of V. Construct a graph with vertex set V by defining two distinct points to be adjacent if the unique line through them meets the hyperplane at infinity in a point of Ω; this graph is $X(\Omega)$.

We will see that there are many interesting connections between the properties of Ω, its code C and its graph $X(\Omega)$. Before we can develop these, we need information about the eigenvalues and eigenvectors of X.

Let tr denote the trace map from the field $\mathbb{F}$ of order q to its prime field (of order p). If θ is a complex primitive p-th root of 1, then the map

$$x \mapsto \theta^{\mathrm{tr}(a^T x)}$$

is a character of the additive group of V, which we denote by ψ_a. If $a \in V$, then

$$a^\perp := \{x : a^T x = 0\}.$$

Usually we will view $a^\perp$ as a subset of $PG(d-1,q)$.

6.5.3 Lemma. *If $\Omega \subseteq PG(d-1,q)$ and ψ_a is as above, then ψ_a is an eigenvector for $X(\Omega)$ with eigenvalue $q|\Omega \cap a^\perp| - |\Omega|$.*

Proof. The connection set $\mathscr{C}$ of $X(\Omega)$ consists of the vectors

$$\gamma x,$$

where γ varies over the non-zero elements of $\mathbb{F}$ and x varies over the columns of M. Then

$$x \mapsto \mathrm{tr}(\gamma a^T x)$$

is a linear map from $\mathbb{F}$ to $GF(p)$. It is onto, and so takes each possible value exactly q/p times as γ varies over $\mathbb{F}$. Since the sum of the distinct powers of θ is zero,

$$\sum_{\gamma\in\mathbb{F}\setminus 0} \theta^{\mathrm{tr}(\gamma a^T x)} = \begin{cases} -1, & x \neq 0; \\ q-1, & x = 0. \end{cases}$$

Therefore $\psi_a(\mathscr{C}) = q|\Omega \cap a^\perp| - |\Omega|$. □

Geometrically $|\Omega \cap a^\perp|$ is the number of points of Ω that lie on the hyperplane of $PG(d-1,q)$ with coordinate vector a^T. If $\gamma \neq 0$, then

$$q|\Omega \cap a^\perp| = q|\Omega \cap (\gamma a)^\perp|,$$

whence we see that each hyperplane gives rise to $q-1$ eigenvectors for $X(\Omega)$, all with the same eigenvalue.

6.6 Language

In this section we develop a set of dictionaries, allowing us to translate between the languages of finite geometry, coding theory and graph theory.

We assume that Ω is a subset of $PG(d-1,q)$ with rank d and size m, represented by a matrix M. We denote the code of Ω by C and its graph by X.

Suppose H is a hyperplane in $PG(d-1,q)$, with coordinate vector h^T. The elements of $\Omega \cap h^T$ index the zero entries of $h^T M$. If $\mathrm{wt}(x)$ denote the weight of the code word x, then

$$|\Omega \cap h^T| = m - \mathrm{wt}(h^T M).$$

Thus a hyperplane of $PG(d-1,q)$ that intersects Ω in exactly i points determines $q-1$ code words of weight $m-i$, and $q-1$ eigenvectors of X with eigenvalue $qi-m$. In particular, the eigenvalues of X and their multiplicities are determined by the weight enumerator of the code of Ω.

6.6.1 Lemma. *Let Ω be a set of m points in $PG(d-1,q)$ and let τ be the least eigenvalue of $X(\Omega)$. Then $\tau \geq -m$, and equality holds if and only if the code of Ω contains a word of weight n.* □

6.6.2 Theorem. *Let Ω be a set of n points in $PG(d-1,q)$ with code C Then $X(\Omega)$ is q-colourable if and only if $C^\perp$ contains a word of weight n.*

Proof. If there is no word of weight n in $C^\perp$, then the least eigenvalue of $X(\Omega)$ is greater than $-n$. The valency of $X(\Omega)$ is $n(q-1)$ and so the ratio bound yields that

$$\alpha(X(\Omega)) < \frac{|V(X)|}{1+\frac{n(q-1)}{n}} = \frac{|V(X)|}{q}.$$

Hence $\chi(X(\Omega)) > q$.

Conversely, let M be a matrix that represents Ω and suppose $a^T M$ is a word of weight n in the code of Ω. If x and y are vertices of $X(\Omega)$ and $a^T x = a^T y$, then $a^T(x-y) = 0$ and therefore x and y are not adjacent in $X(\Omega)$. Hence the map $x \mapsto a^T x$ is a proper colouring of $X(\Omega)$ using the elements of $\mathbb{F}$. □

6.6.3 Corollary. *Let Ω be a set of points in $PG(d-1,q)$. To determine the least eigenvalue of $X(\Omega)$ from Ω is NP-hard.*

Proof. Take M to be the incidence matrix of an orientation of a graph Y. If $a^T M$ has no zero entries, the vector a determines a proper colouring of Y with q colours. If $q = 3$, then Y is 3-colourable if and only if the code over $GF(3)$ generated by M contains a word of weight n. Hence $X(M)$ is 3-colourable if and only if Y is 3-colourable. Since it is NP-hard to decide if a graph is 3-colourable, we are done. □

We also see that it is NP-hard to decide if the adjacency matrix of a Cayley graph for $\mathbb{Z}_2^n$ is invertible (over $\mathbb{R}$).

The connection between eigenvalues of the coset graph and the weight distribution of the code appears to be folk-lore. Some information appears in Delorme and Solé (European J. Comb. 12 (1991)) [***but I have not checked this yet***].

The *covering radius* of a code C is the least integer r such that every word is at distance at most r from a word of C.

6.6.4 Lemma. *The covering radius of $C^\perp(\Omega)$ is equal to the diameter of X.* □

A *cap* in projective space is a set of points such that no three are collinear.

6.6.5 Lemma. *Suppose $\Omega \subseteq PG(d-1,q)$. Then the following are equivalent:*

(a) Ω is a cap.

(b) The minimum distance of $C^\perp$ is at least four.

(c) $X(\Omega)$ is triangle-free. □

Chapter 7

Duality

7.1 The Discrete Fourier Transform

The set $\mathscr{C}_n$ of $n \times n$ circulants over $\mathbb{F}$ is closed under matrix and Schur multiplication and contains I and J, the units for these multiplications. (Thus it is the Bose-Mesner algebra of the association scheme of the cyclic group of order n.) We introduce an important endomorphism of this algebra.

Let $\mathbb{E}$ be an extension field of $\mathbb{F}$ that contains a primitive n-th root of unity. Equivalently, $\mathbb{E}$ is a splitting field for $t^n - 1$. Let θ be a fixed n-th root of unity in $\mathbb{E}$. If $M = p(R)$, define

$$\Theta(M) = \sum_{i=0}^{n-1} p(\theta^i) R^i.$$

Thus Θ is an endomorphism, a linear operator on $\mathscr{C}_n$. We call it a *duality map*.

7.1.1 Lemma. *If $M \in \mathscr{C}_n$ then $\Theta^2(M) = nM^T$.*

Proof. It is enough to show that $\Theta^2(R^k) = nR^T$. We have

$$\begin{aligned}\Theta^2(R^k) = \sum_j \theta^{kj} \Theta(R^j) &= \sum_{i,j} \theta^{kj} \theta^{ij} R^i \\ &= \sum_i \Big(\sum_j \theta^{j(i+k)} \Big) R^i.\end{aligned}$$

The inner sum is zero unless $i = -k$, when it is n. Therefore $\Theta^2(R^k) = R^{-k}$ and since $R^{-1} = R^T$, the result follows. □

7.1.2 Theorem. *If $M, N \in \mathscr{C}_n$ then $\Theta(MN) = \Theta(M) \circ \Theta(N)$ and $\Theta(M \circ N) = \frac{1}{n}\Theta(M)\Theta(N)$.*

Proof. We have

$$\Theta(p(R)q(R)) = \sum_i p(\theta^i)q(\theta^i)R^i = \Big(\sum_i p(\theta^i)R^i\Big) \circ \Big(\sum_i q(\theta^i)R^i\Big),$$

which is the first claim. The second follows from this and the previous lemma. □

7.1.3 Theorem. *If $M^T = \sum_v \mu_i R^i$, then $M\Theta(R^i) = \mu_i \Theta(R^i)$.*

Proof. We have

$$\begin{aligned} M\Theta(R^i) &= v^{-1}\Theta^2(M^T)\Theta(R^i) \\ &= \Theta(\Theta(M^T) \circ R^i) \\ &= \Theta(\mu_i R^i) \\ &= \mu_i \Theta(R^i). \end{aligned}$$ □

It follows from this that the entries of $\Theta(M)$ are eigenvalues of M, and the columns of $\Theta(R_i)$ are eigenvectors for all circulants.

Define the *weight* of a circulant to be the number of non-zero entries in a column.

7.1.4 Lemma. *If $\deg(q(t)) = \ell$, then $\Theta(q(R))$ has weight at least $n - \ell$.*

Proof. If $\deg(q(t)) = \ell$, then at most ℓ distinct powers of θ are zeros of q and so $\Theta(q(R))$ has at most ℓ zero entries in any column. □

The following result is the BCH-bound.

7.1.5 Theorem. *If $M = \varphi(R)$ and $\varphi(t)$ vanishes on k consecutive powers of θ, the minimum distance of the column space of M is at least $k+1$.*

Proof. Suppose $M = p(R)$. If $p(t)$ has k consecutive powers of θ as zeros, then $\Theta(M)$ has k cyclically consecutive zeros in its first column. Hence there is an integer s such that last k entries in $R^s\Theta(M)$ are zero, and therefore there is a polynomial $q(t)$ with degree at most $n-1-k$ such that

$$R^s\Theta(M) = q(R).$$

Consequently

$$\Theta(q(R)) = \Theta(R^s) \circ \Theta^2(M)$$

has weight at least $k+1$. Since $\Theta(R^s) = \Theta(R)^{\circ s}$ has no zero entries and $\Theta^2(M) = M^T$, it follows that M has weight at least $k+1$.

If $g(t)$ is a polynomial, then $g(R)M = g(R)p(R)$ and $g(t)p(t)$ vanishes on k consecutive powers of θ. Therefore $g(R)M$ has weight at least $k+1$, for any polynomial g. This implies that the minimum weight of the column space of M is at least $k+1$. □

M is diagonalisable if and only $n \cdot 1 \neq 0$ in $\mathbb{F}$.

The subset $\{0,3,4,9,11\}$ in $\mathbb{Z}_{21}$ is a cyclic difference set for a projective plane of order four. Hence if

$$\psi(t) = 1 + t^3 + t^4 + t^9 + t^{11}$$

then $N = p(R)$ is the incidence matrix of a plane of order four. Since $\deg(p) = 11$, we see that $\mathrm{rk}(N) \geq 10$ over $\mathbb{Z}_2$. We can check though that ψ divides $t^{21} - 1$: in fact

$$(t-1)\psi(t)\psi^*(t) = t^{21} - 1$$

and consequently $\mathrm{rk}(N) = 10$.

7.2 The Hadamard Transform

In the previous we worked with a duality related to the cyclic group. Here we introduce an analogous duality map related to the elementary abelian group $\mathbb{Z}_2^n$. It may help to view this as the additive group of a vector space of dimension n over $\mathbb{Z}_2$.

When working with the cyclic group we used circulant matrices, which are linear combinations of the powers of R, where R is a cyclic permutation matrix. We introduce the analagous matrices for $\mathbb{Z}_2^n$. First define a matrix P

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

If $u \in \mathbb{Z}_2$, define A_u to be the Kronecker product

$$A_u := P^{u_1} \otimes \cdots \otimes P^{u_n}.$$

Then $A_0 = I$,

$$A_u A_v = A_{u+v}$$

(and in particular $A_u^2 = I$). It follows that the map

$$u \mapsto A_u$$

is a group isomorphism. A simple induction argument on n yields that

$$\sum_u A_u = J.$$

(Partition the sum over the vectors u such that $u_1 = 0$ and the vectors u with $u_1 = 1$.)

It follows that the matrices in

$$\mathscr{A} := \{A_u : u \in \mathbb{Z}_n\}$$

are linearly independent. Define $\mathbb{F}[\mathscr{A}]$ to be vector space over $\mathbb{F}$ spanned by the matrices in $\mathscr{A}$. (For our purposes here, $\mathbb{F} = \mathbb{R}$ will suffice.)

If $u \in \mathbb{Z}_n$, define the function $\psi_u : \mathbb{Z}_n \to \{-1, 1\}$ by

$$\psi_u(v) = (-1)^{u^T v}.$$

Define a duality map Θ on $\mathbb{F}[\mathscr{A}]$ by setting

$$\Theta(A_u) = \sum_v \psi_u(v) A_v$$

and extend Θ to $\mathbb{F}[\mathscr{A}]$ by linearity. We have at once that

$$\Theta(I) = J.$$

Then

$$\begin{aligned}\Theta(A_u)\Theta(A_v) &= \sum_x \sum_y \psi_u(x)\psi_v(y) A_x A_y \\ &= \sum_{x,y} (-1)^{u^T x + v^T y} A_{x+y} \\ &= \sum_{x,x+y} (-1)^{(u-v)^T x} (-1)^{v^T (x+y)} A_{x+y}.\end{aligned}$$

Since

$$\sum_x (-1)^{(u-v)^T x} = \begin{cases} 2^n, & u = v; \\ 0, & \text{otherwise} \end{cases}$$

we conclude that

$$\Theta(A_u)\Theta(A_v) = \delta_{u,v} 2^n \Theta(A_u).$$

Consequently, for all M and N in $\mathbb{F}[\mathscr{A}]$,

$$\Theta(M)\Theta(N) = 2^{-n}\Theta(M \circ N).$$

We also have $\Theta(A_u) \circ \Theta(A_v) = \Theta(A_{u+v})$, whence $\Theta(A_u A_v) = \Theta(A_u) \circ \Theta(A_v)$ and

$$\Theta(MN) = \Theta(M) \circ \Theta(N).$$

Next

$$\begin{aligned} A_u\Theta(A_v) = A_u \sum_w \psi_v(w) A_w &= \sum_w (-1)^{v^T w} A_{u+w} \\ &= (-1)^{v^T u} \sum_w (-1)^{v^T (u+w)} A_{u+w} \\ &= \psi_u(v)\Theta(A_v) \end{aligned}$$

which shows that the columns of $\Theta(A_v)$ are eigenvectors for A_u. Moreover, we see that the entries of $\Theta(M)$ are the eigenvalues of M.

We leave the proof of the next result as an exercise.

7.2.1 Theorem. *If $M \in \mathbb{F}[\mathscr{A}]$, then $\Theta^2(M) = 2^n M$.* □

Since $\Theta(I) = J$, it follows that $\Theta(J) = 2^n I$. The proof of Theorem 7.1.3 is easily modified to yield our next result.

7.2.2 Theorem. *If $M \in \mathbb{F}[\mathscr{A}]$, then the entries of $\Theta(M)$ are the eigenvalues of M.* □

7.2.3 Lemma. *If $M \in \mathbb{F}[\mathscr{A}]$, then* $\operatorname{tr}(\Theta(M)) = \operatorname{sum}(M)$.

Proof. Let ρ denote the sum of a row of M. We have

$$\begin{aligned} I \circ \Theta(M) &= 2^{-n}\Theta^2(I) \circ \Theta(M) \\ &= 2^{-n}\Theta(\Theta(I)M) \\ &= 2^{-n}\Theta(JM) \\ &= 2^{-n}\Theta(\rho J) \\ &= \rho I \end{aligned}$$

Therefore $\operatorname{tr}(\Theta(M)) = \operatorname{sum}(M)$. □

7.3 Two Matrix Duals

Let C be a linear code of length n and let a_i denote the number of words in C of weight i. The *weight enumerator* $W_C(x,y)$ is the polynomial

$$W_C(x,y) = \sum_{i=0}^{n} a_i x^{n-i} y^i.$$

It is a surprising fact that $W_{C^\perp}$ can be obtained from $W_C(x,y)$, and in a simple way.

If C is a linear binary code of length n, define the matrix A_C by

$$A_C := \sum_{u \in C} A_u.$$

7.3.1 Lemma. *If C is a binary linear code, then $\Theta(A_C) = |C| A_{C^\perp}$.*

Proof. If β is a basis for C, then

$$\prod_{u \in \beta} (I + A_u) = A_C.$$

and accordingly $\Theta(A_C)$ is the Schur product of the matrices

$$\Theta(I + A_u) = J + \Theta(A_u),$$

where u runs over β. Now

$$J + \Theta(A_u) = \sum_v (1 + (-1)^{u^T v}) A_v = 2 \sum_{v \in u^\perp} A_v$$

and therefore the Schur product of the matrices $J + \Theta(A_u)$ is $2^{|\beta|} A_{C^\perp}$, as required. □

Let K be the matrix

$$K := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

If M is a matrix $M^{\otimes n}$ denotes the Kronecker product of n copies of M.

7.3.2 Lemma. *We have*

$$\sum_u x^{n-\mathrm{wt}(u)} y^{\mathrm{wt}(u)} A_u = (xI + yK)^{\otimes n}.$$

Proof. Let $e_1, \ldots, e_n$ denote the standard basis for $\mathbb{Z}_2^n$. If $u \in \mathbb{Z}_2^n$, then

$$u = \sum_i u_i e_i.$$

Then

$$A_u = K^{u_1} \otimes \cdots \otimes K^{u_n}$$

and so $x^{n-\mathrm{wt}(u)} y^{\mathrm{wt}(u)} A_u$ is the Kronecker product of the n terms $x^{1-u_i} y^{u_i} K^{u_i}$ for $i = 1, \ldots, n$. This implies the lemma. □

7.3.3 Lemma. *We have* $\Theta(M \otimes N) = \Theta(M) \otimes \Theta(N)$.

Proof. The entries of $\Theta(M) \otimes \Theta(N)$ are the products of the entries of $\Theta(M)$ and $\Theta(N)$, and these are the eigenvalues of M and N. The products of the eigenvalues of M and N are the eigenvalues of $M \otimes N$, and these are the entries of $\Theta(M \otimes N)$. [We have neglected some bookkeeping, you are welcome to supply it. :-)] □

7.3.4 Corollary. *We have*

$$\Theta\Big(\sum_u x^{n-\mathrm{wt}(u)} y^{\mathrm{wt}(u)} A_u\Big) = \sum_u (x+y)^{n-\mathrm{wt}(u)} (x-y)^{wt(u)} A_u.$$

Proof. We have $\Theta(I) = J$. Since $K = J - I$,

$$\Theta(K) = \Theta(J) - \Theta(I) = 2I - J.$$

Therefore

$$\Theta(xI + yK) = xJ + 2yI - yJ = (x-y)(J-I) + (x+y)I = (x+y)I + (x-y)K.$$

We now obtain the result by applying Lemmas 7.3.2 and 7.3.3. □

7.4 MacWilliams Theorem

We apply the results from the previous section to derive MacWilliams theorem, a fundamental result in Coding Theory.

7.4.1 Theorem. *Let C be a binary linear code of length n. Then*

$$W_{C^\perp}(x,y) = \frac{1}{|C|} W_C(x+y, x-y).$$

Proof. Set M equal to $\sum_u x^{n-\mathrm{wt}(u)} y^{\mathrm{wt}(u)} A_u$. Then the diagonal entries of $A_{C^\perp} M$ are each equal to $W_{C^\perp}(x,y)$, whence

$$\mathrm{tr}(A_{C^\perp} M) = 2^n W_{C^\perp}(x,y).$$

Using Lemma 7.2.3, we have

$$\begin{aligned}\mathrm{tr}(A_{C^\perp} M) &= 2^{-n}\,\mathrm{tr}(\Theta^2(A_{C^\perp} M)) \\ &= 2^{-n}\,\mathrm{sum}(\Theta(A_{C^\perp} M)) \\ &= 2^{-n}\,\mathrm{sum}(\Theta(A_{C^\perp}) \circ \Theta(M)) \\ &= 2^{-n}|C^\perp|\,\mathrm{sum}(A_C \circ \Theta(M)) \\ &= |C|^{-1}\,\mathrm{sum}(A_C \circ \Theta(M)).\end{aligned}$$

Since the row sum of $A_C \circ \Theta(M)$ is $W_C(x+y, x-y)$, the last term above is equal to $2^n|C|^{-1} W_C(x+y, x-y)$, and so the theorem is proved. □

By way of example, suppose C is the code of the plane of order four. Our computations in Section **??** yield that the weight enumerator of C is

$$x^{21} + 21x^{16}y^5 + 210x^{13}y^8 + 280x^{12}y^9 + 280x^9y^{12} + 210x^8y^{13} + 21x^5y^{16} + y^{21}.$$

Using MacWilliams theorem, we find the weight enumerator of the dual is

$$x^{21} + 168x^{15}y^6 + 210x^{13}y^8 + 1008x^{11}y^{10} + 280x^9y^{12} + 360x^7y^{14} + 21x^5y^{16}.$$

7.4.2 Theorem. *The length of a doubly even binary self-dual code is divisible by 8.*

Proof. If C is self-dual with length n, then $|C| = 2^{n/2}$ and

$$W_C(x,y) = 2^{-n/2} W_C(x+y, x-y) = W_C\Big(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\Big).$$

Therefore $W_C(x,y)$ is invariant under the substitution represented by the matrix

$$\tau = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Since C is doubly even, it is also invariant when we replace y by iy (with $i^2 = -1$). Equivalently it is invariant under the substitution represented by

$$\sigma = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}.$$

We find that

$$(\tau\sigma)^3 = \frac{1+i}{\sqrt{2}} I.$$

Hence if $\theta := (1+i)/\sqrt{2}$, the substitution

$$x \mapsto \theta x, \quad y \mapsto \theta y$$

leaves $W_C(x,y)$ invariant. But

$$W_C(\theta x, \theta y) = \theta^n W_C(x,y)$$

and as θ is a primitive 8-th root of unity, this implies that $8 \mid n$. □

7.5 Projective Planes

We use the theory at hand to prove that there is no projective plane of order n, where $n \equiv 6$ modulo 8.

We work with linear codes over $GF(2)$. A code is *even* if all its word have even weight, and it is *doubly even* if all words have weight divisible by four. If C is a binary code of length n, the *extended code* by adding an $(n+1)$-th coordinate to each code word, such that the weight of the extended code word is even. (Thus we are adding a parity check; the operation is trivial if C is even.)

7.5.1 Theorem. *Let N be the incidence matrix of a projective plane with order n and let C be the linear code spanned by the rows of N over $GF(2)$. Then the extended code is self-dual and doubly even.*

Proof. Let N be the incidence matrix of our projective plane. Let N_1 denote the matrix we get by adding a final column equal to $\mathbf{1}$ to N. Since n is even and since equal row of N has weight $n+1$, the rows of N_1 have even weight. One consequence is that each word in row(N) has even weight.

Further

$$NN^T = nI + J$$

and hence

$$N_1 N_1^T = (nI + J) + J = 0 \mod 2.$$

It follows that the code generated by the rows of N_1 is self-orthogonal. As $n \equiv 2$ modulo four, each row of N_1 has weight divisible by four, whence it follows that all code words in row(N_1) have weight divisible by four.

Each row of N_1 has length n^2+n+2, and it remains for us to show that

$$\mathrm{rk}(N_1) = \frac{1}{2}(n^2+n+2).$$

Since $\mathbf{1}$ lies in $\mathrm{col}(N)$ over $GF(2)$, we see that N_1 and N have the same rank. We will therefore compute $\mathrm{rk}(N)$.

Let $v = n^2+n+1$ and let H be a parity check matrix for C—in other words, H is a binary matrix with linearly independent rows such that $NH^T = 0$ and

$$\mathrm{rk}(N) + \mathrm{rk}(H) = v.$$

(Or to put it yet another way, the rows of H are a basis for $\ker(N)$.) Permuting columns of N and H if needed, we may assume that H has the form

$$\begin{pmatrix} I_r & K \end{pmatrix}$$

where $r = \mathrm{rk}(H)$. Let H_1 be given by

$$H_1 = \begin{pmatrix} I_r & K \\ 0 & I_{v-r} \end{pmatrix}.$$

Now view N and H_1 as 01-matrices over $\mathbb{Q}$.

Since $\det(H_1) = 1$, we have

$$\det(N) = \det(NH_1^T).$$

Since $NH^T = 0$ modulo two, each entry in the first r columns of NH_1^T is even, and therefore 2^r divides $\det(N)$. Now

$$NN^T = nI + J,$$

from which it follows that

$$\det(N) = (n+1)n^{n(n+1)/2}.$$

As both $n+1$ and $n/2$ are odd, we conclude that $r \le n(n+1)/2$. This implies that

$$\mathrm{rk}(N) = v - r \ge \frac{1}{2}(n^2+n+2);$$

since $\mathrm{rk}(N_1) = \mathrm{rk}(N)$ and since $\mathrm{row}(N_1)$ is self-orthogonal,

$$\mathrm{rk}(N_1) = (n^2+n+2)/2. \qquad \square$$

If $n \equiv 6$ modulo eight, then $n^2+n+2 \equiv 4$ modulo eight. Consequently by Theorem 7.4.2, there is no binary doubly even self-dual code of this length. Thus we have the following result.

7.5.2 Corollary. *If $n \equiv 6$ modulo eight, there is no projective plane of order n.*

This condition is weaker than the Bruck-Ryser-Chowla theorem, but certainly easier to use.

7.6 Duality

We say an association scheme $\mathscr{A}$ is *formally self-dual* if $Q = \overline{P}$.

If $i \in \{0,1,\ldots,d\}$, we define i^T to be the element of $i \in \{0,1,\ldots,d\}$ such that $A_{i^T} = A_i^T$. We recall that $p_j(k) = \overline{p_j(k^T)}$.

7.6.1 Theorem. *Let $\mathscr{A}$ be an association scheme on v vertices such that $\overline{Q} = P$ and let Θ be the linear mapping from $\mathbb{C}[\mathscr{A}]$ to itself such that $\Theta(A_i) = \sum_j p_i(j)A_j$. Then:*

(a) $\Theta(A_i) = v\overline{E_i}$.

(b) $\Theta(I) = J$, $\Theta(J) = vI$.

(c) $\Theta(MN) = \Theta(M) \circ \Theta(N)$ *for all M and N in $\mathbb{C}[\mathscr{A}]$.*

(d) $\Theta(M \circ N) = \frac{1}{v}\Theta(M)\Theta(N)$ *for all M and N in $\mathbb{C}[\mathscr{A}]$.*

(e) If $\mathscr{B}$ is a subscheme of $\mathscr{A}$, then $\Theta(\mathscr{B})$ is also a subscheme.

Proof. Since $\overline{p_i(j)} = q_i(j)$, we have

$$\Theta(A_i) = \sum_{j=0}^{d} \overline{q_i(j)}A_j = v\overline{E_i}.$$

In particular, $\Theta(I) = J$.

Next

$$\Theta(v\overline{E_i})) = \sum_j \overline{q_i(j)}\Theta(A_j) = \sum_{j,k} \overline{q_i(j)}p_j(k)A_k = \sum_{j,k} q_i(j)p_j(k^T)A_k^T.$$

Since $QP = vI$, it follows that

$$\Theta(vE_i) = vA_i^T.$$

Hence

$$\Theta^2(M) = vM^T \tag{7.6.1}$$

for all M in $\mathbb{C}[\mathscr{A}]$. (Note that $\Theta(J) = vI$.)

Since the entries of $\Theta(A_i)$ are the eigenvalues of A_i, we see that $\Theta(A_iA_j) = \Theta(A_i) \circ \Theta(A_j)$ and hence

$$\Theta(MN) = \Theta(M) \circ \Theta(N), \tag{7.6.2}$$

for all M and N in $\mathbb{C}[\mathscr{A}]$.

Finally

$$\Theta(A_i \circ A_j) = \delta_{i,j} v\overline{E_i} = \frac{1}{v}\Theta(A_i)\Theta(A_j).$$

and thus

$$\Theta(M \circ N) = \frac{1}{v}\Theta(M)\Theta(N). \tag{7.6.3}$$

for all M and N in $\mathbb{C}[\mathscr{A}]$. □

If Θ is a map satisfying the conditions of this theorem, we call it a *duality map*. The matrix representing Θ relative to the basis $A_0, \ldots, A_d$ is P.

Suppose $\mathscr{A}$ is the scheme of the cyclic group of order v. If θ is a primitive v-th root of 1 then we may assume that

$$P_{i,j} = \theta^{(i-1)(j-1)}. \tag{7.6.4}$$

It is easy to verify that $P\overline{P} = vI$, so this scheme is formally self-dual. The map Θ is essentially the discrete Fourier transform. We may take θ from any field $\mathbb{F}$ that contains a primitive v-th root of 1, and thus we may define Θ on $\mathbb{F}[\mathscr{A}]$.

It seems reasonable to define an association scheme on v vertices to be *self-dual* if there is an endomorphism Θ of $\mathrm{Mat}_{n\times n}(\mathbb{C})$ such that $\Theta(A_i) = v\overline{E_i}$ for $i = 0, 1, \ldots, d$.

If $\mathscr{A}$ and $\mathscr{B}$ are schemes and the matrix of eigenvalues of $\mathscr{B}$ is the complex conjugate of the matrix of dual eigenvalues of $\mathscr{A}$, we say that $\mathscr{A}$ and $\mathscr{B}$ are *formally dual*. In this case we can define a map Θ as above, and a slightly modified version of Theorem 7.6.1 still holds. If Θ is induced by an endomorphism of $\mathrm{Mat}_{n\times n}(\mathbb{C})$, we say the pair of schemes is *dual*.

De Caen observed that if $\mathscr{A}$ and $\mathscr{B}$ are dual, then the product scheme $\mathscr{A} \otimes \mathscr{B}$ is self-dual. Hence we may choose to view self-duality as the fundamental concept.

Each translation scheme is either self-dual or has a distinct dual translation scheme. The only known examples of dual pairs of non-isomorphic schemes arise in this way. The Higman-Sims scheme is self-dual and is not a translation scheme.

7.7 Duality and Type II Matrices

Consider symmetric formally self-dual schemes. Then P is real and $P^2 = vI$, whence it follows that the eigenvalues of P (or Θ) are $\pm\sqrt{v}$. If we know $\mathrm{tr}(P)$, we can compute the multiplicities of these eigenvalues. If $M \in \mathbb{C}[\mathscr{A}]$ then

$$\Theta(\sqrt{v}M + \Theta(M)) = \sqrt{v}\Theta(M) + vM = \sqrt{v}(\sqrt{v}M + \Theta(M).$$

Thus we have an eigenvector for Θ with eigenvalue $\sqrt{v}$. Similarly $\sqrt{v}M - \Theta(M)$ is an eigenvector for Θ with eigenvalue $-\sqrt{v}$.

7.7.1 Theorem. *Let Θ be a duality map on the symmetric association scheme $\mathscr{A}$ on v vertices, and let M be an eigenvector for Θ with eigenvalue $\sqrt{v}$. The following assertions are equivalent:*

(a) $\Theta(M^{-1}) = \sqrt{v}M^{-1}$.

(b) $\Theta(M^{(-)}) = \sqrt{v}M^{(-)}$.

(c) $vM^{-1} = M^{(-)}$ *(and M is a type II matrix).*

Proof. Assume $\Theta(M) = \sqrt{v}M$. Then we have

$$J = \Theta(I) = \Theta(MM^{-1}) = \Theta(M) \circ \Theta(M^{-1}) = \sqrt{v}M \circ \Theta(M^{-1}).$$

Hence $M^{(-)} = \sqrt{v}\Theta(M^{-1})$. Now M is type II if and only if $vM^{-1} = M^{(-)}$ and so (a) and (c) are equivalent.

Next

$$\Theta(M)\Theta(M^{(-)}) = v\Theta(M \circ M^{(-)}) = v\Theta(J) = v^2 I$$

and therefore

$$v^{-3/2}\Theta(M^{(-)}) = M^{-1}.$$

Hence (b) and (c) are equivalent. □

The matrix of eigenvalues of the scheme of an abelian group Γ can be taken to be the Kronecker product of matrices of the form given by (7.6.4). Hence these schemes are also formally self-dual. Perhaps the most interesting case is when $\Gamma = \mathbb{Z}_2^m$. In this case a matrix M is a *bent function* if both M and $v^{-1/2}\Theta(M)$ are ± 1-matrices. Note that if M is a ± 1-matrix, then it is type II if and only if $v^{-1/2}\Theta(M)$ is a ± 1-matrix. Hence bent functions are the ± 1-matrices which are eigenvectors for Θ.

7.8 Difference Sets

Let $\mathscr{A}$ denote an association scheme on v vertices. A *difference set* in $\mathscr{A}$ is 01-matrix A such that

$$AA^T = nI + \lambda J$$

for some positive integers n and λ. Hence A is an incidence matrix for a symmetric design. It is easy to verify that if A is a difference set then so is $J - A$, and thus we may assume $A \circ I = 0$ is we like. If k is the row sum of A, then $n = k - \lambda$.

Consider the case where A is a difference set and $A = A^T$. Then the squares of the eigenvalues of A are $\lambda v + n$ and n. If k denotes the row sum of A, then k is an eigenvalue of A and

$$k^2 = \lambda(v-1) + k;$$

the remaining eigenvalues of A are $\pm\sqrt{n}$. If $\mathrm{tr}(A) = 0$, there are positive integers a and b such that $1 + a + b = v$ and

$$k + a\sqrt{n} - b\sqrt{n} = \mathrm{tr}(A) = 0.$$

Accordingly

$$b - a = \frac{k}{\sqrt{k-\lambda}},$$

from which it follows that $k - \lambda$ is a perfect square. Since A has exactly three distinct eigenvalues, it is the adjacency matrix of a strongly regular graph with $a = c$.

The case where A is not symmetric is more complex. Since A lies in the Bose-Mesner algebra of the scheme, $AA^T = A^T A$ and therefore A is normal. A normal matrix is symmetric if and only if its eigenvalues are real, consequently some eigenvalues of A are complex. The valency aside, all eigenvalues of A have absolute value $\sqrt{k-\lambda}$. The matrix A is still an incidence matrix of a symmetric design.

Suppose A is a 01-matrix in $\mathscr{A}$ with each row summing to k. Since A is normal, $A = LDL^*$ where L is unitary. Hence

$$A^T = A^* = L\overline{D}L^*$$

and thus if $Az = \theta z$, then $A^T z = \overline{\theta} z$ and $AA^T z = |\theta|^2 z$. If the valency k is a simple eigenvalue of A and its remaining eigenvalues each have absolute value $\sqrt{k-\lambda}$, then $AA^T - (k-\lambda)I$ has rank one. It follows that A is a difference set.

Classical difference sets arise as difference sets in the association scheme of an abelian group Γ. In this case we can view the first row of A as the characteristic function of a subset S of Γ, and the eigenvalues are the complex numbers

$$\psi(S) + \sum_{g \in \Gamma} \psi(g).$$

Chapter 8

Type-II Matrices

In this chapter we present one of the more unusual constructions of association schemes. Its main weakness is that the actual examples is provides can readily be obtained by other methods. But it is closely connected to Vaughan Jones's construction of link invariants, and provides an interesting viewpoint on duality.

8.1 Type-II Matrices

If $M \circ N = J$ we say that N is the *Schur inverse* of M, and denote it $M^{(-)}$. A *type-II matrix* is a Schur invertible $n \times n$ matrix W over $\mathbb{C}$ such that

$$WW^{(-)T} = nI.$$

This condition implies that W^{-1} exists and

$$W^{(-)T} = nW^{-1}.$$

We consider some examples. First

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

is a symmetric type-II matrix. If ω is a primitive cube root of unity then

$$\begin{pmatrix} 1 & 1 & \omega \\ \omega & 1 & 1 \\ 1 & \omega & 1 \end{pmatrix}$$

is also type II. For any complex number t, the matrix

$$W = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & t & -t \\ 1 & -1 & -t & t \end{pmatrix}$$

is type II. Next we have the *Potts models*: if W is $n \times n$ and

$$W = (t-1)I + J,$$

then

$$\begin{aligned} WW^{(-)T} &= ((t-1)I + J)((t^{-1}-1)I + J) \\ &= ((2 - t - t^{-1})I + (n - 2 + t + t^{-1})J, \end{aligned}$$

whence it follows that W is type II whenever $2 - t - t^{-1} = n$, i.e., whenever t is a root of the quadratic

$$t^2 + (n-2)t + 1.$$

As the first example suggests, any Hadamard matrix is a type-II matrix, and it is not unreasonable to view type-II matrices as a generalization of Hadamard matrices.

The Kronecker product of two type-II matrices is a type-II matrix; this provides another easy way to increase the supply of examples. Recall that a *monomial matrix* is the product of a permutation matrix and a diagonal matrix. It is straighforward to verify that if W is type-II and M and N are invertible monomial matrices, then MWN is type II. We say W' is *equivalent* to W if $W' = MWN$, where M and N are invertible monomial matrices.

The transpose W^T is also type II, as is $W^{(-)}$, but these may not be equivalent to W. It would be a useful exercise to prove that any 2×2 type-II matrix is equivalent to the first example above, any 3×3 type-II matrix is equivalent to the second, and any 4×4 type-II matrix is equivalent to a matrix in the third family.

Type-II matrices play a role in the study of von Neumann algebras, but interest there is focussed on those that are unitary. The next result is easy to verify.

8.1.1 Lemma. *For an $n \times n$ matrix, any two of the following statements imply the third:*

(a) W is a type-II matrix.

(b) $n^{-1/2}W$ is unitary.

(c) $|W_{i,j}| = 1$ for all i and j. □

We say a type-II matrix is *flat* if all its entries have the same absolute value. The character table of an abelian group is a flat type-II matrix. A flat real type-II matrix is a Hadamard matrix.

Nomura [**?**] has shown that there are exactly three equivalence classes of 5×5 type-II matrices. One class is represented by the character table of the cyclic group of order five, the other two have representatives of the form $\alpha I + J$ (so here $W^{(-)}$ is not equivalent to W). Haagerup [**?**] has shown that if n is not prime, there are infinitely many equivalence classes of unitary type-II matrices of order n.

8.2 Two Algebras

Let W be a Schur-invertible $n \times n$ matrix. We define $\mathcal{N}_W$ to be the set of matrices for which all the vectors

$$We_i \circ W^{(-)}e_j, \qquad 1 \le i, j \le n$$

are eigenvectors. Clearly this set of matrices is closed under multiplication and contains the identity. Thus it is a matrix algebra, known as the Nomura algebra. Note also that

$$\mathcal{N}_{W^{(-)}} = \mathcal{N}_W.$$

If $M \in \mathcal{N}_W$, we define $\Theta_W(M)$ to be $n \times n$ matrix with ij-entry equal to the eigenvalue of M on $We_i \circ W^{(-)}e_j$. We have

$$\Theta_W(I) = J.$$

We also see that if M and N belong to $\mathcal{N}_W$, then

$$\Theta_W(MN) = \Theta_W(M) \circ \Theta_W(N).$$

It follows that the image of $\mathcal{N}_W$ under Θ is Schur-closed and contains J.

8.2.1 Lemma. *The matrix W is type II if and only if $J \in \mathcal{N}_W$.* □

Let W be a type-II matrix, with rows and columns indexed by the set Ω, where $|\Omega| = n$. We define two families of vectors in $\mathbb{C}^n$, as follows.

$$Y_{a,b} := We_a \circ W^{(-)}e_b, \qquad Y'_{a,b} := W^T e_a \circ W^{(-)T}e_b$$

Suppose

$$F_i := \frac{1}{n} Y_{u,i} Y_{i,u}^T.$$

We verify easily that

$$F_i F_j = \delta_{i,j} F_i,$$

which shows that the F_i's form an orthogonal set of n idempotents. We note that $\mathrm{rk}(F_i) = 1$ and $\mathrm{tr}(F_i) = 1$. As the F_i's commute it follows that $\sum_i F_i$ is an idempotent matrix with trace equal to n; hence

$$\sum_i F_i = I.$$

8.2.2 Lemma. *If $M \in \mathcal{N}_W$ then*

$$M = \sum_i (\Theta(M))_{u,i} F_i.$$

Proof. We have

$$MF_i = \frac{1}{n} MY_{u,i} Y_{i,u}^T = (\Theta(M))_{u,i} F_i.$$

Summing this over i in Ω, recalling that $\sum_i F_i = I$, we get

$$M = \sum_i (\Theta(M))_{u,i} F_i.$$

The following critical result is due to Nomura [**?**].

8.2.3 Theorem. *If $M \in \mathcal{N}_W$ then*

$$\Theta_W(M) Y'_{s,r} = nM_{r,s} Y'_{s,r}.$$

Hence $\Theta_W(M) \in \mathcal{N}_{W^T}$ and $\Theta_{W^T}(\Theta_W(M)) = nM^T \in \mathcal{N}_W$.

Proof. We have

$$(F_i)_{r,s} = \frac{1}{n} \frac{w(r,u)}{w(r,i)} \frac{w(s,i)}{w(s,u)} = \frac{1}{n} \frac{w(r,u)}{w(s,u)} \frac{w(s,i)}{w(r,i)}.$$

Therefore, by 8.2.2,

$$M_{r,s} = \frac{1}{n}\frac{w(r,u)}{w(s,u)}\sum_i (\Theta(M))_{u,i}\frac{w(s,i)}{w(r,i)}$$

and so

$$nM_{r,s}(Y'_{s,r})_u = (\Theta(M)Y'_{s,r})_u.$$

This implies the theorem. □

8.2.4 Corollary. *If W is a type-II matrix, then $\mathcal{N}_W$ and $\mathcal{N}_{W^T}$ are Bose-Mesner algebras of association schemes.* □

Another consequence of the previous theorem is that Θ_W and Θ_{W^T} are bijections.

8.2.5 Lemma. *If W is real then all matrices in $\mathcal{N}_W$ are symmetric.*

Proof. If W is real then the eigenvectors $Y_{a,b}$ are real. Hence the Schur idempotents of the scheme have only real eigenvalues. Since $\mathcal{N}_W$ is closed under transposes and is a commutative algebra, the Schur idempotents are real normal matrices. A real normal matrix is symmetric if and only if its eigenvalues are real. □

If W is a type-II matrix with algebra $\mathcal{N}_W$ then W determines a spin model if and only if some type-II matrix equivalent to W lies in $\mathcal{N}_W$. As any type-II matrix equivalent to W has the same algebra, we may concentrate on the matrices W that lie in their algebra. If $W \in \mathcal{N}_W$ then its diagonal must be constant, and all its row and column sums are equal.

8.3 Eigenspaces

The algebra $\mathcal{N}_W$ determines a scheme consisting of $n \times n$ matrices. We describe how we can determine the eigenmatrix of the scheme. Let us say that vectors $Y_{a,b}$ and $Y_{r,s}$ *overlap* if $Y_{a,b}^T Y_{r,s} \neq 0$.

8.3.1 Lemma. *If $Y_{a,u}$ and $Y_{b,c}$ overlap then $(\Theta(M))_{u,a} = (\Theta(M))_{b,c}$.*

Proof. As the vectors $Y_{u,i}$ for fixed u form a basis, $Y_{b,c}$ lies in their span. In fact

$$Y_{b,c} = \frac{1}{n}\sum_i (Y_{i,u}^T Y_{b,c})\, Y_{u,i}.$$

So

$$(\Theta(M))_{b,c} Y_{b,c} = M Y_{b,c} = \frac{1}{n}\sum_i (Y_{i,u}^T Y_{b,c})(\Theta(M))_{u,i}\, Y_{u,i}.$$

Multiply both sides of this by $Y_{a,u}^T$ to get

$$\begin{aligned}(\Theta(M))_{b,c} Y_{a,u}^T Y_{b,c} &= \frac{1}{n}(Y_{a,u}^T Y_{b,c})(\Theta(M))_{u,a} Y_{a,u}^T Y_{u,a} \\ &= Y_{a,u}^T Y_{b,c}(\Theta(M))_{u,a}.\end{aligned}$$

If $Y_{a,u}^T Y_{b,c} \neq 0$, this implies that $(\Theta(M))_{u,a} = (\Theta(M))_{b,c}$. □

We define a graph with vertex set Ω. Define i and j to be adjacent if there are b and c such that $Y_{b,c}$ overlaps both $Y_{u,i}$ and $Y_{u,j}$. Note u is adjacent to itself, and to no other vertex. Any matrix $\sum F_i$, where i ranges over the vertices in a component of this graph, is a matrix idempotent of the scheme belonging to $\mathscr{N}_W$. (The key point is that this sum lies in $\mathscr{N}_W$.)

We have the following observation, due to Jaeger et al [**?**].

8.3.2 Lemma. *Let W be a Hadamard matrix of order n. If $\mathscr{N}_W$ is non-trivial, then n is divisible by eight.*

Proof. Let w_i denote We_i. Normalize W so that $w_1 = \mathbf{1}$ and assume 1, i, j and k are distinct. Then

$$(w_1 + w_i) \circ (w_1 + w_j) \circ (w_1 + w_j)$$

is the Schur product of three vectors with entries $0, \pm 2$. The sum of the entries of this vector is

$$\begin{aligned}&\langle \mathbf{1}, w_1^{\circ 3}\rangle + \langle \mathbf{1}, w_1^{\circ 2} \circ (w_i + w_j + w_k)\rangle \\ &\qquad + \langle \mathbf{1}, w_1 \circ (w_i \circ w_j + w_i \circ w_k + w_j \circ w_k)\rangle + \langle \mathbf{1}, w_i \circ w_j \circ w_k\rangle\end{aligned}$$

Since W is a Hadamard matrix, the second and third terms here are zero, whence we deduce that, modulo 8,

$$n + \langle \mathbf{1}, w_i \circ w_j \circ w_k\rangle = 0$$

and therefore, if n is not divisible by 8, then w_i cannot be orthogonal to $w_j \circ w_k$. □

Chapter 9

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].

9.1 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.

9.1.1 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 $\mathscr{A}$.

9.2 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 $M = \sum_j a_j E_j$ and we define $M^{\hat{\tau}}$ by

$$M^{\hat{\tau}} = \sum_j a_j E_j^\tau.$$

This is an L-linear map.

9.2.1 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.

9.2.2 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.

9.3 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 $i = 0, 1, \ldots, d$.

9.3.1 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).

9.3.2 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 $\mathscr{A}/\mathbb{Q}$. Therefore 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 $m \geq 3$.

9.4 Multipliers

Let G be an abelian group with elements $g_1, \ldots, g_n$, where g_1 is the identity element, and let A_g denote the $n \times n$ permutation matrix corresponding to the element g of G. The eigenvalues of A_g are all the complex m-th roots of unity; hence if G has exponent m then the splitting field L of $\mathscr{A}$ is $\mathbb{Q}$ extended by a primitive m-th root of unity. (This is true for all finite groups, abelian or not, but is much harder to prove.)

Let L be the splitting field of the association scheme $\mathscr{A}$, let Γ be the Galois group of $L/\mathbb{Q}$ and let α be a primitive m-th root of unity. If $\tau \in \Gamma$, then there is

an integer $t(\tau)$, coprime to m, such that

$$\alpha^\tau = \alpha^{t(\tau)}.$$

Thus we may view t as a map from Γ into $\mathbb{Z}_m^*$, the group of units of the integers mod m, and this mapping is an isomorphism. (For the missing details see, e.g., [frotay: §VI.1].)

Let $\mathscr{A}$ be the association scheme of the abelian group G. If $\tau \in \Gamma$ and $A \in L[\mathscr{A}]$, let $A^{\hat{\tau}}$ denote the image of A under the eigenvalue action of τ, which we introduced in 9.2. If $g \in G$ and $t = t(\tau)$ then

$$(A_g)^{\hat{\tau}} = (A_g)^{t(\tau)} = A_{g^t},$$

consequently we may view Γ as acting as a group of automorphisms of G. (By [ser: §13.1], two elements of G are conjugate under this action if and only if they generate the same cyclic subgroup of G.)

9.4.1 Theorem. *Let $\mathscr{A}$ be the association scheme formed by the elements of the abelian group G, where G has exponent m. If t is an integer coprime to m, then the map $A \mapsto A^{(t)}$ permutes the Schur idempotents of $\mathscr{A}$ and fixes each subscheme in it.*

Proof. The first claim follows from 9.2, where we proved that the Schur idempotents of $\mathscr{A}$ are permuted among themselves under the eigenvalue action of Γ. The remaining claim requires some work though.

It is enough to prove that each subscheme is fixed if t is prime. So assume t is prime and that $\tau \in \Gamma$ such that $t = t(\tau)$. Suppose that $B_0, \ldots, B_d$ are the minimal Schur idempotents of a subscheme $\mathscr{B}$ of $\mathscr{A}$. For each B_i there is a subset C_i of G such that

$$B_i = \sum_{g \in C_i} A_g.$$

As the matrices A_g commute it follows that, modulo t,

$$B_i^t \equiv \sum_{g \in C_i} A_{g^t} = B_i^{\hat{\tau}}. \tag{9.4.1}$$

The right side here is a 01-matrix, because it is the image of B_i under τ. On the other hand, there are integers $b_i(j)$ such that

$$B_i^t = \sum_j b_i(j) B_j$$

and therefore we see that $B^{\hat{\tau}}$ is equal to the sum of the matrices B_j, as j ranges over the set

$$\{j : b_i(j) \equiv 1, \bmod t\}.$$

Accordingly $B_i^{\hat{\tau}} \in \mathscr{B}$. □

We note one consequence of the above proof or, more precisely, of (9.4.1).

9.4.2 Corollary. *Let $\mathscr{A}$ be the association scheme formed by the elements of the abelian group G, where G has exponent m. Let L be the splitting field of $\mathscr{A}$, and let Γ be the Galois group of $L/\mathbb{Q}$. If Γ contains an element of prime order p then $A^{(p)} \equiv A^p$ modulo p.* □

This corollary implies that every subalgebra of the group ring $\mathbb{Z}_p[G]$ is fixed by the map $A \mapsto A^{(p)}$; thus it is a slight strengthening of [Land: Prop. 4.7].

An association scheme $\mathscr{B}$ is called a *translation scheme* if there an abelian group, G say, acting as a regular group of automorphisms on it. It is not hard to show that a translation scheme is the same thing as a subscheme of the scheme formed by the elements of G. In [BCN: Thm 11.1.10], Brouwer et al. use 9.4.1 to show that every metric translation scheme arises from a completely regular linear code in a Hamming scheme. We make a diversion, to show that it implies the standard multiplier theorem in design theory. (For background and terminology, see Lander's book [Land].)

Let D be the incidence matrix of a symmetric (v,k,λ)-design $\mathscr{D}$. This design is determined by a difference set in the abelian group G if and only if D belongs to the Bose-Mesner algebra of the association scheme $\mathscr{A}$ formed by the elements of G. (This is not how a design theorist would express it!) A Schur idempotent D in $\mathbb{C}[\mathscr{A}]$ is the incidence matrix of a symmetric (v,k,λ)-design if and only if

$$D^T D = nI + \lambda J,$$

where $n := k - \lambda$. Let L be the splitting field of $\mathscr{A}$ and suppose that the Galois group of $L/\mathbb{Q}$ contains an element τ of prime order p. We find that

$$D^T D^p = D^T D^{p-1} D = nD^{p-1} + \lambda k^{p-1} J \equiv \lambda k^{p-1} J \quad (mod\ n).$$

Now assume that p divides n. Then, modulo p we have that $D^p \equiv D^{(p)}$ and $k^{p-1} \equiv 1$; whence

$$D^T D^{(p)} \equiv \lambda J \quad (mod\ n). \tag{9.4.2}$$

As the map $A \mapsto A^{(p)}$ preserves Schur multiplication, $D^{(p)}$ is the incidence matrix of a design. If S is a block in this design and $p > \lambda$ then (9.4.2) implies

that S meets each block of $\mathscr{D}$ in at least λ points. By a nice result of Lander [Land: Lemma 5.2], it follows that S must belong to $\mathscr{D}$, and so $D^{(p)}$ differs from D by a permutation of its columns.

The points of $\mathscr{D}$ can be taken to be the elements of G, then τ is permutation of the points of G and we have just shown that this permutation is an automorphism of G. Thus we have proved that p is a multiplier.

9.4.3 Lemma. *Let $\mathscr{A}$ be a primitive subscheme of the association scheme formed by the elements of the abelian group G, and let p be a prime that divides $|G|$. Then, for each minimal Schur idempotent A_i of $\mathscr{A}$, there is an integer c_i such that $A_i^p \equiv c_i I$, mod p.*

Proof. If $C_i \subseteq G$ such

$$A_i = \sum_{g \in C_i} A_g$$

then, modulo p, we have

$$A_i^p \equiv \sum_{g \in C_i} A_{g^p}.$$

The coefficient of A_h in the right side of this expression is equal to

$$|\{g \in C_i : g^p = h\}|,$$

which we denote by n_h. We also have integers $a_i(j)$ such that

$$A_i^p = \sum_j a_i(j) A_j.$$

Consequently

$$\sum_{h : n_h \neq 0} n_h A_h = \sum_{j : a_i(j) \neq 0} a_i(j) A_j \tag{9.4.3}$$

The elements h in the index set of the sum on the left of (9.4.3) all lie in the subgroup formed by the p-th powers of elements of G. Therefore

$$\sum_{h : n_h \neq 0} A_h$$

is the adjacency matrix of a disconnected graph. Since (9.4.3) implies that this sum belongs to $\mathbb{C}[\mathscr{A}]$, it follows that $\mathscr{A}$ is imprimitive if

$$\sum_{g \in C_i} A_{g^p}$$

is not a scalar matrix, mod p. □

9.4.4 Theorem. *Let $\mathscr{A}$ be a primitive translation scheme relative to the abelian group G. If, for some prime p, the Sylow p-subgroup of G is cyclic then G has prime order.*

Proof. If $\mathscr{A}$ is primitive then so is $\mathscr{A}/\mathbb{Q}$. If $\mathscr{A}/\mathbb{Q}$ has only one class then G has prime order. Hence we can assume that $\mathscr{A} = \mathscr{A}/\mathbb{Q}$ and that it has at least two classes. Let $A_0, \ldots, A_d$ be the Schur idempotents of $\mathscr{A}$. As $\mathscr{A}$ is a translation scheme, it is a subscheme of the scheme formed by the elements of G, there are subsets $C_0, \ldots, C_d$ such that $C_0 = \{1\}$ and

$$A_i = \sum_{g \in C_i} A_g.$$

The sets C_i partition the elements of G.

Let U be the subgroup of G formed by the elements of order dividing p. Since the Sylow p-subgroup of G is cyclic, $|U| = p$.

Because $\mathscr{A}$ is primitive, for each index i there is a constant c_i such that $A_i^p \equiv c_i I$, modulo p. But, modulo p,

$$A_i^p \equiv \sum_{g \in C_i} A_{g^p}$$

and, therefore, the number of elements g in C_i such that $g^p = h$ is divisible by p when $h \neq 1$. If $x, y \in G$ and

$$x^p = y^p = h$$

then $(xy^{-1})^p = 1$ and so $xy^{-1} \in U$. This implies that the set of elements g in C_i such that $g^p = h$ is a union of cosets of U, and consequently the set $C_i \setminus U$ is also a union of cosets of U. We also see that $|C_i \cap U|$ must be congruent to c_i modulo p and, as $|U| = p$, this implies that $|C_i \cap U| = c_i$.

Next, $J = \sum_i A_i$ and, as p divides $|G|$, modulo p we have

$$0 \equiv J^p = \Big(\sum_i A_i\Big)^p \equiv \sum_i A_i^p \equiv \sum_i c_i I.$$

Since $A_0 = I$, we have $c_0 = 1$ and therefore, mod p

$$\sum_{i=1}^{d} c_i \equiv p - 1. \tag{9.4.4}$$

Let Γ be the Galois group of the splitting field of the association scheme belonging to G. Then, under the eigenvalue action of Γ, each minimal idempotent A_i of $\mathscr{A}$ is fixed. This implies that each set C_i is fixed by Γ. As remarked just

before the statement of 9.4.1, two elements of G are conjugate under Γ if and only if they generate the same subgroup of G. In particular, the non-identity elements of U must form a single conjugacy class. This implies that, if $i \neq 0$, then c_i is equal to 0 or $p-1$. From (9.4.4) we conclude that one of the sets C_i contains $U \setminus 1$ and that the remaining sets C_i are all disjoint from U. In each case the set $C_i \setminus U$ is a union of cosets of U.

Suppose that $C_i \cap U = \emptyset$ if $i \neq j$. Then it follows that there are 01-matrices $B_1, \ldots, B_d$ such that if $i \neq j$, then

$$A_i = B_i \otimes J_p$$

and

$$A_j = B_j \otimes J_q + I \otimes (J_p - I_p).$$

If we assume $B_0 = I$ then $B_0, \ldots, B_d$ form a translation scheme relative to the quotient group G/U. Hence $\sum_{i \neq j} B_i = J - B_i$ and, therefore,

$$\sum_{i \neq j} A_i = J - A_j.$$

This implies that $I \otimes (J_p - I_p) \in \mathscr{A}$, and so $\mathscr{A}$ is imprimitive. □

The argument in the last part of the proof actually implies that $G/U \times U$ acts as a group of automorphisms of $\mathscr{A}$. It follows inductively that if the Sylow p-subgroup P of G is cyclic and Q is any abelian group such that $|Q| = |P|$, then $G/P \times Q$ acts as group of automorphisms of $\mathscr{A}$. The result itself is best attributed to Wielandt, see Theorem 25.4 in [Wiel].

bi E. Bannai and T. Ito, *Association Schemes, I*. Benjamin-Cummings (London), 1984.

dbg J. de Boer and J. Goeree, Markov traces and II_1 factors in conformal field theory, *Commun. Math. Physics,* **139** (1991), 267–304.

bm2 W. G. Bridges and R. A. Mena, Rational G-matrices with rational eigenvalues, *J. Combinatorial Theory, Ser. A,* **32** (1982), 264–280.

bcn A. E. Brouwer, A. M. Cohen and A. Neumaier, *Distance-Regular graphs*. Springer (Berlin), 1989.

coga A. Coste and T. Gannon, Remarks on Galois symmetry in rational conformal field theories, *Physics Letters B,* **323** (1994), 316–321.

frotay A. Fröhlich and M. J. Taylor, *Algebraic Number Theory*. Cambridge U. P. (Cambridge), 1991.

cgbk C. D. Godsil, *Algebraic Combinatorics*. Chapman and Hall (New York), 1993.

hkcss A. R. Hammons Jr., P. Vijay Kumar, A. R. Calderbank, N. J. A. Sloane and P. Sole, The Z4-Linearity of Kerdock, Preparata, Goethals and Related Codes, *IEEE Trans. Information Theory,* **40** (1994), 301-319,

xdh X. Hou, On the G-matrices with entries and eigenvalues in $Q(i)$, *Graphs and Combinatorics,* **8** (1992), 53–64.

mune A Munemasa, Splitting fields of association schemes, *J. Combinatorial Theory, Ser. A,* **57** (1991), 157–161.

Chapter 10

A Bestiary

10.1 Cyclic Schemes

Let P be the permutation matrix representing a cycle of length n. Thus, if $e_1, \ldots, e_n$ is the standard basis, then $Pe_i = e_{i+1}$ (where the subscripts are computed modulo n. If we define $A_i := P^i$ for $i = 0, 1, \ldots, n-1$, then the matrices A_i form an association scheme with $n-1$ classes, which we are calling the *cyclic scheme*.

Let θ denote a primitive n-th root of unity, and let u_i denote the column vector in $\mathbb{C}^n$ with j-th entry equal to $\theta^{(i-1)(j-1)}$. Then

$$\langle u_i, u_j \rangle = n\delta_{i,j}$$

and

$$Pu_i = \theta^i u_i.$$

It follows that each vector u_i spans an eigenspace of the cyclic scheme. The matrix representing orthogonal projection onto the span of u_i is

$$E_i := \frac{1}{n} u_i u_i^*.$$

The matrices $E_0, \ldots, E_{n-1}$ are the principal idempotents of our scheme. We note that

$$n(E_i)_{r,s} = \theta^{(i-1)(r-1)} \theta^{-(i-1)(s-1)} = \theta^{(i-1)(r-s)}.$$

Further

$$A_i E_j = P^i E_j = \theta^{ji} E_j,$$

whence, if V denotes the matrix of eigenvalues, we have

$$V_{i,j} = \theta^{(i-1)(j-1)}.$$

Thus the columns of the matrix of eigenvalues are the eigenvectors of the scheme; this happens because each eigenspace is 1-dimensional. Note that $V = V^*$ and that V is a type II matrix. This shows that the cyclic scheme is contained in the Nomura algebra of V. Since the dimension of a scheme consisting of $n \times n$ matrices is at most n, we deduce that the cyclic scheme is the Nomura algebra of V.

Let n be an integer. If n is even, let η be a primitive $2n$-th root of unity; if n is odd let it be a primitive n-th root of unity. Let W be the $n \times n$ matrix given by

$$W_{i,j} := \eta^{(i-j)^2}, \qquad i, j \in \mathbb{Z}_n.$$

Then $W^{(-)} = \overline{W}$ and

$$\begin{aligned}(WW^{(-)T})_{a,b} &= \sum_{i \in \mathbb{Z}_n} \eta^{(a-i)^2-(b-i)^2} \\ &= \sum_{i \in \mathbb{Z}_n} \eta^{a^2-b^2-2(a-b)i} \\ &= \eta^{a^2-b^2} \sum_{i \in \mathbb{Z}_n} \eta^{-2(a-b)i} \\ &= n\delta_{a,b}.\end{aligned}$$

This shows that W is a type II matrix. Since W is a circulant, it lies in the Nomura algebra of V. On the other hand

$$(We_a \circ W^{(-)}e_b)_i = \eta^{(a-i)^2-(b-i)^2} = \eta^{a^2-b^2}\eta^{-2(a-b)i},$$

whence $\mathcal{N}_W = \mathcal{N}_V$. Hence $W \in \mathcal{N}_W$ and therefore it determines a spin model.

10.2 Paley Graphs

Let $\mathbb{F}$ be a field of order q, where $q = 1$ modulo 4. Let X denote the graph with the elements of $\mathbb{F}$ as its vertices, where elements x and y of $\mathbb{F}$ are adjacent if and only if $x - y$ is a square in $\mathbb{F}$. We call X the *Paley graph* of order q.

Recall that, if $\mathbb{F}$ has odd order q, then the number of non-zero squares in $\mathbb{F}$ is $(q-1)/2$, and that -1 is a square in $\mathbb{F}$ if and only if $q = 1$ modulo 4. Further,

each element of $\mathbb{F}$ is the sum of two squares. It follows that X is a connected regular graph with valency $(q-1)/2$. It is easy to verify that, if $a \in \mathbb{F}$, then the map τ_a that sends x to $x+a$ is an automorphism of X. The maps τ_a form a regular subgroup of $\mathrm{Aut}(X)$. If s is a non-zero square in $\mathbb{F}$ then the map μ_s that sends x to sx is an automorphism of X that fixes 0. The maps τ_a and μ_s together generate a group of order $\binom{q}{2}$ which acts arc-transitively on X (and on $\overline{X}$). It is not too difficult to use this group to show that X is strongly regular; we will use another approach though.

We first construct eigenvectors for X. Let ψ be an additive character of $\mathbb{F}$ and let S denote the set of non-zero squares in $\mathbb{F}$. Then ψ is a function on the vertices of X. Since

$$\sum_{x\sim a} \psi(x) = \sum_{s\in S} \psi(a+s) = \psi(a) \sum_{s\in S} \psi(s),$$

We deduce that ψ is an eigenvector of X with eigenvalue

$$\sum_{s\in S} \psi(s).$$

If ψ is the trivial character, then this eigenvalue is the valency $(q-1)/2$ of X.

To get the remaining eigenvalues, note that if ψ is a non-trivial additive character and a is a non-zero element of $\mathbb{F}$, then the composition $\psi \circ \mu_a$ is again an additive character, and all additive characters arise in this way. We use ψ_a to denote $\psi \circ \mu_a$. If a is a non-zero square in $\mathbb{F}$ then

$$\sum_{s\in S} \psi_a(s) = \sum_{s\in S} \psi(as) = \sum_{s\in S} \psi(s).$$

If a is not a square then, since $\sum_{x\in\mathbb{F}} \psi(x) = 0$, we have

$$\sum_{s\in S} \psi_a(s) = \sum_{s\in S} \psi(as) = \sum_{s\notin S} \psi(s) = -1 - \sum_{s\in S} \psi(s).$$

It follows that there is a real number θ such that θ and $-1-\theta$ are eigenvalues of X, both with multiplicity $(q-1)/2$.

If $A = A(X)$ then

$$q\frac{q-1}{2} = \mathrm{tr}(A^2) = \frac{q-1}{2}(-1-\theta)^2 + \frac{q-1}{2}\theta^2 + \left(\frac{q-1}{2}\right)^2,$$

whence

$$q = 2\theta^2 + 2\theta + 1 + \frac{q-1}{2}$$

and so θ is a zero of

$$t^2 + t - \frac{q-1}{4}.$$

Thus

$$\theta = \frac{1}{2}(-1 \pm \sqrt{q}).$$

$$P = Q = \begin{pmatrix} 1 & \frac{q-1}{2} & \frac{q-1}{2} \\ 1 & \frac{-1+\sqrt{q}}{2} & \frac{-1-\sqrt{q}}{2} \\ 1 & \frac{-1-\sqrt{q}}{2} & \frac{-1+\sqrt{q}}{2}. \end{pmatrix}$$

A strongly regular graph on q vertices with valency $(q-1)/2$ and remaining eigenvalues $(-1+\pm\sqrt{q})/2$ and $(-1-\pm\sqrt{q})/2$ is called a *conference graph*. The product of the two eigenvalues other than the valency is $(1-q)/4$, so if a conference graph exists then $q \equiv 1$ modulo 4. In fact q must be the sum of two squares (and thus there is no conference graph on 21 vertices, for example). The complement of a conference graph is a conference graph.

10.3 Quasisymmetric Designs

Let $\mathscr{B}$ be the set of blocks of a 2-$(v,k,1)$ design. The block graph X of this design has vertex set v, and two blocks are adjacent if and only if they have exactly one point in common. This is a regular graph with valency

$$k\left(\frac{v-1}{k-1} - 1\right) = \frac{k(v-k)}{k-1}.$$

Let N be the incidence matrix of a 2-$(v,k,1)$ design. Then

$$NN^T = \frac{v-k}{k-1}I + J$$

and

$$N^T N = kI + A(X),$$

where X is the block graph of the design. Since $N^T N$ is positive semidefinite, so is $kI + A(X)$. Hence the least eigenvalue of X is at least $-k$.

The above expressions for NN^T and $N^T N$ imply that X is strongly regular, as we demonstrate now. The key fact is that NN^T and $N^T N$ have the same non-zero eigenvalues, with the same multiplicities. The eigenvalues of

$$\frac{v-k}{k-1}I + J$$

are

$$\frac{v-k}{k-1}+v=\frac{vk-k}{k-1}$$

and

$$\frac{v-k}{k-1},$$

with respective multiplicities 1 and $v-1$. It follows that these are eigenvalues of $kI+A(X)$, with the same multiplicities. Hence the eigenvalues of $A(X)$ are

$$\frac{k(v-k)}{k-1}, \quad \frac{v-k^2}{k-1}, \quad -k \tag{10.3.1}$$

with respective multiplicities 1, $v-1$ and

$$\frac{v(v-1)}{k(k-1)}-v=v\frac{v-1+k-k^2}{k(k-1)}. \tag{10.3.2}$$

The first of these eigenvalues is the valency of X and, since it is simple, X must be connected. (This is easy to prove directly.)

If X is the block graph of a 2-$(v,k,1)$ design then $-k$ is an eigenvalue of X. Since the multiplicity of an eigenvalue is non-negative, (10.3.2) implies that

$$v \geq k^2-k+1.$$

If equality holds, $-k$ has multiplicity zero and, from (10.3.1), the valency of X is k^2-k—hence X is complete. (Of course, if equality holds then our block set is the set of lines of the projective plane of order $k-1$.)

If $v>k^2-k+1$, then X is a connected regular graph with exactly three distinct eigenvalues and is therefore strongly regular. Its eigenmatrices are

$$P=\begin{pmatrix} 1 & \frac{k(v-k)}{k-1} & \frac{(v-k)(v-k^2+k-1)}{k(k-1)} \\ 1 & \frac{v-k^2}{k-1} & -\frac{v-k^2+k-1}{k-1} \\ 1 & -k & k-1 \end{pmatrix}$$

and

$$Q=\begin{pmatrix} 1 & v-1 & \frac{v(v-k^2+k-1)}{k(k-1)} \\ 1 & \frac{(v-1)(v-k^2)}{k(v-k)} & -\frac{v(v-k^2+k-1)}{k(v-k)} \\ 1 & -\frac{(v-1)k}{v-k} & \frac{v(k-1)}{v-k} \end{pmatrix}$$

If N is the incidence matrix of a 2-(v,k,λ) design then $N\mathbf{1}=r\mathbf{1}$ and

$$NN^T=(r-\lambda)I+\lambda J.$$

Here r is, as usual, the number of blocks that contain a given point. So

$$r = \lambda \frac{v-1}{k-1}$$

and thus if $v > k$, then $r > \lambda$ and consequently NN^T is invertible and has exactly two eigenvalues. Consequently we can determine the eigenvalues and their multiplicities for $N^T N$.

We say that a 2-design is *quasisymmetric* if there are distinct integers α and β such that the size of the intersection of any two distinct blocks is α or β. If N is the incidence matrix of a quasisymmetric design, then there are square matrices A_1 and A_2 such that

$$N^T N = kI + \alpha A_1 + \beta A_2.$$

It is not hard to show that A_1 and A_2 are the adjacency matrices of complementary regular graphs with exactly three eigenvalues. Hence each quasisymmetric design gives rise to a strongly regular graph. Note that a 2-$(v,k,1)$ design is quasisymmetric, with $\alpha = 0$ and $\beta = 1$.

10.4 Partial Spreads

Let Z be a complete graph on n^2 vertices. A *parallel class* in Z is a spanning subgraph isomorphic to nK_n. We say two parallel classes S_1 and S_2 are *orthogonal* if they have no edges in common. If A_i denotes the adjacency matrix of A_i, then S_1 and S_2 are orthogonal if and only if

$$(A_1 + I)(A_2 + I) = J.$$

It is also not difficult to verify that

$$(A_i + I)^2 = n(A_i + I).$$

A *partial spread* is a set of pairwise orthogonal parallel classes.

Now suppose that $S_1, \ldots, S_r$ is a partial spread of size r. The graph X formed by the union of (the edges in) the parallel classes is a regular graph with valency $r(n-1)$; we show that it is strongly regular. Let A be the adjacency matrix of X.

Then

$$\begin{aligned}(A+rI)^2 = \left(\sum_{i=1}^{r}(A_i+I)\right)^2 &= \sum_{i=1}^{r}(A_i+I)^2 + \sum_{i\neq j}(A_i+I)(A_j+I)\\ &= n\sum_{i=1}^{r}(A_i+I) + r(r-1)J\\ &= nA + nrI + r(r-1)J\end{aligned}$$

and therefore

$$A^2 - (n-2r)A - (nr-r^2)I = r(r-1)J.$$

This shows that A is strongly regular, with parameters

$$\big(n^2,\ r(n-1);\ r(r-3)+n,\ r(r-1)\big).$$

If $r=1$ then $X = nK_n$, which is a trivial strongly regular graph, and if $r=2$ then X is $L(K_{n,n})$. When $r=3$, the graph X is best known as a *Latin square graph*. The eigenmatrices are

$$P = Q = \begin{pmatrix}1 & r(n-1) & (n+1-r)(n-1)\\ 1 & n-r & -1-n+r\\ 1 & -r & r-1\end{pmatrix}$$

Now set $r=-s$ and $n=-m$. Then, if $m \le s(s+3)$, there could be a strongly regular graph with parameters

$$\big(m^2,\ s(m+1);\ s(s+3)-m,\ s(s+1)\big).$$

and eigenmatrices

$$P = Q = \begin{pmatrix}1 & s(m+1) & (m-1-s)(m+1)\\ 1 & -m+s & -1+m-s\\ 1 & s & -s-1\end{pmatrix}$$

In fact, strongly regular graphs with these parameters do exist in some cases, and are said to be of *negative Latin square* type.

Two especially interesting cases occur when $m = s(s+3)$ and $s = 1$ or 2. The corresponding parameter vectors are

$$(16,\ 5;\ 0,\ 2), \qquad (100,\ 22;\ 0,\ 6).$$

The first is associated to the *Clebsch* graph, the second to the *Higman-Sims* graph. The vertices at distance two from a given vertex in the Clebsch graph form a triangle-free graph on 10 vertices with valency $5-2=3$. Given this hint, it is not hard to construct the Clebsch graph from the Petersen graph.

A partial spread contains at most $n+1$ parallel classes, and a *spread* is a partial spread with exactly $n+1$ classes. If $A_1,\dots,A_{n+1}$ are the corresponding matrices and $A_0 = I$, then $A_0,\dots,A_{n+1}$ is an association scheme. We have

$$\big(A_1 + I - \frac{1}{n}J\big)\big(A_2 + I - \frac{1}{n}J\big) = J - 2J + J = 0,$$

whence we see that the matrices $A_i + I - \frac{1}{n}J$ together with $n^{-2}J$ form the complete set of principal idempotents. Next, if $i \neq j$, then

$$A_i\big(A_j + I - \frac{1}{n}J\big) = -\big(A_j + I - \frac{1}{n}J\big)$$

and

$$A_i\big(A_i + I - \frac{1}{n}J\big) = (n-1)\big(A_j + I - \frac{1}{n}J\big)$$

Hence the eigenmatrices are the $(n+2)\times(n+2)$ matrices

$$P = Q = \begin{pmatrix} 1 & (n-1)\mathbf{1} \\ \mathbf{1} & nI - J \end{pmatrix}.$$

If π denotes a partition of $\{1,\dots,n+1\}$, the matrices

$$\sum_{i\in C} A_i,$$

where C runs over the cells of π, form a subscheme. Each class in this subscheme is strongly regular. It follows that the $n+1$ parallel classes form an amorphic association scheme. Note that spreads correspond to affine planes. Hence examples are only known when n is a prime power; further if $n > 4$ and is a prime power, but not a prime, there are at least two different spreads in K_{n^2}.

10.5 Covers of Complete Bipartite Graphs

We consider distance-regular 2-fold covers of the complete graph $K_{n,n}$.

Suppose $V(K_{n,n})$ is the disjoint union of the sets

$$A = \{a_1,\dots,a_n\}, \quad B = \{b_1,\dots,b_n\}$$

and that each vertex in A is adjacent to each vertex in B. We construct a new graph as follows. Let H be a Hadamard matrix of order n. The vertices of the new graph $X(H)$ are the elements of

$$(A\cup B)\times\{-1,1\}$$

and (a_r, i) is adjacent to (b_s, j) if and only if $H_{r,s}i = j$. We see that $X(H)$ is a regular bipartite graph on $4n$ vertices with valency n. If H and H' are monomially equivalent, then $X(H)$ and $X(H')$ are isomorphic.

10.5.1 Lemma. *Let H be an $n \times n$ Hadamard matrix. The number of vertices at distance two from a fixed vertex in $X(H)$ is $2n-2$.* □

Since $X(H)$ is bipartite with $2n$ vertices in each colour class, it follows that there is a unique vertex at distance four from each vertex. Hence $X(H)$ is antipodal, with diameter four.

The matrix of eigenvalues is

$$P = \begin{pmatrix} 1 & n & 2n-2 & n & 1 \\ 1 & \sqrt{n} & 0 & -\sqrt{n} & -1 \\ 1 & 0 & -2 & 0 & 1 \\ 1 & -\sqrt{n} & 0 & \sqrt{n} & -1 \\ 1 & -n & 2n-2 & -n & 1 \end{pmatrix}.$$

If $H = H_0 - H_1$ where H_0 and H_1 are non-negative and $H_0 \circ H_1 = 0$, then

$$A_1 = \begin{pmatrix} 0 & H_0 \otimes I_2 + H_1 \otimes (J_2 - I_2) \\ H_0^T \otimes I_2 + H_1^T \otimes (J_2 - I_2) & 0 \end{pmatrix}$$

Note that

$$nI = (H_0 - H_1)(H_0 - H_1)^T = (H_0 H_0^T + H_1 H_1^T) - (H_0 H_1^T + H_1 H_0^T)$$

and

$$nJ = (H_0 + H_1)(H_0 + H_1)^T = (H_0 H_0^T + H_1 H_1^T) + (H_0 H_1^T + H_1 H_0^T),$$

whence

$$H_0 H_0^T + H_1 H_1^T = \frac{n}{2}(J + I), \qquad H_0 H_1^T + H_1 H_0^T = \frac{n}{2}(J - I).$$

Note that

$$A_2 = (J_n - I_n) \otimes J_2, \quad A_4 = I_n \otimes (J_2 - I_2)$$

and that $A_3 = A_1 A_4$.

This scheme is formally self-dual.

10.6 Groups

Let Γ be a finite group of order v with conjugacy classes $C_0, \ldots, C_d$. Using the regular permutation representation of Γ, we may view each element of Γ as a $v \times v$ permutation matrix and define A_i to be the sum of the matrices in the i-th conjugacy class. The matrices $A_0, \ldots, A_d$ form an association scheme (and its Bose-Mesner algebra is isomorphic to the centre of the group algebra of Γ). The matrix of eigenvalues of this scheme is determined by the character table of Γ. Since group theorists have determined the character tables of many finite groups, it is useful to be able to translate the information in these tables into the language of association schemes. We show how to do this.

First we must explain what the character table is, which will take some time. A *representation of* Γ *of degree* d over $\mathbb{F}$ is a homomorphism of Γ into the group of $d \times d$ matrices over $\mathbb{F}$. If ϕ is a representation of Γ then $\mathrm{tr}(\phi(g))$ is a function from Γ to $\mathbb{F}$ that is constant on the elements in a conjugacy class of Γ. Such functions are called *characters* of Γ. We will only use representations of Γ over $\mathbb{C}$ here.

The sum of two characters is a character of Γ; a character is *irreducible* if it cannot be written as a sum of two non-zero characters. The number of irreducible characters is equal to the number of conjugacy classes of Γ. The *character table* of Γ is the complex matrix with ij-entry equal to the value of the i-th character of Γ on an element in the j-th conjugacy class. What we have called the character table of an abelian group is a character table in the sense we are using here.

If ψ is a character of Γ, let M_ψ denote the $v \times v$ matrix such that

$$(M_\psi)_{g,h} = \frac{\psi(1)}{|G|}\psi(g^{-1}h).$$

10.6.1 Theorem. *Let Γ be a finite group. The matrices M_ψ, where ψ runs over the distinct irreducible characters of Γ, are the principal idempotents of the conjugacy class scheme of Γ.* □

We offer some comments that go some way towards a proof of this theorem. Suppose C is a conjugacy class of Γ. Let $X(C)$ be the Cayley graph for Γ with connection set C. (Thus g and h are adjacent in $X(C)$ if and only if $hg^{-1} \in C$.) Let A_C denote the adjacency matrix of $X(C)$. If N_i is the $v \times v$ matrix with gh-entry $\psi(g^{-1}h)$, then the gh-entry of $A_C N_\psi$ is equal to

$$\sum_{\{x : xg^{-1} C\}} \psi(x^{-1}h) = \sum_{c \in C} \psi(g^{-1}ch).$$

Assume that ψ is the trace of the irreducible representation Ψ. Then

$$\sum_{c\in C}\psi(g^{-1}ch)=\sum_{c\in C}\psi(chg^{-1})=\sum_{c\in C}\mathrm{tr}(\Psi(chg^{-1})).$$

Since Ψ is a homomorphism, $\Psi(chg^{-1})=\Psi(c)\Psi(hg^{-1})$ and consequently

$$\sum_{c\in C}\mathrm{tr}(\Psi(chg^{-1}))=\mathrm{tr}(\Psi(C)\Psi(hg^{-1})),$$

where $\Psi(C)$ denotes the sum of the values of Ψ over the elements of C. Since Ψ is irreducible, $\Psi(C)$ is a scalar matrix and so there is a constant λ_C such that $\Psi(C)=\lambda_C I$. It follows that

$$(A_C N_\psi)_{g,h}=\mathrm{tr}(\lambda_C\Psi(hg^{-1}))=\lambda_C\psi(hg^{-1})=\lambda_C\psi(g^{-1}h).$$

This shows that each column of N_ψ is an eigenvector for A_C with eigenvalue λ_C.

If $c\in C$, then $\mathrm{tr}(\Phi(C))=|C|\psi(c)$ and therefore

$$\lambda_C=\frac{|C|}{\psi(1)}\psi(c);$$

thus we have the eigenvalues of the matrices A_C in terms of character values. Since the distinct irreducible characters are pairwise orthogonal,

$$N_\psi N_\rho=0$$

if ψ and ρ are distinct irreducible characters of Γ.

This result has several consequences. First, it provides an explicit formula for the dual eigenvalues: if $\psi_0,\ldots,\psi_d$ are the distinct irreducible characters of Γ, then

$$q_i(j)=\psi_i(1)\psi_i(g),$$

where g is any element in the j-th conjugacy class of Γ. Now

$$p_i(j)=\frac{v_i}{m_j}\overline{q_j(i)}$$

Since M_ψ is a projection, the eigenvalues of

$$\frac{|G|}{\psi(1)}M_\psi$$

are 0 and $\frac{|G|}{\psi(1)}$. Since the entries of this matrix are values of characters of Γ, they are algebraic integers. Therefore its eigenvalues are algebraic integers. Consequently the rational number

$$\frac{|G|}{\psi(1)}$$

must be an integer and therefore $\psi(1)$ divides Γ.

[show Krein parameters are rational; give char table for Alt(5) and its rational subscheme]

Chapter 11

Algebra and Modules

11.1 Algebras

A ring $\mathscr{A}$ is an *algebra* over the commutative ring R if there is a homomorphism ψ from R into the centre of A. Any ring $\mathscr{A}$ is an algebra over $\mathbb{Z}$, because the subring of $\mathscr{A}$ generated by 1 is homomorphic image of $\mathbb{Z}$. We will mainly be interested in the case where R is a field, in which case the algebra is a vector space over R.

We offer some examples:

(a) The ring of polynomials $\mathbb{F}[x_1,\ldots,x_n]$ over the field $\mathbb{F}$, in the commuting variables $x_1,\ldots,x_n$.

(b) The ring of all $n\times n$ matrices over the field $\mathbb{F}$. (We will denote this by $\mathrm{Mat}_{n\times n}(\mathbb{F})$.)

(c) The group algebra $\mathbb{F}[G]$ of a finite group G.

(d) The set of all linear mappings of $\mathbb{F}[x]$ to itself is an algebra over $\mathbb{F}$. If M denotes multiplication by x and D denotes differentiation with respect to x, then M and D generate a subalgebra. We note that $DM - MD = I$, from which it follows the algebra generated by D and M is spanned by elements of the form $M^r D^s$. It is possible to show that these elements are linearly independent, whence we see that this algebra has infinite dimension over $\mathbb{F}$.

(e) The ring of polynomials over $\mathbb{F}$ in non-commuting variables $x_1,\ldots,x_n$ is again an algebra over $\mathbb{F}$. Its main use to us will arise from the fact that many interesting algebras are best presented as its quotients:

(i) The *quantum plane* is generated by variables x, y and q subject to the relations
$$yx - qxy = 0, \quad xq - qx = 0, \quad yq - qy = 0.$$

(ii) The *Weyl algebra* is the quotient of the non-commuting algebra generated by variables x and d, subject to the relation $dx - xd = 1$. (The algebra in (d) generated by D and M is a quotient of this.)

(iii) The *enveloping algebra* of the Lie algebra $\mathrm{sl}(2,\mathbb{F})$, generated by elements x, y and h subject to:
$$xy - yx = h, \quad hx - xh = 2x, \quad hy - yh = -2y.$$

(iv) The *Hecke algebra*: this is generated by elements $T_1, \ldots, T_n$ and q, where q commutes with each T_i and we have the additional relations
$$T_i T_j = T_j T_i, \quad |i - j| > 1$$
and
$$T_i T_{i+1} T_i = T_{i+1} T_i T_{i+1}, \quad T_i^2 = (q-1)T_i + qI.$$

There is one problem that arises when we define an algebra as a quotient, as in (e) above, namely that the algebra might be the trivial algebra (with 0 and 1 as its only elements). This does not happen in the examples above.

11.2 Division Algebras

A *division algebra* is an algebra where each non-zero element is invertible. Examples include any field. Division algebras provide basic building blocks for algebras.

11.2.1 Lemma. *If $\mathscr{D}$ is a division algebra over an algebraically closed field $\mathbb{F}$, then $\mathscr{D} = \mathbb{F}$.*

Proof. Suppose $a \in \mathscr{D}$ and let $\psi(t)$ be the minimal polynomial for right multiplication by a. If we have a factorization
$$\psi = \psi_1 \psi_t$$
then $\psi_1(a)\psi_2(a) = 0$. SInce $\mathscr{D}$ is a division algebra either $\psi_1(a) = 0$ or $\psi_2(a) = 0$. Hence one of ψ_1 and ψ_2 is constant and therefore ψ is irreducible. Since $\mathbb{F}$ is algebraically closed, ψ must be linear and therefore $A \in \mathbb{F}$. □

By considering division algebras over $\mathbb{R}$, we arrive at a very important example, the *quaternions*, which we denote by $\mathbb{H}$.

11.2.2 Theorem. *The only finite dimensional division algebras over the reals are the reals, the complex numbers and the quaternions.*

Proof. Suppose $\mathscr{D}$ is a real division algebra. We first prove that if $d \in \mathscr{D} \setminus \mathbb{R}$, then $\mathbb{R}[d] \cong \mathbb{C}$.

Let ψ be the minimal polynomial over $\mathbb{R}$ for right multiplication by d. As above it must be irreducible. Since an irreducible real polynomial has degree at most two and since $d \notin \mathbb{R}$, we conclude that ψ is quadratic, say

$$\psi(t) = t^2 + at + b.$$

Since ψ is irreducible, $a^2 - 4b < 0$ and as $\psi(d) = 0$ we have

$$(d + \frac{1}{2}a)^2 - \frac{1}{4}a^2 + b = 0.$$

Therefore

$$\frac{(2d + a)^2}{4b - a^2} = -1;$$

since $4b - a^2 > 0$, it follows that $R[d]$ contains an element i such that $i^2 = -1$, and so $R[d] \cong \mathbb{C}$.

Next we show that if i, j are both square roots of -1 in $\mathscr{D}$ and $i \neq \pm j$, then i and j do not commute. For suppose $ij = ji$. Then

$$(ij)^2 = ijij = i^2 j^2 = (-1)^2 = 1$$

and so $ij = \pm 1$. Hence $j = \pm i$.

If $\dim_{\mathbb{R}}(\mathscr{D}) = 2$, then $\mathscr{D} \cong \mathbb{C}$. Assume that $\dim(\mathscr{D}) > 2$ and let i be a square root of -1 in $\mathscr{D}$. Let $\mathbb{C}$ denote the subalgebra $\mathbb{R}[i]$. We prove that the centralizer of $\mathbb{C}$ in $\mathscr{D}$ is $\mathbb{C}$ itself. Suppose $d \in \mathscr{D} \setminus \mathbb{C}$. Then $\mathbb{R}[d] \cong \mathbb{R}[j]$ where $j^2 = -1$ and so d commutes with i if and only if j does. But if $ji = ij$ then $j = \pm i$ and so $d \in \mathbb{C}$. We conclude that d and i do not commute.

Next let T denote the operation of conjugation by i on $\mathscr{D}$; then

$$T(d) := -idi$$

for all d in $\mathscr{D}$. Then

$$T^2(d) = (-i)(-i)di^2 = d$$

and so $T^2 = I$. The minimal polynomial of T is $t^2 - 1$ and since this has simple roots T is diagonalizable, with eigenvalues 1 and-1. If $Td = d$, then $id = di$, so d lies in the centralizer of $\mathbb{C}$,that is, $d \in \mathbb{C}$. If $Td = -d$, then $id = -di$, we say that d and i anticommute. Since T is diagonalizable its eigenvectors span and accordingly each element of $\mathscr{D}$ can be written as $a+b$, where a commutes with i and b anticommutes with i.

Suppose $w \in \mathscr{D}$ and $T(w) = -w$. If we have $T(x) = -x$, then

$$T(xw) = T(x)T(w) = (-x)(-w) = xw$$

and therefore, if $U = \ker(T + I)$, then

$$Uw \in \ker(T - I) = \mathbb{C}.$$

Since $\mathscr{D}$ is a division algebra, right multiplication by w is injective, and therefore $\dim(\ker(T+I)) \leq 2$. As $w^2 \in \ker(T-I)$, we see that $\mathbb{C}w \leq U$ and so $Uw = \mathbb{C}$. Since $\mathscr{D}$ is the direct sum of the eigenspaces $\ker(T-I)$ and $\ker(T+I)$, we conclude that $\dim(\mathscr{D}) = 4$.

We leave the rest of the proof as an exercise, but offer that if $\dim(\mathscr{D}) = 4$, then $\mathscr{D}$ must contains elements i and j such that

$$i^2 = j^2 = -1, \ ji = -ij.$$

The rest should not be too difficult. □

The above proof is based on the one in Farenick [], which he ascribes to ???.

If we set k equal to ij, then

$$k^2 = i(ji)j = -i^2 j^2 = -1.$$

Similarly

$$ki = iji = -i^2 j = j, \quad ik = i^2 j = -j.$$

and

$$jk = jij = -ij^2 = i, \quad kj = ij^2 = -i.$$

As

$$(a + bi + cj + dk)(a - bi - cj - dk) = a^2 - (bi + cj + dk)^2 = a^2 + b^2 + c^2 + d^2,$$

the elements 1, i, j and k are linearly independent over $\mathbb{R}$.

11.3 Maps and Modules

If $\mathscr{A}$ and $\mathscr{B}$ are algebras over R, then an algebra homomorphism from $\mathscr{A}$ to $\mathscr{B}$ is a ring homomorphism from $\mathscr{A}$ to $\mathscr{B}$ that commutes with the respective maps ψ_A and ψ_B from R into the centres of $\mathscr{A}$ and $\mathscr{B}$. If R is a field, an algebra homomorphism is a ring homomorphism that is also an R-linear map from $\mathscr{A}$ to $\mathscr{B}$.

We will generally use 'map' and 'homomorphism' as synonyms for 'algebra homomorphism'. Unless explicitly stated otherwise any module is Artinian and Noetherian (in fact an Artinian module must be Noetherian).

If $\mathscr{A}$ is an algebra over R, then a module for $\mathscr{A}$ consists of an R-module, M say, and a homomorphism from $\mathscr{A}$ in $\mathrm{End}(M)$ (which is also an algebra over R). It is conventional to avoid explicit mention of the homomorphism; thus if $x \mathrm{im}\, M$ and $a \in \mathscr{A}$, then the image of x under the action of a is denoted by xa. When ψ is a module homomorphism, we may also use $x\psi$ to denote the image of x under ψ. (The reasons for placing a and ψ on the right are given in the next section.) When R is a field, any module for $\mathscr{A}$ will be a vector space over R. In the absence of an explicit warning, modules over algebras will have finite dimension.

If M and N are R-modules, then $\mathrm{Hom}(M,N)$ is the set of all homomorphisms from M to N. If $M = N$, we write $\mathrm{End}(M)$ in place of $\mathrm{Hom}(M,M)$. If there is a homomorphism ρ from the algebra $\mathscr{B}$ to the algebra $\mathscr{A}$ and M is a module over $\mathscr{A}$, then it is also a module over $\mathscr{B}$: the composite of the homomorphism from $\mathscr{B}$ to $\mathscr{A}$ with the homomorphism from $\mathscr{A}$ into $\mathrm{End}(M)$ makes M into a $\mathscr{B}$-module. For example suppose V is a vector space and $A \in \mathrm{End}(V)$. Then V is a module for the algebra $\mathbb{F}[A]$ of all polynomials in A. This algebra is a homomorphic image of the polynomial ring $\mathbb{F}[t]$, where the homomorphism maps t to A. So V is also a module for $\mathbb{F}[t]$—if $v \in V$ then $vt := vA$.

If M and N are also $\mathscr{A}$-modules then $\mathrm{Hom}_A(M,N)$ denotes the set of homomorphisms from M to N that commute with the action of $\mathscr{A}$. Thus if $\psi \in \mathrm{Hom}(M,N)$, then $\psi \in \mathrm{Hom}_A(M,N)$ if

$$u\psi a = ua\psi$$

for all u in M and all a in $\mathscr{A}$. Both $\mathrm{Hom}(M,N)$ and $\mathrm{Hom}_A(M,N)$ are R-modules, but only $\mathrm{Hom}_A(M,N)$ is an $\mathscr{A}$-module in general.

Any algebra $\mathscr{A}$ can be viewed as a module over itself. The mapping from $\mathscr{A}$ to $\mathrm{End}(\mathscr{A})$ assigns to a in the linear mapping ρ_a, given by

$$x\rho_a := xa.$$

This gives the *regular module* for $\mathscr{A}$. We note that its submodules correspond to the right ideals of $\mathscr{A}$. We will make great use of the regular module of an algebra.

In cases where $\mathscr{A}$ is generated by some subset S, we might write $\mathrm{Hom}_S(M,N)$ or $\mathrm{End}_S(M)$ rather than $\mathrm{Hom}_{\mathscr{A}}(M,N)$ or $\mathrm{End}_{\mathscr{A}}(M)$. The opportunity to do this will arise when $\mathscr{A}$ is $\mathbb{F}[A]$, the algebra of polynomials over $\mathbb{F}$ in the operator A, or when it is $\mathbb{F}[G]$, the group algebra of the group G.

11.4 Opposites

We use $\mathscr{A}^{\mathrm{op}}$ to denote the algebra with the same set of elements as A, and multiplication $*$ defined by

$$a * b := ba,$$

for all a and b. We call it the *opposite* algebra of A. We say that M is a left module over A if it is a right module over $\mathscr{A}^{\mathrm{op}}$. By way of example, if $a \in \mathscr{A}$ we define the element λ_a of $\mathrm{End}(\mathscr{A})$ by

$$x\lambda_a = ax.$$

Then $\lambda_a\lambda_b$ maps x to bax and therefore $\lambda_a\lambda_b = \lambda_{ba}$. Thus we have a homomorphism of $\mathscr{A}^{\mathrm{op}}$ into $\mathrm{End}(\mathscr{A})$.

The point here is that to compute the image of x under $\psi\varphi$, we first apply ψ, then φ. Effectively we are viewing ψ and φ as operations, and as customary we apply the leftmost operation first. It we take this approach, then right modules are more natural than left modules.

An algebra $\mathscr{A}$ is a module over the tensor product $\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}$. Here if

$$a \otimes b \in \mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}$$

and $x \in \mathscr{A}$, then

$$a \otimes b : x \mapsto axb.$$

Note that

$$axb = x\lambda_a\rho_b = x\rho_b\lambda_x;$$

here ρ_a and λ_b commute because $a(xb) = (ax)b$. The submodules of $\mathscr{A}$ correspond to the ideals of $\mathscr{A}$.

11.4.1 Lemma. *If $\mathscr{A}$ is an algebra, then* $\mathrm{End}_{\mathscr{A}}(\mathscr{A}) \cong \mathscr{A}^{\mathrm{op}}$.

Proof. If $\mu \in \mathrm{End}_{\mathscr{A}}(\mathscr{A})$ and $m = 1\mu$, then

$$x\mu = (1x)\mu = (1\mu)x = mx.$$

Thus $\mu = \lambda_{1\mu} \in \mathscr{A}^{\mathrm{op}}$. □

We use $Z(\mathscr{A})$ to denote the centre of $\mathscr{A}$.

11.4.2 Lemma. *If $\mathscr{B} = \mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}$, then $\mathrm{End}_{\mathscr{B}}(\mathscr{A}) = Z(\mathscr{A})$.*

Proof. Exercise. □

The $(\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A})$-submodules of $\mathscr{A}$ are its ideals.

11.5 Schur's Lemma

A module is *simple* if it has no non-zero proper subodules. If M is a module over an algebra and M has dimension one (as a vector space) then M is simple. More generally any if M is a module over and algebra and N is a non-zero submodule with the least possible dimension, then N is simple. Thus simple modules are not hard to find.

Despite its brevity, the next result will prove to be extremely powerful.

11.5.1 Lemma. *If M and N are simple modules over $\mathscr{A}$, then $\mathrm{Hom}_{\mathscr{A}}(M,N)$ is a division ring if M and N are isomorphic, and is zero otherwise.*

Proof. Suppose $x \in \mathrm{Hom}_{\mathscr{A}}(M,N)$. The kernel and range of x are submodules of M. If $x \neq 0$ then the kernel of x must zero, consequently x is injective. Now it follows that x must be surjective. Therefore either $x = 0$ or x is invertible; we conclude that $\mathrm{Hom}_{\mathscr{A}}(M,N)$ is a division ring. □

This result provides the reason why we need information about division algebras. If $\mathscr{A}$ is an algebra over $\mathbb{F}$ then any division ring D obtained by appeal to Schur's lemma will have $\mathbb{F}$ in its centre.

11.5.2 Corollary. *If $e\mathscr{A}$ is a simple submodule of $\mathscr{A}$ and $a \in \mathscr{A}$, then $ae\mathscr{A}$ is either zero or isomorphic to $e\mathscr{A}$.*

Proof. The subset

$$\{x \in e\mathscr{A} : ax = 0\}$$

is a submodule of $e\mathscr{A}$ and, since $e\mathscr{A}$ is simple, it is either zero or equal to $e\mathscr{A}$. In the latter case $ae\mathscr{A} = 0$, in the former left multiplication by a is injective and is therefore an isomorphism. □

If M is an $\mathscr{A}$-module and $m \operatorname{im} M$, then $m\mathscr{A}$ is a submodule of M We say M is *cyclic* if there is an element m in M such that $m\mathscr{A} = M$. A simple module is automatically cyclic. If M is an $\mathscr{A}$-module, then the set

$$\operatorname{ann}(M) := \{a \in \mathscr{A} : Ma = 0\}$$

is called the *annihilator* of M. It is a submodule of $\mathscr{A}$ and, if M is cyclic you may show that $\mathscr{A}/\operatorname{ann}(M) \cong M$. (Note that $\mathscr{A}/\operatorname{ann}(M)$ is a module, and not a quotient algebra.) You might also show that M is simple if and only if $\operatorname{ann}(M)$ is a maximal proper submodule of $\mathscr{A}$.

Chapter 12

Semisimple Modules

12.1 Summands and Idempotents

One of our basic tasks will be to express a module for an algebra as a direct sum of simple modules. You will have already had experience with this, when you found bases of eigenvectors for linear mappings in a linear algebra course. For suppose A is a symmetric real $v \times v$ matrix. Then $V = \mathbb{R}^v$ is a module for the algebra $\mathbb{R}[A]$ of all polynomials in A. Any eigenvector for A spans a 1-dimensional subspace invariant under A, and hence its span is a simple module for $\mathbb{R}[A]$. Now a set of vectors $x_1, \ldots, x_v$ is linearly independent if and only if V is the direct sum of the 1-dimensional subspaces $\langle x_i \rangle$. In other words, A is diagonalizable if and only if V is the direct sum of v simple $\mathbb{R}[A]$-modules.

An element e in a ring $\mathscr{A}$ is *idempotent* if $e^2 = e$. Two idempotents e and f are *orthogonal* if $ef = fe = 0$. If e is an idempotent then $1 - e$ is an idempotent orthogonal to e. Idempotents provide an approach to direct sum decompositions of modules.

Suppose M is a module over $\mathscr{A}$ and $M \cong M_1 \oplus M_2$. If $m \in M$ then there are unique elements m_1 in M_1 and m_2 in M_2 such that $m = m_1 + m_2$. It follows that the maps e_i given by

$$e_i : m \mapsto m_i$$

are endomorphisms of M. Since M_i is a submodule of M, it follows that $m_i a \in M_i$, for all a in A. Therefore

$$m_i a e_i = m_i a = m_i e_i a,$$

from which it follows that $a e_i = e_i a$, for all a. Consequently $e_i \in \mathrm{End}_{\mathscr{A}}(M)$. As

$m_i e_i = m_i$, we have $e_i^2 = e_i$. Also $e_1 + e_2 = 1$. Conversely, each idempotent e distinct from 0 and 1 in $\mathrm{End}_{\mathscr{A}}(M)$ determines a direct sum decomposition of M with summands Me and $M(1-e)$.

A module M is *indecomposable* if it cannot be expressed as a direct sum of proper submodules. Hence M is indecomposable if and only if 0 and 1 are the only idempotents in $\mathrm{End}_{\mathscr{A}}(M)$. If e is an idempotent in $\mathrm{End}_{\mathscr{A}}(M)$ and N is a summand of Me then there must be an idempotent f in $\mathrm{End}_{\mathscr{A}}(M)$ such that $ef = fe = e$. Then $e - f$ is idempotent and is orthogonal to f. We call an idempotent *primitive* if it cannot be written as the sum of two orthogonal non-zero idempotents. Thus primitive idempotents in $\mathrm{End}_{\mathscr{A}}(M)$ correspond to indecomposable summands of M.

12.1.1 Lemma. *If e is an idempotent in the algebra $\mathscr{A}$ and the right ideal $e\mathscr{A}$ is minimal, then e is primitive.*

Proof. Suppose $e = f + g$ where f and g are idempotents and $fg = gf = 0$. Since $fe = f$ we see that $f\mathscr{A} \leq e\mathscr{A}$. Since $f\mathscr{A}$ contains f, it is not zero and since $g(f\mathscr{A}) = 0$, it is not equal to $e\mathscr{A}$. Hence $e\mathscr{A}$ is not minimal. □

As $\mathrm{End}_{\mathscr{A}}(\mathscr{A}) \cong A^{\mathrm{op}}$, it follows that summands of the regular module for $\mathscr{A}$ correspond to idempotents in $\mathscr{A}^{\mathrm{op}}$, and hence to idempotents in $\mathscr{A}$ itself. In concrete terms, if $e \in \mathscr{A}$ and $e^2 = e$ then $\mathscr{A} = e\mathscr{A} \oplus (1-e)A$. Our next result will be useful when we want to find idempotents in an algebra.

12.1.2 Lemma. *Let I be a minimal right ideal of the algebra $\mathscr{A}$. If $I^2 \neq 0$ then there is an idempotent e in $\mathscr{A}$ such that $I = eA$.*

Proof. As $I^2 \neq 0$, there is an element x in I such that $xI \neq 0$. We note that xI is right ideal, as $xI \subseteq I^2 \subseteq I$ and as I is minimal, $xI = I$. Because $x \in I$, there must be an element e in I such that $xe = x$. Then $xe^2 = xe$ and $x(e^2 - e) = 0$.

The set J of elements b in I such that $xb = 0$ is a right ideal contained in I. Since it does not contain y, it is properly contained in I; since I is minimal, $J = 0$. Therefore $e^e - e = 0$.

Finally, eI is a right ideal contained in I and is not zero, accordingly $eI = I$. □

An element x in a ring is *nilpotent* if $x^n = 0$ for some n. An ideal I is *nilpotent* if $I^n = 0$ for some n. Clearly each element of a nilpotent ideal is nilpotent, but the converse is not always true. If I is a nilpotent right ideal then $\mathscr{A}I$ is a nilpotent ideal, because

$$(\mathscr{A}I)^r = \mathscr{A}(I\mathscr{A})^{r-1}I \subseteq AI^r.$$

Thses comments will become significant when we characterize semisimple algebras.

If M is a right ideal in $\mathscr{A}$, you may show that

$$M \cap \mathrm{ann}(M)$$

is a nilpotent right ideal in $\mathscr{A}$.

As $\mathrm{End}_{\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}}(\mathscr{A}) \cong Z(A)$, we see that summands of $\mathscr{A}$, viewed as a module over $\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}$, correspond to idempotents in $Z(\mathscr{A})$; these are often called *central* idempotents. Here submodules of $\mathscr{A}$ are ideals of $\mathscr{A}$; each ideal may be viewed as an algebra, with the central idempotent determining the summand as its identity element.

12.2 Primary Decomposition

Let T be an endomorphism of the vector space V, and suppose that the minimal polynomial of A is $\psi(t)$. We derive a direct sum decomposition of the T-module V for each factorization of ψ into coprime factors.

Suppose

$$\psi = fg,$$

where f and g are coprime. Then there are polynomials a and b such that

$$af + bg = 1$$

and hence

$$a(T)f(T) + b(T)g(T) = I.$$

Multiply both sides of this by $a(T)f(T)$. Then we find that

$$a(T)f(T) = (a(T)f(T))^2 + a(T)b(T)f(T)g(T);$$

since $fg = \psi$ we have $f(T)g(T) = 0$ and therefore $E = a(T)f(T)$ is idempotent. Similarly

$$b(T)g(T) = I - a(T)f(T)$$

is an idempotent, orthogonal to $a(T)f(T)$. Hence EV and $(I-E)V$ are complementary T-submodules of V. You may show that the minimal polynomial of the restriction of T to EV is $f(T)$.

More generally, if ψ factors into k pairwise coprime factor, we may express V as the direct sum of k submodules, one for each factor. (This is easily proved by induction on the number of factors.)

If our underlying field is algebraically closed, we may write

$$\psi(t) = \prod_{i=1}^{k} (t-\theta_i)^{m_i}$$

where $\theta_1, \ldots, \theta_k$ are the zeros of ψ, or equivalently, the eigenvalues of T. We derive orthogonal idempotents $E_1, \ldots, E_k$ such that

$$E_1 + \cdots + E_k = I$$

and V is the direct sum of submodules $E_i V$. This decomposition of V is called the *primary decomposition*. The minimal polynomial of $T \restriction (E_i V)$ is $(t-\theta_i)^{m_i}$; the elements of $E_i V$ are known as *root vectors* or *generalized eigenvectors* for T, with eigenvalue θ_i. Over an algebraically closed field, the existence of the primary decomposition can be paraphrased by the assertion that if $T \in \mathrm{End}(V)$, there is a basis of V that consists of root vectors for T.

If the minimal polynomial of T factors into linear terms over $\mathbb{F}$ and has no repeated factors, then the summands in the primary decomposition are eigenspaces for T. Thus V has a basis of eigenvectors for T in this case (or equivalently T is diagonalizable).

12.3 Group Algebras

We apply Schur's lemma to derive some properties of group algebras.

If M is a module over $\mathscr{A}$, the action of $\mathscr{A}$ on M is given by a homomorphism, φ say, from A into End. If $a \in \mathscr{A}$ then $\varphi(a)$ is a linear mapping and the function that maps a in A to $\mathrm{tr}(\varphi(a))$ is called a *character* of G. The homomorphism φ is called a representation of $\mathscr{A}$.

If M is a module for the group algebra $\mathbb{F}[G]$, we use M^G to denote the subspace of M consisting of the elements m such that $mg = m$ for all g in G. This is clearly a submodule of M.

12.3.1 Lemma. *If M is a module over the group algebra $\mathbb{F}[G]$ determined by the representation φ then*

$$\frac{1}{|G|} \sum_{g \in G} \mathrm{tr}\, \varphi(g) = \dim M^G.$$

Proof. Define σ by

$$\sigma := \frac{1}{|G|} \sum_{g \in G} g.$$

Then $\sigma^2 = \sigma$ and hence M can be written as the direct sum of the subspaces $M\sigma$ and $M(1-\sigma)$. As $m \in M^G$ if and only if $m\sigma = m$, we see that $M\sigma = M^G$. As

$$g\sigma = \sigma g = \sigma,$$

for all g in G, both $M\sigma$ and $M(1-\sigma)$ are submodules of M. Suppose that $n = \dim M$. The characteristic polynomial of σ (or equivalently, of $\varphi(\sigma)$) is $(t-1)^d t^{n-d}$, where d is the rank of $\varphi(\sigma)$. Therefore this rank is equal to the trace of $\varphi(\sigma)$. As the rank of $\varphi(\sigma)$ is the dimension of the range of $\varphi(\sigma)$, the lemma is proved. □

12.3.2 Lemma. *If φ and ψ are representations of G, corresponding to modules M and N, then*

$$\frac{1}{|G|} \sum_{g \in G} \operatorname{tr}\varphi(g) \operatorname{tr}\psi(g^{-1}) = \dim \operatorname{Hom}_G(M, N).$$

Proof. The mapping that sends g to $\varphi(g) \otimes \psi(g^{-1})$ is a homomorphism from A into $\operatorname{End}(M \otimes N^*)$. Apply the previous lemma to this homomorphism. □

Now we make use of Schur's lemma. A representation of a group algebra is *irreducible* if the corresponding module is simple. If M and N are simple modules then $\operatorname{Hom}_{\mathscr{A}}(M, N)$ is zero if $M \not\cong N$. If $M \cong N$ then $\operatorname{Hom}_{\mathscr{A}}(M, N)$ is a division ring containing $\mathbb{F}$. If $\mathbb{F}$ is algebraically closed, $\mathbb{F} = \operatorname{Hom}_{\mathscr{A}}(M, N)$ and so $\operatorname{Hom}_{\mathscr{A}}(M, N)$ is 1-dimensional. Now Lemma 12.3.2 reduces to the orthogonality relation for irreducible characters of a finite group.

Furthermore, the element $\varphi(g) \otimes \psi(g^{-1})$ is an endomorphism of $M \otimes N^*$, and the sum

$$\sum_{g \in G} \varphi(g) \otimes \psi(g^{-1}) \tag{12.3.1}$$

lies in $\operatorname{Hom}_{\mathscr{A}}(N, M)$. Suppose $\mathbb{F}$ is algebraically closed and M and N are simple. Then $\dim \operatorname{Hom}_{\mathscr{A}}(N, M) \leq 1$, with equality if and only if M and N are isomorphic. Hence we infer that the sum in (12.3.1) is zero if $M \not\cong N$, and equals cI if $M \cong N$. Since we have the trace of this sum, it is not too hard to show that

$$c = \frac{|G|}{\dim M}.$$

(And it can be shown that this ratio must be an integer.) One consequence of these deliberations is that, if $M \not\cong N$ then the coordinate functions $\varphi_{i,j}$ are orthogonal to the coordinate functions $\psi_{k,\ell}$; if $M \cong N$ then $\varphi_{i,j}$ and $\varphi_{k,\ell}$ are orthogonal unless $i = k$ and $j = \ell$.

If M is simple and has dimension d, it follows that the space spanned by the matrices $\varphi(g)$, for g in G, has dimension d^2. Therefore the algebra $\varphi(\mathscr{A})$ is the algebra of all $d \times d$ matrices over $\mathbb{F}$. This result will be extended to semisimple algebras in later sections.

12.4 Semisimple Modules

Next we offer a characterization of when a module M may be written as a direct sum of some of its submodules. If $M_1, \ldots, M_r$ are submodules of M, we use

$$\sum_{i=1}^{r} M_r$$

to denote their sum—this is the submodule of M formed by those elements which can be written as finite sum of elements from $\cup_i M_i$.

12.4.1 Lemma. *Let $M_1, \ldots, M_r$ be submodules of M. Then the following assertions are equivalent:*

(a) $M \cong \oplus_{i=1}^{r} M_i$,

(b) $M = \sum_i M_i$ *and* $M_i \cap \sum_{j \neq i} M_j = 0$,

(c) $M = \sum_i M_i$ *and* $M_i \cap \sum_{j<i} M_j = 0$. □

The key to the proof of this result is that if $M = \sum_i M_i$ then there is a homomorphism from $\oplus_i M_i$ onto M; the conditions in (b) and (c) are sufficient for this map to be injective. If M_1 is a submodule of M, we say that a submodule M_2 is a *complement* to M_1 if $M_1 + M_2 = M$ and $M_1 \cap M_2 = 0$. Thus M_2 is a complement to M_1 if and only if $M \cong M_1 \oplus M_2$.

This brings us to the most important concept in this section. A module is *semisimple* if it is a direct sum of simple modules. If M is a direct sum of a finite number of simple modules then it is easy to write down a composition series for it, the Jordan-Hölder theorem then implies that any two presentations of M as a sum of simple modules differ only in the order of the terms.

It will not be clear yet that this concept is important; we begin by offering some characterizations.

12.4.2 Theorem. *The following assertions about the module M are equivalent:*

(a) M is semisimple,

(b) M is the direct sum of simple modules $M_1, \ldots, M_r$,

(c) M is the sum of simple modules $M_1, \ldots, M_s$,

(d) Every submodule of M has a complement,

(e) Every simple submodule of M has a complement. □

12.4.3 Corollary. *If M is semisimple and $N \leq M$ then N and M/N are semisimple.* □

12.5 Semisimple Modules: Examples

A vector space of finite dimension over $\mathbb{F}$ has a basis, and tyhe vector space is the direct sum of the 1-dimensional subspaces spanned by the elements of the given basis. Thus a vector space, viewed as a module over $\mathbb{F}$, is a sum of simple modules; hence it is semisimple.

Another way of proving that a vector space V is a semisimple module is to prove that each subspace is the range of an idempotent element of $\mathrm{End}(V)$. We outline the argument. Suppose that U is a subspace of V, with basis $u_1, \ldots, u_r$. This can be extended to a basis $u_1, \ldots, u_n$ of V; let $v_1, \ldots, v_n$ be the corresponding dual basis. (So $y_i^T x_j = \delta_{i,j}$.) If

$$P := \sum_{i=1}^{r} x_i y_i^T$$

then $P^2 = P$ and, since $P y_i = x_i$, the range of P is V.

Suppose T is a self-adjoint operator on the inner product space V, real or complex. If T fixes a subspace U of V then its adjoint fixes $U^\perp$; thus if T is self-adjoint then V is a semisimple module for $\mathbb{F}[T]$.

Continuing with this example, suppose U is a simple submodule of V. Since T commutes with itself, $T \restriction U \in \mathrm{End}_T(U)$. As T is simple, $\mathrm{End}_T(U)$ is a division algebra. If $\mathbb{F} = \mathbb{C}$, then the only division algebra over $\mathbb{F}$ is $\mathbb{C}$. If $\mathbb{F} = \mathbb{R}$, then the division algebras over $\mathbb{R}$ are $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$; since the image of $\mathbb{R}[T]$ in $\mathrm{End}_T(U)$ is

commutative, it must be isomorphic to $\mathbb{R}$ or $\mathbb{C}$. Thus T acts on U as multiplication by a real or complex scalar. Since U is simple, U is 1-dimensional over the appropriate field, or equivalently it is spanned by an eigenvector for T. Thus the fact that T can be diagonalized rests on the observation that V is semisimple and on the classification of the real and complex division algebras. As a bonus, these arguments go through without change for normal operators, since in this case the algebra generated by T and T^* is commutative and *-closed (and the latter implies that V is semisimple).

Next we consider the case where $\mathscr{A} = \mathbb{F}[G]$ is the group algebra of the finite group G. We will prove that, if $|G|$ is invertible in $\mathbb{F}$, then the regular module over A is semisimple. Assume that V denotes the regular module over the group algebra $\mathscr{A}$ and U is a submodule of it. Then there is an idempotent endomorphism P of V with range equal to U. Suppose now that

$$\widehat{P} := \frac{1}{|G|} \sum_{g \in G} g^{-1} P g.$$

If $u \in U$ then, because U is a submodule $ug^{-1} \in U$. Therefore $ug^{-1}\widehat{P} = ug^{-1}$ and $ug^{-1}\widehat{P}g = u$. If $u \in V$ then $ug^{-1}\widehat{P} \in U$ and so $ug^{-1}\widehat{P}g \in U$. This shows that the range of $\widehat{P}$ is U and that $\widehat{P}^2 = \widehat{P}$. If $h \in G$ then a simple computation yields that $h\widehat{P} = \widehat{P}h$, which implies that $\widehat{P} \in \mathrm{End}_{\mathscr{A}}(V)$. It follows that $V(I - \widehat{P})$ is a submodule of V complementary to U. Hence we have proved that V is semisimple.

If $|G|$ is not invertible in $\mathbb{F}$, then V is not semisimple. For if

$$\theta := \sum_{g \in G} g$$

then $\theta^2 = |G|\theta = 0$, but $\theta \neq 0$. Suppose N is a simple submodule of V. Since $\theta \in Z(A)$, we have

$$N\theta\mathscr{A} = N\mathscr{A}\theta = N\theta$$

and hence $N\theta$ is a submodule of N. If $N\theta \neq 0$ then $N\theta$ must equal N, but now we note that

$$0 = N\theta^2 = N\theta = N.$$

If V is semisimple, it follows that $V\theta = 0$; as V is the regular module we deduce in turn that $\theta = 0$. We are forced to conclude that V is semisimple if and only if the characteristic of $\mathbb{F}$ is coprime with $|G|$.

12.6 Indecomposable Modules

Even if a module is not semisimple, we may still present it as a direct sum of indecomposable modules. In this section we develop some relevant theory. (We will not be using this in later sections.)

The endomorphism algebra of a simple module is a division ring. We start by considering some properties of the endomorphism algebra of an indecomposable module. By the theory in Section 12.1, the only idempotents in such an algebra are 0 and 1. This implies in turn (by Lemma 12.1.2) that all minimal right ideals are nilpotent. To get further, we introduce the radical of an algebra.

We define the *radical* $\operatorname{rad}\mathscr{A}$ of the algebra $\mathscr{A}$ to be the set of elements a in $\mathscr{A}$ such that $Ma = 0$ for every simple module M. It is easy to see that $\operatorname{rad}\mathscr{A}$ is an ideal in $\mathscr{A}$.

If I and J are nilpotent ideals of $\mathscr{A}$ then a simple induction argument shows that $(I+J)^r = 0$ when r is large enough. Therefore $I+J$ is nilpotent and so $\mathscr{A}$ always contains a unique largest nilpotent ideal.

12.6.1 Theorem. *The radical of $\mathscr{A}$ is equal to:*

(a) the intersection of all maximal submodules of $\mathscr{A}$,

(b) the largest nilpotent ideal of $\mathscr{A}$.

Proof. We first show that all nilpotent ideals of $\mathscr{A}$ lie in $\operatorname{rad}\mathscr{A}$. Suppose M is a simple module and J is a right ideal of $\mathscr{A}$. Then MJ is a submodule of M, so either $MJ = 0$ or $MJ = M$. In the latter case $MJ^r = M$ for all r; consequently if M is simple and J is nilpotent then $MJ = 0$. Therefore $\operatorname{rad}\mathscr{A}$ contains all nilpotent right ideals of $\mathscr{A}$. We complete the proof of (b) by showing that $\operatorname{rad}\mathscr{A}$ is nilpotent.

Suppose that B and C are submodules of $\mathscr{A}$ with $B \leq C$ and C/B simple. Then $\operatorname{rad}\mathscr{A}$ acts trivially on C/B; accordingly we have

$$C\operatorname{rad}(\mathscr{A}) \subseteq B.$$

It follows that if $\mathscr{A}$ has a composition series of length r then $\mathscr{A}(\operatorname{rad}\mathscr{A})^r = 0$ and so $\operatorname{rad}\mathscr{A}$ is nilpotent.

To prove (a), we note that if B is a maximal submodule of $\mathscr{A}$ then $\mathscr{A}/B$ is simple, implying that

$$\operatorname{rad}(\mathscr{A}) = \mathscr{A}\operatorname{rad}(\mathscr{A}) \subseteq B.$$

Hence rad$\mathscr{A}$ lies in each maximal submodule of $\mathscr{A}$. To complete the argument, let K be the intersection of all maximal submodules of $\mathscr{A}$ and suppose that rad$\mathscr{A}$ is properly contained in K. Then $K/\operatorname{rad}(\mathscr{A})$ acts non-trivially on some simple module M. Choose m in M such that $mK \neq 0$. Then mK is a non-zero submodule of M and so $mK = M$. Choose x in K such that $mx = -m$. Then $m(1+x) = 0$, which implies that $(1+x)\mathscr{A}$ is a proper submodule of $\mathscr{A}$. But $x \in K$ and so x lies in every maximal submodule of $\mathscr{A}$. This shows that there is a maximal submodule of $\mathscr{A}$ that contains both x and $1+x$. Therefore it contains 1, a contradiction which forces us to conclude that $K = \text{rad}\mathscr{A}$. □

12.6.2 Theorem. *If M is an indecomposable Artinian $\mathscr{A}$-module then every element of* $\text{End}_{\mathscr{A}}(M)$ *is invertible or nilpotent.*

Proof. Let $\mathscr{B}$ equal $\text{End}_{\mathscr{A}}(M)$. We show first that if M is Noetherian and $\varphi \in \mathscr{B}$, there is an integer n such that the intersection of the range and kernel of φ^n is zero. For the modules $\varphi^{-r}(0)$ form an increasing chain of submodules of M. As M is Noetherian, there is a least integer n such that

$$\varphi^{-n-1}(0) = \varphi^{-n}(0).$$

If $y = \varphi^n(x)$ for some x in M and $\varphi^n(y) = 0$ then $\varphi^{2n}(x) = 0$. Hence

$$x \in \varphi^{-2n}(0) = \varphi^{-n}(0)$$

and therefore $y = 0$. (Note: this implies that if φ is surjective the it is injective.)

Next we show that if M is Artinian and $\varphi \in \mathscr{B}$ then there is an integer n such that M is the sum of the kernel and range of φ^n. For consider the decreasing sequence $\varphi^r(M)$; as M is Artinian there is a least integer n such that

$$\varphi^{n+1}(M) = \varphi^n(M).$$

Then $\varphi^{2n}(M) = \varphi^n(M)$ and so, if $y \in M$ there is an element z in M such that $\varphi^n(y) = \varphi^{2n}(z)$. Now

$$\varphi^n(y - \varphi^n(z)) = 0,$$

whence $y - \varphi^n(z)$ lies in the kernel of φ^n. (Note: this shows that φ is injective if it is surjective.)

Finally we prove the theorem. If M is Artinian and $\varphi \in \text{End}_{\mathscr{A}}(M)$, there is an integer m such that M is the direct sum of the range and kernel of φ^m. If M is indecomposable we conclude that either M is the kernel of φ^m (and φ is nilpotent) or M is the range of φ^m (and φ is a surjection, hence a bijection too). □

Note that in the above we only need $\mathscr{A}$ to be a ring.

12.6.3 Corollary. *Suppose M is a Noetherian module and a, b are endomorphisms of M. If ab is surjective then a is surjective.*

Proof. If ab is surjective then b is surjective, and hence injective. If $y \in M$ then there is an element z in M such that $yb = zab$; thererore $(y - za)b = 0$ and thus $y = za$, proving that a is surjective. □

One consequence of this corollary is if a and b are elements of a Noetherian ring and $ab = 1$ then $ba = 1$. Our next result is the analog of Schur's lemma for indecomposable modules.

12.6.4 Theorem. *If A is an Artinian algebra and contains no non-trival idempotents, then $A/\operatorname{rad}(A)$ is a division ring.*

Proof. As a first step we prove that, if I is a nilpotent ideal and $a \in A$ such that $a^2 - a \in I$, then there is an idempotent e in A such that $e - a \in I$.

Assume $b = a^2 - a$ and $a' := a + b - 2ab$. Then a and a' commute, $(a')^2 \in I^2$ and

$$(a')^2 = a^2 + 2ab - 4a^2b = a + a' + 2ab - 4ab.$$

The right side above equals a' modulo I^2. Now a simple induction argument produces an idempotent e as required. (We describe this process as lifting an idempotent mod I to an idempotent of A.)

Next we prove that all ideals in A are nilpotent. Assume that 0 and 1 are the only idempotents in A. Then, by Lemma 12.1.2, any minimal right ideal is nilpotent. Let I be a minimal non-nilpotent ideal of A. Then I is not a minimal ideal, let J be an ideal of A maximal subject to lying in I. Then J is nilpotent and I/J is a minimal ideal in A/J. If I/J is nilpotent then I is nilpotent. If I/J is not nilpotent then it contains an idempotent, by the previous paragraph we deduce that A contains a non-trivial idempotent. In either case we have a contradiction.

Finally, suppose $z \in A$. If zA is a proper right ideal of A then, as we have just seen, it must be nilpotent. Consequently, if $z \notin \operatorname{rad} A$ then $zA = A$ and z has a right inverse. Because A is Artinian, it follows that z is invertible. Hence each element of $A/\operatorname{rad}(A)$ is invertible and therefore it is a division ring. □

We can now state and prove the Krull-Schmidt theorem.

12.6.5 Theorem. *Let M be a A-module. Any two presentations of M as a sum of indecomposable A-modules differ only in the order of the summands.*

Proof. Suppose

$$M \cong \oplus_{i=1}^{r} M_i \cong \oplus_{j=1}^{s} N_j$$

are two presentations of M as a sum of indecomposable A-modules. Let π denote projection of M onto M_1. Then there is an index j such that $\pi(N_j) \neq 0$, we may assume without loss that $j = 1$. Let σ_i denote projection of M onto N_1. All these projections lie in $\mathrm{End}_{\mathscr{A}}(M)$ and $1 = \sum_i \sigma_i$.

The restriction to M_1 of the product $\sigma_i\pi$ lies in $\mathrm{End}_{\mathscr{A}}(M_1)$ and so either it is invertible, or it is nilpotent and lies in $\mathrm{rad}(\mathrm{End}_{\mathscr{A}}(M_1))$. If $\sigma_i\pi\restriction M_1$ lies in $\mathrm{rad}(\mathrm{End}_{\mathscr{A}}(M_1))$ for all i then the sum of these maps, which is the restriction of π to M_1, also lies in $\mathrm{rad}(\mathrm{End}_{\mathscr{A}}(M_1))$. But $\pi\restriction M_1$ is an isomorphism, hence there is an index j such that $\sigma_j\pi\restriction M_1$ is an isomorphism. We may assume without loss that $j = 1$.

If $\sigma_1\pi\restriction M_1$ is an isomorphism then $\sigma_1\restriction M_1$ is surjective, by Corollary 12.6.3. Therefore $\pi\restriction M_1$ must be injective and so, from the proof of Theorem 12.6.2, we see that it must be an isomorphism. The kernel of $\sigma_1\pi$ is the kernel of σ_1, as the restriction of $\sigma_1\pi$ to M_1 is injective, the kernel of $\sigma_1\pi$ is disjoint from its range. Therefore

$$M \cong M_1 \oplus \ker(\sigma_1\pi) = M_1 \oplus N_2 \oplus \cdots \oplus N_s$$

and the theorem follows by induction. □

Chapter 13

Semisimple Algebras

13.1 Semisimple Algebras

An algebra is *semisimple* if its regular module is semisimple. We have seen that the regular module of the group algebra $\mathbb{F}[G]$ is semisimple if $|G|$ is invertible in $\mathbb{F}$, hence in this case the group algebra is semisimple. We begin by presenting some alternative characterizations of semisimple algebras.

13.1.1 Theorem. *Let $\mathscr{A}$ be an Artinian algebra. The following assertions are equivalent:*

(a) $\mathscr{A}$ is semisimple,

(b) Every right ideal of $\mathscr{A}$ is of the form $e\mathscr{A}$, where e is an idempotent,

(c) Every non-zero ideal contains a non-zero idempotent,

(d) $\mathscr{A}$ has no non-zero nilpotent ideals,

(e) $\mathscr{A}$ has no non-zero nilpotent right ideals.

Proof. Suppose $\mathscr{A}$ is semisimple. We prove (a) implies (b). Let N be a minimal right ideal. Then N has a complement I and

$$\mathscr{A} = \mathscr{A}^2 = (I+N)^2 = I^2 + N^2 \subseteq I + N^2.$$

Hence $N^2 \neq 0$ and so, by Lemma 12.1.2, we see that $N = e\mathscr{A}$ for some idempotent e. Now let I be a right ideal of $\mathscr{A}$ and let N be a minimal ideal contained in

I. Then $I = I_1 \oplus N$, where I_1 is a right ideal contained in I. By induction $I_1 = f\mathscr{A}$ for some idempotent f, orthogonal to e. Consequently $e+f$ is idempotent and generates I.

Clearly (b) implies (c). If (c) holds and I is an ideal in $\mathscr{A}$ then I contains a non-zero idempotent e. We have

$$e = e^r \in I^r,$$

whence I is not nilpotent. Thus (d) holds. By our remarks at the end of Section 12.1, if $\mathscr{A}$ contains a nilpotent right ideal it contains a nilpotent ideal. So (d) implies (e). Finally it is not hard to see that (e) implies (b), and that (b) implies (a). □

13.1.2 Lemma. *An algebra is semisimple if and only if it has a faithful semisimple module.*

Proof. Suppose M is a faithful semisimple module for $\mathscr{A}$, and that it is the direct sum of simple modules $M_1, \ldots, M_r$. If I is an ideal in $\mathscr{A}$ then for some i, we have $M_i I \neq 0$. As M_i is simple, $M_i I = M_i$ and therfore $M_i I^r = M_i$ for all r. So no ideal of $\mathscr{A}$ is nilpotent. The converse is immediate. □

By way of example let V be a vector space over $\mathbb{F}$ and suppose $\mathscr{A} = \mathrm{End}(V)$. Then V is a faithful simple module for $\mathscr{A}$ and therefore $\mathscr{A}$ is semisimple. In this case, it is less easy to prove directly that the regular module is semisimple.

An algebra is *simple* if it has no proper non-zero ideals. By the theorem above a simple Artinian algebra is semisimple. But, apparently, the Weyl algebra is Noetherian, simple and not semisimple. (So the terminology lacks perfection.)

The next result can be used to show that a given algebra is not semisimple; it offers the advantage that it allows us to work with a commutative algebra.

13.1.3 Lemma. *If $\mathscr{A}$ is semisimple then its center $Z(\mathscr{A})$ is semisimple.*

Proof. We prove the contrapositive. Suppose that $Z(\mathscr{A})$ is not semisimple. Then by Theorem 13.1.1 it contains a non-zero nilpotent ideal, N say. Since N is central, $N\mathscr{A}$ is a non-zero ideal and is nilpotent. Therefore $\mathscr{A}$ is not semisimple. □

We remark that even if $\mathscr{A}$ is semisimple, the algebra $\mathscr{A} \otimes \mathscr{A}^{\mathrm{op}}$ need not be semisimple. It is if $\mathbb{F}$ is perfect, for example if $\mathbb{F}$ is finite or has characteristic zero. (Look up separable algebras in D&K for details.)

13.2 Simple Artinian Algebras

We characterize simple Artinian algebras. The canonical example is $\mathrm{End}(M)$, where M is a finite-dimensional vector space over $\mathbb{F}$. We verify that $\mathrm{End}(M)$ is simple when M is a vector space over a field $\mathbb{F}$.

13.2.1 Theorem. *The algebra* $\mathrm{End}(V)$ *is simple.*

Proof. We identify $\mathrm{End}(V)$ with $\mathrm{Mat}_{n\times n}(\mathbb{F})$ Suppose N is a non-zero ideal in $\mathrm{Mat}_{n\times n}(\mathbb{F})$. If $u, v, w, x \in V$, then uv^T and wx^T belong to $\mathrm{Mat}_{n\times n}(\mathbb{F})$ and so

$$uv^T N wx^T \in N.$$

If $v^T Nu = 0$ for all v and w, then $v^T N = 0$ for all v. It follows that $AN = 0$ for all A in $\mathrm{Mat}_{n\times n}(\mathbb{F})$, and so $N = 0$. Thus for some v and w we have $v^T Nw \neq 0$ and therefore

$$uv^T Nwx^T = u(v^T Nw)x^T$$

consists of scalar multiples of ux^T. Since u and x are arbitrary we find that N contains all matrices with rank one, and therefore $N = \mathrm{Mat}_{n\times n}(\mathbb{F})$. □

If e is an idempotent in e then $e\mathscr{A}e$ is readily seen to be an algebra with identity element e. (Thus it is not a subalgebra of $\mathscr{A}$, because it does not contain 1 in general. However $e\mathscr{A}e + (1-e)\mathscr{A}(1-e)$ is a subalgebra.)

13.2.2 Lemma. *Let e be an idempotent in $\mathscr{A}$. If $\mathscr{A}$ is simple, so is $e\mathscr{A}e$.*

Proof. Suppose that I is a non-zero ideal in $e\mathscr{A}e$ (which does not mean it is an ideal in A). Then $I = eIe$ and so

$$I = e\mathscr{A}eIe\mathscr{A}e = e\mathscr{A}I\mathscr{A}e.$$

Now $\mathscr{A}I\mathscr{A}$ is a non-zero ideal in $\mathscr{A}$, hence it equals $\mathscr{A}$ and therefore $I = e\mathscr{A}e$. □

The argument in the above proof actually shows that if I is an ideal in $e\mathscr{A}e$ then $I = eJe$ for some ideal J of $\mathscr{A}$.

Our next result implies that any isomorphism between summands of the regular module of $\mathscr{A}$ can be realized by left multiplication by elements of $\mathscr{A}$.

13.2.3 Lemma. *Let e and f be idempotents in the algebra $\mathscr{A}$ and suppose that $e\mathscr{A} \cong f\mathscr{A}$. Then there are elements a and b in $\mathscr{A}$ such that $ae\mathscr{A} = f\mathscr{A}$ and $bf\mathscr{A} = e\mathscr{A}$.*

Proof. Let ψ be an isomorphism from $e\mathscr{A}$ to $f\mathscr{A}$. Then there are elements x and y in $\mathscr{A}$ such that

$$\psi(e) = fx, \quad \psi^{-1}(f) = ey.$$

Then

$$e = \psi^{-1}(fx) = \psi^{-1}(f \cdot fx) = \psi^{-1}(f)fx = eyfx$$

and, similarly, $f = fxey$. If

$$a := fxe, \quad b := eyf,$$

it follows that left multiplication by b is an isomorphism from $e\mathscr{A}$ to $f\mathscr{A}$ and left multiplication by a is an isomorphism from $f\mathscr{A}$ to $e\mathscr{A}$. As $ab = e$ and $ba = f$, these isomorphisms from an inverse pair. □

13.2.4 Lemma. *Let $\mathscr{A}$ be an algebra, let M be a simple submodule of $\mathscr{A}$ and let I be the sum of all submodules of $\mathscr{A}$ isomorphic to M. Then I is an ideal.*

Proof. Let $\mathscr{M}$ denote the set of all submodules of $\mathscr{A}$ isomorphic to M. Let e be a primitive idempotent such that $M = e\mathscr{A}$ and let f be a second idempotent. If $f\mathscr{A} \cong e\mathscr{A}$ then, by Lemma 13.2.3, there is an element a in $\mathscr{A}$ such that $f\mathscr{A} = ae\mathscr{A}$. Consequently every simple submodule in $\mathscr{M}$ lies in the ideal $\mathscr{A}e\mathscr{A}$. Clearly $\mathscr{A}e\mathscr{A}$ is a sum of submodules of the form $ae\mathscr{A}$, and by Corollary 11.5.2 these are all isomorphic to $e\mathscr{A}$. □

The proof of the next result is left as an exercise, it is a modest generalization of the fact that

$$\mathrm{End}(\mathbb{R}^n) \cong \mathrm{Mat}_{n\times n}(\mathbb{R}).$$

13.2.5 Lemma. *If M is an $\mathscr{A}$-module, then* $\mathrm{End}_{\mathscr{A}}(M^n) \cong \mathrm{Mat}_{n\times n}(\mathrm{End}_{\mathscr{A}}(M))$. □

13.2.6 Theorem. *Let $\mathscr{A}$ be a simple Artinian algebra over $\mathbb{F}$. Then $\mathscr{A} \cong \mathrm{Mat}_n(D)$, where D is a division ring over $\mathbb{F}$.*

Proof. If $\mathscr{A}$ is a division algebra we have nothing to do, so assume is not a division algebra. Then it contains a non-zero element u which is not invertible and therefore $u\mathscr{A}$ is proper non-zero right ideal. Since $\mathscr{A}$ is Artinian we conclude that $\mathscr{A}$ contains a minimal right ideal $e\mathscr{A}$, where e is a primitive idempotent. Because A is simple it cannot contain a non-zer nilpotent ideal, and hence it is semisimple. Now Lemma 13.2.4 implies that A is the sum of submodules

isomorphic to eA therefore A is isomorphic to a direct sum of copies of eA. Consequently

$$A^{\mathrm{op}} \cong \mathrm{End}_{\mathscr{A}}(\mathscr{A}) \cong \mathrm{End}_{\mathscr{A}}((e\mathscr{A})^n) = \mathrm{Mat}_{n\times n}(\mathrm{End}_{\mathscr{A}}(e\mathscr{A})).$$

Since eA is simple, $\mathrm{End}_{\mathscr{A}}(e\mathscr{A})$ is a division ring by Schur's lemma. □

Lemma 13.2.4 implies that, if $\mathscr{A}$ is a simple Artinian algebra and M is a simple submodule of $\mathscr{A}$ then $\mathscr{A}$ is isomorphic to a direct sum of copies of M. It follows from the Jordan-Hölder theorem (Theorem 13.3.1) that a simple Artinian algebra has a single isomorphism class of simple modules.

13.2.7 Corollary. *If $\mathscr{A}$ is a simple Artinian algebra and M and N are modules over $\mathscr{A}$ then M and N are isomorphic if and only if they have the same dimension.* □

13.2.8 Lemma. *If $\mathscr{A}$ is semisimple, any simple $\mathscr{A}$-module is isomorphic to a submodule of $\mathscr{A}$.*

Proof. Let M be a non-zero simple A-module. As A is semisimple, it is a direct sum of simple submodules and, as $M \neq 0$, it follows that there is a simple submodule I such that $MI \neq 0$. Hence there is an element m of M such that $mI \neq 0$. Then mI is a non-zero submodule of M and therefore $mI = M$. Consequently $M \cong I$. □

We conclude with some exercises. If e is an idempotent in $\mathscr{A}$ and M is an $\mathscr{A}$-module, show that

$$\mathrm{Hom}_{\mathscr{A}}(e\mathscr{A}, M) \cong Me.$$

(This is an isomorphism of vector spaces.) Consequently, if e and f are idempotents then

$$\mathrm{Hom}_{\mathscr{A}}(e\mathscr{A}, f\mathscr{A}) \cong f\mathscr{A}e.$$

Show further that $\mathrm{Hom}_{\mathscr{A}}(e\mathscr{A}, e\mathscr{A})$ and $(e\mathscr{A}e)^{\mathrm{op}}$ are isomorphic as algebras.

13.3 Composition Series

Let M be a module over A. A chain of submodules of M:

$$M = M_0 \geq M_1 \geq \cdots \geq M_r = 0$$

is a *composition series* of *length* r if each quotient M_i/M_{i+1} is simple and not zero. Thus a module has positive length, and its length is one if and only if it is simple. If $\mathscr{A} = \mathbb{F}$ then the length of M is just its dimension over $\mathbb{F}$.

The following result is known as the Jordan-Hölder theorem. The proof is an exercise in the use of the isomorphism theorems.

13.3.1 Theorem. *Let M be a module over $\mathscr{A}$ that is Artinian and Noetherian. Then M has a composition series; further any two such series have the same length and the number of times a given simple module occurs in a series is independent of the choice of series.* □

13.3.2 Corollary. *Let M be a module over $\mathscr{A}$ with finite length. If N is a submodule of M, the sum of the lengths of M/N and N is the length of M.* □

A module M is Noetherian and Artinian if and only it has a composition series of finite length. It can be shown that any Artinian module is Noetherian. Essentially all the modules we discuss will be Artinian, because they will be vector spaces of finite dimension.

13.3.3 Lemma. *Let $\mathscr{A}$ be a semisimple algebra over $\mathbb{F}$ and let M be a simple $\mathscr{A}$-module such that* $\mathrm{End}_{\mathscr{A}}(M) \cong \mathbb{F}$. *Then the multiplicity of M in a composition series for $\mathscr{A}$ is* $\dim M$.

Proof. Suppose e is an idempotent in $\mathscr{A}$ and $\psi \in \mathrm{Hom}_{\mathscr{A}}(e\mathscr{A}, M)$. If $\psi(e) = m$ then

$$m = \psi(e) = \psi(e^2) = \psi(e)e = me$$

and

$$\psi(ea) = \psi(e)a = mea.$$

Therefore $\psi(e\mathscr{A}) = me\mathscr{A}$.

Now assume that e is a primitive idempotent. Then $e\mathscr{A}$ is a simple submodule of $\mathscr{A}$ and, if $\psi \neq 0$ then $me\mathscr{A}$ is submodule of M isomorphic to $e\mathscr{A}$. As this holds for all m in M such that $me\mathscr{A} \neq 0$, we see that $Me\mathscr{A}$ is the sum of submodules isomorphic to $e\mathscr{A}$. Since $\mathscr{A}$ is semisimple, M is too, and we conclude that $Me\mathscr{A}$ is isomorphic to $(e\mathscr{A})^r$, for some integer r. Consequently

$$\mathrm{Hom}_{\mathscr{A}}(e\mathscr{A}, M) = \mathrm{Hom}_{\mathscr{A}}(e\mathscr{A}, (e\mathscr{A})^r)$$

and, as $Me \cong \mathrm{Hom}_{\mathscr{A}}(e\mathscr{A}, M)$, we infer that

$$\dim(Me) = r \dim \mathrm{End}_{\mathscr{A}}(e\mathscr{A}).$$

Applying this to the regular module, we find that the multiplicity of $e\mathscr{A}$ as a composition factor in $\mathscr{A}$ is $\dim(\mathscr{A}e)/\dim(\mathrm{End}_{\mathscr{A}}(e\mathscr{A}))$, as required. □

One canonical application of this result is to the group ring $\mathscr{A} = \mathbb{C}[G]$. In this case $\mathscr{A}$ is semisimple and, if V is a simple $\mathscr{A}$-module then $\mathrm{End}_{\mathscr{A}}(V) \cong \mathbb{C}$. It follows immediately that $|G|$, the dimension of $\mathscr{A}$, is equal to the sum of the squares of the dimensions of the distinct simple $\mathscr{A}$-modules.

By way of example, suppose G is the symmetric group on three symbols. We see that $\mathscr{A}$ is either the sum of six pairwise non-isomorphic 1-dimensional modules, or two 1-dimensional modules and two 2-dimensional modules. But each element of $\mathscr{A}$ acts on a 1-dimensional module as a scalar, and so in the first case each element of $\mathscr{A}$ can be represented in its action on the regular module as a diagonal matrix. This implies that $\mathscr{A}$ is commutative, but it is not. Therefore the second alternative holds.

13.4 Semisimple Artinian Algebras

We now derive the fundamental theorem on the structure of semisimple Artinian algebras: Wedderburn's theorem.

13.4.1 Theorem. *A semisimple Artinian algebra is isomorphic to a direct sum of matrix algebras over division rings.*

Proof. Suppose $\mathscr{A}$ is semisimple and Artinian. Then there are primitive idempotents $e_1, \ldots, e_n$ such that

$$\mathscr{A} \cong \oplus_i \, e_i \mathscr{A}.$$

Define two idempotents e and f in $\mathscr{A}$ to be equivalent if $e\mathscr{A} \cong fA$. Let $[e_i]$ denote the set of idempotents e_j $(j = 1, \ldots, n)$ that are equivalent to e_i. We define

$$I := \sum_{e_j \in [e_i]} e_j \mathscr{A}.$$

We show that I is an ideal in $\mathscr{A}$. Suppose $K = ae\mathscr{A} \neq 0$. Then by Corollary 11.5.2 it is isomorphic to $e_i\mathscr{A}$. If K is not contained in I then $K \cap I$ is a proper submodule of K, hence $K \cap I = 0$. Thus, if I is the direct sum of m copies of $e_i\mathscr{A}$ then $K + I$ is isomorphic to the direct sum of $m+1$ copies of $e_i\mathscr{A}$. By the Jordan-Hölder theorem, all composition series for $\mathscr{A}$ contain the same number of copies of the simple module $e_i\mathscr{A}$ and, by our definition of I, this number is m. We conclude that $K \subseteq I$. It follows that I is an ideal in $\mathscr{A}$.

Next we prove I is a minimal ideal. Suppose J is a non-zero ideal of $\mathscr{A}$ contained in I. By the Jordan-Hölder theorem, J must contain a submodule N isomorphic to $e_i\mathscr{A}$. By Lemma 13.2.3 we have $e_i\mathscr{A} = aN$ for some a in $\mathscr{A}$, whence $e_i\mathscr{A} \leq J$ and $J = I$.

Now define

$$J := \sum_{e_j \notin [e_i]} e_j \mathscr{A}.$$

Then because it is a sum of ideals, J is an ideal in $\mathscr{A}$. Since $I \cap J = 0$ we see that I is a summand of $\mathscr{A}$, viewed as a module over $\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}$ and hence $I = f\mathscr{A}$ for some central idempotent f. Consequently it is an algebra (with identity f). As $JI = IJ = 0$, each ideal of I is an ideal of $\mathscr{A}$ and therefore I is a simple algebra. □

It follows that, if $\mathscr{A}$ is semisimple, there are primitive central idempotents $f_1, \ldots, f_r$ such that $\mathscr{A}$ is the direct sum of the algebras $f_i\mathscr{A}$. (Since f_i is central $f_i\mathscr{A} = \mathscr{A} f_i \mathscr{A}$.) The above proof shows that if $e\mathscr{A}$ is a simple submodule of $\mathscr{A}$, then $\mathscr{A} e \mathscr{A}$ is a summand of $\mathscr{A}$, and hence $\mathscr{A} e \mathscr{A} = g\mathscr{A}$ for some central idempotent g. It follows that if f is primitive and central, then $f\mathscr{A}$ is a simple algebra. We conclude the decomposition of a semisimple Artinian algebra $\mathscr{A}$ into simple subalgebras can be computed by first determining the decomposition of $Z(\mathscr{A})$ into simple algebras. In practice this is less useful than it appears, because it is usually not easy to determine the center of $\mathscr{A}$.

A semisimple Artinian algebra is said to be *split* if it is isomorphic to a direct sum of full matrix algebras over its underlying field. The cheapest way of arranging this is to have the underlying field algebraically closed. An algebra is split if the minimal polynomial of each element is linear.

13.4.2 Corollary. *A split commutative semisimple algebra has a basis of idempotents.* □

13.5 Representations

The theory we have developed is very powerful, but even so does not go quite far enough. The issue is that in practice we will not be working just with an algebra, but rather an algebra $\mathscr{A}$ and a module M on which it acts. Our module M will be semisimple and faithful, and so $\mathscr{A}$ will be semisimple. Consequently M decomposes as a sum of simple submodules and $\mathscr{A}$ decomposes into a sum of matrix algebras, the problem is to decide how these two decompositions fit together.

Suppose $\mathscr{A}$ is a semisimple Artinian algebra and M is module for $\mathscr{A}$. Let $e_1, \ldots, e_r$ be central primitive idempotents in $\mathscr{A}$ such that

$$\mathscr{A} = \oplus_{i=1}^{r} e_i \mathscr{A}$$

is the decomposition of $\mathscr{A}$ into simple algebras. Then Me_i is a submodule of M and if $x \in Me_i \cap Me_j$, then since $x \in Me_i$ we have $x = xe_i$ and so if $i \neq j$,

$$xe_j = xe_ie_j = 0.$$

Hence M is the direct sum of the submodules Me_i. Now Me_i is a module for the simple Artinian algebra $e_i\mathscr{A}$, and therefore it is isomorphic to the direct sum of copies of some simple module N_i. Note that N_i is a simple module for $\mathscr{A}$ too. Since N_j is annihilated by e_i and N_i is not, N_i and N_j are not isomorphic as $\mathscr{A}$-modules.

13.6 Centralizers

Suppose M is a vector space over $\mathbb{F}$ and $\mathscr{A}$ is a subalgebra of $\mathrm{End}(M)$; equivalently let M be a faithful $\mathscr{A}$-module. Assume $\mathscr{B} = \mathrm{End}_{\mathscr{A}}(M)$. Thus $\mathscr{B}$ is the centralizer of $\mathscr{A}$ in $\mathrm{End}(M)$ and M is a left $\mathscr{B}$-module. Our goal is to further study the relation between $\mathscr{A}$ and $\mathscr{B}$. We have already seen that the idempotents in $\mathscr{B}$ determine the summands of M (viewed as an $\mathscr{A}$-module). If

$$\mathscr{C} := \mathrm{End}_{\mathscr{B}}(M)$$

then $\mathscr{A}$ is contained in $\mathscr{C}$ and M is a right $\mathscr{C}$-module. As $\mathscr{B} \leq \mathrm{End}_{\mathscr{C}}(M)$, the idempotents of $\mathscr{B}$ determine summands of M, viewed as $\mathscr{C}$-module. Thus any summand of M relative to $\mathscr{A}$ is a summand relative to $\mathscr{C}$.

13.6.1 Lemma. *Suppose that M is a faithful semisimple $\mathscr{A}$-module, $\mathscr{B} = \mathrm{End}_{\mathscr{A}}(M)$ and $\mathscr{C} = \mathrm{End}_{\mathscr{B}}(M)$. If $x_1, \ldots, x_m$ are elements of M and $c \in \mathscr{C}$ then there is an element a in $\mathscr{A}$ such that, for all i,*

$$x_i a = x_i c$$

Proof. As M is semisimple, so is M^n. Hence

$$N = (x_1, \ldots, x_m)\mathscr{A}$$

is a summand of M^n. Suppose

$$\mathscr{B}' := \mathrm{End}_{\mathscr{A}}(M^n), \quad \mathscr{C}' := \mathrm{End}_{\mathscr{B}'}(M^n).$$

Then N is a $\mathscr{C}'$-module. As $\mathscr{C}$ imbeds isomorphically in $\mathscr{C}'$ via the diagonal map $c \mapsto (c, \ldots, c)$, we obtain

$$(x_1, \ldots, x_m)\mathscr{A} = N = N\mathscr{C} = (x_1, \ldots, x_m)\mathscr{A}\mathscr{C} = (x_1, \ldots, x_m)\mathscr{C}.$$

This implies the statement of the lemma. □

13.6.2 Corollary. *If M is a finite dimensional faithful semisimple $\mathscr{A}$-module and $\mathscr{B} = \mathrm{End}_{\mathscr{A}}(M)$ then $\mathrm{End}_{\mathscr{B}}(M) \cong \mathscr{A}$.* □

Note that this corollary implies that if M is faithful and semisimple and $\mathrm{End}_{\mathscr{A}}(M) \cong \mathbb{F}$ then $\mathscr{A} \cong \mathrm{End}_{\mathbb{F}}(M)$. The following result is from Goodman and Wallach [**?**, Section 3.3]

13.6.3 Theorem. *If $\mathscr{A}$ is a semisimple subalgebra of $\mathrm{Mat}_{n\times n}(\mathbb{C})$, then*

$$\mathscr{A} \cong \oplus_{i=1}^{k}(I_{m_i} \otimes \mathrm{Mat}_{d_i\times d_i}(\mathbb{C}))$$

and

$$\mathrm{End}_{\mathscr{A}}(\mathbb{C}^n) \cong \oplus_{i=1}^{k}(\mathrm{Mat}_{m_i\times m_i}(\mathbb{C}) \otimes I_{d_i}).$$ □

We note some consequences.

13.6.4 Theorem. *Let V be a vector space over an algebraically closed field. Any proper subalgebra of $\mathrm{End}(V)$ fixes a non-zero proper subspace of V.*

Proof. Suppose $\mathscr{A}$ is a subalgebra of $\mathrm{End}(V)$. If $\mathscr{A}$ does not fix a proper subspace of V, then V is a simple module for $\mathscr{A}$. Hence $\mathrm{End}_{\mathscr{A}}(V) \cong \mathbb{C}$, and therefore by Corollary 13.6.2, it follows that $\mathscr{A} \cong \mathrm{End}(V)$. □

13.6.5 Corollary. *If V is a vector space over an algebraically closed field $\mathbb{F}$ and $\mathscr{A}$ is a commutative subalgebra of $\mathrm{End}(V)$, there is a basis of V with respect to which $\mathscr{A}$ is triangular.*

Proof. Suppose $\dim(V) = d$. We have to show that there are submodules

$$V_0 \leq V_1 \leq \cdots \leq V_d$$

where $\dim(V_i) = i$ (and so $V_0 = 0$ and $V_d = V$).

We first prove that a commutative subalgebra must fix some subspace of dimension one. Assume $\dim(V) \geq 2$. Then $\mathrm{End}(V)$ is not commutative and so $\mathscr{A}$ is a proper subalgebra. By the theorem it fixes a non-zero proper subspace U of V. If $\dim(U) > 1$, then $\mathscr{A}$ acts as a commutative algebra on U and so by induction we may assume that $\dim(U) = 1$. (Hence U is spanned by a common eigenvector for $\mathscr{A}$.)

Now $\mathscr{A}$ acts as a commutative algebra on the quotient module V/U and by induction again we V/U contains a chain of submodules

$$W_0 \leq \cdots \leq W_{d-1}.$$

Together with U we get the required chain of submodules in V. □

13.7 Trace

A *trace* on an algebra $\mathscr{A}$ is a linear map τ from $\mathscr{A}$ to $\mathbb{F}$ such that $\tau(ab) = \tau(ba)$ for all a, b in $\mathscr{A}$. Thus a trace is an element of the dual vector space $\mathscr{A}^*$. If $a \in \mathscr{A}$, then the map

$$\tau_a : x \mapsto \tau(ax)$$

is also in $\mathscr{A}^*$, and the map

$$a \to \tau_a$$

is a linear map from $\mathscr{A}$ to $\mathscr{A}^*$. We say that τ is non-degenerate if this second map is an injection. There is a second way of looking at this. The map that takes (a, b) in $\mathscr{A} \times \mathscr{A}$ to $\tau(ab)$ is a symmetric bilinear form on $\mathscr{A}$ and τ is non-degenerate if and only if this form is non-degenerate. You might also show that τ is non-degenerate if and only if its kernel does not contain a right ideal of $\mathscr{A}$.

A bilinear form $\langle a, b\rangle$ is *associative* if

$$\langle a, xb\rangle = \langle ax, b\rangle$$

for all a, b, x in $\mathscr{A}$.

We leave the proof of the following result as an exercise.

13.7.1 Lemma. *A bilinear form $\langle a, b\rangle$ arises from a trace if and only if it is symmetric and associative.* □

If $a_1, \ldots, a_d$ is a basis for $\mathscr{A}$, we say that $b_1, \ldots, b_d$ is a *dual basis* if

$$\langle a_i, b_j\rangle = \delta_{i,j}.$$

If $c \in \mathscr{A}$ then

$$c = \sum_i \langle c, b_i\rangle a_i;$$

thus a dual basis provides a cheap way of expressing elements of $\mathscr{A}$ as linear combinations of the basis vectors. If $\mathscr{A}$ admits a non-degenerate symmetric bilinear form then each basis has a dual basis. If a dual basis exists, it is unique.

13.8 Maschke

13.8.1 Lemma. *Let $\mathscr{A}$ be an algebra with a non-degenerate trace* tr*, let $a_1, \ldots, a_d$ be a basis for $\mathscr{A}$ and let $a_1^*, \ldots, a_d^*$ be its dual basis. If M and N are modules over*

$\mathscr{A}$ and $\varphi \in \mathrm{Hom}(M,N)$ and

$$[\varphi] := \sum_i a_i \varphi a_i^*$$

then $[\varphi] \in \mathrm{Hom}_{\mathscr{A}}(M,N)$. Further, $[\varphi]$ is independent of the basis.

Proof. For c in $\mathscr{A}$ we have

$$\begin{aligned} c[\varphi] &= \sum_i c a_i \varphi a_i^* \\ &= \sum_i \sum_j \langle c a_i, a_j^* \rangle a_j \varphi a_i^* \\ &= \sum_j a_j \varphi \sum_i \langle c a_i, a_j^* \rangle a_i^* \end{aligned}$$

Now

$$\langle c a_i, a_j^* \rangle = \langle a_j^*, c a_i \rangle = \langle a_j^* c, a_i \rangle$$

and so the last sum equals

$$\sum_j a_j \varphi \sum_i \langle a_j^* c, a_i \rangle a_i^* = \sum_j a_j \varphi a_j^* c = [\varphi] c.$$

Therefore $[\varphi] \in \mathrm{Hom}_{\mathscr{A}}(M,N)$.

Now suppose $b_1, \ldots, b_d$ is a second basis for $\mathscr{A}$ with dual basis $b_1^*, \ldots, b_d^*$. Then

$$a_i = \sum_j \langle a_i, b_j^* \rangle b_j$$

and, since $b_1, \ldots, b_d$ is a dual basis to $b_1^*, \ldots, b_d^*$,

$$a_k^* = \sum_\ell \langle a_k^*, b_\ell \rangle b_\ell^*$$

Hence

$$\begin{aligned} \delta_{i,k} = \langle a_i, a_k^* \rangle &= \sum_{j,\ell} \langle a_i, b_j^* \rangle \langle a_k^*, b_\ell \rangle \langle b_j, b_\ell^* \rangle \\ &= \sum_j \langle a_i, b_j^* \rangle \langle a_k^*, b_j \rangle. \end{aligned}$$

Now

$$b_i = \sum_j \langle b_i, a_j^* \rangle a_j, \qquad b_i^* = \sum_\ell \langle b_i^*, a_\ell \rangle a_\ell^*$$

and consequently

$$\begin{aligned}\sum_i b_i\varphi b_i^* &= \sum_{i,j,\ell}\langle b_i, a_j^*\rangle\langle b_i^*, a_\ell\rangle a_j\varphi a_\ell^* \\ &= \sum_{j,\ell}\delta_{j,\ell}a_j\varphi a_\ell^* \\ &= [\varphi].\end{aligned}$$

Therefore the value of $[\varphi]$ is independent of the choice of basis. □

Our next result is a version of Maschke's theorem; its form follows Halverson and Ram [].

13.8.2 Theorem. *If the trace map on the regular module for $\mathscr{A}$ is non-degenerate, then any finite-dimensional $\mathscr{A}$-module is semisimple.*

Proof. We show that if M and N are $\mathscr{A}$-modules and $M \leq N$, then M has a complement.

Since M is a subspace of N, there is an idempotent p such that $Np = M$ and $yp = y$ for all y in M. If $p \in \mathrm{End}(N)$ we define

$$[p] := \sum_i a_i p a_i^*$$

and we show that $[p] \in \mathrm{Hom}_{\mathscr{A}}(M,N)$ and $[p]$ is an idempotent such that $N[p] = M$. By Lemma 13.8.1, we know that $[p] \in \mathrm{Hom}_{\mathscr{A}}(M,N)$.

But first we consider the case where $p = 1$. Here we have

$$\mathrm{tr}([1]b) = \sum_i \mathrm{tr}(a_i a_i^* b) = \sum_i \langle ba_i, a_i^*\rangle = \mathrm{tr}(b)$$

and accordingly

$$\mathrm{tr}(([1]-1)b) = 0$$

for all b. Since the trace is non-degenerate, this implies that $[1] = 1$.

Suppose $x \in N$. For all i,

$$xa_i p a_i^* \in M$$

and hence

$$x[p] = \sum_i x\, a_i p a_i^* \in M.$$

So $N[p] \leq M$. Now if $y \in M$, then for all i

$$ya_i^* p = ya_i^*$$

and hence

$$y[p] = \sum_i y\, a_i\, p\, a_i^* = \sum_i y\, a_i\, a_i^* = y[1] = y.$$

Next we observe that

$$[1-p] = [1] - [p] = 1 - [p]$$

is an idempotent in $\mathrm{Hom}_{\mathscr{A}}(M,N)$ that is orthogonal to $[p]$. It follows that $N(1-[p])$ is a complement to M in N. We conclude that N is semisimple. □

We work out the trace for the group algebra $\mathbb{F}[G]$. If $x \in \mathbb{F}[G]$ and $g \in G$, then we use

$$\langle g, x\rangle$$

to denote the coefficient of g in x. Then the trace $\mathrm{tr}(x)$ of x is given by

$$\sum_{g\in G} \langle g, xg\rangle.$$

Since

$$\langle g, xg\rangle = \langle 1, x\rangle$$

it follows that

$$\mathrm{tr}(x) = |G|\langle 1, x\rangle.$$

Now if $g \in G$, then

$$\mathrm{tr}(gx) = |G|\langle 1, gx\rangle = |G|\langle g^{-1}, x\rangle$$

It follows that if $|G| \neq 0$ in $\mathbb{F}$, then $\mathrm{tr}(gx) = 0$ for all g in G if and only if $x = 0$. Thus our trace is non-degenerate if $|G| \neq 0$, and this case $\mathbb{F}[G]$ is semisimple.

Chapter 14

Division Algebras

Among other things, we prove that finite division rings are fields.

14.1 Central Simple Algebras

An algebra $\mathscr{A}$ over a field $\mathbb{F}$ is *central simple* if is simple and $Z(\mathscr{A}) = \mathbb{F}$. By way of example, $\mathrm{Mat}_{n\times n}(\mathbb{F})$ is central simple but $\mathbb{C}$, viewed as an algebra over $\mathbb{R}$, is not. The centre of a simple algebra is always a field. To see this, suppose $z \in Z(\mathscr{A})$. Then $z\mathscr{A} = \mathscr{A}z$ is a 2-sided ideal in $\mathscr{A}$ and so equals $\mathscr{A}$. Thus left multiplication by z is a surjection on $\mathscr{A}$. Next, the set

$$\{a \in \mathscr{A} : za = 0\}$$

is also a 2-sided ideal in $\mathscr{A}$. Thus, if $z \neq 0$, then left multiplication by z is injective. It follows that left multiplication by z is an isomorphism, hence z is invertible.

14.1.1 Lemma. *If $\mathscr{A}$ and $\mathscr{B}$ are algebras then $Z(\mathscr{A} \otimes \mathscr{B}) \cong Z(\mathscr{A}) \otimes Z(\mathscr{B})$.*

Proof. Suppose that $x = \sum_i a_i \otimes b_i \in \mathscr{A} \otimes \mathscr{B}$. It is an easy exercise to show that, if this representation of x has as few non-zero terms as possible then the elements a_i are linearly independent, as are the elements b_i. Suppose then that $x \in Z(\mathscr{A} \otimes \mathscr{B})$ and that

$$x = \sum_i a_i \otimes b_i,$$

where the b_i are linearly independent. Then, for any z in $\mathscr{A}$, $x(z \otimes 1) = (z \otimes 1)x$ and consequently

$$\sum_i (a_i z - z a_i) \otimes b_i = 0.$$

Since the b_i's are linearly independent, this implies that $a_i z - z a_i = 0$ for all z in $\mathscr{A}$, and for all i. Thus $a_i \in Z(\mathscr{A})$ and $x \in Z(\mathscr{A}) \otimes \mathscr{B}$.

Now assume that (14.1) holds with $a_i \in Z(\mathscr{A})$ for all i and that the b_i's are linearly independent. If $y \in \mathscr{B}$ then x commutes with $1 \otimes y$ and therefore

$$y b_i - b_i y = 0$$

for all i. This shows that $b_i \in Z(\mathscr{B})$ and $x \in Z(\mathscr{A}) \otimes Z(\mathscr{B})$. It is evident that $Z(\mathscr{A}) \otimes Z(\mathscr{B}) \subseteq Z(\mathscr{A} \otimes \mathscr{B})$, so we are finished. □

If $\mathscr{A}$ is an algebra over $\mathbb{R}$ with $Z(\mathscr{A}) \cong \mathbb{C}$ then

$$Z(\mathscr{A} \times \mathscr{A}) \cong \mathbb{C} \otimes \mathbb{C} \cong \mathbb{C} \oplus \mathbb{C}.$$

Hence $Z(\mathscr{A} \otimes \mathscr{A})$ is not a field.

14.1.2 Lemma. *An algebra $\mathscr{A}$ is central simple over $\mathbb{F}$ if and only if $\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A} \cong \mathrm{End}(A)$.*

Proof. If $\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A} \cong \mathrm{End}(A)$ then, by the previous lemma,

$$\mathbb{F} \cong Z(\mathrm{End}(A)) \cong Z(\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}) \cong Z(\mathscr{A}^{\mathrm{op}}) \otimes Z(\mathscr{A})$$

and therefore $Z(\mathscr{A}) \cong \mathbb{F}$. If I is proper non-zero ideal of $\mathscr{A}$ then $\mathscr{A}^{\mathrm{op}} \otimes I$ is a proper non-zero ideal of $\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}$. Since $\mathrm{End}(A)$ is simple we deduce that $\mathscr{A}$ is simple.

For the converse note first that, if $\mathscr{A}$ is simple, then it is a simple module for $\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}$. We also have

$$\mathrm{End}_{\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}}(\mathscr{A}) \cong Z(\mathscr{A}) \cong \mathbb{F}.$$

We show next that $\mathscr{A}$ is a faithful module. If $xay = 0$ for all a in $\mathscr{A}$ then $x\mathscr{A}y = 0$ and thus $x\mathscr{A}y\mathscr{A} = 0$. Therefore the ideal $\mathscr{A}y\mathscr{A}$ is a proper ideal and so must be the zero ideal. Hence $y = 0$, and $\mathscr{A}$ is faithful.

It folllows now from Corollary 13.6.2 that $\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}$ coincides with the centralizer in $\mathrm{End}(A)$ of $\mathrm{End}_{\mathscr{A}^{\mathrm{op}} \otimes \mathscr{A}}(\mathscr{A})$. As the latter is the centre of $\mathrm{End}(A)$, the lemma is proved. □

14.1.3 Lemma. *If $\mathscr{A}$ is a central simple algebra over $\mathbb{F}$ and $\mathscr{B}$ is an algebra over $\mathbb{F}$ then the ideals of $\mathscr{A} \otimes \mathscr{B}$ are all of the form $\mathscr{A} \otimes I$, for some ideal I of $\mathscr{B}$.*

Proof. Let $a_1, \ldots, a_n$ be a basis for $\mathscr{A}$ and let T_k be the element of $\mathrm{End}(A)$ that maps a_ℓ to $\delta_{k,\ell}$. By the previous lemma there are elements x_i in $\mathscr{A}$ such that, for any u in $\mathscr{A}$ we have

$$T_k(u) = \sum_i x_i u a_i.$$

Each element of $\mathscr{A} \otimes \mathscr{B}$ can be written in the form $\sum_i a_i \otimes b_i$, for suitable b_i. Then we have

$$\left(\sum_j x_j \otimes 1\right)\left(\sum_i a_i \otimes b_i\right)\left(\sum_j a_j \otimes 1\right) = \sum_{i,j} x_j a_i a_j \otimes b_i = 1 \otimes b_k.$$

This shows that any ideal that contains $\sum_i a_i \otimes b_i$ must contain $1 \otimes b_j$ for all j. Hence, if J is an ideal in $\mathscr{A} \otimes \mathscr{B}$ and

$$I := \{b \in \mathscr{B} : 1 \otimes b \in J\},$$

then $J = \mathscr{A} \otimes I$. □

The next result is the Noether-Skolem theorem.

14.1.4 Theorem. *Let $\mathscr{A}$ be a central simple algebra and let $\mathscr{B}$ be a simple algebra over $\mathbb{F}$. If f and g are homomorphisms from $\mathscr{B}$ into $\mathscr{A}$, then there is an element a in $\mathscr{A}$ such that $a^{-1} f(b) a = g(b)$, for all b in $\mathscr{B}$.*

Proof. If h is a homomorphism from $\mathscr{B}$ into $\mathscr{A}$ then we define an action of $\mathscr{A} \otimes \mathscr{B}^{\mathrm{op}}$ on $\mathscr{A}$ by

$$a \otimes b : x \mapsto h(b) x a.$$

Thus $\mathscr{A}$ is a module for $\mathscr{A} \otimes \mathscr{B}^{\mathrm{op}}$, which we denote by $\mathscr{A}^h$.

As $\mathscr{B}$ is simple and $\mathscr{A}$ is central simple, $\mathscr{A} \otimes \mathscr{B}^{\mathrm{op}}$ is simple. Since $\mathscr{A}^f$ and $\mathscr{A}^g$ are modules for this simple algebra and have the same dimension, they are isomorphic. Let ψ be an isomorphism from $\mathscr{A}^f$ to $\mathscr{A}^g$. Then for all b in $\mathscr{B}$ and a, x in $\mathscr{A}$,

$$\psi(f(b) x a) = g(b) \psi(x) a.$$

If $b = 1$ this yields $\psi(xa) = \psi(x)a$, now setting x equal to 1 we see that $\psi(a) = \psi(1)a$ for all a in $\mathscr{A}$. On the other hand if we put a and x equal to 1 here then we find that $\psi(f(b)) = g(b)\psi(1)$. As $f(b) \in \mathscr{A}$, we also have

$$\psi(f(b)) = \psi(1 f(b)) = \psi(1) f(b).$$

Therefore $g(b)\psi(1) = \psi(1) f(b)$. Since ψ is an isomorphism, $\psi(1)$ must be invertible, and thus the proof is complete. □

14.2 Factors

We have

$$\mathrm{Mat}_{mn\times mn}(\mathbb{F}) = \mathrm{Mat}_{m\times m}(\mathbb{F}) \otimes \mathrm{Mat}_{n\times n}(\mathbb{F})$$

and so $I \otimes \mathrm{Mat}_{n\times n}(\mathbb{F})$ is a subalgebra of $\mathrm{Mat}_{mn\times mn}(\mathbb{F})$. Our next result generalizes this.

14.2.1 Theorem. *Let $\mathscr{A}$ be a central simple algebra and let $\mathscr{B}$ be a subalgebra of $\mathscr{A}$. If $\mathscr{B}$ is simple then $C_{\mathscr{A}}(\mathscr{B})$ is simple and*

$$\dim(\mathscr{A}) = \dim(\mathscr{B})\dim(C_{\mathscr{A}}(\mathscr{B}));$$

if $\mathscr{B}$ is central simple then

$$\mathscr{A} \cong \mathscr{B} \otimes C_{\mathscr{A}}(\mathscr{B}).$$

Proof. Our constraints on $\mathscr{A}$ and $\mathscr{B}$ imply that $\mathscr{B}^{\mathrm{op}} \otimes \mathscr{A}$ is simple, let P be a minimal right ideal in it. Then

$$D := \mathrm{End}_{\mathscr{B}^{\mathrm{op}}\otimes\mathscr{A}}(P)$$

is a division ring and $\mathscr{B}^{\mathrm{op}} \otimes \mathscr{A} \cong \mathrm{Mat}_{n\times n}(D)$. Hence

$$\dim(\mathscr{B})\dim(\mathscr{A}) = n^2 \dim(D).$$

We show next that $C_{\mathscr{A}}(\mathscr{B}) \cong \mathrm{End}_{\mathscr{B}^{\mathrm{op}}\otimes\mathscr{A}}(\mathscr{A})$. If $\psi \in \mathrm{End}_{\mathscr{B}^{\mathrm{op}}\otimes\mathscr{A}}(\mathscr{A})$ then for all b in $\mathscr{B}$ and x, a in $\mathscr{A}$, we have

$$\psi(bxa) = b\psi(x)a.$$

Setting b equal to 1 here yields $\psi(xa) = \psi(x)a$, whence $\psi(a) = \psi(1)a$ for all a in $\mathscr{A}$. Setting a equal to 1 and noting that $b \in \mathscr{A}$ we also find that

$$b\psi(1) = \psi(b\cdot 1) = \psi(1\cdot b) = \psi(1)b.$$

It follows that the map $\psi \mapsto \psi(1)$ is an isomorphism from $\mathrm{End}_{\mathscr{B}^{\mathrm{op}}\otimes\mathscr{A}}(\mathscr{A})$ to $C_{\mathscr{A}}(\mathscr{B})$.

As $\mathscr{A}$ is a finite dimensional module over $\mathscr{B}^{\mathrm{op}} \otimes \mathscr{A}$, it is isomorphic to P^k for some k and therefore

$$C_{\mathscr{A}}(\mathscr{B}) \cong \mathrm{End}_{\mathscr{B}^{\mathrm{op}}\otimes\mathscr{A}}(\mathscr{A}) \cong \mathrm{Mat}_{k\times k}(D).$$

Consequently $C_{\mathscr{A}}(\mathscr{B})$ is simple with dimension $k^2\dim(D)$ and, as

$$\dim(\mathscr{A}) = k\dim(P) = kn\dim(D)$$

it follows that

$$\dim(\mathscr{B})\dim(C_{\mathscr{A}}(\mathscr{B})) = \dim(\mathscr{A}).$$

Let $x_1,\dots,x_r$ be a linearly independent subset of $\mathscr{B}$ and suppose $y_1,\dots,y_r$ lie in $C_{\mathscr{A}}(\mathscr{B})$. We will show that if

$$\sum_{i=1}^{r} x_i y_i = 0,$$

then $y_i = 0$ for all i. Given this it follows easily that if $b_1,\dots,b_r$ is a basis for $\mathscr{B}$ and $c_1,\dots,c_s$ is a basis for $C_{\mathscr{A}}(\mathscr{B})$ then the rs products $b_i c_j$ are linearly independent. This implies that $\mathscr{B}C_{\mathscr{A}}(\mathscr{B})$ spans $\mathscr{A}$ and hence $\mathscr{B}\otimes C_{\mathscr{A}}(\mathscr{B}) \cong \mathscr{A}$.

To prove our claim we note that, since $\mathscr{B}$ is central simple, $\mathscr{B}\otimes\mathscr{B}^{\mathrm{op}} \cong \mathrm{End}_{\mathbb{F}}(\mathscr{B})$. So there are elements v_{jk}, w_{jk} in $\mathscr{B}$ such that

$$\sum_k v_{jk} x_i w_{jk} = \delta_{ij} I$$

and therefore

$$\sum_k v_{jk} x_i y_i w_{jk} = \sum_k v_{jk} x_i w_{jk} y_i = y_j,$$

which is all we require. □

It follows from the the first part of this theorem that, if $\mathscr{B}$ is simple then $C_{\mathscr{A}}(\mathscr{B})$ is simple and hence that $C_{\mathscr{A}}(C_{\mathscr{A}}(\mathscr{B}))$ has the same dimension as $\mathscr{B}$. As $C_{\mathscr{A}}(C_{\mathscr{A}}(\mathscr{B}))$ contains $\mathscr{B}$ this provides a second proof of the double centralizer theorem, at least when $\mathscr{B}$ is simple.

14.3 Finite Division Algebras

If F is a division algebra, let F^* denote $F\setminus 0$, viewed as a multiplicative group.

We have another important result, again due to Wedderburn.

14.3.1 Theorem. *A finite division ring is a field.*

Proof. Let D be a finite division ring. Its centre is a field, which we denote by F. If $x \in D \setminus F$ then x and F generate a commutative subalgebra of D, necessarily a field. We assume by way of contradiction that $D \neq F$. It follows that there is a maximal subfield E of D such that $D > E > F$.

If $x \in C_D(E)$ then E and x together generate a commutative subalgebra of D, which is necessarily a field. Hence $E = C_D(E)$. Since E is simple,

$$\dim(E)\dim(C_D(E)) = (\dim(E))^2 = \dim(D),$$

and we have shown that all maximal subfields of D have the same order. Hence they are all isomorphic and, by the Noether-Skolem theorem, it follows that they are all conjugate.

Suppose $m = |D^* : E^*|$. If x and y are elements of D^* in the same coset of E^* then $x^{-1}Ex = y^{-1}Ey$, whence E^* has at most m distinct conjugates in D^*. If $|F| = q$ then $|D^*| = q^\ell - 1$ and $|E^*| = q^k - 1$ for some ℓ and k, and $(q^\ell - 1)/(q^k - 1) = m$. Hence the number of elements of D in a conjugate of E is at most $m(q^k - 2) + 1$, but

$$m(q^k - 2) + 1 < m(q^k - 1) = q^\ell - 1.$$

As every element of D lies in a maximal subfield, this is impossible. □

The argument in the above proof implies that the dimension of a finite-dimensional division ring is a perfect square. Thus the quaternions have dimension four over $\mathbb{R}$, for example.

We add a few comments related to the quaternions. The norm of quaternion $h = a + bi + cj + dj$ is

$$(a^2 + b^2 + c^2 + d^2)^{1/2}.$$

We call h *pure* if $a = 0$, and will use the Noether-Skolem theorem to show that all pure quaternions of norm 1 are conjugate. As we may identify the pure quaternions of norm 1 with the unit sphere in $\mathbb{R}^3$, it follows that we have an action of $\mathbb{H}$ on this sphere. (In fact $\mathbb{H}$ acts as a group of orthogonal transformations, which is not hard to see.)

If $p = bi + cj + dk$ is a pure quaternion, then

$$p^2 = -(b^2 + c^2 + d^2).$$

Thus if p is a pure quaternion with norm 1 then $p^2 = -1$. On the other hand, if $a \in \mathbb{R}$ and p is a pure quaternion then

$$(a + p)^2 = a^2 + p^2 + 2ap;$$

hence $(a+p)^2$ is real if and only if $ap = 0$. Thus the solutions in $\mathbb{H}$ to the equation $t^2 + 1 = 0$ are the pure quaternions of norm 1.

If $p^2 = q^2 = -1$ then $\{1, p\}$ and $\{1, q\}$ are isomorphic subalgebras of $\mathbb{H}$, where the isomorphism maps p to q. So, by the Noether-Skolem theorem, there must be an element a of H such that $a^{-1}pa = q$. This shows that all pure quaternions are conjugate in $\mathbb{H}$.

14.4 Real Algebra

The following comes from; S. H. Kulkarni, A very simple and elementary proof of a theorem of Ingelstam, *American Math. Monthly*, **111** (1), 2004, 54–58.

The norm condition in the statement of the following theorem holds if we have

$$\|xy\| \leq \|x\|\|y\|$$

for all x and y.

14.4.1 Theorem. *Let A be a real algebra (with unit). Assume A is an inner product space such that $\|I\| = 1$ and $\|a^2\| \leq \|a\|^2$ for all a. Then A is isomorphic to $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$.*

Proof. First we show that if $x \in I^{\perp}$ and $\|x\| = 1$, then $x^2 = -1$.

If $t \in \mathbb{R}$, then

$$\|tI + x\|^2 = \|tI\|^2 + \|x\|^2 + 2\langle x, I\rangle = t^2 + 1.$$

By our norm condition,

$$\|(tI + x)^2\| \leq \|tI + x\|^2$$

and therefore

$$\begin{aligned}(t^2+1)^2 \geq \|(tI+x)^2\|^2 &= \langle t^2 I + 2tx + x^2, t^2 I + 2tx + x^2\rangle \\ &= t^4 + 2t^2\langle I, x^2\rangle + 4t^2 + 4t\langle x, x^2\rangle + \langle x^2, x^2\rangle\end{aligned}$$

This implies that, for all real t

$$2t^2(1 + \langle I, x^2\rangle) + 4t\langle x, x^2\rangle + \|x^2\|^2 - 1 \leq 0.$$

Hence the coefficient of t^2 in this quadratic must be non-negative and therefore

$$\langle I, x^2\rangle \leq -1.$$

By Cauchy-Schwarz,

$$\langle I, x^2\rangle^2 \le \|I\|^2\|x^2\|^2 = \|x^2\|^2 \le \|x\|^4 = 1,$$

whence $\langle I, x^2\rangle \ge -1$. Hence equality holds in the Cauchy-Schwarz inequality, and therefore $x^2 = uI$ for some real number u. Therefore $x^2 = -1$.

Next we show that if x and y are orthogonal elements of norm 1 in $I^\perp$, then $xy = -yx$. We have $x^2 = y^2 = -1$. If

$$z := \frac{1}{\sqrt{2}}(x+y)$$

then $\|z\| = 1$. Since $z \in I^\perp$ it follows that

$$-1 = z^2 = \frac{1}{2}(x^2 + y^2 + xy + yx) = -1 + \frac{1}{2}(xy + yx).$$

Therefore $xy + yx = 0$.

This brings us to the proof of the theorem. If $I^\perp = \{0\}$, then A is spanned by I and consequently $A \cong \mathbb{R}$. Assume $I^\perp \ne \{0\}$ and let x be a vector of norm 1 in $I^\perp$. Then $x^2 = -1$; if $\dim(A) = 2$ it follows that $A \cong \mathbb{C}$.

We assume that $\dim(A) \ge 3$ and that $\{I, x, y\}$ is an orthonormal set in A. Let $z = xy$. Then

$$z^2 = xyxy = -yxxy = y^2 = -1.$$

Similarly

$$yz = z,\ zy = -x,\ zx = y,\ xz = -y.$$

If $a, b, c, d \in \mathbb{R}$ and $u = aI + bx + cy + dz$, then

$$(aI + bx + cy + dz)(aI - bx - cy - dz) = (a^2 + b^2 + c^2 + d^2)I.$$

Hence $u = 0$ if and only if $a = b = c = d = 0$, and therefore $\{I, x, y, z\}$ are linearly independent and their span is isomorphic to $\mathbb{H}$.

Thus the theorem holds in $\dim(A) \le 4$, and otherwise there is an orthonormal set $\{I, x, y, z, u\}$. Then $u^2 = -1$ and

$$xu + ux = yu + uy = zu + uz = 0.$$

Consequently

$$uz = uxy - xuy = xyu = zu = -uz$$

and $uz = 0$. But

$$(uz)^2 = uzuz = -uzzu = u^2 = -1,$$

a contradiction. □

Chapter 15

Work

15.1 Classical Parameters

A distance-regular graph of diameter d has *classical parameters* if there are scalars q, α, β such that

$$b_i = ([d] - [i])(\beta - \alpha[i]),$$
$$c_i = [i](1 + \alpha[i-1]).$$

Here

$$[i] := \frac{q^i - 1}{q - 1}.$$

We have

$$a_i = [i](\beta - 1 + \alpha([d] - [i] - [i-1])), \quad i = 1, \ldots, d$$

and

$$k = [d]\beta, \quad c_2 = (1+q)(1+\alpha), \quad a_1 = (\beta - 1 + \alpha([d] - 1)).$$

15.1.1 Theorem. *If X is a distance-regular graph with classical parameters (d, q, α, β), with $d \geq 3$. Then*

$$\alpha = \frac{c_2}{q+1} - 1, \quad \beta = \frac{k}{[d]}.$$

and q is an integer, not 0 or -1. Further, one of the following holds:

(a) $a_i = a_1 c_i$ *for* $i = 2, \ldots, d$ *and* $q = -a_1 - 1$.

(b) $a_i \neq a_1 c_i$ *for some* i, *and*

$$q = \frac{a_2 c_3 - c_2 a_3}{a_1 c_3 - a_3}.$$

Proof. If $q = 0$, then $[2] = [1]$ and $b_2 = 0$. If $q = -1$, then $[i] = 0$ when i is even and $c_2 = 0$.

Next we observe that

$$a_i = [i](a_1 - \alpha([i] + [i-1] - 1))$$

and therefore

$$\begin{aligned} a_i - a_1 c_i &= [i](a_1 - \alpha([i] + [i-1] - 1)) - a_1[i](1 + \alpha[i-1]) \\ &= [i](a_1 - \alpha([i] + [i-1] - 1) - a_1 - a_1\alpha[i-1]) \\ &= -\alpha[i](([i] + [i-1] - 1) + a_1[i-1]) \\ &= -\alpha[i]((q+1)[i-1] + a_1[i-1]) \\ &= -\alpha[i][i-1](q + 1 + a_1). \end{aligned}$$

In particular

$$a_2 - a_1 c_2 = -\alpha(q+1)(q+1+a_1) \tag{15.1.1}$$

and thus

$$a_i - a_1 c_i = \begin{bmatrix} i \\ 2 \end{bmatrix}(a_2 - a_1 c_2).$$

Therefore $a_2 - a_1 c_2$ divides $a_i - a_1 c_i$ and consequently $[i][i-1]$ is an integer, from which it follows that q is an algebraic integer. Next

$$c_3 = [3](1 + \alpha[2]), \quad c_2 - q = [2](1+\alpha) - q = 1 + \alpha[2],$$

implying that $c_3 = [3](c_2 - q)$. Now

$$c_3(a_2 - a_1 c_2) = (c_2 - q)[3](a_2 - a_1 c_2) = (c_2 - q)\begin{bmatrix} 3 \\ 2 \end{bmatrix}(a_2 - a_1 c_2) = (c_2 - q)(a_3 - a_1 c_3)$$

and therefore

$$q(a_1 c_3 - a_3) = c_3(a_2 - a_1 c_2) - c_2(a_3 - a_1 c_3) = c_3 a_2 - c_2 a_3.$$

So if $a_2 - a_1 c_2 \neq 0$, then

$$q = \frac{c_3 a_2 - c_2 a_3}{a_1 c_3 - a_3};$$

since this is rational q must be an integer.

Finally assume that $a_2 - a_1 c_2 = 0$. Then $a_i = a_1 c_i$ for $i = 2, \ldots, d$ and, from eq:a2a1c2, $q + 1 + \alpha = 0$. □

Chapter 16

Adjacency Algebras

16.1 Extending the Adjacency Algebra

Suppose V is an inner product space and $\mathbb{A}$ is a subalgebra of $\mathrm{End}(V)$ that is closed under taking adjoints (i.e., closed under transposes in the real case). If U is a submodule of V, then $U^\perp$ is also a submodule. We see immediately that V is an orthogonal direct sum of irreducible modules (and $\mathbb{A}$ is semisimple). In this context we could refer to $\mathbb{A}$ as a C^*-algebra, or as a finite-dimensional Von Neumann algebra. (Our preference is for the latter.)

In this section we study a special class of self-adjoint algebras. Let X be a graph on v vertices with adjacency matrix A, let h be a nonzero vector in $\mathbb{R}^v$ and let H denote the matrix hh^T. We study the algebra $\mathbb{A} = \langle A, H \rangle$ generated by A and H.

Suppose U is an irreducible module for this algebra. If u is a vector in U such that $Hu \neq 0$, then

$$Hu = (h^T u)h \in U$$

and therefore $h \in U$. Accordingly U contains the $\mathbb{A}$-module generated by h; since U is irreducible it follows that U is generated by h. We call the module generated by h the *standard module* for $\mathbb{A}$.

Now suppose T is an $\mathbb{A}$-module such that $T \cap U = 0$. If $v \in T$ and $Hv \neq 0$, then $h \in T$, a contradiction. Therefore $Hv = 0$ and, since T is a module, $HA^r v = 0$ for all r. Since A is symmetric, it follows that $\langle A^r h, v \rangle = 0$ for all r and therefore $T \leq U^\perp$.

We can summarize our discussion with the following pile of words.

16.1.1 Theorem. *Let X be a graph on v vertices with adjacency matrix A, let h be a nonzero vector in $\mathbb{R}^v$ and let H denote the matrix hh^T. Finally, let $\mathbb{A}$ be the algebra generated by A and H. Then $\mathbb{R}^v$ is the direct sum of the standard module U and a set of 1-dimensional subspaces, each spanned by an eigenvector for A orthogonal to h. If $m = \dim(U)$, then $\dim(\mathbb{A}) = m^2 - m + v$ and $\mathbb{A}$ is the direct sum of $\mathrm{End}(U)$ and $v - m$ copies of $\mathbb{R}$.* □

Here $\mathrm{End}(U)$ is isomorphic to the full algebra of $m \times m$ real matrices. This follows from the general theory, but we can offer a direct argument.

16.1.2 Lemma. *If the standard module U for $\mathbb{A}$ has dimension m and is generated by h, then the m^2 matrices*

$$A^i h h^T A^j, \quad (0 \le i, j \le m-1)$$

are linearly independent.

Proof. First we show that the vectors $A^i h$ for $i = 0, \ldots, m-1$ are linearly independent. If the are not, then they span an A-invariant subspace of U with dimension less than m. Since

$$HA^i h = (h^T A^i h)h$$

this subspace is H-invariant and so it is an $\mathbb{A}$=module. As U is irreducible, we conclude that our set of vectors is linearly independent.

To complete the proof, we need to show that if the vectors $u_1, \ldots, u_m$ are linearly independent elements of U, then

$$u_i u_j^T, \quad (0 \le i, j < m)$$

are linearly independent (elements of $\mathrm{End}(U)$). We leave this as an exercise. □

16.2 Some Applications

Our main application of the theory in the previous section is to the case where the vector h is the characteristic vector of a subset of $V(X)$, perhaps $V(X)$ itself.

16.2.1 Lemma. *Suppose h is the characteristic vector of the subset S of $V(X)$. The permutation matrices that commute with A and hh^T are the automorphisms of X that fix S as a set.* □

16.2.2 Corollary. *Let X be a graph on v vertices and let S be a subset of $V(X)$ with characteristic vector h. If the standard module for $\langle A, hh^T\rangle$ has dimension v, then $\langle A, hh^T\rangle$ is the algebra of all $v \times v$ matrices, and the only automorphism of X that fixes S is the identity.*

The standard module has dimension v if and only if the vectors $A^i h$ for $i \geq 0$ span $\mathbb{R}^v$.

16.2.3 Corollary. *If the vectors $A^i\mathbf{1}$ span $\mathbb{R}^v$, then* $\mathrm{Aut}(X)$ *is trivial.* □

16.2.4 Corollary. *Let e_r be the r-th standard basis vector for $\mathbb{R}^v$. If the vectors $A^i e_r$ span $\mathbb{R}^v$, then the stabilizer in* $\mathrm{Aut}(X)$ *of the r-th vertex of X is trivial.*

Some writes define the *main eigenvalues* of A to be those eigenvalues θ such that there is an eigenvector z for θ which is not orthogonal to $\mathbf{1}$. The number of main eigenvalues equal the dimension of the standard module generated by $\mathbf{1}$. (Here $H = J$.) We can extend this to the general case.

16.2.5 Lemma. *Let A be a symmetric matrix with spectral decomposition $A = \sum \theta E_\theta$. Then the dimension of the standard module relative to h is equal to the number of eigenvalues θ such that $h^T E_\theta h \neq 0$.*

Proof. The non-zero vectors of the form $E_\theta h$ form a basis for U. Since $E_\theta^2 = E_\theta = E_\theta^T$, we have

$$h^T E_\theta h = h^T E_\theta^2 h = h^T E_\theta^T E_\theta h = \|E_\theta h\|^2.$$

Hence $h^T E_\theta h = 0$ if and only if $E_\theta h = 0$. □

16.3 Cospectral Awful Graphs

16.3.1 Theorem. *Suppose X is an awful graph with adjacency matrix A and Y is a graph with adjacency matrix B. If $\mathrm{tr}(A^r J) = \mathrm{tr}(B^r J)$ for all non-negative integers r, then X and Y are cospectral with cospectral complements.*

Proof. The n^2 matrices $A^r J A^s$, where $0 \leq r, s < n$ form a basis for $\mathscr{M} = \mathrm{Mat}_{n\times n}(\mathbb{R})$. Therefore there is a unique linear mapping Ψ from $\mathscr{M}$ to itself such that

$$\Psi(A^r J A^s) = B^r J B^s.$$

Now

$$\begin{aligned}\operatorname{tr}(A^jJA^iA^rJA^s) &= \operatorname{tr}(A^{j+s}JA^{i+r}J)\\ &= \mathbf{1}^TA^{j+s}\mathbf{1}\mathbf{1}A^{i+r}\mathbf{1}\\ &= \operatorname{tr}(A^{j+s}J)\operatorname{tr}(A^{i+r}J).\end{aligned}$$

The last line here is equal to $\operatorname{tr}(B^{j+s}J)\operatorname{tr}(B^{i+r}J)$ and so we deduce that, for all r and s,

$$\langle A^iJA^j, A^rJA^s\rangle = \langle B^iJB^j, B^rJB^s\rangle.$$

The matrices A^rJA^s are linearly independent and therefore their Gram matrix (with respect to the trace inner product) must be invertible. Consequently the Gram matrix of the matrices B^rJB^s is invertible, and therefore these matrices are linearly independent. This implies that Ψ is invertible, since it maps a basis of $\mathscr{M}$ to a basis.

We now show that Ψ is a homomorphism. First we observe that

$$A^iJA^jA^kJA^\ell = (\mathbf{1}^TA^{j+k}\mathbf{1})A^iJA^\ell. \tag{16.3.1}$$

It follows that to prove that Ψ is a homomorphism, it is enough to show that

$$\Psi(A^iJA^jA^kJA^\ell) = \Psi(A^iJA^j)\Psi(A^kJA^\ell).$$

Using (16.3.1) we find that

$$\begin{aligned}\Psi(A^iJA^jA^kJA^\ell) &= \mathbf{1}^TA^{j+k}\mathbf{1}\Psi(A^iJA^\ell)\\ &= \mathbf{1}^TB^{j+k}\mathbf{1}B^iJB^\ell\\ &= B^iJB^jB^kJB^\ell.\end{aligned}$$

It follows that Ψ is an automorphism of $\mathrm{Mat}_{v\times v}(\mathbb{R})$ and, by the Noether-Skolem theorem (Theorem 14.1.4), this implies there is an invertible matrix L such that

$$\Psi(M) = L^{-1}ML,$$

for all matrices M. Since

$$\Phi(A)B^iJB^j = \Phi(A)\Phi(A^iJA^j) = \Phi(A^{i+1}JA^j) = B^{i+1}JB^j = BB^iJB^j$$

and the matrices B^iJB^j form a basis, it follows that $\Phi(A) = B$. Hence we deduce that A and B are cospectral and, since $\Psi(J) = J$, we also see that $J-I-A$ and $J-I-B$ are cospectral. □

We offer two further proofs of this result. For the first, let M be the $n \times n$ matrix with the vectors

$$\mathbf{1}, A\mathbf{1}, \ldots, A^{n-1}\mathbf{1}$$

as its columns and let N be the corresponding matrix based on B. Then

$$M^T M = N^T N,$$

from which it follows that there is an orthogonal matrix Q such that $QM = N$. Thus we have

$$QA^r\mathbf{1} = B^r\mathbf{1},$$

from which it follows that $Q\mathbf{1} = \mathbf{1}$ and so

$$QA^rQ^T\mathbf{1} = B^r\mathbf{1}$$

for all r. Now

$$\begin{aligned} B^rB^s\mathbf{1} &= B^{r+s}\mathbf{1} \\ &= QA^{r+s}Q^T\mathbf{1} \\ &= QA^rQ^TQA^s\mathbf{1} \\ &= QA^rQ^TB^s\mathbf{1} \end{aligned}$$

and since the vectors $B^s\mathbf{1}$ for $s = 0, \ldots, n-1$ form a basis it follows that $B^r = QA^rQ^T$ and, in particular, $B = QAQ^T$.

Our third and final proof uses spectral decomposition and walk generating functions. The generating function all walks in X is

$$\sum_{r\geq 0} (\mathbf{1}^T A^r \mathbf{1})\, t^r = \sum_{r\geq 0} \operatorname{tr}(A^r J)\, t^r.$$

Assuming that A has the spectral decomposition

$$A = \sum_{\theta} \theta E_\theta,$$

we deduce that we have the following expression for our generating function as a rational function:

$$\sum_{r\geq 0} \operatorname{tr}(A^r J) = \sum_{\theta} \frac{\mathbf{1}^T E_\theta \mathbf{1}}{1 - t\theta}.$$

The number of poles in this rational function is equal to the number of eigenvalues θ such that $\mathbf{1}^T E_\theta \mathbf{1} \neq 0$. Since the matrices E_θ are symmetric and idempotent, $\mathbf{1}^T E_\theta \mathbf{1} \neq 0$ if and only if $E_\theta \mathbf{1} \neq 0$ and therefore the number of poles of the generating function is equal o the dimension of the cyclic A-module generated by $\mathbf{1}$. If X is $\mathbf{1}$-full, it follows that the numerator of our rational function is:

$$\prod_\theta (1 - t\theta) = t^n \det(t^{-1} I - A).$$

We conclude that if X is $\mathbf{1}$-full, then its spectrum is determined by the generating function for all walks in X.

Hence if the assumptions of theorem hold, then X and Y are cospectral and, since their generating function for all walks are equal, it follows from [**?**] that their complements are also cospectral.

16.4 Modules and Walks

Let A be the adjacency matrix of a graph X. Then the entries of the powers of A count walks in X. In particular if u and v respectively are the characteristic vectors of subsets S and T of $V(X)$, then

$$u^T A^k v$$

is the number of walks of length k in X that start at a vertex in S and end at a vertex in T. We relate properties of the sequence $(u^T A^k v)_{k\geq 0}$ to properties of the cyclic A-module $\langle v \rangle_A$ generated by v.

If v is a vector in $\mathbb{F}^n$ and $A \in \mathrm{Mat}_{d\times d}(\mathbb{F})$, then the *minimal polynomial of A relative to v* is the monic polynomial ψ of least degree such that $\psi(A)v = 0$.

16.4.1 Lemma. *If v and w are vectors in $\mathbb{F}^n$ and $A \in \mathrm{Mat}_{d\times d}(\mathbb{F})$, then the cyclic A-modules generated by v and w are isomorphic if and only if the minimal polynomials of A relative to v and w are equal.*

Proof. Exercise. □

If $(a_i)_{i\geq 0}$ is a sequence, the *Hankel matrix* H_n of order n is the $n \times n$ matrix such that,

$$(H_n)_{i,j} = a_{i+j-2}.$$

The Hankel matrix H is the infinite matrix whose leading $n \times n$ submatrix is H_n (for all n).

16.4.2 Lemma. *If A is self-adjoint, the sequence $\langle u, A^i u\rangle$ determines the isomorphism class of the module $\langle u\rangle_A$.*

Proof. The first $k+1$ rows of H are linearly independent if and only if there are scalars $a_0, \ldots, a_k$ such that for all non-negative r,

$$\sum_{i=0}^{k} a_i \langle u, A^{i+r} u\rangle = 0$$

Since A is self-adjoint the left side here is equal to

$$\langle u, \sum_{i=0}^{k} a_i A^{i+r} u\rangle = \langle A^r u, p(A)u\rangle$$

where $p(t)$ is a polynomial with degree k. If $U := \langle u\rangle_A$, this shows that $p(A)u \in U^\perp$. Since $p(A)u \in U$, we have

$$p(A)u \in U \cap U^\perp = 0$$

and therefore $p(A)u = 0$. It follows that we can read off the minimal polynomial of A relative to u from the Hankel matrix, and therefore this minimal polynomial is determined by the sequence. □

16.5 An Inner Product on Polynomials

Suppose V is an inner product space and A is a self-adjoint element of $\mathrm{End}(V)$. We develop a connection between cyclic A-submodules of V and sequences of orthogonal polynomials.

Suppose V is an inner product space, $v \in V$ and A is a self-adjoint element of $\mathrm{End}(V)$. We set ourselves the innocent goal of finding an orthogonal basis for the cyclic A-module U generated by v. This is straightforward, if $\dim(U) = d$ then

$$v,\ Ad,\ \ldots, A^{d-1} v$$

is a basis for U and, using Gram-Schmidt, we can convert this basis to an orthogonal basis. The fun starts when we notice that each element of U can be written uniquely in the form $p(A)v$, where p is a real polynomial with degree less than d. So the elements of the orthogonal basis we compute can be written as

$$p_0(A)v, \ldots, p_{d-1}(A)v$$

where $p_0, \ldots, p_{d-1}$ are real polynomials. If we apply Gram-Schmidt in the obvious way and do not normalize the elements of the orthogonal basis we produce, then these polynomials will be monic and p_i will have degree i. We want to determine these polynomials without being forced to use Gram-Schmidt explicitly.

We introduce some machinery. If $p, q \in \mathbb{R}[t]$, define

$$[p, q] := \langle p(A)v, q(A)v \rangle.$$

This is a bilinear form on $\mathbb{R}[t]$ and since A is self-adjoint, this form is symmetric. Let ψ denote the minimal polynomial of A relative to v. Since $\dim(U) = d$, we see that $\deg(\psi) = d$. If p has degree less than d, then $p(A)v \neq 0$, whence

$$[p, p] = \langle p(A)v, p(A)v \rangle > 0.$$

So our form is positive definite on the space of polynomials of degree less than d, and thus is an inner product on this space. In fact we prefer to say that we have an inner product on the quotient ring $\mathbb{R}[t]/(\psi)$ (which is a real algebra of dimension d).

16.6 Spectral Decomposition

To get further we need another description of our inner product. Since A is self-adjoint, it has a spectral decomposition

$$A = \sum_\theta \theta E_\theta,$$

where θ runs over the distinct eigenvalues of A, and matrices E_θ are pairwise orthogonal projections. If p is a polynomial then

$$p(A) = \sum_\theta p(\theta) E_\theta$$

and therefore

$$p(A)v = \sum_\theta p(\theta) E_\theta v.$$

16.6.1 Lemma. *If A is a self-adjoint operator with spectral decomposition $A = \sum_\theta \theta E_\theta$, then the dimension of the cyclic A-module generated by v is equal to the number of eigenvalues θ of A such that $v^T E_\theta v \neq 0$.*

Proof. Assume $U = \langle v \rangle_A$. If $\theta \neq \tau$, then $E_\theta E_\tau = 0$ and therefore the vectors $E_\theta v$ and $E_\tau v$ are orthogonal. Hence the nonzero vectors $E_\theta v$ form a linearly independent set, and so they are a basis for U. To complete the proof we observe that

$$v^T E_\theta v = v^T E_\theta^2 v = v^T E_\theta^T E_\theta v = \langle E_\theta v, E_\theta v \rangle,$$

which shows that $E_\theta v \neq 0$ if and only if $v^T E_\theta v \neq 0$. □

It also follows that if ψ is the minimal polynomial of A relative to v, then

$$\psi(t) = \prod_{\theta : E_\theta v \neq 0} (t - \theta).$$

For the next result, note that the idempotents E_θ are positive semidefinite, and therefore $x^T E_\theta x \geq 0$ for any vector x.

16.6.2 Lemma. *If A is a self-adjoint operator with spectral decomposition $A = \sum_\theta \theta E_\theta$, then*

$$[p, q] = \sum_\theta p(\theta) q(\theta)\, v^T E_\theta v.$$

Proof. We have

$$p(A) v = \sum_\theta p(\theta) E_\theta v, \quad q(A) v = \sum_\theta q(\theta) E_\theta v,$$

and the result is immediate. □

It follows from this that if p is non-negative on on the set of zeros of ψ, then $[1, p] \geq 0$.

16.7 Orthogonal Polynomials

If A is self-adjoint and v generates a cyclic A-module of dimension d, then Gram-Schmidt applied to the sequence

$$v,\ Av, \ldots, A^{d-1} v$$

gives rise to an orthogonal basis of the form

$$p_0(A) v,\ p_1(A) v, \ldots, p_{d-1}(A) v$$

where p_i is a monic polynomial of degree i. Since these vectors are orthogonal, if $i \neq j$, then

$$[p_i, p_j] = 0.$$

Thus our polynomials are orthogonal with respect to this inner product. We use this derive a recurrence for these polynomials.

Since any polynomial of degree at most i is a linear combination of $p_o, \ldots, p_i$, we see that $\deg(q) < j$ then

$$[p_j, q] = 0.$$

Note also that, since p_i is monic,

$$t^i = p_i(t) + r(t),$$

where $\deg(r) < i$ and therefore

$$[p_i, t^i] = [p_i, p_i].$$

We extend our sequence of polynomials defining p_d to be ψ.

16.7.1 Theorem. *Suppose A is a self-adjoint operator on the inner product space V and $v \in V$. Then for $i = 0, \ldots, d-1$, the orthogonal polynomials p_i satisfy the recurrence*

$$tp_i(t) = p_{i+1}(t) + a_i p_i(t) + b_i p_{i-1}(t)$$

where

$$a_i = \frac{[tp_i, p_i]}{[p_i, p_i]}, \qquad b_i = \frac{[p_i, p_i]}{[p_{i-1}, p_{i-1}]}.$$

Proof. We first show that the p_i can be defined by a three-term recurrence as shown, and then we determine the coefficients a_i and b_i.

If $j < i-1$, then $p_i(t)$ and $tp_j(t)$ are orthogonal and therefore

$$0 = [p_i, tp_j] = [tp_i, p_j].$$

If $j < i+1$, then p_j and tp_i are orthogonal and so $[tp_i, p_j] = 0$ in this case too. It follows that tp_i is a linear combination of p_{i+1}, p_i and p_{i-1}.

Suppose

$$tp_i(t) = c_i p_{i+1}(t) + a_i p_i(t) + b_i p_{i-1}(t).$$

Since p_i and p_{i+1} are monic we have $c_i = 1$. To determine a_i we take the inner product of each side with p_i, and obtain the given expression for a_i. If we next take the inner product of each side with p_{i-1} we find that

$$b_i = \frac{[tp_i, p_{i-1}]}{[p_{i-1}, p_{i-1}]}.$$

However $[tp_i, p_{i-1}] = [p_i, tp_{i-1}]$ and since p_i and p_{i-1} are monic,

$$[p_i, tp_{i-1}] = [p_i, p_i].$$

We now have

$$Ap_i(A)v = p_{i+1}(A)v + a_i p_i(A)v + b_i p_{i-1}(A)v,$$

whence the matrix representing the action of A relative to the basis

$$p_0(A)v, \ldots, p_{d-1}(A)v$$

is tridiagonal:

$$\begin{pmatrix} a_0 & 1 & & & \\ b_1 & a_1 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{d-2} & a_{d-2} & 1 \\ & & & b_{d-1} & a_{d-1} \end{pmatrix}$$

If we denote the characteristic polynomial of the leading $i \times i$ submatrix of this by q_i, then we find that

$$tq_i = q_{i+1} + a_i q_i + b_i q_{i-1},$$

whence $p_i = q_i$ and $p_d = \psi$. □

16.7.2 Lemma. *The zeros of p_i are real and distinct.*

Proof. Suppose $p_i(t) = f(t)g(t)$, where g is nonnegative on the zeros of ψ. We show that this implies that g is a constant polynomial, and then deduce the claims of the lemma from this.

Suppose $p_i = fg$ where g is nonnegative and $\deg(f) < i$. Then

$$0 = [p_i, f] = \sum_\theta p_i(\theta) f(\theta)\, v^T E_\theta v = \sum_\theta f(\theta)^2 g(\theta)\, v^T E_\theta v.$$

Since both f^2 and g are nonnegative on the zeros of ψ, this implies that $f(\theta)g(\theta) = 0$ for all θ and so $\deg(p_i) \geq \deg(\psi)$. Hence we conclude that we cannot have a factorization of p_i as described.

If p_i had a repeated root θ, then it would have non-negative factor $(t-\theta)^2$; if it had a complex root $a+bi$ then, since p_i is a real polynomial it would also have $a-bi$ as a root, and hence would have the nonnegative quadratic

$$(t-a)^2 + b^2$$

as a factor. □

16.8 Distance-Regular Graphs

We present an application of the theory from the previous sections. Suppose X is a distance-regular graph with distance matrices $A_0, \ldots, A_d$, where $A_0 = I$. Let $A = A_1$ and consider the cyclic A-module generated by e_1 (the first vector in the standard basis for $\mathbb{R}^n$).

Our first observation is that $p_r(A)e_1$ is a non-zero scalar multiple of $A_r e_1$. To prove this, we note that A_r is a polynomial in A_1 with degree r and thus

$$A_r e_1 = q_r(A)e_1$$

where $\deg(q_r(t)) = r$. There is a nonzero constant γ such that

$$[q_r, q_s] = \langle q_r(A_1), q_s(A_1)\rangle = \gamma\langle A_r e_1, A_s e_1\rangle = 0$$

Hence the polynomials q_i are orthogonal with respect to our inner product on polynomials, and therefore each q_i is a non-zero multiple of p_i.

It follows that the cyclic module generated by e_1 is the space of real functions on $V(X)$ that are constant on the cells of the distance partition with respect to the vertex 1. Hence the matrix representing A_1 relative to the orthogonal vectors $p_r(A)e_1$ is the transpose of the intersection matrix. Equivalently this transpose is equal to the adjacency matrix of the quotient of X relative to the distance partition.

16.9 Locally Distance-Regular Graphs

The theorem in this section is due to Godsil and Shawe-Taylor [] (with a different proof).

If u is a vertex in X, let δ_u denote the distance partition of X relative to u. Then X is distance regular if

(a) For each vertex u, the partition δ_u is equitable, and

(b) The quotient X/δ_u is the same for each vertex.

Let us say a graph is *locally distance regular* if (a) holds. We aim to characterize locally distance-regular graphs.

If $u \in V(X)$, then e_u will denote the characteristic vector of u viewed as a subset of $V(X)$.

16.9.1 Lemma. *Let u be a vertex in X with valency k_u. If δ_u is equitable and $v \sim u$, then*

$$\langle e_u, A^r e_v\rangle = \frac{1}{k_u}\langle e_u, A^{r+1} e_u\rangle.$$

Proof. Let $A = A(X)$ and let U be the cyclic A-module generated by e_u. Let z_i denote the characteristic vector of the i-th cell of δ_v. Then the vectors $z_0, \dots, z_r$ form an orthogonal basis for U. If $w \in \mathbb{R}^n$, then the projection of w on U is

$$\hat{w} = \sum_i \frac{\langle w, z_i\rangle}{\langle z_i, z_i\rangle} z_i.$$

In particular, if $v \in V(X)$ and $d := \mathrm{dist}(u, v)$, then

$$\hat{v} = \frac{1}{\langle z_d, z_d\rangle} z_d.$$

Note that $z_0 = e_u$ and $z_1 = Ae_u$, and so if $v \sim u$ then

$$\hat{v} = \frac{1}{k_u} Ae_u.$$

Therefore

$$\langle e_v - k_u^{-1} Ae_u, A^r u\rangle = 0$$

for all r and thus

$$\langle e_v, A^r u\rangle = \frac{1}{k_u}\langle Ae_u, A^r e_u\rangle = \frac{1}{k_u}\langle e_u, A^{r+1} e_u\rangle. \qquad \square$$

16.9.2 Corollary. *If $u \sim v$ and δ_u and δ_v are equitable and $r \geq 1$, then*

$$\frac{\langle e_u, A^r e_u\rangle}{k_u} = \frac{\langle e_v, A^r e_v\rangle}{k_v}. \qquad \square$$

We see that if u and v have the same valency, then

$$\langle e_u, A^r e_u\rangle = \langle e_v, A^r e_v\rangle$$

for all r and so the cyclic modules generated by e_u and e_v are isomorphic. This implies that the quotients of X relative to the corresponding distance partitions are isomorphic.

16.9.3 Theorem. *If X is locally distance regular, then either X is distance-regular or X is bipartite and semiregular.*

Proof. Suppose $v, w \sim u$ and assume v has exactly a neighbors in common with u. Then v has exactly $k_v - 1 - a$ neighbors at distance two from u. Since δ_u is distance regular each neighbor of u has exactly $k_v - 1 - a$ neighbors at distance two from u. Since the number of neighbors of w at distance two from u must be $k_w - 1 - a$, we conclude that $k_v = k_w$.

A standard argument yields that X is either regular or bipartite and semiregular, and in the first case we see that X is distance-regular. (In the second case you may show that the halved graphs of X are distance-regular.) □

16.10 Coherent Algebras

A *coherent algebra* is a matrix algebra that contains I and J and is closed under Schur multiplication and transpose. The span of I and J provides a trivial example. A commutative coherent algebra is the same thing as the Bose-Mesner algebra of an association scheme (although this is not trivial). The coherent algebra generated by a set of matrices is the smallest coherent algebra that contains the given set. The coherent algebra of a graph X on v vertices is the coherent algebra generated by $A(X)$.

16.10.1 Theorem. *A coherent algebra has a unique basis of* 01*-matrices. If this basis is formed by the matrices* $\mathscr{A} = \{A_0, \ldots, A_d\}$ *then:*

(a) $\sum_i A_i = J$.

(b) Some subset of $\mathscr{A}$ *sums to* I.

(c) $A_i \circ A_j = \delta_{i,j} A_i$.

(d) There are scalars $p_{i,j}(r)$ *such that* $A_i A_j = \sum_r p_{i,j}(r) A_r$.

(e) $A_i^T \in \mathscr{A}$ *for each* i.

(f) All non-zero rows and columns of A_i *have the same sum.*

We saw in Section 1.3 that the centralizer of a set of permutation matrices is a coherent algebra. Association schemes provide an overlapping class of examples.

The basis of 01-matrices of a coherent algebra is known as a *coherent configuration*. Their basic theory was laid out by D. Higman []. Coherent configurations generalize association schemes in two ways. First, the identity matrix

might not be an element of the 01-basis and, second, the coherent algebra need not be commutative. A coherent algebra is *homogeneous* if I belongs to its basis. The commutant of a permutation group is homogeneous if and only if the the group is transitive.

16.10.2 Lemma. *A commutative coherent algebra is homogeneous.*

Proof. If $\mathscr{C}$ is commutative, then each matrix in it commutes with J. Hence any diagonal matrix in $\mathscr{C}$ must be a multiple of I. □

16.10.3 Lemma. *If $\mathscr{C}$ is a homogenous coherent algebra, then any graph whose adjacency matrix lies in $\mathscr{C}$ is walk regular.*

Proof. If $M \in \mathscr{C}$, then the diagonal of M^k is constant. □

If a coherent algebra is not homogeneous, then its diagonal elements determine a partition of its vertex set, and this partition is equitable.

If D is a diagonal matrix in a coherent configuration with coherent algebra $\mathscr{C}$, then the subspace $D\mathscr{C}D$ of $\mathscr{C}$ is an algebra with identity D. It is Schur closed, and hence it it gives rise to a homogeneous coherent configuration.

Chapter 17

Line Digraphs

17.1 Line Digraphs

We now decide to view an edge ij in X as a pair of directed arcs (i, j) and (j, i). (So all our graphs are directed.) The *line digraph* $\mathrm{LD}(X)$ of a directed graph X has the arcs of X as its vertices, and

$$(ij, k\ell)$$

is an arc in the line digraph if ij and $k\ell$ are arcs in X and $j = k$. Generally the line digraph is a directed graph. The out-valency in $LD(X)$ of the arc ij is equal to the out-valency of j in X. Our eventual concern will be with a weighted adjacency matrix for $LD(X)$.

We begin with two incidence matrices D_i and D_o, with rows indexed by the vertices of X and columns by its arcs. If u is a vertex and e an arc of X, then $D_{u,e} = 1$ if u is the initial vertex of e, while $(D_o)_{u,e} = 1$ if e ends on u. Both D_i^T and D_0^T are the characteristic matrices of partitions of the arcs set of X.

17.1.1 Lemma. *If D_i and D_o are the vertex-arc incidence matrices of the graph X, then $D_i D_o^T$ is the adjacency matrix of X and $D_o^T D_i$ is the adjacency matrix of its line digraph.* □

From this we see that $A(X)$ and $A(\mathrm{LD}(X))$ have the same non-zero eigenvalues, with the same multiplicities.

Let P be the permutation matrix corresponding to the permutation of the arc set of X that maps each arc to its reversal. Then $P^2 = I$ and

$$D_i = D_o P, \quad D_o = D_i P.$$

It follows that

$$A(\mathrm{LD}(X)) = D_i^T D_o = D_i^T D_i P = P D_o^T D_o$$

whence

$$PA(\mathrm{LD}(X))P = A(\mathrm{LD}(X))^T.$$

Note also that $A(\mathrm{LD}\, X) - P$ is a 01-matrix, since for each arc uv there is an arc in $\mathrm{LD}(X)$ from uv to vu.

We now turn to weighted matrices. Let $\widehat{D}_o$ and $\widehat{D}_i$ denote the matrices we get from D_o and D_i by scaling each row so that it is a unit vector. We have

$$\widehat{D}_o P = \widehat{D}_i$$

and

$$\widehat{D}_o \widehat{D}_o^T = I.$$

From the latter it follows that

$$(\widehat{D}_o^T \widehat{D}_o)^2 = \widehat{D}_o^T \widehat{D}_o \widehat{D}_o^T \widehat{D}_o = \widehat{D}_o^T \widehat{D}_o;$$

thus $\widehat{D}_o^T \widehat{D}_o$ is symmetric and idempotent, and represents orthogonal projection onto the column space of $\widehat{D}_o^T$. Also

$$\widehat{D}_o^T \widehat{D}_i \mathbf{1} = \widehat{D}_o^T \widehat{D}_o P \mathbf{1} = \widehat{D}_o^T = \mathbf{1},$$

from which we see that $\widehat{D}_o^T \widehat{D}_i$ is the transition matrix of the Markov chain formed by the obvious random walk on $\mathrm{LD}(X)$.

The matrix $\widehat{D}_i^T \widehat{D}_o$ is a weighted adjacency matrix for $\mathrm{LD}(X)$. Let Δ denote the diagonal matrix whose diagonal entries are the valencies of the vertices of X. Then the matrix

$$\Delta^{-1/2} A \Delta^{-1/2}$$

is the *normalized* adjacency matrix of X, which we denote by $\widehat{A}$. We observe that

$$\widehat{D}_i \widehat{D}_o^T = \widehat{D}_o \widehat{D}_i^T = \widehat{A}.$$

So $\widehat{D}_i^T \widehat{D}_o$ and $\widehat{A}$ have the same non-zero eigenvalues with the same multiplicities.

Why we would we consider using the normalized adjacency matrix of X? Assume X has no isolated vertices, which means Δ is invertible. We have

$$\det(tI - \widehat{A}) = \det(\Delta)^{-1} \det(t\Delta - A) = \det(tI - \Delta^{-1} A).$$

Here $\Delta^{-1} A$ is a non-negative matrix with each row sum equal to 1—it is the transition matrix for the obvious random walk on X. The eigenvalues of $\widehat{A}$ are the eigenvalues of this transition matrix, and hence govern the behavior of this random walk.

17.2 Quantum Walks

A random walk on a (possibly directed) graph is represented by a non-negative square matrix whose rows sum to 1. One way of constructing such a matrix is to choose a unitary matrix U, and then take

$$U \circ \overline{U}$$

as our transition matrix. Since quantum physicists prefer unitary operations, they like this approach. A quantum random walk of length n based on U is determined by

$$U^n \circ \overline{U^n}.$$

The problem is to find natural constructions of unitary matrices with given underlying directed graph. For line digraphs there is an easy way to do this.

Suppose M is an orthogonal projection, that is, $M = M^T = M^2$. Then

$$(2M - I)^2 = 4M^2 - 4M + I = 4M - 4M + I = I$$

and, since $2M-I$ is symmetric, $2M-I$ is orthogonal. So referring to the previous section we see that

$$2\widehat{D}_o^T \widehat{D}_o - I$$

is orthogonal. Since the permutation matrix P is orthogonal, so is the product

$$(2\widehat{D}_o^T \widehat{D}_o - I)P = 2\widehat{D}_o^T \widehat{D}_i - P.$$

Thus we have a quantum walk associated with each graph.

Emms et al [] determine the eigenvalues of $2\widehat{D}_o^T \widehat{D}_i - P$. We will do this in a different way.

17.3 Eigenvalues of Quantum Walks

Let U be given by

$$U := 2\widehat{D}_o^T \widehat{D}_i - P.$$

We will determine the eigenvalues of U, but before we do this we show that the answer will be simpler than you might expect.

The matrix U is the product of $2\widehat{D}_o^T \widehat{D}_o - I$ and P, and these two matrices are involutions—their squares are the identity matrix.

17.3.1 Theorem. *Let P and Q be $v \times v$ matrices such that $P^2 = Q^2 = I$. If z is an eigenvector for PQ, then the subspace spanned by $\{z, Pz, Qz\}$ has dimension at most two, and is invariant under the algebra $\langle P, Q\rangle$.*

Proof. Suppose $z \neq 0$ and $PQz = \theta z$. Since $(PQ)^{-1} = QP$,

$$QPz = \theta^{-1} z$$

and therefore

$$P\{z, Pz, Qz\} = \{Pz, z, \theta z\}, \quad Q\{z, Pz, Qz\} = \{Qz, \theta^{-1} z, z\}.$$

This proves that the span of z, Pz and Qz is $\langle P, Q\rangle$-invariant. The image of this subspace under P is spanned by z and Pz, and so its dimension is at most two. Since P is invertible it follows that z, Pz and Qz are linearly dependent. □

If the involutions P and Q here are orthogonal, then the orthogonal complement of a $\langle P, Q\rangle$-invariant subspace is $\langle P, Q\rangle$-invariant. Hence:

17.3.2 Corollary. *If P and Q are orthogonal involutions of order $v \times v$, there is a basis of $\mathbb{R}^n$ with respect to which P and Q are block diagonal, and each block has order at most two.* □

We turn to details.

17.3.3 Theorem. *Let X be a graph with v vertices and let U be the orthogonal matrix representing the quantum walk on* $\mathrm{LD}(X)$. *Then $\mathbb{R}^n$ is the direct sum of the orthogonal subspaces*

$$\operatorname{col} \widehat{D}_i^T + \operatorname{col} \widehat{D}_o^T, \qquad \ker(D_o) \cap \ker(D_i).$$

The first subspace decomposes into an orthogonal direct sum of the space spanned by the constant vectors and 2-dimensional subspaces $C(\lambda)$, where λ runs over the eigenvalues of $\widehat{A}$. The eigenvalues of U on $C(\lambda)$ are $\lambda \pm \sqrt{\lambda^2 - 1}$. The second subspace is the direct sum of two subspaces K_1 and K_{-1}; we have $U \restriction K_1 = I$ and $U \restriction K_{-1} = -I$. The eigenvalues 1 and -1 have equal multiplicity.

Proof. We have

$$U\widehat{D}_i^T = 2\widehat{D}_o^T \widehat{D}_i \widehat{D}_i^T - P\widehat{D}_i^T = 2\widehat{D}_o^T - \widehat{D}_o^T = \widehat{D}_o^T \tag{17.3.1}$$

and

$$U\widehat{D}_o^T = 2\widehat{D}_o^T\widehat{D}_i\widehat{D}_o^T - P\widehat{D}_o^T = 2\widehat{D}_o^T\widehat{A} - \widehat{D}_i^T. \tag{17.3.2}$$

Consequently

$$U^2\widehat{D}_i^T = 2U\widehat{D}_i^T\widehat{A} - \widehat{D}_i^T.$$

Suppose z is an eigenvector for A with eigenvalue λ, and $y := \widehat{D}_i^T z$. Then

$$U^2 y = U^2\widehat{D}_i^T z = 2U\widehat{D}_i^T Az - \widehat{D}_i^T z = 2\lambda Uy - y$$

and so

$$(U^2 - 2\lambda U + I)y = 0. \tag{17.3.3}$$

It follows from Equations (17.3.1) and (17.3.2) that the subspace sum

$$\operatorname{col}\widehat{D}_i^T + \operatorname{col}\widehat{D}_o^T$$

is U-invariant and consequently its orthogonal complement

$$\ker(D_o) \cap \ker(D_i)$$

is also U-invariant. These subspaces are also P-invariant, and since $U = 2\widehat{D}_o^T\widehat{D}_i - P$, the restrictions of U and P to $\ker(D_o) \cap \ker(D_i)$ are equal and so the eigenvalues of U on this subspace are ± 1.

Equation (17.3.3) shows that y and Uy span a U-invariant subspace. If this subspace is 1-dimensional, then y is an eigenvector for U and

$$y \in \operatorname{col}\widehat{D}_i^T \cap \operatorname{col}\widehat{D}_o^T.$$

Therefore y is constant on arcs with a given initial vertex, and constant on arcs with a given final vertex. It follows that y is constant on the arcs in a given component of X, and its eigenvalue is 1.

If y and Uy span a 2-dimensional space, then the minimal polynomial on this subspace is $t^2 - 2\lambda t + 1$ and the eigenvalues of U on this subspace are

$$\lambda \pm \sqrt{\lambda^2 - 1}.$$

Also

$$Uy = U\widehat{D}_i^T z = \widehat{D}_o^T z$$

and therefore y and Uy both lie in $\operatorname{col}\widehat{D}_i^T + \operatorname{col}\widehat{D}_o^T$.

Since the trace of U on a subspace $C(\lambda)$ is zero and

$$\operatorname{tr}(U) = \operatorname{tr}(2\widehat{D}_o^T\widehat{D}_i - P) = 2\operatorname{tr}(\widehat{D}_o^T\widehat{D}_i) = 2\operatorname{tr}(\widehat{D}_i\widehat{D}_o^T) = 2\operatorname{tr}(\widehat{A}) = 0,$$

it follows that 1 and -1 have equal multiplicity. □

Chapter 18

Lie Algebras

We study Lie algebras because they force themselves on us when we study the Terwilliger algebra of the binary Hamming scheme. As we will see, there are other combinatorial applications. Additionally we will work with the universal enveloping algebra of a Lie algebra, which provides a useful example of an infinite dimensional algebra.

18.1 Basics

A *Lie algebra* over a field $\mathbb{F}$ is a vector space with a multiplication $[a,b]$ such that

(a) $[b,a] = -[a,b]$.

(b) For all a, b and c, we have the *Jacobi identity*:

$$[a,[b,c]] + [b,[c,a]] + [c,[a,b]] = 0.$$

The only fields we will use in this context are $\mathbb{R}$ and $\mathbb{C}$, whence we see that $[a,a] = 0$ for all a. We call $[a,b]$ the *Lie bracket* or *commutator* of a and b, and we abbreviate $[a,[b,c]]$ to $[a,b,c]$. A Lie algebra is *abelian* if $[a,b] = 0$ for all a and b.

Note that a Lie algebra is **not** and algebra in the sense we have used elsewhere—the multiplication is not even associative in general.

We offer examples:

(a) $gl(n,\mathbb{F})$, the Lie algebra of all $n \times n$ matrices over $\mathbb{F}$, where

$$[A,B] := AB - BA.$$

(b) The real skew symmetric matrices of order $n \times n$ form a Lie algebra over $\mathbb{R}$.

(c) $\mathbb{R}^3$ with the cross product. We will use $a \wedge b$ to denote the cross product.

(d) A *derivation* of a commutative algebra $\mathscr{A}$ over $\mathbb{F}$ is a map $\delta : \mathscr{A} \to \mathbb{F}$ such that
$$\delta(fg) = \delta(f)g + f\delta(g).$$
You may check that the product of two derivations is not in general a derivation, but their Lie bracket is, and further the set of derivations of $\mathscr{A}$ is a Lie algebra. By way of a more specific example take $\mathscr{A}$ to be the polynomial ring $\mathbb{F}[x_1, \ldots, x_d]$ and note that, for each i, partial differentiation with respect to x_i is a derivation.

The construction in (a) can be usefully generalized: if $\mathscr{A}$ is an algebra over $\mathbb{F}$, then the multiplication
$$[a, b] := ab - ba$$
gives us a Lie algebra. Thus if V is a vector space, then $\mathrm{End}(V)$ is a Lie algebra under this operation. For fixed a in $\mathscr{A}$, the map from $\mathscr{A}$ to itself given by
$$x :\mapsto [a, x]$$
is a derivation (as you should check).

A subspace of a Lie algebra $\mathscr{L}$ is subalgebra if it is closed under the Lie bracket. You could check that the subspace of skew symmetric matrices is a subalgbra of $gl(n, \mathbb{F})$. A subspace U of $\mathscr{L}$ is an *ideal* if $[a, u] \in U$, for all u in U. The subspace of strictly upper triangular matrices is an ideal in the Lie algebra formed by the set of all upper triangular matrices.

If $\mathscr{L}$ is a Lie algebra and S, T are subsets of $\mathscr{L}$, then we define $[S, T]$ to be the subspace of $\mathscr{L}$ spanned by the set
$$\{[x, y] : x \in S, y \in T\}.$$
In particular the subspace $[\mathscr{L}, \mathscr{L}]$ is a subalgebra of $\mathscr{L}$, called the *commutator subalgebra*.

For example, suppose $\mathscr{L} = gl(V)$. Then for any A and B in $\mathscr{L}$, we have
$$\mathrm{tr}[A, B] = \mathrm{tr}(AB) - \mathrm{tr}(BA) = 0.$$
So the commutator of $gl(V)$ consists of matrices with zero trace. It can be shown that it contains all matrices with zero trace. It is known as the special linear Lie algebra and is denoted by $sl(V)$. You may show that $sl(V)$ is equal to its commutator subalgebra.

18.2 Enveloping Algebras

The construction of the Lie algebra $gl(V)$ from the algebra $\mathrm{End}(V)$ can be generalized: if $\mathscr{A}$ is an algebra and $a, b \in \mathscr{L}$, we can define their Lie bracket by

$$[a,b] := ab - ba.$$

This leads us to ask which Lie algebras arise in this way, and the answer is that they all do. Let us denote the Lie algebra we get from $\mathscr{A}$ by $\mathrm{Lie}\mathscr{A}$. The *universal enveloping algebra* of $\mathscr{L}$ is essentially the smallest algebra $\mathscr{U}$ such that $\mathscr{L} = \mathrm{Lie}\mathscr{U}$. Of course the adjective 'universal' indicates that a category theorist has escaped. What we should say is that $\mathscr{U}$ is defined by the condition that if $\psi : \mathscr{L} \to \mathrm{Lie}\mathscr{A}$ for some algebra $\mathscr{A}$, then ψ can be factored into a Lie homomorphism from $\mathscr{L}$ to $\mathrm{Lie}\mathscr{U}$ and a Lie homomorphism from $\mathrm{Lie}\mathscr{U}$ to $\mathrm{Lie}\mathscr{A}$ induced by an algebra homomorphism from $\mathscr{U}$ to $\mathscr{A}$.

We consider a particular example, using the the Lie algebra $sl(2,\mathbb{R})$. The elements of this are the 2×2 matrices of trace zero, which form a vector space of dimension three, with basis

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We note that

$$[X,Y] = H$$

and that

$$[H,X] = 2X, \quad [H,Y] = -2Y.$$

The universal enveloping algebra of $sl(2,\mathbb{F})$ is the quotient of the free polynomial algebra in variables X, Y modulo the relations

$$XY - YX - H = 0, \quad HX - XH - 2H = 0, \quad HY - YH + 2Y = 0.$$

Note that this is an infinite-dimensional algebra—it can be shown that the elements $X^k Y^\ell H^m$ form a basis.

18.3 Posets

A poset is *ranked* if all elements covered by an element have the same height. If P is ranked then the i-th *level number* is the number of elements with height

i. Thus the poset formed by the subsets of $\{1,\ldots,n\}$, ordered by inclusion, is ranked and the i-th level number of $\binom{n}{i}$. If P is ranked with height d and the i-th level number is w_i, we say that P is *rank symmetric* if $w_i = w_{d-i}$ for all i, and we say P is *unimodal* if the sequence of level numbers is unimodal. The lattice of subsets of $\{1,\ldots,n\}$ is rank symmetric and unimodal.

An *antichain* in a poset P is set of elements such that no two are comparable. (Equivalently it is a coclique in the comparability graph of P.) The elements of given height in a poset form an antichain, and we say P is *Sperner* if the maximum size of an antichain is equal to the maximum level number. More generally we call P *strongly Sperner* if the maximum size of a subset that does not contain a chain of length $k+1$ is equal to the sum of the k largest level numbers. A *Peck poset* is a ranked poset that is rank symmetric, unimodal and strongly Sperner. The lattice of subsets of a finite set is Peck.

We use $\mathbb{P}$ to denote the the vector space $\mathbb{R}^P$. We can represent subsets of P by their characteristic vectors, which belong to $\mathbb{P}$. If $a \in P$ we will often denote the characteristic vector of a by a. The subspace of $\mathbb{P}$ spanned by the (characteristic vectors of) the elements of height i will be denoted by $\mathbb{P}(i)$.

Suppose P is a finite ranked poset. An element of $\mathrm{End}(\mathbb{P})$ is a *raising operator* if for each element a of P, the support of Ra is a subset of the elements of P that cover a. Similarly we define *lowering operators*. If R is a raising operator then R^T is lowering. Both raising and lowering operators are nilpotent: if P has height d and R is a raising operator, then $R^{d+1} = 0$.

The following result is due to Stanley and Griggs.

18.3.1 Theorem. *Let P be a rank-symmetric poset with height h. Then P is Peck if and only if there is an order-raising operator R such that the mappings*

$$R^{h-i} \restriction \mathbb{P}(i) : \mathbb{P}(i) \to \mathbb{P}(h-i), \qquad i = 0,\ldots,\left\lfloor \frac{h}{2} \right\rfloor \qquad \square$$

are invertible.

Using the above result, Proctor showed the following.

18.3.2 Theorem. *A ranked poset is Peck if and only if it has raising and lowering operators R and L such that the Lie algebra generated by R and L is isomorphic to $sl(2,\mathbb{C})$.* $\square$

We derive an important consequence of these results.

18.3.3 Corollary. *If P_1 and P_2 are Peck posets, then so is $P_1 \times P_2$.*

Proof. If $_1$ and P_2 are Peck then the vector spaces $\mathbb{P}_1$ and $\mathbb{P}_2$ are modules for $sl(2,\mathbb{C})$. Now

$$\mathbb{C}^{P_1 \times P_2} = \mathbb{P}_1 \otimes \mathbb{P}_2$$

and therefore $\mathbb{C}^{P_1 \times P_2}$ is a module for $sl(2,\mathbb{C})$. We conclude that $P_1 \times P_2$ is Peck. □

If U and V are modules for an algebra $\mathscr{A}$ then $U \otimes V$ is a module for $\mathscr{A} \times \mathscr{A}$, but it is **not** in general a module for $\mathscr{A}$. However it is module for $\mathscr{A}$ when $\mathscr{A}$ is an enveloping algebra of a Lie algebra (and when $\mathscr{A}$ is a group algebra).

18.4 Representations of Lie Algebras

A linear map ψ from a Lie algebra $\mathscr{L}_1$ to a Lie algebra $\mathscr{L}_2$ is a *homomorphism* if

$$\psi([a,b]) = [\psi(a), \psi(b)].$$

A *representation* of a Lie algebra $\mathscr{L}$ is a homomorphism into $gl(n,\mathbb{F})$. More generally ψ could be a homomorphism into $\mathrm{End}(V)$ for some vector space V; in this case we may say that V is a *module* over $\mathscr{L}$. A subspace of V that is invariant under the operators in $\psi(\mathscr{L})$ is a *submodule*. (Calling V a module for $\mathscr{L}$ is a courtesy, since modules are defined over rings—if we wish to be precise, it is a module for the enveloping algebra.)

If $\mathscr{L}$ is a Lie algebra and $A \in \mathscr{L}$, we define the *adjoint map* ad_A by

$$\mathrm{ad}_A(X) := [A, X].$$

This is a linear map, and is a derivation of the enveloping algebra. By Jacobi's identity

$$\begin{aligned} \mathrm{ad}_A([X,Y]) = [A,[X,Y]] &= -[X,[Y,A]] - [Y,[A,X]] \\ &= [X,[A,Y]] + [[A,X],Y]. \end{aligned}$$

We also have, by appeal to Jacobi

$$\begin{aligned} (\mathrm{ad}_X \mathrm{ad}_Y - \mathrm{ad}_Y \mathrm{ad}_X)(Z) &= [X,[Y,Z]] - [Y,[X,Z]] \\ &= [X,[Y,Z]] + [Y,[Z,X]] \\ &= [[X,Y],Z] \\ &= \mathrm{ad}_{[X,Y]}(Z), \end{aligned}$$

which shows that ad_A is a homomorphism from $\mathscr{L}$ into the Lie algebra End(L).

An element A of $\mathscr{L}$ is *ad-nilpotent* if ad_A is nilpotent. We observe that

$$\begin{aligned}\mathrm{ad}_A(X) &= [A,X],\\ (\mathrm{ad}_A)^2(X) &= [A,[A,X]],\\ (\mathrm{ad}_A)^3(X) &= [A,[A,[A,X]]]\end{aligned}$$

and in general, $(\mathrm{ad}_A)^{k+1}(X) = [A,(\mathrm{ad}_A)^k(X)]$. If $A \in gl(n,\mathbb{F})$, then we may represent the linear map ad_A by

$$A \otimes I - I \otimes A.$$

It follows that if $A^k = 0$, then $(\mathrm{ad}_A)^{2k} = 0$. In particular if A in $gl(V)$ is nilpotent then ad_A is nilpotent. Thus we have the fortunate conclusion that nilpotent elements of $gl(V)$ are ad-nilpotent.

18.5 Bilinear Forms

Suppose ψ is a representation of the Lie algebra $\mathscr{L}$ in End(V). A bilinear form β on V is *invariant* if

$$\beta(\psi(X)u,v) + \beta(u,\psi(X)v) = 0$$

for all u and v from V. By way of example, if V is $\mathscr{L}$ itself then

$$\beta(X,Y) := \mathrm{tr}(\mathrm{ad}_X \mathrm{ad}_Y)$$

is a symmetric bilinear form, known as the *Killing form*. We check that it is invariant.

$$\begin{aligned}\beta([A,X],Y) &= \mathrm{tr}(\mathrm{ad}_{[A,X]}\mathrm{ad}_Y)\\ &= \mathrm{tr}([\mathrm{ad}_X,\mathrm{ad}_Y]\mathrm{ad}_Y)\\ &= \mathrm{tr}(\mathrm{ad}_A\mathrm{ad}_X\mathrm{ad}_Y - \mathrm{ad}_X\mathrm{ad}_A\mathrm{ad}_Y)\end{aligned}$$

Similarly

$$\beta(X,[A,Y]) = \mathrm{tr}(\mathrm{ad}_X\mathrm{ad}_A\mathrm{ad}_Y - \mathrm{ad}_X\mathrm{ad}_Y\mathrm{ad}_A)$$

from which we see that β is invariant. (Thus the adjoint of ad_X relative to the Killing form is $-\mathrm{ad}_X$.)

Suppose $\mathscr{L}$ is a Lie algebra with a non-degenerate invariant bilinear form. If $X_1, \ldots, X_d$ is a basis for $\mathscr{L}$, there is a dual basis $Y_1, \ldots, Y_d$ such that

$$\beta(X_i, Y_j) = \delta_{i,j}.$$

The *Casimir element* of the universal enveloping algebra is defined to be

$$\sum_{i=1}^{d} X_i Y_i.$$

18.5.1 Theorem. *Let $\mathscr{L}$ be a Lie algebra with a non-degenerate invariant bilinear form β. Then the Casimir element is independent of the choice of basis for $\mathscr{L}$, and lies in the center of the universal enveloping algebra.*

Proof. Let $X_1, \ldots, X_d$ be a basis for $\mathscr{L}$ with dual basis $Y_1, \ldots, Y_d$ and let Δ be the Casimir element defined using this pair of bases. Let $U_1, \ldots, U_d$ and $V_1, \ldots, V_d$ be a second pair of dual bases. Then there are scalars $\rho_{i,j}$ and $\sigma_{i,j}$ such that

$$U_i = \sum_k \rho_{i,k} X_k,$$
$$V_j = \sum_\ell \sigma_{j,\ell} Y_\ell.$$

We have

$$\sum_i U_i V_i = \sum_{i,k,\ell} \rho_{i,k} \sigma_{i,\ell} X_i Y_i \tag{18.5.1}$$

Since $\beta(X_i, Y_j) = \delta_{i,j}$, we have

$$\delta_{i,j} = \beta(U_i, V_j) = \sum_k \rho_{i,k} \sigma_{j,k}$$

So if we define matrices R and S by $R := (\rho_{i,j})$ and $S := (\sigma_{i,j})$ then $RS^T = 0$. Consequently $SR^T = 0$ and therefore

$$\delta_{k,\ell} = \sum_i \rho_{i,k} \sigma_{i,\ell}.$$

Hence (18.5.1) implies that $\sum_i U_i V_i = \Delta$.

We now prove Δ lies is central. Suppose $A \in \mathscr{L}$. There are scalars $\alpha_{i,j}$ and $\beta_{i,j}$ such that

$$[A, X_i] = \sum_j \alpha_{i,j} X_j$$

and

$$[A, Y_i] = \sum_j \beta_{i,j} Y_j$$

Since β is invariant,

$$0 = \beta([A, X_i], Y_j) + \beta(X_i, [A, Y_j]) = \alpha_{i,j} + \beta_{j,i}.$$

This implies that

$$\sum_i [A, X_i] Y_i = \sum_{i,j} \alpha_{i,j} X_j Y_i = -\sum_{i,j} \beta_{j,i} X_j Y_i = -\sum_i X_i [A, Y_i].$$

Now we compute that

$$A\Delta = \sum_i A X_i Y_i = \sum_i [A, X_i] Y_i + \sum_i X_i A Y_i$$

and

$$\Delta A = \sum_i X_i Y_i A = -\sum_i X_i [A, Y_i] + \sum_i X_i A Y_i,$$

whence we conclude that $A\Delta = \Delta A$. □

18.5.2 Lemma. *If Δ is the Casimir element of the Lie algebra $\mathscr{L}$ and φ is a representation of $\mathscr{L}$, then* $\operatorname{tr}(\varphi(\Delta)) = \dim(\varphi(\mathscr{L}))$. □

18.6 An Example

We compute the Casimir element for $sl(2,\mathbb{C})$, relative to the form

$$\beta(X, Y) := \operatorname{tr}(\operatorname{ad}_X \operatorname{ad}_Y).$$

Recall that X, H and Y form a basis, where

$$X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

and

$$[X, Y] = H, \quad [H, X] = 2X, \quad [H, Y] = -2Y.$$

It follows that

$$\mathrm{ad}_X = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathrm{ad}_Y = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}, \quad \mathrm{ad}_H = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

If

$$A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix},$$

then

$$\mathrm{ad}_A = \begin{pmatrix} 2a & -2b & 0 \\ -c & 0 & b \\ 0 & 2c & -2a \end{pmatrix}$$

and now it is easy verify that if

$$\beta(A,X) = \beta(A,H) = \beta(A,Y) = 0,$$

then $A = 0$. Therefore β is nondegenerate.

Next we calculate that

$$\beta(X,Y) = \beta(Y,X) = 4, \quad \beta(H,H) = 8$$

and all other inner products are zero. So the dual basis to (X,H,Y) is

$$\left(\frac{1}{4}Y, \frac{1}{4}X, \frac{1}{8}H\right)$$

and the Casimir element is

$$\Delta := \frac{1}{4}(XY + YX + \frac{1}{2}H^2).$$

Using the fact that

$$[A,BC] = [A,B]C + B[A,C],$$

it is not hard to verify directly that Δ is central.

18.7 Irreducible Modules

We construct a family of irreducible modules for $sl(2,\mathbb{C})$, by constructing irreducible modules for its enveloping algebra.

18.7.1 Lemma. *Let $\mathscr{U}$ denote the enveloping algebra of $sl(2,\mathbb{C})$, with generators X, Y and H, and suppose V is a module for $\mathscr{U}$ with finite dimension. If v is an eigenvector for H in its action on V, then there are integers k and ℓ such that $X^k v = 0$ and $Y^\ell v = 0$.*

Proof. Suppose $Hv = \lambda v$. Recalling that $[H,X] = 2X$, we have

$$HXv = (XH + 2X)v = \lambda Xv + 2Xv = (\lambda + 2)Xv.$$

Hence if $Xv \neq 0$ and λ is an eigenvalue of H, then $\lambda + 2$ is also an eigenvalue of H. A similar calculation shows that if $Yv \neq 0$, then Yv is an eigenvector for H with eigenvalue $\lambda - 2$. □

Note that XYv is an eigenvector for H with eigenvalue λ, consistent with the fact that H and XY commute.

If V is a module for $\mathscr{U}$, an element v of V has *weight* λ if $Hv = \lambda v$. If $Hv = \lambda v$ and also $Xv = 0$, we say that v is a *highest weight vector* of weight λ. The eigenspaces of H are often called *weight spaces*. We have seen that every finite-dimensional module for $\mathscr{U}$ must contain a highest weight vector; the following theorem completely specifies the structure of the cyclic $\mathscr{U}$-module generated by a highest weight vector.

18.7.2 Theorem. *Suppose V is a module for $\mathscr{U}$ and v is a highest weight vector in V with eigenvalue λ. Let d be the least non-negative integer such that $Y^d v = 0$. Then $\lambda = d-1$, the cyclic $\mathscr{U}$-module generated by v is simple and the vectors*

$$v, Yv, \ldots, Y^{d-1}v$$

form a basis for it. Further, for $k = 0, 1, \ldots, d-1$,

$$HY^k v = (d-1-2k)Y^k v, \qquad XY^k v = k(d-k)Y^{k-1}.$$

Proof. The adjoint map ad_H is a derivation of $\mathscr{U}$ whence

$$[H, Y^k] = [H,Y]Y^{k-1} + Y[H, Y^{k-1}]$$

and a trivial induction yields that

$$[H, Y^k] = -2kY^k.$$

If $Hv = \lambda v$, we have

$$HY^k v = [H, Y^k]v + Y^k Hv = -2kY^k v + \lambda Y^k v = (\lambda - 2k)Y^k v.$$

Let d be the least integer such that $Y^d v = 0$. Then the vector space V_1 spanned by the vectors

$$v, Yv, \ldots, Y^{d-1}v$$

has dimension d, and these vectors for a basis for it. Since these vectors are all eigenvectors for H, we see that V_1 is invariant under both Y and H. We prove that it is X-invariant.

We have

$$[X, Y^k] = [X, Y]Y^{k-1} + Y[X, Y^{k-1}]$$

Since $Xv = 0$, it follows that

$$XY^k v = [X, Y^k]v = HY^{k-1}v + Y[X, Y^{k-1}]v$$

and so by induction we have

$$XY^k v = HY^{k-1}v + YHY^{k-2}v + \cdots + Y^{k-1}Hv.$$

Since the vectors $Y^k v$ are eigenvectors for H, this implies that $XY^k v = c_k Y^{k-1}v$, for some constant c_k and therefore V_1 is a module for $\mathscr{U}$. We have

$$c_k = (\lambda + 2 - 2k) + (\lambda + 4 - 2k) + \cdots + \lambda = k\lambda - (k^2 - k) = k(\lambda - k + 1).$$

We see that c_d is the sum of the eigenvalues of H on V_1 and so $c_d = \mathrm{tr}(H)$. As $H = XY - YX$ we have $\mathrm{tr}(H) = 0$, and therefore $\lambda = d - 1$.

It remains to prove that V is simple. Suppose V_1 is a non-zero submodule of V. Then V_1 contains a highest weight vector u, and since u is an eigenvector for H it must be a non-zero scalar multiple of one of the vectors $Y^i v$. Since $Xu = 0$, we see see that u is a non-zero scalar multiple of v. Hence the cyclic module generate by u is equal to V and therefore $V_1 = V$. □

This result implies that the $\mathscr{U}$ module generated by a highest weight vector v is determined by its dimension (or by the eigenvalue of v). Also note that any simple module is isomorphic to one of the modules described in this theorem, since any module contains a highest weight vector.

18.7.3 Corollary. *If C is the Casimir element of $sl(2,\mathbb{C})$, then $CY^k v = (d^2 - 1)Y^k v$.*

Proof. From above we have

$$\begin{aligned} XY\,Y^k v &= (k+1)(d-k-1)Y^k v \\ YX\,Y^k v &= k(d-k)Y^k v \\ H\,Y^k v &= (d-1-2k)Y^k v \end{aligned}$$

and the claim follows easily from these. □

18.8 Semisimple Elements

We derive two useful identities that hold in the enveloping algebra of $sl(2,\mathbb{C})$. We define

$$H_k := H + kI$$

and we define $H_{k;r}$ recursively by $H_{k;0} = I$ and

$$H_{k;i+1} := H_{k;i} H_{k-i+1}.$$

18.8.1 Lemma. *We have*

$$X^m Y^n = \sum_{r=0}^{m\wedge n} r! \binom{m}{r}\binom{n}{r} Y^{n-r} X^{m-r} H_{n-m;r}$$

Proof. First prove by induction that if $n \geq 1$, then

$$X^n Y = Y X^n + n X^{n-1} H_{n-1} \tag{18.8.1}$$

and then, by a second induction, derive the lemma. □

18.8.2 Lemma. *In a finite dimensional representation of $\mathscr{U}(sl(2,\mathbb{C}))$, if $X^k = 0$ then*

$$\prod_{r=-k+1}^{k-1} (H - rI) = 0.$$

Proof. We do not give a complete proof, but offer a generous hint and leave the details as an exercise.

Suppose V is a finite-dimensional representation for $\mathscr{U}$. The idea is to prove that, if $X^k = 0$, then for $i = 1, \ldots, k$ we have

$$X^{k-i} H_{k-1;2i-1} = 0.$$

Setting $i = k$ in this yields the result.

For convenience we prove the above claim in the case $k = 4$. We have the following equations:

$$X^4 Y = Y X^4 + 4X^3 H_{3;1} \tag{18.8.2}$$

$$X^4 Y^2 = Y^2 X^4 + 8Y X^3 H_{2;1} + 12X^2 H_{2;2} \tag{18.8.3}$$

$$X^4 Y^3 = Y^3 X^4 + 12Y^2 X^3 H_{1;1} + 36Y X^2 H_{1;2} + 24X H_{1;3} \tag{18.8.4}$$

$$X^4 Y^4 = Y^4 X^4 + 16Y^3 X^3 H_{0;1} + 72Y^2 X^2 H_{0;2} + 216Y X H_{0;3} + 24H_{0;4} \tag{18.8.5}$$

Since $X^4 = 0$ we see that (18.8.2) implies

$$X^3 H_{3;1} = 0.$$

Now multiply (18.8.3) on the right by H_3; since $XH_i = H_{i-2}X$, we get

$$0 = 8YX^3 H_3 H_2 + 12X^2 H_1 H_2 H_3$$

and since $YX^3 H_3 = 0$, we deduce that

$$X^2 H_{3;3} = 0.$$

Next multiply (18.8.4) on the right by $H_2 H_3$ and deduce that since

$$YX^2 H_{1;2} H_2 H_3 = YX^2 H_0 H_1 H_2 H_3 = YX^2 H_{3;3} H_0 = 0,$$

that

$$XH_{3;5} = 0.$$

Finally multiply (18.8.5) on the right by $H_{1;3}$ to deduce that

$$H_{3;7} = 0.$$ □

Recall that H, XY and YX all commute.

18.8.3 Lemma. *If $1 \le k \le n$, then*

$$X^n Y^k = \Big(\prod_{i=0}^{k-1} (YX + (n-i)H_{-n+i+1})\Big) X^{n-k}$$

Proof. From (18.8.1) we have

$$X^n Y = YX^n + nX^{n-1} H_{n-1} = YX^n + nH_{-n+1} X^{n-1} = (YX + nH_{-n+1}) X^{n-1}$$

and use induction on k. □

18.8.4 Theorem. *In a finite-dimensional representation of $\mathscr{U}(sl(2,\mathbb{C}))$, the images of H, XY and YX are semisimple.*

Proof. Since H and XY commute and $YX = XY - H$, it is enough to show that H and YX are semisimple. By Lemma 18.7.1, there is an integer k such that $X^k = 0$. From Lemma 18.8.2 it follows that H is semimple, and so the underlying vector space V is a direct sum of eigenspaces of H. Suppose V_λ is one of these eigenspaces, where λ is the eigenvalue of H.

By Lemma 18.8.3 we have

$$0 = X^k Y^k = (YX + k(H - (k-1)I) \cdots (YX + H)$$

and if $z \in V_\lambda$, then

$$0 = (YX + k(\lambda - (k-1)I) \cdots (YX + \lambda)z.$$

Hence the minimal polynomial of YX on V_λ has only simple zeros, and therefore YX is semisimple on V_λ. We conclude that YX must be semisimple. □

18.9 Semisimple Modules

18.9.1 Theorem. *Any finite dimensional module for $\mathcal{U}(sl(2,\mathbb{C}))$ is semisimple.*

Proof. Let $\mathcal{U}$ denote $\mathcal{U}(sl(2,\mathbb{C}))$, let M be a finite-dimensional $\mathcal{U}$-module, and let C be the Casimir element of $\mathcal{U}$. Since C is central and semisimple, M is the direct sum of eigenspaces of C, and so to prove the theorem it will suffice if we show that any eigenspace for C semisimple.

Hence we assume that M itself is an eigenspace for C. Since H also is semisimple, M is the direct sum of weight spaces M_σ and, if $N \leq M$, then N is the direct sum of its weight space N_σ, where

$$N_\sigma = N \cap M_\sigma.$$

We have

$$\dim(M_\sigma) = \dim(N_\sigma) + \dim(M_\sigma / N_\sigma).$$

Note that M/N is a $\mathcal{U}$-module and

$$(M/N)_\sigma = M_\sigma / N_\sigma.$$

Next assume we have the composition series for M:

$$0 = M_0 < M_1 < \cdots < M_r = M.$$

Then

$$\dim(M_\sigma) = \sum_{i=1}^{r} \dim(M_i/M_{i-1})_\sigma$$

but M_i/M_{i-1} is a simple $\mathscr{U}$-module and consequently $\dim(M_i/M_{i-1})_\sigma = 1$. We conclude that $\dim(M_\sigma) = r$ and that $\dim(M)$ is r times the number of eigenvalues of H. The cyclic $\mathscr{U}$-submodule of M generated by a non-zero element is simple and, since all non-zero elements of M are eigenvectors for C with the same eigenvalue, all these simple modules have the same dimension.

Choose a basis $x_1, \ldots, x_d$ for M. Then

$$M = x_1\mathscr{U} + \cdots + x_d\mathscr{U}.$$

where each submodule $x_i\mathscr{U}$ contains a simple submodule S_i (say). (We do not assume that this is a direct sum.) Since $\dim(M_\sigma) = r$, we have $d = r$. Since $x_1, \ldots, x_r$ is a basis, the sum

$$S_1 + \cdots + S_r$$

is direct and therefore $\dim(M)$ is bounded below by r times the number of eigenvalues of H. But we saw that equality holds, and therefore M is a direct sum of simple modules as required. □

This proof follows Jantzen [**?**].

Chapter 19

Terwilliger Algebras

Let $\mathscr{A}$ be an association scheme with d classes and let π be an equitable partition of its vertex set with e classes. Define the diagonal 01-matrix F_i by setting $(F_i)_{u,u} = 1$ if u lies in the i-th class of π. Then the matrices F_i are symmetric idempotents and

$$\sum_i F_i = I.$$

We will study the algebra generated by $\mathscr{A}$ together with the matrices F_i.

If u is a vertex in the scheme and the i-th cell of π consists of the vertices x such that (u,x) lies in the i-th relation, the algebra we get is the *Terwilliger algebra* of the scheme relative to the vertex u.

19.1 Modules

Our basic task is to determine the irreducible modules of the Terwilliger algebra. Suppose $\mathscr{A}$ is an association scheme with d classes $A_0, \ldots, A_d$ and vertex set V, and assume $|V| = v$. Let $\mathbb{T}$ denote the Terwilliger algebra of this scheme and suppose W is an irreducible $\mathbb{T}$-module. Since W is invariant under $\mathscr{A}$, it must have basis that consists of eigenvectors for $\mathscr{A}$. Similarly it must have basis that consists of eigenvectors for the matrices F_i, that is, vectors whose supports are subsets of the cells of the partition π.

The subspace spanned by the characteristic vectors of the cells of π is $\mathbb{T}$-invariant and has dimension equal to $|\pi|$, the number of cells of π. We call it the *standard module* It is a cyclic $\mathbb{T}$-module, generated by $\mathbf{1}$. You may prove that it is irreducible.

This may seem an encouraging start to determining the irreducible modules for the Terwilliger algebra, but unfortunately further progress will require much more effort. Since $\mathbb{T}$ is transpose-closed, $\mathbb{R}^v$ decomposes into an orthogonal sum of irreducible $\mathbb{T}$-modules. Hence if W is irreducible and is not the standard module, we may assume that it is orthogonal to it. Thus each element of W will be orthogonal to the vectors $F_i\mathbf{1}$—it sums to zero on the cells of π.

19.1.1 Lemma. *If W is an irreducible module for an algebra $\mathscr{B}$ and f is an idempotent in $\mathscr{B}$, then Wf is an irreducible module for $f\mathscr{B}f$.*

Proof. We may assume $\dim(W) \geq 2$, or there is nothing to prove. Since

$$Wf\,f\mathscr{B}f = Wf\mathscr{B}f \leq Wf,$$

we see that Wf is a module for $f\mathscr{B}f$.

Suppose U is an $f\mathscr{B}f$-submodule of Wf. Each element of U can be written as wf where $w \in W$ and as $f^2 = f$, it follows that $Uf = U$. Since $Uf\mathscr{B}$ is a $\mathscr{B}$-submodule of W, it is either zero or equal to W. If it is equal to W, then

$$U = Uf\mathscr{B}f = Wf$$

and therefore Wf is irreducible for $f\mathscr{B}f$.

To complete the proof, we show that $Uf\mathscr{B}$ cannot be zero. The key is to note that the set

$$\{u \in W : u\mathscr{B} = 0\}$$

is a $\mathscr{B}$-submodule of W. Since W is simple and not zero, it follows that this set must be the zero module. Consequently $Uf\mathscr{B}$ cannot be zero. □

Note that $f\mathscr{B}f$ is a subspace of $\mathscr{B}$ and is closed under multiplication, but fails to be a subalgebra because it does not contain I (in general). However

$$f\mathscr{B}f + (I-f)\mathscr{B}(I-f)$$

is a subalgebra of $\mathscr{B}$.

When we want to use Lemma 19.1.1, we will have two possible sources of idempotents: the matrices F_i and the principal matrix idempotents E_j.

19.2 Thinness

Let $\mathbb{T}$ be the Terwilliger algebra for an association scheme $\mathscr{A}$ and let W be a $\mathbb{T}$-submodule of $\mathbb{R}^v$. We say that W is *thin* if for each i we have

$$\dim(F_i W) \leq 1.$$

We also say that W is *dual thin* if for each j,

$$\dim(E_j W) \leq 1.$$

We generalise the concept of thinness. Suppose $\mathscr{B}$ is an algebra. We say that a set of idempotents $F_1, \ldots, F_r$ is a *resolution of the identity* if they are pairwise orthogonal ($F_i F_j = 0$ when $i \neq j$) and

$$\sum_i F_i = I.$$

A module W for $\mathscr{B}$ is *thin relative to the resolution* $F_1, \ldots, F_r$ if $\dim(F_i W) \leq 1$ for all i.

Being thin is not easy, but it is a desirable property that holds in many interesting cases.

19.2.1 Lemma. *If $\mathscr{A}$ is an association scheme then the standard modules are thin and dual thin,*

Proof. Exercise. □

19.2.2 Theorem. *If the algebra $\mathscr{B}$ is self-adjoint, then it is thin relative to the resolution $F_1, \ldots, F_r$ if and only if the subalgebra*

$$F_1 \mathscr{B} F_1 + \cdots + F_r \mathscr{B} F_r$$

is commutative.

19.2.3 Lemma. *Suppose $\mathbb{T}$ is the Terwilliger algebra of an association scheme relative to some vertex. If each matrix in*

$$F_0 \mathbb{T} F_0 + \cdots + F_e \mathbb{T} F_e$$

is symmetric, or if $\mathrm{Aut}(X)_1$ is generously transitive on each cell of π, then $\mathbb{T}$ is thin.

Proof. For the first, two symmetric matrices commute if and only if their product is symmetric. The second condition implies that each $F_i \mathbb{T} F_i$ is the Bose-Mesner algebra of a symmetric association scheme. □

19.3 Jaeger Algebras

We define some endomorphisms of $\mathrm{Mat}_{v\times v}(\mathbb{C})$. If A is a $v \times v$ matrix define the operators X_A and Y_A on $\mathrm{Mat}_{v\times v}(\mathbb{C})$ by

$$X_A(M) := AM, \qquad Y_A(M) = MA^*$$

and if B is a $v \times v$ matrix, then we define Δ_B by

$$\Delta_B(M) := B \circ M.$$

Note that

$$Y_A(Y_B(M)) = MB^*A^* = M(AB)^* = Y_{AB}(M),$$

which explains the A^* in the definition of Y_A. Also X_A and Y_B commute, for any A and B.

If $\mathscr{A}$ is an association scheme, we define $\mathscr{J}_2$ to be the algebra generated by the matrices X_A for A in $\mathbb{C}[\mathscr{A}]$. We define $\mathscr{J}_3(\mathscr{A})$ to be the algebra generated by the operators

$$X_A,\ \Delta_B, \qquad A, B \in \mathbb{C}[\mathscr{A}].$$

We obtain $\mathscr{J}_4(\mathscr{A})$ by adjoining the right multiplication operators Y_A as well

The vector space $\mathrm{Mat}_{v\times v}(\mathbb{C})$ is a module M for $\mathscr{J}_3$, and the subspace of matrices with all but the i-th column zero is a submodule, which we denote by $M(i)$. We see that M is the direct sum of the modules $M(i)$.

Our first result shows that $\mathscr{J}_3(\mathscr{A})$ is a kind of global Terwilliger algebra.

19.3.1 Lemma. *The algebra generated by the restriction to $M(i)$ of the operators in $\mathscr{J}_3$ is isomorphic to the Terwilliger algebra of $\mathscr{A}$ relative to the i-th vertex.*

Proof. We have

$$X_A(e_i e_j^T) = (Ae_i)e_j^T$$

and

$$\Delta_B(e_i e_j^T) = (B_{i,j}e_i)e_j^T.$$

So X_A is represented on $M(j)$ by the matrix A, and Δ_B by the diagonal matrix formed from the vector Be_j. □

We say that a $\mathscr{J}_3$-submodule U of $\mathrm{Mat}_{v\times v}(\mathbb{C})$ is *thin* if the subspaces Δ_{A_i} are 1-dimensional, and say that it is *dual thin* if the subspaces $X_{E_j}U$ are 1-dimensional.

19.3.2 Lemma. *If $\mathscr{A}$ is metric then a thin submodule of* $\mathrm{Mat}_{v\times v}(\mathbb{C})$ *is dual thin, if $\mathscr{A}$ is cometric then a dual thin submodule of* $\mathrm{Mat}_{v\times v}(\mathbb{C})$ *is thin.*

Proof. Suppose $\mathscr{A}$ is metric relative to the Schur idempotent A_1. If C is a $v\times v$ matrix, then

$$(A_1(A_i\circ C))\circ A_j = 0$$

if $|i-j|>1$. Hence if M is submodule of $\mathrm{Mat}_{v\times v}(\mathbb{C})$, then

$$A_1(A_i\circ M)\leq A_{i-1}\circ M + A_i\circ M + A_{i+1}\circ M. \tag{19.3.1}$$

Now let r denote the least positive integer such that $A_r\circ M\neq 0$, and let d be the greatest positive integer such that $A_{r+d-1}\circ M\neq 0$. From (19.3.1) it follows that if $r\leq i\leq r+d-1$ then $A_i\circ M\neq 0$. We also see that M is generated by the subspace $A_d\circ M$ as an X_{A_1}-module. In other terms,

$$M = \langle A_1\rangle(A_d\circ M).$$

If E_j is a matrix idempotent, then

$$E_jM = E_j\langle A_1\rangle(A_r\circ M) = E_j(A_r\circ M)$$

If M is thin, then $\dim(A_r\circ M)=1$ and therefore $\dim(E_jM)\leq 1$ for all j. Therefore M is dual thin.

Suppose $\mathscr{A}$ is cometric relative to E_1 and let s be the least integer such that $E_sM\neq 0$. Then each column of a matrix in E_jM lies in $\mathrm{col}(E_j)$, and so if $C\in M$, then each column of $E_1\circ(E_iM)$ is the Schur product of a column of E_1 with a vector in $\mathrm{col}(E_i)$. Hence by ??? we have

$$E_1\circ(E_iM)\leq E_{i-1}M + E_iM + E_{i+1}M.$$

Given this, it is easy to prove the second part of the theorem. □

Chapter 20

Strongly Regular Graphs

20.1 Strongly Regular Graphs

We apply the theory at hand to strongly regular graphs. Assume X is strongly regular with adjacency matrix A, and suppose that A has the partitioned form

$$A = \begin{pmatrix} 0 & \mathbf{1}^T & 0 \\ \mathbf{1} & B_1 & N^T \\ 0 & N & B_2 \end{pmatrix}$$

Thus B_1 is the adjacency matrix of the neighborhood of the vertex 1 in X, and B_2 is the adjacency matrix of the subgraph induced by the vertices at distance two from 1.

20.1.1 Theorem. *If X is a strongly regular graph and $\mathbb{T}$ is its Terwilliger algebra relative to some vertex, then an irrreducible $\mathbb{T}$-module lies in one of the following classes:*

(a) The standard module, with dimension three.

(b) Modules with dimension two, parameterized by eigenvectors w of B_1 such that Nw is an eigenvector for B_2.

(c) Modules with dimension one, arising from an eigenvector of B_1 in $\ker(N^T)$ or from an eigenvector of B_2 in $\ker(N)$. Each of these modules is an eigenspace for A.

Proof. Suppose W is an irreducible $\mathbb{T}$-module. We assume that W is not the standard module, and hence it lies in the orthogonal complement to the standard module. In particular $F_0W = 0$.

First we consider the case where $F_1W = 0$. In this case if $x \in W$, then $\mathrm{supp}(x)$ is a subset of the vertices at distance two from 1. Since

$$\begin{pmatrix} 0 & \mathbf{1}^T & 0 \\ \mathbf{1} & B_1 & N^T \\ 0 & N & B_2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ N^T w \\ B_2 w \end{pmatrix}$$

we conclude that if

$$x = \begin{pmatrix} 0 \\ 0 \\ w \end{pmatrix} \in W,$$

then $N^T w = 0$ and

$$\begin{pmatrix} 0 \\ 0 \\ B_2 w \end{pmatrix} \in W.$$

Since the span W_2 of the vectors $B_2^r w$ is B_2-invariant and since B_2 is symmetric, there is an eigenvector v for B_2 contained in W_2. Hence the vector

$$\begin{pmatrix} 0 \\ 0 \\ v \end{pmatrix}$$

spans a 1-dimensional $\mathbb{T}$-invariant subspace.

Similarly if $F_2WE = 0$, then $\dim(W) = 1$, and W is spanned by an eigenvector for B_1 that lies in $\ker(N)$.

So we assume that neither F_1W nor F_2W are zero. If

$$x = \begin{pmatrix} 0 \\ u \\ v \end{pmatrix} \in W,$$

then

$$\begin{pmatrix} 0 \\ u \\ 0 \end{pmatrix} = F_1 x \in W$$

and so

$$A\begin{pmatrix}0\\u\\0\end{pmatrix}=\begin{pmatrix}0\\B_1u\\Nu\end{pmatrix}\in W.$$

Accordingly the vectors

$$\begin{pmatrix}0\\B_1^ru\\0\end{pmatrix}$$

all lie in W and so there is an eigenvector w for B_1 such that

$$z:=\begin{pmatrix}0\\w\\0\end{pmatrix}\in W.$$

If we assume that $B_1w=\lambda w$, then

$$Az=A\begin{pmatrix}0\\w\\0\end{pmatrix}=\begin{pmatrix}0\\\lambda w\\Nw\end{pmatrix}\in W$$

The matrices I, A and J form a basis for the algebra of polynomials in A, and so the vectors

$$z,\quad Az,\quad Jz$$

span the cyclic A-module generated by z. Since $Jz=0$, this module is spanned by z and Az. It follows that that the span of z and Az contains an eigenvector for A, necessarily of the form

$$\begin{pmatrix}0\\w\\\beta Nw\end{pmatrix},$$

for some scalar β. If the eigenvalue for this eigenvector is θ, then

$$\theta\begin{pmatrix}0\\w\\\beta Nw\end{pmatrix}=A\begin{pmatrix}0\\w\\\beta Nw\end{pmatrix}=\begin{pmatrix}0\\B_1w+\beta N^TNw\\Nw+\beta B_2Nw\end{pmatrix}=\begin{pmatrix}0\\\lambda w+\beta N^TNw\\Nw+\beta B_2Nw\end{pmatrix}$$

whence we see that w is an eigenvector for N^TN and Nw is an eigenvector for B_2. Consequently the span of the vectors

$$\begin{pmatrix}0\\w\\0\end{pmatrix},\qquad\begin{pmatrix}0\\0\\Nw\end{pmatrix}$$

is $\mathbb{T}$-invariant; since these vectors are contained in W and since W is irreducible, they must be a basis for W.

20.1.2 Corollary. *The Terwilliger algebra of a strongly regular graph is thin and dual thin.* □

Proof. It follows from our work above that an irreducible $\mathbb{T}$-module is thin We prove that an irreducible $\mathbb{T}$-module is dual thin. Cases (a) and (c) are easy, so we turn to case (b). Suppose $w \neq 0$ and

$$x = \begin{pmatrix} 0 \\ w \\ 0 \end{pmatrix} \in W.$$

Then $E_0 x = 0$. If $E_1 x = 0$, then $x = E_2 x$ and so x is an eigenvector for A and thus generates a proper submodule of W. We conclude that neither $E_1 x$ nor $E_2 x$ can be zero, and hence W is dual thin. □

20.2 Local Eigenvalues

We apply the theory built up in the previous section to get information about local eigenvalues of strongly regular graphs.

Assume X is a strongly regular graph with parameters (v, k, a, c) and adjacency matrix partitioned as in the previous section. We assume that the eigenvalues of X are

$$k,\ \theta,\ \tau$$

where $\theta \geq 0$ and $\tau < 0$. We denote the multiplicities of θ and τ by $m(\theta)$ and $m(\tau)$. We use δ to denote $a - c$ and recall that

$$A^2 = \delta A + (k - c)I + cJ.$$

Since $\mathbb{T}$ is thin, the matrices

$$F_1 A F_1,\ F_1 A F_2 A F_1$$

commute and therefore the matrices

$$B_1,\ N^T N$$

also commute. Hence we can decompose each eigenspace of B_1 into the subspace of eigenvectors in $\ker(N^T N)$ and the subspace of eigenvectors orthogonal to $\ker(N^T N)$ or, equivalently, eigenvectors in $\ker N$ and eigenvectors in $\mathrm{col}(N^T)$).

If w is an eigenvector for B in $\ker(N)$ then

$$\begin{pmatrix} 0 \\ w \\ 0 \end{pmatrix}$$

is an eigenvector for A (with eigenvalue θ or τ) and spans a 1-dimensional irreducible $\mathbb{T}$-module.

Now suppose w is an eigenvector for B_1 with eigenvalue λ in $\mathbf{1}^\perp \cap \ker(N)$. We have

$$F_2 A^2 F_1 = NB_1 + B_2 N = \delta N + cJ$$

and consequently

$$\delta N w = (\delta N + cJ) w = (NB_1 + B_2 N) w = (\lambda I + B_2) N w.$$

Therefore Nw is an eigenvector for B_2 with eigenvalue $\delta - \lambda$, and the vectors

$$\begin{pmatrix} 0 \\ w \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ Nw \end{pmatrix}$$

span an irreducible $\mathbb{T}$-module with dimension two. It is possible to show in a similar fashion that if z is an eigenvector for B_2 with eigenvalue μ in $\mathbf{1}^\perp \cap \ker(N^T)$, then

$$B_1 N^T z = (\delta - \mu) N^T z$$

and the vectors

$$\begin{pmatrix} 0 \\ N^T z \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$$

span an irreducible $\mathbb{T}$-module with dimension two.

We note one further constraint. Since

$$F_1 A^2 F_1 = J + B_1^2 + N^T N = \delta B_1 + (k - c) I + cJ$$

we have

$$B_1^2 - \delta B_1 - (k - c) I = -N^T N + (c - 1) J$$

and so if $Bw = \lambda w$ and $Jw = 0$, then

$$(\lambda^2 - \delta\lambda - (k-c)) = -N^T N w.$$

Thus w is an eigenvalue for $N^T N$, and since this matrix is positive semidefinite, it follows that

$$\lambda^2 - \delta\lambda - (k-c) \leq 0,$$

or equivalently that

$$\tau \leq \lambda \leq \theta.$$

For later use, note that

$$t^2 - \delta t - (k-c) = (t-\theta)(t-\tau).$$

A *local eigenvalue* of a strongly regular graph s an eigenvalue of B_1 or B_2 not equal to θ or τ.

20.2.1 Lemma. *If X is a triangle-free strongly regular graph with parameters $(v,k;0,c)$, then the eigenvalues of B_2 are its valency $k-c$ and a subset of θ, τ and $-c$.* □

It follows that the second neighborhood of a vertex in a triangle-free strongly regular graph is walk regular.

20.3 Dimensions

We determine the dimension of the Terwilliger algebra of a strongly regular graph. Most the work has already been done, the main task left is to determine the isomorphism classes of the irreducible $\mathbb{T}$-modules.

We deal with the easy cases first. The standard module is the only module with dimension three, and so nothing more need to be said. Two 1-dimensional modules U and V are isomorphic if either both F_1U and F_1V are non-zero, or both F_2U and F_2V are non-zero.

So consider the 2-dimensional $\mathbb{T}$-module spanned by the vectors

$$\begin{pmatrix} 0 \\ w \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ Nw \end{pmatrix}$$

where $B_1 w = \lambda w$. Relative to this basis, F_1 and F_2 are represented by the matrices

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

We have

$$A \begin{pmatrix} 0 \\ w \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda w \\ Nw \end{pmatrix}$$

and

$$A \begin{pmatrix} 0 \\ 0 \\ Nw \end{pmatrix} = \begin{pmatrix} 0 \\ NN^T Nw \\ B_2 NW \end{pmatrix} = \begin{pmatrix} 0 \\ -(\lambda-\theta)(\lambda-\tau)w \\ (\delta-\lambda)Nw \end{pmatrix}.$$

and therefore the matrix representing A is

$$\begin{pmatrix} \lambda & -(\lambda-\theta)(\lambda-\tau) \\ 1 & \delta-\lambda \end{pmatrix}.$$

It follows that this module is determined by the eigenvalue λ. We also see that the structure of the Terwilliger algebra is determined by the eigenvalues of X and the eigenvalues of its first and second neighborhoods.

20.3.1 Theorem. *Let X be a strongly regular graph with v vertices. Let s denote the sum of the number of eigenvalues of B_1 in $\{\theta,\tau\}$ plus and the number of eigenvalues of B_2 in $\{\theta,\tau\}$, and let r denote the number of eigenvalues of B_1 not in this set. Then* $\dim(\mathbb{T}) = 9+4r+s$ *and* $\dim(Z(\mathbb{T})) = 1+r+s$. □

It is clear that $0 \le s \le 4$ but in fact s must be positive if X is not a conference graph. For

$$k + \ell = m_\theta + m_\tau$$

and if $k > \ell$ and $m_\theta > m_\tau$, then A has an eigenvector with eigenvalue θ supported on the neighborhood of a vertex. Similar conclusions hold if we reverse one or both of these inequalities.

By way of example we consider the strongly regular graph $L(K_n)$, where $n \ge 5$. Its parameters are

$$\left(\binom{n}{2}, 2n-4; n-2, 4\right)$$

and its eigenvalues are

$$2n-4,\ n-4,\ -2.$$

The neighborhood of a vertex is isomorphic to $K_2 \square K_{n-2}$, whence its eigenvalues are

$$n-2,\ n-4,\ 0,\ -2$$

and the second neighborhood is isomorphic to $L(K_{n-2})$, with eigenvalues

$$2n-8,\ n-6,\ -2.$$

Thus there are two non-local eigenvalues in the neighborhood, one non-local eigenvalue in the second neighborhood and one local eigenvalue. It follows that the Terwilliger algebra has dimension

$$9+3+4=16$$

while its centre has dimension five.

Chapter 21

Hamming Schemes

21.1 The Binary Hamming Scheme

The Hamming scheme $H(d,2)$ is a metric and cometric association scheme. The matrix $A = A_1$ is the adjacency matrix of the d-cube, and its eigenvalues are the integers

$$d-2i, \quad i = 0,\ldots,d$$

with respective multiplicities

$$\binom{d}{i}.$$

The automorphism group of the Hamming scheme is vertex-transitive, and so the Terwilliger algebra is the same for each vertex.

We can write

$$A = R + L$$

where $L = R^T$ and R is the natural raising operator on the lattice of subsets of $\{1,\ldots,d\}$. (So L is the natural lowering operator.)

21.1.1 Theorem. *The Terwilliger algebra of the binary Hamming scheme is a quotient of the enveloping algebra $U(sl(2,\mathbb{C}))$.*

Proof. View the vertices of the Hamming scheme as subsets of $\{1,\ldots,d\}$. Define

$$H = RL - LR.$$

We note that

$$R_{\alpha,\beta} = 1$$

if and only if $\alpha \subseteq \beta$ and $|\beta| = |\alpha| + 1$. Further $H_{\alpha,\beta} = 0$ if $|\alpha| \neq |\beta|$ and, if $|\alpha| = |\beta| = i$, then

$$H_{\alpha,\beta} = d - 2i.$$

It follows that

$$H = \sum_{i=0}^{d} (d-2i)F_i$$

and hence the the algebra of all polynomials in H is the equal to the algebra generated by the diagonal matrices F_i.

Since

$$[R,L] = H, \quad [H,R] = 2R, \quad [H,L] = -2L$$

the algebra generated by R, L and H is a homomorphic image of $U(sl(2,\mathbb{C}))$.

To complete the proof we must show that R and L generate the Terwilliger algebra of $H(n,d)$. But since the scheme is metric, each element of the Bose-Mesner algebra is a polynomial in A and since the algebra generated by H contains each F_i, we conclude that R and L generate the Terwilliger algebra. □

21.2 Modules

With what we know about the representation theory of $sl(2,\mathbb{C})$, it is easy to determine the irreducible $\mathbb{T}$-modules for the binary Hamming scheme $H(d,2)$. If u is a vertex of $H(d,2)$ with Hamming weight i, then the vectors

$$v, Rv, \ldots, R^{d-2i}v$$

are a basis for an irreducible module of dimension $d-2i+1$. If u and v are binary vectors then the irreducible modules they generate are isomorphic if and only if u and v have the same Hamming weight.

21.2.1 Lemma. *We have*

$$\dim(\mathbb{T}(H(d,2))) = \frac{1}{6}(d+1)(d+2)(d+3).$$

Proof. If $0 \leq 2i \leq d$, then our Terwilliger algebra has one isomorphism class of irreducible module with dimension $d-2i+1$, whence

$$\dim(\mathbb{T}(H(d,2))) = \sum_{i \leq d/2} (d-2i+1)^2 = \frac{1}{6}(d+1)(d+2)(d+3). \qquad \square$$

21.2.2 Lemma. *The Terwilliger algebra of the Hamming scheme is thin and dual thin.*

Proof. If v has Hamming weight i, then the Hamming weight of each vector in $\mathrm{supp}(R^j v)$ is $i+j$. Hence $F_{i+j}R^j v = R_j v$, and therefore the R-module generated by v is thin. Since the Hamming schemes are metric, it follows from Lemma 19.3.2 that this module is also dual thin. □

Chapter 22

Spin

22.1 Braids

The *braid group on n strands* B_n is the group generated by elements

$$\sigma_1, \ldots, \sigma_{n-1}$$

subject to the relations:

$$\sigma_i \sigma_j = \sigma_j \sigma_i, \text{ if } |i-j| \geq 2$$
$$\sigma_i \sigma_j \sigma_i = \sigma_j \sigma_i \sigma_j, \text{ if } |i-j| = 1.$$

[braids, closure]

The map that takes σ_i to the transposition $(i\ i+1)$ in the symmetric group $\mathrm{Sym}(n)$ extends to a homomorphism from B_n. (Its kernel consists of the *pure braids*.)

The Temperley-Lieb algebra $TL_n(\beta)$ contains a homomorphic image of the Braid group.

22.2 Nomura Algebras

Let A and B be $v \times v$ matrices and suppose B is Schur invertible. The *Nomura algebra* $\mathcal{N}_{A,B}$ consists of all $v \times v$ matrices for which all the vectors

$$Ae_i \circ Be_j^{(-)}$$

are eigenvectors. If $M \in \mathcal{N}_{A,B}$, we define $\Theta_{A,B}(M)$ to be the $v \times v$ matrix with ij-entry equal to eigenvalue of M associated to $Ae_i \circ Be_j$. Thus $I \in \mathcal{N}_{A,B}$ and $\Theta_{A,B} = J$.

If $M, N \in \mathcal{N}_{A,B}$, then

$$\Theta_{A,B}(MN) = \Theta_{A,B}(M) \circ \Theta_{A,B}(N).$$

Thus $\mathcal{N}_{A,B}$ is an algebra under matrix multiplication, and the image of $\mathcal{N}_{A,B}$ under $\Theta_{A,B}$ is an algebra under Schur multiplication. We note that

$$\mathcal{N}_{A,B} = \mathcal{N}_{B,A}$$

while

$$\Theta_{B,A}(M) = \Theta_{A,B}(M)^T.$$

22.2.1 Lemma. *If A is invertible and B is Schur invertible, then $\Theta_{A,B}$ is injective and $\mathcal{N}_{A,B}$ is a commutative algebra.*

A $v \times v$ matrix W is a *type-II matrix* if it is Schur invertible and

$$WW^{(-)T} = vI.$$

Hadamard matrices provide one class of examples. If W is a type-II matrix, then W is invertible and

$$W^{-1} = \frac{1}{v} W^{(-)T}.$$

22.2.2 Lemma. *The matrix W is a type-II matrix if and only if $J \in \mathcal{N}_{W,W^{(-)}}$.*

The Nomura algebra $\mathcal{N}_{W,W^{(-)}}$ will play an important role in our work and so we will denote it by $\mathcal{N}_W$. We also write Θ_W for $\Theta_{W,W^{(-)}}$.

22.3 Braids

Let A, B and C be $v \times v$ matrices. We define endomorphisms X_A, Δ_B and Y_C of the vector space $\mathrm{Mat}_{n\times n}(\mathbb{C})$ by

$$X_A(M) := AM, \qquad \Delta_B(M) := B \circ M, \qquad Y_C(M) := MC^T.$$

(We could instead use respectively $A \otimes I$, D_B and $I \otimes C$, where D_B is a diagonal matrix with the entries of B as its diagonal entries and all three matrices are viewed as elements of $\mathrm{End}(\mathrm{Mat}_{v\times v}(\mathbb{C}))$.)

22.3.1 Lemma. *Suppose $A, B \in \mathrm{Mat}_{\nu\times\nu}(\mathbb{C})$. Then $R \in \mathscr{N}_{A,B}$ and $\Theta_{A,B}(R) = S$ if and only*

$$X_R \Delta_B X_A = \Delta_B X_A \Delta_S.$$

We see that $A \in \mathscr{N}_{A,B}$ and $\Theta_{A,B}(A) = B$ if and only if

$$X_A \Delta_B X_A = \Delta_B X_A \Delta_B;$$

we call this the *braid relation*. [If A is invertible and B is Schur invertible and $A \in \mathscr{N}_{A,B}$, does it follow that $\Theta_{A,B}(A) = B$?]

We note the following result, which we call the *exchange identity*.

22.3.2 Theorem. *Let A, B, C, Q, R, S be $\nu \times \nu$ matrices. Then*

$$X_A \Delta_B X_C = \Delta_Q X_R \Delta_S$$

if and only if

$$X_A \Delta_C X_B = \Delta_R X_Q \Delta_{S^T}.$$

Proof. Apply each of the four products to the matrix $e_i e_j^T$. □

The bilinear form $\mathrm{tr}(MN^T)$ on $\mathrm{Mat}_{\nu\times\nu}(\mathbb{C})$ is non-degenerate and hence allows to define the adjoint of elements of $\mathrm{End}(\mathrm{Mat}_{\nu\times\nu}(\mathbb{C}))$. We denote the adjoint by transpose and observe that

$$(X_A)^T = X_{A^T}, \quad (\Delta_B)^T) = \Delta_B.$$

Thus the braid relation implies that

$$X_{A^T} \Delta_B X_{A^T} = \Delta_B X_{A^T} \Delta_B.$$

22.4 Jones Pairs

We say that $\nu \times \nu$ matrices A and B form a *one-sided Jones pair* if A is invertible, B is Schur invertible and $A \in \mathscr{N}_{A,B}$. They form a *Jones pair* if (A, B^T) is also a one-sided Jones pair.

22.4.1 Lemma. *If (A, B) is a one-sided Jones pair, so are each of the following:*

(a) (A^T, B).

(b) $(A^{-1}, B^{(-)})$.

(c) $(D^{-1}AD, B)$, *where* D *is diagonal and invertible.*

(d) (A, BP), *where* P *is a permutation matrix.*

(e) $(\lambda A, \lambda B)$, *for any non-zero complex number* λ.

22.4.2 Lemma. *The matrices* A *and* B *form a one-sided Jones pair if and only if for all* i *and* j *we have*

$$A(Ae_i \circ Be_j) = B_{i,j}(Ae_i \circ Be_j).$$

22.4.3 Corollary. *Let* (A, B) *be a pair of* $v \times v$ *matrices and let* D_j *be the diagonal matrix formed from the* j*-th column of* B. *Then* (A, B) *is a one-side Jones pair if and only if, for* $j = 1, \ldots, v$,

$$AD_jA = D_jAD_j.$$

22.4.4 Lemma. *If* (A, B) *is a one-sided Jones pair, then each column of* B *sums to* $\mathrm{tr}(A)$.

Proof. From the previous result we have

$$A^{-1}D_jA = D_jAD_j^{-1}$$

whence A and D_j are similar and $\mathrm{tr}(A) = \mathrm{tr}(D_j)$. Therefore each column of B sums to $\mathrm{tr}(A)$. □

We say a Jones pair (A, B) is *invertible* if A is Schur invertible and B is invertible.

22.4.5 Theorem. *Suppose* (A, B) *is a one-sided Jones pair and* B *is invertible, then* A *and* B *are type-II matrices and the diagonal of* A *is constant.*

Proof. If $A \in \mathcal{N}_{A,B}$ then $A^{-1} \in \mathcal{N}_{A,B}$ and so

$$\Theta_{B,A}(A^{-1}) = B^{(-)T}.$$

This implies that

$$X_{A^{-1}}\Delta_A X_B = \Delta_A X_B \Delta_{B^{(-)T}} \tag{22.4.1}$$

and taking the transpose of this, we get

$$X_{B^T}\Delta_A X_{A^{-T}} = \Delta_{B^{(-)T}} X_{B^T}\Delta_A.$$

If we apply the right side to I we get $B^T(A\circ A^{-T})$, if we apply the left side to I the result is

$$B^{(-)T}\circ(B^T(A\circ I)) = J(A\circ I)$$

and hence

$$B^T(A\circ A^{-T}) = J(A\circ I).$$

Since B is invertible and its row sums are all equal t some constant β, this implies that

$$A\circ A^{-T} = B^{-T}J(A\circ I) = \beta J(A\circ I).$$

The sum of the entries in the i-th column of $A\circ A^{-T}$ is

$$\sum_r (A^{-1})_{r,i}(A^T)_{r,i} = \sum_r (A^{-1})_{r,i}A_{i,r} = 1$$

and therefore all columns of $J(A\circ I)$ must be equal. It follows that $vA\circ A^{-T} = J$ and so A is a type-II matrix with constant diagonal.

To complete the proof we multiply each side of (22.4.1) on the left by $\Delta_{A^{(-)}}$ and on the right by $X_{B^{-1}}$ to obtain

$$\Delta_{A^{(-)}}X_{A^{-1}}\Delta_A = X_B\Delta_{B^{(-)T}}X_{B^{-1}}.$$

Taking inverses on both sides yields

$$\Delta_{A^{(-)}}X_A\Delta_A = X_B\Delta_{B^T}X_{B^{-1}}$$

and applying each side to I gives

$$A^{(-)}\circ(A(A\circ I)) = B(B^T\circ B^{-1}).$$

Since the diagonal of A is constant, the left side here is equal to aJ for some a and so

$$B^T\circ B^{-1} = aB^{-1}J$$

Arguing as before, the sum of a row of $B^T\circ B^{-1}$ is 1. Therefore $B^{-1}J$ is a multiple of J; from this we see that B is a type-II matrix. □

22.4.6 Lemma. *If (A,B) is Jones pair and A is Schur invertible, then B is invertible.*

Proof. Apply both sides of (22.4.1) to J; this yields

$$A^{-1}(A\circ(BJ)) = A\circ(BB^{(-)T}).$$

Since (A,B^T) is a Jones pair the row sums of B equal $\mathrm{tr}(A)$ and so the left side here is equal to $\mathrm{tr}(A)I$. As A is Schur invertible it follows that $BB^{(-)T}$ is diagonal. However the diagonal entries of $BB^{(-)T}$ are all equal and so it is a scalar matrix. We conclude that B is type II and invertible. □

22.5 Gauge Equivalence

If D is an invertible diagonal matrix we say that $D^{-1}JD$ is a *dual permutation matrix*. The Schur inverse of a dual permutation matrix is a dual permutation matrix.

22.5.1 Lemma. *If A, C and M are Schur invertible and $X_A\Delta_M = \Delta_M X_C$, then $C^{(-)}\circ A$ is a dual permutation matrix. If B, C and M are invertible and $\Delta_B X_M = X_M\Delta_C$, then CB^{-1} is a permutation matrix.*

22.5.2 Corollary. *If (A,B) and (C,B) are one-sided Jones pairs, then $C = D^{-1}AD$ where D is invertible and diagonal.*

22.5.3 Corollary. *If (A,B) and (A,C) are one-sided Jones pairs, then $C = BP$ where P is a permutation matrix.*

22.6 Nomura Algebras of Type-II matrices

A type-II matrix W is called a *spin model* if $(W,W^{(-)})$ is a Jones pair. If $W\in\mathcal{N}_W$, then $(W,W^{(-)})$ need not be a Jones pair, because the columns of $W^{(-)}$ might not sum to $\mathrm{tr}(A)$. If σ denotes the sum of a column of $W^{(-)}$ and we choose γ so that

$$\gamma^2\,\mathrm{tr}(W) = \sigma$$

then γW is a spin model.

22.6.1 Theorem. *Let A be a $\nu \times \nu$ type-II matrix. Then Θ_A is a bijection from $\mathcal{N}_A$ to $\mathcal{N}_{A^T}$ and Θ_{A^T} is a bijection from $\mathcal{N}_{A^T}$ to $\mathcal{N}_A$. If $R \in \mathcal{N}_A$ then $\Theta_{A^T}(\Theta_A(R)) = \nu R^T$.*

Proof. Suppose $R \in \mathcal{N}_A$ and $\Theta_A(R) = S$. Then

$$X_R \Delta_{A^{(-)}} X_A = \Delta_{A^{(-)}} X_A \Delta_S$$

and the transpose of this is

$$X_{A^T} \Delta_{A^{(-)}} X_{R^T} = \Delta_S X_{A^T} \Delta_{A^{(-)}}$$

and applying the exchange identity to this yields

$$X_{A^T} \Delta_{R^T} X_{A^{(-)}} = \Delta_{A^T} X_S \Delta_{A^{(-)T}}.$$

If we multiply both sides of this on the left by $\Delta_{A^{(-)T}}$ and on the right by $X_{A^{(-)-1}}$ we get

$$X_S \Delta_{A^{(-)T}} X_{A^{(-)-1}} = \Delta_{A^{(-)T}} X_{A^T} \Delta_{R^T}.$$

Since $A^{(-)-1} = \frac{1}{\nu} A^T$, this yields

$$X_S \Delta_{A^{(-)T}} X_{A^T} = \Delta_{A^{(-)T}} X_{A^T} \Delta_{\nu R^T}$$

whence $S \in \mathcal{N}_{A^T}$ and $\Theta_{A^T}(S) = \nu R^T$.

As $\Theta_{A^T}(\Theta_A(R)) = \nu R^T$, we see that Θ_A and Θ_{A^T} are bijections. □

This proof shows that the composite map

$$\frac{1}{\nu} \Theta_{A^T} \Theta_A$$

is the transpose map on $\mathcal{N}_A$. Hence $\frac{1}{\nu} \Theta_A \Theta_{A^T}$ is the transpose map on $\mathcal{N}_{A^T}$. In fact Θ_A and Θ_{A^T} commute with the transpose.

22.6.2 Corollary. *If A is a type-II matrix and $R \in \mathcal{N}_A$, then $R^T \in \mathcal{N}_A$ and $\Theta_A(R^T) = \Theta_A(R)^T$.*

Proof. If $R \in \mathcal{N}_A$ then $\nu R^T = \Theta_{A^T}(\Theta_A(R) \in \mathcal{N}_A$ and

$$\Theta_A(\nu R^T) = \Theta_A(\Theta_{A^T}(\Theta_A(R))) = \nu \Theta_A(R)^T.$$ □

22.6.3 Corollary. *If A is a $v \times v$ type-II matrix and $M, N \in \mathcal{N}_A$, then*

$$\Theta_A(M \circ N) = \frac{1}{v}\Theta_A(M)\Theta(N).$$

22.6.4 Corollary. *If A is a type-II matrix then its Nomura algebra is closed under matrix multiplication, Schur multiplication, transpose and complex conjugation.*

Proof. We know that $\mathcal{N}_A$ is closed under matrix multiplication and that

$$\Theta_A(MN) = \Theta_A(M) \circ \Theta_A(N),$$

from which it follows that the image of Θ_A is Schur-closed. Therefore $\mathcal{N}_{A^T}$ is Schur-closed. Swapping A and A^T, we deduce that $\mathcal{N}_A$ is Schur-closed.

We saw above that $\mathcal{N}_A$ is closed under transpose. Since it is Schur-closed it has a basis consisting of 01-matrices, and the complex span of these matrices is closed under complex conjugation. □

This corollary asserts that $\mathcal{N}_A$ is the Bose-Mesner algebra of an association scheme.

22.7 Spin Models

By definition, W is a spin model if $(W, W^{(-)})$ is a one-sided Jones pair.

22.7.1 Lemma. *If A is a type-II matrix and $(A, A^{(-)})$ is a one-sided Jones pair, then it is a Jones pair.*

Proof. Since $(A, A^{(-)})$ is a one-sided Jones pair, we have

$$X_A \Delta_{A^{(-)}} X_A = \Delta_{A^{(-)}} X_A \Delta_{A^{(-)}}$$

and taking the transpose of this yields

$$X_{A^T} \Delta_{A^{(-)}} X_{A^T} = \Delta_{A^{(-)}} X_{A^T} \Delta_{A^{(-)}}.$$

Using the exchange identity we obtain

$$X_{A^T} \Delta_{A^T} X_{A^{(-)}} = \Delta_{A^T} X_{A^{(-)}} \Delta_{A^{(-)T}}$$

and inverting both sides yields

$$X_{A^{(-)-1}}\Delta_{A^{(-)T}}X_{A^{-T}} = \Delta_{A^T}X_{A^{(-)-1}}\Delta_{A^{(-)T}}.$$

If we multiply on the left by $\Delta_{A^{(-)T}}$ and on the right by X_{A^T}, the result is

$$\Delta_{A^{(-)T}}X_{A^{(-)-1}}\Delta_{A^{(-)T}} = X_{A^{(-)-1}}\Delta_{A^{(-)T}}X_{A^T}.$$

We observe that $A^{(-)-1} = \frac{1}{v}A^T$, whence the last equation yields

$$\Delta_{A^{(-)T}}X_{A^T}\Delta_{A^{(-)T}} = X_{A^T}\Delta_{A^{(-)T}}X_{A^T}$$

and therefore $(A^T, A^{(-)T})$ is a one-sided Jones pair. From the transpose of this we see that $(A, A^{(-)T})$ is one-sided Jones pair, and thus it follows that $(A, A^{(-)})$ is a Jones pair. □

22.7.2 Theorem. *If A is spin model, then $\mathcal{N}_A = \mathcal{N}_{A^T}$ and $\Theta_A = \Theta_{A^T}$.*

Proof. We use gauge equivalence. If $(A, A^{(-)})$ and $A, A^{(-)T}$ are one-sided Jones pairs, there is a permutation matrix P such that $A^{(-)T} = A^{(-)}P$, and consequently

$$\mathcal{N}_{A,A^{(-)T}} = \mathcal{N}_{A,A^{(-)}P} = \mathcal{N}_{A,A^{(-)}}$$

Now $A \in \mathcal{N}_A$ if and only if

$$A^T \in \mathcal{N}_{A^T,A^{(-)}} = \mathcal{N}_{A^T,A^{(-)T}}.$$

Since $\mathcal{N}_A$ is closed under transposes, the result holds.

Suppose $R \in \mathcal{N}_A$ and $\Theta_A(R) = S$ then

$$X_R\Delta_{A^{(-)}}X_A = \Delta_{A^{(-)}}X_A\Delta_S$$

and if $R \in \mathcal{N}_{A^T}$ and $\Theta_{A^T} = T$ then

$$X_R\Delta_{A^{(-)T}}X_{A^T} = \Delta_{A^{(-)T}}X_{A^T}\Delta_T.$$

Consequently

$$(\Delta_{A^{(-)T}}X_{A^T}\Delta_T)^{-1}\Delta_{A^{(-)}}X_A\Delta_S = X_{A^{-T}}\Delta_{A^T\circ A^{(-)}}X_A.$$

The left side here equals

$$\Delta_{T^{(-)}}X_{A^{-T}}\Delta_{A^T}\Delta_{A^{(-)}}X_A\Delta_S = \Delta_{T^{(-)}}(X_{A^{-T}}\Delta_{A^T\circ A^{(-)}}X_A)\Delta_S$$

If we define

$$\Xi := X_{A^{-T}} \Delta_{A^T \circ A^{(-)}} X_A$$

then

$$\Xi \Delta_S = \Delta_T \Xi. \tag{22.7.1}$$

We compute $\Xi(M)$, for any $v \times v$ matrix M. Note that

$$\Xi(M) = A^{-T}(A^T \circ A^{-1} \circ (AM))$$

Since $(A, A^{(-)})$ and $(A^T, A^{(-)T})$ are both one-sided Jones pairs, there is an invertible diagonal matrix C such that $A^T = C^{-1}AC$. Therefore

$$A^T \circ A^{-1} = (C^{-1}AC) \circ A^{-1} = C^{-1}JC$$

and so

$$A^T \circ A^{-1} \circ (AM) = (C^{-1}JC) \circ (AM) = C^{-1}AMC = A^T C^{-1}MC$$

and consequently

$$\Xi(M) = A^{-T}(A^T \circ A^{-1} \circ (AM)) = C^{-1}MC.$$

Now apply each side of (22.7.1) to M; we get

$$T \circ (C^{-1}MC) = C^{-1}(S \circ M)C = S\circ)C^{-1}MC).$$

We conclude that $S = T$. □

Chapter 23

Abelian Spin

We study spin models in the Bose-Mesner algebras of abelian groups, aka translation schemes.

23.1 Schemes

Suppose W is a type-II matrix. Its Nomura algebra is the Bose-Mesner algebra of an association scheme $\mathscr{A}$; thus $\mathscr{N}_W = \mathbb{C}[\mathscr{A}]$. We assume $\mathscr{A}$ has d classes and therefore $\dim(\mathscr{N}_W) = d+1$. Let $E_0, \ldots, E_d$ denote the principal idempotents of the scheme. Since $E_i^2 = E_i$ we see that $\Theta_W(E_i)$ must be a Schur idempotent in $\mathscr{N}_{W^T}$ and since

$$\sum_{i=0}^{d} E_i = I$$

we have

$$\sum_{i=0}^{d} \Theta_W(E_i) = J,$$

whence it follows that the Schur idempotents $\Theta_W(E_i)$ are linearly independent. Thus they are the principal Schur idempotents of the scheme determined by W^T.

A similar argument shows that if $A_0, \ldots, A_d$ are the principal Schur idempotents of the scheme determined by W, then the matrices

$$\frac{1}{v}\Theta_W(A_i)$$

are the principal idempotents of the scheme determined by W^T.

Now we specialize to the case where W is a spin model. The eigenvalues $p_i(j)$ of the scheme are defined implicitly by

$$A_i = \sum_{i=0}^{d} p_i(j) E_j.$$

Since $\mathcal{N}_W = \mathcal{N}_{W^T}$, we see that Θ_W maps $\mathbb{C}[\mathscr{A}]$ to itself and the matrix which represents relative to the basis formed by $E_0, \ldots, E_d$ has ij-entry equal to $p_j(i)$. It is traditionally denoted by P. It is invertible and we follow tradition further and set Q equal to vP^{-1}. We have

$$E_j = \frac{1}{v} \sum_{i=0}^{d} q_j(i) A_i$$

where $q_j(i) = Q_{j,i}$. Since

$$\Theta_{W^T}(A_i) = vE_i^T = v\overline{E}_i$$

we see that $\overline{Q}$ is the matrix that represents Θ_{W^T} relative to the basis $A_0, \ldots, A_d$. Therefore $P^{-1}\overline{Q}P$ is the matrix representing Θ_{W^T} relative to $E_0, \ldots, E_d$. Since $\mathcal{N}_{W^T} = \mathcal{N}_W$, the operators Θ_W and Θ_{W^T} commute. Therefore

$$\Theta_W^{-1} \Theta_{W^T} \Theta_W = \Theta_{W^T}$$

and therefore $P^{-1}\overline{Q}P = \overline{Q}$.

23.1.1 Theorem. *If W is a spin model then $P = \overline{Q}$.*

Proof. We have $\Theta_W = \Theta_{W^T}$. □

23.2 Coordinate Matrices

Let W be a spin model, let $\mathscr{A}$ denote the associated set of principal Schur idempotents and let $\mathscr{E}$ be the set of principal matrix idempotents. Since $W \in \mathcal{N}_W$, there are scalars $\lambda_0, \ldots, \lambda_d$ (the eigenvalues of W) such that

$$W = \sum_{i=0}^{d} \lambda_i E_i.$$

Let L be the diagonal matrix with $L_{i,i} = \lambda_i$. If M is an endomorphism of a vector space with basis β we use $[M]_\beta$ to denote the matrix representing M relative to β. Let T represent the transpose map.

23.2.1 Lemma. *We have* $[X_W]_{\mathscr{E}} = L$ *and* $[\Delta_W]_{\mathscr{A}} = L^{-1}$.

Proof. The first claim is immediate from the definition of the coefficients w_i. For the the second note that

$$\lambda_i A_i = \Theta_W(WE_i) = \Theta_W(W) \circ A_i = W^{(-)} \circ A_i$$

and consequently

$$W^{(-)} \circ A_i = \lambda_i A_i$$

as required. □

The last equation above implies that

$$W = \sum_{i=0}^{d} \lambda_i^{-1} A_i.$$

Therefore

$$\lambda_r E_r = WE_r = \sum_{i=0}^{d} \lambda_i^{-1} A_i E_r = \Big(\sum_{i=0}^{d} \lambda_i^{-1} p_i(r)\Big) E_r$$

and hence

$$L\mathbf{1} = PL^{-1}\mathbf{1}.$$

We see that diagonal matrices L such that $L\mathbf{1} = P^{-1}\mathbf{1}$ correspond to the type-II matrices in $\mathbb{C}[\mathscr{A}]$.

23.2.2 Lemma. *We have* $[X_W]_{\mathscr{A}} = PLP^{-1}$ *and* $[\Delta_W]_{\mathscr{E}} = P^{-1}L^{-1}P$.

Proof. The matrix of eigenvalues P represents the change-of-basis map from $\mathscr{E}$ to $\mathscr{A}$ and therefore

$$[\Delta_W]_{\mathscr{E}} = P^{-1}L^{-1}P.$$

The second claim follows similarly. □

The transpose map is an endomorphism of $\mathrm{Mat}_{v\times v}(\mathbb{C})$ that maps $\mathscr{N}_W$ to itself, which we denote by T. Both $[T]_{\mathscr{E}}$ and $[T]_{\mathscr{A}}$ are permutation matrices. Since

$$\Theta(M^T) = \Theta(M)^T,$$

we see that T and Θ commute, and consequently $[T]_{\mathscr{E}} = [T]_{\mathscr{A}}$. Note that

$$\Theta^2 = vT. \tag{23.2.1}$$

23.3 Duality

Suppose W is a spin model and $\mathscr{A}$ is the corresponding association scheme. Then Θ_W maps $\mathbb{C}[\mathscr{A}]$ to itself and swaps matrix with Schur multiplication. (If $\mathscr{A}$ is the scheme of the cyclic group, then Θ_W is better known as the discrete Fourier transform.) We will derive various expressions for Θ_W.

We begin with the observation that $W^T \in \mathscr{N}_W$ and $\Theta_W(W^T) = W^{(-)T}$. Hence

$$X_{W^T}\Delta_{W^{(-)}}X_W = \Delta_{W^{(-)}}X_W\Delta_{W^{(-)T}} \tag{23.3.1}$$

Denote either side of this identity by Λ and note that $\Lambda^T = \Lambda$.

23.3.1 Theorem. *If $R \in \mathscr{N}_W$ and $\Theta_W(R) = S$, then $\Lambda^{-1}X_R\Lambda = \Delta_S$ and $\Lambda^{-1}\Delta_{S^T}\Lambda = X_R$.*

Proof. Since R commutes with W^T, we have

$$\begin{aligned}\Lambda^{-1}X_R\Lambda &= X_{W^{-1}}\Delta_W X_{W^{-T}}X_R X_{W^T}\Delta_{W^{(-)}}X_W\\ &= X_{W^{-1}}\Delta_W X_R\Delta_{W^{(-)}}X_W\\ &= \Delta_S.\end{aligned}$$

Next

$$\begin{aligned}\Lambda^{-1}\Delta_{S^T}\Lambda &= \Delta_{W^T}X_{W^{-1}}\Delta_W\Delta_{S^T}\Delta_{W^{(-)}}X_W\Delta_{W^{(-)T}}\\ &= \Delta_{W^T}X_{W^{-1}}\Delta_{S^T}X_W\Delta_{W^{(-)T}}.\end{aligned}$$

As $R \in \mathscr{N}_{W^T}$ and $\Theta_W = \Theta_{W^T}$, we have

$$X_R\Delta_{W^{-1}}X_{W^T} = \Delta_{W^{-1}}X_{W^T}\Delta_S$$

and therefore by the exchange identity

$$X_R\Delta_{W^T}X_{W^{-1}} = \Delta_{W^T}X_{W^{-1}}\Delta_{S^T}.$$

It follows that

$$\Lambda^{-1}\Delta_{S^T}\Lambda = X_R\Delta_{W^T}X_{W^{-1}}X_W\Delta_{W^{(-)T}} = X_R.$$

23.3.2 Corollary. *If $R \in \mathscr{N}_W$, then $\Lambda^{-2}X_R\Lambda^2 = X_{R^T}$ and $\Lambda^{-2}\Delta_S\Lambda^2 = \Delta_{S^T}$.*

Proof. First

$$\Delta_S = (\Delta_S)^T = (\Lambda^{-1} X_R \Lambda)^T = \Lambda X_{R^T} \Lambda^{-1}$$

and therefore

$$\Lambda^{-2} X_R \Lambda^2 = \Lambda^{-1} \Delta_S \Lambda = X_{R^T}.$$

Second

$$\Lambda^{-2} \Delta_{S^T} \Lambda^2 = \Lambda^{-1} X_R \Lambda = \Delta_S$$

and so taking transposes we get

$$\Lambda^2 \Delta_{S_T} \Lambda^{-2} = \Delta_S,$$

which yields our second claim. □

You may also show that $\Lambda^{-2} X_{R^T} \Lambda^2 = X_R$ and $\Lambda^{-2} \Delta_{S^T} \Lambda^2 = \Delta_S$, from which we get the following.

23.3.3 Corollary. *If W is a spin model, then the map Λ^4 commutes with X_R and Δ_S for all R, S in $\mathcal{N}_W$.* □

It is worth noting that we are dealing here with several algebras. First we have $\mathrm{Mat}_{v\times v}(\mathbb{C})$ and its subalgebra $\mathcal{N}_W = \mathbb{C}[\mathcal{A}]$, which is the Bose-Mesner algebra of $\mathcal{A}$. Then inside the algebra $\mathrm{End}(\mathrm{Mat}_{v\times v}(\mathbb{C}))$ we have the algebra generated by all operators

$$X_R,\ \Delta_R, \qquad R \in \mathcal{N}_W = \mathcal{N}_{W^T}.$$

This may be viewed as an extended version of the Terwilliger algebra of $\mathcal{A}$ (and is not commutative). We will call it the *Jaeger algebra* and denote it by $\mathcal{J}_3$. The map given by conjugation by Λ is an endomorphism of $\mathrm{Mat}_{v\times v}(\mathbb{C})$ which fixes $\mathcal{J}_3$.

Suppose (A, B) is a 1-sided Jones pair and define

$$K := X_A \Delta_B X_A = \Delta_B X_A \Delta_B.$$

Then

$$K^2 = (X_A \Delta_B X_A)^2 = (X_A \Delta_B)^3$$

from which it follows that Λ commutes with X_A and Δ_B. It is easy to verify that

$$K^{-1} X_A K = \Delta_B$$

and consequently that

$$K^{-1}\Delta_B K = X_A.$$

Specializing to the case where A is type II and $B = A^{(-)}$, we see that conjugation by K is a linear map that fixes the algebra generated by X_A and $\Delta_{A^{(-)}}$. This is, in general, a subalgebra of the Jaeger algebra.

23.4 Modular Invariance

We combine the work from the previous two sections.

23.4.1 Lemma. *If W is a spin model with $\delta = W_{i,i}$ and $M \in \mathcal{N}_W$, then $\Theta(M) = v\delta\Lambda^{-1}(M)$.*

Proof. If $R \in \mathcal{N}_W$ and $\Theta(R) = S$, then $\Lambda^{-1}X_R\Lambda = \Delta_S$ and therefore

$$S = \Delta_S(J) = \Lambda^{-1}X_R\Lambda(J).$$

Here

$$\Lambda(J) = \Delta_{W^{(-)}}X_W\Delta_{W^{(-)T}}(J) = \Delta_{W^{(-)}}WW^{(-)T} = vW^{(-)} \circ I.$$

If $\delta := W_{i,i}$, then it follows that

$$S = \Lambda^{-1}X_R(\delta v I) = \delta v\Lambda^{-1}(R). \qquad \square$$

The next result is known as the *modular invariance property* for the pair (P,L).

23.4.2 Theorem. *Let W be a spin model with association scheme $\mathscr{A}$ and $\delta = W_{i,i}$, let P be the matrix of eigenvalues of $\mathscr{A}$ and let L be the diagonal matrix of eigenvalues of W. Then $(PL)^3 = \frac{v}{\delta}I$.*

Proof. From the previous lemma,

$$\Theta = v\delta\Lambda^{-1}v\delta X_{W^T}\Delta_{W^{(-)}}X_W.$$

If F is the matrix that represents transpose, then we have

$$[X_W]_{\mathscr{E}} = L, \quad [\Delta_{W^{(-)}}]_{\mathscr{E}} = P^{-1}LP, \quad [X_{W^T}]_{\mathscr{E}} = FLF.$$

Since $[\Theta]_{\mathscr{E}} = P$, we have

$$P = v\delta FLFP^{-1}LPL$$

and so

$$I = v\delta(P^{-1}F)L(FP^{-1})LPL.$$

As $\Theta^2 = vT$ we have $P^2 = vF$ and so

$$P^{-1}F = FP^{-1}\frac{1}{v}P$$

from which we infer that

$$I = v\delta v^{-2}(PL)^3.$$

23.5 The Cyclic Group

The index set of the matrices in this section start at 0.

Choose the complex number θ so that θ^2 is a primitive v-th root of unity. (Thus is v is odd then θ has order v, but if v is even then its order is $2v$.) Define the $v \times v$ matrix W by

$$W_{i,j} = \theta^{(i-j)^2}.$$

Let S denote the diagonal matrix with $S_{i,i} = \theta^{i^2}$. Then

$$\theta^{(i-j)^2} = \theta^{i^2}\theta^{-2ij}\theta^{j^2}$$

and therefore if V is the Vandermonde matrix given by

$$V_{i,j} = (\theta^{-2})^{ij}$$

then

$$W = SVS.$$

23.5.1 Theorem. *The matrix W is a spin model.*

Proof. First

$$(WW^{(-)T})_{i,j} = \sum_r \theta^{(i-r)^2}\theta^{-(j-r)^2} = \theta^{i^2-j^2}\sum_r \theta^{2(j-i)r}.$$

The last sum equal v if $i = j$ and is otherwise zero. Therefore $WW^{(-)T} = vI$ and so W is a type-II matrix.

The matrix W lies in the Bose-Mesner algebra of the cyclic group of order v, equivalently it is a circulant. If ζ_a is the column vector with i-entry θ^{2ai} then ζ_a is an eigenvector for all circulants of order $v \times v$. Now

$$(We_r \circ We_s^{(-)T})_i = \theta^{(i-r)^2}\theta^{-(i-s)^2} = \theta^{r^2-s^2}\theta^{-2i(r-s)},$$

whence

$$We_r \circ We_s^{(-)T} = \theta^{r^2-s^2}\zeta_{r-s}.$$

This implies that $W \in \mathcal{N}_W$, and therefore we conclude that W is a spin model. □

Bibliography

Index