A Graph-Theoretic Approach to Bounds for Error-Correcting Codes[1]

CIMPA-UNESCO-PHILIPPINES Summer School on Semidefinite Programming in Algebraic Combinatorics

W. J. Martin
Department of Mathematical Sciences
and
Department of Computer Science
Worcester Polytechnic Institute
Worcester, Massachussetts
martin@wpi.edu

July 20, 2009

[1] ©William J. Martin, 2009.

Abstract

In these notes, we address bounds for error-correcting codes. Our approach is from the viewpoint of algebraic graph theory. We therefore begin with a review of the algebraic structure of the Hamming graph, focusing on the binary case. We then derive Delsarte's linear programming bound and explore some applications of it.

In the second part of the notes, we introduce Terwilliger's subconstituent algebra and explore its connection to the semidefinite programming approach of Schrijver. Throughout, our focus is on the binary case. The three steps presented here are a summary of the structure of the Terwilliger algebra as presented by Go, a surprising connection to the biweight enumerator of a binary code, and a full characterization of the positive semidefinite cone of the algebra, given by Visentin and Martin.

Contents

Chapter 1

Introduction and Overview

The theory of error-correcting codes is a gem of twentieth-century mathematics. Motivated by the pressing need for efficient digital communications, mathematicians and engineers developed a body of tools, examples and techniques that was at once elegant and practical. Seemingly archaic facts from mathematics such as the theory of finite fields and invariant theory were dusted off, rejuvinated in their crucial roles in this important new field. Connections to finite group actions, the statistical design of experiments and the theory of lattices enriched the field and facilitated various sorts of "technology transfer". Meanwhile, the applications were often ahead of the theory, first with spacecraft communications, then the compact disc technology, packet networks, and even hard drive storage schemes. New applications continue to emerge, from quantum computing, to cellphone technology to cryptographic protocols. Indeed, mathematicians have identified a fundamental phenomenon in science revolving around the fundamental concept of distinguishability among selected elements in a metric space and the dual concept of approximating a space by a well-chosen representative subset. Now these concepts are extended well beyond their original context (q-ary strings of fixed length n) to include all sorts of spaces, both finite and infinite, with or without a metric, but typically enjoying a rich group action.

From the very beginning of the development, researchers were concerned with efficiency. The codes of Golay and Hamming found in the 1940s were "perfect" but were they "efficient"? By the 1960s, algebraic coding theorists had arrived at the "main question of algebraic coding theory" [31, p222]: What is the maximum size $A(n, d)$ of a binary code C of given length n and given minimum distance $0 < d < n$? Early bounds for this quantity were found to be quite weak. The 1973 thesis of Philippe Delsarte (summarizing several papers, some published as early as 1968) introduced his famous linear programming (LP) bound for q-ary error-correcting codes. In the words of McEliece, this “deceptively powerful” result unified several known bounds and became part of a framework for many future results and extensions. (While Delsarte’s bound applies to many other problems and to all commutative association schemes, our discussion will consider only the Hamming graphs.)

Further developments include the famous MRRW bound (or “bound of the four Americans” as the Russians once called it), powerful and elegant extensions to bounds for spherical codes and designs, powerful bounds of Vladimir Levenshtein, and asymptotic versions of some of the most important bounds.

1.1 Samorodnitsky’s Theorem

As we said, the linear programming bound can be to obtain bounds on vertex subsets in any commutative association scheme. Yet certainly the most studied case is that of the Hamming graphs. In addition to substantial concrete data, the study of the linear programming technique has led to a number of less obvious results. Substantial work has been completed on the asymptotic analysis of dual-feasible solutions to these linear programs. See [33, 25, 8].

To delve a bit further into this issue, let us focus on the binary case. We will use the binary entropy function

$$H_2(x) = -x\log_2 x - (1-x)\log_2(1-x).$$

The rate of a binary code C of length n is $R := \frac{1}{n}\log_2|C|$. Cleverly applying the LP bound in 1977, McEliece et al. proved [31, p222] that any binary code C satisfies the “MRRW bound”:

$$R \leq H_2\left(\frac{1}{2} - \sqrt{\frac{d}{n}\left(1-\frac{d}{n}\right)}\right)$$

where d denotes the minimum distance of C. This gives rise to an asymptotic upper bound on the rate of a binary code. On the other hand, Gilbert [15] and Varshamov [43] proved that, for any $0 \leq \delta < \frac{1}{2}$, there exists a sequence of binary linear codes each with $\frac{d}{n} \geq \delta$ and rate $R \geq 1 - H_2(\frac{d}{n})$.

Coding theorists have been investigating the gap between the MRRW bound and the Gilbert-Varshamov bound for two decades now. Could it be that codes exist with asymptotic rate close to the MRRW bound? Or is it true that the optimal rates are instead closer to the GV bound?[1] A recent result of Samorodnitsky establishes that the linear programming bound cannot resolve this issue. Specifically, he proves in [35] that the optimum of Delsarte's linear program is at least the average of these two bounds. (See also Barg and Jaffe [7].) Many interpret this as evidence that much more efficient codes exist than currently known. But I feel that the true answer may be much closer to the GV bound and that the linear programming bound is simply not strong enough to provide the necessary non-existence results.

Substantial improvements to the LP bound appear in [19], where Jaffe describes software for automating the linear programming bound in an intelligent manner. For instance, if one learns that a desired code C of length n must contain a word of weight k, then one may assume a given support for this codeword and split the original LP into two pieces — one LP for a code of length k containing the all-ones vector and one LP for a code of length $n-k$. This promising line of investigation has answered a number of open questions regarding specific codes [20] but may not be powerful enough to provide asymptotic improvements beyond the known bounds.

The linear programming technique has also been applied to bound the probability of undetected error [1], and to bound quantum error-correcting codes [23, 34], among other problems in coding theory. The linear programming bound for the Johnson scheme was used by Frankl and Wilson in [13] to obtain an extension of the Erdös-Ko-Rado Theorem for intersecting set systems to the setting of vector spaces. This wide applicability motivates us to seek out improvements to the bound.

1.2 The Terwilliger algebra as an extension of the Bose-Mesner algebra

The Bose-Mesner algebra of a d-class association scheme has dimension $d+1$ and the corresponding linear program (LP) has $d+1$ variables and constraints. We suggest that one way to obtain stronger bounding techniques is

[1] For non-binary alphabets, the story may be different. In [41], Tsfasman et al. show that for square-order fields, $\mathbb{F}_q$, $q \geq 49$, codes exceeding the GV bound do indeed exist.

to first identify useful super-algebras of the Bose-Mesner algebra $\mathbb{A}$. Suppose $\mathcal{U}$ is a semi-simple matrix algebra containing $\mathbb{A}$. Suppose $\mathcal{U}$ admits two bases

$$\{B_i : i \in \mathcal{I}\} \qquad \text{and} \qquad \{F_j : j \in \mathcal{J}\}.$$

For a vector $\mathbf{u}$ in the standard module (for example, the characteristic vector of a subset C of X), we consider the two ordered sets of statistics

$$[\mathbf{u}^* B_i \mathbf{u} : i \in \mathcal{I}]$$

and

$$[\mathbf{u}^* F_j \mathbf{u} : j \in \mathcal{J}] \,.$$

In analogy with the linear programming bound of Delsarte, we seek bases $\{B_i\}$ and $\{F_j\}$ with the properties that each complex number $\mathbf{u}^* B_i \mathbf{u}$ has combinatorial interpretation (say, when $\mathbf{u}$ is assumed to be a 01-vector) and the numbers $\mathbf{u}^* F_j \mathbf{u}$ satisfy some simple algebraic conditions. If one hopes to have $\mathbf{u}^* F_j \mathbf{u} \geq 0$, one approach is to rather choose each F_j to be the projection of the standard module onto the sum of all irreducible $\mathcal{U}$-modules of a given isomorphism type; i.e., a central idempotent.

One candidate algebra which has received considerable attention in the recent literature is the subconstituent algebra of Terwilliger [37, 38, 39] (this is now often called the **Terwilliger algebra**). While the full structure of this algebra is currently unclear for P- and Q-polynomial schemes in general, the cases of most practical interest — namely, the Hamming and Johnson schemes — have Terwilliger algebras which are quite well understood.

Let C be a binary code of length n. For a four-tuple $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ of non-negative integers summing to n, let

$$\ell_\alpha = |\{(c, c') \in C \times C : \\ \mathrm{wt}(c) = \alpha_2 + \alpha_3, \mathrm{dist}(c, c') = \alpha_1 + \alpha_3, \mathrm{wt}(c') = \alpha_1 + \alpha_2\}| \,.$$

We consider the biweight enumerator

$$\mathcal{W}_C(y_0, y_1, y_2, y_3) = \sum_{\alpha=(\boldsymbol{\alpha_0},\boldsymbol{\alpha_1},\boldsymbol{\alpha_2},\boldsymbol{\alpha_3})} \ell_\alpha \; y_0^{\alpha_0} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3}.$$

For binary linear codes, MacWilliams identities for this enumerator were given by MacWilliams, et al. in 1972 [30]. We give a new proof of these identities using the Terwilliger algebra of the n-cube. Further, we consider some families of non-linear codes whose dual biweight enumerators necssarily have non-negative coefficients. For unrestricted codes, a characterization of the positive semidefinite cone of the Terwilliger algebra is given which leads to new inequalities for the coefficients of the biweight enumerator of an unrestricted code. Each extremal ray of the positive semidefinite cone corresponds to a projection onto some irreducible module of the S_n action on the free real vector space over the binary n-tuples. It remains to find computationally useful formulae for these inequalities.

1.3 Notation

In these notes, I try to use certain symbols in restricted ways, but I'm sure that some of the rules I am about to lay out are violated repeatedly.

I try to use bold font lower case letters for vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$ and upper case roman letters for matrices. Indices are usually chosen from $\{g, h, i, j, k, \ell\}$. Lower case letters are used for codewords and n-tuples over finite alphabets, with a preference for a, b, c over x, y, z which are used as indeterminates.

I use dist for Hamming distance (whereas Christine Bachoc uses d). I use $\langle \cdot, \cdot \rangle$ for the standard Hermitean inner product on $\mathbb{C}^n$ (or $\mathbb{R}^n$) but I use $a \cdot b$ for the dot product of two vectors over a finite field or ring.

When indexing relations and eigenspaces of an association scheme (such as the n-cube), I try to use h, i for relations (distances) and j, k for eigenspaces, but only when both types of object appear in the same expression.

Chapter 2

The Bose-Mesner Algebra

In this chapter, we consider the Hamming association schemes and their Bose-Mesner algebras. We begin with a fundamental motivating problem in coding theory; this leads naturally to the definition of the Hamming metric and the Hamming graphs. We then work out two fundamental bases for the adjacency algebra (or “Bose-Mesner algebra”) of the Hamming graph and the change-of-basis matrices from one of these bases to the other.

2.1 Block codes

Suppose $\mathcal{Q}$ is an alphabet of size q and we wish to communicate messages over a noisy channel by sending sequences of symbols from $\mathcal{Q}$. In these notes, we will assume that all messages are broken into blocks of a fixed length n. So we are choosing messages from the set $\mathcal{Q}^n$. We assume that, with small probability, any of the symbols in a given message $x \in \mathcal{Q}^n$ may be distorted by noise in the channel, resulting in a somewhat different message $r \in \mathcal{Q}^n$ being read at the receiving end.

The nature of this noise depends on the application (e.g., fingerprints or manufacturing defects in the case of the CD, crosstalk or electromagnetic interference in the case of cellular communications), but we model this using the *q-ary symmetric channel.* We assume that symbols within the message are independently subject to error and that, for some small $\epsilon > 0$, a symbol $\alpha \in \mathcal{Q}$ is received error-free with probability $1-\epsilon$ and is replaced with another symbol β with probability $\epsilon/(q-1)$ for each $\beta \neq \alpha$ in $\mathcal{Q}$[1].

[1] A more realistic channel model, in the binary case, is one in which the probability of a

Under this model[2], a message $x \in \mathcal{Q}^n$ sent over the channel is most likely received as x, but second-most likely received as some q-ary n-tuple that agrees with x in $n-1$ coordinate positions. By the same reasoning, we see that, upon transmission of a fixed n-tuple x, the probability that some other n-tuple y is received depends only on the number of coordinate positions in which x and y agree.

Now for fixed n and q, we define the *Hamming metric* on $\mathcal{Q}^n$. For

$$x = (x_1, x_2, \ldots, x_n), \qquad y = (y_1, y_2, \ldots, y_n),$$

we define the *Hamming distance* from x to y as the number of coordinate positions in which they differ:

$$\mathrm{dist}(x,y) = \left|\{i : 1 \le i \le n,\ x_i \neq y_i\}\right| .$$

We usually assume that our alphabet $\mathcal{Q}$ contains a zero element and we define the *Hamming weight* of x as the distance from x to the all-zero vector:

$$\mathrm{wt}(x) = \left|\{i : 1 \le i \le n,\ x_i \neq 0\}\right| .$$

Exercises:

1. dist is a metric on $\mathcal{Q}^n$:
   (a) $\mathrm{dist}(x,y) = \mathrm{dist}(y,x)$ for all x and y in $\mathcal{Q}^n$;
   (b) $\mathrm{dist}(x,y) \ge 0$ for all x and y in $\mathcal{Q}^n$;
   (c) $\mathrm{dist}(x,y) = 0$ if and only if $x = y$;
   (d) $\mathrm{dist}(x,y) \le \mathrm{dist}(x,z) + \mathrm{dist}(z,y)$ for all x, y and z in $\mathcal{Q}^n$.
2. For any $x \in \mathcal{Q}^n$, there are exactly $\binom{n}{i}(q-1)^i$ elements of $\mathcal{Q}^n$ at distance i from x.
3. If $\mathcal{Q}$ is an abelian group, then $\mathrm{dist}(x,y) = \mathrm{wt}(x-y)$ for all $x, y \in Q^n$.

one getting replaced by a zero is much higher than the probability that a zero is replaced by a one. This applies when bit positions containing ones are represented by higher-energy signals than those containing zeros.

[2] Technically, this is true only for ϵ small and n small. As an exercise, the reader may wish to compute that value $n = n(\epsilon)$ at which the sent codeword is received error-free with less than 50% probability.

2.2 The Hypercube

Before considering general Hamming graphs, we take a look at the simplest case, which happens to be the most important case for several applications.

Let $n > 0$ be fixed and consider $X = \mathbb{Z}_2^n$. We do not necessarily need the structure of an abelian group for the graph theory or (combinatorial) coding theory that we will be doing, but this group structure will play a fundamental role in the spectral analysis we develop here.

The n-cube is the simple undirected graph $Q_n = (X, R_1)$ having vertex set X and adjacency relation

$$R_1 = \{(x,y) \in X \times X : \mathrm{dist}(x,y) = 1\}.$$

Note that we are saying that the vertex set $\underline{\text{is}}$ the set of zero-one strings of length n, not simply that the vertices are labelled by this set.

Exercise: Now draw the n-cube for $1 \le n \le 4$ and be sure to include the names of the vertices.

Assuming that $X = \mathbb{Z}_2^n$, it is easy to see that Q_n is a *vertex transitive graph*: for any two vertices a and b, there is an automorphism of Q_n (i.e., a permutation of the vertices which maps adjacent pairs to adjacent pairs) which maps a to b. For given $a, b \in \mathbb{Z}_2^n$, the required automorphism is simply translation inside the group X by the element $b \oplus a$ (exercise). (Note that X forms an abelian group under coordinatewise sum $\oplus$ modulo two.)

The n-cube is a bipartite distance-regular graph. It has *diameter* n: the maximum distance between any two vertices is n. For example, the (binary) all-ones vector is the unique vertex at distance n from the (binary) zero vector of length n. In fact, there are exactly $\binom{n}{i}$ vertices of Q_n at distance i from any given $x \in X$. Next, consider vertices $x, y \in X$ which are distance i apart in Q_n. The reader may verify that, among the n neighbors of y, i of these are at distance $i-1$ from x and the other $n-i$ neighbors are at distance $i+1$ from x. This implies that Q_n is a *distance-regular* graph[3]. More generally, if $\mathrm{dist}(x,y) = k$, then one checks (exercise!) that the number of elements of X simultaneously at distance i from x and distance j from y is

$$p_{ij}^k = \binom{k}{\ell}\binom{n-k}{i-\ell} \qquad (\ell := (i+k-j)/2). \tag{2.1}$$

[3] A regular graph Γ of diameter n is distance-regular if there exist integers a_i, b_i and c_i $(0 \le i \le n)$ such that, for any two vertices x and y, there are c_i (resp., a_i, b_i) neighbors of y at distance $i-1$ (resp., i, $i+1$) from x where $i = \mathrm{dist}(x,y)$ in Γ.

(To prove this, appeal to the vertex transitivity of Q_n so that you may assume x is the zero vector.)

We will consider not only the n-cube Q_n, but also the distance-i graph of this graph for each $0 \leq i \leq n$. The adjacency matrix of the distance-i graph is the 01-matrix A_i of order 2^n with rows and columns indexed by X and a one in row a column b when $\text{dist}(a,b) = i$ and a zero in this position otherwise. The $n+1$ matrices $A_0, A_1, \ldots, A_n$ form a basis for a vector space $\mathbb{A}$ of symmetric matrices called the *Bose-Mesner algebra* of the n-cube. Indeed, what we showed in the previous paragraph implies that

$$A_i A_j = \sum_{k=0}^{n} p_{ij}^k A_k \tag{2.2}$$

for $0 \leq i, j \leq n$.

Exercises:

1. We find generators for the *automorphism group* of the n-cube Q_n:

   (a) Prove that, for any $c \in \mathbb{Z}_2^n$, the permutation $\phi_c : X \to X$ which sends a to $a + c$ preserves adjacency (i.e., for any $a, b \in X$, a is adjacent to b if and only $a + c$ is adjacent to $b + c$;

   (b) Prove that, for any $\tau \in S_n$ (the symmetric group on n letters), the permutation $\hat{\tau} : X \to X$ which sends $a = (a_1, \ldots, a_n)$ to $(a_{\tau(1)}, \ldots, a_{\tau(n)}$ preserves adjacency.

2. Prove Equation (2.1).

3. Prove Equation (2.2).

4. Show that A_0 is the identity matrix.

5. Show that $A_0 + A_1 + \cdots + A_n = J$, the all-ones matrix.

6. Show that A_i has constant row sum $\binom{n}{i}$.

7. Prove that $A_1 A_i = (n - i + 1)A_{i-1} + (i + 1)A_{i+1}$ and use this, together with a simple induction argument, to prove that A_i is expressible as a polynomial of degree i in A_1.

8. Prove that the vector space $\mathbb{A}$ is closed under entrywise multiplication of matrices: if $M = [m_{ij}]$ and $N = [n_{ij}]$, then $M \circ N$ is the matrix with (i, j)-entry $m_{ij} n_{ij}$.

9. Prove that each A_i and the sum of any subset of $\{A_0, \ldots, A_n\}$ is an idempotent under $\circ$.

Let G be a finite group with identity element e and let $S \subseteq G - \{e\}$ be an inverse-closed subset of G: for all $g \in G$, $g \in S$ if and only if $g^{-1} \in S$. The *Cayley graph* $\Gamma(G; S)$ is the simple undirected graph having vertex set G and having a adjacent to b precisely when $ba^{-1} \in S$. Cayley graphs are obviously vertex transitive graphs and there is a strong connection between their eigenvectors and the characters of the group G. In particular, when G is abelian, every irreducible character of G is an eigenvector for $\Gamma(G; S)$ and these form a basis for $\mathbb{C}^G$. What we work out in the next paragraph is a very special case of this general phenomenon.

A *linear character* of a group X is a homomorphism $\chi : X \to \mathbb{C}^*$ from X to the multiplicative group of non-zero complex numbers. For example, if $X = \mathbb{Z}_2$, then the maps $\chi_0 : b \mapsto 1$ and $\chi_1 : b \mapsto (-1)^b$ are the only two such homomorphisms. If $X = \mathbb{Z}_2^n$, then there are 2^n linear characters, one for each $a \in X$. Define $\chi_a : X \to \mathbb{C}^*$ by

$$\chi_a(b) = (-1)^{a \cdot b}$$

for $b \in X$. In fact, these characters form a group isomorphic to X under multiplication of functions

$$(\chi_a \chi_c)(b) = (-1)^{a \cdot b} (-1)^{c \cdot b} = (-1)^{(a \oplus c) \cdot b};$$

so we have $\chi_a \chi_c = \chi_{a \oplus c}$ and this group, which we call $X^\dagger$ is isomorphic to X.

Henceforth, we view each character χ_a as a vector in $\mathbb{C}^X$. It is not hard to show that, for any $a \in X$ with Hamming weight j and for any $0 \le i \le n$,

$$A_i \chi_a = P_{ji} \chi_a$$

where

$$P_{ji} = \sum_{\ell=0}^{i} (-1)^\ell \binom{j}{\ell} \binom{n-j}{i-\ell}. \tag{2.3}$$

(Exercise: Please try to prove this by summing $\chi_a(b)$ over the n-tuples b at distance i from a fixed $c \in X$ and using the fact that each such b is uniquely expressible in the form $b = c \oplus e$ for some $e \in X$ of Hamming weight i.)

We now let S_j denote the $2^n \times \binom{n}{j}$ matrix with columns χ_a as a ranges over the binary n-tuples of Hamming weight j. With appropriate orderings on rows and columns, S_j has (b,a)-entry $\chi_a(b) = (-1)^{a \cdot b}$. It is straightforward to check that each column of any S_j has norm $\sqrt{2^n}$, but it is also true that all these column vectors are pairwise orthogonal. Indeed, if $a \neq c$ then there is some coordinate h where $a_h \neq c_h$. So we have

$$\begin{aligned}
\langle \chi_a, \chi_c \rangle &= \sum_{b \in X} \chi_a(b) \chi_c(b) \\
&= \sum_{b \in X} (-1)^{(a \oplus c) \cdot b} \\
&= \sum_{b \in X} (-1)^{\sum_{i=1}^n (a_i + c_i) b_i} \\
&= \prod_{i=1}^n \left[(-1)^{(a_i+c_i)0} + (-1)^{(a_i+c_i)1} \right] \\
&= \prod_{i=1}^n \left[1 + (-1)^{a_i+c_i} \right] \\
&= 0
\end{aligned}$$

since the h term in the product, where $a_i + c_i = 1$, is equal to zero. So $S_j^\top S_j = 2^n I$ and $S_j^\top S_k = 0$ when $j \neq k$ for $0 \leq j, k \leq n$.

Now consider the matrix

$$E_j = \frac{1}{2^n} S_j S_j^\top .$$

Since MN and NM have the same eigenvalues (excepting some possible extra zero eigenvalues for the larger one) when they are square matrices, we immediately see that E_j has $\binom{n}{j}$ eigenvalues equal to one and $2^n - \binom{n}{j}$ zero eigenvalues. Moreover, E_j is positive semidefinite since, for any vector $\mathbf{v}$ in $\mathbb{R}^X$,

$$\mathbf{v}^\top E_j \mathbf{v} = \frac{1}{2^n} \left(S_j^\top \mathbf{v} \right)^\top \left(S_j^\top \mathbf{v} \right) \geq 0.$$

So E_j is a matrix representing orthogonal projection onto the column space of S_j, which we denote by V_j.

Since each column of S_j is an eigenvector of A_i with eigenvalue P_{ji}, we see that $A_i S_j = P_{ji} S_j$ and $A_i E_j = P_{ji} E_j$ as well. This shows that each S_j is a common eigenspace for the matrices A_i, $0 \leq i \leq n$.

Now we compute all of the entries of E_j and show that it belongs to the Bose-Mesner algebra $\mathbb{A}$. (This also follows from the spectral decomposition of A_1.) If $\operatorname{dist}(a,c) = i$, then

$$\begin{aligned} (E_j)_{a,c} &= \frac{1}{2^n} \sum_{\mathrm{wt}(b)=j} \chi_b(a)\chi_b(c) \\ &= \frac{1}{2^n} \sum_{\mathrm{wt}(b)=j} (-1)^{b\cdot a}(-1)^{b\cdot c} \\ &= \frac{1}{2^n} \sum_{\mathrm{wt}(b)=j} (-1)^{b\cdot(a\oplus c)} \end{aligned}$$

Now if S is an i-subset of $[n] = \{1, \ldots, n\}$, then there are $\binom{i}{\ell}\binom{n-i}{j-\ell}$ j-subsets T of $[n]$ satisfying $|S \cap T| = \ell$. With S representing the support of $a \oplus c$ and T representing the support of b, we then find

$$\begin{aligned} (E_j)_{a,c} &= \frac{1}{2^n} \sum_{\mathrm{wt}(b)=j} (-1)^{b\cdot(a\oplus c)} \\ &= \frac{1}{2^n} \sum_{\ell=0^i} (-1)^{\ell} \binom{i}{\ell}\binom{n-i}{j-\ell} \end{aligned}$$

which depends only on i and j and not on the choice of a and c themselves. This proves that

$$E_j = \frac{1}{2^n} \sum_{i=0}^{n} P_{ij} A_i$$

where P_{ij} is defined in Equation (2.3) above. In particular, each E_j lies in the Bose-Mesner algebra of Q_n.

Since $S_j^{\top} S_k = 0$ for $j \neq k$, we find

$$E_j E_k = \delta_{j,k} E_j \qquad (0 \leq j, k \leq n)$$

and these $n+1$ matrices are therefore linearly independent inside $\mathbb{A}$ and hence form a basis. The interplay between the two bases $\{A_i\}_{i=0}^n$ and $\{E_j\}_{j=0}^n$ will play a fundamental role in our study.

In the more general setting of an n-class commutative association scheme on a vertex set X, we assume the existence of two such bases $\{A_i\}_{i=0}^n$ and $\{E_j\}_{j=0}^n$ and define the $(n+1)\times(n+1)$ change-of-basis matrices P and Q (often called the first and second eigenmatrices of the association scheme, respectively) defined by

$$A_i = \sum_{j=0}^{n} P_{ji} E_j, \qquad E_j = \frac{1}{|X|} \sum_{j=0}^{n} Q_{ij} A_i. \tag{2.4}$$

In the case of the n-cube we have shown that

$$P_{ij} = Q_{ij} = \sum_{\ell=0}^{i} (-1)^{\ell} \binom{j}{\ell} \binom{n-j}{i-\ell}.$$

The fact that $P = Q$ for the n-cube is rather special. An association scheme is called *formally self-dual* when its two eigenmatrices are equal. But in this case, we have an actually duality coming from the group $X^{\dagger}$ of characters. In short, these characters also determine an association scheme in a natural way and in the case of the n-cube (and the Hamming graphs in general) the two association schemes are in fact isomorphic. So the n-cube is a truly "self-dual" association scheme.

2.3 Hamming Graphs

In this section, we show that everything we said about the n-cube applies more generally to the Hamming graph $H(n,q)$ defined on the n-tuples over any finite alphabet $\mathcal{Q}$ of size $q \geq 2$.

2.4 Eigenspaces and Algebra Bases

Let ω denote a primitive complex q^{th} root of unity. Assume that $\mathcal{Q} = \mathbb{Z}_q = \{0, 1, \ldots, q-1\}$ and let us use the inner product

$$a \cdot b = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \pmod{q}$$

with the occasional abuse of notation where we consider $a \cdot b$ to be an integer. If $\oplus$ denotes addition modulo q of n-tuples in $\mathcal{Q}^n$, then we observe that

$$a \cdot (b \oplus c) = a \cdot b + a \cdot c$$

in $\mathbb{Z}$.

Throughout, we will work with the abelian group $X := \mathcal{Q}^n$.

For $a \in X$, consider the q^n-tuple of complex numbers χ_a with entries

$$\chi_a(b) = \omega^{a \cdot b}$$

for $b \in X$. Viewed as a function on $X = \mathcal{Q}^n$, each χ_a is an irreducible complex character and every character for this abelian group is a non-negative integer linear combination of these fundamental ones. But viewed as a function on vertices, each χ_a is an eigenvector of the Hamming graph $H(n,q)$ and these q^n eigenvectors span $\mathbb{C}^X$, allowing us to diagonalize the Bose-Mesner algebra of the Hamming scheme.

Recall that A_i is the $q^n \times q^n$ matrix with a one in row b, column c precisely when $\text{dist}(b,c) = i$ and a zero in that position otherwise. We compute the b-entry of the vector $A_i \chi_a$. Suppose $\text{wt}(a) = j$. Then

$$\begin{aligned}
(A_i \chi_a)_b &= \sum_{\text{dist}(b,c)=i} \omega^{a \cdot c} \\
&= \sum_{\text{wt}(e)=i} \omega^{a \cdot (b \oplus e)} \\
&= \sum_{\text{wt}(e)=i} \omega^{a \cdot b + a \cdot e} \\
&= \omega^{a \cdot b} \sum_{\text{wt}(e)=i} \omega^{a \cdot e}.
\end{aligned}$$

In order to simplify this last expression, we consider those coordinates in the support of e that lie within the support of a and those which do not. If the supports of a and e have ℓ coordinates in common, then we may fix values for e outside these coordinates (there are $(q-1)^{i-\ell}$ such choices) and sum over all possible nonzero values within those ℓ coordinates to find that these tuples e together contribute $(-1)^\ell$ to the overall sum. There are $\binom{j}{\ell}$ ways to choose the ℓ coordinates within the support of a and $\binom{n-j}{i-\ell}$ ways to choose the remaining coordinates of e. Sorting all this out, one finds

$$(A_i \chi_a)_b = \omega^{a \cdot b} \sum_{\ell=0}^{i} (-1)^\ell (q-1)^{i-\ell} \binom{j}{\ell} \binom{n-j}{i-\ell}.$$

It turns out that the change-of-basis coefficients Q_{ij} satisfying

$$E_j = \frac{1}{q^n} \sum_{i=0}^{n} Q_{ij} A_i$$

are exactly the same as the ones which express the A_i in terms of the matrices E_j:

$$Q_{ij} = P_{ij} \qquad (0 \leq i, j \leq n).$$

This means that the Hamming graph $H(n, q)$ is a *formally self-dual association scheme*: it satisfies $P = Q$. But what's more is that there is an explicit duality here. The characters $X^\dagger = \{\chi_a : a \in X\}$ also for an association scheme: for $0 \leq h \leq n$, join χ_a to χ_b in relation R_h if and only if $\chi_a \chi_{-b}$ lies in eigenspace V_h. It is not hard to show that this happens precisely when $\mathrm{wt}(a \ominus b) = h$. Obviously, this is isomorphic to the Hamming scheme, but it is defined on a vertex set which is an eigenbasis for the Hamming graph. So the *duality map* ψ which maps $a \in \mathbb{Z}_2^n$ to the character χ_a in the "dual group" $(\mathbb{Z}_2^n)^\dagger$ is an isomorphism. When such an explicit isomorphism exists, we say that the association scheme is *self-dual.* (It is quite possible to have a dual scheme on $X^\dagger$ which is not isomorphic to the one one X.)

Before we leave this section, let me just point out a curious graph theoretic property of the n-cube which will arise later. If (X, R) is a simple graph and $a, b, c \in X$, we say a vertex $e \in X$ is an *apex* for a, b, c if e lies simultaneously on some shortest path in the graph from a to b, some shortest path from b to c and some shortest path from c to a. A graph is an "apex graph" if, for every three vertices of it, these vertices have an apex.

Exercises:

1. In any graph (X, R), if one of a, b, c lies on a shortest path joining the other two, then a, b, c have an apex.

2. Every apex graph is triangle free.

3. The n-cube is an apex graph. For $a = 0^n$, $b, c \in \mathbb{Z}_2^n$, explicitly find the apex for a, b, c and prove that it is unique.

Chapter 3

The Linear Programming Bound of Delsarte

Philippe Delsarte was a coding theorist at Philips MBLE Labs in Brussels in the late 1960s. He was authoring papers as early as 1968 and his 1973 dissertation, under the supervision of Vladimir Belevitch, summarized several of these papers and ushered in a new era in algebraic coding theory. In addition to the celebrated linear programming bound that we will discuss here, Delsarte's dissertation established a broad framework for coding theory and design theory, the theory of association schemes. In particular, it was Delsarte who initiated the study of P- and Q-polynomial association schemes.

3.1 Some Simple Codes

We will be deriving bounds on the sizes of various codes. In order to make sense of these bounds, it will help to have a few small examples of codes to work with.

For us, a *code* is simply any subset C of $\mathcal{Q}^n$. The tuples in C are called "codewords". To avoid trivialities, we assume $1 < |C| < q^n$. The *minimum distance* of C is

$$\delta(C) = \min \{ d(x,y) : \ x, y \in C, \ x \neq y \} .$$

It is not hard to see that, using a code with minimum distance δ, one can correct up to $e := \lfloor \frac{\delta}{2} \rfloor$ errors per transmitted codeword. We define the *packing radius* $e(C)$ of C as $\lfloor \frac{\delta}{2} \rfloor$. If C is a q-ary code of length n and size M whose minimum distance is d, then we say C is an $(n, M, d)_q$-code.

It is useful in coding theory to know how far the received vector is from the code (in the Hamming metric). We define

$$d(r, C) = \min\{d(r, c) : c \in C\}.$$

The *covering radius* of C is the maximum of this value over all possible r:

$$\rho(C) = \max\{d(r, C) : r \in \mathcal{Q}^n\}.$$

Exercises:

1. Show that $2e \leq \delta$.
2. Show that $e \leq \rho$.
3. What can you say about the structure of a code in which $e = \rho$?

For any n and q, the *q-ary repetition code* of length n is the code

$$C = \{000\cdots 0,\ 1111\cdots 1,\ 222\cdots 2, \ldots\}$$

of size q which contains one codeword of length n for each symbol $\alpha \in \mathcal{Q}$. This code has minimum distance n; that is, any two distinct codewords differ in <u>all</u> coordinates. So we have $\rho(C) = \lfloor \frac{n}{2} \rfloor$ and $e(C) = \lfloor \frac{n-1}{2} \rfloor$.

For $q = 2$, we obtain the *binary repetition code*, a code of size two. For example, when $n = 5$, the repetition code is $C = \{00000, 11111\}$.

When we treat $\mathcal{Q}$ as a finite field, we can view $X = \mathcal{Q}^n$ as a vector space over $\mathcal{Q}$ and subspaces of this vector space sometimes form particularly useful error-correcting codes. A subspace C of $\mathcal{Q}^n$ of dimension k in which every non-zero codeword has Hamming weight at least d is denoted an $[n, k, d]_q$-code, the square brackets indicated that C is a subspace, a q-ary *linear code* of length n, dimension k and minimum distance d. (Exercise: Show that a linear code C has minimum distance d or larger if and only if the minimum Hamming weight of any non-zero codeword in C is d or larger.) If C is a (linear) $[n, k, d]_q$-code, then its *dual code*

$$C^\perp = \{y \in X : \forall x \in C (x \cdot y = 0)\}$$

is an $[n, n-k, d]_q$-code. The relationship between these two codes — which can be generalized to any subgroup C of $\mathcal{Q}^n$ when $\mathcal{Q}$ takes on the structure of an abelian group — is a fundamental paradigm for our study of the linear programming bound, which applies even when C has no group structure at all.

The dual of the binary repetition code is the linear code consisting of all binary words having even Hamming weight. This code has dimension $n-1$, minimum distance two and covering radius one.

A "single-error-correcting code" is a code with minimum distance three or larger. (We typically use this term only when the minimum distance is either three or four.) If we seek a binary single-error-correcting code of length five, the best we can do is to choose a code of size four[1]:

$$C = \{00000,\ 11100,\ 00111,\ 11011\}\,.$$

A *perfect code* of packing radius e (or "perfect e-code") is a q-ary code of length n with minimum distance $2e+1$ and covering radius e. For such a code, every vector r in X is at distance e or less from exactly one codeword. In the binary case, there is a unique linear perfect 1-code of length n for each length of the form $n = 2^m - 1$ (i.e., the binary Hamming code) and no such perfect 1-code for other lengths n. Strangely, there are a large number of unruly non-linear $(2^m-1, 2^{2^m-m-1}, 3)_2$-codes which seem to be beyond classification. Other than these, there is only one other non-trivial binary perfect code, the Golay code. There is also a perfect ternary Golay code, a $[11, 6, 5]_3$-code as well as q-ary Hamming codes of length $n = (q^m-1)/(q-1)$ for each prime power q and each $m \geq 2$. We'll next briefly review these.

Let $\mathbb{F}_q$ denote the unique finite field of order q. For $m \geq 2$, let N be the $m \times q^m - 1$ matrix having each non-zero vector in $\mathbb{F}_q^m$ as one of its columns. When $q > 2$, this matrix has many pairs of linearly dependent columns. If we define an equivalence relation on the columns of N by $\mathbf{u} \approx \mathbf{v}$ if and only if $\mathbf{u} = \alpha\mathbf{v}$ for some $\alpha in \mathbb{F}_q$, then we retain one representative from each equivalence class and obtain the submatrix H, an $m \times n$ matrix ($n = (q^m-1)/(q-1)$) of rank m in which no two columns are linearly independent. Therefore the null space C of H is an $[n, n-m, 3]_q$-code, a perfect code. To see this, note that the number of tuples in X at distance

[1] While we are not emphasizing applications here, it may be useful to know that this code enables us to encode two information bits in a codeword of length five, thereby attaining an information rate of $2/5 = 0.4$.

one from C is $n(q-1)|C|$ and since

$$|C| = q^{q^m-m-1} = q^n q^{-(m+1)} = \frac{q^n}{1+n(q-1)},$$

we see that C has covering radius one and is therefore perfect.

One way to construct the perfect binary Golay code is to first use the icosahedron to construct the extended binary Golay code. If A is the 12×12 adjacency matrix of the icosahedron and J is the 12×12 all-ones matrix, then $\mathcal{G}_{24}$ is the binary row space of the matrix $G = [I_12|J - A]$. We obtain the famous perfect code $\mathcal{G}_{23}$ by choosing any column of G and deleting it before computing the row space. This is a perfect $[23, 12, 7]_2$-code.

The perfect ternary Golay code is a linear $[11, 6, 5]_3$-code and can be constructed as the mod-3 row space of the matrix

$$G = \begin{bmatrix} 2 & 0 & 1 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 1 & 2 & 1 & 1 \end{bmatrix}.$$

I have nothing more to say about this code at this point.

3.2 Weight enumerators and inner distribution

If C is a q-ary linear code of length n, then we define A_i to be the number of codewords of Hamming weight i for $0 \le i \le n$. While it is not the focus of the present notes, the *weight enumerator* is a generating function that neatly encodes this information in a polynomial. We define

$$W_C(x, y) = \sum_{c \in C} x^{\mathrm{wt}(c)} y^{n-wt(c)} = \sum_{i=0}^{n} A_i x^i y^{n-i}.$$

Clearly the A_i are non-negative integers which sum to $|C|$. Since C is linear, each codeword is at distance i from exactly A_i codewords. So, for example, $A_0 = 1$ and $A_i = 0$ for $1 \le i < \delta(C)$.

In her 1962 doctoral dissertation at Harvard, the British mathematician F. Jessie MacWilliams[2] proved that the weight enumerator of a linear code C and that of its dual code $C^\perp$ are related by a beautiful equation. Indeed,

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

For example, the dual of a perfect Hamming code has a very simple weight enumerator:

$$W(x,y) = y^n + nx^{\frac{n+1}{2}} y^{n-1} 2.$$

This allows us to obtain a quite compact form for the weight enumerator of a Hamming code, even though almost all of the A_i are non-zero for this code.

Exercises:

1. Use the above technique to find the weight enumerator for the binary Hamming code of length seven. Check this by generating all sixteen codewords by hand.

2. True or False?: for $q > 2$, a q-ary Hamming code has $A_i > 0$ for all $i > 2$.

3. Use MacWilliams's identity to show that, if C has A_i codewords of weight i for each i and $C^\perp$ has A'_i codewords of weight i for each i, then
$$A'_j = \frac{1}{|C|} \sum_{i=0}^{n} \sum_{\ell=0}^{i} (-1)^\ell (q-1)^{i-\ell} \binom{j}{\ell} \binom{n-j}{i-\ell} A_i.$$

This result of MacWilliams had substantial impact on coding theory, only one aspect of which we shall consider here. Suppose one is searching for a linear code with given weight distribution $\{A_i\}_{i=0}^n$. Then one may, without knowing the code — or even whether or not it exists —, compute the weight enumerator of its purported dual. If the resulting polynomial has any negative or non-integral coefficients, this is an automatic proof that the sought-after code does not exist. (In the case where all of the A'_j are non-negative integers, one still obtains quite useful information which may help in locating or ruling out the code.) This powerful non-existence result for

[2] It is useful to know that MacWilliams had already been work in computer programming and coding theory at Bell Labs in the 1950s.

linear codes was known to Delsarte as he began his work and there was even some evidence that inequalities similar to $A'_j \geq 0$ could be used to study non-linear codes as well (e.g., the mysterious formal duality between binary Kerdock and Preparata codes).

This brings us to the linear programming bound. Let n and q be fixed and set $X = \mathcal{Q}^n$ without any assumption about its algebraic structure. For an "unrestricted" (i.e., not necessarily linear or additive) code $C \subseteq X$ and an integer $0 \leq i \leq n$, define

$$a_i = \frac{|C \times C \cap R_i|}{|C|} = \frac{1}{|C|} \left| \{(x,y) \in C \times C : \operatorname{dist}(x,y) = i\} \right|,$$

which is the average number of codewords at distance i from a randomly chosen codeword in C. Note that, when C is additive, we have $a_i = A_i$.

Returning to the adjacency algebra of the Hamming graph, we see that, if $\mathbf{x}$ is the characteristic vector of C as a subset of the vertex set X, we have

$$a_i = \frac{1}{|C|} \mathbf{x}^\top A_i \mathbf{x}.$$

One easily checks the following properties

1. $a_i \geq 0$ for all i;

2. $a_0 + a_1 + \cdots + a_n = |C|$;

3. $a_0 = 1$ and $a_i = 0$ for $1 \leq i < \delta(C)$.

These will all be part of the linear programming bound that we derive. The $(n+1)$-tuple $\mathbf{a} = [a_0, a_1, \ldots, a_n]$ is called the *inner distribution* of code C.

Next define, for $0 \leq j \leq n$,

$$b_j = \frac{|C|}{|X|} \mathbf{x}^\top E_j \mathbf{x}.$$

Then, since the projection matrix E_j is positive semidefinite, we have $b_j \geq 0$ for all j and $b_j = 0$ if and only if $E_j \mathbf{x} = \mathbf{0}$. These are the main inequalities of Delsarte's LP. The $(n+1)$-tuple $\mathbf{b} = [b_0, b_1, \ldots, b_n]$ is called the *dual distribution* of code C.

The last piece of the puzzle comes from Equation (2.4) above. Since we have

$$E_j = \frac{1}{|X|} \sum_{i=0}^{n} Q_{ij} A_i$$

where

$$Q_{ij} = \sum_{\ell=0}^{j} (-1)^{\ell} (q-1)^{j-\ell} \binom{i}{\ell} \binom{n-i}{j-\ell},$$

we obtain the crucial equations

$$b_j = \sum_{i=0}^{n} Q_{ij} a_i \qquad (0 \le j \le n).$$

Now we put this all together. Suppose we are given n, q and d and we wish to maximize the size of a q-ary code C of length n and minimum distance d. Any such code gives us an inner distribution $\mathbf{a}$ satisfying all of the following conditions:

- $a_0 = 1$
- $a_i = 0$ for $1 \le i < d$
- $a_i \ge 0$ for $d \le i \le n$
- $\sum_{i=0}^{n} a_i Q_{ij} \ge 0$ for $1 \le j \le n$.

Moreover, the sum $\sum_{i=0}^{n} a_i$ gives the size of C. In other words, for any such code C, its inner distribution is a feasible solution to the linear programming problem (LP)

$$\begin{array}{rcll} \max\ \sum_{i=0}^{n} a_i & & & \\ \mathit{subject\ \ to} & & & \\ \sum_{i=0}^{n} a_i Q_{ij} & \ge & 0 & (1 \le j \le n) \\ a_0 = 1,\ a_i & = & 0 & (1 \le i \le d-1) \\ a_i & \ge & 0 & (d \le i \le n) \end{array}$$

It therefore follows that the optimal objective value of this LP is a valid upper bound on the size of any q-ary code C having length n and minimum distance d or larger.

3.3 LP bounds for general codes and designs in association schemes

From here to the end of the chapter are notes converted from a talk given at a recent Fields Institute workshop in Waterloo, Canada. The presentation is very general, but the reader should assume that $X = \mathcal{Q}^n$ and A_i is the i^{th} adjacency matrix of the Hamming graph and E_j is the matrix representing orthogonal projection onto the j^{th} eigenspace of this graph as above.

3.3.1 Deriving the Bounds

Let $\mathcal{I} = \{1, \ldots, n\}$ be the index set for the non-zero distances in the Hamming graph $H(n, q)$ and let $\mathcal{J} = \{1, \ldots, n\}$ denote the index set for its nontrivial eigenspaces. (In a more general setting, it may be natural to take $\mathcal{I}$ and $\mathcal{J}$ as different sets.)

Suppose we have some subset $C \subseteq X$ in which we are interested. We define the following statistics for C:

$$a_i = \frac{1}{|C|} x_C^\top A_i x_C \qquad (0 \leq i \leq n)$$

and

$$b_j = \frac{v}{|C|} x_C^\top E_j x_C \qquad (0 \leq j \leq n).$$

The vector $\mathbf{a} = [a_0, a_1, \ldots, a_n]$ is called the *inner distribution* of C. If C is a linear code, then a_i is the number of codewords in C of Hamming weight i. For any non-empty subset C of X, the following basic properties are easily verified:

- $a_i \geq 0$ for all i
- $a_0 = 1$
- $\sum_i a_i = |C|$
- $a_i = 0$ iff no edge of graph (X, R_i) has both ends in C

Now the *dual distribution* of C is the vector $\mathbf{b} = [b_0, b_1, \ldots, b_n]$ where

$$b_j = \frac{v}{|C|} x_C^\top E_j x_C \qquad (0 \leq j \leq n).$$

Observe

- $b_j \geq 0$ for all j
- $b_0 = |C|$
- $\sum_j b_j = |X| \ (= v)$
- $b_j = 0$ iff $x_C \perp V_j$ (the j^{th} eigenspace, col E_j)

3.3.2 The LP bound for Codes

For $\mathcal{A} \subseteq \mathcal{I} = \{1, 2, \ldots, n\}$, $C \subseteq X$ is an "$\mathcal{A}$-code" provided $(C \times C) \cap R_i = \emptyset$ for $i \in \mathcal{A}$.

The size of C is bounded above by the optimal objective value to:

$$\begin{array}{rcll} \max \sum_{i=0}^{n} a_i & & & \\ subject\ \ to & & & \\ \sum_{i=0}^{n} a_i Q_{ij} & \geq & 0 & (1 \leq j \leq n) \\ a_0 = 1,\ a_i & = & 0 & (i \in \mathcal{A}) \\ a_i & \geq & 0 & (1 \leq i \leq n) \end{array}$$

This is a *linear programming problem* ("LP"): we are maximizing (or minimizing) some linear function subject to a finite set of linear constraints, each of which may be an equation or an inequality. In an LP, some variables may be restricted to be non-negative or non-positive, while others may be "free" to take on any real values.

Every linear programming problem has a *dual problem.* Optimizing one is equivalent to optimizing the other (although one or the other may be easier to solve in practice). But there is a reason in our case to prefer the dual LP: note that every $\mathcal{A}$-code gives us a feasible solution to the above LP but only the optimal solution gives us a true upper bound.

What if we don't want to (i.e., can't) solve to optimality?

I will assume for now that the reader has some familiarity with LP duality. For simplicity, I'm going to transform this LP into standard form and take its dual. Here is how duality works for general linear programming problems. The dual of the LP

$$\max \quad c^\top x, \qquad \text{subject to} \qquad Ax \leq b, \qquad x \geq 0$$

is

$$\min y^\top b, \qquad \text{subject to} \qquad y^\top A \geq c^\top, \qquad y \geq 0$$

So in order to apply this to our LP formulation for $\mathcal{A}$-codes, we need to write the same LP as above in standard form. We use the following facts to do this: $a_0 = 1$, $Q_{0j} = m_j$ $(0 \leq j \leq n)$. We obtain

$$\begin{array}{rrcll} & 1 + \max \sum_{i \notin \mathcal{A}} a_i & & & \\ \textit{subject to} & & & & \\ & \sum_{i \notin \mathcal{A}} (-Q_{ij}) a_i & \leq & m_j & (1 \leq j \leq n) \\ & a_i & \geq & 0 & (i > 0,\ i \notin \mathcal{A}) \end{array}$$

So we now have the dual LP for $\mathcal{A}$-codes:

$$\begin{array}{rrcll} & 1 + \min \sum_{j=1}^{n} m_j y_j & & & \\ \textit{subject to} & & & & \\ & \sum_{j=1}^{n} (-Q_{ij}) y_j & \geq & 1 & (i \notin \mathcal{A}) \\ & y_j & \geq & 0 & (1 \leq j \leq n) \end{array}$$

This can now be re-written in a more natural form. We apply routine trickery, with the following substitutions

$$b_j := m_j y_j, \qquad b_0 := 1, \qquad \frac{P_{ji}}{v_i} = \frac{Q_{ij}}{m_j}.$$

The first one is just a linear change of variables. The invention of a new variable b_0 which is forced to equal one is for notational convenience. The last identity, where $v_i = m_i = \binom{n}{i}(q-1)^i$ is a standard orthogonality relation from the basic theory of association schemes.

Using these tricks, we may re-write the dual LP as

$$\begin{array}{rrcll} & \min \sum_{j=0}^{n} b_j & & & \\ \textit{subject to} & & & & \\ & \sum_{j=0}^{n} P_{ji} b_j & \leq & 0 & (i \notin \mathcal{A}) \\ b_0 = 1, & b_j & \geq & 0 & (1 \leq j \leq n) \end{array}$$

Let us explore an easy special case: bounding the size of a *clique* (complete subgraph) in our graph. Suppose $\mathcal{A} = \{2, \ldots, n\}$. Write $P_{j1} = \lambda_j$.

$$\begin{array}{rcll} \min b_0 + b_1 + \cdots + b_n & & & \\ \textit{subject to} & & & \\ b_0\lambda_0 + b_1\lambda_1 + \cdots + b_n\lambda_n & \leq & 0 & \text{(one constraint)} \\ b_0 = 1, \qquad b_1,\ b_2,\ \ldots, b_n & \geq & 0 & \end{array}$$

Since the eigenvalues of the Hamming graph $H(n, q)$ (i.e., the eigenvalues of A_1) are $\lambda_j = n(q-1) - qj$, we can find the optimal solution to this LP by inspection:

- $b_0 = 1, \qquad \lambda_0 = n(q-1)$
- $b_n = -\lambda_0/\lambda_n$ (since $\lambda_n = -n$ is smallest)
- $|C| \leq 1 - \frac{\lambda_0}{\lambda_n} = 1 - \frac{n(q-1)}{-n} = 1 + (q-1) = q$

as perhaps you expected. This special solution to Delsarte's LP is called the *Delsarte bound* (or "Hofffman bound") for cliques. It applies to any distance-regular graph and even more generally.

3.3.3 Designs

Let $\mathcal{T}$ be a subset of $\{1, \ldots, n\}$, the index set for all eigenspaces of $H(n, q)$. A subset D of X is called a $\mathcal{T}$*-design* if its characteristic vector is orthogonal to all eigenspaces V_j for $j \in \mathcal{T}$. There are many types of $\mathcal{T}$-designs in different association schemes, but in the Hamming schemes, the obvious ones to study are *orthogonal arrays of strength* t: these coincide with $\mathcal{T}$-designs where $\mathcal{T} = \{1, 2, \ldots, t\}$.

Equivalently, $D \subseteq X$ with characteristic vector x_D is a $\mathcal{T}$-design provided $E_j x_D = 0$ for all $j \in \mathcal{T}$.

The size of D is bounded **below** by the optimal objective value to:

$$\begin{array}{rcll} \min \sum_{i=0}^n a_i & & & \\ \textit{subject to} & & & \\ \sum_{i=0}^n a_i Q_{ij} & \geq & 0 & (j \notin \mathcal{T}) \\ \sum_{i=0}^n a_i Q_{ij} & = & 0 & (j \in \mathcal{T}) \\ a_0 = 1, \quad a_i & \geq & 0 & (1 \leq i \leq n) \end{array}$$

Using the same techniques as before, we obtain the dual LP for $\mathcal{T}$-designs:

$$\max \sum_{j=0}^{n} b_j$$

$$subject \ \ to$$

$$\begin{array}{rcll} \sum_{j=0}^{n} P_{ji} b_j & \geq & 0 & (1 \leq i \leq n) \\ b_0 = 1, \quad b_j & \leq & 0 & (j > 0, j \notin \mathcal{T}) \end{array}$$

(now using $b_j = -m_j y_j$).

The dual object to a "code" is a "design". The dual object, in this sense, to a clique is what I call a "side" of the graph. In the following digression, let us assume we are working in a Q-polynomial association scheme.

Question: What subsets D satisfy $x_D \in V_0 \oplus V_1$? ("sides")

Question: What is the smallest cardinality of D?

Our optimization problem for designs simplifies in this case to

$$\begin{array}{rrcll} & \max \ b_0 + b_1 & & & \\ subject \ \ to & & & & \\ & P_{0i} b_0 + P_{1i} b_1 & \geq & 0 & (i \neq 0) \\ & b_0 & = & 1 & (b_1 \ unrestr.) \end{array}$$

and the optimal solution we get is

$$|D| \geq \frac{v}{1 - \frac{m_1}{Q_{d1}}}$$

where we assume

$$m_1 = Q_{01} > Q_{11} > \cdots > Q_{d1}.$$

So we have the following

Ratio Bound:

$$|D| \geq \frac{v}{1 - \frac{m_1}{Q_{d1}}}$$

where we assume

$$m_1 = Q_{01} > Q_{11} > \cdots > Q_{d1}$$

Examples:

- *Hamming scheme:* For $H(n,q)$, we get $|D| \geq q^{n-1}$ and the optimal solutions are subgraphs isomorphic to the Hamming graph $H(n-1,q)$, obtained by fixing any single coordinate.
- *Johnson scheme:* For $J(n,k)$, we get $|D| \geq \binom{v-1}{k-1}$ and the optimal solutions are subgraphs isomorphic to $J(n-1,k)$, obtained by taking all k-sets containing any fixed symbol.

3.3.4 The Polynomial Case

Let's focus on the binary Hamming association scheme $H(n,2)$, with vertex set $X = \{0,1\}^n$. A binary code of length n and minimum distance δ is just a coclique (independent set) in the graph $(X, R_1 \cup R_2 \cup \ldots \cup R_{\delta-1})$

Recall that each subset C of X gives us a vector $\mathbf{a} = [a_0, a_1, \ldots, a_n]$ of statistics satisfying Delsarte's inequalities. If we restrict to codes having minimum Hamming distance δ or greater, we also have

$$a_1 = a_2 = \cdots = a_{\delta-1} = 0.$$

So the size of any such code C is bounded above by the optimum objective value to the LP

$$\begin{array}{rlll} \textbf{max} & 1 + a_\delta + a_{\delta+1} + \cdots + a_n & & \\ \textbf{s.t.} & m_j + a_\delta Q_{\delta,j} + a_{\delta+1} Q_{\delta+1,j} + \cdots + a_n Q_{nj} & \geq 0 & (1 \leq j \leq n) \\ & a_\delta,\ a_{\delta+1},\ \ldots,\ a_n & \geq 0 & \end{array}$$

We can write this LP in standard form as

$$\begin{array}{rlll} 1 + \textbf{max} & \sum_{i=\delta}^n a_i & & \\ \textbf{s.t.} & & & \\ & \sum_{i=\delta}^n (-Q_{ij}) a_i & \leq m_j & (1 \leq j \leq n) \\ & a_\delta,\ a_{\delta+1},\ \ldots,\ a_n & \geq 0 & \end{array}$$

Now we easily find the dual LP

$$\begin{array}{rlll} 1 + \textbf{min} & \sum_{j=1}^n m_j f_j & & \\ \textbf{s.t.} & & & \\ & \sum_{j=1}^n (-Q_{ij}) f_j & \geq 1 & (\delta \leq i \leq n) \\ & f_1, \ldots, f_n & \geq 0 & \end{array}$$

We are going to modify this dual LP slightly. Next, we introduce $f_0 = 1$ and flip the inequalities

$$\begin{array}{rll} & \textbf{min} & \sum_{j=0}^{n} m_j f_j \\ & \textbf{s.t.} & \\ & \sum_{j=0}^{n} f_j Q_{ij} & \leq 0 \quad (\delta \leq i \leq n) \\ f_0 = 1, & f_1, \ldots, f_n & \geq 0 \end{array}$$

Now we use the fact that, for the binary Hamming scheme, the entries of the matrix Q are given by the *Krawtchouk polynomials*

$$K_j(\lambda) = \sum_{\ell=0}^{j} (-1)^\ell \binom{\lambda}{\ell} \binom{n-\lambda}{j-\ell}$$

with initial values

$$K_0(\lambda) = 1, \qquad K_1(\lambda) = n - 2\lambda.$$

In general, we have

$$Q_{ij} = K_j(i).$$

So we can replace all occurences of Q_{ij} in our dual LP by evaluations of these polynomials. In particular,

$$m_j = Q_{0j} = K_j(0).$$

LP in Polynomial Form

So our strategy now is to optimize over polynomials

$$F(\lambda) = f_0 K_0(\lambda) + f_1 K_1(\lambda) + \cdots + f_n K_n(\lambda).$$

The scalars f_j are called the *Krawtchouk coefficients* of F and since there is one Krawtchouk polynomial of each degree (zero up to n), these coefficients are uniquely determined by F itself.

Now our optimization problem is

$$\begin{array}{rll} \textbf{min} & F(0) & \\ \textbf{s.t.} & F(\lambda) & = \sum_{j=0}^{n} f_j K_j(\lambda) \\ & F(i) & \leq 0 \quad (\delta \leq i \leq n) \\ f_0 = 1, & f_1, \ldots, f_n & \geq 0 \end{array}$$

Note that we can obtain a potentially weaker bound which looks simpler by appealing to the continuity of polynomial F:

$$\begin{array}{rll} \textbf{min} & F(0) & \\ \textbf{s.t.} & F(\lambda) & = \sum_{j=0}^{n} f_j K_j(\lambda) \\ & F(\alpha) & \leq 0 \ \ \forall \alpha \in [\delta, n] \\ f_0 = 1, & f_1, \ldots, f_n & \geq 0 \end{array}$$

even though we only need non-positivity at the integer points in that closed interval $[\delta, n]$.

3.3.5 Beyond Q-Polynomial

In this next part, we extend two well-known bounds of Delsarte to the setting of association schemes with many vanishing Krein parameters.

Originally, these results were proved by Delsarte for cometric association schemes.

Here, we replace the cometric property with certain vanishing conditions for Krein parameters with reference to a partial order $\trianglelefteq$ on the set $\mathcal{J}$ of eigenspaces of the association scheme.

For $\mathcal{E}$ and $\mathcal{F}$, subsets of $\mathcal{J}$, define

$$\mathcal{E} \star \mathcal{F} = \{k \in \mathcal{J} : \sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{F}} q_{ij}^k > 0\}.$$

Krein conditions imply

$$k \in \mathcal{E} \star \mathcal{F}$$

whenever
$q_{ij}^k \neq 0$ for some $i \in \mathcal{E}$ and some $j \in \mathcal{F}$.

Example: In a cometric scheme, if we take $\mathcal{E} = \{0, \ldots, e\}$ and $\mathcal{F} = \{0, \ldots, f\}$, then

$$\mathcal{E} \star \mathcal{F} \subseteq \{0, \ldots, e + f\}.$$

We use this notation to obtain a very general "Fisher-type" inequality.

Theorem 3.3.1 *Let $\mathcal{T} \subseteq \mathcal{J}$. Assume $\mathcal{E} \subseteq \mathcal{J}$ satisfies $\mathcal{E} \star \mathcal{E} \subseteq \mathcal{T}$. Then, for any Delsarte $\mathcal{T}$-design $D \subseteq X$, we have*

$$|D| \geq \sum_{j \in \mathcal{E}} m_j.$$

Moreover, if equality holds, then, for $\ell \neq 0$ in $\mathcal{J}$,

$$\sum_{j\in\mathcal{E}} Q_{\ell j} = 0$$

whenever D contains a pair of ℓ-related elements.

Proof:
Any matrix $M \in \mathbb{A}$ can be expanded in the form

$$M = v\sum_{j\in\mathcal{J}} \beta_j E_j$$

and also as

$$M = \sum_{i\in\mathcal{I}} \alpha_i A_i$$

where $\alpha_i = \sum_j Q_{ij}\beta_j$ for each $i \in \mathcal{I}$.

Restrict to non-negative matrices $M \in \mathbb{A}$ which satisfy the following two conditions:

(a) $\beta_j \leq 0$ for all $j \notin \mathcal{T}$; and

(b) $\beta_0 = 1$.

WOLOG, assume $0 \in \mathcal{T}$.

Let $D \subseteq X$ be a $\mathcal{T}$-design. Abbreviate x_D to x.

Expand $x^\top M x$ in two ways:

$$\begin{aligned}
|D|\alpha_0 &= \alpha_0 x^\top A_0 x \\
&\leq \sum_{\mathcal{I}} \alpha_i x^\top A_i x = v\sum_{\mathcal{J}} \beta_j x^\top E_j x \\
&= v x^\top E_0 x + v\sum_{\mathcal{T}-\{0\}} \beta_j x^\top E_j x + v\sum_{j\notin\mathcal{T}} \beta_j x^\top E_j x \\
&\leq v x^\top E_0 x = |D|^2.
\end{aligned}$$

This gives us the bound $|D| \geq \alpha_0$.

Rather than optimize, we content ourselves with an easy-to-find feasible solution.

Let

$$F = \sum_{j\in\mathcal{E}} E_j.$$

Then $F \circ F$ is a non-negative matrix with spectral decomposition

$$F \circ F = \sum_{k \in \mathcal{J}} \left(\frac{1}{v} \sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{E}} q_{ij}^k \right) E_k. \tag{3.1}$$

Now, by choice of $\mathcal{E}$, we have $q_{ij}^k = 0$ whenever $i, j \in \mathcal{E}$ and $k \notin \mathcal{T}$.

So condition **(a)** is satisfied by any non-negative multiple of $F \circ F$. We scale by

$$\gamma = \frac{v^2}{\sum_{j \in \mathcal{E}} m_j}$$

to obtain a non-negative matrix $M = \gamma(F \circ F)$ which satisfies conditions **(a)** and **(b)**,

It is straightforward to check that the diagonal entries of M are all equal to

$$\alpha_0 = \sum_{j \in \mathcal{E}} m_j.$$

This proves the following theorem.

Theorem 3.3.2 *Let $\mathcal{T} \subseteq \mathcal{J}$. Assume $\mathcal{E} \subseteq \mathcal{J}$ satisfies $\mathcal{E} \star \mathcal{E} \subseteq \mathcal{T}$. Then, for any Delsarte $\mathcal{T}$-design $D \subseteq X$, we have*

$$|D| \geq \sum_{j \in \mathcal{E}} m_j.$$

In algebraic combinatorics, whenever we prove a bound, we can't resist asking "What if Equality Holds?"

Now if $|D| = \alpha_0$, we return to the above string of equations and inequalities to discover that, for each $\ell \neq 0$,

$$\alpha_\ell \left(x^\top A_\ell x \right) = 0$$

must hold.

These are essentially the "Complementary Slackness Conditions" from the theory of linear programming.

Thus, if D contains a pair of ℓ-related elements, we are forced to have

$$\alpha_\ell = \sum_{k \in \mathcal{J}} \beta_k Q_{\ell k} = 0.$$

Now we find

$$\beta_k = \frac{\gamma}{v} \sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{E}} q_{ij}^k$$

so that

$$\alpha_\ell = \frac{\gamma}{v} \sum_{i \in \mathcal{E}} \sum_{j \in \mathcal{E}} \sum_{k \in \mathcal{J}} q_{ij}^k Q_{\ell k}$$

which gives us

$$\alpha_\ell = \gamma \left(\sum_{j \in \mathcal{E}} Q_{\ell j} \right)^2 = 0 \qquad \text{as desired.}$$

Without proof, let me also mention that tight designs (designs D for which equality holds in the above bound) give subschemes of the ambient association scheme.

Theorem 3.3.3 *Let $\mathcal{T} \subseteq \mathcal{J}$ and assume $\mathcal{E} \subseteq \mathcal{J}$ satisfies $\mathcal{E} \star \mathcal{E} \subseteq \mathcal{T}$.*

(a) *if D is any Delsarte $\mathcal{T}$-design in our scheme with degree s, then $s + 1 \geq |\mathcal{E}|$;*

(b) *if $|\mathcal{E}| = s + 1$, then D is a tight design and D is a subscheme;*

(c) *if $|\mathcal{E}| = s$, then either D is a tight design or D is a subscheme.*

3.3.6 Implementation Issues

When we use linear programming to bound the size of codes and designs, our hope is to do things analytically and, on paper, discover solutions to infinitely many linear programming problems at once. But often that is not achievable. And even when it is, it is often a result of substantial computer experimentation. So we are forced to deal with software packages to do our linear programming for us. I'd like to give you a brief comparison of the following three systems that I have used.

- MAPLE
- CPLEX
- Mathematica

Here is what I find:

- MAPLE
  - Easy
  - Able to generate coefficients
  - SLOW simplex, limited to ~ 350 variables
- CPLEX (expensive!)
  - Numerical (64-bit precision)
  - C/C++ library or dumb interactive interface
  - incredibly fast, thousands of variables no problem
  - rounding errors, large condition numbers
- Mathematica
  - I'm just learning it
  - Able to generate coefficients
  - Exact arithmetic, many more variables than MAPLE

Ratio Bound for Cocliques

Before we leave this chapter, let's solve one more LP.

A *coclique* in the Hamming graph $H(n, q)$ is just a code $C \subseteq X$ with minimum distance at least two. We want the linear programming bound for cocliques in (X, R_1).

Our dual LP becomes

$$\min \sum_{j=0}^{n} b_j$$

$$subject \;\; to$$

$$\begin{array}{rcll} \displaystyle\sum_{j=0}^{n} P_{ji} b_j & \leq & 0 & \qquad (i \neq 1) \\ b_0 = 1, \qquad b_j & \geq & 0 & \qquad (1 \leq j \leq n) \end{array}$$

We immediately see two easy solutions to consider:

If we let $b_j = m_j$, we get

$$\mathbf{b}P = [v, 0, 0, \ldots, 0]$$

which is feasible but useless.

If we let $b_j = Q_{1j}$, we get

$$\mathbf{b}P = [0, v, 0, 0, \ldots, 0]$$

which is infeasible but has a great objective value! (zero)

The trick is to combine these two easy solutions to obtain a feasible solution with a reasonable bound.

So we take

$$b_j = sm_j + tQ_{1j}$$

and we get

$$\mathbf{b}P = [sv, tv, 0, \ldots, 0]$$

Our goal is to make s is as small as possible subject to the conditions

$$b_j = sm_j + tQ_{1j} \geq 0$$

for all j.

What are the best values for a and t?

We need $b_0 = 1$, so set $t = 1 - s$.

Now we need

$$b_j = sm_j + (1-s)Q_{1j} \geq 0$$

which is the same as

$$s + (1-s)\frac{P_{j1}}{v_1} \geq 0$$

$$P_{j1} + s\,(v_1 - P_{j1}) \geq 0$$

for all $j = 1, \ldots, n$.

So the smallest eigenvalue will give us the best choice.

The eigenvalues of A_1 are $P_{01}, \ldots, P_{n1}$. Call these

$$k = \lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_n.$$

In order to have

$$s \geq -\lambda_j/(k - \lambda_j)$$

for all j, we can ignore the nonnegative λ and take

$$s = \lambda_n / (\lambda_n - k)$$

which gives the desired bound

$$|S| \leq \frac{v}{1 - \frac{k}{\lambda_n}}.$$

This ratio bound

$$|S| \leq \frac{v}{1 - \frac{k}{\lambda_n}}$$

for cocliques is a feasible solution, but not always optimal.

If we allow some $\sum_j P_{ji} b_j < 0$, then we may be able to do better. But then our optimal coclique must contain no i-related pair.

Chapter 4

The Terwilliger Algebra

The Bose-Mesner algebra introduced in Chapter 2 has very beautiful structure, but one annoying feature: there are two different matrix multiplications and there are very few rules to govern how these two interact. Each matrix in the Bose-Mesner algebra is uniquely determined by its first column. (Why?) So the entrywise product of two matrices M, N in $\mathbb{A}$ is entirely determined by the product of the two diagonal matrices $\Delta(M)$ and $\Delta(N)$ where $\Delta(M)$ has (a,a)-entry equal to the $(a,1)$ entry of M and the diagonal matrix $\Delta(N)$ has (a,a)-entry $(N)_{a,1}$. The linear transformation $M \mapsto \Delta(M)$ is therefore a ring homomorphism that maps entrywise multiplication to ordinary matrix multiplication. The image of the Bose-Mesner algebra $\mathbb{A}$ under this map is called the *dual Bose-Mesner algebra* and is denoted $\mathbb{A}^*$.

Terwilliger now lets the two types of matrix multiplication interact by taking the smallest matrix algebra $\mathbb{T}$ containing both $\mathbb{A}$ and $\mathbb{A}^*$. For the n-cube, this algebra is generated by just two matrices: the adjacency matrix A and the diagonal matrix A^* with (a,a)-entry $n - 2\mathrm{wt}(a)$. For $n = 1$, this is easily seen to be the full matrix algebra $\mathsf{Mat}_2(\mathbb{C})$.

Exericse: Write down A and A^* for $n = 2$ and locate the primitive idempotents for $\mathbb{A}^*$. (More challenging: find a vector space basis for the 10-dimensional Terwilliger algebra $\mathbb{T}$.)

Terwillger defines this algebra for any graph and the Terwilliger algebra has been studied for a great variety of association schemes, but most work in this area has focused on $\mathbb{T}$-algebras of so-called "P- and Q-polynomial association schemes" in the hope of moving toward a full classification of all such association scheme with six or more classes.

But our interest here is only in the Terwilliger algebra of the n-cube.

Fortunately for us, there is a paper by J.T. Go exactly on this topic and it has great tutorial value. So I will simply refer you to Go's paper [21] and briefly summarize its results.

The paper of Go gives the complete decomposition of the "standard module" $\mathbb{C}^X$ into irreducible $\mathbb{T}$-modules. (Here, $X = \{0,1\}^n$.) Let E_i^* denote the diagonal matrix with (a,a)-entry equal to one if $\mathrm{wt}(a) = i$ and zero otherwise. Fixing n and considering the Terwilliger algebra of the n-cube, the *raising operator* for $\mathbb{T}$ is the matrix

$$\sum_{i=0}^{n-1} E_{i+1}^* A E_i^*$$

and the *lowering operator* is the matrix

$$\sum_{i=1}^{n} E_{i-1}^* A E_i^*.$$

Observe that the adjacency matrix is $A = L + R$. Terwilliger observed that these matrices have very nice Lie brackets[1]:

$$LR - RL = A^*, \qquad RA^* - A^*R = 2R, \qquad LA^* - A^*L = -2L.$$

This leads to a natural action of the universal enveloping algebra of the Lie algebra $sl_2(\mathbb{C})$ on the standard module and therefore on every irreducible $\mathbb{T}$-module of it. This is the beginning of a very long story. Terwilliger, Ito, Nomura and others continue to search for a sort of "universal cover", a Lie algebra with the property that every irreducible module of every Terwilliger algebra of every P- and Q-polynomial association scheme admits such an action from this algebra. They think they are close to an answer, but this is ongoing research, and very exciting.

Let me trust that you will pick up Go's very nice paper and at least browse it. For my part, I want to begin with two explicit vector space bases for the Terwilliger algebra of the n-cube and find the change-of-basis matrices between them.

Any three (not necessarily distinct) tuples a, b, c from X give rise to a triple of non-negative integers (i,j,k) via the equations

$$\mathrm{dist}(b,c) = i, \qquad \mathrm{dist}(a,c) = j, \qquad \mathrm{dist}(a,b) = k.$$

[1] The "bracket product" of matrices M and N is $[M,N] = MN - NM$, which is zero iff M and N commute.

Not every triple (i,j,k) of non-negative integers occurs in this way; the triangle inequality must hold, none of i,j,k may exceed n, their sum must not exceed $2n$, and $i+j+k$ must be even. For n fixed, there is a simple bijection between the set of all such triples (i,j,k) and the set of all ordered quadruples $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ of non-negative integers summing to n: it is given by the system

$$\begin{aligned} \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 &= n, \\ \alpha_2 + \alpha_3 &= i, \\ \alpha_1 + \qquad \alpha_3 &= j, \\ \alpha_1 + \alpha_2 \qquad &= k. \end{aligned}$$

(In fact, for any three vertices v_1, v_2, v_3 in the n-cube whose pairwise distances are as above, there exists a unique vertex $u \in X$ (the "apex") with $\mathrm{dist}(u,a) = \alpha_1$, $\mathrm{dist}(u,b) = \alpha_2$ and $\mathrm{dist}(u,c) = \alpha_3$.) Henceforth, we will use $\vartheta(i,j,k) = \alpha$ to indicate that the triple (i,j,k) is related to the quadruple $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ via these equations. By a slight abuse of terminology, we will refer to α as a *composition of n (into four parts).*

4.1 Two bases for the algebra $\mathbb{T}$

We consider the binary Hamming graph Q_n with vertex set X and adjacency matrices $A_0, \ldots, A_n$. The primitive idempotents for the Bose-Mesner algebra $\mathbb{A}$ will be denoted $E_0, \ldots, E_n$. We have

$$E_j = \frac{1}{2^n} \sum_{i=0}^{n} Q_{i,j} A_i$$

where

$$Q_{i,j} = [z^j](1+z)^{n-i}(1-z)^i. \tag{4.1}$$

Here, as below, we use this notation to describe the coefficient of z^j in the expansion of the (finite) power series $(1+z)^{n-i}(1-z)^i$.

In order to construct the dual Bose-Mesner algebra $\mathbb{A}^*$, we first select a base point: here we choose the tuple $\mathbf{0}$. We define A_i^* to be the diagonal matrix with (x,x)-entry equal to $|X|(E_i)_{x,\mathbf{0}}$ and we define E_j^* to be the diagonal matrix with (x,x)-entry equal to $(A_j)_{x,\mathbf{0}}$. These give us two bases for $\mathbb{A}^*$. The subconstituent algebra (or, Terwilliger algebra) of the n-cube is then the algebra generated by $\mathbb{A}$ and $\mathbb{A}^*$. We denote this non-commutative algebra by $\mathbb{T}$.

For a composition α of n, we define matrices

$$L_\alpha = E_i^* A_j E_k^*$$

and

$$M_\alpha = E_i A_j^* E_k$$

where $\vartheta(i,j,k) = \alpha$. It is easy to see that $L_\alpha \neq 0$ if and only if the intersection number $p_{i,j}^k > 0$ and it is well-known $M_\alpha \neq 0$ if and only if the Krein parameter $q_{i,j}^k$ is non-zero. Our goal in this section is to describe the change-of-basis matrix from the basis

$$\mathcal{A} = \left\{ L_\alpha = E_i^* A_j E_k^* : p_{i,j}^k > 0, \vartheta(i,j,k) = \alpha \right\}$$

to the basis

$$\mathcal{B} = \left\{ M_\beta = E_r A_s^* E_t : q_{r,s}^t > 0, \vartheta(r,s,t) = \beta \right\}.$$

We freely use basic facts about this algebra (see [37, 38, 39, 21] for a full treatment).

It is now easy to compute the dimension of $\mathbb{T}$. From the previous section, we know that the triples (i,j,k) for which $p_{i,j}^k > 0$ are in one-to-one correspondence with compositions α of n into four parts. There are $\binom{n+3}{3}$ such compositions. So the dimension of $\mathbb{T}$ is $\binom{n+3}{3}$.

4.1.1 Change-of-basis coefficients

From above, we know that there are unique rational numbers t_α^β which satisfy the following system of equations

$$E_r A_s^* E_t = \frac{1}{2^n} \sum_\alpha t_\alpha^\beta E_i^* A_j E_k^*$$

where $\vartheta(i,j,k) = \alpha$ and $\vartheta(r,s,t) = \beta$. We wish to learn more about these coefficients t_α^β.

Note that the dimension of the Terwilliger algebra $\mathbb{T}$ is also the dimension of the vector space $\mathsf{Hom}_n(y_0, y_1, y_2, y_3)$ of homogeneous polynomials of degree

n in four variables. In fact, we have a natural isomorphism of vector spaces $\varphi : \mathbb{T} \to \mathsf{Hom}_n(y_0, y_1, y_2, y_3)$ given by

$$\varphi : E_i^* A_j E_k^* \mapsto y_0^{\alpha_0} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3} \qquad (\vartheta(i,j,k) = \alpha).$$

This vector space isomorphism can be exploited to solve our problem in an elegant way. For the n-cube, it so happens that the Terwilliger algebra $\mathbb{T}_n$ is coherent. That is, we have a coherent configuration[2] $(X, \mathcal{R})$ with relations corresponding to the zero-one basis

$$\left\{ L_\alpha = E_i^* A_j E_k^* : \ p_{i,j}^k > 0 \right\}. \tag{4.2}$$

These matrices can be extracted from a generating function as follows. Let Z_0, Z_1, Z_2, Z_3 be four commuting indeterminates and consider the $2^n \times 2^n$ matrix

$$\begin{aligned} \Phi_n &= \begin{bmatrix} Z_0 & Z_1 \\ Z_3 & Z_2 \end{bmatrix}^{\otimes n} & (4.3) \\ &= \left(Z_0 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + Z_1 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + Z_2 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + Z_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right)^{\otimes n} & (4.4) \end{aligned}$$

If $\vartheta(i,j,k) = \alpha$ (with n pre-specified), we have

$$E_i^* A_j E_k^* = [Z_0^{\alpha_0} Z_1^{\alpha_1} Z_2^{\alpha_2} Z_3^{\alpha_3}] \ \Phi_n. \tag{4.5}$$

In particular, the standard basis is our basis of Schur idempotents for $\mathbb{T}_1$.

Our second distinguished basis for $\mathbb{T}_1$ is

$$\left\{ E_0 A_0^* E_0, E_0 A_1^* E_1, E_1 A_0^* E_1, E_1 A_1^* E_0 \right\}, \tag{4.6}$$

explicitly

$$\left\{ \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \right\}. \tag{4.7}$$

[2] A coherent configuration is a generalization of an association scheme. Rather than define it here, let me just give a theorem of D. Higman which says that a coherent algebra (the analogue of a Bose-Mesner algebra) is precisely a non-trivial vector space of $n \times n$ complex matrices which is closed under matrix multiplication, Schur multiplication, and conjugate transposition. The Schur idempotents are then 01-matrices which partition the complete graph on n vertices into a nicely behaved set of directed graphs.

The change-of-basis matrix for $\mathbb{T}_1$ from the standard basis (5.2) to the basis (5.7) is given by

$$H = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}. \tag{4.8}$$

This allows us to recover the coefficients t_α^β. If we define

$$\Psi_n = \left(Z_0(E_0A_0^*E_0) + Z_1(E_0A_1^*E_1) + Z_2(E_1A_0^*E_1) + Z_3(E_1A_1^*E_0)\right)^{\otimes n}, \tag{4.9}$$

then

$$M_\beta = E_rA_s^*E_t = [Z_0^{\beta_0}Z_1^{\beta_1}Z_2^{\beta_2}Z_3^{\beta_3}]\ \Psi_n \tag{4.10}$$

where $\vartheta(r,s,t) = \beta$. Using (5.8), we may write

$$\begin{aligned} \Psi_n \;=\; & \frac{1}{2^n}\left((Z_0 + Z_1 + Z_2 + Z_3)(E_0^*A_0E_0^*) + \right. \\ & (Z_0 - Z_1 - Z_2 + Z_3)(E_1^*A_1E_0^*) + \\ & (Z_0 - Z_1 + Z_2 - Z_3)(E_1^*A_0E_1^*) + \\ & \left.(Z_0 + Z_1 - Z_2 - Z_3)(E_0^*A_1E_1^*)\right)^{\otimes n}. \end{aligned} \tag{4.11}$$

Thus, using (5.5) and (5.10), we have the following theorem:

Theorem 4.1.1 *The connection coefficients t_α^β satisfying*

$$M_\beta = \frac{1}{2^n}\sum_\alpha t_\alpha^\beta L_\alpha$$

are given by

$$\begin{aligned} t_\alpha^\beta \;=\; & [Z_0^{\alpha_0}Z_1^{\alpha_1}Z_2^{\alpha_2}Z_3^{\alpha_3}]\,(Z_0 + Z_1 + Z_2 + Z_3)^{\beta_0}\,(Z_0 - Z_1 - Z_2 + Z_3)^{\beta_1}\cdot \\ & (Z_0 - Z_1 + Z_2 - Z_3)^{\beta_2}\,(Z_0 + Z_1 - Z_2 - Z_3)^{\beta_3}. \end{aligned} \tag{4.12}$$

Proof: We use the duality of the algebra to fill in the gaps above. Let $S = 2^{-n/2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{\otimes n}$. Then S diagonalizes the Bose-Mesner algebra. So

$$S^TA_iS = A_i^*, \qquad S^TE_jS = E_j^*.$$

Since S is symmetric, we also have

$$S^TA_i^*S = A_i, \qquad S^TE_j^*S = E_j$$

so that $S^TL_\alpha S = M_\alpha$ and vice versa. Since this holds for $n = 1$, we find $S^T\Phi_nS = \Psi_n$ and $S^T\Psi_nS = \Phi_n$ for all n. Now the calculation follows. □

Chapter 5

The Biweight Enumerator

In this chapter, we introduce the biweight enumerator of a binary error-correcting code and show how this enumerator is naturally associated to the Terwilliger algebra of the n-cube. Using this connection and the relationship between the eigenspaces of the n-cube and the characters of $\mathbb{Z}_2^n$, we give a combinatorial proof of the MacWilliams identities for this enumerator. We finish with an exploration of some simple (but not very strong) inequalities for codes which can be derived from these observations. This material was presented at the Workshop on Asymptotic and Computational Aspects of Coding Theory at the Institute for Advanced Study (Princeton) in March 2001[1].

5.1 The biweight enumerator

Let n be a positive integer and let $X = \{0,1\}^n$. The *Hamming weight* of $u \in X$, denoted as $\mathrm{wt}(u)$, is the number of non-zero coordinates in the tuple u. For $u, v \in X$, the *Hamming distance*, $\mathrm{dist}(u,v)$, is defined as the number of coordinates h ($1 \le h \le n$) with $u_h \neq v_h$. By a *(binary) code* C of length n, we simply mean any non-trivial subset of X. In this paper, we will always assume that C contains the zero tuple $\mathbf{0} = 000\cdots 0$.

Any three (not necessarily distinct) tuples v_1, v_2, v_3 from X give rise to a triple of non-negative integers (i,j,k) via the equations

$$\mathrm{dist}(v_2, v_3) = i, \qquad \mathrm{dist}(v_1, v_3) = j, \qquad \mathrm{dist}(v_1, v_2) = k.$$

[1] A bit of information about this workshop can still be found at
`http://www.math.ias.edu/~huguenin/codingconf.html`

Not every triple (i, j, k) of non-negative integers occurs in this way; the triangle inequality must hold, none of i, j, k may exceed n, their sum must not exceed $2n$, and $i+j+k$ must be even. For n fixed, there is a simple bijection between the set of all such triples (i, j, k) and the set of all ordered quadruples $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ of non-negative integers summing to n: it is given by the system

$$\begin{aligned} \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 &= n, \\ \alpha_2 + \alpha_3 &= i, \\ \alpha_1 + \alpha_3 &= j, \\ \alpha_1 + \alpha_2 &= k. \end{aligned}$$

(In fact, for any three vertices v_1, v_2, v_3 in the n-cube whose pairwise distances are as above, there exists a unique vertex $u \in X$ with $\mathrm{dist}(u, v_i) = \alpha_i$ for $i = 1, 2, 3$.) Henceforth, we will use $\vartheta(i, j, k) = \alpha$ to indicate that the triple (i, j, k) is related to the quadruple $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ via these equations. By a slight abuse of terminology, we will refer to α as a *composition of n (into four parts)*.

Let C be a binary code of length n. For each composition $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ of n into four parts, let

$$\ell_\alpha = |\{(c, c') \in C \times C : \mathrm{wt}(c) = i,\ \mathrm{dist}(c, c') = j,\ \mathrm{wt}(c') = k\}|$$

where $(i, j, k) = \vartheta^{-1}(\alpha)$. Consider the enumerator

$$\mathcal{W}(y_0, y_1, y_2, y_3) = \sum_\alpha \ell_\alpha y_0^{\alpha_0} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3}.$$

Below, we will establish a connection between this enumerator for a linear code and the enumerator for its dual. Moreover, we will examine a linear programming approach based on this enumerator. First, we will need to compute a change-of-basis matrix for an important matrix algebra related to our problem.

5.2 Two bases for the algebra $\mathbb{T}$

We consider the binary Hamming graph Q_n with vertex set X and adjacency matrices $A_0, \ldots, A_n$. The primitive idempotents for the Bose-Mesner algebra

$\mathbb{A}$ will be denoted $E_0, \ldots, E_n$. We have

$$E_j = \frac{1}{2^n} \sum_{i=0}^{n} Q_{i,j} A_i$$

where

$$Q_{i,j} = [z^j](1+z)^{n-i}(1-z)^i. \tag{5.1}$$

Here, as below, we use this notation to describe the coefficient of z^j in the expansion of the (finite) power series $(1+z)^{n-i}(1-z)^i$.

In order to construct the dual Bose-Mesner algebra $\mathbb{A}^*$, we first select a base point: here we choose the tuple $\mathbf{0}$. We define A_i^* to be the diagonal matrix with (x,x)-entry equal to $|X|(E_i)_{x,\mathbf{0}}$ and we define E_j^* to be the diagonal matrix with (x,x)-entry equal to $(A_j)_{x,\mathbf{0}}$. These give us two bases for $\mathbb{A}^*$. The subconstituent algebra (or, Terwilliger algebra) of the n-cube is then the algebra generated by $\mathbb{A}$ and $\mathbb{A}^*$. We denote this non-commutative algebra by $\mathbb{T}$. Terwilliger algebras can be defined for any association scheme and these algebras have been proposed in the analysis of Q-polynomial distance regular graphs and more broadly. But the Terwilliger algebras of the n-cubes are particularly well-behaved and are the focus of several articles.

For a composition α of n, we define matrices

$$L_\alpha = E_i^* A_j E_k^*$$

and

$$M_\alpha = E_i A_j^* E_k$$

where $\vartheta(i,j,k) = \alpha$. It is easy to see that $L_\alpha \neq 0$ if and only if the intersection number $p_{i,j}^k > 0$ and it is well-known $M_\alpha \neq 0$ if and only if the Krein parameter $q_{i,j}^k$ is non-zero. Our goal in this section is to describe the change-of-basis matrix from the basis

$$\mathcal{A} = \left\{ L_\alpha = E_i^* A_j E_k^* : p_{i,j}^k > 0, \vartheta(i,j,k) = \alpha \right\}$$

to the basis

$$\mathcal{B} = \left\{ M_\beta = E_r A_s^* E_t : q_{r,s}^t > 0, \vartheta(r,s,t) = \beta \right\}.$$

We freely use basic facts about this algebra (see [37, 38, 39, 21] for a full treatment).

It is now easy to compute the dimension of $\mathbb{T}$. From the previous section, we know that the triples (i,j,k) for which $p_{i,j}^k > 0$ are in one-to-one correspondence with compositions α of n into four parts. There are $\binom{n+3}{3}$ such compositions. So the dimension of $\mathbb{T}$ is $\binom{n+3}{3}$.

5.2.1 Change-of-basis coefficients

From above, we know that there are unique rational numbers t_α^β which satisfy the following system of equations

$$E_r A_s^* E_t = \frac{1}{2^n} \sum_\alpha t_\alpha^\beta E_i^* A_j E_k^*$$

where $\vartheta(i,j,k) = \alpha$ and $\vartheta(r,s,t) = \beta$. We wish to learn more about these coefficients t_α^β.

Note that the dimension of the Terwilliger algebra $\mathbb{T}$ is also the dimension of the vector space $\mathsf{Hom}_n(y_0, y_1, y_2, y_3)$ of homogeneous polynomials of degree n in four variables. In fact, we have a natural isomorphism of vector spaces $\varphi : \mathbb{T} \to \mathsf{Hom}_n(y_0, y_1, y_2, y_3)$ given by

$$\varphi : E_i^* A_j E_k^* \mapsto y_0^{\alpha_0} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3} \qquad (\vartheta(i,j,k) = \alpha).$$

This vector space isomorphism can be exploited to solve our problem in an elegant way. For the n-cube, it so happens that the Terwilliger algebra $\mathbb{T}_n$ is coherent. That is, we have a coherent configuration $(X, \mathcal{R})$ with relations corresponding to the zero-one basis

$$\left\{ L_\alpha = E_i^* A_j E_k^* : \ p_{i,j}^k > 0 \right\}. \tag{5.2}$$

These matrices can be extracted from a generating function as follows. Let Z_0, Z_1, Z_2, Z_3 be four commuting indeterminates and consider the $2^n \times 2^n$ matrix

$$\begin{aligned} \Phi_n &= \begin{bmatrix} Z_0 & Z_1 \\ Z_3 & Z_2 \end{bmatrix}^{\otimes n} & (5.3) \\ &= \left(Z_0 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + Z_1 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + Z_2 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + Z_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right)^{\otimes n} & (5.4) \end{aligned}$$

If $\vartheta(i,j,k) = \alpha$ (with n pre-specified), we have

$$E_i^* A_j E_k^* = [Z_0^{\alpha_0} Z_1^{\alpha_1} Z_2^{\alpha_2} Z_3^{\alpha_3}] \ \Phi_n. \tag{5.5}$$

In particular, the standard basis is our basis of Schur idempotents for $\mathbb{T}_1$.

Our second distinguished basis for $\mathbb{T}_1$ is

$$\{ E_0 A_0^* E_0, E_0 A_1^* E_1, E_1 A_0^* E_1, E_1 A_1^* E_0 \}, \tag{5.6}$$

explicitly

$$\left\{\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}\right\}. \tag{5.7}$$

The change-of-basis matrix for $\mathbb{T}_1$ from the standard basis (5.2) to the basis (5.7) is given by

$$H = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}. \tag{5.8}$$

This allows us to recover the coefficients t_α^β. If we define

$$\Psi_n = (Z_0(E_0A_0^*E_0) + Z_1(E_0A_1^*E_1) + Z_2(E_1A_0^*E_1) + Z_3(E_1A_1^*E_0))^{\otimes n}, \tag{5.9}$$

then

$$M_\beta = E_rA_s^*E_t = [Z_0^{\beta_0}Z_1^{\beta_1}Z_2^{\beta_2}Z_3^{\beta_3}]\ \Psi_n \tag{5.10}$$

where $\vartheta(r,s,t) = \beta$. Using (5.8), we may write

$$\begin{aligned} \Psi_n \;=\; & \frac{1}{2^n}((Z_0 + Z_1 + Z_2 + Z_3)(E_0^*A_0E_0^*) + \\ & (Z_0 - Z_1 - Z_2 + Z_3)(E_1^*A_1E_0^*) + \\ & (Z_0 - Z_1 + Z_2 - Z_3)(E_1^*A_0E_1^*) + \\ & (Z_0 + Z_1 - Z_2 - Z_3)(E_0^*A_1E_1^*))^{\otimes n}. \end{aligned} \tag{5.11}$$

Thus, using (5.5) and (5.10), we have the following theorem:

Theorem 5.2.1 *The connection coefficients t_α^β satisfying*

$$M_\beta = \frac{1}{2^n}\sum_\alpha t_\alpha^\beta L_\alpha$$

are given by

$$\begin{aligned} t_\alpha^\beta \;=\; & [Z_0^{\alpha_0}Z_1^{\alpha_1}Z_2^{\alpha_2}Z_3^{\alpha_3}]\,(Z_0 + Z_1 + Z_2 + Z_3)^{\beta_0}\,(Z_0 - Z_1 - Z_2 + Z_3)^{\beta_1}\cdot \\ & (Z_0 - Z_1 + Z_2 - Z_3)^{\beta_2}\,(Z_0 + Z_1 - Z_2 - Z_3)^{\beta_3}. \end{aligned} \tag{5.12}$$

Proof: We use the duality of the algebra to fill in the gaps above. Let $S = 2^{-n/2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{\otimes n}$. Then S diagonalizes the Bose-Mesner algebra. So

$$S^T A_i S = A_i^*, \qquad S^T E_j S = E_j^*.$$

Since S is symmetric, we also have

$$S^T A_i^* S = A_i, \qquad S^T E_j^* S = E_j$$

so that $S^T L_\alpha S = M_\alpha$ and vice versa. Since this holds for $n = 1$, we find $S^T \Phi_n S = \Psi_n$ and $S^T \Psi_n S = \Phi_n$ for all n. Now the calculation follows. □

5.3 Linear Programming

Let us briefly review Delsarte's linear programming bound for codes. We seek an upper bound on the size of a code $C \subseteq X$ which has minimum distance at least d. We find that $|C|$ is bounded above by

$$\mathsf{maximize} \sum_{i=0}^{n} a_i$$

$$\mathsf{subject\ to}$$

$$\begin{array}{rcll} \displaystyle\sum_{i=0}^{n} Q_{i,j} a_i & \geq & 0 & \qquad (0 \leq j \leq n) \\ a_i & = & 0 & \qquad (1 \leq i < d) \\ a_0 & = & 1 & \\ a_i & \geq & 0 & \qquad (d \leq i \leq n). \end{array}$$

This is derived as follows. Given a non-empty code C with minimum distance at least d and characteristic vector χ, define

$$a_i = \frac{1}{|C|}\chi^T A_i \chi, \qquad (0 \leq i \leq n).$$

Then, clearly, $a_0 = 1$, each $a_i \geq 0$ and $\sum_i a_i = |C|$. Now define

$$b_j = \frac{|X|}{|C|}\chi^T E_j \chi, \qquad (0 \leq j \leq n).$$

Since each E_j is symmetric and positive semi-definite, we know that each $b_j \geq 0$. Moreover, since

$$E_j = \frac{1}{|X|} \sum_{i=0}^{n} Q_{i,j} A_i,$$

we have $b_j = \sum_i Q_{i,j} a_i$.

We now extend this approach to the Terwilliger algebra of the n-cube. Throughout, we assume that our code contains the zero word.

For each triple (i, j, k) such that $p^j_{i,k} > 0$, let $\alpha = \vartheta(i, j, k)$ and define

$$\ell_\alpha = \chi^T E^*_i A_j E^*_k \chi.$$

Then

$$\ell_\alpha = \left| \{(y, z) \in C \times C : \mathrm{wt}(y) = i,\ \mathrm{wt}(z) = k,\ \mathrm{dist}(y, z) = j\} \right|.$$

Thus ℓ_α is a non-negative integer. As well, the sum of all such ℓ_α is equal to $|C|^2$. These quantities seem deserving of further study. Similarly, for each triple (r, s, t) such that $q^s_{r,t} > 0$, we set $\beta = \vartheta(r, s, t)$ and introduce

$$m_\beta = |X| \cdot \chi^T E_r A^*_s E_t \chi.$$

Since

$$M_\beta = \frac{1}{2^n} \sum_\alpha t^\beta_\alpha L_\alpha,$$

we have

$$m_\beta = \sum_\alpha t^\beta_\alpha \ell_\alpha.$$

It remains to present an efficient strategy for bounding (or interpreting) the parameters m_β of a code. We will first deal with the case of linear codes.

5.3.1 MacWilliams identities

Assume C is a binary linear code. Let $C^\perp_j$ denote the set of dual codewords of C having weight j. A matrix diagonalizing the Bose-Mesner algebra is $2^{-n/2} S$ where

$$S = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{\otimes n}$$

has (x, y)-entry $(-1)^{x \cdot y}$ (here $\cdot$ denotes the binary dot product). Let S_j denote the submatrix of S obtained by restricting to columns indexed by tuples of weight j. Then

$$E_j = \frac{1}{2^n} S_j S_j^T$$

for $0 \leq j \leq d$. For $x \in X$, we use s_x to denote the column of S indexed by x.

We use the following standard fact:

$$\sum_{x \in C} (-1)^{y \cdot x} = \begin{cases} |C|, & \text{if } y \in C^\perp; \\ 0, & \text{otherwise.} \end{cases}$$

So

$$u_j(C) := E_j \chi = \frac{1}{2^n} S_j S_j^T \chi = \frac{|C|}{2^n} \sum_{y \in C_j^\perp} s_y,$$

and if C is an $[n, k, d]$-code, this gives

$$u_j(C) := 2^{k-n} \sum_{y \in C_j^\perp} s_y.$$

Now if $\beta = \vartheta(i, j, k)$,

$$\begin{aligned} m_\beta &= 2^n u_i(C)^T A_j^* u_k(C) = 4^n \langle u_i(C), u_j(0) \circ u_k(C) \rangle \\ &= 4^n \langle u_j(0), u_i(C) \circ u_k(C) \rangle \end{aligned}$$

where $u_j(0)$ denotes the column of E_j indexed by 0. Clearly

$$u_j(0) = \frac{1}{2^n} \sum_{\mathrm{wt}(w) = j} s_w.$$

We need to examine the Schur product of $u_i(C)$ with $u_k(C)$. We have

$$u_i(C) \circ u_k(C) = 4^{k-n} \sum_{y \in C_i^\perp} \sum_{z \in C_k^\perp} s_y \circ s_z = 4^{k-n} \sum_{y \in C_i^\perp} \sum_{z \in C_k^\perp} s_{y \oplus z}$$

where $\oplus$ denotes vector addition over $GF(2)$.

Now we have

$$\begin{aligned} m_\beta = 2^n \chi^T E_i A_j^* E_k \chi & \\ = & \ 4^n \langle u_j(0), u_i(C) \circ u_k(C) \rangle \\ = & \ 4^k \sum_{y \in C_i^\perp} \sum_{z \in C_k^\perp} \langle u_j(0), s_{y \oplus z} \rangle \\ = & \ 2^{2k-n} \sum_{y \in C_i^\perp} \sum_{z \in C_k^\perp} \sum_{\mathrm{wt}(w)=j} \langle s_w, s_{y\oplus z} \rangle . \end{aligned}$$

Now the columns of S are pairwise orthogonal; in fact, $S^T S = 2^n I$. So

$$m_\beta = 2^n \chi^T E_i A_j^* E_k \chi = 4^k \left| \left\{ (y,z) \in C_i^\perp \times C_k^\perp : \mathrm{wt}(y \oplus z) = j \right\} \right| .$$

Written a bit differently, this is

$$m_\beta = 2^{2k} \left| \left\{ (y,z) \in C^\perp \times C^\perp : \mathrm{wt}(y) = i, \ \mathrm{dist}(y,z) = j, \ \mathrm{wt}(z) = k \right\} \right| .$$

Now, using the results of Section 5.2.1, m_β is the coefficient of $y_0^{\beta_0} y_1^{\beta_1} y_2^{\beta_2} y_3^{\beta_3}$ in the expansion of $\mathcal{W}_C(H\mathbf{y})$ where H is as above and $\mathbf{y} = [y_0, y_1, y_2, y_3]^T$. Thus we have a new proof of

Theorem 5.3.1 (MacWilliams' identities for biweight enumerator [30])

$$\begin{aligned} \mathcal{W}_{C^\perp}(y_0, y_1, y_2, y_3) = \frac{1}{|C|^2} \cdot \mathcal{W}_C(y_0 + y_1 + y_2 + y_3, & \\ y_0 - y_1 - y_2 + y_3, y_0 - y_1 + y_2 - y_3, y_0 + y_1 - y_2 - y_3) & \end{aligned} \tag{5.13}$$

Of course, such a theorem always gives us a linear programming bound as a by-product.

Corollary 5.3.2 *If C is a linear code, then, for each triple (i, j, k) with $q_{i,j}^k > 0$,*

$$m_\beta \geq 0.$$

5.3.2 Krein conditions

We note that the inequalities $m_\beta \geq 0$ do not hold for general codes. For example, for the code $C = \{000, 001, 011\}$ in Q_3, we have $m_\beta = -\frac{2}{3}$ for $\beta = \vartheta(0,1,3)$ and for $\beta = \vartheta(1,2,3)$.

Let us assume that code C is distance-invariant: there exist integers $a_0, a_1, \ldots, a_n$ such that, for any $c \in C$,

$$|\{c' \in C : \mathrm{dist}(c,c') = i\}| = a_i.$$

Suppose $\beta = \vartheta(i,j,k)$. We know that

$$m_\beta = 2^n u_i(C)^T A_j^* u_k(C) = 4^n \langle u_j(0), u_i(C) \circ u_k(C) \rangle.$$

Since C is distance-invariant and $u_j(C) = \sum_{c \in C} u_j(c)$, we have

$$m_\beta = \frac{2^n}{|C|} \langle u_j(C), u_i(C) \circ u_k(C) \rangle.$$

Definition: Let C be a q-ary code of length n with characteristic vector χ and let E_j denote the j^{th} primitive idempotent of the Hamming scheme $H(n,q)$. The *Krein parameters* $\bar{q}_{i,j}^k(C)$ of code C are given by

$$\bar{q}_{i,j}^k(C) = \|E_k \chi\| \cdot \langle E_k \chi, (E_i \chi) \circ (E_j \chi) \rangle.$$

We say C satisfies the *Krein conditions* provided all of its Krein parameters are non-negative.

For instance, from above, we see that every binary linear code satisfies the Krein conditions. Some completely regular codes fail. An easy example is any completely regular code of covering radius one with $|C| > \frac{1}{2} q^n$. However, if there is a completely regular partition [9, p351] of $H(n,q)$ with all cells having the same parameters of C, then we have a generalized coset graph and the Krein conditions for C follow from those of this graph, which is guaranteed to be distance-regular.

Theorem 5.3.3 *If C is a binary code of length n which is distance-invariant and satisfies the Krein conditions, then $m_\beta \geq 0$ for all compositions β of n.*

5.3.3 Nonlinear codes

Can we say anything more about the "dual" biweight enumerator of a nonlinear binary code C? Assuming C is distance-invariant, we still have some control on the values m_β as the following proposition shows.

Proposition 5.3.4 *For any non-trivial code C with biweight enumerator*

$$\mathcal{W}(y_0, y_1, y_2, y_3) = \sum_\alpha \ell_\alpha \; y_0^{\alpha_0} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3},$$

the transform $\mathcal{W}(H\mathbf{y}) = \sum_\beta m_\beta y_0^{\beta_0} y_1^{\beta_1} y_2^{\beta_2} y_3^{\beta_3}$ satisfies

$$\sum_{i=0}^{n} m_{\vartheta(i,j,k)} \geq 0$$

and

$$\sum_{j=0}^{n} m_{\vartheta(i,j,k)} \geq 0.$$

Note: Since each composition β corresponds to a unique triple (i, j, k), we are summing over all triples (i, j, k) with fixed values of j and k in the first instance and with fixed values of i and k in the second. By symmetry, there is also such an inequality for the sum of m_β where i and j are fixed and k varies.

Proof: Suppose $\beta = \vartheta(i, j, k)$. Then, we have

$$m_\beta = 2^n \chi^T E_i A_j^* E_k \chi. \tag{5.14}$$

Thus we find

$$\sum_{\substack{\beta_2+\beta_3=i \\ \beta_1+\beta_2=k}} m_\beta = \sum_{j=0}^{n} 2^n \chi^T E_i A_j^* E_k \chi = 4^n \chi^T E_i E_0^* E_k \chi$$

which can be written

$$4^n \langle u_i(C), \mathbf{e}_0 \circ u_k(C) \rangle = 4^n \langle \mathbf{e}_0, u_i(C) \circ u_k(C) \rangle.$$

Now if we assume C is distance-invariant, then for $v \in C$ the v-entry of $u_h(C)$ is independent of v. Moreover, since

$$\langle \chi, u_h(C) \rangle = \sum_{i=0}^{n} \langle u_i(C), u_h(C) \rangle = \langle u_h(C), u_h(C) \rangle,$$

this value $(u_h(C))_v$ is never negative. We are assuming $0 \in C$, so

$$\sum_{\substack{\beta_2+\beta_3=i \\ \beta_1+\beta_2=k}} m_\beta = (u_i(C))_0 \cdot (u_k(C))_0 \geq 0$$

with equality if and only if either $u_i(C) = 0$ or $u_k(C) = 0$.

Similarly,

$$\sum_{\substack{\beta_1+\beta_3=j \\ \beta_1+\beta_2=k}} m_\beta \geq 0$$

since

$$\sum_{\beta} m_\beta = \sum_{i=0}^{n} 2^n \chi^T E_i A_j^* E_k \chi = 2^n \chi^T A_j^* E_k \chi.$$

But

$$\chi^T A_j^* E_k \chi = \langle \chi, u_j(0) \circ u_k(C) \rangle = \langle u_j(0), \chi \circ u_k(C) \rangle = \eta \langle u_j(0), \chi \rangle$$

for some $\eta \geq 0$ since C is assumed to be distance-invariant. Finally, we observe that this last expression simplifies to

$$\eta \langle u_j(0), u_j(C) \rangle$$

which is non-negative since $0 \in C$ by hypothesis. □

Chapter 6

The Positive Semidefinite Cone

In this final chapter, I give a summary of work done jointly with Terry Visentin of the University of Winnipeg. This material was written after the publication of A. Schrijver's landmark paper [36]. Schrijver used the representation theory of C^*-algebras to formulate a semidefinite programming problem which extends Delsarte's linear programming bound for binary codes. This was later extended to a method that applies to q-ary codes by Schrijver together with Giswijt and Tanaka in [17]. In this chapter, our approach is again quite naïve, finding all valid inequalities for codes that arise from the positive semidefinite cone of the Terwilliger algebra.

6.1 Positive semidefinite matrices in the Terwilliger algbra

The Terwilliger algebra $\mathbb{T}_n$ can be described as follows. Of course

$$\mathsf{Mat}_2(\mathbb{R})^{\otimes n} \cong \mathsf{Mat}_{2^n}(\mathbb{R})$$

and we treat these as the same algebra. The symmetric group S_n acts on tensor products via

$$(A_1 \otimes \cdots \otimes A_n)^\sigma = A_{\sigma(1)} \otimes \cdots \otimes A_{\sigma(n)}.$$

This extends linearly to $\mathsf{Mat}_2(\mathbb{R})^{\otimes n}$. We may take $\mathbb{T}_n$ to be the subalgebra consisting of those matrices fixed by each σ in S_n.

It is easy to see that a basis matrix L_α with $\alpha = \vartheta(i,j,k)$ is symmetric if and only if $i = k$, i.e., if $\alpha_1 = \alpha_3$. The number of such α is clearly equal to the number of triples $(\gamma_0, \gamma_1, \gamma_2)$ of non-negative integers summing to n in which γ_1 is even. This, in turn, is the same as the number of pairs from an $(n+2)$-element set whose smaller member is odd. This is given by

$$r_n := (n+1) + (n-1) + (n-3) + \cdots = (\lfloor \frac{n}{2} \rfloor + 1)(\lfloor \frac{n+1}{2} \rfloor + 1),$$

or

$$r_n = \begin{cases} \frac{(n+2)^2}{4}, & \text{if } n \text{ even;} \\ \frac{(n+1)(n+3)}{4}, & \text{if } n \text{ odd.} \end{cases}$$

So the subspace of symmetric matrices in $\mathbb{T}_n$ has dimension

$$\frac{1}{2} \cdot \left[\binom{n+3}{3} + r_n \right] = \begin{cases} \frac{(n+2)(n+4)(2n+3)}{24}, & \text{if } n \text{ even;} \\ \frac{(n+1)(n+3)(2n+7)}{24}, & \text{if } n \text{ odd.} \end{cases}$$

It is clear that the positive semidefinite matrices in $\mathbb{T}_n$ form a cone within this subspace of symmetric matrices. For if E and F are positive semidefinite and $c, d \geq 0$, then for any vector x we have $x^T E x \geq 0$ and $x^T F x \geq 0$ giving

$$x^T(cE + dF)x = cx^T Ex + dx^T Fx \geq 0.$$

Henceforth denote this cone by $\mathcal{C}_T$.

Our next goal is to describe the positive semidefinite cone $\mathcal{C}_T$ of the algebra $\mathbb{T}_n$.

6.2 A symmetrized torus

Before looking at the extreme rays of this cone, we explore an interesting subcone with elementary structure.

We can find samples of p.s.d. matrices as follows. Let $u_1, \ldots, u_n$ be unit vectors in $\mathbb{R}^2$. Let $G_i = u_i u_i^T$. Then each G_i is a rank one projection onto the span of u_i and

$$G = G_1 \otimes \cdots \otimes G_n$$

is a rank one projection onto the span of the vector

$$u_1 \otimes \cdots \otimes u_n.$$

Now sum all the matrices in the S_n orbit of G to obtain a positive semidefinite matrix in $\mathbb{T}_n$. Denote the cone generated by these matrices by $\mathcal{C}_0$.

The rank one projections in $\mathsf{Mat}_2(\mathbb{R})$ are naturally topologized as projective one-space, this being homeomorphic to the unit circle. The foregoing discussion leads us to seek out the topological structure of the space

$$(S^1 \times \cdots \times S^1)/S_n.$$

This is described as follows. We consider the torus $\times_1^n S^1$ with its product topology (or, equivalently, the topology induced from the usual topology on $\mathbb{R}^{2^n}$). We have a natural equivalence relation under the coordinatewise action of the symmetric group S_n. We would like to describe the quotient space.

Problem: Determine the homotopy type (or, at least, the homology) of the symmetrized cartesian product $(\times_1^n S^1)/S_n$.

One can easily answer this question in the case $n = 2$. We wish to take a quotient of the ordinary torus $S^1 \times S^1$ over the two element group acting on the coordinates. With a few cut-and-paste diagrams, we convince ourselves that this space is homeomorphic to a Möbius strip. So this subset of the positive semidefinite cone of $\mathbb{T}_2$ is somehow a pointed cone over a Möbius strip embedded in $\mathbb{R}^8$. But this is only $\mathcal{C}_0$; is this the structure of $\mathcal{C}_T$?

It seems possible to describe the topology of this set of matrices $\times_1^n S^1/S_n$ as a CW-complex. There is a single n-cell which can be viewed as the set

$$C_n = \{(\theta_1, \ldots, \theta_n) : 0 < \theta_1 < \cdots < \theta_n < \pi\}.$$

The various faces are obtained by replacing any subset of the strict inequalities "$<$" with weak inequalities "$\leq$". The gluing together of these cells seems complicated.

By way of comparison, the positive semidefinite cone of $\mathsf{Mat}_k(\mathbb{R})$ has each extreme ray generated by a rank one projection operator. These are in one-to-one correspondence with lines through the origin in $\mathbb{R}^k$. So the boundary of the positive semidefinite cone is a pointed cone over projective $(k-1)$-space $\mathbb{P}^{k-1}$.

6.3 Ranks

If $G_1 = u_1 u_1^T$ and $G_2 = u_2 u_2^T$, then

$$G = (G_1 \otimes G_2) + (G_2 \otimes G_1)$$

has eigenvalues

$$0,\ 0,\ 1+\langle u_1,u_2\rangle^2,\ \langle u_1,v\rangle^2$$

where v is a unit vector orthogonal to u_2. So the matrix G has rank two except when $u_1=\pm u_2$, in which case $\operatorname{rank}(G)=1$. In terms of our topological description, the matrices in the interior of the Möbius strip have rank two and those on the boundary have rank one.

If $G_1=G_2=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$ and $G_3=\begin{bmatrix}x^2 & xy\\ xy & y^2\end{bmatrix}$ with $x^2+y^2=1$, then G has eigenvalues

$$0,\ 0,\ 0,\ 0,\ 0,\ 4x^2+2,\ 2y^2,\ 2y^2.$$

If $G_1=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, $G_2=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$, and $G_3=\begin{bmatrix}x^2 & xy\\ xy & y^2\end{bmatrix}$ with $x^2+y^2=1$, then G has eigenvalues

$$0,\ 0,\ 0,\ 2,\ 1+\gamma,\ 1-\gamma$$

(with each of the latter two eigenvalues appearing with multiplicity two) where $\gamma=|x^2-y^2|$.

If, in the above special case, we replace G_2 by $\frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$, then we obtain an 8×8 matrix G with eigenvalues

$$0,\ 0,\ 0,\ 2x^2+(x+y)^2+1,\ \frac{1}{2}(x^2-xy+2y^2+\sqrt{\gamma}),\frac{1}{2}(x^2-xy+2y^2-\sqrt{\gamma})$$

where

$$\gamma=1-2xy|x^2-y^2|.$$

If $G_1=uu^T$, $G_2=vv^T$ and $G_3=ww^T$ where

$$u=\begin{bmatrix}1\\ 0\end{bmatrix},\qquad v=\begin{bmatrix}x_1\\ y_1\end{bmatrix},\qquad w=\begin{bmatrix}x_2\\ y_2\end{bmatrix}$$

with $x_1^2+y_1^2=x_2^2+y_2^2=1$, then G has eigenvalues

$$0,\ 0,\ 0,\ 2+4\langle u,v\rangle\langle u,w\rangle\langle v,w\rangle,$$

$$2-\langle u,v\rangle\langle u,w\rangle\langle v,w\rangle\pm\Gamma$$

(with each of the latter two eigenvalues appearing with multiplicity two) where $\Gamma^2=x_1^4y_2^4+y_1^4x_2^4+y_1^4y_2^4+4x_1y_1^3x_2y_2^3-2x_1^3y_1x_2y_2^3-2x_1y_1^3x_2^3y_2+3x_1^2y_1^2x_2^2y_2^2-x_1^2y_1^2y_2^4-y_1^4x_2^2y_2^2$.

Of course, when $n = 3$, G has trace equal to six.

In general, G has eigenvalue zero with multiplicity at least three, independent of the three vectors $u_1 = [x_1, y_1]$, $u_2 = [x_2, y_2]$ and $u_3 = [x_3, y_3]$. One of the three remaining eigenvalues is

$$\begin{aligned}\theta \;=\;& 2x_1^2y_2^2x_3^2 + 6y_1^2y_2^2y_3^2 + 2y_1^2y_2^2x_3^2 + 4y_1^2x_2y_2x_3y_3 \\ & +2y_1^2x_2^2y_3^2 + 6x_1^2x_2^2x_3^2 + 2x_1^2x_2^2y_3^2 + 2x_1^2y_2^2y_3^2 \\ & +4x_1^2x_2y_2x_3y_3 + 4x_1y_1x_2^2x_3y_3 + 4x_1y_1x_2y_2x_3^2 \\ & +4x_1y_1x_2y_2y_3^2 + 4x_1y_1y_2^2x_3y_3 + 2y_1^2x_2^2x_3^2.\end{aligned}$$

The other two, each having multiplicity two, have less pleasant expressions. The rank is always at most five since we are summing the projections onto six one-dimensional spaces spanned by

$$u_1 \otimes u_2 \otimes u_3,\; u_1 \otimes u_3 \otimes u_2,\; u_2 \otimes u_1 \otimes u_3,\; u_2 \otimes u_3 \otimes u_1,\; u_3 \otimes u_1 \otimes u_2,\; u_3 \otimes u_2 \otimes u_1$$

which always satisfy the relation

$$u_1 \otimes u_2 \otimes u_3 - u_1 \otimes u_3 \otimes u_2 - u_2 \otimes u_1 \otimes u_3 + u_2 \otimes u_3 \otimes u_1 + u_3 \otimes u_1 \otimes u_2 - u_3 \otimes u_2 \otimes u_1 = 0.$$

Let $x_1, \ldots, x_n$ be chosen from the interval $[-1, 1]$ and let $y_i = \pm\sqrt{1 - x_i^2}$. Let $G_i = \begin{bmatrix} x_i^2 & x_iy_i \\ x_iy_i & y_i^2 \end{bmatrix}$ and let

$$G = \sum_{\sigma \in S_n} G_{\sigma(1)} \otimes \cdots \otimes G_{\sigma(n)}.$$

It should be possible to determine the entries of G and the eigenvalues of G as symmetric functions in the x_i. It would also be interesting to determine the rank of G in terms of the relative position of the vectors $\left\{ \begin{bmatrix} x_i \\ y_i \end{bmatrix} : 1 \leq i \leq n \right\}$. In general, the symmetrized matrix $G \in \mathbb{T}_n$ has trace $n!$ since it is the sum of $n!$ matrices each having trace one.

We now give a partial result regarding the rank of G.

Lemma 6.3.1 *For integers k and n, consider the collection of n-fold tensor products*

$$\{v_i := u_i \otimes \cdots \otimes u_i : 1 \leq i \leq k\}$$

where $u_1, \ldots, u_k$ are projectively distinct unit vectors in $\mathbb{R}^2$. The vectors v_i are linearly independent in $\mathbb{R}^{2^n}$ if and only if $k \leq n + 1$.

Proof: A vector of the form $u \otimes \cdots \otimes u$ has at most $n+1$ distinct entries, the duplications being independent of the entries of u. So the vectors $\{v_i : i\}$ all lie in a space of dimension $n+1$. Now assume $k \leq n+1$ and suppose

$$c_1 v_1 + \cdots + c_k v_k = 0$$

in $\mathbb{R}^{2^n}$. If $u_i = [x_i, y_i]$, then we have $n+1$ equations of the form

$$c_1 x_1^j y_1^{n-j} + \cdots + c_k x_k^j y_k^{n-j} = 0.$$

Applying an orthogonal transformation if necessary, we may assume that each $y_i \neq 0$. Then the above equations can be written

$$(c_1 y_1^n) \left(\frac{x_1}{y_1} \right)^j + \cdots + (c_k y_k^n) \left(\frac{x_k}{y_k} \right)^j = 0.$$

We recognize this as a Vandermonde system. Since the unit vectors u_i are projectively distinct, the ratios x_i/y_i are distinct real numbers so the only solution to this system is $c_i y_i^n = 0$ for all i. □

This allows us to obtain an upper bound on the rank of G. Suppose for the moment that the u_i are distinct. Let w_i be a unit vector orthogonal to u_i. Then for any permutation $\sigma \in S_n$,

$$\langle w_i \otimes \cdots \otimes w_i, u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(n)} \rangle = \langle w_i, u_{\sigma(1)} \rangle \cdots \langle w_i, u_{\sigma(n)} \rangle = 0$$

since $\langle w_i, u_i \rangle = 0$. Using the above lemma, this gives us n linearly independent vectors in the null space of G in the case where the u_i are projectively distinct. In the case where the u_i are not distinct, the rank of G will be smaller.

6.4 Inequalities for codes from cone $\mathcal{C}_0$

In order to obtain inequalities for unrestricted codes from the psd matrices studied in the previous section, we need only express each matrix G as a linear combination of the basis elements $L\alpha$.

Let's first look at $n = 3$. Let

$$u_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, u_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, u_3 = \begin{bmatrix} x_3 \\ y_3 \end{bmatrix},$$

let
$$G_1 = u_1 u_1^T, G_2 = u_2 u_2^T, G_3 = u_3 u_3^T,$$
and let
$$G = G_1 \otimes G_2 \otimes G_3 + \cdots + G_3 \otimes G_2 \otimes G_1$$
(six terms). Then
$$G = \sum_\alpha s_\alpha L_\alpha$$
where s_α is a polynomial in $x_1, x_2, x_3, y_1, y_2, y_3$. It is easy to check that $\deg_{x_i} s_\alpha + \deg_{y_i} s_\alpha = 2$ for any i.

We now give the formula for s_α for arbitrary n. For a triple $\mu = (\mu_0, \mu_1, \mu_2)$ of nonnegative integers summing to n, let f_μ denote the monomial symmetric function of shape $1^{\mu_1} 2^{\mu_2}$ in variables $z_1, \ldots, z_n$; e.g., for $\mu = (1, 0, 2)$ and $n = 3$,
$$f_\mu = z_1^2 z_2^2 + z_1^2 z_3^2 + z_2^2 z_3^2.$$
Then, with $\mu_2 = \alpha_2$ and $\mu_1 = \alpha_1 + \alpha_3$,
$$s_\alpha = \mu_0! \mu_1! \mu_2! \cdot \prod_{i=1}^n x_i^2 f_\mu \left(\frac{y_1}{x_1}, \ldots, \frac{y_n}{x_n} \right).$$

The monomial symmetric function in variables $X_1, \ldots, X_n$ with shape $\lambda = (\lambda_1, \ldots, \lambda_k)$ is the polynomial
$$m_\lambda(\underline{\mathrm{X}}) = \sum_{i_1, \ldots, i_k} X_{i_1}^{\lambda_1} \cdots X_{i_k}^{\lambda_k}$$
where the sum is over all injections $\{1, \ldots, k\} \to \{1, \ldots, n\}$ where j is mapped to i_j. By the notation $[2^k 1^\ell]$, we refer to the partition of $m = 2k + \ell$ having k parts equal to two and ℓ parts equal to one. Now define
$$m_\alpha(x_1, \ldots, x_n, y_1, \ldots, y_n) = y_1^2 \cdots y_n^2 \ \cdot \ m_{[2^{\alpha_0} 1^{\alpha_1 + \alpha_3}]} \left(\frac{x_1}{y_1}, \ldots, \frac{x_n}{y_n} \right).$$
Then
$$s_\alpha = m_\alpha.$$

Theorem 6.4.1 *For any real numbers $\theta_1, \ldots, \theta_n$, we have*
$$\sum_\alpha s_\alpha \ell_\alpha \geq 0$$
where $x_i = \cos(\theta_i)$ and $y_i = \sin(\theta_i)$.

Proof: For these values of x_i and y_i, the matrix

$$M = \sum_{\alpha} s_\alpha L_\alpha$$

is a positive semidefinite matrix in $\mathbb{T}_n$. □

6.5 A family of commutative subalgebras

Observe that the (infinite) set of inequalities given by the theorem includes Delsarte's inequalities. For if we take $u_1 = \cdots = u_j = 2^{-1/2}[1,-1]^T$ and $u_{j+1} = \cdots = u_n = 2^{-1/2}[1,1]^T$, then the symmetrized tensor product G is equal to the j^{th} primitive idempotent, E_j.

Likewise, if we choose $u_i = [0,1]^T$ for j values of i and $u_i = [1,0]^T$ for the remaining values, it is easy to see that $G = E_j^*$. So the primitive idempotents of $\mathbb{A}^*$ are also in the cone $\mathcal{C}_0$. In fact, there are an infinite number of commutative subalgebras of $\mathbb{T}_n$, each isomorphic to $\mathbb{A}$, all of whose primitive idempotents belong to $\mathcal{C}_0$.

We consider matrices G coming from the above construction from unit vectors $u_1, \ldots, u_n$ in $\mathbb{R}^2$ where the u_i take on at most two fixed distinct values. That is, for unit vectors $v, w \in \mathbb{R}^2$, consider the symmetrized tensor products

$$F_k = \sum_{\sigma \in S_n} u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(n)}$$

where each $u_i = v$ for $n-k$ values of i and $u_i = w$ for k values of i. Let $\mathcal{U}_2(v,w)$ denote the subspace of $\mathbb{T}_n$ generated by these matrices.

Proposition 6.5.1 *Given the above definitions,*

- *if $v = \pm w$, then $\mathcal{U}_2(v,w)$ is a 1-dimensional subalgebra of $\mathbb{T}_n$;*

- *if $v \perp w$, then $\mathcal{U}_2(v,w)$ is an $(n+1)$-dimensional commutative subalgebra of $\mathbb{T}_n$ isomorphic to the Bose-Mesner algbra $\mathbb{A}$;*

- *If w is neither parallel nor orthogonal to v, then the matrices F_k do not commute and $\mathcal{U}_2(v,w)$ is not closed under multiplication.*

Proof: The first case is trivial. Now consider two symmetrized tensor products

$$F = \sum_{\sigma} u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(n)}$$

and

$$G = \sum_{\pi} v_{\pi(1)} \otimes \cdots \otimes v_{\pi(n)}.$$

We have

$$FG = \sum_{\sigma} \sum_{\pi} (u_{\sigma(1)} v_{\pi(1)}) \otimes \cdots \otimes (u_{\sigma(n)} v_{\pi(1)})$$

and

$$GF = \sum_{\pi} \sum_{\sigma} (v_{\pi(1)} u_{\sigma(1)}) \otimes \cdots \otimes (v_{\pi(1)} u_{\sigma(n)}).$$

Now if $u_i = w$ and $v_j = v$ then

$$(ww^T)(vv^T) = \langle w, v \rangle wv^T$$

while

$$(vv^T)(ww^T) = \langle w, v \rangle vw^T.$$

So F and G commute if and only if v is orthogonal to w or wv^T is a symmetric matrix. But this will happen if and only if $v = \pm w$. □

So, for each pair of orthogonal unit vectors in $\mathbb{R}^2$, we obtain a "copy" of the Bose-Mesner algebra with its primitive idempotents giving new inequalities for the biweight enumerator. This family of subalgebras interpolates between the Bose-Mesner algebra $\mathbb{A}$ and the dual Bose-Mesner algebra $\mathbb{A}^*$.

6.6 Irreducible S_n-modules

We know that every positive semidefinite matrix M inside $\mathbb{T}_n$ is diagonalizable. So we have

$$M = \sum_{\theta} \theta E_{\theta}$$

where the sum is over all (non-negative) eigenvalues θ of M. So the extreme rays of the positive semidefinite cone $\mathcal{C}_T$ are precisely those generated by these projection matrices E_{θ}.

Now since M lies in the commutant algebra of S_n, $E = E_{\theta}$ must be a projection onto an S_n-invariant subspace. Conversely, any such projection operator E lies in the cone $\mathcal{C}_T$. For if $\operatorname{colsp} E$ is S_n-invariant, then $\operatorname{nullsp} E$, its complement, is also S_n-invariant, So $E^\sigma = E$ for each $\sigma \in S_n$ and E belongs to $\mathbb{T}_n$. This proves

Lemma 6.6.1 *The extreme rays of the positive semidefinite cone are precisely the spans of all projection operators onto all irreducible S_n-submodules of $\mathbb{R}^{2^n}$.* □

Now we know the decomposition of $\mathbb{R}^{2^n}$ as an S_n module into irreducibles up to isomorphism. First, the subconstituents $\mathrm{im} E_k^*$ form an orthogonal decomposition. Then we observe that, for each k, we have $E_k^* \mathbb{T} E_k^*$ isomorphic to the Bose-Mesner algebra of the Johnson scheme $J(n,k)$. The action of S_n on this spaces decomposes into $k+1$ mutually non-isomorphic irreducible S_n modules, one for each partition $\lambda = (n-j, j)$ $(0 \le j \le k)$.

Thus, in the overall decomposition, the irreducible S_n module of isomorphism type indexed by the partition $\lambda = (n-j, j)$ appears with multiplicity $n+1-2j$ and has dimension $\binom{n}{j} - \binom{n}{j-1}$. Our interest, however, goes beyond this statistic. We seek expressions for the projections onto *each* of these irreducible S_n modules. For we know that each is expressible as a linear combination of the basis matrices L_α. To achieve this, we recall a bit of representation theory.

Schur's Lemma tells us that an S_n-homomorphism from any irreducible S_n module to any other irreducible is either an isomorphism or the zero map. Moreover, the only S_n-isomorphisms from an irreducible S_n module to itself are the (non-zero) multiples of the identity map. This essentially proves the following

Lemma 6.6.2 *Let the projection operator F be an extremal element of the positive semidefinite cone. Then F is the projection onto some irreducible S_n-submodule of V of isomorpism type $(n-j, j)$, say. This operator is uniquely determined by a set $\tau_j, \tau_{j+1}, \ldots, \tau_{n-j}$ of non-negative scalars satisfying $\sum_{h=j}^{n-j} \binom{n}{h} \tau_h = 1$. Specifically, the h^{th} diagonal block of the projection $E_h^* F E_h^*$ onto the h^{th} subconstituent is $\tau_h F_j^h$ where F_h^j is the j^{th} primitive idempotent in the standard Q-polynomial ordering for the Johnson scheme $J(n,h)$. The rank of F is $\binom{n}{j} - \binom{n}{j-1}$. Conversely, any projection operator onto any irreducible S_n submodule is an extremal element of the positive semidefinite cone.*

Problem: We have all the entries on the block diagonal for F. Find expressions in terms of Hahn polynomials for the remaining entries.

6.7 The actual inequalities revealed

Let us fix a shape $\lambda = (n-g, g)$ denoting the isomorphism type of our irreducible S_n-submodule. From above, we know that this module W projects trivially onto subcontituents $0, 1, \ldots, g-1$ and $n-g+1, \ldots, n$. Let U be a matrix whose columns form an orthonormal basis for W and write

$$U^\top = \left[0|\cdots|0|U_g^\top|\cdots|U_{n-g}^\top|0|\cdots|0\right],$$

where U_k is the submatrix of U whose rows are indexed by tuples of Hamming weight k. Note that U (and, hence, each U_k) has $\mu_g := \binom{n}{g} - \binom{n}{g-1}$ columns.

Initially, let us assume that the projection $E_g^* W$ is non-zero. By Schur's Lemma, it is then an S_n-isomorphism from W to the ${}^{\text{th}}$ eigenspace of the Johnson scheme $J(n, g)$ on the g^{th} subconstituent. So the columns of U_g form a basis for this eigenspace. If we reverse this process and begin with U_g so that $U_g^\top U_g = I_{\mu_g}$ and $U_g U_g^\top = \mathsf{E}_g^{(g)}$, the g^{th} primitive idempotent in the standard Q-polynomial ordering for the Johnson scheme $J(n, g)$.

Now if R is the raising operator for the n-cube, it has natural block decomposition according to Hamming weight and, in this decomposition, the $(j+1, j)$ block of R is the transpose of the incidence matrix $W_{j,j+1}$ of j-subsets versus $(j+1)$-subsets of an n-set (with appropriate choice of labels). Since R commutes with each element of S_n, we see that both $\varphi : W \mapsto E_{j+1}^* W$ and $\psi : W \mapsto E_{j+1}^* R W$ are S_n-homomorphisms of modules. If neither is the zero map, then both are isomorphisms and hence so is the composition $\varphi \circ \psi^{-1}$. It follows that U_{g+h} is a scalar multiple of $W_{g,g+h}^\top U_g$ for each $h = 0, 1, \ldots, n-2g$. (Note that the $(g+h, g)$ block of R^h is a scalar multiple of $W_{g,g+h}^\top$.)

It now follows that, if we fix U_g as above to have columns forming an orthonormal basis for the g^{th} eigenspace of $J(n, g)$, then we can take

$$U^\top = \left[0|\cdots|0|\tau_g U_g^\top|\tau_{g+1} W_{g,g+1}^\top U_g|\cdots|\tau_{n-g} W_{g,n-g}^\top U_g|0|\cdots|0\right]$$

where $\tau_g, \tau_{g+1}, \ldots, \tau_{n-g}$ are real scalars satisfying a single overall constraint

$$\sum_{k=g}^{n-g} \binom{n-2g}{k-g} \tau_k^2 = 1.$$

This requires some proof, but that will appear elsewhere.

This is all part of ongoing work with Terry Visentin at the University of Winnipeg. The project started with my observations about the Terwilliger algebra and the biweight enumerator in 2000-2001, but the material here is influenced by A. Schrijver's now-famous paper. Here is my attempt to summarize our results.

1. The Terwilliger algebra $\mathbb{T} = \mathbb{T}_n$ is the commutant algebra of the S_n action on the standard module $V = \mathbb{C}^X$

2. Define $\mathbb{T}_{sym}$ to be the subspace of symmetric matrices in $\mathbb{T}$. Note that this is not a subalgebra

3. If E is a positive semidefinite matrix in $\mathbb{T}_{sym}$ with spectral decomposition $E = \sum_\theta \theta F_\theta$, then each $\theta \geq 0$ and each corresponding projection matrix F_θ lies in $\mathbb{T}_{sym}$

4. If F is a matrix representing orthogonal projection onto the real subspace W then $F \in \mathbb{T}_{sym}$ if and only if W is S_n-invariant

5. If $W = W_1 \oplus W_2$ is a decompositon of S_n-submodules of V, then the corresponding orthogonal projection F onto W can be expressed as $F = F_1 + F_2$ where F_i represents orthogonal projection onto the S_n-submodule W_i

6. Define the *positive semidefinite cone* $\mathcal{C}_\mathbb{T}$ to be the set of all positive semidefinite matrices in $\mathbb{T}_{sym}$. This is closed under addition and under multiplication by nonnegative scalars

7. THEOREM: Every E in $\mathcal{C}_\mathbb{T}$ is uniquely expressible as a nonnegative linear combination $E = \sum_i \theta_i F_i$ where the eigenvalues θ_i are not necessarily distinct but each F_i is a matrix representing orthogonal projection onto some irreducible S_n-submodule of V

8. Let W be an irreducible S_n-submodule with orthonormal basis $\{u_1, \ldots, u_d\}$ ($d = \dim W$). Then the matrix $F \in \mathbb{T}_{sym}$ representing orthogonal projection onto W is $F = UU^\top$ where U has i^{th} column u_i

9. The submodule $E_k^* V$ is isomorphic to the standard module of the Johnson graph $J(n,k)$ and the subalgebra $E_k^* \mathbb{T} E_k^*$ is isomorphic to the Terwilliger algebra of the Johnson graph, the first and second isomorphisms being consistent (this can be made precise)

10. Each eigenspace of $J(n,k)$ is an irreducible S_n-submodule. There is exactly one submodule of each isomorphism type (n), $(n-1,1)$, $\ldots$, $(n-k,k)$ for $k \leq n/2$. For $k > n/2$, the types are $(n-\lambda,\lambda)$ for $\lambda = 0,\ldots,n-k$ where, for convenience, we write $(n-0,0)$ for the trivial partition (n)

11. So the submodule $E_k^* V$ decomposes uniquely into $\min(k+1, n-k+1)$ irreducible S_n-submodules, one of each isomorphism type $(n-\lambda,\lambda)$ for each $\lambda = 0,1,\ldots,\min(k,n-k)$

12. So the standard module V decomposes into

$$V = \bigoplus_{\lambda=0}^{\lfloor n/2 \rfloor} \bigoplus_{h=\lambda}^{n-\lambda} W_h^{(n-\lambda,\lambda)}$$

where each $W_h^{(n-\lambda,\lambda)}$ is an irreducible S_n-submodule of isomorphism type $(n-\lambda,\lambda)$ (i.e., this module has multiplicity $n+1-2\lambda$)

13. The irreducible S_n module of type $(n-\lambda,\lambda)$ appearing in the decomposition of $E_k^* V$ is precisely the λ eigenspace (in the natural or Q-polynomial ordering) of the Johnson scheme $J(n,k)$ with projector

$$F = \begin{bmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & \mathsf{E}_\lambda^{(k)} & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

where $\mathsf{E}_\lambda^{(k)}$ is the primitive idempotent of the Bose-Mesner algebra of the Johnson scheme $J(n,k)$ of index λ in the natural ordering. So, in particular, $\dim W = \binom{n}{\lambda} - \binom{n}{\lambda-1}$ for this module. (NOTATION: Symbols in sans serif font correspond to matrices and parameters for the Johnson schemes $J(n,k)$. These are defined on the various subconstituents of the n-cube, so we must distinguish them from the corresponding quantities for Q_n)

14. Now fix $\lambda > 0$ and consider an irreducible S_n module W of isomorphism type $(n-\lambda, \lambda)$. Then $W \mapsto E_k^* W$ is either the zero map or an isomorphism of S_n modules by Schur's Lemma

15. Next consider the raising operator R for the n-cube. If the partial order of the Boolean lattice is extended to a total order, then this can be viewed as the lower-triangular portion of the adjacency matrix of Q_n. More formally, we can write
$$R = \sum_{i=0}^{n-1} E_{i+1}^* A E_i^*$$
where the $(k+1, k)$ block (in the natural partition of rows and columns according to Hamming weight) is the incidence matrix ${W_{k,k+1}^{(n)}}^\top$ (or set-inclusion matrix) of k-subsets versus $(k+1)$-subsets. Clearly, R belongs to the Terwilliger algebra

16. Since both $W \mapsto^{\varphi} E_{k+1}^* W$ and $W \mapsto^{\psi} RE_k^* W$ are zero maps or both are isomorphisms, in the latter case the composition $\psi\varphi^{-1}$ is an isomorphism as well

17. First, let us assume that all projections $E_k^* W$ are non-zero for $k = \lambda, \ldots, n-\lambda$. (The other cases will fall out as degenerate specializations of this.) We know that $\dim W = \binom{n}{\lambda} - \binom{n}{\lambda-1}$. Let U_λ be an $\binom{n}{\lambda} \times \dim W$ matrix whose columns form an orthonormal basis for $E_\lambda^* W$. We claim that $U_k = {W_{\lambda,k}^{(n)}}^\top U_\lambda$ has columns forming an orthogonal basis for $E_k^* W$

18. Indeed, since $U_\lambda^\top U_\lambda = I$, we have
$$U_k^\top U_k = U_\lambda^\top \left(W_{\lambda,k} W_{\lambda,k}^\top \right) U_\lambda$$
where we have omitted the superscript (n) on $W_{\lambda,k}$ for readability. But the matrix in the middle lies in the Bose-Mesner algebra of the Johnson scheme $J(n, \lambda)$. It is easy to check that
$$W_{\lambda,k} W_{\lambda,k}^\top = \sum_{h=0}^{\lambda} \binom{n-\lambda-h}{k-\lambda-h} \mathsf{A}_h^{(\lambda)}$$
where $\mathsf{A}_h^{(\lambda)}$ is the adjacency matrix of the distance-h relation in the Johnson graph $J(n, \lambda)$. Moreover, each column of U_λ is an eigenvector for $\mathsf{A}_h^{(\lambda)}$ with eigenvalue $(-1)^h \binom{\lambda}{h}$

19. By the way, let's write down all the eigenvalues of the Johnson scheme $J(n,\lambda)$. We have
$$\mathsf{A}_h^{(\lambda)}\mathsf{E}_j^{(\lambda)} = \mathsf{P}_{j,h}^{(\lambda)}\mathsf{E}_j^{(\lambda)}$$
where
$$\mathsf{P}_{j,h}^{(\lambda)} = \sum_{\ell=0}^{h}(-1)^\ell\binom{j}{\ell}\binom{\lambda-j}{h-\ell}\binom{n-\lambda-j}{h-\ell}$$

20. Now back to our proof. We now have
$$\begin{aligned} U_k^\top U_k &= \sum_{h=0}^{\lambda}\binom{n-\lambda-h}{k-\lambda-h}U_\lambda^\top \mathsf{A}_h^{(\lambda)}U_\lambda \\ &= \sum_{h=0}^{\lambda}\binom{n-\lambda-h}{k-\lambda-h}\mathsf{P}_{\lambda,h}^{(\lambda)}U_\lambda^\top U_\lambda \\ &= \left[\sum_{h=0}^{\lambda}(-1)^h\binom{\lambda}{h}\binom{n-\lambda-h}{k-\lambda-h}\right]I \end{aligned}$$
So the columns of U_k are indeed pairwise orthogonal, but this also shows that they all have the same norm

21. It now follows that, up to scalar, the columns of U_k form an orthonormal basis for the λ eigensapce of $J(n,k)$. Hence, by Schur's Lemma, they also form an orthonormal basis for E_k^*W provided it is non-zero

22. So for each W of this isomorphism type, there exist scalars τ_k ($\lambda \le k \le n-\lambda$) such that the matrix U defined by
$$U^\top = \left[0|\cdots|0|\tau_\lambda U_\lambda^\top|\tau_{\lambda+1}U_{\lambda+1}^\top|\cdots|\tau_{n-\lambda}U_{n-\lambda}^\top|0|\cdots|0\right]$$
has columns forming an orthonormal basis for W

23. The normaliztion factor here is $\sum_{k=\lambda}^{n-\lambda}\binom{n-2\lambda}{k-\lambda}\tau_k^2 = 1$. (This makes the columns of U orthonormal.) Conversely, for any such scalars τ_k, the column space of the matrix U determined above is an irreducible S_n module of isomorphism type $(n-\lambda,\lambda)$

24. For a shape $\alpha = (\alpha_0,\alpha_1,\alpha_2,\alpha_3)$ (with all $\alpha_i \ge 0$ and summing to n), define
$$L_\alpha = E_i^* A_j E_k^*$$

in $\mathbb{T}$ where $i = \alpha_2 + \alpha_3$, $j = \alpha_1 + \alpha_3$ and $k = \alpha_1 + \alpha_2$. It is well-known that the L_α form a basis for $\mathbb{T}$ (and the appropriate pairwise sums form a basis for $\mathbb{T}_{sym}$). Our last step is to compute coefficients t_α satisfying

$$F = UU^\top = \sum_\alpha t_\alpha L_\alpha.$$

Note that we will use, in our optimization problem, $\ell_\alpha = x^\top L_\alpha x$ where

$$x = \frac{1}{|C|} \sum_{a \in C} \chi_{a+C}$$

and χ_S denotes the characteristic vector of the subset $S \subset \{0,1\}^n$. We now have the inequality $x^\top F x \geq 0$ when F is positive semidefinite and we want to translate this into a condition $\sum_\alpha t_\alpha \ell_\alpha \geq 0$ on the unknowns ℓ_α

25. We return to our fixed projector F with fixed shape $(n-\lambda, \lambda)$. If we view this as a block matrix with (j, ℓ) block being the submatrix obtained by restricting to rows indexed by words of Hamming weight j and to columns indexed by words of Hamming weight ℓ, then the (j, ℓ)-block of F is zero if either $j < \lambda$, $\ell < \lambda$, $j > n - \lambda$ or $\ell > n - \lambda$. In all other cases, the (j, ℓ)-block of F is given by

$$U_j U_\ell^\top = W_{\lambda,j}^\top U_\lambda U_\lambda^\top W_{\lambda,\ell} = W_{\lambda,j}^\top \mathsf{E}_\lambda^{(\lambda)} W_{\lambda,\ell}$$

26. If $\mathrm{wt}(x) = j$ and $\mathrm{wt}(y) = \ell$ and the (j, ℓ)-block is not zero, then

$$F_{xy} = \sum_{\substack{\mathrm{wt}(z_1)=\lambda \\ \mathrm{supp}(z_1)\subseteq \mathrm{supp}(x)}} \; \sum_{\substack{\mathrm{wt}(z_2)=\lambda \\ \mathrm{supp}(z_2)\subseteq \mathrm{supp}(y)}} \left(\mathsf{E}_\lambda^{(\lambda)}\right)_{z_1,z_2}.$$

If we also know that $\mathrm{dist}(x, y) = k$, then we have (with $a = \frac{1}{2}(j+k-\ell)$) $F_{xy} = \omega_{j,k,\ell} :=$

$$\frac{n-2\lambda+1}{n-\lambda+1} \sum_{r=0}^{\lambda} \left[\sum_{s=0}^{r} \binom{j-a}{\lambda-r} \binom{j-a-\lambda+r}{s} \binom{a}{r-s} \binom{\ell-\lambda+r-s}{r} \right].$$

$$\mathsf{P}_{\lambda,r}^{(\lambda)} \binom{\lambda}{r}^{-1} \binom{n-\lambda}{r}^{-1}$$

since, for $\mathrm{dist}_{J(n,\lambda)}(z_1, z_2) = r$ (distance in the Johnson graph),

$$\left(\mathsf{E}_\lambda^{(\lambda)}\right)_{z_1,z_2} = \mathsf{P}_{\lambda,r}^{(\lambda)} \frac{\binom{n}{\lambda} - \binom{n}{\lambda-1}}{\binom{n}{\lambda}\binom{\lambda}{r}\binom{n-\lambda}{r}}$$

27. Now F is a symmetric matrix and $L_\alpha^\top = L_\beta$ for some β, namely $\beta = (\alpha_0, \alpha_3, \alpha_2, \alpha_1)$, so these two matrices have the same coefficient in the expansion of F: $t_\alpha = t_\beta$

28. I think we now have all the ingredients to prove that, for each λ, $0 < \lambda \leq \frac{n}{2}$, and for each choice of scalars $\tau_\lambda, \ldots, \tau_{n-\lambda}$ satisfying

$$\sum_{k=\lambda}^{n-\lambda} \binom{n-2\lambda}{k-\lambda} \tau_k^2 = 1$$

we have a valid inequality

$$\sum_\alpha \tau_{\alpha_2+\alpha_3} \cdot \tau_{\alpha_1+\alpha_2} \cdot \omega_\alpha \cdot \ell_\alpha \geq 0$$

for the parameters ℓ_α of any code C where

$$\omega_\alpha = \frac{n-2\lambda+1}{n-\lambda+1} \sum_{r=0}^{\lambda} \sum_{s=0}^{r} (-1)^r \binom{\alpha_2}{\lambda - r} \cdot$$

$$\binom{\alpha_2 - \lambda + r}{s} \binom{\alpha_3}{r-s} \binom{\alpha_1 + \alpha_2 - \lambda + r - s}{r} \binom{n-\lambda}{r}^{-1}$$

since $\mathsf{P}_{\lambda,r}^{(\lambda)} = (-1)^r \binom{\lambda}{r}$ and $a = \frac{1}{2}(j + k - \ell) = \alpha_3$

29. The idea now is to iteratively optimize, beginning with Delsarte's linear program and using Lagrange multipliers at each stage to choose the best direction $\{\tau_k\}_{k=\lambda}^{n-\lambda}$ for each isomorpism type $(n - \lambda, \lambda)$ and include the corresponding $\lfloor n/2 \rfloor - 1$ inequalities into our system and re-solve.

This ends my terse summary of our approach to bounds for the "biweight distribution" $\{\ell_\alpha\}$. We never planned to use semidefinite programming; we simply iterate through larger and larger linear programming problems allowing the Lagrange multipliers to tell us which new inequalities to bring in to our finite set at each stage.

Chapter 7

Semidefinite Programming

This chapter is extracted from a recent survey paper [32] co-authored with Hajime Tanaka.

7.1 The semidefinite programming bound

Throughout this section, suppose that $(X, \mathcal{R})$ is the binary Hamming scheme $H(n,2) = (S_2 \wr S_n)/S_n$, so that $X = \mathcal{Q}^n$ where $\mathcal{Q} = \{0,1\}$. Let $x = (0,0,\ldots,0)$ be the zero vector and write $\mathbb{T} = \mathbb{T}(x)$, $E_i^* = E_i^*(x)$ $(0 \leq i \leq n)$. Recall that $\mathbb{T}$ coincides with the centralizer algebra of $K = S_n$ acting on X.

Let $Y \subseteq X$ be a code. We consider two subsets Π_1, Π_2 of $G = S_2 \wr S_n$ defined by $\Pi_1 = \{g \in G : x \in gY\}$, $\Pi_2 = \{g \in G : x \notin gY\}$. For $i \in \{1,2\}$, let

$$M_{\mathrm{SDP}}^i = \frac{1}{|Y|n!} \sum_{g \in \Pi_i} \chi_{gY} (\chi_{gY})^{\mathsf{T}} \in \mathbb{C}^{X \times X}$$

where $\chi_{gY} \in \mathbb{C}^X$ denotes the (column) characteristic vector of gY. Since Π_1, Π_2 are unions of right cosets of G by K, it follows that $M_{\mathrm{SDP}}^1, M_{\mathrm{SDP}}^2 \in \mathbb{T}$. Moreover, since the $\chi_{gY}(\chi_{gY})^{\mathsf{T}}$ are nonnegative and positive semidefinite, so are $M_{\mathrm{SDP}}^1, M_{\mathrm{SDP}}^2$. By computing the inner products with the 01-matrices $E_i^* A_j E_k^*$, we readily obtain

$$M_{\mathrm{SDP}}^1 = \sum_{i,j,k} \lambda_{ijk} E_i^* A_j E_k^*, \quad M_{\mathrm{SDP}}^2 = \sum_{i,j,k} (\lambda_{0jj} - \lambda_{ijk}) E_i^* A_j E_k^*,$$

where

$$\lambda_{ijk} = \frac{|X|}{|Y|} \cdot \frac{|\{(y,y',y'') \in Y^3 : (y,y',y'') \text{ satisfies } (*)\}|}{|\{(y,y',y'') \in X^3 : (y,y',y'') \text{ satisfies } (*)\}|},$$

and condition (*) is defined by

$$(y, y') \in R_i, \quad (y', y'') \in R_j, \quad (y'', y) \in R_k. \tag{*}$$

By viewing the λ_{ijk} as variables we get the following *semidefinite programming* (or *SDP*) *bound* established by A. Schrijver:

Theorem 7.1.1 ([36]) *Set*

$$\ell_{\rm SDP} = \ell_{\rm SDP}(n, \delta) = \max \sum_{i=0}^{n} \binom{n}{i} \lambda_{0ii}$$

subject to (i) $\lambda_{000} = 1$*; (ii)* $0 \le \lambda_{ijk} \le \lambda_{0jj}$*; (iii)* $\lambda_{ijk} = \lambda_{i'j'k'}$ *if* (i', j', k') *is a permutation of* (i, j, k)*; (iv)* $\sum_{i,j,k} \lambda_{ijk} E_i^* A_j E_k^* \succcurlyeq 0$*; (v)* $\sum_{i,j,k} (\lambda_{0jj} - \lambda_{ijk}) E_i^* A_j E_k^* \succcurlyeq 0$*; (vi)* $\lambda_{ijk} = 0$ *if* $\{i, j, k\} \cap \{1, 2, \ldots, \delta - 1\} \neq \emptyset$ *(where* $\succcurlyeq$ *means positive semidefinite). Then* $A_2(n, \delta) \le \ell_{\rm SDP}$*.*

It is known that semidefinite programs can be approximated in polynomial time within any specified accuracy by interior-point methods; see [40]. See also [16, §7.2] for a discussion on how to ensure that computational solutions do give valid upper bounds on $A_2(n, \delta)$. While Delsarte's LP bound is a close variant of Lovász's ϑ-bound, Schrijver's SDP bound can be viewed as a variant of an extension of the ϑ-bound based on "matrix cuts" [29]; see also [16, Chapter 6]. In fact, if we define $\mathbf{a} = (a_0, a_1, \ldots, a_n)$ by $a_i = \lambda_{0ii} \binom{n}{i}$ $(0 \le i \le n)$, then the condition that $\mathbf{a}Q$ is nonnegative is equivalent to the positive semidefiniteness of the matrix

$$M_{\rm LP} = \sum_{i=0}^{n} \lambda_{0ii} A_i,$$

but this is in turn a consequence of the positive semidefiniteness of $M^1_{\rm SDP}$ and $M^2_{\rm SDP}$. A hierarchy of upper bounds based on semidefinite programming was later proposed in [24]:

$$\ell_+^{(1)} \ge \ell_+^{(2)} \ge \cdots \ge \ell_+^{(k)} \ge \cdots \ge A_2(n, \delta).$$

It turns out that $\ell_{\rm LP} = \ell_+^{(1)} \ge \ell_{\rm SDP} \ge \ell_+^{(2)}$. Each of the $\ell_+^{(k)}$ can be computed in time polynomial in n, but the program defining $\ell_+^{(2)}$ already contains $O(n^7)$ variables. Two strengthenings of $\ell_{\rm SDP}$ with the same complexity are also given in [24].

The SDP bound was also applied to the problem of finding the stability number of the graph $(X, R_{n/2})$ for even n (known as the *orthogonality graph*) in [22], where it is shown (among other results) that for $n = 16$ the SDP bound gives the exact value 2304, whereas the LP bound only gives much weaker upper bound 4096. This problem arises in connection with quantum information theory [14]; see also [18].

As $M^1_{\mathrm{SDP}}, M^2_{\mathrm{SDP}}$ are $2^n \times 2^n$ matrices, it is in fact absolutely necessary to simplify the program by explicitly describing the Wedderburn decomposition of the semisimple algebra $\mathbb{T}$. The decomposition of $\mathbb{T}$ (as a centralizer algebra) was worked out in [12] in the study of addition theorems for Krawtchouk polynomials, but our discussion below emphasizes the use of $\mathbb{T}$, based on [21].

Let $W \subseteq \mathbb{C}^X$ be an irreducible $\mathbb{T}$-module with endpoint r. Then W has dual endpoint r, and there is a basis $\{w_i\}_{i=r}^{n-r}$ for W such that

$$w_i \in E_i^* W, \quad A_1 w_i = (i - r + 1) w_{i+1} + (n - r - i + 1) w_{i-1} \quad (r \le i \le n - r)$$

where $w_{r-1} = w_{n-r+1} = 0$. Thus, the isomorphism class of W is determined by r. Moreover, it follows that

$$\langle w_i, w_j \rangle = \delta_{ij} \binom{n-2r}{i-r} ||w_r||^2 \quad (r \le i, j \le n - r).$$

See [21] for the details. The actions of the A_i on W may be described from the above information as the A_i are Krawtchouk polynomials in A_1, but our argument goes as follows. For integers i, k, t such that $0 \le k \le i \le n$ and $0 \le t \le \min\{k, n-i\}$, we recall the following normalization of the dual Hahn polynomials found in [10]:

$$\binom{i}{k} Q_t^{i,k}(\lambda^k(z)) = \binom{i}{k-t}\binom{n-i}{t} {}_3F_2\left(\begin{matrix} -t, -z, z-n-1 \\ i-n, -k \end{matrix} \middle| \, 1 \right),$$

where $\lambda^k(z) = k(n-k) - z(n+1-z)$. If $i + j + k$ is odd then $E_i^* A_j E_k^* = 0$ since $H(n, 2)$ is bipartite, so suppose that $i + j + k$ is even. Then it follows that

$$E_i^* A_j E_k^* A_2 E_k^* = \beta_{j+2}^{i,k} E_i^* A_{j+2} E_k^* + \alpha_j^{i,k} E_i^* A_j E_k^* + \gamma_{j-2}^{i,k} E_i^* A_{j-2} E_k^*,$$

where $\beta_{j+2}^{i,k} = (t+1)(i+1-k+t)$, $\alpha_j^{i,k} = (k-t)(i-k+t) + t(n-i-t)$ and $\gamma_{j-2}^{i,k} = (k+1-t)(n+1-i-t)$, with $t = (j+k-i)/2$. Using $2A_2 = A_1^2 - nI$

we find $E_k^* A_2 w_k = \lambda^k(r) w_k$ $(r \leq k \leq n-r)$. Combining these facts with the three-term recurrence relation for the $Q_t^{i,k}$ [10, Theorem 3.1], we obtain

$$E_i^* A_j w_k = Q_t^{i,k}(\lambda^k(r)) E_i^* A_{i-k} w_k = Q_t^{i,k}(\lambda^k(r)) \binom{i-r}{i-k} w_i$$

for $r \leq k \leq i \leq n-r$, $0 \leq j \leq n$ such that $i+j+k$ is even, where $t = (j+k-i)/2$. (The $Q_t^{i,k}$ for $t > \min\{k, n-i\}$ are formally defined by the recurrence relation [10, Theorem 3.1].) Hence, after orthonormalization of the w_i, we get the following algebra isomorphism which preserves the positive-semidefinite cones:

$$\varphi : \mathbb{T} \to \bigoplus_{r=0}^{\lfloor n/2 \rfloor} \mathbb{C}^{(n-2r+1)\times(n-2r+1)}$$

where the r^{th} block of $\varphi(A_j)$ is the symmetric matrix $(a_{i,k}^{j,r})_{i,k=r}^{n-r}$ given by

$$a_{i,k}^{j,r} = a_{k,i}^{j,r} = \begin{cases} Q_{(j+k-i)/2}^{i,k}(\lambda^k(r)) \binom{i-r}{i-k} \binom{n-2r}{i-r}^{1/2} \binom{n-2r}{k-r}^{-1/2} & \text{if } i+j+k \text{ even,} \\ 0 & \text{if } i+j+k \text{ odd,} \end{cases}$$

for $r \leq k \leq i \leq n-r$, $0 \leq j \leq n$. See also [36, 42].

The SDP bound has also been formulated for binary constant weight codes (i.e., codes in $J(v,n)$) in [36] and for nonbinary codes in [17, 16]. The description of the irreducible $\mathbb{T}$-modules becomes more complicated in this case, but this method turns out to improve the LP bound for many parameters. It seems to be an important problem to decide whether it is possible or not to establish a suitable SDP bound for t-designs in $J(v,n)$ or $H(n,q)$. The SDP bound for spherical codes was formulated in [4]; it provides a new proof of $k(3) = 12$ and $k(4) = 24$. See also [6].

Bibliography

[1] A.Ashikhmin and A. Barg. Binomial moments of the distance distribution and the probability of undetected error. *Designs, Codes Crypt.* **16** (1999), 103–116.

[2] C. Bachoc, On harmonic weight enumerators of binary codes, *Des. Codes Cryptogr.* **18** (1999), 11–28.

[3] C. Bachoc, Harmonic weight enumerators of nonbinary codes and MacWilliams identities, in: A. Barg and S. Litsyn (Eds.), *Codes and Association Schemes,* American Mathematical Society, Providence, RI, 2001, pp. 1–23.

[4] C. Bachoc and F. Vallentin, New upper bounds for kissing numbers from semidefinite programming, *J. Amer. Math. Soc.* **21** (2008), 909–924; arXiv:`http://arxiv.org/abs/math/0608426`math/0608426.

[5] E. Bannai and T. Ito, Algebraic combinatorics I: Association schemes, Benjamin/Cummings, Menlo Park, CA, 1984.

[6] E. Bannai and E. Bannai, A survey on spherical designs and algebraic combinatorics on spheres. *Europ. J. Combin.* **30** no. 6 (2009), 1392–1425.

[7] A. Barg and D. B. Jaffe. Numerical results on the asymptotic rate of binary codes. *Preprint*, (2000).

[8] A. Barg and S. Litsyn (eds). *Codes and Association Schemes (New Brunswick, NJ, 1999).* DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Amer. Math. Soc., Providence, RI, in press.

[9] A. E. Brouwer, A. M. Cohen and A. Neumaier. Distance-Regular Graphs. Springer-Verlag, Berlin (1989).

[10] A. R. Calderbank and P. Delsarte, Extending the t-design concept, *Trans. Amer. Math. Soc.* **338** (1993), 941–952.

[11] P. Delsarte. An algebraic approach to the association schemes of coding theory. *Philips Res. Repts. Suppl.* **10** (1973).

[12] C. F. Dunkl, A Krawtchouk polynomial addition theorem and wreath products of symmetric groups, *Indiana Univ. Math. J.* **25** (1976), 335–358.

[13] P. Frankl and R. M. Wilson. The Erdös-Ko-Rado theorem for vector spaces. *J. Combin. Theory Ser. A* **43** (1986), 228–236.

[14] V. Galliard, S. Wolf and A. Tapp, The impossibility of pseudo-telepathy without quantum entanglement, in: *Proceedings, IEEE International Symposium on Information Theory*, Yokohama, Japan, 2003, p. 457; arXiv:`http://arxiv.org/abs/quant-ph/0211011`quant-ph/0211011.

[15] E. N. Gilbert. A comparison of signalling alphabets. *Bell System Tech. J.* **31** (1952), 504–522.

[16] D. Gijswijt, Matrix algebras and semidefinite programming techniques for codes, Ph.D. Thesis, The Universiteit van Amsterdam, 2005.

[17] D. Gijswijt, A. Schrijver and H. Tanaka. New upper bounds for nonbinary codes based on the Terwilliger algebra and semidefinite programming. *J. Combin. Theory Ser. A* **113**, no. 8 (2006), 1719–1731.

[18] C. D. Godsil and M. W. Newman, Coloring an orthogonality graph, *SIAM J. Discrete Math.* **22** (2008), 683–692; arXiv:`http://arxiv.org/abs/math/0509151`math/0509151.

[19] D. B. Jaffe. A brief tour of split linear programming. p164–173 in: *Applied algebra, algebraic algorithms and error-correcting codes (Toulouse, 1997)*, Lecture Notes in Comput. Sci. **1255**, Springer, Berlin, 1997.

[20] D. B. Jaffe. Optimal binary linear codes of length ≤ 30. *Discrete Math.* **223** Issue 1-3 (2000), 135–155.

[21] J. T. Go. The Terwilliger algebra of the hypercube. *European J. Combin.* **23**, no. 4 (2002), 399–429.

[22] E. de Klerk and D. V. Pasechnik, A note on the stability number of an orthogonality graph, *Europ. J. Combin.* **28** (2007), 1971–1979; arXiv:`http://arxiv.org/abs/math/0505038`math/0505038.

[23] E. Knill and R. Laflamme. A theory of quantum error-correcting codes. *Physical Review A* **55** (1997), 900-911.

[24] M. Laurent, Strengthened semidefinite programming bounds for codes, *Math. Program.* **109** Ser. B, (2007), 239–261.

[25] V. I. Levenshtein. On the minimal redundancy of binary error-correcting codes. (Translated by A. M. Odlyzko)

[26] J. H. van Lint, Introduction to Coding Theory (Third edition), Springer-Verlag, Berlin, 1998.

[27] J. H. van Lint and R. M. Wilson, A Course in Combinatorics (Second edition), Cambridge University Press, Cambridge, 2001.

[28] L. Lovász, On the Shannon capacity of a graph, *IEEE Trans. Inform. Th.* **25** (1979), 1–7.

[29] L. Lovász and A. Schrijver, Cones of matrices and set-functions and 0–1 optimization, *SIAM J. Optim.* **1** (1991), 166–190.

[30] F. J. MacWilliams, C. L. Mallows and N. J. A. Sloane. Generalizations of Gleason's theorem on weight enumerators of self-dual codes. *IEEE Trans. Info. Th.* **IT-18** (1972), 794–805.

[31] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. Elsevier-North Holland, Amsterdam, 1977.

[32] W.J. Martin and H. Tanaka. Commutative association schemes. *Europ. J. Combin.* **30** no. 6 (2009), 1497–1525.

[33] R. J. McEliece, E. R. Rodemich, H. Rumsey Jr. and L. R. Welch. New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities. *IEEE Trans. Info. Theory* **IT-23** (1977), 157–166.

[34] E. M. Rains. Quantum weight enumerators. *IEEE Trans. Inform. Theory*, **44** (1998), 1388–1394.

[35] A. Samorodnitsky. On the optimum of Delsarte's linear program (1998), *J. Combin. Theory Ser. A* **96**, no. 2 (2001), 261–287.

[36] A. Schrijver. New code upper bounds from the Terwilliger algebra and semidefinite programming. *IEEE Trans. Inform. Theory* **51**, no. 8 (2005), 2859–2866.

[37] P. Terwilliger. The subconstituent algebra of an association scheme. I. J. Algebraic Combin. **1** (1992), no. 4, 363–388.

[38] P. Terwilliger. The subconstituent algebra of an association scheme. II. J. Algebraic Combin. **2** (1993), no. 1, 73–103.

[39] P. Terwilliger. The subconstituent algebra of an association scheme. III. J. Algebraic Combin. **2** (1993), no. 2, 177–210.

[40] M. J. Todd, Semidefinite optimization, *Acta Numer.* **10** (2001), 515–560.

[41] M. A. Tsfasman, S. G. Vladut and T. Zink. Modular curves, Shimura curves, and Goppa Codes, better than the Varshamov-Gilbert bound.

[42] F. Vallentin, Symmetry in semidefinite programs, *Linear Algebra Appl.* **430** (2009), 360–369; arXiv:`http://arxiv.org/abs/0706.4233`0706.4233.

[43] R. R. Varshamov. Estimate of the number of signals in error correcting codes. *Dokl. Akad. Nauk SSSR* **117** (1957), 739–741. (In Russian.)

Appendix A

Selected Background from Linear Algebra and Graph Theory

This small appendix has been added as a place to put background material that may or may not be known to the reader and is not central to the topic of discussion.

A.1 Linear Algebra and Matrix Theory

Here, we briefly review a few facts from linear algebra.

The *transpose* of matrix $A = [a_{ij}]$ with entry a_{ij} in row i, column j is the matrix $A^\top$ with a_{ij} in row j column i. A matrix A is *symmetric* if it is equal to its transpose: $A^\top = A$. The real $n \times n$ matrices

$$S\mathbb{R}^{n\times n} = \left\{ A \in \mathsf{Mat}_n(\mathbb{R}) : A^\top = A \right\}$$

form a vector space of dimension $\binom{n+1}{2}$. The product of two symmetric matrices is symmetric if and only if they commute (exercise).

If A is an $n \times n$ matrix over the complex numbers $\mathbb{C}$ and $\mathbf{v}$ is a non-zero vector in $\mathbb{C}^n$, then $\mathbf{v}$ is an *eigenvector* for A if there is some scalar $\lambda \in \mathbb{C}$ satisfying $A\mathbf{v} = \lambda\mathbf{v}$. When a non-zero vector $\mathbf{v}$ satisfying this equation exists, we call λ an *eigenvalue* of A.

Let $\langle \cdot, \cdot \rangle$ denote the standard Hermitean inner product on $\mathbb{C}^n$: $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^\top \bar{\mathbf{v}}$ where $\top$ denotes the transpose and $\bar{\mathbf{v}}$ is the entrywise conjugate of vector

$\mathbf{v}$.

The eigenvalues of a real symmetric matrix are all real[1]. Consider a real symmetric $n \times n$ matrix A and an eigenvalue $\lambda \in \mathbb{C}$ for A. Then there is some non-zero complex vector $\mathbf{v}$ of length n satisfying $A\mathbf{v} = \lambda\mathbf{v}$. We have

$$\lambda\langle \mathbf{v}, \mathbf{v}\rangle = (\lambda\mathbf{v})^\top \bar{\mathbf{v}} = (A\mathbf{v})^\top \bar{\mathbf{v}} = \mathbf{v}^\top A\bar{\mathbf{v}} = \mathbf{v}^\top \bar{A\mathbf{v}} = \mathbf{v}^\top \bar{\lambda\mathbf{v}} = \bar{\lambda}\mathbf{v}^\top \bar{\mathbf{v}} = \bar{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle.$$

Since $\mathbf{v} \neq \mathbf{0}$, this shows that λ is equal to its complex conjugate.

If A is a real symmetric matrix, then eigenvectors of A belonging to distinct eigenvalues are orthogonal. To wit, if $A\mathbf{u} = \theta\mathbf{u}$ and $A\mathbf{v} = \tau\mathbf{v}$, then

$$\theta\mathbf{u}^\top \bar{\mathbf{v}} = (A\mathbf{u})^\top \bar{\mathbf{v}} = \mathbf{u}^\top \bar{A\mathbf{v}} = \mathbf{u}^\top \bar{\tau\mathbf{v}} = \tau\mathbf{u}^\top \bar{\mathbf{v}}$$

forcing $\mathbf{u} \perp \mathbf{v}$ when $\theta \neq \tau$.

So a real symmetric $n \times n$ matrix A has the property that $\mathbb{R}^n$ admits an orthonormal basis

$$\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$$

consisting entirely of eigenvectors for A. Ignoring the possibility of multiple eigenvalues for now, let us write $A\mathbf{u}_i = \theta_i \mathbf{u}_i$ (where the θ_i are not necessarily distinct). Letting $E_i = \mathbf{u}_i \mathbf{u}_i^\top$ (a rank one matrix for each i), we have $AE_i = \theta_I E_i$ and $\sum_i E_i = I$ (Exercise[2]). So multiplying both sides of this equation by A, we obtain the *spectral decomposition* of A:

$$A = \sum_{i=1}^{n} \theta_i E_i.$$

Another way to express this is to gather the basis vectors according to common eigenvalues. If matrix A has k distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and F_j is the matrix representing orthogonal projection onto the eigenspace

$$V_j := \{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda_j \mathbf{v}\},$$

then each F_j is a sum of those E_i for which $\theta_i = \lambda_j$ and we obtain

$$A = \sum_{j=1}^{k} \lambda_j F_j.$$

[1] This also holds for complex Hermitean matrices. An $n \times n$ matrix over $\mathbb{C}$ is *Hermitean* if it is equal to its conjugate transpose.

[2] HINT: Write an arbitrary vector $\mathbf{w} \in \mathbb{R}^n$ in terms of the basis $\{\mathbf{u}_i\}_{i=1}^n$ and show that both matrices map $\mathbf{w}$ to itself.

This is a very useful decomposition. Since $F_j F_\ell = \delta_{j,\ell} F_j$, we find

$$A^r = \sum_{j=1}^{k} \lambda_j^r F_j$$

for each non-negative integer r and, more generally,

$$f(A) = \sum_{j=1}^{k} f(\lambda_j) F_j$$

for any rational function $f(t)$ which is not zero at any λ_j. In fact, this extends to power series as well, such as the matrix exponential e^A.

A.1.1 Positive Semidefinite Matrices

An $n \times n$ real symmetric matrix M is *positive semidefinite* (resp., *positive definite*) if $\mathbf{v}^\top M \mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$ (resp., $\mathbf{v}^\top M \mathbf{v} > 0$ for all nonzero $\mathbf{v} \in \mathbb{R}^n$). One easily sees that any non-negative multiple of a positive semidefinite (psd) matrix is also positive semidefinite and that the sum of any number of positive semidefinite matrices of the same size is also psd (as is their direct sum). It is a straightforward exercise to see that every eigenvalue of a positive semidefinite matrix is non-negative and the converse holds as well: If A is a real symmetric matrix, then A is psd if and only if all of its eigenvalues are non-negative. Inside $\mathsf{Mat}_n(\mathbb{R})$, the psd matrices form a cone and the same is true inside any subspace of this vector space.

If $X = \{x_1, \ldots, x_k\}$ is a non-empty finite subset of $\mathbb{R}^n$, then the *Gram matrix* of X is the $k \times k$ matrix G with (i,j)-entry $\langle x_i, x_j \rangle$. Every Gram matrix is positive semidefinite (Proof: It factors as $G = SS^\top$ where S has i^{th} row equal to x_i) and conversely, every $k \times k$ positive semidefinite matrix M of rank m is the Gram matrix of some set of k vectors in $\mathbb{R}^m$. (Exercise: Prove this by building the vectors $x_1, x_2, \ldots$ in turn showing in each case that the square submatrix in the upper left corner of M is the matrix of inner products of the vectors chosen so far.)

A.2 Graph Theory

A simple undirected graph is just a bunch of dots connected by lines. More precisely, a graph $G = (X, R)$ is an ordered pair with X a set ($|X|$ finite for

these notes, but in other applications, X may be infinite), and R a symmetric irreflexive binary relation on X. The elements of X are called *vertices* (or "nodes") and the pairs in R are called *edges* of the graph. So any two vertices $a, b \in X$ are either *adjacent* — i.e., $(a,b), (b,a) \in R$ — or not adjacent — $(a,b), (b,a) \notin R$. A *walk* from a to b in G is a sequence of vertices $a = c_0, c_1, c_2, \ldots, c_r = b$ satisfying $(c_{i-1}, c_i) \in R$ for $1 \leq i \leq r$. The *length* of a walk is r, one less than the number of vertices in the sequence.
Exercise: Prove that, if A is the adjacency matrix of G and $a, b \in X$, then the (a,b)-entry of the matrix A^r is equal to the number of walks of length r from a to b in G. (HINT: Use induction.)

A *path* from a to b in G is a walk from a to b with no repeated vertices except possibly the initial vertex and the terminal vertex (in the case $a = b$). A *cycle* (or "circuit") in G is a path of positive length from a vertex to itself. A graph G is *connected* if for any $a, b \in X$, G contains a walk from a to b. A graph is a *forest* if it contains no cycles. A connected forest is called a *tree.* A graph G is *bipartite* if it contains no cycles of odd length. (Exercise: Prove that this is equivalent to the existence of a bipartition, or "coloring" with two colors, of the vertices such that $(a,b) \in R$ only if a and b are of different colors.

The *girth* of G is the length of a shortest cycle in G (we can take the girth to be infinite if G is a forest). The distance from a to b in G, denoted $\mathrm{dist}(a,b)$, is the length of a shortest path from a to b in G (we can take $\mathrm{dist}(a,b) = \infty$ if no such path exists). When G is connected, the *diameter* of G is the largest possible distance that occurs between vertices of G. For a vertex a and a non-negative integer i, the *sphere of radius* i about a in G is the set $S_i(a) = \{b \in X \mid \mathrm{dist}(a,b) = i\}$.

If a, b are vertices of a graph $G = (X, R)$ and $(a,b) \in R$, we say a is *adjacent* to b or that a and b are *neighbors.* A set of mutually adjacent vertices (a "complete" subgraph) is called a *clique* and a set of mutually non-adjacent vertices is called a *coclique* or "independent set" in G. The *degree* (or "valency") of vertex a is the number of neighbors of a. A graph G is *regular* if every vertex in G has the same degree. We often say "G is a regular graph of valency k" to indicate that every vertex in G has degree k.

Let $G = (X, R)$ be a graph and $\alpha : X \to X$ a permutation of its vertices. We say α is an *automorphism* of G if α "preserves adjacency": for all $a, b \in X$, $(a,b) \in R$ if and only if $(\alpha(a), \alpha(b)) \in R$. The identity permutation is always an automorphism and if α and β are automorphisms of G, then so also are α^{-1} and $\beta\alpha$, their function composition. So the automorphisms of G form a group, the *automorphism group* of the graph G. The vertices are partitioned into orbits by this group (or by any subgroup of it). The graph G is *vertex transitive* if all vertices are in the same orbit: for any $a, b \in X$, there exists an automorphism α of G mapping a to b (i.e., $b = \alpha(a)$). Clearly a vertex transitive graph must be a regular graph, but the converse is not always true.

A graph G is *distance transitive* if, for every two ordered pairs of vertices a, b and a', b' in X, if $\mathrm{dist}(a, b) = \mathrm{dist}(a', b')$, then there is some automorphism α of G satisfying $\alpha(a) = a'$ and $\alpha(b) = b'$. Distance transitive graphs are always vertex transitive (take $\mathrm{dist}(a, b) = 0$). If G is a distance-transitive graph of diameter d, then there exist parameters $\left\{p_{ij}^k\right\}_{0 \le i,j,k \le d}$ such that

$$|S_i(a) \cap S_j(b)| = p_{ij}^k$$

whenever $\mathrm{dist}(a, b) = k$ in G. Indeed, if $\mathrm{dist}(a', b') = k$ also, then the automorphism which maps a to a' and b to b' bijectively maps $S_i(a) \cap S_j(b)$ to $S_i(a') \cap S_j(b')$. One easily checks that all Hamming graphs and Johnson graphs are distance transitive graphs.

Now let us do the same without the group. A graph G of diameter is *distance-regular* provided there exist parameters $\left\{p_{ij}^k\right\}_{0 \le i,j,k \le d}$ such that $|S_i(a) \cap S_j(b)| = p_{ij}^k$ whenever $\mathrm{dist}(a, b) = k$ in G. Every distance transitive graph is distance-regular, but there exist distance-regular graphs with trivial automorphism group. The linear programming bound derived in these notes is valid for any distance-regular graph and more generally for any symmetric association scheme.