Polynomial Spaces

Abhibhav Garg

April 16, 2019

Contents

1 Introduction

Over the course, we saw a number of applications of linear algebra to combinatorics. The most notable application was the Frankl-Wilson theorem, which informally states the following: If we have a set family such that the intersections modulo p are constrained to take atmost s values, then the family can have atmost $\binom{n}{s}$ elements. This was a natural extension of the odd-town problem. In both cases, the argument proceeded by constructing a linearly independent set of vectors, in a space of low dimension, which naturally bounds the size of the family.

In this report, we sketch the ideas of Godsil (1988), who generalised these ideas greatly, by introducing the framework of Polynomial Spaces. This allows one to derive many known results, such as the Frankl Wilson theorem, in a uniform way. While the original paper first introduces the framework completely before instantiating it, here we try and do the two simultaneously, in the hope that this provides more motivation for some of the definitions. The instantiation we pick as the motivating one will be that which allows us to recover the Frankl Wilson theorem.

Notation

Unless stated otherwise, expectations will always be taken with respect to the uniform distribution.

2 Polynomial Spaces

2.1 Basic Definitions

A *polynomial space* is a set Ω and a real valued function $\rho : \Omega \times \Omega \to \mathbb{R}$ that satisfies the following three conditions:

- $\rho(x, y) = \rho(y, x)$ for all $x, y \in \Omega$.
- $\rho(x, x) = \rho(y, y) > 0$ for all $x, y \in \Omega$.
- $\rho(x, x) > \rho(x, y)$ for all $x, y \in \Omega$.

We require (Ω, ρ) to satisfy some additional axioms, which will be stated later. These axioms will always hold when Ω is finite, which is the focus of this report. The function ρ intuitively captures the similarity between two elements of Ω. In many of the example spaces, the space will naturally admit a metric, and ρ will just be the negative of this metric, appropriately translated to satisfy the positivity condition. For the rest of this report, (Ω, ρ) will always denote a polynomial space.

Our main example is the *Johnson Scheme*, $J(n, k)$: Here, Ω is the set of all k sized subsets of a set with n elements, and $\rho(x, y) = |x \cap y|$.

Another example will be the *Hamming Scheme*: Here, Ω is the set of all n length words of a finite alphabet Σ, and $\rho(x, y) = n - |x \Delta y|$, where $|x \Delta y|$ is the usual Hamming distance.

A third example is the unit sphere. Here, Ω is the set of all unit vectors in $\mathbb{R}^n$, and ρ is the usual inner product function. This space is different from the previous two in that it is infinite, but it will satisfy some finiteness conditions that we will see later.

2.2 Codes

We now define the notion of a code. Let Φ be a subset of Ω. Define the distance set, $D(\Phi)$ to as $D(\Phi) = \{\rho(x, y) \mid x, y \in \Phi, x \neq y\}$. Let $d(\Phi) = |D(\phi)|$. This cardinality will also be called the degree of Φ. Finally, if $D(\Phi) \subseteq A$ for some $A \subseteq \mathbb{R}$, then Φ will be called an A-code. Many problems will be reduced to finding large A-codes. In particular, consider the Johnson scheme. Any Φ defines a set family, and the distance set is exactly the set of all possible intersections of members of the family. If we fix an A, we can then bound the size of A-codes, which are equivalent to bounds on set families with restricted intersections. For example, setting A to be all positive natural numbers will recover the hypothesis of EKR, and setting A to be all natural numbers equivalent to some λ_i modulo p will recover the hypothesis of Frankl Wilson.

2.3 Polynomials

We now introduce a class of functions on Ω. Let g be any univariate polynomial from the ring of polynomials over $\mathbb{R}$. For any $a \in \Omega$, define the zonal polynomial $\zeta_a(g)(x) := g(\rho(a, x))$. The

set of functions span $(\{\zeta_a(g) \mid a \in \Omega, \deg(g) \leq r\})$ is a real vector space, which we denote $Z(r)$. We can further also consider the ring generated by the zonal polynomials, namely all functions generated by sums and products of zonal polynomials. We will call this the set of *Polynomials*, and denote it by P. This is also a vector space. Note that Z and P will denote the vector spaces corresponding to Ω. Vector spaces of functions over subsets Φ will be denoted Z_Φ, P_Φ.

We can decompose P into vector spaces $P(r)$ for non-negative integers r, defined recursively as

- $P(0) = Z(0)$.
- $P(1) = Z(1)$.
- $P(k+1) = \text{span}\,(gh \mid g \in P(1), h \in P(k))$.

For a function f, the smallest k such that $f \in Pol(k)$ will be called the degree of f. We also assume the following finiteness condition:

- $\dim P(1) < \infty$.

We now prove a couple of simple finiteness lemmas.

Lemma 2.1. *If Ω is finite, then there is an integer $d \leq |\Omega|$ such that every function on Ω lies in $Z(d)$.*

Proof of lemma 2.1. Pick an element $a \in \Omega$, and let g be the unique monic polynomial whose roots are $D(\Omega \setminus \{a\})$. The zonal polynomial $\zeta_a(g)$ then vanishes on all elements of Ω except a. We can do this for all points, and get a set of $|\Omega|$ zonal polynomials that span the set of all functions on Ω, proving the lemma. □

Lemma 2.2. *If $\dim(P(1)) = k+1$ for finite k, then $\dim(P(r)) \leq \binom{r+k}{r}$. Also, $\dim(Z(r)) \leq \binom{r+k}{r}$.*

Proof of lemma 2.2. Pick a basis of size $k+1$ for $P(1)$. By definition, any element in $P(r)$ is spanned by products of r elements in $P(1)$. Any product of r elements from $P(1)$ can be written as the sum of products of r elements from the basis, of which there are $\binom{r+k}{r}$. The final statement follows from the fact that $Z(r) \subset P(r)$ by definition. □

Going back to the Johnson Scheme, we have the following lemma:

Lemma 2.3. *If (Ω, ρ) is the Johnson Scheme $J(n,k)$, then $\dim(P(r)) = \binom{n}{r}$ for $r = 0, \ldots, k$.*

We defer the full proof of this lemma to a later section. The proof will proceed by showing that the space $P(r)$ is spanned by some indicator vectors (note that since Ω is finite, each vector defines a function, and vice-versa). These vectors will be the rows of the containment matrix[1].

We now state the first application of this setting:

Theorem 2.4. *Let (Ω, ρ) be a polynomial space, and let Φ be a subset of Ω with degree d. Then $|\Phi| \leq \dim(Z(d))$.*

[1] Recall that for $i \leq j$, the containment matrix $N(i,j)$ has rows spanned by all subsets of size i and columns spanned by all subsets of size j, and $N(i,j)_{x,y} = 1$ if and only if $x \subset y$.

Before we prove this, let us see what this implies in the case of Johnson Schemes. As stated before, for a set family Φ, the degree is exactly equal to all possible sizes of intersections of its members. Therefore, the thorem says that if we have a set family whose intersections take values from a set of size s, then the size of the family is atmost $\dim(Z(s))$, which, by 2.3 and 2.2 is upper bounded by $\binom{n}{s}$. In other words, we recover the following theorem by Ray-Chaudhuri and Wilson, which itself generalises EKR:

Theorem 2.5 (Theorem 3 in Ray-Chaudhuri and Wilson (1975))**.** *Let $\mathcal{A}$ be a family of subsets of size k of $[n]$ such that for all $x, y \in \mathcal{A}$, we have $|x \cap y| \in \{\mu_1, \ldots, \mu_s\}$, where the μ_i are distinct. Then $|\mathcal{A}| \leq \binom{n}{s}$.*

We now complete the proof.

Proof of theorem 2.4. Let g be the unique monic polynomial of degree d that has the elements of $D(\Phi)$ as its roots. For all $a \in \Omega$, let $h_a := \zeta_a(g)$. By definition we have that $h_a \in Z(d)$ for every a. Further, each h_a vanishes on all elements of Φ except a itself. Thus the h_a are linearly independent in $Z(d)$. Since we have $|\Phi|$ many linearly independent elements, we must have $|\Phi| \leq \dim(Z(d))$, completing the proof. □

Next, we prove an extension of this theorem.

Theorem 2.6. *Let (Ω, ρ) be a polynomial space such that ρ only takes integer valued. Let p be a prime. Suppose Φ is a subset of Ω such that $D(\Phi)$ has exactly d' distinct elements modulo p, and also these elements are distinct from $\rho(a, a)$ modulo p. Then $|\Phi| \leq \dim(Z(d'))$.*

Again, before we prove the theorem, we see what it implies in the case of Johnson Schemes. Using the exact same argument as before, we can see that this exactly recovers the Frank Wilson theorem.

Proof of theorem 2.6. The proof will be very similar to that of theorem 2.4. The set $D(\Phi)$ splits into d' residue classes modulo p. Pick an element from each residue class, say the representative in $[0, p-1]$, and let g be the unique monic degree d' polynomial (with integer coefficients) that vanishes on these points. We can construct the functions $h_a := \zeta_a(g)$ as before. Each $h_a \in Z(d')$. We will argue that these elements are linearly independent over the reals, and this will complete the proof, since we have $|D(\Phi)|$ linearly independent elements from $Z(d')$.

Each h_a is 0 modulo p on all points in Φ except at a, since $h_a(x) = g(\rho(a, x))$, and going modulo p we have $g(\rho(a, x)) \pmod p = g(\rho(a, x) \pmod p) = 0$. Thus the h_a are independent modulo p. Since they are independent modulo p, they are also independent over the integers: If they satisfied a non-trivial linear relationship over the integers, we could first make sure that the coefficients are relatively prime, and then go modulo p, which will give us a non-trival relationship (modulo p), a contradiction. This further implies that they are also independent over the rationals, since if we had a rational linear combination that vanished, we could have cleared out the denominators. Finally, since the coefficients are integers, this will also imply that the functions are independent over the reals. This completes the proof. □

We will now define the notion of designs. For the Johnson scheme, this notion of designs will coincide with the combinatorial notion, which will allow us to recover another important theorem from Ray-Chaudhuri and Wilson (1975).

2.4 Designs

We first introduce the additional axiom that we require on polynomial spaces: We require that P admits a real valued inner product $\langle \cdot, \cdot \rangle$ that satisfies the following:

- For all $g, h \in P$ and for all $a \in \Omega$ we have $\langle \zeta_a(x) g, h \rangle = \langle g, \zeta_a(x) h \rangle$.
- If $f \in P$ is non-negative, then $\langle f, \mathbf{1} \rangle \geq 0$, where $\mathbf{1}$ is the constant 1 function.

If Ω is finite, the inner product will always be given by

$$\langle f, g \rangle = \mathop{\mathbb{E}}_{x \in \Omega} [fg] = \frac{1}{|\Omega|} \sum_{x \in \Omega} f(x) g(x).$$

A t-design in a polynomial space (Ω, ρ) is a finite subset Φ such that whenevr f, g are functions with $fg \in P(t)$, then we have

$$\langle f, g \rangle = \mathop{\mathbb{E}}_{x \in \Phi} [fg] = \frac{1}{|\Phi|} \sum_{x \in \Phi} f(x) g(x).$$

We have the following, when Φ is a t-design, for f, g such that $fg \in P(t)$:

$$\begin{aligned} \langle f, g \rangle &= \mathop{\mathbb{E}}_{x \in \Phi} [fg] \\ &= \mathop{\mathbb{E}}_{x \in \Phi} [1 \times (fg)] && \text{(Since the sum is finite)} \\ &= \langle \mathbf{1}, fg \rangle \end{aligned}$$

If we assume that $\langle f, \mathbf{1} \rangle$ is the average value of f over Ω, as is the case for the finite inner product, then a t-design allows us to compute this average, which could be over a potentially infinite set, by just averaging over a finite set. Of course there is no guarantee that t-designs will always exist, when Ω is infinite. However, if the polynomial space Ω satisfies a slightly stronger condition, then we can guarantee the existance of weighted t-designs (which are basically t-designs, where distribution over elements is not uniform), even when Ω is infinite. The proof of this is pretty neat, and involves a cone duality argument. However, since these types of polynomial spaces are not the focus of this presentation, we move the formal statement and proof of this result to Appendix A.

The first theorem on t-designs is the following bound, which says that t-designs cannot be too small.

Theorem 2.7. *Let Φ be a t-design in Ω. Then $|\Phi| \geq \dim(Pol(\Omega, \lceil t/2 \rceil))$.*

As before, we first look at the case of Johnson Schemes. For this, we first need a definition. A $t - \{n, k, \lambda\}$ design in the combinatorial sense (abbreviated t-design) is a subset of k-sized subsets of $[n]$ such that every t sized subset of $[n]$ is contained in exactly λ blocks.

We have the following lemma about Johnson schemes, whose proof we defer.

Lemma 2.8. *If (Ω, ρ) is the Johnson Scheme $J(n, k)$, then t-designs in this polynomial space are equivalent to t-designs in the combinatorial sense.*

The above two allows us to recover the following theorem by Ray-Chaudhuri and Wilson

Theorem 2.9 (Theorem 1 in Ray-Chaudhuri and Wilson (1975))**.** *Let $\mathcal{A}$ be a t-design with $t = 2s$, and such that $n \geq k + s$. Then $|\mathcal{A}| \geq \binom{n}{s}$.*

We now prove the theorem.

Proof of theorem 2.7. Pick an orthogonal basis for $P(\lceil t/2 \rceil)$, say $h_1, \ldots, h_n$. By the definition of t-designs, we have that

$$\langle h_i, h_j \rangle = \mathop{\mathbb{E}}_{x \in \Phi} [h_i h_j] = \delta_{ij}.$$

The restrictions of h_a to the set Φ thus form a pairwise orthogonal set in ${P_\Phi}^2$. Since Φ is finite, we have that $|\Phi| = \dim(P(\Phi)) \geq n$, completing the proof. □

In the next section we complete the proofs of the two lemma regarding Johnson Schemes, and in the section after that we will see more instantiations of this framework.

3 Johnson Schemes

We need to prove lemma 2.3 and lemma 2.8. First we introduce a class of indicator functions. For any set $s \subset [n]$, let $f_s(x) = 1$ if and only if $s \subset x$. This is the characteristic function for the k subsets containing s. Further define the space $I(r) = \text{span}\,(\{f_s \mid |s| = r\})$. The next lemma captures some simple properties of $I(r)$.

Lemma 3.1. *The spaces $I(r)$ satisfy $I(r+1) \subseteq \text{span}\,(I(r) \cdot I(1))$. Further, for all $r \leq s$ they also satisfy $I(r) \subseteq I(s)$.*

Proof of lemma 3.1. The first lemma follows since $f_s f_t = f_{s \cup t}$ when s, t are subsets. The second follows from the identity

$$f_t = \frac{1}{\binom{k-|t|}{u-|t|}} \sum_{t \subseteq u, |s|=u} f_u.$$

□

We also have the following lemma that related $I(r)$ and $P(r)$.

Lemma 3.2. *For $J(n,k)$ we have $I(r) = P(r)$.*

Proof of 3.2. Since $I(r)$ and $P(r)$ have the same recursive definition, we only need to prove the result for $r = 1$. Let H be the incidence matrix of points and k subsets: the rows and indexed by points and the columns by subsets. The matrix $H^T H$ then just has entries $\rho(x, y)$, and thus the rows of this matrix give elements of $P(1)$. Infact the columns of $H^T H$ have constant sums, and this value is non-zero. Thus every function in $P(1)$ is in the row span. This gives $P(1) \subseteq I(1)$. Now using the fact that $\text{rank}(H) = \text{rank}(H^T H)$, we get the result $P(1) = I(1)$, which completes the proof. □

[2] Recall that P_Φ is the ring of functions restricted to Φ.

With these definitions, we are ready to prove the lemma. We will first prove lemma 2.8.

Proof of lemma 2.8. By lemma 3.2 and the argument about averages in subsection 2.4 (note that Ω is finite here), we only need to show that

$$\langle \mathbf{1}, f \rangle = \mathop{\mathbb{E}}_{x \in \Phi} [f]$$

for all functions $f \in I(t)$, if and only if Φ is a t-design in the combinatorial sense.

The term on the right counts the number of elements of Φ that contain s, and normalises this number by $|\Phi|$. The term on the left is just $\binom{n-t}{k-t}$, normalised by $|\Omega|$. These two are equal for all t sized subsets of $[n]$ if and only if all of them lie in the same number of elements of Φ. This completes the proof. □

We will finally prove lemma 2.3, which will conclude our discussion on Johnson schemes.

Proof of lemma 2.3. By lemma 3.2, we just have to prove that the indicator functions f_s are linearly independent. This is equivalent to proving that the incidence matrix $N(r, k)$ is full rank. A simple proof of this can be found in Foody and Hedayat (1977) (see lemma 5.1 within). □

We conclude by very briefly discussing the other two mentioned schemes.

4 More Schemes

4.1 Hamming Schemes

A $t - \{q, k, \lambda\}$ orthogonal array over a q-sized alphabet Σ is an $N \times k$ matrix M with entries from Σ such that the following holds: for each $s \leq t$, each s length tuple in Σ^s occurs λ_s times as a row of the $N \times s$ matrix formed by choosing s columns, and also $\lambda = \lambda_t$. It turns out that a t-design is the same thing as a $t - \{q, k, \lambda\}$ orthogonal array. It also turns out that a linear code $\mathcal{C}$ is a t-design if and only if its dual code has minimum distance greater than t. Thus we get an upper bound on the size of linear codes, given a lower bound on the distance.

4.2 Unit Sphere

While the underlying space in the case of the unit sphere is infinite, the dimension of $P(r)$ is still finite, and is equal to $\binom{n+r-1}{r-1} + \binom{n+r-2}{r-2}$. This space is closed and bounded in the sense described in Appendix A, and thus the results of that section apply. The codes and designs in this setting are well studied, see for example Delsarte et al. (1977). In particular, (weighted) t-designs here correspond exactly to numerical integration schemes with precision t.

5 Bibliography

P. Delsarte, J. M. Goethals, and J. J. Seidel. Spherical codes and designs. *Geometriae Dedicata*, 6(3):363–388, Sep 1977. ISSN 1572-9168. doi: 10.1007/BF03187604. URL https://doi.org/10.1007/BF03187604.

W Foody and A Hedayat. On theory and applications of bib designs with repeated blocks. *Ann Statist*, 5, 09 1977. doi: 10.1214/aos/1176343949.

C.D. Godsil. Polynomial spaces. *Discrete Mathematics*, 73(1):71 – 88, 1988. ISSN 0012-365X. doi: https://doi.org/10.1016/0012-365X(88)90134-3. URL http://www.sciencedirect.com/science/article/pii/0012365X88901343.

Dijen K. Ray-Chaudhuri and Richard M. Wilson. On t-designs. *Osaka J. Math.*, 12(3):737–744, 1975. URL https://projecteuclid.org:443/euclid.ojm/1200758175.

A Existence of designs

Here we will sketch a proof of existence of weighted t-designs. First we define a weighted t-design. A non-negative and finitely supported measure ω on Ω is called a weighted t-design if the following holds for all functions f of degree atmost t:

$$\langle \mathbf{1}, f \rangle = \mathop{\mathbb{E}}_{x \sim \omega} [f] .$$

A null design of strength atmost t is a function ν with finite support such that $\langle \nu, f \rangle = 0$ for all functions f of degree atmost t. The difference of two weighted t-designs is always a null design.

We now define some preliminaries. A convex cone in $\mathbb{R}^n$ is a subset that is convex, and closed under scaling by positive constants. Our cones will always be convex, and in the argument that follows, we drop the adjective convex for brevity. An example of a cone is the set of all points with all coordinates non-negative. Given a cone C, its dual cone C^* consists of all vectors x such that $\langle x, y \rangle \geq 0$ for all $y \in C$. For cones that are closed, we have $C = C^{**}$.

We will now define a notion of closedness and boundedness for a polynomial space. This will require a bunch of auxilary definitions. The cone we are interested in is the following. Let $U = P(t)$, and let U' denote the dual space. For every $a \in \Omega$, let $\omega(a) : U \to \mathbb{R}$ be defined as $\omega(a)(p) := p(a)$. Let C be the cone in U' generated by $W := \{\omega(a) \mid a \in \Omega\}$. The dual cone then consists of all non-negative elements of U. Let $C_1 \subset C$ be the subset formed by those elements α such that $\alpha(1) = 1$. Then $\omega(a) \in C_1$ for all $a \in \Omega$.

It is easy to check that C_1 is a convex set. The claim is that the points $\omega(a)$ are extreme points of C_1, that is points that are non themselves linear combinations of other points. To see this, consider the function $\rho(a, a) - \rho(a, x)$. The functional $\omega(b)$ is positive on this, unless $b = a$, and thus the claim follows (Remember that the $\omega(a)$ generate C). We also have that all extreme points of C_1 must lie in W, since C and hence C_1 is generated by W. Thus the set W is exactly the set of all all extreme points of C_1.

If the set C_1 is compact, then every one of its interior points will be linear combinations of extreme points. The cone C will be closed if the set W is closed, and if this happens, the set C_1 will also be closed (since it is an intersection of a closed set and a hyperplane). A polynomial space will be called closed if the set W is closed. We can define a norm on U'. Pick a basis for $P(t)$, and for any $u' \in U'$, define $\|u'\|$ to be maximum value of u' on a basis element. The set U' will be bounded according to this norm if and only if the range of ρ is bounded. A polynomial space is called bounded if this occurs.

Having defined closedness and boundedness, we can now state and prove the existence theorem.

Theorem A.1. *Let (Ω, ρ) be a closed bounded polynomial space. If $\dim(P(t)) = d$, then there are weighted t-designs supported by atmost d points.*

Proof of theorem A.1. Let the map $\lambda : U \to \mathbb{R}$ be defined as $\lambda(p) = \langle \mathbf{1}, p \rangle$. Clearly $\lambda \in C^{**}$. The cone C is closed, and hence $C^{**} = C$, and thus $\lambda \in C$, and also $\lambda \in C_1$. The set C_1 is a compact set in a $d - 1$ dimensional space, and thus we can write λ as a convex combination of $d - 1 + 1 = d$ extreme points. Thus λ is a convex combination of atmost d of the $\omega(a)$, completing the proof. □