Spectral Methods for Intersection Problems

Friedgut's Research Program in Extremal Combinatorics

Yuval Filmus

January 29, 2012

Contents

1 Introduction

One of the classical results in extremal combinatorics is the Erdős-Ko-Rado theorem, which bounds the size of uniform intersecting families of sets. Classical proofs of this theorem as well as related results utilize combinatorial down-to-earth arguments. The advantage of these arguments is that they are very simple and elegant. It is often easy to prove a tight upper bound; identifying the extremal families is more complicated; and analyzing almost-optimal or moderately-big families ("stability") can be challenging.

In the past decade, a new research program, championed by Friedgut and Dinur, is supplanting these combinatorial arguments by a unified spectral approach. Application of the spectral approach is not automatic, and problem-specific arguments are needed for all three types of results. However, the method succeeds where purely combinatorial arguments fail. The most prominent examples to date are multiply intersecting families of permutations and triangle-intersecting families of graphs.

While standing on their own, results in this area have also been applied to theoretical computer science. An extremal result on intersecting families can be viewed as a global structure resulting from local consistency; it is therefore no wonder that such results have been applied to the construction of PCPs [14]. More generally, the same circle of ideas has been very influential in circuit complexity, culminating in the recent spectacular result by Braverman [7].

In this short survey, we provide an overview of the spectral method, concentrating on the upper bound part of the arguments. We describe both the general method and all known applications of it from the literature; these are not described in fullest generality. Due to lack of space, we do not present complete arguments, but attempt to give their gist.

1.1 Structure of the paper

We begin this survey by describing some combinatorial background (§2). Next we describe the spectral method in broad lines (§3). The toy example presented there is followed by four different applications of the method from the literature (§4–7). Some open problems follow (§8). We conclude by outlining some other approaches to these problems (§9).

The applications we present are:

- Intersecting families of colored sets (§4).
- Triangle-intersecting families of graphs (§5).
- Uniform intersecting families (§6).
- Intersecting families of permutations (§7).

More discussion on the particular features of each application can be found in §3.5.

2 Combinatorial Background

In their classical paper from 1961, Erdős, Ko, and Rado [19] proved the following seminal theorem.

Theorem 2.1 (Erdős-Ko-Rado)**.** *Let $\mathcal{F}$ be a k-uniform intersecting family on $[n]$, i.e. $\mathcal{F}$ is a subset of $[n] = \{1, \dots, n\}$ consisting of sets of size k, such that any $A, B \in \mathcal{F}$ are intersecting: $A \cap B \neq \varnothing$. If $k \leq n/2$ then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1}. \tag{1}$$

Moreover, if $k < n/2$ then equality holds in (1) if and only if there exists $i \in [n]$ such that

$$\mathcal{F} = \{S \subset [n] : |S| = k \text{ and } i \in S\};$$

such sets are called dictators[1]*.*

Note that when $k > n/2$, the family consisting of *all* sets of size k is intersecting. When $k = n/2$, dictators are not the unique optimal family; consider for example $n = 4$, $k = 2$ and the following family:

$$\mathcal{F} = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$

While the original proof is not complicated, a truly simple proof was found in 1972 by Katona [40]; see §9.3. Many other proofs are known today, see for example [27].

The original theorem was just the starting point of an entire research program, which branched into two main parts. On the one hand, analogues of the theorem were sought in different settings. On the other hand, research also focused on better understanding of intersecting families. In this section we focus on the latter vein of results; these inspired the results of the former vein, which form the bulk of this survey.

One natural question to ask is what are the alternatives to dictatorship. This was answered by Hilton and Milner [36] in 1967 (upper bound) and Frankl [23] in 1978 (extremal families).

Theorem 2.2 (Hilton-Milner)**.** *Let $\mathcal{F}$ be a k-intersecting family on $[n]$ such that $\bigcap \mathcal{F} = \varnothing$ (such a family is called* non-trivial[2]*). If $k < n/2$ then*

$$|\mathcal{F}| \leq \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1;$$

in the case of equality, $\mathcal{F}$ is obtained from an i-dictatorship by picking a subset I of size k ($i \notin I$), removing all sets not intersecting I, and adding I as compensation.

As an example, here is an extremal non-trivial intersecting family with $n = 7$ and $k = 3$ (here $i = 1$ and $I = \{5, 6, 7\}$):

$$\mathcal{F} = \{1\} \times \{2, 3, 4\} \times \{5, 6, 7\} \cup \{\{1, 5, 6\}, \{1, 5, 7\}, \{1, 6, 7\}, \{5, 6, 7\}\}.$$

The easiest proof of this theorem uses shifting techniques (§9.1–9.2).

A natural generalization is to t-intersecting families: families where the intersection contains at least t elements. Literally decades of research have finally led to the complete solution of this problem by Ahlswede and Khachatrian in 1997 [1–3].

Theorem 2.3 (Ahlswede-Khachatrian)**.** *Let $\mathcal{F}$ be a k-uniform t-intersecting family on $[n]$, where $k \leq n/2 + t - 1$. If $\mathcal{F}$ is extremal (largest possible) then it is (up to isomorphism) one of the following families, for some $s \geq 0$:*

$$\mathcal{F}_s = \{S \subset [n] : |S| = k \text{ and } |S \cap [t + 2s]| \geq t + s\}.$$

For most values of n, k, t there is a unique maximizing s, but for some values there are two consecutive maximizing values.

Here are two extremal families for $n = 6$, $k = 3$, $t = 2$:

$$\mathcal{F}_0 = \{1, 2\} \times \{3, 4, 5, 6\}, \quad \mathcal{F}_1 = \{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 3, 4\}, \{2, 3, 4\}\}.$$

A similar theorem holds for the non-trivial case, where the family is forbidden from containing t common elements. Here the only interesting case is when the global optimum is trivial; in that case the extremal non-trivial family is obtained from a "junta" ($\mathcal{F}_0$) by picking a subset I of size $k - t + 1$ (thus $|I| + t = k + 1$), removing all sets not intersecting I, and adding as compensation all subsets of $I \cup [t]$ of co-size 1 (i.e. with one element removed).

Another natural generalization is to r-wise intersecting families, in which every r sets intersect. Frankl [21] proved the following theorem in 1976.

[1]This intriguing terminology comes from the area of voting, which has also been analyzed using similar methods [31]. In the sequel we also consider juntas.

[2]One is tempted to call it *non-principal* in analogy to ultrafilters. Some call it *non-centered*.

Theorem 2.4 (Frankl)**.** *If $\mathcal{F}$ is a k-uniform r-wise intersecting family on $[n]$, where $k \le (1-1/r)n$, then*

$$|\mathcal{F}| \le \binom{n-1}{k-1}.$$

If equality holds and $k < (1-1/r)n$ then $\mathcal{F}$ is a dictatorship.

This theorem can be easily proved using Katona's method 9.3.

Finally, here is a less natural generalization, which is however important in proving stability results. Instead of looking at one family, we can consider two families, and require that every pair of sets, one per each family, intersect; we call the two families *cross-intersecting*.

Definition 2.1. Two families $\mathcal{F}, \mathcal{G}$ are *cross-intersecting* if any $F \in \mathcal{F}$ and $G \in \mathcal{G}$ have non-empty intersection.

Mastumoto and Tokushige [46] proved in 1989 the following theorem.

Theorem 2.5 (Matsumoto-Tokushige)**.** *Let $\mathcal{F}, \mathcal{G}$ be cross-intersecting families on n which are a-uniform and b-uniform, respectively. If $a, b \le n/2$ then*

$$|\mathcal{F}||\mathcal{G}| \le \binom{n-1}{a-1}\binom{n-1}{b-1}.$$

Furthermore, if equality holds then $\mathcal{F}$ and $\mathcal{G}$ are both dictatorships, unless $a = b = n/2$.

Note that optimizing the *sum* $|\mathcal{F}| + |\mathcal{G}|$ is more problematic; one can choose $\mathcal{F} = \varnothing$ and then take $\mathcal{G}$ to be all $\binom{n}{b}$ sets. Even if we don't allow empty families, we can get a sum which is much larger than $\binom{n-1}{a-1} + \binom{n-1}{b-1}$. Recently Frankl and Tokushige [28] succeeded in combining the two preceding theorems.

Theorem 2.6 (Frankl-Tokushige)**.** *Let $\mathcal{F}_1, \ldots, \mathcal{F}_r$ be r-wise-cross-intersecting k-uniform families on $[n]$, where $k \le (1-1/r)n$; thus for any choice of r sets $S_i \in \mathcal{F}_i$, the sets S_i have non-empty intersection. Then*

$$\prod_{i=1}^{r} |\mathcal{F}_i| \le \binom{n-1}{k-1}^r.$$

The ultimate goal of this line of research would be to generalize the Ahlswede-Khachatrian theorem to r-wise intersecting families.

Conjecture 2.7 (Frankl-Tokushige)**.** *Let $\mathcal{F}$ be a k-uniform r-wise t-intersecting family on $[n]$, where $k < (1-1/r)n$. If $\mathcal{F}$ is extremal then it is of the form*

$$\mathcal{F}_s = \{S \subset [n] : |S| = k \text{ and } |S \cap [t+rs]| \ge t+s\}.$$

If $\mathcal{F}_1, \ldots, \mathcal{F}_r$ are r-wise-cross-t-intersecting k-uniform families on $[n]$, where $k < (1-1/r)n$, then the product $\prod_{i=1}^r |\mathcal{F}_i|$ is maximized when each of the families is an extremal k-uniform r-wise t-intersecting family.

Recent progress on this conjecture can be found in Tokushige [51] and Frankl and Tokushige [28].

Let's summarize the results we've presented:

- Erdős-Ko-Rado: Maximal k-uniform intersecting families on $[n]$ for $k < n/2$ are dictatorships.
- Hilton-Milner: Same setting, maximal non-trivial family is obtained by adding a copy of $[k]$ and removing all sets not intersecting it.
- Ahlswede-Khachatrian: Maximal k-uniform t-intersecting families on $[n]$ for $k < n/2$ are of the form $\binom{t+2s}{t+s}$ for some $s \ge 0$, i.e. intersect a fixed set of size $t+2s$ in at least $t+s$ points.
- Frankl: Maximal k-uniform r-wise intersecting families on $[n]$ for $k < (1-1/r)n$ are dictatorships.
- Matsumoto-Tokushige: The maximum of the product of the sizes of a, b-uniform cross-intersecting families on $[n]$ for $a, b < n/2$ is obtained on dictatorships.
- Frankl-Tokushige: The maximum of the product of the sizes of k-uniform r-wise-cross-intersecting families on $[n][$ for $k < (1-1/r)n$ is obtained on dictatorships.

3 Spectral Method

Extremal families are simple. This initial observation, gleaned from theorems like Ahlswede-Khachatrian's, is the starting point for the spectral method. While extremal families have simple descriptions, if we are given a family in the form of a collection of sets, its simple form is hard to detect. The essential insight is that these families do have simple representations when we look at their spectral decomposition.

Representing the families in a different form presents us with a new problem: expressing the intersection properties of the family. Here we have to use another key property of the spectral decomposition: the effects of convolution on it are easy to describe. A family is intersecting if it is disjoint from the family obtained from complementing each set in it; complementation is just convolution with $[n]$. The original and complemented families are orthogonal, a property which translates directly to the spectral representation.

In order to appreciate the generalities of the method, it is better to start with a straightforward example. We prepare the ground by a short discussion of the relevant Fourier transform in §3.1. We then work out all the details of a bound of 2^{n-1} on the size of an intersecting family on $[n]$ in §3.2. After that, we outline the basic method in §3.3. We then generalize intersection problems to *color agreement* problems in §3.4 and for non-uniform distributions in §6.5. We conclude in §3.5 by discussing the examples.

3.1 A Fourier Transform

The Fourier-Walsh basis is a spectral basis for real-valued[3] functions supported on $2^{[n]}$. Looked at from a different angle, it is a basis for the vector space of all functions $2^{[n]} \longrightarrow \mathbb{R}$.

The basis is composed of 2^n basis functions $\chi_T \colon 2^{[n]} \longrightarrow \mathbb{R}$, indexed by subsets of $[n]$. The basis functions are defined as follows:

$$\chi_T(S) = (-1)^{|S \cap T|}.$$

We also think of subsets of $[n]$ as vectors belonging to $\mathbb{Z}_2^n$. In this language, we have the equivalent formula

$$\chi_T(S) = (-1)^{\langle S,T \rangle}.$$

As an example, here is the Fourier-Walsh basis for $n = 2$:

S	$\varnothing$	$\{1\}$	$\{2\}$	$\{1,2\}$
$\chi_\varnothing$	1	1	1	1
$\chi_{\{1\}}$	1	-1	1	-1
$\chi_{\{2\}}$	1	1	-1	-1
$\chi_{\{1,2\}}$	1	-1	-1	1

The first fact that jumps to mind is that this is an orthonormal basis.

Lemma 3.1. *Endow the vector space of real-valued functions on $2^{[n]}$ with the following inner product:*

$$\langle f, g \rangle = 2^{-n} \sum_{S \subset [n]} f(S)g(S).$$

The vectors χ_T form an orthonormal basis for this space.

The representation of an arbitrary function in this basis is known as the *Fourier-Walsh transform*. Because the Fourier-Walsh basis is orthonormal, there is a very simple formula for this representation.

Definition 3.1. The Fourier-Walsh transform of a function $f \colon 2^{[n]} \longrightarrow \mathbb{R}$ is a function $\hat{f} \colon 2^{[n]} \longrightarrow \mathbb{R}$ defined by

$$\hat{f}(T) = \langle f, \chi_T \rangle = 2^{-n} \sum_{S \subset [n]} f(S)(-1)^{|S \cap T|},$$

[3]Everything also works over the complex numbers, but such generalities will only be necessary in §4.

or, equivalently,
$$f(S) = \sum_{T\subset[n]} \hat{f}(T)\chi_T(S) = \sum_{T\subset[n]} \hat{f}(T)(-1)^{|S\cap T|}.$$

The fact that the domains of both function and transform coincide is not accidental (it happens when we replace $\mathbb{Z}_2^n$ by any other finite abelian group) but also not universal, as we will see in §7.

Orthonormality of the Fourier basis directly implies the all-important Parseval identity.

Lemma 3.2 (Parseval)**.** *For any two* $f, g\colon 2^{[n]} \longrightarrow \mathbb{R}$ *we have*
$$\langle f, g\rangle = \sum_{T\subset[n]} \hat{f}(T)\hat{g}(T).$$

The functions χ_T are not arbitrary. Indeed, they are *characters*.

Definition 3.2. A character of an abelian group G is a non-zero function $\chi\colon G \to \mathbb{C}$ such that for any $x, y \in G$,
$$\chi(x)\chi(y) = \chi(x+y).$$

The functions χ_T are characters of $\mathbb{Z}_2^n$:
$$\chi_T(x)\chi_T(y) = \chi_T(x+y),$$
where we think of x, y as vectors in $\mathbb{Z}_2^n$. The corresponding operation for sets is $A + B = \overline{A\triangle B}$, where $\triangle$ is the symmetric difference. For example, for any T we have $\chi_T(\{1\})\chi_T(\{2\}) = \chi_T(\{1,2\})$.

The fact that the χ_T are characters has a direct corollary on their behavior under *convolution*. Given two $2^{[n]}$-valued random variables X, Y, their convolution $Z = X + Y$ is the random variable corresponding to their sum. Writing this out, we get a definition of convolution for general functions $f\colon 2^{[n]} \longrightarrow \mathbb{R}$.

Definition 3.3. Let $f, g\colon 2^{[n]} \longrightarrow \mathbb{R}$. Their *convolution* $h = f * g$ is defined by
$$h(x) = \sum_{y+z=x} f(y)g(z).$$

For example, convolving a function with $\mathbf{1}_{[n]}$ (a function which is 1 on $[n]$ and 0 elsewhere) is the same as complementation of the domain:
$$(f * \mathbf{1}_{[n]})(x) = f(x + [n]).$$
As sets, $x + [n]$ is just the complementation of x.

The character property of the basis functions, together with their orthonormality, implies the following formula for convolution.

Lemma 3.3. *If* $h = f * g$ *then*
$$\hat{h}(T) = 2^n \hat{f}(T)\hat{g}(T).$$
In particular, if $h = f * \mathbf{1}_S$ *then*
$$\hat{h}(T) = \chi_T(S)\hat{f}(T) = (-1)^{|S\cap T|}\hat{f}(T).$$

Proof. We have
$$\begin{aligned} h(x) &= \sum_y f(y)g(x+y) \\ &= \sum_y \sum_{S,T} \hat{f}(S)\chi_S(y)\hat{g}(T)\chi_T(x+y) \\ &= \sum_{S,T} \chi_T(x)\hat{f}(S)\hat{g}(T) \sum_y \chi_S(y)\chi_T(y) \\ &= 2^n \sum_S \chi_S(x)\hat{f}(S)\hat{g}(S). \end{aligned}$$
For the second statement, $\widehat{\mathbf{1}_S}(T) = \langle \mathbf{1}_S, \chi_T\rangle = 2^{-n}\chi_T(S)$. $\square$

The Fourier-Walsh character χ_T is a function which depends only on the coordinates in T. For each set S, we have $2^{|S|}$ characters which depend only on the coordinates of S, namely χ_T for $T \subset S$. Since the vector space of all functions supported on S has dimension $2^{|S|}$, we have proved the following important property.

Lemma 3.4. *A function $f\colon 2^{[n]} \longrightarrow \mathbb{R}$ depends only on the coordinates in S, i.e. for all R, $f(R) = f(R \cap S)$, if and only if its Fourier-Walsh transform is supported by the coefficients $\{T : T \subset S\}$.*

There is a natural gradation of the space of all functions. For each $S \subset [n]$, consider the set F_S of all functions depending only on S. The lattice structure of $[n]$ is inherited by the F_S, that is $S_1 \subset S_2$ implies $F_{S_1} \subset F_{S_2}$. Thus $F_\varnothing$ is the constant functions, and $F_{[n]}$ is all functions. Now let's form a coarser partition

$$F_k = \bigcup_{|S|=k} F_S.$$

These sets of functions form a chain

$$F_0 \subset F_1 \subset \cdots \subset F_n.$$

The point is that F_S can be defined analogously as functions whose Fourier-Walsh transforms are supported by coefficients T with $T \subset S$. Similarly, F_k consists of all functions whose Fourier-Walsh transforms are concentrated on the first $k+1$ levels (we count $\chi_\varnothing$ as level zero).

Let's recap the important properties of the Fourier-Walsh transform we've described:

- The Fourier-Walsh transform represents an arbitrary function $2^{[n]} \longrightarrow \mathbb{R}$ in some orthonormal basis.
- Parseval's identity: $\langle f, g\rangle = \sum_T \hat{f}(T)\hat{g}(T)$.
- If $h = f * g$ then $\hat{h} = 2^n \hat{f} \cdot \hat{g}$.
- A function depends only on coordinates from S iff its Fourier-Walsh expansion is supported on subsets of S.

3.2 Toy Example

We are now able to prove the following easy theorem.

Theorem 3.5. *Let $\mathcal{F}$ be an intersecting family on $[n]$. Then $|\mathcal{F}| \le 2^{n-1}$.*

The proof will contain virtually all the ingredients of the "upper bound" part of the method.

Theorem 3.5 is trivial since out of any pair of complementary sets, at most one can belong to $\mathcal{F}$. Our proof will consist of writing out this condition in its spectral form, and drawing the necessary conclusions.

We start by replacing $\mathcal{F}$ with its characteristic function f defined by

$$f(S) = \begin{cases} 1 & S \in \mathcal{F}, \\ 0 & S \notin \mathcal{F}. \end{cases}$$

The function $g = f * [n]$ is obtained from f by complementing each set, that is

$$g(S) = f(\overline{S}).$$

Since $\mathcal{F}$ is intersecting, the functions f, g are orthogonal. Using Parseval's identity and the convolution formula for the transform, we get

$$0 = \langle f, g\rangle = \sum_T \hat{f}(T)\hat{g}(T) = \sum_T \chi_T([n])\hat{f}(T)^2.$$

Since $T \cap [n] = T$, we get the all-important

$$\sum_T (-1)^{|T|} \hat{f}(T)^2 = 0.$$

This equation states that the "even" and "odd" squared Fourier coefficients must balance. Since $\chi_\varnothing$ is the constant 1 vector, $\hat{f}(\varnothing)$ is just the relative size of the family. For convenience, denote this by μ; we endeavor to show $\mu \leq 1/2$.

At this point we use the fact that f isn't arbitrary: coming from a bona fide family $\mathcal{F}$, it is 0–1 valued, and so (using Parseval)

$$\sum_T \hat{f}(T)^2 = \langle f, f \rangle = \mu.$$

We now have three constraints:

$$\sum_T (-1)^{|T|} \hat{f}(T)^2 = 0, \tag{2}$$

$$\hat{f}(\varnothing)^2 = \mu^2, \tag{3}$$

$$\sum_T \hat{f}(T)^2 = \mu. \tag{4}$$

This is a linear program with the variables $\hat{f}(T)^2$, which are constrained on being positive. It states that the "even" and "odd" variables must balance, and that one even variable occupies μ of the mass. Clearly it cannot occupy more than half of the mass, and so $\mu \leq 1/2$.

3.3 General Method

We now describe how to generalize the method of proof of Theorem 3.5. Before that, however, let us discuss our goals more completely. Theorem 3.5 only gives an upper bound on the size. However, some of the theorems in §2 are more specific, providing us with the structure of all optimal families (such a nice description is simply not available in the setting of Theorem 3.5). Other goals are describing what happens when a family is *almost* optimal or *somewhat* optimal, and Hilton-Milner type results. Here is the full list.

- Upper bound: Find the maximal size of an intersecting family.
- Uniqueness: Identify all intersecting families achieving the upper bound.
- Stability: Show that if a family is almost maximum, then it is in fact close to a bona fide maximum family.
- Hilton-Milner Gap: Show that if a family is not included in any maximum family, then it is substantially smaller[4]; identify all maximum families of that form.
- Strong stability [12, 13]: Show that if a family is somewhat large, then it is close to a family which depends on few coordinates.

The simple proof of Theorem 3.5 used the fact that no two complementary sets are contained in an intersecting family. This property can be rephrased in the language of graphs [5]. Form a graph G by taking the vertex set $\mathbb{Z}_2^n$ and connecting any two complementary sets. Then *a family is intersecting if and only if it is an independent set in G*. The edges of G denote constraints on the family: pairs of sets that cannot both be in the family at the same time.

The property of G that we used during the proof is that G is a *Cayley graph* for the group $\mathbb{Z}_2^n$. In other words, the edges of G correspond to addition of some element of $\mathbb{Z}_2^n$. Since convolution with a constant

[4]This is not true in all settings, for example it doesn't hold in §6.

element is a diagonal operator over the Fourier-Walsh basis (i.e. it multiplies each basis element by some constant), we could express the fact that a family is intersecting by an equation like (2).

The general analogue of equation (2) is obtained by replacing $(-1)^{|T|}$ with an arbitrary function $L(T)$; we also need to replace $\hat{f}(T)^2$ with $|\hat{f}(T)|^2$, for reasons explained below:

$$\sum_T L(T)|\hat{f}(T)|^2 = 0. \tag{5}$$

In the general case, the spectral transform is not necessarily real (this generality will only be needed in §3.4). In that case, we need to adjust the inner product to adjust for that fact, taking the complex conjugate of one of the sides; Parseval's identity will then result in an expression of the form $|\hat{f}(T)|^2$.

In order to get such an equation, we follow the steps of the proof of Theorem 3.5. In the general case, we have more freedom: we get one equation of the form (5) for each "consistency property". If we have many consistency properties (which will be the case in all the examples below), we can take any linear combination of them and get some equation of the form (5). Alternatively, we can push the linear combination in and construct a *weighted* Cayley graph.

Equation (5) provides us with an upper bound on the measure of f through Hoffman's bound [37], which we present first syntactically.

Lemma 3.6 (Hoffman)**.** *Suppose that*

$$\hat{f}(\varnothing) = \sum_T |\hat{f}(T)|^2 = \mu.$$

Assume $L(\varnothing) = M$ and $\min_T L(T) = m$. Then equation (5) implies

$$\mu \leq \frac{-m}{M-m}. \tag{6}$$

In case of equality, $\hat{f}$ is supported on $\varnothing$ and the coefficients T for which $L(T) = m$. In case of almost equality, most of the weight of $\hat{f}$ is supported on these coefficients[5]*.*

Proof. This is a simple, but instructive, arithmetic computation. Substituting the givens in equation (5),

$$0 = M\mu^2 + \sum_{T \neq \varnothing} L(T)|\hat{f}(T)|^2 \geq M\mu^2 + m(\mu - \mu^2).$$

The bound (6) now follows by isolation of μ. In the case of equality, the inequality in the calculation above must be tight, and so $\hat{f}(T)^2 = 0$ whenever $T \neq \varnothing$ and $L(T) > m$. The case of almost equality is similar. ☐

Consider now an arbitrary Cayley graph G. Applying G to a function f (i.e. multiplying the adjacency matrix G with the vector f) is the same as convolving f with the 'neighborhood function' of the graph. In §3.2 this neighborhood function is given by $\mathbf{1}_{[n]}$. In general we have an arbitrary linear combination of several 'constraints' $\mathbf{1}_S$. The convolution property now implies that G acts diagonally on the spectral coefficients (i.e. it multiplies each of them by a constant); in the case of §3.2, by multiplying $\hat{f}(T)$ by $(-1)^{|T|}$. Looking now at the derivation of equation (5) in that section, we come to the realization that $L(T)$ is the eigenvalue of G corresponding to χ_T. This leads us to the following semantic form of Hoffman's bound.

Lemma 3.7 (Hoffman)**.** *Suppose f is the characteristic function of an independent set in a weighted Cayley graph G. Denote by μ the measure of f, i.e. $\mu = \hat{f}(\varnothing)$. Let $L(T)$ be the eigenvalue of G corresponding to χ_T in the corresponding spectral basis. If $L(\varnothing) = M$ and $\min_T L(T) = m$ then*

$$\mu \leq \frac{-m}{M-m}.$$

[5] What is actually required here is a *spectral gap*. We need that whenever $L(T) > m$, in fact $L(T) \geq m + \gamma$ for some constant γ not depending on n. This will be the case in all our examples.

Proof. Using Parseval's identity, we derive equation (5) from the independence of f following the steps of §3.2. Now apply the previous lemma. $\square$

We can now outline the general method.

1. Represent a family $\mathcal{F}$ with its characteristic function f.
2. Find a group in which the fact that $\mathcal{F}$ is intersecting translates to the fact that it is an independent set in some Cayley graph.
3. Weight the edges of the Cayley graph accordingly.
4. Apply Hoffman's bound to get an upper bound on the measure of $\mathcal{F}$. In many cases the measure corresponds to $|\mathcal{F}|$, but not so in §6.
5. In the case of equality, conclude that the spectral expansion of $\mathcal{F}$ is concentrated on the low-level coefficients. Use some combinatorial argument to deduce the structure of $\mathcal{F}$.
6. In the case of near-equality, approximate f by a function concentrated on the low-level coefficients using some analytical machinery (usually this amounts to applying some hypercontractive inequalities), and deduce stability.
7. Strong stability requires heavier analytical machinery.

Given a group, the proper spectral representation of it is studied by Representation Theory. In the case of abelian groups, the spectral basis consists of all characters. In the general case, these are not enough; see §7 for the solution.

A version of Hoffman's bound applies to cross-intersecting families; recall that these are two families $\mathcal{F}, \mathcal{G}$ such that any $F \in \mathcal{F}$ and $G \in \mathcal{G}$ intersect.

Lemma 3.8 (AKKMS [6])**.** *Let G be a Cayley graph with principal eigenvalue M (corresponding to $\chi_\varnothing$) such that all eigenvalues are bounded* in magnitude *by m. If f, g are families of measures $\mu(f), \mu(g)$ such that no edge of G connects a member of f to a member of g then*

$$\sqrt{\mu(f)\mu(g)} \leq \frac{m}{M+m}.$$

Proof. Straightforward application of Cauchy-Schwartz. $\square$

Lemma 3.8 makes the additional assumption over Lemma 3.7 that not only are all eigenvalues at least $-m$, they are also at most m. In return we get a more general bound. We can now apply all the steps listed above to get a cross-intersecting version of our result, including an upper bound on $\mu(f)\mu(g)$, uniqueness of optimal families and even stability.

3.4 Color agreement problems

A subset of $[n]$ can be viewed as a 2-coloring of $[n]$. Two sets A, B are intersecting if there is some coordinate on which $A(i) = B(i) = 1$. It is natural to generalize this condition and only require that $A(i) = B(i)$. Let's call a family satisfying this requirement — any two sets agree on some coordinate — an *agreeing* family.

Our proof in §3.2 relied on the fact that a set and its complement are not intersecting. They are also not agreeing, and so the whole argument generalizes; not only the trivial $1/2$ upper bound, but also the spectral proof! Reflecting on the general plan layed out in the preceding section, we see that adding an element of a group meshes as perfectly with agreeing families as it does with intersecting families. Therefore all of our arguments will actually work for agreeing families as well as for intersecting ones.

This general phenomenon was already noticed by Chung, Graham, Frankl and Shearer [9], who proved it combinatorially.

Theorem 3.9 (CGFS)**.** *Suppose $\mathcal{F}$ is an $\mathcal{S}$-agreeing family (any two members of $\mathcal{F}$ agree on some $S \in \mathcal{S}$), where $\varnothing \notin \mathcal{S}$. There is an $\mathcal{S}$-intersecting family with the same cardinality.*

Proof. Starting with $\mathcal{F}$, we massage it into a *monotone* family, one which is closed under taking supersets (that is, closed under adding elements $i \in [n]$). Two members of a monotone family cannot agree only on 0-coordinates, since changing all of them to 1 results in a pair of complementary sets.

We monotonize the family by repeatedly trying to monotonize the ith coordinate. To that effect, notice that if we replace all sets S which violate monotonicity at i *at once* by $S \cup i$, then the resulting family is still $\mathcal{S}$-intersecting. This process eventually converges. □

Instead of 2-coloring $[n]$, we can instead color it with an arbitrary number of colors. In §4 we consider the case of constant $c > 2$; it is much nicer than the case $c = 2$ since we have a description of all maximal families. In §7 we consider the case $c = n$ with the additional restriction that all colors are distinct.

3.5 Overview of Examples

In the sequel we present four applications of the general method.

Color-agreeing families are discussed in §4, following [5]. The treatment is straightforward. Shinkar [48] generalized the argument for multiply-intersecting families; we do not discuss this result here.

Triangle-intersecting families of graphs are discussed in §5, following [17]. The starting point [9] is the observation that, even ignoring a bipartite set of edges B, the family is edge-intersecting, i.e. any two graphs contain an edge in common outside of B. This means that in any weighted Cayley graph whose edges are *complements* of bipartite graphs, the family is an independent set. Engineering the weights, we get that triangle-juntas (supergraphs of a fixed triangle) are optimal, proving a conjecture by Simonovits and Sós [49].

Intersecting families of sets are discussed in §6, following [30]. The goal is to prove Erdős-Ko-Rado using spectral methods. The starting point is a smoothing of the problem §6.1, to make it amenable to our methods. Instead of restricting the size of the sets, we employ a weighting scheme which emphasizes sets of the required size. We adapt the Fourier transform accordingly. The method is generalized to t-intersecting problems using the same device of 'ignoring' certain parts of the input (also used in §5)[6], in this case sets of cardinality at most $t - 1$. To wrap things up, we need to translate results from the smooth case to the discrete case.

The proof requires adaptation of the Fourier transform. Our presentation develops this adaptation in an ad-hoc manner. A generalization of the construction appears in §6.5, following [13].

Intersecting families of permutations are discussed in §7, following [15, 16, 18]. The goal is to prove conjectures of Deza and Frankl [11] and Cameron and Ku [8] regarding t-intersecting families of permutations: for large enough n, these are of size at most $(n - t)!$; size $\Omega((n - t)!)$ implies the family is contained in a t-coset. Ellis [15, 16] also proves Hilton-Milner type results by finding the maximal families not contained in a t-coset.

The proofs follow the method of §3.3, only this time the relevant group is S_n, and so recourse is needed to the representation theory of S_n. The crucial observation is that *the eigenvalue corresponding to an irreducible representation is inversely proportional to its dimension.* The spectral transform of optimal families is supported on irreducible representations of small dimension. The converse relation is slightly more complicated due to the fact that S_n is not simple.

For $t = 1$ the proof of the upper bound is relatively simple, since there is no need to weight the Cayley graph: all derangements (permutations without fixed points) look the same for the relevant representations.

The case of larger t is complicated by the fact the relevant representations are no longer blind to the specific cycle types, and it is not clear how to weight them. Let's see what happens in other cases: in §5, an explicit weighting is exhibited; in §6, the weights are chosen in the only possible way that can achieve an optimal bound (the fact that there is only one such choice makes finding it easy). In the case of §7, we have

[6] The original paper presented things differently, but mentions this point of view as an observation of Ryan O'Donnel.

more edge types to play with than constraints. Therefore we choose a matching number of edge types and solve for the weights; two such choices (since $|S_n/A_n| = 2$) are combined to complete the proof.

Proofs of the stability results are more intricate. In a crucial step, the cross-intersecting version of the upper bound is used to guarantee that the family in question cannot contain both a lot of permutations mapping i to j and a lot mapping i to k.

4 Color-Agreeing Families of Sets

A traffic light is operated by n three-way switches in such a way that if we change the value of every switch (all at once) then the light always changes. Show that the traffic light in fact depends only on one switch.

This charming problem, with a classical inductive solution, is one inspiration for the problem we tackle in this section, following Alon, Dinur, Friedgut and Sudakov [5]. Let's look at the subset S_R of $\mathbb{Z}_3^n$ consisting of all settings of the switches that cause the traffic light to be Red. The problem implies that this is a color-agreeing family: every two settings have a switch in common. Similarly, S_G and S_B are color-agreeing. In this section we show that a color-agreeing family can contain at most 3^{n-1} vectors; this implies that $|S_R| = |S_G| = |S_B|$. We further show that a color-agreeing family of maximal size must be a dictatorship; since the families are disjoint, we get that they all depend on the same coordinate.

Our proof will work for d-ary switches, generalizing our work from §3.2. In the sequel, we always assume that $d \geq 3$. We will show three things:

- A color-agreeing family contains at most d^{n-1} vectors.
- A color-agreeing family of size d^{n-1} depends on one coordinate, i.e. it is of the form
$$\{x \in \mathbb{Z}_d^n : x_i = c\}.$$
- A color-agreeing family almost of size d^{n-1} is close to a family depending on only one coordinate.

Our arguments will also work for the cross-color-agreeing case. For example, if $\mathcal{F}, \mathcal{G}$ are cross-color-agreeing (any $S \in \mathcal{F}$, $T \in \mathcal{G}$ share a common coordinate) then $\sqrt{|\mathcal{F}||\mathcal{G}|} \leq d^{n-1}$.

4.1 Upper bound

Let $D \subset \mathbb{Z}_d^n$ consist of all vectors not containing a zero coordinate. A color-agreeing family $\mathcal{F}$ is thus an independent set in the Cayley graph G with edges corresponding to vectors in D. This is a Cayley graph for the group $\mathbb{Z}_d^n$, and so we need to find the analogue of the Fourier-Walsh transform in $\mathbb{Z}_d^n$. The transform is given by the following basis vectors:

$$\chi_x(y) = \omega^{\langle x,y \rangle}, \quad \omega \text{ a primitive } d\text{th root of unity.}$$

In order to apply Hoffman's bound (Lemmas 3.7 and 3.8), we need to calculate the eigenvalues $L(x)$ of G. Using the fact that $\chi_x(\vec{0}) = 1$,

$$\begin{aligned} L(x) &= (D\chi_x)(\vec{0}) \\ &= \sum_{d \in D} \chi_x(d) = \sum_{d_1,\ldots,d_n \neq 0} \chi_x(d_1, \ldots, d_n) \\ &= \sum_{d_1,\ldots,d_n \neq 0} \prod_{i=1}^{n} \omega^{x_i d_i} = \prod_{i=1}^{n} \sum_{d_i \neq 0} \omega^{x_i d_i}. \end{aligned}$$

Since $\sum_i \omega^i = 0$, the inner sum is equal to -1 if $x_i \neq 0$, and to $d-1$ if $x_i = 0$. Denoting the Hamming weight (number of non-zero coordinates) of x by $|x|$, we deduce that

$$L(x) = (d-1)^n \left(\frac{-1}{d-1} \right)^{|x|}.$$

The eigenvalue corresponding to $\vec{0}$ is $M = (d-1)^n$, and all other eigenvalues are bounded in magnitude by $m = (d-1)^{n-1}$. Therefore the measure of a color-agreeing family is bounded by

$$\frac{m}{M+m} = \frac{(d-1)^{n-1}}{(d-1)^n + (d-1)^{n-1}} = \frac{1}{(d-1)+1} = \frac{1}{d}.$$

4.2 Uniqueness

Looking at the argument in §4.1, we see that a family f can achieve the bound only if its Fourier transform is supported by the first two levels (from the uniqueness part in Hoffman's bound). Since the $\vec{0}$th Fourier coefficient captures its measure, we get that the characteristic function of f has the following form:

$$f = \frac{1}{d}\chi_{\vec{0}} + \sum_{i,c} \alpha_{i,c}\chi_{c\delta_i},$$

where $c\delta_i$ is a vector whose only non-zero coordinate, i, is equal to c. If there are $i \neq j$ and c, b such that $\alpha_{i,c}$ and $\alpha_{j,b}$ are both non-zero, then the Fourier expansion of f^2 would contain $\alpha_{i,c}\alpha_{j,b}\chi_{c\delta_i+b\delta_j}$. However, $f^2 = f$, and so this is impossible. Therefore f is of the form

$$f = \frac{1}{d}\chi_{\vec{0}} + \sum_{c} \alpha_c\chi_{c\delta_i}$$

for some i. In other words, f depends only on coordinate i.

4.3 Stability

In order to prove stability, we use an argument from Friedgut, Kalai and A. Naor [32], which is generalized in [5].

Again looking at the argument in §4.1, we see that if a family is almost of size d^{n-1} then most of its Fourier expansion is concentrated on the first two levels. Take the function g formed by removing all other Fourier coefficients. By Parseval's identity, $\|f-g\|_2$ is small. Using a hypercontractive inequality, one can show that g is close to being Boolean, in the sense that $\|g^2 - g\|_2$ is small. Now write g as

$$g = |g|\chi_{\vec{0}} + \sum_i g_i,$$

where g_i is the part depending on coordinate i. On the one hand $\sum_i \|g_i\|^2$ is large (since the family has size close to d^{n-1}), and on the other hand the fact that $\|g^2 - g\|_2$ is small implies that $\sum_{i\neq j}\langle g_i, g_j\rangle$ is small. This implies that there is one coordinate g_i that captures most of g, in the sense that $\|g - |g|\chi_{\vec{0}} - g_i\|_2$ is small.

Let $h = |g|\chi_{\vec{0}} + g_i$. While $\|h - f\|_2$ is small, h is not Boolean and so we are not done yet. However, since h is close to a Boolean function, rounding it to $\{0,1\}$ at most doubles the error[7] and so results in a Boolean function close to f.

5 Triangle-Intersecting Families of Graphs

In 1976, Simonovits and Sós [49] studied intersection problems on graphs. In this setting, we are considering some fixed set of n vertices, and the graphs in questions are all subgraphs of K_n, the complete graph on these n vertices. Thus there are $2^{\binom{n}{2}}$ graphs in total.

One of the problems they considered is triangle-intersecting families of graphs. These are families where every two graphs are triangle-intersecting, that is their intersection contains some triangle. They conjectured

[7] The worst case is when $h = 1/2$ and the values of f are uniformly distributed on $\{0,1\}$. In this case the squared L_2 difference is 1/4, whereas after rounding it is 1/2.

that such a family can contain at most $2^{\binom{n}{2}-3}$ graphs, the optimal families being "triangle-juntas", that is all supergraphs of a fixed triangle.

Simonovits and Sós weren't able to progress much on this question. In 1984, Chung, Graham, Frankl and Shearer [9] proved a bound of $2^{\binom{n}{2}-2}$ using Shearer's entropy lemma (described in §9.5). The conjecture was finally proved by Ellis, Filmus and Friedgut [17] using the general method layed out in §3.3. They showed the following:

- A triangle-intersecting family contains at most $2^{\binom{n}{2}-3}$ graphs.
- If a triangle-intersecting family is of size $2^{\binom{n}{2}-3}$ then it is a triangle-junta.
- A triangle-intersecting family of almost maximal size is close to a triangle-junta.

In fact, the proof goes through for non-bipartite-intersecting graphs. It also holds in a more general non-uniform hypergraph setting: graphs are replaced by subsets of $2^{2^{[n]}}$, and triangles are replaced by Schur triples, that is $X, Y, Z \subset [n]$ satisfying
$$X \triangle Y \triangle Z = \varnothing.$$
Note that this generalizes the graphical setting since a triangle (i, j, k) corresponds to the Schur triple
$$\{i, j\} \triangle \{i, k\} \triangle \{j, k\} = \varnothing.$$

5.1 Upper Bound

If $\mathcal{F}$ is an intersecting family of graphs and $G, H \in \mathcal{F}$ then the intersection $G \cap H$ contains a triangle. In particular, it is not bipartite. Therefore the agreement $\overline{G \triangle H}$ cannot be bipartite. Thus $\mathcal{F}$ is an independent set in the Cayley graph whose edges correspond to all *complements* of bipartite graphs. This is a Cayley graph with respect to the group $\mathbb{Z}_2^{\binom{n}{2}}$ whose spectral transform is the Fourier-Walsh transform described in §3.1.

In contrast to the proof of the bound in §4.1, this time we need to carefully weight the edges. It is a-priori not clear how to do that; there is more than one correct solution. Our approach will be to find what eigenvalues $L(G)$ are achievable, and then find a combination which is of the correct form for an application of Hoffman's bound (Lemma 3.7). We will say that the function $L(G)$ is *spectral* if it is the spectrum of some Cayley graph. The spectral functions form a vector space over $\mathbb{R}$.

Take for a start the Cayley graph formed by the complement of some particular bipartite graph H. The corresponding eigenvalues are
$$L_H(G) = \chi_G(\overline{H}) = (-1)^{|G \cap \overline{H}|} = (-1)^{|G|}(-1)^{|G \cap H|}.$$
All the subgraphs of H are also bipartite, and the functions $\{L_K : K \subset H\}$ are easily seen to span the linear subspace of functions of the form
$$G \mapsto (-1)^{|G|} f(G \cap H).$$
Instead of taking one graph H, we can take a distribution over random graphs H. The expectation of L_H under this distribution is also spectral. We pick the following distribution:

> Pick a random bipartition of $[n]$ by putting each vertex in one of the parts uniformly; output the complete bipartite graph on this bipartition.

For any graph H, let the function f be the indicator for graphs isomorphic to H. We call the resulting spectrum q_H the *Ellis spectrum* for H. The value of $q_H(G)$ is equal to the probability that a random bipartition cuts edges of G which are isomorphic to H.

Similarly, for every integer k we let the function f be the indicator for size k. The resulting Ellis spectrum q_k measures the probability that a random bipartition cuts exactly k edges of G.

We can now describe the required Cayley graph.

Theorem 5.1. *Let the spectral functions L_1, L_2 be given by*

$$L_1(G) = (-1)^{|G|}\left[q_0(G) - \frac{5}{7}q_1(G) - \frac{1}{7}q_2(G) + \frac{3}{28}q_3(G)\right],$$

$$L_2(G) = (-1)^{|G|}\left[\sum_F q_F(G) - q_\square(G)\right],$$

where F runs over all forests of size 4. Furthermore, define $L = L_1 + L_2$.

Then $L(\varnothing) = L_1(\varnothing) = 1$ and $L(G), L_1(G) \geq -1/7$. Furthermore, $L_1(G) = -1/7$ only for the graphs containing one or two edges, for triangles, for forests of size 4 and for diamond graphs (K_4 minus one edge); $L(G) = -1/7$ only for graphs containing one or two edges and for triangles.

The way L_1 was constructed is by listing a small number of graphs and deducing the coefficients of the Ellis functions; here a useful fact is that in any spectrum for which $L(G) \geq -1/7$, we must have $L(G) = -1/7$ for graphs on which the Fourier transform of an optimal family is supported, that is subgraphs of a triangle (this follows from the proof of Hoffman's bound). Given L_1, the properties stated by the theorem are checked by noticing that for large graphs the Ellis functions are small, and so $L_1 \approx 0$. Graphs of middle sizes are dealt with using case-by-case analysis.

The spectrum L_1 is already enough to apply Hoffman's bound and get an upper bound of $1/8$. However, in order to prove uniqueness it is useful to remove the extra graphs on which $L(G) = -1/7$. This is taken care of by the addition of L_2, which distinguishes these graphs from the necessarily-tight subgraphs of the triangle.

5.2 Uniqueness

Suppose a family $\mathcal{F}$ has size exactly $2^{\binom{n}{2}-3}$. Hoffman's bound implies that its Fourier transform is concentrated on the first four levels, since the spectrum L constructed in the previous section is tight only for graphs with at most 3 edges.

The proof is now complete by invoking the following result from Friedgut [30]. It will be useful to use the notation $|f|$ for the *relative* size of f; thus if f is the characteristic function of $\mathcal{F}$ then $|f| = 1/8$.

Lemma 5.2. *Suppose f is a monotone Boolean function on N coordinates whose Fourier-Walsh transform is concentrated on the first $t+1$ levels. Either f is identically zero or $|f| \geq 2^{-t}$; if $|f| = 2^{-t}$ then it is a t-junta (i.e. it consists of all supersets of a fixed set of t points).*

Proof. The proof is by induction on k and N. The base case is $t = 1$. In that case, f is of the form

$$f = |f|\chi_\varnothing + \sum_i c_i \chi_i.$$

Since $f^2 = f$ (due to f begin Boolean), we see that at most one c_i can be non-zero (otherwise a summand of the form $c_i c_j \chi_i \chi_j$ will appear in the Fourier-Walsh expansion of f^2). Thus either f is identically zero, or identically one, or a dictatorship of measure $1/2$.

Now suppose the claim is true for $N-1$ and all t, and for N and $t-1$; we prove it now for N and t. We define two functions f_0, f_1 on $N-1$ coordinates by restricting the last coordinate of f; that is,

$$f_0(x) = f(x, 0), \quad f_1(x) = f(x, 1).$$

Both functions are monotone, and we have $|f| = (|f_0| + |f_1|)/2$ and $|f_0| \leq |f_1|$ (by monotonicity). Assume now that $|f| \leq 2^{-t}$ (otherwise there is nothing to prove). If $|f_0| = 2^{-t}$ then $|f_1| = 2^{-t}$ and the function does not depend on the last coordinate, and so by induction it is a t-junta. Otherwise, by induction f_0 is identically zero. Thus all vectors in x contain the last coordinate, and $|f_1| = 2^{-(t-1)}$. Easy computation shows that for any $S \subset [N-1]$,

$$\frac{1}{2}\hat{f}_1(S) = \hat{f}(S) = -\hat{f}(S \cup N).$$

This implies that the Fourier-Walsh transform of f_1 is supported on the first $(t-1)+1$ levels. The induction hypothesis now implies that either f_1, and so f, is identically zero, or that f_1 is a $(t-1)$-junta, so that f is a t-junta. □

In order to invoke the lemma, $\mathcal{F}$ needs to be monotone. However, a maximal family is always monotone, since otherwise the monotone closure (i.e. all supergraphs of graphs in the family), which is also triangle-intersecting, is larger than the original family. The lemma implies that $\mathcal{F}$ is a junta depending on three edges, which must form a triangle.

Here is an alternative proof of Lemma 5.2 by Shinkar [48], which generalizes to the color-agreeing case.

Proof. Denote by $f_{T,t}$ the restriction of f to the coordinates in T, obtained by substituting t for the other coordinates. The Fourier expansion of $f_{T,t}$ is related to that of f by

$$\begin{aligned}
\hat{f}_{T,t}(S) &= 2^{-|T|} \sum_{U \subset T} f(U,t)\chi_S(U) \\
&= 2^{-n} \sum_{\substack{U \subset T \\ V,W \subset [n]\setminus T}} f(U,V)\chi_S(U)\chi_W(V)\chi_W(t) \\
&= \sum_{W \subset [n]\setminus T} \hat{f}(S,W)\chi_W(t).
\end{aligned}$$

We claim that $\hat{f}$ must have some non-zero coefficient of size t. Indeed, suppose that $\hat{f}(S) \neq 0$ for some $|S| < t$. Consider $f_{S,s}$ for all possible values of s. Since $|f| = 2^{-t}$, there must be some s such that $|f_{S,s}| \leq 2^{-t}$. Now $f_{S,s}$ is a Boolean function on $|S| < t$ coordinates, so either $f_{S,s} = 0$ or $|f_{S,s}| \geq 2^{-|S|}$. In our case the first option must be true, $f_{S,s} = 0$. The formula above implies that there must be some Fourier coefficient corresponding to a superset of S which is non-zero. Continuing this way, we eventually reach a non-zero coefficient $\hat{f}(T) \neq 0$ on level t.

Consider now all possible restrictions $f_{T,t}$. Since $|f| = 2^{-t}$, the average value of $|f_{T,t}|$ is 2^{-t}. Since f is concentrated on the first $t+1$ levels, the formula above guarantees that $\hat{f}_{T,t}(T) \neq 0$. Thus $f_{T,t} \neq 0$, hence $|f_{T,t}| \geq 2^{-t}$. This shows that in fact $|f_{T,t}| = 2^{-t}$. Since f is monotone, $f_{T,t}$ must be the characteristic function of T. Therefore f is a T-junta. □

Shinkar's argument generalizes to the color-agreeing case. In that case, we do not require monotonicity, but use the fact that there are more than two colors to show that $f_{T,t}$ must be the same function for every t.

5.3 Stability

Suppose now we have a triangle-intersecting family $\mathcal{F}$ of measure almost $1/8$. Hoffman's bound implies that most of its Fourier-Walsh expansion is concentrated on the first four levels. A deep theorem of Kindler and Safra [43] implies that $\mathcal{F}$ is well-approximated by a family $\mathcal{G}$ depending on at most K coordinates, for some constant K not depending on n. There is a finite number of such families which are *not* triangle-juntas, and so there is a constant ϵ such that if the measure of $\mathcal{G}$ is at least $1/8-\epsilon$, it is necessarily a triangle-junta. If the measure of $\mathcal{F}$ was close enough to $1/8$ then this condition will be satisfied, so that $\mathcal{F}$ is well-approximated by some triangle-junta.

6 Intersecting Families of Sets

The Erdős-Ko-Rado theorem has many proofs, some of them spectacularly simple (like the one in §9.3). Here we present a simple proof by Friedgut [30] with the added benefit that stability comes out from the proof at no further cost! The proof even generalizes to the setting of Ahlswede-Khachtrian, when the parameters are such that the optimal t-intersecting family is a t-junta.

Our goal in this section is proving the following results:

- Erdős-Ko-Rado: For $k < n/2$, a k-uniform intersecting family on $[n]$ contains at most $\binom{n-1}{k-1}$ sets. The unique families achieving this bound are dictatorships (supersets of a fixed element).
- Ahlswede-Khachtrian (special case): For $t > 1$ and $k < n/(t+1)$, a k-uniform t-intersecting family on $[n]$ contains at most $\binom{n-t}{k-t}$ sets. The unique families achieving this bound are t-juntas (supersets of a fixed set of t elements).
- Stability results of the above: a family of almost maximal measure is close to an optimal family.

More stability results are proved using heavier spectral machinery, as well as some combinatorial arguments, in Dinur and Friedgut [12].

From the spectral perspective, it is more natural to rephrase uniform intersecting questions (ones in which the size of the set is constrained to some set size) in a *smoothed* way. This is discussed in §6.1, where we also show how to convert from the *discrete* case to the smooth case and vice versa. The proof of the upper bound is discussed in §6.2 (Erdős-Ko-Rado) and §6.3 (Ahlswede-Khachtrian). Uniqueness and stability are discussed in §6.4.

6.1 Smoothing

Intersecting results come in two flavors: non-uniform and uniform. Non-uniform results don't care about the size of the sets. Examples are the results of §3.2,4,5, as well as Katona's theorem on maximal (unconstrained) t-intersecting families (see [4]). Other results, such as Erdős-Ko-Rado and Ahlswede-Khachatrian, are uniform: all sets must be of some prescribed size[8].

From a spectral perspective, non-uniform results are much easier to handle than uniform results. The reason is that it is difficult to express the uniformity constraint in the spectral language. Instead on enforcing this constraint explicitly, we will encourage it implicitly by choosing a measure on $2^{[n]}$ which concentrates on sets of size roughly k.

Here are two versions of the Erdős-Ko-Rado theorem: the first one is the *discrete version*, and its counterpart is the *smoothed version*.

Theorem 6.1 (EKR (discrete))**.** *If $k < n/2$ then a k-uniform intersecting family of sets on $[n]$ contains at most k/n of the k-sets; the optimum is achieved only on dictatorships.*

Definition 6.1. For any $p \in (0,1)$, the μ_p measure on $2^{[n]}$ is defined as follows. Let $\mathbf{S}$ be a random subset of $[n]$ obtained by putting any $i \in [n]$ in $\mathbf{S}$ with probability p. Then

$$\mu_p(\mathcal{F}) = \Pr[\mathbf{S} \in \mathcal{F}].$$

In other words, for $T \subset [n]$ we have

$$\mu_p(T) = p^{|T|}(1-p)^{n-|T|},$$

and the measure is extended additively for $2^{[n]}$.

Theorem 6.2 (EKR (smoothed))**.** *If $p < 1/2$ then an intersecting family of sets on $[n]$ has μ_p-measure at most p; the optimum is achieved only on dictatorships.*

The analogue of the parameter k in the discrete version is $p = k/n$ in the smoothed version. The size of the set $\mathbf{S}$ in Definion 6.1 is distributed $\mathrm{Bin}(n,p)$, and so it is concentrated around $nk = p$. So the μ_p measure singles out sets of size roughly k. We will see later that it is easy to handle μ_p spectrally.

The smoothed version doesn't capture the discrete version exactly. Indeed, when $p = 1/2$ then μ_p is just the counting measure, i.e. $\mu_{1/2}(\mathcal{F}) = |\mathcal{F}|/2^n$, so that $\mu_{1/2}$ actually corresponds to the *non-uniform* case. However, we show in §6.1.2 how to convert smoothed results to discrete results incurring minimal loss. The other direction is easier and will be tackled first in §6.1.1.

[8] Another class of results, exemplified by the Ray-Chaudhuri-Wilson theorem, concern families with allowable sizes and allowable intersection sizes.

6.1.1 Discrete to Smooth

In 2005, Dinur and Safra [14] came up with a simple way of converting discrete results to smoothed ones, in the context of constructing a PCP for vertex cover[9]. The result that interested them was the 2-intersecting case of Ahlswede-Khachatrian, but we demonstrate their method using Erdős-Ko-Rado.

Let $p < 1/2$, and consider some intersecting family $\mathcal{F}$ on $[n]$. We are trying to prove Theorem 6.2 (smoothed EKR) using Theorem 6.1 (discrete EKR). Intuitively, discrete EKR allows us to bound the number of k-sets in $\mathcal{F}$ for $k \leq n/2$. Most of the μ_p-measure of $\mathcal{F}$ is concentrated around sets of size $np < n/2$, so we have the correct bound on most of $\mathcal{F}$.

Here's an implementation of our intuition:

$$\begin{aligned}
\mu_p(\mathcal{F}) &= \sum_{k=1}^{n} p^k(1-p)^{n-k} \left|\mathcal{F} \cap \binom{[n]}{k}\right| \\
&\leq \sum_{k=1}^{n/2} p^k(1-p)^{n-k}\binom{n-1}{k-1} + \sum_{k=n/2+1}^{n} p^k(1-p)^{n-k}\binom{n}{k} \\
&= \sum_{k=1}^{n} p^k(1-p)^{n-k}\binom{n-1}{k-1} + \sum_{k=n/2+1}^{n} p^k(1-p)^{n-k}\binom{n-1}{k} \\
&= p + (1-p)\Pr[\mathrm{Bin}(n-1,p) > n/2].
\end{aligned}$$

For large n, the error term $\Pr[\mathrm{Bin}(n-1,p) > n/2]$ is small; in fact, it gets smaller as n increases! Now the original intersecting family $\mathcal{F}$ can be extended to an intersecting family on $[N]$ for any $N > n$ by ignoring the $N-n$ extra points. The μ_p-measure of the new family is the same, and so by taking N to be arbitrarily large, we obtain $\mu_p(\mathcal{F}) \leq p$.

A different interpretation of this idea of 'adding infinitely many points' is in §9.3.

6.1.2 Smooth to Discrete

Freidgut [30] succeeded in translating his results in the smooth case (presented in §6.2–6.4) to the discrete case, with some loss. We present the argument for Erdős-Ko-Rado. More intricate arguments of the same nature are found in Dinur and Friedgut [12], who show that intersecting families of non-negligible measure are approximated by intersecting juntas.

What Friedgut was able to show is the following.

Theorem 6.3. *Let k be arbitrary. There is a probability $p \approx k/n$ (in fact, slightly larger than k/n) such that if $\mathcal{F}$ is a monotone family then*

$$\mu_p(\mathcal{F}) \gtrapprox \frac{\left|\mathcal{F} \cap \binom{[n]}{k}\right|}{\binom{n}{k}}.$$

The proof is very similar to the argument in §6.1.1. Given the number of k-sets in $\mathcal{F}$, we can bound the number of l-sets in $\mathcal{F}$ for $l > k$. Sets of cardinality smaller than k, however, are not accounted for. This is the source of error in the argument, and is also the reason we take p slightly larger than k/n.

Friedgut uses this argument to deduce discrete stability results from smooth stability results. Suppose $\mathcal{D}$ is a k-uniform intersecting family of almost maximal size. Its *up-set* $\mathcal{S}$, consisting of all supersets of sets in $\mathcal{D}$, is a monotone intersecting family of almost maximal measure (first use of Theorem 6.3). Therefore by smooth stability, it is close to some dictatorship $\mathcal{S}'$, which we can project to a k-uniform dictatorship $\mathcal{D}'$. The distance between $\mathcal{D}$ and $\mathcal{D}'$ is bounded by the distance between $\mathcal{S}$ and $\mathcal{S}'$ (second use of Theorem 6.3), and we get discrete stability.

[9]Actually the PCP is for h-independent set, the problem of finding the maximal set of vertices without a clique of size h.

6.2 Erdős-Ko-Rado

Our goal in this section is to prove Theorem 6.2. The general method, as outlined in §3.3, allows us to bound the uniform measure of the characteristic function f of a family, which is a normalized version of the cardinality of the family. Theorem 6.2, however, requires us to use the measure μ_p instead. How do we adapt the general method to handle this case?

Let us review the proof of Hoffman's bound 3.7:

1. If a family f is an independent set in a weighted Cayley graph G then $\langle f, Gf\rangle = 0$.
2. Parseval's identity (following from orthonormality of the spectral basis) implies $\sum_i \hat{f}(i)\widehat{Gf}(i) = 0$.
3. The eigenvectors of Gf are the spectral basis (with eigenvalues $L(i)$), and so $\sum_i L(i)|\hat{f}(i)|^2 = 0$.
4. The 0th basis vector is $\vec{1}$, and so $|f| = \langle f, f\rangle = \langle f, \vec{1}\rangle = \hat{f}(0)$.
5. Simple arithmetic now deduces a bound on $|f|$ from bounds on the $L(i)$.

The measure of f shows up in the identity

$$|f| = \langle f, f\rangle = \langle f, \vec{1}\rangle.$$

In order to get $\mu_p(f)$ instead of $|f|$, we need to replace the inner product with a skewed version. Let us define the p-skewed inner product accordingly:

$$\langle f, g\rangle_p = \sum_x \mu_p(x) f(x) g(x).$$

The usual Fourier basis is not orthonormal (or even orthogonal) with respect to the new inner product, and so we have to replace it. Since we're looking for a tensorial basis, it is enough to consider the case $n = 1$. We are thus looking for two vectors $\chi^p_\varnothing, \chi^p_{\{1\}}$ which form an orthonormal basis with respect to $\langle\cdot,\cdot\rangle_p$.

There are many different ortonormal bases. The proof of Hoffman's bound assumes that

$$\langle f, f\rangle_p = \langle f, \chi^p_\varnothing\rangle_p,$$

and so we can take $\chi^p_\varnothing = \begin{pmatrix} 1 & 1\end{pmatrix}'$; note that this is a vector of norm 1. Given $\chi^p_\varnothing$, the other vector $\chi^p_{\{1\}}$ can be found using simple arithmetic. Suppose that $\chi^p_{\{1\}} = \begin{pmatrix}\alpha & \beta\end{pmatrix}'$. Then

$$(1-p)\alpha + p\beta = 0,$$
$$(1-p)\alpha^2 + p\beta^2 = 1.$$

These equations define α, β up to sign. We choose $\chi^p_{\{1\}}$ so that putting $p = 1/2$ would recover the usual Fourier transform. We obtain the following basis for $n = 1$:

$$\chi^p_\varnothing = \begin{pmatrix} 1 & 1\end{pmatrix}', \quad \chi^p_{\{1\}} = \begin{pmatrix}\sqrt{\frac{p}{1-p}} & -\sqrt{\frac{1-p}{p}}\end{pmatrix}'.$$

The resulting basis for general n is obtained by tensoring:

$$\chi^p_S(T) = \prod_{i=1}^n \left(\sqrt{\frac{p}{1-p}}\right)^{|S\setminus T|}\left(-\sqrt{\frac{1-p}{p}}\right)^{|S\cap T|}.$$

The next step is to find the counterpart of the Cayley graph G. Again, we will first tackle the problem for $n = 1$ and then tensorize. Denote the Cayley graph for this case by G_1. The linear operator G_1 satisfies two properties: its eigenvectors are $\chi^p_\varnothing, \chi^p_{\{1\}}$; and $e'_{\{1\}} G_1 e_{\{1\}} = 0$, where e_1 is the basis vector corresponding to $\{1\}$. The latter property guarantees that $f'G_1f = 0$ for an intersecting family (for $n = 1$ the only way to intersect is through the element 1). Writing G_1 out in matrix form, we get

$$G_1 = \begin{pmatrix} \alpha & \beta \\ \gamma & 0 \end{pmatrix}.$$

Note the 0 at the corner, which guarantees that $e'_{\{1\}}G_1e_{\{1\}} = 0$. Without loss of generality, we can assume that $G_1\chi^p_\varnothing = \chi^p_\varnothing$ (since we can always multiply G_1 by a scalar). Therefore $\alpha + \beta = \gamma = 1$, and the matrix looks as follows:

$$G = \begin{pmatrix} \alpha & 1-\alpha \\ 1 & 0 \end{pmatrix}.$$

Suppose that $\chi^p_{\{1\}}$ is an eigenvector of G_1 with eigenvalue λ. Then

$$\begin{pmatrix} \lambda\sqrt{\frac{p}{1-p}} \\ -\lambda\sqrt{\frac{1-p}{p}} \end{pmatrix} = \lambda\chi^p_{\{1\}} = G_1\chi^p_{\{1\}} = \begin{pmatrix} \alpha\sqrt{\frac{p}{1-p}} - (1-\alpha)\sqrt{\frac{1-p}{p}} \\ \sqrt{\frac{p}{1-p}} \end{pmatrix}.$$

We conclude easily that

$$\lambda = -\frac{p}{1-p}, \quad \alpha = \frac{1-2p}{1-p}.$$

The Cayley graph G in the general case is obtained by taking the n-fold tensor power $G = G_1^{\otimes n}$. Its eigenvalues are

$$G\chi^p_S = \left(-\frac{p}{1-p}\right)^{|S|}\chi^p_S.$$

There is another way of thinking about G. Denote by G_i an application of G_1 on the ith coordinate. As an operator on $\mathbb{R}[\mathbb{Z}_2^n]$, G_i is an operator which acts independently on all 2^{n-1} copies of $\mathbb{R}[\mathbb{Z}_2^n]$ of the form $(x_1, \ldots, x_{i-1}, \mathbb{Z}_2, x_{i+1}, \ldots, x_n)$. The operators G_i all commute, and $G = G_1 \cdots G_n$.

We are now ready to prove the theorem. The maximal eigenvalue of G is $M = 1$, and all other eigenvalues are bounded in magnitude by $m = p/(1-p)$ (since $p < 1/2$). Hoffman's bound thus implies that the μ_p-measure of an intersecting family is at most

$$\frac{m}{M+m} = \frac{p/(1-p)}{1+p/(1-p)} = \frac{p}{1-p+p} = p.$$

Families achieving this bound must have their spectral expansion concentrated on the first two levels.

6.3 Ahlswede-Khachatrian

We extend the argument of the previous section to prove the following special case of smoothed Ahlswede-Khachatrian.

Theorem 6.4 (AK (smoothed))**.** *If $p < 1/(t+1)$ and $\mathcal{F}$ is a t-intersecting family on $[n]$ then $\mu_p(\mathcal{F}) \le p^t$. Moreover, if the bound is achieved then $\mathcal{F}$ is a t-junta.*

The Cayley operator we used in §6.2 is the counterpart of convolution with $[n]$ in the group $\mathbb{Z}_2^n$ (our method in §3.2). In a t-intersecting family, any two sets A, B intersect in at least t points. Thus their symmetric difference $A \triangle B \subset \overline{A \cap B}$ contains at most $n-t$ points. Therefore, a t-intersecting family is an independent set in the Cayley graph where edges correspond to convolution with sets of size more than $n-t$.

What corresponds to convolution with an arbitrary set T in our setting? Convolution with $[n]$ corresponds to applying G_1 in each coordinate. Therefore convolution with T must correspond to applying G_1 only in the coordinates belonging to T! Since coordinates not in T are 'ignored', we see that for every set T we can get a Cayley graph with spectrum

$$L_T(S) = \left(-\frac{p}{1-p}\right)^{|S\cap T|}.$$

We seek a combination of these spectra which will allow us to apply Hoffman's bound, namely:

- Assume for simplicity that $M = L(\varnothing) = 1$.
- We need $m = -p^t/(1-p^t)$ (that's a simple calculation).
- In order to get uniqueness, we need $L(S) > m$ for $|S| > t$.
- Since the spectral expansion of a T-junta is non-zero on all subsets of T, the proof of Hoffman's bound implies that $L(S) = m$ for $0 < |S| \leq t$.

A symmetry argument shows that we can assume that the coefficient of L_T in such a combination depends only on $|T|$. For $|T| = n-1$, summing over all possible T's we get

$$L_1(S) \triangleq \sum_{i \in [n]} L_{[n]\setminus i}(S) = \sum_{i \in [n]} \left(-\frac{p}{1-p}\right)^{|S\setminus i|} = \left(-\frac{p}{1-p}\right)^{|S|} \left[(n-|S|) - \left(\frac{1-p}{p}\right)|S|\right].$$

Similarly, summing over all T such that $|T| = n-s$ we get that

$$L_s(S) \triangleq \sum_{|T|=n-s} L_T(S) = \left(-\frac{p}{1-p}\right)^{|S|} P_s(|S|)$$

for some polynomial P_s of degree s. Thus the span of the functions L_s for $0 \leq s < t$ is exactly

$$\{S \mapsto \left(-\frac{p}{1-p}\right)^{|S|} P(|S|) : \deg P < t\}.$$

We are in good shape, since we know $t+1$ points of the $t-1$-degree polynomial P! In fact, since P only has t coefficients, we might wonder whether any polynomial P at all satisfies the requirements. Notice, however, that if we choose a polynomial P such that the resulting spectrum L satisfies $L(\varnothing) = 1$ and $L(S) = m$ for $|S| < m$, it must also satisfy $L(S) = m$ for $|S| = m$. This follows from substituting the spectral expansion of a t-junta inside the proof of Hoffman's bound.

Having found the one and only polynomial satisfying the first three requirements above, it remains to see whether the last requirement is also satisfied, that is whether $L(S) > m$ for $|S| > t$. Elementary methods (see [30]) show that this is indeed the case, as long as $p < 1/(t+1)$. When $p = 1/(t+1)$, we get extra eigenvalues equal to m. And when $p > 1/(t+1)$, the proof breaks completely, and for a good reason: t-juntas are no longer optimal. In §8 we discuss what happens when we try to generalize the proof to this case.

6.4 Uniqueness and Stability

In §6.2–6.3 we proved that a t-intersecting family has measure at most p^t, as long as $p < 1/(t+1)$. We also proved that a family matching this bound must have its spectral transform concentrated on the first $t+1$ levels. A straightforward generalization of the method of §5.2 shows that this implies that the family is a t-junta. Stability follows from a similar generalization of the method of §5.3.

6.5 General Method

We constructed the p-skewed Fourier transform in §6.2 by listing the properties that the purported transform has to satisfy. In this section we show how to generalize this idea, following Dinur, Friedgut and Regev [13].

Our starting point is some measure μ on $[m]$, which we think of as a set of colors. The measure μ is intended as a generalization of the one-dimensonal version of μ_p, which is a measure on [2]. The meausre μ defines a μ-skewed inner product

$$\langle f, g\rangle_\mu = \sum_{x \in [m]} \mu(x) f(x) g(x).$$

Hoffman's bound requires us to come up with some operator G which has the following two properties:

- The operator G has a complete set of eigenvectors χ_i which are orthonormal with respect to $\langle \cdot, \cdot \rangle_\mu$.
- The constant vector is an eigenvector of G.

We know that a (real) matrix has an orthonormal set of eigenvectors under the usual inner product if it is symmetric. What condition should hold in the present case? Let D denote the diagonal linear operator which multiplies the ith coordinate by $\mu(i)$. Two vectors u, v are μ-orthogonal if $u' D v = 0$. It is thus natural to consider the normalized eigenvectors

$$\psi_i = \sqrt{D}\chi_i.$$

These eigenvectors satisfy the orthogonality relations

$$\psi_i' \psi_j = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

We can think of D as a change-of-basis matrix. It then becomes clear that the ψ_i must be eigenvectors of the matrix

$$H = \sqrt{D} G \sqrt{D}^{-1},$$

with the same eigenvalues. Since the ψ_i form an orthonormal basis, we conclude that $\sqrt{D}G\sqrt{D}^{-1}$ must be symmetric.

So far for the first property. The second property that G has to satisfy is that the constant 1 vector, which we denote by $\chi_\varnothing$, must be an eigenvector (note that $\chi_\varnothing$ has unit norm since μ is a distribution). Without loss of generality, we can assume that the corresponding eigenvalue is 1. We conclude that all rows of G must sum to 1.

Suppose now that all entries of G are non-negative. We can then think of G' as a Markov chain. Since H is symmetric, we get the formula

$$G' = (\sqrt{D}^{-1} H \sqrt{D})' = \sqrt{D} H \sqrt{D}^{-1} = D G D^{-1}.$$

Consider now $\psi_\varnothing$, whose entries are just $\psi_\varnothing(x) = \mu(x)$. We have

$$G' \psi_\varnothing = D G D^{-1} \psi_\varnothing = D G \chi_\varnothing = D \chi_\varnothing = \psi_\varnothing.$$

Therefore μ is the stationary probability of G'. The condition $G' = DGD^{-1}$ just states that G' is a *reversible Markov chain*.

Let's see how all of this works for G_1 of §6.2. The (reversible) Markov chain in question is

$$G_1' = \begin{bmatrix} 1 - \frac{p}{1-p} & 1 \\ \frac{p}{1-p} & 0 \end{bmatrix}.$$

A particle in state 1 switches to state 2 with probability $p/(1-p)$. A particle in state 2 always switches to state 1. The stationary probability is $\begin{pmatrix} 1-p & p \end{pmatrix}'$.

7 Intersecting Families of Permutations

Deza and Frankl [11] considered in 1977 the problem of intersecting families of permutations. A permutation $\pi \in S_n$ is identified with its graph $\{(i, \pi(i)) : i \in [n]\}$, and two permutations are said to intersect if their graphs intersect. Put differently, $\alpha, \beta \in S_n$ intersect if $\alpha(i) = \beta(i)$ for some i.

Using a very simple partition argument (like the linear subspace approach mentioned briefly in §9.5), Deza and Frankl proved that an intersecting family of permutations in S_n can contain at most $(n-1)!$ points: for every π, the family can contain at most one of the n permutation $\pi_i = (1 \cdots n)^i \pi$, since $\pi_i(x) = \pi(x) + i$. They conjectured that the unique optimal families are dictatorships of the form $\{\pi \in S_n : \pi(i) = j\}$, but couldn't prove it. Only in 2003 were Cameron and Ku [8] (and independently Larose and Malvenuto [45]) able to prove Deza and Frankl's conjecture.

Deza and Frankl also conjectured that for every k, a k-intersecting family of permutations in S_n is of size at most $(n-k)!$, *for big enough n*. This conjecture remained open until 2009, when it was proved by Ellis [15] and (independently) Friedgut and Pilpel [18], using spectral methods. Ellis [15] also proved a Hilton-Milner type result, conjectured by Cameron and Ku: if an intersecting family of permutations isn't contained in a dictatorship then, for large enough n, its size is bounded by that of the unique optimal families of the form

$$\{\pi \in S_n : \pi(1) = 1 \text{ and } \pi(i) = i \text{ for some } i > 2\} \cup \{(12)\}.$$

Later, Ellis [16] extended his Hilton-Milner result to the case of t-intersecting permutations, and generalized the entire argument to A_n.

The representation theory of S_n is wildly different than that of the groups used in §4 and §6.5. In those sections, the group in question was abelian, and so the spectral transform was determined by multiplicative characters. The group S_n, however, is not abelian, and so the relevant spectral transform is more complicated, and requires the full power of representation theory.

We motivate and explain the rudiments of representation theory in §7.1. The specifics of the representation theory of S_n are dealt with in §7.2. The case of 1-intersecting families of permutations is described in §7.3–7.5, following [15]. The general case is presented in §7.6, following [16, 17].

7.1 Representation Theory

The general method, as outlined in §3.3, starts by identifying a Cayley graph G of some group H, with the following property: an intersecting family $\mathcal{F} \subset H$ is an independent set in G. Hoffman's bound (Lemma 3.7) then implies some bound on $|\mathcal{F}|$.

The proof of Hoffman's bound uses several properties of the Fourier transform. First, the Fourier transform is an orthonormal linear transformation. Second,

$$\hat{f}(\varnothing) = \frac{1}{|H|} \sum_{s \in H} f(s).$$

Third, if g is obtained from f by convolution with some fixed element $s \in H$, then

$$\hat{g}(t) = \chi_s(t)\hat{f}(t).$$

The function χ_s is known as a character.

In the abelian case, the Fourier transform is constructed by finding an orthonormal basis of multiplicative characters. Multiplicativity of the characters immediately implies the convolution property. In the non-abelian case, however, we don't have enough multiplicative characters. For example, S_n only has two multiplicative characters $\chi^{(n)}, \chi^{(1^n)}$ given by $\chi^{(n)}(\pi) = 1$ and $\chi^{(1^n)}(\pi) = (-1)^\pi$, known as the trivial representation and the sign representation, respectively. Therefore the construction will proceed differently.

We are going to decompose the space $\mathbb{C}[H]$ of all complex-valued functions on H. Call a subspace V of $\mathbb{C}[H]$ *invariant* if it is closed under convolution with an arbitrary function. If V is a minimal invariant subspace, that is if there is no invariant subspace $\{0\} \subsetneq U \subsetneq V$, then we call V a *primitive* invariant subspace. An important basic theorem shows that an invariant subspace is either primitive or can be decomposed as an *orthogonal* sum of two smaller invariant subspaces. Applying this theorem repeatedly, we can decompose $\mathbb{C}[H]$ as an orthogonal sum of primitive invariant subspaces. This decomposition is unique.

If V is an invariant subspace then H acts on V by convolution. The function ρ associating with each $h \in H$ the linear operator of convolution with h is called a *representation* of H. The function ρ is a homomorphism, i.e. $\rho(gh) = \rho(g)\rho(h)$, since convolution with gh is the same as convolution with g followed by convolution with h. The dimension of ρ is $\dim \rho = \dim V$. The *character* of ρ is $\chi = \operatorname{Tr} \rho$. It is easy to see that the character is a class function, i.e. invariant under conjugation.

If V decomposes as $V^1 \oplus V^2$, then the representation ρ also decomposes to the two corresponding representations ρ^1, ρ^2 (each operating on the relevant subspace independently), and moreover $\chi = \chi^1 + \chi^2$. Such a representation is called *reducible*. If V is primitive, then the corresponding representation is *irreducible*. The trivial representation given by $\rho^e \equiv 1$ is always irreducible.

Two invariant subspaces U, V with representations ρ_U, ρ_V are *equivalent* if under some choice of bases for U, V we have $\rho_U(h) = \rho_V(h)$ for all h. More abstractly, ρ_U, ρ_V are equivalent if there is a regular linear transformation $T\colon U \to V$ such that $\rho_U = T\rho_V T^{-1}$.

We can now describe the orthogonal decomposition of $\mathbb{C}[H]$ into primitive invariant subspaces. For each conjugacy class λ of H there corresponds an irreducible representation ρ^λ. The decomposition contains exactly $\dim \rho^\lambda$ primitive invariant subspaces whose representations are equivalent to ρ^λ. The direct sum of these subspaces forms a d^2-dimensional invariant subspace we denote by V^λ.

We can always associate the trivial representation ρ^e with the conjugacy class consisting of the identity element. The corresponding subspace V^e consists of all constant vectors. An orthonormal basis for V^e is the constant one vector $\vec{1}$.

We define $\hat{f}(\rho^\lambda)$ to be the projection of f into V^λ. In particular, for the trivial representation ρ^e we have

$$\hat{f}(\rho^e) = \left(\frac{1}{|H|} \sum_{h \in H} f(h) \right) \vec{1}.$$

If h is obtained from f by convolution with a uniformly random member of a conjugacy class μ, then

$$\hat{h}(\rho^\lambda) = \frac{\chi^\lambda(\mu)}{\dim \rho^\lambda} \hat{f}(\rho^\lambda).$$

The latter property shows that as long as we restrict ourselves to Cayley graphs whose generators are a union of conjugacy classes, the proof of Hoffman's bound goes through. Moreover, all we need to know about the representation theory of H are the characters and the dimensions of the irreducible representations. Since $\rho^\lambda(e) = I$, the dimension can be determined from the character, $\dim \rho^\lambda = \chi^\lambda(e)$.

In order to determine the unique optimal families, the general method proceeds as follows. Hoffman's bound shows that the Fourier transform of the characteristic function f of an optimal family is concentrated on the small Fourier coefficients. It follows that f is in the span of small juntas. An additional argument shows that f must in fact be a small junta.

In our case, we will be able to show that the Fourier transform of f is concentrated on the small irreducible representations, and so f lies in the direct sum of V^λ for "small λ", under a notion of smallness that we will define. The latter will turn out to be the span of small juntas (functions depending only on few coordinates of the permutation). Again, an additional argument will show that f is in fact a small junta.

7.2 Representation Theory of the Symmetric Group

Our goal in this section is to determine the irreducible representations of the symmetric group S_n, as well as some of their properties. We will follow the analogy from the representation theory of the abelian group $\mathbb{Z}_2^n$. As described in §3.1, the irreducible representations (which correspond to basis vectors) can be grouped into levels B_k. If we define $L_k = \bigcup_{l \le k} B_l$, then we get a chain

$$L_0 \subset L_1 \subset \cdots \subset L_n,$$

where L_k is spanned by the k-juntas; thus L_0 consists only of constant functions, and L_n is the entire space. Reversing the process, by determining the correct analog of L_k in the case of S_n we will be able to find all the irreducible representations.

For concreteness, we deal first with S_4. A function on S_4 can depend on 0, 1, 2 or 3 coordinates (values of the permutation); in the latter case, it actually depends on all 4 coordinates. Letting L_k be the linear span of all functions depending on k coordinates, we get a similar chain

$$L_0 \subset L_1 \subset L_2 \subset L_3.$$

Notice that the L_k are invariant subspaces. We can decompose them into primitive invariant subspaces. One can hope that more inequivalent irreducible representations show up in the decomposition of L_{k+1} than in the decomposition of L_k, and this turns out to be the case. However, that gives us only 4 irreducible representations, whereas S_4 has 5 conjugacy classes:

$$e, (12), (12)(34), (123), (1234).$$

The missing link in the chain is the space $L_{1.5}$ spanned by functions depending on two *unordered* coordinates. More concretely, it is spanned by the characteristic functions of sets of the form

$$\{\pi \in S_4 : \{\pi(i), \pi(j)\} = \{k, l\}\}.$$

It is not difficult to show that $L_1 \subset L_{1.5} \subset L_2$. Indeed, to show that $L_{1.5} \subset L_2$, notice that

$$\{\pi \in S_4 : \{\pi(i), \pi(j)\} = \{k, l\}\} = \{\pi \in S_4 : \pi(i) = k,\ \pi(j) = l\} \cup \{\pi \in S_4 : \pi(i) = l,\ \pi(j) = k\}.$$

For the other direction, consider a typical basis family in L_1, $M_{1\to 1} = \{\pi \in S_4 : \pi(1) = 1\}$. Using the notation $M_{\{i,j\}\to\{k,l\}}$ for basis functions of $L_{1.5}$, let us compute

$$N = \sum_{j\neq 1}\sum_{l\neq 1} M_{\{1,j\}\to\{1,l\}}.$$

If $\pi(1) = 1$ then for each $j \neq 1$ there is a unique $l \neq 1$ such that $M_{\{1,j\}\to\{1,l\}}(\pi) = 1$. Thus $N(\pi) = 3$. If $\pi(1) \neq 1$, then we must have $l = \pi(1)$ and $1 = \pi(j)$. Thus $N(\pi) = 1$. Therefore $(N-1)/2$ is the indicator function for $M_{1\to 1}$.

There is a simple connection between conjugacy classes and the spaces L_i. Each conjugacy class corresponds to a partition $\mu = (p_1, \ldots, p_k)$ of n. The corresponding space L^μ is spanned by functions of the form $M_{A_1\to B_1,\ldots,A_k\to B_k}$, where $|A_i| = |B_i| = p_i$, and the A_i's and the B_i's separately form a partition of $[n]$. Thus:

- The partition (4) corresponds to L_0.
- The partition $(3, 1)$ corresponds to L_1.
- The partition (2^2)[10] corresponds to $L_{1.5}$.
- The partition $(2, 1^2)$ corresponds to L_2.
- The partition (1^4) corresponds to L_3.

Suppose a partition μ is obtained from a partition λ by changing two adjacent levels i, j to $i-1, j+1$; the list above is written in the order of such changes. A generalization of the argument above shows that $L^\lambda \subset L^\mu$. Denote the transitive closure of this operation by $\lambda \triangleright \mu$. Clearly $L^\lambda \subset L^\mu$. It turns out that $\lambda \trianglerighteq \mu$ iff the sum of the t largest parts of λ is at least as much as the corresponding sum of μ, for each t. This partial order is not linear, for example (3^2) and $(4, 1^2)$ are incomparable.

A dimension argument shows that if $\lambda \neq \mu$ then indeed $M^\lambda \neq M^\mu$. It turns out that for each μ, the decomposition of M^λ contains an irreducible representation not found in the decompositions of M^μ for $\mu \triangleright \lambda$. This irreducible representation shows up in exactly one primitive invariant subspace of M^λ known as the Specht module S^λ. These can be shown to be inequivalent, and so they form the full complement of irreducible representations of S_n.

The Specht modules can be constructed explicitly, but we are only interested in the characters χ^λ. These can be calculated from the characters ξ^λ of M^λ as soon as we know how many copies of S^μ are contained in the decomposition of M^λ. The relevant coefficient is the *Kostka number* $K_{\lambda\mu}$, and has an explicit combinatorial formula. The character $\xi^\lambda(\pi)$ itself is simply equal to the number of partitions B_i fixed by π.

[10]The notation a^k means that the part a is repeated k times.

We now collect some other properties which will be useful, some of them classical facts from the representation theory of S_n, and some of them proved by bounding some explicit combinatorial formulas.

Partitions as diagrams. We can think of a partition $\lambda = (p_1, \ldots, p_k)$, where $p_1 \geq \cdots \geq p_k$, as a diagram with k lines, where line l contains p_l points, justified to the left. If we transpose the diagram, we obtain the *conjugate partition* λ'. For example, the transpose of $(4,2)$ is $(2^2, 1^2)$:

$$\begin{array}{cccc|cc} \times & \times & \times & \times & \times & \times \\ \times & \times & & & \times & \times \\ & & & & \times & \\ & & & & \times & \end{array}$$

Characters of conjugate permutations. For any partition λ,

$$\chi^\lambda(\pi) = (-1)^\pi \chi^{\lambda'}(\pi).$$

In particular, since $\chi^{(n)} = \xi^{(n)} = 1$ is the trivial representation, $\chi^{(1^n)}(\pi) = (-1)^\pi$ is the sign representation.

Dimension of S^λ. Let $n - t$ be the maximal size of a row or column in the diagram of λ. We put $h(\lambda) = t$. If t is small then $\dim S^\lambda = \Theta(n^t)$. For each s and big enough n, if $t \geq s$ then $\dim S^\lambda = \Omega(n^s)$. So as t increases we get Specht modules of dimensions $\Theta(1), \Theta(n), \Theta(n^2), \ldots$, and at some point the dimensions might stop growing, but are at least $\Omega(n^s)$.

7.3 Upper Bound

If $\mathcal{F}$ is an intersecting family of permutations and $\alpha, \beta \in \mathcal{F}$ then $\alpha^{-1}\beta$ must have a fixed point; in other words, it cannot be a *derangement* (a permutation without fixed points). As we see below, for our purposes all derangements look the same, and so we just take the constraint graph G whose edges correspond to convolution with a derangement. This is a Cayley graph of S_n in which $\mathcal{F}$ is an independent set.

The eigenvalues of G with respect to partitions λ with $h(\lambda) \leq 1$ can be computed explicitly. The eigenvalue $L(n)$ corresponding to the trivial partition (n) is simply the degree of G, which is the number of derangements d_n. It is well known (and easy to prove using inclusion-exclusion) that

$$d_n \approx \left(1 - \frac{1}{e}\right) n!.$$

For the conjugate partition (1^n) it is easy to compute the eigenvalue directly since $\chi^{(1^n)}$ is the sign of the permutation. For the partition $(n-1,1)$ we use the formula $\chi^{(n-1,1)} = \xi^{(n-1,1)} - \chi^{(n)}$ and the interpretation of $\xi^{(n-1,1)}$ as the number of fixed points to obtain the corresponding eigenvalue. For the conjugate partition $(2, 1^{n-1})$, we use the formula $\chi^{(2,1^{n-1})}(\pi) = (-1)^\pi \chi^{(n-1,1)}(\pi)$ to obtain the eigenvalue. The results are summarized in the following table:

λ	$L(\lambda)$
(n)	d_n
(1^n)	$(-1)^{n-1}(n-1)$
$(n-1,1)$	$-d_n/(n-1)$
$(2,1^{n-1})$	$(-1)^n$

For all other partitions $h(\lambda) \geq 2$, and so $\dim S^\lambda = \Omega(n^2)$. It is easy to see that $(G^2)_{\pi,\pi}$ is the degree of the vertex π, namely d_n, and so

$$\sum_\lambda (\dim S^\lambda)^2 L(\lambda)^2 = \operatorname{Tr} G^2 = n! d_n \approx \left(1 - \frac{1}{e}\right)(n!)^2.$$

The first equality follows from the fact that $L(\lambda)$ is the eigenvalue corresponding to each vector in the $\dim S^\lambda$-dimensional primitive invariant subspace, and S^λ appears $\dim S^\lambda$ times. This equation implies that

$$|L(\lambda)| \lesssim \frac{\sqrt{1 - 1/e} \cdot n!}{\dim S^\lambda} = O((n-2)!).$$

Thus for large n, the largest eigenvalue is $M = d_n$ corresponding to the eigenvector $\vec{1}$, and all other eigenvalues are bounded in magnitude by $m = d_n/(n-1)$. Hoffman's bound (Lemma 3.7) now implies that the relative measure of $\mathcal{F}$ is at most

$$\frac{m}{M+m} = \frac{d_n}{d_n/(n-1)+d_n} = \frac{1}{(n-1)+1} = \frac{1}{n}.$$

In fact, we can apply the cross-independent version of Hoffman's bound (Lemma 3.8) to deduce that if $\mathcal{F}, \mathcal{G}$ are cross-intersecting then

$$\sqrt{|\mathcal{F}||\mathcal{G}|} \leq (n-1)!.$$

This will be important in §7.5.

7.4 Uniqueness

Hoffman's bound implies that if an intersecting family has size exactly $(n-1)!$ then its spectral transform is supported by $M^{(n-1,1)}$, which is spanned by $\vec{1}$ and sets of the form $\{\pi \in S_n : \pi(i) = j\}$ (*double cosets*). A short combinatorial argument by Ellis, Friedgut and Pilpel [18] shows that the family is in fact *equal* to a double coset.

Denote the double coset $\{\pi \in S_n : \pi(i) = j\}$ by $C_{i,j}$. For all i, j we have

$$\sum_k C_{i,k} = \sum_l C_{l,j} = \vec{1}.$$

Thus every member of $M^{(n-1,1)}$ has a representation $\sum_{i,j} \alpha_{i,j} C_{i,j}$, which we put in a matrix A. The characteristic function f of the original family can be recovered from A by noting that the only cells contributing to $f(\sigma)$ are $C_{i,\sigma(i)}$, and so

$$f(\sigma) = \sum_i \alpha_{i,\sigma(i)}. \tag{7}$$

The representation is not unique: if we add x_i to each row and y_j to each column then we add in total $(\sum_i x_i + \sum_j y_j)\vec{1}$; thus if $\sum_i x_i + \sum_j y_j = 0$, we get a different representation for the family. We seek coefficients x_i, y_j that will make all entries positive. In other words, we're trying to find a solution to the linear inequalities

$$\alpha_{i,j} + x_i + y_j \geq 0$$

subject to the constraint

$$\sum_i x_i + \sum_j y_j = 0.$$

We assume that there is no solution, and derive a contradiction. If there is no solution, LP duality implies that there is some non-negative combination of the inequality constraints

$$\sum_{i,j} c_{i,j}(\alpha_{i,j} + x_i + y_j) \geq 0$$

which is contradictory, that is the coefficients of x_i and y_j are all the same, and $\sum_{i,j} c_{i,j}\alpha_{i,j} < 0$. Without loss of generality, the coefficients of x_i and y_j are equal to 1. So for all i, j,

$$\sum_k c_{i,k} = \sum_l c_{l,j} = 1,$$

and so the coefficients $c_{i,j}$ form a bistochastic matrix C. A classical theorem of Birkhoff implies that C is a convex combination of permutation matrices. Equation (7) now implies that $\sum_{i,j} c_{i,j}\alpha_{i,j} \geq 0$, and we reach a contradiction.

We conclude that f has a representation as a non-negative combination of double cosets; since f is Boolean and has the cardinality of a double coset, it must be a double coset itself.

7.5 Stability

Our goal in this section is to prove a Hilton-Milner type result: if an intersecting family is not contained in a double coset, then its size is at most the size of the family

$$\{\pi(1) = 1, \pi(i) = i \text{ for some } i > 2\} \cup (12).$$

The proof, due to Ellis [15], goes through a stability result: if an intersecting family has more than roughly $(1 - 1/e)(n - 1)!$ elements, then it is contained in a double coset. In fact, we only describe the proof of this result.

Let $\mathcal{F}$ be a family of size at least $(1 - \delta)(n - 1)!$. Hoffman's bound implies that the distance of $\mathcal{F}$ from $M^{(n-1,1)}$ is at most $\delta|\mathcal{F}|/n! \approx \delta(1 - \delta)/n$. The projection of the characteristic function of $\mathcal{F}$ into $M^{(n-1,1)}$ is

$$P(\sigma) = \frac{1}{n!} \sum_{\pi \in \mathcal{F}} \big((n - 1)|\{i \in [n] : \pi(i) = \sigma(i)\}| - (n - 2)\big).$$

The fact that the error is small implies that

$$\sum_{\sigma \in \mathcal{F}} (1 - P(\sigma))^2 + \sum_{\sigma \notin \mathcal{F}} P(\sigma)^2 \leq \delta(1 - \delta)(n - 1)!.$$

Thus, for almost all permutations $\sigma \notin \mathcal{F}$, $P(\sigma) \approx 0$, and for a significant amount of permutations $\sigma \in \mathcal{F}$, $P(\sigma) \approx 1$ (the difference between the two cases is that most permutations are not in $\mathcal{F}$ to start with). An isoperimetric bound now implies that we can find two permutations $\alpha \notin \mathcal{F}$ and $\beta \in \mathcal{F}$ with extremal values $P(\alpha) \approx 0$ and $P(\beta) \approx 1$ which are close in the Cayley graph generated by all transpositions.

Looking at the formula for $P(\sigma)$, we see that it measures the agreement of permutations in $\mathcal{F}$ with σ. Thus $\mathcal{F}$ agrees a lot with β and not at all with α (beyond the random expected agreement of one point[11]). Since α and β differ only in a few places, there must be a specific coordinate in β with which $\mathcal{F}$ agrees a lot; the isoperimetric bound gives that $\mathcal{F}$ agrees with $\beta(i)$ on $\omega((n - 2)!)$ places.

We now want to claim that $\mathcal{F}$ in fact agrees with β on coordinate i much more strongly. Indeed, for each j let $\mathcal{F}_j$ consist of those permutations in $\mathcal{F}$ which send i to j. The families $\mathcal{F}_j, \mathcal{F}_k$ are cross-intersecting but can't agree on coordinate i, and so satisfy

$$\sqrt{|\mathcal{F}_j||\mathcal{F}_k|} \leq (n - 2)!.$$

Since $|\mathcal{F}_{\beta(i)}| = \omega((n - 2)!)$, we deduce that for $j \neq \beta(i)$, $|\mathcal{F}_j| = o((n - 2)!)$. Since there are $n - 1$ possible such j's, in fact

$$|\mathcal{F}_{\beta(i)}| = |\mathcal{F}| - o((n - 1)!) = (1 - \delta - o(1))(n - 1)!.$$

Running the same argument again, the bounds escalate to $|\mathcal{F}_j| = O((n - 3)!)$ and so

$$\sum_{j \neq i} |\mathcal{F}_j| = O((n - 2)!).$$

If $\mathcal{F}$ is not a subset of the double coset sending i to $\beta(i)$ then there is some permutation $\pi \in \mathcal{F}$ sending i to some $j \neq \beta(i)$. This rules out a significant number of permutations sending i to $\beta(i)$, and because the rest of them are only $O((n-2)!)$, we get that if δ is small enough (in fact, if $|\mathcal{F}| \gtrsim (1 - 1/e)(n - 1)!$), $\mathcal{F}$ must in fact always send i to $\beta(i)$.

A similar, more complicated combinatorial argument proves the Hilton-Milner result.

[11] If $\mathcal{F}$ is a double coset then the expected agreement with $\alpha \notin \mathcal{F}$ is actually slightly lower, namely $(n - 2)/(n - 1)$.

7.6 Multiply Intersecting Families

The arguments of §7.3 generalize to the case of k-intersecting permutations, but they get significantly more complicated. The reason is that in the 1-intersecting case, the edges of the Cayley graph consisted of all derangements, and from the point of view of the critical small-height partitions, all derangements look almost the same (they are only distinguished according to their sign). In the k-intersecting case, there are more partitions to consider, and so our edges will have to be weighted accordingly.

If α, β are k-intersecting permutations then $\alpha^{-1}\beta$ contains at least k fixed points. Thus a k-intersecting family of permutations $\mathcal{F}$ is an independent set in the Cayley graph whose edges correspond to convolutions with *k-derangements*, which are permutations with *fewer than* k fixed points. We are looking for a weighting of the edges such that $L(n) = 1$ and for any other partition π,

$$L(\pi) \geq \omega \triangleq \frac{-(n-k)!}{n! - (n-k)!}.$$

The quantity ω was chosen so that Hoffman's bound would imply that a k-intersecting family has measure at most

$$\frac{-\omega}{1-\omega} = \frac{(n-k)!}{n!}.$$

The optimal families are *k-double-cosets*, i.e. families of the form

$$\{\sigma \in S_n : \sigma(i_1) = j_1, \ldots, \sigma(i_k) = j_k\}.$$

These form the basis for $M^{(n-k,1^k)}$, and so for any $\pi \trianglerighteq (n-k, 1^k)$ we must have $L(\pi) = \omega$.

We will construct the Cayley graph according to the following steps:

1. For each partition $\pi \triangleright (n-k, 1^k)$, choose a cycle type of a k-derangement.
2. By solving linear equations, find a Cayley graph such that $L(n) = 1$ and $L(\pi) = \omega$ for any other $\pi \triangleright (n-k, 1^k)$.
3. By considering a k-double-coset, deduce that $L(n-k, 1^k) = \omega$.
4. Show that the eigenvalues corresponding to partitions of height at least $k+1$ are $o(\omega)$.

The bound in the last step depends on the height of the partitions, and so we won't be able to distinguish between partitions with a long first row and those with a long first column. The latter are dealt with using the fact that the characters of conjugate partitions are related through multiplication by the sign of the permutation.

Choosing a cycle type for each partition $\pi \triangleright (n-k, 1^k)$. We actually do this step in two ways, producing two different Cayley graphs. The first graph will consist only of *even* permutations, and the second graph will consists only of *odd* permutations. The cycle type corresponding to the partition π will be a *refinement* of π, a property which will enable us to solve the linear equations in the following step.

Each partition $\pi \triangleright (n-k, 1^k)$ contains less than k parts of size 1, and so corresponds naturally to a cycle type of a k-derangement. A cycle type of opposite parity arises from splitting the first part: instead of $(p_1, \ldots)$, take $(p_1 - k - 1, k + 1, \ldots)$. Note that $k + 1$ doesn't appear in the rest of the partition, and so all the split partitions are different from all the original partitions[12]. As an example, when $k = 2$ the partition $(6, 2, 1)$ corresponds to the cycle types $(123456)(78)$ and $(123)(456)(78)$, the first odd, the second even.

Solving linear equations. Given the graph, it is easy to find weights in such a way that $L(n) = 1$ and $L(\pi) = \omega$ for $\pi \triangleright (n-k, 1^k)$ by solving linear equations. Arranging both cycle types and partitions in descending order (with respect to $\triangleright$), these take a triangular form, and so it remains to verify that the diagonal is non-zero. That is a consequence of each cycle type being a refinement of the corresponding partition.

[12] The paper actually splits a part of size k, resulting in the identification of $(n-k, k)$ with the split of n, both of which are used together in one of the graphs.

A simple argument shows that for some some N_0, the weights (solutions to the system of linear equations) are the same for any $n \geq N_0$, since the linear equations themselves are the same.

Finding $L(n-k, 1^k)$. The spectral transform of a k-double-coset is concentrated on $M^{(n-k,1^k)}$, and so the proof of Hoffman's bound shows that the fact that its measure is exactly $(n-k)!$ implies that $L(n-k,1^k) = \omega$.

Bounding $L(\pi)$ for partitions of height at least $k+1$. The eigenvalue corresponding to a partition π under convolution with all conjugates of σ is equal to

$$L_\sigma(\pi) = X_\sigma \frac{\chi^\pi(\sigma)}{\dim S^\pi}, \quad X_\sigma = \# \text{ of conjugates of } \sigma.$$

The linear equations we are solving are of the form

$$\sum_\sigma c_\sigma L_\sigma(n) = 1, \quad \sum_\sigma c_\sigma L_\sigma(\pi) = \omega \text{ for } \pi \triangleright (n-k, 1^k),$$

where the sums are over all cycle types of k-derangements chosen in the first step. Opening up L_σ,

$$\sum_\sigma c_\sigma X_\sigma = 1, \quad \sum_\sigma c_\sigma X_\sigma \chi^\pi(\sigma) = \dim S^\pi \omega \text{ for } \pi \triangleright (n-k, 1^k).$$

The characters χ^π do not depend on n, however X_σ, $\dim S^\pi$ and ω all do. Since $\omega = \Theta(n^{-k})$ and $\dim K^\pi = O(n^k)$ (since π is of height at most k), we get that the right-hand sides are all $O(1)$. On the other hand, we have an easy bound $|X_\sigma| \leq (n-1)!$. We conclude that

$$c_\sigma = O\left(\frac{1}{(n-1)!}\right).$$

Now consider some partition π of height at least $k+1$. Let G_σ be the Cayley graph corresponding to convolution with the conjugacy class of σ. We have (see §7.3)

$$\sum_\lambda (\dim S^\lambda)^2 L_\sigma(\lambda)^2 = \operatorname{Tr} G^2 = n! X_\sigma.$$

Therefore

$$L_\sigma(\pi) \leq \frac{\sqrt{n! X_\sigma}}{\dim S^\pi} = O\left(\frac{n!}{\sqrt{n} \dim S^\pi}\right),$$

where the estimate follows from the bound $|X_\sigma| \leq (n-1)!$ mentioned above. Since $c_\sigma = O(1/(n-1)!)$ and the number of partitions $\pi \triangleright (n-k, 1^k)$ is bounded, we have

$$L(\pi) = \sum_\sigma c_\sigma L_\sigma(\pi) = O\left(\frac{\sqrt{n}}{\dim S^\pi}\right).$$

Finally, since π has height at least $k+1$, $\dim S^\pi = \Omega(n^{k+1})$ (for large enough n), and so

$$L(\pi) = O(n^{1/2-k-1}) = O(n^{-k-1/2}) = o(\omega).$$

Wrapping things up. The Cayley graph constructed in the previous steps has the following two features:

- We have $L(n) = 1$ and $L(\pi) = \omega$ for any other $\pi \trianglerighteq (n-k, 1^k)$.
- For every partition π of height larger than k, $L(\pi) = o(\omega)$.

These leave open the status of partitions with a column of length at least $n-k$. Note, however, that if all k-derangements we chose were even, then $L(\pi) = L(\pi')$, and if all were odd, $L(\pi) = -L(\pi')$; this follows from the formula $\chi^\pi(\sigma) = (-1)^\sigma \chi^{\pi'}(\sigma)$. So the eigenvalues corresponding to these partitions get opposite values on the *even graph* and *odd graph*. Taking the average of these two graphs, those partitions get an eigenvalue of zero, and we're done!

Uniqueness and stability. These follow arguments similar to the ones in §7.4–7.5. The uniqueness argument is in [18], and stability is in [16].

8 Open Problems

In this section we gather some problems left open by §4–7, as well as some open problems which seem amenable to spectral methods.

Graphical intersection problems:

- The arguments of §5.1 do not generalize to the cross-intersecting case since one of the positive eigenvalues is too large. We were not able to fix it, but it is very plausible that the cross-intersecting result is true and can be proven by some modification of the argument.

- The arguments of §5.1–5.3 can be combined with the methods of §6 to prove bounds on the μ_p-measures of triangle-intersecting families of graphs (see [17]). These methods fail for $p > 1/2$, but we conjecture that triangle-juntas are optimal up to $p = 3/4$; when p is larger, taking all graphs with enough edges, we get that the intersection is big enough to be bipartite, and on the other hand as n grows the μ_p measure of the family tends to 1.

- The arguments also generalize to the *non-bipartite-intersecting* case, i.e. instead of requiring the intersection to contain a triangle, we make do with any odd-length cycle. An even more natural problem is *cycle-intersection*, in which the intersection can contain any cycle. The arguments completely fail; the critical p in this case is $1/2$.

Ahlswede-Khachatrian beyond $\mathbf{p < 1/(t+1)}$: The arguments of §6 prove a smoothed version of the Ahlswede-Khachatrian theorem for t-intersecting families whenever $p < 1/(t+1)$. The argument breaks for larger p because the optimal family satisfying all the properties used by Hoffman's bound is no longer Boolean. A way around this problem will constitute major progress in the field.

r-wise intersection: Frankl and Tokushige's research problem generalizes Ahlswede-Khachatrian to the r-wise intersecting case. It can be shown that the spectral transform à la §6 of an r-wise intersecting function f satisfies

$$\sum_S \hat{f}(S)^s(-1)^{(s+1)|S|}\left(\frac{p}{1-p}\right)^{|S|/2} = 0 \text{ for } 1 \le s \le r.$$

It's not clear whether an upper bound can be derived from these extra equations.

Open problems from Chung, Graham, Frankl and Shearer [9]. Chung et al. considered four intersection problems, in which they conjectured that appropriate juntas are optimal:

- Families of sets whose intersection contains a cyclic translate of some fixed set. For progress on this question, see §9.4.

- Triangle-intersecting families, solved in §5.

- Families of graphs whose intersection contains a path of length 3 edges: disproved by Christofides, who presented a path-3-intersecting family of measure more than $1/8$.

- Families of integers whose intersection contains arithmetic progressions of length 3. Wide open.

9 Other Approaches

In this section we shortly describe some other approaches to intersecting questions. We focus on methods which obtain tight result and are elegant (in our view).

9.1 Shifting

Shifting is a general technique in the combinatorics of families of sets (for an example different from intersecting families, see the next section). The idea is to reduce a family into a simple form, while maintaining some of its properties. The reduction is accomplished through *shifting*.

Definition 9.1. Let $\mathcal{F}$ be a family of subsets of $[n]$, and let $i, j \in [n]$. The shifted family $\mathcal{G} = S_{i\to j}(\mathcal{F})$ is defined as follows:

- For every $S \in \mathcal{F}$ such that $i \notin S$, add S to $\mathcal{G}$.
- For every $S \in \mathcal{F}$ such that both $i, j \in S$, add S to $\mathcal{G}$.
- For every $S \in \mathcal{F}$ such that $i \in S$ and $j \notin S$, let $S' = (S \setminus i) \cup j$. If $S' \in \mathcal{G}$ then add S to $\mathcal{G}$. Otherwise, add S' to $\mathcal{G}$.

In other words, $S_{i\to j}(\mathcal{F})$ is obtained by shifting element i into element j whenever possible. The shifted family maintains a lot of properties of the original family, in particular the following.

Lemma 9.1. *If $\mathcal{F}$ is a k-uniform intersecting family then so is $S_{i\to j}(\mathcal{F})$, and both families have the same size. Moreover, $\mathcal{F}$ is a dictatorship if and only if $S_{i\to j}(\mathcal{F})$ is.*

Using shifting, we can easily prove the Erdős-Ko-Rado theorem by induction on n (following the original proof). When $n = 2k$, the bound is $\binom{2k-1}{k-1} = \frac{1}{2}\binom{2k}{k}$, which follows trivially from the fact that an intersecting family cannot contain both a set and its complement.

Now consider some k-uniform intersecting family on $[n]$, where $k < n/2$. We are going to compress it towards the lower end of $[n]$ by repeatedly applying shifts $S_{i\to j}$ for $j < i$. The process must stabilize after finitely many applications, and the resulting family $\mathcal{F}$ is invariant under "left" shifts.

We decompose $\mathcal{F}$ into two parts: $\mathcal{F}_n$ consists of those sets containing n, and $\mathcal{F}_{\overline{n}}$ consists of those sets not containing n. Let $\mathcal{G}$ be obtained from $\mathcal{F}_n$ by erasing the point n. We claim that $\mathcal{G}$ is intersecting: if two sets $A, B \in \mathcal{G}$ are disjoint, then there is some element $i \in [n-1]$ which is not in their union (since $2(k-1) < n-1$); then A and $S_{n\to i}(A)$ are disjoint, even though both belong to $\mathcal{F}_n$.

By induction, $|\mathcal{F}_{\overline{n}}| \le \binom{n-2}{k-1}$, and $|\mathcal{F}_n| = |\mathcal{G}| \le \binom{n-2}{k-2}$. Thus $|\mathcal{F}| \le \binom{n-1}{k-1}$. With some more effort, we can prove (using induction on k also) that the bound is achieved only for dictatorships. The second part of Lemma 9.1 now shows that the original family is also a dictatorship.

For more about shifting, consult Frankl's survey [25]. Frankl [24] proved a stability result in the Erdős-Ko-Rado setting using these methods (as borne out by Keevash and Mubayi [42]). The method is used (with appropriate changes) to analyze intersecting families in other settings.

9.2 Kruskal-Katona

The celebrated Kruskal-Katona theorem [39, 44] belongs to the same circle of idea as shifting. It can be viewed as an expansion property of the poset of sets.

The theorem concerns k-uniform families of sets on $[n]$ and their l-shadows ($l < k$), defined as the set of subsets of size l of members of the family. We're trying to construct "economical" families, constructing a family of given size while minimizing the size of the l-shadow. Intuitively, we should try to construct the family by reusing the same points whenever possible. The Kruskal-Katona theorem gives an ordering which is simultaneously optimal for all stages of the construction!

Here is one corollary of the full theorem.

Theorem 9.2. *If $\mathcal{F}$ is a k-uniform family satisfying $|\mathcal{F}| \ge \binom{x}{k}$, then for every $l \le k$ the size of the l-shadow is at least $\binom{x}{l}$.*

Thus projection of the k-uniform family into its l-shadow cannot decimate the family.

The proof of Erdős-Ko-Rado is now very easy, following ideas of Daykin [10]. Let $\mathcal{F}$ be a k-uniform intersecting family of $[n]$, where $k \le n/2$. Form an $(n-k)$-uniform family $\mathcal{G}$ by complementing all sets. If

$|\mathcal{F}| = |\mathcal{G}| > \binom{n-1}{k-1} = \binom{n-1}{n-k}$, then the k-shadow of $\mathcal{G}$ contains more than $\binom{n-1}{k}$ sets, and so must intersect $|\mathcal{F}|$; however, this means that the original family isn't intersecting after all. In the case of equality, the full Kruskal-Katona theorem directly implies that $\mathcal{F}$ is a dictatorship.

Keevash [41] proved a stability result for intersecting families using very similar ideas.

9.3 Katona's Circle Method

In 1972 Katona [40] came up with an extremely simple and revealing proof of the Erdős-Ko-Rado theorem. His idea was to arrange the n points cyclically, and put a sliding window of size k over the circle. If the window is put in place and the points are randomized, then the probability that the contents of the window belong to the family $\mathcal{F}$ in question equals exactly to its relative size.

Now consider a fixed arrangement of points, and move the sliding window instead (this process yields the same random distribution). Any two positions which are in $\mathcal{F}$ must intersect, and so at most k positions of the window can belong to $\mathcal{F}$ (here use is made of $k \leq n/2$). So the relative measure is at most k/n, and

$$|\mathcal{F}| \leq \frac{k}{n}\binom{n}{k} = \binom{n-1}{k-1}.$$

Moreover, in the case of equality, for any arrangement the window must display a family member for k different rotations, which overlap in a unique point. It's not difficult to see that this implies that the family must be a dictatorship.

Similarly, if the size of the family is close to $\binom{n-1}{k-1}$, then it is often the case that an arrangement supports k different positions of the window sporting members of $\mathcal{F}$. Keevash [41] considered what happens when we slightly modify the arrangement of points, and using an expansion property of S_n with respect to certain generating sets, was able to prove a stability result.

In §6.1 we showed how to go from the discrete case to the smoothed case by adding infinitely many points. Taking this idea literally, we arrive at a very simple proof of the smoothed version of Erdős-Ko-Rado. This time we put the points on the unit circle uniformly at random, and use a continuously sliding window of length p. The rest of the proof remains the same.

Katona's argument was generalized to r-wise intersecting families by Frankl [21]. The smoothed argument can also be generalized to this setting.

The main limitation of Katona's method is that it applies only to 1-intersection problems. Howard, Károlyi and Székely [38] tried to extend the argument to the 2-intersecting setting, but were unsuccessful.

9.4 Linear Subspaces

Our starting point in §3.2 was the fact than an intersecting family of sets cannot contain both a set and its complement, and so can contain at most half the possible sets. This seemingly simplistic idea has much deeper repercussions.

In their seminal paper, Chung, Graham, Frankl, and Shearer [9] considered several intersecting problems. One problem they considered was a family of sets, the intersection of which contains a translate of some fixed set S. For example, if $S = \{1, 2, 3\}$ then the intersection must contain three consecutive elements (the elements are ordered cyclically). They conjectured that such families can contain at most $2^{-|S|}$ of the sets, a bound which is achieved by S-juntas, families centered on a certain copy of S.

The arguments used to attack the conjecture in [9] involved constructing a $|S|$-dimensional vector space V with the property that every non-zero vector of V does *not* contain a copy of S. This implies that an S-intersecting family can contain at most one vector from each coset, and so its relative size is at most $2^{-|S|}$.

This argument was successful for sets S which are intervals. While the original argument was more complicated, Russell [47] managed to construct such a vector space for any interval S and every n (the construction depends on n). Griggs and Walker [35] used the same method to prove a similar result for every S and infinitely many n, and Füredi, Griggs, Holzman and Kleitman [33] completely settled the problem for $|S| = 3$ using the same method.

Interestingly, the method also provides an elementary proof for the bound on non-bipartite-intersecting families, but only on up to 8 vertices [20]. This follows from the mysterious existence of a vector space consisting of the zero vector together with 7 complements of cubes.

9.5 Other Approaches

Intersection escalation. Chung, Graham, Frankl and Shearer [9] proved a non-optimal bound of $2^{\binom{n}{2}-2}$ on the size of non-bipartite-intersecting families by projecting them on complements of complete bipartite graphs (similar to the approach taken in §5). The result is intersecting and so contains at most half of the range. Shearer's entropy lemma now escalates this non-trivial upper bound to a non-trivial upper bound on the size of the original family. Smoothed versions of the argument can be obtained using Friedgut's weighted Shearer's lemma [29].

The polynomial method. The polynomial method is a general method used to great power in *restricted intersection* problems, where the allowable sizes of the intersections are specified. The basic result in the field is the Ray-Chaudhury-Wilson theorem. One proof of this theorem involves constructing certain polynomials, and then using the so-called Combinatorial Nullstellensatz. The same method can be used to prove the Erdős-Ko-Rado theorem, as Füredi, Hwang and Weichsel [34] were able to do, but only with some difficulty.

Other spectral arguments. Frankl and Graham's survey [27] presents a mysterious spectral proof of Erdős-Ko-Rado using an altogether different (and intricate) eigenvalue argument due to Lovász.

Random walks. This is a method introduced by Frankl [22, 23] and further pursued by Tokushige. The idea is that after compressing (in the sense of §9.1) a strongly intersecting family, all of its sets are individually compressed, and so they have a prefix with a relatively high density of elements (in some sense).

Frankl associates with each set a random walk by going over the elements in order, moving up for each element in the set, and right for each element not in the set. The density argument shows that this walk intersects a line, and so the smoothed measure of the family is bounded by the probability that an appropriately skewed random walk hits the line. This probability can be calculated explicitly and provides a non-trivial upper bound on the size of the family.

See Tokushige's survey [50], who recommends reading Frankl's [26].

References

[1] Rudolf Ahlswede and Levon H. Khachatrian. The complete nontrivial-intersection theorem for systems of finite sets. *J. Comb. Theory A*, 76(1):121–138, 1996.

[2] Rudolf Ahlswede and Levon H. Khachatrian. The complete intersection theorem for systems of finite sets. *Eur. J. Comb.*, 18(2):125–136, 1997.

[3] Rudolf Ahlswede and Levon H. Khachatrian. A pushing-pulling method: New proofs of intersection theorems. *Combinatorica*, 19(1):1–15, 1999.

[4] Rudolf Ahlswede and Levon H. Khachatrian. Katona's intersection theorem: four proofs. *Combinatorica*, 25(1):105–110, 2005.

[5] Noga Alon, Irit Dinur, Ehud Friedgut, and Benny Sudakov. Graph products, Fourier analysis and spectral techniques. *Geom. Func. Analysis*, 14(5):913–940, 2004.

[6] Noga Alon, Haim Kaplan, Michael Krivelevich, Dalia Malkhi, and Julien Stern. Scalable secure storage when half the system is faulty. In *27th ICALP*, pages 576–587, 2000.

[7] Mark Braverman. Poly-logarithmic independence fools AC0 circuits. In *IEEE Conference on Computational Complexity*, pages 3–8, 2009.

[8] Peter J. Cameron and C.Y. Ku. Intersecting families of permutations. *Euro. J. Comb.*, 24:881–890, 2003.

[9] Fan R. K. Chung, Ronald L. Graham, Peter Frankl, and James B. Shearer. Some intersection theorems for ordered sets and graphs. *J. Comb. Theory, Ser. A*, 43(1):23–37, 1986.

[10] D. E. Daykin. Erdős-Ko-Rado from Kruskal-Katona. *J. Comb. Theory A*, 17(2):254–255, 1974.

[11] Michel Deza and Péter Frankl. On the maximum number of permutations with given maximal or minimal distance. *J. Comb. Theory A*, 22:352–360, 1977.

[12] Irit Dinur and Ehud Friedgut. Intersecting families are essentially contained in juntas. 2008.

[13] Irit Dinur, Ehud Friedgut, and Oded Regev. Independent sets in graph powers are almost contained in juntas. *Geom. and Func. Analysis*, 18(1):77–97, 2008.

[14] Irit Dinur and Shmuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162(1):439–485, 2005.

[15] David Ellis. A proof of the Cameron-Ku conjecture. *Submitted*.

[16] David Ellis. Stability for t-intersecting families of permutations. *J. Comb. Theory, Ser. A*, 118(1):208–227, 2011.

[17] David Ellis, Yuval Filmus, and Ehud Friedgut. Triangle-intersecting families of graphs. *Submitted*.

[18] David Ellis, Ehud Friedgut, and Haran Pilpel. Intersecting families of permutations. *Submitted*.

[19] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford (2)*, 12:313–320, 1961.

[20] Yuval Filmus. Triangle-intersecting families of graphs on eight vertices. *Manuscript*.

[21] Péter Frankl. On Sperner families satisfying an additional condition. *J. Comb. Theory A*, 20:1–11, 1976.

[22] Péter Frankl. The Erdős-Ko-Rado theorem is true for $n = kct$. In *Proc. Fifth Hungarian Colloq., Keszthey, 1976, Vol. 1*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 365–375. North Holland, 1978.

[23] Péter Frankl. On intersecting families of finite sets. *J. Comb. Theory A*, 24:146–161, 1978.

[24] Péter Frankl. Erdős-Ko-Rado theorem with conditions on the maximal degree. *J. Comp. Theory A*, 46:252–263, 1987.

[25] Péter Frankl. The shifting technique in extremal set theory. In *Surveys in Combinatorics*, volume 123 of *Lond. Math. Soc. Lect. Note Ser.*, pages 81–110. 1987.

[26] Péter Frankl. Multiply-intersecting families. *J. Combin. Theory (B)*, 53:195–234, 1991.

[27] Péter Frankl and Ron L. Graham. Old and new proofs of the Erdős-Ko-Rado theorem. *J. of Sichuan Univ. Natural Science Edition*, 26:112–122, 1989.

[28] Péter Frankl and Norihide Tokushige. On r-cross intersecting families of sets. *Submitted*.

[29] Ehud Friedgut. Hypergraphs, entropy and inequalities. *Amer. Math. Monthly*, 111(9):749–760, 2004.

[30] Ehud Friedgut. On the measure of intersecting families, uniqueness and stability. *Combinatorica*, 28(5):503–528, 2008.

[31] Ehud Friedgut, Gil Kalai, Nathan Keller, and Noam Nisan. Quantitative version of the Gibbard-Satterthwaite theorem for three alternatives. *Submitted*.

[32] Ehud Friedgut, Gil Kalai, and Assaf Naor. Boolean functions whose Fourier transform is concentrated on the first two levels. *Adv. in Appl. Math.*, 29:427–437, 2002.

[33] Z. Füredi, J.R. Griggs, R. Holzman, and D.J. Kleitman. Representations of families of triples over GF(2). *J. Comb. Theory, Ser. A*, 53(2):306–315, 1990.

[34] Zoltán Füredi, Kyung-Won Hwang, and Paul M. Weichsel. A proof and generalizations of the Erdős-Ko-Rado theorem using the method of linearly independent polynomials. *J. Combin. Theory Ser. A*, 53(2):306–315, 1990.

[35] J.R. Griggs and J.W. Walker. Anticlusters and intersecting families of subsets. *J. Combin. Theory Ser. A*, 51(1):99–103, 1989.

[36] A.J.W. Hilton and E.C. Milner. Some intersection theorems for systems of finite sets. *Quart. J. Math. Oxford (2)*, 18:369–384, 1967.

[37] A. J. Hoffman. On eigenvalues and colorings of graphs. In *Graph Theory and its Applications*, Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin, Madison, Wis., pages 79–91. Academic Press, New york, 1969.

[38] Ralph Howard, Gyula Károlyi, and László Székely. Towards a Katona type proof for the 2-intersecting Erdős-Ko-Rado theorem. *Elec. J. Comb.*, 8(1):R31, 2001.

[39] Gyula O. H. Katona. A theorem of finite sets. In *Theory of Graphs: Proceedings*, pages 187–207. Academic Press, 1968.

[40] Gyula O. H. Katona. A simple proof of the Erdős-Chao Ko-Rado theorem. *J. Comb. Theory, Ser. B*, 13:183–184, 1972.

[41] Peter Keevash. Shadows and intersections: stability and new proofs. *Adv. Math.*, 218:1685–1703, 2008.

[42] Peter Keevash and Dhruv Mubayi. Set systems without a simplex or a cluster. *Combinatorica*, 30:175–200, 2010.

[43] Guy Kindler and Shmuel Safra. Noise-resistant Boolean functions are juntas. *Manuscript*.

[44] Joseph Kruskal. The optimal number of simplices in a complex. *Math. Opt. Techniques*, pages 251–268, 1963.

[45] Benoit Larose and Claudia Malvenuto. Stable sets of maximal size in Kneser-type graphs. *Euro. J. Comb.*, 25(5):657–673, 2004.

[46] Makoto Matsumoto and Norihide Tokushige. The exact bound in the Erdős-Ko-Rado theorem for cross-intersecting families. *J. Comb. Theory A*, 52(1):90–97, 1989.

[47] Paul A. Russell. Families intersecting on an interval. *Discrete Mathematics*, 309(9):2952–2956, 2009.

[48] Igor Shinkar. Personal Communication, 2011.

[49] Miklós Simonovits and Vera T. Sós. Graph intersection theorems. In *Proc. Colloq. Combinatorics and Graph Theory*, pages 389–391, Orsay, Paris, 1976.

[50] Norihide Tokushige. The random walk method for intersecting families. In *Horizons of combinatorics*, volume 17 of *Bolyai society mathematical studies*, pages 215–224. 2008.

[51] Norihide Tokushige. A multiply intersecting Erdős-Ko-Rado theorem — the principal case. *Disc. Math.*, 310:453–460, 2010.