SRP

Gabriel's Group

January 3, 2026

1 Oddtown/Eventown

Let's start with Oddtown. Oddtown has n inhabitants. The inhabitants have a habit of forming clubs, and there are a few rules governing this:

1. Each club shall have an odd number of members.
2. Each pair of clubs shall share an even number of members.
3. No two clubs are allowed to have identical membership.

We can restate this setting using set theory. Consider Oddtown T as a family of m sets, each set representing a club. We label everyone from 1 to n, so each club set is a subset of $[n]$. The rules can be rephrased as:

1. $\forall A \in T,\ |A| \equiv 1 \pmod 2$.
2. $\forall A, B \in T,\ A \neq B,\ |A \cap B| \equiv 0 \pmod 2$.

To facilitate the later proofs, we introduce the concept of the incidence vector.

Definition. 1.1 (Incidence Vector)**.** *[2] The incidence vector of a set $A \subseteq [n]$ is $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ where*

$$\alpha_i = \begin{cases} 1 & \textit{if } i \in A \\ 0 & \textit{if } i \notin A \end{cases}.$$

This vector representation allows us to utilize linear algebra techniques to analyze the club memberships.

With sufficient background in place, we can now state and prove the main theorem regarding the maximum number of clubs in Oddtown.

Theorem. 1.2 (Oddtown Theorem)**.** *In an Oddtown with n citizens, no more than n clubs can be formed.*

Proof. The idea is to convert the Oddtown rules to vector form and use linear independence to complete our proof. Suppose we have m clubs, $C_1, C_2, \ldots, C_m$. Define the vector $\vec{v}_i$ of C_i as

$$\vec{v}_i = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{where } x_j = \begin{cases} 1 & \text{if } j \in C_i \\ 0 & \text{otherwise} \end{cases}.$$

With the standard inner product, we have

$$\vec{v}_i \cdot \vec{v}_j = |C_i \cap C_j|.$$

Then, combining this with Oddtown rules 1 and 2, we get (where rule 1 states when $i = j$, the inner product must be odd, and rule 2 states it must be even if $i \neq j$):

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} \text{odd} & \text{if } i = j \\ \text{even} & \text{if } i \neq j \end{cases}.$$

We can simplify further by considering the inner product over $\mathbb{F}_2$, resulting in either 0 or 1.

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}. \tag{1}$$

To demonstrate linear independence, consider the linear combination:

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_m \vec{v}_m = 0.$$

By definition of linear dependence, it is necessary to show all $\lambda_i = 0$. Computing the inner product of the linear combination by v_1 will vanish all terms, but the term with v_1 itself.

$$\lambda_1 \vec{v}_1 \cdot \vec{v}_1 + \lambda_2 \vec{v}_2 \cdot \vec{v}_1 + \cdots + \lambda_m \vec{v}_m \cdot \vec{v}_1 = 0$$

Combine with equation (1),

$$\lambda_1 \vec{v}_1 \cdot \vec{v}_1 = \lambda_1 \cdot 1 = 0.$$

Thus, $\lambda_1 = 0$. Conducting analogous process shows $\lambda_i = 0$ for all $i \leq m$. Therefore, all $\vec{v}_i$ are linearly independent, completing the proof that $m \leq n$. □

2 Point sets in $\mathbb{R}^n$ with only two distances

In this section, we illustrate an important application of the linear algebra techniques in a geometric problem. Let us consider a set of m points $a_1, a_2, a_3, ..., a_m$ such that the distance between a_i and a_j can only take 2 values for any $i \neq j$. This is called a two distance set. We want to find out how large such a set can be. We describe the method used by A. Blokhuis (1981) to establish an upper bound on the size of a two distance set [3].

Definition. 2.1. *Let* $a_1, a_2, a_3, ..., a_m$ *be a set of* m *points in* $\mathbb{R}^n$ *such that the distance between* a_i *and* a_j *can only take 2 values for any* $i \neq j$*. Such a collection is called a two distance set.*

Theorem. 2.2. *Let* S *be a two distance set in* $\mathbb{R}^n$*. Then the maximum cardinality of* S *satisfies the following bounds*

$$\frac{n(n-1)}{2} \leq |S| \leq \frac{(n+1)(n+4)}{2}$$

Proof. We can prove the lower bound by constructing a simple point set S in which the distances are either 2 or $\sqrt{2}$. Consider a set of points $a_1, a_2, a_3, ..., a_m$ such that each a_i is of the form $a_i = \{\alpha_1, \alpha_2, \alpha_3 ... \alpha_n\}$, where exactly two positions α_k, α_l are 1, and the others are 0. So the possible distances between any two points are 2 or $\sqrt{2}$. Since there are $\binom{n}{2}$ ways of choosing the two positions for the 1 in a point a_i, there are exactly $\frac{n(n-1)}{2}$ such points in the set.

For the proof for the upper bound, we will make use of some linear algebra techniques. Let us suppose that the two possible distances are a and b. Then for each point $s \in S$, we can consider the polynomial

$$F_s(x) = \frac{1}{a^2 b^2}(||x-s||^2 - a^2)(||x-s||^2 - b^2)$$

in n variables $(x_1, x_2 ..., x_n)$ over the domain $\mathbb{R}^n$. Here, $||x-s||$ denotes the Euclidean distance between points x, s and takes the form

$$||x-s|| = \sqrt{(x_1 - s_1)^2 + (x_2 - s_2)^2 + ...(x_n - s_n)^2}.$$

Now, according to the two distance condition, for any $x \in S$, we have

$$F_s(x) = \begin{cases} 1 & \text{if } x = s \\ 0 & \text{if } x \neq s. \end{cases}$$

Observe for each distinct $s \in S$, there is a unique polynomial $F_s(x)$, so there are exactly $|S|$ such distinct polynomials $F_s(x)$. Moreover, we can show that these polynomials are linearly independent. The notion of linear independence will apply here because the set of polynomials is a vector space. Assume that some linear combination of the polynomials is 0:

$$\lambda_1 F_{s_1}(x) + \ \lambda_2 F_{s_2}(x) + ... + \lambda_m F_{s_m}(x) = 0.$$

Substituting $x = s_1$, we observe that all the terms except $\lambda_1 F_{s_1}(x)$ disappear, showing that $\lambda_1 = 0$. With similar arguments, we can show that $\lambda_i = 0$ for $1 \leq i \leq m$.

Now in order to show an upper bound on the number of polynomials $F_s(x)$, we need to find the dimension of the vector space they are living in. Since these polynomials are linearly independent and there are $|S|$ of them, $|S|$ cannot be greater than the dimension of the vector space.

To find the dimension of the vector space, we expand $F_s(x)$ and write it as a linear combination of polynomials:

$$\begin{aligned} F_s(x) &= \frac{1}{a^2b^2}(||x-s||^2 - a^2)(||x-s|^2 - b^2) \\ &= \frac{1}{a^2b^2}((x_1-s_1)^2 + ...(x_n-s_n)^2 - a^2)((x_1-s_1)^2 + ...(x_n-s_n)^2 - b^2) \\ &= \frac{1}{a^2b^2}(\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} s_i^2 - 2\sum_{i=1}^{n} x_i s_i - a^2)(\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} s_i^2 - 2\sum_{i=1}^{n} x_i s_i - b^2). \end{aligned}$$

Observe that $F_s(x)$ is simply a linear combination of polynomials of the form

$$(\sum_{i=1}^{n} x_i^2)^2, (\sum_{i=1}^{n} x_i^2)x_j, x_i x_j, x_i, 1.$$

Here, all the s_i's can be treated as constants as they are not inputs to the polynomial. There is only one polynomial in n variables of the form $(\sum_{i=1}^{n} x_i^2)^2$. For $(\sum_{i=1}^{n} x_k^2)x_j$, there are n ways to choose j from $[n]$, so n

such polynomials exist. Similarly there are $\frac{n(n+1)}{2}$ choices for $\{i, j\}$ in $x_i x_j$, because there are $\binom{n}{2}$ ways to choose distinct i, j and n ways to choose identical i and j. Lastly, there are n x_i's and 1 constant polynomial. Adding them together, our total number of polynomials is:

$$1 + n + \frac{n(n+1)}{2} + n + 1 = \frac{(n+1)(n+4)}{2}.$$

Every $F_s(x)$ is a linear combination of the $\frac{(n+1)(n+4)}{2}$ linearly independent polynomials, so the polynomials $F_s(x)$ live in an $\frac{(n+1)(n+4)}{2}$ - dimensional vector space. We get our upper bound $|S| \leq \frac{(n+1)(n+4)}{2}$. □

3 Sets with Few Intersection Sizes

In Section 5.4 of Linear Algebra Methods in Combinatorics, Babai and Frankl discuss sets with few intersection sizes mod p [2]. This section outlines important results and applications in combinatorial mathematics, particularly focusing on the extension of the Ray-Chaudhuri–Wilson Theorem by Frankl and Wilson [7].

Let $\mathcal{F}$ be a collection of m subsets of $[n]$. Let L be a set of s non-negative integers i.e. $L = \{l_1, l_2, \ldots, l_s\}$. We call $\mathcal{F}$ an L-intersecting family if, $|A \cap B| \in L$ for every $A, B \in \mathcal{F}, A \neq B$.

We want to prove that that the number of subsets, m, has the following upper bound:

$$m \leq \sum_{k=0}^{s} \binom{n}{k} = \binom{n}{s} + \binom{n}{s-1} + \ldots + \binom{n}{0}.$$

First, we must first define what we mean by a set system $\mathcal{F}$ that is L-intersecting mod p.

Definition. 3.1. *For $L \subset \mathbb{Z}$ and $r, t \in \mathbb{Z}$, we say that*

$$t \in L \pmod{r},$$

if $t \equiv l \pmod{r}$ for some $l \in L$. So, a set system $\mathcal{F}$ is L-intersecting mod r if $|A \cap B| \in L$ mod r for every $A, B \in \mathcal{F}, A \neq B$.

Definition. 3.2. *A multilinear polynomial is a multivariate polynomial that has degree ≤ 1 in each variable.*
For example, $f(x, y, z) = 3xy + 5y - 7z$ is a multilinear polynomial of degree 2 (because of the monomial $3xy$) whereas $g(x, y) = x^2 + 4y$ is not multilinear.

The following result will be useful for our polynomial space proof below.

Proposition. 3.3. *Let $\mathbb{F}$ be a field and $\Omega = \{0,1\}^n \subseteq \mathbb{F}^n$. If f is a polynomial of degree $\leq s$ in n variables over $\mathbb{F}$ then there exists a unique multilinear polynomial $\tilde{f}$ of degree $\leq s$ in the same variables such that*

$$f(x) = \tilde{f}(x) \ \forall \ x \in \Omega.$$

Proof. Since $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \Omega$, any x_i is either 0 or 1. Also, $0^2 = 0$ and $1^2 = 1$ so we have the identity $x_i^2 = x_i$ which gives us the above property. □

Theorem. 3.4 (Non-uniform modular RW Theorem:)**.** *Let p be a prime number and L a set of s integers, $L = \{l_1, l_2, \ldots, l_s\}$. Assume $\mathcal{F} = \{A_1, \ldots, A_m\}$ is a family of subsets of $[n]$ such that*

(a) $|A_i| \notin L \pmod{p}$ for $1 \leq i \leq m$

(b) $|A_i \cap A_j| \in L \pmod{p}$ for $1 \leq j < i \leq m$

then $m \leq \sum_{k=0}^{s} \binom{n}{k}$.

Proof. Let $x, y \in \mathbb{F}_p^n$ and $F(x, y)$ be a polynomial in $2n$ variables over $\mathbb{F}_p$. We set

$$F(x, y) = \prod_{l \in L}(x \cdot y - l) = (x \cdot y - l_1)(x \cdot y - l_2) \dots (x \cdot y - l_s),$$

where $x \cdot y = \sum_{i=1}^{n} x_i y_i$ is the dot product in $\mathbb{F}_p^n$.

Now consider the n-variable polynomials $f_i(x) := F(x, v_i)$, where $v_i \in \mathbb{F}_p^n$ is the incidence vector 1.1 of the set A_i $(i = 1, \dots, m)$.

Recall that A_i is a subset of $[n]$ so we have that the k^{th} entry of v_i will be 1 if $k \in A_i$ and 0 if $k \notin A_i$ for $1 \leq k \leq n$.

We have that for $1 \leq i, j \leq m$,

$$f_i(v_j) \begin{cases} \neq 0 & \text{if } i = j; \\ = 0 & \text{if } i \neq j. \end{cases}$$

This is true because $v_i \cdot v_i = |A_i|$ and $v_j \cdot v_i = |A_j \cap A_i|$. These cases correspond to conditions (a) and (b) of the theorem above. The subset size does not lie in L mod p so all the terms of the product $f_i(v_i) = \prod_{l \in L}(v_i \cdot v_i - l)$ are non-zero. and hence, $f_i(v_i) \neq 0$.

For the other case, the size of the intersection of the subsets always lies in L mod p so at least one of the terms of the product $f_i(v_j) = \prod_{l \in L}(v_j \cdot v_i - l)$ will be 0 and hence, $f_i(v_j) = 0$.

By the proposition, we know that these equations remain valid if we replace f_i by the corresponding polynomials $\tilde{f}_i$. Note that we do this to reduce the dimension of the polynomial space we are working with.

Further, $\tilde{f}_1, \dots, \tilde{f}_m$ are linearly independent.

To see this, let $\sum_{i=1}^{m} \lambda_i \tilde{f}_i(x) = 0$ be a linear relation between the $\tilde{f}_i s$. Substituting $x = v_j$ makes all but the j^{th} term vanish and we get $\lambda_j \tilde{f}_j(v_j) = 0$.

Since $\tilde{f}_j(v_j) \neq 0$, $\lambda_j = 0$ and this must hold for every j. Thus, we get that if $\sum_{i=1}^{m} \lambda_i \tilde{f}_i(x) = 0$, then all λ_i's must be zero. Therefore, $\tilde{f}_1, \dots, \tilde{f}_m$ are linearly independent over $\mathbb{F}_p$.

Finally, all the $\tilde{f}_i$ are multilinear polynomials of degree $\leq s$ and the belong to a space of dimension $\sum_{k=0}^{s} \binom{n}{k}$. To see this, observe that each polynomial has a degree $\leq s$ and that each monomial that forms the basis of the space is a product of distinct variables. Also, the monomial can have only k variables where $0 \leq k \leq s$. An example with $k = s$ is $x_1 x_2 \dots x_s$ and with $k = 2$ is say, $x_3 x_5$. Notice that we have n variables to choose from and so the number of different monomials with k distinct variables is $\binom{n}{k}$. Thus we form a basis by counting the number of monomials which gives us $\sum_{k=0}^{s} \binom{n}{k}$.

We have shown that the m polynomials $\tilde{f}_i$ are linearly independent and live in a space of dimension $\sum_{k=0}^{s} \binom{n}{k}$ which means $m \leq \sum_{k=0}^{s} \binom{n}{k}$. $\square$

A corollary to the theorem above is the following result.

Theorem. 3.5 (Omitted Intersection Theorem)**.** *Let p be a prime number and $\mathcal{F}$ a $(2p-1)$-uniform family of subsets of a set of $4p-1$ elements. If no two members of $\mathcal{F}$ intersect in precisely $p-1$ elements, then*

$$|\mathcal{F}| \leq 2 \cdot \binom{4p-1}{2p-1}.$$

To prove this we must first prove two results.

Proposition. 3.6. *For $n \geq 2s \geq 2k \geq 0$,*

$$\sum_{k=0}^{s} \binom{n}{k} < \binom{n}{s} \cdot \left(1 + \frac{s}{n-2s+1}\right).$$

Proof. Observe that $\frac{\binom{n}{k-1}}{\binom{n}{k}} = \frac{k}{(n-k+1)}$. Since $k \leq s$, we have $\frac{k}{(n-k+1)} \leq \frac{s}{n-s+1}$. We set $\frac{s}{n-s+1} = \alpha$ and get $\frac{k}{(n-k+1)} \leq \alpha < 1$.

For $n \geq 2s$,

$$\sum_{k=0}^{s} \binom{n}{k} \leq \binom{n}{s} \cdot \left(\sum_{i=0}^{\infty} \alpha^i \right) = \frac{\binom{n}{s}}{1-\alpha}.$$

This is true because $\binom{n}{k-i} \leq \alpha^i \binom{n}{s}$ and adding additional terms to complete the infinite geometric series does not change the inequality.

Substituting for α gives us

$$\frac{\binom{n}{s}}{1-\alpha} = \frac{\binom{n}{s}}{1-\frac{s}{n-s+1}} = \binom{n}{s} \cdot \frac{n-s+1}{n-2s+1} = \binom{n}{s} \cdot \left(1 + \frac{s}{n-2s+1}\right).$$

Thus, $\displaystyle\sum_{k=0}^{s} \binom{n}{k} \leq \binom{n}{s} \cdot \left(1 + \frac{s}{n-2s+1}\right)$. □

Proposition. 3.7. *For $n \geq 2s$ and $s \leq \frac{n}{l}$, where $l \in \mathbb{Z}$, we have*

$$\sum_{k=0}^{s} \binom{n}{k} < \binom{n}{s} \cdot \left(1 + \frac{1}{l-2}\right).$$

Proof. If $s \leq \frac{n}{l}$, then

$$\frac{1}{l-2} \geq \frac{1}{l-2+\frac{l}{n}} = \frac{\frac{n}{l}}{n-2\frac{n}{l}+1} \geq \frac{s}{n-2s+1}.$$

This gives us $\frac{s}{n-2s+1} \leq \frac{1}{l-2}$. Thus, $\displaystyle\sum_{k=0}^{s} \binom{n}{k} < \binom{n}{s} \cdot \left(1 + \frac{1}{l-2}\right)$. □

We use the above results to prove the Omitted Intersection Theorem 3.5.

Proof. The universe set is $[n]$, where $n = 4p-1$ and thus, $\mathcal{F} = \{A_1, \ldots, A_m\}$, $|A_i| = 2p-1$. We set $L = \{0, \ldots, p-2\}$ and verify that conditions (a) and (b) of the nonuniform modular RW theorem hold.

(a) hold because $|A_i| = 2p-1 \notin L \pmod{p}$ since $2p-1 \bmod p = p-1$, which is clearly not in L.

(b) holds because $|A_i \cap A_j| < 2p-1, |A_i \cap A_j| \neq p-1$, so $|A_i \cap A_j| \in$

$\{0, 1, \ldots, p-2, p, p+1, \ldots, 2p-2\}$ and thus, $|A_i \cap A_j| \in L$.
Setting $n = 4p-1, s = p-1, l = 4$, where clearly,

$$\frac{n}{l} = \frac{4p-1}{4} = p - \frac{1}{4} \geq p - 1 = s.$$

Thus, we have $|\mathcal{F}| < \binom{4p-1}{p-1} \cdot \left(1 + \frac{1}{4-2}\right) = \frac{3}{2} \cdot \binom{4p-1}{p-1} < 2 \cdot \binom{4p-1}{p-1}$. □

4 Colouring $\mathbb{R}^n$

What is the minimum number of colours needed to colour $\mathbb{R}^n$ so that no two points at unit distance are the same colour?
The case $n = 2$ is known as the Hadwiger-Nelson problem, and is itself a famous open problem. Aubrey de Grey (2018) [4] proved that at least 5 colours are needed so we now know that the lower bound is 5 and the upper bound is 7.
Consider the distance-d graph that has the infinite set $\mathbb{R}^n$ as its vertex set and an edge between vertices if the distance between two points is some $d > 0$.
The chromatic number of a graph is the minimum number of colours needed to colour the vertices of a graph such that no two adjacent vertices have the same colour. Then, the problem is asking us the chromatic number $c(n)$ of the unit-distance graph. It can also be rephrased to ask - what is the minimum number of subsets that $\mathbb{R}^n$ can be divided into such that no two points within a subset are at unit-distance from each other?

Claim (Frankl-Wilson, 1981):
For large n,

$$c(n) > 1.2^n.$$

We will prove a slightly weaker exponential bound.

Since the unit-distance graph on $\mathbb{R}^n$ and the distance-d graph on $\mathbb{R}^n$ are isomorphic - which means that the graph have the same structure but different labels on the vertices - their chromatic number is the same. We can scale a distance-d graph to make it the unit-distance graph without changing any of its other properties.

Our strategy will be to prove that $c(n) > 1.1397^n$ for a distance-d graph of some subset S of the unit cube $\Omega = \{0,1\}^n$ in $\mathbb{R}^n$. Since this graph is a subgraph of the graph in $\mathbb{R}^n$, proving this will prove our claim.

Each subset S corresponds to a set system $\mathcal{H} \subseteq 2^{[n]}$. Then, $S = S(\mathcal{H})$ is a set of the incidence vectors of the members of $\mathcal{H}$.

For example, if $h = \{1, 2, 5\}$ for some $h \in \mathcal{H}$. Then,

$$S(h) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Assume that $\mathcal{H}$ is k-uniform, that is, for all $h \in \mathcal{H}$, $|h| = k$. So, for all $s \in S$, s has k ones.

Let's find the distance $d(A, B)$ between the incidence vectors of some $A, B \in \mathcal{H}$. Then, wherever A and B intersect, the contribution to the distance is 0. We need to find $A \setminus B$ and $B \setminus A$. So we get:

$$d(A, B)^2 = |A| + |B| - 2|A \cap B|$$

$$d(A, B)^2 = 2(k - |A \cap B|) \quad (1)$$

Since k is fixed, our distance depends solely on $|A \cap B|$. Therefore, avoiding a specific intersection size means avoiding a specific distance.

We want to use the Omitted Intersection Theorem, so assume that $n = 4p-1$, and let $k = 2p - 1$, where p is some prime number. Let the intersection size to be avoided be $p - 1$.

We calculate what distance this corresponds to using (1):

$$d(A, B)^2 = 2(2p - 1 - (p - 1))$$

$$d(A, B)^2 = 2p$$

$$d(A, B) = \sqrt{2p}$$

So, avoiding the intersection size $p - 1$ means avoiding the distance $\sqrt{2p}$.

Let's construct a distance -$\sqrt{2p}$ graph $\mathcal{G}_p$ that has vertex set $\mathcal{H}_p = \binom{[4p-1]}{2p-1}$ and has an edge between two vertices $A, B \in \mathcal{H}_p$ if $|A \cap B| = p - 1$. This means that if we prove an exponential lower bound on the chromatic number of $\mathcal{G}_p$, it proves our claim.

The Omitted Intersection Theorem bounds the size of the largest subset that avoids a particular intersection size, which corresponds to the size of the largest set of vertices without an edge between any two of them. This is the independence number $\alpha(\mathcal{G}_p)$. So,

$$\alpha(\mathcal{G}_p) \leq 2 \cdot \binom{4p-1}{p-1}.$$

If we have a graph $\mathcal{J}$ with n vertices, whose independence number is $\alpha(\mathcal{J})$, and whose chromatic number is $\chi(\mathcal{J})$, then

$$\chi(\mathcal{J}) \geq \frac{n}{\alpha(\mathcal{J})}.$$

The number of vertices in $\mathcal{G}_p$, $|\mathcal{H}_p|$ is $\binom{4p-1}{2p-1}$. Therefore,

$$\chi(\mathcal{G}_p) \geq \frac{|\mathcal{H}_p|}{\alpha(\mathcal{G}_p)} \geq \frac{\binom{4p-1}{2p-1}}{\binom{4p-1}{p-1}}.$$

For all large $n = 4p - 1$,

$$2p - 1 \approx n/2,\ \ p - 1 \approx n/4.$$

Also,

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

$$\chi(\mathcal{G}_p) \geq \frac{\binom{4p-1}{2p-1}}{\binom{4p-1}{p-1}} \approx \frac{\left(\frac{n}{4e}\right)^{\frac{n}{4}}\left(\frac{3n}{4e}\right)^{\frac{3n}{4}}}{\left(\frac{n}{2e}\right)^n}\sqrt{\frac{3}{4}}$$

$$\chi(\mathcal{G}_p) \geq \frac{3^{\frac{3n}{4}}}{2^n} = 1.1397^n$$

Since the unit-distance graph on $\mathbb{R}^n$ and the distance-d graph on $\mathbb{R}^n$ are isomorphic, $\chi(\mathcal{G}_p) = c(n) \geq 1.1397^n$ for all large $n = 4p - 1$. This proves our claim.

5 Borsuk's conjecture disproved

Another important application of the results we saw in the previous section is in disproving Borsuk's conjecture. Borsuk had conjectured that every set of diameter one in $\mathbb{R}^n$ can be partitioned into $D+1$ sets of smaller diameter [2]. However, Jeff Kahn and Gil Kalai disproved the conjecture in 1992, showing that in some cases, we need exponentially many subsets so that the diameter of each subset is smaller than one [2].

In this section, we will apply the results from the previous section to disprove Borsuk's conjecture. We will construct a pointset that needs to be partiioned into many subsets for the diameter to be reduced. However, unlike the previous section, where our set lived in $\mathbb{R}^{4p-1}$, we would have to go to higher dimensions to prove the desired result.

Theorem. 5.1. *Let n be of the form $\binom{4p-1}{2}$ for a prime number p and f(n) denote the minimum number such that every set of diameter D in $\mathbb{R}^n$ can be* *partitioned into $f(n)$ sets of smaller diameter. Then for large n, $f(n) > C^{\sqrt{n}}$ for some $C < 2$.*

Proof. We are going to consider the graph $\mathcal{G}_p$ that we created in the previous section, with vertex set $\binom{[4p-1]}{2p-1}$ and an edge between $A, B \in \binom{[4p-1]}{2p-1}$ if $|A \cap B| = p - 1$. Now, let's define another graph $\mathcal{G}$ with the vertex set defined as follows:

$$V = \{h(A) : A \in \binom{[4p-1]}{2p-1}\}$$

where,

$$h(A) = \{\{x, y\} : x \in A, y \in [4p-1] - A\}.$$

By defining $\mathcal{G}$ in this way, we have set up a one to one corresponding between the vertices in G_p and $\mathcal{G}$. Each $h(A)$ is a family of 2-member subsets of $4p-1$ elements. Since choosing a pair $\{x, y\}$ from $[4p-1]$ can be done in $\binom{4p-1}{2}$ ways, for every $h(A)$, we can define an incidence vector in the space $\mathbb{R}^{\binom{4p-1}{2}}$. In this vector, 1 denotes inclusion of the particular pair in $h(A)$, and 0 denotes exclusion. This incidence vector denotes a point in $\{0,1\}^{\binom{4p-1}{2}}$. As a result, the incidence vectors form a set of points $S \subset \{0,1\}^{\binom{4p-1}{2}}$.
In this graph, we define $h(A)$ to be adjacent to $h(B)$ if distance between incidence vectors of $h(A), h(B)$ is the maximum distance between any two incidence vectors in V. In other words, the two points are adjacent if the distance between them is the diameter of the set. As a result, the minimum number of pieces of S to be made such that the diameter of each set is smaller than the diameter of S, is simply the chromatic number of $\mathcal{G}$.
We now prove an exponential lower bound on the chromatic number by showing that $\mathcal{G}$ and $\mathcal{G}_p$ are isomorphic. To show this, we need $h(A), h(B)$ to be adjacent when A, B are adjacent. By construction, we know that

$|A| = 2p-1$, then $|h(A)| = (2p-1)2p$ because we have $2p-1$ ways of choosing the first element of a set and $2p$ ways of choosing the second element. So, $|h(A)| = 4p^2 - 2p$ and the vertex set V is $(4p^2 - 2p)$ - uniform. Furthermore, we can write the distance between the incidence vectors of $h(A), h(B)$ as:

$$d(h(A), h(B))^2 = 2(4p^2 - 2p - |h(A) \cap h(B)|).$$

This distance is maximum when $h(A) \cap h(B)$ is minimised. Suppose, we have $|A \cap B| = r$, then we will find, from the definition of $h(A)$,

- If $x \in A \cap B$ and $y \in [4p-1] - (A \cup B)$, then $\{x, y\} \in h(A) \cap h(B)$.
- If $x \in A - B$, $y \in B - A$, then $\{x, y\} \in h(A) \cap h(B)$.

Moreover, these are the only elements of $h(A) \cap h(B)$. It follows that

$$\begin{aligned}
|h(A) \cap h(B)| &= |A \cap B|(4p - 1 - |A \cup B|) + |A - B||B - A| \\
&= r(4p - 1 - (4p - 2 - r)) + (2p - 1 - r)^2 \\
&= r(1 + r) + (2p - (1 + r))^2 \\
&= 2r^2 + 3r - 4pr + 4p^2 - 4p + 1.
\end{aligned}$$

Let $g(r) = |h(A) \cap h(B)|$ denote the intersection size as a function of r. On differentiating $g(r)$, we get

$$g'(r) = 4r + 3 - 4p = 0.$$

So, the minimum is obtained when $r = p - (3/4)$, i.e when $|A \cap B| = p - (3/4) \approx p - 1$. This shows that that A, B are adjacent in $\mathcal{G}_p$ if and only if $h(A), h(B)$ are adjacent in $\mathcal{G}$.

Thus the two graphs $\mathcal{G}_p$ and $\mathcal{G}$ are isomorphic. So, using the results from the previous section, we obtain

$$f(n) = \chi(\mathcal{G}) = \chi(\mathcal{G}_p) \geq 1.1397^{4p-1}.$$

Additionally, as $4p-1 > \sqrt{2\binom{4p-1}{2}}$,

$$f(n) \geq 1.1397^{4p-1} > 1.1397^{\sqrt{2n}} = 1.203^{\sqrt{n}}.$$

This completes the proof of the theorem. $\square$

6 Sunflower

6.1 Background

In this section, we are going to talk about a kind of set structure called a **sunflower**. Basically, a sunflower is a family of sets where all intersections of its members are the same.

By the way, Deza and Frankl named it sunflower [5].

Definition. 6.1 (sunflower)**.** *A* ***sunflower*** *is a family of sets* $\mathcal{S} = \{F_1,\ F_2,\ ...,\ F_p\}$, *such that, for* $i,\ i',\ j,\ j' \in [n]$, *where* $i \neq i',\ j \neq j', F_i \cap F_j = F_{i'} \cap F_{j'}$.
We call $C := F_i \cap F_j$ *the* ***core*** *(or* ***kernal****), and* $F_i \backslash C$ *the* ***petals*** *of the sunflower.*
Sometimes a sunflower is also called a Δ***-system****.*

Definition. 6.2 (sad)**.** *A sunflower is* ***sad****, if and only if the core of it is an empty set.*

Example. 6.3. $\mathcal{F} = \{\{1\}, \{2\}, \{3\}\}$ *is a sad sunflower with* 3 *petals.*

Definition. 6.4 (k-sunflower-free)**.** *A family of sets* $\mathcal{F}$ *is called* k***-sunflower-free****, if and only if* $\mathcal{F}$ *does not contain a sunflower with* k *petals.*
In short, if $k = 3$*, we say* $\mathcal{F}$ *is* ***sunflower-free****.*

A problem about sunflowers is how large a family of sets can be such that it does not contain a sunflower. In Naslund and Sawin's paper, they introduced two major conjectures about sunflowers [8].

6.2 Erdős-Rado sunflower conjecture

The first conjecture is from a paper by Erdős and Rado. They proved a lemma asserting that a uniform family of sets must contain a sunflower if it is large enough [6].

Lemma. 6.5 (Erdős-Rado sunflower lemma). *Let $r > 3$, and $\mathcal{F}$ be a r-sunflower-free, w-uniform family of sets, then*

$$|\mathcal{F}| \leq w!(r-1)^w.$$

This can be proved by induction[9].
In the same paper, they made this conjecture [6].

Conjecture. 6.6 (Erdős-Rado). *Let $k \geq 3$, let $\mathcal{F}$ be a k-sunflower-free, w-uniform, family of sets. Then*

$$|\mathcal{F}| \leq C_k^w,$$

where $C_k > 0$ is a constant depending on k only [6].

In 2021, Alweiss, Lovett, Wu, and Zhang improved the upper bound of the size of a k-sunflower-free, w-uniform, family of sets [1]. It is about

$$|\mathcal{F}| \lesssim (\log w)^w.$$

6.3 Erdős-Szemerédi sunflower conjecture

The second conjecture is the Erdős-Szemerédi sunflower conjecture.

Conjecture. 6.7 (Erdős-Szemerédi). *Let $\mathcal{F}$ be a k-sunflower-free family of sets which are subsets of $[n]$. Then*

$$|\mathcal{F}| \leq c_k^n,$$

where $c_k < 2$ is a constant depending on k only.

Naslund and Sawin used the slice rank and polynomial method to prove, which is the case $k = 3$ of the Erdős-Szemerédi sunflower problem [8].
First, let's write down some definitions (for 3-dimensional cases).

Definition. 6.8 (slice)**.** *Let A be a finite set, and $\mathbb{F}$ be a field. A function $f : A \times A \times A \mapsto \mathbb{F}$ is called a slice, if and only if there exist two functions $h : A \mapsto \mathbb{F}$ and $g : A \times A \mapsto \mathbb{F}$, such that, f can be written as the product of them*

$$f(x, y, z) = h(x)g(y, z).$$

Definition. 6.9 (slice rank)**.** *Let $f : A \times A \times A \mapsto \mathbb{F}$, the slice rank of f is the minimum number r such that f can be written as a linear combination of r slices.*

Lemma. 6.10 (Tao)**.** *Let A be a finite set, and $\mathbb{F}$ be a field. Let $T : A \times A \times A \mapsto \mathbb{F}$ be a function. We can assert, if T is non-zero if and only if all inputs are the same, then the slice rank of T equals to the size of A.*

Theorem. 6.11 (Naslund-Sawin)**.** *Let $\mathcal{F}$ be a sunflower-free family of sets which are subsets of $[n]$. Then*

$$|\mathcal{F}| \leq 3(n+1) \sum_{k \leq n/3} \binom{n}{k} \leq 3(n+1)c^n, \quad \textit{where } c = (\frac{3}{2})^{\frac{2}{3}} \leq 1.89.$$

Proof. Notice for all set $A \in \mathcal{F}$, because A is a subset of $[n]$, there is a corresponding incidence vector $x \in \{0, 1\}^n$, such that, $x_i = 1$ if $i \in A$; and $x_i = 0$ otherwise. Let's call the set of incidence vectors S.
S has such a property:

Property. 6.12. *For all x, y, z, $\in S$, there exists some index i such that $\{x_i,\, y_i,\, z_i\} = \{0,\, 1,\, 1\}$.*

This must be true, otherwise we will have a 3-sunflower (notice it might be a sad 3-sunflower).

For every $x \in S$, I call the number of 1s is the weight of x (which is also the size of the corresponding A in $\mathcal{F}$). Now we are able to decompose S by the weights of its elements.

$$S = \bigcup_{l=0}^{n} S_l \text{ where we define } S_l := \{x \in S : \mathit{weight}(x) = l\}.$$

Notice, for all $l \in [n]$, $S_l \subset S$. Thus S being sunflower-free implies that S_l is also sunflower-free.

With this decomposition, we can assert another property:

Property. 6.13. *For all $x, y, z \in S_l$, $x + y + z \in \{0, 1, 3\}^n$ implies $x = y = z$.*

Consider the contraposition, if two of them are different, W.L.O.G., assume $x \neq y$, and $x = z$. Because $\mathit{weight}(x) = \mathit{weight}(y) = \mathit{weight}(z)$, then there must exist at least two indexes $i, j \in [n]$, such that $z_i = x_i \neq y_i$ and $z_j = x_j \neq y_j$. W.L.O.G. assume $x_i = 1$, then $y_i = 0$. This implies $x_j = 0$ and $y_j = 1$. Notice $x = z$, hence $z_i = 1$. Hence $x_i + y_i + z_i = 2 \notin \{0, 1, 3\}$.

Now we consider a polynomial $T : S_l \times S_l \times S_l \mapsto \mathbb{R}$, which can precisely encode properties of sunflower-free sets,

$$T(x, y, z) := \prod_{i=0}^{n} (2 - (x_i + y_i + z_i)).$$

Notice T outputs 0 if and only if for some index i, $x_i + y_i + z_i = 2$.

By Property 6.12 we have

$$T(x, y, z) \neq 0 \iff \forall i, x_i + y_i + z_i \neq 2 \iff x = y = z.$$

By Lemma. 6.8, $\mathit{slice\,rank}(T) = |S_l|$. If we can find an upper bound for $\mathit{slice\,rank}(T)$, it means we also find an upper bound for $|S_l|$.

Notice that we can rewrite T as a linear combination of products of monomials of the following form:

$$x_1^{i_1} \dots x_n^{i_n} y_1^{j_1} \dots y_n^{j_n} z_1^{k_1} \dots z_n^{k_n}.$$

By observation,

$$i_1, i_2, ..., i_n, j_1, j_2, j_n, ..., k_1, k_2, ..., k_n \in \{0,\ 1\}.$$

Also notice, T is polynomial that has degree n. Hence

$$i_1 + i_2 + \cdots + i_n + j_1 + j_2 + \cdots + j_n + k_1 + k_2 + \cdots + k_n \leq n.$$

Notice that at least one of

$$i_1 + i_2 + \cdots + i_n, \quad j_1 + j_2 + \cdots + j_n, \quad k_1 + k_2 + \cdots + k_n$$

is $\leq \frac{n}{3}$. W.L.O.G., let's assume $i_1 + i_2 + \cdots + i_n \leq \frac{n}{3}$. Then notice each of such monomial can already be written as a slice

$$f(x)g(y,\ z),$$

where $f(x) = x_1^{i_1} \dots x_n^{i_n}$ and $g(y, z)$ are other terms that only depend on y and z.

Then we can do the counting because the slice rank is at most the number of slices. We only need to count how many monomials we can have. Notice $\sum_{a=1}^{n} i_a \leq \frac{n}{3}$. Then we can choose the degree from 0 to $n/3$, hence for the x case we have at most

$$\sum_{k=0}^{n/3} \binom{n}{k}.$$

Also, if we consider y and z, we will have

$$|S_l| = slice\, rank(T) \leq 3 \sum_{k=0}^{n/3} \binom{n}{k}.$$

Also notice before, $|S| = \sum_{l=0}^{n} |S_l|$, thus,

$$|S| \le \sum_{l=0}^{n} |S_l| \le 3(n+1) \sum_{k=0}^{n/3} \binom{n}{k}.$$

Observe that for $0 < x < 1$, by the binomial theorem,

$$x^{-\frac{n}{3}}(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^{k-\frac{n}{3}} > \sum_{k=0}^{n/3} \binom{n}{k} x^{k-\frac{n}{3}} > \sum_{k=0}^{n/3} \binom{n}{k}.$$

Notice the maximum of $x^{-\frac{1}{3}}(1+x)$ is $(\frac{3}{2})^{\frac{2}{3}}$, at $x = \frac{1}{2}$. Hence

$$3(n+1)((\frac{3}{2})^{\frac{2}{3}})^n \ge 3(n+1)x^{-\frac{n}{3}}(1+x)^n \ge 3(n+1) \sum_{k=0}^{n/3} \binom{n}{k} \ge |S|.$$

As $|S| = |\mathcal{F}|$, the claim follows. □

References

[1] Ryan Alweiss, Shachar Lovett, Kewen Wu, and Jiapeng Zhang. Improved bounds for the sunflower lemma. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, 2021.

[2] László Babai and Peter Frankl. *Linear Algebra Methods in Combinatorics*. 2022.

[3] A. Blokhuis. *A new upper bound for the cardinality of 2-distance sets in Euclidean space*. Eindhoven University of Technology : Dept of Mathematics : memorandum. Technische Hogeschool Eindhoven, 1981.

[4] Aubrey D.N.J. de Grey. The chromatic number of the plane is at least 5, 2018.

[5] Michel Deza and Peter Frankl. Every large set of equidistant $(0, +1, -1)$-vectors forms a sunflower. *Combinatorica*, 1(3):225–231, 1981.

[6] Paul Erdős and Richard Rado. Intersection theorems for systems of sets. *Journal of the London Mathematical Society*, 35(1):85–90, 1960.

[7] Peter Frankl and Richard M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1:357–368, 1981.

[8] E. Naslund and W. Sawin. Upper bounds for sunflower-free sets. *Forum of Mathematics, Sigma*, 5:e15, 2017.

[9] Jonathan Noel. Sunflower lemma. math 492/529 extremal combinatorics, university of victoria. Video, 2020. YouTube.