The Frankl-Wilson theorem and some consequences in Ramsey theory and combinatorial geometry

Lectures 1-5

We first consider one of the most beautiful applications of the linear independence method. Our starting-point is the classical EKR Theorem:

**Theorem 1** (EKR, 1961). Let \( r < n/2 \), and let \( \mathcal{A} \) be an intersecting family of \( r \)-subsets of \([n]\). Then

\[
|\mathcal{A}| \leq \binom{n-1}{r-1}.
\]

Equality holds if and only if \( \mathcal{A} \) consists of all \( r \)-subsets containing some fixed \( i \in [n] \).

What happens if, instead of demanding that \( \mathcal{A} \subseteq [n]^r \) be intersecting, we demand that \( |x \cap y| \) is odd for any two distinct \( x, y \in \mathcal{A} \)? How large can \( |\mathcal{A}| \) be? It turns out that the answer depends heavily on whether \( r \) is even or odd. If \( r \) is odd, we can take \( \mathcal{A} \) to be the family of all \( r \)-sets containing 1 and \((r-1)/2\) of the pairs \( \{2,3\}, \{4,5\}, \ldots, \{(n-1),n\} \), giving

\[
|\mathcal{A}| = \binom{(n-1)/2}{(r-1)/2}.
\]

(if \( n \) is odd). If \( 0 < \alpha < 1 \) is fixed, and \( r = \lfloor \alpha n \rfloor \), this is \( \geq c^n \) for some \( c = c(\alpha) > 1 \): it grows exponentially with \( n \).

Amazingly, though,

**Theorem 2.** If \( r \in \mathbb{N} \) is even, and \( \mathcal{A} \subseteq [n]^r \) is such that \( |x \cap y| \) is odd for any two distinct \( x, y \in \mathcal{A} \), then

\[
|\mathcal{A}| \leq n + 1.
\]

So we have a linear bound on \( |\mathcal{A}| \), completely independent of \( r \) (for \( r \) even).

In order to prove this, we’ll first prove the following

**Theorem 3.** If \( \mathcal{A} \subseteq \mathcal{P}([n]) \) with \( |x| \) odd for all \( x \in \mathcal{A} \), and \( |x \cap y| \) even for any two distinct \( x, y \in \mathcal{A} \), then

\[
|\mathcal{A}| \leq n.
\]

Proof. We’ll find a linearly independent set of size \( |\mathcal{A}| \) in a vector-space of dimension \( n \). For each \( x \subseteq [n] \), let \( \chi_x \) be the characteristic vector of \( x \), i.e.

\[
\chi_x(i) = \begin{cases} 1 & \text{if } i \in x; \\ 0 & \text{if } i \notin x. \end{cases}
\]
For example, if $n = 5$, then $\chi_{123} = (1, 1, 1, 0, 0)$. We think of $\chi_x$ as a vector in $F_2^n$, where $F_2 = \{0, 1\}$ is the two-element field, i.e. we add vectors modulo 2.

We use the 'standard inner-product' on $F_2^n$:

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i v_i.$$  

Note that this is not a genuine inner-product, as it is not positive definite: $\langle \chi_x, \chi_x \rangle = 0$ whenever $|x|$ is even. (We call $\langle , \rangle$ a degenerate inner-product.) But this doesn’t matter for our purposes.

Note that $\langle \chi_x, \chi_y \rangle = |x \cap y|$ for any $x, y \subset [n]$. Hence, for any two distinct $x, y \in A$, $\langle \chi_x, \chi_y \rangle \equiv 0 \pmod{2}$, whereas $\langle \chi_x, \chi_x \rangle \equiv 1 \pmod{2}$ for any $x \in A$. It follows that the $\chi_x$’s are linearly independent as elements of the $F_2$-vector-space $F_2^n$. Indeed, suppose

$$\sum_{x \in A} c_x \chi_x \equiv 0 \pmod{2}$$

for some $c_x$’s in $F_2$; then for any $y$, taking the inner-product with $\chi_y$ gives:

$$0 \equiv \langle \sum_{x \in A} c_x \chi_x, \chi_y \rangle \equiv \sum_{x \in A \setminus \{y\}} c_x \langle \chi_x, \chi_y \rangle + c_y \langle \chi_y, \chi_y \rangle \equiv \sum_{x \in A \setminus \{y\}} c_x |x \cap y| + c_y |y| \equiv c_y |y| \pmod{2},$$

so $c_y = 0$ for all $y \in A$, as required.

The vector space $F_2^n$ has dimension $n$, and $\{\chi_x : x \in A\}$ is a linearly independent subset of size $|A|$, so $|A| \leq n$.

**Remark.** Equality holds in Theorem 3 if $A = \{\{i\} : i \in [n]\}$ is the family of all singletons. For even $n$, equality also holds if $A = [n]^{(n-1)}$. Using ‘products’ of these constructions, one can show that for $n$ sufficiently large, there are at least $2^{n^2/9}$ non-isomorphic extremal examples (exercise).

**Remark.** No purely combinatorial proof of this theorem is known. In a sense, this is unsurprising, as the algebraic proof is making use of the extra ‘structure’ we have imposed upon the ground-set $[n]$. Another reason why it is unsurprising is that there exist many non-isomorphic extremal examples. (Typically, a purely combinatorial proof can be analyzed to give a simple characterization of the extremal examples.)

Theorem 2 follows immediately from Theorem 3:

**Proof of Theorem 2:** Let $r \in \mathbb{N}$ be even, and let $A \subset [n]^{(r)}$ be such that $|x \cap y|$ is odd for any two distinct $x, y \in A$. Let

$$B = \{x \cup \{n+1\} : x \in A\}.$$
Then $|b| = r + 1$ is odd for all $b \in B$, and $|b \cap b'|$ is even for any two distinct $b, b' \in B$. We have

$$|A| = |B| \leq n + 1,$$

by Theorem 3.

If $r$ is even, and $A \subset [n]^{(r)}$ is such that $|x \cap y|$ is even for any two distinct $x, y \in A$, then we can achieve

$$|A| = \binom{(n - 1)/2}{r/2},$$

so we can again get exponential growth. The following table summarizes the situation:

<table>
<thead>
<tr>
<th>$r$ even</th>
<th>$r$ odd</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>x \cap y</td>
</tr>
<tr>
<td>$</td>
<td>x \cap y</td>
</tr>
<tr>
<td>$</td>
<td>A</td>
</tr>
</tbody>
</table>

It is natural to ask: does this phenomenon generalize to other moduli than 2? The answer is yes, for prime moduli (and prime power moduli, though we will not prove this).

**Theorem 4** (Frankl-Wilson). Let $p$ be prime. Let $A \subset [n]^{(r)}$ such that $|x \cap y| \equiv \lambda_1, \lambda_2, \ldots, \lambda_s \pmod{p}$ for any two distinct $x, y \in A$, where $\lambda_i \not\equiv r \pmod{p}$ for all $i$. Then

$$|A| \leq \binom{n}{s}.$$  

**Proof.** Again, we will construct a linearly independent set of size $|A|$ in a vector space $V$ of dimension at most $\binom{n}{s}$. This time, we work over $\mathbb{R}$, unless otherwise stated. We first build a useful ‘store’ of legal matrices by left multiplication.

Let $V$ be the row-space of $N(s, r)$ over $\mathbb{R}$. (This time, we work over $\mathbb{R}$, unless otherwise stated.) Since $N(s, r)$ has $\binom{n}{s}$ rows, we clearly have $\dim(V) \leq \binom{n}{s}$. (In fact, $\dim(V) = \binom{n}{s}$, though we will not need this.)

We will find $|A|$ linearly independent points in $V$; these will be rows of a matrix $M$ whose row-space is in $V$. In fact, our matrix $M$ will be an integer matrix with rows and columns indexed by $[n]^{(r)}$, with row-space in $V$, and entries depending only on $|x \cap y|$. The congruence conditions in the theorem will imply that the square minor $M|_A$ (i.e. the square minor whose rows and columns are indexed by $A$) will be nonsingular over $\mathbb{Z}_p$ (and therefore over $\mathbb{R}$).

We call a matrix $B$ with columns indexed by $[n]^{(r)}$ ‘legal’ if its row space is contained within $V$. Note that multiplying the matrix $N(s, r)$ on the left by a real matrix whose columns are indexed by $[n]^{(s)}$ always produces a legal matrix, since the rows of the product are then real linear combinations of the rows of $N(s, r)$. We first build a useful ‘store’ of legal matrices by left multiplication.
We claim that for any \( i \leq s \), the matrix \( N(i, r) \) is a legal matrix. Indeed, for any \( x \in [n]^{(i)} \) and any \( y \in [n]^{(r)} \), we have

\[
(N(i, s)N(s, r))_{x,y} = \sum_{z \in [n]^{(s)}} 1_{x \subseteq z} 1_{y \subseteq z} = \begin{cases} \binom{r-i}{s-i}, & \text{if } x \subseteq y; \\ 0, & \text{if } x \not\subseteq y. \end{cases}
\]

Hence,

\[
N(i, s)N(s, r) = \binom{r-i}{s-i}N(i, r).
\]

Since \( \binom{r-i}{s-i} \neq 0 \), we have

\[
N(i, r) = \left( N(i, s)/\binom{r-i}{s-i} \right) N(s, r),
\]

so \( N(i, r) \) is a legal matrix, as claimed.

It follows that the \( \binom{n}{s} \times \binom{n}{s} \) matrix \( M(i) = N(i, s) \top N(i, s) \) is also legal. The \((x, y)\) entry of this matrix is a simple function of \(|x \cap y|\). For any \( x, y \in [n]^{(s)} \),

\[
M(i)_{x,y} = \sum_{z \in [n]^{(i)}} 1_{x \subseteq z} 1_{y \subseteq z} = \binom{|x \cap y|}{i}.
\]

Our matrix \( M \) will be an appropriate real linear combination of the \( M(i) \)'s,

\[
M = \sum_{i=0}^{s} M(i),
\]

where \( a_i \in \mathbb{R} \) for each \( i \leq s \). We have

\[
M_{x,y} = \sum_{i=0}^{s} a_i \binom{|x \cap y|}{i} \quad \forall x, y \in [n]^{(r)}.
\]

Since the polynomials \( \{ \binom{T}{i} : 0 \leq i \leq s \} \) are a basis for the space of real polynomials in \( T \) of degree at most \( s \), for any polynomial \( Q(T) \in \mathbb{R}[T] \) of degree at most \( s \), we can choose \( a_i \)'s such that

\[
M_{x,y} = Q(|x \cap y|) \quad \forall x, y \in [n]^{(r)}.
\]

Choose \( Q(T) = (T - \lambda_1)(T - \lambda_2) \ldots (T - \lambda_s) \), and choose the \( a_i \)'s accordingly. Then \( M \) is an integer matrix; we have

\[
M_{x,y} = (|x \cap y| - \lambda_1)(|x \cap y| - \lambda_2) \ldots (|x \cap y| - \lambda_s) \quad \forall x, y \in [n]^{(r)}.
\]

Let \( M|A \) be the submatrix of \( M \) whose rows and columns are indexed by sets in \( A \). We claim that the rows of \( M|A \) are linearly independent over \( \mathbb{R} \). Indeed, if \( x, y \in A \), we have

\[
M_{x,y} \begin{cases} \not\equiv 0 \text{ (mod. } p) & \text{if } x = y, \text{ since no } \lambda_i \equiv r \text{ (mod. } p); \\ \equiv 0 \text{ (mod. } p) & \text{if } x \neq y, \text{ since } |x \cap y| \equiv \text{ some } \lambda_i \text{ (mod. } p). \end{cases}
\]

4
so \( \det(M|_A) \not\equiv 0 \pmod{p} \), so \( \det(M|_A) \not\equiv 0 \), so its rows are linearly independent over \( \mathbb{R} \). It follows that the rows of \( M \) indexed by \( A \) are linearly independent over \( \mathbb{R} \), so \( M \) has rank at least \( |A| \). But the row space of \( M \) (over \( \mathbb{R} \)) is a subspace of \( V \), which has dimension at most \( \binom{n}{s} \), proving the theorem.

**Remark.** For any fixed \( r \) and \( s \), the bound \( \binom{n}{s} \) is essentially best possible. To see this, let \( s \leq r \), and choose any prime \( p > r \). Take \( \lambda_i = r - s + i - 1 \) (1 \( \leq i \leq s \)), and take

\[
A = \{ x \in [n]^{(r)} : \lfloor r - s \rfloor \subset x \}.
\]

Then

\[
|A| = \binom{n - r + s}{s} = (1 + O(1/n)) \binom{n}{s}
\]

if \( r \) and \( s \) are fixed.

**Remark.** The hypothesis \( \lambda_i \not\equiv r \pmod{p} \) is essential: if \( \lambda_1 \equiv r \pmod{p} \), say, then we can get

\[
|A| = \binom{\lfloor (n - \lambda_1)/p \rfloor}{(r - \lambda_1)/p},
\]

which is exponential in \( n \) if \( r = \lfloor \alpha n \rfloor \), for some fixed \( \alpha \in (0, 1) \).

**Remark.** The Frankl-Wilson Theorem also holds if \( p \) is replaced by a prime power. Amazingly, it is false when \( p \) is replaced by a product of at least two distinct primes, e.g. 6. (Grolmusz, 2000.) This indicates that the phenomenon is ‘genuinely’ a number-theoretic / algebraic one, not just a combinatorial one.

**Corollary 5** (Ray-Chaudhury-Wilson). Let \( A \subset [n]^{(r)} \) such that \( |x \cap y| \in L \) for all distinct \( x, y \in A \), where \( L \subset \{0, 1, 2, \ldots, r - 1\} \) with \( |L| = s \). Then

\[
|A| \leq \binom{n}{s}.
\]

**Proof.** Apply the Frankl-Wilson Theorem with any prime \( p > r \). □

**Explicit Ramsey constructions**

In 1930, Ramsey proved the celebrated

**Theorem 6** (Ramsey’s Theorem). For any \( t \in \mathbb{N} \), there exists \( n \in \mathbb{N} \) such that whenever the edges of the complete graph on \( n \) vertices are coloured red and blue, there must be a monochromatic \( K_t \).

This was the birth of Ramsey Theory, a rich and beautiful area of mathematics dealing with ‘finding order in chaos’. The common phenomenon is that ‘total disorder is impossible’: if any sufficiently large structure (e.g. a complete graph) is partitioned into a bounded number of pieces, there must be a largish ‘substructure’ of the original structure (e.g. a complete subgraph) contained within one of the pieces.

Back to Ramsey’s Theorem. For each \( t \in \mathbb{N} \), we define the Ramsey number \( R(t) \) of \( t \) to be the least integer \( n \) such that whenever the edges of \( K_n \) are 2-coloured, there exists a monochromatic \( K_t \).
Estimating the rate of growth of \( R(t) \) accurately is one of the most infamous open problems in combinatorics. Erdős and Szekeres gave a simple argument showing that
\[
R(t) \leq \left(\frac{2t - 2}{t - 1}\right) \leq 2^{2t - 2}.
\]
(To do this, for any \( s, t \in \mathbb{N} \), define \( R(s, t) \) to be the minimum \( n \in \mathbb{N} \) such that whenever the edges of \( K_n \) are coloured red and blue, there exists a red \( K_s \) or a blue \( K_t \). It is easy to see that \( R(s, t) \leq R(s - 1, t) + R(s, t - 1) \); it follows by induction on \( s + t \) that \( R(s, t) \leq \left(\frac{s + t - 2}{s - 1}\right) \).

An exponential lower bound was proved by Erdős, using one of the first applications of the probabilistic method in combinatorics. Namely, if the edges of \( K_n \) are coloured red and blue independently at random, with probability \( \frac{1}{2} \) of each colour, then the expected number of monochromatic \( K_t \)'s is
\[
\left(\frac{n}{t}\right)^2 2^{1 - \frac{t}{2}} \frac{n^t}{t!} = 2^{1 + \frac{t}{2}} \left(\frac{n}{2^{t/2}}\right)^t.
\]
This is \( < 1 \) if \( n \leq 2^{t/2} \), provided \( t \geq 3 \). Hence, under these conditions, there exists at least one red/blue colouring containing no monochromatic \( K_t \), and therefore
\[
R(t) > 2^{t/2} \forall t \geq 3.
\]
Thinking of the red edges in a colouring of \( K_n \) as the edges of a graph \( G \), and the blue edges as the edges of the complement \( \bar{G} \), we make the following

**Definition.** We say that a graph \( G = (V, E) \) is \( t \)-Ramsey if it contains no clique of order \( t \) and no independent set of order \( t \).

**Definition.** If \( G = (V, E) \) is a graph, a subset of \( V \) is said to be homogeneous if it is a clique or an independent set.

Erdős’ argument above implies that if \( n = o(2^{t/2}) \), if we select a graph \( G \) at uniformly at random from the set of all (labelled) graphs on \( \lfloor n \rfloor \), then \( G \) is \( t \)-Ramsey with high probability. (Throughout this course, ‘with high probability’ will mean ‘with probability tending to 1 as \( n \to \infty \).’) To see this, observe that choosing a graph uniformly at random from the set of all labelled graphs on \( \lfloor n \rfloor \) is equivalent to including every edge of \( K_n \) independently with probability \( 1/2 \); this is of course the usual ‘Erdős-Renyi’ random graph\(^1\) \( G(n, 1/2) \). Let the random variable \( X \) denote the number of homogeneous \( t \)-sets in \( G(n, 1/2) \). If \( n = o(2^{t/2}) \), then
\[
\mathbb{E}X < \frac{2^{1 + t/2}}{t!} \left(\frac{n}{2^{t/2}}\right)^t = o(1).
\]
By Markov’s inequality,
\[
\mathbb{P}\{X \geq 1\} \leq \mathbb{E}X,
\]
and therefore \( \mathbb{P}\{X \geq 1\} = o(1) \). In other words, with high probability, \( G(n, 1/2) \) contains no homogeneous \( t \)-set, i.e. is \( t \)-Ramsey.

We now ask the following

**Question.** Can we explicitly construct large \( t \)-Ramsey graphs?

\(^1\)Recall that if \( n \in \mathbb{N} \) and \( 0 \leq p \leq 1 \), we define the Erdős-Renyi random graph \( G(n, p) \) by independently placing each edge of \( K_n \) in \( G \), with probability \( p \).
Explicit constructions of large $t$-Ramsey graphs are notoriously hard to come by. For a while, no explicit constructions of order superlinear in $t$ were known.

Nagy then gave the following explicit construction of a $t$-Ramsey graph on $(t-1)/3$ vertices. Our vertex-set will be $[t-1]^{(3)}$; we join two sets $x, y \in [t-1]^{(3)}$ by an edge if and only if $|x \cap y| = 1$. By Corollary 5, any clique has order at most $t-1$. An independent set $I$ has $|x \cap y|$ even for all distinct $x, y \in I$; so by Theorem 3, $|I| \leq t-1$. This construction is a $t$-Ramsey graph of order $(t-1)/3$.

The best known completely explicit construction to date gives $R(t) > t^{c \log t / \log \log t}$, where $c > 0$ is an absolute constant. This grows faster than any polynomial in $t$, but is still a long way from exponential growth. It is a generalization of Nagy’s construction by Frankl and Wilson, using their theorem above to show that it has the required properties.

Our vertex-set will be $[m]^{(r)}$, so $N = \binom{m}{r}$, for some $m, r$ to be chosen later. Choose a prime $p$, and choose $r \equiv -1 \pmod{p}$. Join $x$ and $y$ if and only if $|x \cap y| \equiv -1 \pmod{p}$. By the Frankl-Wilson theorem, any independent set has order at most $\binom{m}{p-1}$. If we choose $r = p^2 - 1$, then the edges correspond to $|x \cap y| = p - 1, 2p - 1, \ldots, (p-1)p - 1$, so a clique corresponds to an $L$-intersecting family with $|L| = p - 1$, and therefore has size at most $\binom{m}{p-1}$ by Corollary 5. This construction gives

$$R\left(\binom{m}{p-1} + 1\right) > \binom{m}{p^2 - 1}.$$ 

Choosing $m = p^3$ gives

$$R\left(\binom{p^3}{p-1} + 1\right) > \binom{p^3}{p^2 - 1}.$$ 

It follows that $R(t) > t^{c \log t / \log \log t}$ for all $t \in \mathbb{N}$, where $c > 0$ is an absolute constant, as required.

**Remark.** In 2006, Barak, Rao, Shaltiel and Wigderson gave constructions of $t$-Ramsey graphs of higher order, but these constructions are not totally explicit.

It turns out that graphs arising from algebraic structures often mimic the ‘large-scale’ behaviour of random graphs, though on a ‘small-scale’, they are highly non-random. This theme will occur later in the course, when we give algebraic constructions of bounded-degree ‘expander graphs’, graphs which, like random graphs, have a ‘large’ number of edges in every cut.

**The chromatic number of $\mathbb{R}^N$**

Consider the graph $G$ on $\mathbb{R}^N$ where we join $x$ and $y$ if $d(x,y) = 1$. (Here, $d$ denotes the Euclidean distance.) What is the chromatic number, $\chi(\mathbb{R}^N)$, of this graph?
For $N = 2$, it is known only that

$$4 \leq \chi(\mathbb{R}^2) \leq 7.$$  

(Proving that $\chi(\mathbb{R}^2) > 3$ is a nice exercise. One can prove that $\chi(\mathbb{R}^2) \leq 7$ by tiling $\mathbb{R}^2$ with regular hexagons of diameter slightly smaller than 1, and then colouring the hexagons appropriately.)

The value of $\chi(\mathbb{R}^2)$ may in fact depend upon the axiom of choice! Shelah and Soifer [1] constructed a similar ‘distance’ graph $G'$ on $\mathbb{R}^2$, where $xy \in E(G')$ if and only if $d(x, y)$ lies in a certain set of reals, where the value of $\chi(G')$ depends on whether one uses the axiom of choice. (To be precise, there are models of ZF, without the axiom of choice, where $\chi(G')$ is greater than the value when the axiom of choice is assumed.)

Can we estimate $\chi(\mathbb{R}^N)$? We have the following upper bound:

**Theorem 7.** If $N$ is sufficiently large, then $\chi(\mathbb{R}^N) \leq 20^N$.

**Proof.** First, partition $\mathbb{R}^N$ into small $N$-dimensional cubes $C_a$ of side-length $1/\sqrt{N}$:

$$C_a = \{ v \in \mathbb{R}^N : a_j/\sqrt{N} \leq v_j < a_j/\sqrt{N} \ \forall j \in [N] \} \quad (a \in \mathbb{Z}^N).$$

Note any two points in the same small cube are a distance < 1 apart. Now we colour the cubes: formally, we colour the (infinite) graph $H$ with vertex-set $\mathbb{Z}^N$, where $a$ and $b$ are joined if there is a point in $C_a$ which is distance 1 from some point in $C_b$. We colour this graph greedily, in order of increasing $|a|$. Clearly, the graph $H$ is regular; we claim that it has degree less than $20^N$, for $N$ sufficiently large. To see this, simply observe that for fixed $a \in \mathbb{Z}^N$, all cubes $C_b$ containing a point of distance 1 from some point in $C_a$ must lie within a distance 3 from the bottom-left corner $v_a$ of $C_a$. The number of such cubes is at most

$$\frac{\text{vol}(B_N(0, 3))}{\text{vol}(C_b)}.$$  

Recall that the volume of the ball of radius $R$ in $\mathbb{R}^N$ is

$$\frac{\pi^{N/2} R^N}{\Gamma(N/2 + 1)},$$

where $\Gamma(t)$ denotes the Gamma function, which satisfies $\Gamma(t + 1) = t!$ for all $t \in \mathbb{N}$. Hence,

$$\frac{\text{vol}(B_N(0, 3))}{\text{vol}(C_a)} = \frac{\pi^{N/2} 3^N}{\Gamma(N/2 + 1) N^{N/2}} < 20^N,$$

provided $N$ is sufficiently large. So the degree of $H$ is less than $20^N$, as claimed. Hence, the greedy algorithm uses at most $20^N$ colours to properly colour $H$. This gives a proper colouring of $\mathbb{R}^N$, proving the theorem.

What about lower bounds? If we are allowed to use the axiom of choice, then a compactness argument (first given by de Bruijn and Erdős) shows that $\chi(\mathbb{R}^N)$ is equal to the maximum of the chromatic numbers of its finite subgraphs, so if we wish to obtain a lower bound, there is no loss in restricting our attention to finite subgraphs.
Theorem 8 (The de Bruijn-Erdős Theorem). Let $G = (V, E)$ be an arbitrary graph (where $V$ may be infinite, even uncountable). Then

$$\chi(G) = \max \{\chi(H) : H \text{ is a finite subgraph of } G\}.$$  

Proof. It suffices to show that if every finite subgraph of $G$ is $k$-colourable, then so is $G$. Suppose then that every finite subgraph of $G$ is $k$-colourable. Let

$$X = [k]^V$$

be the set of all functions $V \to [k]$. Endow $[k]$ with the discrete topology, and $[k]^V$ with the product topology. Note that the sets of the form

$$\{f \in [k]^V : f|_U = g\},$$

where $U \subset V$ is finite and $g : U \to [k]$, form a basis of open sets in the product topology on $[k]^V$.

Trivially, $[k]$ is compact. Tychonoff’s theorem states that a product of compact spaces is compact, so $[k]^V$ is compact. Recall that a collection of sets is said to have the finite intersection property if any finite subcollection has nonempty intersection, and that a topological space is compact if and only if every collection of closed sets with the finite intersection property has nonempty intersection.

For each finite subset $U \subset V$, let $A_U$ be the set of functions $f : V \to [k]$ such that $f|_U$ defines a proper colouring of the finite subgraph $G[U]$. Observe that each set $A_U$ is closed (and, in fact, open). Moreover, the collection

$$\{A_U : U \subset V, U \text{ finite}\}$$

has the finite intersection property. Indeed, if $(U_i)_{i=1}^m$ is a collection of finite sets, then $G[U_i \cup \cap_{i=1}^m U_i]$ is a finite subgraph of $G$, and therefore $k$-colourable. A corresponding $k$-colouring is an element of $\cap_{i=1}^m A_{U_i}$.

By the compactness of $[k]^V$, Tychonoff’s theorem, and the finite intersection property, we may conclude that

$$\bigcap_{U \subset V, U \text{ finite}} A_U \neq \emptyset.$$  

An element of this intersection is precisely a proper $k$-colouring of $G$, so $G$ is $k$-colourable, as required. $\square$

Recall that (given the other axioms of ZF set theory), Tychonoff’s theorem is equivalent to the axiom of choice. The de Bruijn-Erdős theorem relies upon the axiom of choice: there are models of ZF (without the axiom of choice) in which the de Bruijn-Erdős theorem does not hold (see Shelah and Soifer, [1],).

Frankl and Wilson proved the following:

Theorem 9. If $N$ is sufficiently large, then there exists a finite subset $S \subset \mathbb{R}^N$ such that $\chi(G[S]) \geq 1.05^N$.

They proved this by explicitly constructing a finite subset $S \subset \mathbb{R}^N$ such that any independent set $S' \subset S$ has $|S'|/|S| \leq 1.05^{-N}$.

They used an easy consequence of the Frankl-Wilson theorem. Namely, if $n = 4p$ for some prime $p$, then a family of half-sized subsets of $[n]$ in which we forbid an intersection of size exactly $n/4$ contains an exponentially small fraction of all the half-sized sets.
Corollary 10. Let $p$ be prime. Suppose $A \subset [4p]^{(2p)}$ such that $|x \cap y| \neq p$ for any two distinct $x, y \in A$. Then

$$|A| \leq 2 \left( \frac{4p}{p-1} \right).$$

Proof. First remove one set from each pair $\{x, x^c\} \subset A$ to produce a family $B \subset A$ with $|B| \geq \frac{1}{2}|A|$ and $x \cap y \neq \emptyset$ for any two $x, y \in B$. Then we have $|x \cap y| \not\in \{0, p, 2p\}$ for any two distinct $x, y \in B$, so $|x \cap y| \not\equiv 0 \pmod{p}$ for any two distinct $x, y \in B$. Since $2p \equiv 0 \pmod{p}$, by the Frankl-Wilson Theorem,

$$|B| \leq \left( \frac{4p}{p-1} \right).$$

Hence,

$$|A| \leq 2 \left( \frac{4p}{p-1} \right),$$

as required.

Remark. We have

$$\frac{2 \left( \frac{4p}{p-1} \right)}{\left( \frac{4p}{2p} \right)} = \Theta \left( \frac{\left( \frac{4p}{p} \right)}{\left( \frac{4p}{2p} \right)} \right) = \Theta(2^{-4p(1-H_2(1/4))}) = \Theta((\frac{16}{27})^p),$$

where $H_2(t) = t \log_2(1/t) + (1-t) \log_2(1/(1-t))$ denotes the binary entropy function.

Observe that if $x$ and $y$ are chosen independently and uniformly at random from $[4p]^{(2p)}$, their expected intersection size is $p$. Although two sets chosen independently and uniformly at random are unlikely to have intersection of size exactly $p$, once we have more than an exponentially small fraction of all $(2p)$-sets, an intersection of size exactly $p$ is guaranteed.

Proof of Theorem 9: We use Corollary 10 to construct a finite subset $S \subset \mathbb{R}^N$ in which any independent set $S' \subset S$ (i.e., a set in which no two points have distance 1 apart) has $|S'| \leq 1.05^{-N}|S|$.

First, observe that if $M \leq N$, we can embed the $M$-dimensional discrete cube in $\mathbb{R}^N$ by identifying $P([M])$ with $\{0, 1\}^M$ in the usual way; we have

$$d(x, y) = \sqrt{|x \Delta y|} \quad \forall x, y \in \mathcal{P}([M]).$$

If we restrict ourselves to a subset of $[M]^{(k)}$ for some $k \leq M$, i.e. to a single layer of $\{0, 1\}^M$, then

$$d(x, y) = \sqrt{|x \Delta y|} = \sqrt{2(k - |x \cap y|)} \quad \forall x, y \in [M]^{(k)},$$

so the distance between two points is determined by the size of the intersection of the corresponding sets. Choose $M = 4p$, where $p$ is prime, and choose $k = 2p$. We know that any family $A \subset [4p]^{(2p)}$ in which $|x \cap y| \neq p$ for any $x, y \in A$ has exponentially small fractional size. Since

$$d(x, y) = \sqrt{|x \Delta y|} = \sqrt{2(2p - |x \cap y|)} \quad \forall x, y \in [4p]^{(2p)},$$

10
banning $|x \cap y| = p$ corresponds to banning distance $\sqrt{2p}$. Scaling our set by a factor of $1/\sqrt{2p}$ therefore produces a set $S \subset \mathbb{R}^{4p}$ in which any set with no unit distance has exponentially small fractional size. Formally, let
\[ S = \left\{ \frac{1}{\sqrt{2p}} \chi_x : x \in [4p]^{(2p)} \right\} \subset \mathbb{R}^N; \]
then by Corollary 10, an independent set $S' \subset S$ has size
\[ |S'| \leq 2 \left( \begin{array}{c} 4p \\ p-1 \end{array} \right). \]
Hence,
\[ \frac{|S'|}{|S|} \leq \frac{2 \left( \begin{array}{c} 4p \\ p-1 \end{array} \right)}{\left( \begin{array}{c} 4p \\ p \end{array} \right)} = \Theta((\frac{16}{27})^p). \]
Now choose $p$ to be the largest prime $\leq N/4$. Recall Bertrand's Postulate (actually a theorem of Chebychev): for any real $t \geq 1$, there exists a prime $p$ between $t$ and $2t$. It follows that $p \geq N/8$. Hence,
\[ \frac{|S'|}{|S|} \leq \Theta((\frac{16}{27})^{N/8}). \]
Since each colour-class in a proper colouring of $G[S]$ is an independent set, it follows that the number of colours required on $S$ is at least
\[ \frac{(\frac{4p}{p})}{2(\frac{4p}{p-1})} = \Theta((\frac{27}{16})^p) \geq \Theta((\frac{27}{16})^{N/8}) \geq 1.05^N, \]
provided $N$ is sufficiently large. \(\square\)

**Remark.** The best known bounds are
\[ (1.239... + o(1))^N \leq \chi(\mathbb{R}^N) \leq (3 + o(1))^N. \]
The lower bound is due to Raigorodsky, and the upper bound to Larman and Rogers.

**Borsuk’s Problem**

If $S$ is a bounded subset of $\mathbb{R}^N$, we define the *diameter*
\[ \text{diam}(S) = \sup_{x,y \in S} d(x,y), \]
where $d$ denotes the Euclidean distance. Suppose we wish to decompose a bounded subset $S \subset \mathbb{R}^N$ into as few pieces as possible, such that each piece has diameter strictly less than the diameter of $S$. We write $b(S)$ for the number of pieces required. How large can $b(S)$ be?

Let
\[ a(N) = \max\{b(S) : S \text{ is a bounded subset of } \mathbb{R}^N\} \]
denote the maximum possible number of pieces required to decompose a bounded subsets of $\mathbb{R}^N$ into pieces with strictly smaller diameter.
It is easy to see that \( a(N) \geq N + 1 \) for all \( N \in \mathbb{N} \): just exhibit an equilateral set of \( N + 1 \) points in \( \mathbb{R}^N \), meaning a set in which all the distances are the same. In \( \mathbb{R}^2 \), this is simply an equilateral triangle. It is an easy exercise to construct an equilateral set of size \( N + 1 \) in \( \mathbb{R}^N \).

Borsuk proved in 1932 that \( a(2) = 3 \): in other words, every bounded subset \( S \subset \mathbb{R}^2 \) can be decomposed into at most 3 pieces with diameter strictly smaller than \( S \). He asked whether \( a(N) = N + 1 \) for all \( N \in \mathbb{N} \). Eggleston proved in 1955 that \( a(3) = 4 \), and for a long time, it was believed that \( a(N) = N + 1 \) for all \( N \in \mathbb{N} \). It turns out that if \( S \) is

- smooth and convex,
- centrally symmetric (\( x \in S \Rightarrow -x \in S \)),

then \( S \) can be broken into at most \( N + 1 \) pieces, each with diameter strictly less than the diameter of \( S \).

Amazingly, Kahn and Kalai proved that in general, \( a(N) \geq c \sqrt{N} \) for some absolute constant \( c > 1 \); in fact, they gave an explicit construction of a finite subset \( S \subset \mathbb{R}^N \) which must be broken into at least \( c \sqrt{N} \) pieces, if we wish each piece to have diameter strictly smaller than the diameter of \( S \).

**Theorem 11** (Kahn, Kalai). There exists an absolute constant \( c > 1 \) such that for all \( N \in \mathbb{N} \), \( a(N) \geq c \sqrt{N} \). In fact, there exists a finite subset \( S \subset \mathbb{R}^N \), such that breaking \( S \) into pieces with diameter strictly smaller than the diameter of \( S \) requires at least \( c \sqrt{N} \) pieces.

**Remark.** Our proof, examined carefully, gives a negative answer to Borsuk’s question for \( N \geq 298 \).

**Proof.** We will construct a finite subset \( S \subset \mathbb{R}^N \), such that any set \( S' \subset S \) with \( \text{diam}(S') < \text{diam}(S) \) has \( |S'|/|S| \leq c^{-\sqrt{N}} \). In fact, we’ll go for \( S \subset \{0,1\}^M \) where \( N/4 \leq M \leq N \): our set \( S \) will be a subset of a discrete cube (possibly of dimension slightly less than \( N \)). We have

\[
d (\chi_A, \chi_B) = \sqrt{|A \Delta B|} \quad \forall A, B \in \mathcal{P}([M]),
\]

so again, the distance between two points is determined by (indeed, is a strictly increasing function of) the symmetric difference of the corresponding sets. So we’ll look for a family \( S \subset \mathcal{P}([M]) \) such that any subfamily \( S' \subset S \) with

\[
\max_{A, B \in S'} |A \cap B| < \min_{A, B \in S} |A \Delta B|
\]

has size \( |S'| \leq c^{-\sqrt{N}} |S| \).

As before, we will just use sets in a single layer of the cube, so that distance between points is determined by the size of the intersection of the corresponding sets. If \( S \subset \{0,1\}^M \), then we have

\[
d (\chi_A, \chi_B) = \sqrt{|A \Delta B|} = \sqrt{2(k - |A \cap B|)} \quad \forall A, B \in S,
\]

so \( |A \Delta B| \) is a strictly decreasing function of \( |A \cap B| \), and our task becomes to construct a family of subsets \( S \subset \{0,1\}^M \) such that any subfamily \( S' \subset S \) with

\[
\min_{A, B \in S'} |A \cap B| > \min_{A, B \in S} |A \cap B|
\]

12
has size $|S'| \leq c^{-\sqrt{N}}|S|$.

We know that if we have a family of $n/2$-sized subsets of $[n]$ (where $n = 4p$), and we ‘ban’ intersections of exactly the ‘average’ size $n/4$, the family must have exponentially small fractional size; this was what enabled us to bound the chromatic number of $\mathbb{R}^N$ from below. But now we wish to ‘ban’ intersections of exactly the minimum size. The idea of Kahn and Kalai was to build $S$ out of another set-system $\mathcal{A}$, such that ‘average’-sized intersections in $\mathcal{A}$ correspond to minimum-sized intersections in $S$. To do this, we take $M = \binom{4p}{2}$, where $p$ is the maximal prime such that $\binom{4p}{2} \leq N$, and we identify $\{1, 2, \ldots, M\}$ with $E(K_{4p})$, the edge-set of the complete graph on $\{1, 2, \ldots, 4p\}$. We take our family to be

$$S = \{E(K_{x,x'}) : x \in [4p]^{(2p)}\},$$

the collection of all edge-sets of $2p \times 2p$ complete bipartite subgraphs of $K_{4p}$.

We then have

$$|E(K_{x,x'}) \cap E(K_{y,y'})| = |x \cap y||x^c \cap y^c| + |x \cap y^c||x^c \cap y|.$$

For any $x, y \in [4p]^{(2p)}$, we have $|x \cap y| = |x^c \cap y^c|$, and therefore

$$|E(K_{x,x'}) \cap E(K_{y,y'})| = |x \cap y|^2 + |x \cap y^c|^2 = |x \cap y|^2 + (m/2 - |x \cap y|)^2.$$

This is minimized precisely when $|x \cap y| = p$, where its value is $2p^2$. So ‘average’-sized intersections in $\mathcal{A}$ correspond exactly to minimum-sized intersections in $S$! We just pay a small price: the ‘real’ ground set now has size $4p = \Theta(\sqrt{N})$, so we get a lower bound of $c^{\sqrt{N}}$, rather than $c^N$ as in the chromatic number problem.

Now for the details. Let $S' \subset S$ with

$$\min_{A,B \in S'} |A \cap B| > \min_{A,B \in S} |A \cap B|.$$

Let

$$\mathcal{A} = \{x \in [4p]^{(2p)} : E(K_{x,x'}) \in S'\}.$$

Then $|\mathcal{A}| = 2|S'|$, and we have $x \cap y \neq p$ for any $x, y \in \mathcal{A}$. So by Corollary 10, we have

$$|\mathcal{A}| \leq 2 \left(\frac{4p}{p-1}\right).$$

Since $|S| = \frac{1}{2}\binom{4p}{2}$, we have

$$\frac{|S'|}{|S|} = \frac{|\mathcal{A}|}{\binom{4p}{2}} \leq \frac{2(4p/p-1)}{\binom{4p}{2}} = \Theta((\frac{16}{27})p).$$

Recall that $p$ was chosen to be the maximal prime such that $M = \binom{4p}{2} \leq N$. By Bertrand’s postulate, $p > \sqrt{N}/8$, and therefore

$$|S'|/|S| \leq \Theta((\frac{16}{27})N/8) \leq c^{-\sqrt{N}},$$

where $c > 1$ is an absolute constant.

It follows that we need at least $c^{\sqrt{N}}$ pieces to break $S$ into pieces each of diameter strictly smaller than that of $S$, proving the theorem.
What about upper bounds on $a(N)$? It is easy to obtain an upper bound of the form
\[ a(N) \leq C^N, \]
for some absolute constant $C > 1$:

**Lemma 12.** There exists an absolute constant $C$ such that $a(N) \leq C^N$ for any $N \in \mathbb{N}$.

**Proof.** We must show that any subset $S \subset \mathbb{R}^N$ can be broken into at most $C^N$ pieces, each with diameter strictly less than the diameter of $S$. Without loss of generality, we may assume that $\text{diam}(S) = 1$, and that $0 \in S$. Hence, $S \subset B(0,1)$. Our aim is simply to cover $B(0,1)$ with at most $C^N$ balls of radius 1/4.

To do this, let $Y \subset B(0,1)$ be a maximal subset of $B(0,1)$ such that $d(y,y') > 1/4$ for any two distinct $y,y' \in Y$. Note that the set of all closed balls of radius 1/4 and centre in $Y$ covers $B(0,1)$. Indeed, if there exists $v \in B(0,1) \setminus \bigcup_{y \in Y} B(y,1/4)$, then $d(v,y) > 1/4 \forall y \in Y$, so we could add $v$ to $Y$ while still maintaining the property $d(y,y') > 1/4$ for all distinct $y,y' \in Y$, contradicting the maximality of $Y$. So we have covered $B(0,1)$ by $|Y|$ balls of radius 1/4. A simple volume-packing argument gives us an upper bound on $|Y|$. Observe that the closed balls of radius 1/8 and centre in $Y$ are all pairwise disjoint, and lie in the ball $B(0,1 + 1/8)$. It follows that
\[ |Y| \leq \frac{\text{vol}(B(0,9/8))}{\text{vol}(B(0,1/8))} = 9^N. \]

The sets
\[ \{S \cap B(y,1/4)\} \quad (y \in Y) \]
cover $S$, so there exists a partition of $S$ into at most $9^N$ parts in which each part has diameter at most 1/2, proving the lemma.

The best upper bound is $a(N) \leq (\sqrt[3]{\frac{3}{2}} + o(1))^N$, due to Schramm. It is conjectured that the lower bound can be improved:

**Conjecture 1.** There exists an absolute constant $c > 1$ such that $a(N) \geq c^N$ for all $N \in \mathbb{N}$.

**Grolmusz’ construction.**

Babai and Frankl conjectured that the Frankl-Wilson theorem holds for all composite moduli, not just primes and prime powers. A special case of this is as follows:

**Conjecture 2** (Babai, Frankl). Let $m \geq 2$. If $\mathcal{A} \subset [n]^{(r)}$ is such that $|x \cap y| \neq r \pmod. m$ for any two distinct $x,y \in \mathcal{A}$, then
\[ |\mathcal{A}| \leq \binom{n}{m-1}. \]

Surprisingly, this turned out to be false whenever $m$ is a product of at least 2 distinct primes, demonstrating that the Frankl-Wilson theorem is an intrinsically number-theoretic / algebraic phenomenon, not just a combinatorial one.
Theorem 13 (Grolmusz, 1999). If $m$ is a product of $k \geq 2$ distinct primes, then for infinitely many $n$, there exists $r \equiv 0 \pmod{m}$ and a family $A \subset [n]^{(r)}$ with $|x \cap y| \neq 0 \pmod{m}$ for any two distinct $x, y \in A$, and size

$$|A| \geq \exp(c_m (\log n)^k / (\log \log n)^{k-1}),$$

where $c_m > 0$ is a constant depending upon $m$ alone.

Remark. This grows faster than any polynomial in $n$.

Our strategy is to take a family $A$ of size $l^N$, with sets corresponding to functions from $[N]$ to $[l]$, i.e. $[l]^{[N]}$.

For each subset $T \subset [N]$, we fix a non-negative integer $a_T$ (we’ll choose these later). Now for each $T \subset [N]$, and for each function $g \in [l]^T$, we take $l^{|T|}$ disjoint blocks $B_T(g)$ of size $a_T$, one for each function $g \in [l]^T$. (We call these blocks the $T$-blocks. Note that all the $T$ blocks are disjoint from all the $T'$-blocks, for $T \neq T'$. If we happened to choose $a_T = 0$, we don't take any $T$-blocks.) The union of all the blocks forms our ground-set $X$; it has size

$$n = |X| = \sum_{T \subset [N]} a_T l^{|T|}.$$

We now define our set-system $A$. For each function $f : [N] \to [l]$ let $f|_T \in [l]^T$ denote the restriction of $f$ to $T$. We associate with $f$ the set

$$x_f = \bigcup_{T \subset [N]} B_T(f|_T) \subset X,$$

and we let

$$A = \{x_f : f \in [l]^{[N]}\}.$$

Note that $|A| = l^N$, and each set in $A$ has size

$$\sum_{T \subset [N]} a_T.$$

If we choose $a_T = 0$ for large $|T|$, then $n$ will be much smaller than $|A|$. For any $f, g \in [l]^{[N]}$, $f$ and $g$ contain the same $T$-block if and only if they have the same restriction to $T$, i.e. $f|_T = g|_T$. So

$$|x_f \cap x_g| = \sum_{T : f|_T = g|_T} a_T = \sum_{T \subset S : f(i) = g(i)} a_T.$$

Therefore, the size of the intersection of $x_f$ and $x_g$ is determined purely by the set of coordinates on which $f$ and $g$ agree.

Our task is to choose the $a_T$'s such that $|x_f \cap x_g| \equiv 0 \pmod{m}$ if and only if $f = g$, so we need

$$\sum_{T \subset S} a_T \equiv 0 \pmod{m} \iff S = [N].$$

Since this condition is a modulo-$m$ congruence condition, we may as well choose $a_T \in \{0, 1, \ldots, m-1\}$ for each $T$. In addition, to ensure that $n$ is small compared
to \( A \), we need to ensure that \( a_T = 0 \) for all \( |T| > d \), for some \( d \) not too large. We will then have
\[
n \leq \sum_{j=0}^{d} \binom{N}{j} l^j (m-1) \leq (m-1)N^d l^d.
\]

For ease of calculation, we’ll choose \( l = N \). We then have
\[
n \leq \sum_{j=0}^{d} \binom{N}{j} N^j (m-1) \leq (m-1)N^{2d} \leq O(N^{2d}),
\]

compared with \( |A| = N^N \). Provided we can satisfy the congruence condition with \( d = o(N) \), we will have \( |A| \) growing faster than any polynomial in \( n \). We will in fact satisfy the congruence condition with \( d = \lfloor (mN)^{1/k} \rfloor \).

What we want is a function from \( \mathcal{P}([N]) \) to \( \mathbb{Z}_{\geq 0} \) of the form
\[
S \mapsto \sum_{T \subset S} a_T \quad (S \subset [N]),
\]

which is zero modulo \( m \) if and only if \( S = [N] \), and has \( a_T = 0 \) for all \( |T| > d \). Identifying subsets of \( [N] \) with their characteristic vectors in \( \{0, 1\}^N \), we want to choose \( a_T \)'s such that the function
\[
(z_1, z_2, \ldots, z_N) \mapsto \sum_{T \subset S} a_T \prod_{i \in T} z_i
\]
is zero modulo \( m \) only at \((1, 1, \ldots, 1)\).

In other words, we must construct a multilinear polynomial \( Q(X_1, X_2, \ldots, X_N) \in \mathbb{Z}[X_1, \ldots, X_m] \) with total degree at most \( d = \lfloor (mN)^{1/k} \rfloor \), such that for \( z \in \{0, 1\}^N \),
\[
Q(z) \equiv 0 \pmod{m} \iff z = (1, 1, \ldots, 1).
\]

This leads us to the following general definition:

**Definition.** If \( Q \in \mathbb{Z}[X_1, \ldots, X_N] \) is a multivariate polynomial, and \( F : \{0, 1\}^N \to \{0, 1\} \) is a Boolean function, we say that \( Q \) represents \( F \) modulo \( m \) if the modulo \( m \) zeros of \( Q \) in \( \{0, 1\}^N \) are precisely the zeros of \( F \), i.e.
\[
\forall z \in \{0, 1\}^N, \quad Q(z) \equiv 0 \pmod{m} \iff F(z) = 0.
\]

So our task is to construct a modulo-\( m \) representation of \( \text{NAND} \), of total degree at most \( d = o(N) \). Sanity check: we had better make sure that we cannot do this \( m \) is prime, otherwise the Frankl-Wilson theorem would be false. If \( m = p \) is prime, and \( Q \) is a polynomial of total degree at most \( d \) representing \( \text{NAND} \) modulo \( p \), then we have
\[
\forall z \in \{0, 1\}^N, \quad Q(z) \equiv 0 \pmod{p} \iff z = (1, 1, \ldots, 1).
\]

By Fermat’s Little Theorem, \( a^{p-1} \equiv 1 \pmod{p} \) for all \( a \neq 0 \pmod{p} \), so the polynomial
\[
R = 1 - Q^{m-1}
\]
satisfies
\[ R(z) \equiv \begin{cases} 1 \pmod{p} & \text{if } z = (1, 1, \ldots, 1); \\ 0 \pmod{p} & \text{otherwise.} \end{cases} \]
It is easy to see that such a polynomial \( R \in \mathbb{Z}_p[X_1, \ldots, X_N] \) must have the polynomial \( X_1X_2\cdots X_N \) as a factor, so must have total degree at least \( N \). Hence, \( Q \) must have total degree at least \( N/(p-1) \), i.e. we cannot have \( d = o(N) \).

Now for our construction. Of course, building a low-degree modulo \( m \) representation of \( \text{NAND} \) is equivalent to building a low-degree modulo \( m \) representation of \( \text{OR} \): if \( R(X_1, \ldots, X_N) \) does for \( \text{OR} \), then \( R(1 - X_1, \ldots, 1 - X_N) \) does for \( \text{NAND} \), and has the same total degree. So we will build a low-degree modulo-\( m \) representation of \( \text{OR} \).

Observe that the polynomial
\[
1 - (1 - z_1)(1 - z_2)\cdots(1 - z_N) = \sum_{T \subseteq [N] : |T| \neq 0} (-1)^{|T|} \prod_{i \in T} z_i
\]
is equal to \( OR \), but its total degree, \( N \), is too large. For \( p \) a prime and \( e \in \mathbb{N} \), we form the polynomial
\[
G_p^e(z) = \sum_{T \subseteq [N] : 0 < |T| < p^e} (-1)^{|T|+1} \prod_{i \in T} z_i
\]
by ‘truncating’ \( G \), removing the monomials with total degree \( \geq p^e \). Observe that \( G_p^e \) is really a function of \( s = |\{i \in [N] : z_i = 1\}| \) alone:
\[
G_p^e(s) = \sum_{0 < t < p^e} (-1)^{t+1} \binom{s}{t}.
\]
It has a very useful modulo-\( p \) property, even stronger than being a modulo-\( p \) representation of \( OR \): \( G_p^e(s) \equiv 0 \pmod{p} \) if \( s \equiv 0 \pmod{p^e} \), and \( G_p^e(s) \equiv 1 \pmod{p} \) otherwise. An appropriate linear combination of the \( G_p^e \)'s will be our low-degree representation of \( OR \) modulo \( m \).

**Lemma 14.** Let \( p \) be prime, and let \( e \in \mathbb{N} \). Then the function
\[
G_p^e(s) = \sum_{0 < t < p^e} (-1)^{t+1} \binom{s}{t}
\]
satisfies
\[
G_p^e(s) \equiv \begin{cases} 0 \pmod{p} & \text{if } s \equiv 0 \pmod{p^e}; \\ 1 \pmod{p} & \text{otherwise.} \end{cases}
\]

**Proof.** Clearly, \( G_p^e(0) = 0 \), and if \( 0 < s < p^e \), then \( G_p^e(s) = 1 - (1 - 1)^s = 1 \), so the lemma holds for all \( 0 \leq s < p^e \). Assume now that \( s \geq p^e \). Observe that we may write
\[
\binom{s}{t} = \sum_{j=0}^{t} \binom{p^e}{j} \binom{s - p^e}{t - j}.
\]
We now make the following

**Claim.** If \( 0 < j < p^e \), then
\[
\binom{p^e}{j} \equiv 0 \pmod{p}.
\]
Proof of Claim: We have
\[
\binom{p^e}{j} = \frac{p^e(p^e - 1) \ldots (p^e - j + 1)}{j(j - 1) \ldots (1)}.
\]
Observe that if \( p^a \mid b \) for some \( b < p^e \), then \( p^a \mid p^e - b \). Pairing up the numerator-term \( p^e - b \) with the denominator-term \( b \) for each \( b \), we see that the highest power of \( p \) dividing the numerator is more than the highest power of \( p \) dividing the denominator (\( p^e \) is paired up with \( j < p^e \)). The claim follows.

It follows that 
\[
G_{p^e}(s) \equiv G_{p^e}(s - p^e) \pmod{p^e}
\]
whenever \( s \geq p^e \), so the statement of the lemma holds for all \( s \).

We can now construct our low-degree representation of \( OR \) by taking an appropriate linear combination of the \( G_{p^e} \)'s, using the Chinese Remainder Theorem:

**Lemma 15.** Let \( m = p_1p_2 \ldots p_k \) be a product of \( k \) distinct primes. Then there exists a polynomial \( P \) representing \( OR \) modulo \( m \), with degree at most \( d = \left\lfloor \frac{(mN)^{1/k}}{k} \right\rfloor \).

**Proof.** By the Chinese Remainder Theorem, we can choose \( c_1, \ldots, c_k \) such that 
\[
c_i \equiv 1 \pmod{p_i}
\]
for all \( i \in [k] \), and 
\[
c_i \equiv 0 \pmod{p_j}
\]
for all distinct \( i, j \in [k] \).
Define
\[
R(z) = \sum_{i=1}^{k} c_i G_{p_i^{e_i}}(z),
\]
where the \( e_i \)'s are to be chosen later. Of course, like the \( G_{p^e} \)'s, \( R \) is really a function of \( s = |\{i \in [n] : z_i = 1\}| \) alone:
\[
R(s) = \sum_{i=1}^{k} c_i G_{p_i^{e_i}}(s).
\]
We have \( R(s) \equiv 0 \pmod{m} \) if and only if \( R(s) \equiv 0 \pmod{p_i} \) for each \( i \); this occurs if and only if \( G_{p_i^{e_i}}(s) \equiv 0 \pmod{p_i} \) for each \( i \), which occurs if and only if 
\[
s \equiv 0 \pmod{p_i^{e_i}} \quad \forall i \in [k].
\]
This occurs if and only if \( \prod_{i=1}^{k} p_i^{e_i} \) divides \( s \). But we always have \( s \leq N \), so all we need to do is to choose the \( e_i \)'s such that
\[
\prod_{i=1}^{k} p_i^{e_i} > N;
\]
we will have \( R(s) = 0 \) if and only if \( s = 0 \). For each prime \( p_i \), simply choose \( e_i \in \mathbb{N} \) maximal such that
\[
p_i^{e_i} \leq (mN)^{1/k}.
\]
We then have
\[
p_i^{e_i} > (mN)^{1/k} / p_i,
\]
so
\[
\prod_{i=1}^{k} p_i^{e_i} > N.
\]
Clearly, $R$ is a modulo-$m$ representation of $OR$ with total degree at most $\lfloor(mN)^{1/k}\rfloor$, completing the proof. The ‘magic’ of modular arithmetic has enabled us to eliminate modulo-$m$ zeros using much lower total degree.

For our low-degree representation of $NAND$ modulo $m$, we can now take

$$Q(X_1, \ldots, X_N) = R(1 - X_1, \ldots, 1 - X_N),$$

completing our construction.

Note that we have

$$n \leq (m - 1)N^{2d} \leq (m - 1)N^{2(mN)^{1/k}},$$

whereas

$$|A| = N^N.$$

Hence, if $m$ is fixed,

$$|A| \geq \exp\left( (1 - o(1)) \frac{k}{(2k)^{k/m} (\log n)^{k-1}} \right),$$

which grows faster than any polynomial in $n$, provided $k \geq 2$.

Grolmusz’ construction above can be used to construct another explicit $t$-Ramsey graph of order

$$\exp\left( c \frac{(\log t)^2}{\log \log t} \right),$$

where $c > 0$ is an absolute constant, though the constant $c$ is somewhat smaller than that in the construction of Frankl and Wilson given earlier.

Indeed, let $m = 6$, and consider the graph with vertex-set $\mathcal{A}$, where we join two sets $x, y \in \mathcal{A}$ if and only if $|x \cap y|$ is even. By construction, we have $\mathcal{A} \subset [n]^{(r)}$, where $r \equiv 0 \pmod{6}$, and for any two distinct $x, y \in \mathcal{A}$, we have $|x \cap y| \equiv 0 \pmod{3}$ for any two distinct $x, y \in \mathcal{B}$, whereas $r \equiv 0 \pmod{3}$, so by the Frankl-Wilson Theorem,

$$|C| \leq \binom{n}{2}.$$

If $\mathcal{D} \subset \mathcal{A}$ is an independent set, we have $|x \cap y|$ odd for any two distinct $x, y \in \mathcal{D}$, but $r$ is even, so again by the Frankl-Wilson Theorem, we have

$$|\mathcal{D}| \leq n.$$

Hence, our graph is $\left(\binom{n}{2} + 1\right)$-Ramsey, showing that

$$R\left(\binom{n}{2} + 1\right) \geq \exp\left( c_n \frac{(\log n)^2}{\log \log n} \right),$$

i.e.

$$R(t) \geq \exp\left( c \frac{(\log t)^2}{\log \log t} \right),$$

where $c > 0$ is an absolute constant.
References