Start with a story

Once upon a time there was a town...
Where residents loved forming clubs
To limit # clubs, town council established the following rules:

EVEN TOWN

- Every club must have even # members
- No two clubs can have exactly same set of members
- Every two clubs must share even # members

Did not work too well...

Residents soon found out they could form exponentially many clubs $2^{\Omega(n)}$ for n residents — exercise.

Second attempt

This is a result in EXTREMAL COMBINATORICS, really, but extremal combinatorics and TCS do not seldom intersect

ODD/EVEN TOWN

- Every club must have an odd # members
- Every two clubs must share even # members

(What about clubs being distinct, by the way?)

How many clubs can be formed now?

THEOREM 1 Let $\mathcal{F}$ be a family of subsets of $[n] = \{1, 2, \ldots, n\}$ such that $|A|$ is odd $\forall A \in \mathcal{F}$ and $|A \cap B|$ is even $\forall A, B \in \mathcal{F}$, $A \neq B$.
Then $|\mathcal{F}| \leq n$

Prove this by basic linear algebra
Let us start by recalling some definitions

Let $\mathbb{F}$ field. A VECTOR SPACE over $\mathbb{F}$ (or $\mathbb{F}$-VECTOR SPACE) consists of abelian group V under addition plus scalar multiplication $\mathbb{F} \cdot V \mapsto V$ such that $\forall \alpha, \beta \in \mathbb{F}$ $\forall x, y \in V$

1) $\alpha(\beta x) = (\alpha\beta) x$
2) $(\alpha + \beta) x = \alpha x + \beta x$
3) $\alpha(x+y) = \alpha x + \alpha y$
4) $1 \cdot x = x$

Elements of V VECTORS
Elements of $\mathbb{F}$ SCALARS

Example $\mathbb{R}^n = \{ (x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R} \}$

Write vectors as $\vec{x} = \bar{x} = x = (x_1, x_2, \ldots, x_n)$ typically. (But sometimes subscripts convenient for sets of vectors also – will try to make clear from context)

The SPAN of a set of vectors $\vec{x}_1, \ldots, \vec{x}_m$ is

$$\text{span}(\vec{x}_1, \ldots, \vec{x}_m) = \left\{ \sum_{i=1}^{m} \alpha_i \vec{x}_i \;\middle|\; \alpha_i \in \mathbb{F} \right\}$$

Set of vectors S span V if $\text{span}(S) = V$

$\vec{y} = \sum_{i=1}^{m} \alpha_i \vec{x}_i$ is a LINEAR COMBINATION of the vectors $\vec{x}_i$, $i \in [m]$

The vectors in $S = \{\vec{x}_1, \dots, \vec{x}_m\}$ are LINEARLY INDEPENDENT if

$$\sum \alpha_i \vec{x}_i = \vec{0} \quad \text{implies } \forall i \; \alpha_i = 0$$

(From now on, usually write just 0 for zero vector $(0, 0, \dots, 0)$ which is the additive identity in V)

A set of vectors $\{\vec{x}_1, \dots, \vec{x}_m\}$ form a BASIS for V over $\mathbb{F}$ if they span V and are linearly independent.
The DIMENSION of a vectorspace is the cardinality of any basis (always the same)

Note that linear independence depends on field of scalars $\mathbb{F}$

Over $\mathbb{R}$ $(1, 2)$ and $(2, 1)$ are lin. indep
Over $\mathbb{F}_3$ they are dependent

Recall that for vectors in $\mathbb{R}^n$ we can also take the INNER PRODUCT

$$\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{n} x_i \cdot y_i \qquad \text{sometimes written } \vec{x} \cdot \vec{y}$$

Formally, an inner product is any way to multiply vectors in V to get a scalar in $\mathbb{F}$ such that

1) $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$
2) $\langle \alpha \vec{x}, \vec{y} \rangle = \alpha \langle \vec{x}, \vec{y} \rangle$
3) $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$
4) $\langle \vec{x}, \vec{x} \rangle \geq 0$ with equality iff $\vec{x} = 0$

We can take inner products for vector spaces over finite fields also, except positive definite condition (4) is problematic

$$\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y} = \sum_i x_i y_i$$

with all computations done in $\mathbb{F}$

Ex In $\mathbb{F}_3^3$ we have $\vec{x} = (1,1,1) \neq (0,0,0)$ but $\langle \vec{x}, \vec{x} \rangle = 0$.

In math, seems to be called symmetric bilinear forms
In CS, usually say inner products anyway.
(but keeping in mind that (4) fails to hold)

For vectors in $\mathbb{R}^n$, can define length by using Euclidean norm / L^2-norm

$$\|x\| = \|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Works less well for finite fields ...

Anyway... Back to proving Thm 1

We are going to prove it by studying linearly independent vectors in $\mathbb{F}_2^n$ for $\mathbb{F}_2 = \{0,1\}$

Dimension of vector space: n

Any $n+1$ vectors have to be linearly dependent
(can write $\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \dots + \alpha_{n+1} \vec{x}_{n+1} = 0$ for $\alpha_i \overset{\text{not all}}{=} 0$)

If set of vectors linearly independent $\Rightarrow$ at most n of them

Proof of Thm 1

Consider vector space $\mathbb{F}_2^n$

Represent $A \in \mathcal{F}$ by indicator vector / characteristic vector 1_A (ith coordinate $= 1$ iff $i \in A$; $= 0$ if $i \notin A$)

We claim $\{ 1_A \mid A \in \mathcal{F} \}$ linearly independent, from which theorem immediately follows.

Suppose $\sum_{A \in \mathcal{F}} \alpha_A 1_A = 0$.

Need to prove $\alpha_A = 0 \quad \forall A \in \mathcal{F}$

Fix some arbitrary $B \in \mathcal{F}$ and take inner product with 1_B

$$0 = \left\langle \sum_{A \in \mathcal{F}} \alpha_A 1_A , 1_B \right\rangle$$

$$= \sum_{A \in \mathcal{F}} \alpha_A \langle 1_A , 1_B \rangle \qquad \text{(linearity)}$$

$$= \alpha_B \langle 1_B , 1_B \rangle \qquad |A \cap B| \text{ even for } A \neq B; \text{ cancels mod } 2$$

$$= \alpha_B \qquad |B| \text{ odd}$$

Hence $\{ 1_A \mid A \in \mathcal{F} \}$ linearly independent as claimed, and # vectors $\leq n$.

EVEN/ODD TOWN

- Every club has even # members
- Every two clubs share odd # members

Simple reduction:

even/odd town n residents, m clubs $\Rightarrow$ odd/even $n+1$ residents, m clubs

So no more than $n+1$ clubs. But can do better!

THM 2 Let $\mathcal{F}$ family of subsets of $[n]$ s.t. $|A|$ even $\forall A \in \mathcal{F}$ and $|A \cap B|$ odd $\forall A, B \in \mathcal{F}$, $A \neq B$. Then $|\mathcal{F}| \leq n$.

Proof Translate again to vectors $1_A \in \mathbb{F}_2^n$

Assume towards contradiction $|\mathcal{F}| = n+1$.

Three observations:

- $\langle 1_A, 1_B \rangle = 1$ for $A \neq B$ since $|A \cap B|$ odd
- $\langle 1_A, 1_A \rangle = 0$ since $|A|$ even
- $\{1_A \mid A \in \mathcal{F}\}$ must be linearly dependent, so can find ~~coeffs~~ scalars α_A not all zero st.

$$\sum_{A \in \mathcal{F}} \alpha_A 1_A = 0$$

Look at any $B \in \mathcal{F}$ and take inner product with 1_B to get

$$0 = \left\langle \sum_{A \in \mathcal{F}} \alpha_A 1_A, 1_B \right\rangle =$$

$$= \sum_{A \in \mathcal{F}} \alpha_A \langle 1_A, 1_B \rangle$$

$$= \sum_{A \in \mathcal{F}, A \neq B} \alpha_A$$

so

$$\sum_{A \in \mathcal{F}} \alpha_A = \alpha_B \qquad (*)$$

But B was arbitrary, so for $B' \neq B$ get same equality $(*)$ for $\alpha_{B'}$ and so all α_A are equal. That is, all $\alpha_A = 1$, since not all $\alpha_A = 0$ by assumption

Hence, for any even/odd town with $|\mathcal{F}| = n+1$ clubs it holds that

$$\sum_{A \in \mathcal{F}} 1_A = 0 \qquad (\dagger)$$

and that

$$n = |\mathcal{F}| - 1 = \sum_{\substack{A \in \mathcal{F} \\ A \neq B}} 1 = \left\langle \sum \alpha_A 1_A , 1_B \right\rangle = 0$$

is even (and $|\mathcal{F}|$ is odd).

Now replace each A by complement $\overline{A} = [n] \setminus A$.

Then $|\overline{A}|$ is even and

$$|\overline{A} \cap \overline{B}| = |\overline{A \cup B}| = \cancel{n - |A \cap B|}\ n - |A| - |B| + |A \cap B|$$

is odd. So the $1_{\overline{A}}$ are indicator vectors of an even/odd town. Hence by $(\dagger)$ we have

$$\sum_{A \in \mathcal{F}} 1_{\overline{A}} = 0 \qquad (\ddagger)$$

But now combining $(\dagger)$ and $(\ddagger)$ we get

$$0 = \sum_{A \in \mathcal{F}} 1_A + \sum_{A \in \mathcal{F}} 1_{\overline{A}}$$

$$= \sum_{A \in \mathcal{F}} \left(1_A + 1_{\overline{A}} \right)$$

all-ones vector

$$= \sum_{A \in \mathcal{F}} \underline{1}$$

$$= |\mathcal{F}| \cdot \underline{1}$$

But this implies that $|\mathcal{F}|$ even. Contradiction. $\boxed{\lightning}$

Yet another tweak to the rules...
Every two clubs should share a fixed number of members k
(No conditions on size, but clubs should be different)

THM 3 FISHER'S INEQUALITY
Let $\mathcal{F}$ be a family of nonempty subsets of $[n]$ such that for some fixed k it holds that $|A \cap B| = k$ for all $A, B \in \mathcal{F}$, $A \neq B$.
Then $|\mathcal{F}| \leq n$

Proof Look at indicator vectors 1_A again, but now in real vector space $\mathbb{R}^n$.
We again claim that $\{1_A \mid A \in \mathcal{F}\}$ must be a linearly independent set of vectors.
Suppose $\sum_{A \in \mathcal{F}} \alpha_A 1_A = 0$ (where all calculations are now in $\mathbb{R}$). Need to show $\alpha_A = 0$ $\forall A \in \mathcal{F}$.
Since the norm of the 0-vector is 0 we have

$$0 = \left\| \sum_{A \in \mathcal{F}} \alpha_A 1_A \right\|^2$$

$$= \left\langle \sum_{A \in \mathcal{F}} \alpha_A 1_A , \sum_{B \in \mathcal{F}} \alpha_B 1_B \right\rangle$$

$$= \sum_{A \in \mathcal{F}} \alpha_A^2 |A| + \sum_{\substack{A, B \in \mathcal{F} \\ A \neq B}} \alpha_A \alpha_B k$$

$$= \sum_{A \in \mathcal{F}} \alpha_A^2 (|A| - k) + k \left(\sum_A \alpha_A \right)^2 \quad (**)$$

We want to argue that the only way $(**) = 0$ could hold is that $\alpha_A = 0 \quad \forall A \in \mathcal{F}$.

Note that $|A| \geq k$ for all $A \in \mathcal{F}$ and there can be at most one A^* with $|A^*| = k$ (Why?)

We ~~musta~~ have $\alpha_A = 0 \quad \forall A$ s.t. $|A| > k$ ~~$\forall A \neq A^*$~~, otherwise

$$\sum{}' \alpha_A^2 (|A| - k) > 0$$

~~But~~ If $\exists A^*$ s.t. $|A^*| = k > 0$, then $\sum{}' \alpha_A = \alpha_{A^*}$ and we must have $\alpha_{A^*} = 0$ or else

$$k\left(\sum{}'_A \alpha_A\right)^2 > 0$$

It follows that $\{1_A \mid A \in \mathcal{F}\}$ are indeed linearly independent and hence $|\mathcal{F}| \leq n$ $\square$

Next time, we are going to see how to use these results to construct Ramsey graphs, i.e., graphs without large cliques or independent sets