Additive combinatorics without addition: Ramsey Cayley graphs, information theory, and independence in random graph models

Talk by Jacob Fox

Notes by Sanjana Das

October 11, 2024

This is based on joint work with David Conlon, Huy Tuan Pham, and Liana Yepremyan.

§1 Introduction

The general theme of this talk is that sometimes you have problems you're interested in and you have lots of information, but much of that information turns out to be superfluous.

Here are some basic questions in additive combinatorics:

> **Question 1.1.** Suppose we have a set A living in some ambient group G, and we know $|A+A| = O(|A|)$. What can we say about A?

> **Question 1.2.** If G has size N, how *many* sets $A \subseteq G$ with $|A| = n$ and $|A+A| \leq k|A|$ are there?

You can ask lots of interesting questions along these lines, and understanding these questions has lots of applications. And generally, with these problems, you want to use the group structure to say interesting things about your set A. For example, if $|A+A| = |A|$, then A has to be a subgroup of G. So we can ask, if $A+A$ is not much bigger than this, in what sense is A an *approximate* subgroup? (This corresponds to Question 1.1.)

But the theme of this talk is that for some questions, you actually want to ignore the group structure as much as possible.

§2 Ramsey Cayley graphs

§2.1 Ramsey numbers

As a detour, we'll start with some questions from Ramsey theory. Ramsey theory contains many deep results showing that every large structure contains a very organized substructure.

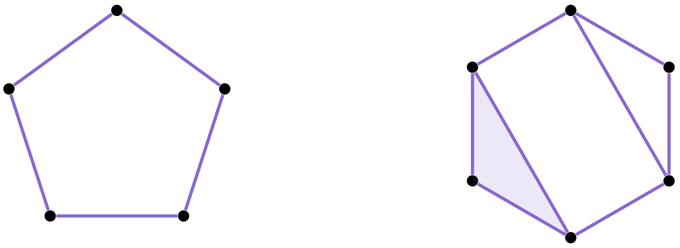
> **Theorem 2.1** (Ramsey)
>
> Every very large graph contains a large clique or independent set.

Definition 2.2. The nth Ramsey number, denoted $r(n)$, is the minimum N such that every graph on N vertices contains a clique or independent set of size n.

As a small example, with $n = 3$:

Example 2.3

We have $r(3) = 6$ — this is because a 5-cycle doesn't contain any triangle or independent set of size 3, but every 6-vertex graph does.



What about larger values of n? For $n = 4$, there's an interesting graph called the *Paley graph*; you can show that the Paley graph P_{17} has no clique or independent set of size 4, while any 18-vertex graph does, so $r(4) = 18$. Regarding $r(5)$ and $r(6)$, there's a famous quote by Paul Erdős — that if aliens threatened to obliterate Earth unless we could find $r(5)$, then we might be able to do so in a year by assembling every mathematician and every computer; but if they wanted $r(6)$, we would have no choice but to launch a preemptive attack. (This was from 1990.) The difficulty is that if you tried a brute-force approach — trying all graphs on 50 vertices — then there would be $\binom{50}{2}$ pairs for which you'd have to decide whether they're an edge or not, so you'd have $2^{\binom{50}{2}}$ graphs; this is way bigger than what we can handle.

Until a week ago, we knew that $43 \le r(5) \le 48$; there was a recent paper which replaced 48 by 46. (Of course, not every mathematician and computer has been used for this problem.) And we know $102 \le r(6) \le 147$.

§2.2 Ramsey graphs

We're mostly interested the *growth* of $r(n)$ as $n \to \infty$. The right way to parametrize this problem is through the notion of *Ramsey graphs.*

Definition 2.4. A N-vertex graph is C-Ramsey if it has no clique or independent set of size $C \log_2 N$.

The reason this is a good parametrization is that it's known the Ramsey number $r(n)$ is exponential in n. In one direction, there's the following bound.

Theorem 2.5 (Erdős–Szekeres 1935)

There is no $\frac{1}{2}$-Ramsey graph.

There was a recent breakthrough on this:

Theorem 2.6 (Campos–Griffiths–Morris–Sahasrabudhe 2023+)

There is some constant $\varepsilon > 0$ such that there is no $(\frac{1}{2} + \varepsilon)$-Ramsey graph.

This problem had been stuck for a long time; there were some really interesting ideas on quasirandom graphs, but we weren't close to getting an exponential improvement until this result. Recently, Gupta–Ndiaye–Norin–Wei improved the small constant ε to a less small constant $\mathcal{E}$.

In the other direction, the following result was an early influential application of the probabilistic method.

Theorem 2.7 (Erdős 1947)
Almost all graphs on N vertices are 2-Ramsey.

Proof. Let $n = 2\log_2 N$ (we won't worry about the fact that this might not be an integer), and consider a random N-vertex graph G. For any subset of n vertices, the probability it forms a clique is $2^{-\binom{n}{2}}$, and so is the probability it forms an independent set. There are $\binom{N}{n}$ such subsets, so by linearity of expectation

$$\mathbb{P}[G \text{ is not 2-Ramsey}] \leq \mathbb{E}[\#(\text{cliques or ISs of size } n)] = 2^{1-\binom{n}{2}} \cdot \binom{N}{n} = o(1). \qquad \square$$

So almost every graph is 2-Ramsey; the natural next question is, can we construct such a graph?

Question 2.8 (Erdős). Explicitly construct C-Ramsey graphs for any constant C.

Even if C is 10^6, we still don't know how to do this. So Ramsey graphs are a sort of dark matter for graphs — they're basically everywhere, but we don't know how to locate them. You may have heard the phrase 'searching for a needle in a haystack'; but here we're searching for *hay* in a haystack. That seems like a much easier task, but we don't know how to do it.

§2.3 Paley graphs

In the search for Ramsey graphs, there has been a proposal to look at a nice family of graphs, called *Paley graphs*, which have some very nice pseudorandom properties.

Definition 2.9. For a prime $N \equiv 1 \pmod 4$, the Paley graph on N vertices, denoted P_N, is the graph on vertex set $\mathbb{Z}_N$ where $x \sim y$ if $x - y$ is a quadratic residue.

We saw earlier that P_5 and P_{17} gave tight lower bounds on $r(3)$ and $r(4)$.

It's known that Paley graphs can't give us a Ramsey graph for *every* N:

Theorem 2.10 (Montgomery 1972)
Assuming the generalized Riemann hypothesis, there are infinitely many N with

$$\omega(P_N) \geq c \log N \log\log N.$$

(We use $\omega(G)$ to denote the size of the largest clique in G.)

Theorem 2.11 (GR 1990)
Unconditionally, there are infinitely many N with

$$\omega(P_N) \geq c \log N \log\log\log N.$$

Still, we might hope that maybe for *most* N, they do give Ramsey graphs.

On the other hand, the best upper bound we have for $\omega(P_N)$ is roughly $\sqrt{N}$, which can be proven using pseudorandom properties of P_N (e.g., the expander mixing lemma). Vinogradov showed such a bound 70 years ago, and the best bound we have today is only a factor-of-2 improvement.

Theorem 2.12 (HP, DBSW 2021)
We have $\omega(P_N) \leq \frac{1}{2}(\sqrt{2N-1}+1)$.

So we're pretty far from estimating $\omega(P_N)$, and this is considered a hard problem in number theory. (We've only talked about cliques, but P_N is self-complementary, so its independence and clique numbers are equal.)

§2.4 Cayley graphs

Paley graphs are examples of a more general family of graphs, called *Cayley graphs*.

Definition 2.13. For a group G and a symmetric subset $S \subseteq G$, the Cayley graph G_S is the graph with vertex set G where $x \sim y$ if $xy^{-1} \in S$.

We don't require G to be abelian. *Symmetric* means that if $x \in S$ then $x^{-1} \in S$; the reason we assume this is so that the graph doesn't have to be directed.

Example 2.14
Paley graphs are Cayley graphs where $G = \mathbb{Z}_N$ and S is the set of quadratic residues.

But we can look for Ramsey graphs among Cayley graphs more generally.

Conjecture 2.15 (Alon) — There exists C such that every finite group has a C-Ramsey Cayley graph.

In discrete geometry and information theory, there are contexts where it's useful to have a C-Ramsey graph which is also Cayley, so this conjecture has a number of interesting applications.

A very natural approach to this conjecture is to look at a uniform random Cayley graph — we look at every element and its inverse, and put them in S with probability $\frac{1}{2}$ (independently of the other elements).

Question 2.16. Are random Cayley graphs typically Ramsey?

To answer this, it's natural to study the clique number of random Cayley graphs.

Theorem 2.17 (Alon)
For any group G of order N, the clique number of a random Cayley graph on G is $O(\log^2 N)$ (asymptotically almost surely).

This is in the right ballpark, but it's not quite $\log N$. However, we *do* have bounds of $\log N$ for *some* groups.

Theorem 2.18 (Green 2005, Green–Morris 2016)
For N prime, the clique number of a uniform random Cayley graph on $\mathbb{Z}_N$ is $(2+o(1))\log_2 N$ (asymptotically almost surely).

So this suggests that maybe looking at random Cayley graphs is the right approach. But it turns out it doesn't work to get Alon's conjecture for certain groups.

Theorem 2.19 (GM)

A uniform random Cayley graph on $\mathbb{F}_2^d$ has clique number $\Theta(\log N \log\log N)$ (asymptotically almost surely, where $N = 2^d$).

The reason for this is that $\mathbb{F}_2^d$ has lots of subspaces, and you'll likely get all the nonnegative elements of one subspace of size $\log N \log\log N$ (this is the size, not the dimension). Then your Cayley graph will have a clique consisting of that subgroup, and each of its cosets. So it's the subspaces that are kind of problematic here — subspaces have a much higher probability of being a clique than a random set would.

The proofs of these results are very specific to these groups. But it turns out that we can get the same upper bound as Theorem 2.19 more generally.

Theorem 2.20

The clique number of a uniform random Cayley graph on any group G of order N is $O(\log N \log\log N)$ (asymptotically almost surely).

You might expect that we don't use the group structure much here; and indeed, the proof does ignore most of the group structure, after doing some finite simple thing.

Meanwhile, regarding Alon's conjecture:

Theorem 2.21

For almost all N, all abelian groups G of order N have a Cayley graph which is C-Ramsey.

So we can't prove the conjecture even for all abelian groups, but we *can* do it for all abelian groups of size N for most N. (The values of N we can't do it for are ones with lots of factors of 2 and 3.)

§2.5 Counting sets with small product set

These questions roughly boil down to counting sets with small product set. Specifically, we're interested in the following object.

Definition 2.22. For a set $A \subseteq G$, we define $AA^{-1} = \{ab^{-1} \mid a, b \in A\}$.

Then A ends up being a clique in G_S if and only if $AA^{-1} \setminus \{1\} \subseteq S$. (This comes from directly translating the condition for something to be an edge — we have $a \sim b$ if and only if $ab^{-1} \in S$.)

This means if AA^{-1} is small, there is a higher probability that A will form a clique — $|AA^{-1}|$ can be anywhere from $|A|$ to $|A|^2$, and we have to worry about the sets A for which it's small. To handle this, we have a counting result.

Theorem 2.23

In any group G of order N, the number of subsets A with $|A| = n$ and $|AA^{-1}| \leq Kn$ is at most

$$N^{C(K+\log n)}(CK)^n$$

(where C is an absolute constant).

This bound looks complicated, but each of the factors is necessary.

- If G has a subgroup of size Kn, then all subsets A of that subgroup will satisfy $|AA^{-1}| \leq Kn$, since AA^{-1} still lives inside that subgroup. So this gives $\binom{Kn}{n}$ subsets, corresponding to the $(CK)^n$ term.
- You could also start with a subgroup of size n and add K arbitrary elements; this gives N^{CK} sets A.
- The $N^{\log n}$ term corresponds to the $\mathbb{F}_2^d$ example, where you have lots of subspaces.

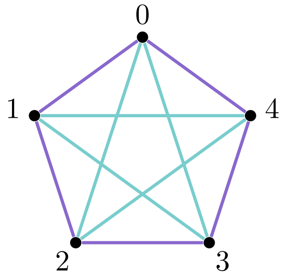
So you need each of these three terms as factors, and it turns out that if you multiply them, then you get an upper bound. (There are more specific bounds for some cases, but this is already quite useful.)

§2.6 From additive combinatorics to edge-colored graphs

How do we go from a question about groups to one where we can ignore group structure? Suppose we start with a group G of order N. We then color the edges of K_G (the complete graph on vertex set G) by assigning each edge xy the color $\{xy^{-1}, yx^{-1}\}$.

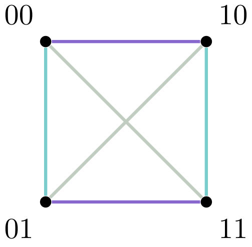
Example 2.24

Suppose $G = \mathbb{Z}_5$. Then we get one color class corresponding to consecutive pairs (pairs which differ by 1), and another color class corresponding to pairs which differ by 2.

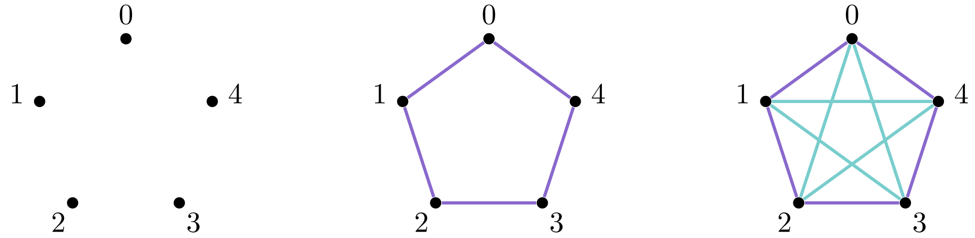


Example 2.25

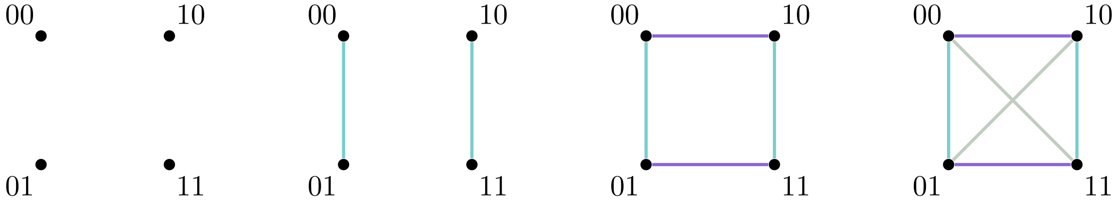
Suppose $G = \mathbb{Z}_2^2$. Then we'll have three colors, where each color class forms a perfect matching (we have one color for the difference 01, one for 10, and one for 11).



If we want to build a Cayley graph on $\mathbb{Z}_5$, then we have two color classes, and we can take zero, one, or two of them. If we take zero, we get an empty graph; if we take one, we get a 5-cycle; and if we take two, we get a complete graph. So there's only three possible Cayley graphs on $\mathbb{Z}_5$ — the empty graph, the complete graph, and a 5-cycle.



We can do the same for $\mathbb{Z}_2^2$ — to get a Cayley graph on $\mathbb{Z}_2^2$, for each of the three color classes, we choose whether to include it or not. If we include none, we get the empty graph; if we include one, we get a perfect matching; if we include two, we get a 4-cycle; and if we include all three, we get the complete graph.



In general, a Cayley graph on G is the edge-union of some color classes. We'll study a more general graph model that captures this:

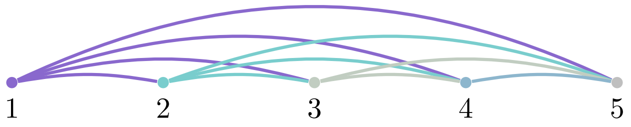
Definition 2.26. Suppose we have an edge-coloring c of a complete graph. Then an entangled graph is the edge-union of some color classes.

Definition 2.27. The random entangled graph, denoted $G_c(p)$, is obtained by including each color class with probability p independently.

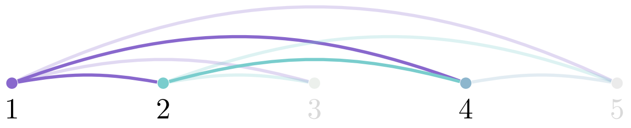
When $p = \frac{1}{2}$, this exactly captures the random Cayley graphs that we were studying earlier (for the special colorings we described).

Definition 2.28. We say c is Δ-bounded if each color class has maximum degree at most Δ.

This is a good condition to impose if we want to study the clique number. To see what goes wrong otherwise, imagine that each color class formed a star (with one star originating from 1, another from 2, and so on).



Then each color class individually is reasonably sparse, but $G_c(p)$ will have a pretty large clique — if we look at all the colors that get picked and take the root vertex of each, then these vertices will form a clique. And the number of color classes that get picked will concentrate at pN, so you'll get an enormous clique.



So it's natural to put a bound on the maximum degree. In the Cayley graph setting, we can take $\Delta = 2$. Now we'll completely forget about the group structure, and only remember the fact that each color class has bounded degree.

> **Question 2.29.** What can we say about $\omega(G_c(p))$ if c is Δ-bounded?

First, we can prove an analog of the small-doubling result (Theorem 2.23).

Theorem 2.30

In a Δ-bounded edge-coloring of K_N, the number of n-vertex subsets with at most Kn colors is at most

$$N^{C\Delta(K+\log n)}(C\Delta K)^n.$$

When proving this, you can actually assume $\Delta = 1$ through an application of Vizing's theorem — Vizing's theorem says that if you have a graph with maximum degree Δ, then you can edge-color it with $\Delta+1$ colors such that eacy color class is a matching. Here, doing this to each of the original color classes means we can multiply the number of colors by $\Delta + 1$ and make the coloring 1-bounded.

This implies the following bound on $\omega(G_c(p))$, the same as for random Cayley graphs.

Theorem 2.31

If an edge-coloring c of K_n is Δ-bounded, then asymptotically almost surely

$$\omega(G_c(p)) = O(\log N \log\log N).$$

§2.7 A simpler version of Theorem 2.30

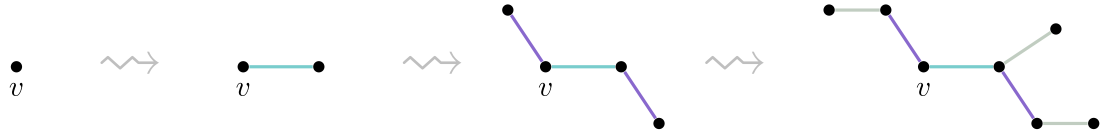
The proof of Theorem 2.30 is too long, but we'll present a proof of a simpler result that captures some of the main ideas.

Proposition 2.32

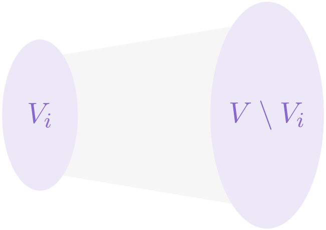
For any proper edge-coloring of K_n which uses at most Kn colors, there is a spanning tree which uses $O(K \log n)$ colors.

This gives Theorem 2.30 with a weaker bound of $K \log n$ in place of $K + \log n$ (here we're assuming $\Delta = 1$) — there are at most N^2 colors in the big graph, so there's $N^{CK \log n}$ ways to choose *which* colors you see in the spanning tree of your n-vertex subset.

Proof. We'll try to grow a connected component one color at a time. So we start off with a single vertex v (which we choose arbitrarily); and then we're going to add one color, and then another color, and so on. Each time we add a new color, we want to grow the component as quickly as possible — so we greedily choose the next color so that the connected subgraph using the chosen colors from our starting vertex v is as large as possible.



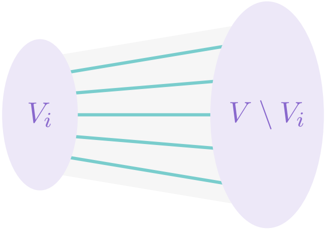
Let V_i be the set of vertices that we have after adding i colors (so $V_0 = \{v\}$, and V_i is a connected set using the i chosen colors). To choose the next color, consider V_i and the rest of the graph $V \setminus V_i$.



There's lots of edges between V_i and $V \setminus V_i$ — specifically, there's $|V_i|\,(n - |V_i|)$ edges. And there's Kn colors in total, so we can find some color such that there are at least

$$\frac{|V_i|\,(n - |V_i|)}{Kn}$$

edges from V_i to $V \setminus V_i$ of that color. And that's the color we'll add as the $(i+1)$st color.



So we've found a color that has lots of *edges* coming out of V_i. This might not immediately mean that adding it grows our component by a lot, but it actually does because of the fact that the edge-coloring is *proper* (i.e., no two edges incident to the same vertex can get the same color). This means each of the edges out of V_i with the chosen color must go to a different vertex, so when we go to V_{i+1}, we add one vertex for each of these edges.

For now, let's assume that $|V_i| \leq \frac{1}{2}n$ (we'll explain what happens past this later). Then we get

$$|V_{i+1}| \geq |V_i| + \frac{|V_i|\,(n - |V_i|)}{Kn} \geq \left(1 + \frac{1}{2K}\right)|V_i|\,.$$

So as long as our original component V_i had less than half the vertices, it grows by a factor of $1 + \frac{1}{2K}$. This means after adding roughly $2K$ colors we'll have roughly doubled its size, so after adding $2K \log n$ steps, we'll have gotten all the way up to having $\frac{1}{2}n$ vertices.

Past $\frac{1}{2}n$, if you do the same calculation, it's no longer the case that V_i grows by a factor of $1 + \frac{1}{2K}$, but we can instead say that its complement is *shrinking* by this factor. So after another $2K \log n$ steps, we'll have gotten every vertex, which means we have a spanning tree. □

It turns out that Proposition 2.32 is sharp (up to the constant). But if you only want 99% of the vertices, rather than 100%, then $K + \log n$ colors is actually enough (and this is also sharp); and this is how we get the exponent of $K + \log n$ in Theorem 2.30. For this, you roughly pick K random colors; then instead of already getting lots of connections, you get subsets of size roughly K that see a constant fraction of the colors, and you can use this to connect things up with roughly $\log n$ extra colors.

§3 Ruzsa's conjecture

One big result from additive combinatorics is the following theorem.

Definition 3.1. The exponent of a group G is the smallest r such that $x^r = 1$ for all $x \in G$.

Theorem 3.2 (Freiman–Ruzsa)

If an abelian group G has exponent r and $A \subseteq G$ satisfies $|A + A| \leq K\,|A|$, then A is contained in a subgroup H with $|H| \leq C(r, K)\,|A|$.

In words, if a set A has bounded doubling, then it has to lie in a small subgroup.

Question 3.3. Can we get good quantitative estimates on $C(r, K)$?

Conjecture 3.4 (Ruzsa) — We can take $C(r, K) = r^{O(K)}$.

It's known that this has to be a *lower* bound on $C(r, K)$ — you can get constructions in $\mathbb{F}_2^d$ that require this. But we'd like an *upper* bound. There's been various improvements on quantitative bounds over the years; these very much used the group structure and lots of Fourier analysis. For $G = \mathbb{F}_p^d$, we know the answer, due to some nice techniques called *compression methods* from extremal set theory. But for general abelian groups, we didn't know it until recently.

Theorem 3.5 (Fox–Pham)

Ruzsa's conjecture is true.

In addition to combinatorial tools like the ones we've discussed, there's an additional tool from a paper of Fox, Pham, Sammy Luo, and Yunkun Zhou; that lemma turns out to be very useful at the end of the proof. So sometimes most of the proof is by combinatorial methods, and it turns out that we can pull in these other tools to finish off.

§4 Information theory

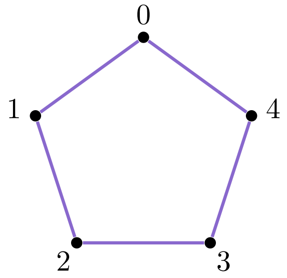
§4.1 Improved communication through repetition

Definition 4.1. The nth power of a graph $G = (V, E)$, denoted G^n, is the graph on vertex set V^n where $u \sim v$ if $u \neq v$ and for each i, we either have $u_i = v_i$ or $(u_i, v_i) \in E$.

This comes up in information theory because the independence number of G^n is the maximum number of messages that a channel with confusion graph G can communicate without error in n uses.

Example 4.2

Suppose that G is a 5-cycle; this means we have 5 elements of our alphabet (which we'll call 0, 1, . . . , 4); and 0 and 1 can be confused, 1 and 2 can be confused, and so on.



If we use the channel only once, then the maximum number of messages you can send is 2, corresponding to the independent set $\{0, 2\}$. But if you can use the channel multiple times, then you can do better. You *could* send at least four messages by taking the independent set $\{0, 2\}^2$ (this corresponds to sending 0 or 2 on the first use, and 0 or 2 again on the second). But there's an even bigger independent set, of size 5; this comes from the following general fact.

> **Fact 4.3** — If G is self-complementary, then $\alpha(G^2) \geq |V(G)|$.

Proof. If we let π be the isomorphism from G to its complement, then

$$\{(x, \pi(x)) \mid x \in V(G)\}$$

is an independent set in G^2. To see why, consider two of its elements $(x, \pi(x))$ and $(y, \pi(y))$. For these to form an edge, you need x and y to form an edge in G; but then $\pi(x)$ and $\pi(y)$ would *not* form an edge in G (because π sends edges to non-edges), so we don't get an edge in G^2. □

This is relevant because it has to do with understanding how efficient such a channel is — we often study efficiency by looking at

$$c_n(G) = \alpha(G^n)^{1/n},$$

which is the maximum number of messages per use of the channel if you can use it n times.

> **Definition 4.4.** The Shannon capacity of G is $c(G) = \lim_{n\to\infty} c_n(G)$.

We really don't understand this — we don't know the Shannon capacity of a 7-cycle or of the random graph $\mathcal{G}(n, \frac{1}{2})$, for instance (the lower and upper bounds are something like $\log n$ and $\sqrt{n}$, respectively).

But Alon and Orlitsky proved that there are self-complementary Ramsey graphs; this means

$$c_1(G) = \Theta(\log |G|) \quad \text{and} \quad c_2(G) \geq \sqrt{|G|}.$$

So by using the channel twice, you can communicate a lot more than what you could by using the channel once — you jump from a logarithmic to square-root number.

§4.2 Self-complementary Ramsey Cayley graphs

> **Conjecture 4.5** (Alon–Orlitsky 1995) — There are self-complementary Ramsey Cayley graphs.

We know that there are self-complementary Ramsey graphs and there are Ramsey Cayley graphs, but we didn't know how to put these together until now.

We'd like to find a construction on $\mathbb{F}_5^d$. But if we take a *random* Cayley graph on $\mathbb{F}_5^d$, it'll have a clique of size $\Theta(\log N \log\log N)$, for the same reason as $\mathbb{F}_2^d$ in Theorem 2.19 (related to subspaces); this means it won't be Ramsey (it also won't be self-complementary).

So choosing a Cayley graph on $\mathbb{F}_5^d$ completely at random doesn't work; but it turns out we *can* still find a construction on $\mathbb{F}_5^d$.

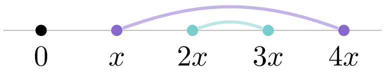
Theorem 4.6

There are self-complementary Ramsey Cayley graphs on $\mathbb{F}_5^d$.

To prove this, the first idea is that from the counting bound (Theorem 2.23), it suffices to pick a random self-complementary Cayley graph on $\mathbb{F}_5^d$ in such a way that:

- Every A that could possibly be a clique satisfies $|A - A| \geq |A| \log |A|$. (The reason for this is that it ensures $K \geq \log n$ in Theorem 2.23.)
- The probability of A being a clique is at most $2^{-\Omega(|A-A|)}$. (This is what you'd have if you chose the graph *uniformly* at random.)

Here's what the construction looks like: if we look at any nonzero $x \in \mathbb{F}_5^d$, then x is on a line through the origin which also contains $2x$, $3x$, and $4x$. If we place x in S then we also have to place $4x$ in S; and similarly, if we place $2x$ in S then we also have to place $3x$ in S. So we have two pairs — $\{x, 4x\}$ and $\{2x, 3x\}$ — where for each pair, either both or neither of its elements gets placed in S. And to generate our random Cayley graph, we'll take exactly *one* of those two pairs to be in S, independently at random for each line.



This gives a random (but not uniform) generating set S, which is symmetric; and this random Cayley graph is self-complementary with isomorphism $\pi(x) = 2x$.

We have that if $x \in S$, then $2x \notin S$; this means that if A is a clique, we have

$$|A + 2 \cdot A| = |A|^2$$

(where $A + 2 \cdot A = \{a + 2b \mid a, b \in A\}$). This is because if A is a clique, then all elements of the form $a + 2b$ have to be distinct — if we had $a + 2b = c + 2d$, then we could rearrange this to $a - c = 2(b - d)$; the fact that A is a clique means that $a - c$ and $b - d$ both have to be in the generating set S, but this is impossible since S can't contain both x and $2x$ for any x.

And this means we have

$$|A|^2 = |A + 2 \cdot A| \leq |A + A + A| \leq |A - A|^3 |A|^{-2},$$

where the last inequality is the Plünnecke–Ruzsa inequality from additive combinatorics. This gives

$$|A - A| \geq |A|^{4/3}.$$

This is a lot bigger than $|A| \log |A|$, so we've achieved our first goal with a lot of wiggle room. And this is the main part of the proof.

§5 General random graph models

Finally, we'll turn to a problem about general random graph models. We've seen various random graph models — for example, the Erdős–Rényi random graph $\mathcal{G}(n, p)$ (where you pick each edge with probability p independently), random Cayley graphs, random entangled graphs, and so on (there's also other examples like random Latin square graphs). All these models have some amount of independence to them; and there's a simple way of describing this independence.

Definition 5.1. Suppose G is a random graph such that each pair of vertices e appears as an edge with probability p_e, and appears independently of all edges apart from those in a graph G_e. We say G is Δ-independent if $\Delta(G_e) \leq \Delta$ for each pair e.

This looks a lot like the condition in the Lóvasz local lemma — there you have a collection of bad events, and you know that each is independent of all but some bounded number of bad events. Here, if we look at the event for whether one edge is included, the *number* of edges it's dependent on is not bounded, but those edges form a bounded-degree graph; so this is a weaker condition in some sense.

It turns out lots of random graph models satisfy this condition; and this is enough to prove bounds similar to the ones we've seen.

Theorem 5.2

For fixed $0 < p < 1$, let G be a Δ-independent random graph on N vertices where $p_e = p$ for all e.

- If $\Delta = N^{o(1)}$, then $\omega(G) \geq (2 - o(1)) \log_{1/p} N$ (asymptotically almost surely).
- If $\Delta = O(1)$, then $\omega(G) \leq O(\log N \log \log N)$.

The first statement says that we can't beat the lower bound on Ramsey numbers (in Theorem 2.7) by using some different Δ-independent random graph model in place of $\mathcal{G}(n, \frac{1}{2})$.

Meanwhile, the second statement gets the same upper bound as the one we saw for Cayley graphs in Theorem 2.20 — so this means $\mathbb{F}_2^d$ is really the worst case.

Alon's conjecture (that all groups have Ramsey Cayley graphs) is still open. You can generalize the ideas from $\mathbb{F}_5^d$ (in Theorem 4.6) to prove that many other groups have Ramsey Cayley graphs; but factors of 2 and 3 are an issue, because they don't have enough room for such a construction. Here's a toy conjecture that's maybe at the heart of the problem.

Conjecture 5.3 — There is a 2-coloring of $\mathbb{F}_2^d$ such that there is no subspace of size Cd whose nonzero elements are monochromatic.

Subspaces are a special case of the possible cliques that you could have in a random Cayley graph. If you just tried to pick a generating set uniformly at random, then you'd only get $d \log d$; so to prove this, you have to do a bit better than that.