# Topics in additive combinatorics


Rudi Mrazović

Balliol College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Trinity 2016

*To my Mum, Dad and brother*
*for their love and support*

# Acknowledgements

I should like to thank my supervisor Ben Green for his support, advice and encouragement.

The first year of my research was funded by Trinity College, Cambridge and Cambridge Overseas Trust. Remaining three years were funded by the University of Oxford. I thank all of these institutions.

I should also like to thank my colleagues and collaborators Freddie Manners and Sean Eberhard for numerous fruitful discussions and suggestions, my colleagues Jakub Konieczny, Przemysław Mazur, Sofia Lindqvist and Aled Walker, and the range of mathematicians with whom I have had useful conversations.

# Abstract

This thesis deals with four problems in additive combinatorics. After giving a brief introduction to the field and overview of the results in Chapter 1, in Chapter 2 we consider the problem of finding the clique number of the Cayley graph on $\mathbb{F}_2^n$ generated by a random subset. We prove a number of results, most notably that for $n$ in a set of density 1, the clique number is concentrated on a single value.

In Chapter 3 we prove that a randomly chosen subset of a finite group is a randomness extractor with high probability. More precisely, we prove that for every fixed $\delta > 0$, given a finite group $G$ and $A \subset G$ a random subset of density $\frac{1}{2}$, we prove that with high probability for all subsets $|X|, |Y| \geq \log^{2+\delta} |G|$ for $(\frac{1}{2} + o(1))|X||Y|$ of the pairs $(x, y) \in X \times Y$ we have $xy \in A$.

In Chapter 4 we give a new probabilistic model for the Paley graph which incorporates some multiplicative structure and as a result captures the Graham-Ringrose phenomenon, namely that the its clique number is sometimes a bit larger than what one might expect when considering the usual random model (random Cayley graph). We prove that if we sample such a random graph independently for every prime, then almost surely (i) for infinitely many primes $p$ the clique number is $\Omega(\log p \log\log p)$, whilst (ii) for almost all primes the clique number is $(2 + o(1)) \log p$.

Whereas in the previous chapters we were mostly concerned with generic sets, in Chapter 5 we consider a rather different problem concerning additive properties of a specific set. We prove an asymptotic for the number of additive triples of bijections $\{1, \dots, n\} \to \mathbb{Z}/n\mathbb{Z}$, that is, the number of pairs of bijections $\pi_1, \pi_2 \colon \{1, \dots, n\} \to \mathbb{Z}/n\mathbb{Z}$ such that the pointwise sum $\pi_1 + \pi_2$ is also a bijection.

# Contents

# Chapter 1

# Introduction

Additive combinatorics is, roughly speaking, an area of mathematics that deals with additive properties of sets of numbers (e.g. integers, or more generally, subsets of abelian groups).

For example, consider the following problem. Let $A \subset \mathbb{Z}$ be a finite set and consider its sumset, i.e. the set

$$A + A = \{a_1 + a_2 \colon a_1, a_2 \in A\}.$$

One could argue that for a *typical* set $A$, the size of $A + A$ should be close to $\binom{|A|}{2}$, that is the number of two element subsets of $A$. However, sets $A$ for which the size of $A + A$ is small (e.g. linear in $|A|$) certainly do exist – one could take for example $A = \{1, \ldots, n\}$ or, more generally, any arithmetic progression with few elements removed. One could ask are there any other, substantially different, examples of sets with the same property. In other words, and more formally, if we know that $|A + A| \leq K|A|$ for some small constant $K > 1$, what can be said about the set $A$? Does $A$ look like an arithmetic progression? An (incomplete) answer to these questions was given by Freĭman and we will actually use a variant of his result in Chapter 2.

Questions about sumsets and their structural properties like the one we have just described are very common in additive combinatorics. In Chapter 2 we will consider the following problem.

The *clique number* of a graph is defined to be the size of its largest clique, i.e. the size of its largest complete subgraph. Given an abelian group $G$ and some set of generators $A$ one can consider corresponding *Cayley sum graph* which has $G$ as the set of vertices and edges joining vertices $x \neq y$ if and only if $x + y \in A$. If $A$ is taken

at random, how does the corresponding Cayley sum graph look like? Specifically, what is its clique number? One can easily see that this is equivalent to asking what is the size of the largest set $X$ such that its *restricted sumset*

$$X\widehat{+}X = \{x_1 + x_2 \colon x_1, x_2 \in X, x_1 \neq x_2\}$$

lies inside $A$.

In Chapter 2 we will address this question in the case of the group $G = \mathbb{F}_2^n$ and prove, among other results, that the clique number of the corresponding Cayley sum graph will almost always be equal to one fixed value. We will also prove an analogous result for the clique number.

In Chapter 3 we will be interested in a different but related problem. We can also motivate it via graph theory.

Let $V$ be a finite set of vertices and for any pair of vertices $x, y \in V, x \neq y$ draw an edge between them with probability $1/2$, independently of all other choices. This way of constructing a random graph is known in the literature as the Erdős–Rényi model.

First of all, we expect that the total number of edges in this graph should be close to $\frac{1}{2}\binom{|V|}{2}$ with high probability. However, we can also expect that for any two large sets of vertices $X, Y \subset V$, the number of pairs of vertices $x \in X$, $y \in Y$ for which $xy$ is an edge is very close to $|X||Y|/2$ and that this is with high probability simultaneously satisfied for all large sets $X$ and $Y$. In the first section of Chapter 3 we will give an easy proof of this statement. However, the analogous statement for random Cayley graphs is significantly harder to prove and most of the Chapter 3 is devoted to this task.

One of the main motivations for considering random Cayley graphs is that they are believed to be a good random model for the Paley graph, that is the graph with the set of vertices $\mathbb{Z}/p\mathbb{Z}$ ($p$ prime) and edges connecting vertices $x$ and $y$ if and only if $x + y$ is a quadratic residue.

Finding the clique number of the Paley graph on $\mathbb{Z}/p\mathbb{Z}$ is closely related to finding the smallest quadratic nonresidue modulo prime $p$ which is an important open problem in analytic number theory. Green and Morris [GM15] proved that if one samples independently an infinite sequence of Cayley graphs, one for each prime $p$, then with high probability the clique number would be $(2+o(1))\log p$ along the whole sequence. For this reason one might think that the clique number of the Paley graph should also be close to this value. Although this might be true for most primes, there are cer-

tainly infinitely many exceptions. Indeed, Graham and Ringrose [GR90] proved that for infinitely many primes $p$ the least quadratic nonresidue $q$ is $\Omega(\log p \log\log\log p)$, that is at least $c \log p \log\log\log p$ for some constant $c > 0$. Obviously, for these primes the set $\{1, 2, \ldots, q/2\}$ forms a large clique.

In Chapter 4 we introduce a new random model for the Paley graph which incorporates some multiplicative structure. As a result, this random model captures the Graham-Ringrose phenomenon which wasn't the case with random Cayley graphs. More precisely, if we sample an infinite sequence of random graphs using this random model (one for each prime $p$), then the clique number would be $\Omega(\log p \log\log p)$ for infinitely many values of $p$. However, we prove that these are actually exceptional cases, that is, with high probability, for primes $p$ in a set of density 1 the clique number is $(2 + o(1)) \log p$.

Finally, in the last chapter we deal with a somewhat different problem. Very often when investigating additive structure of a set, one is interested in linear relations satisfied by the elements of that set. The simplest of these relations is $x + y = z$ and triples satisfying this relation are called *additive triples*. In Chapter 5 we will be interested in the number of additive triples in the set of all bijections from $\{1, \ldots, n\}$ to $G$, where $G$ is a finite abelian group of order $n$.

We will not give a more detailed overview of the results at this point, since each chapter contains much more information and motivation of a particular problem.

Chapter 5 is based on the joint paper with Sean Eberhard and Freddie Manners, and although other chapters were not done in a formal collaboration, I have certainly benefited a lot from conversations with other people. I have tried to acknowledge these particular conversations, advice and suggestions in the corresponding chapters.

## 1.1 Fourier analysis on finite abelian groups

One of our main tools in Chapter 2 and Chapter 5 will be Fourier analysis on finite groups. In this section we give a brief introduction into that field.

### 1.1.1 Characters

In this section, $G$ will always denote a finite abelian group. A character on $G$ is a homomorphism $\gamma\colon G \to \mathbb{C}^\times$. Since every element in $G$ has finite torsion, it is easy to see that $|\gamma(x)| = 1$ for every $x \in G$.

The set of characters on $G$ together with the pointwise multiplication forms a

group; we will call it the *dual group of* $G$ and denote it by $\widehat{G}$.

For example, in the case of the cyclic group $\mathbb{Z}/n\mathbb{Z}$, for any $r \in \mathbb{Z}/n\mathbb{Z}$ we have the character $x \mapsto e(rx/n)$ (we remind the reader of the notation $e(y) = e^{2\pi iy}$). These are actually all characters on $\mathbb{Z}/n\mathbb{Z}$, so the dual group is isomorphic to $\mathbb{Z}/n\mathbb{Z}$.

Let $H$ be another finite abelian group and $\gamma_1$ and $\gamma_2$ characters on $G$ and $H$, respectively. Then, obviously, $(g,h) \mapsto \gamma_1(g)\gamma_2(h)$ is a character on the group $G \oplus H$. Moreover, one can prove that all the characters on $G \oplus H$ are of this form.

From the last two paragraphs and the fact that every finite abelian group $G$ is a product of finite cyclic groups, we see that $G$ is isomorphic to its dual group $\widehat{G}$. To be specific, if $G = \mathbb{Z}/n_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/n_k\mathbb{Z}$, then for any $(r_1, \ldots, r_k) \in G$ we have the character

$$(x_1, \ldots, x_k) \mapsto e\left(r_1x_1/n_1 + \cdots + r_kx_k/n_k\right).$$

### 1.1.2 Orthogonality relations

We equip group $G$ with the uniform probability measure, which induces corresponding scalar product and $L^2$ norm

$$\langle f, g\rangle_{L^2(G)} = \mathbb{E}_{x\in G} f(x)\overline{g(x)} \quad \text{and} \quad \|f\|_{L^2(G)} = \left(\mathbb{E}_{x\in G}|f(x)|^2\right)^{1/2}.$$

The crucial property of characters are the so-called orthogonality relations, that is for any two characters $\gamma_1, \gamma_2 \in \widehat{G}$ we have

$$\langle \gamma_1, \gamma_2\rangle_{L^2(G)} = \begin{cases} 0 & \text{if } \gamma_1 \neq \gamma_2, \\ 1 & \text{if } \gamma_1 = \gamma_2 \end{cases}$$

and dually, for any two elements $x_1, x_2 \in G$

$$\mathbb{E}_{\gamma\in\widehat{G}}\gamma(x_1)\overline{\gamma(x_2)} = \begin{cases} 0 & \text{if } x_1 \neq x_2, \\ 1 & \text{if } x_1 = x_2. \end{cases}$$

### 1.1.3 Fourier coefficients

From the previous subsection we know that the set of characters ($|G|$ of them in total) is an orthonormal set in the space $\mathbb{C}^G$ of all complex functions of $G$. Since this space is $|G|$-dimensional, the set of characters is actually an orthonormal basis.

Coefficients appearing in the expansion of a function $f$ in this basis are called *Fourier coefficients of* $f$. More precisely, for any $\gamma \in \widehat{G}$ we define corresponding

Fourier coefficient by

$$\widehat{f}(\gamma) = \langle f, \gamma \rangle_{L^2(G)} = \mathbb{E}_{x \in G} f(x) \overline{\gamma(x)}.$$

We call the function $\widehat{f}$ the *Fourier transform of* $f$. The expansion of $f$ in the character basis is

$$f(x) = \sum_{\gamma \in \widehat{G}} \widehat{f}(\gamma)\gamma(x), \quad \text{for every } x \in G.$$

This identity is known as the *Fourier inversion formula*.

It is clear from the definition that $f \mapsto \widehat{f}$ is a linear map which sends a function from $\mathbb{C}^G$ to a function from $\mathbb{C}^{\widehat{G}}$. It turns out that there is a natural choice of measure on $\widehat{G}$ which induces the scalar product on $\mathbb{C}^{\widehat{G}}$ for which the map $f \mapsto \widehat{f}$ is an isometry.

The measure in question is the counting measure and the group $\widehat{G}$ equipped with it induces the scalar product and the $l^2$ norm on $\mathbb{C}^{\widehat{G}}$

$$\langle f, g \rangle_{l^2(\widehat{G})} = \sum_{\gamma \in \widehat{G}} f(\gamma)\overline{g(\gamma)} \quad \text{and} \quad \|f\|_{l^2(\widehat{G})} = \left( \sum_{\gamma \in \widehat{G}} |f(\gamma)|^2 \right)^{1/2}.$$

As we already mentioned, map that sends a function to its Fourier transform is isometry (with the choices of measures introduced above), and this is the content of *Parseval identities*

$$\langle f, g \rangle_{L^2(G)} = \langle \widehat{f}, \widehat{g} \rangle_{l^2(\widehat{G})} \quad \text{and} \quad \|f\|_{L^2(G)} = \|\widehat{f}\|_{l^2(\widehat{G})}.$$

### 1.1.4 Convolutions

A very important tool will be that of discrete *convolution* which for two functions $f$ and $g$ on $G$ is defined as

$$f * g(x) = \mathbb{E}_{y \in G} f(y)g(x - y).$$

An important special case will be when functions $f$ and $g$ are indicator functions of some sets $A, B \subset G$. In that case one can easily see that $1_A * 1_B(x)$ is, up to normalization factor of $1/|G|$, equal to the number of representations $x = a + b$ with $a \in A, b \in B$. Hence,

$$\operatorname{supp}(1_A * 1_B) = A + B$$

and so it is hardly surprising that convolutions have a prominent role in the study of sumsets. The following will be the crucial identity when considering Fourier transform of the convolution

$$\widehat{f * g}(\gamma) = \widehat{f}(\gamma)\widehat{g}(\gamma).$$

We should also mention that in Chapter 4 we will use Fourier analysis on $\mathbb{Z}$. We will equip it with the counting measure, and its dual group can be identified with $\mathbb{R}/\mathbb{Z}$. Namely, for any $\theta \in \mathbb{R}/\mathbb{Z}$ we have the character $n \mapsto e(n\theta)$. If we equip the dual group with the uniform probability measure, than all of the formulas that appeared in this section are still valid, and since all of the functions we will work with would be finitely supported, all of the sums and integrals would be convergent for trivial reasons.

## 1.2 Number theoretic background

In this section we introduce some of the basic number theoretical concepts and results we will use later, mostly in Chapter 4. All of the content present in this section can be found e.g. [IK04].

We will say that a function $\chi\colon \mathbb{Z} \to \mathbb{C}$ is a Dirichlet character modulo $m$ if it satisfies the following three properties

(i) it is $m$-periodic, i.e. $\chi(n+m) = \chi(n)$ for every $n$,

(ii) $\chi(n) \neq 0$ if and only if $n$ is coprime to $m$,

(iii) it is completely multiplicative, i.e. $\chi(n_1)\chi(n_2) = \chi(n_1 n_2)$ for all $n_1, n_2$.

Obviously, there is an one-to-one correspondence between Dirichlet characters modulo $m$ and characters on the multiplicative group $(\mathbb{Z}/m\mathbb{Z})^*$. One can also easily see that for coprime numbers $m_1$ and $m_2$, if $\chi_1$ and $\chi_2$ are Dirichlet characters modulo $m_1$ and $m_2$, then $\chi_1\chi_2$ is a Dirichlet character modulo $m_1 m_2$.

We will say that a Dirichlet character modulo $m$ is *principal* if it assumes value 1 for all arguments which are coprime to $m$ (or, in other words, if it corresponds to the trivial character on $(\mathbb{Z}/m\mathbb{Z})^*$). Nonprincipal Dirichlet characters exhibit random-like behaviour and there are lots of results illustrating that phenomenon. The only result of that type that we will use is the content of the following proposition, the proof of which can be found in e.g. [MV07, Theorem 13.7].

**Proposition 1.2.1.** *Let $\chi$ be a nonprincipal Dirichlet character modulo $r$. Then, if the Generalised Riemann Hypothesis holds,*

$$\left|\sum_{p\leq x}\chi(p)\right| \ll x^{1/2}\log rx.$$

Some of the simplest and most commonly used Dirichlet characters are those given by the Legendre symbol. For a prime $p > 2$ we say that an integer $n$ coprime to $p$ is a quadratic residue modulo $p$ if there is an integer $x$ such that $n \equiv x^2 \pmod{p}$. If there is no such integer we say that $n$ is a quadratic nonresidue modulo $p$. We define the appropriate Legendre symbol by

$$\left(\tfrac{n}{p}\right) = \begin{cases} 1 & \text{if } n \text{ is quadratic residue modulo } p, \\ -1 & \text{if } n \text{ is quadratic nonresidue modulo } p, \\ 0 & \text{if } p|n. \end{cases}$$

**Theorem 1.2.2** (Law of quadratic reciprocity)**.** *Let $p, q > 2$ be prime numbers. Then*

$$\left(\tfrac{p}{q}\right)\left(\tfrac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

We will also use the following useful formula for $\left(\frac{2}{\cdot}\right)$.

**Proposition 1.2.3** (Supplementary formula for $\left(\frac{2}{\cdot}\right)$)**.** *For prime $p > 2$ we have*

$$\left(\tfrac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}.$$

We will also need the following foundational result of analytic number theory.

**Theorem 1.2.4** (Prime number theorem)**.** *Let $\pi(n)$ denote the number of primes smaller than $n$. Then*

$$\pi(n) = \frac{n}{(1+o(1))\log n}.$$

Finally, in the course of proof of Proposition 4.4.5 we will use the following standard bound for the number of divisors of an integer.

**Proposition 1.2.5** (Divisor bound)**.** *Every integer $n$ has at most $n^{o(1)}$ divisors.*

## 1.3 Hölder-type and concentration inequalities

Apart from the $L^2$ norm already introduced in the previous section, we will also use $L^p$ norms for other values of $p$. Namely, for $f\colon G \to \mathbb{C}$ we define

$$\|f\|_{L^p(G)} = \left(\mathbb{E}_{x\in G}|f(x)|^p\right)^{1/p}.$$

Most of the time, we will use the simpler notation $\|f\|_p$ when it is clear from the context which group we are referring to.

*Hölder's inequality* is then

$$\|f_1 \dots f_k\|_p \leq \|f_1\|_{p_1} \dots \|f_k\|_{p_k},$$

whenever $1/p = 1/p_1 + \dots + 1/p_k$. The most important special case, and the one we will use most of the time, is when $p = 1$ and $p_1 = p_2 = 2$. In this case the inequality is known as the *Cauchy-Schwarz inequality*

$$|\langle f, g\rangle| \leq \|f\|_2 \|g\|_2.$$

Both of these inequalities are also valid when uniform measure is replaced with any other measure on $G$. Specifically, we can also apply these inequalities when working on the group $\widehat{G}$ with the usual choice of counting measure.

On few occasions we will need the so-called concentration inequalities, i.e. results which say that a random variable is most of the time close to its mean. Although concentration of measure is very rich and active area of research, we will only need following standard inequalities. Detailed proofs of these (and many more) inequalities can be found in [AS08].

Let $X$ be a random variable with finite mean. The first result is the well known inequality by Markov.

**Proposition 1.3.1** (Markov's inequality)**.** *For any $M > 0$ we have*

$$\mathbb{P}(|X| \geq M) \leq \frac{\mathbb{E}|X|}{M}. \tag{1.1}$$

A very common application of this inequality is in the setting where $M = 1$ and $X = 1_{A_1} + \dots + 1_{A_k}$ for some events $A_1, \dots, A_k$. In this case, bound given by Proposition 1.3.1 is also known as the union bound, since $X \geq 1$ if and only if the event $\bigcup_{i=1}^k A_i$ happened.

If we assume that $X$ has finite second moment, the following result, also known as the second moment inequality, follows easily by applying Markov's inequality to the variable $(X - \mathbb{E}X)^2$.

**Proposition 1.3.2** (Chebyshev's inequality)**.** *For any $M > 0$ we have*

$$\mathbb{P}(|X - \mathbb{E}X| \geq M) \leq \frac{\mathrm{Var}X}{M^2}.$$

However, if we assume that $X$ is a sum of independent bounded random variables, we can get an even stronger concentration by applying Markov's inequality to the random variable $e^{tX}$, for suitably chosen $t \in \mathbb{R}$.

**Proposition 1.3.3** (Hoeffding's inequality)**.** *Let $X_1, \ldots, X_n$ be independent bounded random variables with mean 0. Then for every $\lambda \geq 0$ we have*

$$\mathbb{P}\left(\left|\sum_i X_i\right| \geq \lambda \sqrt{\sum_i \|X_i\|_\infty^2}\right) \ll \exp(-\lambda^2/2). \tag{1.2}$$

In Chapter 3 we will need a more precise bound than the one given by Hoeffding's inequality in the special case when $X$ is a sum of independent Bernoulli random variables, and this is the content of the following proposition. Additionally, it also gives a corresponding lower bound that we will also need at some point.

**Proposition 1.3.4.** *Let $Z$ be a $B(n, \frac{1}{2})$ random variable and $0 < \epsilon < \frac{1}{2}$ fixed. If we define $h(x) = (\frac{1}{2} + x)\log(1 + 2x) + (\frac{1}{2} - x)\log(1 - 2x)$, then*

$$e^{-nh(\epsilon) - O(\log n)} \leq \mathbb{P}(Z \geq (\tfrac{1}{2} + \epsilon)n) \leq e^{-nh(\epsilon)}.$$

This proposition is somewhat less known than the previous three so we also provide a simple proof.

*Proof of Proposition 1.3.4.* Let $p = \frac{1}{2} + \epsilon$ and let $Y$ be a $B(n, p)$ random variable. We define $k_{\max}$ to be the value of $k$ for which $\mathbb{P}(Y = k)$ is maximal. Since $Y$ is supported on $n + 1$ values, we obviously have $\mathbb{P}(Y = k_{\max}) \geq \frac{1}{n+1}$. It is well-known that $k_{\max}$ is equal to either $\lceil pn \rceil$ or $\lceil pn \rceil - 1$. However, since

$$\binom{n}{\lceil pn \rceil - 1} \Big/ \binom{n}{\lceil pn \rceil} \leq n,$$

we can conclude that

$$\mathbb{P}(Y = \lceil pn \rceil) \geq \tfrac{1}{n(n+1)},$$

and hence

$$\begin{aligned}\binom{n}{\lceil pn\rceil} &\geq \tfrac{1}{n(n+1)} p^{-\lceil pn\rceil}(1-p)^{\lceil pn\rceil - n} \\ &= e^{-n(p\log p+(1-p)\log(1-p))-O(\log n)} \\ &= 2^n e^{-nh(\epsilon)-O(\log n)}.\end{aligned}$$

Dividing by $2^n$ and noticing that $\mathbb{P}(Z \geq (\frac{1}{2}+\epsilon)n) \geq \mathbb{P}(Z = \lceil(\frac{1}{2}+\epsilon)n\rceil)$ gives us the required lower bound.

For the upper bound, notice that for every fixed $k \geq pn$ we have

$$\mathbb{P}(Y=k) = \binom{n}{k} p^k (1-p)^{n-k} \geq \binom{n}{k} p^{pn}(1-p)^{n-pn} = \binom{n}{k} 2^{-n} e^{nh(\epsilon)}.$$

From this we conclude that

$$\sum_{k=\lceil pn\rceil}^{n} \binom{n}{k} \leq 2^n e^{-nh(\epsilon)} \sum_{k=\lceil pn\rceil}^{n} \mathbb{P}(Y=k) \leq 2^n e^{-nh(\epsilon)}.$$

The required upper bound follows easily from this. □

## 1.4 Notation and conventions

Here we set up some notation that will be used throughout the thesis. Some of the notation was already introduced implicitly in the previous sections, but we will repeat it here for the reader's convenience. Notation and conventions specific to each chapter will be introduced towards its start.

We will use the standard $O$-notation. To be concrete, for functions $f, g\colon \mathbb{N} \to \mathbb{R}$ we will write $f(n) = O(g(n))$ and $|f(n)| \ll g(n)$ if $|f(n)| \leq Cg(n)$ for some constant $C > 0$. More generally, we will write $f(n) = O_{r_1,\ldots,r_k}(g(n))$ and $|f(n)| \ll_{r_1,\ldots,r_k} g(n)$ if $|f(n)| \leq C(r_1,\ldots,r_k)g(n)$ for some constant $C(r_1,\ldots,r_k) > 0$ that, as the notation suggests, depends only on parameters $r_1,\ldots,r_k$.

We will write $f(x) = o_{r_1,\ldots,r_k;x\to x_0}(g(x))$ if $f(x)/g(x) \to 0$ as $x$ tends to $x_0$ and the rate of convergence depends only on parameters $r_1,\ldots,r_k$. On most occasions, the exact dependence on these parameters will not important and $x$ would tend to infinity. In situations like this, we will use the simpler expression $f(x) = o(g(x))$.

For a finite set $S$ and a function $f$, we will denote the average of $f$ on the set $S$

by $\mathbb{E}_{s\in S}f(s)$, i.e.

$$\mathbb{E}_{s\in S}f(s) = \frac{1}{|S|}\sum_{s\in S} f(s).$$

For a set $S$ we will denote its indicator function by $1_S$, so

$$1_S(x) = \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{if } x \notin S \end{cases}.$$

Finally, we will also use the standard expression $e(x) = e^{2\pi ix}$.

# Chapter 2

# One-point concentration of the clique and chromatic numbers of the random Cayley graph on $\mathbb{F}_2^n$

Green [Gre05a] showed that there exist constants $C_1, C_2 > 0$ such that the clique number $\omega$ of the Cayley graph on $\mathbb{F}_2^n$ generated by a random subset satisfies

$$\lim_{n\to\infty} \mathbb{P}(C_1 n \log n < \omega < C_2 n \log n) = 1.$$

In this chapter we find the best possible $C_1$ and $C_2$. Moreover, we prove that for $n$ in a set of density 1, the clique number is actually concentrated on a single value. As a simple consequence of these results, we also prove a one-point concentration result for the chromatic number, thus proving a $\mathbb{F}_2^n$ analogue of the famous conjecture by Bollobás [Bol04] and giving almost the complete answer to the question by Green [Gre15].

This chapter is based on the preprint [Mra15].

## 2.1 Introduction

The clique number of a graph is the size of its largest complete subgraph. It is a classical result of the theory of random graphs that the clique number in the Erdős-Rényi model $G(N, \frac{1}{2})$ is strongly concentrated around $2\log N$ (all the logarithms appearing in this chapter will be with base 2). Furthermore, Bollobás and Erdős [BE76] and Matula [Mat72] independently proved that the clique number is with high probability equal to one of two values, and that for most $N$ it is actually concentrated

on a single value.

There is also a standard more algebraic way of constructing a random graph. Given an abelian group $G$, one constructs a random graph with vertex set $G$ by taking a random subset $A \subset G$ so that each each element lies in $A$ with probability $\frac{1}{2}$, independently of other elements, and joining vertices $x \neq y$ if and only if $x + y \in A$. This graph is called *random Cayley sum graph*. (There is also a related notion of *random Cayley graph* where one takes $A$ to be a random *symmetric* set and joins $x \neq y$ if and only if $x - y \in A$. In $\mathbb{F}_2^n$ these are the same, and for other groups most of the results true for one concept are also true for the other.)

In [Gre05a], Green studied the clique number of the random Cayley sum graph on the groups $\mathbb{Z}/N\mathbb{Z}$ and $\mathbb{F}_2^n$. He proved that with high probability the clique number is less than $160 \log N$ in the $\mathbb{Z}/N\mathbb{Z}$ case and between $\frac{1}{4}n \log n$ and $308n \log n$ in the $\mathbb{F}_2^n$ case. Green and Morris [GM15] improved the former result by proving that for any $\epsilon > 0$ the clique number is with high probability at most $(2 + \epsilon) \log N$. The corresponding lower bound of $(2 - \epsilon) \log N$ is much easier to prove, see [Gre15].

In this chapter, we improve Green's result on $\mathbb{F}_2^n$. Let $\omega_n$ denote the corresponding clique number (we will omit subscript when it is clear from the context). The exact improvement is contained in the following theorem.

**Theorem 2.1.1.** *For every $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(\tfrac{1}{2}n \log n < \omega_n < (1 + \epsilon)n \log n) = 1.$$

Additionally, we also prove that the constants in front of $n \log n$ in the previous theorem are optimal, i.e. we prove the following theorem.

**Theorem 2.1.2.** *The constants $\frac{1}{2}$ and $1 + \epsilon$ in Theorem 2.1.1 are optimal, i.e. for every $\epsilon > 0$ there exist sequences $(n_i)$ and $(n'_j)$ such that*

$$\lim_{i\to\infty} \mathbb{P}(\tfrac{1}{2}n_i \log n_i < \omega_{n_i} < (\tfrac{1}{2} + \epsilon)n_i \log n_i) = 1,$$

*and*

$$\lim_{j\to\infty} \mathbb{P}(n'_j \log n'_j < \omega_{n'_j} < (1 + \epsilon)n'_j \log n'_j) = 1.$$

Moreover, we prove that for most $n$ the clique number is concentrated on a single value. The word *most* here is justified by the following corollary. For arbitrary $A \subset \mathbb{N}$ we define its density as $d(A) = \lim_{n\to\infty} |A \cap \{1, \dots, n\}|/n$ (if the limit exists). If the limit does not exist, we will say that the density of $A$ is not defined.

**Theorem 2.1.3.** *There exists a sequence* $(n_l)$ *of density* 1 *such that*

$$\lim_{l\to\infty} \mathbb{P}(\omega_{n_l} = 2^{\lfloor \log n_l + \log\log n_l \rfloor}) = 1.$$

Theorem 2.1.3 is a $\mathbb{F}_2^n$ analogue of the celebrated two-point concentration result in the Erdős-Rényi model stated above. We note that this is the first result of that type for random Cayley sum graphs. In Section 2.9 we will show that the analogous result is also true for the chromatic number $\chi$, that is the smallest possible number of parts in a partition with the property that no edge in the graph connects vertices that belong to the same part of the partition. The proof will turn out to be a straightforward consequence of the result about the clique number.

**Theorem 2.1.4.** *There exists a sequence* $(n_l)$ *of density* 1 *such that*

$$\lim_{l\to\infty} \mathbb{P}(\chi_{n_l} = 2^{n_l - \lfloor \log n_l + \log\log n_l \rfloor}) = 1.$$

One can consider this result as the $\mathbb{F}_2^n$ analogue of the famous question by Bollobás [Bol04] claiming that the chromatic number in the Erdős-Rényi $G(N, \frac{1}{2})$ model is concentrated on $O(1)$ values which is still wide open. However, some very interesting results have been obtained in models $G(n,p)$ for $p$ somewhat small, see e.g. the work of Coja-Oghlan, Panagiotou and Steger [COPS08] where $p \leq n^{-3/4-\delta}$ (for any $\delta > 0$), and the work of Achlioptas and Naor [AN05] where $p = d/n$ (for arbitrary constant $d > 0$).

Alon [Alo13] studied the chromatic number of random Cayley graphs in various groups with the set of generators ranging from very small (constant) to very large (positive density). Additionally, Green [Gre15] focused on the case of the cyclic group with the set of generators equal to a uniform random subset. In the same paper Green asked what is, asymptotically, the chromatic number of the random Cayley graph of $\mathbb{F}_2^n$, again with the uniform random subset as the set of generators. The previous theorem gives a very precise answer for $n$ in a set of density 1. In the same way we will prove Theorem 2.1.4, it is also possible to use other results for the clique number (namely, Theorem 2.1.1 and Theorem 2.1.2) to get the corresponding information about the chromatic number for every $n$, but unfortunately we were unable to give the complete answer to Green's question.

We will mostly concentrate on the upper bound in Theorem 2.1.1, since the lower bound is essentially already contained in [Gre05a]; we provide the details for the latter in Section 2.6. Notice that the set $X \subset \mathbb{F}_2^n$ spans a clique in the random

Cayley graph generated by the set $A$ if and only if $X\widehat{+}X \subset A$, where $X\widehat{+}X = \{x_1 + x_2 \colon x_1, x_2 \in X, x_1 \neq x_2\}$. This observation, together with Markov's inequality, was used in [Gre05a] to derive the inequality

$$\mathbb{P}(\omega \geq k) \leq \sum_{l \geq k-1} |S_k^l| 2^{-l},$$

where $S_k^l$ is the family of sets $X \subset \mathbb{F}_2^n$ satisfying $|X| = k$ and $|X\widehat{+}X| = l$. If one has a good upper bound for $|S_k^l|$, the right hand side is $o(1)$ and $k$ is an upper bound for the clique number with probability tending to 1. We will use this approach only to deal with sets $X$ of large doubling ($|X\widehat{+}X| \geq 10|X|$); we give details in Section 2.3. The main downside of this approach is that it fails to deal with any dependencies between events $\{X\widehat{+}X \subset A\}$ for different sets $X$.

We have the obvious inequality

$$\mathbb{P}(\omega \geq k) \leq \mathbb{P}\left( \bigcup_{|X|=k, |X\widehat{+}X|<10k} \{X\widehat{+}X \subset A\} \right) + \mathbb{P}\left( \bigcup_{|X|=k, |X\widehat{+}X|\geq 10k} \{X\widehat{+}X \subset A\} \right). \tag{2.1}$$

We denote the two terms on the right hand side by $P_{\text{small}}(n,k)$ and $P_{\text{large}}(n,k)$, respectively. In Section 2.4 we introduce a new approach which deals with some of the dependencies mentioned above, and thus effectively bounds $P_{\text{small}}(n,k)$. We prove the optimality result in Section 2.7 and the one-point concentration of the clique number in Section 2.8.

The author has been recently made aware of an independent, yet unpublished, work by Penman and Wells in which they prove Theorem 2.1.1 by somewhat different methods, using Green's probabilistic argument [Gre05a] to deal with sets of small doubling. However, such approach is not sufficient to prove one-point concentration results contained in Theorem 2.1.3 and Theorem 2.1.4.

*Acknowledgements.* I would like to thank Ben Green for communicating the problem and helpful suggestions and to Sean Eberhard and Freddie Manners for fruitful discussions.

## 2.2 Notation

Although most of the notation was implicitly used in the introduction, we give it here for the reader's convenience. In this chapter $N$ will always denote the size of the group $\mathbb{F}_2^n$, so $N = 2^n$. For sets $X, Y \subset \mathbb{F}_2^n$ we will denote their sumset by

$X+Y$, so $X+Y=\{x+y\colon x\in X, y\in Y\}$. We will also consider their *restriced sumset* $X\widehat{+}Y=\{x+y\colon x\in X, y\in Y, x\neq y\}$. Notice that in the $\mathbb{F}_2^n$ setting we always have $X\widehat{+}Y=(X+Y)\setminus\{0\}$. For a set $A\subset\mathbb{N}$, in addition to the notion of its density defined in the introduction, we also define its lower density as $\underline{d}(A)=\liminf_{n\to\infty}|A\cap\{1,\dots,n\}|/n$. As usual, $\{x\}$ would denote the fractional part of $x$, i.e. $\{x\}=x-\lfloor x\rfloor$.

## 2.3 Sets of large doubling

Our aim in this section is to prove that $P_{\text{large}}(n,k)=o(1)$ for suitable choice of $k$ (we remind the reader of the notation $P_{\text{large}}(n,k)$ introduced at the end of Section 2.1 – it denotes the last term in (2.1)). This is basically a consequence of results obtained in [Gre05a] joined with a new result by Even-Zohar [EZ12].

First of all, note that by the union bound, we have

$$\mathbb{P}\left(\bigcup_{|X|=k,|X\widehat{+}X|\geq 10k}\{X\widehat{+}X\subset A\}\right)\leq\sum_{l\geq 10k}|S_k^l|2^{-l}.$$

Let $X\subset\mathbb{F}_2^n$, $Y\subset\mathbb{F}_2^m$ be fixed sets. We will say that a bijection $f\colon X\to Y$ is a Freĭman isomorphism if for every $x_1,x_2,x_3,x_4\in X$

$$x_1+x_2=x_3+x_4 \text{ if and only if } f(x_1)+f(x_2)=f(x_3)+f(x_4). \tag{2.2}$$

In this case we will say that sets $X$ and $Y$ are Freĭman isomorphic. Let $r$ be the largest integer such that $X$ is Freĭman isomorphic to a subset of $\mathbb{F}_2^r$ which is not contained in a proper affine subspace. This number $r$ is known as the Freĭman dimension over $\mathbb{F}_2$ of the set $X$; we will denote it by $r(X)$.

Proposition 26 from [Gre05a] gives us the following bound for $|S_k^l|$.

**Proposition 2.3.1.** *Let $r_{k,l}=\max_{X\in S_k^l}r(X)$. We have*

$$|S_k^l|\leq N^{r_{k,l}+1}k^{4k}.$$

*For $l\leq k^{31/30}$ there is a better bound*

$$|S_k^l|\leq N^{r_{k,l}+1}\left(\frac{el}{k}\right)^k\exp(k^{31/32}).$$

Although we will use this proposition as a black box we still say few words about the proof for the benefit of intuition. Roughly speaking, the proof naturally breaks into two stages. The first one is about finding an upper bound for the number of sets in $S_k^l$ which are not mutually Freĭman isomorphic. For the first bound, Green finds an upper bound for the number of subsets of $\mathbb{F}_2^n$ of size $k$ which are not mutually Freĭman isomorphic (hence he doesn't use the fact that they must also lie in $S_k^l$). For the second bound he gets a better bound by also requiring that these sets are in $S_k^l$. The second stage is about finding the number of Freĭman isomorphic copies of a fixed set in $\mathbb{F}_2^n$ (the factor $N^{r_{k,l}+1}$ in both bounds). It is not too surprising that small doubling implies some upper bound on the number of these copies. Indeed, small doubling reflects some underlying structure present in the set and this in turn reduces the number of possibilities where we can send some of the points in our set.

We will bound $r_{k,l}$ from Proposition 2.3.1 using the following version of Freĭman's theorem proved by Even-Zohar [EZ12].

**Theorem 2.3.2.** *Suppose $X \subset \mathbb{F}_2^n$ has cardinality $k$ and that $|X+X| \leq Kk$ for some $K \geq 1$. Then there is a subgroup $H$ containing $X$ such that*

$$|H| \leq 4^K k.$$

**Corollary 2.3.3.** *Suppose $X \subset \mathbb{F}_2^n$, $|X| = k$ and $|X + X| = l$. Then*

$$r(X) \leq \log k + 2l/k.$$

*Proof.* Let $f\colon X \to \mathbb{F}_2^r$ be an injection which satisfies (2.2). Notice that $|f(X) + f(X)| = |X + X| \leq l$. By the previous theorem we conclude that $f(X)$ is contained in a subgroup of size at most $4^K k$, for $K = l/k$. All subgroups of $\mathbb{F}_2^n$ are also subspaces and the claim easily follows. □

Let $k \geq \frac{1}{2} n \log n$. Notice that $|X + X| = |X \widehat{+} X| + 1$ for any set $X$. Using the first bound in Proposition 2.3.1 and the bound for the Freĭman dimension given in Corollary 2.3.3 we get

$$\begin{aligned}\sum_{l \geq k^{31/30}} |S_k^l| 2^{-l} &\leq \sum_{l \geq k^{31/30}} N^{\log k + 2(l+1)/k + 1} k^{4k} 2^{-l} \leq \sum_{l \geq k^{31/30}} 2^{l(-1 + \frac{n \log k}{l} + \frac{3n}{k} + \frac{4k \log k}{l})} \\ &\leq \sum_{l \geq k^{31/30}} 2^{l(-1+o(1))} = o(1). \end{aligned} \tag{2.3}$$

Additionally, using the second bound in Proposition 2.3.1, we get

$$\begin{aligned}\sum_{10k\le l<k^{31/30}} |S_k^l|2^{-l} &\le \sum_{10k\le l<k^{31/30}} N^{\log k+2(l+1)/k+1}\left(\frac{el}{k}\right)^k \exp(k^{31/32})2^{-l}\\ &\le \sum_{10k\le l<k^{31/30}} 2^{l(-1+\frac{n\log k}{l}+\frac{3n}{k}+\frac{k}{l}\log\left(\frac{el}{k}\right)+\frac{2k^{31/32}}{l})}\\ &\le \sum_{10k\le l<k^{31/30}} 2^{l(-1+\frac{1}{5}+\frac{k}{l}\log\left(\frac{el}{k}\right)+o(1))} = o(1). \qquad (2.4)\end{aligned}$$

Combining (2.3) and (2.4) we conclude that $P_{\text{large}}(n,k) = o(1)$ for any $k \ge \frac{1}{2}n\log n$.

## 2.4 Sets of small doubling

In this section we deal with sets of small doubling. The strategy is the following. First of all, notice that every set $X$ of size $k$ satisfying $|X\widehat{+}X| < 10k$ is by Theorem 2.3.2 contained in a subspace of size $O(k)$. Now fix a subspace $W$ of this size and consider all sets of size $k$ contained in it. The crucial claim we will prove here is that there exists a small family $\mathcal{F}$ of large sets such that for each $X$, the set $X\widehat{+}X$ contains a set from this family. This will enable us to exploit dependencies between events $\{X\widehat{+}X \subset A\}$ for various sets $X$. More precisely, we will use the following obvious inequality

$$\begin{aligned}\mathbb{P}\left(\bigcup_{|X|=k,X\subset W}\{X\widehat{+}X\subset A\}\right) &= \mathbb{P}\left(\bigcup_{F\in\mathcal{F}}\bigcup_{|X|=k,X\subset W,F\subset X\widehat{+}X}\{X\widehat{+}X\subset A\}\right)\\ &\le \mathbb{P}\left(\bigcup_{F\in\mathcal{F}}\{F\subset A\}\right)\\ &\le \sum_{F\in\mathcal{F}}\mathbb{P}\left(F\subset A\right). \qquad (2.5)\end{aligned}$$

Obviously, the crucial step in this approach is to find the family $\mathcal{F}$ with these properties and this is what we do in the following theorem.

**Theorem 2.4.1.** *Fix $\delta > 0$. For large enough $n \ge n_0(\delta)$ the following holds. Let $k \ge \delta N$ and $k'$ be the power of 2 satisfying $k' < k \le 2k'$. There is a family $\mathcal{F}$ of subsets of $\mathbb{F}_2^n$ with the following properties*

*(i)* $|\mathcal{F}| \le 2^{\delta N}$,

*(ii) every $F \in \mathcal{F}$ has at least $(2-\delta)k'$ elements,*

*(iii) for every $X \subset \mathbb{F}_2^n$ with $k$ elements, there is $F \in \mathcal{F}$ such that $F \subset X + X$.*

A careful reader might object that $k \geq \delta N$ is not the setting we have been working with until now, since $k$ of our interest is of about logarithmic size of the size of the ambient group. However, in the application of this theorem, we will consider our $k$-element subset as being contained in some subgroup of size $O(k)$, which we can do by Theorem 2.3.2. So $N$ here actually represents the size of this subgroup.

We would also like to stress the importance of the term $(2-\delta)k'$ in property (ii). If one replaces this term with $(1-\delta)k$, the proof is much easier and such statement is sufficient to prove Theorem 2.1.1, but not the one-point concentration results we consider to be the main contribution of this chapter.

There are couple of ingredients we need in the proof of this theorem. The first one is a regularity lemma in $\mathbb{F}_2^n$. The variant we will use was first proven by Green in [Gre05b].

Let $V \leqslant \mathbb{F}_2^n$. We give $V$ the uniform probability measure and $V^*$ the counting measure. Hence, for any $f\colon V \to \mathbb{R}$ and $r \in V^*$ we define corresponding Fourier coefficient by

$$\widehat{f}(r) = \mathbb{E}_{v \in V} f(v)(-1)^{\langle v, r\rangle}.$$

Fix a subset $X \subset \mathbb{F}_2^n$. For $i \in \mathbb{F}_2^n$ we consider function $1_{(X-i)\cap V}\colon V \to \mathbb{R}$. We will say that $i$ is an $\epsilon$-regular value for $X$ with respect to $V$ if

$$\sup_{r \neq 0} |\widehat{1_{(X-i)\cap V}}(r)| \leq \epsilon.$$

Notice that $i$ is $\epsilon$-regular value if and only if all the elements in $i + V$ are $\epsilon$-regular values. Indeed, this follows easily by noticing that

$$|\widehat{1_{(X-i)\cap V}}(r)| = |\widehat{1_{(X-(i+v))\cap V}}(r)|$$

for all $v \in V$. Thus, we may also talk about $\epsilon$-regular cosets of $V$. Additionally, we will say that $X$ is $\epsilon$-regular with respect to $V$ if there are less than $\epsilon N$ choices of $i$ which are not $\epsilon$-regular values for $X$ with respect to $V$. Intuitively, $X$ is $\epsilon$-regular (for some small $\epsilon$) if $(X-i)\cap V$ looks like a random subset of $V$ for most choices of $i$.

The following is Theorem 2.1 from [Gre05b], which says that every set is $\epsilon$-regular with respect to some large subspace.

**Theorem 2.4.2** (Regularity lemma in $\mathbb{F}_2^n$)**.** *Let $\epsilon \in (0, \frac{1}{2})$ and $X \subset \mathbb{F}_2^n$. Then there is a subspace $V \leqslant \mathbb{F}_2^n$ of codimension $O_\epsilon(1)$ such that $X$ is $\epsilon$-regular with respect to $V$.*

This result is the main ingredient in the proof of the following proposition which is what we will ultimately need for the proof of Theorem 2.4.1.

**Proposition 2.4.3.** *Let $X \subset \mathbb{F}_2^n$, $\epsilon > 0$ and $|X| \geq \epsilon N$. There is a subspace $V \leqslant \mathbb{F}_2^n$ of codimension $O_\epsilon(1)$ and $I \subset \mathbb{F}_2^n/V$ such that*

$$|X \cap \bigcup_{i+V \in I} (i+V)| \geq (1-\epsilon)|X|$$

*and*

$$|(X \cap (i+V)) + (X \cap (i'+V))| \geq (1-\epsilon)|V|$$

*for all $i+V, i'+V \in I$.*

*Proof.* It is enough to prove the claim only for small enough $\epsilon$. Apply the regularity lemma with parameter $\epsilon^5$. We get at most $\epsilon^5 N$ elements which are not $\epsilon^5$-regular values for $X$. By the comment above, this is equivalent to stating that at most $\epsilon^5|\mathbb{F}_2^n/V|$ cosets of $V$ are not $\epsilon^5$-regular. Let $I$ be the set of all translates $i+V$ of $V$ such that $|X \cap (i+V)| \geq \epsilon^3|V|$ and $i$ is $\epsilon^5$-regular value for $X$ with respect to $V$. It is easy to see that

$$|X \cap \bigcup_{i+V \in I} (i+V)| \geq |X| - \epsilon^3 N - \epsilon^5|\mathbb{F}_2^n/V||V| \geq (1-\epsilon)|X|.$$

Fix $i+V, i'+V \in I$ and let $E = (X-i) \cap V$ and $F = (X-i') \cap V$. To prove the remaining property, it is enough to prove that $|E+F| \geq (1-\epsilon)|V|$.

Let $S = V \setminus (E+F)$. By Plancherel's formula

$$0 = \mathbb{E}_{v \in V}(1_E * 1_F)(v)1_S(v) = |E||F||S|/|V|^3 + \sum_{r \in V^* \setminus \{0\}} \widehat{1_E}(r)\widehat{1_F}(r)\widehat{1_S}(r),$$

and hence

$$\begin{aligned}|E||F||S|/|V|^3 &\leq \epsilon^5 \sum_{r\in V^*\setminus\{0\}} |\widehat{1_E}(r)|^{1/2}|\widehat{1_F}(r)|^{1/2}|\widehat{1_S}(r)| \\ &\leq \epsilon^5 \left(\sum_{r\in V^*} |\widehat{1_E}(r)|^2\right)^{1/4} \left(\sum_{r\in V^*} |\widehat{1_F}(r)|^2\right)^{1/4} \left(\sum_{r\in V^*} |\widehat{1_S}(r)|^2\right)^{1/2} \\ &\leq \epsilon^5 |E|^{1/4}|F|^{1/4}|S|^{1/2}/|V|.\end{aligned}$$

The conclusion is that $|S| \leq \epsilon^{10}|V|^4|E|^{-3/2}|F|^{-3/2}$, which together with hypothesis on the sizes of sets $E$ and $F$ gives us what we want. □

The second major ingredient we will need in the proof of Theorem 2.4.1 is the following classical theorem by Kneser [Kne53].

**Theorem 2.4.4** (Kneser)**.** *Let $G$ be an abelian group. For any set $S \subset G$ define* $\mathrm{Sym}(S) = \{g \in G\colon g + S = S\}$*. Then for any two sets $X, Y \subset G$ we have*

$$|X + Y| \geq |X| + |Y| - |\mathrm{Sym}(X + Y)|.$$

Readers interested in the proof could consult the book by Tao and Vu [TV06]. We use this theorem to derive the following simple consequence.

**Proposition 2.4.5.** *Let $X, Y \subset \mathbb{F}_2^n$ and suppose that $|Y| > m$, where $m$ is some power of 2. Then*

$$|X + Y| \geq \min(|X| + m, 2m).$$

*Proof.* Let $W = \mathrm{Sym}(X + Y)$. Obviously, $W$ is a subgroup, hence $|W|$ is a power of 2. Notice that for any nonempty set $S$ we must have $|S| \geq |\mathrm{Sym}(S)|$. Indeed, for fixed $s \in S$ and $g \in \mathrm{Sym}(S)$, we must have that $t = g + s \in S$. However, $s$ and $t$ uniquely determine $g$, and since there are $|S|$ options for $t$, we are done. Thus, if $|W| \geq 2m$, then the proof is finished.

Suppose now that $|W| \leq m$ and let $\pi\colon \mathbb{F}_2^n \to \mathbb{F}_2^n/W$ be the canonical projection. Notice that $|\pi(X)| \geq |X|/|W|$ and $|\pi(Y)| \geq m/|W| + 1$. Kneser's theorem now gives us that

$$|\pi(X) + \pi(Y)| \geq |\pi(X)| + |\pi(Y)| - 1.$$

It is easy to see that $X + Y$ is equal to some union of cosets of $W$ and hence

$$|X + Y| = |\pi(X + Y)||W| = |\pi(X) + \pi(Y)||W|.$$

Using this and the previous three inequalities we get what we wanted. □

We proceed to the proof of the main theorem in this section.

*Proof of Theorem 2.4.1.* In this proof we will use Proposition 2.4.3 and hence will also use the notation introduced there. Obviously, it is enough to prove the claim only for small enough $\delta$.

Let $\eta = \delta^7$, and $d$ be the maximal codimension given by the regularity lemma when applied with parameter $\eta$. We will say that a set is *almost-coset of* $V$ if it can be obtained by removing at most $\delta^3|V|$ elements from a coset of $V$. Our family $\mathcal{F}$ will consist of all subsets of $\mathbb{F}_2^n$, with at least $(2-\delta)k'$ elements, that can be obtained as a union of almost-cosets of some subspace $V$ of codimension at most $d$. Family $\mathcal{F}$ by its definition satisfies property (ii) from the statement of the theorem. Let's prove that it also satisfies property (i), that is $|\mathcal{F}| \leq 2^{\delta N}$ (for $n$ large enough).

Fix codimension $d' \leq d$ and let $D' = 2^{d'}$. There are at most $N^{d'}$ ways to choose subspace $V$ of codimension $d'$, and $2^{D'}$ ways to choose some of its cosets from which we will remove some elements to obtain a set in $\mathcal{F}$. So the contribution from these two factors is bounded by $2^{o(N)}$. Once we have chosen cosets, we need to bound the number of ways to remove at most $\delta^3 N/D'$ elements from each of them. We bound this by

$$\left(\sum_{s=0}^{\delta^3 N/D'} \binom{N/D'}{s}\right)^{D'} \leq N^{D'} \binom{N/D'}{\delta^3 N/D'}^{D'} \leq N^{D'} \left(e/\delta^3\right)^{\delta^3 N} \leq 2^{o(N)} 2^{\delta^2 N}.$$

Multiplying these bounds, and taking the union over all the possible choices for $d'$ gives the required bound.

Let's prove now the remaining (and most important) property (iii). Take any $X \subset \mathbb{F}_2^n$, $|X| = k$, and apply Proposition 2.4.3 with parameter $\eta$. Let $X_I = X \cap \bigcup_{i+V \in I}(i+V)$.

Let $m$ be the largest power of 2 such that $|I| > m/2$. Then obviously $|I| \leq m$ and by Proposition 2.4.5 also $m \leq |I+I|$.

Let $\alpha$ be the density of $X$ on the union of translates from $I$. Hence

$$(1-\eta)k' < (1-\eta)|X| \leq |X_I| = \alpha|I||V| \leq \alpha m|V|.$$

Notice that if $|X_I + X_I| \geq (2-\delta)k'$, then $X_I + X_I \in \mathcal{F}$ and we are done. If not, then

$$(2-\delta)k' > |X_I + X_I| \geq (1-\eta)|I+I||V| \geq (1-\eta)m|V|.$$

So

$$\frac{1-\eta}{\alpha} < \frac{m|V|}{k'} < \frac{2-\delta}{1-\eta} < 2.$$

However, since the second term is a power of 2, we conclude that $\alpha \geq 1-\eta$.

Let $J$ be the set of all those $j+V \in I$ for which $|X \cap (j+V)| \geq (1-\sqrt{\eta})|V|$. By bounding the size of $X \cap \bigcup_{i+V\in I}(i+V)$ from below and from above we get

$$(1-\eta)|I||V| \leq |J||V| + (1-\sqrt{\eta})(|I|-|J|)|V|,$$

and hence $|J| \geq (1-\sqrt{\eta})|I|$.

If we set $X_J = X \cap \bigcup_{j+V\in J}(j+V)$, then

$$|X_J| \geq |X_I| - \sqrt{\eta}|I||V| \geq (1-\eta)|X| - \sqrt{\eta}N \geq (1-\eta-\sqrt{\eta}/\delta)|X| > (1-\delta)k'.$$

Proposition 2.4.5 now shows that $|X_J + X| \geq (2-\delta)k'$. Given that $X_J$ is a union of almost-cosets of $V$, the same is true for $X_J + X$. Hence, $X_J + X \in \mathcal{F}$ and we are again done. □

By applying Theorem 2.4.1 for, say, $\delta/2$ and removing the element 0 from all sets in $\mathcal{F}$, we get the following corollary.

**Corollary 2.4.6.** *Theorem 2.4.1 is also true if we replace sumsets with restricted sumsets.*

## 2.5 Proof of the upper bound in Theorem 2.1.1

We can now easily finish the proof of the upper bound in Theorem 2.1.1. Let $k = \lfloor (1+\epsilon)n\log n \rfloor$ and $\delta > 0$ to be specified later. As we said at the beginning of this section, any $X$ such that $|X| = k$ and $|X+X| < 10|X|$ is contained in some subspace of size $O(k)$. Since there are at most $N^{\log O(k)}$ such subspaces, applying Corollary 2.4.6 with parameter $\delta$ and from (2.5) we have

$$P_{\text{small}}(n,k) \leq N^{\log O(k)} 2^{\delta O(k)} 2^{-(2-\delta)k'}. \tag{2.6}$$

Since $k \leq 2k'$ we get

$$\begin{aligned} P_{\text{small}}(n,k) &\leq N^{\log O(k)} 2^{\delta O(k)} 2^{-(1-\delta/2)k} \\ &\leq 2^{n\log n(1+\delta O(1)-(1-\delta/2)(1+\epsilon)+o(1))} \end{aligned}$$

and this is $o(1)$ if we choose $\delta$ sufficiently small, which, joined with the appropriate result for $P_{\text{large}}(n,k)$ from Section 2.3 proves the upper bound.

## 2.6 Proof of the lower bound in Theorem 2.1.1

In this section we prove the lower bound stated in Theorem 2.1.1. As we mentioned in the introduction, this is basically already contained in Green's initial paper [Gre05a], so here we just sketch his argument and do a more careful calculation.

Let $m = \lfloor \log n + \log\log n \rfloor$ and let $M$ be the number of $m$-dimensional subspaces $H$ of $\mathbb{F}_2^n$ satisfying $H \setminus \{0\} \subseteq A$. In [Gre05a] it was proven that

$$\mathbb{E}M \geq 2^{nm-m^2-2^m} \quad \text{and} \quad \text{Var}(M) \leq 2\sum_{l=1}^{m} 2^{2mn-2^{m+1}+2^l-nl}.$$

Notice that $2^l - nl \leq 2 - n$ for every $l = 1, \ldots, m$. Indeed, the crucial point is when $l = m$; in that case we have

$$\begin{aligned} 2^m - n(m-1) - 2 &\leq 2^{\log n + \log\log n} - n(\log n + \log\log n - 2) - 2 \\ &\leq -n\log\log n + 2n - 2 \leq 0. \end{aligned} \tag{2.7}$$

Second moment method now gives

$$\mathbb{P}(M = 0) \leq \frac{\text{Var}(M)}{(\mathbb{E}M)^2} \leq 8m2^{2m^2-n} = o(1).$$

Notice now that whenever $H \setminus \{0\} \subset A$ there is a clique of size at least $|H|$, since $H \setminus \{0\} = H \widehat{+} H$, so we get the lower bound

$$\mathbb{P}(\omega > \tfrac{1}{2}n\log n) \geq \mathbb{P}(\omega \geq 2^{\lfloor \log n + \log\log n \rfloor}) = 1 - o(1). \tag{2.8}$$

## 2.7 Optimality

Finally, we prove that the results obtained in the previous two sections are optimal (in terms of constants in front of $n\log n$).

*Proof of Theorem 2.1.2.* First of all, let's prove that for every $\epsilon > 0$, there is a sequence $(n_i)$ for which $k_i = \lfloor (\frac{1}{2} + \epsilon) n_i \log n_i \rfloor$ is the upper bound for the clique number. We already know from Section 2.3 that $P_{\text{large}}(n_i, k_i) = o(1)$.

Let $m_i = \left\lfloor \frac{1}{1+\epsilon} 2^i \right\rfloor$ and $n_i = 2^{m_i}$. Notice that for such $n_i$ we have

$$\left(1 + \frac{\epsilon}{1+\epsilon} - (\tfrac{1}{2}+\epsilon)2^{1-i}\right) 2^{m_i+i-1} - 1 \leq k_i \leq \left(1 + \frac{\epsilon}{1+\epsilon}\right) 2^{m_i+i-1}.$$

We remind the reader that we defined $k'$ as the power of 2 satisfying $k' < k \leq 2k'$. Thus, for $k_i$, by the previous two inequalities, we have $k_i' = 2^{m_i+i-1}$. Moreover, by the second inequality we have $(\frac{1}{2}+\epsilon)k_i'/k_i \geq \frac{1}{2} + \frac{\epsilon}{2}$. By the bound (2.6) for $P_{\text{small}}$ we now have

$$\begin{aligned} P_{\text{small}}(n_i, k_i) &\leq 2^{n_i \log n_i (1+\delta O(1)+o(1)) - (2-\delta)k_i'} \leq 2^{n_i \log n_i (1+\delta O(1)+o(1) - (2-\delta)(\frac{1}{2}+\epsilon)k_i'/k_i)} \\ &\leq 2^{n_i \log n_i (1+\delta O(1)+o(1) - (2-\delta)(\frac{1}{2}+\frac{\epsilon}{2}))} \end{aligned}$$

and this is again $o(1)$ if we choose $\delta$ sufficiently small. Hence

$$\lim_{i\to\infty} \mathbb{P}(\omega \leq (\tfrac{1}{2}+\epsilon) n_i \log n_i) = 1$$

and the lower bound obtained in the previous section is optimal.

To prove that the upper bound from Section 2.4 is also optimal, first notice that the same calculations we did in Section 2.6 would, for some $n$, go through for $m = \lfloor \log n + \log\log n \rfloor + 1$. Indeed, it is easy to see that $2^{1-\{\log x\}} - 1 \leq \frac{4}{x} \leq \frac{\log x}{2x}$ for all large enough integers $x$ of the form $x = 2^s - 1$, so if we set $n_j' = 2^{2^j - 1}$ and $m_j' = \lfloor \log n_j' + \log\log n_j' \rfloor + 1$ we get that (2.7) is now

$$\begin{aligned} 2^{m_j'} - n_j'(m_j' - 1) - 2 &< (2^{1-\{\log\log n_j'\}} - 1) n_j' \log n_j' - n_j' \log\log n_j' + n_j' \\ &\leq \tfrac{\log\log n_j'}{2\log n_j'} \cdot n_j' \log n_j' - n_j' \log\log n_j' + n_j' \\ &= -\tfrac{1}{2} n_j' \log\log n_j' + n_j' < 0. \end{aligned}$$

Proceeding as before, this gives

$$\mathbb{P}(\omega > n_j' \log n_j') \geq \mathbb{P}(\omega \geq 2^{\lfloor \log n_j' + \log\log n_j' \rfloor + 1}) = 1 - o(1). \tag{2.9}$$

□

## 2.8 One-point concentration of the clique number

In this section we prove that we actually have one-point concentration for most $n$, namely Theorem 2.1.3. This will be an easy consequence of the following proposition.

**Proposition 2.8.1.** *Let $\epsilon > 0$ and $(n_l)$ be a sequence satisfying $\{\log n_l + \log\log n_l\} < 1 - \epsilon$. Then*

$$\lim_{l\to\infty} \mathbb{P}(\omega_{n_l} = 2^{\lfloor \log n_l + \log\log n_l \rfloor}) = 1.$$

*Proof.* Let $\epsilon > 0$ and $(n_l)$ be a sequence as in the statement of the proposition. To prove the upper bound, we set $k_l = 2^{\lfloor \log n_l + \log\log n_l \rfloor} + 1$. As before, by the results of Section 2.3 we already have $P_{\text{large}}(n_l, k_l) = o(1)$. On the other hand, after applying Corollary 2.4.6 with parameter $\delta$ and using the bound (2.6) we now get

$$P_{\text{small}}(n_l, k_l) \leq 2^{n_l \log n_l (1 + o(1) + \delta O(1) - (2-\delta)2^{-\{\log n_l + \log\log n_l\}})}.$$

Choosing $\delta$ small enough gives $P_{\text{small}}(n_l, k_l) = o(1)$ and this proves that with high probability the clique number is at most $2^{\lfloor \log n_l + \log\log n_l \rfloor}$. Additionally, from (2.8) we see that the corresponding lower bound is also true, which finishes the proof. □

We can now move to the proof of Theorem 2.1.3.

*Proof of Theorem 2.1.3.* First we prove that for any $\epsilon > 0$ there exists a set $T_\epsilon$ such that $\underline{d}(T_\epsilon) > 1 - \epsilon$ and

$$\lim_{n\to\infty, n\in T_\epsilon} \mathbb{P}(\omega = 2^{\lfloor \log n + \log\log n \rfloor}) = 1.$$

Obviously, it is enough to prove this statement when $\epsilon$ is small enough. By Proposition 2.8.1, it is enough to prove that the set of integers satisfying $\{\log n + \log\log n\} < 1 - \epsilon/16$ has lower density at least $1 - \epsilon$.

To prove this, it is enough to prove that for any large enough $m$ (in terms of $\epsilon$), the number of values of $n$ in the range $2^m \leq n < 2^{m+1}$ for which $\{\log n + \log\log n\} \geq 1 - \epsilon/16$ is at most $\epsilon 2^{m-1}$. Notice that for fixed $m$, the quantity $\log\log n$ varies within an interval of size at most $2/m$. Hence, there exists an interval $I_m$ in $\mathbb{R}/\mathbb{Z}$ of size at most $\epsilon/16 + 2/m < \epsilon/8$ (when $m$ is large enough), such that whenever $\{\log n + \log\log n\} \geq 1 - \epsilon/16$ we must also have that $\{\log n\} \in I_m$.

If $I_m = [t, t + \epsilon/8]$ is contained in $[0, 1]$ (i.e. doesn't wrap around) then the number of values of $n$ between $2^m$ and $2^{m+1}$ for which $\{\log n\} \in I_m$ is at most (for $\epsilon$ sufficiently small)

$$2^{m+t+\epsilon/8} - 2^{m+t} \leq 2^{m+1}(2^{\epsilon/8} - 1) < \epsilon 2^{m-2}$$

If, on the other hand, $I_m$ wraps around (i.e. is of the form $[t,1]\cup[0,s]$), then we can apply the previous bound to both of these smaller intervals and the claim follows.

Let $S_\epsilon = \{n \in T_\epsilon \colon \mathbb{P}(\omega = 2^{\lfloor \log n + \log\log n\rfloor}) > 1-\epsilon\}$. Obviously, $\underline{d}(S_\epsilon) = \underline{d}(T_\epsilon) > 1-\epsilon$. It is now easy to see that the sequence $\bigcup_m S_{1/m}$ satisfies all the requirements needed in the statement. □

Unfortunately, we cannot omit the condition that $n$ goes over a set of density 1 and let $n$ go over all the integers. Indeed by (2.9) there exists a sequence $(n_j)$ such that

$$\lim_{j\to\infty} \mathbb{P}(\omega \geq 2^{\lfloor \log n_j + \log\log n_j\rfloor + 1}) = 1.$$

## 2.9 One-point concentration of the chromatic number

In this final section we prove the one-point concentration result for the chromatic number, that is Theorem 2.1.4. As we will see, this is a very easy consequence of the corresponding result for the clique number – Theorem 2.1.3 and its proof. We note that the argument below first appeared in [Alo13].

*Proof of Theorem 2.1.4.* First of all, notice that, by symmetry, all the results in previous sections would also hold if we considered the independence number (that is, the size of the largest independent set of vertices) instead of the clique number. By Theorem 2.1.3 the independence number is with high probability equal to $2^{\lfloor \log n + \log\log n\rfloor}$, and hence, since every colour forms an independent set, by the pigeonhole principle the chromatic number is at least $2^{n-\lfloor \log n + \log\log n\rfloor}$. On the other hand, the proof of Theorem 2.1.3 reveals that not only that the clique number is as claimed, but also that the largest clique is formed by a subspace $V$ of dimension $\lfloor \log n + \log\log n\rfloor$. Notice that if a subspace forms an independent set, then the same is true for each of its translates, since $(i+V)\widehat{+}(i+V) = V\widehat{+}V$ for any $i \in \mathbb{F}_2^n$. Hence we can define a valid colouring by colouring each translate of $V$ with its own colour, thus using exactly $2^{n-\lfloor \log n + \log\log n\rfloor}$ colours. □

# Chapter 3

# Extractors in Paley graphs: a random model

A well-known conjecture in analytic number theory states that for every pair of sets $X, Y \subset \mathbb{Z}/p\mathbb{Z}$, each of size at least $\log^C p$ (for some constant $C$) we have that for $(\frac{1}{2} + o(1))|X||Y|$ of the pairs $(x, y) \in X \times Y$, $x + y$ is a quadratic residue modulo $p$. We address the probabilistic analogue of this question, that is for every fixed $\delta > 0$, given a finite group $G$ and $A \subset G$ a random subset of density $\frac{1}{2}$, we prove that with high probability for all subsets $|X|, |Y| \geq \log^{2+\delta} |G|$ for $(\frac{1}{2} + o(1))|X||Y|$ of the pairs $(x, y) \in X \times Y$ we have $xy \in A$.

This chapter is based on the paper [Mra16a].

## 3.1 Introduction

A folklore result in analytic number theory states that if we let $Q \subset \mathbb{Z}/p\mathbb{Z}$ be the set of quadratic residues modulo a prime $p$, then for any function $w\colon \mathbb{N} \to \mathbb{R}$ tending to infinity and any pair of sets $X, Y \subset \mathbb{Z}/p\mathbb{Z}$ of size at least $w(p)\sqrt{p}$, for $(\frac{1}{2}+o(1))|X||Y|$ of the pairs $(x, y) \in X \times Y$ we have $x + y \in Q$. Here, the rate of convergence implied by the $o$-notation depends only on $w$.

The proof of this is relatively simple. We refer the reader to Section 3.2 for the notation used below. First of all, after setting $1_X * 1_Y(x) = \mathbb{E}_{z \in \mathbb{Z}/p\mathbb{Z}} 1_X(z)1_Y(x - z)$, it is easy to see that the statement one wants to prove is equivalent to requiring that

$$\left| p \sum_{x \in \mathbb{Z}/p\mathbb{Z}} 1_X * 1_Y(x)\chi(x) \right| = o(|X||Y|),$$

where $\chi$ is the quadratic character (i.e. $\chi(a) = (\frac{a}{p})$). For $r \in \mathbb{Z}/p\mathbb{Z}$ and a function $f\colon \mathbb{Z}/p\mathbb{Z} \to \mathbb{R}$, we define the corresponding Fourier coefficient by

$$\widehat{f}(r) = \mathbb{E}_{x\in\mathbb{Z}/p\mathbb{Z}} f(x) e^{2\pi i x r/p}.$$

Using standard formulas from Fourier analysis, the standard estimate for Gauss sums (see e.g. [IK04], Section 3.5), and Cauchy-Schwarz inequality, we have

$$\begin{aligned}\left| p \sum_{x\in\mathbb{Z}/p\mathbb{Z}} 1_X * 1_Y(x)\chi(x) \right| &= \left| p^2 \sum_{r\in\mathbb{Z}/p\mathbb{Z}} \widehat{1_X}(r)\widehat{1_Y}(r)\bar{\widehat{\chi}}(r) \right| \leq p^{3/2} \sum_{r\in\mathbb{Z}/p\mathbb{Z}} |\widehat{1_X}(r)||\widehat{1_Y}(r)| \\ &\leq \sqrt{p}(|X||Y|)^{1/2},\end{aligned}$$

and this proves the claim. Although this argument was quite straightforward, no significant improvement (in terms of lower bounds for $|X|$ and $|Y|$) is known, although it is widely believed to be true even for sets $X$ and $Y$ of sizes at least $\log^C p$ for some constant $C$ – a conjecture known in some literature [CG88] as the Paley graph conjecture.

Given that the set of quadratic residues is believed to have many properties in common with a random set of the same size, it is natural to ask whether the statement above is true if we replace $Q$ with a genuinely random set. In this chapter we give the positive answer to this question, that is we prove the following theorem.

**Theorem 3.1.1.** *Let $G$ be a group of size $N$ and $w\colon \mathbb{N} \to \mathbb{R}$ some function that tends to infinity. Let $A \subset G$ be a random subset obtained by putting every element of $G$ into $A$ independently with probability $\frac{1}{2}$. Then the following holds with probability $1-o(1)$: for all sets $X, Y \subset G$, $|X|, |Y| \geq w(N)\log^2 N$, the number of pairs $(x,y) \in X \times Y$ such that $xy \in A$ differs from $\frac{1}{2}|X||Y|$ by $o(|X||Y|)$. The rate of convergence implied by the o-notation depends only on $w$.*

If for a pair of subsets $X, Y \subset G$ we have that the number of pairs $(x,y) \in X \times Y$ such that $xy \in A$ differs from $\frac{1}{2}|X||Y|$ by at most $\epsilon|X||Y|$, we will say that it is *$\epsilon$-extracted* by the set $A$. If all the pairs of subsets of size as in the previous theorem are $\epsilon$-extracted by the set $A$, we will say that $A$ is an *$\epsilon$-extractor*. Theorem 3.1.1 shows, in this terminology, that a random subset of $G$ is $o(1)$-extractor with high probability. The reason for this terminology will be explained in Section 3.6.

The most obvious approach to prove this theorem would be to use Fourier analysis as in the proof sketched above. Obviously, one would need to replace the nice property of the Fourier transform of $\chi$ by some equally useful property of the Fourier transform

of a random set. Indeed this is possible to do using some standard Chernoff-type estimates (e.g. Proposition 1.3.3). Unfortunately, this approach is also effective only for sets larger than $\sqrt{N}$. Furthermore, there are going to be problems with non-abelian groups.

In Section 3.3 we will present a different argument that will enable us to prove Theorem 3.1.1. A very important part of the argument is Proposition 3.3.1, which we prove in Section 3.4. Section 3.5 is devoted for proving a bound on the sizes of $X$ and $Y$ for which Theorem 3.1.1 does not hold. Finally, in Section 3.6 we give some open problems left after this chapter and explain the connection with randomness extractors.

Before moving on to the proof of the main theorem, let us first comment on the analogous question in the Erdős–Rényi setting. There, the corresponding result is the following.

**Theorem 3.1.2.** *Let $G = (V, E)$ be a graph sampled as in the Erdős–Rényi model $G(N, \frac{1}{2})$, that is $G$ has a vertex set $V$ of size $N$ and contains each possible edge with probability $\frac{1}{2}$, and these choices are all made independently. Let $w\colon \mathbb{N} \to \mathbb{R}$ be some function that tends to infinity. Then the following holds with probability $1 - o(1)$: for all sets $X, Y \subset V$, $|X|, |Y| \geq w(N) \log N$, for $\frac{1}{2} + o(1)$ of all the pairs $(x, y) \in X \times Y$, there is an edge connecting $x$ and $y$.*

Proof of this theorem follows easily by applying Hoeffding's inequality (1.2) and the union bound (1.1). We provide all the details since some parts of the proof would also appear in the proof of Theorem 3.1.1.

*Proof of Theorem 3.1.2.* Without loss of generality we may assume that $w(N) \leq \log N$. Let $\epsilon(N) = w(N)^{-1/3}$ and $X, Y \subset V$, each of size at least $w(N) \log N$. We will write $e(X, Y)$ to denote the number of pairs $x \in X, y \in Y$ such that there is an edge connecting $x$ and $y$, i.e. $\{x, y\} \in E$. We will say that a pair of sets $(X, Y)$ is *bad* if $e(X, Y)$ deviates too much from what we would expect, that is $|X||Y|/2$. More precisely, $(X, Y)$ would be a bad pair if

$$\left| e(X, Y) - \tfrac{|X||Y|}{2} \right| \geq \epsilon(N)|X||Y|.$$

It is easy to see that

$$e(X, Y) = \sum_{x \in X \setminus Y, y \in Y} 1_{\{x,y\} \in E} + \sum_{x \in X \cap Y, y \in Y \setminus X} 1_{\{x,y\} \in E} + \sum_{x,y \in X \cap Y} 2 \cdot 1_{\{x,y\} \in E}. \qquad (3.1)$$

It is obvious from this that

$$\mathbb{E}e(X,Y) = \frac{|X||Y|}{2} - \frac{|X \cap Y|}{2}.$$

More importantly, all the indicator functions appearing on the right hand side of (3.1) are independent. Thus, by applying Proposition 1.3.3 we get that for any $\lambda \geq 0$ we get

$$\mathbb{P}\left(\left|e(X,Y) - \tfrac{|X||Y|}{2} + \tfrac{|X\cap Y|}{2}\right| \geq \lambda\sqrt{\tfrac{|X||Y|}{4} + \tfrac{|X\cap Y|^2}{4} - \tfrac{|X\cap Y|}{2}}\right) \ll \exp(-\lambda^2/2).$$

Let $\lambda = \frac{\epsilon(N)}{2}\sqrt{|X||Y|}$. We have

$$\begin{aligned}
&\mathbb{P}\left((X,Y) \text{ is a bad pair}\right)\\
&\quad= \mathbb{P}\left(\left|e(X,Y) - \tfrac{|X||Y|}{2}\right| \geq \epsilon(N)|X||Y|\right)\\
&\quad\leq \mathbb{P}\left(\left|e(X,Y) - \tfrac{|X||Y|}{2} + \tfrac{|X\cap Y|}{2}\right| \geq \tfrac{\epsilon(N)}{2}|X||Y|\right)\\
&\quad\leq \mathbb{P}\left(\left|e(X,Y) - \tfrac{|X||Y|}{2} + \tfrac{|X\cap Y|}{2}\right| \geq \lambda\sqrt{\tfrac{|X||Y|}{4} + \tfrac{|X\cap Y|^2}{4} - \tfrac{|X\cap Y|}{2}}\right)\\
&\quad\ll \exp\left(-\lambda^2/2\right) \qquad (3.2)\\
&\quad= \exp\left(-\frac{\epsilon(N)^2|X||Y|}{8}\right), \qquad (3.3)
\end{aligned}$$

where (3.2) follows by applying Proposition 1.3.3.

We now fix a pair of cardinalities $k$ and $l$ ($k, l \geq w(N)\log N$), and do the union bound over all sets $X$ and $Y$ of these sizes. Notice that we can bound the number of such pairs of sets by $N^{k+l}$. Probability that there exists a bad pair of sets $(X,Y)$ such that $|X| = k$ and $|Y| = l$ is then

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{|X|=k,|Y|=l} (X,Y) \text{ is a bad pair}\right) &\ll N^{k+l}\exp\left(-\epsilon(N)^2kl/8\right)\\
&\ll \exp\left((k+l)\log N - w(N)^{-2/3}(k+l)w(N)\log N/16\right)
\end{aligned}$$

and this is obviously $o(1/N^2)$. We can now do the union bound over all possible sizes

$k$ and $l$

$$\begin{aligned}\mathbb{P}\left(\bigcup_{|X|,|Y|\geq w(N)\log N}\{(X,Y)\text{ is a bad pair}\}\right)&\\ \leq\sum_{k,l=w(N)\log N}^{N}\mathbb{P}\left(\bigcup_{|X|=k,|Y|=l}\{(X,Y)\text{ is a bad pair}\}\right)&\\ =o(1).&\end{aligned}$$

and this is what we wanted to prove. □

*Acknowledgements.* I would like to thank Sean Eberhard, Ben Green, Jakub Konieczny and Freddie Manners for helpful discussions, and two anonymous reviewers for carefully reading the paper [Mra16a] this chapter is based on and giving valuable comments.

## 3.2 Notation

Although most of the notation and conventions were implicitly introduced in the previous section, we include them here for the reader's convenience. The logarithm with base 2 will be denoted by $\log_2$, and natural logarithm by log. All the groups we will work with will be finite, and we equip each of them with the uniform probability measure. Of course, this reflects in our definitions of convolution, inner product and $L^2$-norm. To be specific, we will have

$$\begin{aligned}f*g(x)&=\mathbb{E}_{y\in G}f(y)g(y^{-1}x),\\ \langle f,g\rangle&=\mathbb{E}_{x\in G}f(x)g(x),\\ \|f\|_2&=\left(\mathbb{E}_{x\in G}|f(x)|^2\right)^{1/2}.\end{aligned}$$

## 3.3 The main argument

We will use the notation introduced in Theorem 3.1.1.

Let $\epsilon(N)=(10^5/w(N))^{1/6}$. We are going to prove that with probability only $o(1)$ there exist sets $X,Y\subset G$, both of size at least $w(N)\log^2 N$ such that the proportion of pairs $(x,y)\in X\times Y$ satisfying $xy\in A$ differs from $\frac{1}{2}$ by at least $\epsilon(N)$. In order

to make the random variables that will appear later centred, we associate to the set $A$ the function $f_A(x) = 2 \cdot 1_A - 1$. Note that $f_A$ takes values in $\{-1, 1\}$. From this perspective, it is easily seen that we are actually interested in whether

$$\left| \left\langle \frac{N^2 \cdot 1_X * 1_Y}{|X||Y|}, f_A \right\rangle \right| \geq 2\epsilon(N).$$

An obvious approach would be to mimic the proof of Theorem 3.1.2. For fixed $X, Y \subset G$, each of size at least $w(N)\log^2 N$ we can use Proposition 1.3.3 to get

$$\mathbb{P}\left( \left| \left\langle \frac{N^2 \cdot 1_X * 1_Y}{|X||Y|}, f_A \right\rangle \right| \geq 2\epsilon(N) \right) \ll \exp\left( -\frac{2\epsilon(N)^2 |X|^2 |Y|^2}{N^3 \|1_X * 1_Y\|_2^2} \right).$$

Notice that this bound gets weaker as the quantity $\|1_X * 1_Y\|_2^2$ increases. We remind the reader that this quantity is, up to normalization, equal to the number of quadruples $x_1, x_2 \in X$, $y_1, y_2 \in Y$ such that $x_1 + x_2 = y_1 + y_2$.

The next step would then be to simply use the union bound as in the proof of Theorem 3.1.2. Unfortunately, this wouldn't work and the problem is caused by pairs $(X, Y)$ for which $\|1_X * 1_Y\|_2^2$ is close to maximal possible (that is $\frac{1}{N^3}|X||Y|\min(|X|, |Y|)$). To see this, one can take e.g. $G = \mathbb{Z}/N\mathbb{Z}$, $k = \log^3 N$ and consider the family $\mathcal{F}$ of all $k$-element sets of $X \subset G$ such that $\{1, \dots, k/2\} \subset X$. This family contains $\binom{N-k/2}{k/2} \geq N^{\Omega(k)}$ and for each set $X \in \mathcal{F}$ we have $N^3 \|1_X * 1_X\|_2^2 \gg k^3$ (because there are $\gg k^3$ quadruples $x_1, x_2, x_3, x_4 \in \{1, \dots, k/2\}$ such that $x_1 + x_2 = x_3 + x_4$). For this reason the sum

$$\sum_{X \in \mathcal{F}} \exp\left( -\frac{2\epsilon(N)^2 |X|^4}{N^3 \|1_X * 1_X\|_2^2} \right)$$

tends to infinity and the approach from the proof of Theorem 3.1.2 cannot work here.

Our argument for proving Theorem 3.1.1 is the following. We will prove that whenever there are pairs of sets $X$ and $Y$ with an unusual proportion of elements whose product is in $A$, we can find some related sets $S$ and $T$ that share this property, but, crucially, such that $\|1_S * 1_T\|_2^2$ is significantly smaller than the maximum possible so the final union bound would be effective. This rough idea of relating arbitrary sets to unstructured ones and then continue working with these is already present in some previous works and it goes back at least as far as [AAAS94].

The following proposition gives us sets $S$ and $T$ with these nice properties.

**Proposition 3.3.1.** *Let $f\colon G \to \{-1, 1\}$ be any function, and $X, Y \subset G$ both of size at least $w(N)\log^2 N$ and $|X| \leq |Y|$. There exist sets $S, T \subset G$ satisfying the following properties*

(*i*) $|S| \leq |T|$ *and* $|T| \to \infty$,

(*ii*) $\left|\left\langle \frac{N^2 \cdot 1_S * 1_T}{|S||T|}, f \right\rangle - \left\langle \frac{N^2 \cdot 1_X * 1_Y}{|X||Y|}, f \right\rangle\right| \leq \epsilon(N)$,

(*iii*) $\frac{|S|^2|T|^2}{N^3\|1_S * 1_T\|_2^2} \geq \frac{6|T|\log N}{\epsilon(N)^2}$.

We give a proof of this proposition in the next section, and here finish the proof of the main theorem conditional on it.

Suppose that there are sets $X, Y$ of size at least $w(N)\log^2 N$ such that

$$\left|\left\langle \frac{N^2 \cdot 1_X * 1_Y}{|X||Y|}, f_A \right\rangle\right| \geq 2\epsilon(N).$$

Take subsets $S$ and $T$ given by Proposition 3.3.1; by the triangle inequality we have

$$\left|\left\langle \frac{N^2 \cdot 1_S * 1_T}{|S||T|}, f_A \right\rangle\right| \geq \epsilon(N).$$

However, by our choice of $S$ and $T$ this is very unlikely to happen; similar application of Hoeffding's inequality as above gives that the probability is bounded above by

$$\exp\left(-\frac{\epsilon(N)^2|S|^2|T|^2}{2N^3\|1_S * 1_T\|_2^2}\right) \leq \exp\left(-3|T|\log N\right).$$

Now we do the union bound over all sets $S$ and $T$ of fixed sizes satisfying the properties stated in Proposition 3.3.1. We use a crude bound $N^{2|T|}$ for the number of these sets (because $|S| \leq |T|$). The contribution coming from these sets is hence bounded by

$$N^{2|T|}e^{-3|T|\log N} = e^{-|T|\log N}.$$

What remains is to sum over all possible sizes of $S$ and $T$. There are $N^2$ terms and since the size of $T$ tends to infinity, the total sum is still $o(1)$.

## 3.4 Proof of the main proposition

In this section we give the probabilistic proof of Proposition 3.3.1. The idea is to take a random subset $S \subset X$ uniformly from all subsets of some fixed size $s$, and a random subset $T \subset Y$ uniformly from all subsets of some fixed size $t$. Intuitively, we believe that $1_S * 1_T$ should in some way imitate $1_X * 1_Y$. This idea is already present in the work of Croot and Sisask [CS10]. Additionally, if $s$ and $t$ are somewhat smaller than

$|X|$ and $|Y|$, $\|1_S * 1_T\|_2^2$ should be smaller than the theoretical maximum, even if $X$ and $Y$ are highly structured.

We formalise the outlined strategy. In the following, $S$ and $T$ are sets chosen randomly as described, and hence instances of $\mathbb{P}$ (probability) and $\mathbb{E}$ (expectation) notation are over this random choice, unless indicated otherwise. We first prove that the additive energy is on average small.

**Lemma 3.4.1.**

$$\mathbb{E}\|1_S * 1_T\|_2^2 \leq \frac{st}{N^3} + \frac{s^2t^2}{|X|^2|Y|^2}\|1_X * 1_Y\|_2^2.$$

*Proof.* Note that the left hand side is equal to

$$\frac{st}{N^3} + \frac{1}{N^3}\sum \mathbb{P}(x_1, x_2 \in S, y_1, y_2 \in T),$$

where the sum is taken over $x_1, x_2 \in X$, $y_1, y_2 \in Y$ satisfying $x_1 \neq x_2$, $y_1 \neq y_2$, $x_1y_1 = x_2y_2$. This is obviously bounded by

$$\frac{st}{N^3} + \|1_X * 1_Y\|_2^2 \cdot \frac{s(s-1)}{|X|(|X|-1)}\frac{t(t-1)}{|Y|(|Y|-1)},$$

and hence we are done. □

**Lemma 3.4.2.** *For any function $f\colon G \to \{-1, 1\}$ we have*

$$\mathbb{E}\left|\left\langle \frac{N^2 \cdot 1_S * 1_T}{|S||T|}, f\right\rangle - \left\langle \frac{N^2 \cdot 1_X * 1_Y}{|X||Y|}, f\right\rangle\right| \leq 2\sqrt{\frac{|Y|}{st}}$$

*Proof.* Notice that

$$\begin{aligned}
&\mathbb{E}\left|\left\langle \frac{N^2 \cdot 1_S * 1_T}{|S||T|}, f\right\rangle - \left\langle \frac{N^2 \cdot 1_X * 1_Y}{|X||Y|}, f\right\rangle\right| \\
&= \mathbb{E}\left|\sum_{z\in G}\sum_{w\in G} f(z) \cdot \left(\frac{1_S(w)1_T(w^{-1}z)}{st} - \frac{1_X(w)1_Y(w^{-1}z)}{|X||Y|}\right)\right| \\
&= \mathbb{E}\left|\sum_{x\in X}\sum_{y\in Y} f(xy) \cdot \left(\frac{1_S(x)1_T(y)}{st} - \frac{1}{|X||Y|}\right)\right| \\
&= \frac{1}{st}\mathbb{E}\left|\sum_{x\in X}\sum_{y\in Y} f(xy) \cdot (1_S(x)1_T(y) - \mathbb{P}(x \in S, y \in T))\right|
\end{aligned}$$

The idea now is to exploit the fact that lots of the random variables in the sum above are independent, and hence we expect a fair amount of cancellation. To do this we will partition the set $X \times Y$ into few parts, such that corresponding random variables in each part are (very close to) independent. This partition is easy to describe if we identify $X$ and $Y$ with $\{1, \dots, |X|\}$ and $\{1, \dots, |Y|\}$; for $i = 1, \dots, |Y|$, let $\Omega_i = \{(j, j+i) \colon j \in \{1, \dots, |X|\}\}$ (with addition modulo $|Y|$). We first analyse the sum above when taken over all pairs with $f(xy) = 1$; we will denote this set by $\mathcal{F}$. We have

$$\begin{aligned}
&\frac{1}{st}\mathbb{E}\left|\sum_{(x,y)\in\mathcal{F}}(1_S(x)1_T(y) - \mathbb{P}(x \in S, y \in T))\right| \\
&\leq \frac{1}{st}\sum_{i=1}^{|Y|}\mathbb{E}\left|\sum_{(x,y)\in\Omega_i\cap\mathcal{F}}(1_S(x)1_T(y) - \mathbb{P}(x \in S, y \in T))\right| \\
&\leq \frac{1}{st}\sum_{i=1}^{|Y|}\left(\mathbb{E}\left(\sum_{(x,y)\in\Omega_i\cap\mathcal{F}}(1_S(x)1_T(y) - \mathbb{P}(x \in S, y \in T))\right)^2\right)^{1/2} \\
&= \frac{1}{st}\sum_{i=1}^{|Y|}\left(\sum_{(x_j,y_j)\in\Omega_i\cap\mathcal{F}}(\mathbb{P}(x_1, x_2 \in S, y_1, y_2 \in T) - \right. \\
&\qquad\qquad \left. - \mathbb{P}(x_1 \in S, y_1 \in T)\mathbb{P}(x_2 \in S, y_2 \in T))\right)^{1/2}. \qquad (3.4)
\end{aligned}$$

Consider two elements $(x_1, y_1) \neq (x_2, y_2)$ from $\Omega_i \cap \mathcal{F}$. Notice that, by the definition of $\Omega_i$, this implies that $x_1 \neq x_2$ and $y_1 \neq y_2$. We have

$$\mathbb{P}(x_1, x_2 \in S, y_1, y_2 \in T) = \frac{s(s-1)}{|X|(|X|-1)}\frac{t(t-1)}{|Y|(|Y|-1)}$$

and

$$\mathbb{P}(x_1 \in S, y_1 \in T)\mathbb{P}(x_2 \in S, y_2 \in T) = \frac{s^2t^2}{|X|^2|Y|^2},$$

so the contribution to (3.4) coming from nondiagonal terms is nonpositive. Hence

$$\begin{aligned}(3.4) &\leq \frac{1}{st}\sum_{i=1}^{|Y|}\left(\sum_{(x,y)\in\Omega_i\cap\mathcal{F}}\mathbb{P}(x\in S, y\in T)\right)^{1/2}\\ &\leq \frac{1}{\sqrt{st|X||Y|}}\sum_{i=1}^{|Y|}|\Omega_i|^{1/2}\\ &\leq \sqrt{\frac{|Y|}{st}}.\end{aligned}$$

Doing the same thing for the other sum (over $\mathcal{F}^c$), we get the statement of the lemma. □

We can now finish the proof of Proposition 3.3.1 using Markov's inequality. Let $s = \frac{2000\log N}{\epsilon(N)^4}$ and $t = \frac{|Y|\epsilon(N)^2}{50\log N}$. By Lemma 3.4.2 we see that the probability that

$$\left|\left\langle \frac{N^2\cdot 1_S * 1_T}{st}, f\right\rangle - \left\langle \frac{N^2\cdot 1_X * 1_Y}{|X||Y|}, f\right\rangle\right| \leq 6\sqrt{\frac{|Y|}{st}} \leq \epsilon(N) \tag{3.5}$$

is at least $\frac{2}{3}$. Similarly, Lemma 3.4.1 implies that

$$\|1_S * 1_T\|_2^2 \leq \frac{3st}{N^3} + \frac{3s^2t^2}{|X|^2|Y|^2}\|1_X * 1_Y\|_2^2, \tag{3.6}$$

happens with probability at least $\frac{2}{3}$ so with positive probability we have both of the requirements satisfied. Inequality (3.6) implies

$$\frac{s^2t^2}{N^3\|1_S * 1_T\|_2^2} \geq \min\left(st/6, \frac{|X|^2|Y|^2}{6N^3\|1_X * 1_Y\|_2^2}\right) \geq \min(st/6, |Y|/6) \geq \frac{6t\log N}{\epsilon(N)^2},$$

and this finishes the proof.

## 3.5 Bounds for the sizes of extracted sets

Recall that Theorem 3.1.1 states that with high probability all pairs of subsets $(X, Y)$ of size at least $w(N)\log^2 N$ are $o(1)$-extracted by a random subset. In this section we deal with the question of how far the bound $w(N)\log^2 N$ is from optimal. The approach we will use here is conceptually not difficult; it's a basic second moment method. However, some technical work needs to be done to get the result. We note that some of the calculations have already been done in [Gre05a, Section 8]. Indeed,

the main argument there already gives a nontrivial lower bound on the minimum size we must impose on $X$ and $Y$ in order for Theorem 3.1.1 to hold. It is an easy consequence of Green's argument that with probability $\frac{1}{2} - o(1)$ there is a subset $X \subset \mathbb{F}_2^n$, $N = 2^n$ of size $\frac{1}{4}\log_2 N \log_2 \log_2 N$ such that $X + X \subset A$. Obviously, $X + X \subset A$ is a much stronger statement than is required to contradict the conclusion of Theorem 3.1.1. In this section we show that the constant $\frac{1}{4}$ can be made arbitrarily large if we relax this statement so that Theorem 3.1.1 only just fails. The main theorem is as follows.

**Theorem 3.5.1.** *Fix $C > 0$ and let $A$ be a random subset of $\mathbb{F}_2^n$. There exists $\epsilon > 0$ such that, for $n$ large enough, with probability bounded away from 0 there exists a subset $X \subset \mathbb{F}_2^n$ of size at least $C \log N \log\log N$ such that for at least $\frac{1}{2} + \epsilon$ of the pairs $(x_1, x_2) \in X \times X$ we have $x_1 + x_2 \in A$.*

As will be obvious from the proof, we will actually choose $X$ to be a subspace of relatively large dimension. Notice that in this case we have that

$$|\{(x_1, x_2) \colon x_1, x_2 \in X, x_1 + x_2 \in A\}| = |X||A \cap X|,$$

so what we need to prove is that $|A \cap X| \geq (\frac{1}{2} + \epsilon)|X|$. We should also mention that in this setting (i.e. when $X$ must be a subspace), the bound in Theorem 3.5.1 is optimal up to a constant. This is easily proven using the first moment method.

Before moving to the proof, we give two simple counting lemmas we will need. These lemmas are already present in [Gre05a]; as the proofs are very easy, we include them here for completeness.

**Lemma 3.5.2.** *Let $I_k$ be the number of $k$-dimensional subspaces of $\mathbb{F}_2^n$. Then $I_k \geq 2^{nk-k^2}$.*

*Proof.* The number of linearly independent $k$-tuples (i.e. basis of $k$-dimensional subspaces) is equal to

$$(2^n - 1)(2^n - 2)\dots(2^n - 2^{k-1}).$$

However, by the same calculation, each $k$-dimensional subspace was counted exactly

$$(2^k - 1)(2^k - 2)\dots(2^k - 2^{k-1})$$

times, and hence in total there are exactly

$$\frac{(2^n - 1)(2^n - 2)\dots(2^n - 2^{k-1})}{(2^k - 1)(2^k - 2)\dots(2^k - 2^{k-1})}$$

$k$-dimensional subspaces. The lower bound follows after observing that $\frac{2^n-2^i}{2^k-2^i} \geq 2^{n-k}$ for every $i \leq k-1$. □

**Lemma 3.5.3.** *Let $J_k^{(l)}$ be the number of pairs of $k$-dimensional subspaces of $\mathbb{F}_2^n$ whose intersection is a $l$-dimensional subspace. Then $J_k^{(l)} \leq 2^{2nk-nl}$.*

*Proof.* Denote some $k$-dimensional subspaces satisfying the statement of the lemma by $V$ and $W$. First of all, we choose the $l$-dimensional subspace $V \cap W$. By the previous lemma, this can be done in

$$\frac{(2^n-1)(2^n-2)\dots(2^n-2^{l-1})}{(2^l-1)(2^l-2)\dots(2^l-2^{l-1})}$$

ways. Now each of the subspaces $V$ and $W$ is missing $k-l$ linearly independent vectors, and these $2(k-l)$ vectors must all be linearly independent. Hence, there are

$$(2^n-2^l)(2^n-2^{l+1})\dots(2^n-2^{2k-l-1})$$

ways to do this. Of course, many of these choices give the same subspaces $V$ and $W$. Specifically, once $V$ and $V \cap W$ are fixed, there are $(2^k-2^l)(2^k-2^{l+1})\dots(2^k-2^{k-1})$ to choose remaining basis vectors, and the same holds for $W$. The conclusion is that

$$J_k^{(l)} = \frac{(2^n-1)(2^n-2)\dots(2^n-2^{2k-l-1})}{(2^k-2^l)^2\dots(2^k-2^{k-1})^2(2^l-1)\dots(2^l-2^{l-1})}.$$

Bounding each term in the numerator by $2^n$ gives the claimed bound. □

In the course of the proof of Theorem 3.5.1 we will also need rather precise upper bounds for an upper tail of a binomial distribution, more precise than the one provided by Hoeffding's inequality. Additionally, we will need a corresponding lower bound. We proved these bounds in Chapter 1, so here we just state them for the convenience of the reader.

**Proposition 1.3.4.** *Let $Z$ be a $B(n, \frac{1}{2})$ random variable and $0 < \epsilon < \frac{1}{2}$ fixed. If we define $h(x) = (\frac{1}{2}+x)\log(1+2x) + (\frac{1}{2}-x)\log(1-2x)$, then*

$$e^{-nh(\epsilon)-O(\log n)} \leq \mathbb{P}(Z \geq (\tfrac{1}{2}+\epsilon)n) \leq e^{-nh(\epsilon)}.$$

We finally move to the proof of the main result in this section. As we noted above, we will use the second moment method. Fix $\epsilon > 0$ and let $k = \lfloor \log_2 n + \log_2 \log_2 n \rfloor + c$,

where $c > 0$ is a constant to be specified later. Let $\{V_i\}$ be the family of all $k$-dimensional subspaces of $\mathbb{F}_2^n$ and let $X_i$ be the random variable indicating if $|A \cap V_i| \geq (\frac{1}{2} + \epsilon)|V_i|$ and $X = \sum_i X_i$ be the number of $V_i$ such that $|A \cap V_i| \geq (\frac{1}{2} + \epsilon)|V_i|$. We claim that with probability bounded away from zero we have $X > 0$, equivalently, with probability bounded away from one we have $X = 0$.

Let $P_k = \mathbb{P}(|A \cap V_i| \geq (\frac{1}{2} + \epsilon)|V_i|)$ and $C_k^{(l)} = \mathrm{cov}(X_i, X_j)$ for $i \neq j$ and $\dim(V_i \cap V_j) = l$. We will use the notation introduced in the statements of Lemma 3.5.2, Lemma 3.5.3 and Proposition 1.3.4.

Obviously, $\mathbb{E}X = I_k P_k$. On the other hand

$$\mathrm{Var}X = \sum_{i,j} \mathrm{cov}(X_i, X_j) = \sum_{l=0}^{k} J_k^{(l)} C_k^{(l)}.$$

Note that by Chebyshev's inequality we have

$$\mathbb{P}(X = 0) \leq \frac{\mathrm{Var}X}{(\mathbb{E}X)^2} = \sum_{l=0}^{k} I_k^{-2} P_k^{-2} J_k^{(l)} C_k^{(l)}, \tag{3.7}$$

and our aim now is to prove that the quantity on the right is bounded away from 1.

We will inspect three different ranges for $l$. To begin with, let $l = 0$. Obviously, $I_k^{-2} J_k^{(0)} \leq 1$. If we let $K = 2^k$, $H = (\frac{1}{2} + \epsilon)K$ and $Z$ be a $B(K-1, \frac{1}{2})$ random variable, by conditioning on whether 0 (which is the only element of $V \cap W$) is in $A$, it is easy to see that

$$\begin{aligned} C_k^{(0)} &= \tfrac{1}{2}\mathbb{P}(Z \geq H-1)^2 + \tfrac{1}{2}\mathbb{P}(Z \geq H)^2 - \left(\tfrac{1}{2}\mathbb{P}(Z \geq H-1) + \tfrac{1}{2}\mathbb{P}(Z \geq H)\right)^2 \\ &= \tfrac{1}{4}\left(\mathbb{P}(Z \geq H-1) - \mathbb{P}(Z \geq H)\right)^2. \end{aligned}$$

On the other hand, by the same conditioning we get

$$P_k = \tfrac{1}{2}\mathbb{P}(Z \geq H-1) + \tfrac{1}{2}\mathbb{P}(Z \geq H),$$

and hence

$$P_k^{-2} C_k^{(0)} \leq \left(\frac{\mathbb{P}(Z = H-1)}{\mathbb{P}(Z = H-1) + \mathbb{P}(Z = H)}\right)^2 \leq (\tfrac{1}{2} + \epsilon)^2,$$

where the last inequality follows by comparing appropriate binomial coefficients. Combining these observations we get

$$I_k^{-2} P_k^{-2} J_k^{(0)} C_k^{(0)} \leq (\tfrac{1}{2} + \epsilon)^2. \tag{3.8}$$

Suppose now that $1 \leq l \leq \log_2 n$, and let $d = k - l$ and $D = 2^d$. Obviously, $D \geq \log_2 n$. We will bound the $C_k^{(l)}$ by the probability

$$\mathbb{P}(|A \cap V_i| \geq (\tfrac{1}{2} + \epsilon)|V_i|, |A \cap V_j| \geq (\tfrac{1}{2} + \epsilon)|V_j|),$$

where $i \neq j$ and $V_i$ and $V_j$ have $l$-dimensional intersection $W$. Notice that $|V_i \setminus W| = |V_j \setminus W| = \frac{D-1}{D} 2^k$, and if we want both of the events to hold, then certainly $V_1 \setminus W$ and $V_2 \setminus W$ each must contain at least $(\frac{1}{2} + \epsilon - 1/D)2^k$ elements from $A$, since $W$ has $2^k/D$ elements. However, by Proposition 1.3.4, the probability of this happening is at most

$$\begin{aligned}
\exp\left(-2 \cdot \tfrac{D-1}{D} 2^k h(\tfrac{\epsilon D}{D-1} - \tfrac{1}{2(D-1)})\right) &= \exp\left(-2^{k+1} h(\tfrac{\epsilon D}{D-1} - \tfrac{1}{2(D-1)}) + O(2^k/D)\right) \\
&= \exp\left(-2^{k+1} h(\epsilon) + O(2^k/D)\right) \\
&= \exp\left(-2^{k+1} h(\epsilon) + O(2^l)\right).
\end{aligned}$$

The second equality above follows by the intermediate value theorem. After using Lemma 3.5.3 and noticing that for each $1 \leq l \leq \log_2 n$ we have $O(2^l) - nl \leq O(1) - n$ we can conclude

$$\sum_{l=1}^{\log_2 n} J_k^{(l)} C_{k,l} \leq 2^{2nk - n + O(\log_2 \log_2 n)} e^{-2^{k+1} h(\epsilon)}.$$

We can now use the lower bound from Proposition 1.3.4 to bound $P_k$

$$P_k \geq e^{-2^k h(\epsilon) - O(k)}.$$

Using this and Lemma 3.5.2 we get

$$\sum_{l=1}^{\log_2 n} I_k^{-2} P_k^{-2} J_k^{(l)} C_k^{(l)} \leq 2^{2k^2 + O(k) - n}, \tag{3.9}$$

and this is $o(1)$.

Finally, for $l \geq \log_2 n$ we can bound $C_k^{(l)}$ trivially by 1, and get, similarly as in the previous case,

$$\begin{aligned}
\sum_{l=\log_2 n}^{k} I_k^{-2} P_k^{-2} J_k^{(l)} C_k^{(l)} &\leq k 2^{2k^2 - n \log_2 n} e^{2^{k+1} h(\epsilon) + O(k)} \\
&\leq 2^{2k^2 - n \log_2 n} e^{2^{c+1} h(\epsilon) n \log_2 n + O(k)}.
\end{aligned} \tag{3.10}$$

This is $o(1)$ as long as $2^{c+1}h(\epsilon) < \log 2$. Now from (3.7), (3.8), (3.9) and (3.10)

$$P(X = 0) \leq (\tfrac{1}{2} + \epsilon)^2 + o(1),$$

as long as $2^{c+1}h(\epsilon) < \log 2$. Since $h(\epsilon) \to 0$ as $\epsilon \to 0$, we see that we can have $c$ arbitrary large if we choose $\epsilon$ small enough.

## 3.6 Further comments

We have phrased Theorem 3.1.1 in terms of the general finite group $G$. One could, of course, ask the same question for a specific class of groups, say, $\mathbb{Z}/N\mathbb{Z}$ or $\mathbb{F}_2^n$. For example, it is possible to prove that in $\mathbb{Z}/N\mathbb{Z}$ with high probability there exists a subset $X$ of size about $2\log_2 N$ such that $X + X \subset A$. This already gives a bound for the sizes of sets which we can hope to be extracted. One could presumably push this a bit further similarly as we did for $\mathbb{F}_2^n$ in Theorem 3.5.1.

Notice that in the Erdős–Rényi model, the bound $w(N)\log_2 N$ in Theorem 3.1.2 is quite close to optimal. Indeed, the clique number in this graph is with high probability about $2\log_2 N$. One could consider this as a (weak) indication that it might be possible to push the bound in Theorem 3.1.1 further down, maybe even to something like $w(N)\log N \log\log N$. However, methods such as those in this chapter are not sufficient to achieve anything close to that.

Our result can also be considered from the perspective of randomness extractors. We say that a function is a randomness extractor if it transforms every (e.g. pair or sequence of) not too degenerate independent random variables into a Bernoulli random variable which is quite close to uniform. In other words, a randomness extractor takes a bad nonuniform source and creates a nice almost uniform output.

A folklore example of this phenomena is the so-called von Neumann extractor. Suppose that $(X_n)$ is a sequence of independent identically distributed nondegenerate Bernoulli random variables. One would like to construct a function of this sequence which produces a symmetric Bernoulli random variable. Let $f((x_n)) = x_m$, where $m$ is the smallest even index such that $x_m \neq x_{m+1}$. It is easy to see that this function does the job.

Of course, the definition as stated above is far from precise, we give a rigorous definition adapted to the setting we will be interested in. Given a random variable $X$ with values in a finite set $S$, we measure how nondegenerate it is (or, at the other

extreme, how uniform it is) with min-entropy (we will call it entropy from now on)

$$H(X) = \min_{s \in S} \log_2 1/\mathbb{P}(X = s).$$

To gain some intuition, it is worth mentioning that entropy takes values between 0 and $\log_2 |S|$, and that $H(X) = 0$ if and only if $X$ concentrates on a single value (i.e. is completely deterministic), and $H(X) = \log_2 |S|$ if and only if $X$ is uniform on $S$ (i.e. is as unpredictable as one could possibly hope).

Fix $h, c > 0$ (one should think of $h$ as being small compared to $\log_2 |S|$ and $c$ as close to 1). We will say that $f \colon S \times S \to \{0, 1\}$ is a $(h, c)$-randomness extractor if for every pair of independent random variables $X$ and $Y$ with values on $S$ satisfying $H(X), H(Y) > h$, we have $H(f(X, Y)) > c$.

It was proven by Chor and Goldreich [CG88], that it is only necessary to check that $f(X, Y)$ extracts randomness well (i.e. satisfies the required lower bound on the entropy) when $X$ and $Y$ are *flat*, that is to say distribution of each of them is supported and uniformly distributed on a subset of $S$. Having this in mind, our result proves that for any fixed $\epsilon > 0$, for a randomly chosen set $A$ of density $\frac{1}{2}$, the function $(x, y) \to 1_A(xy)$ is $((2+\epsilon) \log_2 \log_2 N, 1-o(1))$-randomness extractor with probability $1 - o(1)$ (we remind the reader that in this case we have $S = \mathbb{Z}/N\mathbb{Z}$). An important area of research is to find explicit $(\delta \log_2 |S|, 1 - o_{|S| \to \infty}(1))$-randomness extractors for small values of $\delta$ ($\delta < 1/2$ seems to be particularly challenging regime). In that respect we should mention a major breakthrough by Chattopadhyay and Zuckerman who have recently found [CZ16] an explicit $((\log_2 \log_2 |S|)^{O(1)}, 1 - o_{|S| \to \infty}(1))$-randomness extractor. For more about this and related issues see e.g. the first section of the paper by Barak, Impagliazzo, and Wigderson [BIW06].

# Chapter 4

# A random model for the Paley graph

For a prime $p$ we define the Paley graph to be the graph with the set of vertices $\mathbb{Z}/p\mathbb{Z}$, and with edges connecting vertices whose sum is a quadratic residue. Paley graphs are notoriously difficult to study, particularly finding bounds for their clique numbers. For this reason, it is desirable to have a random model. A well known result of Graham and Ringrose shows that the clique number of the Paley graph is $\Omega(\log p \log\log\log p)$ (even $\Omega(\log p \log\log p)$, under the generalized Riemann hypothesis) for infinitely many primes $p$ – a behaviour not detected by the random Cayley graph which is hence deficient as a random model for the Paley graph. In this chapter we give a new probabilistic model which incorporates some multiplicative structure and as a result captures the Graham-Ringrose phenomenon. We prove that if we sample such a random graph independently for every prime $p$, then almost surely (i) for infinitely many primes $p$ the clique number is $\Omega(\log p \log\log p)$, whilst (ii) for almost all primes the clique number is $(2 + o(1)) \log p$.

This chapter is based on the preprint [Mra16b].

## 4.1 Introduction

Let $N$ be a prime and $R \subset \mathbb{Z}/N\mathbb{Z}$ be the set of quadratic residues. We define the so-called *Paley sum graph* $\Gamma_R$ to be the graph with the set of vertices $\mathbb{Z}/N\mathbb{Z}$, and edges connecting vertices whose sum is in $R$. It is a well known open problem to find good asymptotics for the clique number $\omega(\Gamma_R)$ of this graph, that is the size of the largest complete subgraph.

Cohen [Coh88] proved the lower bound $(1/2 + o(1)) \log N$ (all logarithms in this

chapter will be with base 2). On the other hand, using basic Fourier analysis and standard estimates for Gauss sums, one can easily prove the upper bound $\sqrt{N}$. There are only minor improvements of this bound. For example, Bachoc, Matolcsi and Ruzsa [BMR13] proved that there are infinitely many primes $N$ for which the clique number is at most $\sqrt{N} - 1$. Their result has been slightly generalized recently by Greaves and Soicher [GS16].

It is widely believed that the set of quadratic residues should have properties similar to a random subset of density $1/2$, which leads to an obvious question of finding the clique number of *the random Cayley sum graph* $\Gamma_A$. As the notation suggests, $\Gamma_A$ is obtained by first choosing a random subset $A \subset \mathbb{Z}/N\mathbb{Z}$ by putting each element in it independently with probability $1/2$, and then joining vertices $x, y \in \mathbb{Z}/N\mathbb{Z}$ if and only if $x + y \in A$. Recently, Green and Morris [GM15] proved that with high probability $\omega(\Gamma_A) = (2 + o(1)) \log N$ which suggests that the clique number of the Paley sum graph might also be close to this value.

However, it is known that the clique number of the Paley sum graph is a little bit bigger than $2 \log N$ for infinitely many primes $N$. Indeed, Graham and Ringrose [GR90] proved that for infinitely many primes $N$ the smallest quadratic nonresidue $q$ is $\Omega(\log N \log\log\log N)$, that is at least $c \log N \log\log\log N$ for some constant $c > 0$. Obviously, for these primes the set $\{1, 2, \ldots, q/2\}$ forms a large clique in $\Gamma_R$. Moreover, Montgomery [Mon71, page 122] proved that this result can, under generalized Riemann hypothesis, be improved to $\Omega(\log N \log\log N)$.

On the other hand, it is easily seen from the method used by Green and Morris that

$$\mathbb{P}(\omega(\Gamma_A) > 10 \log N) \leq 1/N^2.$$

It follows, by the Borel-Cantelli lemma, that if we sample a random Cayley sum graph for each prime $N$, the clique number of only boundedly many of them would be greater than $10 \log N$, which is in contrast to the result of Graham and Ringrose.

In this chapter we introduce a different random graph model for the Paley graph. We show that this model usually gives the same clique number as in the random Cayley sum graphs, but also has the phenomenon present in the result of Graham and Ringrose.

To motivate the construction, notice that the occasional largeness of the clique number of the Paley graph is, by the result of Graham and Ringorse, already caused by the presence of the multiplicative structure in the initial interval $\{1, \ldots, O(\log N \log\log\log N)\}$ (or, in the case of Montgomery's conditional result, $\{1, \ldots, O(\log N \log\log N)\}$). In other words, the multiplicative structure present in the set of quadratic

residues makes it relatively easy to have a large clique. For example, if we know that all the primes up to, say, 100 are quadratic residues, then all the numbers $1, \ldots, 100$ are also quadratic residues, and hence $\{1, \ldots, 50\}$ spans a clique. The idea is then to develop a model which would reflect this multiplicative phenomenon present in the Paley graph, at least in an initial interval (which is what the result of Graham and Ringrose is about).

This suggests the following model.

**Definition 4.1.1.** Let $Q \geq 1$ be an integer. We define a random function $f\colon \mathbb{Z}/N\mathbb{Z} \to \{-1, 1\}$ in the following way. For every prime $p \in [1, Q]$ (here we embed $\mathbb{Z}/N\mathbb{Z}$ inside $\mathbb{Z}$ in an obvious way), we set $f(p)$ to be uniform $\pm 1$ random variable, and we make all these random variables independent. Next, we extend $f$ to be completely multiplicative on $[1, Q]$. Finally, for each $x \notin [1, Q]$ we set $f(x)$ to be again uniform $\pm 1$ random variable, independently of all other choices. We will say that $f$ is a *$Q$-multiplicative random function*. Let $\Gamma_f$ be the random graph with vertex set $\mathbb{Z}/N\mathbb{Z}$ and edges connecting $x \neq y$ if and only if $f(x + y) = 1$. We will say that $\Gamma_f$ is generated by $f$.

Of course, the intuition here is that $f$ is a random model for the quadratic character $\left(\frac{\cdot}{N}\right)$ and thus the set $\{x \in \mathbb{Z}/n\mathbb{Z}\colon f(x) = 1\}$ is a random model for the set of the quadratic residues $R$. In Section 4.3 we will sketch why the decision to take values $f(p)$ (for small primes $p$) independently was sensible.

The majority of the chapter will be devoted to proving that the clique number in our random model is with high probability of the same size as in the random Cayley graph.

**Theorem 4.1.2.** *There exists a positive constant $c < 1/2$, such that for $Q = c \log N \log\log N$ the following holds. Let $\epsilon > 0$ and $f$ be the $Q$-multiplicative random function. Then*

$$\mathbb{P}((2 - \epsilon) \log N \leq \omega(\Gamma_f) \leq (2 + \epsilon) \log N) = 1 - o(1).$$

We prove the upper and lower bounds from this theorem in Sections 4.4 and 4.5, respectively.

Let $R$ be a subset of the primes. We define its relative density in primes to be $\lim_{M\to\infty} \frac{|R\cap[1,M]|}{\pi(M)}$ (if the limit exists), where $\pi(M)$ denotes, as usual, the number of primes up to $M$. As an easy consequence of the previous theorem we will prove the following, which shows that although there will be infinitely many primes $N$ for which the clique number is $\Omega(\log N \log\log N)$, for most primes $N$ it will be about $2 \log N$.

**Theorem 4.1.3.** *Let $\epsilon > 0$ and $c > 0$ be a constant for which Theorem 4.1.2 holds. For each prime $N$ independently sample a random graph $\Gamma_f^{(N)}$. Then, almost surely, the clique number $\omega(\Gamma_f^{(N)})$ will be at least $Q/2$ for infinitely many primes $N$. However, almost surely, for $N$ lying in a set of relative density in primes equal to 1, the clique number $\omega(\Gamma_f^{(N)})$ would be between $(2-\epsilon)\log N$ and $(2+\epsilon)\log N$.*

We give the proof of this theorem in the last section.

## 4.2 Notation

Although most of the notation was implicitly defined in the introduction, we include it here for the reader's convenience. For sets $X, Y \subset \mathbb{Z}/N\mathbb{Z}$ we will denote their sumset by $X+Y$, so $X+Y = \{x+y\colon x \in X, y \in Y\}$. We will also consider their *restriced sumset* $X\widehat{+}Y = \{x+y\colon x \in X, y \in Y, x \neq y\}$. Apart from the standard $O$-notation, for functions $f, g$ on positive integers we will also write $g(n) = \Omega(f(n))$ if there exists $C > 0$ such that $|f(n)| \leq C|g(n)|$ for all large enough $n$. We will write $f(n) = \Theta(g(n))$ if there exist constants $c, C > 0$ such that $c|g(n)| \leq |f(n)| \leq C|g(n)|$ for all large enough $n$. $[A, B]$ will, depending on the context, denote the set of all integers $n$ such that $A \leq n \leq B$, and its image in $\mathbb{Z}/N\mathbb{Z}$. Finally, for an integrable function $g\colon \mathbb{Z} \to \mathbb{R}$ and $\theta \in [0,1]$, we define the corresponding Fourier coefficient by

$$\widehat{g}(\theta) = \sum_n g(n)e(-\theta n).$$

## 4.3 Independence

The purpose of this section is to somehow formalize the intuition that $(\frac{q}{p})$ are *independent* for different primes $q$, which was the motivation for our model. We note that most of the results from this section are already present in the literature (see e.g. [GS03, Proposition 9.1]), although possibly in a slightly different form.

Let $x$ be an integer, and $y$ an integer to be chosen later (one should think of $y$ as being substantially smaller than $x$). For each prime $p \leq x$, let $v_{(p)}$ be the vector of all $(\frac{q}{p})$ where $q$ runs over primes less than $y$ (all instances of $p$ and $q$ in this section will denote primes). We define the counting function $N$ by setting, for each $s \in \{-1,1\}^{\pi(y)}$,

$$N(s) = \#\{p \leq x\colon v_{(p)} = s\}.$$

Notice that this can also be expressed as

$$N(s) = 2^{-\pi(y)} \sum_{p \le x} \prod_{q \le y} \left(1 + (\tfrac{q}{p}) s_q\right). \tag{4.1}$$

Our aim is to prove that $\mathrm{Var}_s N(s)$ is small. Before doing that, we would like to mention that if $(v_{(p)})_{p \le x}$ were independent random vectors, each uniformly distributed on $\{-1,1\}^{\pi(y)}$ (and hence each with independent coordinates), we would have

$$\mathbb{E}_v \mathrm{Var}_s N(s) = \Theta(2^{-\pi(y)} x / \log x). \tag{4.2}$$

For comparison, this is much smaller than the theoretical maximum $O(2^{-\pi(y)} x^2/(\log x)^2)$.

Let $P_y = \prod_{q \le y} q$. Notice that we can identify the characters on $\{-1,1\}^{\pi(y)}$ with the set of divisors of $P_y$ by assigning to each divisor $m$ of $P_y$ the character $s \mapsto \prod_{q|m} s_q$.

By Parseval's formula we have

$$\mathrm{Var}_s N(s) = \mathbb{E}_s N(s)^2 - (\mathbb{E}_s N(s))^2 = \sum_{1 \ne m | P_y} |\widehat{N}(m)|^2. \tag{4.3}$$

Now, for fixed $m \ne 1$ from (4.1) we have

$$\begin{aligned} \widehat{N}(m) &= \mathbb{E}_s N(s) \prod_{q|m} s_q = 2^{-\pi(y)} \sum_{p \le x} \prod_{q|m} \mathbb{E}_{s_q}(s_q + (\tfrac{q}{p})) \prod_{q \nmid m} \mathbb{E}_{s_q}(1 + (\tfrac{q}{p}) s_q) \\ &= 2^{-\pi(y)} \sum_{p \le x} \prod_{q|m} (\tfrac{q}{p}). \end{aligned} \tag{4.4}$$

Assume for a moment that $m$ is even. Notice that by the quadratic reciprocity (Theorem 1.2.2) and the supplementary formula for $(\frac{2}{\cdot})$ (Proposition 1.2.3) we have

$$\prod_{q|m} (\tfrac{q}{p}) = (-1)^{\frac{p^2-1}{8}} \prod_{2 \ne q | m} (\tfrac{p}{q}) (-1)^{\frac{p-1}{2}} (-1)^{\frac{q-1}{2}}.$$

The conclusion is that there exists a Dirichlet character $\chi_m$ of modulus at most $4m$ such that the left-hand side is equal to $\chi_m(p)$ or $-\chi_m(p)$ with, of course, the same choice of $\pm$ sign for all $p$. We may assume that it is the former, since we will only be interested in bounding the absolute value of (4.4). The same conclusion follows for $m$ odd.

Our main tool in bounding $|\widehat{N}(m)|$ will be Proposition 1.2.1. Using it and the

well-known fact that $P_y = e^{O(y)}$, we get from (4.4) that for $m \neq 1$

$$|\widehat{N}(m)| \ll 2^{-\pi(y)} x^{1/2} (\log x + y).$$

Plugging this into (4.3) we get

$$\mathrm{Var}_s N(s) \ll 2^{-\pi(y)} x((\log x)^2 + y^2).$$

For $y = c \log x \log\log x$ this is comparable (within a power of $\log x$) to the situation one has in the random model described above (i.e. (4.2)).

We can now easily recover the conditional result of Montgomery [Mon71] mentioned in Section 4.1.

**Proposition 4.3.1.** *Suppose that the Generalised Riemann Hypothesis is true. Then there exists a constant $c$ and infinitely many primes $p$ such that the smallest quadratic nonresidue modulo $p$ is at least $c \log p \log\log p$.*

*Proof.* Let $y = c \log x \log\log x$, where $c$ is a constant to be chosen later. Obviously, it is enough to prove that for every large enough $x$ there exists a prime $p$ with the desired property between $x$ and $2x$. Moreover, by multiplicativity it is enough to prove that $(\frac{q}{p}) = 1$ for every prime $q \leq y$.

Using the same procedure as above, we can bound $\mathrm{Var}_s N'(s)$ where

$$N'(s) = \#\{x < p \leq 2x \colon v_{(p)} = s\}.$$

We get

$$\mathrm{Var}_s N'(s) \ll 2^{-\pi(y)} x((\log x)^2 + y^2).$$

If, on the other hand, $N'((1, \ldots, 1)) = 0$, then

$$\mathrm{Var}_s N'(s) \gg 2^{-3\pi(y)} x^2 / (\log x)^2.$$

However, these two bounds are incompatible if $c$ is small enough. □

## 4.4 Proof of the upper bound

We now prove the upper bound from Theorem 4.1.2. Throughout this section we will work with a fixed value of $\epsilon > 0$, and consider the cardinality $k = (2 + \epsilon) \log N$. Without loss of generality we may assume that $\epsilon$ is sufficiently small when needed.

We prove that with high probability there is no subset of $\mathbb{Z}/N\mathbb{Z}$ with $k$ elements which spans a clique in $\Gamma_f$. Obviously, this is equivalent to proving that

$$\mathbb{P}\left(\bigcup_{A\subset\mathbb{Z}/N\mathbb{Z}:\ |A|=k}\{f(A\widehat{+}A)=1\}\right)=o_{N\to\infty}(1).$$

It is evident from the definition of the function $f$, that our argument would split into two parts – one dealing with $[1,Q]$ which is number-theoretical and the other dealing with its complement which is more additive-combinatorial. Roughly speaking, for each set $A$ of size $k$ we will decide, using the following proposition, into which of these two cases it falls.

**Proposition 4.4.1** (Trichotomy)**.** *Let $L\geq 1$, $Q=L\log N$, $\delta>0$ and $A$ be a subset of $\mathbb{Z}/N\mathbb{Z}$ of size $k$. Then $A$ must satisfy at least one the following possibilities:*

(*i*) *(type 1 set) $|(A\widehat{+}A)\cap[1,Q]|\leq\delta|A\widehat{+}A|$.*

(*ii*) *(type 2 set) There exists $A'\subset A$ such that $|A'|\geq(1-\delta)|A|$ and $(A'\widehat{+}A')\cap[1,Q]=\emptyset$.*

(*iii*) *(type 3 set) There exist sets $C,D\subset A$, each contained in an interval of size $\log N$, such that $|C|,|D|\geq\frac{\delta^5}{6L^4}\log N$ and $C+D\subset[-Q,2Q]$.*

In the proof of this proposition we will need the following well known result. For the proof see the book by Tao and Vu [TV06].

**Proposition 4.4.2** (Plünnecke-Ruzsa)**.** *Let $A\subset\mathbb{Z}$ be a finite set such that $|A+A|\leq K|A|$ for some $K\geq 1$. Then for all nonnegative integers $m$ and $n$ we have*

$$|mA-nA|\leq K^{m+n}|A|.$$

We will also need the following result proven by Schoen [Sch02] which shows that we can easily shift between sumsets and restricted sumsets.

**Lemma 4.4.3.** *Let $B$ be a subset of $\mathbb{Z}/N\mathbb{Z}$ of size $l$. Then*

$$|B+B|=(1+o_{l\to\infty}(1))|B\widehat{+}B|.$$

*Proof of Proposition 4.4.1.* Suppose $A$ is not a type 1 set. Then

$$2\delta|A\widehat{+}A|\leq 2Q\leq L|A|,$$

and so by Lemma 4.4.3 we have $|A+A| \leq K|A|$ for $K = L/\delta$.

We now split $\mathbb{Z}/N\mathbb{Z}$ into disjoint intervals, each of length $\log N$ (except possibly one). First of all, we prove that only few of these intervals are hit by $A$. Denote this number by $h$ and choose points $a_1, \dots, a_h \in A$, one from each of these $h$ intervals. Notice that we can choose at least $\frac{h}{2L+2}$ of these points (we denote them by $a_1', \dots, a_l'$) such that the distance between any two of them is at least $Q+1$. Indeed, any translate of $[-Q, Q]$ intersects at most $2L+2$ intervals, and hence eliminates at most this many points. Obviously, translates $a_i' + (A\widehat{+}A) \cap [1, Q]$ for $i = 1, \dots, l$ are all disjoint and hence

$$|A + ((A\widehat{+}A) \cap [1, Q])| \geq \frac{h}{2L+2} \cdot \delta |A\widehat{+}A| \geq \frac{\delta h |A|}{3L}.$$

On the other hand, using Proposition 4.4.2 (Plünnecke-Ruzsa) we get

$$|A + ((A\widehat{+}A) \cap [1, Q])| \leq |A + A + A| \leq K^3 |A|.$$

Conclusion is that $h \leq 3K^3 L/\delta$.

Let $\beta = \frac{\delta^2}{3K^3 L}$. We call an interval *good* if it contains at least $\beta \log N$ elements from $A$. Let $A'$ be the intersection of $A$ with the union of all good intervals. Obviously,

$$|A \setminus A'| \leq h \cdot \beta \log N \leq \delta |A|$$

and hence $A$ is a type 2 set if $A'\widehat{+}A'$ is disjoint from $[1, Q]$. If, on the other hand, $(A'\widehat{+}A') \cap [1, Q] \neq \emptyset$, then there exist two good intervals $I_1$ and $I_2$ such that

$$((A \cap I_1)\widehat{+}(A \cap I_2)) \cap [1, Q] \neq \emptyset.$$

Let $C' = A \cap I_1$ and $D' = A \cap I_2$; by the previous line we have $C'\widehat{+}D' \subset [-Q, 2Q]$, and so the only thing left to do to prove that $A$ is a type 3 set is to replace the restricted sumset by a genuine sumset. Without loss of generality, we may assume that $|C'| \leq |D'|$. Let $C \subset C'$ be a set of $|C'|/2$ elements, and let $D \subset D'$ be a set $|D'|/2$ elements, disjoint from $C$. Then $C + D \subset C'\widehat{+}D'$ and we can conlude that $A$ is a type 3 set. □

Our main tool for dealing with type 1 and type 2 sets is the following proposition, the proof of which occupies the majority of [GM15].

**Theorem 4.4.4.** *Let $k = (2+\epsilon)\log N$. For every $m$ define*

$$S_k^m = \{A \subset \mathbb{Z}/N\mathbb{Z} \colon |A| = k \text{ and } |A\widehat{+}A| = m\}.$$

*Then*

$$|S_k^m| \leq 2^{(1-\epsilon^3)m}.$$

We note that in the proof of this theorem different strategies are used depending on the size of $m$. [GM15] deals with the cases $m = O(k)$ and $m = \Omega(k^2)$, whereas the claim for other values of $m$ was already proven in [Gre05a]. We should note that Theorem 4.4.4 was not explicitly stated as such in [GM15], but was implicitly proved in the course of proof of Theorem 1.2 (see pages 21-22 there).

We now shift the attention to our method for dealing with type 3 sets. The rough idea is to first prove that with high probability all of the Fourier coefficients of (the restriction of) function $f$ will be quite small. On the other hand, we will show that the existence of sets $C$ and $D$ as in the definition of type 3 sets, which additionally satisfy $f(C + D) = 1$, implies the existence of a large Fourier coefficient, so we will be able to conclude that this is quite unlikely to happen.

We will work with the function $g\colon \mathbb{Z} \to \mathbb{R}$ defined by

$$g(x) = \begin{cases} f(x \bmod N) & \text{for } -Q \leq x \leq 2Q, \\ 0 & \text{otherwise.} \end{cases}$$

The following proposition covers the first part of the strategy outlined above, namely that it is unlikely that $g$ has a large Fourier coefficient.

**Proposition 4.4.5.** *For any $l \leq Q$ we have*

$$\mathbb{P}\left(\sup_{0\leq\theta\leq 1} |\widehat{g}(\theta)| \geq l\right) \leq Q^{4+o(1)}/l^5.$$

*Proof.* Let $\psi \in [0, 1]$. We have

$$\begin{aligned} \mathbb{E}|\widehat{g}(\psi)|^4 &= \sum_{-Q\leq n_1,n_2,n_3,n_4\leq 2Q} \mathbb{E}g(n_1)g(n_2)g(n_3)g(n_4) \cdot e((n_1 + n_2 - n_3 - n_4)\psi) \\ &\leq \sum_{-Q\leq n_1,n_2,n_3,n_4\leq 2Q} |\mathbb{E}g(n_1)g(n_2)g(n_3)g(n_4)| \end{aligned}$$

Notice that the expectation appearing in the sum would be 0 unless the product of those $n_i$s that take values inside $[1, Q]$ is a square, and no number outside this interval is equal to an odd number of $n_i$s. We will call such quadruples *bad*. Obviously, there are $O(Q^2)$ bad quadruples for which all $n_i$ are outside $[1, Q]$. At the other extreme, consider bad quadruples for which all $n_i$ are inside $[1, Q]$. Their product is a square smaller than $Q^4$, which we can choose in $Q^2$ ways. Additionally, by the divisor bound

(Proposition 1.2.5) we can choose four of its divisors (that is, $n_i$s) in at most $Q^{o(1)}$ ways, giving in total $Q^{2+o(1)}$ bad quadruples. The remaining case is when two of the $n_i$s are equal to a number outside $[1, Q]$, and the product of the remaining two is a square. In the same way as before, we can see that there are at most $Q^{2+o(1)}$ such bad quadruples. We can now conclude that

$$\mathbb{E}|\widehat{g}(\psi)|^4 \leq Q^{2+o(1)}.$$

By Markov's inequality this gives us

$$\mathbb{P}\left(|\widehat{g}(\psi)| \geq l/2\right) \leq Q^{2+o(1)}/l^4. \tag{4.5}$$

Define

$$\theta_j = jl/80Q^2, \quad \text{for } j = 0, \ldots, \lfloor 80Q^2/l \rfloor.$$

For any $\theta \in [0,1]$ there is $j$ such that $|\theta - \theta_j| < l/80Q^2$. For such $j$ we have

$$|\widehat{g}(\theta_j) - \widehat{g}(\theta)| \leq \sum_{-Q \leq n \leq 2Q} |e((\theta_j - \theta)n) - 1| \leq l/2, \tag{4.6}$$

where the last inequality follows from $|e(\alpha) - 1| \leq 2\pi|\alpha|$ which holds for all real $\alpha$.

From (4.5), (4.6), and the union bound we have

$$\mathbb{P}\left(\sup_{0 \leq \theta \leq 1} |\widehat{g}(\theta)| \geq l\right) \leq \mathbb{P}\left(\max_j |\widehat{g}(\theta_j)| \geq l/2\right) \leq Q^{4+o(1)}/l^5,$$

and this is what we wanted to prove. □

The following proposition covers the second part of our strategy, namely the existence of a large Fourier coefficient. The proof is quite standard, but the brevity of the argument allows us to include it here for completeness.

**Proposition 4.4.6.** *Let $C$ and $D$ be two sets of integers such that $g(C + D) = 1$. Then*

$$\sup_{0 \leq \theta \leq 1} |\widehat{g}(\theta)| \geq |C|^{1/2}|D|^{1/2}.$$

*Proof.* From the given condition we have

$$\begin{aligned}
|C||D| &= \sum_{-Q\le x\le 2Q}\sum_{c\in C}\sum_{d\in D} g(x)1_{x=c+d} \\
&= \sum_{-Q\le x\le 2Q}\sum_{c\in C}\sum_{d\in D} g(x)\int e(\theta(x-c-d))\,d\theta \\
&= \int \overline{\widehat{g(\theta)}}\widehat{1_C}(\theta)\widehat{1_D}(\theta)\,d\theta \\
&\le \sup_{0\le\theta\le 1}|\widehat{g}(\theta)|\cdot\int |\widehat{1_C}(\theta)||\widehat{1_D}(\theta)|\,d\theta
\end{aligned}$$

By using Cauchy-Schwarz inequality and Parseval's identity we can bound this further by

$$\sup_{0\le\theta\le 1}|\widehat{g}(\theta)|\cdot\|\widehat{1_C}\|_{L^2([0,1])}\|\widehat{1_D}\|_{L^2([0,1])} = \sup_{0\le\theta\le 1}|\widehat{g}(\theta)|\cdot|C|^{1/2}|D|^{1/2},$$

and this gives the inequality from the statement. □

We now have all the tools needed and are in the position to start the proof.

*Proof of the upper bound in Theorem 4.1.2.* We use the notation introduced in Proposition 4.4.1. Let $L = Q/\log N$ and $\delta = \epsilon^4$. As we mentioned before, our aim is to prove that with high probability $\omega(\Gamma_f) < k = (2+\epsilon)\log N$. Obviously, this is equivalent to proving that

$$\mathbb{P}\left(\bigcup_{A:\ |A|=k}\{f(A\widehat{+}A)=1\}\right) = o(1). \tag{4.7}$$

First of all, we apply Proposition 4.4.1 which shows that every set $A$ of interest is of at least one of the types 1, 2, and 3. Denote these families of sets by $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$, respectively.

First we deal with type 1 sets which is quite easy. Indeed, using the notation introduced in this section, we have

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{A\in\mathcal{F}_1}\{f(A\widehat{+}A)=1\}\right) &\le \sum_{A\in\mathcal{F}_1}\mathbb{P}\left(f(A\widehat{+}A)=1\right) \le \sum_{m\ge k-1}\sum_{A\in S_k^m\cap\mathcal{F}_1}\mathbb{P}\left(f(A\widehat{+}A)=1\right) \\
&\le \sum_{m\ge k-1}|S_k^m|2^{-(1-\delta)m} \le \sum_{m\ge k-1}2^{-(\epsilon^3-\epsilon^4)m} = o(1).
\end{aligned} \tag{4.8}$$

Here we used Theorem 4.4.4 to obtain the last inequality.

Let $A$ now be a type 2 set. After possibly discarding some extra elements, we see that there is a subset $A' \subset A$ of size $k' = (1-\delta)(2+\epsilon)\log N > (2+\epsilon^2)\log N$ such that $A'\widehat{+}A'$ is disjoint from $[1, Q]$. This implies that

$$\begin{aligned}\mathbb{P}\left(\bigcup_{A\in\mathcal{F}_2}\{f(A\widehat{+}A)=1\}\right) &\leq \mathbb{P}\left(\bigcup_{|A'|=k'}\{f(A'\widehat{+}A')=1\}\right) \leq \sum_{|A'|=k'}\mathbb{P}\left(f(A'\widehat{+}A')=1\right)\\ &= \sum_{m\geq k'-1}\sum_{A\in S^m_{k'}}\mathbb{P}\left(f(A\widehat{+}A)=1\right) = \sum_{m\geq k'-1}|S^m_{k'}|\cdot 2^{-m}\\ &\leq \sum_{m\geq k-1} 2^{-\epsilon^6 m} = o(1). \qquad (4.9)\end{aligned}$$

Finally, we deal with type 3 sets. For these the situation is somewhat trickier since one can show that the expected number of type 3 sets that span a clique tends to infinity. Thus, the union bound (i.e. first moment method) used for type 1 and 2 sets wouldn't work.

Let $A$ be a type 3 set. By the definition there are subsets $C, D \subset A$, each of size $\frac{\delta^5}{6L^4}\log N$ such that $C + D \subset [-Q, 2Q]$ and both $C$ and $D$ are contained in intervals of size $\log N$. Because of the last property, after possible translations, we get two subsets of $\mathbb{Z}$ whose sumset is contained in $[-Q, 2Q]$. We will abuse the notation and denote these sets by $C$ and $D$ as well. By Proposition 4.4.6, we get

$$\sup_{0\leq\theta\leq 1}|\widehat{g}(\theta)| \geq \tfrac{\delta^5}{6L^4}\log N.$$

However, by Proposition 4.4.5, we know that the probability of this is quite small. Indeed, taking $l = \frac{\delta^5}{6L^4}\log N$, we have

$$\mathbb{P}\left(\sup_{0\leq\theta\leq 1}|\widehat{g}(\theta)| \geq \tfrac{\delta^5}{6L^4}\log N\right) \leq L^{25}/(\log N)^{1-o(1)}.$$

From this we can finally conclude that

$$\mathbb{P}\left(\bigcup_{A\in\mathcal{F}_3}\{f(A\widehat{+}A)=1\}\right) \leq L^{25}/(\log N)^{1-o(1)} = o(1), \qquad (4.10)$$

for $L = c\log\log N$. Combining (4.8), (4.9), and (4.10) we get (4.7).

□

## 4.5 Proof of the lower bound

In this section we prove the lower bound from Theorem 4.1.2. First of all, we would like to point out that the same bound also holds for random Cayley sum graphs. Although we are not aware of a proof of this fact in the literature, it is certainly a trivial consequence of the much more difficult result about the chromatic number by Green [Gre15]. In this section we will give a proof for our model which works also, with few easy modifications, for random Cayley sum graphs.

We fix $k = (2-\epsilon)\log N$. Let $\mathcal{U}$ be the family of all sets $A \subset [N/4, N/2]$ of size $k$ such that $|A\widehat{+}A| = \binom{k}{2}$. Let $R$ be the number of sets $A \in \mathcal{U}$ such that $f(A\widehat{+}A) = 1$. We will prove that with high probability $R \geq 1$ which, of course, proves that with high probability there is a clique of size $k$.

First we prove that almost all subsets of $[N/4, N/2]$ of size $k$ belong to $\mathcal{U}$.

**Lemma 4.5.1.** $|\mathcal{U}| = (1+o(1))\binom{N/4}{k}$.

*Proof.* Let $A \subset [N/4, N/2]$ be a random subset of size $k$. There are at most $(N/4)^3$ quadruples of different elements $x_1, x_2, x_3, x_4 \in [N/4, N/2]$ such that $x_1+x_2 = x_3+x_4$ and for each of these the probability that it is included in $A$ is $O(k^4/N^4)$. By the union bound, probability that $A$ contains at least one such quadruple is $O(k^4/N) = o(1)$ which is what we had to prove. □

By the previous lemma we have

$$\mathbb{E}R \geq \tfrac{1}{2}\binom{N/4}{k}2^{-\binom{k}{2}} \geq 2^{k\log N - k^2/2 + o(k^2)}. \tag{4.11}$$

We now focus on bounding the variance of $R$. For each $l \leq \binom{k}{2}$, let $\mathcal{V}_l$ be the family of all pairs of sets $A, B \in \mathcal{U}$ such that $|(A\widehat{+}A) \cap (B\widehat{+}B)| = l$. We have

$$\begin{aligned}
\mathrm{Var}R &= \sum_{l=0}^{\binom{k}{2}} \sum_{(A,B)\in\mathcal{V}_l} \mathrm{cov}(1_{f(A\widehat{+}A)=1}, 1_{f(B\widehat{+}B)=1}) \\
&\leq \sum_{l=1}^{\binom{k}{2}} |\mathcal{V}_l| \sup_{(A,B)\in\mathcal{V}_l} \mathbb{P}(f(A\widehat{+}A)=1, f(B\widehat{+}B)=1) \\
&= 2^{-k(k-1)} \sum_{l=1}^{\binom{k}{2}} |\mathcal{V}_l| 2^l.
\end{aligned} \tag{4.12}$$

Here the last equality holds because of our choice of family $\mathcal{U}$ which ensured that for every $A \in \mathcal{U}$ the set $A\widehat{+}A$ is disjoint from $[1, Q]$.

The next step is to bound $\mathcal{V}_l$. For each fixed $A \in \mathcal{U}$ we are going to bound the number of possible $B \in \mathcal{U}$ such that $|(A\widehat{+}A) \cap (B\widehat{+}B)| = l$.

Consider one such set. Let $s$ be the unique integer such that

$$\binom{s}{2} < l \leq \binom{s+1}{2}.$$

We prove that we can order the elements of $B$ and find a subset $B' \subset B$ of $s$ elements such that for each $b' \in B'$ there is a smaller $b \in B$ such that $b + b' \in A\widehat{+}A$. We can do this in the following way.

Form a graph with vertex set $B$ and edges joining $b$ and $b'$ if $b + b' \in A\widehat{+}A$. For each of the connected components $B_1, \ldots, B_r$ of this graph, choose an element $b_i \in B_i$. Now take an arbitrary order of elements of $B$ which satisfies the sole condition that if $b$ and $b'$ come from the same component $B_i$, and distance from $b$ to $b_i$ is less then distance from $b'$ to $b_i$, then $b < b'$. One can easily construct this by ordering component by component. Obviously, $b_1, \ldots, b_r$ are the only elements in this order that don't satisfy the required condition. Given that our graph has $k$ vertices and at least $l$ edges, the problem boils down to finding the maximal possible number of connected components in such a graph. It is intuitively clear (and one can easily show) that a maximizer for this optimization problem is a graph with $k-s$ connected components, all but one of which are singletons, and the remaining one contains $s+1$ vertices and $l$ edges. In this graph, we can take $B'$ to be equal to the nontrivial component, apart from the smallest element in it.

Using this observation, we make an enumerative argument as follows – we can choose set $A$ in at most $(N/4)^k$ ways, ordering in $k!$ ways, positions in the order that would be occupied by elements from $B'$ in $k^s$ ways, elements on positions outside $B'$ in $(N/4)^{k-s}$ and elements from $B'$ in $(k^3)^s$ ways (each element from $B'$ is uniquely determined by choosing one of at most $k^2$ possible sums from $A\widehat{+}A$ and one of at most $k$ predecessors from $B$). Putting all this together we get

$$|\mathcal{V}_l| \leq (N/4)^k k! k^s (N/4)^{k-s} (k^3)^s \leq 2^{(2k-s)\log N + o(k^2)}.$$

Since $l \leq \binom{k}{2}$ we have $s \leq k-1$ and so

$$l - s\log N \leq \binom{s+1}{2} - s\log N \leq (\tfrac{1}{2} - \tfrac{1}{2-\epsilon})k^2 + O(k),$$

and these two bounds together with (4.12) give

$$\mathrm{Var}R \leq 2^{2k\log N-(\frac{1}{2}+\frac{1}{2-\epsilon})k^2+o(k^2)}. \tag{4.13}$$

By Chebyshev's inequality, (4.11) and (4.13) we now have

$$\mathbb{P}(R=0) \leq \frac{\mathrm{Var}R}{(\mathbb{E}R)^2} = o(1), \tag{4.14}$$

which is what we wanted to prove.

## 4.6 Proof of Theorem 4.1.3

Theorem 4.1.3 is a straightforward consequence of Theorem 4.1.2 and the following strengthening of Borel-Cantelli lemma. For the proof of it, see e.g. the book by Chung [Chu01, page 83]

**Proposition 4.6.1** (Generalized second Borel-Cantelli lemma)**.** *Let $(A_n)_{n\geq 1}$ be a sequence of independent events such that $\sum_n \mathbb{P}(A_n) = \infty$. Then*

$$\mathbb{P}(A_n \textit{ happens for infinitely many } n) = 1$$

*and*

$$\lim_{m\to\infty} \frac{\sum_{n=1}^m 1_{A_n}}{\sum_{n=1}^m \mathbb{P}(A_n)} = 1 \quad \textit{almost surely.}$$

*Proof of Theorem 4.1.3.* Notice that

$$\begin{aligned}\mathbb{P}(\omega(\Gamma_f^{(N)}) \geq Q/2) &\geq \mathbb{P}(f([1,Q])=1) = \mathbb{P}\left(f(p)=1, \text{for all primes } p\in[1,Q]\right)\\ &= 2^{-\pi(Q)} \geq 1/N,\end{aligned}$$

where the last inequality follows from the prime number theorem and the assumption $Q \leq \log N \log\log N$. The first part of the claim from the theorem follows by Borel-Cantelli lemma. To prove the second part, for each prime $N$ define the event

$$A_N = \{(2-\epsilon)\log N < \omega(\Gamma_f^{(N)}) < (2+\epsilon)\log N)\}.$$

Since we have proven that $\mathbb{P}(A_N) = 1 - o(1)$, we obviously have that

$$\sum_{N\leq M, N\text{prime}} \mathbb{P}(A_N) = (1-o(1))\pi(M),$$

and the conclusion follows from Proposition 4.6.1. □

*Acknowledgements.* We would like to thank Ben Green for suggesting this project, Freddie Manners for helpful discussions, and Kannan Soundararajan for giving a reference for some of the results contained in Section 4.3. We would also like to thank Sean Eberhard for the idea of ordering the elements of $B$ when bounding $\mathrm{Var}R$ in Section 4.5.

# Chapter 5

# Additive triples of bijections, or the toroidal semiqueens problem

We prove an asymptotic for the number of additive triples of bijections $\{1,\dots,n\} \to \mathbb{Z}/n\mathbb{Z}$, that is, the number of pairs of bijections $\pi_1, \pi_2 \colon \{1,\dots,n\} \to \mathbb{Z}/n\mathbb{Z}$ such that the pointwise sum $\pi_1 + \pi_2$ is also a bijection. This problem is equivalent to counting the number of orthomorphisms or complete mappings of $\mathbb{Z}/n\mathbb{Z}$, to counting the number of arrangements of $n$ mutually nonattacking semiqueens on an $n \times n$ toroidal chessboard, and to counting the number of transversals in a cyclic Latin square. The method of proof is a version of the Hardy–Littlewood circle method from analytic number theory, adapted to the group $(\mathbb{Z}/n\mathbb{Z})^n$.

This chapter is based on the joint work with Sean Eberhard and Freddie Manners [EMM15].

## 5.1 Introduction

Let $S$ be the set of bijections $\{1,\dots,n\} \to \mathbb{Z}/n\mathbb{Z}$, thought of as a subset of the group $(\mathbb{Z}/n\mathbb{Z})^n$. We are interested in counting the number $s_n$ of additive triples in $S$, that is, the number of pairs of bijections $\pi_1, \pi_2 \colon \{1,\dots,n\} \to \mathbb{Z}/n\mathbb{Z}$ such that the pointwise sum $\pi_1 + \pi_2$ is also a bijection.

The number $s_n$ has been studied somewhat extensively, but under a different guise. Since $S$ is invariant under precomposition with permutations of $\{1,\dots,n\}$ it is easy to see by arbitrarily identifying $\{1,\dots,n\}$ with $\mathbb{Z}/n\mathbb{Z}$ that $s_n/n!$ is the number of permutations $\pi$ of $\mathbb{Z}/n\mathbb{Z}$ such that $x \mapsto \pi(x) - x$ is also a permutation. Such maps are called orthomorphisms or complete mappings, and afford the following fun interpretation. Define a semiqueen to be a chess piece which can move any distance

horizontally, vertically, or diagonally in the northeast-southwest direction. Then orthomorphisms represent ways of arranging $n$ mutually nonattacking semiqueens on an $n \times n$ toroidal chessboard. Thus $s_n/n!$ is the number of such arrangements.

In another guise this problem is that of counting the number of transversals of a cyclic Latin square. Here a transversal of an $n \times n$ Latin square is a set of $n$ squares with no two sharing the same row, column, or symbol, and the cyclic Latin square is the Latin square with $(i, j)$ entry given by $i + j \pmod n$. Then as above the number of such transversals is $s_n/n!$.

If $n$ is even then $s_n = 0$. Indeed, in this case for any bijection $\pi\colon \{1, \dots, n\} \to \mathbb{Z}/n\mathbb{Z}$ we have $\sum_{i=1}^{n} \pi(i) = n/2$, so the sum of two bijections is never again a bijection. If $n$ is odd then it is easy to see that $s_n > 0$, but estimating $s_n$ is not easy.

In 1991, Vardi [Var91] made a conjecture equivalent to the following.

**Conjecture 5.1.1.** There are constants $c_1 > 0$ and $c_2 < 1$ such that for any large enough odd number $n$ we have

$$c_1^n n!^2 \leq s_n \leq c_2^n n!^2.$$

The upper bound in this conjecture is known, and various authors have made incremental improvements to the constant $c_2$. Cooper and Kovalenko [CK96] showed that $c_2 = e^{-0.08854}$ is acceptable, and this was later improved by Kovalenko [Kov96] to $c_2 = 1/\sqrt{2}$ and by McKay, McLeod and Wanless [MMW06] to $c_2 = 0.614$. More recently, Taranenko [Tar15] proved that one can take $c_2 = 1/e + o(1)$. Glebov and Luria [GL15] proved the same bound using a somewhat simpler method based on entropy.

There has been much less progress on the lower bound, but nontrivial lower bounds for $s_n$ have been proved under various arithmetic assumptions about $n$. For example, Cooper [Coo00] proved that $s_n \geq e^{\frac{1}{2}\sqrt{n}\log n} n!$ under the hypothesis that $n$ has a divisor of size roughly $\sqrt{n}$, while if, say, $n$ is prime then a lower bound of Rivin, Vardi, and Zimmermann [RVZ94] for the torodial queens problem gives $s_n \geq 2^{\sqrt{(n-1)/2}} n!$. These lower bounds were recently superseded by work of Cavenagh and Wanless [CW10], which showed that $s_n \geq 3.246^n n!$ for all odd $n$. However, this lower bound is still a long way from the one in Conjecture 5.1.1.

Some researchers have also tried investigating the growth rate of $s_n$ numerically: see Cooper, Gilchrist, Kovalenko, and Novaković [CGKN99] and Kuznetsov [Kuz07, Kuz08, Kuz09]. Based on this numerical evidence, Wanless [Wan11] conjectured that

we can take $c_1, c_2 = 1/e + o(1)$; specifically he conjectured that

$$\lim_{n\to\infty} \frac{1}{n} \log(s_n/n!^2) = -1.$$

Finally, we should mention that $s_n/(n \cdot n!)$ is sequence A003111 in [Slo].

In this chapter we prove Vardi's conjecture with the optimal values $c_1, c_2 = 1/e + o(1)$, thus also confirming Wanless's conjecture. In fact our estimate is much more precise – we compute $s_n$ up to a factor of $1 + o(1)$.

**Theorem 5.1.2.** *Let $n$ be an odd integer. Then*

$$s_n = (e^{-1/2} + o(1))n!^3/n^{n-1}.$$

Perhaps the most surprising feature of this estimate is the appearance of the constant $e^{-1/2}$. This constant arises as the sum of the singular series in an argument resembling the circle method from analytic number theory, but it can be rationalized heuristically as follows. If $\pi_1, \pi_2 \colon \{1, \dots, n\} \to \mathbb{Z}/n\mathbb{Z}$ are random bijections then the sum $f = \pi_1 + \pi_2$ is something like a random function subject to $\sum_{x\in\mathbb{Z}/n\mathbb{Z}} f(x) = 0$, and so we might guess that the probability $\pi_1 + \pi_2$ is a bijection is about $n \cdot n!/n^n$. However if we fix two elements $x, y \in \mathbb{Z}/n\mathbb{Z}$ then the difference

$$f(x) - f(y) = (\pi_1(x) - \pi_1(y)) + (\pi_2(x) - \pi_2(y))$$

is the sum of two uniformly random nonzero elements, so $f(x) = f(y)$ with probability $1/(n-1)$, not $1/n$. Thus the probability that $f(x) \neq f(y)$ is smaller than we previously suggested by a factor of

$$\frac{1 - 1/(n-1)}{1 - 1/n} = 1 - 1/n^2 + O(1/n^3).$$

There are a total of $n^2/2 + O(n)$ pairs $x, y$, so we might guess that the probability $\pi_1 + \pi_2$ is a bijection is more like

$$\left(1 - 1/n^2\right)^{n^2/2} n \cdot n!/n^n \approx e^{-1/2} n!/n^{n-1}.$$

Our proof applies with only notational modifications when $\mathbb{Z}/n\mathbb{Z}$ is replaced by any abelian group $G$ of odd order. To be precise, given a finite abelian group $G$, let $s(G)$ be the number of pairs of bijections $\pi_1, \pi_2 : \{1, \dots, |G|\} \to G$ such that $\pi_1 + \pi_2$ is also a bijection. Then by the same method which proves Theorem 5.1.2, we have

the following theorem.

**Theorem 5.1.3.** *Let $G$ be a finite abelian group of odd order $n$. Then*

$$s(G) = (e^{-1/2} + o(1))n!^3/n^{n-1}.$$

There are additional complications however when $G$ is allowed to have even order, say when $G = (\mathbb{Z}/2\mathbb{Z})^d$, and we do not know whether the same method can be made to work in this case.

Like the toroidal semiqueens problem, the toroidal queens problem – the problem of counting the number of ways of arranging $n$ mutually nonattacking queens on an $n \times n$ toroidal chessboard – is equivalent to counting the number of solutions to a particular linear system in the set of bijections, namely the system

$$\pi_1 + \pi_2 = \pi_3,$$
$$\pi_1 - \pi_2 = \pi_4.$$

Linear systems of complexity 2 like this one are known not to be controlled by Fourier analysis alone, but there is some hope that one could use the higher-order theory pioneered by Gowers. This is an interesting avenue which has not yet been fully explored.

## 5.2 Notation

Although we have already implicitly introduced most of the notation and conventions, we include them here for the reader's convenience. We use a primed sum $\sum'_{s_1,\dots,s_k \in S}$ to denote the sum over all $k$-tuples of distinct elements $s_1, \dots, s_k \in S$. All the groups we will work with will be finite, and we equip each of them with the either uniform probability or counting measure, depending on whether we are working on the physical or the frequency side, respectively. Of course, this convention is respected in our definitions of convolution, inner product and $L^2$-norm.

*Acknowledgements.* We would like to thank Ben Green for communicating the problem of estimating $s_n$, Robin Pemantle for a discussion in connection with the method used to prove Theorem 5.5.1, and Fernando Shao for useful conversations.

## 5.3 Outline of the proof

Our approach to proving Theorem 5.1.2 is Fourier-analytic. Write $G = \mathbb{Z}/n\mathbb{Z}$ for $n$ an odd integer, and write $S \subset G^n$ for the set of bijections $\{1, \dots, n\} \to G$. Our goal is to compute the quantity

$$\mathbb{E}_{x,y \in G^n} 1_S(x) 1_S(y) 1_S(x+y) = \langle 1_S * 1_S, 1_S \rangle.$$

A standard application of Parseval's identity shows that this quantity can be expressed in terms of the Fourier transform of $1_S$ as

$$\sum_{\chi \in \widehat{G^n}} |\widehat{1_S}(\chi)|^2 \widehat{1_S}(\chi). \tag{5.1}$$

Here we identify $\widehat{G}$ with $\frac{1}{n}\mathbb{Z}/\mathbb{Z}$ in the usual way, so that for $\chi = (r_1, \dots, r_n) \in (\frac{1}{n}\mathbb{Z}/\mathbb{Z})^n$ we have explicitly

$$\widehat{1_S}(r_1, \dots, r_n) = \frac{1}{n^n} \sum_{x_1, \dots, x_n}{}' e(-r_1 x_1 - \cdots - r_n x_n), \tag{5.2}$$

where, as mentioned in the previous section, we use $\sum'$ to denote the sum over distinct $x_1, \dots, x_n \in G$, i.e., the sum over $S$. In fact it is clear that $\widehat{1_S}$ is real-valued, since $S = -S$, and hence we may drop the absolute value signs in (5.1).

The form of the proof can then be viewed as an analogue of the Hardy–Littlewood circle method in analytic number theory, adapted to the group $G^n$ rather than $\mathbb{Z}$. We borrow some nomenclature from the classical setting.

- There are a small number of characters $\chi \in \widehat{G^n}$, namely the characters $\chi = (r_1, \dots, r_n)$ for which almost all of the $r_i$ are equal, that make a substantial contribution to the sum (5.1). We call the totality of these $\chi$ the *major arcs*. We compute explicitly the contribution of these $\chi$ to (5.1), up to small errors: this is an analogue of the singular series, which accounts for the main term in Theorem 5.1.2, including the bizarre constant $e^{-1/2}$. For all this see Section 5.4.

- We bound the contribution from all other characters using the triangle inequality; that is, we obtain an upper bound for

$$\sum_{\text{all other } \chi} |\widehat{1_S}(\chi)|^3$$

that is smaller than the main term. We call such $\chi$ collectively the *minor arcs*.

A large part of this chapter is therefore devoted to obtaining good bounds for Fourier coefficients $\widehat{1_S}(\chi)$, either pointwise or on average in a suitable sense. In fact we will need to combine several different arguments which are effective in different regimes.

### 5.3.1 Preliminaries on $\widehat{1_S}$

In order to describe how the characters $\widehat{G}^n$ are divided up into different pieces in which different approaches will be effective, we will need some preliminary remarks.

First, note that the function $\widehat{1_S}(r_1, \dots, r_n)$ is invariant under permutation of the $r_i$:

$$\widehat{1_S}(r_1, \dots, r_n) = \widehat{1_S}(r_{\sigma(1)}, \dots, r_{\sigma(n)})$$

for any bijection $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$. This is immediate from the definition (5.2). Hence it makes sense to specify a Fourier coefficient of interest in the form $\widehat{1_S}(r_1^{a_1}, \dots, r_k^{a_k})$, where $r_i \in \widehat{G}$ are distinct, the $a_i$ are positive integers such that $\sum a_i = n$, and the notation $r^a$ means $r$ repeated $a$ times.

Second, observe that $\widehat{1_S}(r_1, \dots, r_n) = 0$ unless $\sum r_i = 0$. This follows from the fact that $S$ is invariant under global shifts $(x_i) \mapsto (x_i + t)$, so one can compute

$$\begin{aligned}
\widehat{1_S}(r_1, \dots, r_n) &= \frac{1}{n^n} \sum_{x_1, \dots, x_n}{}' e(-r_1 x_1 - \cdots - r_n x_n) \\
&= \frac{1}{n^n} \sum_{x_1, \dots, x_n}{}' e(-r_1(x_1 + t) - \cdots - r_n(x_n + t)) \\
&= \widehat{1_S}(r_1, \dots, r_n) e(-(r_1 + \cdots + r_n)t).
\end{aligned}$$

Dually, note that $\widehat{1_S}$ is invariant under global shifts. This follows from the fact that $\sum x_i = 0$ for any $(x_i) \in S$. Hence,

$$\begin{aligned}
\widehat{1_S}(r_1, \dots, r_n) &= \frac{1}{n^n} \sum_{x_1, \dots, x_n}{}' e(-r_1 x_1 - \cdots - r_n x_n) \\
&= \frac{1}{n^n} \sum_{x_1, \dots, x_n}{}' e(-(r_1 + t)x_1 - \cdots - (r_n + t)x_n) \\
&= \widehat{1_S}(r_1 + t, \dots, r_n + t).
\end{aligned}$$

Finally, note the trivial bound

$$|\widehat{1_S}(\chi)| \leq \widehat{1_S}(0) = \frac{n!}{n^n}.$$

### 5.3.2 Entropy ranges for minor arcs

We now explain a straightforward but still fairly powerful bound on $|\widehat{1_S}(\chi)|$ that follows directly from these elementary considerations.

**Proposition 5.3.1** ("Entropy bound")**.** *We have*

$$|\widehat{1_S}(r_1^{a_1}, \dots, r_k^{a_k})| \leq \binom{n}{a_1, \dots, a_k}^{-1/2} \left(\frac{n!}{n^n}\right)^{1/2}.$$

*Proof.* Let $\mathcal{O} \subset \widehat{G}^n$ denote the set of characters obtained by permuting the elements of $(r_1^{a_1}, \dots, r_k^{a_k})$. Hence, $|\mathcal{O}| = \binom{n}{a_1,\dots,a_k}$ and by permutation invariance, $\widehat{1_S}$ takes a constant value on $\mathcal{O}$. Thus by Parseval,

$$\begin{aligned}\frac{n!}{n^n} &= \sum_{\chi \in \widehat{G}^n} |\widehat{1_S}(\chi)|^2 \\ &\geq \sum_{\chi \in \mathcal{O}} |\widehat{1_S}(\chi)|^2 \\ &= |\mathcal{O}||\widehat{1_S}(r_1^{a_1}, \dots, r_k^{a_k})|^2,\end{aligned}$$

and the result follows. □

The term "entropy bound" refers to the fact that the quantity

$$H(\chi) = \frac{1}{n} \log \binom{n}{a_1, \dots, a_k}$$

is roughly the entropy $\sum_{i=1}^k (a_i/n) \log(n/a_i) \in [0, \log n]$ of a random variable taking the value $r_i$ with probability $a_i/n$. In a slight abuse of nomenclature, we refer to this first quantity $H(\chi)$ as the *entropy of* $\chi$. In this language, the bound of Proposition 5.3.1 is precisely $\exp(-Hn/2)(n!/n^n)^{1/2}$ or roughly $\exp(-Hn/2 - n/2)$.

This is already sufficient to control the contribution to (5.1) for most characters $\chi$.

**Corollary 5.3.2.** *We have*

$$\sum_{\chi\colon H(\chi)\geq R} |\widehat{1_S}(\chi)|^3 \leq \exp((3-R)n/2)\left(\frac{n!}{n^n}\right)^3.$$

*Proof.* By Parseval and Proposition 5.3.1,

$$\begin{aligned}
\sum_{\chi\colon H(\chi)\geq R} |\widehat{1_S}(\chi)|^3 &\leq \left(\sum_{\chi\colon H(\chi)\geq R} |\widehat{1_S}(\chi)|^2\right) \sup_{\chi\colon H(\chi)\geq R} |\widehat{1_S}(\chi)| \\
&\leq \left(\sum_{\chi\in\widehat{G^n}} |\widehat{1_S}(\chi)|^2\right) \sup_{\chi\colon H(\chi)\geq R} \exp(-H(\chi)n/2)\left(\frac{n!}{n^n}\right)^{1/2} \\
&\leq \exp(-Rn/2)\left(\frac{n!}{n^n}\right)^{3/2} \\
&\leq \exp((3-R)n/2)\left(\frac{n!}{n^n}\right)^3
\end{aligned}$$

as required, where we have used the estimate $n! > (n/e)^n$. □

This is smaller than the main term for any fixed $R > 3$, so from now on such high-entropy characters need not concern us. We refer to such characters as the *high-entropy minor arcs*. Since a typical character $\chi$ has entropy comparable to $\log n$, the high-entropy case comprises almost all characters in some sense, so this is a good first step.

On the opposite extreme we have characters $\chi$ with entropy $o(1)$, such as $\chi = \left(r^k, -r^k, 0^{n-2k}\right)$ for $k = o(n)$. The characters with entropy $O\left(\frac{\log n}{n}\right)$ are the major arcs, but between $O\left(\frac{\log n}{n}\right)$ and $o(1)$ we have the *low-entropy minor arcs*. Such characters are necessarily of the form $(\ldots, r^{n-o(n)})$, i.e., they take a single value almost all time. By the remarks above we may shift without loss of generality to assume that this value $r$ is zero. Thus the low-entropy minor arcs are closely related to the set of *sparse* characters, characters $\chi$ comprised almost entirely of zeros.

That leaves the characters $\chi$ with $o(1) \leq H(\chi) \leq 3$, which form the *medium-entropy minor arcs*. A good model case are characters such as $\chi = (r^{n/3}, -r^{n/3}, 0^{n/3})$, which are particularly troublesome.

There are two fundamental areas of inefficiency in the arguments of Proposition 5.3.1 and Corollary 5.3.2 that prevent them from giving good bounds for low- or medium-entropy minor arcs. First, the bound on $|\widehat{1_S}|$ given by Proposition 5.3.1 is less effective for smaller $H$, and in particular worse than trivial when $H < 1$.

Clearly there is no hope for this programme unless we can find a bound on $|\widehat{1_S}|$ that is nontrivial throughout the range $0 < H < 1$, and moreover obtains a substantial exponential saving for most of that range.

Second, in the proof of Corollary 5.3.2 we made use of the convenient bound

$$\sum_{\chi \in X} |\widehat{1_S}(\chi)|^3 \leq \left( \sum_{\chi \in \widehat{G}} |\widehat{1_S}(\chi)|^2 \right) \sup_{\chi \in X} |\widehat{1_S}(\chi)|$$

for some appropriate set $X \subset \widehat{G}^n$. It is fairly clear that we cannot afford this luxury for small entropies: since the main term has order $(n!/n^n)^3$, we need $|\widehat{1_S}(\chi)| = o((n!/n^n)^2)$ for all $\chi \in X$ for this to be effective. This is a saving of about $\exp(-n)$ over the trivial bound of $n!/n^n$, and it is simply not reasonable to expect such a bound to hold if, say, $H(\chi) = 1/1000$. One way around this is to pigeonhole the characters $\chi$ into various sets $X_i$, and apply such an $L^2 \cdot L^\infty$ bound on each set. To make this work, we would need a bound on

$$\sum_{\chi \in X_i} |\widehat{1_S}(\chi)|^2$$

which makes a significant saving over the crude estimate

$$\sum_{\chi \in X_i} |\widehat{1_S}(\chi)|^2 \leq \sum_{\chi \in \widehat{G}^n} |\widehat{1_S}(\chi)|^2$$

employed in Corollary 5.3.2. Alternatively, one could try to bound the $L^3$ sum

$$\sum_{\chi \in X_i} |\widehat{1_S}(\chi)|^3$$

directly, using an $L^\infty$ bound and an estimate for the size of $X_i$.

We address both of these inefficiencies in order to reach our final bound. Specifically, we will do each of the following.

- Improving on Proposition 5.3.1, we obtain a good general-purpose bound for $|\widehat{1_S}(\chi)|$ that is nontrivial for values of $H(\chi)$ approaching zero. Roughly speaking, where Proposition 5.3.1 gives the bound $e^{-Hn/2-n/2}$ we will prove the bound $e^{-Hn/2-n}$. This appears in Section 5.5.

- We use this bound to estimate the contribution to (5.1) from $\chi$ with $10^{-10} < H(\chi) < 10$, say, by using a dyadic decomposition and a simple estimate for the number of characters of a given entropy. Thus we dispatch the medium-entropy

minor arcs. This is covered in Section 5.8.

- Separately, we obtain a good estimate for

$$\sum_{m\text{-sparse } \chi} |\widehat{1_S}(\chi)|^2,$$

  i.e., an improvement to Parseval when summing only over those characters $\chi$ having exactly $m$ nonzero terms. Here we might imagine $m \leq n/1000$. We call this a *sparseval* bound, and prove this in Section 5.6.

- Finally, we obtain a slightly different $L^\infty$ bound specialized to the case of sparse characters $\chi$. This appears in Section 5.7. The total contribution from the low-entropy minor arcs is controlled by combining this with the $L^2$ bound above.

In summary, we split the universe of all $\chi$ into several slightly overlapping regions – major arcs, low-entropy minor arcs, medium-entropy minor arcs, and high-entropy minor arcs – and apply a cocktail of different bounds and explicit computations adapted to each region.

### 5.3.3 Interpretations of minor arc bounds

As we have said, a significant part of our effort will be spent proving nontrivial bounds on Fourier coefficients $|\widehat{1_S}(\chi)|$ for $\chi$ in the minor arcs. Although such bounds are at first glance fairly esoteric, in fact they have natural interpretations. For instance, they are intimately related to questions of the following flavour.

**Question 5.3.3.** Suppose the elements of $G = \mathbb{Z}/n\mathbb{Z}$ are divided into $k$ sets $A_1, \dots, A_k$, each of a pre-determined size $a_i \approx n/k$, at random. Consider the random variables

$$T_i = \sum_{r \in A_i} r \in G$$

for $i = 1, \dots, k$. The vector $T = (T_1, \dots, T_k)$ is a random variable taking values in the subgroup $H = \{(x_i) \in G^k \colon \sum x_i = 0\}$ of $G^k$. How close is $T$ to being equidistributed on $H$, in a quantitative sense? In other words how close is the law of $T$ to the uniform distribution on $H$?

Because of the connection of this question to the size of the Fourier coefficients $\widehat{1_S}(r_1^{a_1}, \dots, r_k^{a_k})$, our bounds give a strong answer to this question in many regimes. It is not inconceivable that these kind of bounds may find applications elsewhere.

## 5.4 Major arcs

In this section we compute $\widehat{1_S}(r_1, \dots, r_m, 0, \dots, 0)$ explicitly whenever $m$ is bounded. To this end observe that

$$\widehat{1_S}(r_1, \dots, r_m, 0, \dots, 0) = \frac{(n-m)!}{n^n} \sum_{x_1,\dots,x_m}{}' e(-r_1x_1 - \cdots - r_mx_m). \tag{5.3}$$

The sum over distinct $x_1, \dots, x_m$ can be related to a sum over partitions of $\{1, \dots, m\}$ using a type of Möbius inversion: for any function $F(x_1, \dots, x_m)$ we have

$$\sum_{x_1,\dots,x_m}{}' F(x_1, \dots, x_m) = \sum_{\mathcal{P}} \mu(\mathcal{P}) \sum_{\substack{x_1,\dots,x_m \\ x_i = x_j \text{ whenever } i \overset{\mathcal{P}}{\sim} j}} F(x_1, \dots, x_m),$$

where the outer sum runs over partitions $\mathcal{P}$ of $\{1, \dots, m\}$, and

$$\mu(\mathcal{P}) = (-1)^{m-|\mathcal{P}|} \prod_{P \in \mathcal{P}} (|P| - 1)!.$$

See, e.g., [Bón15] for more information. To apply this to $\widehat{1_S}$, say that a partition $\mathcal{P}$ of $\{1, \dots, m\}$ *kills* $(r_1, \dots, r_m)$ if $\sum_{i \in P} r_i = 0$ for each $P \in \mathcal{P}$, and observe that

$$\sum_{\substack{x_1,\dots,x_m \\ x_i = x_j \text{ whenever } i \overset{\mathcal{P}}{\sim} j}} e(-r_1x_1 - \cdots - r_mx_m) = \begin{cases} n^{|\mathcal{P}|} & \text{if } \mathcal{P} \text{ kills } (r_1, \dots, r_m), \\ 0 & \text{otherwise,} \end{cases}$$

so

$$\widehat{1_S}(r_1, \dots, r_m, 0, \dots, 0) = \frac{(n-m)!}{n^n} \sum_{\mathcal{P} \text{ killing } (r_1,\dots,r_m)} \mu(\mathcal{P}) n^{|\mathcal{P}|}. \tag{5.4}$$

The following two observations are immediate from this formula:

- Suppose every killing partition of $(r_1, \dots, r_m)$ has at most $k$ parts. Then

$$|\widehat{1_S}(r_1, \dots, r_m, 0, \dots, 0)| \leq O_m\left(\frac{1}{n^{m-k}} \frac{n!}{n^n}\right). \tag{5.5}$$

- Suppose that $m$ is even, that $r_1, \dots, r_m \in \widehat{G}$ are nonzero, and that $(r_1, \dots, r_m)$ is killed by a unique partition with $m/2$ parts. In other words suppose that $r_1, \dots, r_m \in \widehat{G}$ are distinct and nonzero, and that, up to a permutation, $r_{2j} =$

$-r_{2j-1}$ for each $j = 1, \dots, m/2$. Then

$$\widehat{1_S}(r_1, \dots, r_m, 0, \dots, 0) = \frac{(-1)^{m/2}}{n^{m/2}} \frac{n!}{n^n} + O_m\left(\frac{1}{n^{m/2+1}} \frac{n!}{n^n}\right). \tag{5.6}$$

**Proposition 5.4.1.** *If $m$ is even then*

$$\sum_{m\text{-sparse } \chi} \widehat{1_S}(\chi)^3 = \frac{(-1)^{m/2}}{2^{m/2}(m/2)!} \frac{n!^3}{n^{3n}} + O_m\left(\frac{1}{n} \frac{n!^3}{n^{3n}}\right),$$

*while if $m$ is odd then*

$$\sum_{m\text{-sparse } \chi} \widehat{1_S}(\chi)^3 = O_m\left(\frac{1}{n} \frac{n!^3}{n^{3n}}\right).$$

*Proof.* First of all note by permutation-invariance that

$$\sum_{m\text{-sparse } \chi} \widehat{1_S}(\chi)^3 = \binom{n}{m} \sum_{r_1, \dots, r_m \neq 0} \widehat{1_S}(r_1, \dots, r_m, 0, \dots, 0)^3. \tag{5.7}$$

For each $(r_1, \dots, r_m)$ choose a maximal-size partition $\mathcal{P}$ which kills $(r_1, \dots, r_m)$, and split the sum up according to $\mathcal{P}$. Suppose $\mathcal{P}$ has $k$ parts. Since $\mathcal{P}$ cannot have singletons we have $k \leq m/2$, and by (5.5) we have

$$|\widehat{1_S}(r_1, \dots, r_m, 0, \dots, 0)| \leq O_m\left(\frac{1}{n^{m-k}} \frac{n!}{n^n}\right).$$

Since the number of $(r_1, \dots, r_m)$ killed by $\mathcal{P}$ is bounded by $n^{m-k}$, the contribution to (5.7) from these $(r_1, \dots, r_m)$ is bounded by

$$\binom{n}{m} n^{m-k} \left(\frac{O_m(1)}{n^{m-k}} \frac{n!}{n^n}\right)^3 = O_m\left(\frac{1}{n^{m-2k}} \frac{n!^3}{n^{3n}}\right),$$

which is satisfactory unless $k = m/2$. Moreover the number of $(r_1, \dots, r_m)$ killed by at least two partitions $\mathcal{P}$ with $m/2$ parts is bounded by $n^{m/2-1}$, so the contribution from these $(r_1, \dots, r_m)$ is bounded by

$$\binom{n}{m} n^{m/2-1} O_m\left(\frac{1}{n^{m/2}} \frac{n!}{n^n}\right)^3 = O_m\left(\frac{1}{n} \frac{n!^3}{n^{3n}}\right),$$

which is again satisfactory. Thus we may restrict our attention to those $(r_1, \dots, r_m)$

which are killed by a unique partition $\mathcal{P}$ with $m/2$ parts, and for these (5.6) gives

$$\widehat{1_S}(r_1, \dots, r_m, 0, \dots, 0) = \frac{(-1)^{m/2}}{n^{m/2}} \frac{n!}{n^n} + O_m\left(\frac{1}{n^{m/2+1}} \frac{n!}{n^n}\right).$$

For a fixed such $\mathcal{P}$ the number of such $(r_1, \dots, r_m)$ is $n^{m/2} + O_m(n^{m/2-1})$, and the number of such $\mathcal{P}$ is

$$(m-1)(m-3)\cdots 1 = \frac{m!}{2^{m/2}(m/2)!},$$

so the total contribution from these $(r_1, \dots, r_m)$ is

$$\begin{aligned}\binom{n}{m} \frac{m!}{2^{m/2}(m/2)!}(n^{m/2} + O_m(n^{m/2-1})) \left(\frac{(-1)^{m/2}}{n^{m/2}} \frac{n!}{n^n} + O_m\left(\frac{1}{n^{m/2+1}} \frac{n!}{n^n}\right)\right)^3 \\ = \frac{(-1)^{m/2}}{2^{m/2}(m/2)!} \frac{n!^3}{n^{3n}} + O_m\left(\frac{1}{n} \frac{n!^3}{n^{3n}}\right).\end{aligned}$$

This proves the proposition. □

## 5.5 Square-root cancellation for general Fourier coefficients

The aim of this section is to prove the following refinement of Proposition 5.3.1.

**Theorem 5.5.1.** *Suppose $\chi = (r_1^{a_1}, \dots, r_k^{a_k})$, where $\sum_{i=1}^k a_i = n$. Then*

$$|\widehat{1_S}(\chi)| \leq \binom{n+k-1}{k-1}^{1/2} \binom{n}{a_1, \dots, a_k}^{-1/2} \frac{n!}{n^n}.$$

**Remark 5.5.2.**

- The terms depending on $k$ are easily seen to be a small error in the regimes we care about. Indeed, if $H(\chi) = O(1)$ then $\chi$ can take at most $O(n/\log n) = o(n)$ distinct values, and the additional term is then $\exp(o(n))$, and thus small compared to the exponential entropy saving.
- The saving over Proposition 5.3.1 is an apparently modest factor of around $e^{-n/2+o(n)}$. This may be uninspiring if $H$ is large, but if $H$ is small it is decisive.
- In general this bound is not sharp, often by a substantial amount. However, it is sometimes attained asymptotically for very special, arithmetically structured

classes of characters. For instance, if we temporarily allow $n$ to be even, then one can compute directly that

$$\log\left[|\widehat{1_S}((1/2)^a, 0^{n-a})|/(n!/n^n)\right] \sim -\frac{1}{2}\log\binom{n}{a}$$

for large $n$. Similarly, numerical evidence strongly suggests that

$$\log\left[|\widehat{1_S}((1/3)^a, (-1/3)^a, 0^{n-2a})|/(n!/n^n)\right] \sim -\frac{1}{2}\log\binom{n}{a, a, n-2a}$$

if $n$ is divisible by 3; for example, when $n = 3003$,

$$\log\left[|\widehat{1_S}((1/3)^{1001}, (-1/3)^{1001}, 0^{1001})|/(n!/n^n)\right] = -1649.01782245...,$$

while

$$-\frac{1}{2}\log\binom{3003}{1001, 1001, 1001} = -1645.46757758....$$

But, for example, the evidence also suggests that

$$\log\left[|\widehat{1_S}((1/3)^a, 0^{n-a})|/(n!/n^n)\right] \sim -\frac{2}{3}\log\binom{n}{a}.$$

From the above remarks, we see that one might hope to improve on Theorem 5.5.1, but only by ruling out the particular characters discussed above and others similar to them. For instance, a strengthening is almost certainly true under the assumption that $n$ is prime. Thus one might think of Theorem 5.5.1 as the best general-purpose bound available.

We should also comment briefly on the term "square-root cancellation". If $\chi = (r_1^{a_1}, \ldots, r_k^{a_k})$ then

$$\widehat{1_S}(\chi) = \frac{n!}{n^n} \mathop{\mathbb{E}}_{\text{partitions } \mathcal{P}} e\left(-r_1\left(\sum_{x\in P_1} x\right) - \cdots - r_k\left(\sum_{x\in P_k} x\right)\right)$$

where the expectation is over all ordered partitions $\mathcal{P} = (P_1, \ldots, P_k)$ of $G$ into $k$ pieces with $|P_i| = a_i$. The number of such partitions is precisely $\binom{n}{a_1,\ldots,a_k}$; assuming that the phases in the average behave randomly then we should expect square-root cancellation, and we heuristically recover the bound in Theorem 5.5.1 up to lower-order terms.

*Proof of Theorem 5.5.1.* Let $\gamma$ be the Gaussian measure on $\mathbb{C}^k$ defined with respect

to Lebesgue measure $\lambda$ by

$$\frac{d\gamma}{d\lambda} = \frac{1}{\pi^k} \exp\left( -\sum_{i=1}^{k} |z_i|^2 \right).$$

We claim that

$$n^n \widehat{1_S}(\chi) = \int_{\mathbb{C}^k} z_1^{a_1} \cdots z_k^{a_k} \prod_{x \in G} \left( \sum_{i=1}^{k} e(-r_i x) \overline{z_i} \right) d\gamma. \tag{5.8}$$

One can check this by expanding the product and using the identity

$$\int_{\mathbb{C}^k} \prod_{i=1}^{k} z_i^{a_i} \overline{z_i}^{b_i} \, d\gamma = \begin{cases} \prod_{i=1}^{k} a_i! & \text{if } a_i = b_i \text{ for each } i, \\ 0 & \text{otherwise.} \end{cases} \tag{5.9}$$

We can now proceed by bounding the integral in (5.8) using various techniques. Applying the Cauchy–Schwarz inequality we bound the right hand side by

$$\left( \int_{\mathbb{C}^k} |z_1|^{2a_1} \cdots |z_k|^{2a_k} \, d\gamma \right)^{1/2} \left( \int_{\mathbb{C}^k} \prod_{x \in G} \left| \sum_{i=1}^{k} e(-r_i x) \overline{z_i} \right|^2 d\gamma \right)^{1/2}. \tag{5.10}$$

The first factor here is exactly $\left( \prod_{i=1}^{k} a_i! \right)^{1/2}$ by (5.9). If we now apply the AM–GM inequality to the second term and evaluate the resulting integral, we get

$$\begin{aligned} \int_{\mathbb{C}^k} \prod_{x \in G} \left| \sum_{i=1}^{k} e(-r_i x) \overline{z_i} \right|^2 d\gamma &\le \int_{\mathbb{C}^k} \left( \frac{1}{n} \sum_{x \in G} \left| \sum_{i=1}^{k} e(-r_i x) \overline{z_i} \right|^2 \right)^n d\gamma \\ &= \int_{\mathbb{C}^k} \left( \sum_{i=1}^{k} |z_i|^2 \right)^n d\gamma \\ &= \sum_{s_1 + \cdots + s_k = n} \binom{n}{s_1, \ldots, s_k} \int_{\mathbb{C}^k} |z_1|^{2s_1} \cdots |z_k|^{2s_k} \, d\gamma \\ &= n! \sum_{s_1 + \cdots + s_k = n} 1 \\ &= n! \binom{n+k-1}{k-1}. \end{aligned}$$

Thus the theorem follows by dividing (5.10) by $n^n$. □

**Remark 5.5.3.** Though the identity (5.8) is easily verified directly, we briefly sketch

here a possible route by which one might be motivated to write down such a formula in the first place.

Let $\{X_x \colon x \in G\}$ be indeterminates, and observe that the left hand side of (5.8) is precisely the coefficient of $\prod_{x \in G} X_x$ in the multivariate polynomial

$$\prod_{i=1}^{k} \left( \sum_{x \in G} X_x e(-r_i x) \right)^{a_i},$$

and so equivalently the value of the derivative

$$\left( \prod_{x \in G} \frac{\partial}{\partial X_x} \right) \prod_{i=1}^{k} \left( \sum_{x \in G} X_x e(-r_i x) \right)^{a_i} \tag{5.11}$$

evaluated at zero (although in fact this evaluation is redundant as this expression is a constant function).

In some generality, it is possible to equate a quantity of the form

$$\frac{\partial^{b_1}}{\partial Y_1^{b_1}} \cdots \frac{\partial^{b_k}}{\partial Y_k^{b_k}} f \big|_{Y_1 = \cdots = Y_k = 0},$$

where $f(Y_1, \ldots, Y_k)$ is a suitable holomorphic function of $k$ complex variables, with the corresponding integral expression

$$\int_{\mathbb{C}^k} \overline{Y_1^{b_1}} \ldots \overline{Y_k^{b_k}} \, f(Y_1, \ldots, Y_k) \, d\gamma,$$

where again $\gamma$ denotes Gaussian measure. This can be verified by applying the case $k = 1$ iteratively, which in turn follows by averaging the usual Cauchy integral formula over different radii. This formula can be considered a particular higher-dimensional variant of the Cauchy integral formula.

The identity (5.8) follows by applying this identity to (5.11) and making an orthogonal change of variables.

## 5.6 Sparseval

We recall from Section 5.3 that one of our goals is to obtain good bounds on

$$\sum_{m\text{-sparse } \chi} |\widehat{1_S}(\chi)|^2, \tag{5.12}$$

where the sum is over all characters $\chi = (r_1, \ldots, r_n)$ with precisely $m$ nonzero entries.

First consider the related sum

$$\sum_{\leq m\text{-sparse } \chi} |\widehat{1_S}(\chi)|^2, \tag{5.13}$$

i.e., the sum over characters with *at most* $m$ non-zero entries. This can be bounded by applying Parseval to the set $S_m$ of injections $\{1, \ldots, m\} \to G$ as a subset of the group $G^m$. Indeed, letting $N(\chi)$ be the set of indices $i$ for which $r_i$ is nonzero, we have

$$\begin{aligned}
\sum_{\leq m\text{-sparse } \chi} |\widehat{1_S}(\chi)|^2 &\leq \sum_{|N|=m} \sum_{N(\chi)\subset N} |\widehat{1_S}(\chi)|^2 \\
&= \binom{n}{m} \sum_{r_1,\ldots,r_m} |\widehat{1_S}(r_1, \ldots, r_m, 0, \ldots, 0)|^2 \\
&= \binom{n}{m} \sum_{r_1,\ldots,r_m} \frac{(n-m)!^2}{n^{2n-2m}} |\widehat{1_{S_m}}(r_1, \ldots, r_m)|^2 \qquad (5.14) \\
&= \binom{n}{m} \frac{(n-m)!^2}{n^{2n-2m}} \frac{n!}{n^m (n-m)!} \\
&= \frac{n^m}{m!} \frac{n!^2}{n^{2n}}.
\end{aligned}$$

Here we used (5.3).

This is our first nontrivial *sparseval* bound. It turns out that overcounting as above is too inefficient in the context of the wider argument, when, say, $m/n$ is a small constant. The purpose of the rest of this section therefore is to improve the bound on (5.12) as much as possible.

**Theorem 5.6.1.** *If $m \leq n/2$ then*

$$\sum_{m\text{-sparse } \chi} |\widehat{1_S}(\chi)|^2 \leq O(m^{1/4}) e^{O(m^{3/2}/n^{1/2})} \binom{n}{m}^{1/2} \frac{n!^2}{n^{2n}}$$

*Proof.* The starting point of our strategy is to apply inclusion–exclusion to obtain an expression for (5.12).

For $m \leq n$ let

$$Q(m, n) = \frac{n^{2n}}{n!^2} \sum_{m\text{-sparse } \chi} |\widehat{1_S}(r)|^2,$$

and observe as in (5.14) that

$$\begin{aligned}\frac{n^m}{m!} &= \frac{n^{2n}}{n!^2}\sum_{|N|=m}\sum_{N(\chi)\subset N}|\widehat{1_S}(\chi)|^2\\ &= \frac{n^{2n}}{n!^2}\sum_{k=0}^{m}\sum_{|N(\chi)|=k}|\widehat{1_S}(\chi)|^2\binom{n-k}{m-k}\\ &= \sum_{k=0}^{m}\binom{n-k}{m-k}Q(k,n).\end{aligned}$$

By inverting this relation we have

$$Q(m,n) = \sum_{k=0}^{m}(-1)^{m-k}\binom{n-k}{m-k}\frac{n^k}{k!} = \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}(-1)^k(n-m+1)\cdots(n-k)n^k. \tag{5.15}$$

This expression exhibits a vast amount of cancellation. To exploit this, we employ a generating function argument, followed by some complex analysis in the spirit of the saddle-point method (see [FS09, Chapter VIII] for background); see also Remark 5.6.2 for some further intuition.

Observe from (5.15) that $Q(m,n)$ is exactly the coefficient of $X^m$ in

$$f(X) = \frac{n^{n+1}e^X}{(n+X)^{n-m+1}}. \tag{5.16}$$

Thus by Cauchy's formula

$$Q(m,n) = \frac{n^{n+1}}{i2\pi}\oint_{|z|=r}\frac{e^z}{(n+z)^{n-m+1}z^{m+1}}\,dz,$$

for any $r$ in the range $0 < r < n$, so

$$Q(m,n) \le \frac{n^{n+1}}{r^m}\max_{|z|=r}\left|\frac{e^z}{(n+z)^{n-m+1}}\right|.$$

We claim that $|e^z/(n+z)^{n-m+1}|$ has only two local maxima on $|z| = r$: one at $z = +r$ and the other at $z = -r$. To see this note that if $|z| = r$ and $\Re z = t$ then

$$\left|\frac{e^z}{(n+z)^{n-m+1}}\right|^2 = \frac{e^{2t}}{(n^2+r^2+2nt)^{n-m+1}},$$

so

$$\frac{d}{dt}\log\left|\frac{e^z}{(n+z)^{n-m+1}}\right|^2 = 2 - \frac{2n(n-m+1)}{n^2+r^2+2nt}.$$

This function has a unique pole at $t = -(n+r^2/n)/2$, which is to the left of the allowed region $-r \le t \le r$, and it has a unique zero somewhere corresponding to a minimum, not a maximum, so the claim holds. Thus $|e^z/(n+z)^{n-m+1}|$ is bounded on $|z| = r$ by

$$\frac{e^{\pm r}}{(n \pm r)^{n-m+1}},$$

so

$$Q(m,n) \le \frac{n^{n+1}e^{\pm r}}{r^m(n\pm r)^{n-m+1}} = \frac{e^{\pm r}}{(1\pm r/n)^{n-m+1}}\frac{n^m}{r^m}.$$

Here

$$\begin{aligned}\log\left(\frac{e^{\pm r}}{(1\pm r/n)^{n-m+1}}\right) &= \pm r - (n-m+1)(\pm r/n - r^2/2n^2 + O(r^3/n^3)) \\ &= r^2/2n + O(r^3/n^2 + rm/n).\end{aligned}$$

Taking $r = (mn)^{1/2}$ we thus have, by Stirling's formula,

$$Q(m,n) \le e^{(1/2+O((m/n)^{1/2}))m}\frac{n^{m/2}}{m^{m/2}} \le O(m^{1/4})e^{O(m^{3/2}/n^{1/2})}\binom{n}{m}^{1/2}. \qquad \square$$

**Remark 5.6.2.** Although it is straightforward to verify from (5.15) that $Q(m,n)$ is the coefficient of $X^m$ in (5.16), this proof is unsatisfying in that one needs to know the formula for $f$ in advance. Here is an alternative proof of this fact which does not have this fault.

Let $\partial$ be the forward difference operator defined on functions $g : \mathbb{N} \to \mathbb{R}$ by

$$\partial g(k) = g(k+1) - g(k),$$

and observe that

$$Q(m,n) = \frac{1}{m!}\partial^m g_{m,n}(0),$$

where $g_{m,n}$ is the polynomial defined by

$$g_{m,n}(k) = n^k(n-k)(n-k-1)\cdots(n-m+1).$$

Intuitively, the sum (5.15) exhibits enormous cancellation, precisely because it is an expression for a high derivative of a very "smooth" function $g_{m,n}$. To take advantage

of this smoothness we investigate the first derivative $\partial g_{m,n}$. We compute

$$\begin{aligned}\partial g_{m,n}(k) &= n^{k+1}(n-k-1)(n-k-2)\cdots(n-m+1)- \\ &\qquad - n^k(n-k)(n-k-1)\cdots(n-m+1) \\ &= kn^k(n-k-1)(n-k-2)\cdots(n-m+1) \\ &= \frac{k}{n}g_{m,n}(k+1).\end{aligned}$$

From this it is possible to control the higher derivatives: using the Leibniz-type formula

$$\partial(gh)(k) = \partial g(k)h(k+1) + g(k)\partial h(k)$$

we derive the recurrence

$$\begin{aligned}\partial^\ell g_{m,n}(k) &= \partial^{\ell-1}\left(\frac{k}{n}g_{m,n}(k+1)\right) \\ &= \frac{k}{n}\partial^{\ell-1}g_{m,n}(k+1) + \frac{\ell-1}{n}\partial^{\ell-2}g_{m,n}(k+2),\end{aligned}$$

which holds for $\ell \geq 2$. In particular, letting

$$a_\ell = \frac{1}{\ell!}\partial^\ell g_{m,n}(m-\ell),$$

we have the recurrence

$$\ell a_\ell = \frac{m-\ell}{n}a_{\ell-1} + \frac{1}{n}a_{\ell-2} \tag{5.17}$$

for $(a_\ell)_{\ell\geq 0}$. Observe that $a_0 = n^m$, $a_1 = (m-1)n^{m-1}$, and that $Q(m,n) = a_m$.

In Section 5.7 we will bound solutions to recurrences like (5.17) in a direct fashion. For (5.17) itself, it is convenient to employ a generating function technique. For an indeterminate $X$ put

$$f(X) = \sum_{\ell=0}^{\infty} a_\ell X^\ell,$$

and observe by multiplying (5.17) by $X^{\ell-1}$ and summing for $\ell \geq 2$ that

$$f'(X) - a_1 = \frac{m-1}{n}(f(X) - a_0) - \frac{1}{n}Xf'(X) + \frac{1}{n}Xf(X)\ .$$

By inputting $a_0 = n^m$ and $a_1 = (m-1)n^{m-1}$ and rearranging we arrive at

$$(n-X)f'(X) = (m-1+X)f(X).$$

The formula (5.16) is obtained by solving this differential equation.

## 5.7 An $L^\infty$ bound for low-entropy minor arcs

The main result of this section is the following proposition.

**Proposition 5.7.1.** *Let $\chi$ be a character with exactly $m$ nonzero coordinates, where $m \leq n/3$. Then*

$$|\widehat{1_S}(\chi)| \leq e^{O(m^{3/2}/n^{1/2}+m^{1/2})} \cdot 2^{-m/2} \binom{n}{m}^{-1/2} \cdot \frac{n!}{n^n}.$$

**Remark 5.7.2.** It is instructive to compare this bound to what one can deduce from Theorem 5.5.1. If $\chi$ has exactly $m$ nonzero coordinates and takes precisely $k$ distinct nonzero values, i.e., $\chi = (r_1^{a_1}, \ldots, r_k^{a_k}, 0^{n-m})$ where $\sum a_i = m$, then Theorem 5.5.1 gives us a bound of

$$|\widehat{1_S}(\chi)| \leq \binom{n+k}{k}^{1/2} \binom{n}{a_1, \ldots, a_k, n-m}^{-1/2} \frac{n!}{n^n}$$

and the worst case of this bound over all choices of $k \leq m$ and $a_1, \ldots, a_k$ can be computed approximately as

$$|\widehat{1_S}(\chi)| \leq \binom{n}{m}^{-1/2} e^{o(m)} \frac{n!}{n^n},$$

provided that $m$ is significantly greater than $\sqrt{n}$. In other words a character similar to $\chi = (r^m, 0^{n-m})$ is asymptotically about as bad as the worst case. Moreover if we allow $n$ to be even and set $r = 1/2$ then the bound from Theorem 5.5.1 is essentially sharp, as remarked in that section.

However, in this setting, Proposition 5.7.1 improves on this bound by a significant factor of $2^{-m/2}$. Clearly then such an improvement is only possible under the assumption that $n$ is odd, and the proof of Proposition 5.7.1 must exploit this assumption in a fundamental way. Since we do not know how to adapt the proof of Theorem 5.5.1 to exploit the absence of 2-torsion, we employ a more specialized approach that suffices in the sparse regime.

We observe that this is the unique stage of the proof in which we need the full strength of the hypothesis that $G$ has odd order. Elsewhere we only need to assume that $\sum_{x \in G} x = 0$.

*Proof of Proposition 5.7.1.* Let $\chi = (r_1, \ldots, r_m, 0^{n-m})$. The key tool in our proof is an exact recursive formula for Fourier coefficients $\widehat{1_S}(\chi)$, in terms of related Fourier coefficients for which $m' < m$. Specifically, as long as $r_m$ is nonzero we have

$$\begin{aligned}
\widehat{1_S}(\chi) &= \frac{(n-m)!}{n^n} \sum_{x_1,\ldots,x_m}{}' e(-r_1x_1 - \cdots - r_mx_m) \\
&= \frac{(n-m)!}{n^n} \sum_{x_1,\ldots,x_{m-1}}{}' \sum_{x_m \neq x_1,\ldots,x_{m-1}} e(-r_1x_1 - \cdots - r_mx_m) \\
&= -\frac{(n-m)!}{n^n} \sum_{x_1,\ldots,x_{m-1}}{}' \sum_{x_m = x_1,\ldots,x_{n-1}} e(-r_1x_1 - \cdots - r_mx_m) \\
&= -\frac{1}{n-m+1} \sum_{i=1}^{m-1} \widehat{1_S}(r_1, \ldots, r_i + r_m, \ldots, r_{m-1}, 0, \ldots, 0). \qquad (5.18)
\end{aligned}$$

As an aside, we remark that this formula can be used for efficient explicit computation of certain kinds of Fourier coefficient.

Let $U_m$ be the maximal value of the $|\widehat{1_S}|$ taken over all characters with exactly $m$ nonzero coordinates. If we apply the triangle inequality to the right hand side of (5.18) and bound each $|\widehat{1_S}(\chi')|$ appearing there by the appropriate value $U_{m'}$ (where $m' < m$), we obtain recursive bounds on the $U_m$.

However, there is a subtlety in this process: the number of nonzero coordinates of

$$(r_1, \ldots, r_i + r_m, \ldots, r_{m-1}, 0, \ldots, 0)$$

might be either $(m-1)$ (if $r_i + r_m \neq 0$) or $(m-2)$ (if $r_i + r_m = 0$), and we do not have much control over which occurs. Since we expect $U_{m-1} < U_{m-2}$, it is in our interests to land in the former case as much as possible. We have the freedom to reorder the $r_i$ before applying (5.18), so our goal is to optimize the recursive bound over all choices of $r_m$.

For each $i \leq m$ let $\mathcal{N}_i = \{j \leq m \colon r_j = -r_i\}$. By reordering, we may assume without loss of generality that $|\mathcal{N}_i|$ is minimized when $i = m$. Suppose $j \in \mathcal{N}_m$. Then $r_j = -r_m$, so since $|G|$ is odd we have $r_j \neq r_m$, so $\mathcal{N}_m \cap \mathcal{N}_j = \emptyset$. Thus by minimality $|\mathcal{N}_m| \leq m/2$.

Hence, applying (5.18) we deduce the bound

$$\begin{aligned}|\widehat{1_S}(\chi)| &\leq \frac{|\mathcal{N}_m|U_{m-2} + (m-1-|\mathcal{N}_m|)U_{m-1}}{n-m+1} \\ &\leq \frac{1}{n-m+1}\max\left\{(m-1)U_{m-1},\ (m/2)U_{m-2} + (m/2-1)U_{m-1}\right\}\end{aligned}$$

where the former term in the maximum covers the case $U_{m-1} > U_{m-2}$ and the latter the more probable case $U_{m-1} \leq U_{m-2}$. Hence

$$U_m \leq \frac{1}{n-m+1}\max\left\{(m-1)U_{m-1},\ (m/2)U_{m-2} + (m/2-1)U_{m-1}\right\}. \tag{5.19}$$

Our task is now to obtain bounds on $U_m$ by solving this recurrence, which – ignoring the maximum – resembles a kind of time-dependent Fibonacci sequence.

In Remark 5.6.2 we dealt with a similar recurrence using generating function methods. Here, the presence of the maximum of two alternatives, and possibly other reasons, make this approach less attractive. Instead, we will bound (5.19) by more hands-on methods.

Specifically, we will phrase (5.19) in terms of a product of $2 \times 2$ matrices (as one might for the Fibonacci sequence), and control the $L^2$ norm of the result by bounding the $L^2 \to L^2$ operator norms of each matrix in the product.

Write $\alpha_m = \frac{m}{2(n-m+1)}$, $\beta_m = \frac{m-2}{2(n-m+1)}$, and $\gamma_m = \frac{m-1}{n-m+1}$. Additionally, we work with a rescaled version $V_m = (\alpha_1\alpha_2\ldots\alpha_m)^{-1/2}U_m$ of $U_m$, for $m \geq 1$. From (5.19) we deduce that

$$V_m \leq \max\left\{\gamma_m\alpha_m^{-1/2}\,V_{m-1},\ \alpha_m^{1/2}\alpha_{m-1}^{-1/2}\,V_{m-2} + \beta_m\alpha_m^{-1/2}\,V_{m-1}\right\} \tag{5.20}$$

holds for any $m \geq 3$.

If we define matrices

$$M_m = \begin{pmatrix} \gamma_m\alpha_m^{-1/2} & 0 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad N_m = \begin{pmatrix} \beta_m\alpha_m^{-1/2} & \alpha_m^{1/2}\alpha_{m-1}^{-1/2} \\ 1 & 0 \end{pmatrix},$$

and vectors $v_m = (V_m, V_{m-1})^T$, we can write (5.20) equivalently as

$$v_m \leq \max\left\{M_m v_{m-1}, N_m v_{m-1}\right\}.$$

Using the easily obtainable bounds $\alpha_m\alpha_{m-1}^{-1} = 1 + O(1/m)$ and $\beta_m^2\alpha_m^{-1} \leq m/n$,

we get that

$$\mathrm{Tr}\left(N_m^T N_m\right) \leq 2 + m/n + O(1/m)$$

and

$$\det\left(N_m^T N_m\right) = 1 + O(1/m),$$

and so by considering the singular values of $N_m$, we get the operator norm bound

$$\|N_m\|_{L^2 \to L^2} \leq 1 + O((m/n)^{1/2} + m^{-1/2}).$$

Completely analogously, using the easy bound $\gamma_m^2 \alpha_m^{-1} \leq 4m/n$, we deduce

$$\|M_m\|_{L^2 \to L^2} \leq 1 + O(m/n).$$

Thus we have

$$\begin{aligned} V_m &\leq \sqrt{V_m^2 + V_{m-1}^2} \\ &\leq \left( \prod_{r=3}^{m} \max\left\{ \|M_r\|_{L^2 \to L^2}, \|N_r\|_{L^2 \to L^2} \right\} \right) \cdot \sqrt{V_2^2 + V_1^2} \\ &\leq \left( \prod_{r=3}^{m} \left(1 + O\left((r/n)^{1/2} + r^{-1/2}\right)\right) \right) \cdot \sqrt{V_2^2 + V_1^2}. \end{aligned}$$

We already know from Section 5.3 that $U_1 = 0$ and it is evident from (5.18) that $U_2 = O(n!/n^{n+1})$. Hence, we can bound the term $\sqrt{V_2^2 + V_1^2}$ in the above inequality by $O(n!/n^n)$, which gives

$$V_m \leq e^{O\left(m^{3/2}/n^{1/2} + m^{1/2}\right)} \cdot \frac{n!}{n^n},$$

and the claim stated in the proposition follows easily from this. □

## 5.8 End of the argument

To estimate the sum of cubes

$$\sum_{\chi \in \widehat{G}^n} \widehat{1_S}(\chi)^3$$

we divide the set of all $\chi \in \widehat{G}^n$ into three regions depending on the value of $H(\chi)$: a high-entropy range $H \geq 10$, a medium-entropy range $\varepsilon \leq H \leq 10$, and a low-entropy range $H \leq \varepsilon$. Here $\varepsilon$ is a small positive constant which will be chosen at the end of

the argument.

In the high-entropy range $H \geq 10$ it is enough to use the bound from Corollary 5.3.2,

$$\sum_{\chi:\, H(\chi)\geq 10} |\widehat{1_S}(\chi)|^3 \leq e^{-3.5n} n!^3/n^{3n}.$$

In the medium-entropy range $\varepsilon \leq H \leq 10$, we first need a bound for the number of characters of a given entropy.

**Lemma 5.8.1.** *If $H \leq 10$ then the number of characters of entropy at most $H$ is bounded by $e^{Hn+o(n)}$.*

*Proof.* Every character of entropy at most $H$ has an orbit under the permutation action of size at most $e^{Hn}$, so it suffices to show that there are only $e^{o(n)}$ such orbits of characters of entropy at most 10. Every orbit is uniquely specified by giving the number $a_r$ of appearances of $r$ for each $r \in \widehat{G}$, so we must show that the number of nonnegative integer vectors $(a_r)_{r\in\widehat{G}}$ with $\sum_r a_r = n$ and satisfying

$$\frac{n!}{\prod_{r\in\widehat{G}} a_r!} \leq e^{10n}$$

is $e^{o(n)}$. Let $\delta > 0$ be a small constant, and let $t$ be the sum of the $a_r$ for which $a_r \leq e^{-11/\delta} n$. Then

$$e^{-11n} n^n \leq e^{-10n} n! \leq \prod_{r\in\widehat{G}} a_r! \leq \prod_{r\in\widehat{G}} a_r^{a_r} \leq (e^{-11/\delta} n)^t n^{n-t} = e^{-11t/\delta} n^n,$$

so $t \leq \delta n$. Since at most $e^{11/\delta}$ of the $a_r$ are bigger than $e^{-11/\delta} n$, the number of $(a_r)_{r\in\widehat{G}}$ is bounded by the number of ways of choosing the set $B$ of at most $e^{11/\delta}$ indices $r$ for which $a_r$ is big, times the number of ways of choosing these $a_r$, times the number of ways of choosing $(a_r)_{r\notin B}$ so that $\sum_{r\notin B} a_r = n - \sum_{r\in B} a_r \leq \delta n$. This is bounded by

$$n^{e^{11/\delta}} n^{e^{11/\delta}} \binom{n + \delta n - 1}{\delta n - 1}.$$

The claimed estimate now follows by applying Stirling's formula and choosing $\delta$ appropriately. □

We conclude by combining Lemma 5.8.1 with Theorem 5.5.1. Note that if the entropy of $\chi$ is $O(1)$ then the number of distinct coordinates of $\chi$ is $O(n/\log n)$, so Theorem 5.5.1 implies

$$|\widehat{1_S}(\chi)| \leq e^{-H(\chi)n/2+o(n)} n!/n^n.$$

Thus for any $H \leq 10$ we have

$$\sum_{\chi\colon 9H/10<H(\chi)\leq H} |\widehat{1_S}(\chi)|^3 \leq |\{\chi\colon H(\chi)\leq H\}| \max_{\chi\colon H(\chi)\geq 9H/10} |\widehat{1_S}(\chi)|^3$$
$$\leq e^{-7Hn/20+o(n)} n!^3/n^{3n}.$$

Decomposing the range $[\varepsilon, 10]$ into ranges of this type and bounding crudely, we obtain

$$\sum_{\chi\colon \varepsilon\leq H(\chi)\leq 10} |\widehat{1_S}(\chi)|^3 \leq O(\log(1/\varepsilon)) e^{-7\varepsilon n/20+o(n)} n!^3/n^{3n}.$$

In the low-entropy range $H \leq \varepsilon$, one can easily verify that $\chi = (r_1, \ldots, r_n)$ must repeat some coordinate $r \in \widehat{G}$ at least $(1-\varepsilon)n$ times. Thus by global shift-invariance of $\widehat{1_S}$ we have

$$\sum_{\chi\colon H(\chi)\leq\varepsilon} \widehat{1_S}(\chi)^3 = n \sum_{m=0}^{\varepsilon n} \sum_{\substack{m\text{-sparse } \chi \\ H(\chi)\leq\varepsilon}} \widehat{1_S}(\chi)^3. \tag{5.21}$$

By combining Proposition 5.7.1 with Theorem 5.6.1 we have that for any $m \leq \varepsilon n$,

$$\sum_{m\text{-sparse } \chi} |\widehat{1_S}(\chi)|^3 \leq \left(\max_{m\text{-sparse } \chi} |\widehat{1_S}(\chi)|\right) \sum_{m\text{-sparse } \chi} |\widehat{1_S}(\chi)|^2$$
$$\leq e^{O\left(m^{3/2}/n^{1/2}+m^{1/2}\right)} 2^{-m/2} \frac{n!^3}{n^{3n}}$$
$$\leq e^{O\left(\varepsilon^{1/2}m+m^{1/2}\right)} 2^{-m/2} \frac{n!^3}{n^{3n}}.$$

As long as $\varepsilon$ is sufficiently small depending on the constant implicit in the $O(\varepsilon^{1/2}m)$ term, this is negligible except when $m$ has size $O(1)$. In this range, we can apply Proposition 5.4.1. We introduce a further parameter $M$ and split the sum (5.21) into the ranges $1 \leq m \leq 2M$ and $2M < m \leq \varepsilon n$, to obtain

$$\sum_{\chi\colon H(\chi)\leq\varepsilon} \widehat{1_S}(\chi)^3 = n \sum_{m=0}^{M} \frac{(-1)^m}{2^m m!} \frac{n!^3}{n^{3n}} + nO_M\left(\frac{1}{n}\frac{n!^3}{n^{3n}}\right)$$
$$+ O\left(n \sum_{m=2M}^{\varepsilon n} e^{O\left(\varepsilon^{1/2}m+m^{1/2}\right)} 2^{-m/2} \frac{n!^3}{n^{3n}}\right).$$

Finally, by combining the three bounds we have just proved for each range $H \leq \varepsilon$,

$\varepsilon \leq H \leq 10$ and $H \geq 10$, we deduce

$$\begin{aligned}\sum_{\chi \in \widehat{G}^n} \widehat{1_S}(\chi)^3 = n \sum_{m=0}^{M} \frac{(-1)^m}{2^m m!} \frac{n!^3}{n^{3n}} + nO_M\left(\frac{1}{n}\frac{n!^3}{n^{3n}}\right) \\ + O\left(n \sum_{m=2M}^{\varepsilon n} e^{O\left(\varepsilon^{1/2}m + m^{1/2}\right)} 2^{-m/2} \frac{n!^3}{n^{3n}}\right) \\ + O\left(\log(1/\varepsilon) e^{-7\varepsilon n/20 + o(n)} \frac{n!^3}{n^{3n}}\right) \\ + O\left(e^{-3.5n} \frac{n!^3}{n^{3n}}\right).\end{aligned}$$

Theorem 5.1.2 now follows by choosing $\varepsilon$ and $M$ appropriately.

## 5.9 Quantitative versions of a result of Hall

In the previous sections we stated two open problems motivated by the work presented in this chapter. The first one was to extend our results to all abelian groups, i.e. address the issue for $n$ even. The second one was the toroidal queens problem, i.e. finding the number of bijections $\pi_1, \pi_2 \in S$ such that $\pi_1 + \pi_2$ and $\pi_1 - \pi_2$ are both in $S$. In this section we give one more interesting open question that arises naturally from our work.

For an arbitrary function $f\colon \{1, \ldots, n\} \to G$ let $r(f)$ be the number of representations $f = \pi_1 + \pi_2$ where $\pi_1, \pi_2 \in S$. Since $r(f) = r(f \circ \tau)$ for any permutation $\tau$ on $\{1 \ldots, n\}$, Theorem 5.1.2 shows that for any $\pi \in S$ we have

$$r(\pi) = (e^{-1/2} + o(1))n!^2/n^{n-1}.$$

So the natural question is what is the value of $r(f)$ for other functions $f\colon \{1, \ldots, n\} \to G$?

First of all, notice that if $f$ is such that $\sum_{i=1}^{n} f(i) \neq 0$, then $r(f) = 0$. Indeed, if $f = \pi_1 + \pi_2$ for $\pi_1, \pi_2 \in S$ then

$$\sum_{i=1}^{n} f(i) = \sum_{i=1}^{n} (\pi_1(i) + \pi_2(i)) = 2 \sum_{g \in G} g = 0,$$

where the last equality can be easily proven by pairing each element with its inverse.

In 1952, Hall [Hal52] proved that that the obstruction stated above is the only

one, i.e. for all other functions $f$ we have $r(f) > 0$. So our main result can be seen as a quantitative version of Hall's theorem in the case when $f$ is a bijection and $n$ is odd.

**Theorem 5.9.1.** *Let $G$ be a finite abelian group of size $n$ and $f\colon \{1,\dots,n\} \to G$ a function such that*

$$\sum_{i=1}^{n} f(i) = 0. \tag{5.22}$$

*There exist two bijections $\pi_1, \pi_2\colon \{1,\dots,n\} \to G$ such that $f = \pi_1 + \pi_2$.*

Here we give a short proof of Hall's theorem.

*Proof.* Let $\pi_1$ and $\pi_2$ be arbitrary permutations and assume that $\pi_1 + \pi_2 \neq f$. Let $I(f, \pi_1, \pi_2)$ be the set of all $i$ for which $f(i) \neq \pi_1(i) + \pi_2(i)$. Since

$$\sum_{i=1}^{n} (f(i) - \pi_1(i) - \pi_2(i)) = 0,$$

we must have that $|I(f, \pi_1, \pi_2)| \geq 2$. Suppose that $|I(f, \pi_1, \pi_2)| = 2$, namely $I(f, \pi_1, \pi_2) = \{i_1, i_2\}$. We define $i_k$ for $k \geq 3$ so that the following is satisfied for $k \geq 2$

$$\pi_1(i_{k-1}) + \pi_2(i_{k+1}) = f(i_k).$$

Let $r < s$ be the smallest indices for which $i_r = i_s$. Suppose that $r \geq 3$. Then

$$\begin{aligned} 0 &= \sum_{k=r}^{s-1} (\pi_1(i_{k-1}) + \pi_2(i_{k+1}) - f(i_k)) \\ &= \sum_{k=r}^{s-1} (\pi_1(i_{k-1}) + \pi_2(i_{k+1}) - \pi_1(i_k) - \pi_2(i_k)) \\ &= \pi_1(i_{r-1}) - \pi_1(i_{s-1}), \end{aligned}$$

and so $i_{r-1} = i_{s-1}$. The conclusion is that $r \leq 2$. However, by our choice of $i_k$ it is easy to see that

$$\pi_1 \cdot (i_{s-1} i_{s-2} \dots i_1) + \pi_2 \cdot (i_r i_{r+1} \dots i_{s-2} i_{s-1}) = f$$

holds at all points except possibly at $i_1$. But then because of the condition (5.22) it must also hold at $i_1$.

Suppose now, on the other hand, that $|I(f,\pi_1,\pi_2)|>2$, and let $i_1,i_2\in I(f,\pi_1,\pi_2)$. We define the function $g\colon\{1,\dots,n\}\to G$ by

$$g(i)=\begin{cases}\pi_1(i)+\pi_2(i) & \text{if } i\neq i_1,i_2,\\ f(i_1) & \text{if } i=i_1,\\ -(\sum_{j\neq i_1,i_2}(\pi_1(j)+\pi_2(j))+f(i_1)) & \text{if } i=i_2.\end{cases}$$

It is easy to see that $I(g,\pi_1,\pi_2)=2$ and hence, by the previous case, there are two bijections $\pi_1'$ and $\pi_2'$ such that $g=\pi_1'+\pi_2'$. But then, by the definition of $g$, we have $|I(f,\pi_1',\pi_2')|\leq|I(f,\pi_1,\pi_2)|-1$. We can now continue with this procedure until we get that $f$ is equal to a sum of two permutations at all points, which is what we had to prove. □

Extending the main result of this chapter so that it also gives $r(f)$ for an arbitrary function $f$ would certainly require some new ideas as can already be seen if one considers the function $f\equiv 0$. Indeed, in that case we know that $r(f)=n!$, but if we want to prove this using our method, then the task boils down to proving that

$$\sum_{\chi\in\widehat{G}^n}|\widehat{1_S}(\chi)|^2=n!/n^n. \tag{5.23}$$

However, one can easily see from our results about major arcs that the minor arcs contribution to the left hand side of (5.23) is $\gg n!/n^n$, so considering only major arcs as we did when $f$ was a bijection would not work in general.

# Bibliography


[AAAS94] Pankaj Kumar Agarwal, Noga Alon, Boris Aronov, and Subhash Suri. Can visibility graphs be represented compactly? *Discrete Comput. Geom.*, 12(3):347–365, 1994. ACM Symposium on Computational Geometry (San Diego, CA, 1993).

[Alo13] Noga Alon. The chromatic number of random Cayley graphs. *European J. Combin.*, 34(8):1232–1243, 2013.

[AN05] Dimitris Achlioptas and Assaf Naor. The two possible values of the chromatic number of a random graph. *Ann. of Math. (2)*, 162(3):1335–1351, 2005.

[AS08] Noga Alon and Joel H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[BE76] Béla Bollobás and Paul Erdős. Cliques in random graphs. *Math. Proc. Cambridge Philos. Soc.*, 80(3):419–427, 1976.

[BIW06] Boaz Barak, Russell Impagliazzo, and Avi Wigderson. Extracting randomness using few independent sources. *SIAM J. Comput.*, 36(4):1095–1118 (electronic), 2006.

[BMR13] C. Bachoc, M. Matolcsi, and I. Z. Ruzsa. Squares and difference sets in finite fields. *Integers*, 13:Paper No. A77, 5, 2013.

[Bol04] Béla Bollobás. How sharp is the concentration of the chromatic number? *Combin. Probab. Comput.*, 13(1):115–117, 2004.

[Bón15] M. Bóna. *Handbook of Enumerative Combinatorics*. Chapman and Hall/CRC, 2015.

[CG88] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput.*, 17(2):230–261, 1988. Special issue on cryptography.

[CGKN99] C. Cooper, R. Gilchrist, I. N. Kovalenko, and D. Novaković. Deriving the number of "good" permutations, with applications to cryptography. *Kibernet. Sistem. Anal.*, (5):10–16, 187, 1999.

[Chu01] Kai Lai Chung. *A course in probability theory.* Academic Press, Inc., San Diego, CA, third edition, 2001.

[CK96] C. Cooper and I. N. Kovalenko. The upper bound for the number of complete mappings. *Theory Probab. Math. Statist.*, 53:77–83, 1996.

[Coh88] Stephen D. Cohen. Clique numbers of Paley graphs. *Quaestiones Math.*, 11(2):225–231, 1988.

[Coo00] C. Cooper. A lower bound for the number of good permutations. *Data Recording, Storage and Processing (Nat. Acad. Sci. Ukraine)*, 213:15–25, 2000.

[COPS08] Amin Coja-Oghlan, Konstantinos Panagiotou, and Angelika Steger. On the chromatic number of random graphs. *J. Combin. Theory Ser. B*, 98(5):980–993, 2008.

[CS10] Ernie Croot and Olof Sisask. A probabilistic technique for finding almost-periods of convolutions. *Geom. Funct. Anal.*, 20(6):1367–1396, 2010.

[CW10] N. J. Cavenagh and I. M. Wanless. On the number of transversals in Cayley tables of cyclic groups. *Discrete Appl. Math.*, 158(2):136–146, 2010.

[CZ16] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 670–683, 2016.

[EMM15] Sean Eberhard, Freddie Manners, and Rudi Mrazović. Additive triples of bijections, or the toroidal semiqueens problem. 2015. Preprint, arXiv:1510.05987.

[EZ12] Chaim Even-Zohar. On sums of generating sets in $\mathbb{Z}_2^n$. *Combin. Probab. Comput.*, 21(6):916–941, 2012.

[FS09] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.

[GL15] R. Glebov and Z. Luria. On the maximum number of latin transversals. 2015. Preprint, arXiv:1506.00983.

[GM15] Ben Green and Robert Morris. Counting sets with small sumset and applications. *Combinatorica*, 2015.

[GR90] Sidney West Graham and C. J. Ringrose. Lower bounds for least quadratic nonresidues. In *Analytic number theory (Allerton Park, IL, 1989)*, volume 85 of *Progr. Math.*, pages 269–309. Birkhäuser Boston, Boston, MA, 1990.

[Gre05a] Ben Green. Counting sets with small sumset, and the clique number of random Cayley graphs. *Combinatorica*, 25(3):307–326, 2005.

[Gre05b] Ben Green. A Szemerédi-type regularity lemma in abelian groups, with applications. *Geom. Funct. Anal.*, 15(2):340–376, 2005.

[Gre15] Ben Green. On the chromatic number of random cayley graphs. *Combin. Probab. Comput., to appear*, 2015.

[GS03] Andrew Granville and Kannan Soundararajan. The distribution of values of $l(1, \chi_d)$. *Geom. Funct. Anal.*, 13(5):992–1028, 2003.

[GS16] Gary R. W. Greaves and Leonard H. Soicher. On the clique number of a strongly regular graph. 2016. Preprint, arXiv:1604.08299.

[Hal52] Marshall Hall, Jr. A combinatorial problem on abelian groups. *Proc. Amer. Math. Soc.*, 3:584–587, 1952.

[IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Kne53] Martin Kneser. Abschätzung der asymptotischen Dichte von Summenmengen. *Math. Z.*, 58:459–484, 1953.

[Kov96] I. N. Kovalenko. On an upper bound for the number of complete mappings. *Kibernet. Sistem. Anal.*, (1):81–85, 188, 1996.

[Kuz07] N. Yu. Kuznetsov. Applying fast simulation to find the number of good permutations. *Cybernet. Systems Anal.*, 43:830–837, 2007.

[Kuz08] N. Yu. Kuznetsov. Estimating the number of good permutations by a modified fast simulation method. *Cybernet. Systems Anal.*, 44:547–554, 2008.

[Kuz09] N. Yu. Kuznetsov. Estimating the number of latin rectangles by the fast simulation method. *Cybernet. Systems Anal.*, 45:69–75, 2009.

[Mat72] David W. Matula. The employee party problem. *Not. Amer. Math. Soc.*, 19(2):A–382, 1972.

[MMW06] B. D. McKay, J. C. McLeod, and I. M. Wanless. The number of transversals in a Latin square. *Des. Codes Cryptogr.*, 40(3):269–284, 2006.

[Mon71] Hugh L. Montgomery. *Topics in multiplicative number theory.* Lecture Notes in Mathematics, Vol. 227. Springer-Verlag, Berlin-New York, 1971.

[Mra15] Rudi Mrazović. One-point concentration of the clique and chromatic numbers of the random cayley graph on $\mathbb{F}_2^n$. 2015. Preprint, arXiv:1510.05991.

[Mra16a] Rudi Mrazović. Extractors in Paley graphs: a random model. *European J. Combin.*, 54:154–162, 2016.

[Mra16b] Rudi Mrazović. A random model for the paley graph. 2016. Preprint, arXiv:1603.00684.

[MV07] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

[RVZ94] I. Rivin, I. Vardi, and P. Zimmermann. The $n$-queens problem. *Amer. Math. Monthly*, 101(7):629–639, 1994.

[Sch02] T. Schoen. The cardinality of restricted sumsets. *J. Number Theory*, 96(1):48–54, 2002.

[Slo] N. J. A. Sloane. The on-line encyclopedia of integer sequences. `https://oeis.org/`.

[Tar15] A. A. Taranenko. Multidimensional permanents and an upper bound on the number of transversals in Latin squares. *J. Combin. Des.*, 23(7):305–320, 2015.

[TV06] Terence Tao and Van Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[Var91] I. Vardi. *Computational recreations in Mathematica*. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, 1991.

[Wan11] I. M. Wanless. Transversals in Latin squares: a survey. In *Surveys in combinatorics 2011*, volume 392 of *London Math. Soc. Lecture Note Ser.*, pages 403–437. Cambridge Univ. Press, Cambridge, 2011.