

**ON THE PROBABILITY THAT A DISCRETE
COMPLEX RANDOM MATRIX IS SINGULAR**

BY PHILIP J. WOOD

A dissertation submitted to the
Graduate School—New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
Graduate Program in Mathematics
Written under the direction of
Van H. Vu
and approved by

New Brunswick, New Jersey

May, 2009

© 2009

Philip J. Wood

ALL RIGHTS RESERVED

ABSTRACT OF THE DISSERTATION

On the probability that a discrete complex random matrix is singular

by Philip J. Wood

Dissertation Director: Van H. Vu

Let n be a large integer and M_n be an n by n complex matrix whose entries are independent (but not necessarily identically distributed) discrete random variables. The main goal of this thesis is to prove a general upper bound for the probability that M_n is singular.

For a constant $0 < p < 1$ and a constant positive integer r , we will define a property *p-bounded of exponent r*. Our main result shows that if the entries of M_n satisfy this property, then the probability that M_n is singular is at most $(p^{1/r} + o(1))^n$. All of the results in this thesis hold for any characteristic zero integral domain replacing the complex numbers.

In the special case where the entries of M_n are “fair coin flips” (taking the values $+1, -1$ each with probability $1/2$), our general bound implies that the probability that M_n is singular is at most $\left(\frac{1}{\sqrt{2}} + o(1)\right)^n$, improving on the previous best upper bound of $\left(\frac{3}{4} + o(1)\right)^n$, proved by Tao and Vu [39].

In the special case where the entries of M_n are “lazy coin flips” (taking values $+1, -1$ each with probability $1/4$ and value 0 with probability $1/2$), our general bound implies that the probability that M_n is singular is at most $\left(\frac{1}{2} + o(1)\right)^n$, which is asymptotically sharp.

Our method is a refinement of those from [22] and [39]. In particular, we make a critical use of the Structure Theorem from [39], which was obtained using tools from additive combinatorics.

One key lemma for extending our results to the complex numbers follows from a more general result about characteristic zero integral domains. We show that any finite system S in a characteristic zero integral domain can be mapped to $\mathbb{Z}/Q\mathbb{Z}$, for infinitely many primes Q , preserving all algebraic incidences in S . This can be seen as a generalization of the well-known Freiman isomorphism lemma, which asserts that any finite subset of a torsion-free group can be mapped into $\mathbb{Z}/Q\mathbb{Z}$, preserving all linear incidences.

As applications, we derive several combinatorial results (such as sum-product estimates) for a finite set in a characteristic zero integral domain. As \mathbb{C} is a characteristic zero integral domain, this allows us to obtain new proofs for some recent results concerning finite sets of complex numbers, without relying on the topology of the plane.

Acknowledgements

This thesis is comprised of work and results from two papers: [7], which is joint work with Jean Bourgain and Van Vu, and [44], which is joint work with Van Vu and Melanie Matchett Wood. All of the chapters except for Chapter 3 (but including Section 3.1.1) are derived from [7], and all of Chapter 3 (except for Section 3.1.1) is derived from [44]. Currently, [7] is in preparation and [44] has been submitted and is being refereed.

I would like to thank Kevin Costello for helpful conversations on the conditioning argument in Subsection 2.2.4 that lead to Corollary 2.2.5. I would also like to thank Ellen Eischen for providing numerous clarifications and simplifications on an early draft of Theorem 3.1.4 and its proof. Thanks is also due to John Bryk, Wei Ho, Terence Tao, and J.B. Tunnell for useful comments on Theorem 3.1.4 and its proof. Finally, I would like to thank the National Defense Science and Engineering Graduate Fellowship program and the National Science Foundation Graduate Research Fellowship program for providing funding for the work in this thesis.

Dedication

This thesis is dedicated to my late grandmother, Margaret Matchett, who earned her Ph.D in mathematics from Indiana University in 1946 under the direction of Emil Artin.

Table of Contents

| | |
|-----------------------------------------------------------------------------------------|------|
| Abstract | ii |
| Acknowledgements | iv |
| Dedication | v |
| List of Figures | viii |
| 1. Introduction | 1 |
| 1.1. The probability that a discrete random matrix is singular | 1 |
| 2. Main result and applications | 7 |
| 2.1. The general theorem | 7 |
| 2.2. Some corollaries of Theorem 2.1.2 | 10 |
| 2.2.1. Proving Inequalities (1.4), (1.5), and (1.6) | 11 |
| 2.2.2. Matrices with entries having symmetric distributions | 12 |
| 2.2.3. Random matrices with entries having arbitrary distributions | 16 |
| 2.2.4. Partially random matrices | 19 |
| 2.2.5. Integer matrices and rational eigenvalues | 21 |
| 3. Reduction to a finite field, with other applications | 24 |
| 3.1. Introduction | 24 |
| 3.1.1. Placing complex matrices in a finite field | 24 |
| 3.1.2. Mapping incidences to a finite field | 26 |
| 3.2. A Szemerédi-Trotter-type result for characteristic zero integral domains | 28 |
| 3.3. A sum-product result for characteristic zero integral domains | 30 |
| 3.4. A matrix product result for $SL_2(D)$ | 31 |
| 3.5. Random matrices with entries in a characteristic zero integral domain | 36 |

| | |
|----------------------------------------------------------------------------------------|-----------|
| 3.6. The density theorem | 38 |
| 3.7. Proof of Theorem 3.1.4 | 39 |
| 4. Proof of main result: reduction to exceptional spaces | 43 |
| 4.1. Proof of the main theorem (Theorem 2.1.2) | 43 |
| 4.1.1. Definitions and preliminaries | 43 |
| 4.1.2. Proof of the medium combinatorial dimension | 47 |
| 4.1.3. The unexceptional medium combinatorial dimension case | 48 |
| 4.2. Analyzing the exceptional medium combinatorial dimension case | 49 |
| 5. The structure of exceptional spaces | 54 |
| 5.1. Halász-type arguments | 54 |
| 5.2. Proof of the Structure Theorem (Theorem 4.2.1) | 60 |
| 6. Extension to the case where some rows are fixed | 69 |
| 6.1. A generalization: f rows have fixed, non-random values | 69 |
| 6.1.1. Outline of the proof of Theorem 6.1.2 | 70 |
| Appendix A. Lemmas in additive combinatorics | 74 |
| A.1. Two background results | 74 |
| A.1.1. A version of the Littlewood-Offord result in $\mathbb{Z}/Q\mathbb{Z}$ | 74 |
| A.1.2. A generalization of a lemma due to Komlós [26] | 75 |
| A.2. The unexceptional case with f fixed rows | 77 |
| References | 80 |
| Vita | 83 |

List of Figures

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1. Asymptotic Upper and Lower Bounds for $\Pr \left(M_{\pm 1, n}^{(\mu)} \text{ is singular} \right)^{1/n}$ for $0 \leq \mu \leq 1$ | 4 |
| 2.1. Asymptotic Upper and Lower Bounds for $\Pr \left(M_{\{\pm 2, \pm 1\}, n}^{(\mu)} \text{ is singular} \right)^{1/n}$ for $0 \leq \mu \leq 1$ | 15 |

Chapter 1

Introduction

1.1 The probability that a discrete random matrix is singular

Let n be a large integer and M_n be an n by n random matrix whose entries are independent (but not necessarily identically distributed) discrete random variables taking values in the complex numbers. The problem of estimating the probability that M_n is singular is a basic problem in the theory of random matrices and combinatorics. The goal of this thesis is to give a bound that applies to a large variety of distributions. The general statement (Theorem 2.1.2) is a bit technical, so we will first discuss a few corollaries concerning special cases.

The most famous special case is when the entries of M_n are independent identically distributed (i.i.d.) Bernoulli random variables (taking values ± 1 with probability $1/2$). The following conjecture has been open for quite some time:

Conjecture 1.1.1. *For $M_{\pm 1, n}$ an n by n matrix with each entry an i.i.d. Bernoulli random variable taking the values $+1$ and -1 each with probability $1/2$,*

$$\Pr(M_{\pm 1, n} \text{ is singular}) = \left(\frac{1}{2} + o(1)\right)^n.$$

It is easy to verify that the singularity probability is at least $(1/2)^n$ by considering the probability that there are two equal rows (or columns).

Even in the case of i.i.d. Bernoulli random variables, proving that the singularity probability is $o(1)$ is not trivial. It was first done by Komlós in 1967 [25] (see also [26]; [32] generalizes Komlós's bound to other integer distributions). The first exponential bound was proven by Kahn, Komlós, and Szemerédi [22], who showed that $\Pr(M_{\pm 1, n} \text{ is singular}) \leq .999^n$. This upper bound was improved upon by Tao and Vu

in [38] to $.958^n$. A more significant improvement was obtained by the same authors in [39]:

$$\Pr(M_{\pm 1, n} \text{ is singular}) \leq \left(\frac{3}{4} + o(1)\right)^n. \quad (1.1)$$

This improvement was made possible through the discovery of a new theorem [39, Theorem 5.2] (which was called the Structure Theorem in [39]), which gives a complete characterization of a set with certain additive properties. The Structure Theorem (to be more precise, a variant of it) will play a critical role in the current thesis as well.

Our general result has the following corollary in the Bernoulli case:

$$\Pr(M_{\pm 1, n} \text{ is singular}) \leq \left(\frac{1}{\sqrt{2}} + o(1)\right)^n, \quad (1.2)$$

which gives a slight improvement over Inequality (1.1) (since $1/\sqrt{2} \approx 0.7071 < .75$).

Let us now discuss a more general class of random matrices. Consider the random variable $\gamma^{(\mu)}$ defined by

$$\gamma^{(\mu)} := \begin{cases} +1 & \text{with probability } \mu/2 \\ 0 & \text{with probability } 1 - \mu \\ -1 & \text{with probability } \mu/2, \end{cases} \quad (1.3)$$

and let $M_{\pm 1, n}^{(\mu)}$ be an n by n matrix with each entry an independent copy of $\gamma^{(\mu)}$. The random variable $\gamma^{(\mu)}$ plays an important role in [22, 38, 39], and the matrices $M_{\pm 1, n}^{(\mu)}$ are of interest in their own right. In fact, giving zero a large weight is a natural thing to do when one would like to (randomly) sparsify a matrix, a common operation used in randomized algorithms (the values of ± 1 , as the reader will see, are not so critical). Our general result implies the following upper bounds:

$$\Pr(M_{\pm 1, n}^{(\mu)} \text{ is singular}) \leq (1 - \mu + o(1))^n \quad \text{for } 0 \leq \mu \leq \frac{1}{2} \quad (1.4)$$

$$\Pr(M_{\pm 1, n}^{(\mu)} \text{ is singular}) \leq \left(\frac{2\mu + 1}{4} + o(1) \right)^n \quad \text{for } \frac{1}{2} \leq \mu \leq 1 \quad (1.5)$$

$$\Pr(M_{\pm 1, n}^{(\mu)} \text{ is singular}) \leq \left(\sqrt{1 - 2\mu + \frac{3}{2}\mu^2} + o(1) \right)^n \quad \text{for } 0 \leq \mu \leq 1. \quad (1.6)$$

Note that Inequality (1.5) implies Inequality (1.1) and that Inequality (1.6) implies Inequality (1.2) (in both cases setting $\mu = 1$).

Figure 1.1 summarizes the upper bounds from Inequalities (1.4), (1.5), and (1.6) and also includes the following lower bounds:

$$(1 - \mu + o(1))^n \leq \Pr(M_{\pm 1, n}^{(\mu)} \text{ is singular}) \quad \text{for } 0 \leq \mu \leq 1 \quad (1.7)$$

$$\left(1 - 2\mu + \frac{3}{2}\mu^2 + o(1) \right)^n \leq \Pr(M_{\pm 1, n}^{(\mu)} \text{ is singular}) \quad \text{for } 0 \leq \mu \leq 1. \quad (1.8)$$

These lower bounds can be derived by computing the probability that one row is all zeros (Inequality (1.7)) or that there is a dependency between two rows (Inequality (1.8)). Note that in the case where $\mu \leq 1/2$, the upper bound in Inequality (1.4) asymptotically equals the lower bound in Inequality (1.7), and thus our result is the best possible in this case. We also used a Maple program to derive the formulas for lower bounds resulting from a dependency between three, four, or five rows; however, these lower bounds were inferior to those in Inequality (1.7) and Inequality (1.8).

We will now present another corollary of the main theorem that has a somewhat different flavor. In this corollary, we treat partially random matrices, which may have many deterministic rows. Our method allows us to obtain exponential bounds so long as there are still at most $c \ln n$ random rows, where $c > 0$ is a particular constant.

Corollary 1.1.2. *Let p be a real constant between 0 and 1, let c be any positive constant less than $1/\ln(1/p)$, and let $S \subset \mathbb{C}$ be a set of complex numbers having cardinality $|S| \leq O(1)$. Let $N_{\mathfrak{f}, n}$ be an n by n complex matrix in which $\mathfrak{f} \leq c \ln n$ rows contain fixed, non-random elements of S and where the other rows contain entries that are independent random variables taking values in S . If the fixed rows are linearly independent and if*

Asymptotic Upper and Lower Bounds for $\Pr\left(M_{\pm 1,n}^{(\mu)} \text{ is singular}\right)^{1/n}$ for $0 \leq \mu \leq 1$

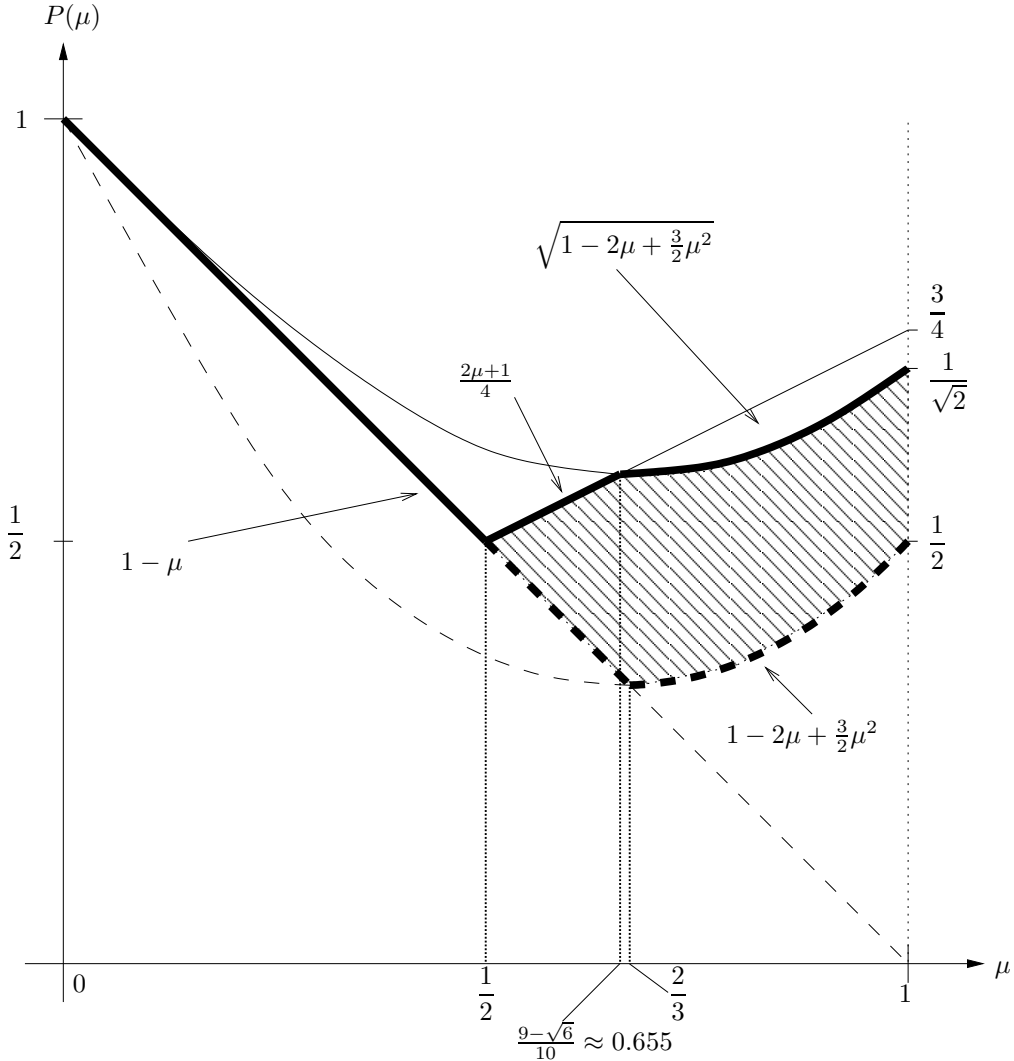


Figure 1.1: Let $P(\mu) := \lim_{n \rightarrow \infty} \Pr\left(M_{\pm 1,n}^{(\mu)} \text{ is singular}\right)^{1/n}$, where $M_{\pm 1,n}^{(\mu)}$ is the n by n matrix with independent random entries taking the value 0 with probability $1 - \mu$ and the values $+1$ and -1 each with probability $\mu/2$. The solid lines denote the upper bounds on $P(\mu)$ given by Inequalities (1.4), (1.5), and (1.6), and the dashed lines denote the lower bounds given by Inequalities (1.7) and (1.8). The upper and lower bounds coincide for $0 \leq \mu \leq \frac{1}{2}$, and the shaded area shows the difference between the best known upper and lower bounds for $\frac{1}{2} \leq \mu \leq 1$. The straight line segments from the point $(0, 1)$ to $(\frac{1}{2}, \frac{1}{2})$ and from the point $(\frac{1}{2}, \frac{1}{2})$ to $(1, \frac{3}{4})$ represent the best upper bounds we have derived using the ideas in [39], and the curve $1 - 2\mu + \frac{3}{2}\mu^2$ for $0 \leq \mu \leq 1$ represents a sometimes-better upper bound we have derived by adding a new idea. Note that the upper bounds given here also apply to the singularity probability of a random matrix with independent entries having arbitrary symmetric distributions in a set S of complex numbers, so long as each entry is 0 with probability $1 - \mu$ and the cardinality of S is $|S| \leq O(1)$ (see Corollary 2.2.1).

for every random entry α , we have $\max_x \Pr(\alpha = x) \leq p$, then

$$\Pr(N_{f,n} \text{ is singular}) \leq (\sqrt{p} + o(1))^n.$$

Notice that the case $f = 0$ and $p = 1/2$ also implies Inequality (1.2).

Remark 1.1.3 (Other exponential bounds). The focus of this thesis is optimizing the base of the exponent in bounds on the singularity probability for discrete random matrices. One main tool in this optimization is the use of a structure theorem similar to [39, Theorem 5.2] (see Theorem 4.2.1 below); however, using such a theorem requires additional assumptions to be placed on the values that can appear as entries, and in particular, this is why we assume in Corollary 1.1.2 that the set S has cardinality $|S| \leq O(1)$ and that $f \leq c \ln n$.

The structure of the rest of the thesis is as follows. In Section 2.1 we define p -bounded of exponent r and state the main theorem of this thesis. In Section 2.2, we discuss some corollaries of Theorem 2.1.2. In particular, we will:

- (A) prove Inequalities (1.4), (1.5), and (1.6);
- (B) prove general bounds on the singularity probability for discrete random matrices with entries that have symmetric distributions and with entries that have asymmetric distributions;
- (C) Prove a version of Corollary 1.1.2 (namely, Corollary 2.2.5) that holds for up to $o(n)$ fixed rows, assuming that the entries in the fixed rows take integer values between $-C$ and C for any positive constant C ; and
- (D) prove that the probability that random matrices with integer entries have a rational eigenvalue is exponentially small.

In Section 3.1.1, we discuss Lemma 3.1.1, a result that is proved in [44] using standard tools from algebraic number theory and algebraic geometry. Lemma 3.1.1 reduces the question of bounding the singularity probability of a random matrix with entries in \mathbb{C} to a question of bounding the singularity probability of a random matrix with entries

in $\mathbb{Z}/Q\mathbb{Z}$ for some large prime Q (in fact, it is possible to replace \mathbb{C} with any characteristic zero integral domain). Chapter 3 is devoted to proving Theorem 3.1.4 and demonstrating various applications of Theorem 3.1.4, including proving Lemma 3.1.1. The proof of Theorem 2.1.2 is outlined in Section 4.1, where we also prove some of the easier lemmas needed for the theorem. In Section 4.2, we state a structure theorem (Theorem 4.2.1) that completes the proof of our Theorem 2.1.2 and that is very similar to [39, Theorem 5.2] (which is the Structure Theorem in [39]). We discuss the proof of Theorem 4.2.1, which uses discrete Fourier analysis and tools from additive combinatorics, in Sections 5.1 and 5.2. Finally, in Section 6.1 we show that the entire argument proving Theorem 2.1.2 can be generalized to random complex matrices with f rows of the matrix containing fixed, non-random entries, so long as $f \leq c \ln n$ for a particular constant $c > 0$ (this leads to Corollary 1.1.2).

Chapter 2

Main result and applications

2.1 The general theorem

To prove the results in Inequalities (1.1) and (1.2) (and also the results in [22] and [38]), one basic idea is to replace entries of a random matrix with independent copies of the random variable $\gamma^{(\mu)}$ or $2\gamma^{(\mu)}$ (see Equation (1.3)). One key idea in proving the more general results of the current thesis is replacing the entries of a random matrix with more complicated symmetric discrete random variables.

A generalized arithmetic progression of rank τ is a set of the form $\{v_0 + m_1v_1 + \dots + m_\tau v_\tau : |m_i| \leq M_i/2\}$, where the v_i are elements of a \mathbb{Z} -module and the m_i and $M_i > 0$ are integers. Note that whenever the term “symmetric” is used in this thesis, it will apply to the distribution of a random variable or to a generalized arithmetic progression; in particular, the term will never apply to matrices. Also, throughout this thesis we will use the notation

$$e(x) := \exp(2\pi ix).$$

The following definition lies at the heart of our analysis.

Definition 2.1.1 (*p*-bounded of exponent *r*). Let *p* be a positive constant such that $0 < p < 1$ and let *r* be a positive integer constant. A random variable α taking values in the integers (or, respectively, the integers modulo some large prime *Q*) is *p*-bounded of exponent *r* if

$$(i) \quad \max_x \Pr(\alpha = x) \leq p, \text{ and}$$

if there exists a constant *q* where $0 < q \leq p$ and a \mathbb{Z} -valued (or, respectively, a $\mathbb{Z}/Q\mathbb{Z}$ -valued) symmetric random variable $\beta^{(\mu)}$ taking the value 0 with probability $1 - \mu = p$

such that the following two conditions hold:

(ii) $q \leq \min_x \Pr(\beta^{(\mu)} = x)$ and $\max_x \Pr(\beta^{(\mu)} = x) \leq p$, and

(iii) the following inequality holds for every $t \in \mathbb{R}$:

$$|\mathbb{E}(e(\alpha t))|^r \leq \mathbb{E}\left(e(\beta^{(\mu)} t)\right)$$

Here, if the values of α and $\beta^{(\mu)}$ are in $\mathbb{Z}/Q\mathbb{Z}$, we view those values as integers in the range $(-Q/2, Q/2)$ (note that each element in $\mathbb{Z}/Q\mathbb{Z}$ has a unique such integer representation).

We will define p -bounded of exponent r for collections of random variables below, but first we note that the conditions above are easy to verify in practice. In particular, if we have a symmetric random variable

$$\beta^{(\mu)} = \begin{cases} b_\ell & \text{with probability } \mu p_\ell/2 \\ \vdots & \vdots \\ b_1 & \text{with probability } \mu p_1/2 \\ 0 & \text{with probability } 1 - \mu \\ -b_1 & \text{with probability } \mu p_1/2 \\ \vdots & \vdots \\ -b_\ell & \text{with probability } \mu p_\ell/2, \end{cases} \quad (2.1)$$

where $b_s \in \mathbb{Z}$ for all s (or, respectively, $b_s \in \mathbb{Z}/Q\mathbb{Z}$ for all s), then condition (iii) becomes

$$|\mathbb{E}(e(\alpha t))|^r \leq \mathbb{E}\left(e(\beta^{(\mu)} t)\right) = 1 - \mu + \mu \sum_{s=1}^{\ell} p_s \cos 2\pi b_s t, \quad (2.2)$$

where the equality on the right-hand side is a simple expected value computation.

We say that a collection of random variables $\{\alpha_{jk}\}_{j,k=1}^n$ is p -bounded of exponent r if each α_{jk} is p -bounded of exponent r with the same constants p , q , and r ; and, importantly, the same value of $\mu = 1 - p$. We also make the critical assumption that the set of all values that can be taken by the $\beta_{jk}^{(\mu)}$ has cardinality $O(1)$ (a relaxation of this assumption is discussed in Remark 5.2.5). However, the definition of $\beta_{jk}^{(\mu)}$ is

otherwise allowed to vary with j and k . Also, we will use S to denote the set of all possible values taken by the random variables α_{jk} , and we will assume that the cardinality of S is at most $|S| \leq n^{o(n)}$.

If α takes non-integer values in \mathbb{C} , we need to map those values to a finite field of prime order so that we may use Definition 2.1.1, and for this task we will apply Lemma 3.1.1, which is proved in Chapter 3. We say that α is p -bounded of exponent r if and only if for each prime Q in an infinite sequence of primes produced by Lemma 3.1.1, we have $\phi_Q(\alpha)$ is p -bounded of exponent r , where ϕ_Q is the ring homomorphism described in Lemma 3.1.1 that maps S , the finite set of all possible values taken by the α_{jk} , into $\mathbb{Z}/Q\mathbb{Z}$ in such a way that for any matrix $N_n := (s_{jk})$ with entries in S , the determinant of N_n is zero if and only if the determinant of $\phi_Q(N_n) := (\phi_Q(s_{jk}))$ is zero.

Theorem 2.1.2. *Let p be a positive constant such that $0 < p < 1$, let r be a positive integer constant, and let S be a generalized arithmetic progression in the complex numbers with rank $O(1)$ (independent of n) and with cardinality at most $|S| \leq n^{o(n)}$. Let N_n be an n by n matrix with entries α_{jk} , each of which is an independent random variable taking values in S . If the collection of random variables $\{\alpha_{jk}\}_{1 \leq j, k \leq n}$ is p -bounded of exponent r , then*

$$\Pr(N_n \text{ is singular}) \leq (p^{1/r} + o(1))^n.$$

In the motivating examples of Section 1.1 (excluding Corollary 1.1.2), we discussed the case where the entries of the matrix are i.i.d.; however, in general the distributions of the entries are allowed to differ (and even depend on n), so long as the entries all take values in the same structured set S described above. The condition that S has additive structure seems to be an artifact of the proof (in particular, at certain points in the proof of Theorem 4.2.1, we need the set $\left\{ \sum_{j=1}^n x_j : x_j \in S \text{ for all } j \right\}$ to have cardinality at most $n^{o(n)}$). The easiest way to guarantee that S has the required structure is to assume that the set of values taken by all the α_{jk} has cardinality at most $O(1)$, and this is the approach we take for the corollaries in Section 2.2, since it also makes it easy to demonstrate that the collection of entries is p -bounded of exponent r .

Remark 2.1.3 (Strict positivity in Inequality (2.2)). Note that the constants μ, p_s, b_s must be such that the right-hand side of Equation (2.2) is non-negative. It turns out for the proof of Theorem 2.1.2 that we will need slightly more. At one point in the proof, we will apply Lemma 5.1.3, for which we must assume that there exists a very small constant $\epsilon_{-1} > 0$ such that $\mathbb{E}(e(\beta_{jk}^{(\mu)} t)) > \epsilon_{-1}$ for all t and for all $\beta_{jk}^{(\mu)}$ used in the definition of p -bounded of exponent r . Of course, if the expectations are not strictly larger than ϵ_{-1} , we can simply reduce μ by $\epsilon_{-1} > 0$. Then, since we are assuming $1 - \mu = p$, we clearly have that all the α_{jk} are $(p + \epsilon_{-1})$ -bounded of exponent r (by using $\beta_{jk}^{(\mu - \epsilon_{-1})}$ instead of $\beta_{jk}^{(\mu)}$) and we have that $\mathbb{E}(e(\beta_{jk}^{(\mu - \epsilon_{-1})} t)) > \epsilon_{-1} > 0$. Since Theorem 2.1.2 would thus yield a bound of $((p + \epsilon_{-1})^{1/r} + o(1))^n$ for every $\epsilon_{-1} > 0$, we can conclude a bound of $(p^{1/r} + o(1))^n$ by letting ϵ_{-1} tend to 0. Thus, without loss of generality, we will assume that $\mathbb{E}(e(\beta_{jk}^{(\mu)} t)) > \epsilon_{-1}$ for all t and for all $\beta_{jk}^{(\mu)}$ used in the definition of p -bounded of exponent r .

2.2 Some corollaries of Theorem 2.1.2

In this section, we will state a number of corollaries of Theorem 2.1.2, starting with short proofs of Inequalities (1.4), (1.5), and (1.6). The two most interesting results in this section will be more general: first (in Section 2.2.2), we will show an exponential bound on the singularity probability for a matrix with independent entries each a symmetric random variable taking values in $S \subset \mathbb{C}$, where $|S| \leq O(1)$ and assuming that each entry takes the value 0 with probability $1 - \mu$; and second (in Section 2.2.3), we will describe a similar (and sometimes better) bound when the condition that the random variables have symmetric distributions is replaced with the assumption that no entry takes a value with probability greater than p . In the first case, the bound will depend only the value of μ , and in the second case, the bound will depend only on the value of p . In Section 2.2.4, we will show an exponential bound on the singularity probability for an n by n matrix with $f = o(n)$ fixed rows containing small integer values and with the remaining rows containing independent random variables taking values in $S \subset \mathbb{C}$, where $|S| \leq O(1)$ (this is similar to Corollary 1.1.2, which is proved in Section 6.1). Finally, in Section 2.2.5, we will prove an exponential upper bound on the probability

that a random integer matrix has a rational eigenvalue.

In each corollary, we will use the definition of p -bounded of exponent 1 and of exponent 2. The definition of p -bounded of exponent 2 is particularly useful, since then the absolute value on the left-hand side of Inequality (2.2) is automatically dealt with; however, when μ is small (for example whenever $\mu \leq 1/2$), one can get better bounds by using p -bounded of exponent 1. We have not yet found an example where the best possible bound from Theorem 2.1.2 is found by using p -bounded of an exponent higher than 2.

2.2.1 Proving Inequalities (1.4), (1.5), and (1.6)

To prove Inequality (1.4), we note for $0 \leq \mu \leq \frac{1}{2}$ that (using the definition in Equation (1.3) of $\gamma^{(\mu)}$)

$$\left| \mathbb{E}(e(\gamma^{(\mu)}t)) \right| = 1 - \mu + \mu \cos(2\pi t),$$

and thus $\gamma^{(\mu)}$ is $(1 - \mu)$ -bounded of exponent 1 (i.e., take $\beta^{(\mu)} := \gamma^{(\mu)}$), and so Inequality (1.4) follows from Theorem 2.1.2.

To prove Inequality (1.5), we note for $\frac{1}{2} \leq \mu \leq 1$ that

$$\left| \mathbb{E}(e(\gamma^{(\mu)}t)) \right| = |1 - \mu + \mu \cos(2\pi t)| \leq \left(\frac{2\mu + 1}{4} \right) + (1 - \mu) \cos(2\pi t) + \left(\frac{2\mu - 1}{4} \right) \cos(4\pi t)$$

(the inequality above may be checked by squaring both sides and expanding as polynomials in $\cos(2\pi t)$). Thus, we can take

$$\beta^{(\mu)} := \begin{cases} +2 & \text{with probability } \frac{2\mu-1}{8} \\ -2 & \text{with probability } \frac{2\mu-1}{8} \\ +1 & \text{with probability } \frac{1-\mu}{2} \\ -1 & \text{with probability } \frac{1-\mu}{2} \\ 0 & \text{with probability } \frac{2\mu+1}{4} \end{cases}$$

to see that $\gamma^{(\mu)}$ is $\left(\frac{2\mu + 1}{4} \right)$ -bounded of exponent 1, and so Inequality (1.5) follows from Theorem 2.1.2.

To prove Inequality (1.6), we note for $0 \leq \mu \leq 1$ that

$$\left| \mathbb{E}(e(\gamma^{(\mu)}t)) \right|^2 = |1 - \mu + \mu \cos(2\pi t)|^2 = 1 - 2\mu + \frac{3}{2}\mu^2 + 2(1-\mu)\mu \cos(2\pi t) + \left(\frac{\mu^2}{2}\right) \cos(4\pi t).$$

Thus, we can take

$$\beta^{(\mu)} := \begin{cases} +2 & \text{with probability } \frac{\mu^2}{4} \\ -2 & \text{with probability } \frac{\mu^2}{4} \\ +1 & \text{with probability } (1-\mu)\mu \\ -1 & \text{with probability } (1-\mu)\mu \\ 0 & \text{with probability } 1 - 2\mu + \frac{3}{2}\mu^2 \end{cases}$$

to see that $\gamma^{(\mu)}$ is $\left(1 - 2\mu + \frac{3}{2}\mu^2\right)$ -bounded of exponent 2, and so Inequality (1.6) follows from Theorem 2.1.2.

2.2.2 Matrices with entries having symmetric distributions

In this subsection, we will prove a singularity bound for an n by n matrix $N_n^{(\mu)}$ for which each entry is a symmetric discrete random variable taking the value 0 with probability $1 - \mu$.

Corollary 2.2.1. *Let S be a set of complex numbers with cardinality $|S| \leq O(1)$. If $N_n^{(\mu)}$ is an n by n matrix in which each entry is an independent symmetric complex random variable taking values in S and taking the value 0 with probability $1 - \mu$, then*

$$\Pr(N_n^{(\mu)} \text{ is singular}) \leq \begin{cases} (1 - \mu + o(1))^n & \text{for } 0 \leq \mu \leq \frac{1}{2} \\ \left(\frac{2\mu+1}{4} + o(1)\right)^n & \text{for } \frac{1}{2} \leq \mu \leq 1 \\ \left(\sqrt{1 - 2\mu + \frac{3}{2}\mu^2} + o(1)\right)^n & \text{for } 0 \leq \mu \leq 1. \end{cases}$$

In particular, the same upper bounds as in Inequalities (1.4), (1.5), and (1.6) (which are shown in Figure 1.1) apply to the singularity probability for $N_n^{(\mu)}$.

Proof. Let α_{ij} be an entry of $N_n^{(\mu)}$. Since α_{ij} is symmetric and takes the value 0 with probability $1 - \mu$, we may write $\alpha_{ij} = \gamma_{ij}^{(\mu)} \eta_{ij}$, where $\gamma_{ij}^{(\mu)}$ is an independent copy of

$\gamma^{(\mu)}$ as defined in Equation (1.3) and η_{ij} is a random variable that shares no values with $-\eta_{ij}$. This description of α_{ij} was inspired by [1], and it allows us to condition on η_{ij} and then use the remaining randomness in $\gamma_{ij}^{(\mu)}$ to get a bound on the singularity probability. In particular,

$$\Pr(N_n^{(\mu)} \text{ is singular}) = \sum_{(c_{ij})} \Pr(N_n^{(\mu)} \text{ is singular} | \{\eta_{ij} = c_{ij}\}) \Pr(\{\eta_{ij} = c_{ij}\}),$$

where the sum runs over all (n^2) -tuples $(c_{ij})_{1 \leq i, j \leq n}$ of possible values taken by random variables η_{ij} . Since $\sum_{(c_{ij})} \Pr(\{\eta_{ij} = c_{ij}\}) = 1$, we can complete the proof by proving an exponential bound on $\Pr(N_n^{(\mu)} \text{ is singular} | \{\eta_{ij} = c_{ij}\})$, and we will use Theorem 2.1.2 for this task.

Consider the random matrix $N_n^{(\mu)} \Big|_{\{\eta_{ij}=c_{ij}\}}$, where the i, j entry is the random variable $c_{ij} \gamma_{ij}^{(\mu)}$ for some constant c_{ij} . Note that the entries of $N_n^{(\mu)} \Big|_{\{\eta_{ij}=c_{ij}\}}$ take values in S , a set with cardinality $O(1)$, and let ϕ_Q be the map from Lemma 3.1.1, which lets us pass to the case where $N_n^{(\mu)} \Big|_{\{\eta_{ij}=c_{ij}\}}$ has entries in $\mathbb{Z}/Q\mathbb{Z}$. Defining $\theta_{ij} := 2\pi\phi_Q(c_{ij})$, we compute

$$\begin{aligned} \left| \mathbb{E} e(\phi_Q(c_{ij} \gamma_{ij}^{(\mu)}) t) \right| &= |1 - \mu + \mu \cos(\theta_{ij} t)| \\ &\leq \begin{cases} 1 - \mu + \mu \cos(\theta_{ij} t) & \text{for } 0 \leq \mu \leq \frac{1}{2}, \\ \frac{2\mu+1}{4} + (1 - \mu) \cos(\theta_{ij} t) + \left(\frac{2\mu-1}{4}\right) \cos(2\theta_{ij} t) & \text{for } \frac{1}{2} \leq \mu \leq 1, \text{ and} \\ \left(1 - 2\mu + \frac{3}{2}\mu^2 + 2(1 - \mu)\mu \cos(\theta_{ij} t) + \frac{\mu^2}{2} \cos(2\theta_{ij} t)\right)^{1/2} & \text{for } 0 \leq \mu \leq 1. \end{cases} \end{aligned}$$

We have thus shown that the entries of $N_n^{(\mu)} \Big|_{\{\eta_{ij}=c_{ij}\}}$ are

$$\begin{aligned} &(1 - \mu)\text{-bounded of exponent 1 for } 0 \leq \mu \leq \frac{1}{2}, \\ &\left(\frac{2\mu+1}{4}\right)\text{-bounded of exponent 1 for } \frac{1}{2} \leq \mu \leq 1, \text{ and} \\ &\left(1 - 2\mu + \frac{3}{2}\mu^2\right)\text{-bounded of exponent 2 for } 0 \leq \mu \leq 1. \end{aligned}$$

Applying Theorem 2.1.2 completes the proof. \square

Corollary 2.2.1 is tight for $0 \leq \mu \leq \frac{1}{2}$, since the probability of a row of all zeroes

occurring is $(1 - \mu + o(1))^n$; however, for any specific case, Theorem 2.1.2 can usually prove better upper bounds than those given by Corollary 2.2.1.

For example, consider the case of a matrix $M_{\{\pm 2, \pm 1\}, n}^{(\mu)}$ with each entry an independent copy of the symmetric random variable

$$\alpha^{(\mu)} := \begin{cases} +2 & \text{with probability } \frac{\mu}{4} \\ -2 & \text{with probability } \frac{\mu}{4} \\ +1 & \text{with probability } \frac{\mu}{4} \\ -1 & \text{with probability } \frac{\mu}{4} \\ 0 & \text{with probability } 1 - \mu \end{cases}$$

Corollary 2.2.2. *For $M_{\{\pm 2, \pm 1\}, n}^{(\mu)}$ as defined above, we have*

$$\Pr(M_{\{\pm 2, \pm 1\}, n}^{(\mu)} \text{ is singular}) \leq \begin{cases} (1 - \mu + o(1))^n & \text{for } 0 \leq \mu \leq \frac{16}{25} \\ \left(\sqrt{1 - 2\mu + \frac{5}{4}\mu^2} + o(1)\right)^n & \text{for } 0 \leq \mu \leq 1. \end{cases}$$

Proof. By the definition of $\alpha^{(\mu)}$ we have

$$\left| \mathbb{E}e(\alpha^{(\mu)}t) \right| = 1 - \mu + \frac{\mu}{2} \cos(2\pi t) + \frac{\mu}{2} \cos(4\pi t), \quad \text{for } 0 \leq \mu \leq \frac{16}{25}$$

(i.e., the right-hand side of the equation above is non-negative for such μ), which proves the first bound.

Also, we have

$$\begin{aligned} \left| \mathbb{E}e(\alpha^{(\mu)}t) \right|^2 &= 1 - 2\mu + \frac{5}{4}\mu^2 + \left(\mu - \frac{3}{4}\mu^2\right) \cos(2\pi t) + \left(\mu - \frac{7}{8}\mu^2\right) \cos(4\pi t) \\ &\quad + \frac{\mu^2}{4} \cos(6\pi t) + \frac{\mu^2}{8} \cos(8\pi t) \end{aligned}$$

for $0 \leq \mu \leq 1$, which proves the second bound. \square

We also have the following lower bounds for the singularity probability of $M_{\{\pm 2, \pm 1\}, n}^{(\mu)}$:

$$(1 - \mu + o(1))^n \quad \text{(from one row of all zeroes)} \quad (2.3)$$

$$(1 - 2\mu + 5\mu^2/4 + o(1))^n \quad \text{(from a two-row dependency)} \quad (2.4)$$

Asymptotic Upper and Lower Bounds for $\Pr\left(M_{\{\pm 2, \pm 1\}, n}^{(\mu)} \text{ is singular}\right)^{1/n}$ for $0 \leq \mu \leq 1$

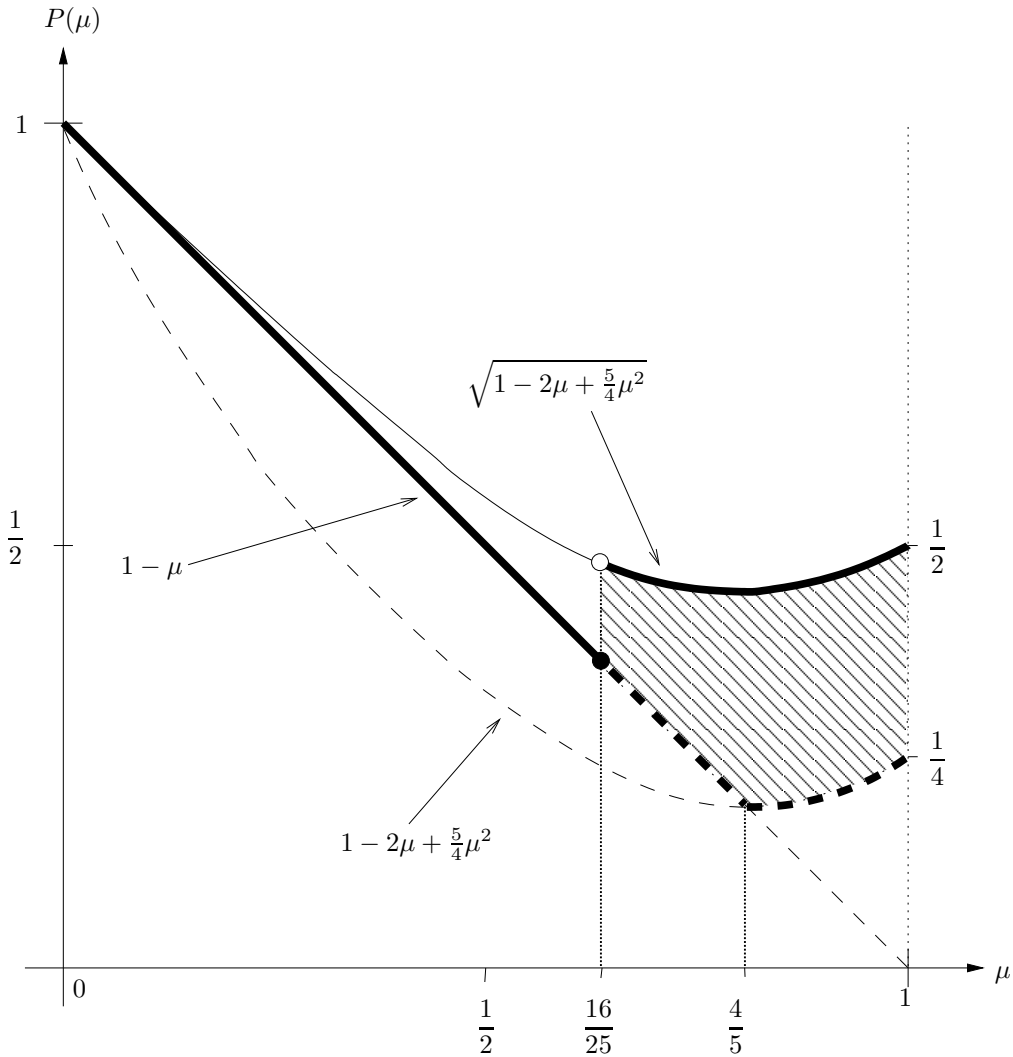


Figure 2.1: Let $P(\mu) := \lim_{n \rightarrow \infty} \Pr\left(M_{\{\pm 2, \pm 1\}, n}^{(\mu)} \text{ is singular}\right)^{1/n}$, where $M_{\{\pm 2, \pm 1\}, n}^{(\mu)}$ is the n by n matrix with independent random entries taking the value 0 with probability $1 - \mu$ and the values $+2, -2, +1, -1$ each with probability $\mu/4$. This figure summarizes the upper bounds on $P(\mu)$ from Corollary 2.2.2 and the lower bounds from Displays (2.3) and (2.4). The best upper bounds (shown in thick solid lines) match the best lower bounds (thick dashed lines) for $0 \leq \mu \leq \frac{16}{25}$; and it is not hard to improve the upper bound a small amount by finding a bound (of exponent 1) to bridge the discontinuity. One should note that even as stated above, the upper bounds are substantially better than those given by Corollary 2.2.1 (which are shown in Figure 1.1). The shaded area represents the gap between the upper and lower bounds.

The results of Corollary 2.2.2 and the corresponding lower bounds are shown in Figure 2.1, and one should note that the upper bounds are substantially better than those guaranteed by Corollary 2.2.1.

2.2.3 Random matrices with entries having arbitrary distributions

A useful feature of the definition of p -bounded of exponent 2 is that it lets one bound the singularity probability of matrices with independent discrete random variables that are asymmetric.

Corollary 2.2.3. *Let p be a constant such that $0 < p \leq 1$ and let $S \subset \mathbb{C}$ be a set with cardinality $|S| \leq O(1)$. If N_n is an n by n matrix with independent random entries taking values in S such that for any entry α , we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(N_n \text{ is singular}) \leq (\sqrt{p} + o(1))^n.$$

We will need the following slightly more general corollary in Section 2.2.4. For a set A and an integer m , we will use the notation $mA := \{\sum_{j=1}^m a_j : a_j \in A\}$ and $A^m := \{\prod_{j=1}^m a_j : a_j \in A\}$.

Corollary 2.2.4. *Let p be a constant such that $0 < p \leq 1$, let $S \subset \mathbb{C}$ be a set with cardinality $|S| \leq O(1)$, and let X_n be an n by n matrix with fixed, non-random entries in $n^{o(n)}(S \cup \{-1, 0, 1\})^{O(1)}$. If N_n is an n by n matrix with independent random entries taking values in S such that for any entry α , we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(X_n + N_n \text{ is singular}) \leq (\sqrt{p} + o(1))^n.$$

Note that that Corollary 2.2.4 implies Corollary 2.2.3 by taking X_n to be the matrix of all zeroes.

Proof of Corollary 2.2.4. Let α_{ij} be an entry in N_n . Our goal is to describe α_{ij} in a two-step random process, condition on one of the steps, and then use the randomness in the other step to bound the singularity probability. The conditioning approach is the same as that used in the symmetric case (Corollary 2.2.1) and was inspired by [1]. The conditioning argument is useful since some entries of the random matrix may take some

values with very small probability (i.e. probability less than any constant); recall that while the entries of the random matrix always take values in a fixed set S of cardinality $O(1)$, the distributions of those random variables within S are allowed to vary with n . (Note that making use of Remark 5.2.5 would provide an alternate way of dealing with entries that take some values with very small probability.)

Say that α_{ij} takes the values v_1, \dots, v_t with probabilities $\varrho_1, \dots, \varrho_t$, respectively, where $\varrho_1 \geq \varrho_2 \geq \dots \geq \varrho_t$. Define new random variables η_{ijk} such that for some i_0 and i_1 , the values taken by η_{ijk} are $v_{i_0}, v_{i_0+1}, \dots, v_{i_0+i_1}$ with corresponding probabilities $\varrho_{i_0}/p_k, \varrho_{i_0+1}/p_k, \dots, \varrho_{i_0+i_1}/p_k$, where $p_k := \sum_{i=1}^{i_1} \varrho_{i_0+i}$. Thus, we can write

$$\alpha_{ij} = \begin{cases} \eta_{ij1} & \text{with probability } p_1 \\ \eta_{ij2} & \text{with probability } p_2 \\ \vdots & \vdots \\ \eta_{ij\ell} & \text{with probability } p_\ell. \end{cases} \quad (2.5)$$

Furthermore, the η_{ijk} can be constructed so that $p_k \leq p$ for every k , so that $p/2 \leq p_k$ for $1 \leq k \leq \ell - 1$, and so that no two η_{ijk} with different k 's ever take the same value.

There are two cases to consider for the technical reason that p_ℓ is not necessarily bounded below by a constant. Let $\epsilon > 0$ be a very small constant, so for example $p/2 > \epsilon$. Case 1 is when $\epsilon \leq p_\ell$, and in this case each p_k is bounded below by ϵ and above by p . We will consider Case 1 first and then discuss the small changes needed to deal with Case 2.

As in the proof of Corollary 2.2.1, we will condition on the values taken by the η_{ijk} in order to prove a bound on the singularity probability. We have that

$$\Pr(X_n + N_n \text{ is singular}) = \sum_{(c_{ijk})} \Pr(X_n + N_n \text{ is singular} | \{\eta_{ijk} = c_{ijk}\}) \Pr(\{\eta_{ijk} = c_{ijk}\}),$$

where the sum runs over all possible values (c_{ijk}) that the η_{ijk} can take. Thus, it is sufficient to prove a bound on the singularity probability for the random matrix

$X_n + N_n \Big|_{\{\eta_{ijk}=c_{ijk}\}}$ which has random entries

$$x_{ij} + \tilde{\alpha}_{ij} = \begin{cases} x_{ij} + c_{ij1} & \text{with probability } p_1 \\ x_{ij} + c_{ij2} & \text{with probability } p_2 \\ \vdots & \vdots \\ x_{ij} + c_{ij\ell} & \text{with probability } p_\ell, \end{cases}$$

where x_{ij} and the c_{ijk} are constants.

Note the entries of $X_n + N_n \Big|_{\{\eta_{ijk}=c_{ijk}\}}$ take values in $n^{o(n)} (S \cup \{-1, 0, 1\})^{O(1)}$, a generalized arithmetic progression with rank $O(1)$ and cardinality at most $n^{o(n)}$, and let ϕ_Q be the map from Lemma 3.1.1, which lets us pass to the case where $X_n + N_n \Big|_{\{\eta_{ijk}=c_{ijk}\}}$ has entries in $\mathbb{Z}/Q\mathbb{Z}$. Defining $\theta_{ijk} := 2\pi\phi_Q(c_{ijk})$ and letting $\tilde{\alpha}'_{ij}$ be an i.i.d. copy of $\tilde{\alpha}_{ij}$, we compute

$$\begin{aligned} |\mathbb{E}e(\phi_Q(x_{ij} + \tilde{\alpha}_{ij})t)|^2 &= \mathbb{E}e\left(\phi_Q(x_{ij} + \tilde{\alpha}_{ij} - x_{ij} - \tilde{\alpha}'_{ij})t\right) = \mathbb{E}e\left(\phi_Q(\tilde{\alpha}_{ij} - \tilde{\alpha}'_{ij})t\right) \\ &= \sum_{k=1}^{\ell} p_k^2 + 2 \sum_{1 \leq k_1 < k_2 \leq \ell} p_{k_1} p_{k_2} \cos((\theta_{ijk_1} - \theta_{ijk_2})t). \end{aligned}$$

Thus, $x_{ij} + \tilde{\alpha}_{ij}$ is $\left(\sum_{k=1}^{\ell} p_k^2\right)$ -bounded of exponent 2 (using the constant $q = \epsilon^2$ in Definition 2.1.1, so q does not depend on n). Given that $0 < p_k \leq p$ for every k , it is not hard to show that $\sum_{k=1}^{\ell} p_k^2 \leq p < p + \epsilon$, and so from Definition 2.1.1, we see that the collection $\{x_{ij} + \tilde{\alpha}_{ij} : \tilde{\alpha}_{ij} \text{ has corresponding probability } p_\ell \geq \epsilon\}$ is $(p + \epsilon)$ -bounded of exponent 2. We are thus finished with Case 1.

Case 2 is when the decomposition of α_{ij} given in Equation (2.5) has $p_\ell < \epsilon$. In this case we need to modify Equation (2.5) slightly, deleting $\eta_{ij\ell}$ and replacing $\eta_{ij(\ell-1)}$ with a new variable $\eta'_{ij(\ell-1)}$ that takes all the values previously taken by $\eta_{ij\ell}$ and by $\eta_{ij(\ell-1)}$ with the appropriate probabilities. Thus, in Case 2, we have that $p/2 \leq p_k < p + \epsilon$ for all $1 \leq k \leq \ell - 1$, showing that each p_k is bounded below by a constant and is bounded above by $p + \epsilon$ (here we are using $p_{\ell-1}$ to denote the probability that α_{ij} draws a value from the random variable $\eta'_{ij(\ell-1)}$).

For Case 2, we use exactly the same reasoning as in Case 1 above to show that such entries of $X_n + N_n \Big|_{\{\eta_{ijk}=c_{ijk}\}}$ are $\left(\sum_{k=1}^{\ell-1} p_k^2\right)$ -bounded of exponent 2 (using the

constant $q = \epsilon^2 < p^2/4$ in Definition 2.1.1, so q does not depend on n). Noting that $\sum_{k=1}^{\ell-1} p_k^2 < p + \epsilon$ and using Definition 2.1.1, we see that the collection $\{x_{ij} + \tilde{\alpha}_{ij} : \tilde{\alpha}_{ij}$ has corresponding probability $p_\ell < \epsilon\}$ is $(p + \epsilon)$ -bounded of exponent 2.

Combining Case 1 and Case 2, we have that the collection $\{x_{ij} + \tilde{\alpha}_{ij}\}$ is $(p + \epsilon)$ -bounded of exponent 2, and so by and by Theorem 2.1.2 we have that $\Pr(X_n + N_n \Big|_{\{\eta_{ijk}=c_{ijk}\}} \text{ is singular}) \leq (\sqrt{p + \epsilon} + o(1))^n$.

The constant $\epsilon > 0$ was chosen arbitrarily, and so letting ϵ tend to zero, we get that

$$\Pr(X_n + N_n \text{ is singular} | \{\eta_{ijk} = c_{ijk}\}) \leq (\sqrt{p} + o(1))^n.$$

□

2.2.4 Partially random matrices

In this subsection, we prove a bound on the singularity probability for partly random matrices where many rows are deterministic.

Corollary 2.2.5. *Let p be a real constant between 0 and 1, let K be a large positive constant, and let $S \subset \mathbb{C}$ be a set of complex numbers having cardinality $|S| \leq K$. Let $N_{\mathfrak{f},n}$ be an n by n matrix in which \mathfrak{f} rows contain fixed, non-random integers between $-K$ and K and where the other rows contain entries that are independent random variables taking values in S . If $\mathfrak{f} \leq o(n)$, if the \mathfrak{f} fixed rows are linearly independent, and if for every random entry α , we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(N_{\mathfrak{f},n} \text{ is singular}) \leq (\sqrt{p} + o(1))^{n-\mathfrak{f}}.$$

Corollary 2.2.5 applies to partly random matrices with $\mathfrak{f} = o(n)$ fixed, non-random rows containing integers bounded by a constant and with random entries taking at most $O(1)$ values in the complex numbers. Corollary 1.1.2, on the other hand, holds with the fixed entries also allowed to take values in the complex numbers and gives a slightly better bound, but additionally requires $\mathfrak{f} \leq O(\ln n)$ (which is far smaller in general than $o(n)$). Proving Corollary 1.1.2 requires mirroring the entire argument used to prove the main theorem (Theorem 2.1.2) in the case where \mathfrak{f} rows contain fixed, non-random entries, and we discuss this argument in Section 6.1. Proving Corollary 2.2.5, however,

can be done directly from Theorem 2.1.2, as we will show below. First, we will state a generalization of Corollary 2.2.5.

Corollary 2.2.6. *Let p be a real constant between 0 and 1, let K be a large positive constant, and let $S \subset \mathbb{C}$ be a set of complex numbers having cardinality $|S| \leq K$. Let $N_{\mathfrak{f},n}$ be an n by n matrix in which \mathfrak{f} rows contain fixed, non-random integers between $-K$ and K and where the other rows contain entries that are independent random variables taking values in S . If $\mathfrak{f} \leq o(n)$, if the fixed rows have co-rank k , and if for every random entry α , we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(N_{\mathfrak{f},n} \text{ has co-rank } > k) \leq (\sqrt{p} + o(1))^{n-\mathfrak{f}}.$$

To obtain Corollary 2.2.6 from Corollary 2.2.5, find a collection \mathcal{C} of $\mathfrak{f} - k$ linearly independent rows among the deterministic rows. Replace the rest of the deterministic rows with a collection \mathcal{C}' of rows containing integer values between $-K$ and K such that \mathcal{C}' is linearly independent from \mathcal{C} . Finally, apply Corollary 2.2.5 to the new partially random matrix whose deterministic rows are from $\mathcal{C} \cup \mathcal{C}'$, thus proving Corollary 2.2.6.

Proof of Corollary 2.2.5. By reordering the rows and columns, we may write

$$N_{\mathfrak{f},n} = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right),$$

where A is an \mathfrak{f} by \mathfrak{f} non-random invertible matrix, B is an \mathfrak{f} by $n - \mathfrak{f}$ non-random matrix, C is an $n - \mathfrak{f}$ by \mathfrak{f} random matrix, and D is an $n - \mathfrak{f}$ by $n - \mathfrak{f}$ random matrix. Note that $N_{\mathfrak{f},n}$ is singular if and only if there exists a vector \mathbf{v} such that $N_{\mathfrak{f},n}\mathbf{v} = 0$. Let \mathbf{v}_1 be the first \mathfrak{f} coordinates of \mathbf{v} and let \mathbf{v}_2 be the remaining $n - \mathfrak{f}$ coordinates. Then $N_{\mathfrak{f},n}\mathbf{v} = 0$ if and only if

$$\begin{cases} A\mathbf{v}_1 + B\mathbf{v}_2 = 0, \text{ and} \\ C\mathbf{v}_1 + D\mathbf{v}_2 = 0. \end{cases}$$

Since A is invertible, these two equations are satisfied if and only if $(-CA^{-1}B + D)\mathbf{v}_2 = 0$, that is, if and only if the random matrix $-CA^{-1}B + D$ is singular.

We want to show that every entry that can appear in $-CA^{-1}B$ is an element of $n^{o(n)}(S \cup \{-1, 0, 1\})^{O(1)}$. By the cofactor formula for A^{-1} , we know that the i, j entry

of A^{-1} is $(-1)^{i+j}(\det A_{ij})/\det A$, where A_{ij} is the $f-1$ by $f-1$ matrix formed by deleting the i -th row and j -th column of A . Thus, $A^{-1} = \frac{1}{\det A} \tilde{A}$, where the i, j entry of \tilde{A} is $(-1)^{i+j} \det A_{ij}$. By the volume formula for the determinant, we know that $|\det A|$ is at most the product of the lengths of the row vectors of A ; and thus $|\det A| \leq n^{o(n)}$ (here we need that A has integer entries between $-K$ and K , where K is a constant, and that $f \leq o(n)$). Similarly, we have $|\det A_{ij}| \leq n^{o(n)}$. Every entry of \tilde{A} is thus in $n^{o(n)}\{-1, 0, 1\}$, every entry of C is in S , and every entry of B is in $O(1)\{-1, 0, 1\}$; thus, every entry of $-C\tilde{A}B$ is an element of $n^{o(n)}(S \cup \{-1, 0, 1\})$.

Conditioning on the values taken by all the entries in C , we have

$$\begin{aligned} \Pr(N_{f,n} \text{ is singular}) &= \Pr(-CA^{-1}B + D \text{ is singular}) \\ &= \sum_{(c_{ij})} \Pr(-CA^{-1}B + D \text{ is singular} | C = (c_{ij})) \Pr(C = (c_{ij})), \end{aligned} \tag{2.6}$$

where the sum runs over all possible matrices (c_{ij}) that C can produce. Considering the entries in $C = (c_{ij})$ to be fixed (note that A and B are fixed by assumption), we now need to bound

$$\Pr(-(c_{ij})A^{-1}B + D \text{ is singular}) = \Pr(-(c_{ij})\tilde{A}B + (\det A)D \text{ is singular}).$$

Note that every entry of $-(c_{ij})\tilde{A}B$ is an element of $n^{o(n)}(S \cup \{-1, 0, 1\})^{O(1)}$ and that the random matrix $(\det A)D$ has entries that take values in the fixed set $\{(\det A)s : s \in S\}$ having cardinality $O(1)$. Thus, by Corollary 2.2.4, we have that

$$\Pr(-(c_{ij})\tilde{A}B + (\det A)D \text{ is singular}) \leq (\sqrt{p} + o(1))^{n-f}.$$

Plugging this bound back into Equation (2.6) completes the proof. \square

2.2.5 Integer matrices and rational eigenvalues

Let η_k be the random variable taking the values $-k, -k+1, \dots, k-1, k$ each with equal probability, and let M_n be the n by n matrix where each entry is an independent copy of η_k . In [27], Martin and Wong show that for any $\epsilon > 0$,

$$\Pr(M_n \text{ has a rational eigenvalue}) \leq \frac{c(n, \epsilon)}{k^{1-\epsilon}},$$

where $c(n, \epsilon)$ is a constant depending on n and ϵ . (One goal in [27] is to study this bound as k goes to ∞ while n is fixed, which is why $c(n, \epsilon)$ is allowed to depend on n .)

Below, we prove a similar result for random integer matrices with entries between $-k$ and k (with k fixed), where we allow each entry to have a different (independent) distribution and we also allow the distributions to be very general.

Corollary 2.2.7. *Fix a positive integer k , and let $M_{k,n}$ be a random integer matrix with independent entries, each of which takes values in the set $\{-k, -k+1, \dots, k-1, k\}$. Let c be a constant such that for every entry α , we have $\max_{-k \leq x \leq k} \Pr(\alpha = x) \leq c/k$. Then*

$$\Pr(M_{k,n} \text{ has a rational eigenvalue}) \leq \left(\frac{c}{k} + o(1)\right)^{n/2},$$

where the $o(1)$ term goes to zero as n goes to ∞ .

For example, in the case where each independent entry has the uniform distribution on $\{-k, -k+1, \dots, k-1, k\}$ (as in [27]), one can set $c = 1/2$ in the corollary above.

Proof. The proof given below follows the same outline as the main theorem of [27], with Corollary 1.1.2 replacing an appeal to [27, Lemma 1].

The characteristic polynomial for $M_{k,n}$ is monic with integer coefficients, and thus the only possible rational eigenvalues are integers (by the rational roots theorem). Every eigenvalue of $M_{k,n}$ has absolute value at most nk (see [27, Lemma 4]); thus, the only possible integer eigenvalues are between $-nk$ and nk .

The matrix $M_{k,n}$ has λ as an eigenvalue if and only if $M_{k,n} - \lambda I$ is singular (where I is the n by n identity matrix). By Corollary 1.1.2 (with $f = 0$), we have

$$\Pr(M_{k,n} - \lambda I \text{ is singular}) \leq \left(\sqrt{\frac{c}{k}} + o(1)\right)^n.$$

Using the union bound, we have

$$\begin{aligned} \Pr(M_{k,n} \text{ has a rational eigenvalue}) &= \Pr(M_{k,n} - \lambda I \text{ is singular, for some } \lambda \in \{-nk, \dots, nk\}) \\ &\leq \sum_{\lambda=-nk}^{nk} \Pr(M_{k,n} - \lambda I \text{ is singular}) \\ &\leq (2nk + 1) \left(\sqrt{\frac{c}{k}} + o(1)\right)^n \\ &\leq \left(\frac{c}{k} + o(1)\right)^{n/2}. \end{aligned}$$



Chapter 3

Reduction to a finite field, with other applications

3.1 Introduction

3.1.1 Placing complex matrices in a finite field

The original work on discrete random matrices in [25, 22, 38, 39] is concerned with matrices having integer entries, which can also be viewed as matrices with entries in $\mathbb{Z}/Q\mathbb{Z}$ where Q is a very large prime. In this section we show that one can pass from a (random) matrix with entries in \mathbb{C} to one with entries in $\mathbb{Z}/Q\mathbb{Z}$ where Q is an arbitrarily large prime number, all without affecting the probability that the determinant is zero, thanks to the following lemma.

Lemma 3.1.1. *Let S be a finite subset of \mathbb{C} . There exist infinitely many primes Q such that there is a ring homomorphism $\phi_Q : \mathbb{Z}[S] \rightarrow \mathbb{Z}/Q\mathbb{Z}$ satisfying the following two properties:*

(i) *the map ϕ_Q is injective on S , and*

(ii) *for any n by n matrix $(s_{ij})_{1 \leq i, j \leq n}$ with entries $s_{ij} \in S$, we have*

$$\det \left((s_{ij})_{1 \leq i, j \leq n} \right) = 0 \quad \text{if and only if} \quad \det \left((\phi_Q(s_{ij}))_{1 \leq i, j \leq n} \right) = 0.$$

In order to apply this lemma, let us point out that the proof of Theorem 2.1.2, which is discussed in Sections 4.1 through 5.2, works exclusively in $\mathbb{Z}/Q\mathbb{Z}$; though at various points, it is necessary to assume Q is extremely large with respect to n and various constants. For this thesis, S will be the set of all possible values taken by the random variables α_{jk} . Recall that by assumption, $|S| \leq n^{o(n)}$, so in particular, S is finite.

Remark 3.1.2 (On the size of Q). When we apply Lemma 3.1.1, we will take $Q > \exp(\exp(Cn))$ for some constant C in order for Freiman-type theorems such as [39, Theorem 6.3] (which is restated in Theorem 5.2.1 below) to apply, and we will also choose Q large enough so that the integral approximation in Inequality (A.1) holds and so that Q is large with respect to various constants. One should note that while Q can be taken arbitrarily large with respect to n , we cannot choose Q so that it is arbitrarily large with respect to $\phi_Q(s)$ for all $s \in S$, where S is the set of all values that could appear in the given random matrix. For example, if $\sqrt{2} \in S$, then the smallest positive integer representative for $\phi_Q(\sqrt{2})$ must be larger than \sqrt{Q} (since $(\phi_Q(\sqrt{2}))^2 = 2$ in $\mathbb{Z}/Q\mathbb{Z}$). Finally, if we were in a situation where $S \subset \mathbb{Q}$, then we could avoid using Lemma 3.1.1 altogether by clearing denominators to pass to \mathbb{Z} and then taking $Q \approx \exp(\exp(Cn))$, as is done in [39].

Lemma 3.1.1 is a corollary of Theorem 3.1.4 which we will state in the next subsection and which we will prove in Section 3.7. We will prove Lemma 3.1.1 in Section 3.5. The remainder of this chapter contains further applications of Theorem 3.1.4, for example proving a sum-product result for the complex numbers and proving a Szemerédi-Trotter-type result for the complex numbers, where the applications follow from the analogous results for \mathbb{Z}/Q where Q is a prime (see [5]). All of the results in the complex numbers that are corollaries of Theorem 3.1.4, including Lemma 3.1.1, go through with the complex numbers replaced by any characteristic zero integral domain. Thus, the results stated in Sections 1.1, 2.1, and 2.2 above for the complex numbers \mathbb{C} also all go through with \mathbb{C} replaced by any characteristic zero integral domain. For example, Corollary 2.2.3 becomes

Corollary 3.1.3. *Let p be a constant such that $0 < p \leq 1$ and let D be a characteristic zero integral domain. Let $S \subset D$ have cardinality $|S| \leq O(1)$. If N_n is an n by n matrix with independent random entries, each taking values in S , such that for every entry α , we have $\max_x \Pr(\alpha = x) \leq p$, then*

$$\Pr(N_n \text{ is singular}) \leq (\sqrt{p} + o(1))^n.$$

3.1.2 Mapping incidences to a finite field

Many problems and results in arithmetic combinatorics deal with algebraic incidences in a finite set S . Classical examples are the Szemerédi-Trotter theorem and sum-product estimates.

A well-studied situation is when S is a subset of $\mathbb{Z}/Q\mathbb{Z}$, the finite field with Q elements where Q is a large prime. In this case, the special structure of the field and powerful techniques such as discrete Fourier analysis provide many tools to attack these problems. These features are not available in other settings and it seems one needs to invent new tricks. For example, when S is a subset of the complex numbers, most studies previous to this thesis relied on some very clever use of properties of the plane. Thus, it seems desirable to have a tool that reduces a problem from a general setting to the special case of $\mathbb{Z}/Q\mathbb{Z}$.

Such a tool exists, if one only cares about the linear relations among the elements of S . In this case, the famous Freiman isomorphism lemma (see, for example, [40, Lemma 5.25]) asserts that any finite subset of an arbitrary torsion-free group can be mapped into $\mathbb{Z}/Q\mathbb{Z}$, given that Q is sufficiently large, preserving all additive (linear) relations in S . Thanks to this result, it has now become a common practice in additive combinatorics to reduce additive problems from a general torsion-free group to $\mathbb{Z}/Q\mathbb{Z}$.

The goal of this chapter is to show that the desired reduction is possible in general. Technically speaking, we prove that any finite set S in a characteristic zero integral domain can be mapped to $\mathbb{Z}/Q\mathbb{Z}$, for infinitely many primes Q , preserving all algebraic incidences in S .

Some notable characteristic zero integral domains include the integers, the complex numbers, and the field of rational functions $\mathbb{C}(t_1, t_2, \dots)$ in any number of formal variables t_i . As applications, we obtain some new results and short proofs of some known results. In particular, it is shown that sum-product estimates and bounds for incidence geometry problems over $\mathbb{Z}/Q\mathbb{Z}$ imply the same bounds for the analogous problems over any characteristic zero integral domain (including the real and complex numbers).

Throughout this chapter, we assume that all rings are commutative with identity 1

and that all ring homomorphisms take 1 to 1. Let D be a characteristic zero integral domain (so D is a commutative ring with identity that has no zero divisors). We will identify the subring of D generated by the identity with the integers \mathbb{Z} (since the two are isomorphic). For a subset S of D , we will use $\mathbb{Z}[S]$ to denote the smallest subring of D containing S .

Theorem 3.1.4 (Reduction Theorem). *Let S be a finite subset of a characteristic zero integral domain D , and let L be a finite set of non-zero elements in the subring $\mathbb{Z}[S]$ of D . There exists an infinite sequence of primes with positive relative density such that there is a ring homomorphism $\phi_Q : \mathbb{Z}[S] \rightarrow \mathbb{Z}/Q\mathbb{Z}$ satisfying $0 \notin \phi_Q(L)$.*

By *positive relative density*, we mean that the sequence has positive density in the sequence of all primes. It is important to note that Theorem 3.1.4 is not true for all primes. For example, if $S = \{i\} \subset \mathbb{C}$, then the desired map does not exist for $Q = -1 \pmod{4}$, since the equation $x^2 = -1$ is not solvable in $\mathbb{Z}/Q\mathbb{Z}$ for these Q . Note that for the applications of Theorem 3.1.4 in this thesis, we only need that there exist infinitely many primes such that a map ϕ_Q exists, which follows from those primes having positive relative density.

The remainder of this chapter is organized as follows. In the next few sections, we present few sample applications of Theorem 3.1.4. Combining arguments from [5] with Theorem 3.1.4, we prove a Szemerédi-Trotter-type result in Section 3.2. In Section 3.3, we use Theorem 3.1.4 to demonstrate a sum-product estimate for characteristic zero integral domains, based on well-known sum-product estimates in $\mathbb{Z}/Q\mathbb{Z}$. Section 3.4 is focused on combining a product result for $\mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})$ from [19] with Theorem 3.1.4 to get an analogous product result for $\mathrm{SL}_2(D)$, where D is a characteristic zero integral domain. In Section 3.5, we show that a random matrix taking finitely many values in a characteristic zero integral domain is singular with exponentially small probability. This extends earlier results on integer matrices to the complex setting. Finally, the proof of Theorem 3.1.4 is given in Section 3.7.

3.2 A Szemerédi-Trotter-type result for characteristic zero integral domains

In this section, we apply Theorem 3.1.4 to the problem of bounding the maximum number of incidences between a finite set of lines and a finite set of points. The well-known Szemerédi-Trotter Theorem [37] solves this problem in the case of points and lines in $\mathbb{R} \times \mathbb{R}$. Recently, in [5], an analogous result was proven for $\mathbb{Z}/q \times \mathbb{Z}/q$ where q is a prime.

Theorem 3.2.1 (Theorem 6.2 in [5]). *Let q be a prime, and let \mathcal{P} and \mathcal{L} be sets of points and lines, respectively, in $\mathbb{Z}/q \times \mathbb{Z}/q$ such that the cardinalities $|\mathcal{P}|, |\mathcal{L}| \leq N \leq q$. Then there exist positive absolute constants c and δ such that*

$$\left| \{(p, \ell) \in \mathcal{P} \times \mathcal{L} : p \in \ell\} \right| \leq cN^{3/2-\delta}. \quad (3.1)$$

Remark 3.2.2. The original version of Theorem 3.2.1 proven in [5] relied on the best known sum-product result at the time (also found in [5]), which worked only for subsets of \mathbb{Z}/q with cardinality between q^α and $q^{1-\alpha}$ for a constant α . In particular, the proof in [5] assumed that Inequality (3.1) was false and used this assumption to construct a subset A of \mathbb{Z}/q with cardinality $N^{1/2-C\delta}$, for some constant C , such that $\max\{|A + A|, |AA|\}$ was small, a contradiction of the sum-product estimate proven in [5]. Thus, the version of Theorem 3.2.1 in [5] required the additional assumption that $N = q^\alpha$ for a constant α . To prove Theorem 3.2.1 as stated above, one can simply replace the sum-product results in [5] by more recent estimates that apply for all subsets of \mathbb{Z}/q (for example, [6, 18, 23]).

In a general ring R , we define a line to be the set of solutions (x, y) in $R \times R$ to an equation $y = mx + b$, where m and b are fixed elements of R . Using Theorem 3.1.4, we prove that the same bound as in Theorem 3.2.1 holds for an arbitrary characteristic zero integral domain:

Theorem 3.2.3. *Let D be a characteristic zero integral domain, and let \mathcal{P} and \mathcal{L} be sets of points and lines (respectively) in $D \times D$ with cardinalities $|\mathcal{P}|, |\mathcal{L}| \leq N$. Then*

there exist positive absolute constants c and δ such that

$$\left| \{(p, \ell) \in \mathcal{P} \times \mathcal{L} : p \in \ell\} \right| \leq cN^{3/2-\delta}.$$

The constants c and δ are the same as those in Theorem 3.2.1. Any improvement to Theorem 3.2.1, for example, better constants or giving a good bound when \mathcal{P} and \mathcal{L} have very different cardinalities, would also immediately translate to Theorem 3.2.3 above. In the case of $\mathbb{R} \times \mathbb{R}$, this theorem is true with δ being replaced with the optimal constant $1/6$ (by the Szemerédi-Trotter Theorem [37]).

Restricting to the case of complex numbers, Solymosi [33, Lemma 1] has proven a Szemerédi-Trotter-type result over \mathbb{C} with $\delta = 1/6$, under the additional assumption that the set of points form a Cartesian product in \mathbb{C}^2 . Our result has a small δ but does not require this additional assumption. It looks plausible that $\delta = 1/6$ holds without any additional assumption.

We conjecture that one can set $\delta = 1/6$ in $\mathbb{Z}/Q\mathbb{Z}$ given that N is sufficiently small compared to p . (This implies $\delta = 1/6$ for the complex case.)

Proof of Theorem 3.2.3. Without loss of generality, assume that $|\mathcal{P}| = |\mathcal{L}| = N$, adding “dummy” points and lines if necessary. By rotating the entire collection of points and lines, we may assume that there are no vertical lines. Say that $\mathcal{P} = \{(x_i, y_i) : i = 1, \dots, N\}$, and, uniquely parameterizing a line $y = mx + b$ by the ordered pair (m, b) , say that $\mathcal{L} = \{(m_i, b_i) : i = 1, \dots, N\}$. Let $S := \bigcup_{i=1}^N \{x_i, y_i, m_i, b_i\}$, set

$$\begin{aligned} L_0 := & \{x_i - x_j : 1 \leq i < j \leq N\} \cup \{y_i - y_j : 1 \leq i < j \leq N\} \cup \\ & \cup \{m_i - m_j : 1 \leq i < j \leq N\} \cup \{b_i - b_j : 1 \leq i < j \leq N\}, \end{aligned}$$

and let $L := L_0 \setminus \{0\}$. By Theorem 3.1.4, there exists a prime $q > N$ and a ring homomorphism $\phi_q : \mathbb{Z}[S] \rightarrow \mathbb{Z}/q$ such that $0 \notin \phi_q(L)$. Define a map $\Phi_q : \mathbb{Z}[S] \times \mathbb{Z}[S] \rightarrow \mathbb{Z}/q \times \mathbb{Z}/q$ by $\Phi_q(a, b) = (\phi_q(a), \phi_q(b))$. Because $0 \notin \phi_q(L)$, we know that $|\Phi_q(\mathcal{P})| = |\Phi_q(\mathcal{L})| = N$. Thus, by Theorem 3.2.1, there exist absolute constants c and δ such that

$$\left| \{(p', \ell') \in \Phi_q(\mathcal{P}) \times \Phi_q(\mathcal{L}) : p' \in \ell'\} \right| \leq cN^{3/2-\delta}.$$

Since ϕ_q is a ring homomorphism, the equation $y = mx + b$ implies that $\phi_q(y) = \phi_q(mx + b) = \phi_q(m)\phi_q(x) + \phi_q(b)$; and thus,

$$|\{(p, \ell) \in \mathcal{P} \times \mathcal{L} : p \in \ell\}| \leq |\{(p', \ell') \in \Phi_q(\mathcal{P}) \times \Phi_q(\mathcal{L}) : p' \in \ell'\}| \leq cN^{3/2-\delta},$$

completing the proof. \square

3.3 A sum-product result for characteristic zero integral domains

Given a subset A of a ring, we define $A + A := \{a_1 + a_2 : a_1, a_2 \in A\}$ and $AA := \{a_1 a_2 : a_1, a_2 \in A\}$. Heuristically, sum-product estimates state that one cannot find a subset A such that both $A + A$ and AA have small cardinality. The first sum-product result was proven in 1983 by Erdős and Szemerédi [14] for the integers, and there have been numerous improvements and generalizations, see for example [28], [16], [13], and [8]. Proving sum-product estimates in $\mathbb{Z}/Q\mathbb{Z}$, where p is a prime, has been the focus of some recent work (see, for example, [5], [4], and [6]), with the best known bound due to Katz and Shen [23], slightly improving a result of Garaev [18]:

Theorem 3.3.1 ([23]). *Let p be a prime and let A be a subset of $\mathbb{Z}/Q\mathbb{Z}$ such that $|A| \leq p^{1/2}$. Then, there exists an absolute constant c such that*

$$c|A|^{14/13-\epsilon} \leq \max\{|A + A|, |AA|\},$$

where ϵ is any positive constant.

Theorem 3.3.2 demonstrates the same lower bound on $\max\{|A + A|, |AA|\}$ for any finite subset A of a characteristic zero integral domain.

Theorem 3.3.2. *There is a positive absolute constant c such that, for every finite subset A of a characteristic zero integral domain,*

$$c|A|^{14/13-\epsilon} \leq \max\{|A + A|, |AA|\},$$

where ϵ is any positive constant.

The constant c in this result is the same as that in Theorem 3.3.1.

Theorem 3.3.2 applies to a very general class of rings; however, our mapping approach requires that the rings be commutative and have characteristic zero. For some results in the non-commutative case, see [8]; and for some results in \mathbb{Z}/m where m is a composite, see [9].

Proof of Theorem 3.3.2. Let

$$L_0 := \{a_1 - a_2 : a_1, a_2 \in A\} \cup \{a_1 + a_2 - (a_3 + a_4) : a_i \in A\} \cup \{a_1 a_2 - a_3 a_4 : a_i \in A\}$$

and let $L := L_0 \setminus \{0\}$.

By Theorem 3.1.4, there exists a prime $p > |A|^2$ and a ring homomorphism $\phi_Q : \mathbb{Z}[A] \rightarrow \mathbb{Z}/Q\mathbb{Z}$ such that

- (i) $|\phi_Q(A)| = |A|$, which follows since $a_1 - a_2 \in L$ for every $a_1, a_2 \in A$ where $a_1 \neq a_2$,
- (ii) $|\phi_Q(A) + \phi_Q(A)| = |A + A|$, which follows by definition of L (if $a_1 + a_2 - (a_3 + a_4) \neq 0$, then $\phi_Q(a_1) + \phi_Q(a_2) = \phi_Q(a_1 + a_2) \neq \phi_Q(a_3 + a_4) = \phi_Q(a_3) + \phi_Q(a_4)$) and because ϕ_Q is a ring homomorphism (if $a_1 + a_2 - (a_3 + a_4) = 0$, then $\phi_Q(a_1) + \phi_Q(a_2) = \phi_Q(a_1 + a_2) = \phi_Q(a_3 + a_4) = \phi_Q(a_3) + \phi_Q(a_4)$),
- (iii) $|\phi_Q(A)\phi_Q(A)| = |AA|$, which follows by the same reasoning as (ii).

We can now apply Theorem 3.3.1 to get that there exists a positive constant c such that

$$c |\phi_Q(A)|^{14/13-\epsilon} \leq \max\{|\phi_Q(A) + \phi_Q(A)|, |\phi_Q(A)\phi_Q(A)|\},$$

for any absolute constant $\epsilon > 0$. Finally, substituting (i), (ii), and (iii) into this inequality gives the desired result. \square

3.4 A matrix product result for $SL_2(D)$

In this section, we will consider finite subsets of the special linear group $SL_2(D)$ of 2 by 2 matrices with determinant 1 and entries in a characteristic zero integral domain D . For A a finite subset of $SL_2(D)$, let $\langle A \rangle$ denote the smallest subgroup of $SL_2(D)$

(under inclusion) that contains A . We will refer to $\langle A \rangle$ as the *group generated by A* . In general, the goal of this section will be to give conditions on $\langle A \rangle$ so that cardinality of the triple product $AAA := \{a_1 a_2 a_3 : a_i \in A\}$ is large.

Helfgott proved the following theorem in [19]:

Theorem 3.4.1 ([19]). *Let p be a prime. Let A be a subset of $\mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})$ not contained in any proper subgroup, and assume that $|A| < p^{3-\epsilon}$ for some fixed $\epsilon > 0$. Then*

$$|AAA| > c |A|^{1+\delta},$$

where $c > 0$ and $\delta > 0$ depend only on ϵ .

In this section, we will prove the following related result by combining Theorem 3.4.1 with Theorem 3.1.4:

Theorem 3.4.2. *Let A be a finite subset of $\mathrm{SL}_2(D)$, where D is a characteristic zero integral domain, and let $\langle A \rangle$ be the subgroup generated by A . If $\langle A \rangle$ has infinite cardinality and $\langle A \rangle$ is not metabelian, then*

$$|AAA| > c |A|^{1+\delta},$$

where $c > 0$ and $\delta > 0$ are absolute constants.

One should note that Chang [10] has already proven a very similar product result for $\mathrm{SL}_2(\mathbb{C})$, in which “metabelian” is replaced by “virtually abelian”.

Theorem 3.4.3 ([10]). *Let A be a finite subset of $\mathrm{SL}_2(\mathbb{C})$, and let $\langle A \rangle$ be the subgroup generated by A . If $\langle A \rangle$ is not virtually abelian (which implies that $\langle A \rangle$ has infinite cardinality), then*

$$|AAA| > c |A|^{1+\delta},$$

where $c > 0$ and $\delta > 0$ are absolute constants.

One major difference between Theorem 3.4.2 and Theorem 3.4.3 is in how the two results are proved. Below, we will prove Theorem 3.4.2 using Helfgott’s Theorem 3.4.1 as a black box along with some group theory and an easy application of Theorem 3.1.4.

On the other hand, Theorem 3.4.3 is proven in [10] by adapting Helfgott's methods in [19] from the case of $\mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})$ to $\mathrm{SL}(\mathbb{C})$ and using tools from additive combinatorics.

The constants $\delta > 0$ in Theorems 3.4.2 and 3.4.3 are not the best possible if one restricts to a subgroup. For example, $\mathrm{SL}_2(\mathbb{Z})$ contains a subgroup isomorphic to F_2 , the free group on 2 generators, and the following product result has recently been shown by Razborov [29]:

Theorem 3.4.4 ([29]). *Let A be a finite subset of a free group F_m (on m generators) with at least two non-commuting elements. Then,*

$$|AAA| \geq \frac{|A|^2}{(\log |A|)^{O(1)}}.$$

One should note that neither Theorem 3.4.2 nor Theorem 3.4.3 fully characterizes finite subsets of $\mathrm{SL}_2(\mathbb{C})$ that have expanding triple product. For example, neither theorem applies when A is contained in an abelian subgroup, but letting

$$A := \left\{ \begin{pmatrix} 1 & 2^j \\ 0 & 1 \end{pmatrix} : 1 \leq j \leq n \right\},$$

we have that $|AAA| \geq |AA| = \binom{n+1}{2} > n^2/2 = |A|^2/2$. One should also note that a sum-product theorem similar to Theorem 3.3.2 does not hold in general for matrices. As pointed out in [11, Remark 0.2], the subset

$$A := \left\{ \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix} : 1 \leq j \leq n \right\}$$

has the property that both the sumset and product sets are small: $|A + A| = |AA| = 2n - 1$. However, it is also shown by Chang [11] that by adding the assumption that the matrices in A are symmetric, one can prove a sum-product result similar to Theorem 3.3.2.

We now turn our attention to the proof of Theorem 3.4.2.

Proof of Theorem 3.4.2. Say that A is a finite subset of $\mathrm{SL}_2(D)$, where D is a characteristic zero integral domain. Let $G := \langle A \rangle$, the subgroup generated by A , and assume that G has infinite cardinality and is not metabelian. Let T be the set of all normal

subgroups N of G such that G/N is abelian (note that we include G in the set T), and define

$$N_0 := \bigcap_{N \in T} N.$$

Then N_0 is a normal subgroup of G and G/N_0 is abelian. Since G is not metabelian by assumption, we know that N_0 is not abelian, and so there exists $B_1, B_2 \in N_0$ such that $B_1 B_2 \neq B_2 B_1$. Also, let $M_1, M_2, M_3, \dots, M_{121}$ be 121 distinct elements of G (note G is infinite by assumption). We may now define a set L_0 as follows:

$$L_0 := \left\{ \begin{array}{l} i, j \in \{1, 2, 3, 4\} \text{ and } b_i \text{ and } c_j \text{ are entries in ma-} \\ b_i - c_j : \text{trices } \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}, \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} \in A \end{array} \right\}$$

$$\cup \left\{ \begin{array}{l} i, j \in \{1, 2, 3, 4\} \text{ and } b_i \text{ and } c_j \text{ are entries in matrices} \\ b_i - c_j : \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \in M_{k_1} \text{ and } \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} \in M_{k_2} \text{ for some } 1 \leq k_1, k_2 \leq 121 \end{array} \right\}$$

$$\cup \left\{ b_1 - 1, b_2, b_3, b_4 - 1 : \text{where } \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = B_1 B_2 B_1^{-1} B_2^{-1} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Let $L := L_0 \setminus \{0\}$, and let S be the set of all entries that appear in matrices in A . By Theorem 3.1.4, there exists $p > |A|$ and $\phi_Q : \mathbb{Z}[S] \rightarrow \mathbb{Z}/Q\mathbb{Z}$ such that $0 \notin \phi_Q(L)$. Let $\Phi_Q : \text{SL}_2(D) \rightarrow \text{SL}_2(\mathbb{Z}/Q\mathbb{Z})$ be defined by $\begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \mapsto \begin{pmatrix} \phi_Q(b_1) & \phi_Q(b_2) \\ \phi_Q(b_3) & \phi_Q(b_4) \end{pmatrix}$. Let $\underline{A} := \Phi_Q(A)$ and let $\underline{G} := \langle \underline{A} \rangle$. Note that by construction $|A| = |\underline{A}|$ and $|AAA| \geq |\underline{A}\underline{A}\underline{A}|$, and also note that $|\underline{G}| \geq 121$.

Assume for the sake of a contradiction that \underline{G} is a proper subgroup of $\text{SL}_2(\mathbb{Z}/Q\mathbb{Z})$. In [36], Suzuki gives the following classification of the proper subgroups of $\text{SL}_2(\mathbb{Z}/Q\mathbb{Z})$:

Theorem 3.4.5 (cf. Theorem 6.17 of [36], page 404). *Let \underline{G} be a proper subgroup of $\text{SL}_2(\mathbb{Z}/Q\mathbb{Z})$ where $p \geq 5$. Then \underline{G} is isomorphic to one of the following groups (or to a subgroup of one of the following groups):*

- (i) a cyclic group,

- (ii) the group with presentation $\langle x, y \mid x^m = y^2, y^{-1}xy = x^{-1} \rangle$, which has order $4m$,
- (iii) a group H of order $p(p-1)$ having a Sylow- p subgroup K such that H/K is cyclic and K is elementary abelian,
- (iv) the special linear group $\mathrm{SL}_2(\mathbb{Z}/3\mathbb{Z})$ on a field of three elements, which has order 24,
- (v) \widehat{S}_4 , the representation group of S_4 (the symmetric group on 4 letters), which has order 48, or
- (vi) the special linear group $\mathrm{SL}_2(\mathbb{Z}/5\mathbb{Z})$ on a field of five elements, which has order 120.

Since $|\underline{G}| > 120$, we may eliminate (iv), (v), and (vi) as possibilities. The remaining possibilities (namely, (i), (ii), and (iii)) are all metabelian; and thus, \underline{G} must have a normal subgroup \underline{N} such that \underline{N} is abelian and $\underline{G}/\underline{N}$ is also abelian.

Let $N := \Phi_Q^{-1}(\underline{N})$. Then N is a normal subgroup of G , and by the third isomorphism theorem $G/N \simeq (G/\ker(\Phi_Q))/(N/\ker(\Phi_Q)) \simeq \underline{G}/\underline{N}$, which is abelian. Thus, N_0 is a subgroup of N , and so $B_1, B_2 \in N$. We know that $B_1 B_2 B_1^{-1} B_2^{-1} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and by the definition of Φ_Q , we also have that

$$\Phi_Q(B_1)\Phi_Q(B_2)\Phi_Q(B_1)^{-1}\Phi_Q(B_2)^{-1} = \Phi_Q(B_1 B_2 B_1^{-1} B_2^{-1}) \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

But, this contradicts the fact that \underline{N} is abelian. Thus, the assumption that \underline{G} is a proper subgroup of $\mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})$ is false, and we have that $\langle \underline{A} \rangle = \underline{G} = \mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})$.

Finally, by Theorem 3.4.1, there exist absolute constants $c > 0$ and $\delta > 0$ such that

$$|AAA| \geq |\underline{A}\underline{A}\underline{A}| \geq c|\underline{A}|^{1+\delta} = c|A|^{1+\delta}.$$

□

Another way to show that $\langle \Phi_Q(A) \rangle$ generates all of $\mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})$ would be to assume that $\langle A \rangle$ is not virtually solvable, which implies by Tits Alternative Theorem [43] that

$\langle A \rangle$ has a non-abelian free subgroup. Then, following [17, Section 2], it is possible to bound the girth of a certain Cayley graph from below in terms of p , eventually showing (via an appeal to Theorem 3.4.5) that $\langle \Phi_Q(A) \rangle = \mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})$.

Also, the proof above uses the following implicit corollary of Theorem 3.4.5: if \underline{G} is a proper subgroup of $\mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})$ and $|\underline{G}| > 120$, then \underline{G} is metabelian. A very similar result for $\mathrm{PSL}_2(\mathbb{Z}/Q\mathbb{Z}) \simeq \mathrm{SL}_2(\mathbb{Z}/Q\mathbb{Z})/(\pm I)$ (where I is the identity matrix) appears in [12, Theorem 3.3.4, page 78].

3.5 Random matrices with entries in a characteristic zero integral domain

In [22, 39], it is shown that a random Bernoulli matrix of size n is singular with probability $\exp(-\Omega(n))$. One may ask what happens for random matrices with complex entries. We are going to give a quick proof for the following:

Theorem 3.5.1. *For every positive number $\rho < 1$, there is a positive number $\delta < 1$ such that the following holds. Let ξ be a random variable with finite support in a characteristic zero integral domain, where ξ takes each value with probability at most ρ . Let M_n be an n by n random matrix whose entries are i.i.d. copies of ξ . Then the probability that M_n is singular is at most δ^n .*

Remark 3.5.2. Here we assume that n is sufficiently large and the size of the support of ξ does not depend on n . In the case when the characteristic zero integral domain is \mathbb{C} , more quantitative bounds are available (see [7, 41]).

Theorem 3.5.1 follows directly from the following two results.

Theorem 3.5.3. *For every positive number $\rho < 1$, there is a positive number $\delta < 1$ such that the following holds. Let n be a large positive integer and $p \geq 2^{n^n}$ be a prime. Let ξ be a random variables with finite support in $\mathbb{Z}/Q\mathbb{Z}$, where ξ takes each value with probability at most ρ . Let M_n be an n by n random matrix whose entries are i.i.d. copies of ξ . Then the probability that M_n is singular is at most δ^n .*

This theorem was implicitly proved in [39]. The bound 2^{n^n} is not essential, we simply want to guarantee that p is much larger than n . The reason that the proof from

[39] does not extend directly to the complex case (or characteristic zero integral domain in general) is that in [39] one relied on the identity

$$\mathbf{I}_{x=0} = \int_0^1 e(xt)dt,$$

where \mathbf{I} is the indicator event and $e(\alpha) := \exp(2\pi i\alpha)$. This identity holds for x an integer, but not true for complex numbers in general. Theorem 3.1.4 provides a simple way to overcome this obstacle. (For other methods, see [41, 42].) We restate Lemma 3.1.1 for the reader's convenience.

Lemma 3.5.4. *Let S be a finite subset of a characteristic zero integral domain. There exist arbitrarily large primes p such that there is a ring homomorphism $\phi_Q : \mathbb{Z}[S] \rightarrow \mathbb{Z}/Q\mathbb{Z}$ satisfying the following two properties:*

- (i) *the map ϕ_Q is injective on S , and*
- (ii) *for any n by n matrix (s_{ij}) with entries $s_{ij} \in S$, we have*

$$\det(s_{ij}) = 0 \quad \text{if and only if} \quad \det(\phi_Q(s_{ij})) = 0.$$

Proof. Let $L := \{\det(s_{ij}) : s_{ij} \in S\} \setminus \{0\}$. Applying Theorem 3.1.4 gives us a ring homomorphism $\phi_Q : \mathbb{Z}[S] \rightarrow \mathbb{Z}/Q\mathbb{Z}$ (for some arbitrarily large prime p) such that $0 \notin \phi_Q(L)$. Since ϕ_Q is a ring homomorphism, $\phi_Q(\det(s_{ij})) = \det(\phi_Q(s_{ij}))$ and also $\phi_Q(0) = 0$; thus, we have satisfied condition (ii).

In this particular case, we will show that (i) follows from (ii). If S contains more than one element, we can find $s \neq t \neq 0$, and thus

$$\det \left(\begin{pmatrix} s & t & \cdots & t & t \\ t & s & t & \cdots & t \\ \vdots & t & \ddots & t & \vdots \\ t & \cdots & t & s & t \\ t & t & \cdots & t & t \end{pmatrix} \right) = \det \left(\begin{pmatrix} s-t & 0 & \cdots & 0 & 0 \\ 0 & s-t & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & s-t & 0 \\ 0 & 0 & \cdots & 0 & t \end{pmatrix} \right) = (s-t)^{n-1}t \neq 0.$$

Thus, by (ii), $0 \neq (\phi_Q(s) - \phi_Q(t))^{n-1} \phi_Q(t)$, and so $\phi_Q(s) \neq \phi_Q(t)$ and we see that ϕ_Q is injective on S . □

The fact that (ii) happens to imply injectivity on S is not important—in fact, for any given finite subset $A \subset \mathbb{Z}[S]$ we can find $\phi_{\tilde{Q}}$ satisfying (ii) above that is also injective on A by adding $\{a_1 - a_2 : a_1 \neq a_2 \text{ and } a_1, a_2 \in A\}$ to L in the proof above. For example, we could find $\phi_{\tilde{Q}}$ that is injective on the set of all determinants of n by n matrices with entries in S .

One should note that it is easy to prove results similar to Lemma 3.1.1 where the determinant is replaced by some polynomial $f(x_1, x_2, \dots, x_k)$ with integer coefficients and one wants a map ϕ_Q such that f evaluated at points in S is zero if and only if f evaluated at points in $\phi_Q(S)$ is zero. This can also easily be extended to the case where f is replaced by a list of polynomials, each of which is evaluated on some subset of S .

3.6 The density theorem

The number 7 is a prime in the ring of integers \mathbb{Z} ; however, if one extends \mathbb{Z} to $\mathbb{Z}[\sqrt{2}]$, the prime 7 splits: $7 = (3 - \sqrt{2})(3 + \sqrt{2})$. This fact has the same mathematical content as the following: the polynomial $x^2 - 2$ is irreducible in $\mathbb{Z}[x]$; however, in $(\mathbb{Z}/7\mathbb{Z})[x]$, where the coefficients of the polynomial are viewed as elements of $\mathbb{Z}/7\mathbb{Z}$, the polynomial splits: $x^2 - 2 = (x - 3)(x + 3)$. The Frobenius Density Theorem describes how frequently such splitting occurs. In modern formulations, the Frobenius Density Theorem quantifies the proportion of primes that split in a given Galois extension of the rational numbers. We will use the following historical version given in [35, page 32], which is phrased in terms of polynomials splitting modulo p . Note that the relative density of a set of primes S is defined to be

$$\lim_{x \rightarrow \infty} \frac{|\{p \leq x : p \in S\}|}{|\{p \leq x : p \text{ is prime}\}|}.$$

Theorem 3.6.1 (Frobenius Density Theorem). *Let $g(z) \in \mathbb{Z}[z]$ be a polynomial of degree k with k distinct roots in \mathbb{C} , and let G be the Galois group of the polynomial g , viewed as a subgroup of S_k (the symmetric group on k symbols). Let n_1, n_2, \dots, n_t be positive integers summing to k . Then, the relative density of the set of primes p for which g modulo p has a given decomposition type n_1, n_2, \dots, n_t exists and is equal to $1/|G|$ times the number of $\sigma \in G$ with cycle pattern n_1, n_2, \dots, n_t .*

For example, since the identity element corresponds to the cycle pattern $1, 1, \dots, 1$ and every group has one identity, the relative density of primes p such that g decomposes into k distinct linear factors modulo p is $1/|G|$.

Theorem 3.6.1 is the version proven by Frobenius in 1880 and published in 1896 [15]. In [35], Stevenhagen and Lenstra give numerous examples and an illuminating discussion of the original motivation for the Frobenius Density Theorem and how it relates to the stronger Chebotarev Density Theorem.

3.7 Proof of Theorem 3.1.4

The first step towards proving Theorem 3.1.4 is proving the following lemma.

Lemma 3.7.1. *Let S be a finite subset of a characteristic zero integral domain D , and let L be a finite set of non-zero elements in the subring $\mathbb{Z}[S]$ of D . Then there exists a complex number θ that is algebraic over \mathbb{Z} and a ring homomorphism $\phi : \mathbb{Z}[S] \rightarrow \mathbb{Z}[\theta]$ such that $0 \notin \phi(L)$.*

By itself, this lemma allows one to extend sum-product and incidence problem results proven in the complex numbers to any characteristic zero integral domain (in much the same way that Theorem 3.1.4 allows one to extend such results proven in $\mathbb{Z}/Q\mathbb{Z}$ to any characteristic zero integral domain).

Lemma 3.7.1 is proved using three main steps: applying the primitive element theorem, applying Hilbert's Nullstellensatz to pass to the case of only algebraic numbers, and applying the primitive element theorem again to get to a ring of the form $\mathbb{Z}[\theta]$. Each of these three steps requires negotiating between the rings we are interested in and their fraction fields. Theorem 3.1.4 is proved by combining Lemma 3.7.1 with the Frobenius Density Theorem (or the stronger Chebotarev Density Theorem) to pass to a quotient isomorphic to $\mathbb{Z}/Q\mathbb{Z}$.

Proof of Lemma 3.7.1. Let S be a finite subset of a characteristic zero integral domain D . Recall that we identify the subring of D generated by the identity with \mathbb{Z} and so we use $\mathbb{Z}[S]$ to denote the smallest subring of D containing S .

We can write $S = \{x_1, x_2, \dots, x_j, \theta_1, \theta_2, \dots, \theta_k\}$, such that $\{x_1, x_2, \dots, x_j\}$ are independent transcendentals over \mathbb{Q} and such that K , the fraction field of $\mathbb{Z}[S]$, is algebraic over $\mathbb{Q}(x_1, x_2, \dots, x_j)$. Using the primitive element theorem, we can find $\tilde{\theta}$ in K also algebraic over $\mathbb{Q}(x_1, x_2, \dots, x_j)$ such that

$$\mathbb{Q}(x_1, x_2, \dots, x_j, \theta_1, \theta_2, \dots, \theta_k) \simeq \mathbb{Q}(x_1, x_2, \dots, x_j, \tilde{\theta}).$$

To get the analogous statement for \mathbb{Z} instead of \mathbb{Q} , one can rationalize the denominators of the θ_i for $i \geq 1$ and rationalize the numerator of $\tilde{\theta}$, and then define θ_0 to be $\tilde{\theta}$ divided by the product of the rationalized denominators and the rationalized numerator. Thus, we can find θ_0 in K also algebraic over $\mathbb{Q}(x_1, x_2, \dots, x_j)$ such that

$$\mathbb{Z}[S] \subset \mathbb{Z}[x_1, x_2, \dots, x_j, \theta_0] \simeq \mathbb{Z}[y_1, y_2, \dots, y_{j+1}]/f_0,$$

where the y_i are formal variables and f_0 is an irreducible element in $\mathbb{Z}[y_1, y_2, \dots, y_{j+1}]$ that is non-constant or zero and that gives zero when evaluated at $y_i = x_i$ for $i = 1, \dots, j$ and $y_{j+1} = \theta_0$.

Let $\overline{\mathbb{Q}}$ be the algebraic closure of the rational numbers, let $\mathcal{L}' := \prod_{\ell \in L} \ell$, and let $\mathcal{L} \in \mathbb{Z}[y_1, \dots, y_{j+1}]$ be the lowest degree representative of the image of \mathcal{L}' under the above inclusion and isomorphism. We will use the following corollary to Hilbert's Nullstellensatz:

Proposition 3.7.2 (c.f. the corollary on page 282 of [31]). *If $\mathcal{L}, f_0 \in \overline{\mathbb{Q}}[y_1, \dots, y_{j+1}]$ and if on points of $\overline{\mathbb{Q}}^{j+1}$ we have that \mathcal{L} is zero whenever f_0 is zero, then there exists $m \geq 0$ and $k \in \overline{\mathbb{Q}}[y_1, \dots, y_{j+1}]$ such that $\mathcal{L}^m = kf_0$.*

Say that $\mathcal{L}^m = kf_0$ for some $k \in \overline{\mathbb{Q}}[y_1, \dots, y_{j+1}]$. Since $\mathcal{L}^m, f_0 \in \mathbb{Z}[y_1, \dots, y_{j+1}]$, we have that k is in $\mathbb{Q}(y_1, \dots, y_{j+1})$ (the fraction field of $\mathbb{Z}[y_1, \dots, y_{j+1}]$). Thus, k is in the ring $\mathbb{Q}[y_1, \dots, y_{j+1}]$, and so there is a positive integer c such that $ck \in \mathbb{Z}[y_1, \dots, y_{j+1}]$. We now have $c\mathcal{L}^m = (ck)f_0$. Since f_0 is irreducible in $\mathbb{Z}[y_1, y_2, \dots, y_{j+1}]$, we must have that f_0 divides \mathcal{L} (f_0 cannot divide the positive integer c since f_0 is either non-constant or zero). But this is impossible since by assumption \mathcal{L} is non-zero in the quotient ring $\mathbb{Z}[y_1, \dots, y_{j+1}]/f_0$. Thus, for every $m \geq 0$ and for every $k \in \overline{\mathbb{Q}}[y_1, \dots, y_{j+1}]$ we must have that $\mathcal{L}^m \neq kf_0$. Therefore, by the contrapositive of Proposition 3.7.2, there exist

algebraic numbers $q_1, \dots, q_{j+1} \in \overline{\mathbb{Q}}$ such that $f_0|_{y_i=q_i} = 0$ while $\mathcal{L}|_{y_i=q_i} \neq 0$. Thus, we have a homomorphism

$$\psi_0 : \mathbb{Z}[y_1, y_2, \dots, y_{j+1}]/f_0 \rightarrow \mathbb{Z}[q_1, \dots, q_{j+1}],$$

defined by $y_i \mapsto q_i$, such that $\psi_0(\mathcal{L}) \neq 0$.

Applying the primitive element theorem again, we have

$$\mathbb{Z}[q_1, \dots, q_{j+1}] \subset \mathbb{Z}[\theta_1] \simeq \mathbb{Z}[z]/f_1,$$

where z is a formal variable and f_1 is an irreducible element in $\mathbb{Z}[z]$ that gives zero when evaluated at $z = \theta_1$. Combining the inclusions and isomorphisms from the applications of the primitive element theorem with ψ_0 completes the proof of Lemma 3.7.1. \square

Recall the statement of Theorem 3.1.4:

Theorem 3.7.3. *Let S be a finite subset of a characteristic zero integral domain D , and let L be a finite set of non-zero elements in the subring $\mathbb{Z}[S]$ of D . There exist arbitrarily large primes Q such that there is a ring homomorphism $\phi_Q : \mathbb{Z}[S] \rightarrow \mathbb{Z}/Q\mathbb{Z}$ satisfying $0 \notin \phi_Q(L)$.*

The proof of Theorem 3.1.4 picks up where the proof of Lemma 3.7.1 left off.

Proof of Theorem 3.1.4. By Lemma 3.7.1, there exists a ring homomorphism $\phi : \mathbb{Z}[S] \rightarrow \mathbb{Z}[z]/f_1$, such that $0 \notin \phi(L)$, where z is a formal variable and f_1 is an irreducible element in $\mathbb{Z}[z]$ that gives zero when evaluated at $z = \theta_1$.

Let $\widehat{L} := \prod_{\ell \in L} \ell$, let $\widetilde{L}(z) \in \mathbb{Z}[z]$ denote the lowest-degree representative of $\phi(\widehat{L})$ in $\mathbb{Z}[z]/f_1$, and let $L_1(z)$ denote the product of all distinct irreducible factors of $\widetilde{L}(z)$ in $\mathbb{Z}[z]$. Note that a ring homomorphism will map $\widetilde{L}(z)$ to zero if and only if it maps $L_1(z)$ to zero. By assumption, $\widetilde{L}(z)$ is non-zero, so we must have that $f_1(z)$ does not divide $\widetilde{L}(z)$ in $\mathbb{Z}[z]$; and thus $f_1(z)$ does not divide $L_1(z)$. Therefore, $L_1(z)$ has no roots (in \mathbb{C} , say) in common with $f_1(z)$, since $f_1(z)$ is irreducible.

By Theorem 3.6.1 (the Frobenius Density Theorem) there exists a sequence of primes (Q_1, Q_2, Q_3, \dots) in \mathbb{Z} (with positive relative density) such that for any prime Q in the

sequence, the polynomial $f_1(z)L_1(z)$ factors completely modulo Q into a product of $\deg(f_1(z)L_1(z))$ distinct linear factors.

Let $(z - a)$ be a linear factor of $f_1(z)$ modulo Q , where Q is any prime in the sequence (Q_1, Q_2, Q_3, \dots) . Since, modulo Q , the linear factors of $f_1(z)$ are all distinct from those of $L_1(z)$, we know that $(z - a)$ does not divide $L_1(z)$ modulo Q . Thus, for infinitely many primes Q , we may quotient out by Q and by $(z - a)$ to get a canonical quotient map

$$\psi_1 : \mathbb{Z}[z]/f_1 \longrightarrow \mathbb{Z}[z]/(Q, z - a) \simeq \mathbb{Z}/Q\mathbb{Z}$$

where $\psi_1(L_1(z)) \neq 0$. One can think of ψ_1 as modding out by Q and then sending z to the element a in $\mathbb{Z}/Q\mathbb{Z}$.

Letting $\phi_Q := \psi_1 \circ \phi$ completes the proof. □

Chapter 4

Proof of main result: reduction to exceptional spaces

4.1 Proof of the main theorem (Theorem 2.1.2)

The proof of Theorem 2.1.2 very closely follows the proof of [39, Theorem 1.2]. Our goal is to highlight the changes that need to be made to generalize the proof in [39] so that it proves Theorem 2.1.2. A reader interested in the details of the proof of Theorem 2.1.2 should read this thesis alongside of [39]. Throughout the proof, we will assume that n is sufficiently large, and we will allow constants hidden in the $o(\cdot)$ and $O(\cdot)$ notation to depend on the constants $\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2, p, q, r, c_{\text{MedDim}}, c_{\text{LgDim}}, c_{\text{LO}}$, and c_m . The constants $\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2$ should be considered very small, and, in fact, we will let them tend to zero to prove the full strength of Theorem 2.1.2. The constants $p, q, r, c_{\text{MedDim}}, c_{\text{LgDim}}, c_{\text{LO}}$, and c_m can be thought of as absolute, except possibly for depending on each other.

4.1.1 Definitions and preliminaries

Given an n by n matrix N_n with entries α_{ij} , we assume that the collection of independent random variables $\{\alpha_{ij}\}_{1 \leq i, j \leq n}$ is p -bounded of exponent r for some fixed constants p, q , and r (here, q is the constant from Definition 2.1.1 which is independent of n). We also assume that each α_{ij} takes at most $n^{o(n)}$ distinct values. Using Lemma 3.1.1, we may assume without loss of generality that each α_{ij} takes values in $\mathbb{Z}/Q\mathbb{Z}$ for some very large prime Q . The entirety of the proof will take place over the field $\mathbb{Z}/Q\mathbb{Z}$, and so terminology such as “linearly independent”, “span”, “dimension”, “rank” and so forth will always be with respect to the field $\mathbb{Z}/Q\mathbb{Z}$.

Let $X_i := (\alpha_{i,1}, \dots, \alpha_{i,n})$ denote the i -th row of N_n . We note that N_n has determinant zero if and only if there is a linear dependency among the rows X_i . It has been shown (see [38, Lemma 5.1] and also [22]) that the dominant contribution to the singularity probability comes from the X_i spanning a hyperplane (of dimension $n - 1$). In particular,

$$\Pr(N_n \text{ is singular}) = p^{-o(n)} \sum_{\substack{V \text{ a non-trivial} \\ \text{hyperplane in } (\mathbb{Z}/Q\mathbb{Z})^n}} \Pr(A_V), \quad (4.1)$$

where A_V denotes the event that X_1, \dots, X_n span V , and *non-trivial* means that V contains the origin, V is spanned by vectors in S^n (where S is the set of all possible values that can occur in N_n), and $\Pr(X_i \in V) > 0$ for all i .

As in [39], we will divide the non-trivial hyperplanes into n^2 classes, since it is then sufficient to show that the sum of $\Pr(A_V)$ over all V in a particular class is at most $(p^{1/r} + o(1))^n$.

Definition 4.1.1 (combinatorial dimension). Let $\mathcal{D} := \{\frac{a}{n} : 0 \leq a \leq n^2, a \in \mathbb{Z}\}$. For any $d_{\pm} \in \mathcal{D}$ such that $d_{\pm} \geq \frac{1}{n}$, we define the *combinatorial Grassmannian* $\text{Gr}(d_{\pm})$ to be the set of all non-trivial hyperplanes V in $(\mathbb{Z}/Q\mathbb{Z})^n$ such that

$$p^{n-d_{\pm}+1/n} < \max_{1 \leq i \leq n} \Pr(X_i \in V) \leq p^{n-d_{\pm}}. \quad (4.2)$$

For $d_{\pm} = 0$, we define $\text{Gr}(0)$ to be the set of all non-trivial hyperplanes such that

$$\max_{1 \leq i \leq n} \Pr(X_i \in V) \leq p^n.$$

We will refer to d_{\pm} as the *combinatorial dimension* of V .

Note that $\text{Gr}(d_{\pm}) = \emptyset$ for $d_{\pm} \geq n - 1 + 1/n$ (by Lemma A.2.1). We will consider hyperplanes V with combinatorial dimension in three main regions: d_{\pm} small, d_{\pm} medium-sized, and d_{\pm} large. The two lemmas and the proposition below suffice to prove Theorem 2.1.2.

Lemma 4.1.2 (Small combinatorial dimension, [22], [38], [39]). *For any $\delta > 0$ we have*

$$\sum_{d_{\pm} \in \mathcal{D} \text{ s.t. } p^{n-d_{\pm}} \leq \delta^n} \sum_{V \in \text{Gr}(d_{\pm})} \Pr(A_V) \leq n\delta^n.$$

In proving Theorem 2.1.2, we will take $\delta = (p + c_{\text{MedDim}}\epsilon_0)^{1/r}$ to take care of all small d_{\pm} not covered by Proposition 4.1.4 below.

Proof. The reasoning here is the same as in [39, Lemma 2.3], making use of fact that $\Pr(X_i \in V) \leq \max_{1 \leq i \leq n} \Pr(X_i \in V) \leq p^{n-d_{\pm}} \leq \delta^n$. In particular,

$$\Pr(A_V) \leq \sum_{i=1}^n \Pr(\{X_j\}_{1 \leq j \leq n} \setminus \{X_i\} \text{ spans } V) \Pr(X_i \in V),$$

which completes the proof since the summing the right-hand side over all V is at most $n \max_i \Pr(X_i \in V)$ (note that an instance of the vectors $\{X_j\}_{1 \leq j \leq n} \setminus \{X_i\}$ can span at most one hyperplane). \square

Lemma 4.1.3 (Large combinatorial dimension, [22],[38],[39]). *We have*

$$\sum_{d_{\pm} \in \mathcal{D} \text{ s.t. } \frac{c_{\text{LgDim}}}{\sqrt{n}} \leq p^{n-d_{\pm}}} \sum_{V \in \text{Gr}(d_{\pm})} \Pr(A_V) \leq (p + o(1))^n$$

Here we choose the constant c_{LgDim} so that $c_{\text{LgDim}} \geq c_{\text{LO}} p^{-1/n} \sqrt{\frac{2r}{q}}$, where c_{LO} is the constant from the Littlewood-Offord inequality (see Lemma A.1.1 in Appendix A.1) and q is the constant from Definition 2.1.1.

Proof. Our proof is essentially the same as [39, Lemma 2.4]. Fix $V \in \text{Gr}(d_{\pm})$, where $\frac{c_{\text{LgDim}}}{\sqrt{n}} \leq p^{n-d_{\pm}}$. Let i_{\max} be an index such that $\Pr(X_{i_{\max}} \in V) = \max_{1 \leq i \leq n} \Pr(X_i \in V)$. By assumption,

$$\Pr(X_{i_{\max}} \in V) \geq p^{n-d_{\pm}+1/n} \geq \frac{c_{\text{LgDim}}}{\sqrt{n}} p^{1/n} \geq c_{\text{LO}} \sqrt{\frac{2r}{qn}}.$$

Noting that $X_{i_{\max}} \in V$ if and only if $X_{i_{\max}}$ is orthogonal to the normal vector for V , we have by Lemma A.1.1 that

$$\Pr(X_{i_{\max}} \in V) \leq c_{\text{LO}} \sqrt{\frac{r}{qk}},$$

where k is the number of nonzero coordinates in the normal vector to V . Combining the two inequalities above shows that $k \leq n/2$.

Thus, we have

$$\begin{aligned} \sum_{d_{\pm} \in \mathcal{D} \text{ s.t. } \frac{c_{\text{LgDim}}}{\sqrt{n}} \leq p^{n-d_{\pm}}} \sum_{V \in \text{Gr}(d_{\pm})} \Pr(A_V) &\leq \Pr \left(\left\{ \begin{array}{l} \text{there exists a vector } \mathbf{v} \text{ with at} \\ \text{most } n/2 \text{ nonzero coordinates} \\ \text{such that } N_n \cdot \mathbf{v} = 0 \end{array} \right\} \right) \\ &\leq (p + o(1))^n \quad (\text{by Lemma A.1.2}) \end{aligned}$$

(Lemma A.1.2 is a natural generalization of [22, Section 3.1]; see also [26], [38, Lemma 5.1], and [3, Lemma 14.10].) \square

Proposition 4.1.4 (Medium combinatorial dimension estimate). *Let $0 < \epsilon_0$ be a constant much smaller than 1, and let $d_{\pm} \in \mathcal{D}$ be such that $(p + c_{\text{MedDim}}\epsilon_0)^{n/r} < p^{n-d_{\pm}} < \frac{c_{\text{LgDim}}}{\sqrt{n}}$. Then*

$$\sum_{V \in \text{Gr}(d_{\pm})} \Pr(A_V) \leq o(1)^n.$$

Here we choose the constant c_{MedDim} so that $c_{\text{MedDim}} > (\frac{1}{100} + c_m)$, where c_m is some absolute constant such that $0 < c_m < 1$ (the $\frac{1}{100}$ here comes from $\underline{\mu}$ as defined in Section 4.1.2 below; in [39], it happens that the constant c_m is also taken to be $\frac{1}{100}$).

To prove Theorem 2.1.2, we can simply combine Lemma 4.1.2 with $\delta = (p + c_{\text{MedDim}}\epsilon_0)^{1/r}$, Lemma 4.1.3, and Proposition 4.1.4. Thus, proving Proposition 4.1.4 will complete the proof of Theorem 2.1.2. To prove Proposition 4.1.4, as in [39, Proposition 2.5], we will separate hyperplanes V of medium combinatorial dimension into two classes, which we will call *exceptional* and *unexceptional* (see Definition 4.1.5). See [39, Section 3] for motivation. The unexceptional case will be proved in the remainder of this section, and the exceptional case will be proved in Sections 4.2, 5.1, and 5.2.

The results in [38] and [22] were derived using the ideas that we will use for the unexceptional medium combinatorial dimension case. The idea of considering the exceptional case separately in [39] (and using tools from additive combinatorics in the exceptional case) is what lead to the improvement of Inequality (1.1), which gives a bound of asymptotically $(\frac{3}{4})^n$, over the $.999^n$ bound in [22].

4.1.2 Proof of the medium combinatorial dimension

Before defining exceptional and unexceptional hyperplanes, we will need some new notation. By assumption, the collection of random variables $\{\alpha_{ij}\}_{1 \leq i, j \leq n}$ is p -bounded of exponent r with a constant $\mu = 1 - p$, with random variables $\beta_{ij}^{(\mu)}$ corresponding to each α_{ij} , and with a constant $0 < q \leq p$ (see Definition 2.1.1). We also define a constant slightly smaller than μ , namely $\underline{\mu} := \mu - \frac{\epsilon_0}{100}$. We will let $Y_i := (y_{i,1}, \dots, y_{i,n}) := (\beta_{i,1}^{(\underline{\mu})}, \dots, \beta_{i,n}^{(\underline{\mu})})$ denote another row vector that corresponds to the row vector X_i ($\beta_{i,j}^{(\underline{\mu})}$ comes from the definition of p -bounded of exponent r). Also, we will let

$$Z_{i,k}^* := \left(\overbrace{0, \dots, 0}^{\text{zeroes}}, y_{i,k_{\text{start}}}, \dots, y_{i,k_{\text{end}}}, \overbrace{0, \dots, 0}^{\text{zeroes}} \right), \quad (4.3)$$

where $k_{\text{start}} := \lfloor (k-1)\frac{n}{r} \rfloor + 1$ and $k_{\text{end}} := \lfloor k\frac{n}{r} \rfloor$. The vector $Z_{i,k}^*$ can be thought of as the k -th segment of Y_i (out of r roughly equal segments). Note that Y_i and $Z_{i,k}^*$ are both defined using $\underline{\mu} := \mu - \frac{\epsilon_0}{100}$, not μ . Finally, let ϵ_1 be a positive constant that is small with respect to ϵ_0 , c_m , and r .

Definition 4.1.5 (exceptional and unexceptional). Consider a hyperplane V of medium combinatorial dimension (that is, d_{\pm} satisfies the condition in Proposition 4.1.4). We say V is *unexceptional* if there exists an i_0 where $1 \leq i_0 \leq n$ and there exists a k_0 where $1 \leq k_0 \leq r$ such that

$$\max_{1 \leq j \leq n} \{\Pr(X_j \in V)\} < \epsilon_1 \Pr(Z_{i_0, k_0}^* \in V).$$

We say V is *exceptional* if for every i where $1 \leq i \leq n$ and for every k where $1 \leq k \leq r$ we have

$$\epsilon_1 \Pr(Z_{i,k}^* \in V) \leq \max_{1 \leq j \leq n} \{\Pr(X_j \in V)\}. \quad (4.4)$$

In particular, there exists i_{max} such that $\Pr(X_{i_{\text{max}}} \in V) = \max_{1 \leq j \leq n} \{\Pr(X_j \in V)\}$; and so if V is exceptional, then

$$\epsilon_1 \Pr(Z_{i_{\text{max}}, k}^* \in V) \leq \Pr(X_{i_{\text{max}}} \in V) \quad \text{for every } k. \quad (4.5)$$

We will refer to $X_{i_{\text{max}}}$ as the *exceptional row*.

Inequality (2.2) following Definition 2.1.1 can be used to give another relationship between $\Pr(Z_{i_{\max},k}^* \in V)$ and $\Pr(X_{i_{\max}} \in V)$ that, together with Inequality (4.5), will be of critical importance in Section 5.1.

Proposition 4.1.4 follows from the two lemmas below, so long as ϵ_1 is chosen suitably small with respect to ϵ_0 , c_m , and r .

Lemma 4.1.6 (Unexceptional space estimate). *We have*

$$\sum_{V \in \text{Gr}(d_{\pm}): V \text{ is unexceptional}} \Pr(A_V) \leq p^{-o(n)} 2^n \epsilon_1^{c_m \epsilon_0 n / r}.$$

Lemma 4.1.7 (Exceptional space estimate). *We have*

$$\sum_{V \in \text{Gr}(d_{\pm}): V \text{ is exceptional}} \Pr(A_V) \leq n^{-\frac{n}{2} + o(n)}.$$

We will prove Lemma 4.1.6 in Section 4.1.3, and we will prove Lemma 4.1.7 in Section 4.2.

4.1.3 The unexceptional medium combinatorial dimension case

The general idea for the case of an unexceptional hyperplane V is to replace some of the rows X_i in the matrix N_n with rows that concentrate more sharply on the subspace V . In the case where the exponent $r = 1$, replacing a row X_i with $Y_i := (\beta_{i,1}^{(\mu)}, \dots, \beta_{i,n}^{(\mu)})$ is successful; however, in the exponent $r = 2$ case, for example, replacing the entire row results in a bound that is off by an exponential factor. We solve this problem by replacing X_i with only half of Y_i (with the other half of the entries being zero). This idea easily extends to any integer $r \geq 2$ and is the motivation for defining the vectors $Z_{i,k}^*$ to have all zeros except for roughly n/r coordinates, as is done in Equation (4.3). The basic utility of Z_{i_0,k_0}^* (from Definition 4.1.5) is that it concentrates more sharply on the unexceptional subspace V than the vector X_i for any i .

Let Z_{i_0,k_0}^* be the vector from the definition of unexceptional (Definition 4.1.5) such that $\Pr(X_i \in V) < \epsilon_1 \Pr(Z_{i_0,k_0}^* \in V)$ for every i , and set $Z := Z_{i_0,j_0}^*$. Let m be the closest integer to $\frac{c_m \epsilon_0 n}{r}$, where c_m is a small positive absolute constant (for example, in [39], c_m is taken to be $\frac{1}{100}$). Finally, let Z_1, \dots, Z_m be copies of Z , independent of each other and of X_1, \dots, X_n .

Lemma 4.1.8 (see Lemma 4.4 in [39]). *Let $B_{V,m}$ be the event that Z_1, \dots, Z_m are linearly independent and lie in V . Then,*

$$\Pr(B_{V,m}) \geq p^{o(n)} \left(\frac{\max_{1 \leq i \leq n} \Pr(X_i \in V)}{\epsilon_1} \right)^m$$

Proof. The argument follows the same reasoning as [39, Lemma 4.4], however, the quantity $2^{d_{\pm} - n}$ in [39] should be replaced by $\max_{1 \leq i \leq n} \Pr(X_i \in V)$. Details are provided in Appendix A.2. \square

To conclude the proof of Lemma 4.1.6, we follow the “row-swapping” argument at the end of [39, Section 4], with the small change of bounding $\Pr(X_i \in V)$ by $\max_{1 \leq i \leq n} \Pr(X_i \in V)$, which we use in place of the quantity $2^{d_{\pm} - n}$. Details are provided in Appendix A.2.

4.2 Analyzing the exceptional medium combinatorial dimension case

The approach for exceptional V in [39] is very different from that used in the unexceptional case or in the large or small combinatorial dimension cases. Using some powerful tools from additive combinatorics, the general idea is to put exceptional hyperplanes V in correspondence with a particular additive structure called a generalized arithmetic progression, and then to show that the number of the particular generalized arithmetic progressions that arise in this way is exceedingly small. The key to this approach is a structure theorem—namely, [39, Theorem 5.3]. In this section, we state a slightly modified structure theorem (Theorem 4.2.1), and then we show how to use Theorem 4.2.1 to prove Lemma 4.1.7. In the beginning of Section 5.1, we outline the changes needed to prove the the structure theorem for our current context, and in Sections 5.1 and 5.2 we provide details.

Before stating the structure theorem, we need some definitions and notation. A *generalized arithmetic progression* of rank \mathfrak{r} is a set of the form

$$P = \{v_0 + m_1 v_1 + \dots + m_{\mathfrak{r}} : |m_i| \leq M_i/2\},$$

where the *basis vectors* $v_0, v_1, \dots, v_{\mathfrak{r}}$ are elements of a \mathbb{Z} -module (here, $\mathbb{Z}/Q\mathbb{Z}$) and where the *dimensions* $M_1, \dots, M_{\mathfrak{r}}$ are positive integers. We say that v_i has *corresponding*

dimension M_i . For a given element $a = v_0 + m_1v_1 + \cdots + m_{\mathfrak{r}}$ in P , we refer to $m_1, \dots, m_{\mathfrak{r}}$ as *coefficients* for a . A generalized arithmetic progression P is *symmetric* if $v_0 = 0$, and P is *proper* if for each $a \in P$, there is a unique \mathfrak{r} -tuple $(m_1, \dots, m_{\mathfrak{r}})$ with $|m_i| < M_i/2$ that gives the coefficients for a . If P is proper and symmetric, we define the P -norm $\|a\|_P$ of an element $a \in P$ to be

$$\|a\|_P := \left(\sum_{i=1}^{\mathfrak{r}} \left(\frac{m_i}{M_i} \right)^2 \right)^{1/2}.$$

We will use the notation mP , where m is a positive integer, to denote the set $\{\sum_{i=1}^m x_i : x_i \in P\}$ and the notation P^m , where m is a positive integer, to denote the set $\{\prod_{i=1}^m x_i : x_i \in P\}$. If P is a generalized arithmetic progression of rank \mathfrak{r} , then so is mP , while P^m , on the other hand, is a generalized arithmetic progression of rank at most \mathfrak{r}^m . Also note that $|mP| \leq m^{\mathfrak{r}}|P|$ and that $|P^m| \leq |P|^m$.

Let V be an exceptional hyperplane of medium combinatorial dimension in $\text{Gr}(d_{\pm})$ and let $X_{i_{\max}} = (\alpha_1, \dots, \alpha_n)$ be the exceptional row (here we are using α_j as shorthand for $\alpha_{i_{\max},j}$). Let $(\beta_1^{(\mu)}, \dots, \beta_n^{(\mu)})$ be the row of random variables corresponding to $X_{i_{\max}}$ from the definition of p -bounded of exponent r , and let $b_{j,s}$ with $1 \leq j \leq n$ and $1 \leq s \leq \ell_j$ be the values taken by $\beta_j^{(\mu)}$ (see Equation (2.1) for the definition of $\beta_j^{(\mu)}$).

Given an exceptional hyperplane V , there exists a representation of the form

$$V = \{(x_1, x_2, \dots, x_n) \in (\mathbb{Z}/Q\mathbb{Z})^n : x_1a_1 + x_2a_2 + \cdots + x_na_n = 0\}$$

for some elements $a_1, a_2, \dots, a_n \in \mathbb{Z}/Q\mathbb{Z}$. We will call a_1, a_2, \dots, a_n the *defining coordinates* of V . Finally, let $\tilde{a}_j := b_{j,1}a_j$. We will refer to $(\tilde{a}_1, \dots, \tilde{a}_n)$ as the *scaled defining coordinates* of V . Note that once i_{\max} is fixed, so are the elements $b_{j,1}$. We should also note that the choice of $b_{j,1}$ among $b_{j,s}$ for $1 \leq s \leq \ell_j$ is arbitrary—since $\beta_j^{(\mu)}$ takes the values $b_{j,s}$ each with probability at least q , any value of s will do; and so we have taken $s = 1$ for convenience.

Let \mathbb{H} denote the *highly rational numbers*, that is, those numbers in $\mathbb{Z}/Q\mathbb{Z}$ of the form $a/b \pmod{Q}$ where a, b are integers such that $|a|, |b| \leq n^{o(n)}$ and $b \neq 0$. The highly rational numbers were defined in [39, Section 8], and we will need a small extension for the current thesis, due to the fact that we are using the scaled defining coordinates

of V instead of simply the defining coordinates of V . If we were to assume that $b_{j,1}$ was an $O(1)$ integer for all j and that every possible value taken by α_{ij} was an $O(1)$ integer for all i, j , then we could still use the same definition of highly rational as in [39]. However, if there is a $b_{j,1}$ or an entry α_{ij} in the matrix N_n that ever takes an irrational value, then when we pass to $\mathbb{Z}/Q\mathbb{Z}$ using Lemma 3.1.1 we have to account for values possibly on the order of Q (see Remark 3.1.2), and the highly rational numbers are not sufficient for this task. We can overcome this difficulty by extending to the highly T -rational numbers, which contain the highly rational numbers along with all the values in a structured set T (described below). We will now give a rigorous definition the highly T -rational numbers.

Let T be a generalized arithmetic progression in $\mathbb{Z}/Q\mathbb{Z}$ with rank $O(1)$ and having cardinality at most $n^{o(n)}$. As in the definition of p -bounded of exponent r (Definition 2.1.1), we will take S to be the generalized arithmetic progression containing all possible values in $\mathbb{Z}/Q\mathbb{Z}$ taken by the random variables α_{ij} that are the entries of N_n ; thus, by assumption $|S| \leq n^{o(n)}$. By the definition of p -bounded of exponent r , we know that all of the random variables $\beta_{ij}^{(\mu)}$ take values in a set with cardinality $O(1)$. Thus, there is a symmetric generalized arithmetic progression T with rank $O(1)$ and cardinality $|T| \leq n^{o(n)}$ such that T contains S , such that T contains the set $\{-1, 0, 1\}$, and such that T contains all the values taken by the $\beta_{ij}^{(\mu)}$. To construct T from S , one can, for example, add each distinct value taken by a $\beta_{ij}^{(\mu)}$ as a new basis vector v' with corresponding dimension $M' := 3$ (say).

A *highly T -rational number* h is any element of $\mathbb{Z}/Q\mathbb{Z}$ of the form a/b , where $a, b \in n^{o(n)}T^{O(1)}$. Note that therefore, the cardinality of the highly T -rational numbers is at most $(n^{do(n)}|T|)^{O(1)} = n^{o(n)}$, where $d = O(1)$ is the rank of T (here we used the fact that $|T| \leq n^{o(n)}$).

Theorem 4.2.1 (Structure Theorem). *There is a constant $C = C(\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2, q, r, \mu)$ such that the following holds. Let V be an exceptional hyperplane and let $\tilde{a}_1, \dots, \tilde{a}_n$ be its scaled defining coordinates (as described above). Then there exist integers*

$$1 \leq \mathfrak{r} \leq C$$

and $M_1, \dots, M_\tau \geq 1$ with the volume bound

$$M_1 \cdots M_\tau \leq C \Pr(X_{i_{\max}} \in V)^{-1}$$

and nonzero elements $v_1, \dots, v_\tau \in \mathbb{Z}/Q\mathbb{Z}$ such that the following holds

- (i) (Scaled defining coordinates lie in a progression) The symmetric generalized arithmetic progression

$$P := \{m_1 v_1 + \cdots + m_\tau v_\tau : -M_i/2 < m_i < M_i/2\}$$

is proper and contains all of the \tilde{a}_j .

- (ii) (Bounded norm) The \tilde{a}_j have small P -norm:

$$\sum_{j=1}^n \|\tilde{a}_j\|_P^2 \leq C.$$

- (iii) (Rational T -commensurability) The set $\{v_1, \dots, v_\tau\} \cup \{\tilde{a}_1, \dots, \tilde{a}_n\}$ is contained in the set

$$\{h v_1 : h \text{ is highly } T\text{-rational}\}.$$

Note that unlike [39], part (iii) above does not necessarily place $\{v_1, \dots, v_\tau\} \cup \{\tilde{a}_1, \dots, \tilde{a}_n\}$ in a simple arithmetic progression.

We will discuss the proof of the structure theorem in Sections 5.1 and 5.2. In the remainder of this section, we will discuss how to use the structure theorem to prove Lemma 4.1.7.

Fix d_\pm of medium combinatorial dimension (see Proposition 4.1.4). Using independence of the rows, we have

$$\begin{aligned} \sum_{\substack{V \in \text{Gr}(d_\pm): \\ V \text{ is exceptional}}} \Pr(A_V) &\leq \sum_{\substack{V \in \text{Gr}(d_\pm): \\ V \text{ is exceptional}}} \prod_{i=1}^n \Pr(X_i \in V) \\ &\leq |\{V \in \text{Gr}(d_\pm) : V \text{ is exceptional}\}| \cdot \left(\max_{1 \leq i \leq n} \Pr(X_i \in V) \right)^n. \end{aligned} \tag{4.6}$$

In [39, Section 5], it is shown using Theorem 4.2.1(i) and (ii) and Gaussian-type methods (and the fact that \mathfrak{r} is bounded by a constant) that

$$|\{V \in \text{Gr}(d_{\pm}) : V \text{ is exceptional}\}| \leq \frac{n^{o(n)}}{Q-1} \sum_{\substack{\mathfrak{r}, \{M_1, \dots, M_{\mathfrak{r}}\} \\ \{v_1, \dots, v_{\mathfrak{r}}\}}} \left(1 + n^{-1/2} M_1 \cdots M_{\mathfrak{r}}\right)^n,$$

where the sum runs over all possible values for \mathfrak{r} , for the M_i , and for $v_1, \dots, v_{\mathfrak{r}}$. By Theorem 4.2.1, we know that $\mathfrak{r} \leq C = O(1)$ and that $M_i \leq M_1 M_2 \cdots M_{\mathfrak{r}} \leq C \Pr(X_{i_{\max}} \in V)^{-1} \leq O(1/p^n)$; thus, there are at most $n^{o(n)}$ choices for \mathfrak{r} and the M_i . Furthermore, there are at most $Q-1$ choices for v_1 (since $v_1 \neq 0$), and once the value for v_1 has been fixed, (iii) tells us that there are at most $n^{o(n)}$ choices for $\{v_2, \dots, v_{\mathfrak{r}}\}$ (since $|n^{o(n)} T^{O(1)}| \leq n^{o(n)}$). Thus, the sum runs over at most $n^{o(n)}$ terms. (This is the point in the proof where it is essential that $n^{o(n)} T^{O(1)}$ has cardinality $n^{o(n)}$.)

Plugging the volume bound on $M_1 \cdots M_{\mathfrak{r}}$ into the previous displayed inequality, we have

$$\begin{aligned} |\{V \in \text{Gr}(d_{\pm}) : V \text{ is exceptional}\}| &\leq n^{o(n)} \left(1 + n^{-\frac{1}{2}} C \Pr(X_{i_{\max}} \in V)^{-1}\right)^n \\ &= n^{-\frac{n}{2} + o(n)} \Pr(X_{i_{\max}} \in V)^{-n}, \end{aligned} \quad (4.7)$$

using the fact that $\Pr(X_{i_{\max}} \in V) \leq \frac{c_{\text{LGDim}}}{\sqrt{n}}$, which is a consequence of d_{\pm} being of medium combinatorial dimension. Plugging in Inequality (4.7) into Inequality (4.6) and summing over all d_{\pm} of medium combinatorial dimension completes the proof of Lemma 4.1.7 (recall that by assumption $\max_{1 \leq i \leq n} \Pr(X_i \in V) = \Pr(X_{i_{\max}} \in V)$).

Chapter 5

The structure of exceptional spaces

5.1 Halász-type arguments

The proof of the structure theorem has two main ingredients: tools from additive combinatorics, and Halász-type arguments using discrete Fourier analysis. Our proof of Theorem 4.2.1 will follow the proof of [39, Theorem 5.2] very closely. We will use results about additive combinatorics from [39, Section 6] directly, and we will discuss below the extent to which the Halász-type arguments of [39, Section 7] need to be modified to work for our current context. The proof of Theorem 4.2.1 will be given in Section 5.2 using results from the current section, [39, Section 6], [39, Section 7], and [39, Section 8]. Our Section 5.2 follows [39, Section 8] closely, with a few modifications to prove rational T -commensurability instead of only rational commensurability.

In this section we discuss modifications to the lemmas in [39, Section 7] that are needed in order to prove Theorem 4.2.1.

We will use $e_Q(\cdot)$ to denote the primitive character

$$e_Q(x) := \exp(2\pi ix/Q).$$

Let i_{\max} be the index of the exceptional row, so for every $1 \leq k \leq r$ we have

$$\epsilon_1 \Pr(Z_{i_{\max},k}^* \in V) \leq \Pr(X_{i_{\max}} \in V), \quad (5.1)$$

and recall that by Definition 4.1.5 we have $\Pr(X_{i_{\max}} \in V) = \max_i \Pr(X_i \in V)$. Let $(\alpha_1, \dots, \alpha_n) := X_{i_{\max}}$ with the corresponding random variables $(\beta_1^{(\mu)}, \dots, \beta_n^{(\mu)})$ from the definition of p -bounded of exponent r (see Definition 2.1.1 and Equation (2.1)), and let (a_1, \dots, a_n) be the defining coordinates of V . Then, using the Fourier expansion,

we can compute

$$\begin{aligned}
\Pr(X_{i_{\max}} \in V) &= \mathbb{E}(\mathbf{1}_{\{X_{i_{\max}} \in V\}}) = \mathbb{E}\left(\frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} e_Q\left(\sum_{j=1}^n \alpha_j a_j \xi\right)\right) \\
&\leq \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^n |\mathbb{E}(e_Q(\alpha_j a_j \xi))| \\
&\leq \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^n \mathbb{E}\left(e_Q(\beta_j^{(\mu)} a_j \xi)\right)^{1/r} \\
&= \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^n \left(1 - \mu + \mu \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r} \tag{5.2}
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^n \left(1 - \underline{\mu} + \underline{\mu} \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r} \\
&\leq \prod_{k=1}^r \Pr(Z_{i_{\max}, k}^* \in V)^{1/r}, \tag{5.3}
\end{aligned}$$

where the last line is an application of Hölder's inequality.

Define

$$f(\xi) := \prod_{j=1}^n \left(1 - \mu + \mu \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r}, \tag{5.4}$$

$$f_j(\xi) := \left(1 - \mu + \mu \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r}, \quad \text{and} \tag{5.5}$$

$$g_k(\xi) := \prod_{(k-1)\frac{n}{r} < j \leq k\frac{n}{r}} \left(1 - \underline{\mu} + \underline{\mu} \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q)\right)^{1/r}, \tag{5.6}$$

where $\underline{\mu} := \mu - \frac{\epsilon_0}{100}$, as defined in Section 4.1.2. Note that $f(\xi) = \prod_{j=1}^n f_j(\xi)$.

We will need the following analog of [39, Lemma 7.1]:

Lemma 5.1.1. *For all $\xi \in \mathbb{Z}/Q\mathbb{Z}$, we have*

$$\prod_{j=1}^n f_j(\xi)^{r\mu/\mu} \leq \prod_{k=1}^r g_k(\xi)$$

Proof. This inequality may be proven pointwise (for each j after expanding out the definition of g_k) using the convexity of the log function, just as in the proof of [39, Lemma 7.1] (see also [38, Lemma 7.1]). \square

Let ϵ_2 be sufficiently small compared to ϵ_1 (we will specify how small in Inequality (5.12) while proving Lemma 5.1.2). Following [39], we define the *spectrum* $\Lambda \subset \mathbb{Z}/Q\mathbb{Z}$ of $\{b_{1,1}a_1, \dots, b_{n,1}a_n\} = \{\tilde{a}_1, \dots, \tilde{a}_n\}$ (the scaled defining coordinates of V) to be

$$\Lambda := \{\xi \in \mathbb{Z}/Q\mathbb{Z} : f(\xi) \geq \epsilon_2\}. \quad (5.7)$$

Let $\|x\|_{\mathbb{R}/\mathbb{Z}}$ denote the distance from $x \in \mathbb{R}$ to the nearest integer. Using the elementary inequality $\cos(2\pi x) \leq 1 - \frac{1}{100} \|x\|_{\mathbb{R}/\mathbb{Z}}^2$, we have

$$\begin{aligned} f(\xi) &\leq \exp\left(-\frac{\mu}{100r} \sum_{j=1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s}a_j\xi/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) \\ &\leq \exp\left(-\frac{q}{50r} \sum_{j=1}^n \|b_{j,1}a_j\xi/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) \end{aligned} \quad (5.8)$$

($\mu p_{j,1} \geq 2q$ since $\min_x \Pr(\beta_j^{(\mu)} = x) \geq q$ by Definition 2.1.1).

Thus, there is a constant $C(\epsilon_2, q, r)$ such that

$$\left(\sum_{j=1}^n \|\tilde{a}_j\xi/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2} = \left(\sum_{j=1}^n \|b_{j,1}a_j\xi/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2} \leq C(\epsilon_2, q, r), \quad (5.9)$$

for every $\xi \in \Lambda$. (E.g., the constant $C(\epsilon_2, q, r) := \left(\frac{50r}{q} \ln\left(\frac{1}{\epsilon_2}\right)\right)^{1/2}$ suffices.)

Lemma 5.1.2. *There exists a constant C depending on $\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2, q, r$, and μ such that*

$$C^{-1}Q \Pr(X_{i_{\max}} \in V) \leq |\Lambda| \leq CQ \Pr(X_{i_{\max}} \in V). \quad (5.10)$$

Furthermore, for every integer $k \geq 4$ we have

$$|k\Lambda| \leq \binom{C+k-3}{k-2} CQ \Pr(X_{i_{\max}} \in V). \quad (5.11)$$

Proof. Our goal is to bound $\sum_{\xi \in \Lambda} f(\xi)$ from above and below, and then pass to bounds on $|\Lambda|$ using the fact that $\epsilon_2 \leq f(\xi) \leq 1$ for all $\xi \in \Lambda$.

Note that

$$\frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} f(\xi) \geq \Pr(X_{i_{\max}} \in V) \quad (\text{by Equation (5.4) and Equation (5.2)}).$$

Also,

$$\begin{aligned}
\frac{1}{Q} \sum_{\xi \notin \Lambda} f(\xi) &= \frac{1}{Q} \sum_{\xi \notin \Lambda} \prod_{j=1}^n f_j(\xi) \leq \epsilon_2^{1-\underline{\mu}/\mu} \frac{1}{Q} \sum_{\xi \notin \Lambda} \prod_{j=1}^n f_j(\xi)^{\underline{\mu}/\mu} \\
&\leq \epsilon_2^{1-\underline{\mu}/\mu} \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{k=1}^r g_k(\xi)^{1/r} && \text{(Lemma 5.1.1)} \\
&\leq \epsilon_2^{1-\underline{\mu}/\mu} \frac{1}{Q} \left(\prod_{k=1}^r \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} g_k(\xi) \right)^{1/r} && \text{(Hölder's inequality)} \\
&\leq \epsilon_2^{1-\underline{\mu}/\mu} \left(\frac{1}{\epsilon_1} \right) \Pr(X_{i_{\max}} \in V) && \text{(by Inequality (5.1))} .
\end{aligned}$$

For the lower bound, we have

$$\begin{aligned}
\sum_{\xi \in \Lambda} f(\xi) &= \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} f(\xi) - \sum_{\xi \notin \Lambda} f(\xi) \\
&\geq Q \Pr(X_{i_{\max}} \in V) - \frac{\epsilon_2^{1-\underline{\mu}/\mu}}{\epsilon_1} Q \Pr(X_{i_{\max}} \in V) \\
&= Q \Pr(X_{i_{\max}} \in V) \left(1 - \frac{\epsilon_2^{1-\underline{\mu}/\mu}}{\epsilon_1} \right) .
\end{aligned}$$

We can choose ϵ_2 sufficiently small with respect to ϵ_1 and $1 - \underline{\mu}/\mu$ so that, for example,

$$1 - \frac{\epsilon_2^{1-\underline{\mu}/\mu}}{\epsilon_1} \geq \frac{1}{2}. \quad (5.12)$$

For the upper bound, we have

$$\begin{aligned}
\sum_{\xi \in \Lambda} f(\xi) &\leq \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} f(\xi) \\
&\leq Q \prod_{k=1}^r \Pr(Z_{i_{\max},k}^* \in V)^{1/r} && \text{(Inequality (5.3))} \\
&\leq Q \frac{1}{\epsilon_1} \Pr(X_{i_{\max}} \in V) && \text{(Inequality (5.1))} .
\end{aligned}$$

Thus, we have shown that $\sum_{\xi \in \Lambda} f(\xi) = \Theta(Q \Pr(X_{i_{\max}} \in V))$. Since $\epsilon_2 \leq f(\xi) \leq 1$ for all $\xi \in \Lambda$, we have proven Inequality (5.10).

Making use of [39, Lemma 6.4], we can prove Inequality (5.11) by showing $|4\Lambda| \leq C|\Lambda|$ for some constant C . Using Lemma 5.1.3 below (for which we need to assume strict positivity of $\mathbb{E}(e(\beta_j^{(\underline{\mu})} t))$)—see Remark 2.1.3), we have that there exists a constant

$c := c(\epsilon_{-1}, \epsilon_2)$ such that

$$f(\xi) \geq c(\epsilon_{-1}, \epsilon_2),$$

for every $\xi \in 4\Lambda$. Thus,

$$\begin{aligned} |4\Lambda| &\leq \frac{1}{c(\epsilon_{-1}, \epsilon_2)} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} f(\xi) \\ &\leq \left(\frac{1}{c(\epsilon_{-1}, \epsilon_2)} \right) \frac{Q}{\epsilon_1} \Pr(X_{i_{\max}} \in V) = C |\Lambda|, \end{aligned}$$

for some constant C . This completes the proof of Lemma 5.1.2. \square

We now state and prove a lemma showing that $f(\xi)$ is at least a constant for all $\xi \in 4\Lambda$. In [39], the lemma below is unnecessary because an inequality following from [39, Inequality (30)] (which corresponds to Inequality (5.9)) and the triangle inequality suffices.

Lemma 5.1.3. *Let Λ and f be defined as in Equation (5.7) and Equation (5.4), respectively. If $\xi \in 4\Lambda$, then*

$$f(\xi) \geq \left(\epsilon_2 \epsilon_{-1}^{\ln(1/\epsilon_2)} \right)^{320000} =: c(\epsilon_{-1}, \epsilon_2).$$

Note that $c(\epsilon_{-1}, \epsilon_2)$ is a constant.

Proof. Note that Inequality (5.8) implies that for any $\xi' \in \Lambda$ we have

$$\left(\sum_{j=1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi' / Q\|_{\mathbb{R}/\mathbb{Z}}^2 \right)^{1/2} \leq \left(\frac{100r}{\mu} \ln \left(\frac{1}{\epsilon_2} \right) \right)^{1/2}.$$

Thus, by the triangle inequality, we have for any $\xi \in 4\Lambda$ that

$$\left(\sum_{j=1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2 \right)^{1/2} \leq 4 \left(\frac{100r}{\mu} \ln \left(\frac{1}{\epsilon_2} \right) \right)^{1/2}. \quad (5.13)$$

Fix $\xi \in 4\Lambda$. Let k_0 be the number of indices j such that

$$100\mu \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2 > \frac{1}{2},$$

and without loss of generality, say that these indices are $j = 1, 2, \dots, k_0$. Squaring Inequality (5.13), we see that $\frac{k_0}{200\mu} \leq \frac{1600r}{\mu} \ln\left(\frac{1}{\epsilon_2}\right)$, and so we have

$$k_0 \leq 320000r \ln\left(\frac{1}{\epsilon_2}\right),$$

which is a constant. Thus, for the vast majority of the indices j , namely $j = k_0 + 1, k_0 + 2, \dots, n$, we have

$$100\mu \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2 \leq \frac{1}{2}. \quad (5.14)$$

We may now compute that

$$\begin{aligned} f(\xi) &:= \prod_{j=1}^n \left(1 - \mu + \mu \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q) \right)^{1/r} \\ &\geq \epsilon_{-1}^{k_0/r} \prod_{j=k_0+1}^n \left(1 - \mu + \mu \sum_{s=1}^{\ell_j} p_{j,s} \cos(2\pi b_{j,s} a_j \xi / Q) \right)^{1/r} && \text{(since } f(\xi') \geq \epsilon_{-1} \text{ for any } \xi' \text{ by the as-} \\ &\geq \epsilon_{-1}^{k_0/r} \prod_{j=k_0+1}^n \left(1 - 100\mu \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2 \right)^{1/r} && \text{sumption of strict positivity—see Re-} \\ &\geq \epsilon_{-1}^{k_0/r} \exp\left(-\frac{200\mu}{r} \sum_{j=k_0+1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) && \text{mark 2.1.3)} \\ &\geq \epsilon_{-1}^{k_0/r} \exp\left(-\frac{200\mu}{r} \sum_{j=k_0+1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) && \text{(since } \cos(2\pi x) \geq 1 - 100 \|x\|_{\mathbb{R}/\mathbb{Z}}^2 \text{ and} \\ &\geq \epsilon_{-1}^{k_0/r} \exp\left(-\frac{200\mu}{r} \sum_{j=k_0+1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) && \text{the factors are all positive by Inequal-} \\ &\geq \epsilon_{-1}^{k_0/r} \exp\left(-\frac{200\mu}{r} \sum_{j=k_0+1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) && \text{ity (5.14))} \\ &\geq \epsilon_{-1}^{k_0/r} \exp\left(-\frac{200\mu}{r} \sum_{j=k_0+1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) && \text{(} 1 - x \geq e^{-2x} \text{ for } 0 \leq x \leq .79) \\ &\geq \epsilon_{-1}^{k_0/r} \exp\left(-\frac{200\mu}{r} \sum_{j=k_0+1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) && \text{(by Inequality (5.13))} \\ &\geq \epsilon_{-1}^{k_0/r} \exp\left(-\frac{200\mu}{r} \sum_{j=k_0+1}^n \sum_{s=1}^{\ell_j} p_{j,s} \|b_{j,s} a_j \xi / Q\|_{\mathbb{R}/\mathbb{Z}}^2\right) && \\ &= \left(\epsilon_2 \epsilon_{-1}^{\ln(1/\epsilon_2)}\right)^{320000}. \end{aligned}$$

This completes the proof. \square

We have shown that the spectrum Λ has small doubling, and the next step is to use this fact to show that a set containing most of the scaled defining coordinates \tilde{a}_j also has small doubling. Towards that end, we will use the Λ -norm from [39], which is defined as follows: for $x \in \mathbb{Z}/Q\mathbb{Z}$, let $\|x\|_{\Lambda}$ be defined by

$$\|x\|_{\Lambda} := \left(\frac{1}{|\Lambda|^2} \sum_{\xi, \xi' \in \Lambda} \|x(\xi - \xi')/Q\|_{\mathbb{R}/\mathbb{Z}}^2 \right)^{1/2}.$$

Note that $0 \leq \|x\|_\Lambda \leq 1$ for all x and that the triangle inequality holds: $\|x + y\|_\Lambda \leq \|x\|_\Lambda + \|y\|_\Lambda$. We also have that

$$\begin{aligned} \|x\|_\Lambda &\leq \left(\frac{1}{|\Lambda|^2} \sum_{\xi, \xi' \in \Lambda} \|x\xi/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2} + \left(\frac{1}{|\Lambda|^2} \sum_{\xi, \xi' \in \Lambda} \|x\xi'/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2} \\ &= 2\left(\frac{1}{|\Lambda|} \sum_{\xi \in \Lambda} \|x\xi/Q\|_{\mathbb{R}/\mathbb{Z}}^2\right)^{1/2}. \end{aligned}$$

Thus, squaring Inequality (5.9) and summing over all $\xi \in \Lambda$, we have

$$\sum_{j=1}^n \|\tilde{a}_j\|_\Lambda^2 \leq 4C(\epsilon_2, q, r) =: C'. \quad (5.15)$$

We will now show that the set of all x with small Λ -norm, which by Inequality (5.15) includes most of the \tilde{a}_j , has small doubling.

Lemma 5.1.4. [39, Lemma 7.4] *There is a constant C such that the following holds.*

Let $A \subseteq \mathbb{Z}/Q\mathbb{Z}$ denote the “Bohr set”:

$$A := \left\{x \in \mathbb{Z}/Q\mathbb{Z} : \|x\|_\Lambda < \frac{1}{100}\right\}.$$

Then we have

$$C^{-1} \Pr(X_{i_{\max}} \in V)^{-1} \leq |A| \leq |A + A| \leq C \Pr(X_{i_{\max}} \in V)^{-1}.$$

The proof of Lemma 5.1.4 is the same as in [39], with the small modification that a_j should be replaced with $\tilde{a}_j := b_{j,1}a_j$ and the quantity $2^{d_\pm - n}$ should be replaced with $\Pr(X_{i_{\max}} \in V)$ (and, of course, the field F in [39] should be replaced with $\mathbb{Z}/Q\mathbb{Z}$). Also, one should note that [39, Inequality (30)], [39, Inequality (31)], and [39, Inequality (32)] correspond to, respectively, Inequalities (5.9), (5.10), and (5.11).

In the next section, we will complete the proof of the structure theorem using the lemma above.

5.2 Proof of the Structure Theorem (Theorem 4.2.1)

The key to proving the structure theorem is an application of Freiman’s Theorem for finite fields.

Theorem 5.2.1 (see Lemma 6.3 in [39]). *For any constant C there are constants τ and δ such that the following holds. Let A be a non-empty subset of $\mathbb{Z}/Q\mathbb{Z}$, a finite field of prime order Q , such that $|A + A| \leq C|A|$. Then, if Q is sufficiently large depending on $|A|$, there is a symmetric generalized arithmetic progression P of rank τ such that $A \subset P$ and $|A|/|P| \geq \delta$.*

Note that by Lemma 3.1.1 we can assume that Q is sufficiently large with respect to $|A| \leq C \Pr(X_{i_{\max}} \in V)^{-1} \leq C(1/p)^n$ (this follows from V being of medium combinatorial dimension).

The set A from Lemma 5.1.4 satisfies $|A + A| \leq C^2|A|$, where $C \leq O(1)$, and also contains all but $O(1)$ of the scaled defining coordinates \tilde{a}_j , since $\tilde{a}_j \notin A$ implies that $\|\tilde{a}_j\|_\Lambda \geq 1/100$ and Inequality (5.15) shows that there can be at most $100C' = O(1)$ such \tilde{a}_j . By Theorem 5.2.1, there exists a symmetric generalized arithmetic progression $P = \{m_1 v_1 + \dots + m_\tau v_\tau : |m_i| < M_i/2\}$ containing A and satisfying the bounds:

$$\text{rank}(P) = \tau \leq O(1) \text{ and} \tag{5.16}$$

$$|P| \leq M_1 M_2 \dots M_\tau \leq O(\Pr(X_{i_{\max}} \in V)^{-1}). \tag{5.17}$$

The symmetric generalized arithmetic progression P is close to what is needed for Theorem 4.2.1, since it satisfies the required volume and rank bounds. We will show below that P can be altered in ways that preserve Inequalities (5.16) and (5.17) (except possibly for changing the implicit constants) so that P satisfies conditions (i), (ii), and (iii) of Theorem 4.2.1.

To show Theorem 4.2.1(i), we will first add the remaining scaled defining coordinates $\{\tilde{a}_1, \dots, \tilde{a}_n\} \setminus P$ (i.e., those \tilde{a}_j such that $\|\tilde{a}_j\|_\Lambda \geq 1/100$) as new basis vectors v'_k with corresponding dimensions M'_k equal to (say) 3. The resulting generalized arithmetic progression, which we will continue to call P by abuse of notation, satisfies both Inequalities (5.16) and (5.17), since there are only $O(1)$ of the \tilde{a}_j with $\|\tilde{a}_j\|_\Lambda \geq 1/100$ (by Inequality (5.15)). Second, we need to ensure that P is proper, for which we will use the following lemma:

Lemma 5.2.2 (cf. Lemma 9.3 in [39]). *There is an absolute constant $C_0 \geq 1$ such that the following holds. Let P be a symmetric progression of rank τ in an abelian group G ,*

such that every nonzero element of G has order at least $\mathfrak{r}^{C_0\mathfrak{r}^3}|P|$. Then there exists a proper symmetric generalized arithmetic progression P' of rank at most \mathfrak{r} containing P such that

$$|P'| \leq \mathfrak{r}^{C_0\mathfrak{r}^3}|P|.$$

Furthermore, if P is not proper and $\mathfrak{r} \geq 2$, then P' can be chosen to have rank at most $\mathfrak{r} - 1$

One can conclude Lemma 5.2.2 from the proof of [39, Lemma 9.3] (the only difference is noting that the rank can be reduced by at least 1 if P is not proper to begin with). Note that we can always choose Q larger than $\mathfrak{r}^{C_0\mathfrak{r}^3}|P| \leq O\left(\frac{1}{p}\right)^n$.

Applying Lemma 5.2.2 gives us a proper symmetric generalized arithmetic progression, which again we call P by abuse of notation, that contains all the \tilde{a}_j and satisfies both Inequalities (5.16) and (5.17).

The next task is to show that P can be further altered so to meet the condition (ii) in Theorem 4.2.1. Note that there are only $O(1)$ scaled defining coordinates \tilde{a}_j such that $\|\tilde{a}_j\|_\Lambda \geq 1/100$, and so these \tilde{a}_j contribute only a constant to the sum $\sum_{j=1}^n \|\tilde{a}_j\|_P^2$. On the other hand, for any \tilde{a}_j with $\|\tilde{a}_j\|_\Lambda < 1/100$, we have that $k\tilde{a}_j \in A \subset P$ for every positive integer $k < \frac{1}{100\|\tilde{a}_j\|_\Lambda}$. We will exploit this fact, and to do so will need the following notation. Let $\Phi_P : P \rightarrow \mathbb{Z}^\mathfrak{r}$ be the map sending a point $m_1v_1 + \dots + m_\mathfrak{r}v_\mathfrak{r}$ in the proper generalized arithmetic progression P to the unique \mathfrak{r} -tuple of coefficients $(m_1, \dots, m_\mathfrak{r})$.

If the representation for \tilde{a}_j in P is $\tilde{a}_j = m_1v_1 + \dots + m_\mathfrak{r}v_\mathfrak{r}$ and $k\tilde{a}_j$ is in P , we would like to be able to say that the representation for $k\tilde{a}_j$ is $km_1v_1 + \dots + km_\mathfrak{r}v_\mathfrak{r}$; i.e., we hope that $\Phi_P(k\tilde{a}_j)$ is equal to $k\Phi_P(\tilde{a}_j)$. If this were true, then we would have $|km_i| \leq M_i$ for $1 \leq i \leq \mathfrak{r}$, which, if k is large, would show that $\|\tilde{a}_j\|_P$ is small. However, at this point we may well have $\Phi_P(k\tilde{a}_j) \neq k\Phi_P(\tilde{a}_j)$. A priori, changing this to equality would require replacing P with kP and then applying Lemma 5.2.2 to get a proper symmetric generalized arithmetic progression, but since k may be large, this would increase the volume of P too much, violating Inequality (5.17). Luckily, the lemma below provides a way around this difficulty. We will say that P is (k_j, x_j) -proper if $\Phi_P(k_jx_j) = k_j\Phi_P(x_j)$.

Lemma 5.2.3. *There exists an absolute constant C_1 such that the following holds. Let P be a symmetric proper generalized arithmetic progression with rank \mathfrak{r} containing elements x_1, \dots, x_m , and let k_1, \dots, k_m be positive integers such that $\ell_j x_j \in P$ for every $1 \leq \ell_j \leq k_j$ and for every j . Then, there exists a proper symmetric generalized arithmetic progression P' of rank at most \mathfrak{r} such that P' contains P ,*

$$|P'| \leq \mathfrak{r}^{C_1 \mathfrak{r}^4} |P|, \text{ and}$$

P is (k_j, x_j) -proper for every j .

Furthermore, if $r \geq 2$ and if there is some j for which P is not (k_j, x_j) -proper, then P' can be chosen to have rank at most $\mathfrak{r} - 1$.

The proof of this lemma relies on an application of Lemma 5.2.2 to $2P$ (which contains P) along with the fact that if $\|\tilde{a}_j\|_\Lambda < 1/100$ then $k\tilde{a}_j \in P$ for every $1 \leq k < \frac{1}{100\|\tilde{a}_j\|_\Lambda}$.

Proof. We proceed by induction on the rank \mathfrak{r} . For the base case, let $\mathfrak{r} = 1$ and consider $x_j \in P$ such that $k_j x_j \in P$. Since P has rank 1 in this case, we have that $x_j = \Phi_P(x_j)v_1$ and $k_j x_j = \Phi_P(k_j x_j)v_1$. Combining these two equations we have $k_j \Phi_P(x_j)v_1 = \Phi_P(k_j x_j)v_1$, and dividing by v_1 (note that we may assume that $v_1 \neq 0$), we see that $k_j \Phi_P(x_j) = \Phi_P(k_j x_j)$. Thus P is (k_j, x_j) -proper for every j .

For $\mathfrak{r} \geq 2$, we may assume that there is some j_0 such that $k_{j_0} \Phi_P(x_{j_0}) \neq \Phi_P(k_{j_0} x_{j_0})$ (i.e., we assume that P is not (k_{j_0}, x_{j_0}) -proper). We may assume that P has the form $\{m_1 v_1 + \dots + m_\tau v_\tau : |m_i| < M_i/2\}$. Let $M := (M_1, \dots, M_\tau)$, and let $(-M/2, M/2)$ denote the box $\{(m_1, \dots, m_\tau) : |m_i| < M_i/2\}$.

Let \underline{k} be the largest integer such that $\Phi_P(\underline{k}x_{j_0}) = \underline{k}\Phi_P(x_{j_0})$, so $1 \leq \underline{k} < k_{j_0}$ and $\Phi_P((\underline{k}+1)x_{j_0}) \neq (\underline{k}+1)\Phi_P(x_{j_0})$. Since $\underline{k}x_{j_0} \in P$ and $x_{j_0} \in P$, we know that $\Phi_P(x_{j_0}) \in (-M/2, M/2)$ and $\Phi_P(\underline{k}x_{j_0}) = \underline{k}\Phi_P(x_{j_0}) \in (-M/2, M/2)$; and thus, $(\underline{k}+1)\Phi_P(x_{j_0}) \in (-M, M)$. This shows that $2P$, which has dimensions $2M = (2M_1, \dots, 2M_\tau)$, is not proper, since it has two distinct representations for $(\underline{k}+1)x_{j_0}$.

We can now apply Lemma 5.2.2 to $2P$, thus finding a proper symmetric generalized arithmetic progression P' of rank at most $\mathfrak{r} - 1$ containing $2P$ (which contains P) such that

$$|P'| \leq \mathfrak{r}^{C_0 \mathfrak{r}^3} |2P| \leq \mathfrak{r}^{2C_0 \mathfrak{r}^3} |P|.$$

Since P' has rank at most $\mathfrak{r} - 1$, we have by induction that there exists P'' a proper symmetric generalized arithmetic progression of rank at most $\mathfrak{r} - 1$ containing P' and such that

$$|P''| \leq (\mathfrak{r} - 1)^{C_1(\mathfrak{r}-1)^4} |P'| \leq \mathfrak{r}^{C_1(\mathfrak{r}-1)^4} \mathfrak{r}^{2C_0\mathfrak{r}^3} |P|,$$

and such that P'' is (k_j, x_j) -proper for every j . Choosing $C_1 \geq 2C_0$ (for example) guarantees that $\mathfrak{r}^{C_1(\mathfrak{r}-1)^4} \mathfrak{r}^{2C_0\mathfrak{r}^3} \leq \mathfrak{r}^{C_1\mathfrak{r}^4}$, which completes the induction. \square

Applying Lemma 5.2.3, we can generate a new proper symmetric generalized arithmetic progression, which again we will call P by abuse of notation, such that P contains the \tilde{a}_j , satisfies Inequalities (5.16) and (5.17), and is (k_j, \tilde{a}_j) -proper for every \tilde{a}_j such that $\|\tilde{a}_j\|_\Lambda < 1/100$, where $k_j := \left\lceil \frac{1}{200\|\tilde{a}_j\|_\Lambda} \right\rceil \geq 1$. We will now show that such P satisfies part (ii) of Theorem 4.2.1. For \tilde{a}_j such that P is (k_j, \tilde{a}_j) -proper, we have that $|k_j m_i| \leq M_i$ for each $1 \leq i \leq \mathfrak{r}$, and so

$$\|\tilde{a}_j\|_P = \sum_{i=1}^{\mathfrak{r}} \left(\frac{m_i}{M_i} \right)^2 \leq \sum_{i=1}^{\mathfrak{r}} \left(\frac{1}{k_j} \right)^2 \leq \sum_{i=1}^{\mathfrak{r}} (200\|\tilde{a}_j\|_\Lambda)^2 = 40000\mathfrak{r} \|\tilde{a}_j\|_\Lambda^2.$$

Thus, part (ii) of Theorem 4.2.1 follows from Inequality (5.15), since P is (k_j, \tilde{a}_j) -proper for all but $O(1)$ of the \tilde{a}_j .

The next step is to make further alterations to P so that we can prove part (iii) of Theorem 4.2.1. The key property that we will use for (iii) is to have the set of vectors $\{\Phi_P(\tilde{a}_j) : 1 \leq j \leq n\}$ span all of $\mathbb{R}^\mathfrak{r}$, and we will use a rank reduction argument on P to produce a new proper symmetric generalized arithmetic progression satisfying this full rank property.

Lemma 5.2.4. [39] *Let P be a proper symmetric generalized arithmetic progression of rank \mathfrak{r} containing a set B such that the set of vectors $\Phi_P(B)$ does not span $\mathbb{R}^\mathfrak{r}$. Then there exists a symmetric generalized arithmetic progression P' containing P such that*

$$\text{rank}(P') \leq \mathfrak{r} - 1 \text{ and}$$

$$|P'| \leq |P|.$$

Note that the resulting P' is not necessarily proper or (k_j, \tilde{a}_j) -proper, even if P had these properties.

Proof. We use the same proof here as appears in [39, Section 8]. If $\{\Phi_P(\tilde{a}_j) : 1 \leq j \leq n\}$ does not have rank \mathfrak{r} , then it is contained in a subspace of $\mathbb{R}^{\mathfrak{r}}$ of dimension $\mathfrak{r} - 1$. Thus, there exists an integer vector $(\alpha_1, \dots, \alpha_{\mathfrak{r}})$ with all the α_i coprime such that $(\alpha_1, \dots, \alpha_{\mathfrak{r}})$ is orthogonal to every vector in $\{\Phi_P(\tilde{a}_j) : 1 \leq j \leq n\}$. Thus, for every $w \in \mathbb{Z}/Q\mathbb{Z}$ and any $\tilde{a}_j = m_1 v_1 + \dots + m_{\mathfrak{r}} v_{\mathfrak{r}}$, we have that

$$\tilde{a}_j = m_1 v_1 + \dots + m_{\mathfrak{r}} v_{\mathfrak{r}} = m_1(v_1 - w\alpha_1) + \dots + m_{\mathfrak{r}}(v_{\mathfrak{r}} - w\alpha_{\mathfrak{r}}).$$

Since not all the α_i are zero, we may assume that $\alpha_{\mathfrak{r}} \neq 0$. Setting $w = v_{\mathfrak{r}}/\alpha_{\mathfrak{r}}$ so that $v_{\mathfrak{r}} - w\alpha_{\mathfrak{r}} = 0$, we see that P is contained in the symmetric generalized arithmetic progression

$$P' := \{m'_1 v'_1 + \dots + m'_{\mathfrak{r}-1} v'_{\mathfrak{r}-1} : |m'_i| < M_i/2\}$$

with rank $\mathfrak{r} - 1$, dimensions $M_1, \dots, M_{\mathfrak{r}-1}$ (which are the same as the corresponding dimensions for P), and basis vectors $v'_i := v_i - \alpha_i v_{\mathfrak{r}}/\alpha_{\mathfrak{r}}$. By construction $|P'| \leq |P|$. \square

We can now run the following algorithm to create a generalized arithmetic progression with all the desired properties. As the input, we take the generalized arithmetic progression P that we arrived at after applying Lemma 5.2.3, thus the input P contains all the \tilde{a}_j , satisfies Inequalities (5.16) and (5.17), and is (k_j, \tilde{a}_j) -proper for every \tilde{a}_j such that $\|\tilde{a}_j\|_{\Lambda} < 1/100$; however, we do not yet know whether $\Phi_P(\{\tilde{a}_j : 1 \leq j \leq n\})$ spans $\mathbb{R}^{\mathfrak{r}}$.

1. If $\Phi_P(\{\tilde{a}_j : 1 \leq j \leq n\})$ spans $\mathbb{R}^{\mathfrak{r}}$, then do nothing; otherwise apply Lemma 5.2.4.
2. If P is proper, then do nothing; otherwise apply Lemma 5.2.2.
3. If for every \tilde{a}_j with $\|\tilde{a}_j\|_{\Lambda} < 1/100$ we have that P is (k_j, \tilde{a}_j) -proper, then do nothing; otherwise apply Lemma 5.2.3.
4. If P satisfies the three properties given in steps 1, 2, and 3, halt; otherwise, return to step 1.

Each application of a lemma in the algorithm may disrupt some property that other two lemmas preserve; however, we also know that each step in the algorithm either does

not change P or reduces the rank of P by at least 1. Since the original input P has rank $O(1)$, the algorithm must terminate in $O(1)$ steps, giving us a generalized arithmetic progression of rank \mathfrak{r} that satisfies Inequalities (5.16) and (5.17), satisfies conditions (i) and (ii) of Theorem 4.2.1, and satisfies the condition that $\Phi_P(\{\tilde{a}_j : 1 \leq j \leq n\})$ spans all of $\mathbb{R}^{\mathfrak{r}}$.

Thus, all that is left to prove is part (iii), the claim of rational T -commensurability. Though we will not need it in the current section, one should recall that Theorem 4.2.1 is only useful when $|n^{o(n)}T^{O(1)}| = n^{o(n)}$, where T is the symmetric generalized arithmetic progression containing $\{-1, 0, 1\}$ and all possible values taken by the $\beta_{ij}^{(\mu)}$ and the α_{ij} (see Section 4.2).

We say that a set W *economically T -spans* a set U if each $u \in U$ can be represented as a highly T -rational linear combination of elements in W , where each coefficient may be expressed as a/b where $a, b \in n^{o(n)}T^{O(1)}$ and where the implicit constants in the $o(\cdot)$ and $O(\cdot)$ notation are uniform over U .

Comparing our definitions with those from [39, Section 8], we note that “highly rational” means the same thing as “highly $\{-1, 0, 1\}$ -rational”, and “economically spans” means the same thing as “economically $\{-1, 0, 1\}$ -spans”. Thus, it is clear that any highly rational number is also highly T -rational for any T containing $\{-1, 0, 1\}$, and also the statement “ W economically spans U ” implies “ W economically T -spans U ” for any set T containing $\{-1, 0, 1\}$. The remainder of this section paraphrases (with some notational changes) the latter portion of [39, Section 8].

We know that $\Phi_P(\{\tilde{a}_j : 1 \leq j \leq n\})$ spans $\mathbb{R}^{\mathfrak{r}}$. Thus, there exists a subset $U \subset \{\tilde{a}_1, \dots, \tilde{a}_n\}$ of cardinality \mathfrak{r} such that $\Phi_P(U)$ spans $\mathbb{R}^{\mathfrak{r}}$. Renumbering if necessary, we can write $U = \{\tilde{a}_1, \dots, \tilde{a}_{\mathfrak{r}}\}$. It will be important later on that U has cardinality $O(1)$.

The set $\{v_1, \dots, v_{\mathfrak{r}}\}$ of basis vectors for P economically $\{-1, 0, 1\}$ -spans $\{\tilde{a}_1, \dots, \tilde{a}_n\}$ by the definition of P (note that $M_i \leq O(\Pr(X_{i_{\max}} \in V)^{-1}) \leq O(p^{-n}) = n^{o(n)}$), and so by Cramer’s rule, the vectors $\Phi_P(U)$ economically $\{-1, 0, 1\}$ -span the standard basis vectors $\{e_1, \dots, e_{\mathfrak{r}}\}$ for $\mathbb{R}^{\mathfrak{r}}$. Applying Φ_P^{-1} (recall that Φ_P is a bijection since P is proper) shows that U economically $\{-1, 0, 1\}$ -spans $\{v_1, \dots, v_{\mathfrak{r}}\}$.

Following this paragraph, we will show that there exists a single vector v_{i_0} where

$1 \leq i_0 \leq \mathfrak{r}$ such that v_{i_0} economically T -spans U , which will show by transitivity that v_{i_0} economically T -spans $\{\tilde{a}_1, \dots, \tilde{a}_n\}$ (since U economically T -spans $\{v_1, \dots, v_{\mathfrak{r}}\}$ which economically T -spans $\{\tilde{a}_1, \dots, \tilde{a}_n\}$; the relation “economically T -spans” is transitive here since the sets U and $\{v_1, \dots, v_{\mathfrak{r}}\}$ have cardinality $O(1)$).

Let s be the smallest integer such that there exists a subset of cardinality s of $\{v_1, \dots, v_{\mathfrak{r}}\}$ (by renumbering, say the set is $\{v_1, \dots, v_s\}$) so that for some nonzero $d \in n^{o(n)}T^{O(1)}$ and some $c_{ij} \in n^{o(n)}T^{O(1)}$ we have

$$d\tilde{a}_i = \sum_{j=1}^s c_{ij}v_j \text{ for every } 1 \leq i \leq n. \quad (5.18)$$

Note that d does not depend on i , and so this statement is slightly stronger than having $\{v_1, \dots, v_s\}$ economically T -span $\{\tilde{a}_1, \dots, \tilde{a}_n\}$. Also, note that Equation (5.18) holds (for example) with $s = \mathfrak{r}$ by the definition of P and since T contains $\{-1, 0, 1\}$.

We now consider two cases:

- The $n \times s$ matrix $C = (c_{ij})$ has rank 1 in $\mathbb{Z}/Q\mathbb{Z}$. In this case, $\tilde{a}_{i_1}/\tilde{a}_{i_2}$ is highly T -rational for all i_1, i_2 (Since all the c_{ij} are highly T -rational). We know that U economically T -spans $\{v_1, \dots, v_{\mathfrak{r}}\}$, and so the numbers v_{i_1}/v_{i_2} are also highly T -rational (note that it is critical here that U has cardinality $O(1)$). This means that v_1 (for example) economically T -spans $\{v_1, \dots, v_{\mathfrak{r}}\}$, and so by transitivity v_1 economically T -spans U .
- The matrix C has rank at least 2. Recall that (a_1, \dots, a_n) is the normal vector for V and that V is spanned by $(n-1)$ linearly independent vectors with entries in S (recall that S contains all possible values taken by the α_{ij}). We can scale the j -th coordinate of each of these vectors by $b_{j,1}^{-1}$ to get a set of $n-1$ linearly independent vectors each of which is orthogonal to $\tilde{a} := (\tilde{a}_1, \dots, \tilde{a}_n)$. Among these $(n-1)$ linearly independent vectors that are orthogonal to $(\tilde{a}_1, \dots, \tilde{a}_n)$, we can find at least one, say $w = (b_{1,1}^{-1}w_1, \dots, b_{n,1}^{-1}w_n)$, that is not orthogonal to every column of C (since C has column rank at least 2). Let $B := \{b_{j,1} : 1 \leq j \leq n\}$, and let $\tilde{w} := w \prod_{b \in B} b = (\tilde{w}_1, \dots, \tilde{w}_n)$. Thus \tilde{w} is orthogonal to \tilde{a} and every coordinate \tilde{w}_i of \tilde{w} is an element of $T^{O(1)}$ (since T contains S and B and $|B| = O(1)$ by the definition of p -bounded of exponent r).

Remark 5.2.5. Note that the line above is the only place in the proof where we use the assumption from the definition of p -bounded of exponent r that the $\beta_{ij}^{(\mu)}$ take values in a set with cardinality $O(1)$. As is evidenced here, the following weaker assumption suffices instead: say that for each $1 \leq i \leq n$ there exists a set B_i such that $|B_i| = O(1)$ and such that $\beta_{i1}^{(\mu)}, \beta_{i2}^{(\mu)}, \dots, \beta_{in}^{(\mu)}$ each take a nonzero value in B with probability at least q . In fact, this weaker assumption also replaces the assumption in the definition of p -bounded of exponent r that $q \leq \min_x \Pr(\beta_{ij}^{(\mu)} = x)$ for every i, j : It suffices for each $\beta_{ij}^{(\mu)}$ to take one value in B_i with probability at least q , instead of taking every value with probability at least q .

We may now compute:

$$0 = d\tilde{a} \cdot \tilde{w} = \sum_{i=1}^n d\tilde{a}_i \tilde{w}_i = \sum_{i=1}^n \sum_{j=1}^s c_{ij} v_j \tilde{w}_i = \sum_{j=1}^s \left(\sum_{i=1}^n c_{ij} \tilde{w}_i \right) v_j.$$

Since \tilde{w} is not orthogonal to every column of $C = (c_{ij})$, we can assume (reordering if necessary), that the coefficient for v_s above is nonzero, and thus we have

$$v_s = \frac{-1}{\sum_{\ell=1}^n c_{\ell s} \tilde{w}_\ell} \sum_{j=1}^{s-1} \left(\sum_{\ell=1}^n c_{\ell j} \tilde{w}_\ell \right) v_j.$$

Plugging this last equation into Equation (5.18), we arrive at

$$d \left(\sum_{\ell=1}^n c_{\ell s} \tilde{w}_\ell \right) \tilde{a}_i = \sum_{j=1}^{s-1} \left(c_{ij} \sum_{\ell=1}^n c_{\ell s} \tilde{w}_\ell - c_{is} \sum_{\ell=1}^n c_{\ell j} \tilde{w}_\ell \right) v_j.$$

Since the coefficient for \tilde{a}_i on the left is an element of $n^{o(n)}T^{O(1)}$ and the coefficient for each v_j on the right is an element of $n^{o(n)}T^{O(1)}$, we have contradicted the minimality of s .

Thus, we have completed the proof of the structure theorem (Theorem 4.2.1). \square

Chapter 6

Extension to the case where some rows are fixed

6.1 A generalization: f rows have fixed, non-random values

In this section, we will give a generalization of Theorem 2.1.2 to the case where the random matrix N_n has $f \leq O(\ln n)$ rows that are assumed to be linearly independent and contain fixed, non-random entries. The proof of the generalized result is very similar to the proof of Theorem 2.1.2, and we will sketch the main differences in the two proofs below.

Definition 6.1.1 (a random matrix $N_{f,n}$ with entries in S). Let f be an integer between 1 and n , let S be a subset of a ring, and let $N_{f,n}$ be an n by n matrix defined as follows. For $1 \leq i \leq f$ and $1 \leq j \leq n$, let the entries s_{ij} of $N_{f,n}$ be fixed (non-random) elements of S such that the rows $(s_{i,1}, \dots, s_{i,n})$ for $1 \leq i \leq f$ are linearly independent. For $f+1 \leq i \leq n$ and $1 \leq j \leq n$, let the entries α_{ij} of $N_{f,n}$ be discrete finite random variables taking values in S . Thus,

$$N_{f,n} := \left(\begin{array}{cccc} s_{1,1} & s_{1,2} & \cdots & s_{1,n} \\ \vdots & \cdots & \cdots & \vdots \\ s_{f,1} & \cdots & \cdots & s_{f,n} \\ \alpha_{f+1,1} & \alpha_{f+1,2} & \cdots & \alpha_{f+1,n} \\ \alpha_{f+2,1} & \alpha_{f+2,2} & \cdots & \alpha_{f+2,n} \\ \alpha_{f+3,1} & \alpha_{f+3,2} & \cdots & \alpha_{f+3,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{array} \right) \left. \begin{array}{l} \right\} \text{Fixed rows; assumed to be lin-} \\ \text{early independent} \\ \\ \left. \right\} \text{Random rows}$$

Theorem 6.1.2. *Let p be a positive constant such that $0 < p < 1$, let r be a positive integer constant, and let S be a generalized arithmetic progression in the complex*

numbers with rank $O(1)$ (independent of n) and with cardinality at most $|S| \leq n^{o(n)}$. Consider the matrix $N_{\mathfrak{f},n}$ with entries in S (see Definition 6.1.1 above), where $\mathfrak{f} \leq \left(\frac{r}{2\ln(1/p)} - o(1)\right) \ln n$. If the collection of random variables $\{\alpha_{jk}\}_{\mathfrak{f}+1 \leq j \leq n, 1 \leq k \leq n}$ is p -bounded of exponent r , then

$$\Pr(N_{\mathfrak{f},n} \text{ is singular}) \leq \max \left\{ (p^{1/r} + o(1))^n, (p + o(1))^{n-\mathfrak{f}} \right\}.$$

Note that the bound on the singularity probability of $N_{\mathfrak{f},n}$ for $r \geq 2$ is the same as in Theorem 2.1.2 (since for $r \geq 2$, we have $n/r \ll n - c \ln n = n - \mathfrak{f}$). This is a reflection of the fact that only the large dimension case uses the randomness in all the rows simultaneously, and in that case the exponential bound does not depend on r . Generally speaking, the best known lower bounds on the singularity probability of a discrete random matrix come from a dependency among at most two random rows, and since $N_{\mathfrak{f},n}$ certainly has more than two random rows, the upper bounds given in Theorem 6.1.2 seem reasonable.

Theorem 6.1.2 leads to Corollary 1.1.2 by following a conditioning argument very similar to that given in Section 2.2.3.

6.1.1 Outline of the proof of Theorem 6.1.2

The proof of Theorem 6.1.2 follows the same lines of reasoning as that of Theorem 2.1.2. In this subsection, we will state the main lemmas with the necessary modifications, and we will mention a few important considerations when making the modifications.

Note that Equation (4.1), which reduces the question of singularity to one of the rows spanning non-trivial hyperplane of dimension $n - 1$ holds in the current context, using the same definition of A_V and “non-trivial hyperplane” (both are defined after Equation (4.1) in Section 4.1.1).

Definition 6.1.3 (combinatorial dimension with \mathfrak{f} fixed rows). Let $\mathcal{D} := \{\frac{a}{n} : 0 \leq a \leq n^2, a \in \mathbb{Z}\}$. For any $d_{\pm} \in \mathcal{D}$, we define the *combinatorial Grassmannian* $\text{Gr}_{\mathfrak{f}}(d_{\pm})$ to be the set of all non-trivial hyperplanes V in $(\mathbb{Z}/Q\mathbb{Z})^n$ such that

$$p^{n-d_{\pm}+1/n} < \max_{\mathfrak{f}+1 \leq i \leq n} \Pr(X_i \in V) \leq p^{n-d_{\pm}}.$$

For $d_{\pm} = 0$, we define $\text{Gr}_f(0)$ to be the set of all non-trivial hyperplanes such that

$$\max_{f+1 \leq i \leq n} \Pr(X_i \in V) \leq p^n.$$

We will refer to d_{\pm} as the *combinatorial dimension* of V .

Lemma 6.1.4 (Small combinatorial dimension, with f fixed rows). *For any $\delta > 0$ we have*

$$\sum_{d_{\pm} \in \mathcal{D} \text{ s.t. } T^{d_{\pm}} q^n \leq \delta^n} \sum_{V \in \text{Gr}_f(d_{\pm})} \Pr(A_V) \leq (n - f)\delta^n.$$

Proof. The proof is the same as that for Lemma 4.1.2; also see [22], [38], [39]. \square

Lemma 6.1.5 (Large combinatorial dimension, with f fixed rows). *We have*

$$\sum_{d_{\pm} \in \mathcal{D} \text{ s.t. } \frac{c_{\text{LgDim}}}{n^{1/2}} \leq T^{d_{\pm}} q^n} \sum_{V \in \text{Gr}_f(d_{\pm})} \Pr(A_V) \leq (p + o(1))^{n-f}$$

Here, c_{LgDim} is the same as in Lemma 4.1.3.

Proof. The proof is the same as that for Lemma 4.1.3, except now we appeal to Lemma A.1.2 with $f > 0$. Note that we must assume $f \leq n/2$ in order to apply Lemma A.1.2. See also [22],[38],[39]. \square

Proposition 6.1.6 (Medium combinatorial dimension estimate, with f fixed rows).

Let $0 < \epsilon_0$ be a constant much smaller than 1, and let $d_{\pm} \in \mathcal{D}$ be such that $(p + c_{\text{MedDim},f\epsilon_0})^{n/r} < T^{d_{\pm}} q^n < \frac{c_{\text{LgDim}}}{\sqrt{n}}$. If $f \leq \left(\frac{r}{2 \ln(1/p)} - o(1)\right) \ln n$, then

$$\sum_{V \in \text{Gr}_f(d_{\pm})} \Pr(A_V) \leq (p + o(1))^{n/r}.$$

Here we choose the constant $c_{\text{MedDim},f}$ so that $c_{\text{MedDim},f} > (c_m + c_f + \frac{1}{100})$, where c_m and c_f are positive absolute constants (in particular, we need c_f such that $f \leq \frac{c_f \epsilon_0 n}{r}$, which is true for any positive constant c_f since $f \leq O(\ln n)$). As before, we will prove this proposition by separating V with medium combinatorial dimension into two cases: exceptional and unexceptional, which are defined below using the definition of $Z_{i,k}^*$ from Equation (4.3) (this definition is the same as in Definition 4.1.5 with the small change that i and j are required to be between $f + 1$ and n instead of between 1 and n).

Definition 6.1.7. Consider a hyperplane V of medium combinatorial dimension (that is, d_{\pm} satisfies the condition in Proposition 6.1.6). We say V is *unexceptional* if there exists an i_0 where $f+1 \leq i_0 \leq n$ and there exists a k_0 where $1 \leq k_0 \leq r$ such that

$$\max_{f+1 \leq j \leq n} \{\Pr(X_j \in V)\} < \epsilon_1 \Pr(Z_{i_0, k_0}^* \in V).$$

We say V is *exceptional* if for every i where $f+1 \leq i \leq n$ and for every k where $1 \leq k \leq r$ we have

$$\epsilon_1 \Pr(Z_{i, k}^* \in V) \leq \max_{f+1 \leq j \leq n} \{\Pr(X_j \in V)\}. \quad (6.1)$$

In particular, there exists i_{\max} such that $\Pr(X_{i_{\max}} \in V) = \max_{f+1 \leq j \leq n} \{\Pr(X_j \in V)\}$; and so if V is exceptional, then

$$\epsilon_1 \Pr(Z_{i_{\max}, k}^* \in V) \leq \Pr(X_{i_{\max}} \in V) \quad \text{for every } k. \quad (6.2)$$

We will refer to $X_{i_{\max}}$ as the *exceptional row*.

Lemma 6.1.8 (Unexceptional space estimate, with f fixed rows). *If $f \leq \frac{c_f \epsilon_0 n}{r}$ for some positive constant c_f , then we have*

$$\sum_{V \in \text{Gr}_f(d_{\pm}): V \text{ is unexceptional}} \Pr(A_V) \leq p^{-o(n)} 2^n \epsilon_1^{c_m \epsilon_0 n / r}.$$

Notice that the bound is the same as in Lemma 4.1.6, except that we replaced c_{MedDim} with $c_{\text{MedDim}, f}$ when defining “unexceptional”.

Proof. The proof follows in the same way as that for Lemma 4.1.6; however, when replacing rows X_i of $N_{f, n}$ with rows Z_i that concentrate more sharply on V , we must take care to only replace random rows of $N_{f, n}$ (i.e., rows X_1, \dots, X_f must not be replaced by Z_i). See Appendix A.2 for details. \square

In the exceptional case, The same structure theorem (Theorem 4.2.1) holds, leading to the following lemma.

Lemma 6.1.9 (Exceptional space estimate, with f fixed rows). *If $f \leq \left(\frac{r}{2 \ln(1/p)} - o(1)\right) \ln n$, then*

$$\sum_{V \in \text{Gr}(d_{\pm}): V \text{ is exceptional}} \Pr(A_V) \leq p^{n/r} \quad (6.3)$$

Note that this upper bound is dramatically worse than the analogous upper bound in Lemma 4.1.7 of $n^{-\frac{n}{2}+o(n)}$.

Proof. As in Lemma 4.1.7, the main step in the proof is applying the structure theorem (Theorem 4.2.1). In the current context, Inequality (4.6) holds with $n - f$ as the exponent instead of n (since there are only $n - f$ random rows). If we combine this modified version of Inequality (4.6) with Inequality (4.7), then we have the bound

$$\begin{aligned} \sum_{\substack{V \in \text{Gr}_f(d_{\pm}): \\ V \text{ is exceptional}}} \Pr(A_V) &\leq n^{-\frac{n}{2}+o(n)} \Pr(X_{i_{\max}} \in V)^{-n} \Pr(X_{i_{\max}} \in V)^{n-f} \\ &= n^{-\frac{n}{2}+o(n)} \Pr(X_{i_{\max}} \in V)^{-f}, \end{aligned}$$

where by assumption $X_{i_{\max}}$ is the random row such that $\Pr(X_{i_{\max}} \in V) = \max_{f+1 \leq i \leq n} \Pr(X_i \in V)$. In order for this upper bound to achieve the desired bound in Inequality (6.3), it is sufficient to have

$$n^{-\frac{n}{2}+o(n)} \Pr(X_{i_{\max}} \in V)^{-f} \leq p^{n/r}. \quad (6.4)$$

Using the assumption that $\Pr(X_{i_{\max}} \in V) \geq (p + c_{\text{MedDim}, f \in 0})^{n/r} > p^{n/r}$ (since V is of medium combinatorial dimension), we see that Inequality (6.4) holds whenever

$$f \leq \left(\frac{r}{2 \ln(1/p)} - o(1) \right) \ln n,$$

which completes the proof. □

Appendix A

Lemmas in additive combinatorics

A.1 Two background results

A.1.1 A version of the Littlewood-Offord result in $\mathbb{Z}/Q\mathbb{Z}$

If $S \subset \mathbb{Q}$, then we can clear denominators and prove (as in [39, Lemma 2.4]) the large combinatorial dimension estimate in \mathbb{R} instead of working in $\mathbb{Z}/Q\mathbb{Z}$, in which case we can also use the Littlewood-Offord result over \mathbb{R} (see [40, Corollary 7.13]), instead of the version over $\mathbb{Z}/Q\mathbb{Z}$ given here in Lemma A.1.1. When working in \mathbb{R} , the integral approximation of Inequality (A.1) can be replaced by a limit going to infinity, and we do not need any extra assumptions on Q . In particular, we may take $Q \approx \exp(\exp(Cn))$ (see Remark 3.1.2).

For Q sufficiently large with respect to q , r , and n , it is clear that we have

$$\frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} (1 - 2q + 2q \cos(2\pi\xi/Q))^{k/r} \leq \int_0^1 (1 - 2q + 2q \cos(2\pi t))^{k/r} dt + \frac{1}{n}, \quad (\text{A.1})$$

for all $1 \leq k \leq n$.

Lemma A.1.1. *Let Q be sufficiently large to satisfy Inequality (A.1), and let $v_1, \dots, v_n \in \mathbb{Z}/Q\mathbb{Z}$ be such that v_1, \dots, v_k are nonzero. Let $\{\alpha_j\}_{j=1}^n$ be a collection of random variables that are p -bounded of exponent r , and let $X_{\mathbf{v}} := \alpha_1 v_1 + \dots + \alpha_n v_n$. Then, for every $x \in \mathbb{Z}/Q\mathbb{Z}$ we have*

$$\Pr(X_{\mathbf{v}} = x) \leq \frac{c_{\text{LO}} \sqrt{r}}{\sqrt{qk}} = O\left(\frac{1}{\sqrt{k}}\right),$$

where c_{LO} is an absolute constant.

Proof. Our proof is closely modeled on the proof of [40, Corollary 7.13]. Let $\beta_j^{(\mu)}$ be the symmetric random variables from the definition of p -bounded of exponent r

corresponding to α_j (see Equation (2.1)). Then, we can compute

$$\begin{aligned}
\Pr(X_{\mathbf{v}} = x) &\leq \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \prod_{j=1}^k |\mathbb{E}(e_Q(\alpha_j a_j \xi))| && \text{(note that } a_j = 0 \text{ for } j > k) \\
&\leq \prod_{j=1}^k \left(\frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} |\mathbb{E}(e_Q(\alpha_j a_j \xi))|^k \right)^{1/k} && \text{(Hölder's inequality)} \\
&\leq \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} |\mathbb{E}(e_Q(\alpha_{j_0} a_{j_0} \xi))|^k && \text{(where } j_0 \text{ corresponds to the largest factor in the previous line)} \\
&\leq \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} \left(1 - \mu + \mu \sum_{s=1}^{\ell_{j_0}} p_{j_0,s} \cos(2\pi b_{j_0,s} v_{j_0} \xi / Q) \right)^{k/r} && \text{(since } \alpha_{j_0} \text{ is } p\text{-bounded of exponent } r) \\
&\leq \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} (1 - 2q + 2q \cos(2\pi b_{j_0,1} v_{j_0} \xi / Q))^{k/r} && \text{(since } \mu p_{j_0,1} \geq 2q) \\
&= \frac{1}{Q} \sum_{\xi \in \mathbb{Z}/Q\mathbb{Z}} (1 - 2q + 2q \cos(2\pi \xi / Q))^{k/r} && \text{(by reordering the sum) .}
\end{aligned}$$

Combining the above inequalities with Inequality (A.1) and following the proof of [40, Corollary 7.13] to bound the integral, we have

$$\begin{aligned}
\Pr(X_{\mathbf{v}} = x) &\leq \int_0^1 (1 - 2q + 2q \cos(2\pi t))^{k/r} dt + \frac{1}{n} \\
&= \frac{c_{\text{LO}} \sqrt{r}}{\sqrt{qk}} = O\left(\frac{1}{\sqrt{k}}\right),
\end{aligned}$$

where c_{LO} is an absolute constant. □

A.1.2 A generalization of a lemma due to Komlós [26]

This lemma is a generalization of the result in [26] (see also [3, Lemma 14.10], [22, Section 3.1], and [38, Lemma 5.3]).

Lemma A.1.2. *Fix n , and let p be a positive constants such that $0 < p < 1$ and let r be a positive integer constant. Consider the matrix $N_{\mathfrak{f},n}$ taking values in $\mathbb{Z}/Q\mathbb{Z}$, where $\mathfrak{f} \leq n/2$ and Q is large enough to satisfy Inequality (A.1). If the collection of random entries in $N_{\mathfrak{f},n}$ is p -bounded of exponent r , then*

$$\Pr\left(\text{there exists } \mathbf{v} \in \Omega_1 \text{ such that } N_{\mathfrak{f},n} \cdot \mathbf{v} = 0\right) \leq (p + o(1))^{n-\mathfrak{f}},$$

where

$$\Omega_1 := \left\{ (v_1, \dots, v_n) \in \mathbb{Z}/Q\mathbb{Z} : \text{at most } (n - \mathfrak{f}) \left(1 - \frac{c}{\ln n}\right) + 1 \text{ of the } v_i \text{ are nonzero} \right\} \setminus \{\underline{0}\},$$

where the constant c can be taken to be $c \geq 2 \ln(100/p)$, and where $\underline{0}$ denotes the zero vector.

Proof. Let $E_k = \{\text{there exists } v \in \Omega_1 \text{ with at most } k \text{ nonzero coordinates such that } N_{\mathfrak{f},n} \cdot v = 0\}$. Clearly,

$$\Pr\left(\text{there exists } v \in \Omega_1 \text{ such that } N_{\mathfrak{f},n} \cdot v = 0\right) \leq \sum_{1 \leq k \leq (n-\mathfrak{f})\left(1-\frac{c}{\ln n}\right)+1} \Pr(E_k \setminus E_{k-1}).$$

Let S be the set of all possible values that could appear as entries in $N_{\mathfrak{f},n}$, and let $N_{\mathfrak{f},n}|_{j_1, \dots, j_k}$ be the n by k matrix consisting of columns j_1, \dots, j_k of $N_{\mathfrak{f},n}$. Following [26, Lemma 2] (see also [3, Lemma 14.10] and [38, Lemma 5.3]) we can write

$$\Pr(E_k \setminus E_{k-1}) \leq \sum_{\substack{1 \leq j_1 < \dots \\ \dots < j_k \leq n}} \sum_{\substack{1 \leq i_1 < \dots \\ \dots < i_{k-1} \leq n}} \sum_{\substack{H \text{ a } (k-1)\text{-} \\ \text{dimensional} \\ \text{hyperplane} \\ \text{spanned by } S^k}} \Pr(\text{RwSpn}_{i_1, \dots, i_{k-1}, H}) \Pr(\text{RwIn}_{i_1, \dots, i_{k-1}, H}),$$

where

$$\begin{aligned} \text{RwSpn}_{i_1, \dots, i_{k-1}, H} &:= \left\{ \text{rows } i_1, \dots, i_{k-1} \text{ of } N_{\mathfrak{f},n}|_{j_1, \dots, j_k} \text{ span } H \right\}, \text{ and} \\ \text{RwIn}_{i_1, \dots, i_{k-1}, H} &:= \left\{ \text{all rows of } N_{\mathfrak{f},n}|_{j_1, \dots, j_k} \text{ except } i_1, \dots, i_{k-1} \text{ are in } H \right\}. \end{aligned}$$

Let $U(k, p, q)$ be a uniform upper bound for $\Pr(\text{row } i \text{ is in } H)$, where $\mathfrak{f} + 1 \leq i \leq n$ and q is the constant from Definition 2.1.1 (here, we mean uniform with respect to the index sets $\{j_1, \dots, j_k\}$ and $\{i_1, \dots, i_k\}$). Then we have

$$\Pr(E_k \setminus E_{k-1}) \leq U(k, p, q)^{n-k-\mathfrak{f}+1} \binom{n}{k} \binom{n}{k-1},$$

since $k-1$ fixed rows of $N_{\mathfrak{f},n}|_{j_1, \dots, j_k}$ can span at most 1 hyperplane H of dimension $k-1$.

For $k \leq \frac{2^8 c_{\text{L}}^2 O^r}{p^2 q}$ (a constant), we can set $U(k, p, q) = p$ by the Weighted Odlyzko Lemma (see Lemma A.2.1), giving us a bound of

$$\Pr(E_k \setminus E_{k-1}) \leq (p + o(1))^{n-\mathfrak{f}}. \quad (\text{A.2})$$

For $\frac{2^8 c_{\text{LO}}^2 r}{p^2 q} < k \leq (n - f) \left(1 - \frac{c}{\ln n}\right) + 1$, we use Lemma A.1.1 to set $U(k, p, q) = \frac{c_{\text{LO}} \sqrt{r}}{\sqrt{qk}}$. Since $\binom{n}{k} \binom{n}{k-1} \leq \frac{2^{2n}}{n}$ we thus have

$$\Pr(E_k \setminus E_{k-1}) \leq \frac{1}{n} 2^{2n} \left(\frac{c_{\text{LO}}^2 r}{qk}\right)^{\frac{n-k-f+1}{2}}.$$

As a function of k , this upper bound has strictly positive second derivative; thus, the largest upper bound will occur at one of the extremal values of $k = \frac{2^8 c_{\text{LO}}^2 r}{p^2 q}$ or $k = (n - f) \left(1 - \frac{c}{\ln n}\right) + 1$, and a bit of computation shows that

$$\Pr(E_k \setminus E_{k-1}) \leq \frac{1}{n} O(p^{n-f}). \quad (\text{A.3})$$

Summing the bounds in Inequalities (A.2) and (A.3) completes the proof. \square

A.2 The unexceptional case with f fixed rows

This section is adapted from the proof of [39, Lemma 4.1], and proves Lemma 4.1.6 by setting $f = 0$. Assume that $f \leq \frac{c_f \epsilon_0 n}{r}$, and let m be the closest integer to $\frac{c_m \epsilon n}{r}$. Let Z_1, \dots, Z_m be i.i.d. copies of the unexceptional row vector Z_{i_0, k_0}^* from Definition 6.1.7, so $\epsilon_1 \Pr(Z_i \in V) > \Pr(X_i \in V)$ for all $f + 1 \leq i \leq n$. We will need the following version of the Weighted Odlyzko Lemma:

Lemma A.2.1. [cf. [39, Lemma 4.3] or [22, Section 3.2]] *For $1 \leq i$, let W_{i-1} be an $(f + i - 1)$ -dimensional subspace containing X_1, \dots, X_f (which are fixed, linearly independent row vectors). Then*

$$\Pr(Z_i \in W_{i-1}) \leq \left(p + \frac{\epsilon_0}{100}\right)^{\frac{n}{r} - f - i + 1}.$$

Proof. Since W_{i-1} has dimension $f + i - 1$, there exists a set of $f + i - 1$ “determining” coordinates such that if a vector $V \in W_{i-1}$, then the $f + i - 1$ “determining” coordinates determine the values of the remaining $n - f - i + 1$ coordinates. Since the maximum probability that any of the n/r random coordinates in Z_i takes a given value is at most $1 - \underline{\mu} = p + \frac{\epsilon_0}{100}$, and since there are at least $\frac{n}{r} - f - i + 1$ of the random coordinates in Z_i that are not among the “determining” coordinates, we have the desired upper bound. \square

Let $V_0 := \text{Span}\{X_1, \dots, X_f\}$, the space spanned by the f fixed rows, and for $1 \leq i \leq m$ let $B_{V,i}$ be the event that Z_1, \dots, Z_m are linearly independent in $V \setminus V_0$. We have the following analog of Lemma 4.1.8 (and also [39, Lemma 4.4]):

Lemma A.2.2 (see Lemma 4.4 in [39]). *Let m, f , and $B_{V,m}$ be as defined above. Then,*

$$\Pr(B_{V,m}) \geq p^{\rho(n)} \left(\frac{\max_{f+1 \leq i \leq n} \Pr(X_i \in V)}{\epsilon_1} \right)^m$$

Proof. Using Bayes' Identity, we have

$$\Pr(B_{V,m}) = \prod_{i=1}^m \Pr(B_{V,i} | B_{V,i-1}), \quad (\text{A.4})$$

where $B_{V,0}$ denotes the full space of the Z_i . Conditioning on a particular instance of Z_1, \dots, Z_{i-1} in $B_{V,i-1}$, we have that

$$\Pr(B_{V,i} | B_{V,i-1}) = \Pr(Z_i \in V) - \Pr(Z_i \in W_{i-1}),$$

where W_{i-1} denotes the $(f + i - 1)$ -dimensional space spanned by X_1, \dots, X_f and Z_1, \dots, Z_{i-1} . We will now establish a uniform bound that does not depend on which particular instance of Z_1, \dots, Z_{i-1} in $B_{V,i-1}$ that we fixed by conditioning. By the definition of unexceptional, we have

$$\Pr(Z_i \in V) > \frac{1}{\epsilon_1} \max_{f+1 \leq i \leq n} \Pr(X_i \in V),$$

and by the Weighted Odlyzko Lemma (see Lemma A.2.1), we have

$$\Pr(Z_i \in W_{i-1}) \leq \left(p + \frac{\epsilon_0}{100} \right)^{\frac{n}{r} - f - i + 1} \leq \left(p + \frac{\epsilon_0}{100} \right)^{\frac{n}{r} (1 - (c_m + c_f) \epsilon_0)}.$$

Using Taylor's Theorem with remainder (for example), one can show that

$$\left(p + \frac{\epsilon_0}{100} \right)^{\frac{n}{r} (1 - (c_m + c_f) \epsilon_0)} \leq \frac{1}{2n} (p + c_{\text{MedDim}} \epsilon_0)^{n/r} \leq \frac{1}{n} \max_{f+1 \leq i \leq n} \Pr(X_i \in V),$$

so long as $c_{\text{MedDim}} > \frac{1}{100} + c_m + c_f > \frac{1}{100} + (c_m + c_f) p \ln \left(\frac{1}{p} \right)$ and n is sufficiently large (the second inequality in the display above is the definition of medium combinatorial dimension).

Thus

$$\Pr(B_{V,i} | B_{V,i-1}) \geq \frac{1}{\epsilon_1} \left(\max_{f+1 \leq i \leq n} \Pr(X_i \in V) \right) \left(1 - \frac{\epsilon_1}{n} \right),$$

and plugging this estimate back into Inequality (A.4) we get

$$\Pr(B_{V,m}) \geq p^{o(n)} \left(\frac{\max_{f+1 \leq i \leq n} \Pr(X_i \in V)}{\epsilon_1} \right)^m.$$

□

To conclude Lemma 6.1.8 (which implies Lemma 4.1.6 by setting $f = 0$), we will proceed as in the proof for [39, Lemma 4.1].

Let Z_1, \dots, Z_m be i.i.d. copies of Z_{i_0, k_0}^* that are independent of the random rows X_{f+1}, \dots, X_n . Using independence and Bayes' Identity we have

$$\Pr(A_V) = \Pr(A_V | B_{V,m}) = \frac{\Pr(A_V \wedge B_{V,m})}{\Pr(B_{V,m})} \leq \Pr(A_V \wedge B_{V,m}) p^{-o(n)} \left(\frac{\epsilon_1}{\max_{f+1 \leq i \leq n} \Pr(X_i \in V)} \right)^m.$$

Because the Z_i are linearly independent in $V \setminus V_0$, we know that there is a subset $I \subset \{f+1, f+2, \dots, n\}$ of cardinality $|I| = m$, such that $\{Z_1, \dots, Z_m\} \cup \{X_i : i \notin I\}$ spans V . Let $C_{V,I}$ be the event that $\{Z_1, \dots, Z_m\} \cup \{X_i : i \notin I\}$ spans V . Then we have

$$\begin{aligned} \Pr(A_V \wedge B_{V,m}) &\leq \sum_{\substack{I \subset \{f+1, \dots, n\} \\ |I|=m}} \Pr(C_{V,I} \wedge \{X_i \in V : i \in I\}) \\ &\leq \left(\max_{f+1 \leq i \leq n} \Pr(X_i \in V) \right)^m \sum_{\substack{I \subset \{f+1, \dots, n\} \\ |I|=m}} \Pr(C_{V,I}). \end{aligned}$$

Summing the above inequality over all unexceptional V (note that $\sum_V \Pr(C_{V,I}) \leq 1$) and combining with the bound for $\Pr(A_V)$ above gives us

$$\begin{aligned} \sum_{\text{unexceptional } V} \Pr(A_V) &\leq \left(\max_{f+1 \leq i \leq n} \Pr(X_i \in V) \right)^m \binom{n-f}{m} p^{-o(n)} \left(\frac{\epsilon_1}{\max_{f+1 \leq i \leq n} \Pr(X_i \in V)} \right)^m \\ &\leq p^{-o(n)} 2^n \epsilon_1^m. \end{aligned}$$

This completes the proof of the estimate for unexceptional V .

References

- [1] Aizenman, Michael; Germinet, Francois; Klein, Abel; Warzel, Simone. On Bernoulli decompositions for random variables, concentration bounds, and spectral localization, *Probab. Theory Related Fields* **143** (2009), no. 1-2, 219–238.
- [2] Aronov, Boris; Pach, János; Sharir, Micha; Tardos, Gábor. Distinct distances in three and higher dimensions. *Combin. Probab. Comput.* **13** (2004), no. 3, 283–293.
- [3] Bollobás, Béla. *Random graphs. Second edition.* Cambridge Studies in Advanced Mathematics, 73. Cambridge University Press, Cambridge, 2001.
- [4] Bourgain, Jean. More on the sum-product phenomenon in prime fields and its applications. *Internat. J. Number Theory* **1** (2005), 1–32.
- [5] Bourgain, Jean, Katz, Nets, Tao, Terence. A sum-product estimate in finite fields, and applications. *Geom. Funct. Anal.* **14** (2004), no. 1, 27–57.
- [6] Bourgain, J.; Glibichuk, A. A.; Konyagin, S. V. Estimates for the number of sums and products and for exponential sums in fields of prime order. *J. London Math. Soc.* (2) **73** (2006), no. 2, 380–398.
- [7] Bourgain, Jean; Vu, Van H.; Wood, Philip Matchett. On the singularity probability of discrete complex random matrices. *in preparation.*
- [8] Chang, Mei-Chu. A sum-product estimate in algebraic division algebras. *Israel J. Math.* **150** (2005), 369–380.
- [9] Chang, Mei-Chu. On sum-product representations in \mathbb{Z}_q . *J. Eur. Math. Soc. (JEMS)* **8** (2006), no. 3, 435–463.
- [10] Chang, Mei-Chu. Product theorems in SL_2 and SL_3 , *J. Inst. Math. Jussieu* **7** (2008), no. 1, 1–25.
- [11] Chang, Mei-Chu. Additive and multiplicative structure in matrix spaces. *Combin. Probab. Comput.* **16** (2007), no. 2, 219–238.
- [12] Davidoff, Giuliana; Sarnak, Peter; Valette, Alain. *Elementary number theory, group theory, and Ramanujan graphs.* London Mathematical Society Student Texts, 55. Cambridge University Press, Cambridge, 2003.
- [13] Elekes, Gy. On the number of sums and products, *Acta Arith.*, **81** (1997) 365–367.
- [14] Erdős, Paul; Szemerédi, Endré. On sums and products of integers. *Studies in Pure Mathematics* (Birkhäuser, Basel, 1983) 213–218.

- [15] Frobenius, Ferdinand Georg. Über Beziehungen zwischen den Primidealen eines algebraischen Körpers und den Substitutionen seiner Gruppe, *Sitzungsberichte Königl. Preußisch. Akad. Wissenschaft. Berlin* (1896), 689–703; *Gesammelte Abhandlungen II*, 719–733. (German) Herausgegeben von J.-P. Serre. Springer-Verlag, Berlin-New York 1968.
- [16] Ford, K. Sums and products from a finite set of real numbers. *Ramanujan J.* **2** (1998) 59–66.
- [17] Gamburd, Alex. On the spectral gap for infinite index “congruence” subgroups of $SL_2(\mathbb{Z})$. *Israel J. Math.* **127** (2002), 157–200.
- [18] Garaev, M.Z. An explicit sum-product estimate in \mathbb{F}_p , *Int. Math. Res. Not. IMRN* **2007**, no. 11, Art. ID rnm035, 11 pp.
- [19] Helfgott, H.A. Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$, *Ann. of Math. (2)* **167** (2008), no. 2, 601–623.
- [20] Iosevich, Alex; Rudnev, Misha. Erdős distance problem in vector spaces over finite fields, *Trans. Amer. Math. Soc.* **359** (2007), no. 12, 6127–6142.
- [21] Janusz, Gerald J. *Algebraic number fields*. Second edition. Graduate Studies in Mathematics, 7. American Mathematical Society, Providence, RI, 1996.
- [22] Kahn, Jeff; Komlós, János; Szemerédi, Endre. On the probability that a random ± 1 -matrix is singular. *J. Amer. Math. Soc.* **8** (1995), no. 1, 223–240.
- [23] Katz, Nets Hawk; Shen, Chun-Yen. A slight improvement to Garaev’s sum product estimate, *Proc. Amer. Math. Soc.* **136** (2008), no. 7, 2499–2504.
- [24] Katz, Nets Hawk; Tardos, Gábor. A new entropy inequality for the Erdős distance problem. Towards a theory of geometric graphs, 119–126, *Contemp. Math.*, 342, *Amer. Math. Soc.*, Providence, RI, 2004.
- [25] Komlós, János. On the determinant of $(0, 1)$ matrices. *Studia Sci. Math. Hungar.* **2** (1967), 7–21.
- [26] Komlós, János. Circulated manuscript, 1977. edited version available online at: <http://www.math.rutgers.edu/~komlos/01short.pdf>
- [27] Martin, Greg; Wong, Erick B.. Almost all integer matrices have no integer eigenvalues. arXiv:0712.3060v1 [math.NT]. December 18, 2007.
- [28] Nathanson, M. B. On sums and products of integers. *Proc. Amer. Math. Soc.* **125** (1997) 9–16.
- [29] Razborov, Alexander A. A Product Theorem in Free Groups, *preprint*.
- [30] Rudelson, M.; Vershynin, R., The Littlewood-Offord problem and invertibility of random matrices, *Adv. Math.* **218** (2008), no. 2, 600–633.
- [31] Shafarevich, Igor R. *Basic algebraic geometry. 1. Varieties in projective space*. Second edition. Translated from the 1988 Russian edition and with notes by Miles Reid. Springer-Verlag, Berlin, 1994.

- [32] Slinko, Arkadii. A generalization of Komlós’s theorem on random matrices *New Zealand J. Math.* **30** (2001), no. 1, 81–86.
- [33] Solymosi, József. On the number of sums and products. *Bull. London Math. Soc.* **37** (2005), no. 4, 491–494.
- [34] Solymosi, Jozsef. On sum-sets and product-sets of complex numbers. (English, French summary) *J. Theor. Nombres Bordeaux* **17** (2005), no. 3, 921–924.
- [35] Stevenhagen, P.; Lenstra, H. W., Jr. Chebotarëv and his density theorem. *Math. Intelligencer* **18** (1996), no. 2, 26–37.
- [36] Suzuki, Michio. *Group theory I*. Translated from the Japanese by the author. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 247. Springer-Verlag, Berlin-New York, 1982.
- [37] Szemerédi, Endré; Trotter, William T., Jr. Extremal problems in discrete geometry. *Combinatorica* **3** (1983), no. 3-4, 381–392.
- [38] Tao, Terence; Vu, Van. On random ± 1 matrices: singularity and determinant. *Random Structures Algorithms* **28** (2006), no. 1, 1–23.
- [39] Tao, Terence; Vu, Van. On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.* **20** (2007), 603–628.
- [40] Tao, Terence; Vu, Van. *Additive Combinatorics*. Cambridge Studies in Advanced Mathematics (No. 105). Cambridge University Press, Cambridge, 2006.
- [41] Tao, Terence; Vu, Van. unpublished manuscript.
- [42] Tao, Terence; Vu, Van. Random matrices: the circular law, *Commun. Contemp. Math.* **10** (2008), no. 2, 261–307.
- [43] Tits, J. Free subgroups in linear groups. *J. Algebra* **20** (1972) 250–270.
- [44] Vu, Van; Wood, Melanie Matchett; Wood, Philip Matchett. Mapping Incidences. *submitted*.

Vita

Philip Joseph Matchett Wood

Education

- 2004–2009** **Rutgers University**, Mathematics Ph.D.
- 2003–2004** **University of Cambridge**, Diploma in Computer Science, first with distinction.
- 2002–2003** **University of Cambridge**, Certificate of Advanced Study in Mathematics (Part III of the Mathematics Tripos), with honors.
- 1998–2002** **Harvard University**, A.B., *magna cum laude* in mathematics.

Employment

- 2004-2009** **National Science Foundation Graduate Research Fellow**. Provides full tuition and stipend for up to three years of graduate study.
- 2007** **Lecturer for Calculus I**, Rutgers University. Lectured a class of 26 students, designed and graded exams, quizzes, and homework, which included assignments in WeBWork, an internet based interactive homework system.
- 2004–2007** **National Defense Science and Engineering Graduate Fellow**. Provides full tuition and stipend for three years of graduate study.
- 2002–2005** **Research Advisor**, University of Minnesota Duluth Research Experience for Undergraduates (REU). Supervised 36 undergraduate researchers in discrete mathematics over 4 summers.
- 2002–2004** **Herchel Smith Harvard Scholar**. Provides full tuition and stipend for two years of graduate study at the University of Cambridge.

Publications

- On the singularity probability of discrete random matrices**, with Jean Bourgain and Van Vu. preprint. 43 pages.
- On a conjecture of Alon**, with Linh Tran and Van Vu. to appear, *Journal of Number Theory*. 8 pages.
- The inverse Erdős-Heilbronn problem**, with Van Vu. submitted. 8 pages.

Mapping incidences, with Van Vu and Melanie Matchett Wood. submitted. arXiv:0711.4407v1 [math.CO]. 17 pages.

A translation method for finding combinatorial bijections, with Doron Zeilberger. to appear, *Annals of Combinatorics*. 19 pages.

Graph pegging numbers, with G. Helleloid, M. Khalid, and D. P. Moulton. *Discrete Mathematics*, In Press, available online 14 May 2008. arXiv:0804.1095v1 [math.CO]. 26 pages.

Bijjective proofs for Fibonacci identities related to Zeckendorf's Theorem. *Fibonacci Quart.* **45** (2007), no. 2, 138–145 (2008).

A bijective proof of $f_{n+4} + f_1 + 2f_2 + \cdots + nf_n = (n+1)f_{n+2} + 3$. *Integers* **6** (2006), A2, 4 pages.

Operations on well-covered graphs and the Roller-Coaster Conjecture. *Electron. J. Combin.* **11** (2004), no. 1, Research Paper 45, 17 pages. (Published under the name Philip Matchett.)