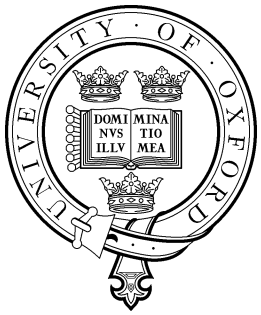
Topics in Analytic and Combinatorial Number Theory



Aled Walker

Magdalen College

University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Hilary Term 2018

Acknowledgements

I am deeply indebted to many people from many different parts of the world, for many things: not least for their assistance towards the preparation of this thesis.

My thanks go to Prof. Olivier Ramaré for giving a talk at the 'Additive Combinatorics in Bordeaux' conference in 2016, which stimulated the joint work that went on to form the second chapter of this thesis; to Dr. Adam Harper for suggesting a simplification to the original proof of Theorem 1.1.5; to Prof. Igor Shparlinski for notifying me of the work in [81]; to Dr. Alex Betts for useful discussions concerning lattices and heights; to Prof. Radhakrishnan Nair for making me aware of the central question of Chapter 3; to Prof. Christoph Aistleitner for comments on an early version of Chapter 3; to Dr. Sam Chow, for several useful email exchanges regarding diophantine inequalities; to Prof. Jan-Christoph Schlage-Puchta for making me aware of his work in [70]; to Dr. Thomas Bloom, Dr. Sam Chow, and Dr. Ayla Gafni, for collaborations on work related to Chapter 4; to Prof. Trevor Wooley for, indirectly, suggesting the topic of diophantine inequalities; to Sofia Lindqvist, Dr. Jakub Konieczny, Dr. Freddie Manners, Dr. Sean Eberhard, Dr. Rudi Mrazovic, James Aaronson, Joni Teräväinen, Luka Rimanić, Pierre-Yves Bienvenu, Prof. Kaisa Matomäki, Dr. Fernando Shao, Dr. Yufei Zhao, Prof. James Maynard, and Dr. Tom Sanders, as well as all those previously acknowledged, for many interesting mathematical discussions; and to the four anonymous journal referees who provided comments on the work that went on to form the core of the various chapters of this thesis.

Thanks to the Mathematical Sciences Research Institute in Berkeley, California, for accepting me as a Programme Associate during Spring 2017, during which time work towards Chapters 3, 4, and 5 was conducted in excellent conditions; to the Engineering and Physical Sciences Research Council for funding this research; to Prof. Kaisa Matomäki and Joni

Teräväinen for hosting and sponsoring my research visit to Finland in 2017; to the Oxford Mathematical Institute and Magdalen College Oxford, for their financial contributions towards my research travel costs throughout my DPhil studies; to Pembroke College Oxford, for being excellent employers throughout the final year of work towards this thesis; and additionally to the organisers of ELAZ 2016, BrOx, BOWL, Additive Combinatorics in Bordeaux, Pseudorandomness at the Simons Institute, MSRI Analytic Number Theory Programme, YRM, the LMS-CMI Bounded Gaps Between Primes school, the Clay Research Workshops on Analytic Number Theory and Harmonic Analysis, and the Regularity and Analytic Methods in Combinatorics workshop (all conferences and workshops that I attended as a D. Phil student).

I would like to especially thank my research supervisor Prof. Ben J. Green, for giving me interesting problems to think about, for our many clarifying discussions, and for letting me go my own way (whilst making sure that I never strayed too far).

Finally, and most importantly, this D. Phil could not have been completed without the endless love and support of my friends and my family.

Topics in Analytic and Combinatorial Number Theory

Aled Walker

Magdalen College
University of Oxford

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Hilary Term 2018

In this thesis we consider three different issues of analytic number theory. Firstly, we investigate how residues modulo q may be expressed as products of small primes. In Chapter 1, we work in the regime in which these primes are less than q, and present some partial results towards an open conjecture of Erdős. In Chapter 2, we consider the kinder regime in which these primes are at most q^C, for some constant C that is greater than 1. Here we reach an explicit version of Linnik's Theorem on the least prime in an arithmetic progression, saving that we replace 'prime' with 'product of exactly three primes'. The results of this chapter are joint with Prof. Olivier Ramaré.

The next two chapters concern equidistribution modulo 1, specifically the notion that an infinite set of integers is metric poissonian. This strong notion was introduced by Rudnick and Sarnak around twenty years ago, but more recently it has been linked with concepts from additive combinatorics. In Chapter 3 we study the primes in this context, and prove that the primes do not enjoy the metric poissonian property, a theorem which, in passing, improves upon a certain result of Bourgain. In Chapter 4 we continue the investigation further, adapting arguments of Schmidt to demonstrate that certain random sets of integers, which are nearly as dense as the primes, are metric poissonian after all.

The major work of this thesis concerns the study of diophantine inequalities. The use of techniques from Fourier analysis to count the number of solutions to such systems, in primes or in other arithmetic sets of interest, is well developed. Our innovation, following suggestions of Wooley and others, is to utilise the additive-combinatorial notion of Gowers norms. In Chapter 5 we adapt methods of Green and Tao to show that, even in an extremely general framework, Gowers norms control the number of solutions weighted by arbitrary bounded functions. We use this result to demonstrate cancellation of the Möbius function over certain irrational patterns.

Statement of Originality

I hereby certify that this thesis is entirely my own original work, except where otherwise indicated. The mathematical works of other authors, in any form, are properly acknowledged at their point of use.

Contents

0 Preliminaries **1**

0.1 Notation 1

0.2 Additive combinatorics 3

0.3 Gowers norms 7

0.4 Lipschitz functions 10

0.5 Probability 13

1 Writing residues as products of few primes **15**

1.1 Introduction 15

1.2 Sieve constructions 21

1.3 Subgroup obstructions 28

1.4 Proofs of main theorems 36

2 An explicit version of Linnik's Theorem **50**

2.1 Introduction 50

2.2 Lemmas 53

2.3 Proof of Theorem 2.1.3 60

3 The primes are not metric poissonian **65**

3.1 Introduction 65

3.2 Proof of Theorem 3.1.4 69

4 On the threshold for the metric poissonian property **75**

4.1 Introduction 75

4.2 Outline of proofs 77

4.3 Preparatory lemmas 83

4.4 The main argument 93

4.5 Proof of Theorem 4.1.3 98

4.6 Proof of Theorem 4.1.4 102

5 Gowers norms control diophantine inequalities **105**
5.1 Introduction 105
5.2 Historical background and the main theorem 113
5.3 Rank matrix and normal form 128
5.4 Upper bounds 135
5.5 Reductions 140
5.6 Transfer from $\mathbb{Z}$ to $\mathbb{R}$ 164
5.7 Degeneracy relations 170
5.8 A Generalised von Neumann Theorem 172
5.9 Constructions 184
5.10 Rank matrix and normal form: proofs 194
5.11 Additional linear algebra 200
5.12 The approximation function in the algebraic case 205

Bibliography **209**

Chapter 0

Preliminaries

Throughout this thesis we will need to refer to some standard results from fields other than analytic number theory. For ease of reference, we devote this preliminary section to recalling the precise statements of these results, as well as to fixing our notation.

0.1 Notation

We will use standard asymptotic notation O, o, and Ω. We do not, as is sometimes the convention, for a function f and a positive function g choose to write $f = O(g)$ if there exists a constant C such that $|f(N)| \leqslant Cg(N)$ *for N sufficiently large*. Rather we require the inequality to hold for all N in some pre-specified range. If N is a natural number, the range is always assumed to be $\mathbb{N}$ unless otherwise specified. (For us, $0 \notin \mathbb{N}$).

It will be a convenient shorthand to use these symbols in conjunction with minus signs. So, by convention, we determine that expressions such as $-O(1), -o(1), -\Omega(1)$ are negative, e.g. $N^{-\Omega(1)}$ refers to a term N^{-c}, where c is some positive quantity bounded away from 0 as the asymptotic parameter tends to infinity.

It will also be convenient to use the Vinogradov symbol $\ll$, where for a function f and a positive function g we write $f \ll g$ if and only if $f = O(g)$. We write $f \asymp g$ if $f \ll g$ and $g \ll f$. We also adopt the κ notation from [38]: $\kappa(x)$ denotes any quantity

that tends to zero as x tends to zero, with the exact value being permitted to change from line to line.

In Chapter 5, all the implied constants may depend on the dimensions of the underlying spaces. These will be obvious in context, and will always be denoted by m, d, h, or s (or, in the case of Proposition 5.3.8, by n). If an implied constant depends on other parameters, throughout the thesis we will denote these by subscripts, e.g. $O_{c,C,\varepsilon}(1)$, or $f \asymp_\varepsilon g$.

If N is a natural number, we use $[N]$ to denote $\{n \in \mathbb{N} : n \leqslant N\}$, whereas $[1, N]$ will be reserved for the closed real interval. For $x \in \mathbb{R}$, we write $[x] := \lfloor x + \frac{1}{2} \rfloor$ for the nearest integer to x, and $\|x\|$ for $|x - [x]|$. This means that there is slight overloading of the notation $[N]$, but the sense will always be obvious in context. When other norms are present, we may write $\|x\|_{\mathbb{R}/\mathbb{Z}}$ for $\|x\|$ to avoid confusion. For $\mathbf{x} \in \mathbb{R}^m$, we let $\|\mathbf{x}\|_{\mathbb{R}^m/\mathbb{Z}^m}$ denote $\sup_i |x_i - [x_i]|$.

We always assume that the vector space $\mathbb{R}^d$ is written with respect to the standard basis. If $X, Y \subset \mathbb{R}^d$ for some d, we define

$$\operatorname{dist}(X, Y) := \inf_{x \in X, y \in Y} \|x - y\|_\infty.$$

If X is the singleton $\{x\}$, we write $\operatorname{dist}(x, Y)$ for $\operatorname{dist}(\{x\}, Y)$. By identifying sets of m-by-d matrices with subsets of $\mathbb{R}^{md}$ simply by identifying the coefficients of the matrices with coordinates in $\mathbb{R}^{md}$, we may also define $\operatorname{dist}(X, Y)$ when X and Y are sets of matrices of the same dimensions.

We let $\partial(X)$ denote the topological boundary of X. If A and B are two sets with $A \subseteq B$, we let $1_A : B \longrightarrow \{0, 1\}$ denote the indicator function of A. (The relevant

set B will usually be obvious from context). The notation for logarithms, log, will always denote the natural log. For $\theta \in \mathbb{R}$, and e being Euler's constant, we also adopt the standard shorthand $e(\theta)$ to mean $e^{2\pi i\theta}$.

In section 5.8, if $\mathbf{x} \in \mathbb{R}^d$ and if a and b are two subscripts with $1 \leqslant a \leqslant b \leqslant d$, we use the notation $\mathbf{x_a^b}$ to denote the vector $(x_a, x_{a+1}, \cdots, x_b)^T \in \mathbb{R}^{b-a+1}$.

The letters μ, λ, φ, σ, χ, etc. will be too important to reserve solely for their number-theoretic meanings (Möbius function, Liouville function, etc.). If they undertake a special number-theoretic meaning, this will be explicitly stated.

Throughout the thesis, and most pertinently in chapter 5, we will consider a linear map $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ to be synonymous with the m-by-d matrix that represents L with respect to the standard bases. The norm $\|L\|_\infty$ will refer to the maximum absolute value of the coefficients of this matrix.

0.2 Additive combinatorics

We begin with some standard definitions.

Definition 0.2.1 (Additive convolution)**.** *Let $f, g : \mathbb{R} \longrightarrow \mathbb{C}$ be two bounded measurable functions with compact support. We define their additive convolution $f * g : \mathbb{R} \longrightarrow \mathbb{C}$ to be the function*

$$(f * g)(x) := \int_{y \in \mathbb{R}} f(x-y)\overline{g(y)}\, dy.$$

In Chapter 1 it will be natural to represent finite abelian groups multiplicatively, and so we record the definition of convolution in this context.

Definition 0.2.2 (Multiplicative convolution)**.** *Let G be a finite abelian group, written multiplicatively. For two functions $f, g : G \longrightarrow \mathbb{C}$, we define their multiplicative convolution $f \star g : G \longrightarrow \mathbb{C}$ to be the function*

$$(f \star g)(x) := \sum_{y \in G} f(xy^{-1})\overline{g(y)}.$$

Of course all this theory may be placed in a common abstract framework of locally compact abelian groups (see [72]), but there is no reason to do so in this thesis.

Definition 0.2.3 (Product set)**.** *Let G be an abelian group, written multiplicatively, and let A be a subset of G. Let k be a natural number. We then define the* k-fold iterated product-set A^k *by*

$$A^k := \{a_1 a_2 \cdots a_k : a_i \in A\}.$$

If B is another subset of G, we also use the notation $A \cdot B$ for the product set

$$\{ab : a \in A, b \in B\}.$$

Definition 0.2.4 (Inverse set)**.** *Let G be an abelian group, written multiplicatively, and let A be a subset of G. We write*

$$A^{-1} := \{a^{-1} : a \in A\}.$$

Definition 0.2.5 (Additive energy)**.** *Let G be an abelian group, written multiplicatively, and let A be a finite subset of G. Define the* additive energy $E(A)$ *to be*

$$|\{(a_1, a_2, a_3, a_4) \in A \times A \times A \times A : a_1 a_2 = a_3 a_4\}|.$$

In Chapter 3 we will need to consider the additive energy of a subset of $\mathbb{Z}$, a group

that is written additively. The corresponding definition of additive energy is as in Definition 0.2.5, *mutatis mutandis*.

We recall the following easy Cauchy-Schwarz argument.

Lemma 0.2.6. *Let G be a finite abelian group, written multiplicatively, and let A be a subset of G. Then $|A \cdot A| \geqslant |A|^4/E(A)$.*

Proof. We have the immediate identities $|A|^2 = \sum_{x \in G}(1_A \star 1_A)(x)$, and $E(A) = \sum_{x \in G}(1_A \star 1_A)(x)^2$. Note that the function $(1_A \star 1_A)$ is only supported on the product set $A \cdot A$. Then by the Cauchy-Schwarz inequality we deduce that $|A|^2 \leqslant |A \cdot A|^{\frac{1}{2}} E(A)^{\frac{1}{2}}$. This rearranges to give the lemma. □

We will use three standard results concerning $|A \cdot A|$, when A is a subset of an abelian group. First, recall the following simple combinatorial lemma of Freiman, which renders precise the notion that being contained in a coset is the only obstruction to a set having reasonable doubling.

Lemma 0.2.7 (Freiman)**.** *Let G be an abelian group, written multiplicatively, and let A be a finite subset of G. Suppose that A is not contained in any proper coset of G. Then either $A \cdot A^{-1} = G$ or $|A \cdot A| \geqslant \frac{3}{2}|A|$.*

A reference for the proof of this lemma is somewhat difficult to locate. It originally appeared in [24], in Russian, and the statement appears as Proposition 1.1 of [76], and as Exercise 2.6.5 of [85]. Fortunately, the brevity of the argument allows us to give it here in full.

Proof. Suppose that $|A \cdot A| < \frac{3}{2}|A|$; we show that $A \cdot A^{-1}$ is closed under multiplication. Indeed, let $w, x, y, z \in A$. The set $\{a \in A : wa \in zA\}$ has size greater than $\frac{1}{2}|A|$, since $|A \cdot A| < \frac{3}{2}|A|$. Similarly the set $\{a \in A : xa \in yA\}|$ has size greater than $\frac{1}{2}|A|$. Therefore these two sets intersect, and we have $a, a_z, a_y \in A$ such that $wa = za_z$ and

$xa = ya_y$. Hence

$$(wx^{-1})(yz^{-1}) = waa^{-1}x^{-1}yz^{-1} = za_z a_y^{-1} y^{-1} yz^{-1} = a_z a_y^{-1}.$$

Therefore $A \cdot A^{-1}$ is a subgroup of G (the other axioms are trivial), and as A is not contained in any proper coset of G we conclude that $A \cdot A^{-1}$ must be the whole of G. □

This lemma is sharp: consider the case where A is the union of a subgroup and a single non-trivial coset of the same subgroup.

We will also need the Ruzsa triangle inequality.

Lemma 0.2.8 (Ruzsa triangle inequality)**.** *Let G be an abelian group, written multiplicatively, and let A, B, and C be finite subsets of G. Then*

$$|A \cdot C^{-1}| \leqslant \frac{|A \cdot B^{-1}||B \cdot C^{-1}|}{|B|}.$$

This is proved in Lemma 2.6 of [85].

Corollary 0.2.9. *Let G be an abelian group, written multiplicatively, and let A be a finite subset of G. Then*

$$|A \cdot A^{-1}| \leqslant \frac{|A \cdot A|^2}{|A|}.$$

This corollary is proved by putting $A = C$ and $B = A^{-1}$ in Lemma 0.2.8.

The final result concerning the size of product sets will be Kneser's Theorem.

Theorem 0.2.10 (Kneser's Theorem)**.** *Let G be a finite abelian group, written multiplicatively, and let A and B be subsets of G. Let H be the stabiliser of $A \cdot B$, i.e. the subgroup of G consisting of those elements h such that $h \cdot A \cdot B = A \cdot B$. Then*

we have

$$|A \cdot B| \geqslant |A \cdot H| + |B \cdot H| - |H|.$$

Proof. Originally from [51], but more easily read in [85, Theorem 5.5] or [63, Theorem 4.3]. □

0.3 Gowers norms

There are several existing accounts of the basic theory of Gowers norms – for example in [35] and [84] – and the reader looking for an introduction to the theory in its full generality should certainly consult these references, as well as Appendices B and C of [38]. However, in the interests of making this thesis as self-contained as possible, we use this section to pick out the central definitions and notions that will be used in the main text.

Definition 0.3.1. *Let N be a natural number. For a function $f : \mathbb{Z}/N\mathbb{Z} \longrightarrow \mathbb{C}$, and a natural number d, define the Gowers U^d norm $\|f\|_{U^d(N)}$ to be the unique non-negative solution to*

$$\|f\|_{U^d(N)}^{2^d} = \frac{1}{N^{d+1}} \sum_{x,h_1,\cdots,h_d} \prod_{\boldsymbol{\omega} \in \{0,1\}^d} \mathcal{C}^{|\boldsymbol{\omega}|} f(x + \mathbf{h} \cdot \boldsymbol{\omega}), \tag{1}$$

where $|\boldsymbol{\omega}| = \sum_i \omega_i$, $\mathbf{h} = (h_1, \cdots, h_d)$, $\mathcal{C}$ is the complex-conjugation operator, and the summation is over $x, h_1, \cdots, h_d \in \mathbb{Z}/N\mathbb{Z}$.

For example,

$$\|f\|_{U^1(N)} = \Big|\frac{1}{N} \sum_x f(x)\Big|,$$

and

$$\|f\|_{U^2(N)} = \left(\frac{1}{N^3} \sum_{x,h_1,h_2} f(x)\overline{f(x+h_1)f(x+h_2)}f(x+h_1+h_2)\right)^{\frac{1}{4}}.$$

It is not immediately obvious that the right-hand side of (1) is always a non-negative real, nor why the U^d norms are genuine norms if $d \geqslant 2$: proofs of both these facts

may be found in [85].

The functions in the main text do not have a cyclic group as a domain but rather the interval $[N]$, but the theory may easily be adapted to this case.

Definition 0.3.2. *Let N, N' be natural numbers, with $N' \geqslant N$. Identify $[N]$ with a subset of $\mathbb{Z}/N'\mathbb{Z}$ in the natural way, i.e. $[N] = \{1, \cdots, N\} \subseteq \{1, \cdots, N'\}$, which we then view as $\mathbb{Z}/N'\mathbb{Z}$. For a function $f : [N] \longrightarrow \mathbb{C}$, and a natural number d, we define the Gowers norm $\|f\|_{U^d[N]}$ to be the unique non-negative real solution to the equation*

$$\|f\|_{U^d[N]}^{2^d} = \frac{1}{|R|} \sum_{x,h_1,\cdots,h_d} \prod_{\boldsymbol{\omega} \in \{0,1\}^d} \mathcal{C}^{|\boldsymbol{\omega}|} f1_{[N]}(x + \mathbf{h} \cdot \boldsymbol{\omega}), \tag{2}$$

where $f1_{[N]}$ is the extension by zero of f to $\mathbb{Z}/N'\mathbb{Z}$, the summation is over $x, h_1, \cdots, h_d \in \mathbb{Z}/N'\mathbb{Z}$, and the set R is the set

$$R := \{x, h_1, \cdots, h_d \in \mathbb{Z}/N'\mathbb{Z} : \text{ for every } \boldsymbol{\omega} \in \{0,1\}^d, x + \mathbf{h} \cdot \boldsymbol{\omega} \in [N]\}.$$

One can immediately see that this definition is equivalent to

$$\|f\|_{U^d[N]} = \|f1_{[N]}\|_{U^d(N')} / \|1_{[N]}\|_{U^d(N')},$$

and is also independent of the choice of N' as long as N'/N is large enough (in terms of d). Taking $N' = O(N)$ we have $\|1_{[N]}\|_{U^d(N')} \asymp 1$, and thus $\|f\|_{U^d[N]} \asymp \|f1_{[N]}\|_{U^d(N')}$. (See [38, Lemma B.5] for more detail on this).

We observe that there is only a contribution to the summand in equation (2) when $x \in [N]$ and for every i we have $h_i \in \{-N, -N+1, \cdots, N-1, N\}$ modulo N'. Further, it may be easily seen that $|R| \asymp N^{d+1}$. Therefore, choosing N'/N sufficiently

large, we conclude that

$$\|f\|_{U^d[N]} \asymp \left(\frac{1}{N^{d+1}} \sum_{x,h_1,\cdots,h_d \in \mathbb{Z}} \prod_{\boldsymbol{\omega} \in \{0,1\}^d} \mathcal{C}^{|\boldsymbol{\omega}|} f(x + \mathbf{h} \cdot \boldsymbol{\omega}) \right)^{\frac{1}{2^d}}. \tag{3}$$

The relation (3) is implicitly assumed throughout the main text.

In order to succinctly state Theorem 5.8.1 in Chapter 5, we will have to refer to a Gowers norm $U^d(\mathbb{R})$, which has been used in some recent work on linear patterns in subsets of Euclidean space (see [15, Lemma 4.2], [19, Proposition 3.3]). This Gowers norm is a less well-studied object, as the theory was originally developed over finite groups. Nevertheless it may be perfectly well defined, and even deep aspects of its inverse theory may be deduced from the corresponding theory of the discrete Gowers norm (see [83]).

Definition 0.3.3. *Let $f : [0,1] \longrightarrow \mathbb{C}$ be a bounded measurable function, and let d be a natural number. Then we define the Gowers norm $\|f\|_{U^d(\mathbb{R})}$ to be the unique non-negative real satisfying*

$$\|f\|_{U^d(\mathbb{R})}^{2^d} = \int_{(x,\mathbf{h}) \in \mathbb{R}^{d+1}} \prod_{\boldsymbol{\omega} \in \{0,1\}^d} \mathcal{C}^{|\boldsymbol{\omega}|} f(x + \sum_{i=1}^{d} h_i \omega_i) \, dx \, dh_1 \cdots dh_d \tag{4}$$

where $|\boldsymbol{\omega}| = \sum_i \omega_i$, and $\mathcal{C}$ is the complex-conjugation operator.

Let N be a positive real, and let $g : [-N, N] \longrightarrow \mathbb{C}$ be a measurable function. Define the function $f : [0,1] \longrightarrow \mathbb{C}$ by $f(x) := g(2Nx - N)$, and then set

$$\|g\|_{U^d(\mathbb{R})} := \|f\|_{U^d(\mathbb{R})}.$$

Explicitly, a change of variables shows that

$$\|g\|_{U^d(\mathbb{R})}^{2^d} \asymp \frac{1}{N^{d+1}} \int\limits_{(x,\mathbf{h})\in\mathbb{R}^{d+1}} \prod_{\boldsymbol{\omega}\in\{0,1\}^d} \mathcal{C}^{|\boldsymbol{\omega}|} g(x+\sum_{i=1}^{d} h_i\omega_i)\, dx\, dh_1 \cdots dh_d. \tag{5}$$

Expression (5) will be used throughout Chapter 5.

We require one further fact about Gowers norms.

Proposition 0.3.4 (Gowers-Cauchy-Schwarz inequality)**.** *Let d be a natural number, and, for each $\boldsymbol{\omega} \in \{0,1\}^d$, let $f_{\boldsymbol{\omega}} : [0,1] \longrightarrow \mathbb{C}$ be a bounded measurable function. Define the Gowers inner-product*

$$\langle (f_{\boldsymbol{\omega}})_{\boldsymbol{\omega}\in\{0,1\}^d} \rangle := \int\limits_{(x,\mathbf{h})\in\mathbb{R}^{d+1}} \prod_{\boldsymbol{\omega}\in\{0,1\}^d} \mathcal{C}^{|\boldsymbol{\omega}|} f_{\boldsymbol{\omega}}(x+\sum_{i=1}^{d} h_i\omega_i)\, dx\, dh_1 \cdots dh_d.$$

Then

$$|\langle (f_{\boldsymbol{\omega}})_{\boldsymbol{\omega}\in\{0,1\}^d} \rangle| \leqslant \prod_{\boldsymbol{\omega}\in\{0,1\}^d} \|f_{\boldsymbol{\omega}}\|_{U^d(\mathbb{R})}.$$

Proof. See [85, Chapter 11] for the proof in the finite group setting. The modification to the setting of the reals is trivial. □

0.4 Lipschitz functions

In Chapter 5 we will use properties of Lipschitz functions.

Definition 0.4.1 (Lipschitz functions)**.** *We say that a function $F : \mathbb{R}^m \longrightarrow \mathbb{C}$ is Lipschitz, with Lipschitz constant at most M, if*

$$M \geqslant \sup_{\substack{\mathbf{x},\mathbf{y}\in\mathbb{R}^m \\ \mathbf{x}\neq\mathbf{y}}} \frac{|F(\mathbf{x})-F(\mathbf{y})|}{\|\mathbf{x}-\mathbf{y}\|_\infty}.$$

We say that a function $G : \mathbb{R}^m/\mathbb{Z}^m \longrightarrow \mathbb{C}$ is Lipschitz, with Lipschitz constant at most M, if

$$M \geqslant \sup_{\substack{\mathbf{x},\mathbf{y}\in\mathbb{R}^m/\mathbb{Z}^m \\ \mathbf{x}\neq\mathbf{y}}} \frac{|F(\mathbf{x}) - F(\mathbf{y})|}{\|\mathbf{x}-\mathbf{y}\|_{\mathbb{R}^m/\mathbb{Z}^m}}.$$

We record the three properties of Lipschitz functions that we will require.

Lemma 0.4.2. *Let N be a positive real, let m be a natural number, let K be a convex subset of $[-N,N]^m$, and let σ be some parameter in the range $0 < \sigma < 1/2$. Then there exist Lipschitz functions $F_\sigma, G_\sigma : \mathbb{R}^m \longrightarrow [0,1]$ supported on $[-2N,2N]^m$, both with Lipschitz constant at most $O(\frac{1}{\sigma N})$, such that*

$$1_K = F_\sigma + O(G_\sigma)$$

and $\int_{\mathbf{x}} G_\sigma(\mathbf{x})\, d\mathbf{x} = O(\sigma N^m)$. Furthermore, $F_\sigma(\mathbf{x}) \geqslant 1_K(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^m$, and G is supported on $\{\mathbf{x} \in \mathbb{R}^m : \mathrm{dist}(\mathbf{x}, \partial(K)) \leqslant \sigma N\}$.

This is [38, Corollary A.3]. It will be used in Lemmas 5.5.9 and 5.5.11 to replace sums with sharp cut-offs by sums with Lipschitz cut-offs.

Lemma 0.4.3. *Let X be a positive real, with $X > 2$. Let $F : \mathbb{R}^m/\mathbb{Z}^m \longrightarrow \mathbb{C}$ be a Lipschitz function such that $\|F\|_\infty \leqslant 1$ and the Lipschitz constant of F is at most M. Then*

$$F(\mathbf{x}) = \sum_{\substack{\mathbf{k}\in\mathbb{Z}^m \\ \|\mathbf{k}\|_\infty \leqslant X}} c_X(\mathbf{k}) e(\mathbf{k}\cdot\mathbf{x}) + O_m\left(M\frac{\log X}{X}\right) \tag{6}$$

for every $\mathbf{x} \in \mathbb{R}^m/\mathbb{Z}^m$, for some function $c_X(\mathbf{k})$ satisfying $\|c_X(\mathbf{k})\|_\infty \ll 1$.

This is [37, Lemma A.9], and will be used in Lemma 5.4.3 as a way of bounding the number of solutions to a certain inequality.

Lemma 0.4.4. *Let X, N, C be positive reals, with $X > 2$ and $N > 1$. Let $F : \mathbb{R}^m \longrightarrow \mathbb{C}$ be a Lipschitz function, supported on $[-CN, CN]^m$, such that $\|F\|_\infty \leqslant 1$ and the*

Lipschitz constant of F is at most M. Then

$$F(\mathbf{x}) = \int_{\substack{\boldsymbol{\xi}\in\mathbb{R}^m \\ \|\boldsymbol{\xi}\|_\infty \leqslant X}} c_X(\boldsymbol{\xi}) e(\frac{\boldsymbol{\xi}\cdot\mathbf{x}}{N})\, d\boldsymbol{\xi} + O_{m,C}\left(MN\frac{\log X}{X}\right) \tag{7}$$

for every $\mathbf{x}\in\mathbb{R}^m$, for some function $c_X(\boldsymbol{\xi})$ satisfying $\|c_X(\boldsymbol{\xi})\|_\infty \ll_{m,C} 1$.

Lemma 0.4.4 is very similar to Lemma 0.4.3, and may be easily proved by adapting that standard harmonic analysis argument found in [37, Lemma A.9] from $\mathbb{R}^m/\mathbb{Z}^m$ to $\mathbb{R}^m$. For completeness, we sketch the proof.

Sketch proof. By rescaling the variable $\mathbf{x}$ by a factor of N, we reduce to the case where F is supported on $[-C,C]^m$ and has Lipschitz constant at most MN.

Let

$$K_X(\mathbf{x}) := \prod_{i=1}^m \frac{1}{X}\left(\frac{\sin(\pi X x_i)}{\pi x_i}\right)^2.$$

Then

$$\widehat{K_X}(\boldsymbol{\xi}) = \prod_{i=1}^m \max(1-\frac{|\xi_i|}{X}, 0).$$

We have

$$(F * K_X)(\mathbf{x}) = \int_{\substack{\boldsymbol{\xi}\in\mathbb{R}^m \\ \|\boldsymbol{\xi}\|_\infty \leqslant X}} \widehat{F}(\boldsymbol{\xi})\widehat{K_X}(\boldsymbol{\xi}) e(\boldsymbol{\xi}\cdot\mathbf{x})\, d\mathbf{x},$$

and, since $|\widehat{F}(\boldsymbol{\xi})| \leqslant \|F\|_1 \ll_C 1$, letting $c_X(\boldsymbol{\xi}) := \widehat{F}(\boldsymbol{\xi})\widehat{K_X}(\boldsymbol{\xi})$ gives a main term of the desired form.

It remains to show that

$$\|F - F * K_X\|_\infty \ll_{m,C} MN\frac{\log X}{X}.$$

By writing

$$|F(\mathbf{x}) - (F * K_X)(\mathbf{x})| = \Big| \int\limits_{\mathbf{y} \in \mathbb{R}^m} (F(\mathbf{x}) - F(\mathbf{y})) K_X(\mathbf{x} - \mathbf{y})\, d\mathbf{y} \Big|,$$

one sees that it suffices to show that

$$\int\limits_{\|\mathbf{z}\|_\infty \leqslant 2C} \|\mathbf{z}\|_\infty K_X(\mathbf{z})\, d\mathbf{z} \ll_{m,C} \frac{\log X}{X}.$$

But this bound follows immediately from a dyadic decomposition. □

We will use Lemma 0.4.4 extensively in the Generalised von Neumann Theorem argument in section 5.8.

0.5 Probability

In Chapter 4 we will need to use a few of the standard tools of probability theory.

Lemma 0.5.1 (The first Borel-Cantelli Lemma)**.** *Let Ω be a probability space, and let $E_1, E_2, \cdots$ be a sequence of events in Ω. Suppose that $\sum_{i=1}^{\infty} \mathbb{P}(E_i) < \infty$. Then*

$$\mathbb{P}(\bigcap_{n=1}^{\infty} \bigcup_{i \geqslant n} E_i) = 0.$$

This is an extremely standard result. See Lemma 1.2 of [85], say.

Lemma 0.5.2 (Chebyshev's inequality)**.** *Let X be a random variable with expectation μ and variance σ^2. Then, for any positive λ,*

$$\mathbb{P}(|X - \mu| \geqslant \lambda) \leqslant \frac{\sigma^2}{\lambda^2}.$$

Again, this result is extremely standard. See Theorem 4.1.1 of [4].

Theorem 0.5.3 (Large deviation bound)**.** *Let Y be the sum of mutually independent indicator random variables, and let μ be the expectation of Y. Then*

$$\mathbb{P}(|Y - \mu| > \varepsilon\mu) < 2\exp(-c_\varepsilon\mu),$$

where c_ε is a positive constant that depends only on ε.

This is Corollary 1.14 of [4], and is a useful version of a Chernoff-type large deviation bound.

Chapter 1

Writing residues as products of few primes

1.1 Introduction

Let q be a large prime. In [22], Erdős made the following conjecture:

Conjecture 1.1.1 (Erdős, 1987)**.** *Every co-prime residue a modulo q may be expressed as the product of exactly two primes, each less than q.*

One may consider Conjecture 1.1.1 as a multiplicative analogy of the famous Goldbach conjecture on sums of two prime numbers, and it is still open, even assuming the Generalised Riemann Hypothesis. The authors of [22] remark that the conjecture, "is almost certainly true, but will be hopelessly difficult to prove," and certainly we do not prove it in this thesis. Rather, the purpose of this chapter is to present our various unconditional partial results. The main theorems may be found in our paper [91], but here we have sharpened several of the bounds and thresholds.

We will be interested in establishing when certain naturally defined subsets of $(\mathbb{Z}/q\mathbb{Z})^\times$ generate that entire group. More precisely, if $S \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$, we will be

interested in when $S^k = (\mathbb{Z}/q\mathbb{Z})^\times$ for some k, where S^k denotes the k-fold iterated product-set (see Definition 0.2.3). Reserving q to denote the fixed large prime introduced above, and with p denoting an arbitrary prime, for a quantity $\eta \leqslant 1$ we define

$$P_\eta := \{p : p < \eta q\}$$

In this language, Erdős conjectured that $P_1^2 = (\mathbb{Z}/q\mathbb{Z})^\times$, if q is large enough. The problem of showing that P_η^k is a large set is obviously the more difficult the smaller the value of η.

When S is a certain kind of subset of $\mathbb{Z}/q\mathbb{Z}$, most notably an interval, the case S^2 has been extensively studied, and is known as the *modular hyperbola* problem (see the survey of Shparlinski [80]). The case of primes has received less attention. The original paper of Erdős, Odlyzko, and Sárközy [22] shows that, under the Generalised Riemann Hypothesis, there are at most $c \log^5 q$ residues a less than q that may not be expressed as the product of two primes less than q (where c is a small absolute positive constant).

Unsurprisingly, the large sieve may be used to recover results of a similar strength, unconditionally, for almost all q. Indeed, consider the follow theorem, which is (part of) Theorem 1 of Friedlander, Kurlberg and Shparlinski [27].

Theorem 1.1.2. *Let $x, M \geqslant 2$. Then*

$$R_\pi(x, M) \ll (M^{-1}x^4 + Mx^2)(\log x)^{-2},$$

where

$$R_\pi(x, M) := \sum_{\substack{M < q \leqslant 2M \\ q \text{ prime}}} \sum_{a=1}^{q-1} \Big(\sum_{\substack{p_1 \leqslant x \\ p_2 \leqslant x \\ p_1 p_2 \equiv a \pmod q}} 1 - \frac{\pi(x)^2}{q-1} \Big)^2.$$

By putting M equal to x, one immediately concludes the following: if M is a large

real parameter, then for all primes q in the dyadic range $M < q \leqslant 2M$ there are at most $\log^4 q$ residues a less than q that may not be expressed as the product of two primes less than q, with the exception of a set of primes q of size $o(M/\log M)$. By putting $M = x^{2-\varepsilon}$, one may deduce the following, which concerns a much smaller value of η: if M is a large real parameter, then for all primes q in the dyadic range $M < q \leqslant 2M$ one has $P_\eta^3 = (\mathbb{Z}/q\mathbb{Z})^\times$, with $\eta = q^{-\frac{1}{2}+\varepsilon}$, apart from an exceptional set of primes q of size $o_\varepsilon(M/\log M)$. Corollary 1 of [48] by Heath-Brown and Li is a similar result about this range of parameters. We do not dwell on these results, as the case of almost all q seems to be a rather different issue to the main work of this chapter.

The most relevant previous work for fixed q, which only came to our attention after the publication of our first preprint [90] on this topic, was undertaken by Shparlinski in [81].

Theorem 1.1.3. [81, Corollary of Theorem 3] *Let m be a large enough natural number, and let r be a residue modulo m satisfying $(r, m) = 1$. Then we may find p and s such that*

$$r \equiv ps \pmod m, \tag{1.1}$$

where $p, s \leqslant m^{0.997}$, p is prime, and s is the product of at most 17 primes.

Applying this theorem to our setting as a black box, it immediately implies the following:

Corollary 1.1.4. *Let q be a large enough prime. Then setting $\eta := q^{-0.003}$, we have*

$$\bigcup_{i=1}^{18} P_\eta^i = (\mathbb{Z}/q\mathbb{Z})^\times.$$

Shparlinski's approach proceeds by employing a bound of Garaev [29] on exponential sums over reciprocals of primes, generalised to composite moduli m. In our work,

we aim to prove results stronger than Corollary 1.1.4, by more direct methods. In this endeavour we are successful, though of course we do not derive any improvement on Theorem 1.1.3 itself, which concerns a different sieving situation.

In response to our publication [91] of the theorems described in this chapter, Shparlinski has sharpened his method, and published a preprint [79] improving on many[1] of these results. Thus, regrettably, not all the work in this chapter represents the current state-of-the-art. Yet, having been an important part of the efforts of the community at large, we believe the work still merits inclusion in this thesis.

We now state the main results of this chapter.

Theorem 1.1.5. *Let q be prime, and let $\varepsilon > 0$. Then:*

(1) $|P_1^2| \geqslant \frac{1}{64}(1 - o(1))q$;

(2) if $\eta = q^{-\frac{1}{2}+\varepsilon}$, *then* $|P_\eta^2| \gg_\varepsilon q$.

To describe in brief, the proof of part (2) of the above theorem is a mixture of two elements: an application of an asymptotic expression for the fourth moment of a multiplicative character sum, given by Ayyad, Cochrane, and Zheng in [6], combined with an introduction of sieve weights. The manoeuvre of switching to a fourth moment means that the implied constant is necessarily rather small. To prove the tighter result for $|P_1^2|$ given in part (1), we employ a different method, using sieve weights to upper-bound the number of solutions to the equation $p_1 p_2 \equiv a \pmod q$, for a fixed a, directly, from which we conclude that the support of P_1^2 cannot be too small. The key feature here, which we believe to be relatively novel, is that while we lose information by switching to sieve weights we also gain information by accessing the stronger L^1 bounds of the sieve weights' Fourier transform. In [81, 79], this positive density question is not considered.

[1] Not all the results have been improved. We will carefully note which of our bounds remain the best known.

Theorem 1.1.6. *Let $\varepsilon > 0$. There then exists a natural number $q_0(\varepsilon)$ such that, if q is a prime number satisfying $q \geqslant q_0(\varepsilon)$, then the following holds:*

(1) every non-zero residue modulo q can be expressed as the product of at most 6 primes, each less than $q^{\frac{15}{16}+\varepsilon}$;

(2) there exist a natural number $k(\varepsilon)$ such that every non-zero residue modulo q can be expressed as the product of at most $k(\varepsilon)$ primes, each less than $q^{\frac{1}{4}+\varepsilon}$.

Note that $\frac{15}{16} = 0.9375$, so part (1) of the above theorem represents an improvement over Corollary 1.1.4. However, several subsequent results of Shparlinski in [79] improve further over part (1). In particular he proves that every non-zero residue modulo q may be expressed as the product of at most 5 primes less than $q^{0.905}$.

Part (2) is a substantial improvement over the analogous result we proved previously in [91], in which the primes were at most $q^{\frac{3}{4}+\varepsilon}$. It is also stronger than the analogous result from [79], in which the primes are at most $q^{\frac{1}{2}+\varepsilon}$. The method is adapted from work of Harman and Shparlinksi [45], though we have not been able to mimic their method for reducing the threshold[2] further to $q^{\frac{1}{4\sqrt{e}}+\varepsilon}$, which they achieve in the case of generation of $(\mathbb{Z}/q\mathbb{Z})^\times$ by small integers.

Finally, we deduce that every residue may be expressed as the product of a fixed number of small primes.

Theorem 1.1.7. *Let $\varepsilon > 0$. Then there exists a natural number $q_0(\varepsilon)$ such that, if q is a prime number satisfying $q \geqslant q_0(\varepsilon)$, the following holds:*

(1) if $\eta = q^{-\frac{1}{16}+\varepsilon}$, then $P_\eta^{20} = (\mathbb{Z}/q\mathbb{Z})^\times$;

(2) if $\eta = q^{-\frac{3}{4}+\varepsilon}$, then there exists a natural number $K(\varepsilon)$ such that $P_\eta^{K(\varepsilon)} = (\mathbb{Z}/q\mathbb{Z})^\times$.

[2]Shparlinski (personal communication) believes that an estimate of Norton might be able to achieve this, but the details are currently unclear.

Part (2) of this theorem, as well as part (2) of the previous theorem, may be viewed as multiplicative analogies of the classical Schnirelmann's theorem that every sufficiently large integer is the sum of at most 37000 primes, proved in [78] (see exposition in [64]). There is an analogy to be made too between the methods of proof: we use Theorem 1.1.5 to establish a positive density result, and then an argument from additive combinatorics to show that this dense set expands.

To facilitate this argument, it will be necessary to preclude the following phenomenon. Consider $q = 5$: we see $P_1 = \{2, 3\}$ consists entirely of quadratic non-residues, and $P_1^2 = \{1, 4\}$, $P_1^3 = \{2, 3\}$, $P_1^4 = \{1, 4\}$ etcetera, and so Theorem 1.1.7 fails to hold. The obstruction arises as P_1 is entirely contained within a coset of a non-trivial subgroup of $(\mathbb{Z}/q\mathbb{Z})^\times$, or equivalently has a non-trivial Fourier coefficient of maximal value. In Lemma 1.3.1 we establish that, for large enough q, the primes less than $q^{\frac{1}{4}+\varepsilon}$ cannot be trapped in such a coset. On the suggestion of Schlage-Puchta we establish this using his version of the large sieve for sequences supported on primes, from [70], which gives the best quantitative results we know of. We also present a very different proof based on the Selberg-Delange method. This was our original approach, and is rather shorter.[3] However, it yields slightly poorer constants in Theorem 1.1.7 than the large sieve method.

Remark 1.1.8. For simplicity we restrict to q prime throughout this chapter. The methods we use to prove Theorems 1.1.5 and 1.1.6 admit a modification to q composite, although the technical details become increasingly complicated. However, it is far from clear to me that the Selberg-Delange argument we use in connection with Lemma 1.3.1 admits such a modification. Shparlinki's results in [81, 79] are framed for q composite, but, in the original motivating paper [22], Conjecture 1.1.1 is only framed for q prime.

[3] We initially thought the approach via the Selberg-Delange method was a little idiosyncratic, yet it was subsequently rediscovered independently by Pollack in the first version of the preprint [68], later published without this result in [69].

Remark 1.1.9. As will become evident in the coming sections, a great many different methods may be applied to attack the problems in this chapter. One can use additive Fourier transforms, multiplicative Fourier transforms, replace expressions by higher moments that are easier to estimate, and all possible combinations of the above. As such it is very difficult to establish whether one has proved an optimal result, or whether some more intricate combination of these tools could improve the various constants. At some point one must balance the length and readability of the argument against the quality of the results that are proved. We try to strike such a balance here.

1.2 Sieve constructions

The proofs of Theorems 1.1.5 and 1.1.6 will be applications of the existence of certain sieve weights. In this section we collect together the precise results we require, and discuss suitable references. For reasons discussed below, we opt for 'off-the-shelf' estimates, rather than constructing bespoke weights for this problem.

Lemma 1.2.1 (Upper-bound sieve)**.** *Let $\gamma > 0$, let ξ be a fixed real number satisfying $0 < \xi < \frac{1}{2} - \frac{\gamma}{2}$, and let x be a large integer, i.e. let x satisfy $x \geqslant x_0(\gamma)$ for some $x_0(\gamma)$. Define $z := x^{\xi}$ and $D := x^{2\xi}$. Let $\nu(d)$ denote the number of distinct prime factors of d. Then there exists a weight function $w^+ : [x] \longrightarrow \mathbb{R}_{\geqslant 0}$ such that:*

(i) if n has no prime factors less than z then $w^+(n) \geqslant 1$;

(ii) $\sum_{n=1}^{x} w^+(n) \leqslant (1 + o(1))\frac{x}{\xi \log x}$*; and*

(iii) $w^+(n) = \sum_{d|n} \lambda_d^+$*, where $(\lambda_d^+)_{d \geqslant 1}$ is a sequence of real numbers satisfying $\lambda_d^+ = 0$ for $d > D$, $\lambda_d^+ = 0$ if d not square-free, and $|\lambda_d^+| \leqslant 3^{\nu(d)}$ for all d.*

Proof. The standard Selberg sieve weights suffice. A comprehensive reference is [26], though the relevant estimates are a little difficult to locate. Specifically, one should

take λ_d^+ to be equal to the weight λ_d constructed on page 92 between expressions (7.25) and (7.26), where one has taken $g(p) \equiv \frac{1}{p}$ for all primes. Then, considering expressions (7.4) and (7.5) of that volume, along with the bound (7.27), if we define $w^+(n)$ to be $\sum_{d|n} \lambda_d$ then w^+ satisfies condition (iii) from the statement of Lemma 1.2.1.

Note that in the hypotheses of Lemma 1.2.1, the sieving level z is equal to $\sqrt{D}$. Letting $P(z)$ be the product of all primes less than z, and considering expressions (7.2), (7.3) and (7.5) from [26], one has

$$w^+(n) = \Big(\sum_{d|n} \rho_d \Big)^2 = \Big(\sum_{\substack{d|n \\ d|P(z)}} \rho_d \Big)^2,$$

which, since $\rho_1 = 1$, is at least 1 if n has no prime factors less than z. So property (i) of Lemma 1.2.1 holds.

For property (ii), we take $\mathcal{A} = (a_n)$ to be the sequence supported on $[x]$ where $a_n \equiv 1$ for all n at most x. Then, in the notation of [26], $\mathcal{A}_d = x/d + O(1)$. Then the sum $\sum_{n \leqslant x} w^+(n)$ is equal to the right-hand side of expression (7.30), where the error $r_d(\mathcal{A})$ is $O(1)$. So it remains to estimate the right-hand side of (7.30). Employ the asymptotic expression for $J(D)$ in the regime $z \geqslant \sqrt{D}$ (for sieving dimension κ equal to 1), which begins page 118 of [26]. Then use (7.27) to estimate the error term $\sum_{d<\sqrt{D}} \lambda_d r_d(\mathcal{A})$ as $O(x^{2\xi+o(1)})$. Since $\xi < \frac{1}{2} - \frac{\gamma}{2}$, property (ii) follows. □

Lemma 1.2.2 (Lower-bound sieve)**.** *Let $\gamma, \delta > 0$, let ξ be a fixed real number satisfying $0 < \xi < \frac{1}{2} - \frac{\gamma}{2} - \frac{\delta}{2}$, and let x be a large integer, i.e. $x \geqslant x_0(\gamma, \delta)$ for some $x_0(\gamma, \delta)$. Define $z := x^\xi$ and $D := x^{2\xi+\delta}$. Then there exists a weight function $w^- : [x] \longrightarrow \mathbb{R}$ such that:*

(i) if n has no prime factors less than z then $w^-(n) \leqslant 1$;

(ii) if n has some prime factor that is less than z, then $w^-(n) \leqslant 0$;

(iii) there exists a positive real $c(\delta)$ such that $\sum_{n=1}^{x} w^-(n) \geqslant c(\delta) \frac{x}{\xi \log x}$; and

(iv) $w^-(n) = \sum_{d|n} \lambda_d^-$, *where* $(\lambda_d^-)_{d\geqslant 1}$ *is a sequence of real numbers satisfying* $\lambda_d^- = 0$ *for* $d > D$ *and* $|\lambda_d^-| \leqslant 1$.

Proof. The weight w^- is essentially the definition of a lower-bound sieve weight. To get a result of the required quality, take $\mathcal{A} = (a_n)$ to be the sequence supported on $[x]$, where $a_n \equiv 1$ for all n at most x, and construct the weights λ_d^- according to the optimal linear sieve[4] (applied to the sequence $\mathcal{A}$). The required results are proved in Chapter 11 of [26] and summarised at the beginning of Chapter 12 of the same volume.

Again, let us try to be more specific about where the relevant estimates may be located. The optimal linear sieve is an optimised version of the beta sieve, which is a combinatorial sieve, meaning that the weights are either equal to $\mu(d)$ or 0. Now, when sieving $\mathcal{A}$ with *any* lower-bound combinatorial linear sieve (λ_d^-), with sieving level z and level of support D as given in the hypotheses of Lemma 1.2.2, by construction the weight $w^-(n) = \sum_{d|n} \lambda_d^-$ immediately satisfies parts (i), (ii) and (iv) of the above theorem. [The closest to an exact reference for this assertion is the general framework of sieve weights described in section 5.3 of [26], in particular expressions (5.20) and (5.21).]

To establish part (iii), we note that the right-hand-side of equation (12.13) of [26] is exactly an estimate for the quantity $\sum_{n=1}^{x} w^-(n)$ with optimised weights. In those authors' notation, with our choice of parameters z and D, we have $s = \frac{2\xi+\delta}{\xi}$. This is at least $2 + 2\delta$, and in particular is at least 2: therefore $f(s) > 0$, and the main term of (12.13) is of the order required in part (iii). As in the proof of property (ii) in Lemma 1.2.1, the errors $|r_d(\mathcal{A})|$ are $O(1)$, and since $D \leqslant x^{1-\gamma}$ the error $R(\mathcal{A}, D)$ is $O(x^{1-\gamma})$. This is negligible compared to the main term (provided x is large enough in terms of γ and δ). The lemma is proved.

Another useful reference for the linear sieve is Chapter 8 of [42], in which Theorem

[4] also known as the Rosser-Iwaniec sieve

8.4 may alternatively be used for this proof. □

One may remark that, in Lemma 1.2.1, we could have used the upper-bound weights in the optimal linear sieve, instead of the Selberg weights, and referred to expression (12.12) of [26] to estimate $\sum_{n\leqslant x} w^+(n)$. We do not rule out the possibility that such a switch may improve[5] the constant 64 appearing in part (1) of Theorem 1.1.5, as one will be able to take the values of the parameters z and D to be much closer together in this framework ($\log D/\log z \approx 1$ as opposed to $\log D/\log z \approx 2$). Computing the optimal result would be rather complicated. By contrast, the Selberg sieve is a significantly simpler object than the optimal linear sieve, and may still be used to derive a perfectly respectable constant[6] in Theorem 1.1.5: we choose to proceed with this method.

Analogously, the replacement of the optimal linear sieve by a weighted linear sieve could potential decrease the constant in Theorem 1.1.6 from 6 to 5, but at the cost of significantly added complications.

In the sequel we shall only use the properties of these weights stated in Lemmas 1.2.1 and 1.2.2. Once x is fixed, we shall freely consider these weights as functions on $\mathbb{N}$, supported on $[x]$.

To finish this section, let us develop some results on the Fourier theory of sieve weights. We recall the usual definitions, if only to fix normalisations. For an arbitrary function $f : \mathbb{Z}/q\mathbb{Z} \longrightarrow \mathbb{C}$ and $r \in \mathbb{Z}/q\mathbb{Z}$, identifying $[q]$ and $\mathbb{Z}/q\mathbb{Z}$ in a harmless

[5] although the complicated piecewise definition of the function $F(s)$ occurring in expression (12.12) of [26] will render rather fiddly the act of calculating the best constant that can be derived from this sieve; a numerical integrator would probably be required.

[6] By contrast, instead of 64, our first drafts ended up with a constant somewhere in the thousands.

manner, we define the additive Fourier coefficient

$$\widehat{f}(r) = \sum_{a \in \mathbb{Z}/q\mathbb{Z}} f(a) e\left(-\frac{ra}{q}\right).$$

Taking $\chi : (\mathbb{Z}/q\mathbb{Z})^\times \longrightarrow \mathbb{C}$ to be a multiplicative character, we define the multiplicative Fourier coefficient

$$\widehat{f}(\chi) = \sum_{x \in (\mathbb{Z}/q\mathbb{Z})^\times} f(x)\overline{\chi(x)}.$$

Sieve weights, being weighted sums of arithmetic progressions, enjoy cancellation in their non-trivial Fourier coefficients, both additive and multiplicative. The following two lemmas formalise this notion; versions hold for the weights coming from either of the two previous propositions but, for ease of application, we state them only for the weights to which they will be applied.

Lemma 1.2.3. *Let w^+ be as in Lemma 1.2.1, with its associated large integer x. Let q be a prime number satisfying $q \geqslant x$. Then*

$$\sum_{r=1}^{q-1} |\widehat{w^+}(r)| \ll q x^{2\xi + o(1)} \log q. \tag{1.2}$$

Proof. The left-hand-side of (1.2) may be written explicitly as

$$\sum_{r=1}^{q-1} \left| \sum_{n=1}^{x} \sum_{d|n} \lambda_d^+ e\left(-\frac{rn}{q}\right) \right|.$$

Swapping the summation over d and n, and using the pointwise bound $3^{\nu(d)} \ll d^{o(1)}$, the above expression is at most

$$x^{o(1)} \sum_{r=1}^{q-1} \sum_{d \leqslant x^{2\xi}} \left| \sum_{y \leqslant \frac{x}{d}} e\left(-\frac{rdy}{q}\right) \right|. \tag{1.3}$$

We denote the inner sum by S. By the standard estimate

$$\left|\sum_{x\leqslant X} e\left(\frac{ax}{q}\right)\right| \ll \max\left(\frac{q}{a}, \frac{q}{q-a}\right)$$

for any a in the range $1 \leqslant a \leqslant q-1$, and for any X, we conclude that

$$|S| \ll \max\left(\frac{q}{(rd \bmod q)}, \frac{q}{q-(rd \bmod q)}\right). \tag{1.4}$$

Here $rd \bmod q$ is the least positive residue congruent to rd modulo q, and we have noted that rd is never a multiple of q. Substituting the bound (1.4) into (1.3) yields

$$\sum_{r=1}^{q-1} |\widehat{w^+}(r)| \ll x^{o(1)} \sum_{r=1}^{q-1} \sum_{d\leqslant x^{2\xi}} \max\left(\frac{q}{(rd \bmod q)}, \frac{q}{q-(rd \bmod q)}\right).$$

Swapping the sums over r and d, we see that for each fixed d the value $rd \bmod q$ achieves each value from 1 to $q-1$ exactly once. Splitting the sum into those r for which $rd \bmod q$ is less than $\frac{q}{2}$, and those for which $rd \bmod q$ is greater than $\frac{q}{2}$, we obtain the lemma. □

We now use the Pólya-Vinogradov theorem to bound the non-trivial multiplicative Fourier coefficients of sieve weights with small support.

Theorem 1.2.4 (Pólya-Vingoradov theorem)**.** *Let q be a natural number and let χ be a non-principal Dirichlet character modulo q. Then, for any natural numbers M and N,*

$$\sum_{n=M}^{M+N-1} \chi(n) \ll q^{\frac{1}{2}} \log q.$$

Proof. This was first proved in 1918 by Pólya and Vinogradov, independently. See Chapter 23 of [16] for details. □

The following short argument was suggested to us by Adam Harper.

Lemma 1.2.5. *Let w^- be the weight from Lemma 1.2.2, with its associated large integer x. Let q be a prime number satisfying $q \geqslant x$. Then for every non-principal character χ we have the bound*

$$|\widehat{w^-}(\chi)| \ll x^{2\xi+\delta} q^{\frac{1}{2}} \log q.$$

Proof. For χ a non-trivial character we have

$$\begin{aligned}
|\widehat{w^-}(\chi)| &= |\sum_{n \leqslant x} w^-(n)\overline{\chi}(n)| \\
&\leqslant \sum_{d \leqslant x^{2\xi+\delta}} |\lambda_d^-| |\sum_{\substack{n \leqslant x \\ d|n}} \overline{\chi}(n)| \\
&\leqslant \sum_{d \leqslant x^{2\xi+\delta}} |\sum_{m \leqslant \frac{x}{d}} \overline{\chi(md)}| \\
&\leqslant x^{2\xi+\delta} q^{\frac{1}{2}} \log q
\end{aligned}$$

with the final line following from the multiplicativity of χ and the Pólya-Vinogradov theorem applied to $\overline{\chi}$. □

In the case where x is equal to $q^{\frac{1}{4}+\varepsilon}$, one may fruitfully employ the Burgess bound in place of Pólya-Vinogradov.

Theorem 1.2.6 (Burgess bound, simplified form)**.** *Let q be prime, and let χ be a non-principal Dirichlet character modulo q. Let ε be a parameter in the range $0 < \varepsilon \leqslant 1/2$. Then there is an absolute positive constant c such that*

$$|\sum_{n \leqslant q^{\frac{1}{4}+\varepsilon}} \chi(n)| \ll q^{\frac{1}{4}+\varepsilon-c\varepsilon^2}.$$

Proof. This is the simplified form of the Burgess bound [13], given in Lemma 12 of [45]. □

Applying Theorem 1.2.6 in the context of Lemma 1.2.5, instead of applying Pólya-Vinogradov, we derive the following lemma.

Lemma 1.2.7. *Let ε be a parameter in the range $0 < \varepsilon \leqslant 1/2$. Let w^- be the weight from Lemma 1.2.2, with its associated large integer x. Suppose further that the parameters ξ and δ from the statement of Lemma 1.2.2 are small enough in terms of ε. Suppose that q is a prime number, with $q^{\frac{1}{4}+\varepsilon} \asymp x$. Then for every non-principal character χ modulo q we have the bound*

$$|\widehat{w^-}(\chi)| \ll q^{\frac{1}{4}+\varepsilon-c\varepsilon^2}, \tag{1.5}$$

where c is an absolute positive constant.

Although both very easy to prove, it turns out that these estimates will suffice for the deductions of Theorems 1.1.5 and 1.1.6.

1.3 Subgroup obstructions

This section will be devoted to the proof and discussion of the following lemma, which will be important in the proof of Theorem 1.1.7.

Lemma 1.3.1. *Let $\varepsilon > 0$, and let q be prime. Suppose that $\eta := q^{-\frac{3}{4}+\varepsilon}$. Then there exists a constant $C(\varepsilon)$ such that if $q \geqslant C(\varepsilon)$ then there does not exist a proper subgroup $H \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$ and an $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that $P_\eta \subseteq xH$.*

This is equivalent to proving that there is no non-principal character of $(\mathbb{Z}/q\mathbb{Z})^\times$ taking constant values on P_η.

Unfortunately, we cannot show that the primes p less than q enjoy any equidistribution in cosets of $(\mathbb{Z}/q\mathbb{Z})^\times$, and so we are restricted to using very general combi-

natorial arguments such as Lemma 0.2.7, in lieu of Fourier analytic techniques. We remark again that Lemma 1.3.1 was later rediscovered independently by Pollack[7] in [68].

Let us briefly discuss why proving equidistribution may be a genuinely difficult problem[8]. Proving equidistribution is equivalent to exhibiting some cancellation for the character sum $\sum_{p<\eta q} \chi(p)$, and the standard method for doing this is to use the zero-free region of the Dirichlet L-function $L(s,\chi)$. For example, consider the result ([49] p. 124)

$$\sum_{p<x} \chi(p) \ll_A \sqrt{q}x(\log x)^{-A} \tag{1.6}$$

for any non-principal character χ modulo q. Regrettably this result is worse than trivial for our applications, as $x \leqslant q$. Though the use of $\sqrt{q}$ in (1.6) by Iwaniec-Kowalksi is a little wasteful (one may do better by referring back directly to the prime number theorem as given in (5.51) of the same volume [49]), nothing better than trivial may be achieved, and this is more than just a phenomenon of Siegel zeros: even the q dependence in the zero-free region $\sigma > 1 - \frac{c}{\log(q(|t|+2))}$ is too poor, and this bound has resisted improvement for 80 years[9]. In Theorem 2 of Chapter 9 of [58], Montgomery gives an upper bound for the length of an interval on which a character is surjective when restricted to primes. This is a stronger conclusion than that of Lemma 1.3.1, but Montgomery's bounds are unfortunately conjectural on a larger zero-free region around $s = 1$ than that which is currently known. Of course, conditional on the Generalised Riemann Hypothesis the left hand side of (1.6) enjoys

[7]The main theorem of Pollack's paper (the final published version is [69]) is a useful result in this general task of finding non-residues in short intervals. But it doesn't contribute directly towards Lemma 1.3.1, as the non-residues found might nonetheless lie in the same coset.

[8]Of course we cannot rule out that we may have missed some trivial argument.

[9]There is some work (some theorems from [28], for example) showing that for sums over short intervals there are only a few exceptional conductors for which a better cancellation fails to hold, but this does not assist with the consideration of fixed conductor q.

almost square-root cancellation, with only logarithmic dependence on q. An easy Fourier argument would then imply that three primes sufficed.

Proposition 1.3.2. *Assume the Generalised Riemann Hypothesis. Then there exists a positive constant C such that for all primes $q \geqslant C$ one has $P_1^3 = (\mathbb{Z}/q\mathbb{Z})^\times$.*

Proof. Assuming GRH, one has the character sum bound $\sum_{p \leqslant x} \chi(p) \ll x^{\frac{1}{2}} (\log xq)^2$ for all non-principal χ. (See [62, Exercise 7.4.8]). Then for any $a \in (\mathbb{Z}/q\mathbb{Z})^\times$ we have

$$\begin{aligned}
1_{P_1} \star 1_{P_1} \star 1_{P_1}(a) &= \frac{1}{q-1} \sum_{\chi} \widehat{1_{P_1}}(\chi)^3 \chi(a) \\
&= \frac{1}{q-1} |P_1|^3 + O\left(q^{-1} \sum_{\chi \neq \chi_0} |\widehat{1_{P_1}}(\chi)|^3 \right) \\
&= \frac{1}{q-1} |P_1|^3 + O\left(q^{-1+\frac{1}{2}} \log^2(q|P_1|) \sum_{\chi \neq \chi_0} |\widehat{1_{P_1}}(\chi)|^2 \right) \\
&\gg \frac{q^2}{\log^3 q} - O(q^{\frac{3}{2}} (\log q)^{\frac{1}{2}}) \\
&> 0
\end{aligned}$$

for large enough q, where the penultimate line follows from Parseval's identity. This immediately implies the proposition. □

An alternative approach to ruling out hypothetical conspiracies of characters at primes is to convert such behaviour into a conspiracy over an interval, obtaining a contradiction to the various known estimates for character sums over intervals. For example, using such an approach, one may show that there are prime quadratic non-residues and residues less than $q^{\frac{1}{4}+\varepsilon}$. The non-residue case is due to Burgess [12], and is an immediate application of his famous character sum bound[10] (Theorem 1.2.6); the residue case is due to Vinogradov and Linnik in [88], although Pintz gave a much

[10] By a further argument Burgess showed that there is a prime quadratic non-residue at most $q^{\frac{1}{4\sqrt{e}}+\varepsilon}$, but this tweak does not seem to be available for prime quadratic residues.

simpler proof in [67]. Pintz's method will be an inspiration for work in the next chapter.

A natural generalisation of this method for n^{th}-power residues, undertaken by Elliott in [21], shows that there are primes p less than $q^{\frac{n-1}{4}+\varepsilon}$ that are n^{th}-power residues modulo q. The equivalent statement for a non-residue follows immediately from Burgess. Collecting these results together, we see that the following result is already known:

Proposition 1.3.3. *Let $\varepsilon > 0$, and let q be prime. Suppose that $\eta := q^{-\frac{1}{4}+\varepsilon}$. Then there exists a constant $C(\varepsilon)$ such that if $q \geqslant C(\varepsilon)$ then the following holds: if $H \subseteq (\mathbb{Z}/q\mathbb{Z})^\times$ is a subgroup with index 2, 3, or 4, then there does not exist $x \in (\mathbb{Z}/q\mathbb{Z})^\times$ such that $P_\eta \subseteq xH$.*

1.3.1 A Selberg-Delange approach

Having discussed at some length the problems surrounding Lemma 1.3.1, let us proceed with the first proof.

Proof of Lemma 1.3.1. Let η be equal to $q^{-\frac{3}{4}+\varepsilon}$ and suppose that P_η is contained in some coset of a non-trivial subgroup of $(\mathbb{Z}/q\mathbb{Z})^\times$: equivalently, some Dirichlet character χ with conductor q (necessarily primitive) is constant on P_η. We may preclude the case when χ is the quadratic character by the result of Pintz [67, Theorem 1] mentioned above, so without loss of generality χ is complex. If $\chi(p) \equiv z$ for all primes p satisfying $p < \eta q$, where z is some root of unity, then $\chi(n)$ agrees with the function $z^{\Omega(n)}$ for all natural numbers n satisfying $n < \eta q$, where $\Omega(n)$ is the number of prime factors of n counted with multiplicity. It will be important that z is bounded uniformly away from -1; w.l.o.g. we may assume, by replacing χ with a suitable power of χ, that $\mathrm{Re}(z) \geqslant \mathrm{Re}(e^{\frac{2\pi i}{3}}) = -\frac{1}{2}$.

We have Burgess's character sum estimate from [13]. We have already stated a

version of this bound in Theorem 1.2.6 , but here we give the more familiar version, namely that for every natural number k one has

$$\sum_{n\leqslant x}\chi(n) \ll x^{1-\frac{1}{k}}q^{\frac{k+1}{4k^2}}(\log q)^{\frac{1}{k}}. \tag{1.7}$$

A proof of this formulation may be found as [49, Theorem 12.6], with the strengthening given in equation (12.58).

We obtain a contradiction by observing that the equivalent sum, with $z^{\Omega(n)}$ replacing $\chi(n)$, does not enjoy the same cancellation as that which is given in (1.7). Indeed, Theorem 5.2 in [86] shows that for any root of unity z (apart from $z=-1$)

$$\sum_{n\leqslant x} z^{\Omega(n)} = x(\log x)^{z-1}\left(\frac{\prod_p\left(1-\frac{z}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^{z}}{\Gamma(z)}+O\left(\frac{1}{\log x}\right)\right)$$

with implicit constant independent of z. The result is proved using the Selberg-Delange method, though is in essence originally due to Sathe [77].

Since $\mathrm{Re}(z)\geqslant -\frac{1}{2}$, z lies on a segment of the unit circle on which $\Gamma(z)$ is uniformly bounded, and hence $\frac{1}{\Gamma(z)}$ is bounded away from zero. Further, as $|z|=1$, the factor $\prod_p\left(1-\frac{z}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^{z}$ is bounded away from zero (immediately seen by taking logarithms). Therefore, for large enough q,

$$\Big|\sum_{n\leqslant \eta q} z^{\Omega(n)}\Big| \gg \frac{\eta q}{\log^{\frac{3}{2}}(\eta q)} \tag{1.8}$$

for all z with $|z|=1$ and $\mathrm{Re}(z)\geqslant -\frac{1}{2}$, where the implicit constant is uniform in all such choices of z. For large enough q and k, depending on ε, (1.8) is contradictory to (1.7) taken with x equal to ηq. This concludes the proof of Lemma 1.3.1. □

1.3.2 A large sieve approach

Some time after the original publication of the above work in [91], it was suggested to us by Prof. Jan-Christoph Schlage-Puchta that one of his results (a large sieve inequality adapted for sequences supported on primes) could be used to derive a stronger version of Lemma 1.3.1. This is indeed the case, and although the resulting bounds do not imply any form of equidistribution, they will enable us to modestly improve Theorem 1.1.7 part (2) over the equivalent result in [91].

Let us state Schlage-Puchta's theorem[11]. This is a special case of Theorem 3 of [70].

Theorem 1.3.4 (Schlage-Puchta, [70]). *Let $(a_p)_{p\in\mathcal{P}}$ be any sequence of complex numbers supported on primes. Let q be a fixed large prime, x a large real number, and R some real parameter satisfying $1 < R < x^{\frac{1}{2}}$. Let k be a natural number, and δ be a positive real number. Finally, let $\mathcal{C}$ be any collection of characters modulo q. Then, if $k \geqslant 2$ is a natural number,*

$$\sum_{\chi\in\mathcal{C}}\left|\sum_{p\leqslant x} a_p\chi(p)\right|^2 \leqslant \left(\frac{x}{\log R} + c_{k,\delta}x^{1-\frac{1}{k}}q^{\frac{k+1}{4k^2}+\delta}|\mathcal{C}|R^{\frac{2}{k}}\right)\sum_{p\leqslant x}|a_p|^2. \tag{1.9}$$

The proof uses a linear-algebraic sieving principle of Bombieri [57, Lemma 1.5] and the stronger decay bounds which are available by passing to sieve weights. In that sense, the method is of a similar spirit to the proofs of this chapter. Using it,

[11] We take this opportunity to fill in a slight gap in the proof of this theorem in [70]. The second display equation of page 147 of that paper does not seem to follow from the author's definition of inner product space and vector $(\widehat{a_p})$, as the interfering factors of $f(p)$ have been miscalculated. Rather, one should take $\widehat{a_p} = f(p)^{-1}e^{\log^2 p/N}a_p$ for p prime in the range $(R^2, N]$, and 0 otherwise. Define $\widehat{\chi}$ to be equal to $\overline{\chi}$, as in the paper. With these definitions, one has $a_p\chi(p) = f(p)e^{-\log^2 p/N}\widehat{a_p}\overline{\widehat{\chi}(p)}$, and so the inner sum of the left-hand side of the second display equation on page 147 can be interpreted as an inner product of $(\widehat{a_p})$ and $(\widehat{\chi}(p))$, as the author intends. Following this alteration through, the right-hand side of the second display equation on page 147 has a different final multiplicative factor, namely $\sum|a_p|^2e^{-\log^2 p/N}f(p)^{-1}$. But then we note that $f(p) \geqslant 1$, as f is an upper bound sieve weight, and so this multiplicative factor may be upper bounded by the multiplicative factor which appears in the paper. The rest of the argument in the paper, building on the second display equation of page 147, is correct.

we prove the following lemma, which strengthens Lemma 1.3.1 in all the regimes of interest.

Lemma 1.3.5. *Let $\varepsilon, \delta > 0$, with δ sufficiently small in terms of ε. Let q be prime, and let $\eta := q^{-\frac{3}{4}+\varepsilon}$. Let C be a positive real and d a natural number, and suppose that $d \leqslant Cq^{\delta}$. Let χ be a non-principal character modulo q, of order d. Then the size of the image $|\chi(P_\eta)|$ is at least $(1 - o_{C,\delta}(1)) \left(\frac{1}{2} - \frac{1}{2+8\varepsilon} - O(\sqrt{\delta})\right) d$, where the errors terms are uniform over all choices of χ satisfying the hypotheses.*

We remark that the bound $\frac{1}{2} - \frac{1}{2+8\varepsilon}$ tends to zero as ε tends to zero, so this lemma has nothing to say about primes p satisfying $p \leqslant q^{\frac{1}{4}}$, as expected. Furthermore, even if $\varepsilon = \frac{3}{4}$ and $d = 2$, the size of the image $|\chi(P_1)|$ guaranteed by this lemma is only $(1 - o(1))\frac{3}{8}d$. Since $\frac{3}{8} < \frac{1}{2}$, even with Lemma 1.3.5 in place we still need the results of Pintz on prime quadratic residues in order to conclude Lemma 1.3.1.

Proof. We apply Theorem 1.3.4. Given the parameters in the statement of Lemma 1.3.5, we then take x to be equal to ηq, a_p to be equal to 1 for all primes, δ to be as given, and the set of characters $\mathcal{C}$ to be the powers of χ, namely $\{\chi_0, \chi, \chi^2, \cdots, \chi^{d-1}\}$ (where χ_0 denotes the principal character). Theorem 1.3.4 then shows that

$$\sum_{a=1}^{d} \left| \sum_{p<\eta q} \chi^a(p) \right|^2 \leqslant (1 + o(1)) \left(\frac{\eta q}{\log R} + c_{k,\delta} C (\eta q)^{1-\frac{1}{k}} q^{\frac{1}{4k}+\frac{1}{4k^2}+2\delta} R^{\frac{2}{k}} \right) \frac{\eta q}{\log \eta q}, \quad (1.10)$$

for any R satisfying $1 < R < (\eta q)^{1/2}$, and for any natural number k at least 2.

Let us find a lower bound for the left-hand side. Indeed, by applying Parseval's identity to a suitable quotient of $(\mathbb{Z}/q\mathbb{Z})^{\times}$, or arguing directly from orthogonality of characters, we have

$$\sum_{a=1}^{d} \left| \sum_{p<\eta q} \chi^a(p) \right|^2 = d \sum_{a=1}^{d} |\{p : p < \eta q, \chi(p) = z^a\}|^2,$$

where z is some fixed primitive d^{th} root of unity. Now, suppose that there are θd values of a for which there exists some prime p less than ηq with $\chi(p) = z^a$. By the Cauchy-Schwarz inequality, we have that

$$d\sum_{a=1}^{d} |\{p < \eta q : \chi(p) = z^a\}|^2 \geqslant \frac{d}{\theta d}\left(\sum_{a=1}^{d} |\{p < \eta q : \chi(p) = z^a\}|\right)^2$$
$$\geqslant (1 - o(1))\frac{1}{\theta}\frac{(\eta q)^2}{(\log \eta q)^2}. \tag{1.11}$$

We now manipulate the right-hand side of (1.10). Let us write $b := \log \eta / \log q + 1$, so that $\eta q = q^b$. Let us also choose some parameter $c < b/2$, and write $R := q^c$ (the condition $c < b/2$ ensures that $R < (\eta q)^{1/2}$). Combining (1.11) with (1.10), and using these variables, we get

$$\frac{1}{\theta}\frac{q^b}{b \log q} \leqslant (1 + o(1))\left(\frac{1}{c}\frac{q^b}{\log q} + c_{k,\delta} C q^b q^{\frac{1}{k}(-\varepsilon + 2c + \frac{1}{4k}) + 2\delta}\right). \tag{1.12}$$

Assume that δ is small enough in terms of ε. Choose $c = \frac{\varepsilon}{2} - 5\sqrt{\delta}$, and choose $k = \lceil 1/\sqrt{\delta} \rceil$. With these choices of parameters, $c < b/2$ (since $b/2 = 1/8 + \varepsilon/2$) and the first term on the right-hand side of (1.12) dominates (if q is large enough in terms of δ and C), i.e.

$$\frac{1}{\theta}\frac{1}{b} \leqslant (1 + o_{\delta,C}(1))\frac{1}{c}.$$

Rearranging, we may conclude that

$$\theta \geqslant (1 - o_{\delta,C}(1))\left(\frac{1}{2} - \frac{1}{2 + 8\varepsilon} - O(\sqrt{\delta})\right)$$

as claimed. □

1.4 Proofs of main theorems

In this section we present the proofs of the three main theorems, using arguments from Fourier analysis.

Proof of Theorem 1.1.5. We begin with part (1). Let $\varepsilon > 0$, and without loss of generality also assume that $\varepsilon \leqslant \frac{1}{4}$. By adjusting the $o(1)$ term in the statement of part (1) of Theorem 1.1.5, we may also assume that q is sufficiently large. Taking η to be equal to $q^{-\frac{1}{4}+\varepsilon}$, we will proceed to show that

$$|P_\eta^2| \geqslant \left(\frac{2\varepsilon}{3+4\varepsilon}\right)^2 q(1-o(1)),$$

which will yield part (1) after the substitution $\varepsilon = \frac{1}{4}$.

To do this, we let a be an element of $(\mathbb{Z}/q\mathbb{Z})^\times$, and let S_a denote the number of solutions to the equation

$$p_1 p_2 \equiv a (\mathrm{mod}\, q)$$

with $p_1, p_2 \in P_\eta$. We proceed to give an upper bound for S_a.

We make an initial reduction. Let S_a^* denote the number of solutions counted by S_a in which $p_1, p_2 > q^{\frac{1}{2}}$. Observe that $|S_a - S_a^*| = O(q^{\frac{1}{2}})$, which is negligible compared to the upper bound for S_a^* which we will eventually give in (1.15). So we proceed to bound S_a^*.

Now let ξ and γ be certain positive real numbers, to be chosen later, satisfying the inequality $\xi < \frac{1}{2} - \frac{\gamma}{2}$. Assume that q is large enough[12] in terms of γ. Let x be defined to be $\lfloor \eta q \rfloor$, and let w^+ be the weight constructed in Lemma 1.2.1 using these constants ξ and γ.

By our above reductions, and property (i) of Lemma 1.2.1, we have the upper

[12]We will eventually let $\gamma := 1/2$, and so in fact all that is required is for q to be sufficiently large.

bound

$$S_a^* \leqslant \sum_{n=1}^{q-1} w^+(n) w^+(an^*),$$

where n^* denotes the multiplicative inverse modulo q. By the additive Fourier inversion formula, this is equal to

$$\frac{1}{q^2} \sum_{r=1}^{q} \sum_{s=1}^{q} \widehat{w^+}(r) \widehat{w^+}(s) \sum_{n=1}^{q-1} e\left(\frac{rn + san^*}{q}\right). \tag{1.13}$$

By property (ii) of Lemma 1.2.1, the contribution from the term where $r = s = 0$ is at most $(1 + o_\gamma(1)) \frac{\eta^2 q}{\xi^2 \log^2 \eta q}$. The remaining contribution is at most $T_1 + T_2 + T_3$, where

$$T_1 := \frac{1}{q^2} \sum_{r=1}^{q-1} \sum_{s=1}^{q-1} |\widehat{w^+}(r)||\widehat{w^+}(s)| \left| \sum_{n=1}^{q-1} e\left(\frac{rn + san^*}{q}\right) \right|;$$
$$T_2 := (1 + o_\gamma(1)) \frac{\eta}{q\xi \log \eta q} \sum_{r=1}^{q-1} |\widehat{w^+}(r)| \left| \sum_{n=1}^{q-1} e\left(\frac{rn}{q}\right) \right|;$$
$$T_3 := (1 + o_\gamma(1)) \frac{\eta}{q\xi \log \eta q} \sum_{s=1}^{q-1} |\widehat{w^+}(s)| \left| \sum_{n=1}^{q-1} e\left(\frac{san^*}{q}\right) \right|.$$

In T_1 the inner sum is the Kloosterman sum $\mathrm{Kl}_2(r, sa; q)$ (see Chapter 11 of [49]) which enjoys the Weil bound

$$\mathrm{Kl}_2(r, sa; q) \leqslant 2\sqrt{q}.$$

The other two exponential sums are trivially of size 1, and the sums of the Fourier coefficients of w^+ are precisely of the form estimated in Lemma 1.2.3. Hence we may conclude that

$$T_1 \ll_\gamma \eta^{4\xi} q^{4\xi + \frac{1}{2} + o(1)};$$
$$T_2 \ll_\gamma \frac{\eta^{1+2\xi} q^{2\xi + o(1)}}{\xi};$$

$$T_3 \ll_\gamma \frac{\eta^{1+2\xi} q^{2\xi+o(1)}}{\xi}.$$

Substituting $\eta = q^{-\frac{1}{4}+\varepsilon}$, a short calculation demonstrates that the term with $r = s = 0$ dominates, as q tends to infinity with γ and ξ fixed, provided that the constant ξ satisfies

$$\xi < \frac{2\varepsilon}{3+4\varepsilon}. \tag{1.14}$$

We choose ξ to satisfy this bound. Note that since $\varepsilon \leqslant \frac{1}{4}$ we have that $\xi < \frac{1}{8}$. Hence we may let $\gamma := \frac{1}{2}$ and satisfy the hypotheses of Lemma 1.2.1. The final conclusion is that

$$S_a^* \leqslant (1 + o_\xi(1)) \frac{q^{\frac{1}{2}+2\varepsilon}}{\xi^2 \log^2 \eta q} \tag{1.15}$$

for any ξ satisfying (1.14). By our initial remarks, the same bound holds for S_a.

Now we sum (1.15) over all a in P_η^2. This yields, using the Prime Number Theorem[13]

$$(1 + o(1)) \frac{\eta^2 q^2}{\log^2 \eta q} \leqslant (1 + o_\xi(1)) |P_\eta^2| \frac{q^{\frac{1}{2}+2\varepsilon}}{\xi^2 \log^2 \eta q}.$$

Making the substitution $\eta = q^{-\frac{1}{4}+\varepsilon}$ and rearranging gives $|P_\eta^2| \geqslant \xi^2 q(1 - o_\xi(1))$ for any ξ satisfying (1.14).

Now we let ξ depend on q, tending to $\frac{2\varepsilon}{3+4\varepsilon}$ from below suitably slowly as q tends to infinity. We then conclude

$$|P_\eta^2| \geqslant \left(\frac{2\varepsilon}{3+4\varepsilon}\right)^2 q(1 - o(1)).$$

After the substitution $\varepsilon = \frac{1}{4}$, this proves part (1).

Now let us consider part (2) of Theorem 1.1.5. First note that we may assume that $\varepsilon \leqslant 1/2$. Then apply Lemma 0.2.6 with $G = (\mathbb{Z}/q\mathbb{Z})^\times$ and $A = P_\eta$, that connects the

[13] In fact Chebyshev's elementary estimates would suffice for the proof

size of the product set $P_\eta \cdot P_\eta$ with the additive energy $E(P_\eta)$. Together, this shows that part (2) of Theorem 1.1.5 may be deduced from the following upper bound on $E(P_\eta)$.

Proposition 1.4.1. *Let q be prime, and let ε be in the range $0 < \varepsilon \leqslant 1/2$. Let $\eta := q^{-\frac{1}{2}+\varepsilon}$. Then $E(P_\eta) \ll_\varepsilon \frac{\eta^4 q^3}{\log^4 q}$.*

We proceed to prove Proposition 1.4.1, noting for future reference that the desired bound has order $q^{1+4\varepsilon}/\log^4 q$.

Proof of Proposition 1.4.1. Assuming the hypotheses of Proposition 1.4.1, let ξ be a small positive constant to be chosen later (depending on ε), γ equal $1/2$, and let $x := \lfloor \eta q \rfloor$. Assuming without loss of generality that q is sufficiently large, let w^+ be the upper-bound sieve weight given in Lemma 1.2.1, with these parameters.

We use the parameter ξ to make an initial reduction similar to the one from the proof of Theorem 1.1.5 part (1). Indeed, by a trivial argument, the number of solutions to $p_1 p_2 \equiv p_3 p_4 (\mathrm{mod}\, q)$, in which all the primes p_i are at most ηq and at least one of them is $O((\eta q)^\xi)$, is $O((\eta q)^{\xi+2})$. This is $O(q^{(\frac{1}{2}+\varepsilon)(\xi+2)})$, which for ξ small enough is of a lower order than the bound claimed in Proposition 1.4.1. Therefore, we may assume that all the primes p_i are in the range $(\eta q)^\xi < p_i \leqslant (\eta q)$. We let $E^*(P_\eta)$ denote this restricted solution count.

Using part (i) of Lemma 1.2.1, we have

$$\begin{aligned} E^*(P_\eta) &\leqslant \sum_{\substack{n_1,\cdots,n_4 \leqslant \eta q \\ n_1 n_2 \equiv n_3 n_4 (\mathrm{mod}\, q)}} w^+(n_1) w^+(n_2) w^+(n_3) w^+(n_4) \\ &= \frac{1}{q-1} \sum_\chi |\widehat{w^+}(\chi)|^4. \end{aligned}$$

Using part (ii) of Lemma 1.2.1, we see that the contribution from the principal char-

acter is at most

$$(1+o(1))\frac{1}{\xi^4}\frac{\eta^4 q^3}{(\log \eta q)^4}, \tag{1.16}$$

which is $\ll_\varepsilon \frac{\eta^4 q^3}{\log^4 q}$ (recalling that ξ will be taken suitably small in terms of ε). The remaining terms are

$$\begin{aligned}
&\frac{1}{q-1}\sum_{\chi\neq\chi_0}\Big|\sum_{n\leqslant \eta q}\Big(\sum_{d|n}\lambda_d^+\Big)\overline{\chi(n)}\Big|^4\\
\ll&\frac{1}{q}\sum_{d_1,\cdots,d_4<(\eta q)^{2\xi}}|\lambda_{d_1}^+\cdots\lambda_{d_4}^+|\sum_{\chi\neq\chi_0}\prod_{i=1}^4\Big|\sum_{m_i\leqslant \eta q/d_i}\overline{\chi(dm_i)}\Big|\\
\ll&q^{o(1)-1}(\eta q)^{8\xi}\max_{d_1,\cdots,d_4<(\eta q)^{2\xi}}\sum_{\chi\neq\chi_0}\prod_{i=1}^4\Big|\sum_{m_i\leqslant \eta q/d_i}\overline{\chi(m_i)}\Big|\\
\ll&q^{o(1)-1}(\eta q)^{8\xi}\max_{d_1,\cdots,d_4<(\eta q)^{2\xi}}\prod_{i\leqslant 4}\Big(\sum_{\chi\neq\chi_0}\Big|\sum_{m_i\leqslant \eta q/d_i}\overline{\chi(m_i)}\Big|^4\Big)^{\frac{1}{4}}.
\end{aligned}\tag{1.17}$$

Here we have used property (iii) of Lemma 1.2.1, in a manner suggested by Harper, reminiscent of the approach in Lemma 1.2.5. The final line follows by two applications of the Cauchy-Schwarz inequality.

Suitable bounds on expressions of this form were first given in [6] (the log factors later sharpened in [30]).

Theorem 1.4.2. *Let q be prime, and let B_1, B_2, B_3, B_4 be less than q. Then the number of solutions to $x_1x_2 \equiv x_3x_4 (\mathrm{mod}\, q)$ with $1 \leqslant x_i \leqslant B_i$ for all i is*

$$\frac{B_1B_2B_3B_4}{q} + O(\sqrt{B_1B_2B_3B_4}\log^2 q).$$

This is Theorem 1 of Ayyad, Cochrane, and Zheng [6]. The method of proof is Fourier analytic, and very much in the spirit of the methods of this chapter[14].

[14] though rather more intricate: the authors find it necessary to combine both additive and multiplicative Fourier expansions.

Now, following on from the final line of (1.17), and using Theorem 1.4.2, one has

$$\sum_{\chi}\Big|\sum_{m_i\leqslant \eta q/d_i}\overline{\chi(m_i)}\Big|^4 = q|\{x_1,x_2,x_3,x_4 : x_1x_2\equiv x_3x_4\,(\mathrm{mod}\,q),\ 1\leqslant x_j\leqslant \eta q/d_i\,\forall j.\}|$$
$$= \frac{(\eta q)^4}{d_i^4} + O(q^{1+o(1)}(\eta q)^2).$$

Therefore,

$$\sum_{\chi\neq\chi_0}\Big|\sum_{m_i\leqslant \eta q/d_i}\overline{\chi(m_i)}\Big|^4 = \frac{(\eta q)^4}{d_i^4} - \left\lfloor\frac{\eta q}{d_i}\right\rfloor^4 + O(q^{1+o(1)}(\eta q)^2)$$
$$\leqslant O((\eta q)^3) + O(q^{1+o(1)}(\eta q)^2)$$
$$\leqslant O(q^{1+o(1)}(\eta q)^2),$$

since $\eta q = O(q)$ by the assumption $\varepsilon\leqslant 1/2$.

In total the above shows that (1.17) is $\ll q^{o(1)}(\eta q)^{8\xi+2}$, which is $\ll q^{o(1)+(\frac{1}{2}+\varepsilon)(8\xi+2)}$. If ξ is small enough, this is of a lower order of magnitude than the term from the principal character, namely (1.16). We therefore conclude Proposition 1.4.1. □

As noted above, part (2) of Theorem 1.1.5 follows immediately from Proposition 1.4.1. □

We proceed to the proof of the second of our three main theorems.

Proof of Theorem 1.1.6 part (1). The proof of Theorem 1.1.6 part (1) is an easy consequence of a standard Fourier analysis idea, namely the use of triple convolutions.

Let ε be a positive constant, and let η be defined to be $q^{-\frac{1}{16}+\varepsilon}$: all as in the statement of the theorem. Without loss of generality we may assume $\varepsilon\leqslant\frac{1}{16}$. Let γ, δ and ξ be constants that satisfy the hypotheses of Lemma 1.2.2; we will choose these constants later, which will only depend on ε. Without loss of generality we assume that q is sufficiently large in terms of γ, δ and ε. Let x be defined to be $\lfloor\eta q\rfloor$, and let w^- be defined to be the weight from Lemma 1.2.2 using all of these parameters.

We proceed by showing that $(w^- \star 1_{P_\eta} \star 1_{P_\eta})(a) > 0$ for all $a \in (\mathbb{Z}/q\mathbb{Z})^\times$. Indeed, by multiplicative Fourier inversion we have the identity

$$\begin{aligned} w^- \star 1_{P_\eta} \star 1_{P_\eta}(a) &= \frac{1}{q-1} \sum_{\chi} \widehat{w^-}(\chi) \widehat{1_{P_\eta}}(\chi)^2 \chi(a) \\ &\geqslant \frac{(1-o(1))c(\delta)\eta^3 q^2}{\xi \log^3 \eta q} - \frac{1}{q-1} \sum_{\chi \neq \chi_0} |\widehat{w^-}(\chi) \widehat{1_{P_\eta}}(\chi)^2 \chi(a)|. \end{aligned} \tag{1.18}$$

by Property (iii) of Lemma 1.2.2. If the claim $w^- \star 1_{P_\eta} \star 1_{P_\eta}(a) > 0$ were false, then we would have

$$\frac{1}{q-1} \sum_{\chi \neq \chi_0} |\widehat{w^-}(\chi) \widehat{1_{P_\eta}}(\chi)^2| \geqslant (1-o(1)) \frac{c(\delta)\eta^3 q^2}{\xi \log^3 \eta q}$$

and therefore

$$\sup_{\chi \neq \chi_0} |\widehat{w^-}(\chi)| \sum_{\chi} |\widehat{1_{P_\eta}}(\chi)|^2 \geqslant (1-o(1)) \frac{c(\delta)\eta^3 q^3}{\xi \log^3 \eta q}.$$

But by Parseval's identity this would imply that

$$\sup_{\chi \neq \chi_0} |\widehat{w^-}(\chi)| \geqslant (1-o(1)) c(\delta) \frac{\eta^2 q}{\xi \log^2 \eta q}. \tag{1.19}$$

From Lemma 1.2.5, we have

$$\sup_{\chi \neq \chi_0} |\widehat{w^-}(\chi)| \ll (\eta q)^{2\xi+\delta} q^{\frac{1}{2}} \log q. \tag{1.20}$$

A short calculation shows that this contradicts (1.19), provided that

$$\xi < \frac{\frac{1}{4} + \log_q \eta}{1 + \log_q \eta} - \frac{\delta}{2} \tag{1.21}$$

For $\eta = q^{-\frac{1}{16}+\varepsilon}$, the condition reads

$$\xi < \frac{1}{5} + \frac{64\varepsilon}{75+80\varepsilon} - \frac{\delta}{2}. \tag{1.22}$$

There are two consequences. Firstly, since $\varepsilon \leqslant \frac{1}{16}$, any value of ξ that satisfies (1.22) automatically satisfies $\xi < \frac{1}{4}$, and so, if δ is sufficiently small and $\gamma = \frac{1}{2}$, the triple (ξ, δ, γ) satisfies the hypotheses for the above application of Lemma 1.2.2, namely $0 < \xi < \frac{1}{2} - \frac{\delta}{2} - \frac{\gamma}{2}$. Secondly, if δ is picked small enough in terms of ε, we may take such a triple in which ξ is greater than $\frac{1}{5}$, in addition to satisfying the inequality (1.22).

Picking such a triple (ξ, δ, γ), and contradicting (1.19) by design, we conclude that $(w^- \star 1_{P_\eta} \star 1_{P_\eta})(a) > 0$ for all a in $(\mathbb{Z}/q\mathbb{Z})^\times$. But trivially we then have $\max(w^-, 0) \star 1_{P_\eta} \star 1_{P_\eta}(a) > 0$ as well. Recalling the statement of Lemma 1.2.2, we observe that $\max(w^-, 0)$ is an arithmetic function supported on the natural numbers less than ηq, and furthermore only supported on numbers all of whose prime factors are at least $(\eta q)^\xi$. Since $\xi > \frac{1}{5}$, we see that $\max(w^-, 0)$ is only supported on numbers with at most 4 prime factors. Part (1) of Theorem 1.1.6 is then immediate. □

The following is a slight strengthening of Theorem 1.1.6 part (1), which we will use later.

Corollary 1.4.3. *Under the same hypotheses as Theorem 1.1.6 part (1),*

$$(\mathbb{Z}/q\mathbb{Z})^\times = P_\eta^3 \cup P_\eta^4 \cup P_\eta^5 \cup P_\eta^6.$$

Proof. Proceed as in the proof of Theorem 1.1.6 part (1), but instead of using the

sieve weight $w^-(n)$ use the function $w(n)$, where

$$w(n) := \begin{cases} w^-(n) & \text{if } n \neq 1 \\ 0 & \text{if } n = 1. \end{cases}$$

Altering w^- at a single point doesn't affect the bound on Fourier coefficients (1.20), so the proof proceeds identically. But in the final step of the proof the function $\max(w, 0)$ is only supported on numbers with 1, 2, 3, or 4 prime factors, and so the function $\max(w, 0) \star 1_{P_\eta} \star 1_{P_\eta}$ is only supported on $P_\eta^3 \cup P_\eta^4 \cup P_\eta^5 \cup P_\eta^6$. Corollary 1.4.3 follows. □

To prove part (2) of Theorem 1.1.6, if one were to take $\eta := q^{-\frac{1}{4}+\varepsilon}$, one would need very little additional argument. Indeed, one could proceed identically until (1.21). By picking δ small enough in terms of ε one may ensure that the upper bound in (1.21) is positive, and so there is some natural number $k(\varepsilon)$ and some ξ satisfying (1.21) such that $\xi > \frac{\frac{3}{4}+\varepsilon}{k(\varepsilon)-1}$. Then $\max(w^-, 0)$ is supported on numbers with at most $k(\varepsilon) - 2$ prime factors, and part (2) would be proved.

However, to prove part (2) in the form stated (with $\eta = q^{-\frac{3}{4}+\varepsilon}$ being substantially smaller than above) we have had to modify this idea to include repeated convolution, taking inspiration from [45].

Proof of Theorem 1.1.6 part (2). Let ε be positive and suppose $\eta := q^{-\frac{3}{4}+\varepsilon}$, as in the statement of the theorem. Let x be defined to be $\lfloor \eta q \rfloor$. Let γ be defined to be $\frac{1}{2}$, and let ξ and δ be small positive parameters depending on ε (to be chosen later). We may assume that q is sufficiently large in terms of ε. Let w^- be the lower-bound sieve weight from Lemma 1.2.2.

Now, for a some fixed coprime residue modulo q, we define

$$W(a) := \sum_{\substack{n \leqslant q^{\frac{1}{4}+\varepsilon} \\ p_1,\cdots,p_8 \leqslant q^{\frac{1}{4}} \\ np_1\cdots p_8 \equiv a (\mathrm{mod}\, q)}} w^-(n).$$

We will show that $W(a) > 0$, whereupon Theorem 1.1.6 part (2) immediately follows, since $W(a)$ is at most the number of ways of expressing a as the product of at most $\xi^{-1}(\frac{1}{4}+\varepsilon)$ primes, with each prime less than $q^{\frac{1}{4}+\varepsilon}$.

The other sum of interest, other than W, will be

$$\widehat{V}(\chi) := \sum_{p \leqslant q^{\frac{1}{4}}} \chi(p).$$

One observes that

$$\sum_{\chi} |\widehat{V}(\chi)|^8 \ll q^2.$$

Indeed, $|\widehat{V}(\chi)|^8$ is equal to

$$\left| \sum_{n \leqslant q} a_n \chi(n) \right|^2,$$

where

$$|a_n| \ll \begin{cases} 1 & \text{if } n \text{ is a product of four primes, each at most } q^{\frac{1}{4}}; \\ 0 & \text{otherwise.} \end{cases}$$

By Parseval,

$$\begin{aligned} \sum_{\chi} |\widehat{V}(\chi)|^8 &\leqslant q \sum_{n \leqslant q} |a_n|^2 \\ &\ll q^2. \end{aligned}$$

Now, by Fourier inversion,

$$W(a) = \frac{1}{q-1}\sum_{\chi} \widehat{w^-}(\chi)\widehat{V}^8(\chi)\chi(a).$$

The contribution from the principal character is at least $c(\delta)\xi^{-1}q^{\frac{5}{4}+\varepsilon-o(1)}$, where $c(\delta)$ is the parameter in the statement of Lemma 1.2.2. Provided that ξ and δ are small enough in terms of ε, the contribution from the remaining characters is

$$\begin{aligned}
&\ll \frac{1}{q}\sum_{\chi\neq\chi_0} |\widehat{w^-}(\chi)||\widehat{V}(\chi)|^8 \\
&\ll \frac{1}{q}\sup_{\chi\neq\chi_0}|\widehat{w^-}(\chi)|\sum_{\chi}|\widehat{V}(\chi)|^8 \\
&\ll q^{\frac{5}{4}+\varepsilon-c\varepsilon^2},
\end{aligned}$$

for some positive constant c, by the Burgess bound as used in (1.5).

If q is large enough in terms of ε (and in particular large enough in terms of ξ and δ, which depend on ε), the contribution from the principal character dominates the contribution from the remaining characters, and so $W(a) > 0$.

As already noted, this proves part (2) of Theorem 1.1.6. □

We finish this chapter with a proof of our final main theorem.

Proof of Theorem 1.1.7. We begin by detailing a simple argument that proves a weaker version of Theorem 1.1.7 part (1), namely that, taking η to be $q^{-\frac{1}{16}+\varepsilon}$, if q is large enough in terms of ε then $P_\eta^{48} = (\mathbb{Z}/q\mathbb{Z})$.

We will apply Lemma 0.2.7. Let G be an abelian group, written multiplicatively, and let A be a finite subset of G. Suppose that A is not contained within any proper coset of G. Let us investigate the second case of Lemma 0.2.7 in more detail. If $A \cdot A^{-1} = G$ it follows from the Ruzsa triangle inequality and its corollary (see Lemma 0.2.8 and Corollary 0.2.9) that $|A \cdot A| \geqslant \left(\frac{|G|}{|A|}\right)^{\frac{1}{2}}|A|$. Hence $|A \cdot A| \geqslant \frac{3}{2}|A|$,

provided that $|A| \leqslant \frac{2^2}{3^2}|G|$. In the case when $\frac{2^2}{3^2}|G| \leqslant |A| \leqslant \frac{1}{2}|G|$, we have the estimate $|A \cdot A| \geqslant \sqrt{2}|A|$. In particular we have $|A \cdot A| > \frac{1}{2}|G|$, and hence[15] $A^4 = G$. We summarise these observations thus:

Corollary 1.4.4. *Let G be an abelian group, written multiplicatively, and let A be a finite subset of G. Suppose that A is not contained in any proper coset of G. Then:*

$$\begin{cases} |A \cdot A| \geqslant \frac{3}{2}|A| & \text{if } |A| \leqslant \frac{4}{9}|G| \\ |A \cdot A| \geqslant \sqrt{2}|A| & \text{if } \frac{4}{9}|G| \leqslant |A| \leqslant \frac{1}{2}|G| \\ |A \cdot A| = |G| & \text{otherwise.} \end{cases}$$

In particular, if $|A| > \frac{1}{3}|G|$ then $A^4 = G$.

We let G be $(\mathbb{Z}/q\mathbb{Z})^\times$. To prove the weaker version of part (1) of Theorem 1.1.7, write $\eta = q^{-\frac{1}{16}+\varepsilon}$ and apply Corollary 1.4.4 iteratively starting with $A = P_\eta^6$. Lemma 1.3.1 ensures that the hypotheses of Corollary 1.4.4 are satisfied. By Corollary 1.4.3,

$$G = P_\eta^3 \cup P_\eta^4 \cup P_\eta^5 \cup P_\eta^6. \tag{1.23}$$

If k and k' are natural numbers and $k > k'$, it is clear that $|P_\eta^k| \geqslant |P_\eta^{k'}|$. Therefore we may conclude that $|P_\eta^6| \geqslant \frac{1}{4}|G|$. Since $\frac{1}{4} > \frac{2}{3} \times \frac{2}{3} \times \frac{1}{2}$ we may apply Corollary 1.4.4 three times, to conclude that $A^8 = G$. In other words, $P_\eta^{48} = (\mathbb{Z}/q\mathbb{Z})^\times$. (Recall that the claim of Theorem 1.1.7 part (1) is that $P_\eta^{20} = (\mathbb{Z}/q\mathbb{Z})^\times$.)

The above weak argument is enough to prove part (2) of Theorem 1.1.7 in full. Indeed, let η equal $q^{-\frac{3}{4}+\varepsilon}$. Note from part (2) of Theorem 1.1.6 that there exists some natural number r at most $k(\varepsilon)$ such that $|P_\eta^r| \geqslant k(\varepsilon)^{-1}(q-1)$. Iterating Corollary 1.4.4, starting with $A = P_\eta^{k(\varepsilon)}$, gives the result (noting that Lemma 1.3.1 implies that

[15] by the usual pigeonhole argument which shows that $B \cdot B = G$ if $|B| > \frac{1}{2}|G|$.

the hypotheses of Corollary 1.4.4 hold at each iteration).

Now let us introduce a sharper argument, which will enable us to prove part (1) of Theorem 1.1.7 as it is stated. We will replace the use of Lemma 1.3.1 by the use of the stronger Lemma 1.3.5, and this strength will be focused through Kneser's Theorem (Theorem 0.2.10).

Fix δ to be a positive constant, independent of all the other parameters, and chosen to be sufficiently small.[16] We assume that q is sufficiently large in terms of δ and ε.

We begin with a general consideration. Let $H \leqslant (\mathbb{Z}/q\mathbb{Z})^\times$ be any proper subgroup satisfying $|H| \geqslant q^{1-\delta}$. Let $d := (q-1)/|H|$. By considering a character χ with kernel equal to H, from Lemma 1.3.5 we conclude[17] that P_η contains elements in at least $(1 - o_\delta(1))(\frac{11}{30} - O(\sqrt{\delta}))d$ cosets of H. If δ is small enough, this implies that P_η contains elements in at least $\frac{21}{60}d$ cosets of H.

We know that $P_\eta \cdot H$ is not contained in a proper subgroup of $(\mathbb{Z}/q\mathbb{Z})^\times$, by Lemma 1.3.1. Hence, by Corollary 1.4.4, we conclude that $P_\eta^4 \cdot H = (P_\eta \cdot H)^4 = (\mathbb{Z}/q\mathbb{Z})^\times$.

Now we come to specifics. Let H be the stabiliser of P_η^{10}. If $|H| \geqslant q^{1-\delta}$, then we may proceed very crudely, combining the observations above with Kneser's Theorem (Theorem 0.2.10) to yield

$$\begin{aligned} |P_\eta^{10}| &\geqslant 2|P_\eta^5 \cdot H| - |H| \\ &\geqslant 2(q-1) - (q-1) \\ &= q - 1. \end{aligned}$$

So $P_\eta^{10} = (\mathbb{Z}/q\mathbb{Z})^\times$, which certainly implies part (1) of Theorem 1.1.7.

[16] A detailed analysis shows that $\delta = 10^{-6}$ suffices for use in the argument to follow.

[17] Remember that here we are taking $\eta = q^{-\frac{1}{16}+\varepsilon}$, whereas Lemma 1.3.5 was stated for $\eta = q^{-\frac{3}{4}+\varepsilon}$.

Alternatively, $|H| \leqslant q^{1-\delta}$. Then, for $i \geqslant 1$, let $\alpha_i = |P_\eta^i|/(q-1)$. Kneser's Theorem implies that

$$\alpha_{10} \geqslant \max(\alpha_4 + \alpha_6, 2\alpha_5) - O(q^{-\delta}). \tag{1.24}$$

Since q is large enough in terms of ε, from equation (1.23) and the following remarks we have $\alpha_4 + \alpha_6 \geqslant \alpha_3 + \alpha_5$ and $\alpha_3 + \alpha_4 + \alpha_5 + \alpha_6 \geqslant 1$. Hence $\alpha_4 + \alpha_6 \geqslant \frac{1}{2}$, and so $\alpha_{10} \geqslant \frac{1}{2} - O(q^{-\delta})$.

But in fact a stronger statement holds, namely that $\alpha_{10} > \frac{1}{2}$. Indeed, let ρ be a fixed small positive constant to be chosen later[18] (independent of all other parameters) and assume that q is large enough in terms of ρ. If either $\alpha_4+\alpha_6 > \frac{1}{2}+\rho$ or $\alpha_5 > \frac{1}{4}+\rho$, then (1.24) implies that $\alpha_{10} > \frac{1}{2}$ (provided q is large enough in terms of ρ and δ).

Otherwise, $\alpha_5 \leqslant \frac{1}{4} + \rho$ and $\alpha_6 \leqslant \alpha_5 + 2\rho$. (Indeed, if $\alpha_6 > \alpha_5 + 2\rho$ then $\alpha_4+\alpha_6 > \alpha_3+\alpha_5+2\rho$, so $2(\alpha_4+\alpha_6) > 1+2\rho$, which is the previous case). Therefore $\alpha_6 \leqslant \frac{1}{4}+3\rho$. Since the stabilisers of different P_η^i are nested, we know that the stabiliser of P_η^6 has size at most $q^{1-\delta}$. Then from Kneser's Theorem we have $\alpha_6 \geqslant 2\alpha_3 - O(q^{-\delta})$, and hence $\alpha_3 < \frac{1}{8} + 2\rho$ (since q is large enough). If ρ is a small enough constant, this implies that $\alpha_3+\alpha_4+\alpha_5+\alpha_6 \leqslant \frac{7}{8}+11\rho < 1$. This is a contradiction to equation (1.23).

Therefore $\alpha_{10} > \frac{1}{2}$, and hence $\alpha_{20} = 1$. This proves Part (1) of Theorem 1.1.7 to the desired strength. □

We conclude this chapter by remarking that since Theorem 1.1.6 establishes that P_η^6 is very large, the proof of the weaker form of Theorem 1.1.7 part (1) did not in fact require the full generality of Lemma 1.3.1. Elliott's bounds [21] for the least prime quadratic, cubic, and quartic residues suffice in this case.

[18] $\rho = 10^{-3}$ suffices

Chapter 2

An explicit version of Linnik's Theorem

2.1 Introduction

Throughout this chapter, φ will denote the Euler φ function, and ζ will denote the Riemann zeta function.

The main difficulties in the previous chapter stemmed from a single source, that the primes under consideration were less than the modulus q. If this constraint is slackened, other theorems from sieve theory start to become powerful tools. In this short chapter we use the Brun-Titchmarsh Theorem, in combination with a variety of explicit estimates and additive combinatorial techniques, to conclude that, in the notation of the previous chapter, $P_\eta^3 = (\mathbb{Z}/q\mathbb{Z})^\times$ for all[1] $q \geqslant 2$, where $\eta = q^{\frac{13}{3}}$. The results that we prove in this chapter are the product of joint work with Prof. Olivier Ramaré, and are mostly contained in [71].

[1] q is no longer restricted to be prime.

The regime of primes less than q^L (for some $L > 1$) is the setting of the classical Linnik's Theorem.

Theorem 2.1.1 (Linnik [53], [54])**.** *There exists an effective absolute constant L such that, for all $q \geqslant 2$ and for all residues a co-prime to q, there exists a prime p at most q^L such that $p \equiv a (mod\ q)$.*

Over the years a string of explicit versions of this theorem have been proved, with the best bounds currently being due to Xylouris [94]. He states that $L = 5$ is permissible, provided that q is sufficiently large. Under the Generalised Riemann Hypothesis the bound of q^5 may be reduced[2] slightly to $\varphi(q)^2(\log q)^2$.

Xylouris' proof relies on intricate estimates of zero-free regions of Dirichlet L-functions, and though his bound for 'sufficiently large' is effective, no explicit version has been shown. Our main result indicates that one can, by rather different methods, access a fully explicit result with respectable constants, provided one replaces 'primes' by 'products of three primes'.

Theorem 2.1.2. *If q is a natural number at least 2, then for all residues a co-prime to q there exists a number b such that $b \leqslant q^{16}$, b is congruent to a modulo q, and b is equal to the product of exactly three primes. Furthermore, each of these primes may be taken to be at most $q^{\frac{16}{3}}$.*

Trivially, this result admits the following reformulation.

Theorem 2.1.3. *Let x be a positive real number and suppose q is a natural number satisfying $2 \leqslant q \leq x^{1/16}$. Then for any invertible residue class a modulo q, there exists a product of three primes, all of which are below $x^{1/3}$, that is congruent to a modulo q.*

[2]This is a remark from the first page of [46].

The parameter x will be useful throughout, and so it is this formulation that we will actually prove.

It goes without saying that Theorem 2.1.3 is of a qualitatively weaker form than the result of Xylouris, but, as already remarked, qualitative control is not our main concern. Consider the following quantitative result concerning primes in arithmetic progressions:[3]

Theorem 2.1.4 (Kadiri, [50], as quoted by Ramaré)**.** *If q is a natural number such that $q \geqslant 10^{30}$ and q is non-exceptional*[4]*, and a is a residue co-prime to q, then there exists a prime p at most $3q^{5\log q}$ such that p is congruent to a modulo q.*

It was certainly not immediately obvious to us, before embarking upon the proof of Theorem 2.1.2, that passing from one prime to products of exactly three primes could yield such a strong quantitative improvement over results such as Theorem 2.1.4.

Our treatment is certainly not completely optimal[5], but we have instead sought the simplest possible argument. The main surprise is that we use sieve techniques, but are not blocked by the parity principle. A lower bound for $L(1,\chi)$ is employed, but certainly we do not rely on Siegel's Theorem, and in particular our bound is not strong enough to push a possible Siegel zero away from 1 (see [25] for more on this issue).

[3]Theorem 2.1.4 is a result that is oft-quoted by Ramaré, though we have found it rather difficult to extract this precise result from [50].

[4]i.e. no character modulo q has an exceptional Siegel zero, for some explicit quantitative notion of Siegel zero (that we have found rather hard to extract from [50])

[5]Indeed, Ramaré and Srivastav (personal communication) have adapted the methods, improving the bound $q \leqslant x^{1/16}$ in Theorem 2.1.3 to $900q \leqslant x^{1/9}$.

2.2 Lemmas

Throughout this section, let q be as in the statement of Theorem 2.1.3. We begin with some crude bounds. Let us define

$$f_0(q) := \prod_{p|q} \left(1 - \frac{1}{\sqrt{p}}\right)^{-1}. \tag{2.1}$$

Lemma 2.2.1. *We have* $f_0(q) \leq 3.32\sqrt{q}$.

Proof. For all primes p we have $\left(1 - \frac{1}{\sqrt{p}}\right)^{-1} \leqslant \alpha_p \sqrt{p}$, where

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{2}-1} & p = 2 \\ \frac{1}{\sqrt{3}-1} & p = 3 \\ 1 & \text{otherwise,} \end{cases}$$

and since $\alpha_2 \leqslant 2.42$ and $\alpha_3 \leqslant 1.37$, we obtain the inequalities

$$f_0(q) \leqslant 2.42 \cdot 1.37 \cdot \sqrt{q} \leqslant 3.32 \cdot \sqrt{q}. \qquad \square$$

We will also require a rudimentary estimate on $\varphi(q)$.

Lemma 2.2.2. *If* $q \geqslant 31$, *then* $\varphi(q) > 8$.

Proof. Recall that

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Therefore if $\varphi(q) \leqslant 8$, the only prime factors of q are $2, 3, 5, 7$. By performing an easy case analysis on which of these primes divides q, one sees that the only q for which $\varphi(q) \leqslant 8$ are $1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 16, 20, 24$, and 30. $\square$

Now we move on to proving bounds on quantities involving non-principal Dirichlet characters modulo q. The classic reference for these objects is Davenport [16].

In the previous chapter, we used both the Pólya-Vinogradov inequality and the Burgess bound for sums of Dirichlet characters over an interval. Here, in keeping with the elementary feel of the enterprise, we will settle for the following trivial bound.

Lemma 2.2.3. *Let χ be a non-principal Dirichlet character modulo q. Let I be a subset of $\{1, \cdots, q\}$. We have*

$$\left| \sum_{n \in I} \chi(n) \right| \leqslant \varphi(q)/2.$$

The same bound holds true for any finite interval instead of I.

Proof. Trivial from the definitions. □

As mentioned in the introduction, we will require an effective lower bound on $L(1, \chi)$ when χ is a non-principal quadratic character. The most direct argument is to use the Dirichlet class number formula .

Lemma 2.2.4. *Let χ be a non-principal quadratic character modulo q, and $L(s, \chi)$ the corresponding Dirichlet L-function. We have*

$$L(1, \chi) \geqslant 0.3 \frac{1}{\sqrt{q}}.$$

Proof. Let q^* be the conductor of χ and let χ^* be the associated primitive character. By comparing the Euler products, one gets

$$L(1, \chi) = L(1, \chi^*) \prod_{\substack{p|q, \\ p \nmid q^*}} \left(1 - \frac{\chi^*(p)}{p}\right) \geqslant L(1, \chi^*) \prod_{\substack{p|q, \\ p \nmid q^*}} \left(1 - \frac{1}{p}\right)$$

The Dirichlet class number formula (see for instance equations (15) and (16) of [16,

Chapter 6]) states that

$$L(1,\chi^*) = \begin{cases} \frac{2\pi h(d)}{w|d|^{\frac{1}{2}}} & \text{for } d<0 \\ \frac{h(d)\log\varepsilon}{d^{\frac{1}{2}}} & \text{for } d>0, \end{cases}$$

where $d = \pm q^*$ is the fundamental discriminant for which $\chi^*(n) = (\frac{d}{n})$, $h(d)$ is the number of equivalence classes of binary quadratic forms with discriminant d, w is the number of automorphs of binary quadratic forms with discriminant d, and $\varepsilon = \frac{1}{2}(t_0 + u_0\sqrt{d})$, where (t_0, u_0) is the solution of the Pell equation $t^2 - du^2 = 4$ in positive integers for which u_0 is least.

We now split into two cases. If $d < 0$, since the number of automorphs is at most 6 (see equation (3) of [16, Chapter 6]), we have $L(1,\chi^*) \geqslant 1/\sqrt{q^*}$, and hence

$$\begin{aligned} \sqrt{q}L(1,\chi) &\geqslant \prod_{\substack{p|q,\\ p\nmid q^*}} \left(1-\frac{1}{p}\right)\frac{\sqrt{q}}{\sqrt{q^*}} \\ &\geqslant \prod_{\substack{p|q,\\ p\nmid q^*}} \left(p^{\frac{1}{2}} - p^{-\frac{1}{2}}\right) \\ &\geqslant 2^{\frac{1}{2}} - 2^{-\frac{1}{2}}, \end{aligned}$$

which, absurdly crudely, is at least 0.3.

If $d > 0$, i.e. $d = q^*$, we need some bounds on the fundamental unit. Proceeding extremely crudely, one has $\varepsilon^2 \geqslant q^*/2$ and, since q^* is at least 5, this yields $\log\varepsilon \geqslant 0.458$.

We repeat the estimation from the previous case, leading to the bound

$$\sqrt{q}\,L(1,\chi) \geqslant (2^{\frac{1}{2}} - 2^{-\frac{1}{2}})0.458,$$

which is also at least 0.3. □

Using the class number formula and bounds on fundamental units seems rather out of the main spirit of this chapter. For the final version of our work [71], Ramaré found a modification of an idea of Gel'fond from [31][6] that provides a lower bound without resorting to such theorems. The motivation for this argument runs as follows[7]. The underlying reason that $L(1,\chi)$ is non-zero is that $\zeta(s)L(s,\chi)$ is the Dedekind zeta function of a quadratic number field. Both this zeta function and $\zeta(s)$ have a simple pole at $s=1$, and hence $L(s,\chi)$ is non-zero. This immediately suggests that, if one is seeking a quantitative argument, one should consider the arithmetic function $1\star\chi$ (where $\star$ denotes Dirichlet convolution).[8]

Lemma 2.2.5 (Ramaré)**.** *Let χ be a non-principal quadratic character modulo q, and $L(s,\chi)$ the corresponding Dirichlet L-function. Then*

$$L(1,\chi) \geq \frac{\pi}{4\varphi(q)} - \frac{\pi}{\varphi(q)^2}.$$

Proof. One considers the sum $S(\alpha) = \sum_{n\geqslant 1}(1\star\chi)(n)e^{-n\alpha}$ for real positive α. Since $(1\star\chi)(m^2) \geqslant 1$ for every integer m, and $(1\star\chi)(n) \geqslant 0$ in general, one has

$$1 + S(\alpha) \geqslant \sum_{m\geqslant 0} e^{-m^2\alpha} \geqslant \int_0^\infty e^{-\alpha t^2}\,dt = \frac{\Gamma(1/2)}{2\sqrt{\alpha}} = \frac{\sqrt{\pi}}{2\sqrt{\alpha}}.$$

On the other hand we can expand the definition $(1\star\chi)(n) = \sum_{d|n}\chi(d)$ and get

$$S(\alpha) = \sum_{d\geqslant 1}\frac{\chi(d)}{e^{\alpha d}-1} = \frac{L(1,\chi)}{\alpha} - \sum_{d\geqslant 1}\chi(d)g(\alpha d)$$

by using the non-negative non-increasing function $g(x) = \frac{1}{x} - \frac{1}{e^x-1}$. We find that, by

[6] Ramaré (personal communication) feels that [32] is the reference that is easiest to read.

[7] This was explained to us by Ben Green.

[8] Adam Harper informs us that the idea of using this manner of convolution argument to prove non-vanishing of $L(1,\chi)$ goes back at least as far as Mertens.

Lemma 2.2.3,

$$\begin{aligned}\sum_{d\geqslant 1}\chi(d)g(\alpha d) &= -\sum_{d\geqslant 1}\chi(d)\int_{\alpha d}^{\infty} g'(t)\,dt\\ &= -\int_0^{\infty}\sum_{d\leq t/\alpha}\chi(d)g'(t)\,dt\\ &\geq \frac{\varphi(q)}{2}\int_0^{\infty} g'(t)\,dt\\ &= -\varphi(q)/4\end{aligned}$$

since

$$\lim_{x\to 0^+} g(x) = 1/2.$$

By comparing both upper and lower estimate for $S(\alpha)$, we reach

$$L(1,\chi) \geq \frac{\sqrt{\pi\alpha}}{2} - \alpha - \frac{\alpha\varphi(q)}{4}.$$

We select $\alpha = \pi/\varphi(q)^2$. The lemma then follows. □

It is this bound on $L(1,\chi)$ that we shall use in the rest of the chapter. It will be used in the proof of the following lemma, which is an elementary method for finding primes p such that $\chi(p) = 1$. We adapt the proof of J. Pintz taken from [67].

Lemma 2.2.6. *Let q be a natural number at least 3, χ be a non-principal quadratic character modulo q. Then there is a prime p at most q^4 such that $\chi(p) = 1$.*

We remark that much stronger results are known, and indeed the cited paper of Pintz [67] shows that such a prime may be taken at most $q^{\frac{1}{4}+\varepsilon}$. It was this result which we mentioned in connection with parity breaking in the previous chapter. However, Pintz's proof uses the bound on $L(1,\chi)$ that comes from Siegel's theorem. The point here, as with all of this chapter, is that a respectable bound may be achieved much more easily.

Proof. Assume that all primes less than x satisfy $\chi(p) \neq 1$. We use the notation $d|q^\infty$ to say that all the prime factors of d divide q. Then on the one hand we have

$$\sum_{n\leqslant x}(1\star\chi)(n) = \sum_{n\leqslant x}\prod_{p^\alpha||n}(1+\chi(p)+\cdots+\chi^\alpha(p)) = \sum_{d|q^\infty}\sum_{\substack{m^2\leqslant x/d,\\(m,q)=1}} 1 \leqslant \sum_{d|q^\infty}\sqrt{\frac{x}{d}} \leqslant \sqrt{x}f_0(q)$$

where f_0 is the function defined in (2.1). On the the other hand, we can approximate this sum by $L(1,\chi)$ as follows:

$$\sum_{n\leqslant x}(1\star\chi)(n) = \sum_{d\leqslant x}\chi(d)\left[\frac{x}{d}\right] = x\sum_{d\leqslant x}\frac{\chi(d)}{d} - \sum_{d\leqslant x}\chi(d)\left\{\frac{x}{d}\right\}.$$

The first summation over d is an approximation of $L(1,\chi)$ (recall Lemma 2.2.3):

$$\begin{aligned} L(1,\chi) &= \sum_{d\leq x}\frac{\chi(d)}{d} + \int_x^\infty \sum_{x<d\leq t}\chi(d)dt/t^2 \\ &= \sum_{d\leq x}\frac{\chi(d)}{d} + O\left(\frac{\varphi(q)}{2x}\right), \end{aligned}$$

where the implied constant in the error term may be taken to be 1.

We treat the second summation in d above by Axer's method from [5] (see also [60, Theorem 8.1]):

$$\begin{aligned} \left|\sum_{d\leq x}\chi(d)\left\{\frac{x}{d}\right\}\right| &\leqslant \sum_{d\leqslant y}1 + \sum_{m\leqslant x/y}\left|\sum_{d:[x/d]=m}\chi(d)\left\{\frac{x}{d}\right\}\right| \\ &\leqslant y + \frac{\varphi(q)x}{2y} \\ &\leqslant \sqrt{2\varphi(q)x} \end{aligned}$$

by selecting $y = \sqrt{\varphi(q)x/2}$, the second inequality following by Abel summation. All

of this implies that

$$\sqrt{x}L(1,\chi) \leq f_0(q) + \sqrt{2\varphi(q)} + \varphi(q)/(2\sqrt{x}).$$

However, the previous lemma gives us a lower bound for $L(1,\chi)$, and therefore

$$\frac{\pi}{4\varphi(q)} - \frac{\pi}{\varphi(q)^2} \leq \frac{f_0(q)}{\sqrt{x}} + \sqrt{\frac{2\varphi(q)}{x}} + \frac{\varphi(q)}{2x} \tag{2.2}$$

We substitute $x = q^4$, and use the upper bound for $f_0(q)$ provided by Lemma 2.2.1 to try to derive a contradiction. Assume first that $q \geqslant 31$. Then replace the left-hand side of equation (2.2) by $\pi/(8\varphi(q))$: this is permissible by Lemma 2.2.2. Applying the bound $\varphi(q) \leq q$, one infers that

$$\frac{\pi}{8} \leqslant \frac{3.32}{\sqrt{q}} + \frac{\sqrt{2}}{\sqrt{q}} + \frac{1}{2q^2}.$$

After a short calculation we derive a contradiction for all $q \geq 146$.

For the remaining q it is easy enough to find primes $p \leqslant q^4$ such that $p \equiv 1$ modulo q: these may be found in the table at the end of the chapter. □

The final ingredients in the argument are three standard results. The first is a strong form of the Brun-Titchmarsh inequality, proved by Montgomery and Vaughan in [59].

Lemma 2.2.7. *For a natural number q at most x, we have*

$$\sum_{\substack{y<p\leqslant y+x,\\ p\equiv a(\bmod q)}} 1 \leqslant \frac{2x}{\varphi(q)\log(x/q)}.$$

for any positive y.

The second is an explicit lower bound in the prime number theorem.

Lemma 2.2.8. *We have $\pi(x) \geq x/(\log x - 1)$ when $x \geq 5\,393$. Furthermore, if $q < x$, the number of primes p at most x such that $(p, q) = 1$ is at least $x/\log x$, again when $x \geq 5\,393$.*

Proof. The first inequality is taken from Corollary 5.3 of [20]. For the second, we simply note that the number of prime factors of q is at most $(\log x)/\log 2$ and that

$$\frac{x}{\log x - 1} - \frac{\log x}{\log 2} \geq \frac{x}{\log x}$$

when $x \geq 5\,000$. □

The third is Kneser's Theorem, which we stated as Theorem 0.2.10.

2.3 Proof of Theorem 2.1.3

Let us first treat the case $x \geq 10^{16}$.

Let $X = x^{1/3}$. Since X is at least 10^5, Lemma 2.2.8 tells us that the number $\pi_q(X)$ of primes below X that are coprime to q is at least $X/\log X$. The Brun-Titchmarsh inequality in the form given by Montgomery and Vaughan, recalled in Lemma 2.2.7, tells us that the number of prime numbers less than X that are congruent to $a \mod q$, for a co-prime to q, is at most $\frac{32}{13}X/(\varphi(q)\log X)$. (Here we have used the inequality $q \leqslant X^{3/16}$). This implies, when compared to the total number of prime numbers that are coprime to q given by Lemma 2.2.8, that at least $\frac{13}{32}\varphi(q)$ such residue classes contain a prime that is at most X. Let us call this set of classes A.

As in the previous chapter, we will apply Kneser's Theorem (Lemma 0.2.10) to the group $G := (\mathbb{Z}/q\mathbb{Z})^\times$. Let H be the stabilizer of $A \cdot A$. We divide into cases according to the index of H.

If H is equal to G then, since $A \cdot A \cdot H = A \cdot A$, we have $A \cdot A = G$ and of course $A \cdot A \cdot A = G$.

If H has index 2, then it is the kernel of some quadratic character χ. Because A generates G multiplicatively, there is a point a in A such that $\chi(a) = -1$. By Lemma 2.2.6, there is another one, say a', such that $\chi(a') = 1$. Hence $A \cdot A$ also has a point b such that $\chi(b) = 1$ and one, say b', such that $\chi(b') = -1$. This implies that $A \cdot A \cdot H = G$, i.e. $A \cdot A = G$.

When H is of index 3, then $A \cdot H$ covers at least 2 H-cosets (since $\frac{13}{32} > \frac{1}{3}$) and is thus of cardinality at least $2\varphi(q)/3$. Kneser's Theorem ensures that $|A \cdot A| \geq \varphi(q)$, i.e. that again $A \cdot A = G$.

When H is of index 4, then $A \cdot H$ covers at least 2 H-cosets (since $\frac{13}{32} > \frac{1}{4}$) and is thus of cardinality at least $\varphi(q)/2$. By Kneser's Theorem,

$$|A \cdot A| \geq 2|A \cdot H| - |H| \geq \frac{3}{4}\varphi(q).$$

When H is of index Y say, with Y at least 5, let us write $|A|/\varphi(q) = 1/U$. The set $A \cdot H$ is made out of at least $\lceil Y/U \rceil$ cosets modulo H. Using the same manipulation as above, Kneser's Theorem ensures that $|A \cdot A|/\varphi(q) \geq (2\lceil Y/U \rceil - 1)/Y$. A quick computation shows that the minimum of $(2\lceil Y/U \rceil - 1)/Y$ as Y ranges over the set $\{5, 6, 7, 8, 9\}$ is attained at $Y = 7$ and has value $5/7$. When Y is larger than 10, we directly check that $(2\lceil Y/U \rceil - 1)/Y \geq \frac{2}{U} - \frac{1}{Y} \geq \frac{13}{16} - \frac{1}{10} \geq \frac{7}{10}$.

Combining these final two cases, we have proved in either instance that $|A \cdot A| \geq \frac{7}{10}\varphi(q)$. Since $\frac{13}{32}$ is greater than $\frac{3}{10}$, we have that $|G| - |A \cdot A|$ is less than $|A|$. Then, by the usual pigeonhole argument, we have $A \cdot A \cdot A = G$.

It remains to deal with $x < 10^{16}$, which is done by explicit calculation. The inclusion of this addendum was kindly suggested to us by an anonymous referee. Indeed, when $x < 10^{16}$, the modulus q is restricted to be not more than 10, implying

that only a limited number of congruence classes are to be looked at. We proceed by hand, verifying the threshold K for which $x \geqslant K$ implies $A = G$ (so certainly $A \cdot A \cdot A = G$).

When $q = 2$, we only need $x \geq 3^3$.

When $q = 3$, we only need $x \geq 7^3$.

When $q = 4$, we only need $x \geq 5^3$.

When $q = 5$, we only need $x \geq 19^3$.

When $q = 6$, we only need $x \geq 11^3$.

When $q = 7$, we only need $x \geq 29^3$.

When $q = 8$, we only need $x \geq 23^3$.

When $q = 9$, we only need $x \geq 23^3$.

When $q = 10$, we only need $x \geq 19^3$.

This takes care of the situation when $x \geq 29^3$. However, when x is below 29^3, the bound $x^{1/16}$ is less than 2, so the statement of the theorem is degenerate. This ends the proof. $\square$

Table 2.1: Primes p at most q^4 satisfying $p \equiv 1 \bmod q$

q	**p**	**q**	**p**	**q**	**p**	**q**	**p**
2	3	21	43	40	41	59	709
3	7	22	23	41	83	60	61
4	5	23	47	42	43	61	367
5	11	24	73	43	173	62	311
6	7	25	101	44	89	63	127
7	29	26	53	45	181	64	193
8	17	27	109	46	47	65	131
9	19	28	29	47	283	66	67
10	11	29	59	48	97	67	269
11	23	30	31	49	197	68	137
12	13	31	311	50	101	69	139
13	53	32	97	51	103	70	71
14	29	33	67	52	53	71	569
15	31	34	103	53	107	72	73
16	17	35	71	54	109	73	293
17	103	36	37	55	331	74	149
18	19	37	149	56	113	75	151
19	191	38	191	57	229	76	229
20	41	39	79	58	59	77	463

Table 2.2: Primes p at most q^4 satisfying $p \equiv 1 \bmod q$

q	**p**	**q**	**p**	**q**	**p**	**q**	**p**
78	79	97	389	116	233	135	271
79	317	98	197	117	937	136	137
80	241	99	199	118	709	137	823
81	163	100	101	119	239	138	139
82	83	101	607	120	241	139	557
83	167	102	103	121	727	140	281
84	337	103	619	122	367	141	283
85	1021	104	313	123	739	142	569
86	173	105	211	124	373	143	859
87	349	106	107	125	251	144	433
88	89	107	643	126	127	145	1451
89	179	108	109	127	509	146	293
90	181	109	1091	128	257		
91	547	110	331	129	1033		
92	277	111	223	130	131		
93	373	112	113	131	263		
94	283	113	227	132	397		
95	191	114	229	133	1597		
96	97	115	461	134	269		

Chapter 3

The primes are not metric poissonian

3.1 Introduction

The next two chapters of this thesis have a different focus. However, the general philosophy – the fruitful combination of techniques and notions from additive combinatorics and from analytic number theory – remains constant.

Let $\mathcal{A}$ be an infinite increasing sequence of natural numbers, and let A_N denote the first N elements of $\mathcal{A}$. For $\alpha \in [0,1]$, we consider the sequence $\alpha\mathcal{A}$, taken modulo 1. Recall that the sequence $\alpha\mathcal{A}$ is said to be *equidistributed* in $\mathbb{R}/\mathbb{Z}$ if for every interval $I \subset \mathbb{R}/\mathbb{Z}$ one has

$$\lim_{N\to\infty} \frac{1}{N} \sum_{x\in A_N} 1_I(\alpha x) = |I|. \tag{3.1}$$

For many arithmetic sequences $\mathcal{A}$ of interest, the sequence $\alpha\mathcal{A}$ is equidistributed in $\mathbb{R}/\mathbb{Z}$ for all irrational α. This is true for $\mathcal{A} = \mathbb{N}$ itself, or more generally the set of k^{th} powers for any natural number k, and, most pertinently for us, the set of primes.

In this chapter we will consider a strictly stronger notion of equidistribution. With

notation as above, we define the pair correlation function

$$F(\mathcal{A}, \alpha, s, N) := \frac{1}{N} \sum_{\substack{x_i, x_j \in A_N \\ x_i \neq x_j}} 1_{[-s/N, s/N]}(\alpha(x_i - x_j)), \tag{3.2}$$

where both the interval $[-s/N, s/N]$ and the sequence $\alpha\mathcal{A}$ are considered modulo 1.

Informally, $F(\mathcal{A}, \alpha, s, N)$ counts the number of pairs $(\alpha x_i, \alpha x_j)$ such that the distance $\alpha x_i - \alpha x_j \bmod 1$ is approximately s times the average gap length of the sequence $\alpha A_N \bmod 1$. Analysing the behaviour of $F(\mathcal{A}, \alpha, s, N)$ for a specific α can require delicate diophantine information about α (see [47], [74]), but one may instead settle for results which hold for almost all α.

In the setting of (3.1), *any* $\mathcal{A}$ satisfies the equidistribution property for almost all α (the sharpest results in this direction are due to Baker [8]). However, in the setting of pair correlations, the situation is more subtle.

Definition 3.1.1 (Metric poissonian property)**.** *Let $\mathcal{A}$ be an infinite increasing sequence of natural numbers. We say that $\mathcal{A}$ is* metric poissonian[1] *if for almost all $\alpha \in [0, 1]$, and for all fixed positive s, we have*

$$F(\mathcal{A}, \alpha, s, N) = 2s(1 + o_{\mathcal{A},\alpha,s}(1)) \tag{3.3}$$

as $N \to \infty$.

Notice that if we had picked N i.i.d. random variables $(X_n)_{n \in [N]}$ uniformly distributed on $\mathbb{R}/\mathbb{Z}$, instead of the sequence $\alpha A_N \bmod 1$, then as N tends to infinity the equivalent pair correlation function would tend to $2s$ with high probability. Therefore (3.3) may be viewed as some strong indication that $\alpha\mathcal{A} \bmod 1$ exhibits the behaviour of a random sequence. The connection to the equidistribution property (3.1) was

[1]See the introduction to [10] for an explanation of this terminology.

recently made rigorous: indeed, three simultaneous papers [1, 41, 82] recently showed that if (3.3) holds, for some fixed α and for all s, then for the same α one has that $\alpha\mathcal{A}$ is equidistributed in $\mathbb{R}/\mathbb{Z}$.

One might expect that, for the classical sequences $\mathcal{A}$ where $\alpha\mathcal{A}$ is equidistributed for irrational α, one could prove that these sequences $\mathcal{A}$ are metric poissonian. Indeed, for k at least 2 and $\mathcal{A}$ the set of k^{th} powers, this was shown by Rudnick and Sarnak [73]. However, the sequence $\mathcal{A} = \mathbb{N}$ is *not* metric poissonian. This follows from a consideration of the continued fraction expansion of α, but is in fact a special case of a more general phenomenon, connected to the large additive energy[2] of the truncations A_N.

Why should the additive energy $E(A_N)$ be connected with the metric poissonian property? Consider an additive quadruple (b_1, b_2, b_3, b_4) satisfying $b_1 + b_2 = b_3 + b_4$. Obviously, if $\|\alpha(b_1 - b_3)\| \leqslant \frac{s}{N}$ then $\|\alpha(b_2 - b_4)\| \leqslant \frac{s}{N}$. This is extremely different behaviour to that which would be seen if αb_1, αb_2, αb_3, αb_4 were genuinely independent uniform random variables on $\mathbb{R}/\mathbb{Z}$.

Note the trivial bounds $N^2 \ll E(A_N) \leqslant N^3$. We have the following result of Bourgain, which shows that all sets of nearly maximal energy fail to have the metric poissonian property.

Theorem 3.1.2. *[3, Appendix] Let $\mathcal{A}$ be an infinite increasing sequence of natural numbers. Suppose*

$$\text{limsup}_{N\to\infty} \frac{E(A_N)}{N^3} > 0.$$

Then $\mathcal{A}$ is not metric poissonian.

It is clear that the sequence $\mathcal{A} = \mathbb{N}$ satisfies the hypotheses of this theorem, and

[2] see Definition 0.2.5.

therefore this sequence is not metric poissonian.

Remarkably, a near-converse to this theorem has also been proved to be true.

Theorem 3.1.3. *Let $\mathcal{A}$ be an infinite increasing sequence of natural numbers. Suppose that $E(A_N) \ll N^{3-\delta}$, for some fixed positive δ and for every natural number N. Then $\mathcal{A}$ is metric poissonian.*

This theorem first appears as stated[3] in the recent work of Aistleitner, Larcher and Lewko [3]. It immediately implies the theorem of Rudnick-Sarnak on k^{th} powers, and also earlier work on lacunary sequences [75].

It is natural to wonder whether there is a tight energy threshold[4] for this problem, and we will return to this issue in the next chapter. Although the case of general sets is still rather mysterious, it is certainly interesting to consider the behaviour of specific sets $\mathcal{A}$ that satisfy $N^{3-\varepsilon} \ll_\varepsilon E(A_N) \ll o(N^3)$ for all positive ε. In this chapter, we prove the following result (answering a question posed by Nair[5]).

Theorem 3.1.4. *The primes are not metric poissonian.*

When $\mathcal{A}$ is the set of primes one has $E(A_N) \asymp N^3(\log N)^{-1}$ (this follows from Theorem 3.2.2 below, and also from [38, Theorem 1.8]). So certainly the primes are not included in the range of applicability of either Theorem 3.1.2 or Theorem 3.1.3.

In [3], Bourgain constructs a sequence $\mathcal{A}$ that is not metric poissonian but nonetheless has $E(A_N) = o(N^3)$, thereby showing that the converse to Theorem 3.1.2 is false. A quantitative analysis of his argument shows that the bound $E(A_N) \ll_\varepsilon N^3(\log\log N)^{-\frac{1}{4}+\varepsilon}$ is achievable, for any positive ε. So, as an immediate corollary to Theorem 3.1.4, we have an improved bound[6] for the smallest energy $E(A_N)$ of the

[3]The same result may be deduced from Theorem 3.2 of Harman's earlier book [44], combined with the relevant modification of the variance estimate from page 69 of the same volume.

[4]A very recent pre-print [2] has shown that there is no tight threshold.

[5]at the ELAZ 2016 conference in Strobl.

[6]This bound has since been improved in [52].

initial segments of a set $\mathcal{A}$ which is not metric poissonian.

3.2 Proof of Theorem 3.1.4

The plan of the proof is as follows, following our paper [92]. For each fixed α, we will try to find infinitely many n such that $\|\alpha n\|$ is extremely small. Using such an n we will be able to construct a scale N and a small constant s such that $\|\alpha mn\| \leqslant s/N$ for some initial segment of integers m. By a variant of a well known result concerning the exceptional set for the Goldbach problem, we may show that many such mn are represented many times as $p_i - p_j$ for two primes p_i, p_j that are at most p_N (the N^{th} prime). Combining all these observations, we will conclude, provided s is small enough, that $F(\mathcal{P}, \alpha, s, N) \geqslant c$ for some constant c that satisfies $c > 2s$. Since this inequality holds for infinitely many N, we cannot have $F(\mathcal{P}, \alpha, s, N) = 2s(1+o_{\alpha,s}(1))$ for all almost all α and for all $s > 0$. In fact, we will show that, for almost all α, this asymptotic *fails* to hold.

We now begin to consider the details of this argument. We will use the following result of Harman on diophantine approximation ([44, Theorem 4.2], [43]).

Theorem 3.2.1. *Let $\psi(n)$ be a non-increasing function with $0 < \psi(n) \leqslant \frac{1}{2}$. Suppose that*

$$\sum_n \psi(n) = \infty.$$

Let $\mathcal{B}$ be an infinite set of integers, and let $S(\mathcal{B}, \alpha, N)$ denote the number of n at most N, with $n \in \mathcal{B}$, such that $\|n\alpha\| < \psi(n)$. Then for almost all α we have

$$S(\mathcal{B}, \alpha, N) = 2\Psi(N, \mathcal{B}) + O_\varepsilon(\Psi(N)^{\frac{1}{2}}(\log \Psi(N))^{2+\varepsilon}) \tag{3.4}$$

for all $\varepsilon > 0$, with implied constant uniform in α, where

$$\Psi(N) = \sum_{n \leqslant N} \psi(n)$$

and

$$\Psi(N, \mathcal{B}) = \sum_{\substack{n \leqslant N \\ n \in \mathcal{B}}} \psi(n).$$

This theorem may be thought of as a flexible version of Khintchines's theorem on Diophantine approximation, in which one can further pass to approximations coming from a set $\mathcal{B}$, provided the density of $\mathcal{B}$ is large enough. The quality of the error term in this theorem is much better than we need in our application, although it is important that there is no dependence on N except through $\Psi(N)$. Earlier results of this type include a $\log N$ factor in the error, which would not have been adequate.

The other technical tool will be the standard bound on the size of the exceptional set in a Goldbach-like problem.

Theorem 3.2.2. *Let X be a real number satisfying $X > 2$. For a natural number n at most X, define*

$$r(n) := \sum_{\substack{p_i, p_j \leqslant X \\ p_i - p_j = n}} \log p_i \log p_j.$$

Then for any positive B, for all but $O_B(\frac{X}{\log^B X})$ exceptional values of n we have the approximation

$$r(n) = \mathfrak{S}(n) J(n) + O_B\left(\frac{X}{\log^B X}\right), \tag{3.5}$$

where

$$\mathfrak{S}(n) := \begin{cases} 2 \prod_{p \geqslant 3} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{\substack{p \mid n \\ p \geqslant 3}} \frac{p-1}{p-2} & n \text{ even}, \\ 0 & n \text{ odd} \end{cases}$$

is the singular series, and

$$J(n) := (1_{[0,X]} * 1_{[-X,0]})(n).$$

Proof. This result follows by trivial modifications of the usual argument for the binary Goldbach problem, originally due (independently) to van der Corput, Estermann, and Čudakov. The clearest reference is [87, Chapter 3.2], or, for a more modern approach, one may consider the proof of Theorem 19.1 in [49, Chapter 19]. □

We combine these two key ingredients in the following proposition.

Proposition 3.2.3. *There exists a small absolute positive constant c, such that for almost all $\alpha \in [0,1]$, and for all fixed positive s, there exist infinitely many n satisfying:*

(1) $\|\alpha n\| < \frac{s}{n \log n}$;

(2) At least $c \log n$ of the numbers n, $2n$, $\cdots$, $\lfloor \frac{1}{10} \log n \rfloor n$ are expressible in at least $c\frac{n}{\log n}$ ways as the difference $p_1 - p_2$ of two primes p_1, p_2 in the range $p_1, p_2 \leqslant \frac{1}{2} n \log n$.

Proof of Proposition 3.2.3. Let c be a positive quantity to be specified later. With this c, let $\mathcal{B}$ be the set of natural numbers n that satisfy (2), and let $\psi(n) = \min(\frac{1}{2}, \frac{s}{n \log n})$. It is to this $\mathcal{B}$ and this ψ that we will apply Theorem 3.2.1. We claim that $\mathcal{B}$ is quite a dense set. More precisely, we claim that $\Psi(\mathcal{B}, N) \gg_s \Psi(N)$ for N large enough.

To prove this assertion, let K be a natural number and assume that K is sufficiently large. Let n and m be natural numbers restricted to the ranges $K \leqslant n < 2K$ and $1 \leqslant m \leqslant \lfloor \frac{1}{10} \log 2K \rfloor$. For notational convenience we let X denote the quantity $\frac{1}{2} K \log K$, and we consider Theorem 3.2.2 with this X. Since K is sufficiently large, we observe that $|nm| \leqslant \frac{1}{4} X$ for all n and m in these ranges.

We say that the pair (n, m) is *exceptional* if nm lies in the exceptional set from Theorem 3.2.2 for which the asymptotic formula (3.5) for $r(nm)$ fails to hold. [Note that certainly $nm \leqslant X$, so Theorem 3.2.2 applies in this setting.]

The map $(n, m) \mapsto nm$ is at most $\frac{1}{10}\log 2K$-to-1, due to the restricted range of m. Since the exceptional set from Theorem 3.2.2 has size at most $O(\frac{K}{\log^2 K})$, we conclude that there are at most $O(\frac{K}{\log K})$ exceptional pairs (n, m). So certainly there are at least $(1 - O(\log^{-1} K))K$ values of $n \in [K, 2K)$ such that the asymptotic formula for $r(nm)$ holds for all m that satisfy $m \leqslant \frac{1}{10}\log 2K$. Let D_K denote this set of n.

We have shown, therefore, that D_K is a large subset of $[K, 2K)$. We claim further that $D_K \subset \mathcal{B}$, provided we choose c small enough. Indeed, let us analyse the asymptotic formula for $r(nm)$. When nm is even, the singular series $\mathfrak{S}(nm)$ is always $\Omega(1)$. Furthermore, $(1_{[0,X]} * 1_{[-X,0]})(nm)$ is always $\Omega(X)$, since $|nm| \leqslant \frac{1}{4}X$. So, for all $n \in D_K$ and for all $m \leqslant \frac{1}{10}\log n$ such that nm is even, $r(nm)$ is always $\Omega(X)$.

Removing the log weights on the primes, and recalling the definition of X, in particular we notice that there is some small absolute constant c_1 such that the following holds: if $n \in D_K$ and if nm is even, there are at least $c_1 K/\log K$ pairs of primes (p_i, p_j) with $p_i, p_j \leqslant \frac{1}{2}K\log K$ such that $p_i - p_j = nm$. By the definition of $\mathcal{B}$, provided c is chosen smaller than $\min(c_1, \frac{1}{100})$, we have that $D_K \subset \mathcal{B}$.

We may now prove that $\Psi(\mathcal{B}, N) \gg_s \Psi(N)$ for N large enough. Indeed, for some integer k_0 that satisfies $k_0 = O_s(1)$,

$$\begin{aligned}\Psi(\mathcal{B}, N) &= \sum_{\substack{n \leqslant N \\ n \in \mathcal{B}}} \min(\frac{1}{2}, \frac{s}{n\log n}) \\ &\gg -O_s(1) + \sum_{k=k_0}^{\lfloor \log_2 N \rfloor - 1} \sum_{\substack{2^k \leqslant n < 2^{k+1} \\ n \in \mathcal{B}}} \frac{s}{2^k \log(2^k)}\end{aligned}$$

$$\gg -O_s(1) + \sum_{k=k_0}^{\lfloor \log_2 N \rfloor - 1} \sum_{2^k \leqslant n < 2^{k+1}} (1 - O(k^{-1})) \frac{s}{k2^k}$$

$$\gg -O_s(1) + \sum_{k=1}^{\lfloor \log N \rfloor} \frac{s}{k}$$

$$\gg_s \log\log N$$

$$\gg_s \Psi(N).$$

Therefore, applying Theorem 3.2.1 to this set $\mathcal{B}$ and this function ψ, the main term from the conclusion of Theorem 3.2.1 dominates the error term, and we conclude that for almost all α there are infinitely many $n \in \mathcal{B}$ satisfying $\|\alpha n\| < \frac{s}{n \log n}$. The proposition is proved. □

With this moderately technical proposition proved, the deduction of Theorem 3.1.4 is extremely short.

Proof of Theorem 3.1.4. Let $\Omega \subset [0,1]$ be the full-measure set of α for which Proposition 3.2.3 holds. Let c be the constant from Proposition 3.2.3, and fix some s satisfying $0 < 2s < c^2$. Let $\alpha \in \Omega$, and fix a large N to be one of the infinitely many natural numbers that satisfy the conclusions of Proposition 3.2.3.

By construction, we know that

$$\|\alpha N\| < \frac{s}{N \log N}.$$

Therefore, for all $d \leqslant \frac{1}{10} \log N$, we have

$$\|\alpha d N\| < \frac{s}{N}.$$

But by the second conclusion of Proposition 3.2.3, this implies that there are at least

$c^2 N$ pairs of distinct primes $p_i, p_j \leqslant \frac{1}{2} N \log N$ such that

$$\|\alpha(p_i - p_j)\| < \frac{s}{N}.$$

Since $P_N \sim N \log N$, and N is large, this certainly implies that

$$F(\mathcal{P}, \alpha, s, N) \geqslant c^2 > 2s.$$

This holds for infinitely many N, and therefore for all $\alpha \in \Omega$ we have

$$F(\mathcal{P}, \alpha, 2s, N) \neq 2s(1 + o(1))$$

as $N \to \infty$. Since Ω has measure 1, Theorem 3.1.4 is proved. $\square$

Chapter 4

On the threshold for the metric poissonian property

4.1 Introduction

The property of a set of natural numbers being metric poissonian seems intimately connected with the notion of additive energy. Though we briefly discussed this relationship in Chapter 3, we spend this chapter[1] investigating the phenomenon further. Choosing to focus on the more tractable setting of 'randomly generated' sequences, we manage to establish the metric poissonian property under weaker assumptions on the additive energy than those required in [3], albeit in this more restricted setting.

Let us define the class of sequences $\mathcal{A}$ that we will consider.

Definition 4.1.1. *Let $\psi : \mathbb{N} \longrightarrow (0, 1]$ be a non-increasing function. Then let $\mathcal{A}_\psi$ denote the random set formed by, for each natural number n, placing n in $\mathcal{A}_\psi$ independently at random with probability $\psi(n)$. For a natural number N, we let $A_{\psi,N}$ denote the (random) set consisting of the least N elements of $\mathcal{A}$.*

[1]The material presented here went on to become part of a collaboration [10].

We prove three main theorems in this chapter. The first two are easy to state:

Theorem 4.1.2. *Let $\psi = \min(1, (\log n)^{-1})$. Then, with probability 1, $\mathcal{A}_\psi$ is not metric poissonian.*

Theorem 4.1.3. *Let δ be positive, and let $\psi = \min(1, (\log n)^{-(1+\delta)})$. Then, with probability 1, $\mathcal{A}_\psi$ is metric poissonian.*

These two theorems nearly isolate the exact threshold for which the metric poissonian property holds for the random set $\mathcal{A}_\psi$. These theorems have been improved in our collaborative work [10] (the term $(\log n)^{-1-\delta}$ replaced by $(\log n)^{-1}(\log\log n)^{-2-\delta}$, for example), but in this thesis we choose to present our original single-author work.

Our third theorem shows that, despite the result of Theorem 4.1.2, when $\psi(n) = \min(1, (\log n)^{-1})$ the metric poissonian property is extremely close to holding for $\mathcal{A}_\psi$. In order to be able to state the theorem, we need to introduce an item of notation.

Recall the definition of $F(\mathcal{A}, \alpha, s, N)$ from the previous chapter, namely expression (3.2). In this random setting, it will be easier for us to truncate to intervals of the form $[1, X]$ rather than to the first N elements of $\mathcal{A}_\psi$. To that end, for a real number X, define

$$N := |\mathcal{A}_\psi \cap [1, X]| \tag{4.1}$$

and assume that $N \geqslant 1$. Then define

$$F_1(\mathcal{A}_\psi, \alpha, s, X) := F(\mathcal{A}_\psi, \alpha, s, N). \tag{4.2}$$

Suppose that, with probability 1, $N \to \infty$ as $X \to \infty$. Then it is clear that, with probability 1, $F(\mathcal{A}_\psi, \alpha, s, N) \to 2s$ as $N \to \infty$ if and only if $F_1(\mathcal{A}_\psi, \alpha, s, X) \to 2s$ as $X \to \infty$.

Now we may state our third result:

Theorem 4.1.4. *Let $\eta > 0$, and let $\psi = \min(1, (\log n)^{-1})$. For each natural number j, define $X_j := e^{j^{1+\eta}}$. Then, with probability 1, for almost all α and for all s one has*

$$F_1(\mathcal{A}_\psi, \alpha, s, X_j) = 2s(1 + o_{\mathcal{A}_\psi,\alpha,s}(1)),$$

as $j \to \infty$.

Notation: Whenever X and N appear in the same expression, it is understood that N is defined by (4.1): the same with X_j and N_j, etcetera.

4.2 Outline of proofs

Theorem 4.1.2 may be proved by the same methods as we used in Chapter 3. With probability 1, the random set $\mathcal{A}_\psi$ satisfies all the properties of the primes that we used there, the only difference being the absence of the singular series in Theorem 3.2.2. As far as this thesis is concerned, we leave the matter there: full details may be found in [10] and the proof of Theorem 1.6 therein.

For Theorems 4.1.3 and 4.1.4, we will proceed by adapting an argument of Schmidt, which we learned from Harman's book [44]. In this section we present an extremely broad overview of this approach, and of the difficulties encountered when adapting it to our setting. Our hope is to render the rather intricate estimates occurring towards the end of the chapter less confusing, though, of course, the reader might find the omission of details in this section to be confusing in itself.

In Schmidt's work, he was interested in estimating the following quantity:

$$G(g,\alpha,X)=\sum_{n\leqslant X}1_{[0,g(n)]}(\|\alpha n\|),$$

where $g(n)$ is any non-increasing positive function of n with $\sum g(n)$ divergent. One may write $F_1(\mathcal{A}_\psi,\alpha,s,X)$ in a form that is superficially similar to this, by noting that

$$F_1(\mathcal{A}_\psi,\alpha,s,X)=\sum_{n\in\mathbb{Z}}\left(\frac{R_{\psi,N}(n)}{N}\right)1_{[0,s/N]}(\|\alpha n\|), \tag{4.3}$$

with $R_{\psi,N}(n)$ (the 'representation function') being the number of ways n can be expressed as x_i-x_j with $x_i,x_j\in A_{\psi,N}$ and $x_i\neq x_j$.

Schmidt wished to show that, for almost all α,

$$G(g,\alpha,X)\sim 2\sum_{n\leqslant X}g(n) \tag{4.4}$$

as $X\to\infty$. There were two parts to Schmidt's argument. Initially he showed that, for all X, there exists a set $\Omega_X\subseteq[0,1]$ with the following properties: the measure of Ω_X is $o(1)$, and for all $\alpha\notin\Omega_X$ one has $|G(g,\alpha,X)-2\sum_{n\leqslant X}g(n)|=o(1)$. This worked by estimating the 'variance'

$$\int_0^1|G(g,\alpha,X)-2\sum_{n\leqslant X}g(n)|^2\,d\alpha, \tag{4.5}$$

after the clever manoeuvre of having eliminating certain[2] α from the integral. This part of his argument may be adapted to the framework of estimating $F_1(\mathcal{A}_\psi,\alpha,s,X)$ with very little modification, since the representation function $R_{\psi,N}$ is never too large. We give the details later in the chapter.

[2]As a heuristic, one removes the very well-approximable α. This procedure is not considered in [3], which is ultimately why we are able to prove stronger results in this chapter.

However, the second aspect of his argument (in which he deduced his desired asymptotic behaviour (4.4) from the above preliminary results), presents the would-be adapter with greater problems. Firstly, Schmidt picked an increasing sequence of scales $(X_j)_{j=1}^{\infty}$ that is sparse enough so that

$$\sum_j \text{meas}(\Omega_{X_j}) < \infty.$$

Then he used the first Borel-Cantelli lemma to conclude that, for almost all α,

$$|G(g, \alpha, X_j) - 2\sum_{n \leqslant X_j} g(n)| \leqslant E_j,$$

as $j \to \infty$, with some explicit and acceptable error term E_j. This idea, at least, may be exactly transplanted to the setting of $F_1(\mathcal{A}_\psi, \alpha, s, X)$. Yet his next observation, based on monotonicity of G, seems entirely inaccessible to us, namely the observation that if $X_j \leqslant X \leqslant X_{j+1}$ then one has

$$G(g, \alpha, X_j) - 2\sum_{n \leqslant X_{j+1}} g(n) \leqslant G(g, \alpha, X) - 2\sum_{n \leqslant X} g(n) \leqslant G(g, \alpha, X_{j+1}) - 2\sum_{n \leqslant X_j} g(n). \tag{4.6}$$

With this, provided $\sum_{n \leqslant X_j} g(n)$ is not too far from $\sum_{n \leqslant X_{j+1}} g(n)$ (which turns out to be an extremely weak condition on the speed of growth of X_j) Schmidt managed to conclude that

$$G(g, \alpha, X) \approx 2\sum_{n \leqslant X} g(n)$$

for almost all α and for every large X. Made rigorous, this gives the asymptotic (4.4).

Consider the setting of $F_1(\mathcal{A}_\psi, \alpha, s, X)$. It is certainly not obvious, or even necessarily true, that if $X < Y$, and $N := \mathcal{A}_\psi \cap [1, X]$ and $M := A_\psi \cap [1, Y]$, we have the inequality

$$NF_1(\mathcal{A}_\psi, \alpha, s, X) \leqslant MF_1(\mathcal{A}_\psi, \alpha, s, Y).$$

This is despite the fact that, replacing each side with the asymptotic sizes we predict, this inequality holds. The sum for $F_1(\mathcal{A}_\psi, \alpha, s, Y)$ contains more terms, but the 'Bohr-set condition' $1_{[0,s/M]}(\|\alpha(x_i - x_j)\|)$ is a stronger one for those pairs with $x_i, x_j \leqslant X$ than that which is imposed in the definition of $F_1(\mathcal{A}_\psi, \alpha, s, X)$. These are two conflicting influences.

The paper [3] suggests a possible fix[3]. By being careful with the dependence of the error terms on s, one may show a cancellation in the 'variance' expression (the analogous expression to (4.5) in the case of $F_1(\mathcal{A}_\psi, \alpha, s, X)$) that is uniform in s, provided that $s \asymp 1$. So it is possible to allow s to depend mildly on X. In particular, we will manage to show that

$$F_1(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j) \to 2s$$

as $j \to \infty$, provided that the ratio of $N_{j+1}/N_j \to 1$ as $j \to \infty$. This turns out to be enough to implement a sandwiching argument in the manner of (4.6). The need for the ratio of N_{j+1} to N_j to tend to 1 is why we require a density of $(\log n)^{-1-\delta}$ in Theorem 4.1.3 rather than a density closer to $(\log n)^{-1}$. The details of all this may be found in the main argument.

We finish the introduction with a long list of definitions of quantities that will be considered later in the chapter.

As described already, for integers n and natural numbers N we define

$$R_{\psi,N}(n) := \sum_{\substack{x_i, x_j \in A_{\psi,N} \\ x_i \neq x_j \\ x_i - x_j = n}} 1,$$

[3]This device is one of the interesting improvements of [3] over [73] (the foundational paper in the area).

and note that

$$F_1(\mathcal{A}_\psi, \alpha, s, X) = 2\sum_{n\in\mathbb{N}} \left(\frac{R_{\psi,N}(n)}{N}\right) 1_{[0,s/N]}(\|\alpha n\|). \tag{4.7}$$

If n is a natural number and s a positive real, we let $\mathcal{E}_n^X \in [0,1]$ be the set

$$\mathcal{E}_n^X = \bigcup_{\substack{m\leqslant n \\ (m,n)\leqslant s^2X^2/N^2+1}} \left(\frac{m-\frac{s}{N}}{n}, \frac{m+\frac{s}{N}}{n}\right).$$

Here, as is the standard notation, (m,n) denotes the greatest common divisor of m and n and $\left(\frac{m-\frac{s}{N}}{n}, \frac{m+\frac{s}{N}}{n}\right)$ denotes the real open interval with the given endpoints. (We hope the reader will forgive our dropping s from the notation). Then we define an auxiliary count

$$F_1^*(\mathcal{A}_\psi, \alpha, s, X) = 2\sum_{n\leqslant X} \left(\frac{R_{\psi,N}(n)}{N}\right) 1_{\mathcal{E}_n^X}(\alpha).$$

Roughly speaking, $F_1^*(\mathcal{A}_\psi, \alpha, s, X)$ is equal to $F_1(\mathcal{A}_\psi, \alpha, s, X)$ except on very small sets of α where $F_1(\mathcal{A}_\psi, \alpha, s, X)$ is particularly large.

The choice to raise X/N to the second power in the cut-off for $\mathcal{E}_n^X$ is not critical for the application (any power greater than 1 could be substituted). The addition of '+1' in the cut-off is also purely cosmetic, as soon we will restrict to a case in which the ratio X^2/N^2 tends to infinity. We make a technical note, pointing out that we take the cut-off s^2X^2/N^2+1 to be independent of n. This is due to the fact that the size of the approximation intervals $[0, s/N]$ depend only on N, and not on n as in Schmidt's setting. Though this change causes problems towards the end of the proof, as discussed regarding the sandwiching argument above, here it offers a great simplification, and we may dispense with several technical lemmas from Harman's account.

If n is a natural number, and k some positive real independent of n, we define

$$\varphi(k,n) = \sum_{\substack{m \leqslant n \\ (m,n) \leqslant k}} 1.$$

A particular instance of $\varphi(k,n)$ will be distinguished. If X is a positive real, let N be defined by (4.1). Then denote

$$\Phi^X(n) = \sum_{\substack{m \leqslant n \\ (m,n) \leqslant s^2X^2/N^2+1}} 1.$$

We also define

$$A^X(n,m) = \sum_{\substack{u \leqslant n \\ v \leqslant m \\ um-vn=0 \\ (u,n) \leqslant s^2X^2/N^2+1 \\ (v,m) \leqslant s^2X^2/N^2+1}} 1,$$

and

$$\tau^X(n) = \sum_{\substack{d \mid n \\ d \leqslant s^2X^2/N^2+1}} 1.$$

As a clarifying note, we observe that the condition $v \leqslant m$ in the definition $A^X(m,n)$ implies the condition $u \leqslant n$, since $\frac{vn}{m}$ is at most n automatically; despite this, we feel it is more enlightening to present the conditions in a symmetric form.

Remark 4.2.1. As discussed above, it will be necessary for us to have some control over estimates when s has a weak dependence upon N. For our purposes, it will always be the case that $s = s(N)$ will satisfy $s \asymp 1$, for some absolute implied constants.

4.3 Preparatory lemmas

We begin with some easy bounds on $R_{\psi,N}(n)$, coming from large deviation bounds. We had an early draft of this lemma in our personal notes, but the details of the $\mathcal{S}_0, \mathcal{S}_1$ splitting were worked out in collaboration with Bloom, Chow, and Gafni in [10].

Lemma 4.3.1. *Let $\psi : \mathbb{N} \longrightarrow (0,1]$ be a non-increasing function, and assume that $\sum_{x \geqslant 1} \psi(x)^2$ diverges. Let X be a positive real number, and let $N = |A_\psi \cap [1, X]|$ as always. Let ε be positive. Then there is a positive constant c_ε, depending only on ε, such that:*

(1) With probability greater than $1 - O(\exp(-c_\varepsilon \sum_{x \leqslant X} \psi(x)))$, we have

$$(1-\varepsilon)\sum_{x \leqslant X} \psi(x) \leqslant N \leqslant (1+\varepsilon)\sum_{x \leqslant X} \psi(x).$$

(2) Let

$$K_X := \sum_{X - \frac{1}{4}\sum_{y \leqslant X} \psi(y)^2 < x \leqslant X} \psi(x)^2.$$

Then with probability greater than $1 - O(X\exp(-c_\varepsilon K_X))$ we have, for all $n \leqslant X$,

$$R_{\psi,N}(n) \leqslant (1+\varepsilon)\sum_{x \leqslant X} \psi(x)^2.$$

Proof. For each $x \in \mathbb{N}$, let ξ_x denote the Bernoulli random variable where $\mathbb{P}(\xi_x = 1) = \psi(x)$. Then $N = \sum_{x \leqslant X} \xi_x$, which is a random variable with expectation $\sum_{x \leqslant X} \psi(x)$. Since the ξ_x are independent by assumption, one may apply the large deviation bound Lemma 0.5.3 to conclude part (1).

For part (2), by adjusting the implied constants we may assume that X is greater than some absolute fixed constant. We first consider each n separately. For each x

in the range $1 \leqslant x \leqslant X - n$, let $\omega_{x,n}$ denote the random variable $\xi_x \xi_{n+x}$. We have

$$R_{\psi,N}(n) = \sum_{x \leqslant X-n} \omega_{x,n},$$

which is a random variable with expectation $\sum_{x \leqslant X-n} \psi(x)\psi(n+x)$. Suppose first that $n \geqslant X - (1+\varepsilon) \sum_{x \leqslant X} \psi(x)^2$. Then by the trivial triangle inequality bound we have $R_{\psi,N}(n) \leqslant (1+\varepsilon) \sum_{x \leqslant X} \psi(x)^2$ as required.

It remains to consider the case $n \leqslant X - (1+\varepsilon) \sum_{x \leqslant X} \psi(x)^2$. We wish to apply large deviation bounds for sums of independent random variables, and, indeed, the family of random variables $\{\omega_{x,n} : x \leqslant X - n\}$ is very close to being independent. Methods of splitting this family into groups of genuinely independent random variables are alluded to in the discussion in the appendix of [3]. Here we describe an extremely coarse decomposition which nonetheless is strong enough for our purposes.

Split $\{x \in \mathbb{N} : x \leqslant X - n\}$ into two sets, $\mathcal{S}_0$ and $\mathcal{S}_1$, constructed as follows. If $n \leqslant X/3$, let $x \in \mathcal{S}_j$ if $\lfloor x/n \rfloor \equiv j \bmod 2$. If $n > X/3$, instead let

$$\mathcal{S}_0 = \Big\{ x \in \mathbb{N} : x \leqslant \frac{X-n}{2} \Big\}, \qquad \mathcal{S}_1 = \Big\{ x \in \mathbb{N} : \frac{X-n}{2} < x \leqslant X - n \Big\}.$$

For each $j \in \{0,1\}$ the family $\{\omega_{x,n} : x \in \mathcal{S}_j\}$ is independent, as no two indices differ by n. Applying the union bound and the large deviation bound Lemma 0.5.3 once more, we have

$$\begin{aligned}
&\mathbb{P}\Big(R_{\psi,N}(n) \geqslant (1+\varepsilon) \sum_{x \leqslant X-n} \psi(x)\psi(n+x) \Big) \\
&\leqslant \mathbb{P}\Big(\sum_{x \in \mathcal{S}_0} \omega_{x,n} \geqslant (1+\varepsilon) \sum_{x \in \mathcal{S}_0} \psi(x)\psi(n+x) \Big) \\
&+ \mathbb{P}\Big(\sum_{x \in \mathcal{S}_1} \omega_{x,n} \geqslant (1+\varepsilon) \sum_{x \in \mathcal{S}_1} \psi(x)\psi(n+x) \Big)
\end{aligned}$$

$$\ll \exp\Big(-c_\varepsilon \sum_{x\in\mathcal{S}_0}\psi(x)\psi(n+x)\Big) + \exp\Big(-c_\varepsilon \sum_{x\in\mathcal{S}_1}\psi(x)\psi(n+x)\Big).$$

It remains to estimate these final quantities. Note that by construction we have $\min(|\mathcal{S}_0|, |\mathcal{S}_1|) \geqslant (X-n)/4$, since X is assumed to be large. So, since $\psi(x)$ is non-increasing, we deduce that

$$\begin{aligned}\mathbb{P}\Big(R_{\psi,N}(n) \geqslant (1+\varepsilon)\sum_{x\leqslant X-n}\psi(x)\psi(n+x)\Big) &\ll \exp\Big(-c_\varepsilon \sum_{X-\frac{X-n}{4}<x\leqslant X}\psi(x)^2\Big)\\ &\ll \exp(-c_\varepsilon K_X).\end{aligned}$$

By a similar monotonic principle, $\mathbb{P}\Big(R_{\psi,N}(n) \geqslant (1+\varepsilon)\sum_{x\leqslant X-n}\psi(x)\psi(n+x)\Big)$ is at least $\mathbb{P}(R_{\psi,N}(n) \geqslant (1+\varepsilon)\sum_{x\leqslant X}\psi(x)^2)$. This finishes the proof of part (2). □

In the specific case when $\psi(n) = \min(1, (\log n)^{-(1+\delta)})$, for some fixed non-negative δ, these expressions are easy to compute. Indeed,

$$\begin{aligned}\sum_{x\leqslant X}\psi(x)^2 &\sim \frac{X}{\log^{2+2\delta} X}\\ K_X &\gg \frac{X}{\log^{4+4\delta} X}\\ \sum_{x\leqslant X}\psi(x) &\sim \frac{X}{\log^{1+\delta} X}.\end{aligned}$$

The sums

$$\begin{aligned}&\sum_{X\geqslant 1}\exp(-c_\varepsilon\sum_{x\leqslant X}\psi(x))\\ &\sum_{X\geqslant 1} X\exp(-c_\varepsilon K_X)\end{aligned}$$

are both convergent in these cases, and so one may conclude the following corollary.

Corollary 4.3.2. *Let δ be non-negative, and let $\psi(n) = \min(1, (\log n)^{-1-\delta})$. Then*

with probability 1, for all positive ε there exists some X_ε such that, for all X satisfying $X \geqslant X_\varepsilon$, one has

$$(1-\varepsilon)\frac{X}{\log^{1+\delta} X} \leqslant N \leqslant (1+\varepsilon)\frac{X}{\log^{1+\delta} X},$$

and, for all $n \leqslant X$,

$$R_{\psi,N}(n) \leqslant (1+\varepsilon)\frac{X}{\log^{2+2\delta} X}.$$

Proof. Fix some positive ε. Then, with probability 1, a suitable X_ε exists: this follows immediately from the Lemma 4.3.1, the subsequent asymptotics, and the first Borel-Cantelli lemma (Lemma 0.5.1).

The corollary as stated then follows by intersecting these probability 1 events over all rational ε. □

Now we move on to some simple number-theoretic lemmas. The first such lemma, which is Lemma 4.1 from [44], is a simple estimation of the function $\varphi(k,n)$.

Lemma 4.3.3. *Let X and Y be real numbers, and suppose that $1 \leqslant Y < X$. Let k be a real number that is at least 1. Then*

$$\sum_{Y \leqslant n \leqslant X} \frac{\varphi(k,n)}{n} = X - Y + 1 + O\left(\frac{X-Y}{k} + \log(X+1)\right).$$

Proof. Without loss of generality we may assume that X and Y are natural numbers. Then, since $\varphi(k,n) \leqslant n$, we need only prove a one-sided lower-bound estimate for $\sum_{n=Y}^{X} \frac{\varphi(k,n)}{n} - (X-Y+1)$. Observe the obvious inequality

$$\begin{aligned} \varphi(k,n) &\geqslant n - \sum_{\substack{d|n \\ d>k}} \sum_{\substack{m \leqslant n \\ d|m}} 1 \\ &\geqslant n - \sum_{\substack{d|n \\ d>k}} \frac{n}{d}. \end{aligned} \tag{4.8}$$

(Indeed, every m counted in the sum defining $\varphi(k,n)$ survives the cull on the right-

hand side of (4.8)). Thus

$$\begin{aligned}\sum_{n=Y}^{X}\frac{\varphi(k,n)}{n} &\geqslant \sum_{n=Y}^{X}\left(1-\frac{1}{n}\sum_{\substack{d|n\\ d>k}}\frac{n}{d}\right)\\ &= X-Y+1-\sum_{\substack{d>k\\ d\leqslant X}}\frac{1}{d}\sum_{\substack{n=Y\\ d|n}}^{X}1\\ &= X-Y+1-\sum_{\substack{d>k\\ d\leqslant X}}\frac{X-Y}{d^2}+O\left(\sum_{\substack{d>k\\ d\leqslant X}}\frac{1}{d}\right)\\ &= X-Y+1+O\left(\frac{X-Y}{k}+\log(X+1)\right)\end{aligned}$$

as claimed. □

The main use of this argument will be a particular corollary.

Corollary 4.3.4. *Let X be a real number satisfying $X > 2$, and suppose that N in (4.1) is non-zero. Then*

$$\sum_{n\leqslant X}\frac{\Phi^X(n)}{n} = X + O\left(\frac{X}{s^2X^2/N^2+1}+\log X\right).$$

Now for a very standard style of lemma.

Lemma 4.3.5. *Let X be a real number satisfying $X > 2$, and suppose that N in (4.1) is non-zero. Then*

$$\sum_{n\leqslant X}\tau^X(n) \ll X\log(s^2X^2/N^2+2).$$

Proof. We have

$$\begin{aligned}\sum_{n\leqslant X}\tau^X(n) &= \sum_{n\leqslant X}\sum_{\substack{d|n\\ d\leqslant s^2X^2/N^2+1}} 1\\ &= \sum_{d\leqslant s^2X^2/N^2+1}\sum_{\substack{n\leqslant X\\ d|n}} 1\\ &\leqslant X\sum_{d\leqslant s^2X^2/N^2+1}\frac{1}{d}\\ &\ll X\log(s^2X^2/N^2+2)\end{aligned}$$

as claimed. □

Now for the real meat of the matter, which is a strong overlap estimate on $\mathrm{meas}(\mathcal{E}_n^X\cap\mathcal{E}_m^X)$. Almost all this approach is once again due to Harman and Schmidt.

Lemma 4.3.6. *Let X be a real number satisfying $X>2$, and suppose that N in (4.1) is non-zero. Then, then for all natural numbers n and m satisfying $n\geqslant m\geqslant 1$ we have*

$$\mathrm{meas}(\mathcal{E}_n^X\cap\mathcal{E}_m^X)\leqslant 4s^2\frac{1}{N^2}+2\frac{s}{N}\frac{A^X(n,m)}{n}. \tag{4.9}$$

The presence of the constant 4 on the right-hand side of (4.9) is crucial, as in the main argument we will need the first term above to cancel exactly with other large terms.

The term $4s^2\frac{1}{N^2}$ represents roughly what the expected size of $\mathrm{meas}(\mathcal{E}_n^X\cap\mathcal{E}_m^X)$ would be, were the events of being in $\mathcal{E}_n^X$ and being in $\mathcal{E}_m^X$ truly independent subsets. The presence of $A^X(n,m)$ on the right-hand side of (4.9), a quantity which becomes smaller as the greatest common divisor bound imposed in its definition becomes stricter, suggests why imposing this greatest common divisor bound was a useful manoeuvre. This leverage is not present in [3].

Proof. Now, $\mathrm{meas}(\mathcal{E}_n^X \cap \mathcal{E}_m^X) = B_1 + B_2$, where B_2 is the contribution from those intervals

$$\left(\frac{u-\frac{s}{N}}{n}, \frac{u+\frac{s}{N}}{n}\right) \cap \left(\frac{v-\frac{s}{N}}{m}, \frac{v+\frac{s}{N}}{m}\right) \tag{4.10}$$

in which the centres of the intervals coincide, i.e.

$$\frac{u}{n} = \frac{v}{m}.$$

B_1 is the contribution from such intersections in which the centres do not coincide.

B_2 is very straightforward to estimate. Indeed, since $n \geqslant m$, we have

$$\begin{aligned} B_2 &= \frac{2s}{Nn} \sum_{\substack{u\leqslant n, v\leqslant m \\ \frac{u}{n}=\frac{v}{m} \\ (u,n)\leqslant s^2X^2/N^2+1 \\ (v,m)\leqslant s^2X^2/N^2+1}} 1 \\ &= 2\frac{s}{N}\frac{A^X(n,m)}{n} \end{aligned}$$

by definition. So it remains to show that the contribution B_1 is at most $4s^2/N^2$.

Suppose $\frac{v}{m} > \frac{u}{n}$ but that the intersection (4.10) is non-empty, i.e.

$$\frac{v-\frac{s}{N}}{m} < \frac{u+\frac{s}{N}}{n}.$$

Then

$$0 < vn - um < (n+m)\frac{s}{N},$$

and so we conclude two things. Firstly there is some non-zero integer h satisfying $|h| < (n+m)\frac{s}{N}$ such that $vn - um = h$. Secondly, the size of the intersection (4.10) is at most

$$\frac{u+\frac{s}{N}}{n} - \frac{v-\frac{s}{N}}{m},$$

which equals

$$\frac{s}{N}\left(\frac{1}{n}+\frac{1}{m}\right)-\frac{|h|}{mn}.$$

But clearly the intersection (4.10) is also at most $2s/Nn$. Considering analogously the case $\frac{v}{m}<\frac{u}{n}$, all these comments may be packaged into the inequality

$$B_1 \leqslant \sum_{\substack{h\in\mathbb{Z}\setminus\{0\}\\ |h|\leqslant(n+m)\frac{s}{N}}} S(h)\min\left(\frac{2s}{Nn},\frac{s}{N}\left(\frac{1}{n}+\frac{1}{m}\right)-\frac{|h|}{mn}\right), \tag{4.11}$$

where $S(h)$ be the number of solutions to $vn-um=h$ with $u\leqslant n$ and $v\leqslant m$.

We now consider $S(h)$ more carefully, and note that in fact

$$0\leqslant S(h)\leqslant\begin{cases}(m,n) & \text{if } (m,n)|h\\ 0 & \text{otherwise.}\end{cases} \tag{4.12}$$

Indeed, for notational simplicity let us write $d:=(m,n)$. Clearly $S(h)=0$ unless $d|h$. In the case where d does divide h, we divide both sides of the equation by d to reduce matters to counting the number of solutions $\{u,v\}$ to the equation

$$vn'-um'=h', \tag{4.13}$$

with $u\leqslant n$ and $v\leqslant m$, where $n'=n/d$, $m'=m/d$, and $h'=h/d$.

Suppose that $\{u_1,v_1\}$ and $\{u_2,v_2\}$ were two different solutions to (4.13). Then

$$(v_1-v_2)n'-(u_1-u_2)m'=0.$$

Since $(n',m')=1$, unique prime factorisation implies that $n'|(u_1-u_2)$. Since $n/n'=d$, this means that there are at most d possible values of u such that $u\leqslant n$ and there exists a v with $\{u,v\}$ being a solution to (4.13).

But u determines v in (4.13), so this means that there are at most d solutions $\{u, v\}$ to (4.13), thus proving the bound (4.12).

Substituting the bound (4.12) into (4.11), and letting y denote $(n+m)\frac{s}{N}$, we obtain

$$\begin{aligned} B_1 &\leqslant \sum_{\substack{h\in\mathbb{Z}\setminus\{0\}\\ 0<|h|\leqslant y\\ (m,n)|h}} (m,n)\min\left(\frac{2s}{Nn}, \frac{s}{N}\left(\frac{1}{n}+\frac{1}{m}\right) - \frac{|h|}{mn}\right) \\ &\leqslant \sum_{\substack{h\in\mathbb{Z}\setminus\{0\}\\ 0<|h|\leqslant \frac{y}{(m,n)}}} (m,n)\min\left(\frac{2s}{Nn}, \frac{s}{N}\left(\frac{1}{n}+\frac{1}{m}\right) - \frac{|h|(m,n)}{mn}\right). \end{aligned}$$

Now the function in the summand is non-increasing in $|h|$, so we may upper bound it by the integral

$$\begin{aligned} & 2\int_0^{\frac{y}{(m,n)}} (m,n)\min\left(\frac{2s}{Nn}, \frac{s}{N}\left(\frac{1}{n}+\frac{1}{m}\right) - \frac{h(m,n)}{mn}\right) dh \\ =& 2\int_0^{y} \min\left(\frac{2s}{Nn}, \frac{s}{N}\left(\frac{1}{n}+\frac{1}{m}\right) - \frac{h}{mn}\right) dh \\ \leqslant& 2\int_0^{(n-m)\frac{s}{N}} \frac{2s}{Nn}\,dh + 2\int_{(n-m)\frac{s}{N}}^{(n+m)\frac{s}{N}} \frac{s}{N}\left(\frac{1}{n}+\frac{1}{m}\right) - \frac{h}{mn}\,dh \\ =& 4s^2\frac{1}{N^2}\frac{n-m}{n} + 4s^2\frac{1}{N^2}m\left(\frac{1}{n}+\frac{1}{m}\right) - \frac{1}{mn}\frac{s^2}{N^2}((n+m)^2-(n-m)^2) \\ =& 4s^2\frac{1}{N^2}\left(1-\frac{m}{n}+\frac{m}{n}+1-1\right) \\ =& 4s^2\frac{1}{N^2} \end{aligned}$$

as claimed. □

We complete the list of lemmas by upper bounding sums of the quantity $A^X(n,m)$.

Lemma 4.3.7.

$$\sum_{m=1}^{n} A^X(n,m) \leqslant n\tau^X(n).$$

Proof. We upper bound $\sum_{m=1}^{n} A^X(n,m)$ by neglecting the constraint $(v,m) \leqslant s^2X^2/N^2+1$, i.e.

$$\sum_{m=1}^{n} A^X(n,m) \leqslant \sum_{\substack{1\leqslant m,u,v\leqslant n \\ \frac{u}{n}=\frac{v}{m} \\ (u,n)\leqslant s^2X^2/N^2+1}} 1.$$

We make some trivial observations. For fixed u, there are unique coprime postive integers a and b such that $\frac{u}{n} = \frac{a}{b}$. The denominator b is a divisor of n. Further, if a and b are fixed coprime positive integers where $a \leqslant b$ and b is a divisor of n, the number of solutions to

$$\frac{a}{b} = \frac{v}{m}, \qquad \text{with } 1 \leqslant v, m \leqslant n$$

is exactly $\frac{n}{b}$. Therefore

$$\sum_{m=1}^{n} A^X(n,m) \leqslant \sum_{b|n} \sum_{\substack{u\leqslant n \\ \exists a:\, \frac{u}{n}=\frac{a}{b},\, (a,b)=1 \\ (u,n)\leqslant s^2X^2/N^2+1}} \frac{n}{b}.$$

If a pair u and b are counted in the double sum, then $b = \frac{n}{(u,n)}$. Therefore we get the expression

$$\begin{aligned}\sum_{m=1}^{n} A^X(n,m) &\leqslant \sum_{b|n} \sum_{\substack{u\leqslant n \\ (u,n)=\frac{n}{b} \\ (u,n)\leqslant s^2X^2/N^2+1}} \frac{n}{b} \\ &\leqslant \sum_{\substack{b|n \\ \frac{n}{b}\leqslant s^2X^2/N^2+1}} b\frac{n}{b}\end{aligned}$$

$$= \sum_{\substack{c|n \\ c \leqslant s^2X^2/N^2+1}} n$$
$$= n\tau^X(n)$$

as claimed. □

4.4 The main argument

Firstly we make rigorous the idea that, for almost all α, $F_1^*(\mathcal{A}_\psi, \alpha, s, X)$ is a good approximation to $F_1(\mathcal{A}_\psi, \alpha, s, X)$.

Lemma 4.4.1. *Let δ be non-negative, and let $\psi(n) = \min(1, (\log n)^{-1-\delta})$ as usual. Let c, C be two absolute positive constants with $c < C$. Then with probability 1 there exists a positive real $X_{1/2}$ (that may depend on c, C, δ) such that the following holds: if $X \geqslant X_{1/2}$, and if s is a positive real that satisfies $c < s < C$, then N (as defined in (4.1)) is non-zero, and there exists a set $E_{X,s} \subseteq [0,1]$ that has measure $O((\log X)^{-\frac{3}{2}-2\delta})$ and such that, if $\alpha \in [0,1] \setminus E_{X,s}$,*

$$F_1(\mathcal{A}_\psi, \alpha, s, X) - F_1^*(\mathcal{A}_\psi, \alpha, s, X) \ll_{c,C} (\log X)^{-\frac{1}{2}}.$$

The implied constant is independent of α and δ.

Proof. Let $\varepsilon = 1/2$, and consider Corollary 4.3.2 with this value of ε. This shows that, with probability 1, there exists a positive real number $X_{1/2}$ such that, for all $X \geqslant X_{1/2}$,

$$N \asymp \sum_{x \leqslant X} \psi(x) \asymp \frac{X}{\log^{1+\delta} X} \tag{4.14}$$

and for all $n \leqslant X$

$$R_{\psi,N}(n) \ll \sum_{x \leqslant X} \psi(x)^2 \asymp \frac{X}{\log^{2+2\delta} X}. \tag{4.15}$$

We henceforth assume that we are operating in the probability 1 event in which such an $X_{1/2}$ exists. Making $X_{1/2}$ larger as necessary (in terms of C and δ), we may also assume that if $X \geqslant X_{1/2}$ then $s/N < 1/2$.

So, let X be a real number with $X \geqslant X_{1/2}$. Note first that

$$\begin{aligned}\int_0^1 F_1(\mathcal{A}_\psi, \alpha, s, X) - F_1^*(\mathcal{A}_\psi, \alpha, s, X)\, d\alpha \\ = 4\frac{s}{N}\sum_{n\leqslant X}\left(\frac{R_{\psi,N}(n)}{N}\right) - 2\sum_{n\leqslant X}\left(\frac{R_{\psi,N}(n)}{N}\right)\mathrm{meas}(\mathcal{E}_n^X). \qquad (4.16)\end{aligned}$$

The intervals that are used to define $\mathcal{E}_n^X$ are disjoint (since $s/N < 1/2$), and so

$$\begin{aligned}\mathrm{meas}(\mathcal{E}_n^X) &= 2\frac{s}{Nn}\sum_{\substack{m\leqslant n\\ (m,n)\leqslant s^2X^2/N^2+1}} 1 \\ &= 2\frac{s}{N}\frac{\Phi^X(n)}{n}.\end{aligned}$$

Substituting into (4.16) we get that

$$\int_0^1 F(\mathcal{A}_\psi, \alpha, s, N) - F^*(\mathcal{A}_\psi, \alpha, s, N)\, d\alpha = 4\frac{s}{N}\sum_{n\leqslant X}\left(\frac{R_{\psi,N}(n)}{N}\right)\left(1 - \frac{\Phi^X(n)}{n}\right).$$

This is

$$\begin{aligned}&\ll s\frac{\sum_{x\leqslant X}\psi(x)^2}{(\sum_{x\leqslant X}\psi(x))^2}\sum_{n\leqslant X}\left(1 - \frac{\Phi^X(n)}{n}\right) \\ &\ll s\left(\frac{X}{s^2X^2/N^2+1} + \log X\right)\frac{\sum_{x\leqslant X}\psi(x)^2}{(\sum_{x\leqslant X}\psi(x))^2},\end{aligned}$$

by an application of Corollary 4.3.4, along with bounds (4.14) and (4.15).

Replacing the expressions involving ψ with their asymptotic form, we conclude that

$$\int_0^1 F_1(\mathcal{A}_\psi, \alpha, s, X) - F_1^*(\mathcal{A}_\psi, \alpha, s, X)\, d\alpha \ll s\left(\frac{1}{s^2(\log X)^{2+2\delta}+1} + \frac{\log X}{X}\right)$$
$$\ll_{c,C} \frac{1}{(\log X)^{2+2\delta}}.$$

The integrand is always positive, so the lemma follows by Markov's inequality. □

Now we approach the important variance estimate.

Lemma 4.4.2. *Let δ be non-negative, and let $\psi(n) = \min(1, (\log n)^{-1-\delta})$ as usual. Let c, C be two absolute positive constants with $c < C$. Then with probability 1 there exists a positive real $X_{1/2}$ (that may depend on c, C, δ) such that the following holds: if $X \geqslant X_{1/2}$, and if s is a positive real that satisfies $c < s < C$, then N (as defined in (4.1)) is non-zero, and*

$$\int_0^1 (F_1^*(\mathcal{A}_\psi, \alpha, s, X) - 2s)^2\, d\alpha \ll_{c,C} (2+2\delta)(\log\log X)(\log X)^{-1-\delta}.$$

In other words, viewing α as a uniform random variable taking values in $[0, 1]$, the variance of $F_1^*(\mathcal{A}_\psi, \alpha, s, X)$ enjoys the bound $(2+2\delta)(\log\log X)(\log X)^{-1-\delta}$. For any positive δ, this bound tends to zero faster than $\log X$, and that will be our key leverage.

Proof. We proceed as in the proof of Lemma 4.4.1, picking the same $X_{1/2}$ (and assuming that we are operating in the probability 1 event in which such an $X_{1/2}$ exists).

Let X be a real number with $X \geqslant X_{1/2}$. Then we have that

$$\int_0^1 (F_1^*(\mathcal{A}_\psi, \alpha, s, X) - 2s)^2 \, d\alpha \tag{4.17}$$

is equal to

$$\begin{aligned}
&4 \sum_{n,m \leqslant X} \frac{R_{\psi,N}(n) R_{\psi,N}(m)}{N^2} \operatorname{meas}(\mathcal{E}_n^X \cap \mathcal{E}_m^X) - 8s \sum_{n \leqslant X} \frac{R_{\psi,N}(n)}{N} \operatorname{meas}(\mathcal{E}_n^X) + 4s^2 \\
\leqslant &16 \frac{s^2}{N^2} \sum_{n,m \leqslant X} \frac{R_{\psi,N}(n) R_{\psi,N}(m)}{N^2} + 16 \frac{s}{N} \sum_{\substack{n \leqslant X \\ m \leqslant n}} \frac{R_{\psi,N}(n) R_{\psi,N}(m)}{N^2} \left(\frac{A^X(n,m)}{n} \right) \\
&- 8s \sum_{n \leqslant X} \frac{R_{\psi,N}(n)}{N} \operatorname{meas}(\mathcal{E}_n^X) + 4s^2 \\
\leqslant &16 \frac{s^2}{N^4} \sum_{n,m \leqslant X} R_{\psi,N}(n) R_{\psi,N}(m) + 16 \frac{s}{N^3} \sum_{\substack{n \leqslant X \\ m \leqslant n}} R_{\psi,N}(n) R_{\psi,N}(m) \left(\frac{A^X(n,m)}{n} \right) \\
&- 16 \frac{s^2}{N^2} \sum_{n \leqslant X} R_{\psi,N}(n) \frac{\Phi^X(n)}{n} + 4s^2.
\end{aligned} \tag{4.18}$$

We may now recombine the terms, without needing to assume any regularity in the sizes of the $R_{\psi,N}(n)$, simply the fact that $\sum_{m \leqslant X} R_{\psi,N}(m) = \frac{1}{2} N(N-1)$. It is at this stage we see how vital it was that, in the statement of Lemma 4.9, the constant in the first term on the right-hand side was exactly 4. Indeed, since

$$16 \frac{s^2}{N^4} \sum_{n,m \leqslant X} R_{\psi,N}(n) R_{\psi,N}(m) + 4s^2 = 16 \frac{s^2}{N^2} \sum_{n \leqslant X} R_{\psi,N}(n) + O(\frac{s^2}{N}),$$

we have that (4.18) is

$$\begin{aligned}
= 16 \frac{s^2}{N^2} \sum_{n \leqslant X} R_{\psi,N}(n) \left(1 - \frac{\Phi^X(n)}{n} \right) + 16 \frac{s}{N^3} \sum_{\substack{n \leqslant X \\ m \leqslant n}} R_{\psi,N}(n) R_{\psi,N}(m) \left(\frac{A^X(n,m)}{n} \right) \\
+ O(\frac{s^2}{N}).
\end{aligned} \tag{4.19}$$

Consider the first term of (4.19). Since $R_{\psi,N}(n) \ll \sum_{x \leqslant X} \psi(x)^2$ for every n (by (4.15)) and $N \asymp \sum_{x \leqslant X} \psi(x)$ (by (4.14), we may use Corollary 4.3.4 and the fact that $s < C$ to deduce that

$$16\frac{s^2}{N^2}\sum_{n \leqslant x} R_{\psi,N}(n)\left(1 - \frac{\Phi^X(n)}{n}\right) \ll_C \frac{\sum_{x \leqslant X}\psi(x)^2}{(\sum_{x \in X}\psi(x))^2}\left(\frac{X}{s^2X^2/N^2+1} + \log X\right).$$

The second term of (4.19) may be bounded above by

$$\begin{aligned}
&\ll \frac{(\sum_{x \leqslant X}\psi(x)^2)^2}{(\sum_{x \leqslant X}\psi(x))^3}\sum_{\substack{n \leqslant X \\ m \leqslant n}}\frac{A^X(m,n)}{n} \\
&\ll \frac{(\sum_{x \leqslant X}\psi(x)^2)^2}{(\sum_{x \leqslant X}\psi(x))^3}\sum_{n \leqslant X}\tau^X(n) \\
&\ll \frac{(\sum_{x \leqslant X}\psi(x)^2)^2}{(\sum_{x \leqslant X}\psi(x))^3} X\log(s^2X^2/N^2+2)
\end{aligned}$$

Replacing the expressions involving ψ with their asymptotic form, we conclude that the variance (4.17) is

$$\begin{aligned}
\ll_{c,C} \frac{X/(\log X)^{2+2\delta}}{(X/(\log X)^{1+\delta})^2}&\left(\frac{X}{(\log X)^{2+2\delta}} + \log X\right) \\
&+ \frac{(X/(\log X)^{2+2\delta})^2}{(X/(\log X)^{1+\delta})^3} X\log((\log X)^{2+2\delta}),
\end{aligned}$$

which is $O_{c,C}((2+2\delta)(\log\log X)(\log X)^{-1-\delta})$. The lemma is proved. □

The fact that one can get a bound for the variance of $F_1^*(\mathcal{A}_\psi, \alpha, s, X)$ that decays faster than $(\log X)^{-1}$, even when δ is very small, is at the heart of our method. We remark that, had we applied the methods of [3] to estimate the variance of $F_1(\mathcal{A}_\psi, \alpha, s, X)$ directly, one would have obtained a bound of approximately $(\log X)^{-\delta}$. This would have sufficed for $\delta > 1$, but not for smaller δ.

4.5 Proof of Theorem 4.1.3

Assume that $\psi(n) = \min(1, (\log n)^{-1-\delta})$ for some fixed positive δ. To show that $\mathcal{A}_\psi$ is metric poissonian, it will be enough to prove the following claim.

Claim 4.5.1. *For all positive ε, s, with probability 1 there exists a set $\Omega_{\varepsilon,s} \subseteq [0,1]$ with measure 1 and with the follow property: for all $\alpha \in \Omega_{\varepsilon,s}$, there exists a real number $X(\alpha, \varepsilon, s)$ such that for all $X \geqslant X(\alpha, \varepsilon, s)$ one has*

$$|F_1(\mathcal{A}_\psi, \alpha, s, X) - 2s| \leqslant \varepsilon.$$

Proof that Claim 4.5.1 implies Theorem 4.1.3. Consider all rational values of ε and s, and the probability 1 events and sets $\Omega_{\varepsilon,s}$ given by Claim 4.5.1. By intersecting the sets $\Omega_{\varepsilon,s}$ over all rational values of ε and s, and intersecting all the probability 1 events, we conclude that with probability 1 there exists a set $\Omega \subseteq [0,1]$ of measure 1 such that for all $\alpha \in \Omega$ and for all positive rational numbers s we have $F_1(\mathcal{A}_\psi, \alpha, s, X) \to 2s$ as $X \to \infty$.

We claim that the same conclusion is true for irrational s. Indeed, let ε be positive and let s_1, s_2 be rational numbers such that $0 < s_1 < s < s_2$ and $|s - s_1|, |s - s_2| \leqslant \varepsilon/4$. Let X be large enough so that

$$|F_1(\mathcal{A}_\psi, \alpha, s_1, X) - 2s_1| \leqslant \frac{\varepsilon}{2}$$
$$|F_1(\mathcal{A}_\psi, \alpha, s_2, X) - 2s_2| \leqslant \frac{\varepsilon}{2}.$$

Then

$$2s - \varepsilon \leqslant 2s_1 - \frac{\varepsilon}{2} \leqslant F_1(\mathcal{A}_\psi, \alpha, s_1, X) \leqslant F_1(\mathcal{A}_\psi, \alpha, s, X)$$

$$\leqslant F_1(\mathcal{A}_\psi, \alpha, s_2, X) \leqslant 2s_2 + \frac{\varepsilon}{2} \leqslant 2s + \varepsilon.$$

The quantity ε was arbitrary, so

$$F_1(\mathcal{A}_\psi, \alpha, s, X) \longrightarrow 2s$$

as $X \to \infty$, and so $\mathcal{A}_\psi$ is metric poissonian. Theorem 4.1.3 is proved. □

It remains to prove Claim 4.5.1. To this end, assume that we are in the probability 1 scenario described in Corollary 4.3.2, i.e. that for all positive ε there exists some X_ε such that, that for all X satisfying $X \geqslant X_\varepsilon$, one has

$$(1-\varepsilon)\frac{X}{\log^{1+\delta} X} \leqslant N \leqslant (1+\varepsilon)\frac{X}{\log^{1+\delta} X}, \tag{4.20}$$

and, for all $n \leqslant X$,

$$R_{\psi,N}(n) \leqslant (1+\varepsilon)\frac{X}{\log^{2+2\delta} X}. \tag{4.21}$$

Fix the values of $\varepsilon, s > 0$ in the statement of Claim 4.5.1 and let η be a positive constant, picked small enough in terms of δ. Let

$$X_j = e^{j^{1-\eta}}.$$

We may assume that X is sufficiently large in terms of ε and s, and write $X_j \leqslant X < X_{j+1}$ for some j.

One has the inequality

$$N_j F_1(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j) \leqslant N F_1(\mathcal{A}_\psi, \alpha, s, X) \leqslant N_{j+1} F_1(\mathcal{A}_\psi, \alpha, s\frac{N_{j+1}}{N_j}, X_{j+1}).$$

Observe the critical fact that $X_{j+1}/X_j \to 1$ as $j \to \infty$. Therefore, by expression (4.20), $N_{j+1}/N_j \to 1$ as $j \to \infty$. In particular there exist positive absolute constants

c and C such that, if j is large enough, $c < s, s\frac{N_j}{N_{j+1}}, s\frac{N_{j+1}}{N_j} < C$.

We make two deductions from the lemmas in section 4.4. Firstly, apply Lemma 4.4.1 to the expression $F_1(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j)$. Therefore, with probability 1, if X is large enough in terms of C, c, δ then there exists a set $E_{X_j, s\frac{N_j}{N_{j+1}}}$ with all the properties from Lemma 4.4.1.

Certainly we have

$$\text{meas}(E_{X_j, s\frac{N_j}{N_{j+1}}}) \ll_{c,C} \frac{1}{\log^{\frac{3}{2}+2\delta} X_j} \ll j^{-\frac{3}{2}},$$

if η is small enough. Hence

$$\sum_{j=1}^{\infty} \text{meas}(E_{X_j, s\frac{N_j}{N_{j+1}}}) < \infty.$$

So by the first Borel-Cantelli lemma (Lemma 0.5.1),

$$\text{meas} \bigcup_{J \geqslant 1} \bigcap_{j \geqslant J} \left([0,1] \setminus E_{X_j, s\frac{N_j}{N_{j+1}}} \right) = 1.$$

This means that, provided η is small enough, for almost all $\alpha \in [0,1]$ there exists a value $j(\alpha)$ such that for all $j \geqslant j(\alpha)$

$$|F_1(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j) - F_1^*(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j)| \ll_{c,C} j^{-\frac{1}{2}+\frac{\eta}{2}} \ll_{c,C} j^{-\frac{1}{4}}. \tag{4.22}$$

Arguing analogously for $s\frac{N_{j+1}}{N_j}$, for all $j \geqslant j(\alpha)$ one has

$$|F_1(\mathcal{A}_\psi, \alpha, s\frac{N_{j+1}}{N_j}, X_{j+1}) - F_1^*(\mathcal{A}_\psi, \alpha, s\frac{N_{j+1}}{N_j}, X_{j+1})| \ll_{c,C} j^{-\frac{1}{4}}. \tag{4.23}$$

Secondly, by applying Chebyshev's inequality (Lemma 0.5.2) to the variance esti-

mate in Lemma 4.4.2, we conclude that

$$\begin{aligned}\mathrm{meas}\{\alpha : (|F_1^*(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j) - 2s\frac{N_j}{N_{j+1}}| \geqslant 0.1\varepsilon\} &\ll_{\delta,c,C} \varepsilon^{-2}(\log X_j)^{-1-\delta/2} \\ &\ll_{\delta,c,C} \varepsilon^{-2} j^{-1-\delta/4}\end{aligned}$$

if η is small enough.

Therefore, applying the first Borel-Cantelli lemma again, for almost all $\alpha \in [0,1]$ there exists a positive real $j(\alpha,\varepsilon)$ such that

$$|F_1^*(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j) - 2s\frac{N_j}{N_{j+1}}| < 0.1\varepsilon \tag{4.24}$$

for all $j \geqslant j(\alpha,\varepsilon)$. By an analogous argument we may also assume that

$$|F_1^*(\mathcal{A}_\psi, \alpha, s\frac{N_{j+1}}{N_j}, X_{j+1}) - 2s\frac{N_{j+1}}{N_j}| < 0.1\varepsilon \tag{4.25}$$

for all $j \geqslant j(\alpha,\varepsilon)$. Without loss of generality, we may assume that $j(\alpha,\varepsilon) \geqslant j(\alpha)$.

We now conclude the proof. Indeed, let $K_{c,C}$ be a sufficiently large constant, depending on c and C, and let $K_{c,C,\varepsilon}$ be a sufficiently large constant, depending on c, C and ε. Assume that X is large enough so that, if X is in the range $X_j \leqslant X < X_{j+1}$, then $|N_{j+1}/N_j - 1| \leqslant K_{c,C,\varepsilon}^{-1}$ and j is large enough such that both $j \geqslant j(\alpha,\varepsilon)$ and $j^{-\frac{1}{2}} \leqslant \varepsilon K_{c,C}^{-1}$. Then, by combining (4.22) and (4.24), if $K_{c,C}$ is large enough one has

$$F_1(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j) \geqslant 2s\frac{N_j}{N_{j+1}} - 0.2\varepsilon.$$

But

$$F_1(\mathcal{A}_\psi, \alpha, s\frac{N_j}{N_{j+1}}, X_j) \leqslant \frac{N}{N_j} F_1(\mathcal{A}_\psi, \alpha, s, X),$$

so we conclude, provided that $K_{c,C,\varepsilon}$ is large enough, that

$$F_1(\mathcal{A}_\psi, \alpha, s, X) \geqslant 2s\frac{N_j^2}{NN_{j+1}} - 0.5\varepsilon.$$

Arguing similarly for the upper bound, we get

$$F_1(\mathcal{A}_\psi, \alpha, s, X) \leqslant 2s\frac{N_{j+1}^2}{NN_j} + 0.5\varepsilon.$$

Since $|N_{j+1}/N_j - 1| \leqslant K_{c,C,\varepsilon}^{-1}$, we conclude that

$$|F_1(\mathcal{A}_\psi, \alpha, s, X) - 2s| \leqslant \varepsilon.$$

This proves Claim 4.5.1. □

As noted above, this concludes the proof of Theorem 4.1.3. □

4.6 Proof of Theorem 4.1.4

This proof is very similar to the previous one, and in fact a little simpler: we give a sketch.

Let η be a small positive parameter, and define

$$X_j := e^{j^{1+\eta}}$$

(a subtle though important difference in definition from the previous proof). The manoeuvre in Claim 4.5.1, of considering for intersection over all rational s and ε, goes through unaltered, and therefore it suffices to prove, for all fixed positive s and ε, that, with probability 1, for almost all α there exists a real number $j(\alpha, \varepsilon, s)$ such

that

$$|F_1(\mathcal{A}_\psi, \alpha, s, X_j) - 2s| \leqslant \varepsilon$$

for all $j \geqslant j(\alpha, \varepsilon, s)$.

We stress at this point that we will just consider fixed s, rather than utilising the expression sN_{j+1}/N_j in a sandwiching argument as in the previous proof[4].

Letting $E_{X_j,s}$ be the exceptional set coming from Lemma 4.4.1, one has the bound

$$\sum_{j=1}^{\infty} \text{meas}(E_{X_j,s}) < \infty,$$

and so, applying the first Borel-Cantelli lemma (Lemma 0.5.1) as previously, we conclude that for almost all α there exists a value $j(\alpha)$ such that for all $j \geqslant j(\alpha)$ one has

$$|F_1(\mathcal{A}_\psi, \alpha, s, X_j) - F_1^*(\mathcal{A}_\psi, \alpha, s, X_j)| \ll_s j^{-\frac{1}{2}}$$

Now, by applying Chebyshev's inequality (Lemma 0.5.2) to the variance estimate for F_1^* given in Lemma 4.4.2, we conclude that

$$\text{meas}(\alpha : |F_1^*(\mathcal{A}_\psi, \alpha, s, X_j) - 2s| \geqslant 0.1\varepsilon) \ll_{s,\varepsilon} (\log\log X_j)(\log X_j)^{-1} \ll_{s,\varepsilon} j^{-1-\eta/2}.$$

Since η is positive, the sum of these measures converges, and so we conclude (from the first Borel-Cantelli lemma again) that for almost all α there exists a $j(\alpha, \varepsilon)$ such that

$$|F_1^*(\mathcal{A}_\psi, \alpha, s, X_j) - 2s| < 0.1\varepsilon$$

for all $j \geqslant j(\alpha, \varepsilon)$.

Combining these two statements, we conclude that for almost all α there exists a

[4]This is because the sandwiching argument no longer works, as the ratio N_{j+1}/N_j does not tend to 1. Even worse, this ratio isn't bounded.

$j(\alpha, \varepsilon, s)$ such that

$$|F_1(\mathcal{A}_\psi, \alpha, s, X_j) - 2s| \leqslant \varepsilon$$

for all $j \geqslant j(\alpha, \varepsilon, s)$. The theorem is proved. □

Chapter 5

Gowers norms control diophantine inequalities

5.1 Introduction

Diophantine inequalities are a vast and varied topic in analytic number theory (see [9], say). We will focus on a particular class of problems, which are of the following general form. Let A be a set of integers, let ε be a positive parameter, and let L be an m-by-d real matrix, with $d \geqslant m+1$. One may ask whether there are infinitely many solutions to

$$\|L\mathbf{a}\|_\infty \leqslant \varepsilon \tag{5.1}$$

with all the coordinates of $\mathbf{a}$ lying in A. Further, letting N be a natural number, one might seek an asymptotic formula (as N tends to infinity) for the number of such solutions that satisfy $\|\mathbf{a}\|_\infty \leqslant N$. Much is known about this problem for certain special sets A (see [7, 18, 55, 61, 65, 66]), in particular for the image sets of polynomials. This is discussed in section 5.2. However, as far we are aware, the situation has not before been considered in such generality.

In section 0.3 we recorded the basic properties of Gowers norms. These norms were introduced around twenty years ago, as part of Gowers' proof of Szemerédi's Theorem [33], and since then they have become a fundamental tool in additive combinatorics and in analytic number theory. One particular application is to the study of linear equations with rational coefficients. Indeed, the study of such systems was greatly enhanced by the introduction, by Green and Tao in [36, 38], of a powerful and wide-ranging technique, known as a 'Generalised von Neumann Theorem', which can be used to show that Gowers norms are, in some sense, 'universal' over all such linear systems: this is Theorem 7.1 of [38], and we recall a similar version in Theorem 5.1.1. It was using this technique, in combination with a deep study of the inverse theory of Gowers norms, that those authors and Ziegler managed to prove that, generically, $m+2$ prime variables are adequate to obtain an asymptotic formula for the number of prime solutions to m linear equations with rational coefficients, rather than the $2m+1$ variables required by the circle method.

Our motivation for embarking upon this entire project was to try to use these ideas of Green and Tao to understand (5.1) when A is the set of primes. We have a theorem in this direction, but the proof is still the subject of ongoing work, and we choose not to present it in this thesis. Many additional technical difficulties occur for the primes, stemming from the well-studied irksome fact that the von Mangoldt function is unbounded. The purpose of this chapter is rather to develop a theory for diophantine inequalities weighted by bounded functions.

Consider the following theorem of Green and Tao, which applies Gowers norms to bound the number of solutions to equations with integer coefficients. A more general version of this theorem is a critical tool in those authors' work on linear equations in primes ([38]).

Theorem 5.1.1 (Generalised von Neumann Theorem for rational forms (non-quan-

titative)). *Let N, m, d be natural numbers, satisfying $d \geqslant m+2$. Let L be an m-by-d real matrix with integer coefficients, with rank m. Suppose that there does not exist any non-zero row-vector in the row-space of L that has two or fewer non-zero coordinates. Then there is some natural number s at most $d-2$ that satisfies the following. Let $f_1, \cdots, f_d : [N] \longrightarrow [-1,1]$ be arbitrary functions, and suppose that*

$$\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho$$

for some parameter ρ in the range $0 < \rho \leqslant 1$. Then

$$\frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n} \in [N]^d \\ L\mathbf{n}=\mathbf{0}}} \prod_{j=1}^{d} f_j(n_j) \ll_L \rho^{\Omega(1)} + o_\rho(1).$$

Theorem 5.1.1 is implicit in [38], but it is not explicitly stated in that paper (the authors' focus being on results over primes). We will later require a quantitative version (Theorem 5.5.2), at which point we will describe fully how to extract these statements from [38].

At first sight the non-degeneracy condition in the statement of Theorem 5.1.1, concerning the row-space of L, may seem a little unnatural. However, it is actually a necessary condition for Gowers norms to be used in this way (as we show later in Theorem 5.2.12).

The main result of this chapter (Theorem 5.2.10) will generalise Theorem 5.1.1 to the setting of diophantine inequalities. Because we take care to record the quantitative dependencies of the error terms, Theorem 5.2.10 is rather technical to state. Fortunately, it admits a corollary that is much more transparent. This corollary is strong enough to give our main application (an application to cancellation of the

Möbius function, see Corollary 5.12).

Corollary 5.1.2 (Generalised von Neumann Theorem for diophantine inequalities (non-quantitative))**.** *Let N, m, d be natural numbers, satisfying $d \geqslant m+2$, and let ε be a positive parameter. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be an m-by-d real matrix, with rank m. Suppose that there does not exist any non-zero row-vector in the row-space of L that has two or fewer non-zero coordinates. Then there is some natural number s at most $d-2$, independent of ε, such that the following is true. Let $f_1, \dots, f_d : [N] \longrightarrow [-1, 1]$ be arbitrary functions, and suppose that*

$$\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho,$$

for some parameter ρ in the range $0 < \rho \leqslant 1$. Then

$$\Big|\frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n}\in[N]^d \\ \|L\mathbf{n}\|_\infty \leqslant \varepsilon}} \Big(\prod_{j=1}^{d} f_j(n_j)\Big)\Big| \ll_{L,\varepsilon} \rho^{\Omega(1)} + o_{\rho,L}(1).$$

We can provide detailed information about how the implied constant and the $o_{\rho,L}(1)$ term depend on L, but we defer those technicalities to Theorem 5.2.10.

Note how, by picking ε small enough, Corollary 5.1.2 immediately implies Theorem 5.1.1.

Let us illustrate Corollary 5.1.2 with some examples.

Example 5.1.3 (Three-term irrational AP)**.** The first example could have been proved by Davenport and Heilbronn using the methods of [18], but we include it here to demonstrate the simplest case where Corollary 5.1.2 applies. Let

$$L := \begin{pmatrix} 1 & -\sqrt{2} & -1+\sqrt{2} \end{pmatrix}.$$

Then $m = 1$ and $d = 3$, and obviously there does not exist any non-zero row-vector in the row-space of L that has two or fewer non-zero coordinates.

Therefore Corollary 5.1.2 applies, and so, if $f_1, f_2, f_3 : [N] \longrightarrow [-1, 1]$ are three functions satisfying $\min_j \|f_j\|_{U^2[N]} \leqslant \rho$ for some ρ in the range $0 < \rho \leqslant 1$, we have

$$\Big|\frac{1}{N^2} \sum_{\substack{n_1,n_2,n_3 \leqslant N \\ |n_1 - \sqrt{2}n_2 + (-1+\sqrt{2})n_3| \leqslant \varepsilon}} f_1(n_1) f_2(n_2) f_3(n_3)\Big| \ll_\varepsilon \rho^{\Omega(1)} + o_\rho(1). \tag{5.2}$$

The statement (5.2) admits a different interpretation, which some readers may find more natural, that of counting the number of occurrences of a certain irrational pattern: a 'three-term irrational arithmetic progression'. Indeed, recall that for $\theta \in \mathbb{R}$ we let $[\theta]$ denote $\lfloor \theta + \frac{1}{2} \rfloor$, i.e. the nearest integer to θ. Then for any three functions $f_1, f_2, f_3 : [N] \longrightarrow [-1, 1]$, we make the definition

$$T(f_1, f_2, f_3) := \frac{1}{N^2} \sum_{x,d \in \mathbb{Z}} f_3(x) f_2(x + d) f_1([x + \sqrt{2}d]). \tag{5.3}$$

Informally speaking, T counts the number of near-occurrences of the pattern $(x, x + d, x + \sqrt{2}d)$, weighted by the functions f_j. By a simple change of variables $n_1 = [x + \sqrt{2}d]$, $n_2 = x + d$, $n_3 = x$, and noting that $x + \sqrt{2}d \notin \frac{1}{2}\mathbb{Z}$, we see

$$T(f_1, f_2, f_3) = \frac{1}{N^2} \sum_{\substack{n_1,n_2,n_3 \leqslant N \\ |n_1 - \sqrt{2}n_2 + (-1+\sqrt{2})n_3| \leqslant \frac{1}{2}}} f_1(n_1) f_2(n_2) f_3(n_3). \tag{5.4}$$

By (5.2), this means

$$|T(f_1, f_2, f_3)| \ll \rho^{\Omega(1)} + o_\rho(1), \tag{5.5}$$

provided $\min_j \|f_j\|_{U^2[N]} \leqslant \rho$.

One can use these results to count the number of near-occurrences of the pattern

$(x, x+d, x+\sqrt{2}d)$ in a Fourier-uniform set, which we do in Corollary 5.1.4 below. Indeed, suppose that A is a subset of $[N]$ with $|A| = \alpha N$. Let

$$f_A := 1_A - \alpha 1_{[N]} \tag{5.6}$$

be its so-called 'balanced function'. By the usual telescoping trick, $T(1_A, 1_A, 1_A)$ is equal to

$$T(\alpha 1_{[N]}, \alpha 1_{[N]}, \alpha 1_{[N]}) + T(f_A, \alpha 1_{[N]}, \alpha 1_{[N]}) + T(1_A, f_A, \alpha 1_{[N]}) + T(1_A, 1_A, f_A). \tag{5.7}$$

Bounding the final three terms using $\|f_A\|_{U^2[N]}$, and using the relation (5.4), one may establish that

$$\frac{1}{N^2} \sum_{x,d \in \mathbb{Z}} 1_A(x) 1_A(x+d) 1_A([x+\sqrt{2}d]) = C\alpha^3(1+o(1)) + O(\rho^{\Omega(1)}) + o_\rho(1) \tag{5.8}$$

for some positive constant C, provided $\|f_A\|_{U^2[N]} \leqslant \rho$. If $\|f_A\|_{U^2[N]} = o(1)$ then, by picking $\rho = \rho(N)$ to be a quantity that tends to zero suitably slowly as N tends to infinity, (5.8) implies that the number of occurrences of the configuration $(x, x+d, [x+\sqrt{2}d])$ in A is asymptotically $C\alpha^3 N^2$.

For bounded functions, the U^2-norm is closely related to the Fourier transform. We say that A is Fourier-uniform if its balanced function f_A satisfies

$$\sup_{\theta \in [0,1]} \frac{1}{N} \sum_{n \leqslant N} f_A(n) e(n\theta) = o(1),$$

and it is a standard result (see [84, Exercise 1.3.18]) that A is Fourier uniform if and only if $\|f_A\|_{U^2[N]} = o(1)$. Therefore expression (5.8), and the remarks following it, imply the following corollary.

Corollary 5.1.4. *Let N be a natural number. If A is a Fourier-uniform subset of*

$[N]$, *with* $|A| = \alpha N$, *then*

$$\sum_{x,d\in\mathbb{Z}} 1_A(x)1_A(x+d)1_A([x+\sqrt{2}d]) = C\alpha^3 N^2 + o(N^2),$$

for some positive constant C.

Example 5.1.5 (Four-term irrational APs)**.** Let

$$L := \begin{pmatrix} 1 & 0 & -\sqrt{2} & -1+\sqrt{2} \\ 0 & 1 & -\sqrt{3} & -1+\sqrt{3} \end{pmatrix}. \tag{5.9}$$

We verify that the non-degeneracy condition from Corollary 5.1.2 is satisfied. Indeed, when L is an m-by-$m+2$ matrix, elementary linear algebra shows that there exists a non-zero row-vector in the row-space of L that has two or fewer non-zero coordinates if and only if there exists some m-by-m submatrix of L that has determinant zero. With L as in (5.9), we see that none of the 6 determinants of the 2-by-2 submatrices are zero, and hence Corollary 5.1.2 applies.

Let N be a natural number, and let $f_1, f_2, f_3, f_4 : [N] \longrightarrow [-1,1]$ be arbitrary functions. Then

$$\frac{1}{N^2} \sum_{\substack{\mathbf{n}\in[N]^4 \\ \|L\mathbf{n}\|_\infty \leqslant \frac{1}{2}}} \Big(\prod_{j=1}^{4} f_j(n_j) \Big) = \frac{1}{N^2} \sum_{x,d\in\mathbb{Z}} f_3(x) f_4(x+d) f_1([x+\sqrt{2}d]) f_2([x+\sqrt{3}d]).$$

Corollary 5.1.2 controls the left-hand side of this expression, and the pattern on the right-hand side, namely

$$(x, x+d, [x+\sqrt{2}d], [x+\sqrt{3}d]), \tag{5.10}$$

we call a four-term irrational progression.

Let A be a subset of $[N]$, with $|A| = \alpha N$. Suppose that$\|f_A\|_{U^3[N]} \leqslant \rho$ (where f_A is as in (5.6)), for some parameter ρ in the range $0 < \rho \leqslant 1$. By telescoping as in (5.7) we may derive

$$\frac{1}{N^2} \sum_{x,d \in \mathbb{Z}} 1_A(x)1_A(x+d)1_A([x+\sqrt{2}d])1_A([x+\sqrt{3}d]) = C\alpha^4(1+o(1)) + O(\rho^{\Omega(1)}) + o_\rho(1) \tag{5.11}$$

for some positive constants C.

One may construct similar results for any pattern

$$(x, x+d, [x+\theta_1 d], \cdots, [x+\theta_k d])$$

where $\theta_i \notin \mathbb{Q}$ for all i.

We comment that the infinitary theory of patterns such as (5.10) was previously considered in [56], albeit in the different language of ergodic theory. In particular, an easy deduction from [56, Theorem B] shows that all sets of natural numbers with positive upper Banach density contain infinitely many copies of the pattern $(x, x+d, [x+\sqrt{2}d], [x+\sqrt{3}d])$. Yet from [56] one cannot recover any statement that has the generality of Corollary 5.1.2, nor an asymptotic formula such as (5.11).

Corollary 5.1.2 has immediate consequences for counting solutions to diophantine inequalities weighted by explicit bounded pseudorandom functions. In particular there is the following natural analogue of [38, Proposition 9.1].

Corollary 5.1.6 (Möbius orthogonality)**.** *Let N, m, d be natural numbers satisfying $d \geqslant m+2$, and let ε be a positive parameter. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be an m-by-d real matrix, with rank m. Suppose that there does not exist any non-zero row-vector in the row-space of L that has two or fewer non-zero coordinates. Let μ denote the Möbius*

function. Then

$$\sum_{\substack{\mathbf{n}\in[N]^d\\ \|L\mathbf{n}\|_\infty\leqslant\varepsilon}} \mu(n_1)\Big(\prod_{j=2}^{d} f_j(n_j)\Big) = o_{L,\varepsilon}(N^{d-m})$$

for any bounded functions $f_2,\cdots,f_d : [N] \longrightarrow [-1,1]$. *The same is true with* μ *replaced by the Liouville function* λ.

Proof. This follows immediately from Corollary 5.1.2 and the deep facts (stated in [38], proved in [39] and [40]) that $\|\mu\|_{U^k[N]} = o_k(1)$ and $\|\lambda\|_{U^k[N]} = o_k(1)$. □

For example, Corollary 5.1.6 implies that

$$\sum_{\substack{\mathbf{n}\in[N]^4\\ n_1-n_2=n_2-n_3\\ |(n_2-n_3)-\sqrt{2}(n_3-n_4)|\leqslant\frac{1}{2}}} \mu(n_1)\mu(n_2)\mu(n_3)\mu(n_4) = o(N^2). \tag{5.12}$$

There are of course many such examples; we chose (5.12) to emphasise that one can choose configurations that combine rational and irrational relations.

5.2 Historical background and the main theorem

The aim of this section is to state Theorem 5.2.10, which is a fully quantitative version of Corollary 5.1.2. We will also state a partial converse to this theorem; this is Theorem 5.2.12.

Before doing this, let us take this opportunity to recall some of the main classical results in the area. As we have already remarked, much is known about the inequality (5.1) for certain special sets A, particularly when $m = 1$. If A is the set of squares, say, it was shown by Davenport and Heilbronn in [18] that there are infinitely many

solutions to (5.1) for $m = 1$ and $d = 5$, i.e. infinitely many solutions to

$$|\lambda_1 n_1^2 + \lambda_2 n_2^2 + \lambda_3 n_3^2 + \lambda_4 n_4^2 + \lambda_5 n_5^2| \leqslant \varepsilon,$$

provided the coefficients λ_i are non-zero, not all of the same sign, and not all in pairwise rational ratio. Their work also proves the same result for k^{th} powers, provided that the number of variables is at least $2^k + 1$. The method is Fourier-analytic, replacing the interval $[-\varepsilon, \varepsilon]$ with a smooth cut-off and expressing the solution count via the inversion formula. See [17, Chapter 20], [87, Chapter 11]. Freeman [23] refined the minor-arc analysis from [18] to obtain asymptotic formulas for the number of solutions where $n_i \leqslant N$ for every i. The number of variables required was subsequently reduced by Wooley in [93].

Of course there is much more work on such polynomial questions, only tangentially related to this chapter, i.e. Margulis' solution to the Oppenheim Conjecture [55], and the subsequent quantitative versions given by Bourgain [11]. Regarding questions with $m \geqslant 2$, Parsell [66] considered the case of A being the k^{th} powers, with Müller [61] developing a refined result in the case of inequalities for general real quadratics. Müller's main result imposes a technical hypothesis on the so-called 'real pencil' of the quadratic forms under consideration: we will return to this issue when considering the technical details of our main theorem (Theorem 5.2.10).

These questions have also been asked when A is the set of prime numbers, and may be tackled using similar analytic tools. A result first claimed in [7] by Baker[1] states that for any fixed positive ε there exist infinitely many triples of primes (p_1, p_2, p_3) satisfying

$$|\lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3| \leqslant \varepsilon, \tag{5.13}$$

[1]In fact Baker proved a slightly different result, writing in [7] that the result we quote here followed easily from the then-existing methods. A proof does not seem to have been written down until Parsell [65].

assuming again that the coefficients λ_i are non-zero, not all of the same sign, and not all in pairwise rational ratio. Parsell [65] then used a similar refinement to that of Freeman to prove a lower bound[2] on the number of solutions to (5.13) satisfying $p_1, p_2, p_3 \leqslant N$. This bound had the expected order of magnitude, namely $\varepsilon N^2 (\log N)^{-3}$.

These analytic approaches ultimately rely on establishing tight mean-value estimates for certain exponential sums, and thus require a large enough number of variables for such estimates to hold. In the case of primes, say, for m inequalities the method of Parsell will yield an asymptotic for the number of solutions to (5.1) in prime variables provided $d \geqslant 2m+1$ (at least for generic L). In preparation, we have a paper [89] that reaches the same conclusion under the weaker hypothesis $d \geqslant m+2$ (provided L has algebraic coefficients).

Having introduced the background of this work, we can begin to build up the necessary notation required in order to state the main theorem (Theorem 5.2.10). First, let us introduce a multilinear form that will count solutions to a general version of (5.1).

Definition 5.2.1. *Let N, m, d be natural numbers. Let ε be positive, and let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a linear map. Let $F : \mathbb{R}^d \longrightarrow [0,1]$ and $G : \mathbb{R}^m \longrightarrow [0,1]$ be two functions, with F supported on $[-N, N]^d$ and G supported on $[-\varepsilon, \varepsilon]^m$. Let $f_1, \cdots, f_d : [N] \longrightarrow [-1, 1]$ be arbitrary functions. We define*

$$T^L_{F,G}(f_1, \cdots, f_d) := \frac{1}{N^{d-m}} \sum_{\mathbf{n} \in \mathbb{Z}^d} \Big(\prod_{j=1}^{d} f_j(n_j) \Big) F(\mathbf{n}) G(L\mathbf{n}).$$

The normalisation factor of N^{d-m} is appropriate; we will show in Lemma 5.4.1 that $T^L_{F,G}(f_1, \cdots, f_d) \ll_\varepsilon 1$.

[2] An asymptotic formula for the number of solutions follows very easily from Parsell's work, though does not appear to be present in the literature.

Now let us introduce the appropriate notions of 'non-degeneracy'. These will be needed in order to appropriately quantify the Gowers norm relations in the main theorem (Theorem 5.2.10).

Definition 5.2.2 (Rank varieties)**.** *Let m, d be natural numbers satisfying $d \geqslant m+1$. Let $V_{\text{rank}}(m,d)$ denote the set of all linear maps $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ whose rank is less than m. We call $V_{\text{rank}}(m,d)$ the* rank variety*.*

Let $V^{\text{global}}_{\text{rank}}(m,d)$ denote the set of all linear maps $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ for which there exists a standard basis vector of $\mathbb{R}^d$, say $\mathbf{e_i}$, for which $L|_{\text{span}(\mathbf{e_j}:j\neq i)}$ has rank less than m. We call $V^{\text{global}}_{\text{rank}}(m,d)$ the global rank variety*.*

We remark that $V^{\text{global}}_{\text{rank}}(m,d)$ contains $V_{\text{rank}}(m,d)$.

The next notion is a rephrasing of the non-degeneracy condition that appeared in Corollary 5.1.2 and Theorem 5.1.1.

Definition 5.2.3 (Dual degeneracy variety)**.** *Let m, d be natural numbers satisfying $d \geqslant m+2$. Let $\mathbf{e_1}, \cdots, \mathbf{e_d}$ denote the standard basis vectors of $\mathbb{R}^d$, and let $\mathbf{e_1^*}, \cdots, \mathbf{e_d^*}$ denote the dual basis of $(\mathbb{R}^d)^*$. Then let $V^*_{\text{degen}}(m,d)$ denote the set of all linear maps $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ for which there exist two indices $i, j \leqslant d$, and some real number λ, such that $\mathbf{e_i}^* - \lambda \mathbf{e_j}^*$ is non-zero and $(\mathbf{e_i}^* - \lambda \mathbf{e_j}^*) \in L^*((\mathbb{R}^m)^*)$. We call $V^*_{\text{degen}}(m,d)$ the* dual degeneracy variety*.*

Though defined as sets of linear maps, by fixing bases we can view $V_{\text{rank}}(m,d)$ and $V^*_{\text{degen}}(m,d)$ as sets of matrices. In that language, one can easily verify that an m-by-d matrix L is in $V^*_{\text{degen}}(m,d)$ precisely when there exists a non-zero row-vector in the row-space of L that has two or fewer non-zero coordinates. The formulation in terms of dual spaces will be particularly convenient for some of the algebraic manipulations

in section 5.5, however. We remark that $V^{*}_{\text{degen}}(m,d)$ contains $V^{\text{global}}_{\text{rank}}(m,d)$.

If $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ is a surjective linear map, it is certainly true that $\text{span}(L(\mathbb{Z}^d)) = \mathbb{R}^m$. But $L(\mathbb{Z}^d)$ needn't be dense in $\mathbb{R}^m$, as it may satisfy some rational relations.

Definition 5.2.4 (Rational dimension, rational map, purely irrational)**.** *Let m and d be natural numbers, with $d \geqslant m+1$. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map. Let u denote the largest integer for which there exists a surjective linear map $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^u$ for which $\Theta L(\mathbb{Z}^d) \subseteq \mathbb{Z}^u$. We call u the* rational dimension *of L, and we call any map Θ with the above property a* rational map *for L. We say that L is* purely irrational *if $u = 0$.*

For example, suppose that $L : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ is the linear map represented by the matrix

$$L := \begin{pmatrix} 1 & 0 & -\sqrt{2} & -\sqrt{3}+1 \\ 0 & 1 & 5\sqrt{2} & 5\sqrt{3} \end{pmatrix}.$$

If $\Theta : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is given by the matrix

$$\Theta := \begin{pmatrix} 5 & 1 \end{pmatrix},$$

then $\Theta L(\mathbb{Z}^4) \subseteq \mathbb{Z}$, and in fact $\Theta L(\mathbb{Z}^4) = \mathbb{Z}$. So the rational dimension of L is at least 1. But the rational dimension of L cannot be 2, as if there were a surjective map $\Theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ such that $\Theta L(\mathbb{Z}^4) \subseteq \mathbb{Z}^2$, then $L(\mathbb{Z}^4)$ would be the subset of a 2-dimensional lattice, which it is not. So the rational dimension of L is equal to 1.

Earlier in this section we remarked that Müller, in the work [61], imposed a technical hypothesis on the so-called 'real pencil' of the quadratic forms under consideration. In our language, Müller was trying find conditions for when $T^{L}_{F,G}(f_1, \cdots, f_d) > 0$ in the case where the functions f_j are supported on the image of quadratic monomials.

The hypothesis he imposed on L was exactly that L should be purely irrational.[3] We work in a more general framework, considering all L, including those that are not purely irrational. As will become apparent, this is significantly more complex.

In our definition of rational dimension, there is some flexibility over the exact choice of map Θ. The next lemma identifies an invariant.

Lemma 5.2.5. *Let m and d be natural numbers, with $d \geqslant m+1$. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and let u be the rational dimension of L. Then if $\Theta_1, \Theta_2 : \mathbb{R}^m \longrightarrow \mathbb{R}^u$ are two rational maps for L, $\ker \Theta_1 = \ker \Theta_2$.*

Proof. Suppose that $\Theta_1, \Theta_2 : \mathbb{R}^m \longrightarrow \mathbb{R}^u$ are two rational maps for L for which $\ker \Theta_1 \neq \ker \Theta_2$. Then consider the map $(\Theta_1, \Theta_2) : \mathbb{R}^m \longrightarrow \mathbb{R}^{2u}$. The kernel of this map has dimension at most $m-u-1$, as it is the intersection of two different subspaces of dimension $m-u$. Therefore the image has dimension at least $u+1$.

Also, $((\Theta_1, \Theta_2) \circ L)(\mathbb{Z}^d) \subseteq \mathbb{Z}^{2u}$. Let Φ be any surjective map from $\mathrm{im}((\Theta_1, \Theta_2))$ to $\mathbb{R}^{u+1}$ for which $\Phi(\mathbb{Z}^{2u} \cap \mathrm{im}((\Theta_1, \Theta_2))) \subseteq \mathbb{Z}^{u+1}$. Then $\Phi \circ (\Theta_1, \Theta_2) : \mathbb{R}^m \longrightarrow \mathbb{R}^{u+1}$ is surjective and $(\Phi \circ (\Theta_1, \Theta_2) \circ L)(\mathbb{Z}^d) \subseteq \mathbb{Z}^{u+1}$. This contradicts the definition of u as the rational dimension. □

The quantitative aspects of such relations will be required in order to properly state the main theorem (Theorem 5.2.10). Recall that for all linear maps between vector spaces of the form $\mathbb{R}^a$, we identify them with their matrix representation with respect to the standard bases. Also recall that for a linear map $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^u$, we use $\|\Theta\|_\infty$ to denote the maximum absolute value of the coefficients of its matrix.

Definition 5.2.6 (Rational complexity)**.** *Let m and d be natural numbers, with $d \geqslant m+1$. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and let u denote the rational*

[3] He also has conditions on the rank of the quadratic form obtained by combining the monomials on which the f_j are supported with the matrix L.

dimension of L. We say that L has rational complexity at most C *if there exists a map Θ that is a rational map for L and for which $\|\Theta\|_\infty \leqslant C$.*

If L is purely irrational, then L has rational complexity 0.

A linear map with maximal rational dimension is equivalent to a linear map with integer coefficients, in the following sense:

Lemma 5.2.7. *Let m and d be natural numbers, with $d \geqslant m+1$. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and suppose that L has rational dimension m and rational complexity at most C. Then there exists an invertible m-by-m matrix Θ and an m-by-d matrix S with integer coefficients such that, as matrices, $\Theta L = S$. Furthermore, $\|\Theta\|_\infty \leqslant C$.*

Proof. Let $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ be a rational map for L for which $\|\Theta\|_\infty \leqslant C$. □

We will use this lemma in section 5.5, to reduce the study of maps L with maximal rational dimension to the study of maps L with integer coefficients, which were considered in [38] (see Theorems 5.1.1 and 5.5.2).

We must quantify the rational relations in a second way. Indeed, L might have rational dimension u but be extremely close to having rational dimension at least $u+1$, in the sense that there might exist some surjective linear map $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^{u+1}$ such that the matrix of ΘL is very close to having integer coefficients. This phenomenon, essentially a notion of diophantine approximation, will also have a quantitative effect on our final bounds. We introduce the following definition:

Definition 5.2.8 (Approximation function)**.** *Let m and d be natural numbers, with $d \geqslant m+1$. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and let u denote the rational dimension of L. Let $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^u$ be any rational map for L. Suppose*

that $u \leqslant m-1$. *We then define the* approximation function of L, *denoted* $A_L : (0,1] \times (0,1] \longrightarrow (0,\infty)$ *by*

$$A_L(\tau_1,\tau_2) := \inf_{\substack{\varphi\in(\mathbb{R}^m)^* \\ \mathrm{dist}(\varphi,\Theta^*((\mathbb{R}^u)^*))\geqslant\tau_1 \\ \|\varphi\|_\infty\leqslant\tau_2^{-1}}} \mathrm{dist}(L^*\varphi,(\mathbb{Z}^d)^T),$$

where $(\mathbb{Z}^d)^T$ *denotes the set of those* $\varphi \in (\mathbb{R}^d)^*$ *that have integer coordinates with respect to the standard dual basis.*

If $u = m$, *we define* $A_L(\tau_1,\tau_2)$ *to be identically equal to* τ_1.

Let us unpack this definition, before giving some examples. Firstly, note that the definition is independent of the choice of Θ. Indeed, $\Theta^*((\mathbb{R}^u)^*) = (\ker\Theta)^\circ$ which, by Lemma 5.2.5, is independent of Θ. Regarding the notion 'dist', we remind the reader that we consider a-by-b matrices M as elements of $\mathbb{R}^{ab}$, simply by identifying the coefficients of M with coordinates in $\mathbb{R}^{ab}$. The ℓ^∞ norm and the dist operator may then be defined on matrices, i.e. if V is a set of a-by-b matrices, and L is an a-by-b matrix, then

$$\mathrm{dist}(L,V) := \inf_{L'\in V}\|L-L'\|_\infty.$$

In this instance we are working with 1-by-d matrices, i.e. elements of $(\mathbb{R}^d)^*$.

Let us consider a simple example. Suppose that, as a matrix,

$$L := \begin{pmatrix} 1 & -\sqrt{2} & -1+\sqrt{2} \end{pmatrix}, \tag{5.14}$$

as in Example 5.1.3. Then $A_L(\tau_1,\tau_2)$ is equal to

$$\inf_{k\in\mathbb{R}:\tau_1\leqslant|k|\leqslant\tau_2^{-1}} \max(\|k\|_{\mathbb{R}/\mathbb{Z}}, \|-k\sqrt{2}\|_{\mathbb{R}/\mathbb{Z}}, \|-k+k\sqrt{2}\|_{\mathbb{R}/\mathbb{Z}}).$$

We claim that

$$A_L(\tau_1, \tau_2) \gg \min(\tau_1, \tau_2).$$

Indeed, we know that, for all $q \in \mathbb{N}$, $\|q\sqrt{2}\|_{\mathbb{R}/\mathbb{Z}} \geqslant 1/(10q)$. This is the statement that $\sqrt{2}$ is a badly approximable irrational. The proof is straightforward: if there were some natural number p for which $|q\sqrt{2} - p| < 1/(10q)$, then

$$1 \leqslant |2q^2 - p^2| < \frac{\sqrt{2}}{10} + \frac{p}{10q} < \frac{\sqrt{2}}{5} + \frac{1}{10},$$

which is a contradiction.

Suppose first that $\|k\|_{\mathbb{R}/\mathbb{Z}} \leqslant \tau_2/100$ and $1/2 \leqslant |k| \leqslant \tau_2^{-1}$. Then, replacing k by $[k]$ (the nearest integer to k), we can conclude that

$$\begin{aligned}
\max(\| - k\sqrt{2}\|_{\mathbb{R}/\mathbb{Z}}, \| - k + k\sqrt{2}\|_{\mathbb{R}/\mathbb{Z}}) &\geqslant \|[k]\sqrt{2}\|_{\mathbb{R}/\mathbb{Z}} - \frac{\tau_2}{50} \\
&\geqslant \frac{1}{10[k]} - \frac{\tau_2}{50} \\
&\geqslant \frac{1}{10\tau_2^{-1} + 10} - \frac{\tau_2}{50} \\
&\gg \tau_2.
\end{aligned}$$

Otherwise, one has

$$\|k\|_{\mathbb{R}/\mathbb{Z}} \gg \min(\tau_1, \tau_2).$$

Therefore,

$$A_L(\tau_1, \tau_2) \gg \min(\tau_1, \tau_2)$$

as claimed.

Such a function is clearly rather tame. In fact, it is not difficult[4] to show that if

[4]If L is not purely irrational then one needs to employ the dimension reduction argument from Lemma 5.5.10 in addition to an easy diophantine approximation argument. This lemma is a lengthy piece of elementary linear algebra.

L is an m-by-d matrix with rank m and with algebraic coefficients, then

$$A_L(\tau_1, \tau_2) \gg_L \min(\tau_1, \tau_2^{O_L(1)}), \tag{5.15}$$

where the $O_L(1)$ term in the exponent depends on the algebraic degree of the coefficients[5] of L. We shall sketch a proof of this statement in section 5.12. In general, however, $A_L(\tau_1, \tau_2)$ could tend to zero arbitrarily quickly as τ_2 tends to zero, for example in the case when $L = \begin{pmatrix} 1 & -\lambda & -1+\lambda \end{pmatrix}$ and λ is a Liouville number (an irrational number that may be very well approximated by rationals).

Yet, however fast $A_L(\tau_1, \tau_2)$ decays, we have the following critical claim:

Claim 5.2.9. *For all permissible choices of L, τ_1 and τ_2 in Definition 5.2.8, $A_L(\tau_1, \tau_2)$ is positive.*

Proof. Let u be the rational dimension of L. Without loss of generality we may assume that $u \leqslant m-1$. Then, for all $\varphi \in (\mathbb{R}^m)^* \backslash \Theta^*((\mathbb{R}^u)^*)$ we have that $\mathrm{dist}(L^*\varphi, (\mathbb{Z}^d)^T) > 0$. (If this were not the case then the map $(\Theta, \varphi) : \mathbb{R}^m \longrightarrow \mathbb{R}^{u+1}$ would contradict the definition of u.) Therefore, as the definition of $A_L(\tau_1, \tau_2)$ involves taking the infimum of a positive continuous function over a compact set, $A_L(\tau_1, \tau_2)$ is positive. □

One might ask why we chose to formulate Definition 5.2.8 in terms of a general $\varphi \in (\mathbb{R}^m)^*$, instead of one with integer coordinates, when in practice the calculation of $A_L(\tau_1, \tau_2)$ quickly reduces to considering those φ with integer coordinates. This will certainly be true in the one lemma of this chapter where A_L plays a significant role, namely Lemma 5.4.3. Our first reason is that we find the definition as stated more natural, in that it does not presuppose that any of the coordinates of L are integers; our second reason is that, when one comes to apply these ideas to the setting of

[5]One could perhaps remove this dependence by using the Schmidt subspace theorem, though as there are power losses throughout the rest of the argument there does not seem to be a great advantage in doing so.

the primes, one is drawn to estimate certain sieve expressions using the Davenport-Heilbronn method. This method involves estimating an integral over $\varphi \in (\mathbb{R}^m)^*$, where one wishes to control the minor arc contribution by $A_L(\tau_1, \tau_2)$, and so it is natural that the variable φ should be allowed to vary continuously. More details will appear in [89].

Having laid the necessary groundwork, we may now state the main theorem of this chapter.

Theorem 5.2.10 (Main Theorem)**.** *Let N, m, d be natural numbers, satisfying $d \geqslant m+2$, and let ε, c, C, C' be positive reals. Let $L = L(N) : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map that satisfies $\|L\|_\infty \leqslant C$. Let $A_L : (0,1] \times (0,1] \longrightarrow (0,\infty)$ be the approximation function of L. Suppose further that*
$\mathrm{dist}(L, V^*_{\mathrm{degen}}(m,d)) \geqslant c$, *and that L has rational complexity at most C'. Then there exists a natural number s at most $d-2$ such that the following is true. Let $F : \mathbb{R}^d \longrightarrow [0,1]$ be the indicator function of $[1,N]^d$, and let $G : \mathbb{R}^m \longrightarrow [0,1]$ be the indicator function of a convex domain contained in $[-\varepsilon, \varepsilon]^m$. Let $f_1, \cdots, f_d : [N] \longrightarrow [-1,1]$ be arbitrary functions, and suppose that*

$$\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho,$$

for some parameter ρ in the range $0 < \rho \leqslant 1$. Then

$$T^L_{F,G}(f_1, \cdots, f_d) \ll_{c,C,C',\varepsilon} \rho^{\Omega(1)} + o_{\rho,A_L,c,C,C'}(1) \tag{5.16}$$

as N tends to infinity. The $o_{\rho,A_L,c,C,C'}(1)$ term may be bounded above by

$$N^{-\Omega(1)} \rho^{-O(1)} A_L(\Omega_{c,C,C'}(1), \rho)^{-1}.$$

We remind the reader that the implied constants may depend on the dimensions m and d. Also note that in the above statement one may replace C and C' by a single constant C, without weakening the conclusion. We proceed with this assumption. Observe also that the non-degeneracy condition $\mathrm{dist}(L, V^*_{\mathrm{degen}}(m,d)) \geqslant c$ is a quantitative refinement of the non-degeneracy condition on the row-space of L in Theorem 5.1.1 and Corollary 5.1.2.

Since $A_L(\Omega_{c,C}(1), \rho)^{-1}$ is finite (by Claim 5.2.9), Theorem 5.2.10 immediately implies Corollary 5.1.2 from the start of this chapter. From (5.15), or rather our full quantitative version Lemma 5.12.1, we also have the following corollary:

Corollary 5.2.11. *Assume the same hypotheses as Theorem 5.2.10, and assume further that L has algebraic coefficients with algebraic degree at most k. Let H denote the maximum absolute value of all of the coefficients of all of the minimal polynomials of the coefficients of L. Then*

$$T^L_{F,G}(f_1, \cdots, f_d) \ll_{c,C,\varepsilon,H} \rho^{\Omega(1)} + N^{-\Omega(1)}\rho^{-O_k(1)}$$

as N tends to infinity.

At this juncture, it might not be clear why so many quantitative non-degeneracy conditions were required in the statement of Theorem 5.2.10. To try to illuminate this issue, we will also prove the following partial converse to Theorem 5.2.10, demonstrating that the non-degeneracy condition involving $V^*_{\mathrm{degen}}(m,d)$ is necessary in order to use Gowers norms in this way.

Theorem 5.2.12. *Let m, d be natural numbers, satisfying $d \geqslant m+2$, and let ε, c, C be positive constants. For each natural number N, let $L = L(N) : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a linear map satisfying $\|L\|_\infty \leqslant C$. Let $F : \mathbb{R}^d \longrightarrow [0,1]$ denote the indicator function*

of $[1,N]^d$ *and* $G : \mathbb{R}^m \longrightarrow [0,1]$ *denote the indicator function of* $[-\varepsilon,\varepsilon]^m$. *Assume further that* $\mathrm{dist}(L, V_{\mathrm{rank}}(m,d)) \geqslant c$ *and that* $T^L_{F,G}(1,\cdots,1) \gg_{c,C,\varepsilon} 1$ *for large enough* N.

Suppose that

$$\liminf_{N\to\infty} \mathrm{dist}(L, V^*_{\mathrm{degen}}(m,d)) = 0.$$

Let s *be a natural number, and let* $H : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ *be any function satisfying* $H(\rho) = \kappa(\rho)$, *and let* $E_\rho(N)$ *denote some error term depending on a parameter* ρ, *such that* $E_\rho(N) = o_\rho(1)$. *Then one can find infinitely many natural numbers* N *such that there exist functions* $f_1,\cdots,f_d : [N] \to [-1,1]$ *and some* ρ *at most* 1 *such that both*

$$\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho$$

and

$$|T^L_{F,G}(f_1,\cdots,f_d)| > H(\rho) + E_\rho(N). \tag{5.17}$$

In other words, the conclusion of Theorem 5.2.10 cannot possibly hold if $\mathrm{dist}(L, V^*_{\mathrm{degen}}(m,d))$ is arbitrarily close to 0, even if one replaces the $\rho^{\Omega(1)}$ dependence on ρ with a function $H(\rho)$ that could potentially decay to zero arbitrarily slowly as ρ tends to zero.

Example 5.2.13. Suppose

$$L = \begin{pmatrix} 1+N^{-1} & \sqrt{3}+N^{-\frac{1}{2}} & \pi & -\pi+\sqrt{2} \\ 2 & 2\sqrt{3}+N^{-\frac{1}{2}} & -\sqrt{5} & e \end{pmatrix}.$$

Then L has rank 2 and $L \notin V^*_{\mathrm{degen}}(2,4)$. If one considers Theorem 5.1.1, one might therefore hope to apply the theory of Gowers norms to bound the number of solutions to inequalities given by L. However, by considering perturbations of the first two columns, we see that $\mathrm{dist}(L, V^*_{\mathrm{degen}}(2,4)) = o(1)$. (Indeed, one may perturb L

by $O(N^{-1/2})$ such that there is a vector $(0, 0, x_3, x_4)$ in the row space). Therefore Theorem 5.2.10 does not apply in this case, despite the fact that $L \notin V^*_{\text{degen}}(2, 4)$. Furthermore, Theorem 5.2.12 shows that, in fact, we cannot possibly use Gowers norms to control inequalities given by such an L. This example is informative, as it shows us that whatever methods we use to prove Theorem 5.2.10, these methods must break down when applied to such an L, despite the fact that $L \notin V^*_{\text{degen}}$.

The proof of Theorem 5.2.10 will be rather involved. It is tempting to think that the result would follow more easily from taking rational approximations of the coefficients of L, and then using the existing Generalised von Neumann Theorem (a quantitative version of Theorem 5.1.1) as a black box. Though of course we cannot completely rule out an alternative approach to that of this chapter, it seems that such an argument will only quickly succeed if the coefficients of L are all extremely well-approximable, else the height of the rational approximations becomes too great to apply [38, Theorem 7.1]. One must find an alternative method for other maps L.

To finish this introduction, and to assist the reader, we now describe the overall structure of the chapter, and also indicate our proof strategy.

If in the statement of Theorem 5.2.10 one replaces the convex cut-offs F and G with Lipschitz cut-offs, then the expression $T^L_{F,G}(f_1, \cdots, f_d)$ may be bounded by Gowers norms by a relatively straightforward argument, which we present in sections 5.6 through 5.8. In section 5.6 we introduce a new approximation argument, in which we replace the solution count $T^L_{F,G}(f_1, \cdots, f_d)$ by a related solution count $\widetilde{T}^L_{F,G}(\widetilde{f}_1, \cdots, \widetilde{f}_d)$, which, rather than being a discrete summation over $\mathbb{Z}^d$, is an integral over $\mathbb{R}^d$. The expression $\widetilde{T}^L_{F,G}(\widetilde{f}_1, \cdots, \widetilde{f}_d)$ may be analysed using the Cauchy-Schwarz inequality in a way that is almost identical to the proof of the usual Generalised von

Neumann Theorem [38, Theorem 7.1], and we perform this manipulation in section 5.8. This argument makes no mention of the approximation function A_L.

So it remains to reduce Theorem 5.2.10 to the version with Lipschitz cut-offs (we explicitly state this version in Theorem 5.5.6). Unfortunately, if L is not purely irrational then there are substantial technical difficulties in replacing G with a Lipschitz cut-off. To circumvent these difficulties, in section 5.5 we give an intricate (though ultimately elementary) linear algebraic argument that reduces Theorem 5.2.10 to the case where L is purely irrational, at which point one may replace the functions F and G with Lipschitz cut-offs with relative ease. This argument thus resolves the main theorem (Theorem 5.2.10), and all the associated corollaries.

The other sections contain supporting material, discussions, and lemmas. In section 5.3, we introduce two particular pieces of quantitative linear algebra that will be required in the main argument. Though some of these notions are a little delicate (such as our quantitative notion of Cauchy-Schwarz complexity), the proofs reduce to standard arguments, and we defer them to section 5.10. The attack on Theorem 5.2.10 proper begins in earnest in section 5.4, in which we give three upper bounds for $T^L_{F,G}(1, \cdots, 1)$, each valid under different regimes. The properties of L relating to diophantine approximation, in particular to the approximation function A_L, become apparent in Lemma 5.4.3.

Section 5.9 deals solely with the proof of the partial converse, namely Theorem 5.2.12, and may be read largely independently of the other sections. Using a semi-random method, we explicitly construct functions $f_1, \cdots, f_d$ that satisfy (5.17).

The final three sections may be viewed as an appendix. The first contains the proofs of the statements from section 5.3; the second contains an assortment of other short linear algebraic lemmas; and the third illustrates how one may control the approximation function A_L in the case when L has algebraic coefficients.

Remark 5.2.14. Many of the implied constants throughout the chapter will depend on the parameter ε from the statement of Theorem 5.2.10. Ultimately, the implied constant in (5.16) tends to infinity as ε tends to zero, as our approximation argument in section 5.6 will not be efficient in powers of ε. Yet, to save our notation from becoming unreadable, we choose not to keep track of the precise behaviour of implied constants involving ε.

5.3 Rank matrix and normal form

Here we introduce some technical notions, of a linear algebraic nature, that will help us to quantify the manipulations to come. We will state the main propositions required later, but, as the proofs reduce to a close analysis of known algorithms and would no doubt be obvious to the expert, we defer them to section 5.10.

Firstly, to prove Theorem 5.2.10 it will be useful to introduce a concept that we will refer to as the *rank matrix* of L. Consider the following basic fact: if m and d are natural numbers with $d \geqslant m+1$, and if an m-by-d matrix L is assumed to have row rank m, L also has column rank m, and thus we may find m linearly independent columns in L which determine an m-by-m submatrix M with non-zero determinant. When L depends on the quantity N, we use the term *rank matrix* to refer to a quantitative refinement of this idea.

Proposition 5.3.1 (Rank matrix)**.** *Let m, d be natural numbers, with $d \geqslant m+1$. Let c, C be positive constants. For a natural number N, let $L = L(N) : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and assume that $\|L\|_\infty \leqslant C$ and $\mathrm{dist}(L, V_{\mathrm{rank}}(m,d)) \geqslant c$. Let the coefficients of L be denoted $(\lambda_{ij})_{i\leqslant m, j\leqslant d}$. Then there exists a matrix M, an m-by-m submatrix of L, with the following properties:*

(1) $|\det M| = \Omega_{c,C}(1)$*;*

(2) $\|M^{-1}\|_\infty = O_{c,C}(1)$*;*

(3) Let $\mathbf{v} \in \mathbb{R}^d$ *be a vector such that* $\mathbf{v}^T$ *is in the row-space of* L*. Suppose that* $\|\mathbf{v}\|_\infty \leqslant C_1$*. Then for all* i *in the range* $1 \leqslant i \leqslant m$ *there exist coefficients* a_i *satisfying* $|a_i| = O_{c,C,C_1}(1)$ *such that* $\sum_{i=1}^{m} a_i \lambda_{ij} = v_j$ *for all* j *in the range* $1 \leqslant j \leqslant d$*.*

We call such a matrix M *a* rank matrix *of* L*.*

If L *satisfies the stronger hypothesis* $\mathrm{dist}(L, V_{\mathrm{rank}}^{\mathrm{global}}(m,d)) \geqslant c$*, then, for each* j*, there exists a rank matrix of* L *that doesn't include the* j^{th} *column of* L*.*

We defer the proof to section 5.10.

The second technical condition that we will describe in this section involves reparametrising certain linear forms. This reparametrisation will become important in Proposition 5.8.3, when we come to apply the Cauchy-Schwarz inequality.

Let $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map. Putting the standard coordinates on $\mathbb{R}^n$ and $\mathbb{R}^m$, we may write $(\psi_1, \cdots, \psi_m) := \Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ as a system of homogeneous linear forms. The crux of the theory from [38] is that, provided Ψ is of 'finite Cauchy-Schwarz complexity', Ψ admits an extension which is in so-called 'normal form'. Below we will give a brief overview of this terminology, before introducing our own quantitative versions; a much fuller discussion may be found in [38, Section 1] and [34].

In words, a reparametrisation into normal form is one in which each linear form is the only one that mentions all of its particular collection of variables. For example, the forms

$$\psi_1(t,u,v) = u + v$$

$$\psi_2(t,u,v) = v + t$$
$$\psi_3(t,u,v) = u + t$$
$$\psi_4(t,u,v) = u + v + t \tag{5.18}$$

are in normal form with respect to ψ_4, since ψ_4 is the only form to utilise all three of the variables. However, this system is not in normal form with respect to ψ_3, say. However, the system

$$\psi_1(t,u,v,w) = u + v + 2w$$
$$\psi_2(t,u,v,w) = v + t - w$$
$$\psi_3(t,u,v,w) = u + t - w$$
$$\psi_4(t,u,v,w) = u + v + t, \tag{5.19}$$

that parametrises the same subspace of $\mathbb{R}^4$, *is* in normal form for all i.

We repeat the precise definition from [38].

Definition 5.3.2. *Let m, n be natural numbers, and let $(\psi_1, \cdots, \psi_m) = \Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a system of homogeneous linear forms. Let $i \in [m]$. We say that Ψ is in normal form with respect to ψ_i if there exists a non-negative integer s and a collection $J_i \subseteq \{\mathbf{e_1}, \cdots, \mathbf{e_n}\}$ of the standard basis vectors, satisfying $|J_i| = s + 1$, such that*

$$\prod_{\mathbf{e} \in J_i} \psi_{i'}(\mathbf{e})$$

is non-zero when $i' = i$ and vanishes otherwise. We say that Ψ is in normal form if it is in normal form with respect to ψ_i for every i.

Let us also recall what it means for a certain system of forms Ψ' to extend the system of forms Ψ.

Definition 5.3.3. *For a system of linear forms $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, an extension Ψ' is a system of linear forms on $\mathbb{R}^{n'}$, for some n' with $n' \geqslant n$, such that*

(1) $\Psi'(\mathbb{R}^{n'}) = \Psi(\mathbb{R}^n)$*;*

(2) if we identify $\mathbb{R}^n$ with the subset $\mathbb{R}^n \times \{0\}^{n'-n}$ in the obvious manner, then Ψ is the restriction of Ψ' to this subset.

The paper [38] includes a result (Lemma 4.4) on the existence of extensions in normal form, but we will need a quantitative refinement of this analysis.

The reader will note from examples (5.18) and (5.19) that the property of 'being in normal form' is a property of the parametrisation, and not of the underlying space that is being parametrised. It is natural to wonder whether there is some property of a space that can enable one to find a parametrisation in normal form, even if one's original parametrisation is not. Fortunately there is such a notion, and it is the finite (Cauchy-Schwarz) complexity introduced in [38]. We introduce this notion in the following definitions, which we have phrased in such a way as to help us formulate a quantitative version.

Definition 5.3.4 (Suitable partitions)**.** *Let m, n be natural numbers, with $m \geqslant 2$, and let $(\psi_1, \cdots, \psi_m) = \Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a system of homogeneous linear forms. Fix $i \in [m]$. Let $\mathcal{P}_i$ be a partition of $[m] \setminus \{i\}$, i.e.*

$$[m] \setminus \{i\} = \bigcup_{k=1}^{s+1} \mathcal{C}_k$$

for some s satisfying $0 \leqslant s \leqslant m-2$ and some disjoint sets $\mathcal{C}_k$. We say that $\mathcal{P}_i$ is suitable *for Ψ if*

$$\psi_i \notin \operatorname{span}_{\mathbb{R}}(\psi_j : j \in \mathcal{C}_k)$$

for any k.

Definition 5.3.5 (Degeneracy varieties)**.** *Let m, n be natural numbers, with $m \geqslant 2$. Let $\mathcal{P}_i$ be a partition of $[m] \setminus \{i\}$. We define the* $\mathcal{P}_i$-degeneracy variety *$V_{\mathcal{P}_i}$ to be the set of all the systems of homogeneous linear forms $\Psi : \mathbb{R}^n \to \mathbb{R}^m$ for which $\mathcal{P}_i$ is not suitable for Ψ. Finally, the* degeneracy variety *V_{degen} is given by*

$$V_{\text{degen}}(n, m) := \bigcup_{i=1}^{m} \bigcap_{\mathcal{P}_i} V_{\mathcal{P}_i},$$

where the inner intersection is over all possible partitions $\mathcal{P}_i$.

It is easy to observe that $\Psi \in V_{\text{degen}}$ if and only if, for some $i \neq j$, ψ_i is a real multiple of ψ_j.

In [38, Definition 1.5], the authors refer to those $\Psi \in V_{\text{degen}}(n, m)$ as having infinite (Cauchy-Schwarz) complexity, and develop their theory for $\Psi \notin V_{\text{degen}}(n, m)$. As we did for describing degeneracy properties of L, we need to quantify such a notion.

Definition 5.3.6 (c_1-Cauchy-Schwarz complexity)**.** *Let m, n be natural numbers, with $m \geqslant 3$, and let c_1 be a positive constant. Let $(\psi_1, \cdots, \psi_m) = \Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a system of homogeneous linear forms. For $i \in [m]$, we define a quantity s_i either by defining $s_i + 1$ to be the minimal number of parts in a partition $\mathcal{P}_i$ of $[m] \setminus \{i\}$ such that $\text{dist}(\Psi, V_{\mathcal{P}_i}) \geqslant c_1$, or by $s_i = \infty$ if no such partition exists. Then we define $s := \max(1, \max_i s_i)$. We say that s is the c_1-Cauchy-Schwarz complexity of Ψ.*

We remark, for readers familiar with [38], that we preclude the 'complexity 0' case. This is for a mundane technical reason, that occurs when absorbing the exponential phases in section 5.8, when it will be convenient that $s + 1 \geqslant 2$. This is why we need the condition $m \geqslant 3$ in the above definition. We also take this opportunity to note that if s satisfies the above definition, and $s \neq \infty$, then $2 \leqslant s + 1 \leqslant m - 1$.

We note an easy consequence of these definitions.

Lemma 5.3.7. *Let m, n be natural numbers, with $m \geqslant 3$, and let c_1 be a positive constant. Let $\Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a system of homogeneous linear forms. Suppose that* $\mathrm{dist}(\Psi, V_{\mathrm{degen}}(n,m)) \geqslant c_1$. *Then Ψ has finite c_1-Cauchy-Schwarz complexity.*

Proof. It is easy to observe that $\Psi \in V_{\mathrm{degen}}(n,m)$ if and only if, for some $i \neq j$, ψ_i is a real multiple of ψ_j. From now until the end of the proof, fix $\mathcal{P}_i$ to be the partition of $[m]\backslash\{i\}$ in which every form ψ_k is in its own part. Our initial observation then implies that $\Psi \in V_{\mathrm{degen}}(n,m)$ if and only if $\Psi \in V_{\mathcal{P}_i}$ for some i. So $\mathrm{dist}(\Psi, V_{\mathrm{degen}}(n,m)) \geqslant c_1$ implies that $\mathrm{dist}(\Psi, V_{\mathcal{P}_i}) \geqslant c_1$ for all i. Therefore, by using these partitions $\mathcal{P}_i$ in Definition 5.3.6, we conclude that Ψ has finite c_1-Cauchy-Schwarz complexity. □

After having built up these definitions, we state the key proposition on the existence of normal form extensions to systems of real linear forms. We remind the reader that all implied constants may depend on the dimensions of the underlying spaces.

Proposition 5.3.8 (Normal form algorithm)**.** *Let m, n be natural numbers, with $m \geqslant 3$, and let c_1, C_1 be positive constants. Let $(\psi_1, \cdots, \psi_m) = \Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a system of homogeneous linear forms, and suppose that the coefficients of Ψ are bounded above in absolute value by C_1. Furthermore, suppose that Ψ has c_1-Cauchy-Schwarz complexity s, for some finite s. Then, for each $i \in [m]$, there is an extension $\Psi' : \mathbb{R}^{n'} \longrightarrow \mathbb{R}^m$ such that:*

(1) $n' = n + s + 1 \leqslant n + m - 1$;

(2) Ψ' is of the form

$$\Psi'(\mathbf{u}, w_1, \cdots, w_{s+1}) := \Psi(\mathbf{u} + w_1\mathbf{f_1} + \cdots + w_{s+1}\mathbf{f_{s+1}})$$

for some vectors $\mathbf{f_k} \in \mathbb{R}^n$, such that $\|\mathbf{f_k}\|_\infty = O_{c_1,C_1}(1)$ for every k;

(3) Ψ' is in normal form with respect to ψ_i';

(4) $\psi_i'(\mathbf{0}, \mathbf{w}) = w_1 + \cdots + w_{s+1}$.

The proof is deferred to section 5.10.

We conclude this discussion of normal form by noting an example of a system of homogeneous linear forms which may be reparametrised in normal form, but without quantitative control over the resulting extension.

Indeed, take $\iota(N)$ to be some function such that $\iota(N) \to \infty$ as $N \to \infty$. Consider the forms

$$\psi_1(u_1, u_2, u_3) = (1 + \iota(N)^{-1})u_1 + u_2$$
$$\psi_2(u_1, u_2, u_3) = u_1 + u_2$$
$$\psi_3(u_1, u_2, u_3) = u_3.$$

Notice that $\mathrm{dist}(\Psi, V_{\mathrm{degen}}(3,3)) \to 0$ as $N \to \infty$, so Ψ does not have finite c_1-Cauchy-Schwarz complexity for any positive absolute constant c_1. One may nonetheless construct a normal form reparametrisation

$$\psi_1'(u_1, u_2, u_3, w_1, w_2) = (1 + \iota(N)^{-1})u_1 + u_2 + w_1$$
$$\psi_2'(u_1, u_2, u_3, w_1, w_2) = u_1 + u_2 + w_2$$
$$\psi_3'(u_1, u_2, u_3, w_1, w_2) = u_3.$$

The system Ψ does have all its non-zero coefficients bounded away from 0 and ∞, but

$$\Psi'(u_1, u_2, u_3, w_1, w_2) = \Psi(u_1 + \iota(N)w_1 - \iota(N)w_2, u_2 - \iota(N)w_1 + (\iota(N) + 1)w_2, u_3),$$

so Ψ' is not obtained by bounded shifts of the u_i variables. Such an extension would

not be suitable for our requirements in section 5.8.

One final remark: in [38], the simple algorithm that constructs normal form extensions with respect to a fixed i may easily be iterated, and so the authors work with systems that are in normal form with respect to every index i. A careful analysis of the proof in Appendix C of [38] demonstrates that it is sufficient for Ψ merely to admit, for each i separately, an extension that is in normal form with respect to ψ_i, but this is of little consequence in [38]. Yet certain quantitative aspects of the iteration of the normal form algorithm, critical to our application of these ideas, are not immediately clear to us. We have stated Proposition 5.3.8 for normal forms only with respect to a single i, in order to avoid this technical annoyance.

5.4 Upper bounds

This section is devoted to proving three upper bounds on the expressions $T^L_{F,G}(1, \cdots, 1)$. (For the definition of this quantity, the reader may refer to Definition 5.2.1). The first is exceptionally crude, but will nonetheless be useful in section 5.6.

Lemma 5.4.1. *Let N, m, d be natural numbers, satisfying $d \geqslant m+1$, and let c, C, ε be positive constants. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and suppose that $\|L\|_\infty \leqslant C$ and $\mathrm{dist}(L, V_{\mathrm{rank}}(m,d)) \geqslant c$. Let $F : \mathbb{R}^d \longrightarrow [0,1]$ and $G : \mathbb{R}^m \longrightarrow [0,1]$ be two functions, with F supported on $[-N, N]^d$ and G supported on $[-\varepsilon, \varepsilon]^m$. Then*

$$T^L_{F,G}(1, \cdots, 1) \ll_{c,C,\varepsilon} \|G\|_\infty.$$

Proof. Let M be a rank matrix of L (Proposition 5.3.1), and suppose without loss of generality that M consists of the first m columns of L. For j in the range $m+1 \leqslant j \leqslant d$, let the vector $\mathbf{v_j} \in \mathbb{R}^m$ be the j^{th} column of the matrix $M^{-1}L$. Then

$N^{d-m}T^L_{F,G}(1,\cdots,1) \leqslant \|G\|_\infty \cdot Z$, where Z is the number of solutions to

$$\begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix} + \sum_{j=m+1}^{d} \mathbf{v_j} n_j \in M^{-1}([-\varepsilon,\varepsilon]^m)$$

in which $n_1,\cdots,n_d$ are integers satisfying $|n_1|,\cdots,|n_d| \leqslant N$. Fixing a choice of the variables $n_{m+1},\cdots,n_d$ forces the vector $(n_1,\cdots,n_m)^T$ to lie in a convex region of diameter $O_{c,C,\varepsilon}(1)$. There are at most $O_{c,C,\varepsilon}(1)$ such points, so $Z \ll_{c,C,\varepsilon} N^{d-m}$. The claimed bound follows. □

Our second estimate is a slight strengthening of the above, albeit under stronger hypotheses.

Lemma 5.4.2. *Let N,m,d be natural numbers, with $d \geqslant m+1$, and let c,C,ε be positive constants. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and suppose that $\|L\|_\infty \leqslant C$ and $\mathrm{dist}(L, V^{\mathrm{global}}_{\mathrm{rank}}(m,d)) \geqslant c$. Let σ be a real number in the range $0 < \sigma < 1/2$. Let $F : \mathbb{R}^d \longrightarrow [0,1]$ and $G : \mathbb{R}^m \longrightarrow [0,1]$ be two functions, with F supported on*

$$\{\mathbf{x} \in \mathbb{R}^d : \mathrm{dist}(\mathbf{x}, \partial([1,N]^d)) \leqslant \sigma N\}$$

and G supported on $[-\varepsilon,\varepsilon]^m$. Then

$$T^L_{F,G}(1,\cdots,1) \ll_{c,C,\varepsilon} \sigma\|G\|_\infty.$$

Proof. Without loss of generality, we may assume that F is supported on

$$\{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_\infty \leqslant 2N,\ |x_d| \leqslant \sigma N\}$$

or

$$\{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_\infty \leqslant 2N,\ |x_d - (N-1)| \leqslant \sigma N\}$$

Consider the first case. By Proposition 5.3.1 there exists a rank matrix M that does not contain the column d. By reordering columns, we can assume without loss of generality that M consists of the first m columns of L. Continuing as in the proof of Lemma 5.4.1, for j in the range $m+1 \leqslant j \leqslant d$, let the vector $\mathbf{v_j} \in \mathbb{R}^m$ be the j^{th} column of the matrix $M^{-1}L$. Then the expression $N^{d-m}T^L_{F,G}(1,\cdots,1)$ may be bounded above by $\|G\|_\infty$ times the number of solutions to

$$\begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix} + \sum_{j=m+1}^{d} \mathbf{v_j} n_j \in M^{-1}([-\varepsilon,\varepsilon]^m)$$

satisfying $|n_1|,\cdots,|n_{d-1}| \leqslant 2N$ and $|n_d| \leqslant \sigma N$. We conclude as in the previous proof.

In the second case, the relevant equation is

$$\begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix} + \sum_{j=m+1}^{d} \mathbf{v_j} n_j + (N-1)\mathbf{v_d} \in M^{-1}([-\varepsilon,\varepsilon]^m),$$

in which we count solutions satisfying $|n_1|,\cdots,|n_{d-1}| \leqslant 2N$ and $|n_d| \leqslant \sigma N$. We conclude as in the previous proof. □

Our third estimate is more refined, and will be needed in section 5.5 when we deduce the main result (Theorem 5.2.10) from Theorem 5.5.6. It will help us to replace the sharp cut-off $1_{[-\varepsilon,\varepsilon]^m}$ with a Lipschitz cut-off. For the definition of the approximation function A_L, we refer the reader to Definition 5.2.8.

Lemma 5.4.3. *Let N, m, d be natural numbers, with $d \geqslant m+1$. Let c, C, ε be positive constants, and let σ_G be a parameter in the range $0 < \sigma_G < 1/2$. Suppose that $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ is a purely irrational surjective linear map, satisfying $\|L\|_\infty \leqslant C$ and $\mathrm{dist}(L, V_{\mathrm{rank}}(m,d)) \geqslant c$. Let A_L denote the approximation function of L. Let*

$F : \mathbb{R}^d \longrightarrow [0,1]$ be supported on $[-N,N]^d$, and let $G : \mathbb{R}^m \longrightarrow [0,1]$ be a Lipschitz function, with Lipschitz constant $O(1/\sigma_G)$, supported on $[-\varepsilon,\varepsilon]^m$. Assume further that $\int_{\mathbf{x}} G(\mathbf{x})\,d\mathbf{x} = O_\varepsilon(\sigma_G)$. Then for all τ_2 in the range $0 < \tau_2 \leqslant 1$,

$$T^L_{F,G}(1,\cdots,1) \ll_{c,C,\varepsilon} \sigma_G + \frac{\tau_2^{1-o(1)}}{\sigma_G} + \frac{\tau_2^{-O(1)} A_L(\Omega_{c,C}(1),\tau_2)^{-1}}{N}. \tag{5.20}$$

Proof. Following the proof of Lemma 5.4.1 verbatim, we arrive at the bound

$$T^L_{F,G}(1,\cdots,1) \ll_{c,C,\varepsilon} \frac{1}{N^{d-m}} \sum_{\substack{n_{m+1},\cdots,n_d \in \mathbb{Z} \\ |n_{m+1}|,\cdots,|n_d| \leqslant N}} \widetilde{G}\Big(\sum_{j=m+1}^{d} \mathbf{v_j} n_j\Big), \tag{5.21}$$

where $\mathbf{v_j}$ denotes the j^{th} column of the matrix $M^{-1}L$, and $\widetilde{G} : \mathbb{R}^m \longrightarrow [0,1]$ denotes the function

$$\widetilde{G}(\mathbf{x}) = \sum_{\mathbf{a} \in \mathbb{Z}^m} (G \circ M)(\mathbf{a} + \mathbf{x}).$$

It remains to estimate the right-hand side of (5.21).

We may consider $\widetilde{G}$ as a function on $\mathbb{R}^m/\mathbb{Z}^m$. With respect to the metric $\|\mathbf{x}\|_{\mathbb{R}^m/\mathbb{Z}^m}$, $\widetilde{G}$ is Lipschitz with Lipschitz constant $O_{c,C,\varepsilon}(1/\sigma_G)$. Also,

$$\int\limits_{\mathbf{x} \in [0,1)^m} \widetilde{G}(\mathbf{x})\,d\mathbf{x} = \int\limits_{\mathbf{x} \in \mathbb{R}^m} (G \circ M)(\mathbf{x})\,d\mathbf{x} = O_{c,C,\varepsilon}(\sigma_G).$$

By [37, Lemma A.9], for any X at least 2 we may write

$$\widetilde{G}(\mathbf{x}) = \sum_{\substack{\mathbf{k} \in \mathbb{Z}^m \\ \|\mathbf{k}\|_\infty \leqslant X}} b_X(\mathbf{k}) e(\mathbf{k} \cdot \mathbf{x}) + O_{c,C,\varepsilon}\Big(\frac{\log X}{\sigma_G X}\Big), \tag{5.22}$$

where $b_X(\mathbf{k}) \in \mathbb{C}$ and satisfies $|b_X(\mathbf{k})| = O(1)$. Moreover[6] $b_X(\mathbf{0}) = \int_{\mathbf{x} \in [0,1)^m} \widetilde{G}(\mathbf{x})\,d\mathbf{x}$.

[6]This final fact is not given explicitly in the statement of [37, Lemma A.9], although it is given in the proof. In any case, it may be immediately deduced from (5.22), by letting X tend to infinity

Returning to (5.21), we see that for any X at least 2 we may write

$$T^L_{F,G}(1,\cdots,1) \ll_{c,C,\varepsilon} \sigma_G + \frac{\log X}{\sigma_G X} + X^{O(1)} \max_{\substack{\mathbf{k}\in\mathbb{Z}^m \\ 0<\|\mathbf{k}\|_\infty\leqslant X}} \Big(\prod_{j=m+1}^{d} \min(1, N^{-1}\|\mathbf{k}\cdot\mathbf{v_j}\|^{-1}_{\mathbb{R}/\mathbb{Z}})\Big), \tag{5.23}$$

where the final error term comes from summing over the arithmetic progressions $[-N,N]\cap\mathbb{Z}$.

It remains to relate the final error term of (5.23) to the approximation function A_L. Since L is purely irrational,

$$A_L(\tau_1,\tau_2) = \inf_{\substack{\varphi\in(\mathbb{R}^m)^* \\ \tau_1\leqslant\|\varphi\|_\infty\leqslant\tau_2^{-1}}} \mathrm{dist}(L^*\varphi, (\mathbb{Z}^d)^T).$$

Let τ_2 be in the range $0<\tau_2\leqslant 1$. Then there is a suitable choice of parameter X, which satisfies $X\asymp_{c,C}\tau_2^{-1}$, such that

$$\begin{aligned} A_L(\Omega_{c,C}(1),\tau_2) &\leqslant \inf_{\substack{\mathbf{k}\in\mathbb{R}^m \\ 1\ll_{c,C}\|\mathbf{k}\|_\infty\ll_{c,C}\tau_2^{-1}}} \mathrm{dist}(\mathbf{k}^T M^{-1}L, (\mathbb{Z}^d)^T) \\ &\leqslant \min_{\substack{\mathbf{k}\in\mathbb{Z}^m \\ 1\ll_{c,C}\|\mathbf{k}\|_\infty\ll_{c,C}\tau_2^{-1}}} \max(\{\|\mathbf{k}\cdot\mathbf{v_j}\|_{\mathbb{R}/\mathbb{Z}} : m+1\leqslant j\leqslant d\}) \\ &\leqslant \min_{\substack{\mathbf{k}\in\mathbb{Z}^m \\ 0<\|\mathbf{k}\|_\infty\leqslant X}} \max(\{\|\mathbf{k}\cdot\mathbf{v_j}\|_{\mathbb{R}/\mathbb{Z}} : m+1\leqslant j\leqslant d\}). \end{aligned} \tag{5.24}$$

Substituting this bound into (5.23), one derives

$$T^L_{F,G}(1,\cdots,1) \ll_{c,C,\varepsilon} \sigma_G + \frac{\tau_2^{1-o(1)}}{\sigma_G} + \frac{\tau_2^{-O(1)} A_L(\Omega_{c,C}(1),\tau_2)^{-1}}{N}$$

as required. □

In Lemma 5.4.3, it was vitally important that L was assumed to be purely irrational. This was manifested in the relations (5.24), when one could upper-bound

and integrating (5.22) over all $\mathbf{x}\in\mathbb{R}^m/\mathbb{Z}^m$.

$A_L(\Omega_{c,C}(1), \tau_2)$ by a minimum taken over all $\mathbf{k} \in \mathbb{Z}^m$ of a certain size. Although one can attempt such an estimate when L is not purely irrational, the integral $\int_{\mathbf{x}\in\mathbb{R}^m} G(\mathbf{x})\, d\mathbf{x}$ is no longer the relevant object. Rather, one must take some rational map for L, denoted Θ, and consider $\int_{\mathbf{x}\in\ker\Theta+\mathbf{y}} G(\mathbf{x})\, d\mathbf{x}$ for some shift $\mathbf{y}$ (where $\ker\Theta + \mathbf{y}$ receives the natural Lebesgue measure). It could be that $\int_{\mathbf{x}\in\mathbb{R}^m} G(\mathbf{x})\, d\mathbf{x}$ is controlled but $\int_{\mathbf{x}\in\ker\Theta+\mathbf{y}} G(\mathbf{x})\, d\mathbf{x}$ is not (consider the case where G is the indicator of thin domain that has a flat side parallel to $\ker\Theta$, say). We opt to avoid these technicalities, creating instead a dimension reduction argument, that reduces all cases to the purely irrational case.

5.5 Reductions

In this section we reduce the main result (Theorem 5.2.10) to a different result, namely Theorem 5.5.6. This theorem will be simpler in one key respect: the replacement of sharp cut-offs by Lipschitz cut-offs.

We begin by dismissing the case of maximal rational dimension.

Proposition 5.5.1. *Theorem 5.2.10 holds under the additional assumption that L has rational dimension m.*

Proof. We appeal to a quantitative version of Theorem 5.1.1.

Theorem 5.5.2 (Generalised von Neumann Theorem for rational forms (quantitative version))**.** *Let N, m, d be natural numbers, satisfying $d \geqslant m+2$, and let C_1, C_2 be positive constants. Let $S = S(N)$ be an m-by-d matrix with integer coefficients, $\|S\|_\infty \leqslant C_1$, and let $\mathbf{r} \in \mathbb{Z}^m$ be some vector with $\|\mathbf{r}\|_\infty \leqslant C_2 N$. Suppose S has rank m, and $S \notin V^*_{\mathrm{degen}}(m,d)$. Let $K \subseteq [-N,N]^d$ be convex. Then there exists some natural number s at most $d-2$ that satisfies the following. Let $f_1, \cdots, f_d : [N] \longrightarrow \mathbb{C}$*

be arbitrary functions with $\|f_j\|_\infty \leqslant 1$ *for all* j*, and assume that*

$$\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho$$

for some ρ *in the range* $0 < \rho \leqslant 1$*. Then*

$$\frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n} \in \mathbb{Z}^d \cap K \\ S\mathbf{n} = \mathbf{r}}} \prod_{j=1}^{d} f_j(n_j) \ll_{C_1,C_2} \rho^{\Omega(1)} + o_\rho(1).$$

Furthermore, the $o_\rho(1)$ *term may be bounded above by* $\rho^{-O(1)}N^{-1}$*.*

In the proof, a certain familiarity with the methods and terminology of [38] will be assumed.

Proof Sketch. This theorem may be proved by following the proof of Theorem 1.8 of [38]. (In our language, the non-degeneracy condition in the statement of Theorem 1.8 of [38] is exactly $S \notin V^*_{\text{degen}}(m,d)$). One follows the same linear algebraic reductions as those used in section 4 of [38] to reduce Theorem 1.8 to Theorem 7.1 of the same paper (the Generalised von Neumann Theorem).

Theorem 7.1 may then be considered solely in the case of bounded functions f_j, as in [84, Exercise 1.3.23], rather than in the more general case of functions bounded by a pseudorandom measure. It is clear from the proof that, in this more restricted setting, the $\kappa(\rho)$ term that appears in the statement may be replaced by a polynomial dependence, and the $o_\rho(1)$ term may be bounded above by $\rho^{-O(1)}N^{-1}$.

This settles Theorem 5.5.2, where s is the Cauchy-Schwarz complexity of some system of forms $(\psi_1, \cdots, \psi_d)$ that parametrises $\ker S$. But s is at most $d-2$, as any system of d forms with finite Cauchy-Schwarz complexity has Cauchy-Schwarz complexity at most $d-2$. Therefore Theorem 5.5.2 is proved. □

Now let us use Theorem 5.5.2 to resolve Proposition 5.5.1. Indeed, let L be as in

Theorem 5.2.10, and assume that L has rational dimension m and rational complexity at most C. Let $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ be some linear isomorphism satisfying $\Theta L(\mathbb{Z}^d) \subseteq \mathbb{Z}^m$ and $\|\Theta\|_\infty \leqslant C$. Let M be a rank matrix of L (Proposition 5.3.1). Then the matrix $M^{-1}L$ satisfies $\|M^{-1}L\|_\infty \ll_{c,C} 1$ and has rational dimension m, since $((\Theta M) \circ (M^{-1}L))(\mathbb{Z}^d) = \Theta L(\mathbb{Z}^d) \subseteq \mathbb{Z}^m$. The matrix $M^{-1}L$ also has rational complexity $O_{c,C}(1)$. Therefore, replacing L with $M^{-1}L$, we may assume that the first m columns of L form the identity matrix.

As in Lemma 5.2.7, we write $\Theta L = S$, where S has integer coefficients and $\|\Theta\|_\infty \ll_{c,C} 1$. Hence $\|S\|_\infty \ll_{c,C} 1$. But Θ must also have integer coefficients, as the first m columns of L form the identity matrix, and hence $\|\Theta^{-1}\|_\infty \ll_{c,C} 1$ as well. Note finally that $S \notin V^*_{\text{degen}}(m,d)$, since $L \notin V^*_{\text{degen}}(m,d)$.

Now, suppose that $G : \mathbb{R}^m \longrightarrow [0,1]$ is the indicator function of some convex domain D, with $D \subseteq [-\varepsilon,\varepsilon]^m$. Then there are at most $O_{c,C,\varepsilon}(1)$ possible vectors $\mathbf{r} \in \mathbb{Z}^m$ such that $\mathbf{r} \in S(\mathbb{Z}^d) \cap \Theta(D)$. Let R be the set of all such vectors. Therefore, with F being the indicator function of the set $[1,N]^d$, we have

$$T^L_{F,G}(f_1,\cdots,f_d) = \sum_{\mathbf{r}\in R} \sum_{\substack{\mathbf{n}\in[N]^d \\ S\mathbf{n}=\mathbf{r}}} \prod_{j=1}^{d} f_j(n_j) \ll_{c,C,\varepsilon} \rho^{\Omega(1)} + o_\rho(1), \tag{5.25}$$

by Theorem 5.5.2. The $o_\rho(1)$ term may be bounded above by $\rho^{-O(1)}N^{-\Omega(1)}$. This is the desired conclusion of Theorem 5.2.10 in the case when L has rational dimension m. □

Having dismissed this case, we prepare to state Theorem 5.5.6 (the theorem that will imply the remaining cases). We begin with a definition that generalises Definition 5.2.1.

Definition 5.5.3. *Let N, m, d, h be natural numbers, with $d \geqslant h \geqslant m+2$. Let ε be positive. Let $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ and $L : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ be linear maps. Let $F : \mathbb{R}^h \to [0,1]$*

and $G : \mathbb{R}^m \to [0,1]$ *be two functions, with* F *supported on* $[-N,N]^h$ *and* G *supported on* $[-\varepsilon,\varepsilon]^m$. *Let* $\widetilde{\mathbf{r}} \in \mathbb{Z}^d$ *be some vector, and let* $f_1, \cdots, f_d : \mathbb{R} \longrightarrow [-1,1]$ *be arbitrary functions. We then define*

$$T_{F,G}^{L,\Xi,\widetilde{\mathbf{r}}}(f_1,\cdots,f_d) := \frac{1}{N^{h-m}} \sum_{\mathbf{n}\in\mathbb{Z}^h} \Big(\prod_{j=1}^{d} f_j(\xi_j(\mathbf{n}) + \widetilde{\mathbf{r}}_j) \Big) F(\mathbf{n}) G(L\mathbf{n}). \tag{5.26}$$

In the chapter so far we have introduced many non-degeneracy relations (Definitions 5.2.2, 5.2.3, 5.3.5). In order to state Theorem 5.5.6, we must introduce another.

Definition 5.5.4 (Dual pair degeneracy variety)**.** *Let* m, d, h *be natural numbers satisfying* $d \geqslant h \geqslant m+2$. *Let* $\mathbf{e_1}, \cdots, \mathbf{e_d}$ *denote the standard basis vectors of* $\mathbb{R}^d$, *and let* $\mathbf{e_1}^*, \cdots, \mathbf{e_d}^*$ *denote the dual basis of* $(\mathbb{R}^d)^*$. *Then let* $V^*_{\text{degen},2}(m,d,h)$ *denote the set of all pairs of linear maps* $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ *and* $L : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ *for which there exist two indices* $i, j \leqslant d$, *and some real number* λ, *such that* $(\mathbf{e_i}^* - \lambda \mathbf{e_j}^*)$ *is non-zero and* $\Xi^*(\mathbf{e_i}^* - \lambda \mathbf{e_j}^*) \in L^*((\mathbb{R}^m)^*)$. *We call* $V^*_{\text{degen},2}(m,d,h)$ *the* dual pair degeneracy variety.

Definition 5.5.5 (Distance metric for pairs of matrices)**.** *Let* m, d, h *be natural numbers, with* $d \geqslant h \geqslant m+2$, *and let* $V^*_{\text{degen},2}(m,d,h)$ *be the dual pair degeneracy variety. Let* $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ *and* $L : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ *be linear maps. We say that* $\text{dist}((\Xi, L), V^*_{\text{degen},2}(m,d,h)) \geqslant c$ *if* $(\Xi + Q, L) \notin V^*_{\text{degen},2}(m,d,h)$ *for all* $Q : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ *with* $\|Q\|_\infty < c$.

Although this is no great subtlety, we should emphasise that in the above definition we only consider perturbations to Ξ, and not perturbations to L as well.

We are now ready to state Theorem 5.5.6, the theorem to which we will reduce the main result (Theorem 5.2.10).

Theorem 5.5.6 (Lipschitz case)**.** *Let* N, m, d, h *be natural numbers, with* $d \geqslant h \geqslant m+2$, *and let* c, C, ε *be positive constants. Let* $\Xi = \Xi(N) : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ *be an*

injective linear map with integer coefficients, and assume that $\Xi(\mathbb{Z}^h) = \mathbb{Z}^d \cap \operatorname{im}\Xi$. *Let* $L = L(N) : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ *be a surjective linear map. Assume that* $\|\Xi\|_\infty \leqslant C$, $\|L\|_\infty \leqslant C$, $\operatorname{dist}(L, V_{\mathrm{rank}}(m,d)) \geqslant c$ *and* $\operatorname{dist}((\Xi, L), V^*_{\mathrm{degen},2}(m,d,h)) \geqslant c$. *Then there exists a natural number* s *at most* $d-2$ *such that the following holds. Let* σ_F, σ_G *be any two parameters in the range* $0 < \sigma_F, \sigma_G < 1/2$. *Let* $F : \mathbb{R}^h \longrightarrow [0,1]$ *be a Lipschitz function supported on* $[-N,N]^h$ *with Lipschitz constant* $O(1/\sigma_F N)$, *and let* $G : \mathbb{R}^m \longrightarrow [0,1]$ *be a Lipschitz function supported on* $[-\varepsilon,\varepsilon]^m$ *with Lipschitz constant* $O(1/\sigma_G)$. *Let* $\widetilde{\mathbf{r}}$ *be a fixed vector in* $\mathbb{Z}^d$, *satisfying* $\|\widetilde{\mathbf{r}}\|_\infty = O_{c,C,\varepsilon}(1)$. *Suppose that* $f_1, \cdots, f_d : [N] \longrightarrow [-1,1]$ *are arbitrary bounded functions satisfying*

$$\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho,$$

for some ρ *in the range* $0 < \rho \leqslant 1$. *Then*

$$T^{L,\Xi,\widetilde{\mathbf{r}}}_{F,G}(f_1, \cdots, f_d) \ll_{c,C,\varepsilon} \rho^{\Omega(1)}(\sigma_F^{-O(1)} + \sigma_G^{-O(1)}) + \sigma_F^{-O(1)} N^{-\Omega(1)}. \tag{5.27}$$

Although the above theorem contains more technical conditions than even Theorem 5.2.10 did, it does represent a significant reduction in complexity from the original problem. Note in particular that the approximation function A_L does not feature in the estimate (5.27).

The remainder of this section will be devoted to proving the main theorem (Theorem 5.2.10), assuming the result of Theorem 5.5.6.

We begin with two lemmas: one concerning lattices, and the other concerning a quantitative decomposition of the dual space $(\mathbb{R}^d)^*$. Their proofs are entirely standard, but we state them prominently, as we will need to refer to them often in the dimension reduction argument of Lemma 5.5.10.

Lemma 5.5.7 (Parametrising the image lattice)**.** *Let u, d be integers with $d \geqslant u+1$. Let $S : \mathbb{R}^d \longrightarrow \mathbb{R}^u$ be a surjective linear map with $S(\mathbb{Z}^d) \subseteq \mathbb{Z}^u$, and suppose that $\|S\|_\infty \leqslant C$. Then there exists a set $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\} \subset \mathbb{Z}^u$ that is a basis for the lattice $S(\mathbb{Z}^d)$ and for which $\|\mathbf{a_i}\|_\infty = O_C(1)$ for every i. Furthermore there exist $\mathbf{x_1}, \cdots, \mathbf{x_u} \in \mathbb{Z}^d$ such that, for every i, $S(\mathbf{x_i}) = \mathbf{a_i}$ and $\|\mathbf{x_i}\|_\infty = O_C(1)$.*

Proof. The lattice $S(\mathbb{Z}^d)$ is u dimensional, as S is surjective. If $\{\mathbf{e_j} : j \leqslant d\}$ denotes the standard basis of $\mathbb{R}^d$ then integer combinations of elements from the set $\{S(\mathbf{e_j}) : j \leqslant d\}$ span $S(\mathbb{Z}^d)$. Since $\|S\|_\infty \leqslant C$, these vectors also satisfy $\|S(\mathbf{e_j})\|_\infty = O_C(1)$. Therefore the u successive minima of the lattice $S(\mathbb{Z}^d)$ are all $O_C(1)$, and so, by Mahler's theorem ([85, Theorem 3.34]) the lattice $S(\mathbb{Z}^d)$ has a basis $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\}$ of the required form.

Note that S has integer coefficients. The construction of suitable $\mathbf{x_1}, \cdots, \mathbf{x_u}$ may be achieved by applying any of the standard algorithms. For example, using Gaussian elimination one may find a basis for $\ker S$ that, by inspection of the algorithm, consists of vectors with rational coordinates of naive height $O_C(1)$. By clearing denominators, one gets vectors $\mathbf{v_1}, \cdots, \mathbf{v_{d-u}}$ that span a full-dimensional sublattice of the $d-u$ dimensional lattice $\mathbb{Z}^d \cap \ker S$, and that satisfy $\|\mathbf{v_i}\|_\infty = O_C(1)$ for all i. Given some $\mathbf{a_i}$, by its construction there must be some $\mathbf{x_i} \in \mathbb{Z}^d$ that satisfies $S(\mathbf{x_i}) = \mathbf{a_i}$. Write $\mathbf{x_i} = \mathbf{x_i}|_{\ker S} + \mathbf{x_i}|_{(\ker S)^\perp}$ as the sum of the obvious projections. By adding a suitable combination of the vectors $\mathbf{v_1}, \cdots, \mathbf{v_{d-u}}$ to $\mathbf{x_i}$ one may ensure that $\|\mathbf{x_i}|_{\ker S}\|_\infty = O_C(1)$. Furthermore, $\operatorname{dist}(S, V_{\text{rank}}(m,d)) = \Omega_C(1)$, since S has integer coordinates, and so (by Lemma 5.11.1) $\|\mathbf{x_i}|_{(\ker S)^\perp}\|_\infty = O_C(1)$. Hence $\|\mathbf{x_i}\|_\infty = O_C(1)$. □

Having established that such a basis $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\}$ exists, we can now use it to quantitatively decompose $(\mathbb{R}^d)^*$.

Lemma 5.5.8 (Dual space decomposition)**.** *Let u, d, be integers with $d \geqslant u+1$, and let C, η be constants. Let $S : \mathbb{R}^d \longrightarrow \mathbb{R}^u$ be a surjective linear map with $S(\mathbb{Z}^d) \subseteq \mathbb{Z}^u$,*

and suppose that $\|S\|_\infty \leqslant C$. *Let* $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\}$ *be a basis for the lattice* $S(\mathbb{Z}^d)$ *that satisfies* $\|\mathbf{a_i}\|_\infty = O_C(1)$ *for every* i. *Let* $\mathbf{x_1}, \cdots, \mathbf{x_u} \in \mathbb{Z}^d$ *be vectors such that, for every* i, $S(\mathbf{x_i}) = \mathbf{a_i}$ *and* $\|\mathbf{x_i}\|_\infty = O_C(1)$. *Suppose that* $\Xi : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^d$ *is an injective linear map such that* $\operatorname{im}\Xi = \ker S$ *and such that* $\Xi(\mathbb{Z}^{d-u}) = \mathbb{Z}^d \cap \operatorname{im}\Xi$. *Suppose further that* $\|\Xi\|_\infty \leqslant C$.

Let $\mathbf{w_1}, \cdots, \mathbf{w_{d-u}}$ *denote the standard basis vectors in* $\mathbb{R}^{d-u}$. *Then*

(1) the set $\mathcal{B} := \{\mathbf{x_i} : i \leqslant u\} \cup \{\Xi(\mathbf{w_j}) : j \leqslant d-u\}$ *is a basis for* $\mathbb{R}^d$, *and a lattice basis for* $\mathbb{Z}^d$;

(2) writing $\mathcal{B}^* := \{\mathbf{x_i^*} : i \leqslant u\} \cup \{\Xi(\mathbf{w_j})^* : j \leqslant d-u\}$ *for the dual basis, the change of basis matrix between the standard dual basis and* $\mathcal{B}^*$, *and the inverse of this matrix, both have integer coordinates. The coefficients of both of these matrices are bounded in absolute value by* $O_C(1)$.

Write $V := \operatorname{span}(\mathbf{x_i^*} : i \leqslant u)$ *and* $W := \operatorname{span}(\Xi(\mathbf{w_j})^* : j \leqslant d-u)$. *Then*

(3) $V = S^*((\mathbb{R}^u)^*)$;

(4) Suppose that $\varphi \in (\mathbb{R}^d)^*$ *satisfies* $\|\Xi^*(\varphi)\|_\infty \leqslant \eta$. *Then, writing* $\varphi = \varphi_V + \varphi_W$ *with* $\varphi_V \in V$ *and* $\varphi_W \in W$, *we have* $\|\varphi_W\|_\infty = O_C(\eta)$.

Proof. For part (1), the fact that $\mathcal{B}$ is a basis for $\mathbb{R}^d$ is just a manifestation of the familiar principle $\mathbb{R}^d \cong \ker S \oplus \operatorname{im} S$. To show that $\mathcal{B}$ is a lattice basis for $\mathbb{Z}^d$, let $\mathbf{n} \in \mathbb{Z}^d$ and write

$$\mathbf{n} = \sum_{i=1}^{u} \lambda_i \mathbf{x_i} + \sum_{j=1}^{d-u} \mu_j \Xi(\mathbf{w_j})$$

for some $\lambda_i, \mu_j \in \mathbb{R}$. Applying S, we see $S(\mathbf{n}) = \sum_i^u \lambda_i \mathbf{a_i}$, and hence $\lambda_i \in \mathbb{Z}$ for all i, as $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\}$ is a basis for the lattice $S(\mathbb{Z}^d)$. But this implies $\sum_{j=1}^{d-u} \mu_j \Xi(\mathbf{w_j}) \in \mathbb{Z}^d \cap \operatorname{im}(\Xi)$. Therefore, as $\Xi(\mathbb{Z}^{d-u}) = \mathbb{Z}^d \cap \ker S$, $\mu_j \in \mathbb{Z}$ for all j.

Part (2) follows immediately from part (1). Part (3) is immediate from the definitions.

For part (4), let j be at most $d-u$. Then the assumption $\|\Xi^*(\varphi)\|_\infty \leqslant \eta$ means that $|\Xi^*(\varphi)(\mathbf{w_j})| \leqslant \eta$. Hence $|\varphi(\Xi(\mathbf{w_j}))| \leqslant \eta$. But, writing $\varphi_W = \sum_{j=1}^{d-u} \mu_j \Xi(\mathbf{w_j})^*$, this implies that $|\mu_j| \leqslant \eta$. Since the coefficients of the change of basis matrix between $\mathcal{B}^*$ and the standard dual basis are bounded in absolute value by $O_C(1)$, this implies that $\|\varphi_W\|_\infty \leqslant O_C(\eta)$. □

We now begin the attack on Theorem 5.2.10 in earnest. Assume the hypotheses of Theorem 5.2.10. As a reminder, we have natural numbers m, d satisfying $d \geqslant m+2$, and positive reals ε, c, C. For a natural number N, we have $L = L(N) : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ being a surjective linear map with approximation function A_L, with $\mathrm{dist}(L, V^*_{\mathrm{degen}}(m,d)) \geqslant c$, and with rational complexity at most C. We have $F : \mathbb{R}^d \longrightarrow [0,1]$ being the indicator function of $[1,N]^d$ and $G : \mathbb{R}^m \longrightarrow [0,1]$ being the indicator function of a convex domain contained in $[-\varepsilon, \varepsilon]^m$, and functions $f_1, \cdots, f_d : [N] \longrightarrow [-1,1]$ that satisfy $\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho$ for some ρ in the range $0 < \rho \leqslant 1$.

The proof has four parts:

- replacing the indicator function of $[1,N]^d$ with a Lipschitz cut-off;
- replacing L by a purely irrational map;
- replacing the function G by a Lipschitz cut-off, using Lemma 5.4.3;
- applying Theorem 5.5.6.

The second of these steps is by far the most technically intricate: this is Lemma 5.5.10.

Lemma 5.5.9 (Replacing variable cut-off)**.** *Assume the hypotheses of Theorem 5.2.10 (in particular let F be the indicator function $1_{[1,N]^d}$), and let σ_F be any parameter in the range $0 < \sigma_F < 1/2$. Then there exists a Lipschitz function $F_{1,\sigma_F} : \mathbb{R}^d \longrightarrow [0,1]$,*

supported on $[-2N, 2N]^d$ *and with Lipschitz constant* $O(1/\sigma_F N)$*, such that*

$$|T^L_{F,G}(f_1, \cdots, f_d)| \ll |T^L_{F_{1,\sigma_F},G}(f_1, \cdots, f_d)| + O_{c,C}(\sigma_F).$$

Proof. By Lemma 0.4.2, for any parameter σ_F in the range $0 < \sigma_F < 1/2$ we may write

$$1_{[1,N]^d} = F_{1,\sigma_F} + O(F_{2,\sigma_F}),$$

where $F_{1,\sigma_F}, F_{2,\sigma_F}$ are Lipschitz functions supported on $[-2N, 2N]^d$, with Lipschitz constants $O(1/\sigma_F N)$, and with $\int_{\mathbf{x}} F_{2,\sigma_F}(\mathbf{x})\, d\mathbf{x} = O(\sigma_F N^d)$. Moreover, F_{2,σ_F} is supported on

$$\{\mathbf{x} \in \mathbb{R}^d : \mathrm{dist}(\mathbf{x}, \partial([1, N]^d)) = O(\sigma_F N)\}.$$

Therefore

$$T^L_{F,G}(f_1, \cdots, f_d) \ll |T^L_{F_{1,\sigma_F},G}(f_1, \cdots, f_d)| + |T^L_{F_{2,\sigma_F},G}(1, \cdots, 1)|.$$

Recall, from the remark after Definition 5.2.3, that $V^*_{\mathrm{degen}}(m, d)$ contains $V^{\mathrm{global}}_{\mathrm{rank}}(m, d)$. Therefore, since we assume that $\mathrm{dist}(L, V^*_{\mathrm{degen}}(m, d)) \geqslant c$, we have $\mathrm{dist}(L, V^{\mathrm{global}}_{\mathrm{rank}}(m, d)) \geqslant c$. Hence, by Lemma 5.4.2,

$$|T^L_{F_{2,\sigma_F},G}(f_1, \cdots, f_d)| = O_{c,C}(\sigma_F).$$

This gives the lemma. □

Next comes the critical lemma, in which we successfully replace the map L by a purely irrational map L'. For the definition of the approximation function A_L, one may consult Definition 5.2.8.

Lemma 5.5.10 (Generating a purely irrational map)**.** *Let* σ_F *be a parameter in the range* $0 < \sigma_F < 1/2$*. Assume the hypotheses of Theorem 5.2.10, with the exception*

that $F : \mathbb{R}^d \longrightarrow [0,1]$ *is now a Lipschitz function supported on* $[-2N, 2N]^d$ *and with Lipschitz constant* $O(1/\sigma_F N)$. *Let* u *be the rational dimension of* L, *and assume that* $u \leqslant m-1$. *Then there exists a surjective linear map* $L' : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^{m-u}$, *an injective linear map* $\Xi : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^d$, *a finite subset* $\widetilde{R} \subset \mathbb{Z}^d$, *and, for each* $\widetilde{\mathbf{r}} \in \widetilde{R}$, *functions* $F_{\widetilde{\mathbf{r}}} : \mathbb{R}^{d-u} \longrightarrow [0,1]$ *and* $G_{\widetilde{\mathbf{r}}} : \mathbb{R}^{m-u} \longrightarrow [0,1]$, *that together satisfy the following properties:*

(1) Ξ *has integer coefficients,* $\|\Xi\|_\infty = O_{c,C}(1)$, *and* $\Xi(\mathbb{Z}^{d-u}) = \mathbb{Z}^d \cap \operatorname{im} \Xi$;

(2) $|\widetilde{R}| = O_{c,C}(1)$, *and* $\|\widetilde{\mathbf{r}}\|_\infty = O_{c,C}(1)$ *for all* $\widetilde{\mathbf{r}} \in \widetilde{R}$;

(3) $F_{\widetilde{\mathbf{r}}}$ *is supported on* $[-O_{c,C}(N), O_{c,C}(N)]^{d-u}$, *with Lipschitz constant* $O_{c,C}(1/\sigma_F N)$, *and* $G_{\widetilde{\mathbf{r}}}$ *is the indicator function of a convex domain contained in* $[-O_{c,C,\varepsilon}(1), O_{c,C,\varepsilon}(1)]^{m-u}$;

(4) $T^L_{F,G}(f_1, \cdots, f_d) = \sum_{\widetilde{\mathbf{r}} \in \widetilde{R}} T^{L', \Xi, \widetilde{\mathbf{r}}}_{F_{\widetilde{\mathbf{r}}}, G_{\widetilde{\mathbf{r}}}}(f_1, \cdots, f_d)$;

(5) L' *is purely irrational;*

(6) $\|L'\|_\infty = O_{c,C}(1)$ *and* $\operatorname{dist}(L', V_{\mathrm{rank}}(m-u, d-u)) = \Omega_{c,C}(1)$;

(7) $\operatorname{dist}((\Xi, L'), V^*_{\mathrm{degen},2}(m-u, d, d-u)) = \Omega_{c,C}(1)$;

(8) for all $\tau_1, \tau_2 \in (0,1]$, $A_{L'}(\tau_1, \tau_2) \gg_{c,C} A_L(\Omega_{c,C}(\tau_1), \Omega_{c,C}(\tau_2))$;

(9) for all $\tau_1, \tau_2 \in (0,1]$, $A_{L'}(\tau_1, \tau_2) \ll_{c,C} A_L(\Omega_{c,C}(\tau_1), \Omega_{c,C}(\tau_2))$.

The fundamental aspect of this lemma is part (4), of course, as this directly concerns how we control the number of solutions to the diophantine inequality itself when passing from L to L'. However, we do need to establish parts (1) - (8), in order to be able to ensure that the hypotheses of Lemma 5.4.3 and Theorem 5.5.6 are satisfied. Part (9) is included for completeness, and to assist the calculations in section 5.12.

Before giving the full details of the proof, we sketch the idea. Let $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^u$ be a rational map for L. The space $\ker(\Theta L)$ has dimension $d-u$, and so we may parametrise it by some injective map $\Xi : \mathbb{R}^{d-u} \longrightarrow \ker(\Theta L)$. Without too much difficultly, Ξ can be chosen to satisfy $\Xi(\mathbb{Z}^{d-u}) = \mathbb{Z}^d \cap \operatorname{im} \Xi$. Then

$$L\Xi : \mathbb{R}^{d-u} \longrightarrow \ker \Theta,$$

is a map from a $d-u$ dimensional space to an $m-u$ dimensional space, and it turns out that $L\Xi$ is purely irrational, and $L' = L\Xi$ may be used in Lemma 5.5.10.

Of course this isn't quite possible, as we only defined the notion of purely irrational maps between vector spaces of the form $\mathbb{R}^a$. But it is true after choosing a judicious isomorphism from $\ker \Theta$ to $\mathbb{R}^{m-u}$ (though this does complicate the notation).

Let us complete the details.

Proof. First we note that the lemma is obvious when $u = 0$, since one may take $\Xi : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ to be the identity map, $\widetilde{\mathbf{r}}$ to be $\mathbf{0}$, and L' to be L. So assume that $u > 1$.

We proceed with a general reduction, familiar from our proof of Proposition 5.5.1, in which we may assume that the first m columns of L form the identity matrix.

Indeed, let $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^u$ be a rational map for L with $\|\Theta\|_\infty \leqslant C$. Now let $\widetilde{L} := M^{-1}L$, where M is a rank matrix of L (Proposition 5.3.1), which, without loss of generality, consists of the first m columns of L. Let $\widetilde{\Theta} := \Theta M$ and let $\widetilde{G} := G \circ M$. Then

$$T^L_{F,G}(f_1, \cdots, f_d) = T^{\widetilde{L}}_{F,\widetilde{G}}(f_1, \cdots, f_d),$$

and, considering $\widetilde{\Theta}$, $\widetilde{L}$ has rational complexity $O_{c,C}(1)$. Furthermore, $\widetilde{G}$ is the indicator function of a convex domain contained in $[-O_{c,C}(\varepsilon), O_{c,C}(\varepsilon)]^m$. We also

have $\mathrm{dist}(\widetilde{L}, V^*_{\mathrm{degen}}(m,d)) = \Omega_{c,C}(1)$. Finally, for all $\tau_1, \tau_2 \in (0,1]$, $A_{\widetilde{L}}(\tau_1,\tau_2) \asymp_{c,C} A_L(\Omega_{c,C}(\tau_1), \Omega_{c,C}(\tau_2))$.

Therefore, by replacing L with $\widetilde{L}$ and G with $\widetilde{G}$, we may assume throughout the proof of Lemma 5.5.10 that the first m columns of L form the identity matrix. This is at the cost of replacing ε by $O_{c,C}(\varepsilon)$, C by $O_{c,C}(1)$, and c by $\Omega_{c,C}(1)$.

Now let $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^u$ be a rational map for L with $\|\Theta\|_\infty = O_{c,C}(1)$. Since the first m columns of L form the identity matrix, Θ must have integer coefficients.

Part (1): By rank-nullity $\ker(\Theta L)$ is a $d-u$ dimensional subspace of $\mathbb{R}^d$. The matrix of ΘL has integer coefficients and $\|\Theta L\|_\infty = O_{c,C}(1)$. Combining these two facts, we see that $\ker(\Theta L) \cap \mathbb{Z}^d$ is a $d-u$ dimensional lattice, and by the standard algorithms one can find a lattice basis $\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(d-u)}} \in \mathbb{Z}^d$ that satisfies $\|\mathbf{v^{(i)}}\|_\infty = O_{c,C}(1)$ for every i. Define $\Xi : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^d$ by

$$\Xi(\mathbf{w}) := \sum_{i=1}^{d-u} w_i \mathbf{v^{(i)}}.$$

Then Ξ satisfies property (1) of the lemma. Note that image of the map $L\Xi : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^m$ is exactly $\ker \Theta$.

Part (2): Since $\|\Theta\|_\infty = O_{c,C}(1)$, if $\mathbf{y} \in \mathbb{R}^m$ and $\Theta(\mathbf{y}) = \mathbf{r}$ then $\|\mathbf{y}\|_\infty \gg_{c,C} \|\mathbf{r}\|_\infty$. Recall that the support of G is contained within $[-O_{c,C,\varepsilon}(1), O_{c,C,\varepsilon}(1)]^m$, and that $\Theta L(\mathbb{Z}^d) \subseteq \mathbb{Z}^u$. It follows that there are at most $O_{c,C,\varepsilon}(1)$ possible vectors $\mathbf{r} \in \mathbb{Z}^u$ for which there exists a vector $\mathbf{n} \in \mathbb{Z}^d$ for which both $G(L\mathbf{n}) \neq 0$ and $\Theta L\mathbf{n} = \mathbf{r}$. Let R denote the set of all such vectors $\mathbf{r}$.

For each $\mathbf{r} \in R$, there exists a vector $\widetilde{\mathbf{r}} \in \mathbb{Z}^d$ such that $\Theta L\widetilde{\mathbf{r}} = \mathbf{r}$ and $\|\widetilde{\mathbf{r}}\|_\infty = O_{c,C,\varepsilon}(1)$. Let $\widetilde{R}$ denote the set of these $\widetilde{\mathbf{r}}$. Then $\widetilde{R}$ satisfies part (2).

Before proceeding to prove part (3) of the lemma, we pause to emphasise a particular consequence of Lemmas 5.5.7 and 5.5.8. Applying these lemmas to the map $S := \Theta L$, there exists a set $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\} \subset \mathbb{Z}^u$ that is a basis for the lattice $\Theta L(\mathbb{Z}^d)$ and for which $\|\mathbf{a_i}\|_\infty = O_{c,C}(1)$ for each i. Also, there exists a set of vectors $\{\mathbf{x_1}, \cdots, \mathbf{x_u}\} \subset \mathbb{Z}^d$ such that $\Theta L(\mathbf{x_i}) = \mathbf{a_i}$ for each i, and $\|\mathbf{x_i}\|_\infty = O_{c,C}(1)$. By Lemma 5.5.8,

$$\mathcal{B} := \{\mathbf{x_i} : i \leqslant u\} \cup \{\Xi(\mathbf{w_j}) : j \leqslant d-u\} \tag{5.28}$$

is a basis for $\mathbb{R}^d$ and a lattice basis for $\mathbb{Z}^d$, where $\mathbf{w_1}, \cdots, \mathbf{w_{d-u}}$ denotes the standard basis of $\mathbb{R}^{d-u}$.

Part (3): By the definition of $\widetilde{R}$, and the fact that $\Xi(\mathbb{Z}^{d-u}) = \mathbb{Z}^d \cap \ker(\Theta L)$, we have

$$T^L_{F,G}(f_1, \cdots, f_d) = \sum_{\widetilde{\mathbf{r}} \in \widetilde{R}} \frac{1}{N^{d-m}} \sum_{\mathbf{n} \in \mathbb{Z}^{d-u}} \Big(\prod_{j=1}^{d} f_j(\xi_j(\mathbf{n}) + \widetilde{\mathbf{r}}_j) \Big) F(\Xi(\mathbf{n}) + \widetilde{\mathbf{r}}) G(L\Xi(\mathbf{n}) + L\widetilde{\mathbf{r}}), \tag{5.29}$$

where $\widetilde{\mathbf{r}}_j$ denotes the j^{th} coordinates of $\widetilde{\mathbf{r}}$. Now by an easy linear algebraic argument (recorded in Lemma 5.11.4),

$$\mathbb{R}^m = \text{span}(L\mathbf{x_i} : i \leqslant u) \oplus \ker \Theta \tag{5.30}$$

as an algebraic direct sum, and there exists an invertible linear map $P : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ such that

$$P((\text{span}(L\mathbf{x_i} : i \leqslant u))) = \mathbb{R}^u \times \{0\}^{m-u}, \tag{5.31}$$

$$P(\ker \Theta) = \{0\}^u \times \mathbb{R}^{m-u}, \tag{5.32}$$

and both $\|P\|_\infty = O_{c,C}(1)$ and $\|P^{-1}\|_\infty = O_{c,C}(1)$.

We have

$$G(L\Xi(\mathbf{n}) + L\widetilde{\mathbf{r}}) = (G \circ P^{-1})(PL\Xi(\mathbf{n}) + PL\widetilde{\mathbf{r}}),$$

and we note that $PL\Xi(\mathbf{n}) \in \{0\}^u \times \mathbb{R}^{m-u}$ for every $\mathbf{n} \in \mathbb{Z}^{d-u}$. Define $G_{\widetilde{\mathbf{r}}} : \mathbb{R}^{m-u} \longrightarrow [0,1]$ by

$$G_{\widetilde{\mathbf{r}}}(\mathbf{x}) := (G \circ P^{-1})(\mathbf{x_0} + PL\widetilde{\mathbf{r}}),$$

where $\mathbf{x_0}$ is the extension of $\mathbf{x}$ by 0 in the first u coordinates. Then the function $G_{\widetilde{\mathbf{r}}}$ is the indicator function of a convex set contained in $[-O_{c,C,\varepsilon}(1), O_{c,C,\varepsilon}(1)]^{m-u}$.

Define

$$F_{\widetilde{\mathbf{r}}}(\mathbf{n}) := F(\Xi(\mathbf{n}) + \widetilde{\mathbf{r}}).$$

Then $F_{\widetilde{\mathbf{r}}}$ has Lipschitz constant $O_{c,C}(1/\sigma_F N)$ and $F_{\widetilde{\mathbf{r}}}$ is supported on $[-O_{c,C,\varepsilon}(N), O_{c,C,\varepsilon}(N)]^{d-u}$. (For a full proof of this fact, apply Lemma 5.11.3 to the map Ξ). So $F_{\widetilde{\mathbf{r}}}$ and $G_{\widetilde{\mathbf{r}}}$ satisfy part (3).

Part (4): Writing $\pi_{m-u} : \mathbb{R}^m \longrightarrow \mathbb{R}^{m-u}$ for the projection onto the final $m-u$ coordinates, expression (5.29) is equal to

$$\sum_{\widetilde{\mathbf{r}} \in \widetilde{R}} \frac{1}{N^{d-m}} \sum_{\mathbf{n} \in \mathbb{Z}^{d-u}} \Big(\prod_{j=1}^{d} f_j(\xi_j(\mathbf{n}) + \widetilde{\mathbf{r}}_j) \Big) F_{\widetilde{\mathbf{r}}}(\mathbf{n}) G_{\widetilde{\mathbf{r}}}(\pi_{m-u} PL\Xi(\mathbf{n})). \tag{5.33}$$

Let

$$L' := \pi_{m-u} PL\Xi. \tag{5.34}$$

Then $L' : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^{m-u}$ is surjective, and

$$T^L_{F,G}(f_1, \cdots, f_d) = \sum_{\widetilde{\mathbf{r}} \in \widetilde{R}} T^{L',\Xi,\widetilde{\mathbf{r}}}_{F_{\widetilde{\mathbf{r}}}, G_{\widetilde{\mathbf{r}}}}(f_1, \cdots, f_d).$$

This resolves part (4).

Part (5): We wish to show that L' is purely irrational. Suppose for contradiction that there exists some surjective linear map $\varphi : \mathbb{R}^{m-u} \longrightarrow \mathbb{R}$ with $\varphi L'(\mathbb{Z}^{d-u}) \subseteq \mathbb{Z}$, i.e. with $\varphi \pi_{m-u} PL\Xi(\mathbb{Z}^{d-u}) \subseteq \mathbb{Z}$. Then define the map $\Theta' : \mathbb{R}^m \longrightarrow \mathbb{R}^{u+1}$ by

$$\Theta'(\mathbf{x}) := (\Theta(\mathbf{x}), \varphi \pi_{m-u} P(\mathbf{x})).$$

Then Θ' is surjective, and $\Theta' L(\mathbb{Z}^d) \subseteq \mathbb{Z}^{u+1}$. (This second fact is immediately seen by writing $\mathbb{Z}^d$ with respect to the lattice basis $\mathcal{B}$ from (5.28)). This contradicts the assumption that L has rational dimension u. So L' is purely irrational.

Part (6): The bound $\|L'\|_\infty = O_{c,C}(1)$ follows immediately from the bounds on the coefficients of Ξ, L, P, and π_{m-u} separately.

We wish to prove that $\mathrm{dist}(L', V_{\mathrm{rank}}(m-u, d-u)) \gg_{c,C} 1$, i.e. that $\mathrm{dist}(\pi_{m-u} PL\Xi, V_{\mathrm{rank}}(m-u, d-u)) \gg_{c,C} 1$. Suppose for contradiction that, for a small parameter η, there exists a linear map $Q : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^{m-u}$ such that $\|Q\|_\infty < \eta$ and $\pi_{m-u} PL\Xi + Q$ has rank less than $m-u$. Recall that $PL\Xi(\mathbb{R}^{d-u}) = \{0\}^u \times \mathbb{R}^{m-u}$. So, extending Q by zeros to a map $Q : \mathbb{R}^{d-u} \longrightarrow \{0\}^u \times \mathbb{R}^{m-u}$, and applying P^{-1}, there is a map $Q' : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^m$ such that $\|Q'\|_\infty = O_{c,C}(\eta)$ and $L\Xi + Q'$ has rank less than $m-u$.

We may factorise $Q' = H\Xi$ for some m-by-d matrix H. Indeed let

$$\mathcal{B} := \{\mathbf{x_i} : i \leqslant u\} \cup \{\Xi(\mathbf{w_j}) : j \leqslant d-u\}$$

be the basis of $\mathbb{R}^d$ from (5.28), i.e. the basis formed by applying Lemma 5.5.8 to the map $S := \Theta L$. Define the linear map H by $H(\Xi(\mathbf{w_j})) := Q'(\mathbf{w_j})$ for each j and $H(\mathbf{x_i}) := \mathbf{0}$ for each i. Since the change of basis matrix between $\mathcal{B}$ and the standard basis of $\mathbb{R}^d$ has integer coefficients with absolute values at most $O_{c,C}(1)$, it

follows that the matrix representing H with respect to the standard bases satisfies $\|H\|_\infty = O_{c,C}(\eta)$.

So we know that $(L+H)\Xi$ has rank less than $m-u$. But $\Xi : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^d$ is injective, so this implies that the rank of $L+H$ is less than m. Hence $\text{dist}(L, V_{\text{rank}}(m,d)) = O_{c,C}(\eta)$, which contradicts the assumptions of the lemma (if η is small enough). So $\text{dist}(L', V_{\text{rank}}(m-u, d-u)) \gg_{c,C} 1$ as required.

Part (7): We wish to show that $\text{dist}((\Xi, L'), V^*_{\text{degen},2}(m-u, d, d-u)) = \Omega_{c,C}(1)$. Suppose for contradiction that, for a small parameter η, there exists a linear map $Q : \mathbb{R}^{d-u} \longrightarrow \mathbb{R}^d$ such that $\|Q\|_\infty \leqslant \eta$ and $\text{dist}((\Xi+Q, L'), V^*_{\text{degen},2}(m-u, d, d-u)) \leqslant \eta$. In other words, we suppose there exist two indices $i, j \leqslant d$, and a real number λ, such that

$$(\Xi + Q)^*(\mathbf{e_i}^* - \lambda \mathbf{e_j}^*) \in (L')^*((\mathbb{R}^{m-u})^*),$$

where $\{\mathbf{e_1}, \cdots, \mathbf{e_d}\}$ denotes the standard basis of $\mathbb{R}^d$ and $\{\mathbf{e_1}^*, \cdots, \mathbf{e_d}^*\}$ denotes the dual basis. Expanding out the definition of L', this means that there exists some $\varphi \in (\mathbb{R}^{m-u})^*$ such that

$$\Xi^*(\mathbf{e_i}^* - \lambda \mathbf{e_j}^* - L^*(P^* \pi^*_{m-u}(\varphi))) = -Q^*(\mathbf{e_i}^* - \lambda \mathbf{e_j}^*).$$

Because $\|Q\|_\infty \leqslant \eta$, this means that

$$\|\Xi^*(\mathbf{e_i}^* - \lambda \mathbf{e_j}^* - L^*(P^* \pi^*_{m-u}(\varphi)))\|_\infty = O(\eta). \tag{5.35}$$

Let

$$\mathcal{B}^* := \{\mathbf{x_i}^* : i \leqslant u\} \cup \{\Xi(\mathbf{w_j})^* : j \leqslant d-u\} \tag{5.36}$$

denote the basis of $(\mathbb{R}^d)^*$ that is dual to the basis $\mathcal{B}$ from (5.28). It follows from part

(4) of Lemma 5.5.8 and (5.35) that

$$\mathbf{e_i}^* - \lambda \mathbf{e_j}^* - L^*(P^* \pi^*_{m-u}(\varphi)) = \omega_V + \omega_W,$$

where $\omega_V \in L^* \Theta^*((\mathbb{R}^u)^*)$, $\omega_W \in \text{span}(\Xi(\mathbf{w_j})^* : j \leqslant d-u)$, and $\|\omega_W\|_\infty = O_{c,C}(\eta)$. So therefore

$$\mathbf{e_i}^* - \lambda \mathbf{e_j}^* = L^*(\alpha) + \omega_W,$$

for some $\alpha \in (\mathbb{R}^m)^*$.

This is enough to derive a contradiction. Indeed, without loss of generality one may assume that $\|\mathbf{e_i}^* - \lambda \mathbf{e_j}^*\|_\infty \geqslant 1$ (this is obvious if $i \neq j$, and if $i = j$ we may just pick $\lambda = 0$ at the outset). Therefore $\|\mathbf{e_i}^* - \lambda \mathbf{e_j}^* - \omega_W\| \geqslant 1/2$, provided η is small enough. Since $\|L^*\|_\infty = O_{c,C}(1)$, we conclude that $\|\alpha\|_\infty = \Omega_{c,C}(1)$.

This means that there exists a linear map $E : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ with $\|E\|_\infty = O_{c,C}(\eta)$ for which $E^*(\alpha) = \omega_W$. Then

$$\mathbf{e_i}^* - \lambda \mathbf{e_j}^* \in (L+E)^*((\mathbb{R}^m)^*),$$

and hence $\text{dist}(L, V^*_{\text{degen}}(m,d)) = O_{c,C}(\eta)$. This is a contradiction to the hypotheses of Theorem 5.2.10, provided η is small enough, and hence $\text{dist}((\Xi, L'), V^*_{\text{degen},2}(m - u, d, d-u)) = \Omega_{c,C}(1)$.

Part (8): Let $\tau_1, \tau_2 \in (0,1]$. We desire to prove the relationship

$$A_{L'}(\tau_1, \tau_2) \gg_{c,C} A_L(\Omega_{c,C}(\tau_1), \Omega_{c,C}(\tau_2)), \tag{5.37}$$

where L' is as in (5.34).

We have already proved that L' is purely irrational (that was part (5) of the lemma). So, if $A_{L'}(\tau_1, \tau_2) < \eta$, for some η, there exists some $\varphi \in (\mathbb{R}^{m-u})^*$ for which

$\tau_1 \leqslant \|\varphi\|_\infty \leqslant \tau_2^{-1}$ and for which

$$\mathrm{dist}((\pi_{m-u}PL\Xi)^*(\varphi), (\mathbb{Z}^{d-u})^T) < \eta,$$

where, one recalls, we use $(\mathbb{Z}^{d-u})^T$ to denote the set of those functions in $(\mathbb{R}^{d-u})^*$ that have integer coordinates with respect to the standard dual basis.

We claim that

$$\mathrm{dist}(L^*(P^*\pi_{m-u}^*(\varphi)), (\mathbb{Z}^d)^T) \ll_{c,C} \eta; \tag{5.38}$$

$$\|P^*\pi_{m-u}^*(\varphi)\|_\infty \ll_{c,C} \tau_2^{-1}; \tag{5.39}$$

$$\mathrm{dist}(P^*\pi_{m-u}^*(\varphi), \Theta^*((\mathbb{R}^u)^*)) \gg_{c,C} \tau_1, \tag{5.40}$$

from which (5.37) immediately follows.

Let us prove (5.38). Indeed, we already know that
$\mathrm{dist}(\Xi^*L^*P^*\pi_{m-u}^*(\varphi), (\mathbb{Z}^{d-u})^T) < \eta$, i.e. that

$$\|\Xi^*L^*P^*\pi_{m-u}^*(\varphi) - \alpha\|_\infty < \eta, \tag{5.41}$$

for some $\alpha \in (\mathbb{Z}^{d-u})^T$. Let us write $\alpha = \sum_{j=1}^{d-u} \lambda_j \mathbf{w_j}^*$ for some $\lambda_j \in \mathbb{Z}$, where $\mathbf{w_1}, \cdots, \mathbf{w_{d-u}}$ denotes the standard basis for $\mathbb{R}^{d-u}$ and $\mathbf{w_1}^*, \cdots, \mathbf{w_{d-u}}^*$ denotes the dual basis. Let $\mathcal{B}^*$ be as in (5.36). Then $\mathbf{w_j}^* = \Xi^*((\Xi(\mathbf{w_j})^*)$, and so

$$\alpha = \Xi^*(\sum_{j=1}^{d-u} \lambda_j \Xi(\mathbf{w_j})^*).$$

So from (5.41) and the final part of Lemma 5.5.8,

$$L^*P^*\pi_{m-u}^*(\varphi) - \sum_{j=1}^{d-u} \lambda_j \Xi(\mathbf{w_j})^* = \omega_V + \omega_W, \tag{5.42}$$

where $\omega_V \in \text{span}(\mathbf{x_i^*} : i \leqslant u)$, $\omega_W \in \text{span}(\Xi(\mathbf{w_j})^* : j \leqslant d-u)$, and $\|\omega_W\|_\infty = O_{c,C}(\eta)$.

But $L^*P^*\pi^*_{m-u}(\varphi) \in \text{span}(\Xi(\mathbf{w_j})^* : j \leqslant d-u)$ too. Indeed, for every i at most $d-u$,

$$L^*P^*\pi^*_{m-u}(\varphi)(\mathbf{x_i}) = \varphi(\pi_{m-u}PL\mathbf{x_i}) = \varphi(\mathbf{0}) = 0,$$

by the properties of P (see (5.31)). Therefore $\omega_V = \mathbf{0}$, and so

$$\|L^*P^*\pi^*_{m-u}(\varphi) - \sum_{j=1}^{d-u} \lambda_j \Xi(\mathbf{w_j})^*\|_\infty = O_{c,C}(\eta).$$

Since $\sum_{j=1}^{d-u} \lambda_j \Xi(\mathbf{w_j})^* \in (\mathbb{Z}^d)^T$, this implies (5.38) as claimed.

The bound (5.39) is immediate from the bounds on the coefficients of P^* and π^*_{m-u}, so it remains to prove (5.40). Suppose for contradiction that, for some small parameter δ,

$$P^*\pi^*_{m-u}(\varphi) = \alpha_1 + \alpha_2,$$

where $\alpha_1 \in \Theta^*((\mathbb{R}^u)^*)$ and $\|\alpha_2\|_\infty \leqslant \delta\tau_1$. We know that $\|\varphi\|_\infty \geqslant \tau_1$, which means that there is some standard basis vector $\mathbf{f_k} \in \mathbb{R}^{m-u}$ for which $|\varphi(\mathbf{f_k})| \geqslant \tau_1$. Let $\mathbf{b_{k+u}}$ be the standard basis vector of $\mathbb{R}^m$ for which $\pi_{m-u}(\mathbf{b_{k+u}}) = \mathbf{f_k}$. Recall the properties of P (given in (5.31) and (5.32)), in particular recall that $P : \ker\Theta \longrightarrow \{0\}^u \times \mathbb{R}^{m-u}$ is an isomorphism. Then

$$|P^*\pi^*_{m-u}(\varphi)(P^{-1}(\mathbf{b_{k+u}}))| = |\pi^*_{m-u}(\varphi)(\mathbf{b_{k+u}})| = |\varphi(\mathbf{f_k})| \geqslant \tau_1.$$

Note that $\Theta^*((\mathbb{R}^u)^*) = (\ker\Theta)^\circ$, and so

$$|P^*\pi^*_{m-u}(\varphi)(P^{-1}(\mathbf{b_{k+u}}))| = |(\alpha_1+\alpha_2)(P^{-1}(\mathbf{b_{k+u}}))| = |\alpha_2(P^{-1}(\mathbf{b_{k+u}}))| \ll_{c,C} \delta\tau_1,$$

as $P^{-1}(\mathbf{b_{k+u}}) \in \ker\Theta$ and satisfies $\|P^{-1}(\mathbf{b_{k+u}})\|_\infty = O_{c,C}(1)$. This is a contradiction

if δ is small enough, and so (5.40) holds. This resolves part (8).

Part (9): Let $\tau_1, \tau_2 \in (0, 1]$. We desire to prove the relationship

$$A_{L'}(\tau_1, \tau_2) \ll_{c,C} A_L(\Omega_{c,C}(\tau_1), \Omega_{c,C}(\tau_2)), \tag{5.43}$$

where L' is as in (5.34).This inequality is the reverse inequality of part (8), and in fact it will not be required in the proof of any of our main theorems. However, it will be required in order to analyse $A_L(\tau_1, \tau_2)$ when L has algebraic coefficients (in section 5.12), so we choose to state and prove it here, close to our argument for part (8).

Suppose that $A_L(\tau_1, \tau_2) < \eta$, for some parameter η. Then there exists some $\varphi \in (\mathbb{R}^m)^*$ such that $\text{dist}(\varphi, \Theta^*((\mathbb{R}^u)^*)) \geqslant \tau_1$, $\|\varphi\|_\infty \leqslant \tau_2^{-1}$, and $\text{dist}(L^*\varphi, (\mathbb{Z}^d)^T) < \eta$. In particular there exists some $\omega \in (\mathbb{Z}^d)^T$ for which

$$\|L^*\varphi - \omega\|_\infty < \eta.$$

We expand both $L^*\varphi$ and ω with respect to the dual basis $\mathcal{B}^*$ from (5.36). So,

$$\begin{aligned} L^*\varphi &= \sum_{i=1}^{u} \lambda_i \mathbf{x_i^*} + \sum_{j=1}^{d-u} \mu_j \Xi(\mathbf{w_j})^* \\ \omega &= \sum_{i=1}^{u} \lambda_i' \mathbf{x_i^*} + \sum_{j=1}^{d-u} \mu_j' \Xi(\mathbf{w_j})^*. \end{aligned}$$

Since $\mathcal{B}^*$ is a lattice basis for $(\mathbb{Z}^d)^T$, we have $\lambda_i' \in \mathbb{Z}$ and $\mu_j' \in \mathbb{Z}$ for each i and j. Since the change of basis matrix between $\mathcal{B}^*$ and the standard dual basis has integer coefficients that are bounded in absolute value by $O_{c,C}(1)$ (part (2) of Lemma 5.5.8), one has $|\lambda_i - \lambda_i'| = O_{c,C}(\eta)$ and $|\mu_j - \mu_j'| = O_{c,C}(\eta)$ for each i and j.

Let $\mathbf{w_1^*}, \cdots, \mathbf{w_{d-u}^*}$ denote the standard dual basis of $(\mathbb{R}^{d-u})^*$, and define

$$\omega' := \sum_{j=1}^{d-u} \mu'_j \mathbf{w_j^*}.$$

Certainly $\omega' \in (\mathbb{Z}^{d-u})^T$. We claim that there exists a map $\varphi' \in (\mathbb{R}^{m-u})^*$ such that $\tau_1 \ll_{c,C} \|\varphi'\|_\infty \ll_{c,C} \tau_2^{-1}$ and $\|(L')^*\varphi' - \omega'\|_\infty \ll_{c,C} \eta$, which will immediately resolve (5.43) and part (9).

Indeed, recall the decomposition $\mathbb{R}^m = (\mathrm{span}(L\mathbf{x_i} : i \leqslant u)) \oplus \ker \Theta$ as an algebraic direct sum from (5.30). Let $\varphi = \varphi_1 + \varphi_2$, where $\varphi_1 \in (\mathrm{span}(L\mathbf{x_i} : i \leqslant u))^\circ$ and $\varphi_2 \in (\ker \Theta)^\circ$. Since $\mathrm{dist}(\varphi, (\ker \Theta)^\circ) \geqslant \tau_1$, we have $\|\varphi_1\|_\infty \geqslant \tau_1$. By the properties of the matrix P ((5.31) and (5.32)) there exists some $\varphi' \in (\mathbb{R}^{m-u})^*$ such that

$$\varphi_1 = P^* \pi_{m-u}^* \varphi'.$$

Furthermore, by evaluating φ' at the standard basis vectors, one sees that

$$\tau_1 \ll_{c,C} \|\varphi'\|_\infty \ll_{c,C} \tau_2^{-1}.$$

(We laid out the full details of such an argument when proving (5.40) during the proof of part (8) of the present lemma). We shall use this φ'.

By evaluating $L^*\varphi_1$ at the elements of $\mathcal{B}$ one immediately sees that

$$L^*\varphi_1 = \sum_{j=1}^{d-u} \mu_j \Xi(\mathbf{w_j})^*.$$

Hence

$$\Xi^* L^* P^* \pi_{m-u}^* \varphi' = \sum_{j=1}^{d-u} \mu_j \mathbf{w_j^*},$$

in other words $(L')^*\varphi' = \sum_{j=1}^{d-u} \mu_j \mathbf{w}_\mathbf{j}^*$. But since $|\mu_j - \mu'_j| = O_{c,C}(\eta)$ for each j, one has $\|(L')^*\varphi' - \omega'\|_\infty = O_{c,C}(\eta)$ as required. This settles part (9).

The entire lemma is settled. □

The final lemma we need in order to deduce Theorem 5.2.10 involves removing the sharp cut-off G.

Lemma 5.5.11 (Removing image cut-off)**.** *Let m, d, h be natural numbers, satisfying $d \geqslant h \geqslant m+1$. Let c, C, ε be positive, and let σ_G be any parameter in the range $0 < \sigma_G < 1/2$. Let $L' : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ be a purely irrational surjective map, and let $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ be an injective map. Suppose that $\|L'\|_\infty \leqslant C$ and that $\mathrm{dist}(L', V_{\mathrm{rank}}(m,h)) \geqslant c$. Let $F_{\widetilde{\mathbf{r}}} : \mathbb{R}^h \longrightarrow [0,1]$ be any function supported on $[-N,N]^h$, and let $G_{\widetilde{\mathbf{r}}} : \mathbb{R}^m \longrightarrow [0,1]$ be the indicator function of a convex set contained within $[-\varepsilon,\varepsilon]^m$. Then, for any parameter σ_G in the range $0 < \sigma_G < 1/2$, there exists a Lipschitz function $G_{\widetilde{\mathbf{r}},\sigma_G,1}$ supported on $[-O_{c,C,\varepsilon}(1), O_{c,C,\varepsilon}(1)]^m$, and with Lipschitz constant $O_{c,C,\varepsilon}(1/\sigma_G)$, such that, for any parameter τ_2 in the range $0 < \tau_2 \leqslant 1$ and for any functions $f_1, \cdots, f_d : [N] \longrightarrow [-1,1]$,*

$$\begin{aligned}&|T^{L',\Xi,\widetilde{\mathbf{r}}}_{F_{\widetilde{\mathbf{r}}},G_{\widetilde{\mathbf{r}}}}(f_1,\cdots,f_d)|\\ &\ll_{c,C,\varepsilon} |T^{L',\Xi,\widetilde{\mathbf{r}}}_{F_{\widetilde{\mathbf{r}}},G_{\widetilde{\mathbf{r}},\sigma_G,1}}(f_1,\cdots,f_d)| + \sigma_G + \frac{\tau_2^{1-o(1)}}{\sigma_G} + \frac{\tau_2^{-O(1)} A_L(\Omega_{c,C}(1),\tau_2)^{-1}}{N}.\end{aligned}$$

Proof. Applying Lemma 0.4.2 to the function $G_{\widetilde{\mathbf{r}}}$, we have

$$G_{\widetilde{\mathbf{r}}} = G_{\widetilde{\mathbf{r}},\sigma_G,1} + O(G_{\widetilde{\mathbf{r}},\sigma_G,2}),$$

where $G_{\widetilde{\mathbf{r}},\sigma_G,1}, G_{\widetilde{\mathbf{r}},\sigma_G,2} : \mathbb{R}^m \longrightarrow [0,1]$ are Lipschitz functions with Lipschitz constant $O_{c,C,\varepsilon}(1/\sigma_G)$, both supported on $[-O_{c,C,\varepsilon}(1), O_{c,C,\varepsilon}(1)]^m$, and with $\int_{\mathbf{x}} G_{\widetilde{\mathbf{r}},\sigma_G,2}(\mathbf{x})\, d\mathbf{x} = O_{c,C,\varepsilon}(\sigma_G)$.

By the triangle inequality,

$$|T^{L',\Xi,\widetilde{\mathbf{r}}}_{F_{\widetilde{\mathbf{r}}},G_{\widetilde{\mathbf{r}},\sigma_G,2}}(1,\cdots,1)| \leqslant T^{L'}_{F_{\widetilde{\mathbf{r}}},G_{\widetilde{\mathbf{r}},\sigma_G,2}}(1,\cdots,1).$$

We now apply Lemma 5.4.3, with linear map L' and Lipschitz function $G_{\widetilde{\mathbf{r}},\sigma_G,2}$. Inserting the bound from Lemma 5.4.3, the present lemma follows. □

We conclude this section by combining the three previous lemmas, along with Theorem 5.5.6, to deduce our main result.

Proof of Theorem 5.2.10 assuming Theorem 5.5.6. Assume the hypotheses of Theorem 5.2.10. Let σ_F and σ_G be any parameters satisfying $0 < \sigma_F, \sigma_G < 1/2$, and let τ_2 be any parameter satisfying $0 < \tau_2 \leqslant 1$.

By Lemma 5.5.9,

$$|T^L_{F,G}(f_1,\cdots,f_d)| \leqslant |T^L_{F_{1,\sigma_F},G}(f_1,\cdots,f_d)| + O_{c,C}(\sigma_F),$$

for some function $F_{1,\sigma_F} : \mathbb{R}^d \longrightarrow [0,1]$ supported on $[-2N,2N]^d$ and with Lipschitz constant $O(1/\sigma_F N)$. By part (4) of Lemma 5.5.10, writing F_{1,σ_F} for F, we have

$$|T^L_{F_{1,\sigma_F},G}(f_1,\cdots,f_d)| \leqslant \sum_{\widetilde{\mathbf{r}}\in\widetilde{R}} |T^{L',\Xi,\widetilde{\mathbf{r}}}_{F_{\widetilde{\mathbf{r}}},G_{\widetilde{\mathbf{r}}}}(f_1,\cdots,f_d)|,$$

where the objects $F_{\widetilde{\mathbf{r}}}$, $G_{\widetilde{\mathbf{r}}}$, L', Ξ and $\widetilde{R}$ satisfy all the conclusions of that lemma.

Parts (1), (5) and (6) of Lemma 5.5.10 show that Ξ and L' satisfy the hypotheses of Lemma 5.5.11, where in the notation of Lemma 5.5.11 we take $h := d - u$ and rewrite m for $m - u$. So, applying Lemma 5.5.11, there are some Lipschitz functions $G_{\widetilde{\mathbf{r}},\sigma_G,1} : \mathbb{R}^{m-u} \longrightarrow [0,1]$ supported on $[-O_{c,C,\varepsilon}(1), O_{c,C,\varepsilon}(1)]^{m-u}$ and with Lipschitz

constant $O_{c,C,\varepsilon}(1/\sigma_G)$ such that

$$|T^L_{F,G}(f_1,\cdots,f_d)|$$
$$\ll_{c,C,\varepsilon} \sum_{\widetilde{\mathbf{r}}\in\widetilde{R}} |T^{L',\Xi,\widetilde{\mathbf{r}}}_{F_{\widetilde{\mathbf{r}}},G_{\widetilde{\mathbf{r}},\sigma_G,1}}(f_1,\cdots,f_d)| + \sigma_G + \frac{\tau_2^{1-o(1)}}{\sigma_G} + \frac{\tau_2^{-O(1)}A_{L'}(\Omega_{c,C}(1),\tau_2)^{-1}}{N} + \sigma_F. \tag{5.44}$$

(Recall that $|\widetilde{R}| = O_{c,C,\varepsilon}(1)$, by part (2) of Lemma 5.5.10).

By conclusion (8) of Lemma 5.5.10, we may replace the term $A_{L'}(\Omega_{c,C}(1),\tau_2)^{-1}$ with the term $A_L(\Omega_{c,C}(1),\Omega_{c,C}(\tau_2))^{-1}$.

Since $F_{\widetilde{\mathbf{r}}}$, L', Ξ, and $\widetilde{R}$ together satisfy conclusions (1), (2), (3), (6), and (7) of Lemma 5.5.10, the hypotheses are satisfied so that we may apply Theorem 5.5.6 to the expression $T^{L',\Xi,\widetilde{\mathbf{r}}}_{F_{\widetilde{\mathbf{r}}},G_{\widetilde{\mathbf{r}},\sigma_G,1}}(f_1,\cdots,f_d)$. (We take $h = d-u$ and rewrite m for $m-u$, as above). Therefore there exists an s at most $d-2$, independent of $F_{\widetilde{\mathbf{r}}}$, $G_{\widetilde{\mathbf{r}}}$ and $\widetilde{\mathbf{r}}$, such that, if

$$\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho,$$

for some ρ in the range $0 < \rho \leqslant 1$ then $|T^L_{F,G}(f_1,\cdots,f_d)|$ is

$$\ll_{c,C,\varepsilon} \rho^{\Omega(1)}(\sigma_F^{-O(1)} + \sigma_G^{-O(1)}) + \sigma_F^{-O(1)}N^{-\Omega(1)}$$
$$+ \sigma_G + \frac{\tau_2^{1-o(1)}}{\sigma_G} + \frac{\tau_2^{-O(1)}A_L(\Omega_{c,C}(1),\Omega_{c,C}(\tau_2))^{-1}}{N} + \sigma_F. \tag{5.45}$$

It remains to pick appropriate parameters. Let C_1 be a constant that is suitably large in terms of c, C, and all $O(1)$ constants, and let c_1 be a constant that is suitably small in terms of all $O(1)$ constants. Pick $\sigma_F := \sigma_G := \rho^{c_1}$ and $\tau_2 := C_1\rho$. Then

$$|T^L_{F,G}(f_1,\cdots,f_d)| \ll_{c,C,\varepsilon} \rho^{\Omega(1)} + o_{\rho,A_L,c,C}(1),$$

where, after the combining the various error terms from (5.45), the $o_{\rho,A_L,c,C}(1)$ term may be bounded above by

$$N^{-\Omega(1)}\rho^{-O(1)}A_L(\Omega_{c,C}(1),\rho)^{-1},$$

as $A_L(\tau_1,\tau_2)$ is monotonically decreasing as τ_2 decreases. This is the desired conclusion of Theorem 5.2.10. □

In order to resolve our main result, then, it suffices to prove[7] Theorem 5.5.6.

5.6 Transfer from $\mathbb{Z}$ to $\mathbb{R}$

As remarked above, our present task is to prove Theorem 5.5.6. Any reader only wishing to consider the case of diophantine inequalities with Lipschitz cut-offs may begin here, and eschew section 5.5.

We devote this section to the formulation and proof of a certain 'transfer' argument, whereby we replace the discrete summation in the definition of $T_{F,G}^{L,\Xi,\widetilde{\mathbf{r}}}(f_1,\cdots,f_d)$ with an integral.

Let us introduce some notation for the integral in question.

Definition 5.6.1. *Let N,m,d,h be natural numbers, with $d \geqslant h \geqslant m+2$. Let ε be positive. Let $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ and $L : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ be linear maps. Let $F : \mathbb{R}^h \to [0,1]$ and $G : \mathbb{R}^m \to [0,1]$ be two functions, with F supported on $[-N,N]^h$ and G supported*

[7]The reader may have noticed from the proof above that, in fact, it suffices to prove Theorem 5.5.6 in the case that L is purely irrational, but the general version is no harder to prove.

on $[-\varepsilon,\varepsilon]^m$. *Let* $g_1,\cdots,g_d:\mathbb{R}\longrightarrow[-1,1]$ *be arbitrary functions. We define*

$$\widetilde{T}_{F,G}^{L,\Xi,\widetilde{\mathbf{r}}}(g_1,\cdots,g_d):=\frac{1}{N^{h-m}}\int\limits_{\mathbf{x}\in\mathbb{R}^h}\Big(\prod_{j=1}^{d}g_j(\xi_j(\mathbf{x})+\widetilde{\mathbf{r}}_j)\Big)F(\mathbf{x})G(L\mathbf{x})\,d\mathbf{x}. \tag{5.46}$$

Next, we determine a particular class of measurable functions that will be useful to us.

Definition 5.6.2. *Let* $\chi:\mathbb{R}\longrightarrow[0,1]$ *be a measurable function, and let* η *be a positive parameter. We say that* χ *is* η-supported *if* χ *is supported on* $[-\eta,\eta]$ *and* $\chi(x)\equiv 1$ *for all* $x\in[-\eta/2,\eta/2]$.

If $f:\mathbb{Z}\longrightarrow\mathbb{R}$ has finite support, and $\chi:\mathbb{R}\longrightarrow[0,1]$ is a measurable function, we may define the (rather singular) convolution $(f*\chi)(x):\mathbb{R}\longrightarrow\mathbb{R}$ by

$$(f*\chi)(x):=\sum_{n\in\mathbb{Z}}f(n)\chi(x-n).$$

If χ is η-supported, for small enough η, then there is only one possible integer n that makes a non-zero contribution to the sum.

We now state the key lemma.

Lemma 5.6.3. *Let* N,m,d,h *be natural numbers, with* $d\geqslant h\geqslant m+2$, *and let* c,C,ε,η *be positive constants. Let* $\Xi:\mathbb{R}^h\longrightarrow\mathbb{R}^d$ *be an injective linear map with integer coefficients, and assume that* $\Xi(\mathbb{Z}^h)=\mathbb{Z}^d\cap\operatorname{im}\Xi$. *Let* $L:\mathbb{R}^h\longrightarrow\mathbb{R}^m$ *be a surjective linear map. Assume that* $\|\Xi\|_\infty\leqslant C$, $\|L\|_\infty\leqslant C$, *and* $\operatorname{dist}(L,V_{\mathrm{rank}}(m,h))\geqslant c$. *Let* $F:\mathbb{R}^h\longrightarrow[0,1]$ *be a Lipschitz function supported on* $[-N,N]^h$ *with Lipschitz constant* $O(1/\sigma_F N)$, *and let* $G:\mathbb{R}^m\longrightarrow[0,1]$ *be a Lipschitz function supported on* $[-\varepsilon,\varepsilon]^m$ *with Lipschitz constant* $O(1/\sigma_G)$. *Let* $\widetilde{\mathbf{r}}$ *be a fixed vector in* $\mathbb{Z}^d$, *satisfying* $\|\widetilde{\mathbf{r}}\|_\infty=O_{c,C,\varepsilon}(1)$. *Let* $\chi:\mathbb{R}\longrightarrow[0,1]$ *be an* η-*supported measurable function. Then, if* η *is small enough (in terms of the dimensions* m,d,h, C, *and* ε*) there exists some*

positive real number $C_{\Xi,\chi}$ such that, if $f_1, \cdots, f_d : [N] \longrightarrow [-1,1]$ are arbitrary functions,

$$T_{F,G}^{\Xi,L,\widetilde{\mathbf{r}}}(f_1,\cdots,f_d) = \frac{1}{C_{\Xi,\chi}\eta^h}\widetilde{T}_{F,G}^{\Xi,L,\widetilde{\mathbf{r}}}(f_1*\chi,\cdots,f_d*\chi) + O_{C,c,\varepsilon}(\eta/\sigma_G) + O_{C,c,\varepsilon}(\eta/\sigma_F N). \tag{5.47}$$

Moreover, $C_{\Xi,\chi} \asymp_C 1$ for some absolute implied constants.

Proof. Let $\boldsymbol{\chi} : \mathbb{R}^d \longrightarrow [0,1]$ denote the function $\mathbf{x} \mapsto \prod_{i=1}^{d} \chi(x_i)$. We choose

$$C_{\Xi,\chi} := \frac{1}{\eta^h} \int\limits_{\mathbf{x}\in\mathbb{R}^h} \boldsymbol{\chi}(\Xi(\mathbf{x}))\, d\mathbf{x}.$$

Since χ is η-supported, $C_{\Xi,\chi} \asymp_C 1$.

Then, expanding the definition of the convolution,

$$\frac{1}{C_{\Xi,\chi}\eta^h}\widetilde{T}_{F,G}^{L,\Xi,\widetilde{\mathbf{r}}}(f_1*\chi,\cdots,f_d*\chi)$$

equals

$$\frac{1}{N^{h-m}} \sum_{\mathbf{n}\in\mathbb{Z}^d} \Big(\prod_{j=1}^{d} f_j(n_j)\Big) \frac{1}{C_{\Xi,\chi}\eta^h} \int\limits_{\mathbf{y}\in\mathbb{R}^h} F(\mathbf{y})G(L\mathbf{y})\boldsymbol{\chi}(\Xi(\mathbf{y}) + \widetilde{\mathbf{r}} - \mathbf{n})\, d\mathbf{y}. \tag{5.48}$$

Note that any vector $\mathbf{n} \in \mathbb{Z}^d$ that gives a non-zero contribution to expression (5.48) satisfies $\|\mathbf{n} - \Xi(\mathbf{y}) - \widetilde{\mathbf{r}}\|_\infty \ll \eta$, for some $\mathbf{y} \in \mathbb{R}^h$. Therefore, $\mathbf{n}$ must be of the form $\Xi(\mathbf{n}') + \widetilde{\mathbf{r}}$ for some unique $\mathbf{n}' \in \mathbb{Z}^h$. (This is proved in full in Lemma 5.11.2). Therefore, writing $\Xi = (\xi_1, \cdots, \xi_d)$, we may reformulate (5.48) as

$$\frac{1}{N^{h-m}} \sum_{\mathbf{n}\in\mathbb{Z}^h} \Big(\prod_{j=1}^{d} f_j(\xi_j(\mathbf{n}) + \widetilde{\mathbf{r}}_j)\Big) \frac{1}{C_{\Xi,\chi}\eta^h} \int\limits_{\mathbf{y}\in\mathbb{R}^h} F(\mathbf{y})G(L\mathbf{y})\boldsymbol{\chi}(\Xi(\mathbf{y}-\mathbf{n}))\, d\mathbf{y},$$

which is equal to

$$\frac{1}{N^{h-m}}\sum_{\mathbf{n}\in\mathbb{Z}^h}\Big(\prod_{j=1}^{d}f_j(\xi_j(\mathbf{n})+\widetilde{\mathbf{r}}_j)\Big)\frac{1}{C_{\Xi,\chi}\eta^h}\int\limits_{\mathbf{y}\in\mathbb{R}^h}(F(\mathbf{n})+O_C(\eta/\sigma_F N))G(L\mathbf{y})\boldsymbol{\chi}(\Xi(\mathbf{y}-\mathbf{n}))\,d\mathbf{y}. \tag{5.49}$$

Indeed, the inner integral is only non-zero when $\|\Xi(\mathbf{y})-\Xi(\mathbf{n})\|_\infty \ll \eta$, and this implies that $\|\mathbf{y}-\mathbf{n}\|_\infty \ll C^{-O(1)}\eta$. (This is proved in full in Lemma 5.11.3). Then recall that F has Lipschitz constant $O(1/\sigma_F N)$.

Continuing, expression (5.49) is equal to

$$\frac{1}{N^{h-m}}\sum_{\mathbf{n}\in\mathbb{Z}^h}\Big(\prod_{j=1}^{d}f_j(\xi_j(\mathbf{n})+\widetilde{\mathbf{r}}_j)\Big)F(\mathbf{n})H(L\mathbf{n})+E \tag{5.50}$$

where

$$H(\mathbf{x})=\frac{1}{C_{\Xi,\chi}\eta^h}\int\limits_{\mathbf{y}\in\mathbb{R}^h}\boldsymbol{\chi}(\Xi(\mathbf{y}))G(\mathbf{x}+L\mathbf{y})\;\;d\mathbf{y}$$

and E is a certain error, that may be bounded above by

$$\ll_C \frac{\eta}{\sigma_F N}\frac{1}{N^{h-m}}\sum_{\mathbf{n}\in[-N,N]^h}H(L\mathbf{n}). \tag{5.51}$$

Let us deal with the first term of (5.50), in which we wish to replace H with G. We therefore consider

$$\Big|\frac{1}{N^{h-m}}\sum_{\mathbf{n}\in\mathbb{Z}^h}\Big(\prod_{j=1}^{d}f_j(\xi_j(\mathbf{n})+\widetilde{\mathbf{r}}_j)\Big)F(\mathbf{n})(G(L\mathbf{n})-H(L\mathbf{n}))\Big|,$$

which is

$$\leqslant \frac{1}{N^{h-m}}\sum_{\mathbf{n}\in\mathbb{Z}^h}F(\mathbf{n})|G-H|(L\mathbf{n}). \tag{5.52}$$

Using Lemma 5.11.3 again, the function H is supported on $[-\varepsilon-O_C(\eta),\varepsilon+O_C(\eta)]^m$.

Thus, if η is small enough in terms of ε, the function $|G-H| : \mathbb{R}^m \longrightarrow \mathbb{R}$ is supported on $[-O_C(\varepsilon), O_C(\varepsilon)]^m$. Furthermore, $\|G-H\|_\infty = O_C(\eta/\sigma_G)$. Indeed,

$$\begin{aligned} & G(\mathbf{x}) - \frac{1}{C_{\Xi,\chi}\eta^h} \int\limits_{\mathbf{y}\in\mathbb{R}^h} G(\mathbf{x}+L\mathbf{y})\chi(\Xi(\mathbf{y}))\,d\mathbf{y} \\ &= G(\mathbf{x}) - \frac{1}{C_{\Xi,\chi}\eta^h} \int\limits_{\mathbf{y}\in\mathbb{R}^h} (G(\mathbf{x}) + O_C(\eta/\sigma_G))\boldsymbol{\chi}(\Xi(\mathbf{y}))\,d\mathbf{y} \\ &= O_C(\eta/\sigma_G), \end{aligned}$$

by the definition of $C_{\Xi,\chi}$. So, by the crude bound given in Lemma 5.4.1, (5.52) is $O_{c,C,\varepsilon}(\eta/\sigma_G)$.

Turning to the error E from (5.50), we've already remarked that it may be bounded above by expression (5.51). Applying Lemma 5.4.1 again, expression (5.51) is $O_{c,C,\varepsilon}(\eta/\sigma_F N)$.

Lemma 5.6.3 follows immediately upon substituting the estimates on (5.51) and (5.52) into (5.50). □

We finish this section by noting a simple relationship between the Gowers norms $\|f*\chi\|_{U^{s+1}(\mathbb{R})}$ and the Gowers norms $\|f\|_{U^{s+1}[N]}$.

Lemma 5.6.4 (Relating different Gowers norms)**.** *Let s be a natural number, and assume that η is a positive parameter that is small enough in terms of s. Let $\chi : \mathbb{R} \longrightarrow [0,1]$ be an η-supported measurable function. Let N be a natural number, and let $f : [N] \longrightarrow [-1,1]$ be an arbtirary function. View $f*\chi$ as a function supported on $[-2N, 2N]$. Then we have*

$$\|f*\chi\|_{U^{s+1}(\mathbb{R})} \ll \eta^{\frac{s+2}{2^{s+1}}} \|f\|_{U^{s+1}[N]}. \tag{5.53}$$

The definition of the real Gowers norm $\|f * \chi\|_{U^{s+1}(\mathbb{R})}$ is recorded in Definition 0.3.3.

Proof. From expression (5), we have

$$\|f * \chi\|_{U^{s+1}(\mathbb{R})}^{2^{s+1}} \ll \frac{1}{N^{s+2}} \int\limits_{(x,\mathbf{h})\in\mathbb{R}^{s+2}} \prod_{\boldsymbol{\omega}\in\{0,1\}^{s+1}} (f * \chi)(x + \mathbf{h}\cdot\boldsymbol{\omega})\, dx\, d\mathbf{h}.$$

Substituting in the definition of $f * \chi$, this is equal to

$$\frac{1}{N^{s+2}} \sum_{(n_{\boldsymbol{\omega}})_{\boldsymbol{\omega}\in\{0,1\}^{s+1}}\in\mathbb{Z}^{\{0,1\}^{s+1}}} \Big(\prod_{\boldsymbol{\omega}\in\{0,1\}^{s+1}} f(n_{\boldsymbol{\omega}})\Big) \int\limits_{(x,\mathbf{h})\in\mathbb{R}^{s+2}} \boldsymbol{\chi}(\Psi(x,\mathbf{h}) - \mathbf{n})\, dx\, d\mathbf{h}, \quad (5.54)$$

where $\Psi : \mathbb{R}^{s+2} \longrightarrow \mathbb{R}^{2^{s+1}}$ has coordinate functions $\psi_{\boldsymbol{\omega}}$, indexed by $\boldsymbol{\omega} \in \{0,1\}^{s+1}$, where $\psi_{\boldsymbol{\omega}}(x,\mathbf{h}) := x + \mathbf{h}\cdot\boldsymbol{\omega}$. In similar notation to that used in the previous proof, for $\mathbf{x} \in \mathbb{R}^{2^{s+1}}$, we let $\boldsymbol{\chi}(\mathbf{x}) := \prod_{i=1}^{2^{s+1}} \chi(x_i)$. Note that Ψ is injective, $\Psi(\mathbb{Z}^{s+2}) = \mathbb{Z}^{2^{s+1}} \cap \operatorname{im}\Psi$, and $\|\Psi\|_\infty = O_s(1)$.

The contribution to the inner integral of (5.54) from a particular $\mathbf{n}$ is zero unless $\|\mathbf{n} - \Psi(x,\mathbf{h})\|_\infty \ll \eta$, for some $(x,\mathbf{h}) \in \mathbb{R}^{s+2}$. Therefore, if η is small enough we can conclude that $\mathbf{n}$ must be of the form $\Psi(p,\mathbf{k})$, for some unique $(p,\mathbf{k}) \in \mathbb{Z}^{s+2}$. (To spell it out, apply Lemma 5.11.2 with the map Ψ in place of the map Ξ). So (5.54) is equal to

$$\frac{1}{N^{s+2}} \sum_{(p,\mathbf{k})\in\mathbb{Z}^{s+2}} \Big(\prod_{\boldsymbol{\omega}\in\{0,1\}^{s+1}} f(p + \mathbf{k}\cdot\boldsymbol{\omega})\Big) \int\limits_{(x,\mathbf{h})\in\mathbb{R}^{s+2}} \boldsymbol{\chi}(\Psi(x-p,\mathbf{h}-\mathbf{k}))\, dx\, d\mathbf{h}, \quad (5.55)$$

which, after a change of variables , is equal to

$$\frac{C}{N^{s+2}} \sum_{(p,\mathbf{k})\in\mathbb{Z}^{s+2}} \prod_{\boldsymbol{\omega}\in\{0,1\}^{s+1}} f(p + \mathbf{k}\cdot\boldsymbol{\omega}), \quad (5.56)$$

where

$$C := \int_{(x,\mathbf{h})\in\mathbb{R}^{s+2}} \boldsymbol{\chi}(\Psi(x,\mathbf{h}))\,dx\,d\mathbf{h}.$$

Since $\boldsymbol{\chi}$ has support contained within $[-\eta,\eta]^{2^{s+1}}$, a vector $(x,\mathbf{h})$ only makes a non-zero contribution to the above integral if $\|\Psi(x,\mathbf{h})\|_\infty \ll \eta$. This implies that $\|(x,\mathbf{h})\|_\infty \ll \eta$. (To prove this is full, apply Lemma 5.11.3 to the linear map Ψ). Since $\|\boldsymbol{\chi}\|_\infty = O(1)$, this means $C = O(\eta^{s+2})$. The lemma then follows from (5.56). □

5.7 Degeneracy relations

Our aim for this short section is to establish a quantitative relationship between the dual pair degeneracy variety $V^*_{\text{degen},2}(m,d,h)$ and the dual degeneracy variety $V_{\text{degen}}(h-m,d)$ (see Definitions 5.5.4 and 5.2.3 respectively), which will be needed in the next section. To introduce the ideas, we first prove a non-quantitative proposition.

Lemma 5.7.1. *Let m,d,h be natural numbers, with $d \geqslant h \geqslant m+2$. Let $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ be an injective linear map, and let $L : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ be a surjective linear map, and suppose that $(\Xi, L) \notin V^*_{\text{degen},2}(m,d,h)$. Let $\Phi : \mathbb{R}^{h-m} \longrightarrow \ker L$ be any surjective linear map. Then the linear map $\Xi\Phi : \mathbb{R}^{h-m} \longrightarrow \mathbb{R}^d$, viewed as a system of homogenous linear forms, is not in $V_{\text{degen}}(h-m,d)$.*

Proof. Let $\mathbf{e_1},\cdots,\mathbf{e_d}$ denote the standard basis vectors in $\mathbb{R}^d$, and let $\mathbf{e_1^*},\cdots,\mathbf{e_d^*}$ denote the dual basis of $(\mathbb{R}^d)^*$. Suppose for contradiction that $\Xi\Phi \in V_{\text{degen}}(h-m,d)$. Then by definition there exist two indices $i,j \leqslant d$, and a real number λ, such that $\Xi\Phi(\mathbb{R}^{h-m}) \subset \ker(\mathbf{e_i^*} - \lambda\mathbf{e_j^*})$.

But then $\Phi(\mathbb{R}^{h-m}) \subset \ker(\Xi^*(\mathbf{e_i^*}-\lambda\mathbf{e_j^*}))$, i.e. $\Xi^*(\mathbf{e_i^*}-\lambda\mathbf{e_j^*}) \in (\ker L)^\circ$. But $(\ker L)^\circ = L^*((\mathbb{R}^m)^*)$, and so $\Xi^*(\mathbf{e_i^*} - \lambda\mathbf{e_j^*}) \in L^*((\mathbb{R}^m)^*)$.

Then, by the definition of $V^*_{\text{degen},2}(m,d,h)$, we have $(\Xi,L) \in V^*_{\text{degen},2}(m,d,h)$, which is a contradiction. □

The ideas having been introduced, we state the quantitative version we require.

Lemma 5.7.2. *Let m, d, h be natural numbers, with $d \geqslant h \geqslant m+2$, and let c, C be positive constants. Let $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ be a linear map, and let $L : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ be a surjective linear map. Suppose that $\|\Xi\|_\infty \leqslant C$, and $\operatorname{dist}((\Xi, L), V^*_{\mathrm{degen},2}(m,d,h)) \geqslant c$. Let K denote $\ker L$, choose any orthonormal basis $\{\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(h-m)}}\}$ for K, and let $\Phi : \mathbb{R}^{h-m} \longrightarrow K$ denote the associated parametrisation, i.e. $\Phi(\mathbf{x}) := \sum_{i=1}^{h-m} x_i \mathbf{v^{(i)}}$. Then $\|\Xi\Phi\|_\infty = O(C)$ and $\operatorname{dist}(\Xi\Phi, V_{\mathrm{degen}}(h-m,d)) = \Omega(c)$.*

For the definition of $\operatorname{dist}((\Xi, L), V^*_{\mathrm{degen},2}(m,d,h))$, consult Definition 5.5.5.

Proof. Certainly $\|\Phi\|_\infty = O(1)$, as the chosen basis $\{\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(h-m)}}\}$ is orthonormal. Therefore $\|\Xi\Phi\|_\infty = O_C(1)$.

Let $\mathbf{e_1}, \cdots, \mathbf{e_d}$ denote the standard basis vectors in $\mathbb{R}^d$, and let $\mathbf{e_1^*}, \cdots, \mathbf{e_d^*}$ denote the dual basis of $(\mathbb{R}^d)^*$. Suppose for contradiction that $\operatorname{dist}(\Xi\Phi, V_{\mathrm{degen}}(h-m,d)) \leqslant \eta$ for some small parameter η. In other words, assume that there exists a linear map $P : \mathbb{R}^{h-m} \longrightarrow \mathbb{R}^d$ with $\|P\|_\infty \leqslant \eta$ such that $\Xi\Phi + P \in V_{\mathrm{degen}}(h-m,d)$. By definition, this means that

$$(\Xi\Phi + P)(\mathbb{R}^{h-m}) \subset \ker(\mathbf{e_i^*} - \lambda \mathbf{e_j^*}),$$

for some two indices $i, j \leqslant d$, and some real number λ.

We can factorise $P = Q\Phi$, for some linear map $Q : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ with $\|Q\|_\infty \ll \eta$. Indeed, let $\mathbf{f_1}, \cdots, \mathbf{f_{h-m}}$ denote the standard basis vectors in $\mathbb{R}^{h-m}$, and for all k at most $h-m$ define

$$Q(\mathbf{v^{(k)}}) := P(\mathbf{f_k}).$$

(If the notation for the indices seems odd here, it is designed to match the notation in Proposition 5.8.2, in which having superscript on the vectors $\mathbf{v^{(k)}}$ is natural). Complete $\{\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(h-m)}}\}$ to an orthonormal basis $\{\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(h)}}\}$ for $\mathbb{R}^h$ and, for k in the range $h-m+1 \leqslant k \leqslant h-m$, define $Q(\mathbf{v^{(k)}}) := \mathbf{0}$. Then $P = Q\Phi$, and $\|Q\|_\infty = O(\eta)$, since $\{\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(h)}}\}$ is an orthonormal basis.

Thus,

$$(\Xi\Phi + Q\Phi)(\mathbb{R}^{h-m}) \subset \ker(\mathbf{e}_\mathbf{i}^* - \lambda\mathbf{e}_\mathbf{j}^*).$$

So

$$\Phi(\mathbb{R}^{h-m}) \subset \ker((\Xi + Q)^*(\mathbf{e}_\mathbf{i}^* - \lambda\mathbf{e}_\mathbf{j}^*)).$$

Like the previous proof, we conclude that

$$(\Xi + Q)^*(\mathbf{e}_\mathbf{i}^* - \lambda\mathbf{e}_\mathbf{j}^*) \in L^*((\mathbb{R}^m)^*).$$

Hence $((\Xi + Q), L) \in V^*_{\text{degen},2}(m, d, h)$, which, if η is small enough, contradicts the assumption that $\text{dist}((\Xi, L), V^*_{\text{degen},2}(m, d, h)) \geqslant c$. □

We remark that a non-effective version of Lemma 5.7.2, would have followed from Lemma 5.7.1 by soft analysis.

5.8 A Generalised von Neumann Theorem

In this section we complete the proof of Theorem 5.5.6, and therefore complete the proof of our main result (Theorem 5.2.10). It will be enough to prove the following statement.

Theorem 5.8.1. *Let N, m, d, h be natural numbers, with $d \geqslant h \geqslant m + 2$, and let c, C, ε be positive reals. Let $\Xi = \Xi(N) : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ be an injective linear map with integer coefficients, and let $L = L(N) : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map. Suppose further that $\|L\|_\infty \leqslant C$, $\|\Xi\|_\infty \leqslant C$, $\text{dist}(L, V_{\text{rank}}(m, d)) \geqslant c$ and $\text{dist}((\Xi, L), V^*_{\text{degen},2}(m, d, h)) \geqslant c$. Then there is some natural number s at most $d - 2$ such that the following holds. Let $\widetilde{\mathbf{r}} \in \mathbb{Z}^d$ be some vector with $\|\widetilde{\mathbf{r}}\|_\infty = O_{c,C,\varepsilon}(1)$, and let σ_F be a parameter in the range $O < \sigma_F < 1/2$. Let $F : \mathbb{R}^h \longrightarrow [0, 1]$*

be a Lipschitz function supported on $[-N,N]^h$*, with Lipschitz constant* $O(1/\sigma_F N)$*, and let* $G : \mathbb{R}^m \longrightarrow [0,1]$ *be any function supported on* $[-\varepsilon,\varepsilon]^m$*. Let* $g_1,\cdots,g_d : [-2N,2N]^d \longrightarrow [-1,1]$ *be arbitrary measurable functions. Suppose*

$$\min_{j\leqslant d} \|g_j\|_{U^{s+1}(\mathbb{R})} \leqslant \rho$$

for some ρ *at most* 1*. Then*

$$|\widetilde{T}^{L,\Xi,\widetilde{\mathbf{r}}}_{F,G}(g_1,\cdots,g_d)| \ll_{c,C,\varepsilon} \rho^{\Omega(1)}\sigma_F^{-1} \tag{5.57}$$

Proof that 5.8.1 implies Theorem 5.5.6. Assume the hypotheses of Theorem 5.5.6. This gives natural numbers N,m,d,h, linear maps $L : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ and $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$, and functions $F : \mathbb{R}^h \longrightarrow [0,1]$ and $G : \mathbb{R}^m \longrightarrow [0,1]$. Let $f_1,\cdots,f_d : [N] \longrightarrow [-1,1]$ be arbitrary functions, and for ease of notation let

$$\delta := T^{L,\Xi,\widetilde{\mathbf{r}}}_{F,G}(f_1,\cdots,f_d).$$

From Lemma 5.4.1 and the triangle inequality, we have the crude bound $\delta = O_{c,C,\varepsilon}(1)$. Let $\eta := c_1\delta\sigma_G$, where c_1 is small enough depending on m,d,h,c,C,ε, and let $\chi : \mathbb{R} \longrightarrow [0,1]$ be an η-supported measurable function (see Definition 5.6.2). For all j at most d, let $g_j := f_j * \chi$. Finally, suppose $\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho$, for some parameter ρ in the range $0 < \rho \leqslant 1$.

We proceed by bounding $\widetilde{T}^{L,\Xi,\widetilde{\mathbf{r}}}_{F,G}(g_1,\cdots,g_d)$. Indeed, by Lemma 5.6.4, if c_1 is small enough

$$\min_j \|g_j\|_{U^{s+1}(\mathbb{R})} \ll \eta^{\frac{s+2}{2^{s+1}}} \min_j \|f_j\|_{U^{s+1}[N]} \ll_{c,C,\varepsilon} \rho.$$

Applying Theorem 5.8.1 to these function $g_1,\cdots,g_d$, this implies

$$\widetilde{T}^{L,\Xi,\widetilde{\mathbf{r}}}_{F,G}(g_1,\cdots,g_d) \ll_{c,C,\varepsilon} \rho^{\Omega(1)}\sigma_F^{-1}. \tag{5.58}$$

Now we use this to bound δ by Gowers norms. Indeed, by Lemma 5.6.3, we have

$$\delta \ll_{c,C,\varepsilon} \frac{1}{(c_1\delta\sigma_G)^h}\widetilde{T}^{L,\Xi,\widetilde{\mathbf{r}}}_{F,G}(g_1,\cdots,g_d) + c_1\delta + c_1\delta\sigma_G\sigma_F^{-1}N^{-1}.$$

Picking c_1 small enough, we may move the $c_1\delta$ term to the left-hand side to get an $\Omega(\delta)$ term. The bound (5.58) then yields

$$\delta^{h+1} \ll_{c,C,\varepsilon} \rho^{\Omega(1)}\sigma_F^{-1}\sigma_G^{-h} + \sigma_F^{-1}N^{-1},$$

and so

$$\delta \ll_{c,C,\varepsilon} \rho^{\Omega(1)}(\sigma_F^{-O(1)} + \sigma_G^{-O(1)}) + \sigma_F^{-O(1)}N^{-\Omega(1)}.$$

This yields the desired conclusion of Theorem 5.5.6. □

So it remains to prove Theorem 5.8.1. The bulk of the work will be done in the following two propositions.

Proposition 5.8.2 (Separating out the kernel)**.** *Let N, m, d, h be natural numbers, with $d \geqslant h \geqslant m+2$, and let c, C, ε be positive constants. Let σ_F be a parameter in the range $0 < \sigma_F < 1/2$. Let $\Xi : \mathbb{R}^h \to \mathbb{R}^d$ be an injective linear map with integer coefficients, and let $L : \mathbb{R}^h \longrightarrow \mathbb{R}^m$ be a surjective linear map. Assume further that $\|L\|_\infty \leqslant C$, $\|\Xi\|_\infty \leqslant C$, $\mathrm{dist}(L, V_{\mathrm{rank}}(m,h)) \geqslant c$ and $\mathrm{dist}((\Xi, L), V^*_{\mathrm{degen},2}(m,d,h)) \geqslant c$. Let $F : \mathbb{R}^h \longrightarrow [0,1]$ be a Lipschitz function supported on $[-CN, CN]^h$, with Lipschitz constant $O_C(1/\sigma_F N)$, and let $G : \mathbb{R}^m \longrightarrow [0,1]$ be a Lipschitz function supported on $[-\varepsilon, \varepsilon]^m$. Let $\widetilde{\mathbf{r}}$ be a fixed vector in $\mathbb{Z}^d$, satisfying $\|\widetilde{\mathbf{r}}\|_\infty = O_C(1)$. Then there exists a system of linear forms $(\psi_1, \cdots, \psi_d) = \Psi : \mathbb{R}^{h-m} \longrightarrow \mathbb{R}^d$, and a Lipschitz function $F_1 : \mathbb{R}^{h-m} \longrightarrow [0,1]$ supported on $[-O_{c,C,\varepsilon}(N), O_{c,C,\varepsilon}(N)]^{h-m}$ with Lipschitz constant*

$O(1/\sigma_F N)$, *such that, if* $g_1, \cdots, g_d : [-2N, 2N] \longrightarrow [-1, 1]$ *are arbitrary functions,*

$$|\widetilde{T}_{F,G}^{L,\Xi,\widetilde{\mathbf{r}}}(g_1, \cdots, g_d)| \ll_{c,C,\varepsilon} \Big|\frac{1}{N^{h-m}} \int_{\mathbf{x}} \prod_{j=1}^{d} g_j(\psi_j(\mathbf{x}) + a_j) F_1(\mathbf{x})\, d\mathbf{x}\Big|, \tag{5.59}$$

where, for each j, a_j *is some real number that satisfies* $a_j = O_C(1)$.

Furthermore, there exists a natural number s *at most* $d-2$ *such that the system* Ψ *has* $\Omega_{c,C}(1)$*-Cauchy-Schwarz complexity at most* s*, in the sense of Definition 5.3.6.*

Proof of Proposition 5.8.2. For ease of notation, let

$$\beta := |\widetilde{T}_{F,G}^{L,\Xi,\widetilde{\mathbf{r}}}(g_1, \cdots, g_d)|.$$

Noting that $\ker L$ is a vector space of dimension $h-m$, define $\{\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(h-m)}}\} \subset \mathbb{R}^h$ to be an orthonormal basis for $\ker L$. Then the map $\Phi : \mathbb{R}^{h-m} \longrightarrow \mathbb{R}^h$, defined by

$$\Phi(\mathbf{x}) := \sum_{i=1}^{h-m} x_i \mathbf{v^{(i)}}, \tag{5.60}$$

is an injective map that parametrises $\ker L$. (This is reminiscent of Lemma 5.7.2).

Now, extend the orthonormal basis $\{\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(h-m)}}\}$ for $\ker L$ to an orthonormal basis $\{\mathbf{v^{(1)}}, \cdots, \mathbf{v^{(h)}}\}$ for $\mathbb{R}^h$. By implementing a change of basis, we may rewrite

$$\beta = \frac{1}{N^{h-m}} \int_{\mathbf{x}\in\mathbb{R}^h} F(\sum_{i=1}^{h} x_i \mathbf{v^{(i)}}) G(L(\sum_{i=1}^{h} x_i \mathbf{v^{(i)}}))(\prod_{j=1}^{d} g_j(\xi_j(\Phi(\mathbf{x}) + \sum_{i=h-m+1}^{h} x_i \mathbf{v^{(i)}}) + \widetilde{\mathbf{r}}_j))\, d\mathbf{x}. \tag{5.61}$$

We wish to remove the presence of the variables $x_{h-m+1}, \cdots, x_h$. To set this up, note that, by the choice of the vectors $\mathbf{v^{(i)}}$,

$$G(L(\sum_{i=1}^{h} x_i \mathbf{v^{(i)}})) = G(L(\sum_{i=h-m+1}^{h} x_i \mathbf{v^{(i)}})).$$

The vector $\sum_{i=h-m+1}^{h} x_i \mathbf{v^{(i)}}$ is in $(\ker L)^{\perp}$. Hence, due to the limited support of G,

there is a domain D, contained in $[-O_{\varepsilon,c,C}(1), O_{\varepsilon,c,C}(1)]^m$, such that $G(L(\sum_{i=h-m+1}^{h} x_i\mathbf{v^{(i)}}))$ is equal to zero unless $(x_{h-m+1}, \cdots, x_h)^T \in D$. (This is proved in full in Lemma 5.11.1).

We can use this observation to bound the right-hand side of (5.61). Indeed, using $\mathbf{x_1^{h-m}}$ to refer to the vector in $\mathbb{R}^{h-m}$ given by the first the first $h-m$ coordinates of $\mathbf{x}$, we have

$$\beta \ll \operatorname{vol} D \times \sup_{\mathbf{x_{h-m+1}^h} \in D} \frac{1}{N^{h-m}} \Big| \int\limits_{\mathbf{x_1^{h-m}} \in \mathbb{R}^{h-m}} F(\sum_{i=1}^{h} x_i\mathbf{v^{(i)}}) G(L(\sum_{i=h-m+1}^{h} x_i\mathbf{v^{(i)}})) \\ (\prod_{j=1}^{d} g_j(\xi_j(\Phi(\mathbf{x_1^{h-m}}) + \sum_{i=h-m+1}^{h} x_i\mathbf{v^{(i)}}) + \widetilde{\mathbf{r}}_j))\, d\mathbf{x_1^{h-m}} \Big|. \quad (5.62)$$

So there exists some fixed vector $(x_{h-m+1}, \cdots, x_h)^T$ in D such that

$$\beta \ll_{c,C,\varepsilon} \frac{1}{N^{h-m}} \Big| \int\limits_{\mathbf{x_1^{h-m}} \in \mathbb{R}^{h-m}} F(\sum_{i=1}^{h} x_i\mathbf{v^{(i)}}) G(L(\sum_{i=h-m+1}^{h} x_i\mathbf{v^{(i)}})) \\ (\prod_{j=1}^{d} g_j(\xi_j(\Phi(\mathbf{x_1^{h-m}}) + \sum_{i=h-m+1}^{h} x_i\mathbf{v^{(i)}}) + \widetilde{\mathbf{r}}_j))\, d\mathbf{x_1^{h-m}} \Big|. \quad (5.63)$$

Define the function $F_1 : \mathbb{R}^{h-m} \longrightarrow [0,1]$ by

$$F_1(\mathbf{x_1^{h-m}}) := F(\Phi(\mathbf{x_1^{h-m}}) + \sum_{i=h-m+1}^{h} x_i\mathbf{v^{(i)}})$$

and for each j at most d, a shift

$$a_j := \xi_j(\sum_{i=h-m+1}^{h} x_i\mathbf{v^{(i)}}) + \widetilde{\mathbf{r}}_j.$$

Then

$$\beta \ll_{c,C,\varepsilon} \Big|\frac{1}{N^{h-m}} \int\limits_{\mathbf{x}\in\mathbb{R}^{h-m}} F_1(\mathbf{x}) \prod_{j=1}^{d} g_j(\xi_j(\Phi(\mathbf{x})) + a_j)\, d\mathbf{x}\Big|, \tag{5.64}$$

and F_1 and a_j satisfy the conclusions of the proposition.

Finally, since $\mathrm{dist}((\Xi, L), V^*_{\mathrm{degen},2}(m,d,h)) \geqslant c$ and $\|\Xi\|_\infty, \|L\|_\infty \leqslant C$, Lemma 5.7.2 tells us that $\Xi\Phi : \mathbb{R}^{h-m} \longrightarrow \mathbb{R}^d$ satisfies $\mathrm{dist}(\Xi\Phi, V_{\mathrm{degen}}(h-m,d)) \gg_{c,C} 1$. (One may consult Definitions 5.3.5 and Definition 5.5.4 for the definitions of $V_{\mathrm{degen}}(h-m,d)$ and $V^*_{\mathrm{degen},2}(m,d,h)$). Thus, by Lemma 5.3.7, there exists some s at most $d-2$ for which $\Xi\Phi$ has $\Omega_{c,C}(1)$-Cauchy-Schwarz complexity at most s.

Writing Ψ for $\Xi\Phi$, the proposition is proved. □

We now proceed to the second proposition, which is a standard Cauchy-Schwarz argument.

Proposition 5.8.3 (Cauchy-Schwarz argument)**.** *Let s, d be natural numbers, with $d \geqslant 3$, and let C be a positive constant. Let $(\psi_1, \cdots, \psi_d) = \Psi : \mathbb{R}^{s+1} \longrightarrow \mathbb{R}^d$ be a linear map, and suppose that $\psi_1(\mathbf{e_k}) = 1$, for all the standard basis vectors $\mathbf{e_k} \in \mathbb{R}^{s+1}$. Suppose that, for all j in the range $2 \leqslant j \leqslant s+1$, there exists some k such that $\psi_j(\mathbf{e_k}) = 0$. Let $N \geqslant 1$ be real, and let $g_1, \cdots, g_d : [-N, N] \longrightarrow [-1, 1]$ be arbitrary measurable functions, and, for each j at most d, let a_j be some real number with $|a_j| \leqslant CN$. Let $F : \mathbb{R}^{s+1} \longrightarrow [0,1]$ be any Lipschitz function, supported on $[-CN, CN]^{s+1}$ with Lipschitz constant $O(1/\sigma_F N)$. Suppose that $\|g_1\|_{U^{s+1}(\mathbb{R})} \leqslant \rho$, for some parameter ρ in the range $0 < \rho \leqslant 1$. Then*

$$\Big|\frac{1}{N^{s+1}} \int\limits_{\mathbf{w}\in\mathbb{R}^{s+1}} \prod_{j=1}^{d} g_j(\psi_j(\mathbf{w}) + a_j) F(\mathbf{w})\, d\mathbf{w}\Big| \ll_C \rho^{-\Omega(1)} \sigma_F^{-1}. \tag{5.65}$$

We stress again that implied constants may depend on the implicit dimensions (so the $\Omega(1)$ term in (5.65) may depend on s).

Proof. This theorem is very similar to the usual Generalised von Neumann Theorem

(see [84, Exercise 1.3.23]), and the proof is very similar too. A few extra technicalities arise from our dealing with the reals rather than with a finite group, but these are easily surmountable.

We begin with some simple reductions. First, we assume that C is large enough in terms of all other $O(1)$ parameters. For notational convenience, we will also allow C to vary form line to line. Next, since $\psi_1(\mathbf{w}) = w_1 + w_2 + \cdots + w_{s+1}$, by shifting w_1 we can assume that $h_1 = 0$ in (5.65). Due to the restricted support of F, we may restrict the integral over $\mathbf{w}$ to $[-CN, CN]^{s+1}$. By Lemma 0.4.4, for any $Y > 2$ there is a function $\mathbf{c}_Y : \mathbb{R}^{s+1} \longrightarrow \mathbb{C}$ satisfying $\|\mathbf{c}\|_\infty \ll 1$ such that we may replace $F(\mathbf{w})$ by

$$\int\limits_{\substack{\boldsymbol{\theta}\in\mathbb{R}^{s+1}\\ \|\boldsymbol{\theta}\|_\infty \leqslant Y}} c_Y(\boldsymbol{\theta}) e(\frac{\boldsymbol{\theta}\cdot\mathbf{w}}{N})\, d\boldsymbol{\theta} + O_C\left(\frac{\log Y}{\sigma_F Y}\right).$$

We will determine a particularly suitable Y later (that will depend on ρ).

This means that

$$\begin{aligned}&\Big|\frac{1}{N^{s+1}} \int\limits_{\mathbf{w}\in\mathbb{R}^{s+1}} \prod_{j=1}^{d} g_j(\psi_j(\mathbf{w}) + a_j) F(\mathbf{w})\, d\mathbf{w}\Big| \\ &\ll \int\limits_{\substack{\boldsymbol{\theta}\in\mathbb{R}^{s+1}\\ \|\boldsymbol{\theta}\|_\infty \leqslant Y}} \Big|\frac{1}{N^{s+1}} \int\limits^{*}_{\mathbf{w}\in\mathbb{R}^{s+1}} e(\frac{\boldsymbol{\theta}}{N}\cdot\mathbf{w})\Big(\prod_{j=1}^{d} g_j(\psi_j(\mathbf{w}) + a_j)\Big)\, d\mathbf{w}\Big|\, d\boldsymbol{\theta} + O_C\left(\frac{\log Y}{\sigma_F Y}\right),\end{aligned} \tag{5.66}$$

where $\int^*$ indicates the limits $\mathbf{w} \in [-CN, CN]^{s+1}$. Fix $\boldsymbol{\theta}$. The inner integral of (5.66) will be our primary focus.

Firstly, we wish to 'absorb' the exponential phases $e(\frac{\boldsymbol{\theta}}{N}\cdot\mathbf{w})$. To do this, we write $e(\frac{\boldsymbol{\theta}}{N}\cdot\mathbf{w})$ as a product of functions $\prod_{k=1}^{s+1} b_k(\mathbf{w})$, where, for each k, the function $b_k : \mathbb{R}^{s+1} \longrightarrow \mathbb{C}$ is bounded and does not depend on the variable w_k. Since $s+1 \geqslant 2$,

this is clearly possible. Therefore we may rewrite the inner integral of (5.66) as

$$\frac{1}{N^{s+1}} \int\limits_{\mathbf{w}\in\mathbb{R}^{s+1}}^{*} g_1(\psi_1(\mathbf{w})) \prod_{k=1}^{s+1} b_k(\mathbf{w})\, d\mathbf{w}, \tag{5.67}$$

where the functions $b_k : \mathbb{R}^{s+1} \longrightarrow \mathbb{C}$ are (possibly different) functions, satisfying $\|b_k\|_\infty \leqslant 1$ for all k, and such that b_k does not depend on the variable w_k.

A brief aside: readers familiar with the arguments of [38, Appendix C] (which motivate the present proof) may note that a different device is used in that paper to absorb the exponential phases. Those authors work in the setting of the finite group $\mathbb{Z}/N\mathbb{Z}$, and there the exponential phases can be absorbed simply by twisting the functions $g_j : \mathbb{Z}/N\mathbb{Z} \longrightarrow [-1, 1]$ by a suitable linear phase function (witness the discussion surrounding expression (C.7) from [38]). The key point there is that, if the linear form $\mathbf{w} \mapsto \boldsymbol{\theta} \cdot \mathbf{w}$ fails to be in the set $\mathrm{span}(\psi_j : 1 \leqslant j \leqslant d)$, then a Fourier expansion of g_j demonstrates that a certain expression, analogous to the inner integral of (5.66), is equal to zero. This clean argument isn't quite so easy to apply here, as the linear phases are not integrable over all of $\mathbb{R}$, which is why we chose a different approach.

Returning to (5.67), recall that $\psi_1(\mathbf{w}) = w_1 + w_2 + \cdots + w_{s+1}$. Therefore, applying the Cauchy-Schwarz inequality in each of the variables w_1 through w_{s+1} in turn, one establishes that the absolute value of expression (5.67) is at most

$$\ll_C \Big(\frac{1}{N^{2s+2}} \int\limits_{\mathbf{w}\in\mathbb{R}^{s+1}}^{*} \int\limits_{\mathbf{z}\in\mathbb{R}^{s+1}}^{*} \prod_{\boldsymbol{\alpha}\in\{0,1\}^{s+1}} g_1\Big(\sum_{\substack{k\leqslant s+1\\ \boldsymbol{\alpha}_k=0}} w_k + \sum_{\substack{k\leqslant s+1\\ \boldsymbol{\alpha}_k=1}} z_k\Big)\, d\mathbf{w}\, d\mathbf{z}\Big)^{\frac{1}{2^{s+1}}}. \tag{5.68}$$

This expression may be immediately related to the real Gowers norm as given in Definition 0.3.3, by the change of variables $m_k := z_k - w_k$, for all k at most $s + 1$,

and $u := w_1 + \cdots + w_{s+1}$. Performing this change of variables shows that(5.68) is

$$\ll \Big(\frac{1}{N^{2s+2}} \int\limits_{(u,\mathbf{m},\mathbf{z_2^{s+1}})\in D} \prod_{\boldsymbol{\alpha}\in\{0,1\}^{s+1}} g_1(u+\boldsymbol{\alpha}\cdot\mathbf{m})\, du\, d\mathbf{m}\, d\mathbf{z_2^{s+1}}\Big)^{\frac{1}{2^{s+1}}}, \tag{5.69}$$

where D is convex domain contained within $[-CN, CN]^{2s+2}$. It remains to replace D by a Cartesian box.

By Lemma 0.4.2 we may write

$$1_D = F_\sigma + O(G_\sigma),$$

for any σ in the range $0 < \sigma < 1/2$, where $F_\sigma, G_\sigma : \mathbb{R}^{2s+2} \longrightarrow [0,1]$ are Lipschitz functions supported on $[-CN, CN]^{2s+2}$, with Lipschitz constant $O_C(1/\sigma N)$, such that $\int_{\mathbf{x}} G_\sigma(\mathbf{x})\, d\mathbf{x} = O_C(\sigma N^{2s+2})$. Then, since $\|g_1\|_\infty \leqslant 1$, we may bound (5.69) above by

$$\Big(\frac{1}{N^{2s+2}} \int\limits^{*}_{u,\mathbf{m},\mathbf{z_2^{s+1}}} F_\sigma(u,\mathbf{m},\mathbf{z_2^{s+1}}) \prod_{\boldsymbol{\alpha}\in\{0,1\}^{s+1}} g_1(u+\boldsymbol{\alpha}\cdot\mathbf{m}) du\, d\mathbf{m}\, d\mathbf{z_2^{s+1}} + O_C(\sigma)\Big)^{\frac{1}{2^{s+1}}}, \tag{5.70}$$

where $\int^*$ now refers to the domain of integration $[-CN, CN]^{2s+2}$.

By applying Lemma 0.4.4 to F_σ, for any $X > 2$ the absolute value of expression (5.70) is

$$\ll_C \Big(\Big(\frac{1}{N^{2s+2}} \int\limits_{\substack{\boldsymbol{\xi}\in\mathbb{R}^{2s+2}\\ \|\boldsymbol{\xi}\|_\infty\leqslant X}} \Big| \int\limits^{*}_{u,\mathbf{m},\mathbf{z_2^{s+1}}} e\Big(\frac{\boldsymbol{\xi}}{N}\cdot(u,\mathbf{m},\mathbf{z_2^{s+1}})\Big) \prod_{\boldsymbol{\alpha}\in\{0,1\}^{s+1}} g_1(u+\boldsymbol{\alpha}\cdot\mathbf{m}) du\, d\mathbf{m}\, d\mathbf{z_2^{s+1}}\Big| d\boldsymbol{\xi}\Big) + O(\sigma) + O\Big(\frac{\log X}{\sigma X}\Big)\Big)^{\frac{1}{2^{s+1}}}. \tag{5.71}$$

Integrating over the variables $z_2, \cdots, z_{s+1}$, and splitting the exponential phase

amongst the different functions, expression (5.71) is

$$\ll_C \Big(\Big(\frac{1}{N^{s+2}} \int\limits_{\substack{\boldsymbol{\xi} \in \mathbb{R}^{2s+2} \\ \|\boldsymbol{\xi}\|_\infty \leqslant X}} \Big| \int\limits_{(u,\mathbf{m}) \in [-CN,CN]^{s+2}} \prod_{\boldsymbol{\alpha} \in \{0,1\}^{s+1}} g_{\boldsymbol{\alpha}}(u + \boldsymbol{\alpha} \cdot \mathbf{m})\, du\, d\mathbf{m} \Big| \, d\boldsymbol{\xi} \Big) + O_C(\sigma) + O_C\Big(\frac{\log X}{\sigma X} \Big) \Big)^{\frac{1}{2^{s+1}}}, \quad (5.72)$$

where each function $g_{\boldsymbol{\alpha}}$ is of the form

$$g_{\boldsymbol{\alpha}}(u) := g_1(u) e(k_{\boldsymbol{\alpha}} u)$$

for some real $k_{\boldsymbol{\alpha}}$. Note that $\|g_{\boldsymbol{\alpha}}\|_{U^{s+1}(\mathbb{R})} = \|g_1\|_{U^{s+1}(\mathbb{R})}$.

Recall that g_1 is supported on $[-2N, 2N]$. Therefore, if $\prod_{\boldsymbol{\alpha} \in \{0,1\}^{s+1}} g_{\boldsymbol{\alpha}}(u + \boldsymbol{\alpha} \cdot \mathbf{m}) \neq 0$ then $(u, \mathbf{m}) \in [-O(N), O(N)]^{s+2}$. So, if C is large enough in terms of s, we may replace the restriction $(u, \mathbf{m}) \in [-CN, CN]^{s+2}$ in (5.72) with the condition $(u, \mathbf{m}) \in \mathbb{R}^{s+2}$, without changing the value of (5.72).

Then, by the Gowers-Cauchy-Schwarz inequality (Proposition 0.3.4) and the triangle inequality, (5.72) is

$$\begin{aligned} &\ll_C (X^{O(1)} \|g_1\|_{U^{s+1}(\mathbb{R})}^{2^{s+1}} + \sigma + \frac{\log X}{\sigma X})^{\frac{1}{2^{s+1}}} \\ &\ll_C (X^{O(1)} \rho^{2^{s+1}} + \sigma + \frac{\log X}{\sigma X})^{\frac{1}{2^{s+1}}} \end{aligned} \quad (5.73)$$

Choosing $X = \rho^{-c_1}$, with c_1 suitably small in terms of s, and $\sigma = \rho^{c_1/2}$, expression (5.73) is $O_C(\rho^{\Omega(1)})$.

Putting this estimate into (5.66), we get a bound on (5.66) of

$$\ll_C Y^{O(1)} \rho^{\Omega(1)} + O(\frac{\log Y}{\sigma_F Y}). \quad (5.74)$$

Picking $Y = \rho^{-c_1}$, with c_1 suitably small in terms of s, we may ensure that (5.74) is $O_C(\rho^{\Omega(1)}\sigma_F^{-1})$, thus proving the proposition. □

With these propositions in hand, Theorem 5.8.1 follows quickly.

Proof of Theorem 5.8.1. Assuming all the hypotheses of Theorem 5.8.1, apply the result of Proposition 5.8.2 to $\widetilde{T}_{F,G}^{L,\Xi,\tilde{\mathbf{r}}}(g_1, \cdots, g_d)$. Thus

$$|\widetilde{T}_{F,G}^{L,\Xi,\tilde{\mathbf{r}}}(g_1, \cdots, g_d)| \ll_{c,C,\varepsilon} \Big|\frac{1}{N^{h-m}} \int\limits_{\mathbf{x}\in\mathbb{R}^{h-m}} F_1(\mathbf{x}) \prod_{j=1}^{d} g_j(\psi_j(\mathbf{x}) + a_j)\, d\mathbf{x}\Big|, \tag{5.75}$$

where $\Psi : \mathbb{R}^{h-m} \longrightarrow \mathbb{R}^d$ has $\Omega_{c,C}(1)$-Cauchy-Schwarz complexity at most s, for some s at most $d-2$, $F_1 : \mathbb{R}^{h-m} \longrightarrow [0,1]$ is a Lipschitz function supported on $[-O_{c,C,\varepsilon}(N), O_{c,C,\varepsilon}(N)]^{h-m}$ with Lipschitz constant $O(1/\sigma_F N)$, and $a_j = O_C(N)$.

We apply Proposition 5.3.8 to Ψ. Therefore, for *any* real numbers $w_1, \cdots, w_{s+1}$,

$$|\widetilde{T}_{F,G}^{L,\Xi,\tilde{\mathbf{r}}}(g_1, \cdots, g_d)| \ll \Big|\frac{1}{N^{h-m}} \int\limits_{\mathbf{x}\in\mathbb{R}^{h-m}} F_1(\mathbf{x} + \sum_{k=1}^{s+1} w_k \mathbf{f_k}) \prod_{j=1}^{d} g_j(\psi_j'(\mathbf{x}, \mathbf{w}) + a_j)\, d\mathbf{x}\Big|, \tag{5.76}$$

where

- for each j at most d, $\psi_j' : \mathbb{R}^{h-m} \times \mathbb{R}^{s+1} \longrightarrow \mathbb{R}$ is a linear form;
- $\psi_1'(\mathbf{0}, \mathbf{w}) = w_1 + \cdots + w_{s+1}$;
- $\mathbf{f_1}, \cdots, \mathbf{f_{s+1}} \in \mathbb{R}^{h-m}$ are some vectors that satisfy $\|\mathbf{f_k}\|_\infty = O_{c,C}(1)$ for each k at most $s+1$;
- the system of forms $(\psi_1', \cdots, \psi_d')$ is in normal form with respect to ψ_1'.

We remark that the right-hand side of expression (5.76) is independent of $\mathbf{w}$, as it was obtained by applying the change of variables $\mathbf{x} \mapsto \mathbf{x} + \sum_{k=1}^{s+1} w_k \mathbf{f_k}$ to expression (5.75).

Now, let $P : \mathbb{R}^{s+1} \longrightarrow [0,1]$ be some Lipschitz function, supported on $[-N,N]^{s+1}$, with Lipschitz constant $O(1/N)$. Also suppose that $P(\mathbf{x}) \equiv 1$ if $\|\mathbf{x}\|_\infty \leqslant N/2$. Integrating over $\mathbf{w}$, we have that $|\widetilde{T}_{F,G}^{L,\Xi,\widetilde{\mathbf{r}}}(g_1,\cdots,g_d)|$ is

$$\ll_{c,C,\varepsilon} \frac{1}{N^{h-m+s+1}} \int\limits_{\mathbf{w}\in\mathbb{R}^{s+1}} P(\mathbf{w}) \Big| \int\limits_{\mathbf{x}\in\mathbb{R}^{h-m}} F_1(\mathbf{x}+\sum_{k=1}^{s+1} w_k\mathbf{f_k}) \prod_{j=1}^{d} g_j(\psi_j'(\mathbf{x},\mathbf{w})+a_j)\, d\mathbf{x}\Big| \, d\mathbf{w}$$

$$\ll_{c,C,\varepsilon} \Big| \frac{1}{N^{h-m+s+1}} \int\limits_{\substack{\mathbf{x}\in\mathbb{R}^{h-m}\\ \mathbf{w}\in\mathbb{R}^{s+1}}} H(\mathbf{x},\mathbf{w}) \prod_{j=1}^{d} g_j(\psi_j'(\mathbf{x},\mathbf{w})+a_j)\, d\mathbf{x}\, d\mathbf{w}\Big|, \tag{5.77}$$

where the function $H : \mathbb{R}^{h-m+s+1} \longrightarrow [0,1]$ is defined by

$$H(\mathbf{x},\mathbf{w}) := F_1(\mathbf{x}+\sum_{k=1}^{s+1} w_k\mathbf{f_k})P(\mathbf{w}).$$

Since the vectors $\mathbf{f_k}$ satisfy $\|\mathbf{f_k}\|_\infty = O_{c,C}(1)$, H is a Lipschitz function supported on $[-O_{c,C,\varepsilon}(N), O_{c,C,\varepsilon}(N)]^{h-m+s+1}$, with Lipschitz constant $O_{c,C}(1/\sigma_F N)$. Notice in (5.77) that we were able to move the absolute value signs outside the integral, as P is positive and the integral over $\mathbf{x}$ is independent of $\mathbf{w}$ (so in particular has constant sign).

Fix $\mathbf{x}$. Then the integral over $\mathbf{w}$ in (5.77) satisfies the hypotheses of Proposition 5.8.3. Applying Proposition 5.8.3 to this integral, and then integrating over $\mathbf{x}$, one derives

$$|\widetilde{T}_{F,G}^{L,\Xi,\widetilde{\mathbf{r}}}(g_1,\cdots,g_d)| \ll_{c,C,\varepsilon} \rho^{\Omega(1)}\sigma_F^{-1}.$$

Theorem 5.8.1 is proved. □

By our long series of reductions, this means that both Theorem 5.5.6 and the main result (Theorem 5.2.10) are proved. □

5.9 Constructions

In this section we prove Theorem 5.2.12, which, we remind the reader, is the partial converse of main result (Theorem 5.2.10). In other words, we show that L being bounded away from $V^*_{\text{degen}}(m,d)$ is a necessary hypotheses for Theorem 5.2.10 to be true.

Proof of Theorem 5.2.12. Recall the hypotheses of Theorem 5.2.12. In particular, we suppose that

$$\liminf_{N\to\infty} \text{dist}(L, V^*_{\text{degen}}(m,d)) = 0,$$

i.e. we assume that $\text{dist}(L, V^*_{\text{degen}}(m,d)) = \omega(N)^{-1}$, for some function $\omega(N)$ such that

$$\limsup_{N\to\infty} \omega(N) = \infty.$$

Let η be a small positive quantity, picked small enough in terms of c and C, and let N be a natural number that is large enough so that $\omega(N) \geqslant \eta^{-1}$ and $\eta N \geqslant \max(1,\varepsilon)$. All implied constants to follow will be independent of η.

Since F is the indicator function of $[1,N]^d$ and G is the indicator function of $[-\varepsilon,\varepsilon]^m$, one has

$$T^L_{F,G}(f_1,\cdots,f_d) = \frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n}\in[N]^d \\ \|L\mathbf{n}\|_\infty \leqslant \varepsilon}} \prod_{j=1}^{d} f_j(n_j).$$

Our aim is to construct functions $f_1,\ldots,f_d : [N] \longrightarrow [-1,1]$ such that

$$\min_j \|f_j\|_{U^{s+1}[N]} \leqslant \rho$$

for some ρ at most 1 and that

$$T^L_{F,G}(f_1, \cdots, f_d) > H(\rho) + E_\rho(N). \tag{5.78}$$

We begin by observing that the condition $\|L\mathbf{n}\|_\infty \leqslant \varepsilon$ implies certain constraints on two of the variables n_i. Indeed, let $L' \in V^*_{\text{degen}}(m,d)$ be such that $\|L - L'\|_\infty = \text{dist}(L, V^*_{\text{degen}}(m,d))$. Write λ'_{ij} for the coefficients of L'. By reordering columns, without loss of generality we may assume that there exist real numbers $\{a_i\}_{i=1}^m$ not all 0 s.t. for all j in the range $3 \leqslant j \leqslant d$ we have

$$\sum_{i=1}^m a_i \lambda'_{ij} = 0, \tag{5.79}$$

and further we may assume that for all i we have $\lambda'_{i1} = \lambda_{i1}$ and $\lambda'_{i2} = \lambda_{i2}$ (else $L' \in V^*_{\text{degen}}(m,d)$ is not one of the closest matrices to L). By reordering rows and rescaling, we may assume that a_1 has maximal absolute value amongst all the a_i, and that $|a_1| = 1$.

Define

$$b_1 := \sum_{i=1}^m a_i \lambda_{i1}, \quad b_2 := \sum_{i=1}^m a_i \lambda_{i2},$$

and let $\mathbf{n} \in [N]^d$ be some solution to $\|L\mathbf{n}\|_\infty \leqslant \varepsilon$. The critical observation is that (5.79), combined with the assumptions on the a_i, implies that

$$|b_1 n_1 + b_2 n_2| \ll \eta N. \tag{5.80}$$

Indeed, for j in the range $3 \leqslant j \leqslant d$ we have

$$\begin{aligned} \left| \sum_{i=1}^m a_i \lambda_{ij} \right| &= \left| \sum_{i=1}^m a_i (\lambda_{ij} - \lambda'_{ij}) \right| \\ &\ll \eta. \end{aligned}$$

Since $\|L\mathbf{n}\|_\infty \leqslant \varepsilon$, we certainly have that

$$\left| b_1 n_1 + b_2 n_2 + \sum_{j=3}^{d} n_j \sum_{i=1}^{m} a_i \lambda_{ij} \right| \ll \varepsilon,$$

and then (5.80) follows by the triangle inequality and the fact that $\eta N \geqslant \varepsilon$.

The constraint (5.80) will turn out to be enough for the proof. We consider various cases, constructing different counterexample functions f_1 and f_2 based on the size and sign of b_1 and b_2. To facilitate this, we let c_1 be a suitably small positive constant, depending on c and C, but independent of η. All constants C_1 and C_2 to follow will be assumed to satisfy $O_{c,C}(1)$.

Case 1: $|b_1|, |b_2| \leqslant c_1$**.**

Under the assumptions of Theorem 5.2.12, this case is actually precluded. Indeed, consider the matrix L'', defined by taking

$$\lambda''_{ij} = \lambda'_{ij}$$

for all pairs $(i.j) \in [m] \times [d]$, except for $(1,1)$ and $(1,2)$. In these cases we let

$$\lambda''_{11} = \lambda'_{11} - \frac{b_1}{a_1}$$
$$\lambda''_{12} = \lambda'_{12} - \frac{b_2}{a_1}.$$

Then

$$\sum_{i=1}^{m} a_i \lambda''_{ij} = 0$$

for all j in the range $1 \leqslant j \leqslant d$. In other words we have shown that $\|L - L''\|_\infty \leqslant \eta + c_1$

for some matrix L'' with rank less than m. Since $\eta + c_1 < c$ (if c_1 is small enough), this implies that $\mathrm{dist}(L, V_{\mathrm{rank}}(m,d)) < c$, which contradicts the assumptions of Theorem 5.2.12. Therefore this case is indeed precluded.

Case 2: b_1, b_2 both of the same sign, and $b_1, b_2 \geqslant c_1$.

In this case, (5.80) implies[8] that $n_1 \leqslant C_1 \eta N$ for some constant C_1. Now, define $f_1 : [N] \longrightarrow [-1,1]$ to be the indicator function of the interval $[\lceil C_1 \eta N \rceil, N] \cap \mathbb{N}$. We then have

$$\begin{aligned} \|f_1 - 1\|_{U^{s+1}[N]} &\ll \Big(\frac{1}{N^{s+2}} \sum_{x, h_1, \cdots, h_{s+1} \ll C_1 \eta N} 1 \Big)^{\frac{1}{2^{s+1}}} \\ &\leqslant C_2 (C_1 \eta)^{\frac{s+2}{2^{s+1}}} \end{aligned}$$

for some constant C_2. However, observe that

$$\begin{aligned} |T^L_{F,G}(f_1 - 1, 1, \cdots, 1)| &= |T^L_{F,G}(f_1, 1, \cdots, 1) - T^L_{F,G}(1, 1, \cdots, 1)| \\ &= |0 - T^L_{F,G}(1, 1, \cdots, 1)| \gg_{c,C,\varepsilon} 1 \end{aligned}$$

by the hypotheses of Theorem 5.2.12. If $T^L_{F,G}(f_1 - 1, 1, \cdots, 1)$ did not satisfy (5.78), then

$$1 \ll_{c,C,\varepsilon} H(\rho) + E_\rho(1),$$

where $\rho := C_2 (C_1 \eta)^{\frac{s+2}{2^{s+1}}}$. Picking η small enough, then N large enough, this inequality cannot possibly hold, and we have a contradiction. So $T^L_{F,G}(f_1 - 1, 1, \cdots, 1)$ satisfies (5.78).

Case 3: b_1, b_2 of opposite signs, and $b_1, b_2 \geqslant c_1$.

This is the most involved case, although the central idea is very simple. The con-

[8] The same conclusion is true for n_2, but this will not be needed.

dition (5.80) confines n_2 to lie within a certain distance of a fixed multiple of n_1. By constructing functions f_1 and f_2 using random choices of blocks of this length, but coupled in such a way that condition (5.80) is very likely to hold, we can guarantee that $T^L_{F,G}(f_1 - p, f_2 - p, 1, \cdots, 1)$ is bounded away from zero, where p is the probability used to choose the random blocks. However, despite the block construction and the coupling, the functions f_1 and f_2 still individually exhibit enough randomness to conclude that $\|f_1 - p\|_{U^{s+1}[N]} = o(1)$, and the same for f_2.

We now fill in the technical details. Relation (5.80) implies that

$$|b_1 n_1 + b_2 n_2| \leqslant C_1 \eta N, \tag{5.81}$$

for some C_1 satisfying $C_1 = O(1)$, and without loss of generality assume that b_1 is positive, b_2 is negative, and $|b_1|$ is at least $|b_2|$. Let C_2 be some parameter, chosen so that $(C_1 C_2 \eta)^{-1}$ is an integer. Such a C_2 will of course depend on η, but in magnitude we may pick $C_2 \asymp 1$. We consider the real interval $[0, N]$ modulo N, and for $x \in [0, N]$ and i in the range $0 \leqslant i \leqslant (C_1 C_2 \eta)^{-1} - 1$ we define the half-open interval modulo N

$$I_i := [x + iC_1 C_2 \eta N, x + (i+1) C_1 C_2 \eta N).$$

This choice guarantees that

$$[0, N] = \bigcup_{i=0}^{(C_1 C_2 \eta)^{-1} - 1} I_i, \tag{5.82}$$

and the union is disjoint. Now, for δ a small constant to be chosen later[9], we define

$$I_i^\delta := [x + (i + \frac{1}{2} - \delta) C_1 C_2 \eta N, x + (i + \frac{1}{2} + \delta) C_1 C_2 \eta N).$$

[9]This δ is unrelated to the notation $\delta = T^L_{F,G}(f_1, \cdots, f_d)$ used in previous sections.

We will use the partition (5.82) to construct a function f_1, using an averaging argument to choose an x so that the I_i^δ intervals capture a positive proportion of the solution density of the linear inequality system. Indeed, for $n_1 \in [N]$ let the weight $u(n_1)$ denote the number of $d-1$-tuples $n_2, \cdots, n_d \leqslant N$ that together with n_1 satisfy the inequality $\|L\mathbf{n}\|_\infty < \varepsilon$. The weight $u(n_1)$ could be zero, of course. Let

$$E_\delta := \cup_i I_i^\delta.$$

Then

$$\begin{aligned}\frac{1}{N}\int_0^N \sum_{n\in[N]} u(n) 1_{E_\delta}(n)\, dx &= \frac{1}{N}\sum_{n\in[N]} u(n) \int_0^N 1_{E_\delta}(n)\, dx \\ &= \sum_{n\in[N]} u(n) 2\delta \\ &= 2\delta N^{d-m} T^L_{F,G}(1, \cdots, 1)\end{aligned}$$

Therefore, by the assumptions of Theorem 5.2.12, we may fix an x such that

$$\sum_{n\in[N]} u(n) 1_{E_\delta}(n) \gg_{c,C} \delta N^{d-m} T^L_{F,G}(1, \cdots, 1). \tag{5.83}$$

Let us finally define the function f_1. Let p be a small positive constant (to be decided later). Fix a value of x such that (5.83) holds. Then we define a random subset $A \subseteq [N]$ by picking all of $I_i \cap \mathbb{N}$ to be members of A, with probability p, or none of $I_i \cap \mathbb{N}$ to be members of A, with probability $1-p$. We then make this same choice for each i in the range $0 \leqslant i \leqslant (C_1C_2\eta)^{-1} - 1$, independently. Observe immediately that for each $n \in [N]$ the probability that $n \in A$ is always p (though these events are not always independent). We let $f_1(n)$ be the indicator function $1_A(n)$.

The function f_2 is defined in terms of f_1. Indeed, let

$$J_i = \frac{b_1}{|b_2|} I_i \cap (0, N],$$

where the dilation is not considered modulo N but rather just as an operator on subsets of $\mathbb{R}$. Since $b_1 \geqslant |b_2|$ we have that these J_i also form a disjoint partition of $[0, N]$. [NB: If $b_1 > |b_2|$ it may be that certain J_i are empty, since the dilate of the corresponding I_i may land entirely outside $[0, N]$.] Then let B be the subset of $[N]$ defined so that for each i with J_i non-empty we have $J_i \cap \mathbb{N} \subseteq B$ if and only if $I_i \cap \mathbb{N} \subseteq A$. Note again that for each individual $n \in [N]$ the probability that $n \in B$ is always p. We let $f_2(n)$ be the indicator function $1_B(n)$.

Our first claim is that, if p is small enough in terms of δ,

$$|\mathbb{E}T^L_{F,G}(f_1, f_2, 1 \cdots, 1) - T^L_{F,G}(p, p, 1 \cdots, 1)| \gg_{c,C,\varepsilon} \delta^2. \tag{5.84}$$

Indeed, suppose that I_i is included in the set A, and suppose that $n_1 \in I_i^\delta$. If $n_2 \in [N]$ satisfies $|\frac{b_1}{|b_2|} n_1 - n_2| \leqslant \frac{1}{b_2} C_1 \eta N$ and if δ is small enough in terms of b_1 and b_2, then[10] $n_2 \in J_i$. Thus, by the observation (5.81), $n_2 \in B$, for every integer n_2 that is the second coordinate of a solution vector[11] $\mathbf{n}$ for which the first coordinate is n_1. Therefore

$$\begin{aligned}\mathbb{E}T^L_{F,G}(f_1, f_2, 1, \cdots, 1) &= \frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n} \in [N]^d \\ \|L\mathbf{n}\|_\infty \leqslant \varepsilon}} \mathbb{P}(n_1 \in A \wedge n_2 \in B) \\ &\geqslant \frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n} \in [N]^d \\ \|L\mathbf{n}\|_\infty \leqslant \varepsilon}} \mathbb{P}(n_1 \in A \wedge n_1 \in I_i^\delta \text{ for some } i \wedge n_2 \in B)\end{aligned}$$

[10] This fact is the reason why we introduced the parameter δ.
[11] i.e a vector $\mathbf{n}$ such that $\|L\mathbf{n}\|_\infty \leqslant \varepsilon$.

$$\begin{aligned}
&\geqslant \frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n}\in[N]^d \\ \|L\mathbf{n}\|_\infty \leqslant \varepsilon}} \mathbb{P}(n_1 \in A \wedge n_1 \in I_i^\delta \text{ for some } i)\\
&= \frac{1}{N^{d-m}} \sum_{n_1\in[N]} u(n_1) p 1_{E_\delta}(n_1)\\
&\geqslant 2\delta p T^L_{F,G}(1,\cdots,1),
\end{aligned}$$

where the final line follows from (5.83). On the other hand $T^L_{F,G}(p,p,1,\cdots,1) = p^2 T^L_{F,G}(1,\cdots,1)$, and hence

$$\mathbb{E}T^L_{F,G}(f_1,f_2,1\cdots,1) - T^L_{F,G}(p,p,1\cdots,1) \geqslant (2\delta p - p^2) T^L_{F,G}(1,\cdots,1). \tag{5.85}$$

Picking p small enough in terms of δ, and using the assumption that $T^L_{F,G}(1,\cdots,1) = \Omega_{c,C,\varepsilon}(1)$, this proves the relation (5.84).

Our second claim is that

$$\mathbb{E}\|f_1 - p\|_{U^{s+1}[N]}, \mathbb{E}\|f_2 - p\|_{U^{s+1}[N]} \ll \eta^{\frac{1}{2^{s+1}}}. \tag{5.86}$$

We first consider f_1. Then

$$\mathbb{E}\|f_1 - p\|^{2^{s+1}}_{U^{s+1}[N]} \ll \frac{1}{N^{s+2}} \sum_{(x,\mathbf{h})\in\mathbb{Z}^{s+2}} \mathbb{E}\Big(\prod_{\boldsymbol{\omega}\in\{0,1\}^{s+1}} (f_1 - p1_{[N]})(x+\mathbf{h}\cdot\boldsymbol{\omega})\Big).$$

Observe that for fixed $(x,\mathbf{h})$ the random variables $(f_1 - p1_{[N]})(x+\mathbf{h}\cdot\boldsymbol{\omega})$ each have mean zero and, unless some two of the expressions $x+\mathbf{h}\cdot\boldsymbol{\omega}$ lie in the same block I_i, these random variables are independent. Hence, apart from those exceptional cases, we may factor the expectation and conclude that

$$\mathbb{E}\Big(\prod_{\boldsymbol{\omega}\in\{0,1\}^{s+1}} (f_1 - p1_{[N]})(x+\mathbf{h}\cdot\boldsymbol{\omega})\Big) = \prod_{\boldsymbol{\omega}\in\{0,1\}^{s+1}} \mathbb{E}((f_1 - p1_{[N]})(x+\mathbf{h}\cdot\boldsymbol{\omega})) = 0.$$

Therefore,

$$\begin{aligned}\mathbb{E}\|f_1 - p\|_{U^{s+1}[N]}^{2^{s+1}} &\ll \frac{1}{N^{s+2}} \sum_{(x,\mathbf{h})\in[-N,N]^{s+2}} 1_R(\mathbf{h}) \\ &\ll \eta,\end{aligned}$$

where

$$R = \{\mathbf{h} : |\mathbf{h}\cdot(\boldsymbol{\omega_1} - \boldsymbol{\omega_2})| \leqslant C_1C_2\eta N \text{ for some } \boldsymbol{\omega_1}, \boldsymbol{\omega_2} \in \{0,1\}^{s+1},\ \boldsymbol{\omega_1} \neq \boldsymbol{\omega_2}\}.$$

Thus by Jensen's inequality we have

$$\mathbb{E}\|f_1 - p\|_{U^{s+1}[N]} \ll \eta^{\frac{1}{2^{s+1}}}, \tag{5.87}$$

as claimed in (5.86).

The calculation for f_2 is essentially identical, noting that the length of the blocks J_i is also $O(\eta N)$.

It is possible that one could finish the argument here by considering a second moment, and choosing some explicit f_1 and f_2. To avoid calculating a second moment, we argue as follows. Suppose for contradiction that there were no functions $f_1, \cdots, f_d$ that satisfied (5.78). Then, by (5.84), if we pick p to be small enough in terms of δ we have

$$\begin{aligned}\delta^2 &\ll_{c,C,\varepsilon} |\mathbb{E}T_{F,G}^L(f_1, f_2, 1, \cdots, 1) - T_{F,G}^L(p, p, 1\cdots, 1)| \\ &\ll |\mathbb{E}T_{F,G}^L(f_1 - p, f_2, 1, \cdots, 1)| + |\mathbb{E}T_{F,G}^L(p, f_2 - p, 1, \cdots, 1)| \\ &\ll \mathbb{E}(H(\rho_1) + E_{\rho_1}(N)) + \mathbb{E}(H(\rho_2) + E_{\rho_2}(N)),\end{aligned} \tag{5.88}$$

where ρ_1 (resp. ρ_2) is any chosen upper-bound on $\|f_1 - p\|_{U^{s+1}[N]}$ (resp. $\|f_2 -$

$p\|_{U^{s+1}[N]}$). Note that the values ρ_i may be random variables themselves.

We claim that the random variables ρ_1 and ρ_2 may be chosen so that the right-hand side of (5.88) is $\kappa(\eta) + o_\eta(1)$. To prove this, we make two observations. Note first that by Markov's inequality

$$\mathbb{P}(\|f_1 - p\|_{U^{s+1}[N]} \geqslant \eta^{\frac{1}{2^{s+2}}}) \ll \eta^{\frac{1}{2^{s+2}}}$$

We choose the (random) upper-bound ρ_1 satisfying

$$\rho_1 = \begin{cases} 1 & \text{if } \|f_1 - p\|_{U^{s+1}[N]} \geqslant \eta^{\frac{1}{2^{s+2}}} \\ \eta^{\frac{1}{2^{s+2}}} & \text{otherwise .} \end{cases}$$

Secondly, we may upper-bound H by a concave envelope, so without loss of generality we may assume that H is concave.

Then by Jensen's inequality,

$$\begin{aligned} \mathbb{E}(H(\rho_1) + E_{\rho_1}(N)) &\ll H(\mathbb{E}\rho_1) + \mathbb{E}(E_{\rho_1}(1)) \\ &\ll \kappa(\eta^{\frac{1}{2^{s+2}}}) + o_\eta(1) \\ &\ll \kappa(\eta) + o_\eta(1). \end{aligned} \tag{5.89}$$

We do the same manipulation for f_2. Combining (5.89) with (5.88) we conclude that

$$\delta^2 \ll_{c,C,\varepsilon} \kappa(\eta) + o_\eta(1). \tag{5.90}$$

The only condition on δ occurred in the proof of (5.84), in which we assumed that δ was small enough in terms of b_1 and b_2. Therefore there exists a suitable δ that satisfies $\delta = \Omega_{c,C}(1)$. Picking such a δ, and then picking η small enough and N large enough, (5.90) is a contradiction. So there must be some functions $f_1, \cdots, f_d$ that

satisfy (5.78).

Case 4: Exactly one of b_1, b_2 satisfies $b_i \geqslant c_1$.

Without loss of generality we may assume that $b_1 \geqslant c_1$. But then, as in Case 2, (5.80) implies that $n_1 \leqslant C_1 \eta N$ for some constant C_1. The same construction as in Case 2 then applies.

We have covered all cases, and thus have concluded the proof of Theorem 5.2.12. □

5.10 Rank matrix and normal form: proofs

In this section we prove the two quantitative statements from section 5.3, namely Propositions 5.3.1 and 5.3.8.

We begin with a simple proposition concerning points bounded away from algebraic varieties.

Proposition 5.10.1. *Let n be a natural number, and let $I \trianglelefteq \mathbb{R}[X_1, \cdots, X_n]$ be an ideal with generators $q_1, \cdots, q_l$. Let $V(I) \subset \mathbb{R}^n$ denote the affine variety generated by I. Suppose that $\mathbf{x} \in \mathbb{R}^n$ is a point with $\|\mathbf{x}\|_\infty \leqslant C$ and with $\mathrm{dist}(\mathbf{x}, V(I)) \geqslant c$, for some absolute positive constants c and C. Then, there is some q_j such that $|q_j(\mathbf{x})| = \Omega_{c,C,I}(1)$.*

Proof. This is nothing more than the Heine-Borel theorem. To spell it out, suppose for contradiction that, for all positive ε, there exists an $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_\infty \leqslant C$ and $\mathrm{dist}(\mathbf{x}, V(I)) \geqslant c$, but with $|q_j(\mathbf{x})| < \varepsilon$ for every j. Taking a sequence of ε tending to 0, we get a corresponding sequence of $\mathbf{x}_\varepsilon$. Since all $\mathbf{x}_\varepsilon$ lie in a compact set, there exists a convergent subsequence tending to some limit point $\mathbf{x}$. But then

$\operatorname{dist}(\mathbf{x}, V(I)) \geqslant c$ (by the continuity of the dist function and the fact that $V(I)$ is closed in the Euclidean topology). Yet $q_j(\mathbf{x}) = 0$ for every j, and hence $\mathbf{x} \in V(I)$, which is a contradiction. □

From Proposition 5.10.1 it is easy to deduce the existence of rank matrices.

***Proof of* Proposition 5.3.1**. Suppose that $d \geqslant m$. Let k be equal to $\binom{d}{m}$, and identify $\mathbb{R}^{md}$ with the space of m-by-d real matrices. Then let $q_1, \cdots, q_k$ be the k polynomials on $\mathbb{R}^{md}$ that are given by the k determinants of m-by-m submatrices. Let $I \trianglelefteq \mathbb{R}[X_1, \cdots, X_{md}]$ be the ideal generated by the polynomials q_i. One then sees that V_{rank} is exactly the variety generated by I. This is since row rank equals column rank, and linear independence of columns in a square matrix can be detected by the determinant.

Since we assume that $\|L\|_\infty \leqslant C$ and $\operatorname{dist}(L, V_{\text{rank}}) \geqslant c$ we can fruitfully apply Proposition 5.10.1 to deduce that there exist some positive δ, depending only on c and C, and some natural number j such that $|q_j(L)| \geqslant \delta$. The matrix M whose determinant corresponds to the polynomial q_j is exactly the claimed rank matrix.

This settles the first part of Proposition 5.3.1. The second part then follows immediate by the construction of M^{-1} as the adjugate matrix of M divided by $\det M$.

The third part, namely the statement about linear combinations of rows, follows quickly from the others. Indeed, without loss of generality, assume that the rank matrix M is realised by columns 1 through m. Then, the fact that the rows of L are linearly independent means that there are unique real numbers a_i such that $\sum_{i=1}^{m} a_i \lambda_{ij} = v_j$ for all j in the range $1 \leqslant j \leqslant d$. (Recall that $(\lambda_{ij})_{i \leqslant m, j \leqslant d}$ denotes the coefficients of L). Restricting to j in the range $1 \leqslant j \leqslant m$, we observe that the a_i are

forced to satisfy

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = (M^T)^{-1} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}.$$

Since $\|(M^{-1})^T\|_\infty = \|M^{-1}\|_\infty = O_{c,C}(1)$, we conclude that $a_i = O_{c,C,C_1}(1)$ for all i.

The final part of the proposition is to show that if $\mathrm{dist}(L, V^{\mathrm{global}}_{\mathrm{rank}}(m,d)) \geqslant c$ then, for each j, there exists a rank matrix of L that doesn't include the j^{th} column. But this statement follows immediately from the above, after having deleted the j^{th} column. □

We now consider the quantitative normal form algorithm, and prove Proposition 5.3.8. We remind the reader that, in the proof, the implied constants may depend on the dimensions of the underlying spaces, namely m and n. For the definition of the variety $V_{\mathcal{P}_i}$, which consists of all systems of linear forms for which the partition $\mathcal{P}_i$ is not 'suitable', the reader may consult Definition 5.3.5.

***Proof of* Proposition 5.3.8**. Fix i, and let $\mathcal{P}_i$ be a partition of $[m] \setminus \{i\}$ such that $\mathrm{dist}(\Psi, V_{\mathcal{P}_i}) \geqslant c_1$ (such a $\mathcal{P}_i$ exists by the definition of c_1-Cauchy-Schwarz complexity, i.e. by Definition 5.3.6). The partition $\mathcal{P}_i$ has $s_i + 1$ parts, for some s_i at most s, but it is clear from Definition 5.3.6 (and the definition of the degeneracy varieties themselves) that we may, without loss of generality, further subdivide the partition and assume that the partition $\mathcal{P}_i$ has exactly $s+1$ parts. Call the parts $\mathcal{C}_1$ through $\mathcal{C}_{s+1}$.

Via Gaussian elimination, for each $k \in [s+1]$ we may find a vector $\mathbf{f_k} \in \mathbb{R}^n$ that witnesses the fact that $\mathrm{dist}(\Psi, V_{\mathcal{P}_i}) > 0$, i.e. for which $\psi_i(\mathbf{f_k}) = 1$ but $\psi_j(\mathbf{f_k}) = 0$ for all $j \in \mathcal{C}_k$. (If given a free choice for one of the coordinates of $\mathbf{f_k}$, we set it to be 0). We claim further that Gaussian elimination may be applied in such a way so that the

form

$$\Psi'(\mathbf{u}, w_1, \cdots, w_{s+1}) := \Psi(\mathbf{u} + w_1\mathbf{f_1} + \cdots + w_{s+1}\mathbf{f_{s+1}})$$

satisfies the conclusions of the proposition.

Indeed, if $\Psi' = (\psi'_1, \cdots, \psi'_m)$, the form $\psi'_i(\mathbf{u}, w_1, \cdots, w_{s+1})$ is the only one that uses all of the w_k variables. Furthermore, $\psi'_i(\mathbf{0}, \mathbf{w}) = w_1 + \cdots + w_{s+1}$. Also, $n' = n + s + 1$, which is at most $n + m - 1$.

So Proposition 5.3.8 is proved if we can find such $\mathbf{f_k}$ satisfying $\|\mathbf{f_k}\|_\infty \leqslant O_{c_1,C_1}(1)$. This is an issue of some subtlety, as, even having assumed the bound $\mathrm{dist}(\Psi, V_{\mathrm{degen}}) \geqslant c_1$, there could be unbounded $\mathbf{f_k}$ that satisfy $\psi_i(\mathbf{f_k}) = 1$ and $\psi_j(\mathbf{f_k}) = 0$ for all $j \in \mathcal{C}_k$.

Consider a fixed k, and fix some choice of implementation of the Gaussian elimination algorithm from above. The coordinates of the claimed solution vector $\mathbf{f_k}$ are the evaluations of certain rational functions taken at the coefficients of Ψ. [We now identify Ψ with the coordinate vector in $\mathbb{R}^{mn}$ of its coefficients.] It could be that Ψ is a pole of some of these functions, although we know that there is at least one implementation of the algorithm in which it is not.

Let Γ be the set of possible implementations of Gaussian elimination. The size $|\Gamma|$ is essentially $(1 + |\mathcal{C}_k|)!$, but for us it will be enough that $|\Gamma| = O(1)$. Now, for each $\gamma \in \Gamma$, let the rational functions

$$\frac{p_{\gamma,1}(\Psi)}{q_{\gamma,1}(\Psi)}, \cdots, \frac{p_{\gamma,n}(\Psi)}{q_{\gamma,n}(\Psi)}$$

be the n rational functions defining the claimed coefficients of $\mathbf{f_k}$. One may assume without loss of generality that, for all j, $p_{\gamma,j}, q_{\gamma,j} \in \mathbb{Z}[X_1, \cdots, X_n]$ are co-prime polynomials, with coefficients of size $O(1)$. Now let

$$Q_\gamma := \prod_{j \leqslant n} q_{\gamma,j}.$$

We claim that $V(I) \subseteq V_{\mathcal{P}_i}$, where I is the ideal generated by the set of polynomials $\{Q_\gamma : \gamma \in \Gamma\}$ and $V(I)$ is the affine variety generated by I. (See Definition 5.3.5 for the definition of $V_{\mathcal{P}_i}$.) This claim may be proved in one line: indeed, if $Q_\gamma(\Psi) = 0$ for all $\gamma \in \Gamma$ then there is no Gaussian elimination implementation that finds a solution $\mathbf{f_k}$, and this in turn implies that $\mathcal{P}_i$ is not suitable[12] for Ψ. Since $V(I) \subseteq V_{\mathcal{P}_i}$, the assumptions of Proposition 5.3.8 imply that $\mathrm{dist}(\Psi, V(I)) \geqslant c_1$.

Applying Proposition 5.10.1 to the ideal I, we conclude that there is some $\gamma \in \Gamma$ such that $|Q_\gamma(\Psi)| = \Omega_{c_1,C_1}(1)$. In particular, we conclude that the solution vector $\mathbf{f_k}$ obtained by the implementation γ has coefficients that are $O_{c_1,C_1}(1)$. This concludes the proof of Proposition 5.3.8. □

Let us illustrate the above proof with an instructive example. Consider $n = 3$, $m = 2$, $i = 1$, and denote

$$\Psi = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \end{pmatrix}.$$

Then the partition $\mathcal{P}_i$ consists of the singleton $\{2\}$, and implementing Gaussian elimination a certain way we have

$$\mathbf{f_1} = \begin{pmatrix} s_{22}/(s_{11}s_{22} - s_{12}s_{21}) \\ -s_{21}/(s_{11}s_{22} - s_{12}s_{21}) \\ 0 \end{pmatrix}$$

as a solution, in the case where $s_{11}s_{22} - s_{12}s_{21}$ is non-zero. Of course if $s_{11}s_{23} - s_{13}s_{21}$

[12] See Definition 5.3.5 for the term 'suitable'.

is non-zero too, we have another solution

$$\mathbf{f_1} = \begin{pmatrix} s_{23}/(s_{11}s_{23} - s_{13}s_{21}) \\ 0 \\ -s_{21}/(s_{11}s_{23} - s_{13}s_{21}) \end{pmatrix}.$$

So, if one applied Gaussian elimination idly, one might end up with either of these two solutions. Unfortunately it could be the case that $\mathrm{dist}(\Psi, V_{\mathcal{P}_i}) \geqslant c_1$ whilst one of these determinants, $s_{11}s_{22} - s_{12}s_{21}$ say, was non-zero yet $o(1)$ (as the unseen variable N, on which Ψ will ultimately depend, tends to infinity). In this instance, applying the first implementation of the algorithm would not give a desirable solution vector $\mathbf{f_1}$. Therefore we need some subtlety to ensure that we pick the correct implementation.

It is worth including a brief discussion on why these quantitative subtleties do not arise in the setting of [38]. Indeed, assume that Ψ has rational coefficients of naive height at most C_1. Proceed with all the linear algebra from the previous proof, over $\mathbb{Q}$, and choose *any* implementation of Gaussian elimination that is valid for Ψ. As previously remarked, the coordinates of the solution vector $\mathbf{f_k}$ are the evaluations of certain rational functions $\frac{p_j}{q_j}$ with $p_j, q_j \in \mathbb{Z}[X_1, \cdots, X_n]$ co-prime, taken at the coefficients of Ψ. [We now once more identify Ψ with the coordinate vector in $\mathbb{R}^{mn}$ of its coefficients.] By the construction of the algorithm,

$$\Psi \notin \bigcup_{j=1}^{n} \{\Psi' : q_j(\Psi') = 0\}.$$

Yet now we observe a key distinction from the situation over the reals, namely that there are only $O_{C_1}(1)$ many possible choices of Ψ (since Ψ has rational coordinates of bounded height). Therefore, with the above information, we can immediately

conclude that

$$\mathrm{dist}(\Psi, \bigcup_{j=1}^{n} \{\Psi' : q_j(\Psi') = 0\}) \gg_{C_1} 1,$$

without needing to assume this as an extra hypothesis.

5.11 Additional linear algebra

In this section, we collect together the assortment of standard linear algebra lemmas that we used at various points throughout the chapter. We also give the linear algebra argument used to construct the matrix P during the proof of Lemma 5.5.10.

This first lemma demonstrates the intuitive fact, that if $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ is a linear map then $L : (\ker L)^{\perp} \longrightarrow \mathbb{R}^m$ has bounded inverse.

Lemma 5.11.1. *Let m, d be natural numbers, with $d \geqslant m+1$, and let c, C, l be positive constants. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and suppose $\|L\|_{\infty} \leqslant C$ and $\mathrm{dist}(L, V_{\mathrm{rank}}(m, d)) \geqslant c$. Let K denote $\ker L$. Let R be a convex set contained in $[-l, l]^m$. Then, if $\mathbf{v} \in K^{\perp}$, $L\mathbf{v} \in R$ only when $\mathbf{v} \in R'$, where R' is some convex region that satisfies $R' \subseteq [-O_{c,C}(l), O_{c,C}(l)]^d$.*

Proof. Writing L as a m-by-d matrix with respect to the standard bases, let $\boldsymbol{\lambda_i} \in \mathbb{R}^d$ denote the column vector such that $\boldsymbol{\lambda_i}^T$ is the i^{th} row of L. Since $\mathrm{dist}(L, V_{\mathrm{rank}}(m, d)) \geqslant c$, the vectors $\boldsymbol{\lambda_i}$ are linearly independent. Moreover, we may extend the set $\{\boldsymbol{\lambda_i} : i \leqslant m\}$ by orthogonal vectors of unit length to form a basis $\{\boldsymbol{\lambda_i} : i \leqslant d\}$ for $\mathbb{R}^d$.

We claim that for all $k \in [d]$ we have

$$\sum_{i=1}^{d} a_{ki} \boldsymbol{\lambda_i} = \mathbf{e_k},$$

for some coefficients a_{ki} satisfying $|a_{ki}| = O_{c,C}(1)$, where $\mathbf{e_k} \in \mathbb{R}^d$ is the k^{th} standard

basis vector. Indeed, fix k, and note that $\mathbf{e_k} = \mathbf{x_k} + \mathbf{y_k}$, where $\mathbf{x_k} \in \text{span}(\boldsymbol{\lambda_i} : i \leqslant m)$ and $\mathbf{y_k} \in \text{span}(\boldsymbol{\lambda_i} : m+1 \leqslant i \leqslant d)$. The vectors $\mathbf{x_k}$ and $\mathbf{y_k}$ are orthogonal by construction, so in particular $\|\mathbf{x_k}\|_2^2 + \|\mathbf{y_k}\|_2^2 = 1$, and hence $\|\mathbf{x_k}\|_\infty, \|\mathbf{y_k}\|_\infty \ll 1$. By the third part of Proposition 5.3.1 applied to $\mathbf{x_k}$ we get $|a_{ki}| = O_{c,C}(1)$ when $i \leqslant m$, and the orthonormality of $\{\boldsymbol{\lambda_i} : m+1 \leqslant i \leqslant d\}$ implies that $|a_{ki}| = O(1)$ when i is in the range $m+1 \leqslant i \leqslant d$.

Now notice that $\text{span}(\boldsymbol{\lambda_i} : m+1 \leqslant i \leqslant d)$ is exactly equal to K. Let $\mathbf{v} \in K^\perp$, and suppose $L\mathbf{v} \in R$. Letting L' be the d-by-d matrix whose rows are $\boldsymbol{\lambda_i}^T$, we have that $L'\mathbf{v} = \mathbf{w}$ for some vector $\mathbf{w}$ satisfying $\|\mathbf{w}\|_\infty \ll l$. Pre-multiplying by the matrix $A = (a_{ki})$, we immediately get $\mathbf{v} = A\mathbf{w}$, and hence $\|\mathbf{v}\|_\infty = O_{c,C}(l)$. The region $R' := (L^{-1}R) \cap K^\perp$ is therefore bounded. R' is clearly convex, and so the proposition is proved. □

This next lemma concerns vectors, with integer coordinates, that lie near to a subspace.

Lemma 5.11.2. *Let h, d be natural numbers, with $h \leqslant d$, and let C, η be positive reals. Let $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ be an injective linear map, with $\|\Xi\|_\infty \leqslant C$. Suppose further that $\Xi(\mathbb{Z}^h) = \mathbb{Z}^d \cap \Xi(\mathbb{R}^h)$. Let $\mathbf{n}, \widetilde{\mathbf{r}} \in \mathbb{Z}^d$. Suppose that*

$$\text{dist}(\mathbf{n}, \Xi(\mathbb{R}^h) + \widetilde{\mathbf{r}}) \leqslant \eta. \tag{5.91}$$

Then, if η is small enough in terms of C, h and d, $\mathbf{n} = \Xi(\mathbf{m}) + \widetilde{\mathbf{r}}$, for some unique $\mathbf{m} \in \mathbb{Z}^h$.

Proof. By replacing $\mathbf{n}$ with $\mathbf{n} - \widetilde{\mathbf{r}}$, we can assume without loss of generality that $\widetilde{\mathbf{r}} = \mathbf{0}$. It will also be enough to show that $\mathbf{n} \in \Xi(\mathbb{R}^h)$, as the injectivity of Ξ and the assumption that $\Xi(\mathbb{Z}^h) = \mathbb{Z}^d \cap \Xi(\mathbb{R}^h)$ immediately go on to imply the existence of a unique $\mathbf{m}$.

Suppose for contradiction then that $\mathbf{n} \notin \Xi(\mathbb{R}^h)$. In matrix form, Ξ is a d-by-h matrix with linearly independent columns, all of whose coefficients are integers with absolute value at most C. We can extend this matrix to a d-by-d matrix $\widetilde{\Xi}$, with linearly independent columns, all of whose coefficients are integers with absolute value at most C. Then $(\widetilde{\Xi})^{-1}$ is a d-by-d matrix with rational coefficients of naive height at most $C^{O(1)}$, and $(\widetilde{\Xi})^{-1}(\Xi(\mathbb{R}^h)) = \mathbb{R}^h \times \{0\}^{d-h}$.

Since $\mathbf{n} \notin \Xi(\mathbb{R}^h)$, we have $(\widetilde{\Xi})^{-1}(\mathbf{n}) \notin \mathbb{R}^h \times \{0\}^{d-h}$. But $(\widetilde{\Xi})^{-1}(\mathbf{n}) \in \frac{1}{K}\mathbb{Z}^d$, for some natural number K satisfying $K = O(C^{O(1)})$. Therefore

$$\operatorname{dist}((\widetilde{\Xi})^{-1}(\mathbf{n}), (\widetilde{\Xi})^{-1}(\Xi(\mathbb{R}^h))) \gg C^{-O(1)}.$$

Applying $\widetilde{\Xi}$, we conclude that

$$\operatorname{dist}(\mathbf{n}, \Xi(\mathbb{R}^h)) \gg C^{-O(1)},$$

which is a contradiction to (5.91) if η is small enough. □

The construction of the matrix $\widetilde{\Xi}$ in the above proof also has an even more basic consequence, namely that $\Xi^{-1} : \operatorname{im}\Xi \longrightarrow \mathbb{R}^h$ is bounded.

Lemma 5.11.3. *Let h, d be natural numbers, with $h \leqslant d$, and let C, η be positive reals. Suppose that $\Xi : \mathbb{R}^h \longrightarrow \mathbb{R}^d$ is an injective linear map, with $\|\Xi\|_\infty \leqslant C$. Suppose further that $\Xi(\mathbb{Z}^h) \subseteq \mathbb{Z}^d \cap \Xi(\mathbb{R}^h)$. Then if $\|\Xi(\mathbf{y})\|_\infty \leqslant \eta$, we have $\|\mathbf{y}\|_\infty \ll C^{-O(1)}\eta$.*

Proof. Construct the matrix $\widetilde{\Xi}$ as in the previous proof. Then $\|(\widetilde{\Xi})^{-1}(\Xi(\mathbf{y}))\|_\infty \ll C^{O(1)}\eta$, by the bound on the size of the coefficients of $\widetilde{\Xi}$. But $(\widetilde{\Xi})^{-1}(\Xi(\mathbf{y})) \in \mathbb{R}^d$ is nothing more than the vector $\mathbf{y} \in \mathbb{R}^h$ extended by zeros. So $\|\mathbf{y}\|_\infty \ll C^{O(1)}\eta$ as claimed. □

Finally, we give the linear algebra argument used to construct the matrix P during the proof of Lemma 5.5.10.

Lemma 5.11.4. *Let m, d be natural numbers, with $d \geqslant m+1$. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map with rational dimension u, and let $\Theta : \mathbb{R}^m \longrightarrow \mathbb{R}^u$ be a rational map for L. Suppose that $\|L\|_\infty \leqslant C$ and $\|\Theta\|_\infty \leqslant C$. Equating L with its matrix, suppose that the first m columns of L form the identity matrix. Let $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\}$ be a basis for the lattice $\Theta L(\mathbb{Z}^d)$ that satisfies $\|\mathbf{a_i}\|_\infty = O_C(1)$ for every i. Let $\mathbf{x_1}, \cdots, \mathbf{x_u} \in \mathbb{Z}^d$ be vectors such that, for every i, $\Theta L(\mathbf{x_i}) = \mathbf{a_i}$ and $\|\mathbf{x_i}\|_\infty = O_C(1)$. Then*

$$\mathbb{R}^m = \operatorname{span}(L\mathbf{x_i} : i \leqslant u) \oplus \ker \Theta \tag{5.92}$$

and there is an invertible linear map $P : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ such that

$$P((\operatorname{span}(L\mathbf{x_i} : i \leqslant u))) = \mathbb{R}^u \times \{0\}^{m-u},$$
$$P(\ker \Theta) = \{0\}^u \times \mathbb{R}^{m-u},$$

and both $\|P\|_\infty = O_C(1)$ and $\|P^{-1}\|_\infty = O_C(1)$.

Note that both $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\}$ and $\mathbf{x_1}, \cdots, \mathbf{x_u} \in \mathbb{Z}^d$ exist by applying Lemma 5.5.7 to the map $S := \Theta L$.

Proof. The expression (5.92) is immediate from the definitions, so it remains to construct P. We may assume, since the first m columns of L form the identity matrix, that Θ has integer coefficients.

As $\|\Theta\|_\infty = O_C(1)$, we may pick a basis $\{\mathbf{y_1}, \cdots, \mathbf{y_{m-u}}\}$ for $\ker \Theta$ in which $\mathbf{y_j} \in \mathbb{Z}^m$ and $\|\mathbf{y_j}\|_\infty = O_C(1)$ for all j. Let $\mathbf{b_1}, \cdots, \mathbf{b_m}$ denote the standard basis of $\mathbb{R}^m$, and define P by letting

$$\begin{aligned} P(L\mathbf{x_i}) &:= \mathbf{b_i}, & 1 \leqslant i \leqslant u \\ P(\mathbf{y_j}) &:= \mathbf{b_{j+u}}, & 1 \leqslant j \leqslant m-u, \end{aligned} \tag{5.93}$$

and then extending linearly to all of $\mathbb{R}^m$. Clearly $P((\text{span}(L\mathbf{x_i} : i \leqslant u))) = \mathbb{R}^u \times \{0\}^{m-u}$ and $P(\ker \Theta) = \{0\}^u \times \mathbb{R}^{m-u}$. It is also immediate that $\|P^{-1}\|_\infty = O_C(1)$, since $\|L\mathbf{x_i}\|_\infty = O_C(1)$ and $\|\mathbf{y_j}\|_\infty = O_C(1)$ for all i and j. It remains to bound $\|P\|_\infty$. If $L\mathbf{x_i}$ were all vectors with integer coordinates then this bound would be immediate as well, as then P^{-1} would have integer coordinates and hence $|\det P^{-1}| \geqslant 1$. As it is, we have to proceed more slowly.

To this end, for a standard basis vector $\mathbf{b_k}$ write

$$\mathbf{b_k} = \sum_{i=1}^{u} \lambda_i L\mathbf{x_i} + \sum_{j=1}^{d-u} \mu_j \mathbf{y_j}.$$

It will be enough to show that $|\lambda_i|, |\mu_j| = O_C(1)$ for all i and j. First note that, since the first m columns of L form the identity, $\mathbf{b_k} \in L(\mathbb{Z}^d)$. Also $\Theta(\mathbf{b_k}) = \sum_{i=1}^{u} \lambda_i \mathbf{a_i}$. So $\mathbf{a} := \sum_{i=1}^{u} \lambda_i \mathbf{a_i}$ is an element of $\Theta L(\mathbb{Z}^d)$ that satisfies $\|\mathbf{a}\|_\infty = O_C(1)$. Since $\|\mathbf{a_i}\|_\infty = O_C(1)$ for every i, and $\{\mathbf{a_1}, \cdots, \mathbf{a_u}\}$ is a basis for the lattice $\Theta L(\mathbb{Z}^d)$, this implies that $|\lambda_i| = O_C(1)$ for every i.

So then $\sum_{j=1}^{d-u} \mu_j \mathbf{y_j}$ is a vector in $\ker \Theta$ satisfying $\| \sum_{j=1}^{d-u} \mu_j \mathbf{y_j}\|_\infty = O_C(1)$. Since $\{\mathbf{y_1}, \cdots, \mathbf{y_{m-u}}\}$ is a set of linearly independent vectors, each of which has integer coordinates with absolute value $O_C(1)$, this implies that $|\mu_j| = O_C(1)$ for every j.

Therefore P satisfies the conclusions of the lemma. □

Remark 5.11.5. We note the effects of the above construction in the case when L has algebraic coefficients. We use a rudimentary version of height: if $Q \in \mathbb{Z}[X]$ we define

$$H(Q) := \max(|q_i| : q_i \text{ a coefficient of } Q)$$

to be the *height* of Q, and we say that the height of an algebraic number is the height of its minimal polynomial. (So there are $O_{k,H}(1)$ algebraic numbers of degree at most k and height at most H). Then, if in the statement of Lemma 5.11.4 all the coefficients of L are algebraic numbers with degree at most k and height at most H,

all the coefficients of P are algebraic numbers of degree $O_k(1)$ and height $O_{C,k,H}(1)$.

5.12 The approximation function in the algebraic case

We use this final section to give the proof of relation (5.15). The following lemma makes this relation quantitatively precise.

Lemma 5.12.1. *Let m, d be natural numbers, with $d \geqslant m+1$, and let c, C be positive constants. Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a surjective linear map, and suppose that the matrix of L has algebraic coefficients of algebraic degree at most k and algebraic height at most H (see Remark 5.11.5 for definitions). Suppose that $\|L\|_\infty \leqslant C$, that $\mathrm{dist}(L, V_{\mathrm{rank}}(m,d)) \geqslant c$, and that L has rational complexity at most C. Let τ_1, τ_2 be two parameters in the range $0 < \tau_1, \tau_2 \leqslant 1$. Then*

$$A_L(\tau_1, \tau_2) \gg_{k,H,c,C} \min(\tau_1, \tau_2^{O_k(1)}).$$

Proof. We begin by reducing to the case when L is purely irrational. Indeed, consider Lemma 5.5.10 and replace L by the map L' (expression (5.34)). By part (9) of Lemma 5.5.10, $A_{L'}(\tau_1, \tau_2) \ll_{c,C} A_L(\Omega_{c,C}(\tau_1), \Omega_{c,C}(\tau_2))$. Also, using Remark 5.11.5, it follows that L' has algebraic coefficients of algebraic degree at most $O_k(1)$ and algebraic height at most $O_{c,C,k,H}(1)$. So, replacing L with L', without loss of generality we may assume that L is purely irrational.

Suppose for contradiction that for all choices of constants c_1 and C_2, there exist parameters τ_1 and τ_2 such that $A_L(\tau_1, \tau_2) < c_1 \min(\tau_1, \tau_2^{C_2})$, i.e. there exists a map

$\alpha \in (\mathbb{R}^m)^*$ and a map $\varphi \in (\mathbb{Z}^d)^T$ such that $\tau_1 \leqslant \|\alpha\|_\infty \leqslant \tau_2^{-1}$ and

$$\|L^*\alpha - \varphi\|_\infty < c_1 \min(\tau_1, \tau_2^{C_2}). \tag{5.94}$$

Fix α and φ so that they satisfy (5.94). We will obtain a contradiction if c_1 is small enough in terms of c, C, k, H, and if and C_2 is large enough in terms of k.

In the first part of the proof, we apply various reductions to enable us to replace α with a map that has integer coordinates with respect to the standard dual basis of $(\mathbb{R}^m)^*$.

Let M be a rank matrix of L (Proposition 5.3.1), and assume without loss of generality that M consists of the first m columns of L. Then there exists a map $\beta \in (\mathbb{R}^m)^*$, namely $\beta := M^*\alpha$, such that $\tau_1 \ll_{c,C} \|\beta\|_\infty \ll_{c,C} \tau_2^{-1}$ and

$$\|L^*(M^{-1})^*\beta - \varphi\|_\infty < c_1 \min(\tau_1, \tau_2^{C_2}). \tag{5.95}$$

Since the first m columns of $M^{-1}L$ form the identity matrix,(5.95) implies that

$$\operatorname{dist}(\beta, (\mathbb{Z}^m)^T) < c_1 \min(\tau_1, \tau_2^{C_2}). \tag{5.96}$$

We know that $\|\beta\|_\infty = \Omega_{c,C}(\tau_1)$. Also, considering (5.96), by perturbing β by a suitable element $\gamma \in (\mathbb{R}^m)^*$ with $\|\gamma\|_\infty < c_1 \min(\tau_1, \tau_2^{C_2})$ we may obtain a map $\rho \in (\mathbb{Z}^m)^T$. Combining these facts, note how

$$\begin{aligned} \|\rho\|_\infty &\geqslant \|\beta\|_\infty - c_1 \min(\tau_1, \tau_2^{C_2}) \\ &\gg_{c,C} \tau_1 \end{aligned}$$

if c_1 is small enough, and so certainly $\rho \neq 0$.

From (5.95), we therefore conclude that there exists some $\rho \in (\mathbb{Z}^m)^T \setminus \{0\}$, satisfying $\|\rho\|_\infty = O_{c,C}(\tau_2^{-1})$, such that

$$\|L^*(M^{-1})^*\rho - \varphi\|_\infty < c_1 C_3 \tau_2^{C_2} \tag{5.97}$$

where C_3 is some constant that depends on c and C. Referring back to (5.94), we see that we have achieved our goal of replacing α with a map that has integer coefficients.

Expression (5.97) leads to a contradiction. Morally this follows from Liouville's theorem on the diophantine approximation of algebraic numbers, but we couldn't find exactly the statement we needed in the literature, so we include a short argument here.

Indeed, let $\varphi = (\varphi_1 \quad \cdots \quad \varphi_d)$ be the representation of φ with respect to the standard dual basis of $(\mathbb{R}^d)^*$ (with analogous notation for $L^*(M^{-1})^*\rho$). Since L is assumed to be purely irrational, so is $M^{-1}L$. Therefore, since $\rho : \mathbb{R}^m \longrightarrow \mathbb{R}$ is surjective (since it is non-zero), we may pick some co-ordinate i at most d for which $(L^*(M^{-1})^*\rho)_i - \varphi_i \neq 0$. So there are algebraic numbers $\lambda_1, \cdots, \lambda_m$ with algebraic degree $O_k(1)$ and algebraic height $O_{c,C,k,H}(1)$ for which

$$0 < |\sum_{j=1}^{m} \lambda_j \rho_j - \varphi_i| < c_1 C_3 \tau_2^{C_2}, \tag{5.98}$$

where $(\rho_1 \quad \cdots \quad \rho_m)$ is the representation of ρ with respect to the standard dual basis. Note that if c_1 is small enough, by (5.98) and the fact that $\|\rho\|_\infty = O_{c,C}(\tau_2^{-1})$ one has $|\varphi_i| = O_{c,C}(\tau_2^{-1})$.

Our aim will be to find a suitable polynomial Q for which $Q(\sum_{j=1}^{m} \lambda_j \alpha_j) = 0$, and then to apply Liouville's original argument.

Assume without loss of generality that each $\lambda_j\rho_j$ is non-zero. For each j at most m, let $Q_j \in \mathbb{Z}[X]$ denote the minimal polynomial of $\lambda_j\rho_j$. Note that the degree of Q_j is $O_k(1)$ (since $\rho_j \in \mathbb{Z}$). By the bounds on the degree and height of λ_j, and since $\|\rho\|_\infty = O_{c,C}(\tau_2^{-1})$, we have $H(Q_j) = O_{c,C,k,H}(\tau_2^{-O_k(1)})$.

By using the standard construction based on resultants (see [14, section 4.2.1]), this implies that there is a polynomial $Q \in \mathbb{Z}[X]$ with degree $O_k(1)$ such that $Q(\sum_{j=1}^{m} \lambda_j\rho_j) = 0$ and $H(Q) = O_{c,C,k,H}(\tau_2^{-O_k(1)})$.

Now, it could be that φ_i is a root of Q. If this is the case, we use the factor theorem and Gauss' Lemma to replace Q by the integer-coefficient polynomial $Q \cdot (X - \varphi_i)^{-1}$. In this case, $H(Q \cdot (X - \varphi_i)^{-1}) \ll_{c,C,k,H} (\varphi_i + 1)^{O_k(1)}\tau_2^{-O_k(1)}$. By repeating this process as necessary, since $|\varphi_i| = O_{c,C}(\tau_2^{-1})$ we may assume therefore that φ_i is not a root of Q and that there exists a constant C_L depending on L such that $H(Q) = O_{c,C,k,H}(\tau_2^{-O_k(1)})$.

This immediately implies a bound on the derivative of Q, namely that, for any θ,

$$|Q'(\theta)| \ll_{c,C,k,H} \tau_2^{-O_k(1)} \sum_{0 \leqslant a \leqslant O_k(1)} \theta^a.$$

But then the mean value theorem implies that for some θ in the interval $[\sum_j \lambda_j\alpha_j, \varphi_i]$ one has

$$1 \leqslant |Q(\varphi_i)| = |Q(\sum_{j=1}^{m} \lambda_j\rho_j) - Q(\varphi_i)| \leqslant |Q'(\theta)||\sum_{j=1}^{m} \lambda_j\rho_j - \varphi_i| \ll_{c,C,k,H} c_1C_3\tau_2^{-O_k(1)}\tau_2^{C_2}.$$

If C_2 is large enough in terms of k, this implies that $c_1 = \Omega_{c,C,k,H}(1)$, which is a contradiction if c_1 is small enough. Therefore the lemma holds. □

Bibliography

[1] C. Aistleitner, T. Lachmann, and F. Pausinger. Pair correlations and equidistribution. *Journal of Number Theory*, 182:206–220, 2017.

[2] C. Aistleitner, T. Lachmann, and N. Technau. There is no Khintchine threshold for metric pair correlations. Preprint available at `https://arxiv.org/abs/1802.02659`.

[3] C. Aistleitner, G. Larcher, and M. Lewko. Additive energy and the Hausdorff dimension of the exceptional set in metric pair correlation problems. *Israel Journal of Mathematics*, 222(1):463–485, 2017.

[4] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016.

[5] A. Axer. Über einige Grenzwertsätze. *Wien. Ber.*, 120:1253–1298, 1911.

[6] A. Ayyad, T. Cochrane, and Z. Zheng. The congruence $x_1x_2 \equiv x_3x_4 \pmod{p}$, the equation $x_1x_2 = x_3x_4$, and mean values of character sums. *J. Number Theory*, 59(2):398–413, 1996.

[7] A. Baker. On some diophantine inequalities involving primes. *J. Reine Angew. Math.*, 228:166–181, 1967.

[8] R. C. Baker. Metric number theory and the large sieve. *J. London Math. Soc. (2)*, 24(1):34–40, 1981.

[9] R. C. Baker. *Diophantine inequalities*, volume 1 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1986. Oxford Science Publications.

[10] T. F. Bloom, S. Chow, A. Gafni, and A. Walker. Additive energy and the metric poissonian property. *Mathematika*, 64(3):679–700, 2018.

[11] J. Bourgain. A quantitative Oppenheim theorem for generic diagonal quadratic forms. *Israel J. Math.*, 215(1):503–512, 2016.

[12] D. A. Burgess. The distribution of quadratic residues and non-residues. *Mathematika*, 4:106–112, 1957.

[13] D. A. Burgess. On character sums and primitive roots. *Proc. London Math. Soc. (3)*, 12:179–192, 1962.

[14] H. Cohen. *A course in computational algebraic number theory*, volume 138 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1993.

[15] B. Cook, Á. Magyar, and M. Pramanik. A Roth-type theorem for dense subsets of $\mathbb{R}^d$. *Bulletin of the London Mathematical Society*, 49(4):676–689, 2017.

[16] H. Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2000.

[17] H. Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2005. With a foreword by R. C. Vaughan, D. R. Heath-Brown and D. E. Freeman, Edited and prepared for publication by T. D. Browning.

[18] H. Davenport and H. Heilbronn. On indefinite quadratic forms in five variables. *J. London Math. Soc.*, 21:185–193, 1946.

[19] P. Durcik, V. Kovač, and L. Rimanić. On side lengths of corners in positive density subsets of the Euclidean space. *International Mathematics Research Notices*, page rnx093, 2017.

[20] P. Dusart. Explicit estimates of some functions over primes. *The Ramanujan Journal*, 45(1):227–251, Jan 2018.

[21] P. D. T. A. Elliott. The least prime k^{th}-power residue. *J. London Math. Soc. (2)*, 3:205–210, 1971.

[22] P. Erdős, A. M. Odlyzko, and A. Sárközy. On the residues of products of prime numbers. *Period. Math. Hungar.*, 18(3):229–239, 1987.

[23] D. E. Freeman. Asymptotic lower bounds and formulas for Diophantine inequalities. In *Number theory for the millennium, II (Urbana, IL, 2000)*, pages 57–74. A K Peters, Natick, MA, 2002.

[24] G. A. Freĭman. Groups and the inverse problems of additive number theory. In *Number-theoretic studies in the Markov spectrum and in the structural theory of set addition (Russian)*, pages 175–183. Kalinin. Gos. Univ., Moscow, 1973.

[25] J. Friedlander and H. Iwaniec. A note on Dirichlet L-functions. Preprint available at `https://arxiv.org/abs/1701.03771`.

[26] J. Friedlander and H. Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2010.

[27] J. B. Friedlander, P. Kurlberg, and I. E. Shparlinski. Products in residue classes. *Math. Res. Lett.*, 15(6):1133–1147, 2008.

[28] M. Z. Garaev. Character sums in short intervals and the multiplication table modulo a large prime. *Monatsh. Math.*, 148(2):127–138, 2006.

[29] M. Z. Garaev. An estimate for Kloosterman sums with primes and its application. *Mat. Zametki*, 88(3):365–373, 2010.

[30] M. Z. Garaev and V. C. Garcia. The equation $x_1x_2 = x_3x_4 + \lambda$ in fields of prime order and applications. *J. Number Theory*, 128(9):2520–2537, 2008.

[31] A. O. Gel′fond. On the arithmetic equivalent of analyticity of the Dirichlet L-series on the line $\operatorname{Re} s = 1$. *Izv. Akad. Nauk SSSR. Ser. Mat.*, 20:145–166, 1956.

[32] A. O. Gel′fond and U. V. Linnik. *Elementary methods in the analytic theory of numbers*. Translated from the Russian by D. E. Brown. Translation edited by I. N. Sneddon. International Series of Monographs in Pure and Applied Mathematics, Vol. 92. Pergamon Press, Oxford-New York-Toronto, Ont., 1966.

[33] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[34] W. T. Gowers and J. Wolf. The true complexity of a system of linear equations. *Proc. Lond. Math. Soc. (3)*, 100(1):155–176, 2010.

[35] B. Green. Montréal notes on quadratic Fourier analysis. In *Additive combinatorics*, volume 43 of *CRM Proc. Lecture Notes*, pages 69–102. Amer. Math. Soc., Providence, RI, 2007.

[36] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.

[37] B. Green and T. Tao. Quadratic uniformity of the Möbius function. *Ann. Inst. Fourier (Grenoble)*, 58(6):1863–1935, 2008.

[38] B. Green and T. Tao. Linear equations in primes. *Ann. of Math. (2)*, 171(3):1753–1850, 2010.

[39] B. Green and T. Tao. The Möbius function is strongly orthogonal to nilsequences. *Ann. of Math. (2)*, 175(2):541–566, 2012.

[40] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Ann. of Math. (2)*, 176(2):1231–1372, 2012.

[41] S. Grepstad and G. Larcher. On pair correlation and discrepancy. *Archiv der Mathematik*, 109(2):143–149, Aug 2017.

[42] H. Halberstam and H.-E. Richert. *Sieve methods*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974. London Mathematical Society Monographs, No. 4.

[43] G. Harman. Metric Diophantine approximation with two restricted variables. I. Two square-free integers, or integers in arithmetic progressions. *Math. Proc. Cambridge Philos. Soc.*, 103(2):197–206, 1988.

[44] G. Harman. *Metric number theory*, volume 18 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1998.

[45] G. Harman and I. E. Shparlinski. Products of small integers in residue classes and additive properties of Fermat quotients. *Int. Math. Res. Not. IMRN*, (5):1424–1446, 2016.

[46] D. R. Heath-Brown. Zero-free regions for Dirichlet L-functions, and the least prime in an arithmetic progression. *Proc. London Math. Soc. (3)*, 64(2):265–338, 1992.

[47] D. R. Heath-Brown. Pair correlation for fractional parts of αn^2. *Math. Proc. Cambridge Philos. Soc.*, 148(3):385–407, 2010.

[48] D. R. Heath-Brown and X. Li. Prime values of a^2+p^4. *Invent. Math.*, 208(2):441–499, 2017.

[49] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[50] H. Kadiri. Short effective intervals containing primes in arithmetic progressions and the seven cubes problem. *Math. Comp.*, 77(263):1733–1748, 2008.

[51] M. Kneser. Abschätzung der asymptotischen Dichte von Summenmengen. *Math. Z.*, 58:459–484, 1953.

[52] T. Lachmann and N. Technau. On exceptional sets in the metric poissonian pair correlations problem. Preprint available at `https://arxiv.org/abs/1708.08599`,.

[53] U. V. Linnik. On the least prime in an arithmetic progression. I. The basic theorem. *Rec. Math. [Mat. Sbornik] N.S.*, 15(57):139–178, 1944.

[54] U. V. Linnik. On the least prime in an arithmetic progression. II. The Deuring-Heilbronn phenomenon. *Rec. Math. [Mat. Sbornik] N.S.*, 15(57):347–368, 1944.

[55] G. A. Margulis. Discrete subgroups and ergodic theory. In *Number theory, trace formulas and discrete groups (Oslo, 1987)*, pages 377–398. Academic Press, Boston, MA, 1989.

[56] R. McCutcheon. FVIP systems and multiple recurrence. *Israel J. Math.*, 146:157–188, 2005.

[57] H. L. Montgomery. *Topics in multiplicative number theory.* Lecture Notes in Mathematics, Vol. 227. Springer-Verlag, Berlin-New York, 1971.

[58] H. L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.

[59] H. L. Montgomery and R. C. Vaughan. Hilbert's inequality. *J. London Math. Soc. (2)*, 8:73–82, 1974.

[60] H. L. Montgomery and R. C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

[61] W. Müller. Systems of quadratic Diophantine inequalities. *J. Théor. Nombres Bordeaux*, 17(1):217–236, 2005.

[62] M. R. Murty. *Problems in analytic number theory*, volume 206 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2008. Readings in Mathematics.

[63] M. B. Nathanson. *Additive number theory: Inverse problems and the geometry of sumsets*, volume 165 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996.

[64] M. B. Nathanson. *Additive number theory: the classical bases*, volume 164 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996.

[65] S. T. Parsell. Irrational linear forms in prime variables. *J. Number Theory*, 97(1):144–156, 2002.

[66] S. T. Parsell. On simultaneous diagonal inequalities. III. *Q. J. Math.*, 53(3):347–363, 2002.

[67] J. Pintz. Elementary methods in the theory of L-functions. VI. On the least prime quadratic residue (mod ρ). *Acta Arith.*, 32(2):173–178, 1977.

[68] P. Pollack. Bounds for the first several prime character nonresidues, version 1. Preprint available at `http://arxiv.org/abs/1508.05035`.

[69] P. Pollack. Bounds for the first several prime character nonresidues. *Proc. Amer. Math. Soc.*, 145(7):2815–2826, 2017.

[70] J.-C. Puchta. Primes in short arithmetic progressions. *Acta Arith.*, 106(2):143–149, 2003.

[71] O. Ramaré and A. Walker. Products of primes in arithmetic progressions: a footnote in parity breaking. *J. Théor. Nombres Bordeaux*, 30(1):219–225, 2018.

[72] W. Rudin. *Fourier analysis on groups*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1990. Reprint of the 1962 original, A Wiley-Interscience Publication.

[73] Z. Rudnick and P. Sarnak. The pair correlation function of fractional parts of polynomials. *Comm. Math. Phys.*, 194(1):61–70, 1998.

[74] Z. Rudnick, P. Sarnak, and A. Zaharescu. The distribution of spacings between the fractional parts of $n^2\alpha$. *Invent. Math.*, 145(1):37–57, 2001.

[75] Z. Rudnick and A. Zaharescu. A metric result on the pair correlation of fractional parts of sequences. *Acta Arith.*, 89(3):283–293, 1999.

[76] T. Sanders. The structure theory of set addition revisited. *Bull. Amer. Math. Soc. (N.S.)*, 50(1):93–127, 2013.

[77] L. G. Sathe. On a problem of Hardy on the distribution of integers having a given number of prime factors. I. - II. *J. Indian Math. Soc. (N.S.)*, 17:63–82, 83–141, 1953.

[78] L. Schnirelmann. Über additive Eigenschaften von Zahlen. *Math. Ann.*, 107(1):649–690, 1933.

[79] I. Shparlinski. *On short products of primes in arithmetic progressions*. Preprint available at `https://arxiv.org/abs/1705.06087`.

[80] I. E. Shparlinski. Modular hyperbolas. *Jpn. J. Math.*, 7(2):235–294, 2012.

[81] I. E. Shparlinski. On products of primes and almost primes in arithmetic progressions. *Period. Math. Hungar.*, 67(1):55–61, 2013.

[82] S. Steinerberger. Localized quantitative criteria for equidistribution. *Acta Arith.*, 180(2):183–199, 2017.

[83] T. Tao. An inverse theorem for the continuous gowers uniformity norm. Blog post, available at `https://terrytao.wordpress.com/2015/07/22/an-inverse-theorem-for-the-continuous-gowers-uniformity-norm/`.

[84] T. Tao. *Higher order Fourier analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[85] T. Tao and V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[86] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 46 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Translated from the second French edition (1995) by C. B. Thomas.

[87] R. C. Vaughan. *The Hardy-Littlewood method*, volume 125 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 1997.

[88] A. I. Vinogradov and U. V. Linnik. Hypoelliptic curves and the least prime quadratic residue. *Dokl. Akad. Nauk SSSR*, 168:259–261, 1966.

[89] A. Walker. Linear inequalities in primes. In preparation.

[90] A. Walker. A multiplicative analogue of Schnirelmann's theorem, version 1. Preprint available at `http://arxiv.org/abs/1505.03328v1`.

[91] A. Walker. A multiplicative analogue of Schnirelmann's theorem. *Bull. Lond. Math. Soc.*, 48(6):1018–1028, 2016.

[92] A. Walker. The primes are not metric poissonian. *Mathematika*, 64(1):230–236, 2018.

[93] T. D. Wooley. On Diophantine inequalities: Freeman's asymptotic formulae. In *Proceedings of the Session in Analytic Number Theory and Diophantine Equations*, volume 360 of *Bonner Math. Schriften*, page 32. Univ. Bonn, Bonn, 2003.

[94] T. Xylouris. *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression*, volume 404 of *Bonner Mathematische Schriften [Bonn Mathematical Publications]*. Universität Bonn, Mathematisches Institut, Bonn, 2011. Dissertation for the degree of Doctor of Mathematics and Natural Sciences at the University of Bonn, Bonn, 2011.