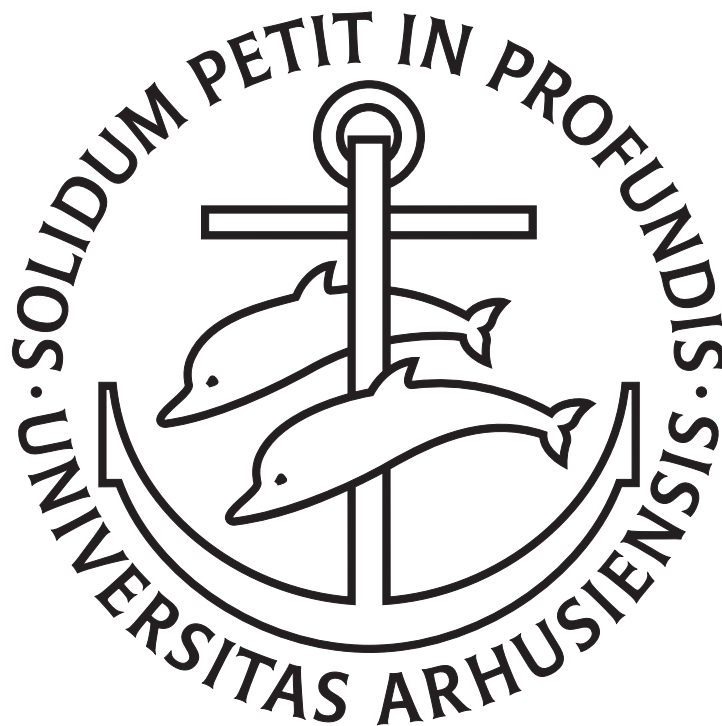


On the existence of long arithmetic progressions in the primes

The Green-Tao theorem



Master Thesis in Mathematics

by Jonas Lindstrøm Jensen, 2003 3834

Supervisors: Simon Kristensen and Jørgen Brandt

University of Aarhus – Department of Mathematical Sciences, March 2009

Abstract

The main part of this thesis is a proof of the Green-Tao theorem which states that the primes contains arithmetic progressions of any length. The proof follows the original proof of Ben Green and Terrence Tao closely but contributes with quite a lot of technical details that Green and Tao left out.

In the last chapter we consider sets with arbitrarily long arithmetic progressions, AP-sets, and prove that a homogenous system of linear equations has solutions in any AP-set if $(1, 1, \dots, 1)$ is a solution and the solution space has dimension at least 2. This gives us a new characterization of AP-sets.

Contents

| | |
|--|-----------|
| Contents | 2 |
| 1 Introduction | 5 |
| 2 The theorem and basic concepts | 7 |
| 2.1 The theorem | 7 |
| 2.2 Basic definitions and notation | 7 |
| 2.3 Measure and measure conditions | 9 |
| 2.4 Some useful results about expected values | 11 |
| 2.5 Function spaces | 12 |
| 3 Proof strategy | 13 |
| 3.1 Szemerédi’s theorem in pseudorandom measures | 13 |
| 3.2 The proof | 14 |
| 3.3 Overview of the rest of the proof | 16 |
| 4 Gowers uniformity and von Neumanns theorem | 17 |
| 4.1 The Gowers uniformity norm | 17 |
| 4.2 A special Cauchy-Schwarz | 24 |
| 4.3 Von Neumann’s theorem | 27 |
| 5 Dual functions and basic Gowers anti-uniformity | 37 |
| 5.1 The dual space of U^d | 37 |
| 5.2 Gowers anti-uniform functions | 40 |
| 6 σ-algebras on \mathbb{Z}_N | 51 |
| 6.1 Definitions and notation | 51 |
| 6.2 Two propositions | 54 |
| 7 Furstenberg Tower | 61 |
| 7.1 The Furstenberg tower | 61 |
| 7.2 Proof of Szemerédi in pseudorandom measures | 69 |
| 8 Construction of a pseudorandom measure | 73 |

| | |
|--|-----------|
| <i>CONTENTS</i> | 3 |
| 8.1 Definitions and notation | 73 |
| 8.2 A proof that ν is pseudorandom | 76 |
| 9 Solutions in arithmetic progression to linear equations | 87 |
| 9.1 Introduction | 87 |
| 9.2 APs and GAPS | 87 |
| 9.3 Finding solutions in an AP-set | 89 |
| 9.4 Prime-like sets | 90 |
| 9.5 The Erdős-Turan conjecture | 91 |
| Bibliography | 93 |

Chapter 1

Introduction

Additive patterns in the primes is a field of research that has drawn much attention. Famous open problems in this field are the existence of infinitely many twin primes and Goldbach's conjecture. It is believed that both of these conjectures are true, and heuristic arguments suggest that they are true if the primes are *randomly distributed* in the right sense. Of course the primes are not randomly distributed since there are no even primes (except 2) and no primes divisible by 3, 5, 7, The general conjecture is that these obstructions are the only ones that stops the heuristic argument, so there are no "secret" patterns in the primes except for these obvious ones.

This thesis is considering a specific additive pattern namely arithmetic progressions which are configurations of the form

$$a, a + d, \dots, a + (k - 1)d$$

for some $a, d, k \in \mathbb{N}$. The heuristic arguments mentioned above also suggest that there are arbitrarily long arithmetic progressions in the primes. Van der Waerden proved in 1927 [13] that if the integers are coloured with finitely many colours then there is a colour such that there are arithmetic progressions of any length using numbers of this colour – this result is not so interesting when considering the primes as a colouring, because then van der Waerden only gives us long arithmetic progressions in either the primes or the composite numbers and the composite numbers contains infinitely long arithmetic progressions, for instance the even numbers.

A stronger result was proven by Szemerédi in 1975 [12], namely that any subset of positive (upper) density contains arithmetic progressions of any length. This theorem implies Van Der Waerden's theorem since the numbers with at least one of the finitely many colours must have positive density, but the primes have density 0 and so Szemerédi's theorem says nothing about the primes.

Ben Green and Terence Tao proved in 2004 [6] that the primes contain arbitrarily long arithmetic progressions and it is their proof of this theorem that this thesis will present. The great open problem in this context is the Erdős-Turan Conjecture which states that

we for $A \subseteq \mathbb{N}$ has

$$\sum_{a \in A} \frac{1}{a} = \infty \Rightarrow A \text{ contains arithmetic progressions of any length.}$$

This conjecture would imply all the above results, and will be discussed shortly in chapter 9 of this thesis, where we find a new equivalent formulation of it.

The proof of Green and Tao consists of several parts. In chapter 2 of this thesis we will introduce the necessary notation and some basic concepts of the proof, and in chapter 3 we will present the strategy and structure of the proof. The main part is the proof of a generalization of Szemerédi's theorem, and chapters 4-7 will be dedicated to proving this. Chapter 8 will deal with the construction of a certain function, a *pseudorandom measure*, which is concentrated on the primes and satisfies some bounds and other criteria necessary for the proof. The proof as it is presented in this thesis is more or less as in [6] but I have contributed with quite a lot of technical details that they have left out in their paper, and I have changed the order and structure of some of the proofs.

Chapter 9 is my personal contribution, which can be seen as an application of the Green-Tao theorem that proves the existence of prime solutions to linear equations. More generally it is a proof of the existence of solutions to certain systems of linear equations in any subset of the integers that contain arbitrarily long arithmetic progressions. Furthermore I will present a new arithmetic way of characterizing such subsets which gives a new formulation of the Erdős-Turan Conjecture. An article [8] on the results of this chapter has on the 16th of February been submitted to the journal 'INTEGERS: Electronic Journal of Combinatorial Number Theory'. The methods in this chapter are quite elementary and the chapter can be read alone without reading the rest of the thesis.

I would like to thank my supervisors Simon and Jørgen who have helped me through the writing process, Jimi Lee Truelsen for reading, correcting and commenting on the thesis and Andrew Granville for commenting on the results in chapter 9. I would furthermore thank my fellow students for good company, and finally thanks to my family for all the support I have received during my many years of studying.

Chapter 2

The theorem and basic concepts

2.1 The theorem

The theorem I want to prove in this thesis is the following theorem of Green and Tao [6].

Theorem 2.1.1. *The primes contain infinitely many arithmetic progressions of length k for any k .*

Recall that an arithmetic progression of length k is a set of the form

$$\{x, x + r, x + 2r, \dots, x + (k - 1)r\}.$$

where $x, r \in \mathbb{N}$. So the theorem states that the primes contain infinitely many of these of any given length. Notice that we actually just need one arithmetic progression of any length to prove the existence of infinitely many arithmetic progressions of any given length.

2.2 Basic definitions and notation

Definition 2.2.1. Let $N \in \mathbb{N}$ be a prime. We denote $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$.

Remark 2.2.2. Throughout the proof, k will denote the size of the arithmetic progression we are looking for. So in all proofs we assume that N is a very large prime and that k is a natural number. Sometimes we will assume that k is bounded from below but it does not really matter, since a progression of length k contains progressions of length m for all $m \leq k$.

Several times during the proof it is necessary to assume that N is sufficiently large depending on some parameters, but we will not discuss the size of N – just make sure that the proof is valid.

In the rest of the thesis $o(1)$ will denote a term that tends to zero as $N \rightarrow \infty$, and we will write $O(1)$ for a bounded quantity. If the convergence to zero or the coefficient in the bound depends on some parameters, we will write these parameters in the subscript. The constant will always depend on k (the length of the arithmetic progression we are looking for), so we will usually not write that. We write $O(X) = O(1)X$ and $o(X) = o(1)X$.

We define the expected value of real functions on \mathbb{Z}_N as follows.

Definition 2.2.3 (Expected value). Let $n \geq 1$ and $f : A \rightarrow \mathbb{R}$ where $A \subseteq \mathbb{Z}_N^n$ with $A \neq \emptyset$. Then we define

$$\mathbb{E}(f(x) \mid x \in A) = \frac{\sum_{x \in A} f(x)}{\#A}.$$

If f is defined on all of \mathbb{Z}_N^n we write

$$\mathbb{E}(f) = \mathbb{E}(f(x) \mid x \in \mathbb{Z}_N^n).$$

Remark 2.2.4. We will need to take expectations with respect to more than one variable. The generalisation to this case is done by taking average over the product set. All the results here are valid for several variables because two variables over two sets can be written as one variable over the product set.

The expected value can be extended to be taken over any function $f : A \rightarrow \mathbb{R}$ where $\#A < \infty$ by the same definition, but usually A will be a subset of \mathbb{Z}_N .

The following proposition can be proved by straightforward calculations, and will be quite useful in calculations involving expected values.

Proposition 2.2.5 (Properties of the expected value). Let $f : A \rightarrow \mathbb{R}, g : B \rightarrow \mathbb{R}$ where $A \subseteq \mathbb{Z}_N^n, B \subseteq \mathbb{Z}_N^m$ with $A, B \neq \emptyset$.

1. If $m = n, A = B$ and we let $a, b \in \mathbb{R}$ then

$$\mathbb{E}(af(x) + bg(x) \mid x \in A) = a\mathbb{E}(f(x) \mid x \in A) + b\mathbb{E}(g(x) \mid x \in A).$$

2. We furthermore have

$$\begin{aligned} \mathbb{E}(f(x)g(y) \mid x \in A, y \in B) &= \mathbb{E}(f(x)\mathbb{E}(g(y) \mid y \in B) \mid x \in A) \\ &= \mathbb{E}(f(x) \mid x \in A)\mathbb{E}(g(y) \mid y \in B), \end{aligned}$$

3. and as a special case we have

$$\mathbb{E}(f(x) \mid x \in A) = \mathbb{E}(f(x) \mid x \in A, y \in B).$$

4. If $\phi, \psi : A \times B \rightarrow \mathbb{R}$. Then

$$\mathbb{E}(\phi(x, y)\psi(x, y) \mid x \in A, y \in B) = \mathbb{E}(\mathbb{E}(\phi(x, y)\psi(x, y) \mid y \in B) \mid x \in A).$$

5. If I is some finite set and we have $h_i : A \rightarrow \mathbb{R}$ for all $i \in I$ then

$$\mathbb{E}\left(\prod_{i \in I} h_i(x) \mid x \in A\right)^2 = \mathbb{E}\left(\prod_{i \in I, j \in I} h_i(x)h_j(y) \mid x, y \in A\right).$$

2.3 Measure and measure conditions

The proof uses the notion of a measure on $\mathbb{Z}/N\mathbb{Z}$ but it is not a measure in the measure theoretic sense since, the only reasonable measure to consider on such a set is the counting measure. The only thing we require of a measure is that it is positive and that the expected value of a measure should tend to 1 as N goes to infinity.

Definition 2.3.1 (Measure). A function $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ is called a *measure* if

$$\mathbb{E}(\nu) = 1 + o(1).$$

We want to construct a specific kind of measure. This measure has to satisfy the following two conditions.

Definition 2.3.2 (The linear forms condition). A measure ν is said to satisfy the (m_0, t_0, L_0) -*linear forms condition* if the following is satisfied. Let $m \leq m_0$ and $t \leq t_0$. Let $L = (L_{ij}) \in \text{Mat}_{m,t}(\mathbb{Q})$ be such that no row is a rational multiple of another row, and such that all entries has height $\leq L_0$ and full rank. Now let $b_1, \dots, b_m \in \mathbb{Z}_N$ and define for $i = 1, \dots, m$ linear forms $\psi_i : \mathbb{Z}_N^t \rightarrow \mathbb{Z}_N$ by

$$\psi(x_1, \dots, x_t) = \sum_{j=1}^t L_{ij} x_j + b_i$$

where L_{ij} is considered as a member of \mathbb{Z}_N (so $N \geq L_0$ and $a/b = ab^{-1} \in \mathbb{Z}_N$). Then

$$\mathbb{E}(\nu(\psi_1(x)) \dots \nu(\psi_m(x)) \mid x \in \mathbb{Z}_N^t) = 1 + o_{L_0, m_0, t_0}(1).$$

Remark 2.3.3. Recall that the height of a rational number $\frac{a}{b}$ with $\gcd(a, b) = 1$ is $H(\frac{a}{b}) = \max(|a|, |b|)$.

Notice also that the (m_0, t_0, L_0) -linear forms condition implies the (m'_0, t'_0, L'_0) -linear forms condition when $m'_0 \leq m_0, t'_0 \leq t_0$ and $L'_0 \leq L_0$.

Definition 2.3.4 (The correlation condition). A measure $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ is said to satisfy the m_0 -*correlation condition* for $m_0 \in \mathbb{Z}^+$ if for every $1 < m \leq m_0$ there is $\tau : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ such that $\mathbb{E}(\tau^q) = O_{m,q}(1)$ for all $q \geq 1$ and such that

$$\mathbb{E}(\nu(x + h_1)\nu(x + h_2) \cdots \nu(x + h_m) \mid x \in \mathbb{Z}_N) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j) \quad (2.3.1)$$

for all $h_1, \dots, h_m \in \mathbb{Z}_N$.

A measure satisfying both of these conditions with suitable parameters will be called *pseudorandom*. We define it as follows.

Definition 2.3.5 (Pseudorandom measure). A measure $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ is called k -*pseudorandom* if it satisfies the $(k2^{k-1}, 3k - 4, k)$ -linear forms condition and the 2^{k-1} -correlation condition.

Example 2.3.6 (The constant measure). Define $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ by $\nu(x) = 1$ for all $x \in \mathbb{Z}_N$. Then ν is definitely a measure. It satisfies the linear forms condition with any parameters since

$$\mathbb{E}(\nu(\psi_1(x)) \cdots \nu(\psi_m(x)) \mid x \in \mathbb{Z}_N^t) = 1,$$

no matter what the ψ_i 's are. It also satisfies the correlation condition with any parameter, since the LHS in (2.3.1) is 1, and we can thus define $\tau = 1$ to get the desired inequality. So ν is k -pseudorandom for any k .

The following theorem tells us that the pseudorandom measures are star-shaped around the constant measure.

Lemma 2.3.7 (Pseudorandom measures are star-shaped around 1). *Let ν be k -pseudorandom. Then $\nu' = (\nu + 1)/2$ is k -pseudorandom.*

Proof. Since ν is a measure, we have that ν' is non-negative and

$$\mathbb{E}(\nu') = \frac{1}{2}(\mathbb{E}(\nu) + 1) = 1 + o(1),$$

so ν' is a measure.

To prove that ν' satisfies the $(k2^{k-1}, 3k - 4, k)$ -linear forms condition we let $m \leq k2^{k-1}$, $t \leq 3k - 4$ and let ψ_1, \dots, ψ_m be defined as in the definition of the linear forms condition. Then

$$\begin{aligned} \mathbb{E}(\nu'(\psi_1(x)) \cdots \nu'(\psi_m(x)) \mid x \in \mathbb{Z}_N^t) &= 2^{-m} \mathbb{E}((\nu(\psi_1(x)) + 1) \cdots (\nu(\psi_m(x)) + 1) \mid x \in \mathbb{Z}_N^t) \\ &= 2^{-m} \sum_{A \subseteq \{1, \dots, m\}} \mathbb{E} \left(\prod_{i \in A} \nu(\psi_i(x)) \mid x \in \mathbb{Z}_N^t \right). \end{aligned}$$

Now since ν is pseudorandom, it satisfies the $(k2^{k-1}, 3k - 4, k)$ -linear forms condition, and we see that the terms in the sum are of the form considered in the definition of the linear forms condition so each of the terms is $1 + o(1)$, and since the sum has 2^m terms (one for each subset $A \subseteq \{1, \dots, m\}$) the whole thing is $1 + o(1)$ and we are done.

We now need to prove that ν' satisfies the 2^{k-1} -correlation condition. Let $m \leq 2^{k-1}$ and $h_1, \dots, h_m \in \mathbb{Z}_N$. We have

$$\mathbb{E}(\nu'(x + h_1) \cdots \nu'(x + h_m) \mid x \in \mathbb{Z}_N) = 2^{-m} \sum_{A \subseteq \{1, \dots, m\}} \mathbb{E} \left(\prod_{i \in A} \nu(x + h_i) \mid x \in \mathbb{Z}_N \right). \quad (2.3.2)$$

There is a τ such that ν satisfies the correlation condition, so let us try to use this τ again. It certainly satisfies $\mathbb{E}(\tau^q) = O_{m,q}(1)$ for all $q \geq 1$. Let us now consider one of the terms in the sum. For any $A \subseteq \{1, \dots, m\}$ we have

$$\mathbb{E} \left(\prod_{i \in A} \nu(x + h_i) \mid x \in \mathbb{Z}_N \right) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j)$$

because ν satisfies the correlation condition. There are 2^m terms in the sum on the RHS of (2.3.2), so

$$\mathbb{E}(\nu'(x+h_1)\cdots\nu'(x+h_m) \mid x \in \mathbb{Z}_N) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j)$$

as desired. □

2.4 Some useful results about expected values

It will be necessary to perform linear changes of variables in expectations which we will formalize in the following way.

Definition 2.4.1 (A uniform cover). A map $\phi : A \rightarrow B$ where $A \subseteq \mathbb{Z}_N^n$ and $B \subseteq \mathbb{Z}_N^m$ is said to be a *uniform cover of B by A* if ϕ is surjective and if

$$\#\phi^{-1}(b) = \frac{\#A}{\#B}$$

for all $b \in B$.

Lemma 2.4.2 (Uniform covers preserve expected values). *If $\phi : A \rightarrow B$ is a uniform cover of B by A and $f : B \rightarrow \mathbb{R}$, then*

$$\mathbb{E}(f(\phi(a)) \mid a \in A) = \mathbb{E}(f(b) \mid b \in B).$$

Proof. We have

$$\mathbb{E}(f(\phi(a)) \mid a \in A) = \frac{1}{\#A} \sum_{a \in A} f(\phi(a)) = \frac{1}{\#A} \sum_{b \in B} \sum_{a \in \phi^{-1}(b)} f(b),$$

since ϕ is surjective. Now recall that $\#\phi^{-1}(b) = \#A/\#B$ so

$$\sum_{a \in \phi^{-1}(b)} f(b) = \frac{\#A}{\#B} f(b),$$

and hence

$$\mathbb{E}(f(\phi(a)) \mid a \in A) = \frac{1}{\#A} \sum_{b \in B} \frac{\#A}{\#B} f(b) = \mathbb{E}(f(b) \mid b \in B)$$

as desired. □

2.5 Function spaces

Definition 2.5.1 (L^q spaces on \mathbb{Z}_N^n). Let $1 \leq q \leq \infty$ and let $f : \mathbb{Z}_N^n \rightarrow \mathbb{R}$. We define

$$\|f\|_{L^q} = \mathbb{E}(|f|^q)^{1/q}$$

for $q < \infty$ and

$$\|f\|_{L^\infty} = \sup_{x \in \mathbb{Z}_N^n} |f(x)|.$$

We can consider the functions $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ as a Hilbert space with inner product $(f, g) \mapsto \mathbb{E}(fg)$. This gives us the triangle inequality, Jensens inequality, Hölders inequality and Cauchy-Schwarz for averages.

Lemma 2.5.2 (Jensens inequality). Let $f : A \rightarrow \mathbb{R}$ where $A \subseteq \mathbb{Z}_N$. If $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is convex then

$$\phi(\mathbb{E}(f(x) \mid x \in A)) \leq \mathbb{E}(\phi(f(x)) \mid x \in A).$$

Lemma 2.5.3 (Hölders inequality for averages). Let $f_i : A \rightarrow \mathbb{R}, i = 1, \dots, m$ for $A \subseteq \mathbb{Z}_N^n$ and let $p_1, \dots, p_m > 0$ and $q \geq 1$ be given such that

$$\frac{1}{p_1} + \dots + \frac{1}{p_m} = \frac{1}{q},$$

then

$$\mathbb{E} \left(\prod_{i=1}^m |f_i(x)|^q \mid x \in A \right)^{\frac{1}{q}} \leq \prod_{i=1}^m \mathbb{E} (|f_i(x)|^{p_i} \mid x \in A)^{\frac{1}{p_i}}.$$

Corollary 2.5.4 (Cauchy-Schwarz for averages). Let $f : A \rightarrow \mathbb{R}$ and $g : A \rightarrow \mathbb{R}$ for $A \subseteq \mathbb{Z}_N^n$. Then

$$\mathbb{E} (|f(x)g(x)| \mid x \in A)^2 \leq \mathbb{E} (f(x)^2 \mid x \in A) \mathbb{E} (g(x)^2 \mid x \in A).$$

Chapter 3

Proof strategy

3.1 Szemerédi's theorem in pseudorandom measures

Let us first recall Szemerédi's theorem. The formulation used here is a bit different from the usual formulation, since it uses the notion of expected values on \mathbb{Z}_N , but it is equivalent to the common version. See [12], [3] or [4] for 3 different proofs of this theorem.

Theorem 3.1.1 (Szemerédi's theorem). *Let $0 < \delta \leq 1$ and $k \geq 1$ be fixed. Let $f : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ such that*

$$0 \leq f(x) \leq 1$$

for all $x \in \mathbb{Z}_N$ and

$$\mathbb{E}(f) \geq \delta.$$

Then we have

$$\mathbb{E}(f(x)f(x+r) \cdots f(x+(k-1)r) \mid x, r \in \mathbb{Z}_N) \geq c(k, \delta) - o_{k, \delta}(1) \quad (3.1.1)$$

for some constant $c(k, \delta) > 0$, not depending on f nor N .

Note that since the RHS of (3.1.1) does not depend on N this theorem gives us arbitrarily long arithmetic progressions in the support of f , and since we require $f \leq 1$ and $\mathbb{E}(f) > 0$ this is only possible if the support of f has positive density. We now want to generalize Szemerédi's theorem in the following way. Chapter 4-7 of this thesis will be dealing with proving this theorem, and we use the original version of Szemerédi's theorem in the proof.

Theorem 3.1.2. *Let $k \geq 3$ and $0 < \delta \leq 1$. Let $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ be a k -pseudorandom measure and let $f : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ satisfy*

$$0 \leq f(x) \leq \nu(x)$$

for all $x \in \mathbb{Z}_N$ and

$$\mathbb{E}(f) \geq \delta.$$

Then we have

$$\mathbb{E}(f(x)f(x+r)\cdots f(x+(k-1)r) \mid x, r \in \mathbb{Z}_N) \geq c(k, \delta) - o_{k, \delta}(1),$$

for some constant $c(k, \delta)$.

Note that the only thing that has changed from theorem 3.1.1 to this theorem, is that we replace $f(x) \leq 1$ by $f(x) \leq \nu(x)$ for a pseudorandom measure ν , and since the constant function 1 is k -pseudorandom for any k this is a generalization. Now in this theorem it is possible to use a function f such that the support of f does not necessarily have positive density, since f is allowed to be larger than 1 on some values as long as it is bounded by a pseudorandom measure.

This turns out to be exactly what we need since we want f to have support in the primes, which does not have positive density. We now need to construct such a function and a pseudorandom measure that bounds it. We will do this in the following way.

Let

$$w = w(N) = \log \log N$$

and let

$$W = W(N) = \prod_{p \leq w} p.$$

Now define a function $\tilde{\Lambda} : \mathbb{N} \rightarrow \mathbb{R}^+$ by

$$\tilde{\Lambda}(n) = \begin{cases} \frac{\phi(W)}{W} \log(Wn+1) & \text{if } Wn+1 \text{ is a prime} \\ 0 & \text{otherwise} \end{cases}$$

where ϕ is Euler's ϕ function. Now define $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} \frac{1}{k2^{k+5}} \tilde{\Lambda}(x) & \text{if } \varepsilon N \leq x \leq 2\varepsilon N \\ 0 & \text{otherwise} \end{cases}$$

In chapter 8 we will prove the following proposition.

Proposition 3.1.3. *Let $\varepsilon_k = \frac{1}{2^k(k+4)!}$ and let N be a sufficiently large prime number. Then there exists a k -pseudorandom measure $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ such that*

$$0 \leq f \leq \nu(x)$$

for all $x \in \mathbb{Z}_N$.

3.2 The proof

proposition 3.1.3 states the existence of a function with support in the primes (actually in the primes of the form $Wx+1$) which is bounded by a k -pseudorandom measure and theorem 3.1.2 uses this function to provide us with arithmetic progressions in the primes. Using these we can actually conclude the proof of the Green-Tao theorem in the following way.

Proof of theorem 2.1.1. Let $k \in \mathbb{N}$ be given. We want to show that there exists infinitely many arithmetic progressions of length k consisting of primes. Define $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ as we did above. For a sufficiently large N we get by proposition 3.1.3 that there exists a k -pseudorandom measure ν such that

$$0 \leq f(x) \leq \nu(x)$$

for all $x \in \mathbb{Z}_N$. In chapter 8 we will prove that

$$\sum_{n \leq N} \tilde{\Lambda}(n) = N(1 + o(1)),$$

This gives us

$$\mathbb{E}(f) = \frac{1}{k2^{k+5}}\varepsilon(1 + o(1)).$$

For sufficiently large k we have

$$\varepsilon \frac{1}{k2^{k+5}} \leq 1,$$

so if we let

$$\delta < \frac{1}{k2^{k+5}}\varepsilon,$$

then $\mathbb{E}(f) \geq \delta$ for sufficiently large N . theorem 3.1.2 now gives us

$$\mathbb{E}(f(x)f(x+r) \cdots f(x+(k-1)r) \mid x, r \in \mathbb{Z}_N) \geq c(k, \frac{A}{k2^{k+5}}\varepsilon) - o(1). \quad (3.2.1)$$

The RHS goes to $c(k, \frac{A}{k2^{k+5}}\varepsilon) > 0$, which does not depend on N , as N goes to infinity. The interior of the average on the LHS is positive whenever f is positive on all elements in a configuration

$$\{x, x+r, \dots, x+(k-1)r\}$$

but the LHS is greater than $c(k, \frac{1}{k2^{k+5}}\varepsilon) > 0$, which does not depend on N , so letting $N \rightarrow \infty$ we see that there must be infinitely many such configurations such that f is positive on all elements.

Now notice that

$$f(x) > 0 \iff Wx + 1 \text{ is prime and } \varepsilon N \leq x \leq 2\varepsilon N,$$

if $x+ir, i = 0, \dots, k-1$ is an arithmetic progression then so is $W(x+ir)+1, i = 0, \dots, k-1$ since $W(x+ir)+1 = (Wx+1)+i(Wr)$, so each of the configurations gives us an arithmetic progression consisting of primes by the definition of f .

There are two problems remaining. First of all the configuration is not an arithmetic progression of length k if $r = 0$, and secondly we can be in the situation that the progression is a progression in \mathbb{Z}_N but not in \mathbb{Z} , which will happen if the progression at some point exceeds N .

Let us consider the first problem. The contribution to the average on the LHS when $r = 0$ is

$$\frac{1}{N^2} \sum_{x \in \mathbb{Z}_N} f(x)^k.$$

And since $f(x) = O(\log N)$ this is

$$O\left(\frac{(\log N)^k}{N}\right) = o(1)$$

so we can ignore the contribution of $r = 0$.

Now we need to prove that the arithmetic progressions we find in \mathbb{Z}_N also are progressions in \mathbb{Z} . Recall that in the configurations $\{x, x+r, \dots, x+(k-1)r\}$ we found, we have $\varepsilon N \leq x+ir \leq 2\varepsilon N$ for all $i = 0, \dots, k-1$ because f has its support in this interval.

If the arithmetic progression exceeds N we must have $r \geq (1-\varepsilon)N$ so the arithmetic progression must wrap around on every step. This will also give us an arithmetic progression – it will be one we have counted already but that does not matter since we have infinitely many. \square

3.3 Overview of the rest of the proof

The rest of the proof is rather technical and to help the reader through it I try to give a basic overview of what we will be doing.

In Chapter 4 we will introduce the Gowers uniformity norm on functions $\mathbb{Z}_N \rightarrow \mathbb{R}$ and use this to prove a theorem we will call a new von Neumann theorem. This name is inspired by the similarity with von Neumann's mean ergodic theorem.

In Chapter 5 and 6, dual functions and σ -algebras on \mathbb{Z}_N will be introduced. We will need these in Chapter 7 where we will use a technique inspired by Furstenberg to prove a theorem that together with the new von Neumann theorem and Szemerédi's theorem will be used to prove Szemerédi's theorem in pseudorandom measures.

As seen above this new version of Szemerédi's theorem is used in the proof of the Green-Tao theorem. We also need a specific pseudorandom measure, that has its support in the primes. Chapter 8 will be used to construct this measure and prove that it actually is pseudorandom.

Chapter 4

Gowers uniformity and von Neumanns theorem

In Gowers' proof of Szemerédi's theorem [4] Gowers uses a theorem (theorem 3.2 in [4]) that gives a bound on expressions of the form

$$\mathbb{E} (f_1(x+r)f_2(x+2r)\cdots f_k(x+kr) \mid x, r \in \mathbb{Z}_N)$$

where the f_i 's are bounded in a certain way. Expressions of this form are of course relevant because they are counting arithmetic progressions. We will need a generalisation of this, namely theorem 4.3.1, which in [6] is called a generalized von Neumann theorem. The difference from Gowers' proof is the introduction of pseudorandom measures and the Gowers uniformity norm, which gives the theorem a more measure-theoretical flavour, as in the ergodic theory proof of Szemerédi's theorem [3].

First we will need some notation. We will define a Gowers inner product and a Gowers norm and then prove some properties for these.

4.1 The Gowers uniformity norm

Definition 4.1.1 (Gowers inner product). Let $d \geq 1$ be a dimension. For $\omega = (\omega_1, \dots, \omega_d) \in \{0, 1\}^d$ and $h = (h_1, \dots, h_d) \in \mathbb{Z}_N^d$ we define

$$\omega \cdot h = \omega_1 h_1 + \cdots + \omega_d h_d.$$

Let $(f_\omega)_{\omega \in \{0,1\}^d}$ be a $\{0, 1\}^d$ -tuple of functions $\mathbb{Z}_N \rightarrow \mathbb{R}$. Then we define the *d-dimensional Gowers inner product* of these functions by

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d} = \mathbb{E} \left(\prod_{\omega \in \{0,1\}^d} f_\omega(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right).$$

From the Gowers inner product, we define the *Gowers uniformity norm*:

Definition 4.1.2 (Gowers uniformity norm). Let $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ and $d \geq 1$, then we define the *Gowers uniformity norm of f* as

$$\|f\|_{U^d} = \langle (f)_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/2^d}.$$

We state some properties of this norm in the following proposition. Liu [9] handles the case where the f_ω 's are complex but we will only need the case where the f_ω 's are real.

Proposition 4.1.3 (Properties of the Gowers uniformity norm). *Let $(f_\omega)_{\omega \in \{0,1\}^d}$ be an $\{0,1\}^d$ -tuple of functions $\mathbb{Z}_N \rightarrow \mathbb{R}$ and let $f, g : \mathbb{Z}_N \rightarrow \mathbb{R}$.*

1. *The norm is positive:*

$$\|f\|_{U^d} \geq 0.$$

2. *Cauchy-Schwarz:*

$$|\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d}| \leq \prod_{\omega \in \{0,1\}^d} \|f_\omega\|_{U^d}.$$

3. *Triangle inequality:*

$$\|f + g\|_{U^d} \leq \|f\|_{U^d} + \|g\|_{U^d}.$$

4. *Increasing by dimension:*

$$\|f\|_{U^{d-1}} \leq \|f\|_{U^d}.$$

5. *For $d = 1$, the norm is the absolute expected value:*

$$\|f\|_{U^1} = |\mathbb{E}(f)|.$$

6. *The Gowers uniformity norm is a norm for $d \geq 2$.*

Proof. 1. First assume that $(f_\omega)_{\omega \in \{0,1\}^d}$ is a sequence of functions such that f_ω is independent of the last digit, ω_d of ω . So for $\omega = \omega_1, \dots, \omega_d$ we have

$$f_{\omega_1, \dots, \omega_{d-1}, 0} = f_{\omega_1, \dots, \omega_{d-1}, 1},$$

and this function can thus just be denoted $f_{\omega'}$ where $\omega' = \omega_1, \dots, \omega_{d-1}$. We will use this notation throughout the proof of this proposition. Note that this certainly is the case when $f_\omega = f$ for all ω . We now have

$$\begin{aligned} \langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d} &= \mathbb{E} \left(\prod_{\omega \in \{0,1\}^d} f_\omega(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) \\ &= \mathbb{E} \left(\prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(x + \omega' \cdot h') f_{\omega'}(x + h_d + \omega' \cdot h') \mid x \in \mathbb{Z}_N, h' \in \mathbb{Z}_N^{d-1}, h_d \in \mathbb{Z}_N \right). \end{aligned}$$

Now this can be rewritten as

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d} = \mathbb{E} \left(\mathbb{E} \left(\prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega'}(y + \omega' \cdot h') \mid y \in \mathbb{Z}_N \right)^2 \mid h' \in \mathbb{Z}_N^{d-1} \right). \quad (4.1.1)$$

This might be a bit hard to see, but if we consider the expected values as sums we see that they have the same number of terms, and that the terms are pairwise equal. We see that this last expression clearly is positive, since we are taking expected value over positive numbers, so we have

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d} \geq 0,$$

when f_ω is independent of the last digit. As mentioned this is certainly the case when $f_\omega = f$ for all ω , so the norm is positive.

2. Let us now consider the general case, where f_ω depends on the last digit. Then we can do the same trick as before to get

$$\begin{aligned} \langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d} &= \mathbb{E} \left(\prod_{\omega \in \{0,1\}^d} f_\omega(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) \\ &= \mathbb{E} \left(\prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega',0}(x + \omega' \cdot h') f_{\omega',1}(x + h_d + \omega' \cdot h') \mid x \in \mathbb{Z}_N, h' \in \mathbb{Z}_N^{d-1}, h_d \in \mathbb{Z}_N \right), \end{aligned}$$

and doing as before we rewrite this as

$$\begin{aligned} \langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d} &= \mathbb{E} \left(\mathbb{E} \left(\prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega',0}(y + \omega' \cdot h') \mid y \in \mathbb{Z}_N \right) \right. \\ &\quad \left. \times \mathbb{E} \left(\prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega',1}(y + \omega' \cdot h') \mid y \in \mathbb{Z}_N \right) \mid h' \in \mathbb{Z}_N^{d-1} \right), \end{aligned}$$

and hence by using Cauchy-Schwarz (corollary 2.5.4) we get

$$\begin{aligned} |\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d}| &\leq \mathbb{E} \left(\mathbb{E} \left(\prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega',0}(y + \omega' \cdot h') \mid y \in \mathbb{Z}_N \right)^2 \mid h' \in \mathbb{Z}_N^{d-1} \right)^{1/2} \\ &\quad \times \mathbb{E} \left(\mathbb{E} \left(\prod_{\omega' \in \{0,1\}^{d-1}} f_{\omega',1}(y + \omega' \cdot h') \mid y \in \mathbb{Z}_N \right)^2 \mid h' \in \mathbb{Z}_N^{d-1} \right)^{1/2}, \end{aligned}$$

which by (4.1.1) is

$$|\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d}| \leq \langle (f_{\omega',0})_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/2} \langle (f_{\omega',1})_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/2} \quad (4.1.2)$$

because $f_{\omega',0}$ and $f_{\omega',1}$ depends only on $d-1$ digits. We could repeat this argument with an arbitrary digit instead of the last one to obtain

$$|\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d}| \leq \langle (f_{\omega_1, \dots, \omega_{i-1}, 0, \omega_{i+1}, \dots, \omega_d})_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/2} \langle (f_{\omega_1, \dots, \omega_{i-1}, 1, \omega_{i+1}, \dots, \omega_d})_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/2}$$

for $1 \leq i \leq d$. We can now take the two factors from the RHS of (4.1.2), which are both non-negative, and use the above formula with $i = d-1$ on both of them to get

$$|\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d}| \leq \langle (f_{\omega'',0,0})_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/4} \langle (f_{\omega'',0,1})_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/4} \\ \langle (f_{\omega'',1,0})_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/4} \langle (f_{\omega'',1,1})_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/4}$$

where $\omega'' = \omega_1, \dots, \omega_{d-2}$. So repeating this for all the d digits we get

$$|\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d}| \leq \prod_{\tau \in \{0,1\}^d} \langle (f_\tau)_{\omega \in \{0,1\}^d} \rangle_{U^d}^{1/2^d} = \prod_{\tau \in \{0,1\}^d} \|f_\tau\|_{U^d}$$

by the definition of the Gowers uniformity norm.

3. Let $f, g : \mathbb{Z}_N \rightarrow \mathbb{R}$ and $d \geq 1$. We now have

$$\|f + g\|_{U^d}^{2^d} = \langle (f + g)_{\omega \in \{0,1\}^d} \rangle_{U^d} \\ = \mathbb{E} \left(\prod_{\omega \in \{0,1\}^d} (f + g)(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) \\ = \mathbb{E} \left(\sum_{A \subseteq \{0,1\}^d} \prod_{\omega \in A} f(x + \omega \cdot h) \prod_{\omega \notin A} g(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) \\ = \sum_{A \subseteq \{0,1\}^d} \mathbb{E} \left(\prod_{\omega \in A} f(x + \omega \cdot h) \prod_{\omega \notin A} g(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right)$$

Now let $A \subseteq \{0,1\}^d$ be fixed and define F_ω as

$$F_\omega = \begin{cases} f, & \text{if } \omega \in A, \\ g, & \text{if } \omega \notin A, \end{cases}$$

then we have

$$\mathbb{E} \left(\prod_{\omega \in A} f(x + \omega \cdot h) \prod_{\omega \notin A} g(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) = \langle (F_\omega)_{\omega \in \{0,1\}^d} \rangle$$

and by the Gowers uniformity Cauchy-Schwarz we have

$$\langle (F_\omega)_{\omega \in \{0,1\}^d} \rangle \leq \prod_{\omega \in \{0,1\}^d} \|F_\omega\|_{U^d},$$

which by the definition of F_ω is equal to $\|f\|_{U^d}^{|A|} \|g\|_{U^d}^{2^d-|A|}$. Thus

$$\begin{aligned} \|f + g\|_{U^d}^{2^d} &\leq \sum_{A \subseteq \{0,1\}^d} \|f\|_{U^d}^{|A|} \|g\|_{U^d}^{2^d-|A|} \\ &= \sum_{k=0}^{2^d} \sum_{|A|=k} \|f\|_{U^d}^{|A|} \|g\|_{U^d}^{2^d-|A|} \\ &= \sum_{k=0}^{2^d} \binom{2^d}{k} \|f\|_{U^d}^k \|g\|_{U^d}^{2^d-k} \\ &= (\|f\|_{U^d} + \|g\|_{U^d})^{2^d} \end{aligned}$$

as desired, where we in the last equality use the binomial formula.

4. Let $d \geq 2$ and let $f : \mathbb{Z}_N \rightarrow \mathbb{R}$. We see that

$$\|1\|_{U^d} = 1.$$

For each $\omega \in \{0,1\}^d$ we define

$$f_\omega = \begin{cases} 1, & \text{if } \omega_d = 1 \\ f, & \text{if } \omega_d = 0 \end{cases}.$$

By using the Cauchy-Schwarz inequality on this sequence we now see that

$$|\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d}| \leq \prod_{\omega \in \{0,1\}^d} \|f_\omega\|_{U^d} = \|f\|_{U^d}^{2^{d-1}}, \quad (4.1.3)$$

since exactly 2^{d-1} of the ω 's have a 0 in the last digit. But we also have

$$\begin{aligned} |\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d}| &= \mathbb{E} \left(\prod_{\omega \in \{0,1\}^d} f_\omega(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) \\ &= \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^d \\ \omega_d = 0}} f(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) \\ &= \mathbb{E} \left(\prod_{\omega' \in \{0,1\}^{d-1}} f(x + \omega' \cdot h') \mid x \in \mathbb{Z}_N, h' \in \mathbb{Z}_N^{d-1} \right), \end{aligned}$$

and by the definition of the Gowers uniformity norm this is equal to $\|f\|_{U^{d-1}}^{2^{d-1}}$. Combining this with (4.1.3) now yields

$$\|f\|_{U^{d-1}} \leq \|f\|_{U^d}$$

as claimed.

5. Now let $d = 1$ and let $f : \mathbb{Z}_N \rightarrow \mathbb{R}$. Then

$$\begin{aligned} \|f\|_{U^1} &= \mathbb{E} \left(\prod_{\omega \in \{0,1\}} f(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N \right)^{1/2} \\ &= \mathbb{E} (f(x)f(x+h) \mid x, h \in \mathbb{Z}_N)^{1/2} \\ &= |\mathbb{E}(f)|, \end{aligned}$$

where we in the third equality use proposition 2.2.5 and a change of variables.

6. By the definition of a norm we need to show that for $f, g : \mathbb{Z}_N \rightarrow \mathbb{R}$ and $a \in \mathbb{R}$ we have

$$(i) \quad \|af\|_{U^d} = |a| \|f\|_{U^d},$$

$$(ii) \quad \|f+g\|_{U^d} \leq \|f\|_{U^d} + \|g\|_{U^d},$$

$$(iii) \quad \|f\|_{U^d} = 0 \Rightarrow f = 0.$$

We see that (ii) is proved already in 3. of this proposition. We can prove (i) by straight forward calculation:

$$\begin{aligned} \|af\|_{U^d} &= \mathbb{E} \left(\prod_{\omega \in \{0,1\}^d} af(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right)^{1/2^d} \\ &= \left(a^{2^d} \mathbb{E} \left(\prod_{\omega \in \{0,1\}^d} af(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d \right) \right)^{1/2^d} \\ &= |a| \|f\|_{U^d}. \end{aligned}$$

We now want to prove (iii). Notice that by 5. we see that it actually is a necessary condition to require $d \geq 2$ since when $d = 1$ we have $\|f\|_{U^1} = |\mathbb{E}(f)|$, so it is possible to have $\|f\|_{U^1} = 0$ for non-zero f .

By 4. it is enough to prove (iii) for $d = 2$ since

$$\|f\|_{U^d} = 0 \Rightarrow \|f\|_{U^2} = 0$$

for $d > 2$ since the norm is non-negative, so let $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ be given and assume that $\|f\|_{U^2} = 0$. Now define a sequence $(f_\omega)_{\omega \in \{0,1\}^2}$ by $f_{00} = f$ and $f_{01} = f_{10} = f_{11} = \mathbb{1}_{\{x_0\}}$ for some given $x_0 \in \mathbb{Z}_N$. By the Gowers Cauchy-Schwarz we now have

$$\begin{aligned} \left| \mathbb{E} (f(x) \mathbb{1}_{\{x_0\}}(x+h_1) \mathbb{1}_{\{x_0\}}(x+h_2) \mathbb{1}_{\{x_0\}}(x+h_1+h_2) \mid x, h_1, h_2 \in \mathbb{Z}_N) \right| \\ \leq \|f\|_{U^2} \|\mathbb{1}_{\{x_0\}}\|_{U^2}^3. \end{aligned} \quad (4.1.4)$$

The LHS of (4.1.4) can be computed by noticing that all the 3 indicator functions have to be 1 in order to contribute to the expected value, and this only happens once, namely when $h_1 = h_2 = 0$ and $x = x_0$, so the LHS of (4.1.4) is equal to $\frac{f(x_0)}{N^3}$.

Now consider the RHS of (4.1.4). We have

$$\|\mathbb{1}_{\{x_0\}}\|_{U^d} = \mathbb{E} \left(\mathbb{1}_{\{x_0\}}(x) \mathbb{1}_{\{x_0\}}(x + h_1) \mathbb{1}_{\{x_0\}}(x + h_2) \mathbb{1}_{\{x_0\}}(x + h_1 + h_2) \mid x, h_1, h_2 \in \mathbb{Z}_N \right)^{1/4},$$

and once again they all have to be 1 in order to contribute, which only happens when $h_1 = h_2 = 0$ and $x = x_0$, so

$$\|\mathbb{1}_{\{x_0\}}\|_{U^d} = \left(\frac{1}{N^3} \right)^{1/4},$$

and we can therefore rewrite (4.1.4) to

$$\frac{|f(x_0)|}{N^3} \leq \left(\frac{1}{N^3} \right)^{3/4} \|f\|_{U^d}$$

which implies $f(x_0) = 0$ since we assumed $\|f\|_{U^d} = 0$. This does not depend on our choice of $x_0 \in \mathbb{Z}_N$, so it is valid for all $x_0 \in \mathbb{Z}_N$ and hence $f = 0$. \square

In our new version of Szemerédi's theorem we want to replace 1 with a pseudorandom measure ν in the original Szemerédi's theorem. The following lemma shows that ν and 1 are not far from each other in the U^d -norm.

Lemma 4.1.4. *Let ν be k -pseudorandom. Then we have*

$$\|\nu - 1\|_{U^d} = o(1)$$

for all $1 \leq d \leq k - 1$.

Proof. By item 4. in proposition 4.1.3 we see that it is sufficient to prove the lemma for $d = k - 1$, since we have

$$\|\nu - 1\|_{U^{d'}} \leq \|\nu - 1\|_{U^d}$$

for all $d' \leq d$. By the definition of the U^d norm, we thus need to show that

$$\mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} (\nu(x + \omega \cdot h) - 1) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) = o(1).$$

The product in the expected value on the LHS can be expanded to obtain

$$\sum_{A \subseteq \{0,1\}^{k-1}} (-1)^{|A|} \mathbb{E} \left(\prod_{\omega \in A} \nu(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right).$$

So now consider one of the terms (ignoring the sign for a moment) of this expression,

$$\mathbb{E} \left(\prod_{\omega \in A} \nu(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \quad (4.1.5)$$

for some $A \subseteq \{0, 1\}^{k-1}$. Now number the elements of A and define for $1 \leq i \leq |A|$ the linear forms $\phi_i : \mathbb{Z}_N^k \rightarrow \mathbb{Z}_N$ by

$$\phi_i(y) = y_1 + \omega \cdot (y_2, \dots, y_k)$$

where ω is the i 'th element of A and $y = (y_1, \dots, y_k)$. Then (4.1.5) can be written as

$$\mathbb{E} (\nu(\phi_1(y)) \cdots \nu(\phi_{|A|}(y)) \mid y \in \mathbb{Z}_N^k).$$

Now we want to use the $(2^{k-1}, k, 1)$ -linear forms condition, which is valid because ν is k -pseudorandom, but to use this we need to prove that none of the linear forms is a multiple of any other. So assume that $\phi_i = a\phi_j$ for some $a \in \mathbb{Z}_N \setminus \{0\}$ and for some $1 \leq i, j \leq |A|$, $i \neq j$, and let ω and τ be the corresponding members of A . Now let $y = (0, 1, 0, \dots, 0)$. Then

$$\omega_1 = \phi_i(y) = a\phi_j(y) = a\tau_1.$$

This can be done for all coordinates, so we have $\omega_k = a\tau_k$ for all $1 \leq k \leq |A|$. Since the ω_k 's and τ_k 's can be 0 or 1, and $a \neq 0$ we have only a few possibilities for the choice of a . Either $a = 1$ and we have $\phi_i = \phi_j$ and hence $i = j$ which is a contradiction, or $a \neq 1$ and then we must have $\omega = \tau = 0$ and hence $i = j$ which again is a contradiction.

So we can now use the $(2^{k-1}, k, 1)$ -linear forms condition which states that

$$\mathbb{E} (\nu(\phi_1(y)) \cdots \nu(\phi_{|A|}(y)) \mid y \in \mathbb{Z}_N^k) = 1 + o(1),$$

so

$$\sum_{A \subseteq \{0, 1\}^{k-1}} (-1)^{|A|} \mathbb{E} \left(\prod_{\omega \in A} \nu(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) = \sum_{A \subseteq \{0, 1\}^{k-1}} (-1)^{|A|} + o(1),$$

because the size of A only depends on k , and $o(1)$ is allowed to depend on k . But there are as many subsets of A with an even number of elements as there are subsets with an odd number of elements, so the above expression is just $o(1)$ as we wanted. \square

4.2 A special Cauchy-Schwarz

In the proofs of this chapter we will need the following definition, where we from two vectors y, y' define a new vector $y^{(S)}$ where some of the coordinates are from y' and the rest from y and where S is an index set that indicates which coordinates come from y' .

Definition 4.2.1. Let $k > 1$ and $0 \leq d \leq k$, and let $y = (y_1, \dots, y_{k-1}) \in \mathbb{Z}_N^{k-1}$ and $y' = (y'_{k-d}, \dots, y'_{k-1}) \in \mathbb{Z}_N^d$ and $S \subseteq \{k-d, \dots, k-1\}$. Then we define the vector $y^{(S)} = (y_1^{(S)}, \dots, y_{k-1}^{(S)}) \in \mathbb{Z}_N^{k-1}$ by

$$y_i^{(S)} = \begin{cases} y'_i & \text{if } i \in S \\ y_i & \text{otherwise.} \end{cases}$$

The following lemma is again a kind of Cauchy-Schwarz, and this one we will use to prove theorem 4.3.1.

Lemma 4.2.2 (Yet another Cauchy-Schwarz). *Let $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ be a measure and $k > 1$. Let $\phi_0, \phi_1, \dots, \phi_{k-1} : \mathbb{Z}_N^{k-1} \rightarrow \mathbb{Z}_N$ such that we for all i have that ϕ_i does not depend on the i -th coordinate. Let $f_0, \dots, f_{k-1} : \mathbb{Z}_N \rightarrow \mathbb{R}$ such that*

$$|f_i(x)| \leq \nu(x)$$

for all $x \in \mathbb{Z}_N$ and $1 \leq i \leq k-1$. Define for $0 \leq d \leq k-1$

$$J_d = \mathbb{E} \left(\prod_{S \subseteq \{k-d, \dots, k-1\}} \left(\prod_{i=0}^{k-d-1} f_i(\phi_i(y^{(S)})) \right) \left(\prod_{i=k-d}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \right) \middle| y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d \right)$$

and

$$P_d = \mathbb{E} \left(\prod_{S \subseteq \{k-d, \dots, k-1\}} \nu(\phi_{k-d-1}(y^{(S)})) \middle| y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d \right).$$

Then we have for all $0 \leq d \leq k-2$ that

$$|J_d|^2 \leq P_d J_{d+1}.$$

Proof. Consider J_d . Since ϕ_{k-d-1} does not depend on the $k-d-1$ 'th coordinate, we may use proposition 2.2.5 to split the average into two parts – one depending on ϕ_{k-d-1} and one not depending on ϕ_{k-d-1} . So we can write

$$J_d = \mathbb{E} \left(G(y, y') H(y, y') \middle| y_1, \dots, y_{k-d-2}, y_{k-d}, \dots, y_{k-1} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d \right),$$

where

$$G(y, y') = \prod_{S \subseteq \{k-d, \dots, k-1\}} f_{k-d-1}(\phi_{k-d-1}(y^{(S)})) \nu^{-1/2}(\phi_{k-d-1}(y^{(S)}))$$

and

$$H(y, y') = \mathbb{E} \left(\prod_{S \subseteq \{k-d, \dots, k-1\}} \prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)})) \prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \middle| y_{k-d-1} \in \mathbb{Z}_N \right).$$

Notice that we have added a factor $\nu^{1/2}(\phi_{k-d-1}(y^{(S)}))$ to H and divided by it in G . By Cauchy-Schwarz (corollary 2.5.4) we have

$$|J_d|^2 \leq \mathbb{E} \left(G(y, y')^2 \mid y_1, \dots, y_{k-d-2}, y_{k-d}, \dots, y_{k-1} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d \right) \\ \cdot \mathbb{E} \left(H(y, y')^2 \mid y_1, \dots, y_{k-d-2}, y_{k-d}, \dots, y_{k-1} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d \right). \quad (4.2.1)$$

Consider the first factor. Since $f_{k-d-1}(x) \leq \nu(x)$ for all $x \in \mathbb{Z}_N$ we have that

$$G(y, y')^2 = \left(\prod_{S \subseteq \{k-d, \dots, k-1\}} f_{k-d-1}(\phi_{k-d-1}(y^{(S)})) \nu^{-1/2}(\phi_{k-d-1}(y^{(S)})) \right)^2 \\ \leq \prod_{S \subseteq \{k-d, \dots, k-1\}} \nu(\phi_{k-d-1}(y^{(S)})),$$

which exactly is the factors in P_d . Now notice that by 3. in corollary 2.2.5 it makes no difference if we in the first factor on the RHS of (4.2.1) take average over all variables or omit y_{k-d-1} since $G(y, y')$ does not depend on y_{k-d-1} . So we get

$$\mathbb{E} \left(G(y, y')^2 \mid y_1 \dots y_{k-d-2}, y_{k-d}, \dots, y_{k-1} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d \right) \leq P_d.$$

Now consider the second factor on the RHS of (4.2.1). We want to prove that this is equal to J_{d+1} . With the d increased by 1, the y' vector will be one dimension bigger, so we need to add an extra variable y'_{k-d-1} . We also have to take products over more subsets S and alter the two inner products. All in all we have to show that

$$H(y, y')^2 = \mathbb{E} \left(\prod_{S \subseteq \{k-d-1, \dots, k-1\}} \left(\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)})) \right) \right. \\ \left. \left(\prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \right) \mid y_{k-d-1}, y'_{k-d-1} \in \mathbb{Z}_N \right). \quad (4.2.2)$$

Now

$$H(y, y')^2 = \left[\sum_{y_{k-d-1} \in \mathbb{Z}_N} \frac{1}{N} \prod_{S \subseteq \{k-d, \dots, k-1\}} \prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)})) \prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \right]^2 \\ = \sum_{y_{k-d-1}, y'_{k-d-1} \in \mathbb{Z}_N} \frac{1}{N^2} \prod_{S \subseteq \{k-d, \dots, k-1\}} \prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)})) \prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \\ \prod_{S \subseteq \{k-d, \dots, k-1\}} \prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S \cup \{k-d-1\})})) \prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S \cup \{k-d-1\})})) \\ = \mathbb{E} \left(\prod_{S \subseteq \{k-d-1, \dots, k-1\}} \left(\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)})) \right) \left(\prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \right) \mid y_{k-d-1}, y'_{k-d-1} \in \mathbb{Z}_N \right)$$

The second and third equality are a little bit tricky. Consider the second equality – since we are squaring the sum ranging over $y_{k-d-1} \in \mathbb{Z}_N$, we get N^2 pairs, represented by the two variables, denote them y_{k-d-1} and y'_{k-d-1} , both ranging over \mathbb{Z}_N . So in the y'_{k-d-1} part of the pair, y_{k-d-1} should be replaced by y'_{k-d-1} , which can be done by adding the element $k-d-1$ to the set S .

Now consider the third equality. We now see that the expression can be greatly simplified, since we in the first component are taking product over all $S \subseteq \{k-d, \dots, k-1\}$, and in the second we are taking product over the same subsets with $k-d-1$, so we are actually considering all $S \subseteq \{k-d-1, \dots, k-1\}$, and we can thus rewrite the whole thing as (4.2.2) and then return to the expected value notation. This finishes the proof. \square

Corollary 4.2.3. *With J_d and P_d defined as before we have for all $k > 1$*

$$|J_0|^{2^{k-1}} \leq J_{k-1} \prod_{d=0}^{k-2} |P_d|^{2^{k-2-d}}.$$

Proof. From the lemma we have

$$|J_{k-2}|^2 \leq P_{k-1} J_{k-1}$$

and

$$|J_{k-3}|^2 \leq P_{k-2} J_{k-2},$$

and putting this together we get

$$|J_{k-3}|^{2^2} \leq |P_{k-2}|^2 |J_{k-2}|^2 \leq J_{k-1} P_{k-1} P_{k-2}^2.$$

Doing this $k-2$ times gives the desired result. \square

4.3 Von Neumann's theorem

We are now ready to begin the proof of the main theorem of this chapter. In the ergodic theoretic proof of Szemerédi's theorem a version of von Neumann's ergodic mean theorem (theorem 3.1 in [3]) is used. The proof of Green and Tao is inspired by the ergodic theoretical proof of Szemerédi's theorem, and so they need a version of von Neumann's theorem in their setting.

The original von Neumann's ergodic mean theorem (see for instance [14]) works in quite a different setting than ours, but we still have an average converging to a constant.

Theorem 4.3.1 (Von Neumann in pseudorandom measures). *Let ν be a k -pseudorandom measure, and let $f_0, \dots, f_{k-1} : \mathbb{Z}_N \rightarrow \mathbb{R}$ be functions such that*

$$|f_j(x)| \leq \nu(x) + 1$$

for all $x \in \mathbb{Z}_N$ and $0 \leq j \leq k-1$. Let c_0, \dots, c_{k-1} be some permutation of some k consecutive elements of $\{-k+1, \dots, -1, 0, 1, \dots, k-1\}$. Then

$$\mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + c_j r) \mid x, r \in \mathbb{Z}_N \right) = O \left(\inf_{0 \leq j \leq k-1} \|f_j\|_{U^{k-1}} \right) + o(1).$$

Remark 4.3.2.

Proof. We can use lemma 2.3.7 to replace ν with $\nu' = (\nu + 1)/2$ in the theorem, and thus reduce $|f_j(x)| \leq \nu(x) + 1$ to $|f'_j(x)| \leq \nu'(x)$ where $f'_j = f_j/2$. By permuting the c_j and f_j we may also assume that

$$\inf_{1 \leq j \leq k-1} \|f_j\|_{U^{k-1}} = \|f_0\|_{U^{k-1}},$$

and by shifting x by $c_0 r$ we can assume that $c_0 = 0$.

To use the lemma we proved before we need to define some ϕ_i 's such that ϕ_i is independent of the i -th coordinate for all i . Define $\phi_0, \dots, \phi_{k-1} : \mathbb{Z}_N^{k-1} \rightarrow \mathbb{Z}_N$ by

$$\phi_i(y) = \sum_{j=1}^{k-1} \left(1 - \frac{c_i}{c_j} \right) y_j$$

for $i = 0, \dots, k-1$ where $y = (y_1, \dots, y_{k-1})$. Now notice that

$$\phi_0(y) = y_1 + \dots + y_{k-1}$$

and

$$\phi_i(y) = x(y) + c_i r(y)$$

for $i > 0$ where $x(y) = y_1 + \dots + y_{k-1}$ and

$$r(y) = - \sum_{j=0}^{k-1} \frac{y_j}{c_j}.$$

Notice also that $\phi_i(y)$ does not depend on y_i for all i since the term involving y_i will have the coefficient $(1 - c_i/c_i) = 0$ and hence cancel out, so we are in a situation where we can use lemma 4.2.2. Now define the J_d 's and P_d 's like in that lemma.

We now want to prove the following 4 statements.

1.

$$J_0 = \mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + c_j r) \mid x, r \in \mathbb{Z}_N \right),$$

2.

$$|J_0|^{2^{k-1}} \leq (1 + o(1)) J_{k-1},$$

3.

$$J_{k-1} = \mathbb{E} \left(W(x, h) \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right),$$

where

$$W(x, h) = \mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} \prod_{i=0}^{k-1} \nu^{1/2}(\phi_i(y + \omega h)) \mid y_1, \dots, y_{k-2} \in \mathbb{Z}_N \right)$$

and $y = (y_1, \dots, y_{k-2}, x - y_1 - \dots - y_{k-1})$.

4.

$$\mathbb{E} \left((W(x, h) - 1) \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) = o(1).$$

Let us first prove these 4 statements, and then afterwards show that this is enough to prove the theorem.

1. Define $\Phi : \mathbb{Z}_N^{k-1} \rightarrow \mathbb{Z}_N^2$ by

$$\Phi(y) = (x(y), -r(y)).$$

We now want to prove that Φ is a uniform cover (see definition 2.4.1). It is certainly surjective so we just need to prove that $\#\Phi^{-1}(z_1, z_2) = \#\mathbb{Z}_N^{k-1} / \#\mathbb{Z}_N^2 = N^{k-3}$ for all $z_1, z_2 \in \mathbb{Z}_N$. This can be done using linear algebra since $\Phi^{-1}(z_1, z_2)$ is the solution (y_1, \dots, y_{k-1}) to the following system of equations:

$$\begin{aligned} z_1 &= y_1 + \dots + y_{k-1} \\ z_2 &= c_1^{-1}y_1 + \dots + c_{k-1}^{-1}y_{k-1}. \end{aligned}$$

Since all the c_i 's are non-zero and different this systems has full rank and the solution-space will have dimension $k - 1 - 2 = k - 3$ and hence have N^{k-3} elements.

Define for $j = 0, \dots, k-1$ the functions $g_j : \mathbb{Z}_N^2 \rightarrow \mathbb{Z}_N$ by $g_j(x, r) = f_j(x + c_j r)$. Since Φ is a uniform cover we get

$$\mathbb{E} \left(\prod_{j=0}^{k-1} g_j(x, r) \mid x, r \in \mathbb{Z}_N \right) = \mathbb{E} \left(\prod_{j=0}^{k-1} g_j(\Phi(y)) \mid y \in \mathbb{Z}_N^{k-1} \right),$$

and using the definition of the g_j 's and Φ we then obtain

$$\mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + c_j r) \mid x, r \in \mathbb{Z}_N \right) = \mathbb{E} \left(\prod_{j=0}^{k-1} f_j(\phi_j(y)) \mid y \in \mathbb{Z}_N^{k-1} \right). \quad (4.3.1)$$

Now notice that by the definition of J_d we have

$$J_0 = \mathbb{E} \left(\prod_{i=0}^{k-1} f_i(\phi_i(y)) \mid y \in \mathbb{Z}_N^{k-1} \right).$$

Combining this with (4.3.1) we get

$$\mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + c_j r) \mid x, r \in \mathbb{Z}_N \right) = J_0,$$

which concludes **1**.

2. Recall that

$$P_d = \mathbb{E} \left(\prod_{S \subseteq \{k-d, \dots, k-1\}} \nu(\phi_{k-d-1}(y^{(S)})) \mid y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d \right)$$

for $d \leq k-1$. By assumption ν is k -pseudorandom, so it satisfies the $(k2^{k-1}, 3k-4, k)$ -linear forms condition, and by remark 2.3.3 it also satisfies the $(2^d, k-1+d, k)$ -linear forms condition for all $d \leq k-1$ since all the parameters are smaller. Now if $d \geq 1$ we enumerate the subsets $S \subseteq \{k-d, \dots, k-1\}$ with $1, 2, 3, \dots, 2^d$ and define $\psi_i : \mathbb{Z}_N^{k-1+d} \rightarrow \mathbb{Z}_N$ as

$$\psi_i(y, y') = \phi_{k-d-1}(y^{(S_i)})$$

for all i . If $d=0$ we say that S can only be the empty set. Using the $(2^d, k-1+d, k)$ -linear forms condition on the ψ_i 's gives us

$$P_d = \mathbb{E} \left(\prod_{S \subseteq \{k-d, \dots, k-1\}} \nu(\phi_{k-d-1}(y^{(S)})) \mid y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d \right) = 1 + o(1)$$

for all $0 \leq d \leq k-2$. corollary 4.2.3 now gives us

$$|J_0|^{2^{k-1}} \leq J_{k-1} \prod_{d=0}^{k-2} |P_d|^{2^{k-2-d}} = J_{k-1}(1 + o(1)),$$

since the $o(1)$ term is allowed to depend on k .

3. Recall that $W : \mathbb{Z}_N^2 \rightarrow \mathbb{R}$ is defined by

$$W(x, h) = \mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} \prod_{i=0}^{k-1} \nu^{1/2}(\phi_i(y + \omega h)) \mid y_1, \dots, y_{k-2} \in \mathbb{Z}_N \right)$$

where ωh is the coordinatewise product, which is the vector in \mathbb{Z}_N^{k-1} with $(\omega h)_i = \omega_i h_i$, and where $y = (y_1, \dots, y_{k-2}, x - y_1 - \dots - y_{k-2})$. Now let $y \in \mathbb{Z}_N^{k-1}$ be fixed and consider the expression

$$\begin{aligned} & \mathbb{E} \left(W(x, h) \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \in \mathbb{Z}_N \right) \\ &= \mathbb{E} \left(\mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} \prod_{i=0}^{k-1} \nu^{1/2}(\phi_i(y + \omega h)) \mid y_1, \dots, y_{k-2} \in \mathbb{Z}_N \right) \right. \\ & \quad \left. \cdot \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \end{aligned}$$

By 4. in proposition 2.2.5 we see that we can put the whole thing in one expected value. Now we rewrite the product to only take the product over $\omega \in \{0, 1\}^{k-1}$ once instead of twice to get

$$\mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} \left(f_0(x + \omega \cdot h) \prod_{i=0}^{k-1} \nu^{1/2}(\phi_i(y + \omega h)) \right) \mid x, y_1, \dots, y_{k-2} \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right). \quad (4.3.2)$$

We now want to prove the following claim.

Claim. We have

$$\{\phi_0(y^{(S)}) \mid S \subseteq \{1, \dots, k-1\}, y' \in \mathbb{Z}_N^{k-1}\} = \{x + \omega \cdot h \mid \omega \in \{0, 1\}^{k-1}, h \in \mathbb{Z}_N^{k-1}\}. \quad (4.3.3)$$

where $x = \phi_0(y)$.

Let $\phi_0(y^{(S)})$ be an element of the LHS. Now let $\omega \in \{0, 1\}^{k-1}$ be defined by $\omega_i = \mathbb{1}_S(i)$ and let $h \in \mathbb{Z}_N^{k-1}$ be defined by $h_i = y'_i - y_i$. Then

$$x + \omega \cdot h = y_1 + \dots + y_{k-1} + \mathbb{1}_S(i)(y'_i - y_i) = \sum_{i \in S} y'_i + \sum_{i \in S^c} y_i = \phi_0(y^{(S)}),$$

so $\phi_0(y^{(S)})$ is also an element in the RHS of (4.3.3) with the specified ω and h .

Now let $x + \omega \cdot h$ be an element in the RHS. Let $S \subseteq \{1, \dots, k-1\}$ be defined by

$$i \in S \iff \omega_i = 1$$

and let $y' \in \mathbb{Z}_N^{k-1}$ be defined by $y'_i = h_i + y_i$. Now the exact same calculations as above gives us that $x + \omega \cdot h = \phi_0(y^{(S)})$ which proves the claim.

In the same way we can prove that

$$\{y^{(S)} \mid S \subseteq \{1, \dots, k-1\}, y' \in \mathbb{Z}_N^{k-1}\} = \{y + \omega h \mid \omega \in \{0, 1\}^{k-1}, h \in \mathbb{Z}_N^{k-1}\}.$$

Due to these one-to-one correspondences we can now rewrite (4.3.2) as

$$\mathbb{E} \left(\prod_{S \subseteq \{1, \dots, k-1\}} \left(f_0(\phi_0(y^{(S)})) \prod_{i=0}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \right) \middle| x, y_1, \dots, y_{k-2} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^{k-1} \right)$$

which is equal to J_{k-1} .

4. Since we have assumed that $|f_0(x)| \leq \nu(x)$ for all $x \in \mathbb{Z}_N$ it is enough to prove that

$$\mathbb{E} \left((W(x, h) - 1) \prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) = o(1).$$

If we consider the LHS squared, we see by Cauchy-Schwarz (corollary 2.5.4) that it is smaller than

$$\begin{aligned} \mathbb{E} \left(|W(x, h) - 1|^2 \prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \\ \times \mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right), \end{aligned}$$

and it is hence enough to prove that this is $o(1)$. So if we can prove that

$$\mathbb{E} \left(|W(x, h) - 1|^2 \prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) = o(1) \quad (4.3.4)$$

and

$$\mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) = 1 + o(1), \quad (4.3.5)$$

we are done, since $(1 + o(1))o(1) = o(1)$.

Now consider the LHS of (4.3.4). Expanding out the square we get that this expression can be rewritten as

$$\begin{aligned} \mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \\ - 2 \mathbb{E} \left(W(x, h) \prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \\ + \mathbb{E} \left(W(x, h)^2 \prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right), \quad (4.3.6) \end{aligned}$$

so it is enough to prove that each of the three expected values is $1 + o(1)$ because then the above expression is $1 + o(1) - 2(1 + o(1)) + 1 + o(1) = o(1)$. Since the last term is equal to the LHS of (4.3.5), this also takes care of proving (4.3.5) as well.

Consider the first term in (4.3.6). Since ν is k -pseudorandom, it satisfies the $(k2^{k-1}, 3k-4, k)$ -linear forms condition and hence also the $(2^{k-1}, k, 1)$ -linear forms condition. Now define 2^{k-1} linear forms $\mathbb{Z}_N^k \rightarrow \mathbb{Z}_N$ (one for each $\omega \in \{0, 1\}^{k-1}$) by

$$(x, h_1, \dots, h_{k-1}) \mapsto x + \omega \cdot h,$$

where $h = (h_1, \dots, h_{k-1})$. Then the linear forms condition gives us that the first term in (4.3.6) is $1 + o(1)$ as desired.

Now consider the second expected value in (4.3.6). First we need to notice that we can write $W(x, h)$ in another way, namely as

$$W(x, h) = \mathbb{E} \left(\prod_{i=1}^{k-1} \prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega_i=0}} \nu(\phi_i(y + \omega h)) \mid y_1, \dots, y_{k-2} \in \mathbb{Z}_N \right), \quad (4.3.7)$$

because if we first interchange the two products, and then recall that ϕ_i is independent of the i -th variable we see that it does not matter if $\omega_i = 0$ or $\omega_i = 1$, so if we only take product over the ω 's with $\omega_i = 0$ we just need to take $\nu(\phi_i(y + \omega h))^{1/2}$ squared each time.

Once again we now use the fact that ν is k -pseudorandom, so in particular it satisfies the $(2^{k-2}(k+1), 2k-2, k)$ -linear form condition. We now want to define $2^{k-2}(k+1)$ linear forms and use the linear forms condition on these. The variables on which we want to define the forms are $x, h_1, \dots, h_{k-1}, y_1, \dots, y_{k-2}$. Define the first 2^{k-1} linear forms (one for each $\omega \in \{0, 1\}^{k-1}$) by

$$(x, h_1, \dots, h_{k-1}, y_1, \dots, y_{k-2}) \mapsto x + \omega \cdot h,$$

and define $(k-1)2^{k-2}$ linear forms (one for each $1 \leq i \leq k-1$ and each $\omega \in \{0, 1\}^{k-1}$ with $\omega_i = 0$) by

$$(x, h_1, \dots, h_{k-1}, y_1, \dots, y_{k-2}) \mapsto \phi_i(y + \omega h)$$

where $h = (h_1, \dots, h_{k-1})$, the ϕ_i 's are as defined in the beginning of the proof and $y = (y_1, \dots, y_{k-1})$ where $y_{k-1} = x - y_1 - \dots - y_{k-2}$. Note that these are linear forms because the ϕ_i 's are linear forms. This gives us $2^{k-1} + (k-1)2^{k-2} = (k+1)2^{k-1}$ linear forms in total, so now we just need to see that there linear forms actually is what we are looking for. Using (4.3.7) and 2. in proposition 2.2.5 we now get that

$$\begin{aligned} & \mathbb{E} \left(W(x, h) \prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \\ &= \mathbb{E} \left(\prod_{i=1}^{k-1} \prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega_i=0}} \nu(\phi_i(y + \omega h)) \prod_{\omega \in \{0,1\}^{k-1}} \nu(x + \omega \cdot h) \mid x, y_1, \dots, y_{k-2} \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right), \end{aligned}$$

so with the linear forms we picked above and using the linear forms condition we see that this is $1 + o(1)$ as we wanted.

Now consider the last term of (4.3.6). Here we will need the full magnitude of the parameters of the linear forms condition, namely the $(k2^{k-1}, 3k-4, k)$ -linear forms condition. First define $(k-1)2^{k-2}$ linear forms (one for each $1 \leq i \leq k-1$ and each $\omega \in \{0, 1\}^{k-1}$ with $\omega_i = 0$) by

$$(x, h_1, \dots, h_{k-1}, y_1, \dots, y_{k-2}, y'_1, \dots, y'_{k-2}) \mapsto \phi_i(y + \omega h).$$

where $h = (h_1, \dots, h_{k-1})$, the ϕ_i 's are as defined in the beginning of the proof and $y = (y_1, \dots, y_{k-1})$ where $y_{k-1} = x - y_1 - \dots - y_{k-2}$. Then define $(k-1)2^{k-2}$ linear forms (one for each $1 \leq i \leq k-1$ and each $\omega \in \{0, 1\}^{k-1}$ with $\omega_i = 0$) by

$$(x, h_1, \dots, h_{k-1}, y_1, \dots, y_{k-2}, y'_1, \dots, y'_{k-2}) \mapsto \phi_i(y' + \omega h)$$

where $y' = (y'_1, \dots, y'_{k-1})$ and $y'_{k-1} = x - y'_1 - \dots - y'_{k-2}$. Finally define 2^{k-1} linear forms (one for each $\omega \in \{0, 1\}^{k-1}$) by

$$(x, h_1, \dots, h_{k-1}, y_1, \dots, y_{k-2}, y'_1, \dots, y'_{k-2}) \mapsto x + \omega \cdot h.$$

Now we have $(k-1)2^{k-2} + (k-1)2^{k-2} + 2^{k-1} = k2^{k-1}$ linear forms and these are actually exactly the ones we needed. The term we are considering contains $W(x, h)^2$, so let us consider this for a moment. Using (4.3.7) and proposition 2.2.5, 5., we get that

$$W(x, h)^2 = \mathbb{E} \left(\prod_{i=1}^{k-1} \prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega_i=0}} \nu(\phi_i(y + \omega h)) \prod_{j=1}^{k-1} \prod_{\substack{\omega' \in \{0,1\}^{k-1} \\ \omega'_i=0}} \nu(\phi_j(y' + \omega' h)) \right. \\ \left. \left| y_1, \dots, y_{k-2}, y'_1, \dots, y'_{k-2} \in \mathbb{Z}_N \right. \right).$$

Using this and 2. in proposition 2.2.5 to put the whole thing under one expected value we get that the third term of (4.3.6) is the expected value over the product of the linear forms we have defined, so using the linear forms condition we get that this is $1 + o(1)$, which concludes the proof of 4.

Now we have proven the 4 claims, and all we need to do now to finish the proof is to show that these 4 statements together implies the generalised von Neumann theorem. Recall that we need to prove that

$$\mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + c_j r) \left| x, r \in \mathbb{Z}_N \right. \right) = O \left(\inf_{0 \leq j \leq k-1} \|f_j\|_{U^{k-1}} \right) + o(1).$$

By 1. we get that the LHS is equal to J_0 so by 2. we get

$$\mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + c_j r) \mid x, r \in \mathbb{Z}_N \right)^{2^{k-1}} = J_0^{2^{k-1}} \leq J_{k-1} + o(1). \quad (4.3.8)$$

Now by the definition of the U^{k-1} -norm we have that

$$\|f_0\|_{U^{k-1}}^{2^{k-1}} = \mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) \mid x, r \in \mathbb{Z}_N \right).$$

so using 3. and 4. we get that

$$\begin{aligned} o(1) &= \mathbb{E} \left((W(x, h) - 1) \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \\ &= \mathbb{E} \left(W(x, h) \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) - \|f_0\|_{U^{k-1}}^{2^{k-1}} \\ &= J_{k-1} - \|f_0\|_{U^{k-1}}^{2^{k-1}}. \end{aligned}$$

Inserting this in (4.3.8) gives us

$$\mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + c_j r) \mid x, r \in \mathbb{Z}_N \right)^{2^{k-1}} \leq \|f_0\|_{U^{k-1}}^{2^{k-1}} + o(1), \quad (4.3.9)$$

but recall that we in the beginning of the proof permuted the c_j and f_j such that

$$\|f_0\|_{U^{k-1}} = \inf_{1 \leq j \leq k-1} \|f_j\|_{U^{k-1}},$$

so taking 2^{k-1} 'th roots on both sides of (4.3.9) concludes the proof, because $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ so

$$\left| \mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + c_j r) \mid x, r \in \mathbb{Z}_N \right) \right| \leq \left(\|f_0\|_{U^{k-1}}^{2^{k-1}} + o(1) \right)^{1/2^{k-1}} \leq \|f_0\|_{U^{k-1}} + o(1).$$

This finishes the proof. \square

Chapter 5

Dual functions and basic Gowers anti-uniformity

The results and concepts of this and the following chapter will be used to prove proposition 7.1.2 which will then be used to prove Szemerédi's theorem in pseudorandom measures (theorem 3.1.2).

5.1 The dual space of U^d

When $d \geq 2$ we let U^d denote the normed vector space of the functions $\mathbb{Z}_N \rightarrow \mathbb{R}$ with norm $\|\cdot\|_{U^d}$. We showed that this is a norm in the preceding chapter when $d \geq 2$. Now let $(U^d)^*$ be the dual space which consists of linear mappings $U^d \rightarrow \mathbb{R}$. I now claim that each function $\phi \in (U^d)^*$ can be considered as a unique function $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ by the equality

$$\phi(g) = \mathbb{E}(fg)$$

for $g : \mathbb{Z}_N \rightarrow \mathbb{R}$. Note that this claim actually is the Riesz representation theorem in our setting, if we see $g \mapsto \mathbb{E}(fg)$ as an 'integral' with respect to a 'measure' f . See [11] for several formulations and proofs of the Riesz representation theorem.

Let us now prove the claim. Each $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ certainly gives us a $\phi \in (U^d)^*$ in this way since $g \mapsto \mathbb{E}(fg)$ is linear. Let $\phi \in (U^d)^*$ be given. Now define $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ by

$$f(i) = N\phi(e_i)$$

for $i \in \mathbb{Z}_N$ where $e_i : \mathbb{Z}_N \rightarrow \mathbb{R}$ is the function defined by

$$e_i(x) = \begin{cases} 1 & \text{if } x = i \\ 0 & \text{otherwise} \end{cases} .$$

Since both $g \mapsto \phi(g)$ and $g \mapsto \mathbb{E}(fg)$ are linear it is enough to prove that they are equal for $g = e_i$ for all $i = 1, \dots, N$ because the e_i 's span the vector space of all functions $\mathbb{Z}_N \rightarrow \mathbb{R}$. We see that

$$\mathbb{E}(fe_i) = \frac{1}{N}f(i) = \phi(e_i)$$

and we are done. So there is a bijection from the dual $(U^d)^*$ the functions $\mathbb{Z}_N \rightarrow \mathbb{R}$. The dual norm is usually (see for instance [10]) defined as

$$\|\phi\| = \sup\{|\phi(g)| \mid g : \mathbb{Z}_N \rightarrow \mathbb{R}, \|g\|_{U^d} \leq 1\}.$$

Using the bijection given by our new version of the Riesz representation theorem this norm can be seen as a new norm on the functions $\mathbb{Z}_N \rightarrow \mathbb{R}$ defined as follows.

Definition 5.1.1 (The $(U^d)^*$ norm). Let $f : \mathbb{Z}_N \rightarrow \mathbb{R}$. Define the $(U^d)^*$ norm by

$$\|f\|_{(U^d)^*} = \sup\{|\mathbb{E}(fg)| \mid g : \mathbb{Z}_N \rightarrow \mathbb{R}, \|g\|_{U^d} \leq 1\}.$$

This is a genuine norm, which can be proved by straightforward calculations. We now proceed with the definition of the dual function.

Definition 5.1.2 (The dual function). Let $F : \mathbb{Z}_N \rightarrow \mathbb{R}$. We define the *dual function* $F^* : \mathbb{Z}_N \rightarrow \mathbb{R}$ of F by

$$F^*(x) = \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1} \right),$$

where $\omega \neq 0$ means that $\omega \neq (0, \dots, 0)$.

The following lemma give us some properties of the dual norm and the dual function.

Lemma 5.1.3. *Let $F : \mathbb{Z}_N \rightarrow \mathbb{R}$. Then we have*

$$\mathbb{E}(F F^*) = \|F\|_{U^{k-1}}^{2^{k-1}}, \quad (5.1.1)$$

and

$$\|F^*\|_{(U^{k-1})^*} = \|F\|_{U^{k-1}}^{2^{k-1}-1}. \quad (5.1.2)$$

And if we assume that $|F(x)| \leq \nu(x) + 1$ for all $x \in \mathbb{Z}_N$ where ν is k -pseudorandom then

$$\|F^*\|_{L^\infty} \leq 2^{2^{k-1}-1} + o(1). \quad (5.1.3)$$

Proof. First we want to prove (5.1.1). Consider the LHS. By the definition of F^* we have

$$\mathbb{E}(F F^*) = \mathbb{E} \left(F(x) \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1} \right) \mid x \in \mathbb{Z}_N \right),$$

and since the factor we omit in the product is the one where $\omega = 0$, $F(x)$ can be seen as this omitted factor and then put the whole thing in one expected value to get

$$\mathbb{E}(F F^*) = \mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} F(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1}, x \in \mathbb{Z}_N \right)$$

which by definition is $\|F\|_{U^{k-1}}^{2^{k-1}}$.

Now consider (5.1.2). If $F = 0$ we are done since $0^* = 0$ so we may assume that $F \neq 0$. If we can prove that

$$|\mathbb{E}(fF^*)| \leq \|F\|_{U^{k-1}}^{2^{k-1}-1} \|f\|_{U^{k-1}} \quad (5.1.4)$$

for all $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ we have

$$\|F^*\|_{(U^{k-1})^*} = \sup\{|\mathbb{E}(fF^*)| \mid f : \mathbb{Z}_N \rightarrow \mathbb{R}, \|f\|_{U^{k-1}} \leq 1\} \leq \|F\|_{U^{k-1}}^{2^{k-1}-1},$$

so let us prove (5.1.4). Define a $\{0, 1\}^{k-1}$ -sequence of functions $\mathbb{Z}_N \rightarrow \mathbb{R}$ by

$$f_\omega = \begin{cases} f & \text{if } \omega = 0 \\ F & \text{otherwise.} \end{cases}$$

Then the Gowers inner-product of this sequence is

$$\begin{aligned} \langle (f_\omega)_{\omega \in \{0,1\}^{k-1}} \rangle &= \mathbb{E} \left(f(x) \prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \\ &= \mathbb{E} \left(f(x) \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1} \right) \mid x \in \mathbb{Z}_N \right) \\ &= \mathbb{E}(f(x)F^*(x)), \end{aligned}$$

and by the Gowers Cauchy-Schwarz (proposition 4.1.3) we thus get that

$$\mathbb{E}(f(x)F^*(x)) = \langle (f_\omega)_{\omega \in \{0,1\}^{k-1}} \rangle \leq \|f\|_{U^{k-1}} \|F\|_{U^{k-1}}^{2^{k-1}-1}$$

as desired. Now we need to prove that

$$\|F^*\|_{(U^{k-1})^*} \geq \|F\|_{U^{k-1}}^{2^{k-1}-1}.$$

If we in the set of which we are taking supremum pick $f = \|F\|_{U^{k-1}}^{-1} F$ then $\|f\|_{U^{k-1}} = 1$ and

$$\mathbb{E}(fF^*) = \|F\|_{U^{k-1}}^{2^{k-1}-1}$$

by (5.1.1) so

$$\sup\{|\mathbb{E}(fF^*)| \mid f : \mathbb{Z}_N \rightarrow \mathbb{R}, \|f\|_{U^{k-1}} \leq 1\} \geq \|F\|_{U^{k-1}}^{2^{k-1}-1}$$

which concludes the proof.

Finally we want to prove (5.1.3). Let $\nu' = (\nu + 1)/2$. We have

$$|F(x)| \leq 2\nu'$$

since $2\nu' = \nu + 1$. Now it is enough to prove that $\nu^*(x) \leq 1 + o(1)$ uniformly for all $x \in \mathbb{Z}_N$, because then

$$\begin{aligned} \|F^*\|_{L^\infty} &= \sup_{x \in \mathbb{Z}_N} |F^*(x)| \\ &= \sup_{x \in \mathbb{Z}_N} \left| \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1} \right) \right| \\ &\leq \sup_{x \in \mathbb{Z}_N} \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} 2\nu'(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1} \right), \end{aligned}$$

and by (5.1.2) this is equal to

$$2^{2^{k-1}-1} \sup_{x \in \mathbb{Z}_N} \nu^*(x)$$

which is smaller than $2^{2^{k-1}-1} + o(1)$. As seen before $\nu^*(x)$ can be written as

$$\mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} \nu'(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1} \right), \quad (5.1.5)$$

but since ν is k -pseudorandom, lemma 2.3.7 gives us that ν' is k -pseudorandom so it satisfies the linear forms condition. The $(2^{k-1} - 1, k - 1, 1)$ -linear forms condition with the $2^{k-1} - 1$ linear forms (one for each $\omega \neq 0$)

$$h \mapsto \omega \cdot h + x$$

then gives us that (5.1.5) is $1 + o(1)$. □

5.2 Gowers anti-uniform functions

Definition 5.2.1 (Gowers anti-uniform functions). Let ν be k -pseudorandom. A function $G : \mathbb{Z}_N \rightarrow \mathbb{R}$ is called a *Gowers anti-uniform function with respect to ν* if there is a function $F : \mathbb{Z}_N \rightarrow \mathbb{R}$ such that $|F(x)| \leq \nu(x) + 1$ for all $x \in \mathbb{Z}_N$ and $G = F^*$.

Remark 5.2.2. When we assume the existence of a Gowers anti-uniform function G , we will often just denote it F^* right away and omit the G .

Lemma 5.2.3. *Let $d \geq 1$ and let P be a polynomial of degree d with real coefficients on K variables, and let F_1^*, \dots, F_K^* be Gowers anti-uniform functions with respect to a k -pseudorandom measure ν . Then*

$$\|P(F_1^*, \dots, F_K^*)\|_{(U^{k-1})^*} = O_{K,d,P}(1).$$

Proof. First we may WLOG replace the F_j with $F_j/2$ for all j and because of lemma 2.3.7 replace ν with $(\nu + 1)/2$, and we may therefore assume

$$|F_j(x)| \leq \nu(x). \quad (5.2.1)$$

By linearity and the triangle-inequality we get that it is enough to prove the lemma when P is a monomial, and if some variables appear with higher power than 1, we can repeat that variable and thus increase the number of variables in P , so

$$P(x_1, \dots, x_K) = x_1 \cdots x_K.$$

So we need to show that

$$\|P(F_1^*, \dots, F_K^*)\|_{(U^{k-1})^*} = O_K(1),$$

since both the degree of P and P itself now only depends on K . By definition of the dual norm we have to show that

$$\mathbb{E}(fF_1^* \cdots F_K^*) = O_K(1)$$

for all $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ with $\|f\|_{U^{k-1}} \leq 1$. Using the definition of the dual function we can write the LHS as

$$\mathbb{E} \left(f(x) \prod_{i=1}^K \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F_i(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1} \right) \mid x \in \mathbb{Z}_N \right). \quad (5.2.2)$$

Now define a map $\Phi : \mathbb{Z}_N^{k-1} \times \mathbb{Z}_N^{k-1} \rightarrow \mathbb{Z}_N^{k-1}$ by $\Phi(h, H) = h + H$. This map is surjective, and

$$\#\Phi^{-1}(x) = \#\{(h, H) \in \mathbb{Z}_N^{k-1} \times \mathbb{Z}_N^{k-1} \mid h + H = x\} = N^{k-1},$$

since for each h we can find exactly one H such that $h + H = x$, so Φ is a uniform cover and (5.2.2) can be rewritten to

$$\mathbb{E} \left(f(x) \prod_{i=1}^K \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F_i(x + \omega \cdot h + \omega \cdot H) \mid h, H \in \mathbb{Z}_N^{k-1} \right) \mid x \in \mathbb{Z}_N \right)$$

by lemma 2.4.2. We can now move the variable h to the outer expected value by 2. in proposition 2.2.5. For each i in the product we denote H by H^i . This gives us

$$\mathbb{E} \left(f(x) \prod_{i=1}^K \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F_i(x + \omega \cdot h + \omega \cdot H^i) \mid H^i \in \mathbb{Z}_N^{k-1} \right) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right). \quad (5.2.3)$$

Now define for each $(H^1, \dots, H^K) = \hat{H} \in (\mathbb{Z}_N^{k-1})^K$ and $\omega \in \{0, 1\}^{k-1}$ a function $f_{\omega, \hat{H}}$ by

$$f_{\omega, \hat{H}} = \begin{cases} f & \text{if } \omega = 0 \\ g_{\omega \cdot \hat{H}} & \text{otherwise} \end{cases}$$

where $\omega \cdot \hat{H} = (\omega \cdot H^1, \dots, \omega \cdot H^K)$ and

$$g_{u_1, \dots, u_K}(x) = \prod_{j=1}^K F_j(x + u_j).$$

Now we want to prove that (5.2.3) is equal to

$$\mathbb{E} \left(\langle (f_{\omega, \hat{H}})_{\omega \in \{0, 1\}^{k-1}} \rangle_{U^{k-1}} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right).$$

By definition of the Gowers inner product and 2. in proposition 2.2.5 this is equal to

$$\mathbb{E} \left(\prod_{\omega \in \{0, 1\}^{k-1}} f_{\omega, \hat{H}}(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}, \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right).$$

Recall that when $\omega = 0$ we have $f_{\omega, \hat{H}} = f$ and $f_{\omega, \hat{H}} = g_{\omega \cdot \hat{H}}$ otherwise, so the above expression is

$$\mathbb{E} \left(f(x) \prod_{\substack{\omega \in \{0, 1\}^{k-1} \\ \omega \neq 0}} g_{\omega \cdot \hat{H}}(x + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}, \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right),$$

and recalling the definition of g this is

$$\mathbb{E} \left(f(x) \prod_{\substack{\omega \in \{0, 1\}^{k-1} \\ \omega \neq 0}} \prod_{i=1}^K F_i(x + \omega \cdot H^i + \omega \cdot h) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}, \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right).$$

Changing the order of the products this is equal to

$$\mathbb{E} \left(f(x) \prod_{i=1}^K \prod_{\substack{\omega \in \{0, 1\}^{k-1} \\ \omega \neq 0}} F_i(x + \omega \cdot H^i + \omega \cdot h) \mid x \in \mathbb{Z}_N, h, H^1, \dots, H^K \in \mathbb{Z}_N^{k-1} \right).$$

Using 2. in proposition 2.2.5 once more we get

$$\mathbb{E} \left(f(x) \prod_{i=1}^K \mathbb{E} \left(\prod_{\substack{\omega \in \{0, 1\}^{k-1} \\ \omega \neq 0}} F_i(x + \omega \cdot h + \omega \cdot H^i) \mid H^1, \dots, H^K \in \mathbb{Z}_N^{k-1} \right) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right),$$

but notice that the expression

$$\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F_i(x + \omega \cdot h + \omega \cdot H^i)$$

is independent of H^j when $j \neq i$, so by 3. in proposition 2.2.5 we can omit all the H^j where $j \neq i$ and get

$$\mathbb{E} \left(f(x) \prod_{i=1}^K \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} F_i(x + \omega \cdot h + \omega \cdot H^i) \mid H^i \in \mathbb{Z}_N^{k-1} \right) \mid x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right),$$

so we have the equality

$$\mathbb{E} (f F_1^* \cdots F_K^*) = \mathbb{E} \left(\langle (f_{\omega, \hat{H}})_{\omega \in \{0,1\}^{k-1}} \rangle_{U^{k-1}} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right).$$

Recall that what we need to prove is that the LHS is $O_K(1)$. By the Gowers Cauchy-Schwarz (2. in proposition 4.1.3) we can bound the RHS by

$$\mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}} \|f_{\omega, \hat{H}}\|_{U^{k-1}} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right),$$

and recalling the definition of $f_{\omega, \hat{H}}$ this can be written as

$$\mathbb{E} \left(\|f\|_{U^{k-1}} \prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} \|g_{\omega, \hat{H}}\|_{U^{k-1}} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right).$$

So to finish the proof we need to show that this is $O_K(1)$. But $\|f\|_{U^{k-1}} \leq 1$ so it is enough to prove that

$$\mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} \|g_{\omega, \hat{H}}\|_{U^{k-1}} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right) = O_K(1).$$

Now notice that the LHS can be seen as the $L^1(\mathbb{Z}_N^K)$ -norm of

$$\hat{H} \mapsto \prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} \|g_{\omega, \hat{H}}\|_{U^{k-1}},$$

so by using Hölders inequality (lemma 2.5.3) we get

$$\begin{aligned} \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} \|g_{\omega \cdot \hat{H}}\|_{U^{k-1}} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right) \\ \leq \prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} \mathbb{E} \left(\|g_{\omega \cdot \hat{H}}\|_{U^{k-1}}^{2^{k-1}-1} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right)^{\frac{1}{2^{k-1}-1}}, \end{aligned}$$

and it is therefore enough to prove that each of the factors on the RHS is $O_K(1)$ because the number of factors only depends on k and $O_K(1)$ is allowed to depend on k . So let $\omega \in \{0,1\}^{k-1}$ be given. We want to prove that

$$\mathbb{E} \left(\|g_{\omega \cdot \hat{H}}\|_{U^{k-1}}^{2^{k-1}-1} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right) = O_K(1),$$

where we can omit the outer exponent because it only depends on k . Since

$$x \mapsto |x|^{\frac{2^{k-1}}{2^{k-1}-1}}$$

is convex, we get by Jensens inequality (lemma 2.5.2) that

$$\mathbb{E} \left(\|g_{\omega \cdot \hat{H}}\|_{U^{k-1}}^{2^{k-1}-1} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right) \leq \mathbb{E} \left(\|g_{\omega \cdot \hat{H}}\|_{U^{k-1}}^{2^{k-1}} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right)^{\frac{2^{k-1}-1}{2^{k-1}}}$$

so it is enough to prove that

$$\mathbb{E} \left(\|g_{\omega \cdot \hat{H}}\|_{U^{k-1}}^{2^{k-1}} \mid \hat{H} \in (\mathbb{Z}_N^{k-1})^K \right) = O_K(1). \quad (5.2.4)$$

Now consider the map $(\mathbb{Z}_N^{k-1})^K \rightarrow \mathbb{Z}_N^K$ defined by

$$\hat{H} \mapsto \omega \cdot \hat{H}. \quad (5.2.5)$$

I now claim that this is a uniform covering. The preimage of a $M = (m_1, \dots, m_K) \in \mathbb{Z}_N^K$ is

$$\{\hat{H} \in (\mathbb{Z}_N^{k-1})^K \mid \hat{H} \cdot \omega = M\} = \{(H^1, \dots, H^K) \in (\mathbb{Z}_N^{k-1})^K \mid H^i \cdot \omega = m_i \forall i\}$$

and we need to show that this set has cardinality

$$\#(\mathbb{Z}_N^{k-1})^K / \#\mathbb{Z}_N^K = (N^{(k-1)})^K / N^K = (N^{k-2})^K.$$

Now write $H^i = (H_1^i, \dots, H_K^i)$. The size of the pullback of M is the number of solutions (H^1, \dots, H^K) to the following system of equations

$$\begin{aligned} m_1 &= \omega \cdot H^1 = \omega_1 H_1^1 + \omega_2 H_2^1 + \dots + \omega_{k-1} H_{k-1}^1 \\ &\vdots \\ m_K &= \omega \cdot H^K = \omega_1 H_1^K + \omega_2 H_2^K + \dots + \omega_{k-1} H_{k-1}^K. \end{aligned}$$

But since $\omega \neq 0$ there is one i such that $\omega_i = 1$. So for each $l = 1, \dots, K$ we can pick H_j^l freely in \mathbb{Z}_N when $j \neq i$, and H_i^l has to be

$$H_i^l = m_l - \omega_1 H_1^l - \omega_2 H_2^l - \dots - \omega_{i-1} H_{i-1}^l - \omega_{i+1} H_{i+1}^l - \dots - \omega_{k-1} H_{k-1}^l,$$

so we can pick each H^l in N^{k-2} different ways and thus pick $\hat{H} = (H^1, \dots, H^K)$ in $(N^{k-2})^K$ different ways. So the map defined in (5.2.5) is a uniform covering and we can rewrite the LHS of (5.2.4) as

$$\mathbb{E} \left(\left\| g_{u_1, \dots, u_K} \right\|_{U^{k-1}}^{2^{k-1}} \left| u_1, \dots, u_K \in \mathbb{Z}_N \right. \right).$$

By the definition of the Gowers norm and the definition of g_{u_1, \dots, u_K} this can be written as

$$\mathbb{E} \left(\mathbb{E} \left(\prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \prod_{i=1}^K F_i(x + u_i + h \cdot \tilde{\omega}) \left| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right. \right) \left| u_1, \dots, u_K \in \mathbb{Z}_N \right. \right),$$

which can be written in one expected value. If we also change the order of the products we get

$$\mathbb{E} \left(\prod_{i=1}^K \prod_{\tilde{\omega} \in \{0,1\}^{k-1}} F_i(x + u_i + h \cdot \tilde{\omega}) \left| x, u_1, \dots, u_K \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right. \right).$$

Noticing by \mathcal{B} . in proposition 2.2.5 that we can omit all the u_j where $i \neq j$ because $F_i(x + u_i + h \cdot \tilde{\omega})$ is independent of those we now get

$$\mathbb{E} \left(\prod_{i=1}^K \mathbb{E} \left(\prod_{\tilde{\omega} \in \{0,1\}^{k-1}} F_i(x + u + h \cdot \tilde{\omega}) \left| u \in \mathbb{Z}_N \right. \right) \left| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right. \right).$$

By (5.2.1) we now get that it is enough to prove that the above expression is $O_K(1)$ with all the F_j 's replaced by ν , i.e. that

$$\mathbb{E} \left(\mathbb{E} \left(\prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(x + u + h \cdot \tilde{\omega}) \left| u \in \mathbb{Z}_N \right. \right)^K \left| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right. \right) = O_K(1).$$

Now notice that for fixed x , $x + u$ runs through \mathbb{Z}_N when u runs through \mathbb{Z}_N so we can replace $x + u$ with a new variable y which runs through \mathbb{Z}_N , and then omit the x . So it is enough to prove

$$\mathbb{E} \left(\mathbb{E} \left(\prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(y + h \cdot \tilde{\omega}) \mid y \in \mathbb{Z}_N \right)^K \mid h \in \mathbb{Z}_N^{k-1} \right) = O_K(1).$$

Now we want to apply the correlation condition (this is actually the only place in the proof that the correlation condition is needed) on ν which gives us that there is a function $\tau : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ such that

$$\mathbb{E}(\tau^q) = O_q(1)$$

for all $q \geq 1$ and

$$\mathbb{E} \left(\prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(y + h \cdot \tilde{\omega}) \mid y \in \mathbb{Z}_N \right) \leq \sum_{\substack{\tilde{\omega}, \tilde{\omega}' \in \{0,1\}^{k-1} \\ \tilde{\omega} \neq \tilde{\omega}'}} \tau(h \cdot (\tilde{\omega} - \tilde{\omega}')),$$

so we have

$$\begin{aligned} \mathbb{E} \left(\mathbb{E} \left(\prod_{\tilde{\omega} \in \{0,1\}^{k-1}} \nu(y + h \cdot \tilde{\omega}) \mid y \in \mathbb{Z}_N \right)^K \mid h \in \mathbb{Z}_N^{k-1} \right) \\ \leq \mathbb{E} \left(\left(\sum_{\substack{\tilde{\omega}, \tilde{\omega}' \in \{0,1\}^{k-1} \\ \tilde{\omega} \neq \tilde{\omega}'}} \tau(h \cdot (\tilde{\omega} - \tilde{\omega}')) \right)^K \mid h \in \mathbb{Z}_N^{k-1} \right), \end{aligned}$$

but then the RHS is the K -th power of the $L^K(\mathbb{Z}_N^{k-1})$ -norm of the sum, so using the triangle inequality we get that

$$\begin{aligned} \mathbb{E} \left(\left(\sum_{\substack{\tilde{\omega}, \tilde{\omega}' \in \{0,1\}^{k-1} \\ \tilde{\omega} \neq \tilde{\omega}'}} \tau(h \cdot (\tilde{\omega} - \tilde{\omega}')) \right)^K \mid h \in \mathbb{Z}_N^{k-1} \right)^{\frac{1}{K}} \\ \leq \sum_{\substack{\tilde{\omega}, \tilde{\omega}' \in \{0,1\}^{k-1} \\ \tilde{\omega} \neq \tilde{\omega}'}} \mathbb{E} \left((\tau(h \cdot (\tilde{\omega} - \tilde{\omega}')))^K \mid h \in \mathbb{Z}_N^{k-1} \right)^{\frac{1}{K}}. \end{aligned}$$

So it suffices to show that

$$\mathbb{E} \left((\tau(h \cdot (\tilde{\omega} - \tilde{\omega}')))^K \mid h \in \mathbb{Z}_N^{k-1} \right) = O_K(1)$$

for all $\tilde{\omega}, \tilde{\omega}' \in \{0, 1\}^{k-1}$ with $\tilde{\omega} \neq \tilde{\omega}'$. We now define a map $\mathbb{Z}_N^{k-1} \rightarrow \mathbb{Z}_N$ by

$$h \mapsto h \cdot (\tilde{\omega} - \tilde{\omega}').$$

We now claim that this is a uniform covering. If that is the case then

$$\mathbb{E} \left((\tau(h \cdot (\tilde{\omega} - \tilde{\omega}')))^K \mid h \in \mathbb{Z}_N^{k-1} \right) = \mathbb{E} (\tau(x)^K \mid x \in \mathbb{Z}_N) = \mathbb{E} (\tau^K) = O_K(1)$$

and we are done, so it only remains to prove that this map is a uniform covering. Let $\tilde{\omega}, \tilde{\omega}' \in \{0, 1\}^{k-1}$ with $\tilde{\omega} \neq \tilde{\omega}'$ and $x \in \mathbb{Z}_N$ be given. We now need to prove that the preimage of this x has cardinality N^{k-2} . But the $h \in \mathbb{Z}_N^{k-1}$ such that $h \cdot (\tilde{\omega} - \tilde{\omega}') = x$ is exactly the $h \in \mathbb{Z}_N^{k-1}$ such that

$$x = h_1 \bar{\omega}_1 + \cdots + h_{k-1} \bar{\omega}_{k-1}$$

where $\tilde{\omega} - \tilde{\omega}' = (\bar{\omega}_1, \dots, \bar{\omega}_{k-1}) \neq 0$ because $\tilde{\omega} \neq \tilde{\omega}'$. If we consider this as a system of linear equations it has non-zero determinant so the solution space has dimension $k-2$, and it has N^{k-2} solutions as we claimed. \square

Now we are ready to prove the following proposition.

Proposition 5.2.4. *Let ν be k -pseudorandom. Let $K \geq 1$ and let $\Phi : I^K \rightarrow \mathbb{R}$ be a continuous function, where $I = [-2^{2^{k-1}}, 2^{2^{k-1}}]$. Let F_1^*, \dots, F_K^* be basic Gowers anti-uniform functions and let $\psi : \mathbb{Z}_N \rightarrow \mathbb{R}$ be defined by*

$$\psi(x) = \Phi(F_1^*(x), \dots, F_K^*(x)).$$

Then

$$\mathbb{E}((\nu - 1)\psi) = o_{K, \Phi}(1).$$

Furthermore, if Φ ranges over a compact set $E \subset C^0(I^K)$ then

$$\mathbb{E}((\nu - 1)\psi) = o_{K, E}(1).$$

Proof. As before, we may replace ν with $(\nu + 1)/2$ by lemma 2.3.7 if we replace F_j with $F_j/2$ for all j and hence get that

$$|F_j(x)| \leq \nu(x)$$

for all j and all $x \in \mathbb{Z}_N$. Now let $\varepsilon > 0$ be given. By the Weierstrass approximation theorem we get that there is a polynomial $P : I \rightarrow \mathbb{R}$ on K variables, depending on K, Φ and ε such that

$$\|\Phi(x_1, \dots, x_K) - P(x_1, \dots, x_K)\|_{L^\infty} < \varepsilon.$$

By (5.1.3) we see that

$$\left| \sup_{x \in \mathbb{Z}_N} F_j^*(x) \right| \leq 2^{2^{k-1}-1} + o(1),$$

so for large N we have $F_j^*(x) \in I$ for all $x \in \mathbb{Z}_N$ and we may insert the F_j^* 's in both Φ and P to get that

$$\|\Phi(F_1^*, \dots, F_K^*) - P(F_1^*, \dots, F_K^*)\|_{L^\infty} < \varepsilon.$$

But then

$$\begin{aligned} |\mathbb{E}((\nu - 1)(\Phi(F_1^*, \dots, F_K^*) - P(F_1^*, \dots, F_K^*)))| &\leq \mathbb{E}(|\nu| |\Phi(F_1^*, \dots, F_K^*) - P(F_1^*, \dots, F_K^*)|) \\ &\quad + \mathbb{E}(|\Phi(F_1^*, \dots, F_K^*) - P(F_1^*, \dots, F_K^*)|) \\ &\leq (2 + o(1))\varepsilon, \end{aligned}$$

since $\mathbb{E}(\nu) = 1 + o(1)$. And using lemma 5.2.3 we get that

$$\|P(F_1^*, \dots, F_K^*)\|_{(U^{k-1})^*} = O_{K, \Phi, \varepsilon}(1)$$

because P and hence also the degree of P depends on Φ, K and ε . By lemma 4.1.4 we have that

$$\|\nu - 1\|_{U^{k-1}} = o(1).$$

And combining these two we get that

$$\begin{aligned} \mathbb{E}((\nu - 1)P(F_1^*, \dots, F_K^*)) &= o(1)\mathbb{E}(P(F_1^*, \dots, F_K^*)) \leq o(1)\|P(F_1^*, \dots, F_K^*)\|_{(U^{k-1})^*} \\ &= o(1)O_{K, \Phi, \varepsilon}(1) = o_{K, \Phi, \varepsilon}(1) \end{aligned}$$

So using the estimates we have found so far we get that

$$\begin{aligned} |\mathbb{E}((\nu - 1)\Phi(F_1^*, \dots, F_K^*))| &\leq |\mathbb{E}((\nu - 1)(\Phi(F_1^*, \dots, F_K^*) - P(F_1^*, \dots, F_K^*)))| \\ &\quad + |\mathbb{E}((\nu - 1)P(F_1^*, \dots, F_K^*))| = (2 + o(1))\varepsilon + o_{K, \Phi, \varepsilon}(1) \end{aligned}$$

and for sufficiently large N , this is smaller than 4ε so we are done since $\varepsilon > 0$ was picked arbitrarily.

Now let $E \subseteq C^0(I^K)$ be a compact set (in the uniform topology) and let $\eta > 0$ be given. We have

$$E \subseteq \bigcup_{\Phi \in E} B_\Phi(\eta)$$

where $B_x(\eta)$ is the open ball of radius η with center x . Since E is compact we also have

$$E \subseteq \bigcup_{\Phi \in X} B_\Phi(\eta)$$

where $X \subseteq E$ is a finite set. Now for each $\Phi \in X$ we have

$$|\mathbb{E}((\nu - 1)\Phi(F_1^*, \dots, F_K^*))| = o_{K, \Phi}(1)$$

Now let $\Phi' \in X$ be the Φ that makes the LHS largest. Then

$$|\mathbb{E}((\nu - 1)\Phi(F_1^*, \dots, F_K^*))| = o_{K, \Phi'}(1) = o_{K, E}(1)$$

for all $\Phi \in X$ because our choice of Φ' only depends on E . Now let $\Psi \in E$ be given. Then there is $\Phi \in X$ such that

$$\|\Psi - \Phi\|_{L^\infty} < \eta.$$

Then

$$\begin{aligned} |\mathbb{E}((\nu - 1)\Psi(F_1^*, \dots, F_K^*))| &\leq |\mathbb{E}((\nu - 1)\Psi(F_1^*, \dots, F_K^*)) - \mathbb{E}((\nu - 1)\Phi(F_1^*, \dots, F_K^*))| \\ &\quad + |\mathbb{E}((\nu - 1)\Phi(F_1^*, \dots, F_K^*))| \\ &< |\eta\mathbb{E}(\nu - 1)| + o_{K,E}(1) \\ &= o_{K,E}(1) \end{aligned}$$

because $\mathbb{E}(\nu - 1) = o(1)$. □

Chapter 6

σ -algebras on \mathbb{Z}_N

6.1 Definitions and notation

Let us define what we consider as a σ -algebra in this context.

Definition 6.1.1 (σ -algebras and atoms). A collection of subsets $\mathcal{B} \subseteq 2^{\mathbb{Z}_N}$ is called a σ -algebra if

1. $\emptyset, \mathbb{Z}_N \in \mathcal{B}$,
2. $A \in \mathcal{B} \Rightarrow \mathbb{Z}_N \setminus A = A^C \in \mathcal{B}$,
3. $A, B \in \mathcal{B} \Rightarrow A \cup B \in \mathcal{B}$,
4. $A, B \in \mathcal{B} \Rightarrow A \cap B \in \mathcal{B}$.

An element $A \in \mathcal{B}$ is called an *atom* of \mathcal{B} if $A \neq \emptyset$ and A is minimal in \mathcal{B} with respect to inclusion

$$(B \subseteq A \text{ and } B \in \mathcal{B} \setminus \{\emptyset\}) \Rightarrow B = A.$$

Note that we only need to consider finite unions and intersections, since the set \mathbb{Z}_N is finite and so every σ -algebra will be finite.

The following lemma states that the atoms of a σ -algebra forms a partition of \mathbb{Z}_N .

Lemma 6.1.2. *Let \mathcal{B} be a σ -algebra and let $A_1, \dots, A_n \in \mathcal{B}$ be the distinct atoms of \mathcal{B} . Then $A_i \cap A_j = \emptyset$ if $i \neq j$,*

$$A_1 \cup \dots \cup A_n = \mathbb{Z}_N,$$

and each $B \in \mathcal{B}$ can be written as

$$B = \bigcup_{i \in I} A_i$$

for some $I \subseteq \{1, \dots, n\}$.

Proof. There are three things to prove. Assume that there are $i \neq j$ such that $A_i \cap A_j \neq \emptyset$. Now let $B = A_i \cap A_j$. By the definition of σ -algebras we have that

$$\emptyset \neq B = A_i \cap A_j \in \mathcal{B}$$

and $B \subseteq A_i$ and $B \subseteq A_j$. But by the definition of atoms this implies that $A_j = B = A_i$ which is a contradiction.

Now we want to prove that the atoms form a partition of \mathbb{Z}_N . Let $x \in \mathbb{Z}_N$ be given, we will now find an atom that contains x . Let $C_1 = \mathbb{Z}_N$.

Now let $i \geq 1$. If C_i is an atom we are done, otherwise there is a non-trivial subset $C'_i \subset C_i$. If $x \in C'_i$, let $C_{i+1} = C'_i$, otherwise let $C_{i+1} = C_i \setminus C'_i = C_i \cap C_i^c$. Then $x \in C_{i+1} \in \mathcal{B}$, and we can increase i by 1 and perform the step again.

If we continue like this, we will at some point end up with a C_j which is an atom, because if not, the algorithm would go on forever and this would give an infinitely long inclusion sequence

$$C_1 \supset C_2 \supset C_3 \supset \dots,$$

but there only are finitely many elements of \mathcal{B} , so the algorithm will terminate with some atom C_j with $x \in C_j$, and we are done.

The last thing we need to prove is that each element $B \in \mathcal{B}$ can be written as a union of atoms. This can also be done by a recursive algorithm. Let $B \in \mathcal{B}$ and define a tree with B being the top node. Now if B is an atom, we are done. Otherwise there is a subset $B' \subset B$ such that $B' \in \mathcal{B}$. Notice also that $B'' = B \setminus B' = B' \cap B \in \mathcal{B}$ by the definition of σ -algebras. Let B' and B'' be child nodes of B . Now if B' and B'' both are atoms, we are done. If one of them is not, we can start the algorithm again with this one being B . This gives a tree, which must be finite by the same argument as before, and where the leafs are atoms whose union is exactly B since everywhere in the tree the child nodes form a partition of the parent node. \square

Notice also that we can define a σ -algebra by giving a partition of \mathbb{Z}_N – the elements of this partition will then be the atoms of the new σ -algebra. Now we define the concept of measurability of functions.

Definition 6.1.3 (Measurable function). Let \mathcal{B} be a σ -algebra. A function $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ is *measurable with respect to \mathcal{B}* if

$$f^{-1}(\{x\}) \in \mathcal{B}.$$

This is equivalent to saying that for all atoms $A \in \mathcal{B}$ we have $f(x) = f(y)$ if $x, y \in A$.

This gives us the possibility to define conditional expectations.

Definition 6.1.4 (Conditional expectations). For a σ -algebra \mathcal{B} and a function $f : \mathbb{Z}_N \rightarrow \mathbb{R}$, we define the function $\mathbb{E}(f | \mathcal{B}) : \mathbb{Z}_N \rightarrow \mathbb{R}$ by

$$\mathbb{E}(f | \mathcal{B})(x) = \mathbb{E}(f(y) | y \in \mathcal{B}(x))$$

for $x \in \mathbb{Z}_N$, where $\mathcal{B}(x)$ is the unique atom containing x .

Proposition 6.1.5. *The conditional expectation is linear and it preserves positivity and constant functions. If $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ and $A \in \mathcal{B}$ we have*

$$\mathbb{E}(\mathbb{1}_A \mathbb{E}(f | \mathcal{B})) = \mathbb{E}(\mathbb{1}_A f), \quad (6.1.1)$$

and if f is \mathcal{B} -measurable for some σ -algebra \mathcal{B} then

$$\mathbb{E}(f | \mathcal{B})(x) = f(x)$$

for all $x \in \mathbb{Z}_N$.

Proof. The linearity comes from the fact that the expected value is linear, and the preservation of positivity and constant functions is clear.

Now let $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ and $A \in \mathcal{B}$. It is enough to prove (6.1.1) for $A \in \mathcal{B}$ being an atom because if $A_1 \cup \dots \cup A_i$ is a partition of A then

$$\begin{aligned} \mathbb{E}(\mathbb{1}_A \mathbb{E}(f | \mathcal{B})) &= \mathbb{E}(\mathbb{1}_{A_1} \mathbb{E}(f | \mathcal{B})) + \dots + \mathbb{E}(\mathbb{1}_{A_i} \mathbb{E}(f | \mathcal{B})) \\ &= \mathbb{E}(\mathbb{1}_{A_1} f) + \dots + \mathbb{E}(\mathbb{1}_{A_i} f) = \mathbb{E}(\mathbb{1}_A f). \end{aligned}$$

By definition we have

$$\mathbb{E}(\mathbb{1}_A \mathbb{E}(f | \mathcal{B})) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \mathbb{1}_A(x) \frac{1}{\#\mathcal{B}(x)} \sum_{y \in \mathcal{B}(x)} f(y).$$

Now notice that if $x \in A$ then $\mathcal{B}(x) = A$. So the above expression can be rewritten to

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} \mathbb{1}_A(x) \frac{1}{\#A} \sum_{y \in A} f(y) = \mathbb{E}(\mathbb{1}_A \mathbb{E}(f(x) | x \in A)).$$

By proposition 2.2.5 this is equal to $\mathbb{E}(\mathbb{1}_A(x)f(y) | x \in \mathbb{Z}_N, y \in A)$. Notice now that the map $\mathbb{Z}_N \times A \rightarrow \mathbb{Z}_N$ defined by $(x, y) \mapsto x$ is a uniform cover because the pullback of an element $x \in \mathbb{Z}_N$ is $\{x\} \times A$ which has cardinality $\#A$. This gives us

$$\mathbb{E}(\mathbb{1}_A(x)f(y) | x \in \mathbb{Z}_N, y \in A) = \mathbb{E}(\mathbb{1}_A f)$$

as desired.

Now assume that f is \mathcal{B} -measurable. Then

$$\mathbb{E}(f | \mathcal{B})(x) = \mathbb{E}(f(y) | y \in \mathcal{B}(x))$$

and for all $x \in \mathbb{Z}_N$ we have that $f(y) = f(x)$ for all $y \in \mathcal{B}(x)$ so the above expression is equal to

$$\mathbb{E}(f(x) | y \in \mathcal{B}(x)) = f(x).$$

□

We now prove that measurable functions act as constants in conditional expectations.

Lemma 6.1.6. *Let $f, g : \mathbb{Z}_N \rightarrow \mathbb{R}$ and assume that f is \mathcal{B} -measurable. Then*

$$f\mathbb{E}(g \mid \mathcal{B}) = \mathbb{E}(fg \mid \mathcal{B}).$$

Proof. Since f is constant on all atoms of \mathcal{B} we have for a given $x \in \mathbb{Z}_N$ that $f(y) = c$ for all $y \in \mathcal{B}(x)$ and hence

$$\mathbb{E}(fg \mid \mathcal{B})(x) = \mathbb{E}(f(y)g(y) \mid y \in \mathcal{B}(x)) = \mathbb{E}(cg(y) \mid y \in \mathcal{B}(x)) = f(x)\mathbb{E}(g \mid \mathcal{B}).$$

□

The following corollary tells us that any \mathcal{B} -measurable function is orthogonal to its own orthogonal complement.

Corollary 6.1.7. *Let $f, g : \mathbb{Z}_N \rightarrow \mathbb{R}$ and assume that f is \mathcal{B} -measurable. Then*

$$\mathbb{E}(fg - f\mathbb{E}(g \mid \mathcal{B}_K)) = 0.$$

Definition 6.1.8 (Union of σ -algebras). Let $\mathcal{B}_1, \dots, \mathcal{B}_n$ be σ -algebras. Then we define

$$\bigvee_{i=1}^n \mathcal{B}_i = \mathcal{B}_1 \vee \dots \vee \mathcal{B}_n$$

to be the σ -algebra generated by $\mathcal{B}_1, \dots, \mathcal{B}_n$, and we define it by saying that the atoms of this σ -algebra are the sets

$$\bigcap_{i=1}^n A_i,$$

where A_i is an atom of \mathcal{B}_i for all $i = 1, \dots, n$. Not all expressions of this form gives an atom though since some of the intersections can be empty and \emptyset is by definition not an atom.

6.2 Two propositions

Proposition 6.2.1. *Let ν be a k -pseudorandom measure, let $0 < \varepsilon < 1, 0 < \eta < 1/2$ and let $G : \mathbb{Z}_N \rightarrow I = [-2^{2^{k-1}}, 2^{2^{k-1}}]$. Then there exists a σ -algebra $\mathcal{B}_{\varepsilon, \eta}(G)$ such that*

1. *For any σ -algebra we have*

$$\|G - \mathbb{E}(G \mid \mathcal{B} \vee \mathcal{B}_{\varepsilon, \eta}(G))\|_{L^\infty} \leq \varepsilon.$$

2. *$\mathcal{B}_{\varepsilon, \eta}(G)$ has $O(\varepsilon^{-1})$ atoms.*

3. *If A is an atom of $\mathcal{B}_{\varepsilon, \eta}(G)$ then there is a continuous function $\Psi_A : I \rightarrow [0, 1]$ such that*

$$\mathbb{E}((\mathbb{1}_A - \Psi_A(G))(\nu + 1)) = O(\eta),$$

and $\Psi_A \in E_{\varepsilon, \eta} \subset C^0(I)$ where $E_{\varepsilon, \eta}$ is compact and independent of G, ν, N and A .

Proof. Let ε, η and G be given as in the proposition. Consider the following expression

$$\int_0^1 \sum_{n \in \mathbb{Z}} \mathbb{E} \left(\mathbb{1}_{[\varepsilon(n-\eta+\alpha), \varepsilon(n+\eta+\alpha)]}(G(x))(\nu(x) + 1) \mid x \in \mathbb{Z}_N \right) d\alpha. \quad (6.2.1)$$

Now define

$$I_{n,\alpha} = G^{-1}([\varepsilon(n - \eta + \alpha), \varepsilon(n + \eta + \alpha)]).$$

Then (6.2.1) can be rewritten as

$$\int_0^1 \sum_{n \in \mathbb{Z}} \mathbb{E} \left(\mathbb{1}_{I_{n,\alpha}}(x)(\nu(x) + 1) \mid x \in \mathbb{Z}_N \right) d\alpha.$$

For fixed α and sufficiently large n we have $I_{n,\alpha} = \emptyset$, so the sum is finite and we can therefore interchange expected value and sum. Since $\eta < 1/2$ the $I_{n,\alpha}$'s are disjoint for fixed α and we can write the sum of indicator functions as an indicator function over a union to get that the above expression is

$$\int_0^1 \mathbb{E} \left(\mathbb{1}_{M_\alpha}(x)(\nu(x) + 1) \mid x \in \mathbb{Z}_N \right) d\alpha, \quad (6.2.2)$$

where

$$M_\alpha = \bigcup_{n \in \mathbb{Z}} I_{n,\alpha}.$$

Now let

$$E_x = \bigcup_{n \in \mathbb{Z}} [\varepsilon(G(x) - n - \eta), \varepsilon(G(x) - n + \eta)]$$

for $x \in \mathbb{Z}_N$. Then

$$\mathbb{1}_{M_\alpha}(x) = \mathbb{1}_{E_x}(\alpha)$$

for all $0 \leq \alpha \leq 1$ and $x \in \mathbb{Z}_N$. So (6.2.2) is equal to

$$\mathbb{E} \left(\left(\nu(x) + 1 \right) \int_0^1 \mathbb{1}_{E_x}(\alpha) d\alpha \mid x \in \mathbb{Z}_N \right). \quad (6.2.3)$$

Notice now that if we fix $x \in \mathbb{Z}_N$, then α can only be in a set of the form $[\varepsilon(G(x) - n - \eta), \varepsilon(G(x) - n + \eta)]$ for at most one $n \in \mathbb{Z}$ because $0 \leq \alpha \leq 1$ and hence

$$\int_0^1 \mathbb{1}_{E_x}(\alpha) d\alpha \leq 2\varepsilon\eta \leq 2\eta$$

for each $x \in \mathbb{Z}_N$. So (6.2.3) is smaller than

$$2\eta \mathbb{E}(\nu(x) + 1 \mid x \in \mathbb{Z}_N),$$

and since ν is a measure we have $\mathbb{E}(\nu) = 1 + o(1)$ so (6.2.1) is $O(\eta)$. Now if the interior function of the integral is not $O(\eta)$ for any $0 \leq \alpha \leq 1$ then the integral could not be $O(\eta)$ so there must exist $0 \leq \alpha \leq 1$ such that

$$\sum_{n \in \mathbb{Z}} \mathbb{E}(\mathbb{1}_{[\varepsilon(n-\eta+\alpha), \varepsilon(n+\eta+\alpha)]}(G(x))(\nu(x) + 1) \mid x \in \mathbb{Z}_N) = O(\eta).$$

Let such an α be fixed. Now define the σ -algebra $\mathcal{B}_{\varepsilon, \eta}(G)$ to be the σ -algebra with atoms

$$G^{-1}([\varepsilon(n + \alpha), \varepsilon(n + 1 + \alpha))), n \in \mathbb{Z}.$$

Some of these might be empty (because we do not require that G is surjective and G only takes its values in I) and will hence be discarded. The non-empty ones form a partition of \mathbb{Z}_N because the intervals $[\varepsilon(n + \alpha), \varepsilon(n + 1 + \alpha)), n \in \mathbb{Z}$ forms a partition of the real line.

If we let \mathcal{B} be a σ -algebra then any atom of $\mathcal{B} \vee \mathcal{B}_{\varepsilon, \eta}(G)$ must be the intersection of an atom from \mathcal{B} and one from $\mathcal{B}_{\varepsilon, \eta}(G)$ and hence be a subset of $G^{-1}([\varepsilon(n + \alpha), \varepsilon(n + 1 + \alpha)))$ for some $n \in \mathbb{Z}$, so on this atom G takes its values on the interval $[\varepsilon(n + \alpha), \varepsilon(n + 1 + \alpha))$ and so

$$\begin{aligned} \|G - \mathbb{E}(G \mid \mathcal{B} \vee \mathcal{B}_{\varepsilon, \eta}(G))\|_{L^\infty} &= \sup_{x \in \mathbb{Z}_N} \|G(x) - \mathbb{E}(G \mid \mathcal{B} \vee \mathcal{B}_{\varepsilon, \eta}(G))(x)\| \\ &= \sup_{x \in \mathbb{Z}_N} \|\mathbb{E}(G(x) - G(y) \mid y \in (\mathcal{B} \vee \mathcal{B}_{\varepsilon, \eta}(G))(x))\| \\ &\leq \varepsilon \end{aligned}$$

which proves the first claim in the proposition.

Now let

$$A = G^{-1}([\varepsilon(n + \alpha), \varepsilon(n + 1 + \alpha)))$$

be an atom of $\mathcal{B}_{\varepsilon, \eta}(G)$. Notice that $G(x) \in I$ for all $x \in \mathbb{Z}_N$ so we may assume that $|\varepsilon(n + 1 + \alpha)| \leq 2^{2^{k-1}}$, so $n = O(\varepsilon^{-1})$ since A would be empty for bigger $|n|$. This proves the second claim in the proposition.

Let A and n be given as above. Now define $\psi_\eta : \mathbb{R} \rightarrow [0, 1]$ such that ψ_η is continuous and such that

$$\psi_\eta(x) = \begin{cases} 1 & \text{if } \eta \leq x \leq 1 - \eta \\ 0 & \text{if } x \leq -\eta \text{ or } x \geq 1 + \eta. \end{cases}$$

Now define $\Psi_A : \mathbb{R} \rightarrow [0, 1]$ by

$$\Psi_A(x) = \psi_\eta\left(\frac{x}{\varepsilon} - n - \alpha\right). \quad (6.2.4)$$

Recall that

$$\sum_{m \in \mathbb{Z}} \mathbb{E}(\mathbb{1}_{[\varepsilon(m-\eta+\alpha), \varepsilon(m+\eta+\alpha)]}(G(x))(\nu(x) + 1) \mid x \in \mathbb{Z}_N) = O(\eta).$$

Now if we can prove that

$$\mathbb{E}((\mathbb{1}_A - \Psi_A(G))(\nu + 1)) \leq \sum_{n \in \mathbb{Z}} \mathbb{E}(\mathbb{1}_{[\varepsilon(n-\eta+\alpha), \varepsilon(n+\eta+\alpha)]}(G(x))(\nu(x) + 1) \mid x \in \mathbb{Z}_N) \quad (6.2.5)$$

we have

$$\mathbb{E}((\mathbb{1}_A - \Psi_A(G))(\nu + 1)) = O(\eta),$$

which proves the last claim in the proposition. To prove (6.2.5) it is enough to prove that there is an $m \in \mathbb{Z}$ such that

$$\mathbb{1}_A(x) - \Psi_A(G(x)) \leq \mathbb{1}_{[\varepsilon(m-\eta+\alpha), \varepsilon(m+\eta+\alpha)]}(G(x))$$

for all $x \in \mathbb{Z}_N$ because all the terms in the sum on the RHS are non-negative.

If the LHS is 0 there is nothing to prove. The LHS is positive but ≤ 1 when $x \in A$ and $\Psi_A(G(x)) \notin [\eta, 1 - \eta]$. In that case we need the RHS to be 1 so we need $G(x) \in [\varepsilon(m - \eta + \alpha), \varepsilon(m + \eta + \alpha)]$ for some $m \in \mathbb{Z}$. But since $x \in A$ we have

$$G(x) \in [\varepsilon(n + \alpha), \varepsilon(n + 1 + \alpha)],$$

so we have $G(x) \in [\varepsilon(m - \eta + \alpha), \varepsilon(m + \eta + \alpha)]$ for $m = n$ or $m = n - 1$. This proves the claim.

Now recall that $n = O(\varepsilon^{-1})$ and $0 \leq \alpha \leq 1$ and recall that Ψ_A only depends on n, α, ε and η . Now all of these parameters are bounded and their bounds depend only on ε and η so we can find a closed and bounded (and hence compact) set $\hat{E}_{\varepsilon, \eta} \subseteq \mathbb{Z} \times \mathbb{R}^3$ such that $(n, \alpha, \varepsilon, \eta) \in \hat{E}_{\varepsilon, \eta}$. Now we can define a map $\hat{E}_{\varepsilon, \eta} \rightarrow C^0(I)$ such that $(n, \alpha, \varepsilon, \eta) \mapsto \Psi_A$ as in (6.2.4). This map is continuous (if we use the product topology on the domain where we on \mathbb{Z} can use the trivial topology where all subsets are open) so the image is also compact. Denote this image $E_{\varepsilon, \eta} \in C^0(I)$, then we have $\Psi_A \in E_{\varepsilon, \eta}$. \square

The following proposition is quite technical, but it states that given a pseudorandom measure ν and some Gowers anti-uniform functions we can define a σ -algebra \mathcal{B} from the σ -algebras we defined in proposition 6.2.1 from each of the functions. The conditional expectation of each of these functions with respect to \mathcal{B} will then not be far from the functions themselves, and furthermore there is a set in \mathcal{B} such that $\nu + 1$ will be small on this set and such that the conditional expectation of $\nu - 1$ will be small on the complement of the set.

Proposition 6.2.2. *Let ν be a k -pseudorandom measure. Let $K \geq 1$ and let F_1^*, \dots, F_K^* be Gowers anti-uniform functions. Let $0 < \varepsilon < 1$ and $0 < \eta < 1/2$ and let $\mathcal{B}_{\varepsilon, \eta}(F_j^*), j = 1, \dots, K$ be the σ -algebras given by proposition 6.2.1. Let*

$$\mathcal{B} = \bigvee_{j=1}^K \mathcal{B}_{\varepsilon, \eta}(F_j^*).$$

If $\eta < \eta_0(\varepsilon, K)$ is sufficiently small and $N > N_0(\varepsilon, K, \eta)$ is sufficiently large then

$$\|F_j^* - \mathbb{E}(F_j^* | \mathcal{B})\|_{L^\infty} \leq \varepsilon$$

for $j = 1, \dots, K$, and there exists a set $\Omega \in \mathcal{B}$ such that

$$\mathbb{E}((\nu + 1)\mathbb{1}_\Omega) = O_{K,\varepsilon}(\eta^{1/2})$$

and

$$\|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 | \mathcal{B})\|_{L^\infty} = O_{K,\varepsilon}(\eta^{1/2}).$$

Proof. The first claim follows directly from the first claim of proposition 6.2.1 since we can write

$$\mathcal{B} = \mathcal{B}_{\varepsilon,\eta}(F_j^*) \vee \left(\bigvee_{i \neq j} \mathcal{B}_{\varepsilon,\eta}(F_i^*) \right).$$

Now we proceed to the second claim. From proposition 6.2.1 we see that each $\mathcal{B}_{\varepsilon,\eta}(F_j^*)$ has $O(\varepsilon^{-1})$ atoms. Now recall that each of \mathcal{B} 's atoms is an intersection of one atom from each $\mathcal{B}_{\varepsilon,\eta}(F_j^*)$, $j = 1, \dots, K$, so we can safely claim that \mathcal{B} has no more than $O_{K,\varepsilon}(1)$ atoms. For the remainder of this proof we call an atom $A \in \mathcal{B}$ *small* if

$$\mathbb{E}((\nu + 1)\mathbb{1}_A) \leq \eta^{1/2}.$$

Now let Ω be the union of all small atoms. By the definition of σ -algebras we then have $\Omega \in \mathcal{B}$ and

$$\mathbb{E}((\nu + 1)\mathbb{1}_\Omega) = \mathbb{E} \left((\nu + 1) \sum_{A \text{ small}} \mathbb{1}_A \right) = \sum_{A \text{ small}} \mathbb{E}((\nu + 1)\mathbb{1}_A)$$

which is smaller than the number of small atoms times $\eta^{1/2}$, and since the number of atoms is $O_{K,\varepsilon}(1)$ the above expression is $O_{K,\varepsilon}(\eta^{1/2})$ which proves the second claim of the proposition.

To prove the last claim it suffices to prove that

$$\mathbb{E}(\nu(x) - 1 | x \in A) = o_{K,\varepsilon,\eta}(1) + O_{K,\varepsilon}(\eta^{1/2}), \quad (6.2.6)$$

for all atoms A which are not small, because if we let Ω' be the union of all the non small atoms then

$$\begin{aligned} \|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 | \mathcal{B})\|_{L^\infty} &= \sup_{x \in \mathbb{Z}_N} |\mathbb{1}_{\Omega'}(x)\mathbb{E}(\nu(y) - 1 | y \in \mathcal{B}(x))| \\ &= \sup_{x \in \mathbb{Z}_N} \left| \sum_{A \text{ not small}} \mathbb{1}_A(x)\mathbb{E}(\nu(y) - 1 | y \in \mathcal{B}(x)) \right| \\ &\leq \sum_{A \text{ not small}} \sup_{x \in \mathbb{Z}_N} |\mathbb{E}(\mathbb{1}_A(x)(\nu(y) - 1) | y \in \mathcal{B}(x))|. \end{aligned}$$

Now notice that the terms are zero when $x \notin A$ so we might as well take average over A instead of over $\mathcal{B}(x)$, because when $x \in A$ then $\mathcal{B}(x) = A$. Then the above expression is equal to

$$\sum_{A \text{ not small}} \sup_{x \in \mathbb{Z}_N} |\mathbb{E}((\nu(y) - 1) \mid y \in A)|$$

and we now see that if we have (6.2.6) then this is smaller than the number of non small atoms times $o_{K,\varepsilon,\eta}(1) + O_{K,\varepsilon}(\eta^{1/2})$, and since there is $O_{K,\varepsilon}(1)$ atoms we get that the above expression is $O_{K,\varepsilon}(\eta^{1/2})$ so

$$\|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 \mid \mathcal{B})\|_{L^\infty} = O_{K,\varepsilon}(\eta^{1/2}),$$

which proves the claim. So let us prove (6.2.6). Notice that

$$\mathbb{E}(\nu(x) - 1 \mid x \in A) = \frac{\mathbb{E}((\nu(x) - 1)\mathbb{1}_A(x) \mid x \in \mathbb{Z}_N)}{\#A/N} \quad (6.2.7)$$

$$= \frac{\mathbb{E}((\nu(x) - 1)\mathbb{1}_A(x) \mid x \in \mathbb{Z}_N)}{\mathbb{E}(\mathbb{1}_A)}, \quad (6.2.8)$$

and

$$\mathbb{E}((\nu(x) - 1)\mathbb{1}_A(x) \mid x \in \mathbb{Z}_N) + 2\mathbb{E}(\mathbb{1}_A) = \mathbb{E}((\nu(x) + 1)\mathbb{1}_A(x) \mid x \in \mathbb{Z}_N) > \eta^{1/2} \quad (6.2.9)$$

because A is *not* small. Now we claim that we can use these two statements to prove (6.2.6) if we have

$$\mathbb{E}((\nu(x) - 1)\mathbb{1}_A(x) \mid x \in \mathbb{Z}_N) = o_{K,\varepsilon}(1) + O_{K,\varepsilon}(\eta) \quad (6.2.10)$$

because then we can use (6.2.7) to get

$$\mathbb{E}(\nu(x) - 1 \mid x \in A) = \frac{o_{K,\varepsilon}(1) + O_{K,\varepsilon}(\eta)}{\mathbb{E}(\mathbb{1}_A)},$$

and then use (6.2.9) to see that

$$\mathbb{E}(\mathbb{1}_A) > 2(\mathbb{E}(\nu(x) - 1 \mid x \in A) + \eta^{1/2}),$$

and combining these we get

$$\mathbb{E}(\nu(x) - 1 \mid x \in A) < \frac{o_{K,\varepsilon}(1) + O_{K,\varepsilon}(\eta)}{2(\mathbb{E}(\nu(x) - 1 \mid x \in A) + \eta^{1/2})} = (o_{K,\varepsilon,\eta}(1) + O_{K,\varepsilon}(\eta^{1/2}))$$

when N is sufficiently large and η is sufficiently small. So now we need to prove (6.2.10) to finish the proof.

Since A is an atom of \mathcal{B} we know that $A = A_1 \cap \dots \cap A_K$ where A_i is an atom in $\mathcal{B}_{\varepsilon,\eta}(F_i^*)$ for $i = 1, \dots, K$. So for each i we get from proposition 6.2.1 that there is a continuous

function $\Psi_{A_i} : I \rightarrow [0, 1]$, because we get from lemma 5.1.3 that F_i^* takes its values in I , such that

$$\mathbb{E} (|(\mathbb{1}_{A_i} - \Psi_{A_i}(F_i^*))(\nu + 1)|) = O(\eta).$$

Now define $\Psi_A : I^K \rightarrow [0, 1]$ by

$$\Psi_A(x_1, \dots, x_K) = \Psi_{A_1}(x_1) \cdots \Psi_{A_K}(x_K).$$

Then

$$\mathbb{E} (|(\mathbb{1}_A - \Psi_A(F_1^*, \dots, F_K^*))(\nu + 1)|) = \mathbb{E} (|(\mathbb{1}_{A_1} \cdots \mathbb{1}_{A_K} - \Psi_{A_1}(F_1^*) \cdots \Psi_{A_K}(F_K^*))(\nu + 1)|).$$

We now want to prove that this is $O_K(\eta)$. By induction it is enough to prove that it is true for $K = 2$. Completing the square we get that

$$\begin{aligned} \mathbb{E} (|(\mathbb{1}_{A_1} \mathbb{1}_{A_2} - \Psi_{A_1}(F_1^*) \Psi_{A_2}(F_2^*))(\nu + 1)|) &\leq \mathbb{E} (|(\mathbb{1}_{A_1} - \Psi_{A_1}(F_1^*))(\mathbb{1}_{A_2} - \Psi_{A_2}(F_2^*))|(\nu + 1)) \\ &+ \mathbb{E} (|\mathbb{1}_{A_1} - \Psi_{A_1}(F_1^*)| \Psi_{A_2}(F_2^*)) + \mathbb{E} (|\mathbb{1}_{A_2} - \Psi_{A_2}(F_2^*)| \Psi_{A_1}(F_1^*)). \end{aligned} \quad (6.2.11)$$

The first term can be bounded using Cauchy-Schwarz by

$$(\mathbb{E} ((\mathbb{1}_{A_1} - \Psi_{A_1}(F_1^*))^2) \mathbb{E} ((\mathbb{1}_{A_2} - \Psi_{A_2}(F_2^*))^2))^{1/2},$$

And since $|\mathbb{1}_{A_i} - \Psi_{A_i}(F_i^*)| \leq 1$ for $i = 1, 2$ this is less than

$$(\mathbb{E} (|\mathbb{1}_{A_1} - \Psi_{A_1}(F_1^*)|) \mathbb{E} (|\mathbb{1}_{A_2} - \Psi_{A_2}(F_2^*)|))^{1/2},$$

which is $O(\eta)$. We now need to consider the last two terms of (6.2.11). Let us consider the second term. Since $\Psi_{A_2}(F_2^*) \leq 1$ we have

$$\mathbb{E} (|\mathbb{1}_{A_1} - \Psi_{A_1}(F_1^*)| \Psi_{A_2}(F_2^*)) \leq \mathbb{E} (|\mathbb{1}_{A_1} - \Psi_{A_1}(F_1^*)|).$$

This is $O(\eta)$ and the same argument shows that the third term of (6.2.11) also is $O(\eta)$. This finishes the induction step and hence the proof of the claim.

Since $\mathbb{1}_A - \Psi_A(F_1^*, \dots, F_K^*) \geq 0$ we have

$$\mathbb{E} (|\nu - 1| |\mathbb{1}_A - \Psi_A(F_1^*, \dots, F_K^*)|) = O_K(\eta).$$

As all Ψ_{A_i} ranges over a compact set in $C^0(I^K)$ independent of A_i , Ψ_A must also range over a compact set so from proposition 5.2.4 we get that

$$\mathbb{E} ((\nu - 1) \Psi_A(F_1^*, \dots, F_K^*)) = o_{K,E}(1),$$

where $E = E_{\varepsilon, \eta} \subset C^0(I^K)$ is the compact subset Ψ_A ranges over. So we get

$$\mathbb{E} ((\nu - 1) \mathbb{1}_A) = O_K(\eta) - o_{K,E}(1) = O_K(\eta) + o_{\varepsilon, \eta, K}(1),$$

as we wanted. Now η can be picked arbitrarily small depending on ε and K , so this is

$$O_{K, \varepsilon}(\eta) + o_{\varepsilon, K}(1)$$

which finishes the proof. \square

Chapter 7

Furstenberg Tower

7.1 The Furstenberg tower

First we will prove the following rather technical proposition. This will be used to prove proposition 7.1.2 which will be just what we need to finish the proof of Szemerédi's theorem in pseudorandom measures. The propositions of this chapter are, as the von Neumann theorem in chapter 4, inspired by the ergodic theoretical proof of Szemerédi's theorem [3].

Proposition 7.1.1. *Let ν be a k -pseudorandom measure, and let $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ such that*

$$0 \leq f(x) \leq \nu(x)$$

for all $x \in \mathbb{Z}_N$. Let $0 < \varepsilon < 1$ be small and let $K \geq 0$ be an integer. Assume that $\eta \leq \eta_0(\varepsilon, K)$ is sufficiently small and that $N > N_0(\varepsilon, K, \eta)$ is sufficiently large. Let $F_1, \dots, F_K : \mathbb{Z}_N \rightarrow \mathbb{R}$ satisfy

$$|F_j(x)| \leq (1 + O_{K,\varepsilon}(\eta^{1/2}))(\nu(x) + 1) \quad (7.1.1)$$

for all $1 \leq j \leq K$ and $x \in \mathbb{Z}_N$. Let

$$\mathcal{B}_K = \mathcal{B}_{\varepsilon,\eta}(F_1^*) \vee \dots \vee \mathcal{B}_{\varepsilon,\eta}(F_K^*). \quad (7.1.2)$$

where $\mathcal{B}_{\varepsilon,\eta}(F_j^*)$, $1 \leq j \leq K$ is as defined in proposition 6.2.1 (note that for this to be well-defined, we need to prove that F_1^*, \dots, F_K^* all take values in the interval $[-2^{2^{k-1}}, 2^{2^{k-1}}]$ – this will be proved in 3. of the proof of this proposition). Suppose that there exists $\Omega_K \in \mathbb{Z}_N$ such that

$$\mathbb{E}((\nu + 1)\mathbb{1}_{\Omega_K}) = O_{K,\varepsilon}(\eta^{1/2}) \quad (7.1.3)$$

and

$$\|(1 - \mathbb{1}_{\Omega_K})\mathbb{E}(\nu - 1 \mid \mathcal{B}_K)\|_{L^\infty} = O_{K,\varepsilon}(\eta^{1/2}). \quad (7.1.4)$$

Now define

$$F_{K+1} = (1 - \Omega_K)(f - \mathbb{E}(f \mid \mathcal{B}_K))$$

and

$$\mathcal{B}_{K+1} = \mathcal{B}_K \vee \mathcal{B}_{\varepsilon,\eta}(F_{K+1}^*),$$

and suppose that $\|F_{K+1}\|_{U^{k-1}} > \varepsilon^{1/2^k}$. Then there exists a set $\Omega_{K+1} \supseteq \Omega_K$ such that the following is true.

1. $\|(1 - \mathbb{1}_{\Omega_K})\mathbb{E}(f \mid \mathcal{B}_K)\|_\infty \leq 1 + O_{K,\varepsilon}(\eta^{1/2})$,
2. $|F_{K+1}(x)| \leq (1 + O_{K,\varepsilon}(\eta^{1/2}))(\nu(x) + 1)$,
3. $\mathbb{E}((\nu + 1)\mathbb{1}_{\Omega_{K+1}}) = O_{K,\varepsilon}(\eta^{1/2})$,
4. $\|(1 - \mathbb{1}_{\Omega_{K+1}})\mathbb{E}(\nu - 1 \mid \mathcal{B}_{K+1})\|_{L^\infty} = O_{K,\varepsilon}(\eta^{1/2})$,
5. $\mathbb{E}\left(|(1 - \mathbb{1}_{\Omega_{K+1}})\mathbb{E}(f \mid \mathcal{B}_{K+1})|^2\right) \geq \mathbb{E}\left(|(1 - \mathbb{1}_{\Omega_K})\mathbb{E}(f \mid \mathcal{B}_K)|^2\right) + 2^{-2^k+1}\varepsilon$.

Proof. Let everything be given as in the theorem. Let us prove the 5 statements one at a time.

1. From the triangle inequality and (7.1.4) we get that

$$\begin{aligned} \|(1 - \mathbb{1}_{\Omega_K})\mathbb{E}(\nu \mid \mathcal{B}_K)\|_{L^\infty} &\leq \|(1 - \mathbb{1}_{\Omega_K})\mathbb{E}(\nu - 1 \mid \mathcal{B}_K)\|_{L^\infty} + \|1 - \mathbb{1}_{\Omega_K}\|_{L^\infty} \\ &= 1 + O_{K,\varepsilon}(\eta^{1/2}). \end{aligned}$$

We have $0 \leq f(x) \leq \nu(x)$ for all $x \in \mathbb{Z}_N$ so

$$\|(1 - \mathbb{1}_{\Omega_K})\mathbb{E}(f \mid \mathcal{B}_K)\|_{L^\infty} \leq \|(1 - \mathbb{1}_{\Omega_K})\mathbb{E}(\nu \mid \mathcal{B}_K)\|_{L^\infty} = 1 + O_{K,\varepsilon}(\eta^{1/2}),$$

which concludes **1**.

2. From the definition of F_{K+1} we get

$$\begin{aligned} |F_{K+1}(x)| &= |(1 - \mathbb{1}_{\Omega_K}(x))(f(x) - \mathbb{E}(f \mid \mathcal{B}_K)(x))| \\ &\leq |(1 - \mathbb{1}_{\Omega_K}(x))f(x)| + |(1 - \mathbb{1}_{\Omega_K}(x))\mathbb{E}(f \mid \mathcal{B}_K)(x)| \end{aligned}$$

for all $x \in \mathbb{Z}_N$. Since $0 \leq f(x) \leq \nu(x)$ and $\nu(x) \geq 0$ for all $x \in \mathbb{Z}_N$ we get

$$|(1 - \mathbb{1}_{\Omega_K}(x))f(x)| \leq |(1 - \mathbb{1}_{\Omega_K}(x))\nu(x)| = (1 - \mathbb{1}_{\Omega_K}(x))\nu(x)$$

and from **1**. we get that

$$|(1 - \mathbb{1}_{\Omega_K}(x))\mathbb{E}(f \mid \mathcal{B}_K)(x)| \leq 1 + O_{K,\varepsilon}(\eta^{1/2})$$

for all $x \in \mathbb{Z}_N$. Putting these together we get

$$|F_{K+1}(x)| \leq (1 - \mathbb{1}_{\Omega_K}(x))\nu(x) + 1 + O_{K,\varepsilon}(\eta^{1/2}).$$

And since $1 - \mathbb{1}_{\Omega_K}(x) \leq 1 + O_{K,\varepsilon}(\eta^{1/2})$ we get

$$|F_{K+1}(x)| \leq (1 + O_{K,\varepsilon}(\eta^{1/2}))(1 + \nu(x))$$

as desired.

3. First we need to see that if $a \in \mathbb{R}$ we have

$$(aF)^* = \mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k-1}, \omega \neq 0} aF(x + \omega \cdot h) \mid h \in \mathbb{Z}_N^{k-1} \right) = a^{2^{k-1}-1} F^*. \quad (7.1.5)$$

From the assumptions and **2.** we get that

$$|F_j(x)| \leq (1 + O_{K,\varepsilon}(\eta^{1/2}))(1 + \nu(x))$$

for $0 \leq j \leq K+1$. Now let $0 \leq j \leq K+1$ be given. We define $\tilde{F}_j = F_j / (1 + O_{K,\varepsilon}(\eta^{1/2}))$ such that

$$|\tilde{F}_j(x)| \leq 1 + \nu(x)$$

and by lemma 5.1.3 we then get

$$\|\tilde{F}_j^*\|_{L^\infty} \leq 2^{2^{k-1}-1} + o(1)$$

and (7.1.5) then gives us

$$\|F_j^*\|_{L^\infty} \leq (1 + O_{K,\varepsilon}(\eta^{1/2}))^{2^{k-1}-1} (2^{2^{k-1}-1} + o(1)).$$

Now all constants are allowed to depend on k , so we have

$$(1 + O_{K,\varepsilon}(\eta^{1/2}))^{2^{k-1}-1} = 1 + O_{K,\varepsilon}(\eta^{1/2})$$

since $\eta < 1$. We can pick N as large as we want depending on K, ε and η , so the term $o(1)$ can be put inside the $O_{K,\varepsilon}(\eta^{1/2})$ -term, and we have

$$2^{2^{k-1}-1} O_{K,\varepsilon}(\eta^{1/2}) = O_{K,\varepsilon}(\eta^{1/2})$$

again because the constant is allowed to depend on k . So all in all we get

$$\|F_j^*\|_{L^\infty} \leq 2^{2^{k-1}-1} + O_{K,\varepsilon}(\eta^{1/2}), \quad (7.1.6)$$

and this is true for all $0 \leq j \leq K+1$. So for sufficiently small η and ε we can apply proposition 6.2.2 on these functions – recall that they had to take values in the interval $[-2^{2^{k-1}}, 2^{2^{k-1}}]$, so this ensures that we can define the σ -algebras like in proposition 6.2.1. From proposition 6.2.2 we now get a set $\Omega \in \mathcal{B}_{K+1}$ such that

$$\mathbb{E}((\nu + 1)\mathbb{1}_\Omega) = O_{K,\varepsilon}(\eta^{1/2})$$

so if we now define $\Omega_{K+1} = \Omega_K \cup \Omega$, then this finishes the proof of **3.** because

$$\mathbb{E}((\nu + 1)\mathbb{1}_{\Omega_{K+1}}) \leq \mathbb{E}((\nu + 1)\mathbb{1}_\Omega) + \mathbb{E}((\nu + 1)\mathbb{1}_{\Omega_K}) = O_{K,\varepsilon}(\eta^{1/2}).$$

4. The application of proposition 6.2.2 in the proof of **3.** also gives us

$$\|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 \mid \mathcal{B}_{K+1})\|_{L^\infty} = O_{K,\varepsilon}(\eta^{1/2}).$$

Since the above expression can be seen as the maximum of $\mathbb{E}(\nu - 1 \mid \mathcal{B}_{K+1})$ on the set $\mathbb{Z}_N \setminus \Omega$ and we have

$$\mathbb{Z}_N \setminus \Omega_{K+1} \subseteq \mathbb{Z}_N \setminus \Omega$$

we get that

$$\|(1 - \mathbb{1}_{\Omega_{K+1}})\mathbb{E}(\nu - 1 \mid \mathcal{B}_{K+1})\|_{L^\infty} = O_{K,\varepsilon}(\eta^{1/2}).$$

5. From the definition of F_{K+1} we have

$$|\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})(f - \mathbb{E}(f \mid \mathcal{B}_K))F_{K+1}^*)| = |\mathbb{E}(F_{K+1} \cdot F_{K+1}^*)|$$

and by lemma 5.1.3 we get that

$$|\mathbb{E}(F_{K+1}F_{K+1}^*)| = \|F_{K+1}\|_{U^{k-1}}^{2^{k-1}}.$$

Recall that we assumed $\|F_{K+1}\|_{U^{k-1}} > \varepsilon^{1/2^k}$, which all in all gives us

$$|\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})(f - \mathbb{E}(f \mid \mathcal{B}_K))F_{K+1}^*)| \geq \varepsilon^{1/2}. \quad (7.1.7)$$

On the other hand we have

$$\begin{aligned} & |\mathbb{E}((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})(f - \mathbb{E}(f \mid \mathcal{B}_K))F_{K+1}^*)| \\ & \leq \mathbb{E}((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})|f - \mathbb{E}(f \mid \mathcal{B}_K)||F_{K+1}^*|) \\ & \leq \|F_{K+1}^*\|_{L^\infty} \mathbb{E}((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})|f - \mathbb{E}(f \mid \mathcal{B}_K)|). \end{aligned}$$

Since η is small (in particular $\eta < 1$) we get from (7.1.6) that $\|F_{K+1}^*\|_{L^\infty} = O_{K,\varepsilon}(1)$, so the above expression is equal to

$$O_{K,\varepsilon}(1)\mathbb{E}((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})|f - \mathbb{E}(f \mid \mathcal{B}_K)|). \quad (7.1.8)$$

Since $\Omega_{K+1} \setminus \Omega_K \subseteq \mathbb{Z}_N \setminus \Omega_K$, we have

$$f(x) - \mathbb{E}(f \mid \mathcal{B}_K)(x) = F_{K+1}(x)$$

for all $x \in \Omega_{K+1} \setminus \Omega_K$, so for all $x \in \mathbb{Z}_N$ we have

$$(\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})|f(x) - \mathbb{E}(f \mid \mathcal{B}_K)(x)| = (\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})|F_{K+1}(x)|$$

which by **2.** is $(1 + O_{K,\varepsilon}(\eta^{1/2}))(\nu(x) + 1)$, and since $\eta < 1$ this is $O_{K,\varepsilon}(1)(\nu(x) + 1)$. So (7.1.8) can be rewritten as

$$O_{K,\varepsilon}(1)\mathbb{E}((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})(\nu + 1)),$$

and from **3.** and (7.1.3) we get that this is $O_{K,\varepsilon}(\eta^{1/2})$. Putting everything together we get

$$|\mathbb{E}((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})(f - \mathbb{E}(f | \mathcal{B}_K))F_{K+1}^*)| \leq O_{K,\varepsilon}(\eta^{1/2}). \quad (7.1.9)$$

We also have

$$\begin{aligned} & |\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})(f - \mathbb{E}(f | \mathcal{B}_K))(F_{K+1}^* - \mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1})))| \\ & \leq \|F_{K+1}^* - \mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1})\|_{L^\infty} \mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})|f - \mathbb{E}(f | \mathcal{B}_K)|), \end{aligned}$$

and when we in the proof of **3.** applied proposition 6.2.2 on F_{K+1}^* we also got that for any σ -algebra \mathcal{B} we have

$$\|F_{K+1}^* - \mathbb{E}(F_{K+1}^* | \mathcal{B} \vee \mathcal{B}_{\varepsilon,\eta}(F_{K+1}^*))\|_{L^\infty} \leq \varepsilon,$$

so if we let $\mathcal{B} = \mathcal{B}_K$ we get

$$\|F_{K+1}^* - \mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1})\|_{L^\infty} \leq \varepsilon. \quad (7.1.10)$$

As before we can bound $f - \mathbb{E}(f | \mathcal{B}_K)$ by $(O_{\varepsilon,K}(\eta^{1/2}) + 1)(\nu + 1)$ and thus get that

$$\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})|f - \mathbb{E}(f | \mathcal{B}_K)|) \leq (O_{\varepsilon,K}(\eta^{1/2}) + 1)\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})(\nu + 1)).$$

Now since ν is a measure, this is smaller than $(O_{\varepsilon,K}(\eta^{1/2}) + 1)(2 + o(1))$, so using this and (7.1.10) we get that

$$\begin{aligned} & |\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})(f - \mathbb{E}(f | \mathcal{B}_K))(F_{K+1}^* - \mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1})))| \\ & \leq \varepsilon(O_{\varepsilon,K}(\eta^{1/2}) + 1)(2 + o(1)) = O(\varepsilon), \end{aligned} \quad (7.1.11)$$

if we pick η sufficiently small compared to ε and K .

Combining (7.1.7), (7.1.9) and (7.1.11) with the triangle inequality we get

$$|\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})(f - \mathbb{E}(f | \mathcal{B}_K))\mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1}))| \geq \varepsilon^{1/2} - O_{K,\varepsilon}(\eta^{1/2}) - O(\varepsilon).$$

Now the LHS here is equal to

$$\mathbb{E}(\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})(f - \mathbb{E}(f | \mathcal{B}_K))\mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1}) | \mathcal{B}_{K+1}))$$

by proposition 6.1.5, and since $1 - \mathbb{1}_{\Omega_{K+1}}$, $\mathbb{E}(f | \mathcal{B}_K)$ and $\mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1})$ all are \mathcal{B}_{K+1} measurable, this is equal to

$$\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})(\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K))\mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1})).$$

By Cauchy-Schwarz (corollary 2.5.4) this is smaller than

$$\mathbb{E}((1 - \mathbb{1}_{\Omega_{K+1}})^2(\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K))^2)^{1/2} \mathbb{E}(\mathbb{E}(F_{K+1}^* | \mathcal{B}_{K+1})^2)^{1/2},$$

and since

$$\mathbb{E} \left(\mathbb{E} (F_{K+1}^* | \mathcal{B}_{K+1})^2 \right)^{1/2} = \mathbb{E} (F_{K+1}^*)^{1/2} \leq \|F_{K+1}^*\|_{L^\infty} \leq 2^{2^{k-1}-1} + O_{K,\varepsilon}(\eta^{1/2})$$

by (7.1.6), so we get

$$\mathbb{E} \left((1 - \mathbb{1}_{\Omega_{K+1}})^2 (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K))^2 \right)^{1/2} \geq 2^{-2^{k-1}+1} \varepsilon^{1/2} - O_{K,\varepsilon}(\eta^{1/2}) - O(\varepsilon).$$

Notice that the LHS is an L^2 norm as we defined it in the section about function spaces in chapter 1. So we can write the above expression as

$$\| (1 - \mathbb{1}_{\Omega_{K+1}}) (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K)) \|_{L^2} \geq 2^{-2^{k-1}+1} \varepsilon^{1/2} - O_{K,\varepsilon}(\eta^{1/2}) - O(\varepsilon). \quad (7.1.12)$$

If η is sufficiently small relative to K and ε then $\mathbb{E}(f | \mathcal{B}_K) \leq 2$ outside Ω_K by **1**. and we get

$$\begin{aligned} \| (\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K) \|_{L^2}^2 &\leq 2 \| \mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K} \|_{L^2}^2 \\ &= 2 \mathbb{E} (\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) \\ &\leq 2 \mathbb{E} (\mathbb{1}_{\Omega_{K+1}}) \\ &\leq 2 \mathbb{E} ((\nu + 1) \mathbb{1}_{\Omega_{K+1}}) \\ &= O_{K,\varepsilon}(\eta^{1/2}), \end{aligned}$$

where we in the last equality use **3**. Using this it is enough to prove that

$$\begin{aligned} \| (1 - \mathbb{1}_{\Omega_{K+1}}) \mathbb{E}(f | \mathcal{B}_{K+1}) \|_{L^2}^2 \\ \geq \| (1 - \mathbb{1}_{\Omega_{K+1}}) \mathbb{E}(f | \mathcal{B}_K) \|_{L^2}^2 + 2^{-2^k+2} \varepsilon - O_{K,\varepsilon}(\eta^{1/2}) - O(\varepsilon^{3/2}) \end{aligned} \quad (7.1.13)$$

to conclude **5**. and hence the proof of the proposition. The difference between this and **5**. is the L^2 notation, that we have Ω_{K+1} instead of Ω_K on the RHS and the extra error terms. The L^2 notation does not make any difference, but will be a helpful way of writing it, and the error terms can be absorbed by the $2^{-2^k+2} \varepsilon$ term, if we pick ε sufficiently small with respect to k and η sufficiently small with respect to K and ε . The last difference can be overcome by using the calculations from before and the triangle inequality to get

$$\begin{aligned} \| (1 - \mathbb{1}_{\Omega_{K+1}}) \mathbb{E}(f | \mathcal{B}_K) \|_{L^2}^2 &\geq \| ((1 - \mathbb{1}_{\Omega_{K+1}}) + (\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K})) \mathbb{E}(f | \mathcal{B}_K) \|_{L^2}^2 \\ &\quad - \| (\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K) \|_{L^2}^2 = \| (1 - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K) \|_{L^2}^2 - O_{K,\varepsilon}(\eta^{1/2}), \end{aligned}$$

so let us prove (7.1.13). The LHS can be written as

$$\| (1 - \mathbb{1}_{\Omega_{K+1}}) \mathbb{E}(f | \mathcal{B}_K) + (1 - \mathbb{1}_{\Omega_{K+1}}) (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K)) \|_{L^2}^2$$

and then expanded as

$$\begin{aligned} \| (1 - \mathbb{1}_{\Omega_{K+1}}) \mathbb{E}(f | \mathcal{B}_K) \|_{L^2}^2 &+ \| (1 - \mathbb{1}_{\Omega_{K+1}}) (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K)) \|_{L^2}^2 \\ &+ 2 \mathbb{E} \left((1 - \mathbb{1}_{\Omega_{K+1}}) \mathbb{E}(f | \mathcal{B}_K) (1 - \mathbb{1}_{\Omega_{K+1}}) (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K)) \right). \end{aligned}$$

The first term is the term we want and the second term has been proven to be greater than

$$(2^{-2^{k-1}+1}\varepsilon^{1/2} - O_{K,\varepsilon}(\eta^{1/2}) - O(\varepsilon))^2 = 2^{-2^k+2}\varepsilon - O_{K,\varepsilon}(\eta^{1/2}) - O(\varepsilon),$$

so if we can prove that the last term is $O_{K,\varepsilon}(\eta^{1/2})$ we are done. Notice that $(1 - \mathbb{1}_{\Omega_{K+1}})^2 = (1 - \mathbb{1}_{\Omega_{K+1}})$ so we can rewrite the last term as

$$\mathbb{E} \left((1 - \mathbb{1}_{\Omega_{K+1}}) \mathbb{E}(f | \mathcal{B}_K) (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K)) \right). \quad (7.1.14)$$

Now we see that $(1 - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K)$ is \mathcal{B}_K -measurable and hence in particular \mathcal{B}_{K+1} -measurable, since \mathcal{B}_K is a sub- σ -algebra of \mathcal{B}_{K+1} so from corollary 6.1.7 we get

$$\mathbb{E} \left((1 - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K) (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K)) \right) = 0$$

because we have

$$\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K) = \mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(\mathbb{E}(f | \mathcal{B}_K) | \mathcal{B}_{K+1}).$$

Subtracting this from (7.1.14) we get

$$\mathbb{E} \left((\mathbb{1}_{\Omega_K} - \mathbb{1}_{\Omega_{K+1}}) \mathbb{E}(f | \mathcal{B}_K) (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K)) \right).$$

Since $\Omega_K \subseteq \Omega_{K+1}$ this might be negative, so we continue our calculations WLOG (since we are only interested in bounding it) with the negative of this, namely

$$\mathbb{E} \left((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K) (\mathbb{E}(f | \mathcal{B}_{K+1}) - \mathbb{E}(f | \mathcal{B}_K)) \right). \quad (7.1.15)$$

Now $(\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K)$ is \mathcal{B}_{K+1} -measurable, so if we once again use corollary 6.1.7 we get

$$\mathbb{E} \left((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K) (f - \mathbb{E}(f | \mathcal{B}_{K+1})) \right) = 0,$$

and adding this to (7.1.15) gives us

$$\mathbb{E} \left((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) \mathbb{E}(f | \mathcal{B}_K) (f - \mathbb{E}(f | \mathcal{B}_K)) \right). \quad (7.1.16)$$

If $x \notin \Omega_K$ we get from **1.** that if η is sufficiently small with respect to ε and K we have $\mathbb{E}(f | \mathcal{B}_K)(x) = O_{K,\varepsilon}(\eta^{1/2}) \leq 2$, so the above expression can in absolute value be bounded from above by

$$2\mathbb{E} \left((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) |f - \mathbb{E}(f | \mathcal{B}_K)| \right).$$

Now $0 \leq f(x) \leq \nu(x)$ by assumption so this is smaller than

$$2\mathbb{E} \left((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) (\nu + \mathbb{E}(\nu | \mathcal{B}_K)) \right),$$

and from (7.1.4) we get that if $x \notin \Omega_K$ then $\mathbb{E}(\nu | \mathcal{B}_K)(x) = 1 + O_{K,\varepsilon}(\eta^{1/2}) \leq 2$ when η is sufficiently small with respect to ε and K , so the above expression is bounded by

$$4\mathbb{E} \left((\mathbb{1}_{\Omega_{K+1}} - \mathbb{1}_{\Omega_K}) (\nu + 1) \right)$$

since ν is positive. This expression is $O_{\varepsilon,K}(\eta^{1/2})$ by (7.1.3) and **3.** which concludes the proof. \square

The above proposition will be used to prove the following proposition in an algorithmic way. As mentioned this proof is inspired by Furstenberg's ergodic theoretical proof of Szemerédi's theorem [3], but in that case the algorithm was not guaranteed to terminate and they needed the axiom of choice to conclude the proof. But in our case the algorithm is guaranteed to terminate in a bounded number of steps.

Proposition 7.1.2. *Let ν be a k -pseudorandom measure, and let $f : \mathbb{Z}_N \rightarrow \mathbb{R}$ satisfy*

$$0 \leq f(x) \leq \nu(x)$$

for all $x \in \mathbb{Z}_N$. Let $0 < \varepsilon < 1$ and assume that $N > N_0(\varepsilon)$ is sufficiently large. Then there exist a σ -algebra \mathcal{B} and a set $\Omega \in \mathcal{B}$ such that

1. $\mathbb{E}(\nu \mathbb{1}_\Omega) = o_\varepsilon(1)$,
2. $\|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 \mid \mathcal{B})\|_{L^\infty} = o_\varepsilon(1)$,
3. $\|(1 - \mathbb{1}_\Omega)(f - \mathbb{E}(f \mid \mathcal{B}))\|_{U^{k-1}} \leq \varepsilon^{1/2^k}$.

Proof. We will prove this by constructing an algorithm. It starts out with the trivial σ -algebra and a set in it (actually \emptyset) that satisfies **1.** and **2.**, and if it does not satisfy **3.** we iterate using the preceding proposition to extend our σ -algebra and find a set in it such that **1.** and **2.** are still satisfied, and this time maybe also **3.**. We continue like this until all conditions are satisfied.

So we need to specify the technical details of the algorithm, and argue that it terminates. Let $\varepsilon > 0$ be given, and let $K_0 = \lceil 2^{2^k} / \varepsilon + 1 \rceil$. We will pick the parameter $0 < \eta < \varepsilon$ later, but it is assumed to be sufficiently small depending on ε and K_0 .

Step 1 Initialize $K = 0, \Omega_0 = \emptyset, \mathcal{B}_0 = \{\emptyset, \mathbb{Z}_N\}$ and $F_1 = f - \mathbb{E}(f)$.

Step 2 If we have $\|F_{K+1}\|_{(U^{k-1})^*} > \varepsilon^{1/2^k}$ then define

$$\mathcal{B}_{K+1} = \mathcal{B}_K \vee \mathcal{B}_{\varepsilon, \eta}(F_{K+1}^*)$$

and use proposition 7.1.1 to get a set $\Omega_{K+1} \supseteq \Omega_K$ in \mathcal{B}_{K+1} . Otherwise let $\Omega = \Omega_K$ and $\mathcal{B} = \mathcal{B}_K$ and terminate the algorithm. Define $F_{K+2} = (1 - \mathbb{1}_{\Omega_{K+1}})(f - \mathbb{E}(f \mid \mathcal{B}_{K+1}))$.

Step 3 Now increment K to $K + 1$. If $K > K_0$, we terminate with an error. Otherwise goto **Step 2**.

Let us consider the conditions **1.** and **2.** with error term $O_{\varepsilon, K}(\eta^{1/2})$ on the RHS instead of $o_\varepsilon(1)$. Because if we let η be sufficiently small then we can replace $O_{\varepsilon, K}(\eta^{1/2})$ by $o_\varepsilon(1)$. We can ignore K since $K \leq K_0$ which only depends on k and ε . So we are considering the alternative conditions

- 1.' $\mathbb{E}(\nu \mathbb{1}_\Omega) = O_{\varepsilon, K}(\eta^{1/2})$,

$$\mathbf{2.}' \quad \|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 \mid \mathcal{B})\|_{L^\infty} = O_{\varepsilon, K}(\eta^{1/2}).$$

First of all we need to prove that conditions $\mathbf{1.}'$ and $\mathbf{2.}'$ valid with the initial values, and then that it is preserved after an iteration. When $K = 0$ we have $\mathbb{E}(\nu \mathbb{1}_\Omega) = 0$ which certainly is $O_{\varepsilon, K}(\eta^{1/2})$ and

$$\|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 \mid \mathcal{B}_0)\|_{L^\infty} = \|\mathbb{E}(\nu - 1)\|_{L^\infty} = \mathbb{E}(\nu - 1) = o(1) = O_{\varepsilon, K}(\eta^{1/2})$$

for sufficiently large N by the definition of a measure. So both conditions $\mathbf{1.}$ and $\mathbf{2.}$ are satisfied after the initialization. Now assume that the conditions $\mathbf{1.}'$ and $\mathbf{2.}'$ are satisfied for K . We now want to prove that they are satisfied for $K + 1$. We got the Ω_{K+1} and \mathcal{B}_{K+1} from proposition 7.1.1, and this proposition also gives us, that $\mathbf{1.}'$ and $\mathbf{2.}'$ are satisfied for $K + 1$.

We also see that if the algorithm terminates without an error we have proven the proposition, because then $\mathbf{3.}$ is also satisfied. So what is left to be proven is that the algorithm terminates without an error – so it terminates after at most K_0 steps. So assume that we have reached the K_0 'th step. Now define for $0 \leq K \leq K_0$

$$E_K = \|(1 - \mathbb{1}_{\Omega_K})\mathbb{E}(f \mid \mathcal{B}_K)\|_{L^2}^2.$$

From $\mathbf{5.}$ in proposition 7.1.1 we get that

$$E_{K+1} \geq E_K + 2^{-2^k+1}\varepsilon$$

for all $0 \leq K \leq K_0$ so

$$E_{K_0} \geq K_0 2^{-2^k+1}\varepsilon \geq 2$$

But from $\mathbf{1.}$ in proposition 7.1.1 we also get

$$E_K \leq 1 + O_{K, \varepsilon}(\eta^{1/2})$$

for all K , so for sufficiently small η depending on N and ε (again we can ignore the dependency on K because $K \leq K_0$) these two bounds contradict each other. This concludes the proof. \square

7.2 Proof of Szemerédi in pseudorandom measures

Now let us return to the proof of theorem 3.1.2. Let us first restate the theorem

Theorem 7.2.1 (Szemerédi's theorem in pseudorandom measures). *Let $k \geq 3$ and $0 < \delta \leq 1$. Let $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ be a k -pseudorandom measure and let $f : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ satisfy*

$$0 \leq f(x) \leq \nu(x)$$

for all $x \in \mathbb{Z}_N$ and

$$\mathbb{E}(\langle \cdot, \cdot \rangle f) \leq \delta.$$

Then we have

$$\mathbb{E}(\prod_{i=0}^{k-1} f(x+ir) \mid x, r \in \mathbb{Z}_N) \geq c(k, \delta) - o_{k, \delta}(1),$$

for some constant $c(k, \delta)$.

Proof. Let f and δ be as in the theorem. Let $0 < \varepsilon < \delta$ be a parameter which we will specify later. From proposition 7.1.2 we now get a σ -algebra \mathcal{B} and a set Ω . Now let

$$f_U = (1 - \mathbb{1}_\Omega)(f - \mathbb{E}(f \mid \mathcal{B}))$$

and

$$f_{U^\perp} = (1 - \mathbb{1}_\Omega)\mathbb{E}(f \mid \mathcal{B}).$$

Since $\mathbb{Z}_N \setminus \Omega \in \mathcal{B}$ we can use proposition 6.1.5 to get

$$\mathbb{E}(f_{U^\perp}) = \mathbb{E}((1 - \mathbb{1}_\Omega)\mathbb{E}(f \mid \mathcal{B})) = \mathbb{E}((1 - \mathbb{1}_\Omega)f)$$

and from the assumptions on f and 1. in proposition 7.1.2 we get that

$$\mathbb{E}((1 - \mathbb{1}_\Omega)f) \geq \mathbb{E}(f) - \mathbb{E}(\nu \mathbb{1}_\Omega) \geq \delta - o_\varepsilon(1)$$

so

$$\mathbb{E}(f_{U^\perp}) \geq \delta - o_\varepsilon(1). \quad (7.2.1)$$

From the triangle inequality and proposition 7.1.2 we get that

$$\begin{aligned} \|f_{U^\perp}\|_{L^\infty} &\leq \|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu \mid \mathcal{B})\|_{L^\infty} \\ &\leq \|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 \mid \mathcal{B})\|_{L^\infty} + \|\mathbb{E}(1 \mid \mathcal{B})\|_{L^\infty} = 1 + o_\varepsilon(1). \end{aligned} \quad (7.2.2)$$

Now we want to use Szemerédi's theorem, theorem 3.1.1, but f_{U^\perp} does not satisfy the conditions since we might have $\mathbb{E}(f_{U^\perp}) < \delta$ and $\|f_{U^\perp}\|_{L^\infty} > 1$, but the difference is no more than $o_\varepsilon(1)$, so from the triangle inequality there is $g : \mathbb{Z}_N \rightarrow \mathbb{R}$ such that g satisfies the conditions of theorem 3.1.1 and

$$\|f_{U^\perp} - g\|_{L^\infty} = o_\varepsilon(1)$$

and

$$\mathbb{E}(f_{U^\perp} - g) = o_\varepsilon(1).$$

Using Szemerédi's theorem (theorem 3.1.1, and this is the only place we use this theorem) we now get that there is a constant $c(k, \delta) > 0$ such that

$$\mathbb{E}(g(x)g(x+r)\cdots g(x+(k-1)r) \mid x, r \in \mathbb{Z}_N) \geq c(k, \delta) - o_{k, \delta}(1).$$

And since $\|f_{U^\perp} - g\|_{L^\infty} = o_\varepsilon(1)$ and $\|g\|_{L^\infty} = 1 + o_\varepsilon(1)$ we have (recall that the constant in the $o(1)$ terms are allowed to depend on k)

$$\begin{aligned} & f_{U^\perp}(x)f_{U^\perp}(x+r)\cdots f_{U^\perp}(x+(k-1)r) - g(x)g(x+r)\cdots g(x+(k-1)r) \\ &= f_{U^\perp}(x)f_{U^\perp}(x+r)\cdots f_{U^\perp}(x+(k-1)r) - (g(x) + o_\varepsilon(1))g(x+r)\cdots g(x+(k-1)r) \\ &\quad + o_\varepsilon(1)g(x+r)\cdots g(x+(k-1)r) \\ &= f_{U^\perp}(x)(f_{U^\perp}(x+r)\cdots f_{U^\perp}(x+(k-1)r) - g(x+r)\cdots g(x+(k-1)r)) + o_\varepsilon(1). \end{aligned}$$

Now we can use the same trick to get

$$\begin{aligned} & f_{U^\perp}(x+r)\cdots f_{U^\perp}(x+(k-1)r) - g(x+r)\cdots g(x+(k-1)r) \\ &= f_{U^\perp}(x+r)(f_{U^\perp}(x+2r)\cdots f_{U^\perp}(x+(k-1)r) - g(x+2r)\cdots g(x+(k-1)r)) + o_\varepsilon(1), \end{aligned}$$

and then insert this. If we continue like this we end up with

$$f_{U^\perp}(x)f_{U^\perp}(x+r)\cdots f_{U^\perp}(x+(k-2)r)(f_{U^\perp}(x+(k-1)r) - g(x+(k-1)r)) + o_\varepsilon(1)$$

which is $o_\varepsilon(1)$ because $\|f_{U^\perp} - g\|_{L^\infty} = o_\varepsilon(1)$ and f_{U^\perp} is bounded. So

$$\mathbb{E}(f_{U^\perp}(x)f_{U^\perp}(x+r)\cdots f_{U^\perp}(x+(k-1)r) \mid x, r \in \mathbb{Z}_N) \geq c(k\delta) - o_\delta(1) - o_\varepsilon(1). \quad (7.2.3)$$

We know that

$$\|(1 - \mathbb{1}_\Omega)f\|_{L^\infty} \leq \|f\|_{L^\infty} \leq \|\nu\|_{L^\infty}$$

for all $x \in \mathbb{Z}_N$ and from 2. in proposition 7.1.2 we get

$$\|f_{U^\perp} - 1\|_{L^\infty} \leq \|(1 - \mathbb{1}_\Omega)\mathbb{E}(\nu - 1 \mid \mathcal{B})\|_{L^\infty} = o_\varepsilon(1),$$

so

$$\|f_{U^\perp}\|_{L^\infty} \leq 1 + o_\varepsilon(1).$$

Combining these estimates we get that

$$\|f_U\|_{L^\infty} \leq \|(1 - \mathbb{1}_\Omega)f_U\|_{L^\infty} = \|(1 - \mathbb{1}_\Omega)f + f_{U^\perp}\|_{L^\infty} \leq \|\nu\|_{L^\infty} + 1 + o_\varepsilon(1).$$

Here we want to use the new von Neumann theorem, theorem 4.3.1, with some $f_i = f_U$ and some $f_i = f_{U^\perp}$. The f_i 's that are equal to f_{U^\perp} satisfies the conditions for theorem 4.3.1 since

$$\|f_{U^\perp}\|_{L^\infty} \leq 1 + o_\varepsilon(1) \leq \|\nu\|_{L^\infty} + 1$$

for sufficiently small epsilon. But once again the estimates are not good enough for the f_i 's that are equal to f_U . But we can use the same trick as before to give us a $h : \mathbb{Z}_N \rightarrow \mathbb{R}$ such that $\|f_U - h\|_{L^\infty} = o_\varepsilon(1)$ and then use h . By the same argument as before this gives us an error term $o_\varepsilon(1)$ in the final result, so we have

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x+jr) \mid x, r \in \mathbb{Z}_N\right) = O\left(\inf_{0 \leq j \leq k-1} \|f_j\|_{U^{k-1}}\right) + o(1) + o_\varepsilon(1) \quad (7.2.4)$$

where each f_j is either f_U or f_{U^\perp} . Now recall that from 3. in proposition 7.1.2 we get that

$$\|f_U\|_{U^{k-1}} \leq \varepsilon^{1/2^k}$$

so if at least one of the f_j 's are equal to f_U , the infimum on the RHS is $\leq \varepsilon^{1/2^k}$ and the RHS of (7.2.4) can be rewritten to

$$O(\varepsilon^{1/2^k}) + o_\varepsilon(1)$$

because the two o -terms can be written together. This gives us

$$\mathbb{E} \left(\prod_{j=0}^{k-1} f_j(x + jr) \mid x, r \in \mathbb{Z}_N \right) = O(\varepsilon^{1/2^k}) + o_\varepsilon(1) \quad (7.2.5)$$

whenever each of the f_j 's is equal to either f_U or f_{U^\perp} and at least one is equal to f_U .

Now let $\tilde{f} = f_U + f_{U^\perp} = (1 - \mathbb{1}_\Omega)f$. Consider the expression

$$\mathbb{E} \left(\tilde{f}(x)\tilde{f}(x+r) \cdots \tilde{f}(x+(k-1)r) \mid x, r \in \mathbb{Z}_N \right).$$

This can be written as

$$\sum_{S \subseteq \{0,1,\dots,k-1\}} \mathbb{E} \left(\prod_{j \in S} f_U(x + jr) \prod_{i \notin S} f_{U^\perp}(x + ir) \right)$$

where $i \notin S$ means $i \in \{0,1,\dots,k-1\} \setminus S$. When $S = \emptyset$ we get a term that is equal to the LHS of (7.2.3) and for all other S at least one of the factors in the expected value is f_U so we can use (7.2.5) on all of these terms. So we get

$$\mathbb{E} \left(\tilde{f}(x)\tilde{f}(x+r) \cdots \tilde{f}(x+(k-1)r) \mid x, r \in \mathbb{Z}_N \right) \geq c(k, \delta) + o_\delta(1) + O(\varepsilon^{1/2^k}) + o_\varepsilon(1).$$

By the definition of \tilde{f} we get that

$$0 \leq \tilde{f}(x) \leq f(x)$$

for all $x \in \mathbb{Z}_N$ so

$$\mathbb{E} (f(x)f(x+r) \cdots f(x+(k-1)r) \mid x, r \in \mathbb{Z}_N) \geq c(k, \delta) - o_\delta(1) + O(\varepsilon^{1/2^k}) + o_\varepsilon(1).$$

The term $O(\varepsilon^{1/2^k})$ can be picked arbitrarily small because ε can be picked arbitrarily small and $o_\varepsilon(1)$ can also be arbitrarily small when N is sufficiently large, so the whole thing can be written as

$$\mathbb{E} (f(x)f(x+r) \cdots f(x+(k-1)r) \mid x, r \in \mathbb{Z}_N) \geq c(k, \delta) - o_\delta(1)$$

which is what we wanted to prove. □

Chapter 8

Construction of a pseudorandom measure

In this chapter we want to construct a k -pseudorandom measure ν and a function $\tilde{\Lambda}$ with support in the primes such that ν majorises $\tilde{\Lambda}$. As mentioned in chapter 3, this is what we need in order to use the modified version of Szemerédi's theorem to give us arithmetic progressions of arbitrary length in the primes.

8.1 Definitions and notation

Recall that we defined $\tilde{\Lambda} : \mathbb{N} \rightarrow \mathbb{R}^+$ in the following way.

Definition 8.1.1 (The modified Von Mangoldt function). Let $w(N)$ be defined by

$$w = w(N) = \log \log N$$

and let

$$W = W(N) = \prod_{p \leq w} p.$$

Now define $\tilde{\Lambda} : \mathbb{N} \rightarrow \mathbb{R}^+$ by

$$\tilde{\Lambda}(n) = \begin{cases} \frac{\phi(W)}{W} \log(Wn + 1) & \text{if } Wn + 1 \text{ is a prime} \\ 0 & \text{otherwise} \end{cases}$$

for $n \in \mathbb{N}$.

We now have the following lemma, which we have used in the proof of the theorem.

Lemma 8.1.2. *With $\tilde{\Lambda}$ defined as above we have*

$$\sum_{n \leq N} \tilde{\Lambda}(n) = N(1 + o(1)).$$

Proof. We define

$$\psi(x, W, 1) = \sum_{\substack{n \leq x, n \text{ prime power} \\ n \equiv 1 \pmod{W}}} \log n,$$

and

$$\theta(x, W, 1) = \sum_{\substack{n \leq x, n \text{ prime} \\ n \equiv 1 \pmod{W}}} \log n.$$

We now rewrite $\sum \tilde{\Lambda}(n)$ by adding and subtracting the same term

$$\begin{aligned} \sum_{n \leq N} \tilde{\Lambda}(n) &= \sum_{n \leq N} \tilde{\Lambda}(n) - \frac{\phi(W)}{W} \psi(NW + 1, W, 1) + \frac{\phi(W)}{W} \psi(NW + 1, W, 1) \\ &= \frac{\phi(W)}{W} \theta(NW + 1, W, 1) - \frac{\phi(W)}{W} \psi(NW + 1, W, 1) + \frac{\phi(W)}{W} \psi(NW + 1, W, 1). \end{aligned}$$

As it is common in analytic number theory we define

$$\theta(x) = \sum_{p \leq x} \log p,$$

and

$$\psi(x) = \sum_{p^m \leq x} \log p.$$

We now have

$$\psi(x) = \theta(x) + \theta(x^{1/2}) + \cdots + \theta(x^{1/m}),$$

where m is the largest integer such that $x^{1/m} \geq 2$. So $m = \lfloor \log_2 x \rfloor$ and we have

$$0 \leq \psi(x) - \theta(x) = \sum_{i=2}^{\log_2 x} \theta(x^{1/i}) \leq \log_2(x) \theta(x^{1/2}).$$

By the prime number theorem we have

$$\theta(x) = x + O\left(\frac{x}{(\log x)^A}\right) \tag{8.1.1}$$

for any $A > 0$, so $\theta(x^{1/2}) \leq x^{1/2} \log(x^{1/2})$ for sufficiently large x . Now

$$\log_2(x) \theta(x^{1/2}) \leq \frac{\sqrt{x} (\log x)^2}{2 \log 2},$$

and we have thus proved the bound

$$0 \leq \psi(x) - \theta(x) \leq \frac{\sqrt{x} (\log x)^2}{2 \log 2}.$$

This is also valid when we consider $\psi(x, W, 1)$ and $\phi(x, W, 1)$ instead, so using this we get

$$\begin{aligned} \frac{\phi(W)}{W}\theta(NW + 1, W, 1) - \frac{\phi(W)}{W}\psi(NW + 1, W, 1) \\ \leq O\left(\frac{\phi(W)}{W}\frac{\sqrt{NW}(\log NW)^2}{2\log 2}\right) \leq O(\sqrt{NW}\log(NW)^2). \end{aligned}$$

since $\phi(W) \leq W$. Now $W = e^{\theta(w)}$ by the definition of w and W , and by (8.1.1) we have $\theta(x) \leq 2x$ for sufficiently large x . So

$$W \leq e^{2w} = (\log \log N)^2 \leq (\log N)^2.$$

Using this we see that

$$\frac{\phi(W)}{W}\theta(NW + 1, W, 1) - \frac{\phi(W)}{W}\psi(NW + 1, W, 1) \leq O(\sqrt{NW}\log(NW)^2) = o(N).$$

Now we need to prove that

$$\frac{\phi(W)}{W}\psi(NW + 1, W, 1) = N + o(N).$$

We have by Siegel-Walfisz (see p. 124 in [7]) that

$$\psi(WN + 1, W, 1) = \frac{WN}{\phi(W)} + O\left(\frac{WN}{\log(WN)^A}\right)$$

for any $A > 0$. Letting $A = 2$ we get

$$\frac{\phi(W)}{W}\psi(NW + 1, W, 1) = N + O\left(\frac{N}{\log(NW)^2}\right) = N + o(N)$$

as desired. □

Recalling that the Möbius function $\mu : \mathbb{N} \rightarrow \{-1, 0, 1\}$ is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n \text{ is square-free with an even number of distinct prime factors} \\ -1 & \text{if } n \text{ is square-free with an odd number of distinct prime factors} \\ 0 & \text{otherwise} \end{cases},$$

we define the Goldston-Yıldırım truncated divisor sum in the following way.

Definition 8.1.3 (Goldston-Yıldırım truncated divisor sum). Let $R \in \mathbb{R}^+$. We define $\Lambda_R : \mathbb{N} \rightarrow \mathbb{Z}$ by

$$\Lambda_R(n) = \sum_{d|n, d \leq R} \mu(d) \log(R/d)$$

for $n \in \mathbb{N}$.

Remark 8.1.4. Note that if we define $\log(x)_+ = \max(\log(x), 0)$ we have the identity

$$\Lambda_R(n) = \sum_{d|n} \mu(d) \log(R/d)_+$$

because the terms with $d > R$ will not contribute to the sum.

Now we are ready to define a measure ν , which we in this chapter will prove is k -pseudorandom.

Definition 8.1.5. Let $R = N^{k^{-1}2^{-k-4}}$ and $\varepsilon = \frac{1}{2^k(k+4)!}$. We define $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ by

$$\nu(n) = \begin{cases} \frac{\phi(W)\Lambda_R(Wn+1)^2}{W \log R} & \text{when } \varepsilon N \leq n \leq 2\varepsilon N \\ 1 & \text{otherwise} \end{cases},$$

where ϕ is Euler's ϕ function.

Remark 8.1.6. For the rest of this chapter ν, R, Λ_R, W and ε will be as defined above.

8.2 A proof that ν is pseudorandom

Now proposition 3.1.3 follows from the following proposition, and this is hence the last thing we need to prove to conclude the proof.

Proposition 8.2.1. *Let N be a sufficiently large prime number. Then ν is a k -pseudorandom measure*

$$\nu(x) \geq \frac{1}{k2^{k+5}} \tilde{\Lambda}(x)$$

for all $\varepsilon N \leq x \leq 2\varepsilon N$.

To prove this proposition, we need to prove the following 4 claims.

1. We have $\nu(x) \geq 0$ for all $x \in \mathbb{Z}_N$ and

$$\nu(x) \geq \frac{1}{k2^{k+5}} \tilde{\Lambda}(x)$$

for all $\varepsilon N \leq x \leq 2\varepsilon N$.

2. ν is a measure, ie. it satisfies $\mathbb{E}(\nu) = 1 + o(1)$.
3. The measure ν satisfies the $(k2^{k-1}, 3k - 4, k)$ -linear forms condition.
4. The measure ν satisfies the 2^{k-1} -correlation condition.

Let us prove them one at a time.

1. ν majorises $\tilde{\Lambda}$

It is clear that $\nu(x) \geq 0$ for all $x \in \mathbb{Z}_N$, so we just need to prove that

$$\nu(x) \geq \frac{1}{k2^{k+5}} \tilde{\Lambda}(x)$$

for all $\varepsilon N \leq x \leq 2\varepsilon N$. If $Wx + 1$ is not a prime then the RHS is zero and we are done, so let us assume that $Wx + 1$ is a prime. Then we need to prove

$$\frac{\phi(W)\Lambda_R(Wx + 1)^2}{W \log R} \geq \frac{\phi(W)}{Wk2^{k+5}} \log(Wx + 1)$$

which is the same as proving

$$\frac{\Lambda_R(Wx + 1)^2}{\log R} \geq \frac{1}{k2^{k+5}} \log(Wx + 1).$$

Since $\varepsilon N \leq x \leq 2\varepsilon N$ we have $Wx + 1 = O(N \log N)$ which is greater than R when N is sufficiently large since $R = N^a$ with $a < 1$. So when $Wx + 1$ is a prime there is only one divisor d of $Wx + 1$ such that $d \leq R$ namely $d = 1$, so

$$\Lambda_R(Wx + 1) = \log R = k^{-1}2^{-k-4} \log N$$

and dividing by $k^{-1}2^{-k-4}$ we have to prove that

$$\log N \geq \frac{1}{2} \log(Wx + 1)$$

which is true for sufficiently large N because $W = O(\log N)$.

2. ν is a measure

To prove this we need the following proposition. For the proof of this proposition, we refer to the appendix of [6].

Proposition 8.2.2. *Let $m, t \in \mathbb{N}$. For each $1 \leq i \leq m$, define linear forms $\psi_i : \mathbb{R}^t \rightarrow \mathbb{R}$ by*

$$\psi_i(x_1, \dots, x_t) = \sum_{j=1}^t L_{ij}x_j + b_i$$

where (L_{ij}) is an integer $m \times t$ matrix such that no row is a rational multiple of another row and such that

$$|L_{ij}| \leq \frac{\sqrt{w(N)}}{2}$$

for all $i = 1, \dots, m$ and $j = 1, \dots, t$. Let $\theta_i = W\psi_i + 1$ for $i = 1, \dots, m$. Assume that $B \subseteq \mathbb{R}^t$ is a product of t intervals $I_j \subseteq \mathbb{R}$ such that each of the intervals has length at least R^{10m} . Then

$$\mathbb{E} (\Lambda_R(\theta_1(x))^2 \cdots \Lambda_R(\theta_m(x))^2 \mid x \in B) = (1 + o_{m,t}(1)) \left(\frac{W \log R}{\phi(W)} \right)^m.$$

To prove **2.** we apply proposition 8.2.2 with $m = t = 1$ and $B = [\varepsilon N, 2\varepsilon N]$ for sufficiently large N . This gives us

$$\mathbb{E} (\Lambda_R(\theta_1(x))^2 \mid x \in [\varepsilon N, 2\varepsilon N]) = (1 + o(1)) \left(\frac{W \log R}{\phi(W)} \right)$$

and recalling the definition of ν this is equivalent to

$$\mathbb{E} (\nu(x) \mid x \in [\varepsilon N, 2\varepsilon N]) = 1 + o(1).$$

But since $\nu(x) = 1$ for $x \notin [\varepsilon N, 2\varepsilon N]$ we get

$$\mathbb{E} (\nu(x) \mid x \notin [\varepsilon N, 2\varepsilon N]) = 1,$$

and combining these we get

$$\mathbb{E} (\nu(x) \mid x \in \mathbb{Z}_N) = 1 + o(1)$$

which is what we wished to prove.

3. ν satisfies the linear forms condition

We now want to prove that ν satisfies the $(k2^{k-1}, 3k-4, k)$ -linear forms condition. So let

$$\psi_i(x_1, \dots, x_t) = \sum_{j=1}^t L_{ij}x_j + b_i$$

for $i = 1, \dots, m$ be m linear forms such that $m \leq k2^{k-1}$, $t \leq 3k-4$ and such that the L_{ij} 's are rational numbers with height $\leq k$. Furthermore no row of the $m \times t$ -matrix (L_{ij}) is a rational multiple of any other row. By the definition of the linear forms condition (definition 2.3.2) we need to prove that

$$\mathbb{E} (\nu(\psi_1(x)) \cdots \nu(\psi_m(x)) \mid x \in \mathbb{Z}_N^t) = 1 + o_{m,t}(1). \quad (8.2.1)$$

The denominators of the L_{ij} 's are all smaller than k in absolute value, so if we multiply all the L_{ij} 's by $k!$ we clear the denominators and get the new bound $|L_{ij}| \leq k \cdot k! \leq (k+1)!$, and taking N sufficiently large we get

$$|L_{ij}| \leq (k+1)! < \frac{\sqrt{w(N)}}{2}$$

which is required in proposition 8.2.2.

Now assume that $Q = Q(N)$ is a function such that $Q(N) \rightarrow \infty$ for $N \rightarrow \infty$, and $Q(N) < N$. Let for $(u_1, \dots, u_t) \in \mathbb{Z}_Q = \mathbb{Z}/Q\mathbb{Z}$

$$B_{u_1, \dots, u_t} = \{(x_1, \dots, x_t) \in \mathbb{Z}_N^t \mid \lfloor u_j \frac{N}{Q} \rfloor \leq x_j < \lfloor (u_j + 1) \frac{N}{Q} \rfloor \text{ for all } j = 1, 2, \dots, t\}.$$

We will now need the following lemma. It states that we can partition the set we are taking average over into almost equally sized sets, take the average of these averages, and then get the right result with a multiplicative error of $1 + o(1)$.

Lemma 8.2.3. *Let $f : \mathbb{Z}_N^n \rightarrow \mathbb{R}$ and let $B_i^N, i \in I$ be a partition of \mathbb{Z}_N^n such that $\#B_i^N \rightarrow \infty$ for $N \rightarrow \infty$ for all $i \in I$ and*

$$\#B_i^N - \#B_j^N = O(1)$$

for all $i, j \in I$. Then

$$\mathbb{E}(\mathbb{E}(f(x) \mid x \in B_i) \mid i \in I) = (1 + o(1))\mathbb{E}(f). \quad (8.2.2)$$

Proof. We will just write $B_i = B_i^N$. Let $b = \mathbb{E}(\#B_i \mid i \in I)$. Then $\#B_i = b + O(1)$ for all $i \in I$ and

$$\left| \frac{1}{\#B_i} - \frac{1}{b} \right| = \frac{O(1)}{b(b + O(1))} = \frac{1}{b}o(1)$$

because $b \rightarrow \infty$ as $N \rightarrow \infty$, so

$$\frac{1}{\#B_i} = (1 + o(1))\frac{1}{b}.$$

Using this on the LHS of (8.2.2) we get

$$\frac{1}{\#I} \sum_{i \in I} \frac{1}{\#B_i} \sum_{x \in B_i} f(x) = \frac{1}{\#I} \sum_{i \in I} \frac{1}{b} (1 + o(1)) \sum_{x \in B_i} f(x)$$

which is equal to $(1 + o(1))\mathbb{E}(f)$. \square

Now notice that $\{B_{u_1, \dots, u_t} \mid u_1, \dots, u_t \in \mathbb{Z}_Q\}$ is a partition of \mathbb{Z}_N^t , so due to lemma 8.2.3 we have that

$$\mathbb{E}(\mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x)) \cdots \nu(\psi_m(x)) \mid x \in B_{u_1, \dots, u_t}) \mid u_1, \dots, u_t \in \mathbb{Z}_Q) \quad (8.2.3)$$

is $1 + o(1)$ times the LHS (8.2.1) since two boxes differ at most 2 in size, and since we can pick Q such that $Q \rightarrow \infty$ for $N \rightarrow \infty$ sufficiently slow, such that the sizes of the boxes will go to ∞ as $N \rightarrow \infty$.

So we need to show that (8.2.3) is $1 + o_{m,t}(1)$. We will need the following definition and lemma.

Definition 8.2.4 (Nice t -tuples). A t -tuple $(u_1, \dots, u_t) \in \mathbb{Z}_Q^t$ is called *nice* if for all $1 \leq i \leq m$ we have

$$\psi_i(B_{u_1, \dots, u_t}) \subseteq [\varepsilon N, 2\varepsilon N] \quad \text{or} \quad \psi_i(B_{u_1, \dots, u_t}) \cap [\varepsilon N, 2\varepsilon N] = \emptyset.$$

Lemma 8.2.5. *When $(u_1, \dots, u_t) \in \mathbb{Z}_Q^t$ is nice we have*

$$\mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x)) \cdots \nu(\psi_m(x)) \mid x \in B_{u_1, \dots, u_t}) = 1 + o_{m,t}(1).$$

When $(u_1, \dots, u_t) \in \mathbb{Z}_Q^t$ is not nice we have

$$\mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x)) \cdots \nu(\psi_m(x)) \mid x \in B_{u_1, \dots, u_t}) = O_{m,t}(1) + o_{m,t}(1)$$

and the proportion of non-nice tuples $(u_1, \dots, u_t) \in \mathbb{Z}_Q^t$ is $O_{m,t}(1/Q)$.

Proof. Suppose that (u_1, \dots, u_t) is nice. Then due to proposition 8.2.2 we have

$$\mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x)) \cdots \nu(\psi_m(x)) \mid x \in B_{u_1, \dots, u_t}) = 1 + o_{m,t}(1)$$

because we can replace each $\nu(\psi_i(x))$ by either

$$\frac{\phi(W)}{W \log R} \Lambda_R(\omega_i(x))^2$$

or 1, and because if Q is sufficiently slowly growing in N , then $N/Q \geq R^{10m}$ for sufficiently large N , so the conditions in proposition 8.2.2 are met.

Now assume that (u_1, \dots, u_t) is not nice. We now use the bound

$$\|\nu\|_{L^\infty} \leq 1 + \frac{\phi(W)}{W \log R} \Lambda_R(\omega_i(x))^2$$

to obtain

$$\begin{aligned} & \mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x)) \cdots \nu(\psi_m(x)) \mid x \in B_{u_1, \dots, u_t}) \\ & \leq \sum_{A \subseteq \{1, \dots, m\}} \mathbb{E} \left(\prod_{i \in A} \frac{\phi(W)}{W \log R} \Lambda_R(\omega_i(x))^2 \mid x \in B_{u_1, \dots, u_t} \right). \end{aligned}$$

Using proposition 8.2.2 again gives us

$$\mathbb{E} \left(\prod_{i \in A} \frac{\phi(W)}{W \log R} \Lambda_R(\omega_i(x))^2 \mid x \in B_{u_1, \dots, u_t} \right) = 1 + o_{m,t}(1),$$

so summing over all A gives us

$$\mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x)) \cdots \nu(\psi_m(x)) \mid x \in B_{u_1, \dots, u_t}) = O_{m,t}(1) + o_{m,t}(1)$$

because there are $O_m(1)$ such A 's.

We now need to prove that the proportion of non-nice tuples are $O_{m,t}(1/Q)$. Suppose that (u_1, \dots, u_t) is not nice. Then one of the $\psi_i(B_{u_1, \dots, u_t})$ has non-empty intersection with the interval $[\varepsilon N, 2\varepsilon N]$, but is not completely contained in the interval, so there is $x, y \in B_{u_1, \dots, u_t}$ such that

$$\psi(x) \in [\varepsilon N, 2\varepsilon N] \quad \text{and} \quad \psi(y) \notin [\varepsilon N, 2\varepsilon N].$$

But since $L_{ij} = O(1)$ both $\psi(x)$ and $\psi(y)$ can be estimated by $\psi(u_1, \dots, u_t)$, and we hence get

$$\psi(x) = \sum_{j=1}^t L_{ij} \lfloor u_j \frac{N}{Q} \rfloor + b_i + O_{m,t}(N/Q)$$

and

$$\psi(y) = \sum_{j=1}^t L_{ij} \lfloor u_j \frac{N}{Q} \rfloor + b_i + O_{m,t}(N/Q).$$

where the error-term is of magnitude N/Q because that is the size of the box and is hence the maximal distance the i 'th coordinate of x can be from $\lfloor u_j \frac{N}{Q} \rfloor$, and the dependence on m and t is due to the size of the sum.

If $\psi(y) < \varepsilon N$ then

$$\psi(y) < \varepsilon N \leq \psi(x)$$

and hence

$$\varepsilon N = \sum_{j=1}^t L_{ij} \lfloor u_j \frac{N}{Q} \rfloor + b_i + O_{m,t}(N/Q)$$

by the expressions for $\psi(x)$ and $\psi(y)$ we found before. If $\psi(y) > 2\varepsilon N$ we get a similar expression with $2\varepsilon N$ on the LHS, so we have

$$a\varepsilon N = \sum_{j=1}^t L_{ij} \lfloor u_j \frac{N}{Q} \rfloor + b_i + O_{m,t}(N/Q)$$

for either $a = 1$ or $a = 2$. Dividing by N/Q we get

$$a\varepsilon Q = \sum_{j=1}^t L_{ij} u_j + \frac{b_i Q}{N} + O_{m,t}(1).$$

None of the tuples $(L_{ij})_{j=1}^t$ are zero, so at most $O_{m,t}(Q^{t-1})$ tuples (u_1, \dots, u_t) satisfy this equation, and this is $O_{m,t}(1/Q)$ of the possible tuples. \square

We now proceed with the proof. We have

$$(1 - O_{m,t}(1/Q))(1 + o_{m,t}(1)) + O_{m,t}(1/Q)(O_{m,t}(1) + o_{m,t}(1)) = 1 + o_{m,t}(1),$$

because we picked Q such that $Q(N) \rightarrow \infty$ when $N \rightarrow \infty$ so $O_{m,t}(1/Q) = o_{m,t}(1)$, and due to the lemma we have

$$\mathbb{E}(\mathbb{E}(\nu(\psi_1(x)) \cdots \nu(\psi_m(x)) \mid x \in B_{u_1, \dots, u_t}) \mid u_1, \dots, u_t \in \mathbb{Z}_Q) = 1 + o_{m,t}(1),$$

which proves the claim.

4. ν satisfies the correlation condition

We now want to prove that ν satisfies the 2^{k-1} -correlation condition. To do this we will need the following two lemmas. For the proof of the first one, we refer to the appendix of [6].

Lemma 8.2.6. *Let $m \geq 1$ be an integer and let B be an interval of length $\geq R^{10m}$. Let $h_1, \dots, h_m \in \mathbb{Z}$ be distinct such that $|h_i| \leq N^2$ for all i . Now define*

$$\Delta = \prod_{1 \leq i < j \leq m} |h_i - h_j|.$$

Then we have for sufficiently large N that

$$\begin{aligned} \mathbb{E} (\Lambda_R(W(x + h_1) + 1)^2 \cdots \Lambda_R(W(x + h_m) + 1)^2 \mid x \in B) \\ \leq (1 + o_m(1)) \left(\frac{W \log R}{\phi(W)} \right)^m \prod_{p|\Delta} (1 + O_m(p^{-1/2})) \end{aligned}$$

where the product is over all primes p that divides Δ .

Lemma 8.2.7. *Let $m \geq 1$ be an integer. There is a function $\tau_m : \mathbb{Z} \rightarrow \mathbb{R}^+$ such that $\tau_m(n) \geq 1$ for all $n \neq 0$, and such that*

$$\prod_{p|\Delta} (1 + O_m(p^{-1/2})) \leq \sum_{1 \leq i < j \leq m} \tau_m(h_i - h_j)$$

for all distinct $h_1, \dots, h_m \in [\varepsilon N, 2\varepsilon N]$, where Δ is as in lemma 8.2.6 and

$$\mathbb{E} (\tau_m^q(n) \mid 0 < |n| \leq N) = O_{m,q}(1)$$

for all $q \geq 1$.

Proof. By the definition of Δ we have

$$\prod_{p|\Delta} (1 + O_m(p^{-1/2})) \leq \prod_{1 \leq i < j \leq m} \prod_{p||h_i - h_j|} (1 + O_m(p^{-1/2}))$$

because we might include a prime p several times on the RHS. We can now use the following bound

$$\prod_{p||h_i - h_j|} (1 + O_m(p^{-1/2})) \leq \left(\prod_{p||h_i - h_j|} (1 + p^{-1/2}) \right)^{O_m(1)}$$

to get

$$\prod_{p|\Delta} (1 + O_m(p^{-1/2})) \leq \prod_{1 \leq i < j \leq m} \left(\prod_{p||h_i - h_j|} (1 + p^{-1/2}) \right)^{O_m(1)}.$$

Now define $\tau_m : \mathbb{Z} \rightarrow \mathbb{R}^+$ such that

$$\tau_m(n) = \left(\prod_{p|n} (1 + p^{-1/2}) \right)^{O_m(1)}$$

for all $n \in \mathbb{Z}$. By the arithmetic mean-geometric mean inequality we have

$$\frac{1}{\binom{m}{2}} \sum_{1 \leq i < j \leq m} \tau_m(h_i - h_j) \geq \prod_{1 \leq i < j \leq m} \tau_m(h_i - h_j)^{\binom{m}{2}^{-1}} = \prod_{1 \leq i < j \leq m} \tau_m(h_i - h_j)^{O_m(1)}.$$

Using the definition of τ_m , this is

$$\prod_{1 \leq i < j \leq m} \left(\prod_{p \mid |h_i - h_j|} (1 + p^{-1/2}) \right)^{O_m(1)}$$

so τ_m satisfies the desired inequality.

Now we need to show that

$$\mathbb{E} \left(\prod_{p \mid n} (1 + p^{-1/2})^{O_m(q)} \mid 0 < |n| \leq N \right) = O_{m,q}(1)$$

for all $0 < q$. Now for sufficiently large p (for all but $O_{m,q}(1)$ primes), we have

$$(1 + p^{-1/2})^{O_m(q)} \leq 1 + p^{-1/4},$$

so

$$\mathbb{E} \left(\prod_{p \mid n} (1 + p^{-1/2})^{O_m(q)} \mid 0 < |n| \leq N \right) \leq O_{m,q}(1) \mathbb{E} \left(\prod_{p \mid n} (1 + p^{-1/4}) \mid 0 < |n| \leq N \right).$$

But multiplying out we get the inequality

$$\prod_{p \mid n} (1 + p^{-1/4}) \leq \sum_{d \mid n} d^{-1/4},$$

where we on the RHS only includes positive d . Notice that it is not an equality since the RHS also includes divisors which are divisible by a prime power. Using this we get

$$\mathbb{E} \left(\prod_{p \mid n} (1 + p^{-1/2})^{O_m(q)} \mid 0 < |n| \leq N \right) \leq O_{m,q}(1) \frac{1}{2N} \sum_{1 \leq |n| \leq N} \sum_{d \mid n} d^{-1/4}.$$

We see that $d^{-1/4}$ appears in this double sum $2N/d$ times for each $1 \leq d \leq N$, because it appears N/d times for the positive n and N/d times for the negative n , so the above expression is

$$O_{m,q}(1) \sum_{d=1}^N d^{-5/4} = O_{m,q}(1),$$

which concludes the proof. \square

Recall that in order to prove that ν satisfies the 2^{k-1} -correlation condition, we need to prove that for any $1 \leq m \leq 2^{k-1}$ there is a function $\tau : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ such that $\mathbb{E}(\tau^q) = O_{m,q}(1)$ for all $q \geq 1$ and such that

$$\mathbb{E}(\nu(x+h_1) \cdots \nu(x+h_m) \mid x \in \mathbb{Z}_N) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j)$$

for all $h_1, \dots, h_m \in \mathbb{Z}_N$.

Let m be given. Then we define $\tau : \mathbb{Z}_N \rightarrow \mathbb{R}^+$ by

$$\tau(n) = \begin{cases} \tau_m(n - \lfloor N/2 \rfloor) & \text{if } n \neq 0 \\ \exp(Cm \log N / \log \log N) & \text{if } n = 0 \end{cases}$$

where $C > 0$ is some constant we pick later, and we consider $n - \lfloor N/2 \rfloor \in \mathbb{Z}_N$ in the obvious way by identifying \mathbb{Z}_N with the integers between $-N/2$ and $N/2$. From lemma 8.2.7 we see that

$$\mathbb{E}(\tau^q(x) \mid x \neq 0) = O_{m,q}(1),$$

and since

$$\frac{\exp(Cm \log N / \log \log N)}{N} = N^{Cm / \log \log N - 1} = o_{m,q}(1),$$

the case $x = 0$ only contributes with $o_{m,q}(1)$, so we have $\mathbb{E}(\tau^q) = O_{m,q}(1)$.

Let us consider the case where at least two of the h_i 's are equal. Now it is enough to show that

$$\mathbb{E}(\nu(x+h_1) \cdots \nu(x+h_m) \mid x \in \mathbb{Z}_N) \leq \exp(Cm \log N / \log \log N)$$

by the definition of $\tau(0)$. We have

$$\mathbb{E}(\nu(x+h_1) \cdots \nu(x+h_m) \mid x \in \mathbb{Z}_N) \leq \|\nu\|_{L^\infty}^m.$$

We now need to prove that

$$\|\nu\|_{L^\infty} \leq \exp(C \log N / \log \log N),$$

and from the definition of ν we get

$$\|\nu\|_{L^\infty} \leq O(\Lambda_R(Wx+1)^2 / \log R)$$

since $\phi(n)/n \leq 1$ for all n . Now $\log R = O(\log N)$ and

$$|\Lambda_R(Wx+1)| \leq d(Wx+1) \log(R)$$

where $d(\cdot)$ is the divisor function. From [7] (in their bound they use 2 as base for the exponential function, but here we use e instead) have

$$d(Wx+1) \leq \exp(C' \log(Wx+1) / \log \log(Wx+1))$$

for some constant C' , and since $Wx + 1 = O(N \log N)$ this is less than

$$\exp(C'' \log(N \log N) / \log \log(N \log N)) \leq \exp(C \log N / \log \log N)$$

for some constants C'' and C .

Now suppose that all h_i 's are distinct. Now define $g : \mathbb{Z}_N \rightarrow \mathbb{R}$ by

$$g(x) = \frac{\phi(W)}{W} \frac{\Lambda_R(Wx + 1)^2}{\log R} \mathbb{1}_{[\varepsilon N, 2\varepsilon N]}(x).$$

Then

$$\nu(x) \leq 1 + g(x)$$

for all $x \in \mathbb{Z}_N$ by definition of ν and hence

$$\mathbb{E}(\nu(x + h_1) \cdots \nu(x + h_m) \mid x \in \mathbb{Z}_N) \leq \mathbb{E}((1 + g(x + h_1)) \cdots (1 + g(x + h_m)) \mid x \in \mathbb{Z}_N).$$

The RHS can be written as

$$\sum_{A \subseteq \{1, \dots, m\}} \mathbb{E} \left(\prod_{i \in A} g(x + h_i) \mid x \in \mathbb{Z}_N \right).$$

Notice that if $i, j \in A$ with $|h_i - h_j| > \varepsilon N$, then either $g(x + h_i) = 0$ or $g(x + h_j) = 0$, so in the above expression we can assume in the sum that $|h_i - h_j| \leq \varepsilon N$ for all $i, j \in A$. By lemma 8.2.6, where we might have to increase N in order to have $N \geq R^{10m}$, we have

$$\mathbb{E} \left(\prod_{i \in A} g(x + h_i) \mid x \in \mathbb{Z}_N \right) \leq (1 + o_m(1)) \prod_{p|\Delta} (1 + O_m(p^{-1/2}))$$

and since $|h_i - h_j| \leq \varepsilon N$ we get

$$o_m(1) \prod_{p|\Delta} (1 + O_m(p^{-1/2})) = o_m(1).$$

Using lemma 8.2.7 we now get

$$\mathbb{E} \left(\prod_{i \in A} g(x + h_i) \mid x \in \mathbb{Z}_N \right) \leq \sum_{1 \leq i < j \leq m} \tau(h_i - h_j) + o_m(1).$$

We now want to sum over all A , but in order to get the desired bound we need to scale τ by a bounded factor, but this factor only depends on m , which is allowed due to the construction of τ in the proof of lemma 8.2.7 as

$$O_m(1) \prod_{p|n} (1 + p^{-1/2})^{O_m(1)}.$$

This concludes the proof of ν satisfying the correlation condition, and hence the proof of ν being k -pseudorandom.

Chapter 9

Solutions in arithmetic progression to linear equations

9.1 Introduction

In this chapter we will prove that we have infinitely many solutions to a system of linear equations in a set with arbitrarily long arithmetic progressions if the null space of the matrix has dimension at least 2 and contains $(1, 1, \dots, 1)$. Due to the Green-Tao theorem the primes is such a set and due to Szemerédi's theorem [12] all subsets of the integers with positive density are such sets. The results in this chapter are work of my own and an article on the results [8] has the 16th of February been submitted to the journal 'INTEGERS: Electronic Journal of Combinatorial Number Theory'.

The main result in this chapter is that the method here also gives us that for each of the solutions the coordinates are in the same arithmetic progression, and that the sets that contain arithmetic progressions of any length are exactly the sets that have infinitely many solutions to a homogeneous system of linear equations whenever the sum of the columns is zero. This gives us a new arithmetic structure on such sets, which gives a new formulation of the Erdős-Turan conjecture.

Of already existing results on prime solutions to linear equations we mention two. Balog [1] gave a lower bound on the number of prime solutions to a homogeneous system of linear equations $M\underline{x} = 0$ if the matrix M has a certain admissible structure, the null space contains a vector with positive coordinates and $M\underline{x} \equiv 0 \pmod{p^\alpha}$ has integer solutions coprime to p for all prime powers p^α . In particular he proved that if M is admissible and $(1, 1, \dots, 1)$ is a solution, then $M\underline{x} = 0$ has prime solutions. Choi, Liu and Tsang [2] has considered upper bounds for prime solutions to ternary linear equations.

9.2 APs and GAPs

In this chapter we will see when a system of linear equations has solutions in a set that contains arbitrarily long arithmetic progressions. Notice that this is in particular true for

the primes.

Since we are considering arithmetic progressions, the following notation will come in handy.

Definition 9.2.1 (Arithmetic progressions). Let $k, d \geq 1$ and $a \geq 0$ be integers. Then an *arithmetic progression* (AP) of length k , base a and step d is the set

$$\text{AP}(k, a, d) = \{a + \lambda d \mid 0 \leq \lambda < k\}.$$

Definition 9.2.2 (AP-set). Let $A \subseteq \mathbb{N}$. We will call A an *AP-set* if there for any $k \geq 1$ exists a pair $(a, d) \in \mathbb{N}^2$ such that

$$\text{AP}(k, a, d) \subseteq A.$$

Remark 9.2.3. Notice that an AP-set contains infinitely many APs of any length.

Definition 9.2.4 (Generalized arithmetic progressions). Let $d \geq 1$, $a \geq 0$, $b_1, \dots, b_d \geq 1$ and $N_1, \dots, N_d \geq 1$ be integers. Then a *generalized arithmetic progression* (GAP) of dimension d , base a , step (b_1, \dots, b_d) and volume (N_1, \dots, N_d) is the set

$$\{a + n_1 b_1 + \dots + n_d b_d \mid 0 \leq n_i < N_i \text{ for all } i\}.$$

Remark 9.2.5. Notice that a GAP of dimension d , base a , step (b_1, \dots, b_d) and volume $(2N_1 - 1, \dots, 2N_d - 1)$ can be written as

$$\{a' + n_1 b_1 + \dots + n_d b_d \mid -N_i < n_i < N_i \text{ for all } i\} \tag{9.2.1}$$

where $a' = a + (N_1 - 1)b_1 + \dots + (N_d - 1)b_d$.

Note also that the elements in a GAP might not be distinct, since we could have $n_j r_j = n_i r_i$ for some i, j with $0 \leq n_j < N_j$ and $0 \leq n_i < N_i$.

We can construct a GAP of any dimension and volume from a sufficiently long AP, so in particular an AP-set contains infinitely many GAPs of any given dimension and volume. The following lemma is taken from [5] and gives us a little more than just GAPs in AP-sets.

Lemma 9.2.6. *Any AP-set contains infinitely many GAPs of any given dimension and volume such that each GAP is contained in an AP.*

Proof. Let d and N_1, \dots, N_d be given. Let

$$N = \max_{1 \leq i \leq d} N_i$$

and let $k = N^d$. Now let an AP-set A be given. Then there are infinitely many $(a, m) \in \mathbb{N}^2$ such that $\text{AP}(k, a, m) \subseteq A$. Let one of these pairs (a, m) be given. Define r_i for each $1 \leq i \leq d$ by $r_i = N^{i-1}m$. Then for any integers n_1, \dots, n_d such that $0 \leq n_i < N_i$ for all i we have

$$\begin{aligned} a + n_1 r_1 + \dots + n_d r_d &= a + n_1 m + n_2 N m + \dots + n_d N^{d-1} m \\ &= a + m(n_1 + n_2 N + \dots + n_d N^{d-1}). \end{aligned}$$

And since $1 \leq n_i < N_i \leq N$ we have

$$n_1 + n_2N + \cdots + n_dN^{d-1} < N^d$$

since this is the base N representation of the number. So now we have that each element of the GAP is in $\text{AP}(k, a, m)$. No two elements of the GAP are equal since that would yield

$$n_1 + n_2N + \cdots + n_dN^{d-1} = n'_1 + n'_2N + \cdots + n'_dN^{d-1}$$

which implies $n_i = n'_i$ for all $i = 1, \dots, d$ due to the uniqueness of the base N representation. \square

9.3 Finding solutions in an AP-set

Using the existence of GAPs in AP-sets we can now find infinitely many solutions to systems of linear equations in any AP-set.

Theorem 9.3.1. *Let $n \geq 3$, $m \geq 1$, $M \in \text{Mat}_{m,n}(\mathbb{Z})$ and let A be an AP-set. Assume that the solution space of*

$$M\mathbf{x} = 0 \tag{9.3.1}$$

has dimension $d \geq 2$ and contains $(1, 1, \dots, 1)$. Then (9.3.1) has infinitely many solutions $\mathbf{x} = (x_1, \dots, x_n)$ such that $x_i \in A$ for all i , $x_i \neq x_j$ for some i, j and all x_i are elements in the same AP.

Proof. The solution space of (9.3.1) can be written as

$$m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_d\mathbf{r}_d, \quad m_i \in \mathbb{R}$$

where $\mathbf{r}_1 = (1, 1, \dots, 1)$, $\mathbf{r}_i = (r_{i1}, \dots, r_{in}) \in \mathbb{Z}^n$ for $2 \leq i \leq d$ and $\mathbf{r}_1, \dots, \mathbf{r}_d$ are linearly independent over \mathbb{R} . Now let $N = \max_{i,j} |r_{ij}| + 1$ and take a GAP of dimension $d - 1$ and volume $(2N - 1, \dots, 2N - 1)$. According to lemma 9.2.6 we can construct GAPs of any given size such that it is contained in an AP. Now take such a GAP, and as we did in (9.2.1) we write it as

$$\{a + n_1b_1 + \cdots + n_{d-1}b_{d-1} \mid -N < n_i < N \text{ for all } i\}. \tag{9.3.2}$$

Now

$$a\mathbf{r}_1 + b_1\mathbf{r}_2 + \cdots + b_{d-1}\mathbf{r}_d$$

is a solution to (9.3.1) and each coordinate is an element in the GAP given in (9.3.2). Now assume that the solution we have found has all coordinates equal. Then it is equal to $c\mathbf{r}_1$ for some $c \in \mathbb{N}$ so

$$(a - c)\mathbf{r}_1 + b_1\mathbf{r}_2 + \cdots + b_{d-1}\mathbf{r}_d = 0.$$

This is not possible since $\mathbf{r}_1, \dots, \mathbf{r}_d$ are linearly independent. \square

9.4 Prime-like sets

theorem 9.3.1 gives us a sufficient condition to be able to find infinitely many solutions in an AP-set. Let us now examine to what extent it also is a necessary condition. To examine this we need to require a bit more from our AP-set.

Definition 9.4.1 (Prime-like sets). A set $A \subseteq \mathbb{N}$ is called *prime-like* if for each $AP(k, a, d) \subseteq A$ with $k \geq 3$ we have $\gcd(a, d) = 1$.

Notice that the primes is prime-like because if we have a progression $AP(k, a, d)$ in the primes, then a is prime and d is even and not divisible by a , because if $a \mid d$ then $\gcd(a + d, a) = a$ so $a + d$ is not prime..

Theorem 9.4.2. Let A be a prime-like AP-set, $M \in \text{Mat}_{m,n}(\mathbb{Z})$ and $k \geq 3$. Assume that

$$M\mathbf{x} = 0 \tag{9.4.1}$$

has infinitely many solutions such that for each solution (x_1, \dots, x_n) there is $(a, d) \in \mathbb{N}^2$ such that

$$\{x_1, \dots, x_n\} \subseteq AP(k, a, d) \subseteq A.$$

Then $(1, 1, \dots, 1)$ is a solution to (9.4.1).

Proof. Let $1 \leq i \leq m$ be given. Assume for contradiction that $a_{i1} + \dots + a_{in} \neq 0$. Let $\{(x_1^{(j)}, \dots, x_n^{(j)}) \mid j \in \mathbb{N}\}$ be the infinitely many solutions given in the lemma. For each $j \in \mathbb{N}$ there exist b_j and d_j such that $x_l^{(j)} = b_j + \lambda_l^{(j)}d_j$ with $0 \leq \lambda_l^{(j)} < k$ for all $l = 1, \dots, n$ since each $x_l^{(j)}$ is an element of $AP(k, b_j, d_j)$. Inserting this in (9.4.1) we get that we for each $j \in \mathbb{N}$ have

$$b_j(a_{i1} + \dots + a_{in}) = -d_j(a_{i1}\lambda_1^{(j)} + \dots + a_{in}\lambda_n^{(j)}).$$

Since $\gcd(b_j, d_j) = 1$, b_j must divide $a_{i1}\lambda_1^{(j)} + \dots + a_{in}\lambda_n^{(j)}$ so if we let $C = |a_{i1}| + \dots + |a_{in}|$ we have $b_j \leq Ck$. Now

$$|d_j| = \left| b_j \frac{a_{i1} + \dots + a_{in}}{a_{i1}\lambda_1^{(j)} + \dots + a_{in}\lambda_n^{(j)}} \right| \leq Ck$$

so the set $\{d_j \mid j \in \mathbb{N}\}$ is also finite. The solutions $\{(x_1^{(j)}, \dots, x_n^{(j)}) \mid j \in \mathbb{N}\}$ are therefore taken from only finitely many APs of length k , and there can hence be only finitely many of them. This is a contradiction against the assumption, and this finishes the proof. \square

Combining this with theorem 9.3.1 we get the following.

Theorem 9.4.3. Let A be a prime-like AP-set and let $M \in \text{Mat}_{m,n}(\mathbb{Z})$ such that the null space of M has dimension at least 2. Then there is a $k \in \mathbb{N}$ such that the equation

$$M\mathbf{x} = 0$$

has infinitely many solutions where for each solution, all coordinates are elements of the same AP of length k in A if and only if $(1, 1, \dots, 1)$ is a solution.

We now give an example of an application of theorem 9.3.1. This is a known result, see for instance [5].

Corollary 9.4.4. *Let an AP-set A and $n \geq 1$ be given. Then there exists infinitely many n -tuples $x_1, \dots, x_n \in A$ with $x_i \neq x_j$ for some i, j such that*

$$\frac{x_1 + \dots + x_n}{n} \in A.$$

Proof. When $n = 1$ it is trivial so let $n \geq 2$ be given. Consider the linear equation

$$x_1 + \dots + x_n - nx_{n+1} = 0.$$

From theorem 9.3.1 we know that this equation has infinitely many solutions $x_1, \dots, x_n, x_{n+1} \in A$ with $x_i \neq x_j$ for some i, j . Now for each of these we have

$$\frac{x_1 + \dots + x_n}{n} = x_{n+1} \in A,$$

which finishes the proof. □

9.5 The Erdős-Turan conjecture

We have proven that in any AP-set we can find infinitely many solutions to any system of linear equation, as long as the sum of the columns of the matrix is zero. This motivates the following definition.

Definition 9.5.1 (Zero-solution sets). Let $M \in \text{Mat}_{m,n}(\mathbb{Z})$ such that the sum of the columns is zero and the null space of M has dimension at least 2. A set $A \subseteq \mathbb{N}$ is a *zero-solution set* if the following is true for all such M . The equation

$$M\mathbf{x} = 0$$

has infinitely many solutions $\mathbf{x} = (x_1, \dots, x_n)$ with $x_1, \dots, x_n \in A$ and $x_i \neq x_j$ for some i, j .

Now theorem 9.3.1 can be formulated as follows: If A is an AP-set then A is a zero-solution set. We now want to prove that zero-solution sets and AP-sets are the same.

Theorem 9.5.2. *Let $A \subseteq \mathbb{N}$. Then A is a zero-solution set if and only if A is an AP-set.*

Proof. The 'if' part we get from theorem 9.3.1. Let $n \geq 3$ be an integer and let $M \in \text{Mat}_{n-2,n}(\mathbb{Z})$ be given such that the solution space of $M\mathbf{x} = 0$ is given by

$$m_1(1, 1, \dots, 1) + m_2(0, 1, 2, \dots, n-1), \quad m_1, m_2 \in \mathbb{R}.$$

Since A is a zero-solution set there are infinitely many solutions in A with $m_2 \neq 0$. We also see that such a solution is in A so it is integer and both m_1 and m_2 are hence integer. Each of these solutions gives us an AP of length n . □

This might be useful in proving or disproving the Erdős-Turan conjecture since we can now formulate it as

$$\sum_{a \in A} \frac{1}{a} = \infty \Rightarrow A \text{ is a zero-solution set.}$$

Bibliography

- [1] Antal Balog (1992), Linear Equations in Primes, *Mathematika*, vol. 39, pp. 367-378.
- [2] Choi, Lui & Tsang (1992), Conditional bounds for small prime solutions of linear equations, *Manuscripta math. col 74*, pp. 321-340.
- [3] H. Furstenberg, Y. Katznelson and D. Ornstein (1982), The Ergodic Theoretical Proof of Szemerédi's Theorem, *Bull. of the AMS*, volume 7, number 3, pp. 527-552
- [4] W. T. Gowers (2001), A New Proof of Szemerédi's Theorem, *GAFSA 11*, pp. 465-588.
- [5] Andrew Granville (2008), Prime Number Patterns, *American Mathematical Monthly*, vol. 115, pp. 279-296.
- [6] Ben Green & Terence Tao (2008), The Primes Contain Arbitrarily Long Arithmetic Progressions, *Ann. of Math. (2) vol. 167 no. 2*, pp. 481-547.
- [7] Henryk Iwaniec & Emmanuel Kowalski (2004), *Analytic Number Theory*, American Mathematical Society, Colloquium Publications, vol. 53.
- [8] Jonas Lindstrøm Jensen (2009), *On solutions in arithmetic progressions to homogeneous systems of linear equations*, ISSN: 1397-4076, Preprint Series, Department of Mathematical Sciences, University of Aarhus. Can be obtained from <http://www.imf.au.dk/publs?id=697>.
- [9] Wenxin Liu, *Gowers Uniformity Norms and the Generalized Von Neuman Theorem*, taken from <http://ergodicpnt.googlepages.com/GeneralizedvonNeumannthm0115.pdf>.
- [10] W. Rudin, *Functional Analysis*, Second Edition, McGraw-Hill.
- [11] W. Rudin, *Real and complex analysis*, McGraw-Hill.
- [12] Endre Szemerédi (1975), On sets of integers containing no k elements in arithmetic progression, *Acta Arith. 27*, pp. 299-345.

- [13] B. L. van der Waerden (1927), Beweis einder Baudetschen Vermutung, *Nieuw. Arch. Wisk.* 15, pp. 212-216.
- [14] Peter Walters, *An Introduction to Ergodic Theory*, Springer Verlag.