ACTA ARITHMETICA
147.2 (2011)

Green–Tao theorem in function fields

by

Thái Hoàng Lê (Los Angeles)

1. Introduction. In [9], Green and Tao proved the following celebrated theorem now bearing their name:

Theorem 1 (Green–Tao). *The primes contain arithmetic progressions of arbitrary length. Furthermore, the same conclusion is true for any set of primes of positive relative upper density.*

Subsequently, other variants of this theorem have been proved. Tao and Ziegler [20] proved the generalization for polynomial progressions $a + p_1(d), \dots, a + p_k(d)$, where $p_i \in \mathbb{Z}[x]$ and $p_i(0) = 0$. Tao [19] proved the analog in the Gaussian integers. It is well known that the integers and the polynomials over a finite field share a lot of similarities in many aspects relevant to arithmetic combinatorics. Therefore, it is natural, as Green and Tao did, to suggest that the analog of this theorem should hold in the setting of function fields:

Conjecture 1. *For any finite field $\mathbb{F}$, the monic irreducible polynomials in $\mathbb{F}[t]$ contain affine spaces of arbitrarily high dimension.*

We will confirm this conjecture. More precisely, we will prove:

Theorem 2 (Green–Tao for function fields). *Let $\mathbb{F}_q$ be a finite field over q elements. Then for any $k > 0$, we can find polynomials $f, g \in \mathbb{F}_q[t]$, $g \neq 0$, such that the polynomials $f + Pg$, where P runs over all polynomials $P \in \mathbb{F}_q[t]$ of degree less than k, are all irreducible. Furthermore, such configurations can be found in any set of irreducible polynomials of positive relative upper density.*

Here we define the *upper density* of a set $\mathcal{A} \subset \mathbb{F}_q[t]$ to be

$$\overline{d}(\mathcal{A}) = \overline{\lim_{N\to\infty}} \frac{\#\{f \in \mathcal{A} : \deg(f) < N\}}{q^N},$$

2010 *Mathematics Subject Classification*: Primary 11B30; Secondary 11T55.
Key words and phrases: Green–Tao theorem, irreducible polynomials, finite fields.

DOI: 10.4064/aa147-2-3

and the *relative upper density* of $\mathcal{A}$ in the set $\mathcal{P}$ of all irreducible polynomials to be

$$\overline{d}_{\mathcal{P}}(\mathcal{A}) = \overline{\lim_{N\to\infty}} \frac{\#\{f \in \mathcal{A} : \deg(f) < N\}}{\#\{f \in \mathcal{P} : \deg(f) < N\}}.$$

The conjecture then follows since the monic polynomials have positive density in all the polynomials.

Our arguments follow Green–Tao's very closely. In many places the adaptation to the $\mathbb{F}_q[t]$ setting is immediate, so we will often omit the proof and refer the reader to Green–Tao's original paper. The readers interested in details can find the unabridged version of this paper at [11]. We have also chosen to incorporate some modifications that considerably simplify some major steps in the original arguments. Therefore, the paper may prove to be helpful to those who want to understand the ideas of the proof of Green–Tao's theorem.

2. Notation and outline of the proof

2.1. Notation. Throughout the paper, we will be working with a fixed field $\mathbb{F}_q$ on q elements, where q is a prime power. Let $\mathbb{F}_q[t]$ be the ring of polynomials with coefficients in $\mathbb{F}_q$. Let $\mathbb{F}_q(t)$ be the quotient ring of $\mathbb{F}_q[t]$, i.e. $\mathbb{F}_q(t) = \{f/g : f, g \in \mathbb{F}_q[t],\ g \neq 0\}$. The value of k will be kept fixed throughout the paper.

Denote by $\mathbb{G}_N$ the set of all polynomials in $\mathbb{F}_q[t]$ of degree less than N. A priori, $\mathbb{G}_N$ is an additive group. We also, for each N, fix a monic irreducible polynomial $f_N \in \mathbb{F}_q[t]$. Then the additive group $\mathbb{G}_N$ can be endowed with a field structure isomorphic to $\mathbb{F}_{q^N}$, the field on q^N elements, via multiplication modulo f_N. The need for the field structure arises in the same way as when we convert $\{1, \dots, N\}$ into $\mathbb{Z}_N$, the main reason being that we can freely perform divisions. Let $K = |\mathbb{G}_k| = q^k$, the number of polynomials of degree less than k. Let us call a set $\{f + Pg : P \in \mathbb{G}_k\}$ a *k-configuration*. If $g \neq 0$ then it is called a *nontrivial* k-configuration. A k-configuration in $\mathbb{G}_N$ is necessarily a k-configuration in $\mathbb{F}_{q^N}$, but not vice versa.

For a nonzero polynomial $f \in \mathbb{F}_q[t]$ define the norm of f to be $|f| = q^{\deg(f)}$. Also, let $|0| = 0$. Then the norm $|\cdot|$ defines a distance on $\mathbb{F}_q[t]$. Often, when dealing with the wraparound effect, we will make use of cylinder sets. A *cylinder set* of radius r is simply the set of all $f \in \mathbb{F}_q[t]$ whose distance to a given point is at most r. The cylinder sets are the analog of intervals in $\mathbb{R}$, with the more pleasant property that for any two cylinder sets, either they are disjoint or one is contained in the other.

If ϕ is a function on a finite set A, we write $\mathbb{E}(\phi(x) \,|\, x \in A)$, or $\mathbb{E}_{x\in A}\phi(x)$, or simply $\mathbb{E}_A\phi$ to denote the expectation of ϕ on A, in other words the

average value of ϕ on A. We denote the inner product of two functions ϕ, ψ on A as $\langle \phi, \psi \rangle = \mathbb{E}_{x \in A} \phi(x)\psi(x)$.

For two quantities A, B, we write $A = O(B)$, or $A \ll B$, or $B \gg A$ if there is an absolute constant C such that $|A| \leq CB$. If A and B are functions of the same variable x, we write $A = o_{x \to \infty}(B)$ if A/B tends to 0 as x tends to infinity. If the constant C (respectively, the rate of convergence of A/B) depends on a parameter, e.g. m, then we write $A = O_m(B)$ (respectively, $A = o_{m;x \to \infty}$). Dependence on fixed quantities such as q or k will often be omitted. Most of the time we will be dealing with functions of N, and when it is clear we will remove it from the notation. Thus $O(1)$ stands for a bounded quantity (independent of N) and $o(1)$ stands for a function that goes to 0 as N tends to infinity.

2.2. Outline of the proof. The starting point of Green–Tao is Szemerédi's theorem, which states that any set of integers of positive density contains arbitrarily long arithmetic progressions. Actually, they needed a stronger form of Szemerédi's theorem, obtained by incorporating an argument known as Varnavides's trick [21]. In the setting of function fields, an analog of Szemerédi's theorem is readily available [2], [1]. Coupled with the Varnavides argument, this gives the following result, which we will prove in Section 3:

THEOREM 3 (Szemerédi for function fields). *For every $\delta > 0$, there exists a constant $c(\delta) > 0$ such that, for every function $\phi : \mathbb{F}_{q^N} \to \mathbb{R}$ with $0 \leq \phi(x) \leq 1$ for all x and $\mathbb{E}(\phi \,|\, \mathbb{F}_{q^N}) \geq \delta$, we have $\mathbb{E}(\prod_{P \in \mathbb{G}_k} \phi(f + Pg) \,|\, f, g \in \mathbb{F}_{q^N}) \geq c(\delta)$.*

Following Green and Tao, our next step is a transference principle, which allows us to generalize Szemerédi's theorem to larger classes of functions ϕ, not necessarily bounded. Let us call a function $\nu : \mathbb{G}_N \to \mathbb{R}$ a *measure*. A *pseudorandom measure* is a measure satisfying two technical conditions (to be defined later in Section 4), called the linear forms condition and the correlation condition.

THEOREM 4 (Green–Tao–Szemerédi for function fields). *Let $\nu : \mathbb{F}_{q^N} \to \mathbb{R}$ be a pseudorandom measure. Then for every $\delta > 0$, there exists a constant $c'(\delta) > 0$ such that, for every function $\phi : \mathbb{F}_{q^N} \to \mathbb{R}$ with $0 \leq \phi(f) \leq \nu(f)$ for all f and $\mathbb{E}(f \,|\, \mathbb{F}_{q^N}) \geq \delta$, we have $\mathbb{E}(\prod_{P \in \mathbb{G}_k} \phi(f + Pg) \,|\, f, g \in \mathbb{F}_{q^N}) \geq c'(\delta) - o(1)$.*

This is obtained by means of a decomposition result, namely any function ϕ bounded by a pseudorandom measure can be decomposed as $\phi = \phi_1 + \phi_2$, where ϕ_1 is a nonnegative, bounded function whose average is bounded from below, and ϕ_2 is uniform in the sense that it is small in a norm (the Gowers norm to be defined later) that is relevant to counting k-configurations. Thus

the contribution of f_2 to $\mathbb{E}(\prod_{P\in\mathbb{G}_k}\phi(f+Pg)\,|\,f,g\in\mathbb{F}_{q^N})$ is small, so that the latter is close to $\mathbb{E}(\prod_{P\in\mathbb{G}_k}\phi_1(f+Pg)\,|\,f,g\in\mathbb{F}_{q^N})$, which is bounded from below by the usual Szemerédi theorem. The proof of the decomposition result in Green–Tao [9] and later in [20] is quite involved. Recently Gowers [5] and Reingold–Trevisan–Tulsiani–Vadhan [14], [15] have found much simpler proofs of this result, the main tool being the Hahn–Banach theorem. Moreover, their formulations of the result are very general and directly applicable to our setting of function fields.

Once Theorem 4 is established, the final step is to show that ν can be constructed in such a way that ν majorizes functions supported on irreducible polynomials, such as (variants of) the von Mangoldt function Λ, where

$$\Lambda(f)=\begin{cases}\deg(P) & \text{if } f=cP^k, \text{ where } P \text{ is irreducible and } c\in\mathbb{F}_q,\\ 0 & \text{otherwise.}\end{cases}$$

To this end, in Section 7 we will adapt the truncated divisor sum of Goldston and Yıldırım from their work on short gaps between primes [3].

THEOREM 5 (Goldston–Yıldırım for function fields). *For any $\mathcal{A}\subset\mathcal{P}$ such that $\overline{d}_{\mathcal{P}}(\mathcal{A})>0$, there exist a constant $\delta>0$, a pseudorandom measure $\nu:\mathbb{F}_{q^N}\to\mathbb{R}$, a function $\phi:\mathbb{F}_{q^N}\to\mathbb{R}$ and $W,b\in\mathbb{F}_{q^N}$ such that the following are true for infinitely many N:*

- *ϕ is 0 outside of $\{h\in\mathbb{G}_N : Wh+b\in\mathcal{A}\}$.*
- *$0\le\phi\le\nu$.*
- *$\mathbb{E}(\phi\,|\,\mathbb{F}_{q^N})\ge\delta$.*
- *$\|\phi\|_\infty\ll N$.*

REMARKS. The introduction of W, known as the "W-trick" and first used by Green in [7], is quite common in this situation in arithmetic combinatorics, when we want to transfer results about dense sets to the primes. We will need the irreducible polynomials to be distributed sufficiently uniformly in congruence classes, and for this purpose, we will take W to be a product of small irreducible polynomials.

Proof of Theorem 2 using Theorems 4 and 5. Suppose N is such that the conclusions of Theorem 5 holds. We partition $\mathbb{G}_N$ into q^k disjoint cylinder sets $\mathcal{C}_i$ of radius q^{N-k}, so that $|f_1-f_2|<q^{N-k}$ for any two polynomials f_1,f_2 in the same cylinder set. There must be a cylinder set $\mathcal{C}_i$ on which the average of ϕ is at least δ. Let $\psi=\phi 1_{\mathcal{C}_i}$. Applying Theorem 4 to the function ψ, we have $\mathbb{E}(\prod_{P\in\mathbb{G}_k}\psi(f+Pg)\,|\,f,g\in\mathbb{F}_{q^N})\ge c'(\delta/q^k)-o(1)$. Because of the bound on the magnitude of ϕ, the contribution of the products corresponding to trivial k-configurations is $o(1)$. Thus for N sufficiently large, ψ is nonzero on some nontrivial k-configuration $\{f+Pg : P\in\mathbb{G}_k\}\subset\mathbb{F}_{q^N}$. A priori, this is a k-configuration in $\mathbb{F}_{q^N}$. By the definition of ψ, $f+Pg\in\mathcal{C}_i$ for every $P\in\mathbb{G}_k$.

In particular $q^{N-k} > |(f+g)-f| = |g|$, so that the above k-configuration is indeed a k-configuration in $\mathbb{G}_N$. Thus $\{W(f+Pg)+b : P \in \mathbb{G}_k\}$ is a nontrivial k-configuration that lies entirely in $\mathcal{A}$, since ψ is supported in $\{h \in \mathbb{G}_N : Wh+b \in \mathcal{A}\}$. ∎

Remarks. The techniques here not only give infinitely many k-configurations, but also show that the number of such configurations is $\gg q^{2N}/N^K$, which is of correct magnitude in the context of the Hardy–Littlewood conjecture on tuples of primes. We remark that while more algebraic methods can generate configurations of irreducibles, e.g. the analog of the twin prime conjecture ([12, Theorem 4]), such methods do not give the correct bound (up to a constant).

The next sections are organized as follows. In Section 3 we establish Theorem 3. In Section 4, we define the machinery of Green–Tao including pseudorandom measures, Gowers norms and dual functions. In Section 5, we prove the decomposition result mentioned earlier. In Section 6, we introduce arithmetic functions in $\mathbb{F}_q[t]$. We give in Section 7 the construction of a measure ν that majorizes the irreducible polynomials. Section 8 is devoted to establishing the pseudorandomness of ν, thus finishing our proof of Theorem 2.

3. Szemerédi's theorem in function fields. As we mentioned in the last section, we need an analog of Szemerédi's theorem in $\mathbb{F}_q[t]$, namely that we can find nontrivial k-configurations inside any subset of $\mathbb{F}_q[t]$ of positive upper density:

Proposition 1. *Let $\delta > 0$. Then for N sufficiently large, $N \geq N_0 = N_0(q,k,\delta)$, in every subset A of size δq^N of $\mathbb{G}_N$, we can find polynomials $f, g \in \mathbb{F}_q[t]$, $g \neq 0$, such that $f + Pg \in A$ for every polynomial $P \in \mathbb{G}_k$.*

There are at least two ways to see this. It is an immediate consequence of a far more general result of Bergelson–Leibman–McCutcheon [1]:

Theorem 6 (Polynomial Szemerédi for countable integral domains). *Let K be a countable integral domain, M be a finitely generated K-module, and $p_1, \dots, p_n$ be polynomials $K \to M$ such that $p_i(0) = 0$ for every i. Then for any set $A \subset M$ of upper Banach density $d^*(A) > 0$, there exist $d \in K$, $d \neq 0$, and $a \in M$ such that $a + p_i(d) \in A$ for every $i = 1, \dots, n$.*

When $K = M = \mathbb{F}_q[t]$ and $p_1, \dots, p_n$ are the linear polynomials $g \mapsto Pg$, where $P \in \mathbb{G}_k$, then we have the desired result. Proposition 1 can also be proved using the density Hales–Jewett theorem [2] ([1]). For the details, see [11].

([1]) These two approaches are not unrelated. In fact, the Hales–Jewett density theorem is one of the ingredients in the proof of the Bergelson–Leibman–McCutcheon theorem.

A Varnavides argument [21] shows that not only is there such a k-configuration, but there are in fact many of them:

PROPOSITION 2. *Let $\delta > 0$. Then there is a constant $c(\delta) > 0$ such that, for N sufficiently large, in every subset A of size δq^N of $\mathbb{G}_N$, we can find at least $c(\delta)q^{2N}$ k-configurations.*

Proof. Let $m = N_0(q, k, \delta/2)$, where N_0 is the function of Proposition 1, and suppose $N \geq m$. We claim that among the q^{2N-m} m-configurations in $\mathbb{G}_N$, there are at least $(\delta/2)q^{2N-2m}$ on which the density of A is at least $\delta/2$. Indeed, let us count the number of pairs (V, h) where V is an m-configuration in $\mathbb{G}_N$ and $h \in A \cap V$. On the one hand, since for every given point in $\mathbb{G}_N$ there are q^{N-m} m-configurations in $\mathbb{G}_N$ containing it, the number of such pairs is $\delta q^N q^{N-m} = \delta q^{2N-m}$. Each k-configuration V on which the density of A is at most $\delta/2$ contributes at most $(\delta/2)q^m q^{2N-2m} = (\delta/2)q^{2N-m}$ pairs. Thus the contribution of those V on which the density of A is at least $\delta/2$ is at least $\delta q^{2N-m} - (\delta/2)q^{2N-m} = (\delta/2)q^{2N-m}$. Therefore, the number of m-configurations V on which the density of A is $\geq \delta/2$ is at least $(\delta/2)q^{2N-2m}$.

From the definition of m, each such m-configuration (with the exception of at most q^N trivial m-configurations) contains a nontrivial k-configuration. Any given k-configuration is counted at most q^{m-k} times. Thus the number of nontrivial k-configurations in $\mathbb{G}_N$ is at least $q^{k-m}((\delta/2)q^{2N-m} - q^N) \gg_\delta q^{2N}$, as desired. ∎

From this, Theorem 3 easily follows.

Proof of Theorem 3. Suppose $\mathbb{E}(\phi \,|\, \mathbb{G}_N) \geq \delta$. Let $B \subset \mathbb{G}_N$ be the set on which $\phi \geq \delta/2$. Then $|B| + (\delta/2)(q^N - |B|) \geq \delta q^N$, so that $|B| \geq (\delta/2)q^N$. Proposition 2 implies that B contains at least $c(\delta/2)q^{2N}$ k-configurations. Thus $\mathbb{E}(\prod_{P \in \mathbb{G}_k} \phi(f + Pg) \,|\, f, g \in \mathbb{F}_{q^N}) \geq (\delta/2)^k c(\delta/2)$. ∎

REMARK. In contrast with the usual Szemerédi theorem for the integers, where we have a quantitative proof due to Gowers [4], [5], the proof of the Bergelson–Leibman–McCutcheon theorem uses ergodic theory and therefore does not give any bound of $c(\delta)$ in terms of δ. As for the density Hales–Jewett theorem, very recently, researchers in the Polymath collaborative project initiated by Gowers managed to find combinatorial proofs [13], from which some bounds might be extracted, but still far weaker than what is known for the integers. Thus we do not seek to find a bound for the first occurrence in the irreducible polynomials of the configurations $\{f + Pg : P \in \mathbb{G}_k\}$.

4. Pseudorandom measures, Gowers norms, and Gowers anti-uniformity. In this section we will introduce the machinery necessary for our proof of the generalized Szemerédi theorem (Theorem 3). At this stage

there are hardly any differences between the function field setting and the integer setting, and the proofs are almost identical, so we will skip the proofs and refer the reader to the unabridged version of this paper [11], or Sections 3, 5, 6 of Green–Tao's original paper [9].

4.1. Pseudorandom measures. A *measure* is a function ([2]) $\nu : \mathbb{F}_{q^N} \to [0,\infty)$. A *pseudorandom measure* is a measure satisfying the two conditions defined below:

DEFINITION 1 (Linear forms condition). We say that a measure $\nu : \mathbb{F}_{q^N} \to [0,\infty)$ satisfies the (m_0, n_0, k_0)-*linear forms condition* if whenever we have $m \le m_0$ linear forms in $n \le n_0$ variables $\psi_1, \dots, \psi_m : (\mathbb{F}_{q^N})^n \to \mathbb{G}_N$ of the form $\psi_i(\mathbf{f}) = \sum_{j=1}^{n} L_{ij} f_i + b_i$ such that all the coefficients L_{ij} are in the set ([3]) $\{f/g : f, g \in \mathbb{G}_{k_0}\}$, and no two of the vectors $(L_{ij})_{1\le j\le n}$, $i = 1, \dots, m$, are proportional, then we have

$$\mathbb{E}(\nu(\psi_1(\mathbf{f})) \cdots \nu(\psi_m(\mathbf{f})) \,|\, \mathbf{f} \in (\mathbb{F}_{q^N})^n) = 1 + o_{N\to\infty}(1).$$

In particular, if ν satisfies the linear forms condition then $\mathbb{E}(\nu(f) \,|\, f \in \mathbb{G}_N) = 1 + o(1)$. Note that we require the $o_{N\to\infty}(1)$ to be uniform in all choices of $b_1, \dots, b_m \in \mathbb{F}_{q^N}$.

DEFINITION 2 (Correlation condition). We say that a measure $\nu : \mathbb{F}_{q^N} \to [0,\infty)$ satisfies the l_0-*correlation condition* if whenever we have $l \le l_0$ linear forms of the form $f + h_1, \dots, f + h_l$ with $h_1, \dots, h_l \in \mathbb{F}_{q^N}$, then

$$\mathbb{E}(\nu(f + h_1) \cdots \nu(f + h_l) \,|\, f \in \mathbb{F}_{q^N}) \le \sum_{1\le i\le j\le q} \tau(h_i - h_j)$$

where τ is a function $\mathbb{G}_N \to \mathbb{R}^+$ such that $\mathbb{E}(\tau(f)^p \,|\, f \in \mathbb{F}_{q^N}) = O_p(1)$ for every $p > 1$.

The point is that the function τ is not necessarily bounded as N tends to infinity, but its L^p-norm is always bounded.

DEFINITION 3 (Pseudorandom measures). A measure $\nu : \mathbb{F}_{q^N} \to \infty$ is called k-*pseudorandom* if it satisfies the $(K2^{K-1}, 3K - 4, k)$-linear forms condition and the 2^{K-1}-correlation condition (recall that $K = q^k$).

REMARK. The exact values of the parameters m_0, n_0, k_0, l_0 are not too important. However, it is essential in the construction that these values are finite. From now on we will refer to k-pseudorandom measures as *pseudorandom measures*.

([2]) More precisely, it is a family $\{\nu_N\}_{N\in\mathbb{Z}^+}$ such that for each N, ν_N is a function from $\mathbb{F}_{q^N} \to \mathbb{R}$.

([3]) Note that this set can be embedded in $\mathbb{F}_{q^N}$ for every $N \ge k_0$. Of course, the embedding depends on our choice of the irreducible polynomial f_N underlying $\mathbb{F}_{q^N}$.

LEMMA 1. *If ν is pseudorandom, then so is $\nu_{1/2} = (\nu + 1)/2$. More generally, for any $0 < \alpha < 1$, $\nu_\alpha = (1 - \alpha)\nu + \alpha$ is also pseudorandom.*

In practice we will be dealing with $\nu_{1/2}$ and $\nu_{1/4}$.

4.2. Gowers norms. One efficient tool in counting linear patterns is Gowers norms. The Gowers norms have been first used by Gowers in his proof of Szemerédi's theorem [4], [5]. They have a counterpart in ergodic theory, known as the Host–Kra seminorm [10].

DEFINITION 4 (Gowers norm). Let G be a finite abelian group, and ϕ a complex-valued function defined on G. For $\omega = (\omega_1, \dots, \omega_d) \in \{0,1\}^d$, let $|\omega| = \omega_1 + \cdots + \omega_d$. Also, let C be the complex conjugation. We define the dth *Gowers norm* of ϕ to be

$$\|\phi\|_{U^d(G)} = \Big(\mathbb{E}_{x,h_1,\dots,h_k \in G} \prod_{\omega \in \{0,1\}^d} C^{|\omega|}\phi(x + \omega_1 h_1 + \cdots + \omega_d h_d)\Big)^{1/2^d}.$$

Alternatively, the Gowers norms $\|\cdot\|_{U^d(G)}$ can be defined recursively as follows:

$$\|\phi\|_{U^1(G)} = |\mathbb{E}(\phi \,|\, G)|, \qquad \|\phi\|^{2^{d+1}}_{U^{d+1}(G)} = \mathbb{E}(\|\phi \cdot \phi_t\|^{2^d}_{U^d(G)} \,|\, t \in G),$$

where ϕ_t is the function $\phi_t(x) = \phi(t + x)$. The following facts about the Gowers norms are standard and the proofs can be found in [9] or [8]:

PROPOSITION 3.

- *Suppose ϕ_ω, for $\omega \in \{0,1\}^d$, are 2^d functions $G \to \mathbb{C}$. Then*
$$\mathbb{E}\Big(\prod_{\omega \in \{0,1\}^d} \phi_\omega(x + \omega_1 h_1 + \cdots + \omega_d h_d)\Big) \le \prod_{\omega \in \{0,1\}^d} \|\phi_\omega\|_{U^d(G)}.$$
- *For every ϕ, the sequence $\|\phi\|_{U^d(G)}$, $d = 1, 2, \dots$, is an increasing sequence. In particular for every $d \ge 1$, $\|\phi\|_{U^d(G)} \ge \|\phi\|_{U^1(G)} = |\mathbb{E}(\phi \,|\, G)|$.*
- *For every $d \ge 2$, $\|\cdot\|_{U^d(G)}$ is indeed a norm on $\mathbb{C}^G$, the space of complex functions on G.*

Henceforth, if the context is clear, we will assume $G = \mathbb{F}_{q^N}$ and omit the group in the notation of the Gowers norm. In practice we will be dealing with the U^{K-1} norm, where $K = q^k$. Our first observation is that a pseudorandom measure is close to the uniform measure in the U^{K-1} norm.

LEMMA 2. *Let ν be a pseudorandom measure on $\mathbb{F}_{q^N}$. Then $\|\nu - 1\|_{U^{K-1}} = o(1)$. Consequently, $\|\nu\|_{U^{K-1}} = 1 + o(1)$.*

As mentioned before, Gowers norms are effective in counting linear patterns, as witnessed by the following:

PROPOSITION 4 (Generalized von Neumann ([4])). *Suppose ν is pseudorandom. Let $(\phi_P)_{P\in\mathbb{G}_k}$ be functions bounded in absolute value by ν. Then*

$$\mathbb{E}\Big(\prod_{P\in\mathbb{G}_k}\phi_P(f+Pg)\;\Big|\; f,g\in\mathbb{F}_{q^N}\Big)\le\min_P\|\phi_P\|_{U^{K-1}}+o_{N\to\infty}(1).$$

COROLLARY 1. *If the ϕ_P are bounded by $3+\nu$ then*

$$\mathbb{E}\Big(\prod_{P\in\mathbb{G}_k}\phi_P(f+Pg)\;\Big|\; f,g\in\mathbb{F}_{q^N}\Big)\le 4^K\min_P\|\phi_P\|_{U^{K-1}}+o_{N\to\infty}(1).$$

4.3. Gowers anti-uniformity

DEFINITION 5. For a real function ϕ on $\mathbb{F}_{q^N}$, define its U^d *dual function* $\mathcal{D}_d\phi$ by

$$\mathcal{D}_d\phi=\mathbb{E}\Big(\prod_{\omega\in\{0,1\}^d,\,\omega\neq 0}\phi(x+\omega_1h_1+\cdots+\omega_dh_d)\;\Big|\; x,h_1,\dots,h_d\in\mathbb{F}_{q^N}\Big).$$

From a functional analytic point of view $\mathcal{D}_d\phi$ may be regarded as a "support functional" of ϕ, with the difference that $\mathcal{D}\phi$ is not linear. From now on we will be working with the U^{K-1} dual functions. We will be particularly interested in the dual functions of functions bounded by a pseudorandom measure ν. It can be shown that the dual functions have the following properties:

PROPOSITION 5.

- $\langle\phi,\mathcal{D}_d\phi\rangle=\ \|\phi\|_{U^d(G)}^{2^d}$.
- *If $0\le\phi\le\nu$, then $\langle\phi,\mathcal{D}_{K-1}\phi\rangle=1+o(1)$.*
- *For every m there is a constant $C(m)$ such that if $0\le\phi_1,\dots,\phi_m\le\nu$, then*
$$\|\mathcal{D}_{K-1}\phi_1\cdots\mathcal{D}_{K-1}\phi_m\|_{U^{K-1}}^*\le C(m)$$
where $\|\cdot\|_{U^{K-1}}^*$ is the dual norm of $\|\cdot\|_{U^{K-1}}$ (defined in the usual way, $\|f\|_{U^{K-1}}^*=\sup\{|\langle f,g\rangle|:\|g\|_{U^{K-1}}\le 1\}$).

REMARKS. The third property is by far the most important property of the dual functions, and perhaps surprising, since m is not bounded, while the number of forms in the linear forms condition and correlation condition is bounded. However, this comes from the fact that the exponent p of the function τ in the correlation condition is not bounded. In Reingold–Trevisan–Tulsiani–Vadhan's language this means that ν is indistinguishable from the uniform measure according to the family $\{\mathcal{D}_{K-1}\phi_1\cdots\mathcal{D}_{K-1}\phi_m:0\le\phi_i\le\nu\}$.

([4]) Green and Tao call this type of inequalities generalized von Neumann theorems to emphasize their connection with the classical von Neumann theorem in ergodic theory.

5. A decomposition and a transference principle. In this section we reproduce Gowers' proof [6] of Green–Tao's structure theorem and use the latter to derive Theorem 4. The reader is however referred to the original paper for a survey about the interplay between decomposition results and the use of the Hahn–Banach theorem in arithmetic combinatorics.

We first forget for a moment the definitions of Gowers norms and dual functions, but instead axiomatize their properties in Proposition 5. Consider a finite set G and let $\mathbb{R}^G$ be the set of all real functions on G with the inner product $\langle f, g\rangle = \mathbb{E}_{x\in G} f(x)g(x)$.

DEFINITION 6. We say that a norm $\|\cdot\|$ on $\mathbb{R}^G$ is a *quasi-algebra predual norm* with respect to a convex, compact set $\mathcal{F} \subset \mathbb{R}^G$ if there is a function $c : \mathbb{R}^+ \to \mathbb{R}^+$, a function $C : \mathbb{Z}^+ \to \mathbb{R}^+$, and an operator $\mathcal{D} : \mathbb{R}^G \to \mathbb{R}^G$ such that the following hold:

- $\langle f, \mathcal{D}f\rangle \le 1$ for every $f \in \mathcal{F}$.
- $\langle f, \mathcal{D}f\rangle \ge c(\epsilon)$ for every $f \in \mathcal{F}$ with $\|f\| \ge \epsilon$.
- $\|\mathcal{D}f_1 \cdots \mathcal{D}f_m\|^* \le C(m)$, where $\|\cdot\|^*$ is the dual norm of $\|\cdot\|$.
- The set $\{\mathcal{D}f : f \in \mathcal{F}\}$ is compact and spans $\mathbb{R}^G$.

The reason why $\|\cdot\|$ is called a quasi-algebra predual norm is that the dual norm $\|\cdot\|^*$ is "close" to being an algebra norm (this will be made precise in Lemma 4). The application we have in mind is when $G = \mathbb{F}_{q^N}$, $\|\cdot\|$ is the (normalized) U^{K-1} Gowers norm, the $\mathcal{D}f$ are the (normalized) U^{K-1} dual functions, and $\mathcal{F}$ is the space of nonnegative functions bounded by a pseudorandom measure ν.

Associated to the norm $\|\cdot\|$, we will also consider the norm $\|g\|_{\mathrm{BAC}} = \max\{|\langle g, \mathcal{D}f\rangle| : f \in \mathcal{F}\}$ and its dual $\|\cdot\|^*_{\mathrm{BAC}}$ (since the set $\{\mathcal{D}f : f \in \mathcal{F}\}$ is compact, $\|\cdot\|_{\mathrm{BAC}}$ is indeed a norm). Here BAC stands for Basic Anti-uniform Correlation. Thus $\|\cdot\|$ and $\|\cdot\|_{\mathrm{BAC}}$ are equivalent in the sense that if $f \in \mathcal{F}$ and $\|f\| \ge \epsilon$ then $\|f\|_{\mathrm{BAC}} \ge c(\epsilon)$. The following gives a simple characterization of the $\|\cdot\|^*_{\mathrm{BAC}}$ norm.

LEMMA 3. $\|f\|^*_{\mathrm{BAC}} = \inf\{\sum_{i=1}^k |\lambda_i| : f = \sum_{i=1}^k \lambda_i \mathcal{D}f_i, f_1, \ldots, f_k \in \mathcal{F}\}$.

Proof. This is [6, Corollary 3.5], and can also be proven using Farkas' lemma [17, Section 1.16] (which is another incarnation of the Hahn–Banach theorem!). ∎

We now see that the name "quasi-algebra predual" is justified by the following:

LEMMA 4. *If $\psi \in \mathbb{R}^G$ is such that $\|\psi\|^*_{\mathrm{BAC}} \le 1$, then $\|\psi^m\|^* \le C(m)$. More generally, for any polynomial P, $\|P(\psi)\|^* \le C(P)$ for some constant $C(P)$ depending on P alone.*

Proof. This is [6, Lemma 4.7]. ∎

Specializing to the case where $\mathcal{F}$ is the set of all nonnegative functions bounded by a function $\nu \in \mathbb{R}^G$, we claim that any function from $\mathcal{F}$ can be written as the sum of a bounded function and another function small under $\|\cdot\|$. This was essentially the key result in the paper of Green and Tao [9]. The slight generalization given here was later given in the paper of Tao and Ziegler [20]:

THEOREM 7 (Green–Tao structure theorem). *For every $\eta > 0$, there is $\epsilon = \epsilon(\eta, C, c) > 0$ such that the following holds: Let ν be a measure on G such that $\|\nu - 1\| < \epsilon$, $\mathbb{E}_G \nu \le 1 + \eta$, and all properties in Definition 6 hold for $\mathcal{F} = \{f : 0 \le f \le \nu\}$. Then every function $f \in \mathcal{F}$ can be decomposed as $f = g + h$, where $0 \le g \le 1 + \eta$ and $\|h\| \le \eta$.*

Proof. Suppose such a decomposition does not exist. Since $\|h\|_{\mathrm{BAC}} \le c(\eta)$ implies $\|h\| \le \eta$, this implies that f cannot be expressed as the sum of elements from two convex sets $X_1 = \{0 \le g \le 1 + \eta\}$ and $X_2 = \{\|h\|_{\mathrm{BAC}} \le c(\eta)\}$ in $\mathbb{R}^G$.

CLAIM 1. *There is a function $\psi \in \mathbb{R}^G$ such that $\langle f, \psi\rangle > 1$, but $\langle g, \psi\rangle \le 1$ and $\langle h, \psi\rangle \le 1$ for all $g \in X_1$ and $h \in X_2$.*

Proof. Let $X = X_1 + X_2$; then X is convex and closed. We invoke the following form of the Hahn–Banach theorem: if $f \notin X$, then there is a linear functional $\langle \cdot, \psi\rangle$ on $\mathbb{R}^G$ such that $\langle f, \psi\rangle > 1$ and $\langle g, \psi\rangle \le 1$ for every $g \in X$. Since X_1 and X_2 both contain 0, they are contained in X and the claim follows. ∎

The condition $\langle g, \psi\rangle \le 1$ for every $g \in X_1$ implies that $\mathbb{E}_G \psi_+ \le 1/(1 + \eta)$, where $\psi_+(x) = \max(0, \psi(x))$. The condition $\langle h, \psi\rangle \le 1$ for every $h \in X_2$ implies that $\|\psi\|^*_{\mathrm{BAC}} \le c(\eta)^{-1}$.

CLAIM 2. *For any $\eta' > 0$, there is a polynomial $P = P(\eta, \eta', C, c)$ and a constant $R = R(\eta, \eta', C, c)$ such that $\|P\psi - \psi_+\|_\infty \le \eta'$ and $\|P\psi\|^* \le R$.*

Proof. Since $\|\psi\|^*_{\mathrm{BAC}} \le c(\eta)^{-1}$, by Lemma 4 we have $\|\psi\|^* \le C_1 = C(1)c(\eta)^{-1}$. By Weierstrass' approximation theorem, there is a polynomial $P(x) = a_n x^n + \cdots + a_0$ such that $|P(x) - \max(0, x)| \le \eta'$ for every $x \in [-C_1, C_1]$. Then clearly $\|P\psi - \psi_+\|_\infty \le \eta'$. The claim that $\|P\psi\|^*$ is bounded (independently of ψ) follows from Lemma 4. ∎

We now have $1 < \mathbb{E}_G f\psi_+ \le \mathbb{E}_G \nu\psi_+$. We split the latter as

$$\mathbb{E}_G \nu\psi_+ = \mathbb{E}_G \psi_+ + \mathbb{E}_G(\nu - 1)P\psi + \mathbb{E}_G(\nu - 1)(\psi_+ - P\psi).$$

Also, $|\mathbb{E}_G(\nu - 1)P\psi| \le \|\nu - 1\|\,\|P\psi\|^* \le \epsilon R$ and

$$|\mathbb{E}_G \nu(\psi_+ - P\psi)| \le (\mathbb{E}_G \nu)\|P\psi - \psi_+\|_\infty \le \eta'(1 + \eta).$$

Thus $1 \le 1/(\eta + 1) + \eta'(1 + \eta) + \epsilon R$. If we fix a small value of η' (e.g. $\eta' = \eta/12$ will do), then this is a contradiction for ϵ small enough. ∎

Let us now formulate the result in the setting of Gowers norms and pseudorandom measures on $\mathbb{F}_{q^N}$:

COROLLARY 2. *Let ν be a pseudorandom measure on $\mathbb{F}_{q^N}$. Then for every $\eta > 0$, for N sufficiently large, every function ϕ on $\mathbb{F}_{q^N}$ such that $0 \le \phi \le \nu$ can be decomposed as $\phi = \phi_1 + \phi_2$, where $0 \le \phi_1 \le 2 + \eta$ and ϕ_2 is uniform in the sense that $\|\phi_2\|_{U^{K-1}} \le \eta$.*

Proof. If $0 \le \phi \le \nu$, then $0 \le \phi/2 \le \nu_{1/2} = (\nu + 1)/2$. We already know that $\nu_{1/2}$ is also pseudorandom. Let $G = \mathbb{F}_{q^N}$ and $\mathcal{F}$ be the space of all nonnegative functions bounded by $\nu_{1/2}$. Let us check that the normalized Gowers U^{K-1} norm $\|\phi\| = \frac{1}{2}\|\phi\|_{U^{K-1}}$ is quasi-algebra predual with respect to $\mathcal{F}$, where $\mathcal{D}\phi = \frac{1}{2}\mathcal{D}_{K-1}\phi$. Thanks to Proposition 5, the first three conditions in Definition 6 are met. The only thing left to check is the fourth condition, that the set of dual functions $\mathcal{D}\phi$ spans $\mathbb{R}^G$. Note that if ϕ is a point mass, then $\mathcal{D}\phi$ is also a point mass (at the same point). Since $\nu_{1/2}$ is pointwise positive ([5]), $\{\mathcal{D}\phi : \phi \in \mathcal{F}\}$ contains masses at every point of G, hence spans $\mathbb{R}^G$. By Theorem 7, there is $\epsilon = \epsilon(\eta) > 0$ such that we have a decomposition $\phi/2 = \phi_1 + \phi_2$, where $0 \le \phi_1 \le 1 + \eta/2$ and $\|\phi_2\|_{U^{K-1}} \le \eta/2$ as soon as $\mathbb{E}_G \nu_1 \le 1 + \eta$ and $\|\nu_1 - 1\| \le \epsilon$. But this is always true since ν_1 is a pseudorandom measure. Such a decomposition for $\phi/2$ gives the desired decomposition for ϕ. ∎

With this in hand, we can now prove Theorem 4:

Proof of Theorem 4 using the Green–Tao structure theorem. We know that for every $\eta > 0$, for N sufficiently large (depending on η), every function ϕ bounded by a pseudorandom measure on $\mathbb{F}_{q^N}$ can be decomposed as $\phi = \phi_1 + \phi_2$, where $0 \le \phi_1 \le 2 + \eta$ and $\|\phi_2\|_{U^{K-1}} \le \eta$. In particular $|\mathbb{E}\phi_2| \le \eta$, so that if $\mathbb{E}\phi \ge \delta$, then $\mathbb{E}\phi_1 \ge \delta - \eta$. Write

$$\mathbb{E}\Big(\prod_{P \in \mathbb{G}_k} \phi(f + Pg) \;\Big|\; f, g \in \mathbb{F}_{q^N} \Big) = \mathbb{E}\Big(\prod_{P \in \mathbb{G}_k} \phi_1(f + Pg) \;\Big|\; f, g \in \mathbb{F}_{q^N} \Big) + (2^K - 1 \text{ other terms}).$$

The other terms are of the form $\mathbb{E}(\prod_{P \in \mathbb{G}_k} \phi_P(f + Pg) \,|\, f, g \in \mathbb{F}_{q^N})$ where each ϕ_P is ϕ_1 or ϕ_2, and not all ϕ_P are equal to ϕ_1.

Since ϕ_1 is pointwise bounded by $2 + \eta$ and ϕ_2 is pointwise bounded by $\max(\nu, 2 + \eta) \le 3 + \nu$ in absolute value, by Corollary 1, these terms are at most $4^K \|\phi_2\|_{U^{K-1}} + o(1)$ in absolute value. On the other hand, by Theorem 3,

$$\mathbb{E}\Big(\prod_{P \in \mathbb{G}_k} \phi_1(f + Pg) \;\Big|\; f, g \in \mathbb{F}_{q^N} \Big) \ge (2 + \eta)^K c\Big(\frac{\delta - \eta}{2 + \eta} \Big).$$

([5]) This is the sole reason why we work with $\nu_{1/2}$ rather than with ν.

Hence

$$\mathbb{E}\Big(\prod_{P\in\mathbb{G}_k}\phi(f+Pg)\;\Big|\; f,g\in\mathbb{F}_{q^N}\Big)\geq(2+\eta)^K c\Big(\frac{\delta-\eta}{2+\eta}\Big)-(2^K-1)4^K\eta-o(1).$$

By choosing η appropriately small, the main term on the right hand side is positive, so that there is a positive constant $c'(\delta)$ such that

$$\mathbb{E}\Big(\prod_{P\in\mathbb{G}_k}\phi(f+Pg)\;\Big|\; f,g\in\mathbb{F}_{q^N}\Big)\geq c'(\delta)-o(1)$$

for every function ϕ on $\mathbb{F}_{q^N}$ bounded by a pseudorandom measure. ∎

6. Elementary arithmetic in $\mathbb{F}_q[t]$. We assume from now on that symbols involving the letter P (such as P, P', or P_i) will stand for monic, irreducible polynomials. Every $f\in\mathbb{F}_q[t]$ can be uniquely written as $f=cP_1^{\alpha_1}\cdots P_m^{\alpha_m}$, where $c\in\mathbb{F}_q$ and $\alpha_i\in\mathbb{Z}^+$. We can now introduce arithmetic functions on $\mathbb{F}_q[t]$:

- The Euler totient function $\Phi(f)$ is the number of polynomials of degree less than $\deg(f)$ which are relatively prime to f. Then we have the following formula for $\Phi(f)$ in terms of the prime factorization of f:
$$\Phi(f)=|f|\prod_{P|f}\Big(1-\frac{1}{|P|}\Big)=\prod_{i=1}^{m}\frac{|P_i^{\alpha_i+1}|-1}{|P_i|-1}.$$
- The Möbius function
$$\mu(f)=\begin{cases}(-1)^m & \text{if } \alpha_i=1 \text{ for every } i=1,\dots,m,\\ 0 & \text{otherwise.}\end{cases}$$
- The von Mangoldt function
$$\Lambda(f)=\begin{cases}\deg(f) & \text{if } m=1,\\ 0 & \text{otherwise.}\end{cases}$$
- $d(f)$, the number of monic divisors of f. We have the formula
$$d(f)=\prod_{i=1}^{m}(\alpha_i+1).$$
- For $d_1,\dots,d_m\in\mathbb{F}_q[t], d_i\neq 0$, denote by $[d_1,\dots,d_m]$ the least common divisor of $d_1,\dots,d_m$, in other words, the polynomial of smallest degree that is divisible by d_i for every $i=1,\dots,m$ (which is defined up to multiplication by an element of $\mathbb{F}_q\setminus\{0\}$).

The zeta function ζ_q of $\mathbb{F}_q[t]$ is defined by

$$\zeta_q(s)=\sum_{f\text{ monic}}\frac{1}{|f|^s}$$

for any $s \in \mathbb{C}$ such that $\Re s > 1$. We have the following closed form for the zeta function:

$$\zeta_q(s) = \frac{1}{1 - q^{1-s}} \quad \text{for } \Re s > 1.$$

Thus it can be analytically continued on the whole plane, with a simple pole at $s = 1$ with residue $1/\log q$. Similarly to the Riemann zeta function, ζ_q admits a factorization as an Euler product:

$$\zeta_q(s) = \prod_P \left(1 - \frac{1}{|P|}\right)^{-1}.$$

We have the following analogs of the prime number theorem and Dirichlet's theorem on primes in arithmetic progressions in function fields [16]. Note that the error terms are much better than their integer counterparts, thanks to the Riemann hypothesis for curves over a finite field [16].

Proposition 6.

- *Let $\pi_q(N)$ be the number of irreducible polynomials of degree N in $\mathbb{F}_q[t]$. Then*

$$\pi_q(N) = (q-1)\frac{q^N}{N} + O\left(\frac{q^{N/2}}{N}\right).$$

- *Let $a, r \in \mathbb{F}_q[t]$ be relatively prime, and $\deg(m) > 0$. Let $\pi_q(N; a, r)$ be the number of irreducible polynomials of degree N in $\mathbb{F}_q[t]$ which are congruent to r (modulo a). Then*

$$\pi_q(N; a, r) = (q-1)\frac{1}{\Phi(m)}\frac{q^N}{N} + O\left(\frac{q^{N/2}}{N}\right).$$

We will need the following two lemmas in our construction of the function τ in the correlation condition.

Lemma 5 (Divisor bound). *Let $f \in \mathbb{F}_q[t]$ with $\deg(f) = N$. Then*

$$d(f) \leq q^{O_q(N/\log N)}.$$

Proof. If f has the factorization $f = c\prod_{i=1}^m P_i^{\alpha_i}$, then we have $d(f) = \prod_{i=1}^m (\alpha_i + 1)$. Therefore,

$$\frac{d(f)}{|f|^\epsilon} = \prod_{i=1}^m \frac{\alpha_i + 1}{|P_i|^{\epsilon\alpha}},$$

where ϵ is to be chosen later, possibly depending on f. Note that if $\deg(P_i) \geq 1/\epsilon$, then

$$\frac{\alpha_i + 1}{|P_i|^{\epsilon\alpha_i}} \leq \frac{\alpha_i + 1}{q^{\alpha_i}} \leq 1.$$

If $\deg(P_i) < 1/\epsilon$, then

$$\frac{\alpha_i+1}{|P_i|^{\epsilon\alpha_i}} \le \frac{\alpha_i+1}{q^{\epsilon\alpha_i}} \le \frac{q^{2\sqrt{\alpha_i}}}{q^{\epsilon\alpha_i}} \le q^{1/\epsilon}.$$

Since the second case can occur for at most $q^{1/\epsilon}$ values of P_i, we have

$$\frac{d(f)}{q^{N\epsilon}} \le (q^{1/\epsilon})^{q^{1/\epsilon}} = q^{N\epsilon+\frac{1}{\epsilon}q^{1/\epsilon}}.$$

Thus $d(f) \le q^{N\epsilon+(1/\epsilon)q^{1/\epsilon}}$ for every $\epsilon > 0$. If we choose $\epsilon = 1/\log N$, then we obtain $d(f) \le q^{O_q(N/\log N)}$, as required. ∎

LEMMA 6. *Let $\mathcal{S}$ be a finite set of irreducible polynomials in $\mathbb{F}_q[t]$. Then for every K,*

$$\exp\Big(\sum_{P\in\mathcal{S}} \frac{1}{|P|}\Big) = O_K\Big(\sum_{P\in\mathcal{S}} \frac{\log^K |P|}{|P|}\Big).$$

Proof. The proof is identical to that of [20, Lemma E.1]. ∎

In our proof of the Goldston–Yıldırım estimates (Propositions 7–9) in the next sections, we will be concerned with Euler products in several variables, i.e. of the form $\prod_P (1-\sum_{j=1}^n c_{P,j}/|P|^{1+s_j})$, as $\Re s_j > 0$ and $s_j \to 0$ uniformly. The following lemma gives an asymptotic formula for such Euler products.

LEMMA 7. *Let P range over monic irreducible polynomials in $\mathbb{F}_q[t]$. For every P let $c_{P,1},\dots,c_{P,n}$ be real numbers such that $|c_{P,j}| \le 1$ and $c_{P,j} = c_j$ for P outside a finite set $\mathcal{S}$. Let $s_1,\dots,s_n \in \mathbb{C}$ be such that $\Re s_j > 0$ and $s_j = o(1)$ uniformly. Then*

$$\prod_P \Big(1-\sum_{j=1}^n \frac{c_{P,j}}{|P|^{1+s_j}}\Big) = G(1+o_n(1)) \prod_{P\in\mathcal{S}} \Big(1+O_n\Big(\frac{1}{|P|}\Big)\Big) \prod_{j=1}^n (s_j \log q)^{c_j}$$

where $G = \prod_P (1-\sum_{j=1}^n c_{P,j}/|P|)(1-1/|P|)^{-(c_1+\cdots+c_n)}$.

Note that the O_n and o_n depend only on n and the rate $s_1,\dots,s_n \to 0$ and not on the exceptional set $\mathcal{S}$.

Proof. The proof goes along the lines of [18, Lemma 1.2]. ∎

7. A pseudorandom measure that majorizes the irreducible polynomials. In this section we prove Theorem 5 by constructing a pseudorandom measure ν. The proof of its pseudorandomness is however deferred to the next section. For simplicity we assume that $\mathcal{A} = \mathcal{P}$; the construction in this particular case extends to the general case with ease. Throughout this whole section and the next two, polynomials denoted by the letter d (such as d, d' or d_i) will stand for monic polynomials.

Let us fix once and for all:

- $R = \alpha N$, where α is a small constant depending only on k.
- $w = w(N)$, a function tending sufficiently slowly to infinity. We may take $w(N) \ll \log N$.
- $W = \prod_{\deg(P)<\omega} P$. We have $W(t) = t^{q^w} - t$, so that ([6]) $\deg(W) \ll N$. We will see that eventually we can take w to be a large number, hence W to be a large degree polynomial.
- $\chi : \mathbb{R} \to \mathbb{R}$ a smooth function ([7]) supported on $[-1, 1]$ such that $\chi(0) > 0$ and $\int_0^\infty (\chi'(x))^2\, dx = 1$.
- $\Lambda_R(f) = \sum_{d|f,\, \deg(d)<R} \mu(d)\chi(\deg(d)/R)$, the Goldston–Yıldırım divisor sum.
- $\nu(f) = R(\Phi(W)/|W|)\Lambda_R(Wf + 1)^2$. ([8])

Proof of Theorem 5 under the assumption that ν is pseudorandom. Notice that if f is irreducible and $\deg(f) \geq R$ then $\Lambda_R(f) = \chi(0) = 1$. For $f \in \mathbb{G}_N$ let

$$\phi(f) = \begin{cases} \chi(0)^2 \dfrac{\Phi(W)}{|W|} R & \text{if } Wf + b \text{ is irreducible and } \deg(Wf + b) \geq R, \\ 0 & \text{otherwise.} \end{cases}$$

Then clearly $0 \leq \phi \leq \nu$ and $\|\phi\|_\infty \ll N$. Using the prime number theorem in $\mathbb{F}_q[t]$ (Proposition 6), with the observation that $N + \deg(W)$ increases at most linearly in N, we can verify that $\mathbb{E}_{\mathbb{G}_N} \phi(f)$ is bounded from below. ∎

The only thing missing from the conclusions of Theorem 5 is to check that ν is indeed a pseudorandom measure, i.e. it satisfies the linear forms condition and the correlation condition. This will be done in the next section. In order to do so, we will need estimates on sums of the form $\sum \Lambda_R(\psi_1) \cdots \Lambda_R(\psi_n)$ where the ψ_i are linear forms. The following proposition shows how to deal with sums of this kind.

Proposition 7 (Goldston–Yıldırım estimates). *Let $J_1, \ldots, J_n \in \mathbb{F}_q[t]$, not necessarily distinct. Let r be the number of distinct elements in $\{J_1, \ldots, J_n\}$. Also, for every monic, irreducible $P \in \mathbb{F}_q[t]$, let α_P be the number of distinct residue classes modulo P occupied by $J_1, \ldots, J_n$. Put $\Delta = \Delta(J_1, \ldots, J_n) = \prod (J_i - J_{i'})$, where the product is taken over all couples*

([6]) The introduction of W, alluded to earlier as the W-trick, is meant to absorb small irreducible polynomials arising in the linear forms condition. Except for this technical reason, for the most part we can go through the arguments pretending that $W = 1$ without losing the general idea.

([7]) Goldston–Yıldırım used a truncated sum corresponding to $\chi(x) = \max(1 - |x|, 0)$. As observed by Tao [18], the use of a smooth function allows us to perform Fourier analysis.

([8]) In the general case, $\nu(f) = R\frac{\Phi(W)}{|W|}\Lambda_R(Wf + b)^2$ for some b appropriately chosen by the pigeonhole principle.

$(J_i, J_{i'})$ such that $J_i \neq J_{i'}$. Then as $N \to \infty$,

$$\sum_{f \in \mathbb{G}_N} \Lambda_R(f+J_1)\cdots\Lambda_R(f+J_n) = CGH(1+o_n(1))q^N\left(\frac{\log q}{R}\right)^r \tag{1}$$

where C is a computable constant (not depending on $N, J_1, \dots, J_n$ but only on the multiplicities of the J_i and χ), G is the "arithmetical factor" $\prod_P(1-1/|P|)^{-r}(1-\alpha_P/|P|)$, and $H = \prod_{P|\Delta}(1+O_n(1/|P|))$.

Remarks. This proposition illustrates how the Goldston–Yıldırım method works. We will not apply this proposition directly (since we will be incorporating the W-trick), but rather its variants (Propositions 8 and 9), for which only minor modifications are needed.

Proof of Proposition 7. The proof goes along the lines of [18, Section 2], so we will outline the ideas briefly. Writing out the definition of Λ_R, we see that the left hand side of (1) is

$$\sum_{d_1,\dots,d_n \in \mathbb{G}_N} \left(\prod_{i=1}^n \mu(d_i)\chi\left(\frac{\deg(d_i)}{R}\right)\right) \sum_{f \in \mathbb{G}_N} 1_{d_i | f+J_i \,\forall i=1,\dots,n}. \tag{2}$$

Note that since χ is supported on $[-1,1]$, the summation over $d_1, \dots, d_n \in \mathbb{G}_N$ is the same as the summation over $d_1, \dots, d_n \in \mathbb{G}_R$. Also, because of the appearance of the function μ, only squarefree d_i are involved. Suppose R is sufficiently small compared to N, say $nR < N$. Then (2) is equal to

$$q^N \sum_{d_1,\dots,d_n} \frac{g(d_1,\dots,d_n)}{|[d_1,\dots,d_n]|} \prod_{i=1}^k \mu(d_i)\chi\left(\frac{\deg(d_i)}{R}\right) \tag{3}$$

where $g(d_1, \dots, d_n)$ is the number of solutions in $\mathbb{G}_{\deg([d_1,\dots,d_n])}$ of the system of congruences $f + J_i \equiv 0 \pmod{d_i}$ for every $i = 1, \dots, k$.

Write ([9])

$$\chi(x) = \int_{-\infty}^{\infty} q^{-(1+it)x}\psi(t)\,dt,$$

where ψ is rapidly decreasing, i.e. $\psi(t) = O_A((1+|t|)^{-A})$ for every A. An easy calculation shows that for every $m = 1, \dots, n$, we have

$$\chi\left(\frac{\deg(d_m)}{R}\right) = \int_{-\sqrt{R}}^{\sqrt{R}} |d_m|^{-(1+it_m)/R}\psi(t_m)\,dt_m + O_A(|d_m|^{-1/R}R^{-A}).$$

([9]) The appearance of q here is for mere aesthetic reasons.

We split the expression (3) into a main term of

$$q^N \sum_{d_1,\dots,d_n} \frac{g(d_1,\dots,d_n)}{|[d_1,\dots,d_n]|} \prod_{m=1}^{n} \mu(d_m) \int_{-\sqrt{R}}^{\sqrt{R}} |d_m|^{-(1+it_m)/R} \psi(t_m)\, dt_m$$

plus an error term, which is

$$\begin{aligned}(4)\qquad & q^N \sum_{d_1,\dots,d_n} \frac{g(d_1,\dots,d_n)}{|[d_1,\dots,d_n]|} O_A(R^{-A}|d_1\cdots d_n|^{-1/R}) \\ &\ll_A q^N R^{-A} \sum_{d_1,\dots,d_n} \frac{|d_1\cdots d_n|^{-1/R}}{|[d_1,\dots,d_n]|}.\end{aligned}$$

It is easy to see that when $A \ge n$, the error term (4) is $o(q^N)$. Consequently, the sum in (3) converges absolutely.

Therefore, it suffices to show that

$$\begin{aligned}&\sum_{d_1,\dots,d_n} \frac{g(d_1,\dots,d_n)}{[d_1,\dots,d_n]} \prod_{m=1}^{k} \mu(d_m) \prod_{m=1}^{k} \int_{-\sqrt{R}}^{\sqrt{R}} |d_m|^{-(1+it_m)/R} \psi(t_m)\, dt_m \\ &\qquad = CGH(1+o_n(1)) \Big(\frac{\log q}{R}\Big)^r.\end{aligned}$$

Note that all of our expressions are in terms of R, and we have eliminated the role of N. We now switch the orders of the sums and the integral (which is legitimate since the sum is absolutely convergent) and the expression under the integration.

Lemma 8. *For every $I \subset \{1,\dots,n\}$, $I \ne \emptyset$, let $c_I = 1$ if $J_i = J_{i'}$ for every $i, i' \in I$, and 0 otherwise. Then for every $t_1,\dots,t_n \in [-\sqrt{R},\sqrt{R}]$ we have*

$$\begin{aligned}(5)\qquad & \sum_{d_1,\dots,d_n} \frac{g(d_1,\dots,d_n)}{[d_1,\dots,d_n]} \prod_{m=1}^{n} \mu(d_m) |d_m|^{-(1+it_m)/R} \\ &= GH(1+o_{R\to\infty}(1)) \Big(\frac{\log q}{R}\Big)^r \prod_{I\subset\{1,\dots,n\},\, I\ne\emptyset} \Big(\sum_{m\in I}(1+it_m)\Big)^{(-1)^{|I|+1}c_I}.\end{aligned}$$

Proof. Recall that in the expression on the left hand side of (5), only square-free $d_1,\dots,d_n$ are involved. We note that $g(d_1,\dots,d_n)$ only takes two values, 0 or 1. More precisely, by the Chinese remainder theorem, $g(d_1,\dots,d_n) = \prod_P c_{P,\{i:P|d_i\}}$, where the $c_{P,I}$ are "local factors" defined by $c_{P,I} = \sharp\{\deg(f) < \deg(P) : P \,|\, f + J_i \text{ for every } i \in I\}$ for every $I \subset$

$\{1,\dots,n\}$, $I \neq \emptyset$. We have the following explicit formula:

$$c_{P,I} = \begin{cases} 1 & \text{if } J_i \equiv J_{i'} \pmod P \text{ for every } i, i' \in I, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathcal{S}$ be the set of all irreducible divisors of Δ. Then for P outside of $\mathcal{S}$, we have $c_{P,I} = c_I$. The left hand side of (5) factors as

$$\prod_P \Bigg(1 - \sum_{I \subset \{1,\dots,n\},\, I \neq \emptyset} (-1)^{|I|+1} \frac{c_{P,I}}{|P|^{1+\sum_{m\in I}(1+it_m)/R}}\Bigg).$$

An application of Lemma 7 gives the desired asymptotic form for this product. ∎

By integrating over all $t_1,\dots,t_n \in [-\sqrt{R},\sqrt{R}]$, the estimate (1) easily follows, with

$$C = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{I \subset \{1,\dots,k\},\, I\neq\emptyset} \Big(\sum_{m\in I}(1+it_m)\Big)^{(-1)^{|I|+1}c_I} \prod_{m=1}^{k} \psi(t_m)\, dt_1 \cdots dt_k.$$

This expression can be simplified a little. Let $a_1,\dots,a_r$ be the multiplicities of $J_1,\dots,J_n$. Then $C = \prod_s C_{a_s}$, where

(6)
$$C_a = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{I \subset \{1,\dots,a\},\, I\neq\emptyset} \Big(\sum_{m\in I}(1+it_m)\Big)^{(-1)^{|I|+1}} \prod_{m=1}^{a} \psi(t_m)\, dt_1 \cdots dt_a.$$

In our applications we will be able to compute these constants explicitly in terms of χ. ∎

8. The correlation condition and the linear forms condition

8.1. The correlation condition. As mentioned before, we will need a variant of Proposition 7:

PROPOSITION 8. *Let $J_1,\dots,J_n \in \mathbb{F}_q[t]$, not necessarily distinct. Let $\Delta = \Delta(J_1,\dots,J_n)$ be as in Proposition 7. Then we have the asymptotic formula*

$$\sum_{f\in\mathbb{G}_N} \Lambda_R(W(f+J_1)+1)\cdots\Lambda_R(W(f+J_n)+1) = CGH(1+o_n(1))q^N\left(\frac{\log q}{R}\right)^r \tag{7}$$

where

$$G = \prod_{\deg(P)<w} \left(1-\frac{1}{|P|}\right)^{-r} \prod_{\deg(P)\geq w} \left(1-\frac{\alpha_P}{|P|}\right)\left(1-\frac{1}{|P|}\right)^{-r},$$

$$H = \prod_{P|\Delta} \left(1 + O_n\left(\frac{1}{|P|}\right)\right),$$

and C is the same constant as in Proposition 7.

Proof. The proof follows the lines of that of Proposition 7. The only difference is that we have a different formula for local factors $c_{P,I} = \sharp\{\deg(f) < \deg(P) : P \,|\, W(f + J_i) + 1 \ \forall i \in I\}$: for any $I \subset \{1, \dots, n\}$, $I \neq \emptyset$, we have

$$c_{P,I} = \begin{cases} 1 & \text{if } \deg(P) \geq w \text{ and } J_i \equiv J_{i'} \pmod P \text{ for every } i, i' \in I, \\ 0 & \text{otherwise.} \end{cases}$$

By incorporating this change into the proof, we will find the desired expression for G. ∎

Proof of the correlation condition. We are interested in expressions of the form

$$\mathbb{E}(\nu(f + h_1) \cdots \nu(f + h_l) \,|\, f \in \mathbb{G}_N)$$

where $h_1, \dots, h_l \in \mathbb{G}_N$ and the number l of forms is bounded by l_0 which depends only on k. Recall that our goal is to find a function τ on $\mathbb{G}_N$ such that

$$\mathbb{E}(\nu(f + h_1) \cdots \nu(f + h_l) \,|\, f \in \mathbb{G}_N) \leq \sum_{1 \leq i \leq j \leq l} \tau(h_i - h_j). \tag{8}$$

Moreover, for every $1 \leq p < \infty$,

$$\mathbb{E}(\tau(f)^p) = O_p(1). \tag{9}$$

In the event where two of the h_i are equal, we bound the left hand side of (8) by $\|\nu\|_\infty^l = q^{O(N/\log N)}$, thanks to Lemma 5. If we choose $\tau(0) = q^{O(N/\log N)}$ then clearly the inequality (8) is satisfied. Moreover, since $q^{O(N/\log N)} = O_\epsilon(q^{N\epsilon})$ for every $\epsilon > 0$, the addition of $q^{O(N/\log N)}$ to $\tau(0)$ does not affect the boundedness of $\mathbb{E}(\tau^p)$ for every $p > 1$. Therefore, we have to find a function $\tau \in L^p$ for every $p > 1$ so that the inequality (8) is satisfied when all the h_i are distinct. From the definition of ν, we have

$$\begin{aligned}&\mathbb{E}(\nu(f + h_1) \cdots \nu(f + h_l) \,|\, f \in \mathbb{G}_N) \\ &\quad = R^l \left(\frac{\Phi(W)}{|W|}\right)^l \mathbb{E}(\Lambda_R(W(f + h_1) + 1)^2 \cdots \Lambda_R(W(f + h_l) + 1)^2 \,|\, f \in \mathbb{G}_N).\end{aligned}$$

Thanks to Proposition 8, we know that

$$\begin{aligned}&\mathbb{E}(\Lambda_R(W(f + h_1) + 1)^2 \cdots \Lambda_R(W(f + h_l) + 1)^2 \,|\, f \in \mathbb{G}_N) \\ &\qquad\qquad = CGH(1 + o(1))\left(\frac{\log q}{R}\right)^l,\end{aligned}$$

where

$$G = \prod_{\deg(P)<w}\Big(1-\frac{1}{|P|}\Big)^{-l}\prod_{\deg(P)\ge w}\Big(1-\frac{l}{|P|}\Big)\Big(1-\frac{1}{|P|}\Big)^{-l}$$
$$= \Big(\frac{\Phi(W)}{W}\Big)^{-l}\prod_{\deg(P)\ge w}\Big(1-\frac{l}{|P|}\Big)\Big(1-\frac{1}{|P|}\Big)^{-l} = \Big(\frac{\Phi(W)}{W}\Big)^{-l}(1+o(1)).$$

Let us compute C explicitly. In this case, each h_i has multiplicity 2, hence (6) gives

$$C = \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{(1+it_1)(1+it_2)}{2+it_1+it_2}\psi(t_1)\psi(t_2)\,dt_1\,dt_2\Big)^l$$
$$= \Big(\int_0^{\infty}\log q\Big(\int_{-\infty}^{\infty}q^{-(1+it)x}\psi(t)\,dt\Big)^2\Big)^l = \Big(\int_0^{\infty}\log q\Big(\frac{\chi'(x)}{\log q}\Big)^2dx\Big)^l$$
$$= (\log q)^{-l}.$$

Therefore, $\mathbb{E}(\nu(f+h_1)\cdots\nu(f+h_l)\,|\,f\in\mathbb{G}_N) = (1+o(1))H$. Recall that $H=\prod_{P|\Delta}(1+O_l(1/|P|))$, where $\Delta=\prod_{1\le i\le j\le m}(h_i-h_j)$. Let us bound $(1+o(1))H$ by $\exp(M\sum_{P|\Delta}1/|P|)$, where M is a constant depending only on l_0, hence on k.

For $f\neq 0$, put $\tau(f)=\exp(K\sum_{P|f}1/|P|)$ for K sufficiently large depending on M. Then clearly the inequality (8) is satisfied. The only thing left is to verify (9). By Lemma 6 we have

$$\mathbb{E}(\tau(f)^p\,|\,\mathbb{G}_N) = \mathbb{E}\Big(\exp\Big(pK\sum_{P|f}\frac{1}{|P|}\Big)\;\Big|\;f\in\mathbb{G}_N\Big)$$
$$\ll_{K,p}\mathbb{E}\Big(\sum_{P|f}\frac{\log^{Kp}|P|}{|P|}\;\Big|\;f\in\mathbb{G}_N\Big).$$

For every P, the number of $f\in\mathbb{G}_N$ that are divisible by P is at most $q^N/|P|$. Thus

$$\mathbb{E}\Big(\sum_{P|f}\frac{\log^{Kp}|P|}{|P|}\;\Big|\;f\in\mathbb{G}_N\Big)\le\frac{1}{q^N}\sum_P\frac{q^N}{|P|}\frac{\log^{Kp}|P|}{|P|}$$
$$=\sum_P\frac{\log^{Kp}|P|}{|P|^2}=O_{K,p}(1)$$

as required. ∎

8.2. The linear forms condition. For the linear forms condition, the following variant of Proposition 7 is needed:

PROPOSITION 9. *Let $\psi_1, \dots, \psi_m$ be nonzero linear forms $(\mathbb{G}_N)^n \to \mathbb{F}_q[t]$, not necessarily distinct, of the form $\psi_i(\mathbf{f}) = \sum_{j=1}^n L_{ij} f_j + b_i$ for every $\mathbf{f} = (f_1, \dots, f_n) \in (\mathbb{G}_N)^n$, where*

- *$m \le m_0$ and $n \le n_0$, where m_0, n_0 are constants depending on k,*
- *$L_{ij}, b_i \in \mathbb{F}_q[t]$ and $\deg(L_{ij}) < w/2$ for every i, j,*
- *for any $i, i' \in \{1, \dots, m\}$, the vectors $(L_{ij})_{j=1}^n$ and $(L_{i'j})_{j=1}^n$ are either nonproportional (over $\mathbb{F}_q(t)$), or identical and $\psi_i = \psi_{i'}$.*

Let r be the number of distinct forms among $\psi_1, \dots, \psi_m$. Then (assuming that $R = \alpha N$ and α is sufficiently small depending only on m_0, n_0) we have the following asymptotic formula as $N \to \infty$:

$$\sum_{\mathbf{f} \in (\mathbb{G}_N)^n} \Lambda_R(W\psi_1(f)+1) \cdots \Lambda_R(W\psi_m(f)+1) = C(1+o(1)) \left(\frac{\Phi(W)}{W}\right)^{-r} \left(\frac{\log q}{R}\right)^r q^{Nn} \tag{10}$$

where C is a computable constant (depending only on χ and the multiplicities of the ψ_i).

Proof. We skip the proof since it is similar to that of Proposition 7. Note that similarly to the constant C in Proposition 7, C factors as $\prod_{s=1}^r C_{a_s}$, where $a_1, \dots, a_r$ are multiplicities of the ψ_i, and

$$C_a = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{I \subset \{1,\dots,a\},\, I \neq \emptyset} \Big(\sum_{j \in I} (1+it_j)\Big)^{(-1)^{|I|+1}} \prod_{j=1}^a \psi(t_j)\, dt_1 \cdots dt_a. \ \blacksquare$$

Proof of the linear forms condition. We are interested in expressions of the form

$$\mathbb{E}(\nu(\psi_1(\mathbf{f})) \cdots \nu(\psi_m(\mathbf{f})) \mid \mathbf{f} \in (\mathbb{F}_{q^N})^n)$$

where $\psi_i(\mathbf{f}) = \sum_{j=1}^n L_{ij} f_j + b_i$ is a linear form in n variables, $m \le m_0$, $n \le n_0$, no two homogeneous parts are proportional, and the coefficients L_{ij} are in the set $\{P/Q : \deg(P), \deg(Q) < k\}$. Recall that we want to bound these expressions by $1 + o(1)$. By the definition of ν, the above expression is equal to

$$\left(\frac{\Phi(W)}{|W|}\right)^m R^m \mathbb{E}(\Lambda_R(W\psi_1(\mathbf{f})+1)^2 \cdots \Lambda_R(W\psi_m(\mathbf{f})+1)^2 \mid \mathbf{f} \in (\mathbb{F}_{q^N})^n). \tag{11}$$

Our first reduction is to replace the assumption that all the coefficients L_{ij} are in $\{f/g \mid f, g \in \mathbb{G}_k\}$ by $L_{ij} \in \mathbb{G}_M$ for some sufficiently large M depending on k. Indeed, via a change of variables $\mathbf{f} \mapsto (\prod_{h \in \mathbb{G}_k,\, h \text{ monic}} h)\mathbf{f}$, the ψ_i become linear forms with coefficients in $\mathbb{F}_q[t]$ and of degrees still bounded,

$$\deg\Big(\prod_{h \in \mathbb{G}_k,\, h \text{ monic}} h\Big) + k = \sum_{d=1}^{k-1} dq^d + k = M.$$

We are tempted to apply Proposition 9 right away. However, a priori the ψ_i are linear forms from $(\mathbb{F}_{q^N})^n$ to $\mathbb{F}_{q^N}$, which are different from the linear forms Ψ_i from $(\mathbb{F}_q[t])^n$ to $\mathbb{F}_q[t]$ given by the same formula $\Psi_i(\mathbf{f}) = \sum_{j=1}^n L_{ij} f_j + b_i$ for every $\mathbf{f} = (f_1, \dots, f_n) \in (\mathbb{F}_q[t])^n$. More precisely, $\psi_i(\mathbf{f})$ is the residue of $\Psi_i(\mathbf{f})$ upon division by f_N, the irreducible polynomial underlying $\mathbb{F}_{q^N}$. To remedy this, let us divide $(\mathbb{F}_q[t])^n$ into q^{nM} "boxes" such that for $\mathbf{f}, \mathbf{f}'$ in the same box B, we have $\max_i |f_i - f_i'| < q^{N-M}$ (in other words, each box is a product of cylinder sets of radius q^M in $\mathbb{F}_q[t]$). Then for $\mathbf{f}, \mathbf{f}'$ in the same box, we have $\Psi_i(\mathbf{f}) - \Psi_i(\mathbf{f}') < q^{N-M+M} = q^N$. This means that the residues of $\Psi_i(\mathbf{f})$ and $\Psi_i(\mathbf{f})$ upon division by f_N are the same. In other words, for any box B, we have a formula

$$\psi_i(\mathbf{f}) = \Psi_{i,B}(\mathbf{f}) = \sum_{j=1}^n L_{ij} f_j + b_{i,B}$$

for every $\mathbf{f} \in B$, and $b_{i,B}$ depends only on the box B. We now rewrite the expression (11) as

$$\left(\frac{\Phi(W)}{|W|}\right)^m R^m \frac{1}{q^{nM}} \sum_B \mathbb{E}(\Lambda_R(W\psi_{1,B}(\mathbf{f})+1)^2 \cdots \Lambda_R(W\psi_{m,B}(\mathbf{f})+1)^2 \,|\, \mathbf{f} \in B).$$

Note that for N sufficiently large, we have $\deg(L_{ij}) < M < w/2$. For each box B, Proposition 9 tells us that

$$\mathbb{E}(\Lambda_R(W\psi_{1,B}(\mathbf{f}) + 1)^2 \cdots \Lambda_R(W\psi_{m,B}(\mathbf{f}) + 1)^2 \,|\, \mathbf{f} \in B) = C(1 + o(1))\left(\frac{\Phi(W)}{W}\right)^{-m} \left(\frac{\log q}{R}\right)^m.$$

Similarly to the calculations in the proof of the correlation condition, we see that $C = (\log q)^m$. Summing this up over all the q^M boxes yields the linear forms condition. ∎

Acknowledgements. I am grateful to my advisor Terence Tao for suggesting this project, frequent consultation and assistance throughout the preparation of this paper, and giving helpful comments on the manuscript. I would also like to thank the referee for suggesting changes to an early version of this paper.

References

[1] V. Bergelson, A. Leibman and R. McCutcheon, *Polynomial Szemerédi theorem for countable modules over integral domains and finite fields*, J. Anal. Math. 95 (2005), 243–296.

[2] H. Furstenberg and Y. Katznelson, *A density version of the Hales–Jewett theorem*, ibid. 57 (1991), 64–119.

[3] D. Goldston and C. Y. Yıldırım, *Small gaps between primes, I*, arXiv:math/0504336v1.
[4] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. 8 (1998), 529–551.
[5] —, *A new proof of Szemerédi's theorem*, ibid. 11 (2001), 465–588.
[6] —, *Decompositions, approximate structure, transference, and the Hahn–Banach theorem*, Bull. London Math. Soc. 42 (2010), 573–606.
[7] B. Green, *Roth's Theorem in the primes*, Ann. of Math. 161 (2005), 1609–1636.
[8] B. Green and T. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinburgh Math. Soc. 51 (2008), 73–153.
[9] —, —, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. 167 (2008), 481–547.
[10] B. Host and B. Kra, *Non-conventional ergodic averages and nilmanifolds*, ibid. 161 (2005), 397–488.
[11] T. H. Lê, *Green–Tao theorem in function fields*, unabridged version, arXiv:0908.2642v2.
[12] P. Pollack, *Simultaneous prime specializations of polynomials over finite fields*, Proc. London Math. Soc. (3) 97 (2008), 545–567.
[13] D. H. J. Polymath, *A new proof of the density Hales–Jewett theorem*, arXiv:0910.3926.
[14] O. Reingold, L. Trevisan, M. Tulsiani and S. Vadhan, *Dense subsets of pseudorandom sets*, Technical Report TR08-045, ECCC, 2008.
[15] —, —, —, —, *New proofs of the Green–Tao–Ziegler dense model theorem: An exposition*, arXiv:0806.0381v1.
[16] M. Rosen, *Number Theory in Function Fields*, Grad. Texts in Math. 210, Springer, New York, 2002.
[17] T. Tao, *Structure and Randomness: Pages from Year One of a Mathematical Blog*, Amer. Math. Soc., 2008.
[18] —, *A remark on Goldston–Yıldırım correlation estimates*, preprint.
[19] —, *The Gaussian primes contain arbitrarily shaped constellations*, J. Anal. Math. 99 (2006), 109–176.
[20] T. Tao and T. Ziegler, *The primes contain arbitrarily long polynomial progressions*, Acta Math. 201 (2008), 213–305.
[21] P. Varnavides, *On certain sets of positive density*, J. London Math. Soc. 34 (1959), 358–360.

Thái Hoàng Lê
Department of Mathematics
UCLA
Los Angeles, CA 90095-1596, U.S.A.
E-mail: leth@math.ucla.edu

Received on 27.4.2009
and in revised form on 17.6.2010 (6010)