

This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

doi:10.1017/etds.2019.27

# Nilsystems and ergodic averages along primes

TANJA EISNER

*Institute of Mathematics, University of Leipzig, P.O. Box 100 920, 04009 Leipzig, Germany*

*(e-mail: eisner@math.uni-leipzig.de)*

*(Received 15 July 2018 and accepted in revised form 1 March 2019)*

Dedicated to Vitaly Bergelson on the occasion of his 65th birthday.

*Abstract.* A celebrated result by Bourgain and Wierdl states that ergodic averages along primes converge almost everywhere for  $L^p$ -functions,  $p > 1$ , with a polynomial version by Wierdl and Nair. Using an anti-correlation result for the von Mangoldt function due to Green and Tao, we observe everywhere convergence of such averages for nilsystems and continuous functions.

**Key words:** ergodic averages along primes, nilsystems, everywhere convergence

**2010 Mathematics Subject Classification:** 37A30, 37A45 (Primary); 11B30, 37B05 (Secondary)

## 1. Introduction

Ergodic theorems, originally motivated by physics, have found applications in and connections to many areas of mathematics. A prominent example is the result on almost everywhere convergence of ergodic averages along primes by Bourgain [6, 8] (for  $p > (1 + \sqrt{3})/2$ ) and subsequently Wierdl [37] (for all  $p > 1$ ).

**THEOREM 1.1.** *Let  $(X, \mu, T)$  be a measure-preserving system,  $p > 1$  and  $f \in L^p(X, \mu)$ . Then the ergodic averages along primes*

$$\frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} T^p f \tag{1}$$

*converge almost everywhere.*

The proof is based on the Carleson transference principle to the discrete model ( $\mathbb{Z}$ , Shift), the Hardy–Littlewood circle method and estimates of prime number exponential sums. An analogous result for polynomials instead of primes was proved by Bourgain [7, 8], see also Thouvenot [36], with variation estimates by Krause [26] showing

that the averages converge rapidly. For analogous estimates for averages (1), see Zorin-Kranich [42]. Moreover, Theorem 1.1 has been generalized to polynomials of primes by Wierdl [38] and Nair [31, 32].

Since the proof of Bourgain and Wierdl does not give any information on the set of points where the convergence holds, the following natural question arises.

*Question 1.2.* For which systems and functions do the ergodic averages along primes (1) converge *everywhere*?

We give a partial answer to this question and show that ergodic averages along polynomials of primes converge everywhere for all nilsystems and all continuous functions. For the definitions of a nilsystem and a polynomial sequence, see Section 2.

**THEOREM 1.3.** *Let  $G/\Gamma$  be a nilmanifold,  $g : \mathbb{N} \rightarrow G$  be a polynomial sequence and  $F \in C(G/\Gamma)$ . Then the averages*

$$\frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} F(g(p)x)$$

*converge for every  $x \in G/\Gamma$ . Moreover, if  $G$  is connected and simply connected,  $g(n) = g^n$  and the system  $(G/\Gamma, \mu, g)$  is ergodic, then the limit equals  $\int_{G/\Gamma} F d\mu$ .*

The key to this result is the powerful theory developed by Green and Tao [18–20], partially together with Ziegler [21], in their study of arithmetic progressions and linear equations in the primes, in particular the asymptotic orthogonality of the modified von Mangoldt function to nilsequences; see Theorem 2.1 below.

Note that nilsystems and nilsequences have been playing a fundamental role in the study of other kinds of ergodic averages, namely the norm convergence of (linear and polynomial) multiple ergodic averages, motivated by Furstenberg’s ergodic-theoretic proof [16] of Szemerédi’s theorem [35] on the existence of arithmetic progressions in large sets of integers. Here is a list of relevant works: Conze and Lesigne [10], Furstenberg and Weiss [17], Host and Kra [23], Lesigne [30], Ziegler [40], Host and Kra [24], Ziegler [41], Bergelson, Host and Kra [2], Bergelson, Leibman and Lesigne [4], Bergelson and Leibman [3], Leibman [28, 29], Frantzikinakis [12], Host and Kra [25], Chu [9], Eisner and Zorin-Kranich [11] and Zorin-Kranich [43]. For other applications of the Green–Tao–Ziegler theory to ergodic theorems, see, e.g., Frantzikinakis, Host and Kra [14, 15], Wooley and Ziegler [39], Bergelson, Leibman and Ziegler [5] and Frantzikinakis and Host [13].

Our argument is similar to (but simpler than) the one in Wooley and Ziegler [39] in the context of the norm convergence of multiple polynomial ergodic averages along primes.

## 2. Preliminaries and the $W$ -trick

Let  $G$  be an  $s$ -step Lie group and  $\Gamma$  be a discrete cocompact subgroup of  $G$ . The homogeneous space  $G/\Gamma$  together with the Haar measure  $\mu$  is called an  $s$ -step nilmanifold. For every  $g \in G$ , the left multiplication by  $g$  is an invertible  $\mu$ -preserving transformation on  $G/\Gamma$ , and the triple  $(G/\Gamma, \mu, g)$  is called a nilsystem. Nilsystems enjoy remarkable algebraic and ergodic properties making them an important class of systems in the classical ergodic theory; see Auslander, Green and Hahn [1], Green [22], Parry [33, 34] and

Leibman [27]. For example, single and multiple ergodic averages converge everywhere for such systems and continuous functions.

For a continuous function  $F$  on  $G/\Gamma$  and  $x \in G/\Gamma$ , the sequence  $(F(g^n x))_{n \in \mathbb{N}}$  is called a (*basic linear*) *nilsequence* as introduced by Bergelson, Host and Kra [2]. A nilsequence in their definition is a uniform limit of basic nilsequences (being allowed to come from different systems and functions). Note that the property of Cesàro convergence along primes is preserved by uniform limits, so Theorem 1.3 implies in particular that every nilsequence is Cesàro convergent along primes.

Rather than linear sequences  $(g^n)$ , following Leibman [27], Green and Tao [19] and Green, Tao and Ziegler [21], we will consider polynomial sequences  $(g(n))$ , where  $g : \mathbb{N} \rightarrow G$  is called a *polynomial sequence* if it is of the form  $g(n) = g_1^{p_1(n)} \cdot \dots \cdot g_m^{p_m(n)}$  for some  $m \in \mathbb{N}$ ,  $g_1, \dots, g_m \in G$  and some integer polynomials  $p_1, \dots, p_m$ . For an abstract equivalent definition, see [19]. A sequence of the form  $(F(g(n)x))$  for a continuous function  $F$  on  $G/\Gamma$  is called a *polynomial nilsequence*. Although this notion seems to be more general than the one of linear basic nilsequences, it is not; see the references at the beginning of the proof of Theorem 1.3 in the following section.

Note that a nilsequence does not determine  $G, \Gamma, F$  etc uniquely, giving room for reductions. For example, we can assume without loss of generality that  $x = \text{id}_G \Gamma$ . Moreover, denoting by  $G^0$  the connected component of the identity in  $G$ , since we are only interested in the orbit of  $x$  under  $g(n)$ , we can assume without loss of generality that  $G = \langle G^0, g_1, \dots, g_m \rangle$ .

We use the notation  $o_{a,b}(1)$  and  $O_{a,b}(1)$  to denote a function which converges to zero or is bounded, respectively, for fixed parameters  $a, b$  uniformly in all other parameters.

We now introduce the  $W$ -trick as in Green and Tao [18]. Consider

$$\Lambda'(n) := \begin{cases} \log n & \text{if } n \in \mathbb{P}, \\ 0 & \text{otherwise.} \end{cases}$$

For  $\omega \in \mathbb{N}$ , define

$$W = W_\omega := \prod_{p \in \mathbb{P}, p \leq \omega} p$$

and for  $r < W$  coprime to  $W$  define the modified  $\Lambda'$ -function by

$$\Lambda'_{r,\omega}(n) := \frac{\phi(W)}{W} \Lambda'(Wn + r), \quad n \in \mathbb{N},$$

where  $\phi$  denotes the Euler totient function.

The key to our result is the following anti-correlation property of  $\Lambda'_{r,\omega}$  with nilsequences due to Green and Tao [18] conditional to the ‘Möbius and nilsequences conjecture’ proven by them later in [20]. Here,  $\omega : \mathbb{N} \rightarrow \mathbb{N}$  is an arbitrary function with  $\lim_{N \rightarrow \infty} \omega(N) = \infty$  satisfying  $\omega(N) \leq (1/2) \log \log N$  for all large  $N \in \mathbb{N}$ . Note that the corresponding function  $W : \mathbb{N} \rightarrow \mathbb{N}$  is then  $O(\log^{1/2} N)$ .

**THEOREM 2.1.** (Green–Tao [18, Proposition 10.2]) *Let  $\omega(\cdot)$  and  $W(\cdot)$  be as above,  $G/\Gamma$  be an  $s$ -step nilmanifold with a smooth metric,  $G$  being connected and simply connected*

and let  $(F(g^n x))$  be a bounded nilsequence on  $G/\Gamma$  with Lipschitz constant  $M$ . Then

$$\max_{r < W(N), (r, W(N))=1} \left| \frac{1}{N} \sum_{n=1}^N (\Lambda'_{r, \omega(N)}(n) - 1) F(g^n x) \right| = o_{M, G/\Gamma, s}(1)$$

as  $N \rightarrow \infty$ .

An immediate corollary is the following, cf. Theorem 2.2 (and the discussion afterwards) in Frantzikinakis, Host and Kra [15]. Here,  $\omega$  and  $W = W_\omega$  are again numbers, not functions.

**COROLLARY 2.2.** *Let  $G/\Gamma$  be an  $s$ -step nilmanifold with a smooth metric,  $G$  being connected and simply connected and let  $(F(g^n x))$  be a bounded nilsequence on  $G/\Gamma$  with Lipschitz constant  $M$ . Then*

$$\lim_{\omega \rightarrow \infty} \limsup_{N \rightarrow \infty} \max_{r < W, (r, W)=1} \left| \frac{1}{N} \sum_{n=1}^N (\Lambda'_{r, \omega}(n) - 1) F(g^n x) \right| = 0,$$

where the convergence is uniform in  $F, g$  and  $x$ .

*Proof.* We call a triple  $(F, g, x)$  *admissible* if  $(F(g^n x))$  is a bounded nilsequence on  $G/\Gamma$  with Lipschitz constant  $M$ . Define for  $\omega, N \in \mathbb{N}$  and admissible  $(F, g, x)$

$$a_{\omega, (F, g, x)}(N) := \max_{r < W, (r, W)=1} \left| \frac{1}{N} \sum_{n=1}^N (\Lambda'_{r, \omega}(n) - 1) F(g^n x) \right|$$

and assume that the claimed uniform convergence does not hold. Then there exist  $\varepsilon > 0$ , a subsequence  $(\omega_j)$  of  $\mathbb{N}$  and a sequence of admissible  $(F_j, g_j, x_j)$  so that

$$\limsup_{N \rightarrow \infty} a_{\omega_j, (F_j, g_j, x_j)}(N) > \varepsilon \quad \text{for all } j \in \mathbb{N}.$$

In particular, there exists a subsequence  $(N_j)$  of  $\mathbb{N}$  such that  $a_{\omega_j, (F_j, g_j, x_j)}(N_j) > \varepsilon$  for every  $j \in \mathbb{N}$ .

Define now the function  $\omega : \mathbb{N} \rightarrow \mathbb{N}$  by

$$\omega(N) := \omega_j \quad \text{if } N \in [N_j, N_{j+1}),$$

which grows sufficiently slowly if  $(N_j)$  grows sufficiently fast. Then we have

$$a_{\omega(N_j), (F_j, g_j, x_j)}(N_j) = a_{\omega_j, (F_j, g_j, x_j)}(N_j) > \varepsilon \quad \text{for all } j \in \mathbb{N},$$

contradicting Theorem 2.1, which states that  $\lim_{N \rightarrow \infty} a_{\omega(N), (F, g, x)}(N) = 0$  uniformly in admissible  $(F, g, x)$ . □

### 3. Proof of Theorem 1.3

We first need several standard simple facts.

**LEMMA 3.1.** (See, e.g., [14]) *For a bounded sequence  $(a_n) \subset \mathbb{C}$ , one has*

$$\lim_{N \rightarrow \infty} \left| \frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} a_p - \frac{1}{N} \sum_{n=1}^N \Lambda'(n) a_n \right| = 0.$$

LEMMA 3.2. Let  $(b_n) \subset \mathbb{C}$  satisfy  $b_n = o(n)$ . Then the following assertions hold.

(a) The sequence  $(b_n)$  is Cesàro convergent if and only if for every  $\varepsilon > 0$ , there exist  $W, N_0 \in \mathbb{N}$  such that

$$\left| \frac{1}{WN} \sum_{n=1}^{WN} b_n - \frac{1}{WM} \sum_{n=1}^{WM} b_n \right| < \varepsilon \quad \text{for all } N, M \geq N_0. \tag{2}$$

(b) If  $(b_n)$  is supported on the primes, then, for every  $W \in \mathbb{N}$ ,

$$\frac{1}{WN} \sum_{n=1}^{WN} b_n = \frac{1}{W} \sum_{r < W, (r,W)=1} \frac{1}{N} \sum_{n=1}^N b_{Wn+r} + o_W(1). \tag{3}$$

*Proof.* (a) The ‘only if’ implication is clear. To show the ‘if’ implication, let  $\varepsilon > 0$  and take  $W, N_0$  satisfying (2). Let further  $N_1 \in \mathbb{N}$  be such that  $|b_n| < \varepsilon n/W$  holds for every  $n \geq N_1$ . We can assume without loss of generality that  $WN_0 \geq N_1$ . For  $k \geq WN_0$ , let  $N = N(k) \geq N_0$  be such that  $k \in [WN, W(N + 1))$ . By the triangle inequality, it suffices to show that

$$\left| \frac{1}{k} \sum_{n=1}^k b_n - \frac{1}{WN} \sum_{n=1}^{WN} b_n \right| < 4\varepsilon \tag{4}$$

if  $k$  is large enough.

We first observe that

$$\begin{aligned} \left| \frac{1}{k} \sum_{n=1}^k b_n - \frac{1}{WN} \sum_{n=1}^k b_n \right| &= \frac{k - WN}{kWN} \sum_{n=1}^k |b_n| \leq \frac{W}{k(k - W)} \left( O(1) + \sum_{n=N_1}^k \frac{k\varepsilon}{W} \right) \\ &\leq o(1) + \frac{k}{k - W} \varepsilon < 2\varepsilon \end{aligned}$$

for large enough  $k$ . On the other hand,

$$\left| \frac{1}{WN} \sum_{n=1}^k b_n - \frac{1}{WN} \sum_{n=1}^{WN} b_n \right| \leq \frac{1}{WN} \sum_{n=WN+1}^k \frac{n\varepsilon}{W} \leq \frac{k}{k - W} \varepsilon < 2\varepsilon$$

for large enough  $k$ , proving (4).

(b) The growth condition implies that

$$\frac{1}{WN} \sum_{n=1}^{WN} b_n = \frac{1}{WN} \sum_{r=1}^W \sum_{n=0}^{N-1} b_{Wn+r} = \frac{1}{W} \sum_{r=1}^W \frac{1}{N} \sum_{n=1}^N b_{Wn+r} + o_W(1).$$

If  $(b_n)$  is supported on the primes, (3) follows. □

The following property of connected nilsystems is well known.

LEMMA 3.3. Let  $X := G/\Gamma$  be a connected nilsystem with Haar measure  $\mu$  and  $g \in G$ . Then  $(X, \mu, g)$  is ergodic if and only if  $(X, \mu, g)$  is totally ergodic.

*Proof.* Since ergodicity of a nilsystem is equivalent to ergodicity of its Kronecker factor (also called maximal factor torus or ‘horizontal’ torus)  $G/([G, G]\Gamma)$ , see Leibman [27], we can assume without loss of generality that  $X$  is a compact connected abelian group.

Let  $(X, \mu, g)$  be ergodic,  $m \in \mathbb{N}$  and let  $F \in L^2(X, \mu)$  be an  $g^m$ -invariant function, i.e.,  $F(g^m x) = F(x)$  for every  $x \in X$ . Consider the Fourier decomposition

$$F = \sum_{\chi \in \widehat{X}} c_\chi \chi.$$

By the assumption, we have

$$F = \sum_{\chi \in \widehat{X}} c_\chi (\chi(g))^m \chi.$$

By the uniqueness of the decomposition, we obtain

$$c_\chi = c_\chi (\chi(g))^m \quad \text{for all } \chi \in \widehat{X}.$$

Assume that  $c_\chi \neq 0$ . Then  $(\chi(g))^m = 1$ , i.e.,  $\chi(g)$  is an  $m$ th root of unity. Since  $(X, \mu, g)$  is ergodic,  $\{g^n : n \in \mathbb{Z}\}$  is dense in  $X$ . Since  $\chi$  is a character and  $X$  is connected,  $\chi(g)$  has to be equal to 1—otherwise  $X$  would have two clopen components  $\overline{\{g^n : m_0 | n\}}$  and  $\{g^n : m_0 \nmid n\}$ , where  $m_0$  is the smallest period of  $\chi(g)$ . Thus,  $F = c_1 1$  and  $(X, \mu, g)$  is totally ergodic. □

*Proof of Theorem 1.3.* As mentioned above, we can assume that  $x = \text{id}_G \Gamma \in G^0$ , where  $G^0$  is the connected component of the identity in  $G$ , and  $G = \langle G^0, g_1, \dots, g_m \rangle$ .

Every polynomial nilsequence can be represented as a linear nilsequence on a larger nilmanifold; see Leibman [27, Proposition 3.14], Chu [9, Proposition 2.1 and its proof] and, in the context of connected groups, Green, Tao and Ziegler [21, Proposition C.2]. Thus, we can assume that  $g(n) = g^n$  for some  $g \in G$ .

By the argument in Wooley and Ziegler [39, text between Corollary 3.14 and Proposition 3.15], the nilsequence  $(F(g^n x))$  can be written as a finite sum of (linear) nilsequences coming from a connected and simply connected Lie group. Thus, we can assume without loss of generality that  $G$  is connected and simply connected.

We first assume that  $F$  is Lipschitz and define  $b_n := \Lambda'(n)F(g^n x)$ . To show convergence of

$$\frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} F(g^p x), \tag{5}$$

by Lemma 3.1 it is enough to show that  $(b_n)$  satisfies the condition in Lemma 3.2(a).

For every  $\omega \in \mathbb{N}$ , we have by Lemma 3.2(b),

$$\begin{aligned} \frac{1}{WN} \sum_{n=1}^{WN} b_n &= \frac{1}{W} \sum_{r < W, (r,W)=1} \frac{1}{N} \sum_{n=1}^N b_{Wn+r} + o_W(1) \\ &= \frac{1}{\phi(W)} \sum_{r < W, (r,W)=1} \frac{1}{N} \sum_{n=1}^N \Lambda'_{r,\omega}(n) F(g^{Wn+r} x) + o_W(1) \\ &= \frac{1}{\phi(W)} \sum_{r < W, (r,W)=1} \frac{1}{N} \sum_{n=1}^N (\Lambda'_{r,\omega}(n) - 1) F(g^{Wn+r} x) \\ &\quad + \frac{1}{\phi(W)} \sum_{r < W, (r,W)=1} \frac{1}{N} \sum_{n=1}^N F(g^{Wn+r} x) + o_W(1) \\ &=: I(N) + II(N) + o_W(1). \end{aligned} \tag{6}$$

Let  $\varepsilon > 0$  and take a large  $\omega$  such that  $\limsup_{N \rightarrow \infty} |I(N)| < \varepsilon$ , which exists by Corollary 2.2. Since the sequence  $(F(g^{Wn+r}x))_{n \in \mathbb{N}}$  is Cesàro convergent for every  $r$ , see Leibman [27] and Parry [33, 34], there is  $N_0 \in \mathbb{N}$  such that  $|II(N) - II(M)| < \varepsilon$  for every  $N, M \geq N_0$ . Thus,  $(b_n)$  satisfies the condition in Lemma 3.2(a).

Take now  $F \in C(G/\Gamma)$  arbitrary,  $x \in G/\Gamma$  and  $\varepsilon > 0$ . By the uniform continuity of  $F$ , there exists  $G \in C(G/\Gamma)$  Lipschitz with  $\|F - G\|_\infty \leq \varepsilon$ . We then have

$$\begin{aligned} & \left| \frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} F(g^p x) - \frac{1}{\pi(M)} \sum_{p \in \mathbb{P}, p \leq M} F(g^p x) \right| \\ & \leq \left| \frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} F(g^p x) - \frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} G(g^p x) \right| \\ & \quad + \left| \frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} G(g^p x) - \frac{1}{\pi(M)} \sum_{p \in \mathbb{P}, p \leq M} G(g^p x) \right| \\ & \quad + \left| \frac{1}{\pi(M)} \sum_{p \in \mathbb{P}, p \leq M} G(g^p x) - \frac{1}{\pi(M)} \sum_{p \in \mathbb{P}, p \leq M} F(g^p x) \right| \\ & \leq 2\varepsilon + \left| \frac{1}{\pi(N)} \sum_{p \in \mathbb{P}, p \leq N} G(g^p x) - \frac{1}{\pi(M)} \sum_{p \in \mathbb{P}, p \leq M} G(g^p x) \right|, \end{aligned}$$

which is less than  $3\varepsilon$  for large enough  $N, M$  by the above, finishing the argument.

The last assertion of the theorem follows analogously from the decomposition (6) using Lemma 3.3, the fact that a nilsystem is ergodic if and only if it is uniquely ergodic, see Parry [33, 34], and the uniform convergence of Birkhoff’s ergodic averages to the space mean for uniquely ergodic systems. The last step (for non-Lipschitz functions) should be modified by showing that the difference  $(1/\pi(N)) \sum_{p \in \mathbb{P}, p \leq N} F(g^p x) - (1/N) \sum_{n=1}^N F(g^n x)$  converges to zero.  $\square$

*Acknowledgements.* The author thanks Vitaly Bergelson for correcting the references and is deeply grateful to the referee for careful reading and suggestions which have considerably improved the paper.

REFERENCES

[1] L. Auslander, L. Green and F. Hahn. *Flows on Homogeneous Spaces (Annals of Mathematics Studies, 53)*. Princeton University Press, Princeton, NJ, 1963.  
 [2] V. Bergelson, B. Host and B. Kra. Multiple recurrence and nilsequences. *Invent. Math.* **160** (2005), 261–303.  
 [3] V. Bergelson and A. Leibman. Distribution of values of bounded generalized polynomials. *Acta Math.* **198** (2007), 155–230.  
 [4] V. Bergelson, A. Leibman and E. Lesigne. Intersective polynomials and the polynomial Szemerédi theorem. *Adv. Math.* **219** (2008), 369–388.  
 [5] V. Bergelson, A. Leibman and T. Ziegler. The shifted primes and the multidimensional Szemerédi and polynomial van der Waerden theorems. *C. R. Math. Acad. Sci. Paris* **349** (2011), 123–125.  
 [6] J. Bourgain. An approach to pointwise ergodic theorems. *Geometric Aspects of Functional Analysis (1986/87) (Lecture Notes in Mathematics, 1317)*. Springer, Berlin, 1988, pp. 204–223.

- [7] J. Bourgain. On the pointwise ergodic theorem on  $L^p$  for arithmetic sets. *Israel J. Math.* **61** (1988), 73–84.
- [8] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Publ. Math. Inst. Hautes Études Sci.* **169** (1989), 5–45.
- [9] Q. Chu. Convergence of weighted polynomial multiple ergodic averages. *Proc. Amer. Math. Soc.* **137** (2009), 1363–1369.
- [10] J. P. Conze and E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *C. R. Acad. Sci. Paris Sér. I Math.* **306** (1988), 491–493.
- [11] T. Eisner and P. Zorin-Kranich. Uniformity in the Wiener–Wintner theorem for nilsequences. *Discrete Contin. Dyn. Syst.* **33** (2013), 3497–3516.
- [12] N. Frantzikinakis. Multiple correlation sequences and nilsequences. *Invent. Math.* **202** (2015), 875–892.
- [13] N. Frantzikinakis and B. Host. Weighted multiple ergodic averages and correlation sequences. *Ergod. Th. & Dynam. Sys.* **38** (2018), 81–142.
- [14] N. Frantzikinakis, B. Host and B. Kra. Multiple recurrence and convergence for sequences related to the prime numbers. *J. Reine Angew. Math.* **611** (2007), 131–144.
- [15] N. Frantzikinakis, B. Host and B. Kra. The polynomial multidimensional Szemerédi theorem along shifted primes. *Israel J. Math.* **194** (2013), 331–348.
- [16] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31** (1977), 204–256.
- [17] H. Furstenberg and B. Weiss. A mean ergodic theorem for  $\frac{1}{N} \sum_{n=1}^N f(T^n x)g(T^{n^2} x)$ . *Convergence in Ergodic Theory and Probability (Ohio State University Mathematical Research Institute Publications, 5)*. De Gruyter, Berlin, 1996, pp. 193–227.
- [18] B. Green and T. Tao. Linear equations in primes. *Ann. of Math. (2)* **171** (2010), 1753–1850.
- [19] B. Green and T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math. (2)* **175** (2012), 465–540.
- [20] B. Green and T. Tao. The Möbius function is strongly orthogonal to nilsequences. *Ann. of Math. (2)* **175** (2012), 541–566.
- [21] B. Green, T. Tao and T. Ziegler. An inverse theorem for the Gowers  $U^{s+1}$ -norm. *Ann. of Math. (2)* **176** (2012), 1231–1372.
- [22] L. W. Green. Spectra of nilmanifolds. *Bull. Amer. Math. Soc. (N.S.)* **67** (1961), 414–415.
- [23] B. Host and B. Kra. An odd Furstenberg–Szemerédi theorem and affine systems. *J. Anal. Math.* **86** (2002), 183–220.
- [24] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)* **161** (2005), 397–488.
- [25] B. Host and B. Kra. Uniformity seminorms on  $L^\infty$  and applications. *J. Anal. Math.* **108** (2009), 219–276.
- [26] B. Krause. Polynomial ergodic averages converge rapidly: variations on a theorem of Bourgain. *Preprint*, 2014, arXiv:1402.1803.
- [27] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergod. Th. & Dynam. Sys.* **25** (2005), 201–213.
- [28] A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables. *Israel J. Math.* **146** (2005), 303–315.
- [29] A. Leibman. Nilsequences, null-sequences, and multiple correlation sequences. *Ergod. Th. & Dynam. Sys.* **35** (2015), 176–191.
- [30] E. Lesigne. Sur une nil-variété, les parties minimales associées à une translation sont uniquement ergodiques. *Ergod. Th. & Dynam. Sys.* **11** (1991), 379–391.
- [31] R. Nair. On polynomials in primes and Bourgain’s circle method approach to ergodic theorems. *Ergod. Th. & Dynam. Sys.* **11** (1991), 485–499.
- [32] R. Nair. On polynomials in primes and Bourgain’s circle method approach to ergodic theorems II. *Studia Math.* **105** (1993), 207–233.
- [33] W. Parry. Ergodic properties of affine transformations and flows on nilmanifolds. *Amer. J. Math.* **91** (1969), 757–771.
- [34] W. Parry. Dynamical systems on nilmanifolds. *Bull. Lond. Math. Soc.* **2** (1970), 37–40.
- [35] E. Szemerédi. On sets of integers containing no  $k$  elements in arithmetic progression. *Acta Arith.* **27** (1975), 199–245 a collection of articles in memory of Jurii Vladimirovič Linnik.
- [36] J.-P. Thouvenot. La convergence presque sûre des moyennes ergodiques suivant certaines sous-suites d’entiers (d’après Jean Bourgain) [Almost sure convergence of ergodic means along some subsequences of integers (after Jean Bourgain)]. *Sémin. Bourbaki*, Vol. 1989/90, *Astérisque* No. 189–190 (1990), Exp. No. 719, 133–153 (in French).
- [37] M. Wierdl. Pointwise ergodic theorem along the prime numbers. *Israel J. Math.* **64** (1988), 315–336.



- [38] M. Wierdl. Almost everywhere convergence and recurrence along subsequences in ergodic theory. *PhD Thesis*, Ohio State University, 1989.
- [39] T. Wooley and T. Ziegler. Multiple recurrence and convergence along the primes. *Amer. J. Math.* **134** (2012), 1705–1732.
- [40] T. Ziegler. A non-conventional ergodic theorem for a nilsystem. *Ergod. Th. & Dynam. Sys.* **25** (2005), 1357–1370.
- [41] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.* **20** (2007), 53–97.
- [42] P. Zorin-Kranich. Variation estimates for averages along primes and polynomials. *J. Funct. Anal.* **268** (2015), 210–238.
- [43] P. Zorin-Kranich. A double return times theorem. *Preprint*, 2015, arXiv:1506.05748v1.pdf.