4. Polynomial Recurrence

In this section we prove a polynomial recurrence theorem, which in view of the Correspondence Principle implies Sàrközy's theorem (Theorem 2.10). To prove it we introduce an important tool in Ergodic Ramsey Theory – the van der Corput trick.

Another idea that is briefly explored in this section is that of a dichotomy between "structure" and "randomness", albeit in a very embryonic form. In this context, structure is captured by periodic functions, and randomness (or "mixing") is captured by the notion of total ergodicity. This kind of dichotomy will become more clear (and useful) in the following sections.

4.1. The van der Corput trick.

If the Correspondence Principle is the soul of Ergodic Ramsey Theory, its beating heart is the so-called van der Corput trick. There are many variations of this technique (the interested reader may read the expository article [3]), catered for specific applications throughout Ergodic Ramsey Theory.

The original lemma due to van der Corput [5] is concerned with uniform distribution in the unit interval.

Definition 4.1. *A sequence $(x_n)_{n=1}^\infty$ taking values in $[0,1]$ is said to be* **uniformly distributed** *or* **equidistributed** *if for every interval $(a,b)\subset[0,1]$,*

$$\lim_{N\to\infty}\frac{1}{N}\Big|\big\{n\in[1,N]:x_n\in(a,b)\big\}\Big|=b-a. \tag{4.1}$$

Due to the fact that there are uncountably many intervals (a,b) inside $[0,1]$, it is not clear that uniformly distributed sequences even exist. However, we have the following criterion by Weil [22] (for a proof, see [15, Theorems 1.1.1 and 1.2.1]).

Lemma 4.2 (Weyl criterion)**.** *Let $(x_n)_{n=1}^\infty$ be a sequence taking values in $[0,1]$. The following are equivalent.*

1. *$(x_n)_{n=1}^\infty$ is uniformly distributed.*
2. *The sequence of measures $\mu_N=\frac{1}{N}\sum_{n=1}^N\delta_{x_n}$ converges in the weak* topology to the Lebesgue measure.*
3. *For every continuous function $f\in C[0,1]$,*
$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N f(x_n)=\int_0^1 f(t)\,\mathsf{d}t.$$
4. $$\forall\ h\in\mathbb{N}\qquad \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N e^{2\pi ihx_n}=0. \tag{4.2}$$

Example 4.3. *Let $\alpha\in\mathbb{R}\setminus\mathbb{Q}$. Then the sequence $x_n=n\alpha \bmod 1$ is uniformly distributed. Indeed, for every $h,N\in\mathbb{N}$ we have*

$$\frac{1}{N}\sum_{n=1}^N e^{2\pi ihx_n}=\frac{1}{N}\sum_{n=1}^N\left(e^{2\pi ih\alpha}\right)^n=\frac{1}{N}\cdot\frac{e^{2\pi ih\alpha(N+1)}-e^{2\pi ih\alpha}}{e^{2\pi ih\alpha}-1}$$

and the last expression converges to 0 as $N\to\infty$.

Exercise 4.4. *Show that the sequence $x_n=\sqrt{n}\bmod 1$ is uniformly distributed.*

Exercise 4.5. *Show that the sequence $x_n=\log n\bmod 1$ is* **not** *uniformly distributed.*

Here is the original version of the van der Corput trick.

Lemma 4.6. *Let $(x_n)_{n=1}^\infty$ be a sequence taking values in $\mathbb{R}$. If for every $m\in\mathbb{N}$ the sequence $n\mapsto x_{n+m}-x_n\bmod 1$ is uniformly distributed, then also the sequence $n\mapsto x_n\bmod 1$ is uniformly distributed.*

We will prove a more general result below. As a corollary of Lemma 4.6 we obtain Weyl's equidistribution theorem.

Corollary 4.7. *Let $f\in\mathbb{R}[t]$ be a polynomial with real coefficients. If at least one of the coefficients of f, other than the constant term, is irrational, then $f(n)\bmod 1$ is uniformly distributed.*

Proof. We proceed by induction on the degree $d = d(f)$ of the largest degree term of f with an irrational coefficient. If $d = 1$, then the sequence $f(n) \bmod 1$ is the sum of a periodic sequence (say of period p) and the sequence $n \mapsto n\alpha \bmod 1$ where α is the irrational coefficient of degree 1. Since $p\alpha$ is still irrational, one can adapt the argument in Example 4.3 to show that $f(n) \bmod 1$ is indeed uniformly distributed when $d = 1$.

Next suppose that $d > 1$. For each $m \in \mathbb{N}$, the sequence $g_m : n \mapsto f(n+m) - f(n)$ is itself a polynomial with $d(g_m) = d(f) - 1$ by induction, $g_m \bmod 1$ is uniformly distributed, and in view of Lemma 4.6, so is $f(n) \bmod 1$. □

The most useful versions of the van der Corput trick for Ergodic Ramsey Theory deal with sequences of vectors in a Hilbert space; here is a simple formulation that will be useful later.

Lemma 4.8. *Let H be a Hilbert space and let $(x_n)_{n=1}^{\infty}$ be a bounded sequence taking values in H. If for every $d \in \mathbb{N}$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_{n+d}, x_n \rangle = 0 \tag{4.3}$$

then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x_n = 0.$$

There is also a version for uniform Cesàro averages, which can be proved in the same way (see Exercise 4.9 below).

Proof of Lemma 4.8. For any $\epsilon > 0$ and any $D \in \mathbb{N}$, if $N \in \mathbb{N}$ is large enough we have

$$\left\| \frac{1}{N} \sum_{n=1}^{N} x_n - \frac{1}{D} \sum_{d=1}^{D} \frac{1}{N} \sum_{n=1}^{N} x_{n+d} \right\| < \frac{\epsilon}{2}$$

Hence it suffices to show that, if D is large enough,

$$\limsup_{N\to\infty} \left\| \frac{1}{D} \sum_{d=1}^{D} \frac{1}{N} \sum_{n=1}^{N} x_{n+d} \right\| < \frac{\epsilon}{2}$$

Using the Cauchy-Schwarz inequality we have

$$\begin{aligned}
\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} \frac{1}{D} \sum_{d=1}^{D} x_{n+d} \right\|^2 &\leq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left\| \frac{1}{D} \sum_{d=1}^{D} x_{n+d} \right\|^2 \\
&= \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \frac{1}{D^2} \sum_{d_1,d_2=1}^{D} \langle x_{n+d_1}, x_{n+d_2} \rangle \\
&\leq \frac{1}{D^2} \sum_{d_1,d_2=1}^{D} \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_{n+d_1}, x_{n+d_2} \rangle
\end{aligned} \tag{4.4}$$

Note that, for $d_1 \neq d_2$, it follows from (4.5) that $\frac{1}{N}\sum_{n=1}^{N} \langle x_{n+d_1}, x_{n+d_2} \rangle \to 0$ as $N \to \infty$. We conclude that the quantity in (4.4) is bounded by $\frac{D}{D^2} = \frac{1}{D}$ which is arbitrarily small for large enough D. □

Exercise 4.9. *Adapt the proof of Lemma 4.8 to the following version for uniform Cesàro averages (see Remark 2.18): Let H be a Hilbert space and let $(x_n)_{n=1}^{\infty}$ be a bounded sequence taking values in H. If for every $d \in \mathbb{N}$,*

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N} \langle x_{n+d}, x_n \rangle = 0 \tag{4.5}$$

then

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N} x_n = 0.$$

Exercise 4.10. $(*)$

Let $p \in \mathbb{R}[x]$ have at least one irrational coefficient (other than the constant term) and let $U \subset [0,1]$ be open and non-empty. Is it true that the set $\{n \in \mathbb{N} : p(n) \bmod 1 \in U\}$ is syndetic? [Hint: Use Exercise 4.9 to obtain versions of Lemma 4.6 and Corollary 4.7 for uniform Cesàro averages and then use a similar argument as for Exercise 2.23.]

4.2. **Totally ergodicity.**

Definition 4.11. *A measure preserving system $(X, \mathcal{B}, \mu, T)$ is* **totally ergodic** *if for every $n \in \mathbb{N}$, the measure preserving system $(X, \mathcal{B}, \mu, T^n)$ is ergodic.*

A convenient notation we will often use from now on is the following: given a m.p.s. $(X, \mathcal{B}, \mu, T)$ and a function $f \in L^2(X)$, we denote by Tf the composition $f \circ T$ (another way to think about this is, as an abuse of language, to denote by T the associated Koopman operator).

Example 4.12. *Recall the circle rotation $(X, \mathcal{B}, \mu, T)$ described in Example 2.4, where $X = [0,1]$, $\mathcal{B}$ is the Borel σ-algebra, μ is the Lebesgue measure and $T : x \mapsto x + \alpha \bmod 1$. This system is totally ergodic if and only if α is irrational. Indeed, if α is rational, say $\alpha = p/q$, then $q\alpha$ is an integer and hence T^q is the identity map on $[0,1]$, which is trivially not ergodic.*

On the other hand, if α is irrational, then the system is ergodic. To see this we use the ergodic theorem. Then we need to show that for every $f \in L^2$ the average

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^n f$$

is a constant function. But this is easy to check for functions $t \mapsto e(nt)$ with $n \in \mathbb{Z}$, and finite linear combinations of functions of this kind form a dense subset of L^2.

Finally, for every $n \in \mathbb{N}$, the measure preserving system $(X, \mathcal{B}, \mu, T^n)$ is the circle rotation by $n\alpha$; since $n\alpha$ is also irrational when α is, the system $(X, \mathcal{B}, \mu, T)$ is totally ergodic in this case.

When a system $(X, \mathcal{B}, \mu, T)$ is totally ergodic, we obtain from the ergodic theorem the following corollary.

Corollary 4.13. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. Then it is totally ergodic if and only if for every $f \in L^2(X)$ and every $q, r \in \mathbb{N}$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{qn+r} f = \int_X f \, \mathsf{d}\mu. \qquad \text{in } L^2(X) \tag{4.6}$$

Proof. If the system is not totally ergodic, then there exists $q \in \mathbb{N}$ and a non-constant $f \in L^2(X)$ such that $T^q f = f$. Thus (4.6) implies that the system is totally ergodic.

To prove the converse direction, let $(X, \mathcal{B}, \mu, T)$ be totally ergodic and let $f \in L^2(X)$ and $q, r \in \mathbb{N}$ be arbitrary. Applying the ergodic theorem (Theorem 2.17) to the (ergodic) system $(X, \mathcal{B}, \mu, T^q)$ we conclude that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{qn+r} f = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} (T^q)^n (T^r f) = \int_X T^r f \, \mathsf{d}\mu = \int_X f \, \mathsf{d}\mu.$$

□

Remark 4.14. *A measure preserving system $(X, \mathcal{B}, \mu, T)$ is called* **invertible** *if T is invertible a.e. and the inverse is measurable and measure preserving. In this situation we can allow q and r in Corollary 4.13 to be negative, but if the system is not invertible, then the expression $T^n f$ does not make sense for a negative value of n.*

Nevertheless, Corollary 4.13 still makes sense when $r < 0$, even if the system is not invertible. Indeed, in this case the expression $qn + r$ is positive for all but finitely many values of n, and since we take an average over $\mathbb{N}$ we can just ignore those finitely many values.

One could interpret the expression T^{qn+r} appearing in (4.6) as $T^{p(n)}$ where p is a linear polynomial. The following theorem reveals the power of the van der Corput trick, which allows one to upgrade Corollary 4.13 to general polynomials.

Theorem 4.15. *Let $(X, \mathcal{B}, \mu, T)$ be a totally ergodic system and let $p \in \mathbb{Z}[x]$ be such that either the system is invertible or the polynomial has a positive leading coefficient. Then for every $f \in L^2(X)$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{p(n)} f = \int_X f \, \mathsf{d}\mu. \qquad \text{in } L^2(X) \tag{4.7}$$

Proof. We proceed by induction on the degree of p. If p is linear, then the result follows from Corollary 4.13, so assume that p has degree at least 2. Eq. (4.7) holds for f if and only if it holds for $f - c$ where c is a constant; therefore, after subtracting $\int_X f \, \mathsf{d}\mu$ from f we can assume that $\int_X f \, \mathsf{d}\mu = 0$. Letting $x_n = T^{p(n)} f$, we need to show that $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x_n = 0$, and to this end we will invoke the van der Corput lemma (Lemma 4.8). Fixing $d \in \mathbb{N}$ we can compute

$$\langle x_{n+d}, x_n \rangle = \int_X T^{p(n+d)} f \cdot T^{p(n)} \overline{f} \, \mathsf{d}\mu = \int_X T^{p(n+d)-p(n)} f \cdot \overline{f} \, \mathsf{d}\mu = \left\langle T^{p(n+d)-p(n)} f, f \right\rangle.$$

Since $n \mapsto p(n+d) - p(n)$ is a polynomial of degree smaller than the degree of p, we can use the induction hypothesis (together with the fact that convergence in $L^2(X)$ implies convergence in the weak topology) to conclude

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \langle x_{n+d}, x_n \rangle = \left\langle \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{p(n+d)-p(n)} f, f \right\rangle = 0.$$

This establishes the hypothesis (4.5) of the van der Corput lemma, so we conclude that $\displaystyle\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x_n = 0$, as desired. □

Remark 4.16. *Both Corollary 4.13 and Theorem 4.15 have versions for uniform Cesàro averages, which can be proved in the exact same way. The choice to present the regular Cesàro versions was made with the hope that the main ideas became more transparent.*

4.3. **Total ergodicity and finite factors.** This subsection is not necessary to the proof of Sàrközy's theorem, but it leads to some important ideas that will appear in later sections.

Here is another example of an ergodic system that is not totally ergodic.

Example 4.17. *Let $X = \{0, 1\}$, $\mathcal{B}$ the discrete σ-algebra, μ the normalized counting measure and $T : x \mapsto x + 1 \bmod 2$. In other words $(X, \mathcal{B}, \mu, T)$ is a transposition on 2 points. Then this system is ergodic, since the only sets with measure in $(0, 1)$ are the singletons $\{0\}$ and $\{1\}$, and neither of them is invariant. However, the system is not totally ergodic, since T^2 is the identity map and leaves both singletons (which have positive measure) invariant.*

While Example 4.17 seems rather trivial, it turns out that finite systems are in some sense the only obstruction to total ergodicity. To better capture this, we need the notion of factor maps.

Definition 4.18 (Factor map)**.** *Let $(X, \mathcal{A}, \mu, T)$ and $(Y, \mathcal{B}, \nu, S)$ be m.p.s. and let $\phi : X \to Y$. Then ϕ is a* **factor map** *if it is surjective, preserves the measure (i.e. $\mu(\phi^{-1}B) = \nu(B)$ for every $B \in \mathcal{B}$) and intertwines T and S, in the sense that $S \circ \phi = \phi \circ T$.*

More generally, one can allow ϕ to be a surjective map between full measure sets $X_0 \in \mathcal{A}$ and $Y_0 \in \mathcal{B}$ such that $T^{-1}X_0 = X_0$ and $S^{-1}Y_0 = Y_0$, and the relation $S \circ \phi = \phi \circ T$ only needs to hold in X_0.

We say that the system $(Y, \mathcal{B}, \nu, S)$ is a ***factor*** of $(X, \mathcal{A}, \mu, T)$ if there is a factor map $\phi : X \to Y$. We will also say that, in this case, $(X, \mathcal{A}, \mu, T)$ is an ***extension*** of $(Y, \mathcal{B}, \nu, S)$.

Theorem 4.19. *Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system. Then it is totally ergodic if and only if it does not allow for any non-trivial finite factor.*

Proof. Let $(Y, \mathcal{B}, \nu, S)$ be a non-trivial finite system and suppose that there is a factor map $\pi : X \to Y$. Let $y \in Y$ be such that $\nu(\{y\}) \in (0, 1)$ and let $A = \pi^{-1}(\{y\})$. Then $\mu(A) = \nu(\{y\}) \in (0, 1)$. Let $k = |Y|!$. Then S^k acts trivially on Y, and in particular $S^{-k}\{y\} = \{y\}$. Therefore $T^{-k}A = A$ and we conclude that $(X, \mathcal{A}, \mu, T^k)$ is not ergodic.

To prove the converse direction, suppose that $(X, \mathcal{A}, \mu, T)$ is not totally ergodic. Let $n \in \mathbb{N}$ be such that T^n is not ergodic and let $A \in \mathcal{A}$ be such that $\mu(A) \in (0,1)$ and $T^{-n}A = A$. It follows that the σ-algebra $\mathcal{B}$ generated by the sets $A, T^{-1}A, \ldots, T^{-(n-1)}A$ is invariant under T, finite and non-trivial. Let Y be the (finite) set of atoms of $\mathcal{B}$, and let $\pi : X \to Y$ be the containment map (i.e. $\pi(x)$ is the atom of $\mathcal{B}$ that contains x; more explicitly $\pi(x) = \bigcap_{\substack{B\in\mathcal{B}\\ x\in B}} B$). It is easy to check that π is indeed a factor map. □

Exercise 4.20. *Finish the proof of Theorem 4.19 by explicitly describing the measure preserving system structure of Y and showing that π is indeed a factor map.*

Exercise 4.21. *Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system and let $(Y, \mathcal{B}, \nu, S)$ be a factor. Prove that:*

- *If $(X, \mathcal{A}, \mu, T)$ is ergodic, then so is $(Y, \mathcal{B}, \nu, S)$.*
- *If $(X, \mathcal{A}, \mu, T)$ is totally ergodic, then so is $(Y, \mathcal{B}, \nu, S)$.*

4.4. **Proof of Sárközy's theorem.** Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. Consider the following subspaces of $L^2(X)$:

$$H_{rat} := \overline{\{f \in L^2(X) : T^k f = f \text{ for some } k \in \mathbb{N}\}};\ H_{te} := \left\{ f \in L^2(X) : \forall\, k \in \mathbb{N},\ \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T^{kn} f = 0 \right\}$$

Exercise 4.22. *Show that if $f, g \in H_{rat}$ are bounded, then their product $f \cdot g$ is also in H_{rat}. Can you find an example showing that the same is not tru for H_{te}?*

Exercise 4.23. $(*)$
Show that the collection $\{A \in \mathcal{B} : 1_A \in H_{rat}\}$ is a σ-algebra.

Observe that in a totally ergodic system the space H_{rat} consists only of constant functions, while the space H_{te} contains every function with 0 integral. The following proposition generalizes this observation.

Proposition 4.24. *For any measure preserving system $(X, \mathcal{B}, \mu, T)$, the spaces H_{rat} and H_{te} are orthogonal and $L^2(X) = H_{rat} \oplus H_{te}$.*

Proof. Let $f \in L^2(X)$ be such that $T^k f = f$ for some $k \in \mathbb{N}$ and let $g \in H_{te}$. Then $\langle f, g\rangle = \langle T^k f, T^k g\rangle = \langle f, T^k g\rangle$. Iterating this observation we deduce that $\langle f, g\rangle = \langle f, T^{kn} g\rangle$ for every $n \in \mathbb{N}$. Averaging over n we then deduce

$$\langle f, g\rangle = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \langle f, T^{kn} g\rangle = \left\langle f, \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T^{kn} g \right\rangle = 0,$$

showing that H_{rat} and H_{te} are orthogonal.

Now suppose that $f \in L^2(X)$ is orthogonal to H_{rat}, we need to show that $f \in H_{te}$. But for every $k \in \mathbb{N}$, the space H_{rat} contains the invariant subspace I_k for the system $(X, \mathcal{B}, \mu, T^k)$. It follows that f is orthogonal to I_k for every k, and in view of the mean ergodic theorem, $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} T^{kn} f = 0$, so that indeed $f \in H_{te}$. □

We are now ready to prove Sárközy's theorem (Theorem 1.13). Using the correspondence principle, or more precisely, using Proposition 2.11, our task is reduced to establishing polynomial recurrence, formulated in Theorem 2.10. The proof we provided for the Poincaré recurrence theorem (Theorem 2.7) does not extend far beyond the scope of Theorem 2.7. However, we saw a different proof of Poincaré's recurrence when proving the stronger Khintchine's recurrence (Theorem 2.24) using the ergodic theorem. Our proof of Theorem 2.10 follows this second strategy, replacing the ergodic theorem with the "polynomial ergodic theorem" for totally ergodic systems that we obtained in Eq. (4.7).

We will in fact establish a stronger version of Theorem 2.10.

Definition 4.25. *A polynomial $p \in \mathbb{Z}[x]$ is called* **divisible** *or* **intersective** *if for every $k \in \mathbb{N}$ there exists $n \in \mathbb{N}$ such that $p(n)$ is a multiple of k.*

If $p(0) = 0$ or, more generally, p has an integer root, then it is divisible. However there are polynomials, such as $p(x) = (x^2 - 3)(x^2 - 5)(x^2 - 15)$ which have no integer root but are divisible. It is easy to see that if p is not divisible, then there exists a finite system where recurrence does not occur at times of the form $p(n)$. In other words, if p is not divisible, then the set $\{p(n) : n \in \mathbb{N}\}$ is not a set of recurrence. The converse of this observation is the content of the following theorem, which significantly extends Theorem 2.10.

Theorem 4.26. *Let $(X,\mathcal{B},\mu,T)$ be a measure preserving system, let $A\in\mathcal{B}$, let $\epsilon>0$ and let $p\in\mathbb{Z}[x]$ be a divisible polynomial with a positive leading coefficient. Then there exists $n\in\mathbb{N}$ such that $\mu(A\cap T^{-p(n)}A)>\mu^2(A)-\epsilon$.*

Proof. Decompose $1_A=f+g$ with $f\in H_{rat}$ and $g\in H_{te}$. Since H_{rat} contains the constant functions, using the Cauchy-Schwarz inequality we have $\langle 1_A,f\rangle=\|f\|^2\geq\langle f,1\rangle^2=\mu(A)^2$. Find $h\in H_{rat}$ such that $T^kh=h$ for some $k\in\mathbb{N}$, and such that $\|f-h\|<\epsilon/2$. In particular it follows that $\langle 1_A,h\rangle>\mu(A)^2-\epsilon/2$.

Using divisibility of p, find $a\in\mathbb{N}$ such that $p(a)\equiv 0\bmod k$ and consider the polynomial $q(n)=p(a+kn)$. Then $T^{q(n)}h=h$ for all $n\in\mathbb{N}$. As in the proof of Theorem 4.15, an application of the van der Corput trick implies that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}T^{q(n)}g=0.$$

Finally, we have

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(A\cap T^{-q(n)}A) &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\left\langle 1_A,h+T^{q(n)}(f-h)+T^{q(n)}g\right\rangle\\ &= \left\langle 1_A,h+\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}T^{q(n)}(f-h)+T^{q(n)}g\right\rangle\\ &\geq \langle 1_A,h\rangle-\epsilon/2 \quad\geq\quad \mu(A)^2-\epsilon.\end{aligned}$$

□

Exercise 4.27. *Adapt the proof of Theorem 4.26 to obtain that, under the same conditions, if additionally $\mu(A)>0$, then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(A\cap T^{-p(n)}A)>0.$$

Exercise 4.28. $(*)$ *Using Exercise 4.9 in the proof of Theorem 4.26, show that for any set $E\subset\mathbb{N}$ with $\bar{d}(E)>0$, the set $\{n\in\mathbb{N}:n^2\in E-E\}$ is syndetic.*

References

[1] V. Bergelson. Ergodic Ramsey theory–an update. In *Ergodic theory of $\mathbb{Z}^d$ actions*, volume 228 of *London Math. Soc. Lecture Note Ser.*, pages 1–61. Cambridge Univ. Press, Cambridge, 1996.

[2] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.*, 9(3):725–753, 1996.

[3] V. Bergelson and J. Moreira. Van der Corput's difference theorem: some modern developments. *Indag. Math. (N.S.)*, 27(2):437–479, 2016.

[4] I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinaĭ. *Ergodic theory*, volume 245 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1982. Translated from the Russian by A. B. Sosinskiĭ.

[5] J. G. van der Corput. Diophantische Ungleichungen. I. Zur Gleichverteilung Modulo Eins. *Acta Math.*, 56(1):373–456, 1931.

[6] Manfred Einsiedler and Thomas Ward. *Ergodic theory with a view towards number theory*, volume 259 of *Graduate Texts in Mathematics*. Springer-Verlag London, Ltd., London, 2011.

[7] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. d'Analyse Math.*, 31:204–256, 1977.

[8] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton University Press, Princeton, N.J., 1981.

[9] E. Glasner. *Ergodic theory via joinings*, volume 101 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003.

[10] T. Gowers. A new proof of Szemerédi's theorem. *GAFA*, 11:465–588, 2001.

[11] R. L. Graham, B. L. Rothschild, and J. H. Spencer. *Ramsey theory.* John Wiley & Sons, Inc., New York, second edition, 1990.

[12] B. Host and B. Kra. *Nilpotent structures in ergodic theory*, volume 236 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2018.

[13] T. Kamae and M. Mendès France. Van der Corput's difference theorem. *Israel J. Math.*, 31(3-4):335–342, 1978.

[14] A. Y. Khinchin. *Three pearls of number theory.* Graylock Press, Rochester, N. Y., 1952.

[15] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* John Wiley & Sons, New York-London-Sydney, 1974.

[16] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:245–252, 1953.
[17] A. Sárközy. On difference sets of sequences of integers. I. *Acta Math. Acad. Sci. Hungar.*, 31(1–2):125–149, 1978.
[18] I. Schur. Über die kongruenz $x^m + y^m \equiv z^m (\text{mod } p)$. *Jahresbericht der Deutschen Math. Verein.*, 25:114–117, 1916.
[19] E. Szemerédi. On the sets of integers containing no k elements in arithmetic progressions. *Acta Arith.*, 27:299–345, 1975.
[20] B. L. van der Waerden. Beweis einer baudetschen vermutung. *Nieuw. Arch. Wisk.*, 15:212–216, 1927.
[21] P. Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.
[22] H. Weyl. Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.*, 77(3):313–352, 1916.