ERGODIC AVERAGING SEQUENCES

MICHAEL BOSHERNITZAN, GRIGORI KOLESNIK, ANTHONY QUAS,
AND MÁTÉ WIERDL

Abstract. We consider generalizations of the pointwise and mean ergodic theorems to ergodic theorems averaging along different subsequences of the integers or real numbers.

The Birkhoff and Von Neumann ergodic theorems give conclusions about convergence of average measurements of systems when the measurements are made at integer times.

We consider the case when the measurements are made at times $a(n)$ or $(\lfloor a(n) \rfloor)$where the function $a(x)$ is taken from a class of functions called a Hardy field, and we also assume that $|a(x)|$ goes to infinity slower than some positive power of x. A special, well-known Hardy field is Hardy's class of logarithmico-exponential functions.

The main theme of the paper is to point out that for a function $a(x)$ as described above, a complete characterization of the ergodic averaging behavior of the sequence $(\lfloor a(n) \rfloor)$is possible in terms of the distance of $a(x)$ from (certain) polynomials.

Contents

1. Introduction

Let $(T^s)_{s\in\mathbb{R}}$ be a measure preserving flow on the probability space $(\Omega, \mathcal{B}, \mu)$, and let (a_n) be a sequence of real numbers. Consider the ergodic averages along the

This research was supported by grants from the NSF.

sequence (a_n)

$$A_t f = A_t((a_n), f) = \frac{1}{t}\sum_{n\le t} f\circ T^{a_n}$$

We say the sequence (a_n) is *pointwise good*, if the following property holds: for any measure preserving flow $(T^s)_{s\in\mathbb{R}}$ on any probability space $(\Omega, \mathcal{B}, \mu)$ and $f \in L^2(\Omega)$, the averages $A_t f$ converge a.e. as $t \to \infty$.

We say the sequence (a_n) is *norm good* if the corresponding convergence in norm occurs: for any measure preserving flow $(T^s)_{s\in\mathbb{R}}$ on any probability space $(\Omega, \mathcal{B}, \mu)$ and $f \in L^2(\Omega)$, the averages $A_t f$ converge in L^2 norm as $t \to \infty$.

In this paper, we consider the case when $a_n = a(n)$ or $a_n = \lfloor a(n)\rfloor$ for some real valued function a defined for every large enough real number x.

In the case where $a(x)$ is a polynomial with real coefficients, J. Bourgain proved in a series of papers that $(a(n))$ and $(\lfloor a(n)\rfloor)$ are good. That polynomial sequences are norm good is much simpler to prove, and the proof is essentially Von Neumann's method used to prove the mean ergodic theorem combined with Weyl's estimates on trigonometric sums. The paper [12] by M. Lacey and the paper [18] by J-P Thouvenot give simplified accounts of Bourgain's method.

But what happens for other functions such as $a(x) = x^{3/2}$ or $a(x) = x^2 \cdot \log x$? In general, what conditions on the function $a(x)$ guarantee that $(a(n))$ or $(\lfloor a(n)\rfloor)$ is (norm) good?

Let us immediately make the following remarks

- There appears to be no simple (say, monotonicity) condition on the derivatives of $a(x)$ that can characterize those $a(x)$ for which $(a(n))$ or $(\lfloor a(n)\rfloor)$ is good.
- In our context, there is no essential difference between the sequences $(a(n))$ and $(\lfloor a(n)\rfloor)$ regarding norm convergence. For example, both sequences are norm good when $a(x) = x^{3/2}$. The theorems used to prove these facts are essentially identical.
- There is a fundamental difference between the sequences $(a(n))$ and $(\lfloor a(n)\rfloor)$ for pointwise convergence. For example, the sequence $(n^{3/2})$ in not good, while one of our results is that the sequence $(\lfloor n^{3/2}\rfloor)$ is pointwise good.
- A function $a(x)$ which goes to infinity faster than any polynomial is usually very difficult to handle. For example, if $a(x) = e^{\log^{4/3} x}$ nobody knows if the sequence $(a(n))$ is uniformly distributed $\mod 1$, let alone whether it is norm good.

Because of the last remark, we make the restriction that our function $a(x)$ is subpolynomial in the sense that for some positive integer k, the ratio $a(x)/x^k$ goes to 0 as x goes to ∞.

As for the first remark above, recall that the traditional way of specifying conditions for the function $a(x)$ to guarantee that the sequence $(a(n))$ is uniformly distributed $\mod 1$ is to give estimates on higher order derivatives of $a(x)$. A typical example is the classical Fejér-Van der Corput estimates; for a taste, see Lemma 7.5. Then, for a given function $a(x)$ one has to check these often very complicated conditions. This checking is done in an *ad hoc* manner, so one checks the conditions for, say, $x^3/\log x$ then for $x^3/\log\log x$, then for $x^3/(\log x/\log\log x)$, etc. Also this checking always involves taking higher order derivatives for each function $a(x)$. For illustration of this, see [15, Exercises 2.23-25, 3.9-15].

Our idea to avoid this problem is to consider only a special class of functions which nevertheless includes a lot of "interesting" functions, and provide simple conditions which are also easy to check for this class of functions. Specifically, the conditions for a sequence to be good are that the functions are essentially equal to polynomials or are sufficiently far away from polynomials. In the case that the function are essentially equal to polynomials, the claims hold using the results of Bourgain [6]. In this paper, we provide proofs in the case that the functions are far from polynomials.

In the next section, we give the precise definition of the class of functions we consider. To illustrate our ideas here, let us consider a subclass of these functions, namely the so-called *logarithmico-exponential* functions of Hardy.

These are all the functions one can get by combining real constants, the variable x, and the symbols exp, log, $\cdot$, $+$. This class certainly includes all rational functions, the log and exp functions. But it also includes $\sqrt{x}$ since $\sqrt{x} = \exp(1/2 \cdot \log x)$ or the function

$$\sqrt{2}x^3 - \frac{x^2 \cdot \log x}{\log\log x \cdot e^{\sqrt{\log x}}}. \tag{1.1}$$

Let us denote the class of logarithmico-exponential functions by L. Once we restrict the $a(x)$ to be from L, transparent characterizations of functions $a(x)$ for which both $(a(n))$ and $(\lfloor a(n) \rfloor)$ is norm good is possible. Let $C\mathbb{Q}[x]$ denote the set of real constant multiples of rational polynomials.

Theorem A. *Suppose $a(x)$ is a logarithmico-exponential function, it is subpolynomial, and $a(x)/x \to \infty$ as $x \to \infty$. Then either of $(a(n))$ and $(\lfloor a(n) \rfloor)$ is norm good if and only if*

(1) *There exists $p \in C\mathbb{Q}(x)$ such that $a(x) - p(x) \to K$; or*
(2) *For all $p \in C\mathbb{Q}(x)$, $|a(x) - p(x)|/\log x \to \infty$.*

The reader now *easily* checks that if $a(x)$ is any of the functions

$$x^\pi, \quad x^7/\log x, \quad \sqrt{3} \cdot x^5 - \log x \cdot \log\log\log x, \quad \sqrt{3}x^5 - \log\log x + x$$

then $(a(n))$ and $(\lfloor a(n) \rfloor)$ are norm good, and if $a(x)$ is any of the functions

$$\sqrt{3}x^5 - \sqrt{\log x} \cdot e^{\sqrt{\log\log x}}, \quad x^2 \cdot \sqrt{1 + \frac{\log x}{x^2}}$$

then $(a(n))$ and $(\lfloor a(n) \rfloor)$ are not norm good. We leave it to the reader to decide whether $(\lfloor a(n) \rfloor)$ is norm good if $a(x)$ is the function in (1.1) ....

Now the situation for pointwise convergence is more complicated, although again the distance of the function $a(x)$ from certain polynomials seems to be the key to deciding whether $(\lfloor a(n) \rfloor)$ is good. First we have

Theorem B. *Suppose $a(x)$ is a logarithmico-exponential function, it is subpolynomial, and $a(x)/x \to \infty$. Then $(\lfloor a(n) \rfloor)$ is pointwise good for convergence of L^2 functions if for all real polynomials p, we have either*

(1) *$a(x) - p(x) \to K$; or*
(2) *for some $\epsilon > 0$, $(a(x) - p(x))/\log^{q(p)+\epsilon} x \to \infty$, where the exponent $q(p)$ depends only on the degree ∂p and is given by $q(p) = 2^{\partial p+1} - 1$.*

So this theorem discusses sufficient conditions for pointwise convergence, and these conditions are in terms of the distance from *all* real polynomials while the conditions for norm convergence were in terms of the distance of $a(x)$ from constant multiples of rational polynomials.

As for necessary conditions, we have

Theorem C. *Suppose $a(x)$ is a logarithmico-exponential function, and that for some polynomial $p \in C\mathbb{Q}(x)$ we have*

- $a(x) - p(x) \to \infty$*; and*
- *for some* $0 < \epsilon < 1$, $a(x) - p(x)/(\log x\ e^{(\log\log x)^{\epsilon}}) \to 0$.

Then $(\lfloor a(n) \rfloor)$ *is pointwise bad.*

To highlight one of differences between Theorem C and Theorem A, note that while $\sqrt{3}x^3 + \log x \cdot \log\log x$ is norm good, it is not pointwise good.

To address the second problem with Theorem B, the issue is that it does not say anything about functions which are close to polynomials not in $C\mathbb{Q}(x)$. As an example, take $a(x) = \sqrt{2}x^2 + x + \log x$. According to Theorem A the sequence $(\lfloor a(n) \rfloor)$ is norm good.

In the paper, we address this problem in detail. Since it would be difficult to describe the results without further definitions, in this introductory section, let us just make the following remarks

- If $a(x) = \sqrt{2}x^2 + x + \log x$ then $(\lfloor a(n) \rfloor)$ is good pointwise;
- For some irrational θ, if $a(x) = \theta x^2 + x + \log x$ then $(\lfloor a(n) \rfloor)$ is not good pointwise.

These results imply that it is not possible to characterize those $a(x)$ for which $(\lfloor a(n) \rfloor)$ is pointwise good in terms of the distance from contant multiples of rational polynomials. Our results indicate that instead, one should try a characterization based on the distance from polynomials whose coefficients are well approximated by rationals. One of the main questions we leave open is the exact description of this set of polynomials.

2. Definitions and Notation

Definition 2.1. Denote by B the set of germs at $+\infty$ of continuous, real functions of a real variable x. (A germ is an equivalence class of functions that are identical for large x.) Endowed with pointwise multiplication and addition B forms a ring.

A subfield of B that is closed under differentiation is called a *Hardy field.* The union of all Hardy fields is denoted by U.

On B we define the relation $\trianglelefteq$ as follows: $a \trianglelefteq b$ iff $a(x) \le b(x)$ eventually.

It turns out that every subfield (and hence a Hardy field) is totally ordered by $\trianglelefteq$. Since each subfield of B has to contain $\mathbb{Q}$, the field of rational numbers, it follows that if a is in some subfield of B then $\lim_{x\to\infty} a(x)$ exists (possibly equal to $\pm\infty$).

The class U is large enough to include the class L of logarithmico-exponential functions mentioned above. Further, given any Hardy field F, and letting $\bar{F}$ be a maximal Hardy field extension of F (by inclusion), it may be shown that $L \subset \bar{F}$. In fact, more is true: for example, the function $\operatorname{li} x = \int_2^x 1/\log x\, dx$ is also an element of $\bar{F}$.

Given two functions f and g belonging to a Hardy field, we write $f \prec g$ or $f \preceq g$ respectively if $\lim_{x\to\infty} f(x)/g(x) = 0$ or $\lim_{x\to\infty} f(x)/g(x) < \infty$.

Definition 2.2. A function a from U will be called *subpolynomial* if $a(x) \prec x^k$ for some k.

Definition 2.3. A function r from U will be called *non-polynomial* if it is subpolynomial and for each $k \in \mathbb{Z}^+$, either $r(x) \prec x^k$ or $r(x) \succ x^k$.

Definition 2.4. A sequence $(a(n))$ of integers will be called *universally good for norm convergence of L^p functions* if for each probability measure-preserving system $(\Omega, \mathcal{B}, \mu, T)$ and $f \in L^p(\mu)$,

$$\frac{1}{t} \sum_{n \leq t} f \circ T^{a(n)}$$

is convergent in L^p. The definition may be extended in the obvious way to flows in the case where $(a(n))$ is a sequence of reals.

Definition 2.5. A sequence $(a(n))$ of integers will be called *universally good for pointwise convergence of L^p functions* if for each probability measure-preserving system $(\Omega, \mathcal{B}, \mu, T)$ and $f \in L^p(\mu)$,

$$\frac{1}{t} \sum_{n \leq t} f(T^{a(n)}(\omega))$$

is convergent for μ-almost every $\omega \in \Omega$.

Since we shall mainly be dealing with functions belonging to L^2, we shall sometimes use the term *pointwise good* to mean universally good for pointwise convergence of L^2 functions.

We observe that if a sequence is good for pointwise convergence of L^∞ functions then it is automatically good for norm convergence. To prove this, note that if we write $A_t f = \frac{1}{t} \sum_{n \leq t} f \circ T^{a(n)}$, then we are trying to prove $A_t f$ is Cauchy for $f \in L^2$. Observe that if $|g| \leq M$, then $|A_t g| \leq M$ and by the bounded convergence theorem, pointwise convergence of $A_t g$ implies L^2 convergence of $A_t g$. For $f \in L^2$, let $f = g + h$, where $|g| \leq M$ and $\|h\|_2 < \epsilon$. Then $\|A_t h\|_2 \leq \epsilon$ for all t and $A_t g$ is a Cauchy sequence under the L^2 norm so it follows that for t and t' sufficiently large, $\|A_t f - A_{t'} f\|_2 \leq 2\epsilon$. Since ϵ is arbitrary, it follows that $A_t f$ is Cauchy in L^2 as required.

A similar argument shows that a sequence is universally good for norm convergence of L^2 functions, if and only if it also it is universally good for norm convergence of L^p functions (for any $1 \leq p < \infty$). Consequently, we will omit reference to the space and speak simply of a sequence being *universally good for norm convergence* or *norm good*.

Remark 2.6. One can see that if a sequence is norm good and good for pointwise convergence of L^p functions, then the limits must agree.

In this language, the Birkhoff ergodic theorem states that the sequence of natural numbers is universally good for pointwise convergence of L^1 functions and the Von Neumann ergodic theorem states that the sequence of natural numbers are universally good for norm convergence. These theorems also identify the common limit for a function f as $\mathbf{E}_\mu(f|\mathcal{I})$, where $\mathcal{I}$ is the σ-algebra of measurable T-invariant subsets of Ω. In particular, if the measure μ is ergodic with respect to T, the limit is just $\int f \, d\mu$. This motivates the following definition.

Definition 2.7. A sequence $(a(n))$ is called *ergodic* if for each $f \in L^2$, the L^2-limit of $\frac{1}{t}\sum_{n\le t} f\circ T^{a(n)}$ is $\mathbf{E}_\mu(f|\mathcal{I})$.

Definition 2.8. We denote by $C\mathbb{Q}[x]$ the set of all real multiples of rational polynomials: $C\mathbb{Q}[x] = \{\alpha p(x)\colon \alpha \in \mathbb{R},\, p(x) \in \mathbb{Q}[x]\}$. Similarly, $C\mathbb{A}[x]$ denotes the set of all real multiples of polynomials with algebraic coefficients.

3. Statement of Results

In this paper, we mostly consider sequences of the form $(\lfloor a(n)\rfloor)$ where a is a subpolynomial member of U and conditions for the sequence to be norm and pointwise good and bad.

The remainder of the paper is organized as follows: In Section 4, a number of preparatory lemmas including Lemma 3.1 below are proved. Section 5 contains a key lemma concerning exponential sums which deals with the comparison between exponential sums containing terms of the form $e(a(n)\beta)$ and $e(\lfloor a(n)\rfloor\beta)$. Section 6 concerns norm convergence and contains the proofs of Theorem 3.2 and 3.3. Section 7 contains the results establishing pointwise convergence: Theorems 3.4, 3.5 and 3.8, while Section 8 contains results demonstrating the failure of pointwise convergence: Theorems 3.6, 3.7 and 3.9. The paper concludes with sections on open problems in the area and bibliographic notes.

Lemma 3.1. *Let $a \in U$ be subpolynomial. Then a may be uniquely expressed as $p + r$ where p is a polynomial, r is non-polynomial and satisfies $r(x) \prec x^s$ for each non-zero term $c_s x^s$ of p.*

Theorem 3.2. *Let $a \in U$ be subpolynomial and satisfy $a(x) \succ x$. The sequences $(a(n))$ and $(\lfloor a(n)\rfloor)$ are norm good if and only if*

1. *There exists $p \in C\mathbb{Q}(x)$ such that $a(x) - p(x) \to K$; or*
2. *For all $p \in C\mathbb{Q}(x)$, $|a(x) - p(x)| \succ \log x$.*

In the latter case, $(\lfloor a(n)\rfloor)$ is ergodic.

Theorem 3.3. *Let $a \in U$ be of the form $a(x) = cx + r(x)$ where $r(x) \prec x$ and $c > 0$. Then the sequence $(\lfloor a(n)\rfloor)$ is norm good if and only if one of the following holds:*

1. *$r(x) \succ \log x$ as $x \to \infty$.*
2. *$c = 1/m$;*
3. *$r(x) \to K$ as $x \to \infty$;*

In cases 1 and 2, $(\lfloor a(n)\rfloor)$ is ergodic.

In the case of a sequence of the form $(a(n))$ where $a(x) = cx+r(x)$, the conditions for norm convergence are the same as those in Theorem 3.2.

Theorem 3.4. *Let $a \in U$ be subpolynomial and have decomposition $p + r$. If there is an ϵ such that $r(x) \succ x^\epsilon$, then $(\lfloor a(n)\rfloor)$ is pointwise good and is ergodic.*

Theorem 3.5. *Let $a \in U$ be subpolynomial and have a decomposition $p + r$ where $\partial p = n \ge 2$. If for some $\epsilon > 0$, $r(x) \succ (\log x)^{2^{n+1}-1+\epsilon}$ then $(\lfloor a(n)\rfloor)$ is pointwise good and is ergodic.*

In the case of Theorems 3.4 and 3.5, it follows from Theorem 3.2 that the sequences are ergodic.

Theorem 3.6. *Let $p \in C\mathbb{Q}(x)$. Then if $1 \prec r(x) \prec \log x \exp((\log\log x)^m)$, for some $0 \le m < 1$, we have $(\lfloor p(n) + r(n) \rfloor)$ is pointwise bad.*

Theorem 3.7. *Let p be a polynomial in $C\mathbb{Q}(x)$ whose lowest order term is of degree $k \ge 3$. Then if $1 \prec r(x) \prec (\log x)^m$, where $m < 2k/(k+2)$, then $(\lfloor p(n) + r(n) \rfloor)$ is pointwise bad.*

It will be noticed that in the above four results, giving two sufficient conditions and two necessary conditions for $(\lfloor a(n) \rfloor)$ to be pointwise good, the sufficient conditions have a different character from the necessary conditions: The sufficient conditions deal with functions which are 'sufficiently far from any polynomial', whereas the necessary conditions deal with functions which are 'sufficiently far from any polynomial in $C\mathbb{Q}(x)$'. This contrasts with the situation for norm convergence, where the necessary and sufficient conditions dealt with 'distance' from polynomials in $C\mathbb{Q}(x)$. The next two theorems show that this situation is not artificial.

We first need the following definition: A number α is said to be *badly approximable by rationals* if there exist an $m \ge 2$ and a $C > 0$ such that any rational p/q satisfies $|\alpha - p/q| \ge C/q^m$. We note that the numbers which are badly approximable by rationals form a set of full measure.

Theorem 3.8. *Let $a \in U$ be subpolynomial, satisfy $a(x) \succ x$ and have a decomposition $p+r$ where p is a polynomial with the property that the ratio of two of its non-constant coefficients is badly approximable by rationals (e.g. $p \in C\mathbb{A}[x] \setminus C\mathbb{Q}[x]$). Then $(\lfloor a(n) \rfloor)$ is pointwise good and is ergodic.*

Theorem 3.9. *Let $p \in C\mathbb{Q}(x)$. Then any of the coefficients of p may be perturbed by an arbitrarily small irrational amount to give a polynomial $\tilde{p}$ such that $(\lfloor \tilde{p}(n) + \log n \rfloor)$ is pointwise bad.*

4. Preliminary Lemmas

Lemma 4.1. *Suppose $a \in U$ satisfies $1 \prec a(x) \prec x^k$. Then there is a $\delta \ge 0$ such that for every $\epsilon > 0$, $x^{\delta-\epsilon} \prec a(x) \prec x^{\delta+\epsilon}$.*

Proof. Set $\delta = \inf\{r : a(x) \prec x^r\}$. □

In establishing sufficient conditions for the sequence $(\lfloor a(n) \rfloor)$ to be pointwise good, we will later make use of Van der Corput's lemma where an estimate is made in terms of the derivatives of a. In the lemma below, we use the properties of Hardy fields to control the derivatives of a, where a is a function belonging to a Hardy field.

Lemma 4.2. *Suppose $a \prec b \in U$ are non-polynomial then a' and b' are non-polynomial and $a' \prec b'$. In particular, if $x^{\delta-\epsilon} \prec a(x) \prec x^{\delta+\epsilon}$ for every $\epsilon > 0$. Then for each $k \in \mathbb{N}$ and $\epsilon > 0$, $x^{\delta-k-\epsilon} \prec a^{(k)}(x) \prec x^{\delta-k+\epsilon}$.*

Proof. First note that if a' fails to be non-polynomial (i.e. $a'(x)/x^k$ converges to a finite non-zero limit for some $k \in \mathbb{Z}^+$), then a also fails to be non-polynomial. This shows that if $a \in U$ is non-polynomial, then so is a'.

If $a'(x) \succeq b'(x)$, then integrating, we see that $a(x) \succeq b(x) + C$ for some constant C. Since $a(x)$ and $b(x)$ are non-polynomial, it follows that $a(x) \succeq b(x)$ which is a contradiction. It follows that $a'(x) \prec b'(x)$ as required.

Clearly we can apply the lemma repeatedly to deduce that if $x^{\delta-\epsilon} \prec a(x) \prec x^{\delta+\epsilon}$ for every $\epsilon > 0$, then for each $k \in \mathbb{N}$ and $\epsilon > 0$, $x^{\delta-k-\epsilon} \prec a^{(k)}(x) \prec x^{\delta-k+\epsilon}$. □

Proof of Lemma 3.1. Let $a \in U$ be subpolynomial. If a is non-polynomial, the decomposition is $a = 0 + a$. Otherwise, there exists an $n \geq 0$ such that $a(x)/x^n$ does not converge to either 0 or ∞. Since $a(x)/x^n \in U$, it follows that the limit l exists (and is neither 0 nor ∞). Then set $b(x) = a(x) - lx^n$. Clearly $b(x) \prec x^n$. Iterating this construction gives the required decomposition. Uniqueness is clear. □

Lemma 4.3. *Let $h > 0$ be fixed and let $r \in U$ satisfy $r(x) \succ 1$ and define for sufficiently large x*

$$R(x) = \frac{r^{-1}(x) - r^{-1}(x-h)}{r^{-1}(x)}.$$

(1) *If for some $a \geq 1$, $r(x) \preceq (\log x)^a$, then $R(x) \succeq x^{\frac{1}{a}-1}$.*
(2) *If for some $0 \leq a < 1$, $r(x) \preceq \log x \exp((\log\log x)^a)$, then we have the bound $R(x) \succeq 1/\exp((\log x)^a)$.*

Proof. We first deal with the case $a = 1$ of statement 1. The hypothesis implies that there exists a C such that for sufficiently large x, $r(x) \leq C \log x$. Using L'Hôpital's rule, we see that $xr'(x) \leq C$. Since $(\log r^{-1})'(x) = 1/(r^{-1}(x) r'(r^{-1}(x)))$, we see that for sufficiently large x, $(\log r^{-1})'(x) \geq 1/C$. It follows that $r^{-1}(x-h)/r^{-1}(x) \leq e^{-h/C}$ giving

$$\frac{r^{-1}(x) - r^{-1}(x-h)}{r^{-1}(x)} \geq 1 - e^{-h/C} \succeq 1$$

as required.

To complete the proof of statement 1, we let $a > 1$ and set $s(x) = r(x)^{1/a}$. Then s^{-1} satisfies $(\log s^{-1})' \geq 1/C$ by the above. We note that $r^{-1}(x) = s^{-1}(x^{1/a})$ so

$$\frac{r^{-1}(x-h)}{r^{-1}(x)} = \frac{s^{-1}\left((x-h)^{1/a}\right)}{s^{-1}\left(x^{1/a}\right)} \leq \exp\left(\frac{(x-h)^{1/a} - x^{1/a}}{C}\right).$$

We then see

$$\frac{r^{-1}(x) - r^{-1}(x-h)}{r^{-1}(x)} \geq 1 - \exp\left(\frac{(x-h)^{1/a} - x^{1/a}}{C}\right) \geq (x^{1/a} - (x-h)^{1/a})C \geq Cx^{\frac{1}{a}-1},$$

where C denotes at each stage a constant which is independent of x (possibly different constants at different stages of the proof).

To prove statement 2, we note that the case $a = 0$ follows from the above. Suppose $0 < a < 1$ and set $s(x) = r(x)/\exp((\log r(x))^a)$ so that $s = \phi \circ r$, where $\phi(x) = x/\exp((\log x)^a)$. We note that $s(x) \preceq \log x$ so by the first part of the proof, $(\log s^{-1})' \geq 1/C$. We then argue as above: $r^{-1}(x) = s^{-1}(\phi(x))$ so

$$\frac{r^{-1}(x-h)}{r^{-1}(x)} = \frac{s^{-1}\left((x-h)/(\exp(\log(x-h))^a)\right)}{s^{-1}\left(x/(\exp(\log x)^a)\right)} \leq \exp\left(\frac{1}{C}\left(\frac{x-h}{\exp((\log(x-h))^a)} - \frac{x}{\exp((\log x)^a)}\right)\right).$$

Finally,

$$\begin{aligned}\frac{r^{-1}(x)-r^{-1}(x-h)}{r^{-1}(x)} &\geq 1-\exp\left(\frac{1}{C}\left(\frac{x-h}{\exp((\log(x-h))^a)}-\frac{x}{\exp((\log x)^a)}\right)\right)\\ &\geq \left(\frac{x-h}{\exp((\log(x-h))^a)}-\frac{x}{\exp((\log x)^a)}\right)C \geq \frac{C}{\exp((\log x)^a)}.\end{aligned}$$

□

Lemma 4.4. *Let (f_n) be a sequence of nonnegative numbers. For $\rho \in \mathbb{Q}$ with $\rho > 1$ denote $I_\rho = \{\rho^l : l \in \mathbb{N}\}$. Suppose that for each $\rho \in \mathbb{Q}$ with $\rho > 1$ the averages $\frac{1}{t}\sum_{n\leq t} f_n$ converge to some finite limit as t runs through the sequence I_ρ.*

Then $\frac{1}{t}\sum_{n\leq t} f_n$ is convergent as $t \to \infty$.

Proof. Set $F_t = \frac{1}{t}\sum_{n\leq t} f_n$ and let $\rho \in \mathbb{Q}$ satisfy $\rho > 1$. By assumption, $\frac{1}{t}\sum_{n\leq t} f_n$ is convergent to some finite limit, L say, as t runs through the sequence I_ρ.

For an arbitrary $t > 1$ let l be the positive integer such that $\rho^{l-1} \leq t < \rho^l$. Since the sequence (f_n) is nonnegative, we can estimate

$$\frac{\rho^{l-1}}{t}F_{\rho^{l-1}} \leq F_t \leq \frac{\rho^l}{t}F_{\rho^l}. \tag{4.1}$$

We see then that $L/\rho \leq \liminf_{t\to\infty} F_t \leq \limsup_{t\to\infty} F_t \leq L\rho$. If $L = 0$, then we have $\liminf_{t\to\infty} F_t = \limsup_{t\to\infty} F_t = 0$ and we are done. Otherwise, we see that $\limsup_{t\to\infty} F_t / \liminf_{t\to\infty} F_t \leq \rho^2$. Since ρ can be chosen arbitrarily close to 1, the desired conclusion follows. □

5. A key inequality

In this section, we prove an inequality which will be of key importance in the later sections where we establish pointwise convergence. We start with a lemma estimating the Fourier coefficients of certain continuous versions of step functions.

Lemma 5.1 (Vinogradov). *Let c, d, Δ be real numbers satisfying*

$$0 < \Delta < 1/2, \qquad \Delta \leq d - c \leq 1 - \Delta.$$

Let the 1-periodic function ϕ satisfy $\phi(\alpha) = 1$ for $c+\Delta/2 \leq \alpha \leq d-\Delta/2$; $\phi(\alpha) = 0$ for $d + \Delta/2 \leq \alpha \leq 1 + c - \Delta/2$; and ϕ is the linear interpolation between these values on the remainder of the real line.

Then the Fourier series $\phi(\alpha) = d - c + \sum_{|\ell|\geq 1} c_\ell e(\ell\alpha)$ of ϕ satisfies

$$|c_\ell| \leq 2\min\{d-c, \frac{1}{|\ell|}, \frac{1}{\ell^2\Delta}\}.$$

The proof of this lemma is a straightforward (though tedious) computation, and is safely left for the reader.

Theorem 5.2. *Suppose that the numbers t, Q, r, S are all greater than 10 and $r^2 \leq S - 1$. Let $b(1), b(2), \ldots, b(\lfloor t\rfloor)$ be real numbers. Define $U(\alpha)$ by*

$$U(\alpha) = \frac{1}{t}\left|\sum_{n\leq t} e(b(n)\alpha)\right|. \tag{5.1}$$

Suppose finally that

$$U(\alpha) \leq \frac{1}{r} \text{ for each } 1/Q \leq |\alpha| \leq S. \tag{5.2}$$

Then there is an absolute constant C (< 100) so that for every β with $1/Q \leq |\beta| \leq 1/2$ we have

$$\left|\frac{1}{t}\sum_{n\leq t} e(\lfloor b(n)\rfloor \beta)\right| \leq C\sqrt{\frac{\log r}{r}}. \tag{5.3}$$

Before embarking upon a proof of the theorem, we first prove a lemma which will be used in two places giving some estimates based on Fourier analysis.

Lemma 5.3. *There exists a constant C such that for any sequence $(b(n))$ satisfying the conditions of Theorem 5.2 and any piecewise linear function ϕ as in Lemma 5.1 with $\Delta^{-2} \leq S - 1$ one has for each $|\beta| \leq 1/2$,*

$$S(\beta, \phi) < U(\beta)\|\phi\|_1 + C \cdot \left(\frac{\log \Delta^{-2}}{r} + \Delta\right), \tag{5.4}$$

where

$$S(\beta, \phi) = \left|\frac{1}{t}\sum_{n\leq t} e(\beta b(n))\phi(b(n))\right|.$$

Proof. Since ϕ is piecewise smooth and continuous, it has an absolutely convergent Fourier series with coefficients satisfying the conclusions of Lemma 5.1. We have

$$\begin{aligned} S(\beta,\phi) &= \left|\frac{1}{t}\sum_{n\leq t} e(b(n)\beta)\sum_{\ell} c_\ell e(\ell b(n))\right| = \left|\sum_{\ell} c_\ell \frac{1}{t}\sum_{n\leq t} e\left((\beta+\ell)b(n)\right)\right| \\ &\leq \sum_{\ell} |c_\ell| \left|\frac{1}{t}\sum_{n\leq t} e\left((\beta+\ell)b(n)\right)\right| = \sum_{\ell} |c_\ell| U(\beta+\ell) \\ &\leq |c_0|U(\beta) + \sum_{1\leq |l|\leq \Delta^{-2}} |c_\ell|\cdot U(\beta+\ell) + \sum_{|l|>\Delta^{-2}} |c_\ell| U(\beta+\ell). \end{aligned}$$

In the first summation, by (5.2), the $U(\beta+\ell)$ terms may be bounded above by $1/r$ and by Lemma 5.1, the $|c_\ell|$ terms may be bounded above by $1/\ell$. In the second summation, we have the trivial bound $U(\beta+\ell) \leq 1$ and $|c_\ell|$ can be bounded above by $1/(\Delta\ell^2)$.

Putting this together, we see

$$\begin{aligned} S(\beta,\phi) &\leq \|\phi\|_1 U(\beta) + \sum_{1\leq |l|\leq \Delta^{-2}} \frac{1}{lr} + \sum_{|l|>\Delta^{-2}} \frac{1}{\Delta\ell^2} \\ &\leq \|\phi\|_1 U(\beta) + C\left(\frac{\log\Delta^{-2}}{r} + \Delta\right) \end{aligned}$$

This proves the lemma. □

Proof of Theorem 5.2. Let us fix $\beta \in [-1/2, 1/2)$, $1/Q \leq |\beta|$. We will find a connection between the sums on the left of the inequalities (5.1) and (5.3) by describing the location of $\langle b(n)\rangle$ [1] in the interval $[0, 1)$. Let q be a natural number to be specified later, and divide the interval $[0, 1)$ into q equal parts. Let us denote $I_j = [(j-1)/q, j/q)$, $j = 1, 2, \ldots, q$. We then let ϕ_j be the piecewise linear function described in Lemma 5.1 where c is taken to be $(j-1)/q$ and $d = j/q$. The

[1] Here, $\langle x\rangle$ denotes the fractional part of x, that is, $\langle x\rangle = x - \lfloor x\rfloor$.

constant Δ will be chosen later so as to optimize certain upper bounds (and will in fact be given by $1/r$). In particular, we will have $\Delta < 1/q$. The function ϕ_j is then a smoothed version of the characteristic function of I_j. It can be checked that $\sum_j \phi_j \equiv 1$. Set $\psi = \phi_q + \phi_1$. We then have

$$\frac{1}{t}\sum_{n\le t} e(\lfloor b(n)\rfloor\beta) = \sum_{1<j<q}\frac{1}{t}\sum_{n\le t} e(\lfloor b(n)\rfloor\beta)\phi_j(b(n)) + \frac{1}{t}\sum_{n\le t} e(\lfloor b(n)\rfloor\beta)\psi(b(n)).$$

Notice that for $1 < j < q$ we have $\phi_j(x) > 0 \Rightarrow \lfloor x\rfloor = x - j/q + \theta/q$, where $|\theta| \le 2$ and therefore, we can write for $1 < j < q$,

$$\begin{aligned}\frac{1}{t}\sum_{n\le t} e(\lfloor b(n)\rfloor\beta)\phi_j(b(n)) &= \frac{1}{t}\sum_{n\le t} e(b(n)\beta - j/q\beta + \theta/q)\phi_j(b(n)) \\ &= e(-j/q\beta)\frac{1}{t}\sum_{n\le t} e(b(n)\beta + \theta/q)\phi_j(b(n)).\end{aligned}$$

where $\theta = \theta_{j,n}$ satisfies $|\theta| \le 2$.

Now we see

$$\begin{aligned}&\left|\frac{1}{t}\sum_{n\le t} e(b(n)\beta + \theta/q)\phi_j(b(n)) - \frac{1}{t}\sum_{n\le t} e(b(n)\beta)\phi_j(b(n))\right| \\ &= \left|\frac{1}{t}\sum_{n\le t} e(b(n)\beta)(1 - e(\theta/q))\phi_j(b(n))\right| \\ &\le C\cdot\frac{2}{q}\cdot\frac{1}{t}\sum_{n\le t}\phi_j(b(n)),\end{aligned}$$

where for the inequality, we used the fact that $|1 - e(\gamma)| \le C|\gamma|$. Since

$$\sum_{1<j<q}\frac{1}{t}\sum_{n\le t}\phi_j(b(n)) \le \frac{1}{t}\sum_{n\le t} 1 = 1,$$

we see on adding up the above that

$$\begin{aligned}&\left|\frac{1}{t}\sum_{n\le t} e(\lfloor b(n)\rfloor\beta)\right| \\ &\le \sum_{1<j<q}\left|\frac{1}{t}\sum_{n\le t} e(b(n)\beta)\phi_j(b(n))\right| + \left|\frac{1}{t}\sum_{n\le t} e(\lfloor b(n)\rfloor\beta)\psi(b(n))\right| + \frac{C}{q} \\ &\le \sum_{1<j<q} S(\beta,\phi_j) + \frac{1}{t}\sum_{n\le t}\psi(b(n)) + \frac{C}{q} \\ &\le \sum_{1<j<q} S(\beta,\phi_j) + S(0,\psi) + \frac{C}{q}\end{aligned}$$

Using Lemma 5.3, we see that $S(\beta,\phi_j)\le 1/qr+C(\log\Delta^{-2}/r+\Delta)$ and $S(0,\psi)\le 2/q+C(\log\Delta^{-2}/r+\Delta)$. Combining this, we see that

$$\left|\frac{1}{t}\sum_{n\le t}e(\lfloor b(n)\rfloor\beta)\right|\le\frac{1}{r}+\frac{2}{q}+Cq\left(\frac{\log\Delta^{-2}}{r}+\Delta\right)+\frac{C}{q}.$$

Then choosing q to be $\lfloor\sqrt{r/\log r}\rfloor$ and Δ to be $1/r^2$, we see that

$$\left|\frac{1}{t}\sum_{n\le t}e(\lfloor b(n)\rfloor\beta)\right|\le C\sqrt{\frac{\log r}{r}}.$$

□

6. Norm convergence

We start off with a simple lemma characterizing those sequences $a(n)$ which are norm good. This lemma appeared as an exercise in [16].

Lemma 6.1. *Let $(b(n))$ be a sequence of real numbers. Then the sequence $(b(n))$ is norm good if and only if for each $\beta\in\mathbb{R}$, $\frac{1}{t}\sum_{n\le t}e(b(n)\beta)$ is convergent as $t\to\infty$.*

In the case where $b(n)$ is an integer sequence, the sequence is ergodic if and only if

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n\le t}e(b(n)\beta)=\begin{cases}1 & \text{if } \beta\in\mathbb{Z}\\ 0 & \text{otherwise}\end{cases}$$

In the case where $b(n)$ is an integer sequence, it is clearly sufficient to consider β belonging to the interval $[0,1)$.

We give the proof in the case of a sequence of integers. The proof for real numbers is similar, but uses the Fourier transform on $\mathbb{R}$ rather that S^1.

Proof. We observe that if T is a measure-preserving transformation of a probability space (Ω,μ) and $f\in L^2(\Omega,\mu)$, then using the spectral theorem,

$$\begin{aligned}&\left\|\frac{1}{t}\sum_{n\le t}f\circ T^{b(n)}-\frac{1}{t'}\sum_{n=1}^{t'}f\circ T^{b(n)}\right\|_2^2\\ &=\left\langle\frac{1}{t}\sum_{n\le t}f\circ T^{b(n)}-\frac{1}{t'}\sum_{n\le t'}f\circ T^{b(n)},\frac{1}{t}\sum_{n\le t}f\circ T^{b(n)}-\frac{1}{t'}\sum_{n\le t'}f\circ T^{b(n)}\right\rangle\\ &=\int\left|\frac{1}{t}\sum_{n\le t}e(b(n)\beta)-\frac{1}{t'}\sum_{n\le t'}e(b(n)\beta)\right|^2 d\mu_f(\beta),\end{aligned}$$

where μ_f is the scalar spectral measure of f. Setting

$$\Lambda_{t,t'}(\beta)=\left|\frac{1}{t}\sum_{n\le t}e(b(n)\beta)-\frac{1}{t'}\sum_{n\le t'}e(b(n)\beta)\right|^2,$$

we see that $\Lambda_{t,t'}(\beta)\le 1$ for each β.

If $(b(n))$ has the property that $\frac{1}{t}\sum_{n\le t} e(b(n)\beta)$ is convergent for each β, then $\Lambda_{t,t'}(\beta) \to 0$ as $t, t' \to \infty$. By the bounded convergence theorem, it then follows that

$$\left\| \frac{1}{t}\sum_{n\le t} f\circ T^{b(n)} - \frac{1}{t'}\sum_{n\le t'} f\circ T^{b(n)} \right\|_2^2 \to 0 \text{ as } t, t' \to \infty.$$

as required.

Conversely if for some β, $\frac{1}{t}\sum_{n\le t} e(b(n)\beta)$ is not convergent, then letting T be the map $T(x) = x + \beta \bmod 1$ of the circle and $f(x) = e(x)$, we see that $(b(n))$ is not norm good.

The condition for ergodicity is a consequence of the spectral theorem. □

The bulk of the work in proving Theorem 3.2 is contained in results of Boshernitzan [3] and Niederreiter [14].

Theorem 6.2 ([3])**.** *Let $a \in U$ be sub-polynomial. Then the sequence $\langle a(n)\rangle$ is uniformly distributed modulo 1 if and only if for every $p \in \mathbb{Q}[x]$, we have*

$$\lim_{x\to\infty} \left| \frac{a(x) - p(x)}{\log x} \right| = \infty.$$

The following is a self-contained version of a theorem which appeared in [14]. The proof resembles that of Theorem 5.2.

Theorem 6.3. *If $a(n)$ is a sequence of real numbers with the property that for each number $\alpha \neq 0$, $\langle \alpha a(n)\rangle$ is uniformly distributed modulo 1, then $(\lfloor a(n)\rfloor)$ is norm good and is ergodic.*

Proof. Let $\theta \notin \mathbb{Z}$ and $0 < \epsilon < 1$ be given. We want to show that for all sufficiently large t,

$$\left| \frac{1}{t}\sum_{n\le t} e(\lfloor a(n)\rfloor\theta) \right| < \epsilon.$$

Let A_t denote the left hand side of the inequality. We then choose an integer $M > 0$ such that $|e(2\theta/M) - 1| < \epsilon/3$ and choose $\delta < \epsilon/6M$. Now define $\phi_j(x)$ to be the function on the circle such that ϕ_j takes the value 1 on $[\frac{j}{M} + \delta, \frac{j+1}{M} - \delta]$, 0 outside $[\frac{j}{M} - \delta, \frac{j+1}{M} + \delta]$ and is linear in between. We note that $\phi_0 + \ldots + \phi_{M-1} = 1$.

Now we have

$$A_t = \left| \frac{1}{t}\sum_{n\le t}\sum_{j<M} \phi_j(\langle a(n)\rangle) e(\lfloor a(n)\rfloor\theta) \right|.$$

Since $\phi_j(\langle a(n)\rangle) > 0$ implies $|\lfloor a(n)\rfloor - (a(n) - j/M)| < 2/M$, we see that in this case, $|e(\lfloor a(n)\rfloor\theta) - e((a(n) - j/M)\theta)| < \epsilon/3$. It follows that we have

$$A_t \le \epsilon/3 + \left| \frac{1}{t}\sum_{n\le t}\sum_{j<M} \phi_j(\langle a(n)\rangle) e((a(n) - j/M)\theta) \right|.$$

We now approximate the functions $\phi_j(x)$ by trigonometric polynomials $\psi_j(x) = \sum_{|k|<K} b_{j,k}e(kx)$ in such a way that $\|\phi_j - \psi_j\|_\infty < \epsilon/3M$. We then have

$$\begin{aligned} A_t &\le 2\epsilon/3 + \left|\frac{1}{t}\sum_{n\le t}\sum_{j<M}\psi_j(\langle a(n)\rangle)e((a(n)-j/M)\theta)\right| \\ &\le 2\epsilon/3 + \sum_{j<M}\sum_{|k|<K}|b_{j,k}|\left|\frac{1}{t}\sum_{n\le t}e((k+\theta)a(n))\right|. \end{aligned}$$

By hypothesis (since $k+\theta \notin \mathbb{Z}$), we see that all of the terms inside the absolute values converge to 0 as $t\to\infty$. Since there are only finitely many such terms, we conclude that for sufficiently large t, $A_t \le \epsilon$ as required. ☐

Theorem 6.4. *The sequence $(a(n))$ has the property that for each $s \in \mathbb{R}$, and each $f \in C([0,1]\times S^1)$,*

$$\frac{1}{t}\sum_{n\le t} f(\langle a(n)\rangle, \langle sa(n)\rangle) \tag{6.1}$$

is convergent if and only if the following conditions hold:

1. *$(a(n))$ and $(\lfloor a(n)\rfloor)$ are norm good; and*
2. *$\frac{1}{t}\sum_{n\le t}\langle a(n)\rangle$ is convergent.*

Proof. First suppose that for each $s \in \mathbb{R}$ and each $f \in C([0,1]\times S^1)$, the sequence $\frac{1}{t}\sum_{n\le t} f(\langle a(n)\rangle, \langle sa(n)\rangle)$ is convergent. Letting $f(x,y) = e(y)$, we see that $f(\langle a(n)\rangle, \langle sa(n)\rangle) = e(sa(n))$ so from Lemma 6.1, we see that $a(n)$ is norm good. Now taking the function $f(x,y)$ to be $f(x,y) = e(-sx)e(y)$, we see that $f(\langle a(n)\rangle, \langle sa(n)\rangle) = e(sa(n))e(-s\langle a(n)\rangle) = e(s\lfloor a(n)\rfloor)$, thus showing that $(\lfloor a(n)\rfloor)$ is norm good. To prove the second assertion, we take $f(x,y) = x$.

For the converse, fix $s \in \mathbb{R}$. Define

$$\mathcal{F} = \left\{ f \in C([0,1]\times S^1) \colon \frac{1}{t}\sum_{n\le t} f(\langle a(n)\rangle, \langle sa(n)\rangle) \text{ is convergent} \right\}.$$

This is a closed subspace of $C([0,1]\times S^1)$. We aim to show that it is all of $C([0,1]\times S^1)$. We first observe that if $f(x,y) = e(lx)e(my)$, then $f(\langle a(n)\rangle, \langle sa(n)\rangle) = e((l+ms)a(n))$ so since $(a(n))$ is assumed to be norm good, we have $f \in \mathcal{F}$. Since linear combinations of functions of this form are dense in $C(S^1\times S^1)$, it follows that $C(S^1\times S^1) \subset \mathcal{F}$.

Since linear combinations of functions of the form $g(x)e(my)$ form a dense subset of $C([0,1]\times S^1)$, it is sufficient to show that all such functions are contained in $\mathcal{F}$. If $g(0) = g(1)$, this follows from the above.

Fix $m \in \mathbb{Z}$ and suppose that g and g' are two functions in $C[0,1]$ such that $g(0) \ne g(1)$ and $g'(0) \ne g'(1)$. There then exists a $\lambda \ne 0$ such that $g(0)+\lambda g'(0) = g(1)+\lambda g'(1)$. Since $(g(x)+\lambda g'(x))e(my) \in C(S^1\times S^1)$, we see that $g(x)e(my) \in \mathcal{F}$ if and only if $g'(x)e(my) \in \mathcal{F}$. Accordingly, it is sufficient to show for each $m \in \mathbb{Z}$ that there is at least one function $g \in C([0,1])$ satisfying $g(0) \ne g(1)$ such that $g(x)e(my) \in \mathcal{F}$.

If $ms \in \mathbb{Z}$, then taking $g(x) = xe(-msx)$, we see that $g(\langle a(n)\rangle)e(m\langle sa(n)\rangle) = \langle a(n)\rangle e(-msa(n))e(msa(n)) = \langle a(n)\rangle$. By hypothesis 2, we see that $g(x)e(my) \in \mathcal{F}$.

If $ms \notin \mathbb{Z}$, then taking $g(x) = e(-msx)$, we see that $g(\langle a(n)\rangle)e(m\langle sa(n)\rangle) = e(-ms\langle a(n)\rangle)e(msa(n)) = e(ms\lfloor a(n)\rfloor)$. By hypothesis 1, it follows that $g(x)e(my) \in \mathcal{F}$.

□

Corollary 6.5. *Suppose that $(a(n))$ is norm good. If for each $\epsilon > 0$, there is a function g in $C([0,1])$ such that $|g(0) - g(1)| = 1$ and*

$$\limsup_{t\to\infty} \frac{1}{t}\sum_{n\le t} |g(\langle a(n)\rangle)| < \epsilon,$$

then $(\lfloor a(n)\rfloor)$ is norm good.

Proof. Let $s > 0$ be fixed. We define $\mathcal{F}$ as in the proof of Theorem 6.4. Since $(a(n))$ is assumed to be norm good, as shown before $C(S^1 \times S^1) \subset \mathcal{F}$. To show that there is convergence in (6.1) for all $f \in C([0,1] \times S^1)$, it is sufficient to show convergence for functions of the form $f(x,y) = g(x)e(my)$, where $|g(0) - g(1)| = 1$. Further, if $g(0) = g'(0)$ and $g(1) = g'(1)$, then $(g(x) - g'(x))e(my) \in C(S^1 \times S^1)$, so there is automatically convergence in (6.1). It follows that if one lets the quantity $\mathrm{osc}_n(s)$ be defined by the expression

$$\lim_{t\to\infty}\left|\sup_{t'>t}\left|\frac{1}{t}\sum_{n\le t} g(\langle a(n)\rangle)e(m\langle sa(n)\rangle) - \frac{1}{t'}\sum_{n\le t'} g(\langle a(n)\rangle)e(m\langle sa(n)\rangle)\right|\right|,$$

then $\mathrm{osc}_n(s)$ is independent of g provided $|g(0) - g(1)| = 1$. From the hypothesis and the triangle inequality, we see that $\mathrm{osc}_n(s) \le 2\epsilon$ for any $\epsilon > 0$ thus giving convergence for f of the given form in (6.1). The conclusion then follows from Theorem 6.4. □

Proof of Theorem 3.2. Suppose first that for all $p \in \mathbb{Q}[x]$ and $c \ne 0$, $|a(x) - cp(x)| \succ \log x$. Then for $\alpha \ne 0$ and any $p \in \mathbb{Q}[x]$, we observe that $|\alpha a(x) - p(x)| = |\alpha||a(x) - \frac{1}{\alpha}p(x)| \succ \log x$. By Theorem 6.2, $\langle \alpha a(n)\rangle$ is uniformly distributed modulo 1. By Lemma 6.1, $(a(n))$ is norm good. To show that $(\lfloor a(n)\rfloor)$ is norm good, note that since α is an arbitrary non-zero number, $(\lfloor a(n)\rfloor)$ is norm good and ergodic by Theorem 6.3.

Next, consider the case where $a(x) - cp(x) \to K$. If $c \notin \mathbb{Q}$, then $a(n)$ is uniformly distributed modulo 1, so by choosing appropriate g (0 on $[0, 1-\epsilon]$ and linearly increasing to 1 on the remainder of the interval) in Corollary 6.5, to show that $(\lfloor a(n)\rfloor)$ is norm good it is sufficient to show that $(a(n))$ is norm good. Clearly by Lemma 6.1, $a(n)$ is norm good if and only if $cp(n) + K$ is norm good. However, for any polynomial q, $1/t\sum_{n\le t} e(q(n))$ is convergent (to 0 if $q(x)$ has an irrational coefficient of x^k for some $k > 0$ by Weyl's theorem and because the sequence is periodic if all the non-constant coefficients of q are rational). It follows that $1/t\sum_{n\le t} e(\beta(cp(n) + K))$ is convergent for all t so by Lemma 6.1, we see that $(a(n))$ and hence $(\lfloor a(n)\rfloor)$ is norm good.

If $c \in \mathbb{Q}$, then since $r(x) = a(x) - cp(x)$ belongs to U, it is eventually monotonic. It follows that $\langle a(n)\rangle$ eventually takes no values in $[0, \epsilon]$ or $\langle a(n)\rangle$ eventually takes no values in $[1-\epsilon, 1)$. It follows that there exists a constant k such that $|\lfloor a(n)\rfloor -$

$(a(n)+k)| \to 0$ so that by Lemma 6.1, $(\lfloor a(n) \rfloor)$ is norm good if and only if $(a(n))$ is norm good. Arguing as above, we see that $(a(n))$ is norm good and hence $(\lfloor a(n) \rfloor)$ is also norm good.

It remains to examine the case $1 \prec |a(x) - cp(x)| \preceq \log x$ for $c \neq 0$ and $p \in \mathbb{Q}[x]$. By absorbing any denominators into c, we may assume $p \in \mathbb{Z}[x]$. Write $r(x) = a(x) - cp(x)$. We assume that $r(x) \to \infty$, but simple modifications would deal with the case $r(x) \to -\infty$. To demonstrate that $(a(n))$ is not norm good, we note that $e(a(c)/c) = e(r(x)/c)$. Since $r(x) \preceq \log x$, we see that the tail dominates the Césaro averages of $e(r(x)/c)$ showing that the sequence of averages in Lemma 6.1 is not convergent. It follows that $(a(n))$ is not norm good as required. It remains to show that $(\lfloor a(n) \rfloor)$ is not norm good.

For a large number N, let $q = p'(N)$. Then expanding $p(n)$ in powers of $n - N$, we see that $p(N + jq) = p(N) + k(j)q^2$, where $k(j)$ is an integer-valued function. Consider the map $T\colon [0, cq^2) \to [0, cq^2)$ given by $T(x) = x + 1 \bmod cq^2$. This preserves the normalized Lebesgue measure on the interval $[0, cq^2)$. Let $f(x) = \chi_{[0,3)}(x)$.

Let J be any interval $[x, x + 1)$ of length 1 in $[0, cq^2)$. Let s_1 be a number such that $r(s_1) + x + cp(N) = 1 \pmod{cq^2}$ and s_2 be the smallest number larger than s_2 such that $r(s_2) + x + cp(N) = 2 \pmod{cq^2}$. Then for $y \in J$ and a number of the form $N + jq$ with $s_1 < N + jq < s_2$, we see

$$\begin{aligned} T^{\lfloor a(N+jq) \rfloor}(y) &= y + \lfloor a(N + jq) \rfloor \bmod cq^2 \\ &= y + \lfloor cp(N) + k(j)cq^2 + r(N + jq) \rfloor \bmod cq^2 \\ &= y + cp(N) + r(N + jq) - \zeta \bmod cq^2, \end{aligned}$$

where $0 \le \zeta < 1$. It follows that $T^{\lfloor a(N+jq) \rfloor}(y) \in [0, 3)$ so that $f(T^{\lfloor a(N+jq) \rfloor}(y)) = 1$. We see that for $y \in J$,

$$\frac{1}{s_2} \sum_{n \le s_2} f(T^{\lfloor a(N+jq) \rfloor}(y)) \ge \left\lfloor \frac{s_2 - s_1}{q s_2} \right\rfloor .$$

By hypothesis, $1 \prec r(x) \preceq \log x$ so by Lemma 4.3, there is a $C > 0$ such that for any choice of x, $(s_2 - s_1)/s_1 \ge C$. We then see

$$\frac{1}{R_m} \sum_{n \le R_m} f(T^{\lfloor a(n) \rfloor}(y)) \ge \frac{C}{q}$$

for each $y \in J$.

Suppose then for a contradiction that $(1/R_m) \sum_{n \le R_m} f \circ T^{\lfloor a(n) \rfloor}$ is convergent in L^2 to a limit f^*. The limit must be at least C/q on the interval J. But since J is arbitrary, the limit must be at least C/q everywhere. We rule this out however because convergence in L^2 implies convergence in L^1, but $\|f\|_1 = 3/(cq^2)$, while $\|f^*\|_1 \ge C/q$. Since we would have $\|f\|_1 = \|f^*\|_1$, taking N large gives a large value of q and establishes the contradiction. □

We note that the above proof applies verbatim to the case of $(a(n))$ where $a(x) = cx + r(x)$ and $r(x) \prec x$.

Proof of Theorem 3.3. Let $a(x) = cx + r(x)$. If $r(x) \succ \log x$, then using the results of Boshernitzan and Niederreiter as above, we see that $(\lfloor a(n) \rfloor)$ is norm good.

We now show that we have convergence in norm if $c = \frac{1}{m}$ and $r(x) \preceq \log x$. We note the sequence $(\lfloor n/m \rfloor)$ consists of m 0's followed by m 1's etc. By Birkhoff's theorem, $(\lfloor n/m \rfloor)$ is a good averaging sequence for L^1 functions. Since $r(x) \in U$, it is eventually monotonic. If $r(x)$ is eventually increasing, then we see that the sequence $(\lfloor n/m + r(n) \rfloor)$ eventually takes each value at most m times. Since $a(n) = n/m + O(\log n)$, we see that there are at most $O(\log t)$ values between 0 and $a(t)$ which are taken less than m times. The sequences $(\lfloor a(n) \rfloor)$ and $(\lfloor n/m \rfloor)$ differ on a set of density 0 so that for L^∞ functions, $(\lfloor a(n) \rfloor)$ is a good averaging sequence for pointwise convergence (with convergence to the usual ergodic limit). Since these functions are dense in L^2, it follows that $(\lfloor a(n) \rfloor)$ is norm good and is ergodic.

If $r(x)$ converges to a finite limit, we argue as in Theorem 3.2 to show that $(\lfloor a(n) \rfloor)$ is norm good.

To show that the conditions are necessary and sufficient, it remains to show that for any $c > 0$ not of the form $1/m$ and any function r with $1 \prec r(x) \preceq \log x$, $(\lfloor cn + r(n) \rfloor)$ is not norm good.

We deal first with the case $c \notin \mathbb{Q}$. Set $a(n) = cn + r(n)$ where $1 \prec r(n) \preceq \log n$. Then we will show that

$$\frac{1}{t} \sum_{n \le t} e\left(\frac{\lfloor a(n) \rfloor}{c} \right)$$

is not convergent.

Set $\alpha = c \int_0^{1/c} e(x)\, dx$. Since c is not the reciprocal of an integer, it follows that $\alpha \neq 0$. We now choose ϵ so as to ensure that for all sufficiently large t, $|s(\lfloor t + \epsilon t \rfloor) - s(\lfloor t \rfloor)| < |1/(12\pi)|$. Define

$$D_t = \frac{1}{\epsilon t} \sum_{t < n \le t(1+\epsilon)} e\left(\frac{\lfloor a(n) \rfloor}{c} \right).$$

Then we see

$$\begin{aligned} D_t &= \frac{1}{\epsilon t} \sum_{t < n \le t(1+\epsilon)} e(s(n)) e\left(- \frac{cn + r(n) \bmod 1}{c} \right) \\ &= \frac{1}{\epsilon t} \sum_{t < n \le t(1+\epsilon)} e(s(n)) e\left(-(n + s(n)) \bmod \tfrac{1}{c} \right). \end{aligned}$$

Letting $U(n) = \sum_{j \le n} e(-(j + s(j)) \bmod \frac{1}{c})$, we get

$$\begin{aligned} D_t &= \frac{1}{\epsilon t} \sum_{t < n \le t(1+\epsilon)} e(s(n))(U(n) - U(n-1)) \\ &= \frac{1}{\epsilon t} \Big(e(s(\lfloor t + \epsilon t \rfloor + 1)) U(\lfloor t + \epsilon t \rfloor) - e(s(\lfloor t \rfloor + 1)) U(\lfloor t \rfloor) \\ &\quad - \sum_{t < n \le t(1+\epsilon)} \big(e(s(n+1)) - e(s(n)) \big) U(n) \Big) \\ &= \frac{1}{\epsilon} \Big(\frac{U(\lfloor t + \epsilon t \rfloor)}{t + \epsilon t} (1+\epsilon) e(s(\lfloor t + \epsilon t \rfloor + 1)) - \frac{U(\lfloor t \rfloor)}{t} e(s(\lfloor t \rfloor + 1)) \Big) - \\ &\quad - \frac{1}{\epsilon} \sum_{t < n \le t(1+\epsilon)} \big(e(s(n+1)) - e(s(n)) \big) \frac{U(n)}{n} \frac{n}{t} \end{aligned}$$

Since $n+s(n)$ is uniformly distributed modulo $1/c$, we see from Boshernitzan's result that $U(n)/n \to \alpha$ as $n \to \infty$. It follows that

$$\begin{aligned} D_t &= \frac{\alpha}{\epsilon}(e(s(\lfloor t+\epsilon t\rfloor+1)) - e(s(\lfloor t\rfloor+1))) + \alpha e(s(\lfloor t+\epsilon t\rfloor+1)) + o(1) \\ &\quad - \frac{1}{\epsilon}\sum_{t<n\le t(1+\epsilon)} (e(s(n+1)) - e(s(n)))\alpha \\ &\quad + \frac{1}{\epsilon}\sum_{t<n\le t(1+\epsilon)} (e(s(n+1)) - e(s(n)))\left(\alpha - \frac{U(n)}{n}\frac{n}{t}\right) \\ &= \alpha e(s(\lfloor t+\epsilon t\rfloor+1)) + \frac{1}{\epsilon}\sum_{t<n\le t(1+\epsilon)} (e(s(n+1)) - e(s(n)))\left(\alpha - \frac{U(n)}{n}\frac{n}{t}\right) \\ &\quad + o(1) \end{aligned}$$

Let t be sufficiently large that $|U(n)/n - \alpha| \le \alpha\epsilon$ for all $n \ge t$. Since $n/t < 1+\epsilon$, we have for all n in the summation

$$\left|\alpha - \frac{U(n)}{n}\frac{n}{t}\right| \le 3\epsilon\alpha.$$

We see also that $|e(s(n+1)) - e(s(n))| \le 2\pi|s(n+1)-s(n)|$. It follows that the absolute value of the summation is at most $6\pi\alpha|s(\lfloor t+\epsilon t\rfloor+1) - s(\lfloor t\rfloor)|$. By the choice of ϵ, we see

$$|D_t - \alpha e(s(\lfloor t+\epsilon t\rfloor+1))| < \frac{|\alpha|}{2}$$

It follows that D_t is not convergent as $t \to \infty$. Defining

$$A_t = \frac{1}{t}\sum_{n\le t} e\left(\frac{\lfloor a(n)\rfloor}{c}\right),$$

we have

$$A_{t+\epsilon t} - \frac{1}{1+\epsilon}A_t = \frac{\epsilon}{1+\epsilon}D_t$$

We see that it follows that A_t is not a convergent sequence (if A_t had been convergent, then taking limits D_t would be too).

For the second case, we suppose $c = \frac{p}{q}$. Set $a(n) = cn + r(n)$ where $1 \prec r(n) \preceq \log n$. Then we show

$$\frac{1}{t}\sum_{n\le t} e\left(\frac{\lfloor a(n)\rfloor}{p}\right)$$

is not convergent.

Define

$$C(n) = \sum_{j=n}^{n+q-1} e\left(\frac{\lfloor a(j)\rfloor}{p}\right)$$

Given an interval on which $r(x)$ does not attain any values of the form i/q for $i \in \mathbb{Z}$, we note that $\lfloor a(j+q)\rfloor = \lfloor a(j)\rfloor + p$ on this interval, so $C(n)$ is constant provided n and $n+q$ both lie in the interval. At a point where $r(x_0)$ attains a value i/q however, there is a transition: since p and q are coprime, there exists a unique $0 \le m < q$ such that $mp+i$ is a multiple of q. Terms $a(n)$ with n in the residue class of m modulo q now have $e(\lfloor a(n)\rfloor/p)$ differing before and after the transition by a factor $e(1/p)$. In particular, if n and n' are values before and after the transition, then $|C(n) - C(n')| = |1 - e(1/p)|$. Since $r(x) \preceq \log x$, by Lemma 4.3 the intervals between transitions have the property that the ratio of their endpoints is bounded away from 1 for sufficiently large x. Defining D_t as in the previous lemma, we see that D_t takes the value $C(\lfloor t \rfloor)/q$ for t such that $\lfloor t \rfloor$ and $\lfloor t(1+\epsilon) \rfloor$ stay in the same interval. The above shows that the sequence D_t is not convergent, so arguing as before, it follows that the sequence A_t is not convergent.

□

7. Pointwise convergence: sufficient conditions

We note that in all the theorems which we prove in this section, one conclusion is that $(\lfloor a(n) \rfloor)$ is ergodic. We note that in each case, this follows from Remark 2.6 and Theorem 3.2.

The central object of study in this section is the sequence of functions defined by

$$A_t f(x) = \frac{1}{t} \sum_{n \leq t} f(T^{\lfloor a(n) \rfloor}(x)).$$

We will make a standing assumption that $a(n)$ is defined for all positive integers. This does not constitute a loss of generality because the truth of the theorems for a sequence $a(n)$ is equivalent to the the truth of the theorems for the sequence $a'(n)$ defined by $a'(n) = a(n+k)$.

We will first provide a sufficient condition for the above sequence of functions to be convergent and then establish that the condition holds in the cases of Theorems 3.4, 3.5 and 3.8. In all of these theorems, we will make the assumptions that $a(x) \succ x$, $a \in U$ and $a(x)$ is subpolynomial. By Lemma 4.1, it follows for such an $a(x)$, that there exists a $\delta \geq 1$ such that for any $\epsilon > 0$, $x^{\delta-\epsilon} \prec a(x) \prec x^{\delta+\epsilon}$ (and in the case $\delta = 1$, $x \prec a(x) \prec x^{1+\epsilon}$).

Define

$$S_t(\alpha) = \frac{1}{t} \left| \sum_{n \leq t} e(a(n)\alpha) \right|.$$

Theorem 7.1. *Let $a \in U$ be subpolynomial, satisfying $a(x) \succ x$. Let δ be such that for each $\epsilon > 0$, $x^{\delta-\epsilon} \prec a(x) \prec x^{\delta+\epsilon}$. Suppose further that there are constants ϵ and C and a decreasing function $\sigma(t)$ such that*

(1) *$S_t(\alpha) \leq \sigma(t)$ for all $t^{\frac{1}{2}-\delta} \leq |\alpha| < t^{\epsilon}$;*
(2) *$\sum_{n=0}^{\infty} (\sigma(2^n))^{1-\epsilon} < C$.*

Then $(\lfloor a(n) \rfloor)$ is good for pointwise convergence of L^2 functions.

It will be noticed that the sufficient conditions given above for pointwise convergence of L^2 functions resemble quantitative versions of the conditions which appeared in Lemma 6.1 for norm convergence.

The proof of this theorem will proceed by using spectral theory to compare the desired quantities $A_t f(\omega)$ with quantities $V_t f(\omega)$, which are shown to be pointwise convergent on a set of full measure. To define $V_t f(\omega)$, first let $F(t) = a^{-1}(t)$, (i.e. the compositional inverse function of a). Then $V_t f(\omega)$ is defined by

$$V_t f(\omega) = \frac{1}{t} \sum_{n \leq a(t)} F'(n) f(T^n \omega).$$

The V_t operator is like the A_t operator with the averaging taken over all times up to $\lfloor a(t) \rfloor$ with suitable weights, rather than just over the times occurring in the sequence $a(n)$.

Lemma 7.2. *Suppose a is as in the statement of Theorem 7.1. Then for all $f \in L^1$, $V_t f(\omega)$ is pointwise convergent almost everywhere and the limit coincides with the limit of the Césaro averages of f, $1/n \sum_{j=0}^{n-1} f(T^j \omega)$.*

Since we work using spectral theory, we will make use of estimates of the Fourier transforms of these operators:

$$\widehat{A}_t(\beta) = \frac{1}{t} \sum_{n \le t} e(\lfloor a(n) \rfloor \beta); \text{ and}$$

$$\widehat{V}_t(\beta) = \frac{1}{t} \sum_{n \le a(t)} F'(n) e(n\beta)$$

Note that $\widehat{A}_t(\beta)$ differs from the quantity $S_t(\beta)$ which appears in the statement of Theorem 7.1 as the former is the Fourier transform of the sequence $(\lfloor a(n) \rfloor)$, while the latter is the Fourier transform of the sequence $a(n)$. We will use Theorem 5.2 to compare the two.

The following lemmas are combined with the hypotheses in Theorem 7.1 to get the required spectral estimates.

Lemma 7.3. *Let a and δ be as described above. Then for any $\epsilon > 0$, there is a constant C_ϵ such that for each β,*

$$|\widehat{A}_t(\beta) - \widehat{V}_t(\beta)| \le C_\epsilon (t^{-1} + t^{\delta - 1 + \epsilon} |\beta|).$$

We use this lemma to control $|\widehat{A}_t(\beta) - \widehat{V}_t(\beta)|$ for small β. When β is larger, we show separately that $|\widehat{V}_t(\beta)|$ and $|\widehat{A}_t(\beta)|$ are small.

Lemma 7.4. *Let a and δ be as described above. Then for $0 < |\beta| < 1/2$, we have*

$$|\widehat{V}_t(\beta)| \le C_\epsilon |\beta|^{-\frac{1}{\delta} - \epsilon} t^{-1}.$$

The proof of Theorem 7.1 is then as follows:

Proof of Theorem 7.1. It is clearly sufficient to prove that the theorem holds for an arbitrary non-negative function $f \in L^2$. Let f be non-negative and let $\rho \in \mathbb{Q}$ satisfy $\rho > 1$.

Fix $t > 1$. We will estimate $|\widehat{A}_t(\beta) - \widehat{V}_t(\beta)|$. For $0 \le |\beta| \le t^{\frac{1}{2} - \delta}$, Lemma 7.3 gives us

$$|\widehat{A}_t(\beta) - \widehat{V}_t(\beta)| \le C_\epsilon t^{\delta - 1 + \epsilon + \frac{1}{2} - \delta} = C_\epsilon t^{-\frac{1}{2} + \epsilon}. \tag{7.1}$$

For $t^{\frac{1}{2} - \delta} < |\beta| < \frac{1}{2}$, from Lemma 7.4, we see

$$|\widehat{V}_t(\beta)| \le C_\epsilon |\beta|^{-\frac{1}{\delta} - \epsilon} t^{-1 + \epsilon} \le C_\epsilon t^{-\frac{1}{2\delta} + \frac{\epsilon}{2} + \epsilon\delta}.$$

Since an equation like this is valid for any ϵ, we see that there is a $C_\epsilon > 0$ such that

$$|\widehat{V}_t(\beta)| \le C_\epsilon t^{-\frac{1}{2\delta} + \epsilon}. \tag{7.2}$$

We finish by bounding $\widehat{A}_t(\beta)$ for t in this range. For this we use Theorem 5.2. By hypothesis, we know that $S_t(\alpha) \le \sigma(t)$ for $t^{\frac{1}{2} - \delta} < |\alpha| < t^\epsilon$. Letting $Q = t^{\delta - \frac{1}{2}}$,

$S = t^\epsilon$ and $r = \min(\sqrt{S-1}, 1/\sigma(t))$, we apply the theorem to deduce that for $t^{\delta-\frac{1}{2}} \le |\beta| \le \frac{1}{2}$,

$$|\widehat{A}_t(\beta)| \le C\sqrt{\frac{\log r}{r}} \le C\max\left(\sigma(t)^{\frac{1}{2}-\epsilon}, t^{-\frac{\epsilon}{3}}\right). \tag{7.3}$$

Combining equations (7.1), (7.2) and (7.3), we see that for all β in the range under consideration ($|\beta| \le \frac{1}{2}$), we have

$$|\widehat{A}_t(\beta) - \widehat{V}_t(\beta)|^2 \le \max\left(C_\epsilon t^{-1+\epsilon}, C_\epsilon t^{-\frac{1}{\delta}+\epsilon}, C\sigma(t)^{1-\epsilon}, t^{-\frac{2\epsilon}{3}}\right). \tag{7.4}$$

We note that when any of the terms is summed over t taking values in J_ρ for $\rho > 1$, the sum is bounded above independently of β, where for the $\sigma(t)^{1-\epsilon}$ term, we are using

$$\sum_{t\in J_\rho} \sigma(t)^{1-\epsilon} \le \lceil \log 2/\log\rho \rceil \sum_{n=0}^{\infty} \sigma(2^n)^{1-\epsilon}.$$

It follows then that $\sum_{t\in J_\rho} |\widehat{A}_t(\beta) - \widehat{V}_t(\beta)|^2$ is bounded above by a constant C independently of β.

By the spectral theorem, we have

$$\int \sum_{t\in J_\rho} |A_t f(\omega) - V_t f(\omega)|^2\, d\mu = \int \sum_{t\in J_\rho} |\widehat{A}_t(\beta) - \widehat{V}_t(\beta)|^2\, d\nu_f,$$

where ν_f is the spectral measure of f. Since $\int 1\, d\nu_f = \|f\|_2^2$, we see that

$$\int \sum_{t\in J_\rho} |A_t f(\omega) - V_t f(\omega)|^2\, d\mu \le C\|f\|^2.$$

In particular, it follows that for almost every ω, $A_t f(\omega) - V_t f(\omega)$ is convergent along J_ρ to 0. Since by Lemma 7.2, $V_t f(\omega)$ is convergent almost everywhere, it follows that for each $\rho > 1$, $A_t f(\omega)$ is convergent for almost every ω along J_ρ. Since there are countably many $\rho > 1$ in $\mathbb{Q}$, it follows that for a set of ω of full measure, $A_t f(\omega)$ is convergent along each J_ρ with $\rho > 1$ in $\mathbb{Q}$. Now by Lemma 4.4, we see that $A_t f(\omega)$ is convergent on this set of full measure as required.

□

Proof of Lemma 7.2. Let $f \in L^1$. Then $V_t f(\omega)$ is defined as $\frac{1}{t}\sum_{j<a(t)} F'(j) f(T^j\omega)$. Replacing t by $F(n)$, we see that establishing convergence of $V_t f(\omega)$ as $t \to \infty$ is equivalent to establishing convergence of $\frac{1}{F(n)}\sum_{j=0}^{n-1} F'(j) f(T^j\omega)$ as $n \to \infty$.

This will be a consequence of Birkhoff's theorem which states that for points ω belonging to a set of full measure, $1/n\sum_{i=0}^{n-1} f(T^i\omega)$ is convergent. Let ω be any such point and let α be the limit of the sequence. Write $f_n = f(T^n(\omega))$ and $\alpha_n = \frac{1}{n}\sum_{i=0}^{n-1} f_i$ so that $\alpha_n \to \alpha$ as $n \to \infty$. Write $c_n = F'(n)$ and note that for sufficiently large n, c_n is a decreasing sequence of numbers converging to 0. Let $\epsilon > 0$ be given and let K be chosen so that $|\alpha_n - \alpha| < \epsilon$ for $n \ge K$ and $(c_n)_{n\ge K}$ is a decreasing sequence converging to 0.

Write $C_m = \sum_{n=0}^{m-1} c_n$ and note that $F(m) \to \infty$ and $C_m/F(m) \to 1$ as $m \to \infty$. In order to establish the claim, it is then sufficient to establish convergence of the modified sequence $\frac{1}{C_m}\sum_{n=0}^{m-1} c_n f(T^n\omega)$.

By summation by parts, we see

$$\frac{1}{C_m}\sum_{n=0}^{m-1} c_n f_n = \frac{1}{C_m}\left(mc_{m-1}\alpha_m + \sum_{n=1}^{m-1} n(c_{n-1}-c_n)\alpha_n\right).$$

Applying this to the case where $f_n \equiv 1$, we see

$$1 = \frac{1}{C_m}\left(mc_{m-1} + \sum_{n=1}^{m-1} n(c_{n-1}-c_n)\right). \tag{7.5}$$

Combining these two equations, we get

$$\begin{aligned}
&\left|\frac{1}{C_m}\sum_{n=0}^{m-1} c_n f_n - \alpha\right| \\
=&\frac{1}{C_m}\left|mc_{m-1}(\alpha_m-\alpha) + \sum_{n=0}^{m-1} n(c_{n-1}-c_n)(\alpha_n-\alpha)\right| \\
\le&\frac{1}{C_m}\left|\sum_{n=0}^{K-1} n(c_{n-1}-c_n)(\alpha_n-\alpha)\right| + \epsilon\frac{1}{C_m}\left|mc_{m-1} + \sum_{n=K}^{m-1} n(c_{n-1}-c_n)\right|
\end{aligned}$$

Taking the limit as $m\to\infty$, we see that the first term converges to 0 and by (7.5), the second term converges to ϵ giving

$$\limsup_{m\to\infty}\left|\frac{1}{C_m}\sum_{n=0}^{m-1} c_n f_n - \alpha\right| \le \epsilon.$$

Since ϵ is arbitrary we see that $1/C_m\sum_{n=0}^{m-1} c_n f_n$ is convergent to the same limit as the Césaro averages $1/m\sum_{n=0}^{m-1} f(T^n\omega)$. □

Proof of Lemma 7.3. For simplicity, we will make the assumption that $a(0)=0$ and a is increasing on $[0,\infty)$. This will ensure that the inverse function $F(x)$ is defined on $[0,\infty)$ also. Since the initial terms do not affect the convergence or otherwise of ergodic averages and the Hardy field assumption only governs the behavior of a on a neighborhood of ∞, this does not entail a loss of generality.

We need to show that

$$\left|\sum_{n\le t} e(\lfloor a(n)\rfloor\beta) - \sum_{n\le a(t)} F'(n)e(n\beta)\right| \le C(1+t^{\delta+\epsilon}|\beta|).$$

Since $\delta\ge 1$ and $|e(\lfloor a(n)\rfloor\beta) - e(a(n)\beta)| \le C\beta$, it is sufficient to show

$$\left|\sum_{n\le t} e(a(n)\beta) - \sum_{n\le a(t)} F'(n)e(n\beta)\right| \le C(1+t^{\delta+\epsilon}|\beta|). \tag{7.6}$$

We then observe by the change of variables formula, setting $u=F(x)$ that

$$\int_0^t e(\beta a(u))\,du = \int_0^{a(t)} F'(x)e(\beta x)\,dx.$$

To show (7.6), it is therefore sufficient to bound the difference between the sums and the integrals. We observe that if $n\le u<n+1$, then $|F'(u)e(\beta u) - F'(n)e(\beta n)| \le$

$|F'(u)-F'(n)|+|F'(n)||e(\beta u)-e(\beta n)|$. It follows that

$$\begin{aligned}&\left|\sum_{n\le a(t)} F'(n)e(n\beta)-\int_0^{a(t)} F'(x)e(\beta x)\,dx\right|\\ &\le C+\sum_{n\le a(t)}|F'(n+1)-F'(n)|+C|\beta|\sum_{n\le a(t)}|F'(n)|\\ &\le C+C|\beta|F(a(t))|=C(1+\beta t).\end{aligned}$$

Similarly, for $n\le x<n+1$, $|e(\beta a(x))-e(\beta a(n))|\le C|\beta|\sup_{n\le x<n+1} a'(x)$ so we see

$$\left|\sum_{n\le t} e(\beta a(n))-\int_0^t e(\beta a(u))\,du\right|\le C|\beta|a(t).$$

Combining the above gives the desired proof. □

Proof of Lemma 7.4. We begin by noting the following properties of $F(x)$:

(1) $F(x)\to\infty$ as x approaches ∞;
(2) For sufficiently large x, F is convex down and is twice continuously differentiable. In particular $F'(x)$ decreases to 0 as x approaches ∞ and $F''(x)$ is negative;
(3) For some constant C, $xF'(x)\le CF(x)$; This property is a consequence of the first two.
(4) $x^{\frac{1}{\delta}-\epsilon}\prec F(x)\prec x^{\frac{1}{\delta}+\epsilon}$ for each $\epsilon>0$.

From property 4, we see that it is sufficient to prove

$$|\sum_{n\le t} F'(n)e(n\beta)|\le C\cdot F(1/|\beta|). \tag{7.7}$$

We can assume that $|\beta|\ge 1/t^{\delta}$, since the above estimate is nontrivial only in this case. To ease our notation, we introduce, for $s\ge 1$,

$$G_s(\beta)=\sum_{n\le s} e(n\beta);\qquad\text{and}$$
$$W_s(\beta)=\sum_{n\le s} F'(n)e(n\beta).$$

Using the fact that the numbers $e(\beta),e(2\beta),\dots,e(\lfloor s\rfloor\beta)$ form a geometric progression and the estimate $|1-e(\beta)|\ge 2|\beta|$ (which is valid for $|\beta|\le 1/2$), we obtain

$$|G_s(\beta)|\le\frac{1}{|\beta|}. \tag{7.8}$$

Let s satisfy $1\le s\le t$ and estimate

$$|W_t(\beta)|\le|W_s(\beta)|+|W_t(\beta)-W_s(\beta)|.$$

We trivially have

$$|W_s(\beta)|\le C\cdot F(s).$$

Next, we will show that

$$|W_t(\beta)-W_s(\beta)|\le CF'(s)\cdot\frac{1}{|\beta|}. \tag{7.9}$$

Choosing $s = 1/|\beta|$, by property 3, the last two estimates above imply the one in (7.7).

The inequality in (7.9) is obtained using summation by parts:

$$
\begin{aligned}
&|W_t(\beta) - W_s(\beta)| \\
&= \left| \sum_{\lfloor s \rfloor < n \le t} F'(n) e(n\beta) \right| = \left| \sum_{\lfloor s \rfloor < n \le t} F'(n) \left(G_n(\beta) - G_{n-1}(\beta) \right) \right| \\
&= \left| \sum_{\lfloor s \rfloor < n \le \lfloor t \rfloor} \left(F'(n) - F'(n+1) \right) G_n(\beta) + F'\left(\lfloor t \rfloor \right) G_t(\beta) - F'\left(\lfloor s \rfloor + 1 \right) G_s(\beta) \right|
\end{aligned}
$$

using the inequality in (7.8) and property 2,

$$
\begin{aligned}
&\le \sum_{\lfloor s \rfloor < n \le \lfloor t \rfloor} \left(F'(n) - F'(n+1) \right) \frac{1}{|\beta|} + F'\left(\lfloor t \rfloor \right) \frac{1}{|\beta|} + F'\left(\lfloor s \rfloor + 1 \right) \frac{1}{|\beta|} \\
&= 2F'\left(\lfloor s \rfloor + 1 \right) \frac{1}{|\beta|} \le 2F'(s) \frac{1}{|\beta|},
\end{aligned}
$$

where in the last inequality we again used that $F'(x)$ is decreasing. □

We will now use a lemma of Van der Corput to show that the hypotheses of Theorem 7.1 are satisfied under the hypotheses of Theorems 3.4 and 3.5, thus proving these theorems.

Lemma 7.5 (Van der Corput [20])**.** *Let Y, X, ℓ be integers. Suppose that $Y < X$, $\ell \ge 2$ and set $s = 2^\ell$. Suppose that the real function b is ℓ-times differentiable in the interval $[Y, X]$ and $|b^{(\ell)}| \ge \varrho$ in $[Y, X]$, where ϱ is a positive number.*

Then, letting

$$
R = \frac{1}{X - Y} \left| b^{(\ell-1)}(X) - b^{(\ell-1)}(Y) \right|,
$$

we have

$$
\begin{aligned}
&\left| \sum_{Y \le n \le X} e(b(n)) \right| \le \\
&\le 21(X - Y) \left(\left(\frac{R^2}{\varrho} \right)^{1/(s-2)} + \left(\frac{1}{\varrho (X - Y)^\ell} \right)^{2/s} + \left(\frac{R}{\varrho (X - Y)} \right)^{2/s} \right).
\end{aligned}
$$

We now use the lemma to prove Theorems 3.4 and 3.5. In these proofs, we will use the symbol C to denote an upper bound which is independent of X and α in the specified range.

Proof of Theorem 3.4. Since $r \in U$ is sub-polynomial, it follows by Lemma 4.1 that there exists a $\delta > 0$ such that for all $\epsilon > 0$, $x^{\delta-\epsilon} \prec r(x) \prec x^{\delta+\epsilon}$. Since r is assumed to be non-polynomial, it follows by Lemma 4.2 that $x^{\delta-k-\epsilon} \prec r^{(k)} \prec x^{\delta-k+\epsilon}$. Choose $l > \max(1 + \delta, \partial p)$ so that $a^{(l)} = r^{(l)}$. We then apply Van der Corput's Lemma to $b(n) = \alpha a(n)$ with $X = t$, $Y = t^{1-\tau}$ and $s = 2^l$. For sufficiently large x, we have $x^{\delta-\epsilon} < a(x) < x^{\delta+\epsilon}$. We then compute $R \le C b^{(l-1)}(t^{1-\tau})/t \le \alpha C t^{-(1-\tau)(l-(1+\delta)+\epsilon)-1}$ and $\rho \ge \alpha C t^{-(l-\delta)-\epsilon}$. If we replaced τ and ϵ with 0, R^2/ρ

would be given by $C\alpha t^{-(l-\delta)}$, so in particular, choosing sufficiently small τ and ϵ (independently of t and α), we see that for sufficiently large t (independently of α), $R^2/\rho \leq C\alpha t^{-1}$.

Similarly, we see that $1/(\rho(X-Y)^l) \leq C/(\alpha t^{\delta-\epsilon})$ where C only depends on ϵ.

Finally, taking again ϵ and τ to be 0, we would be able to estimate the term $R/(\rho(X-Y))$ above by C/t. So for small positive values of ϵ and τ, we see that $R/(\rho(X-Y)) \leq C/\sqrt{t}$.

Putting this together, we see that for sufficiently small ϵ and τ and for sufficiently large t, we have

$$\left|\frac{1}{t}\sum_{n\leq t} e(\alpha a(n))\right| \leq \frac{t^{1-\tau}}{t} + C\left(\frac{\alpha}{t}\right)^{\frac{1}{s-2}} + C\left(\frac{1}{\alpha t^{\delta-\epsilon}}\right)^{\frac{2}{s}} + C\left(\frac{1}{\sqrt{t}}\right)^{\frac{2}{s}},$$

where the first term comes from the trivial estimate $|e(\alpha a(n))| \leq 1$ for $n \leq t^{1-\tau}$. We therefore see that there is a $\epsilon > 0$ such that for sufficiently large t and α lying between $t^{-\delta+\frac{1}{2}}$ and $t^{\frac{1}{2}}$,

$$\left|\frac{1}{t}\sum_{n\leq t} e(\alpha a(n))\right| \leq \frac{C}{t^\epsilon}.$$

It is then clear that the sequence $a(n)$ satisfies the hypotheses of Theorem 7.1 so we see that $(\lfloor a(n)\rfloor)$ is a good sequence for pointwise convergence of L^2 functions as claimed. □

Proof of Theorem 3.5. Suppose $a(x) = p(x) + r(x)$ where $p(x)$ is a polynomial of degree $n \geq 2$ and $r(x)$ is non-polynomial satisfying $r(x) \succ (\log x)^{2^{n+1}-1+\eta}$. We may assume that $r(x) \prec x^\epsilon$ for each $\epsilon > 0$ as if not, the conclusion follows from Theorem 3.4. We make two applications of Van der Corput's lemma: one with $l = n$ and one with $l = n+1$ to give control of the quantity $S_t(\alpha)$ appearing in Theorem 7.1 in two distinct ranges of α. As above, let $b(x) = \alpha a(x)$.

First with $l = n$, we see that $a^{(l-1)}(x) = cx + d + r^{(l-1)}(x)$ for constants c and d so taking $Y = \sqrt{t}$ and $X = t$, that $R \leq C\alpha A$ for a constant A independent of t and α. Also, $a^{(l)} = c + r^{(l)}(x)$ so that $\rho \geq C\alpha$. For $t^{-n+\frac{1}{2}} \leq |\alpha| \leq (\log t)^{-(2^n-2)-\frac{\eta}{2}}$, we have as in Theorem 3.4

$$\begin{aligned}\left|\frac{1}{t}\sum_{n\leq t} e(\alpha a(n))\right| &\leq \frac{\sqrt{t}}{t} + C\alpha^{1/(2^n-2)} + C\left(\frac{1}{\alpha t^n}\right)^{1/2^{n-1}} + Ct^{-1/2^{n-1}} \\ &\leq C(\log t)^{-1-\eta'},\end{aligned}$$

where $\eta' = \eta/(2^{n+1}-4)$.

With $l = n+1$, we have $a^{(l-1)}(x) = c + r^{(n)}(x)$ and $a^{(l)}(x) = r^{(n+1)}(x)$. Setting $X = t$ and $Y = t^{1-\tau}$, we get $R \leq \alpha C r^{(n)}(t^{1-\tau})/t$ and $\rho = \alpha r^{(n+1)}(t)$. Using Lemma 4.2, we see that $R^2/\rho \leq C\alpha t^{-n-1+3\epsilon+n\tau}$. Letting τ and ϵ be small, we get $R^2/\rho \leq C\alpha t^{-n}$. Also,

$$\left(\frac{1}{\rho(X-Y)^l}\right)^{2/2^l} \leq C\left(\frac{1}{\alpha r^{(n+1)}(t)t^{n+1}}\right)^{1/2^n}.$$

Since $r(x) \succ (\log x)^{2^{n+1}-1+\eta}$, it follows from an application of Lemma 4.2 that $r^{(n+1)}(x) \succ (\log x)^{2^{n+1}-2+\eta}/x^{n+1}$.

If $|\alpha| \geq (\log t)^{-2^n+2-\eta/2}$, then we see

$$\left(\frac{1}{\rho(X-Y)^l}\right)^{2/2^l} \leq C\left(\frac{1}{(\log t)^{2^n+\eta/2}}\right)^{1/2^n}.$$

Finally, the term $(R/(\rho(X-Y)))^{1/2^n}$ is bounded above by a negative power of t so we see that for $(\log t)^{-2^n+2-\eta/2} \leq |\alpha| \leq t$,

$$\left|\frac{1}{t}\sum_{n\leq t} e(\alpha a(n))\right| \leq C(\log t)^{-1-\eta/2^{n+1}}.$$

It may then be verified that the sequence $a(n)$ satisfies the conditions of Theorem 7.1 completing the proof of the theorem. □

We finish this section with the proof of Theorem 3.8. We first quote Weyl's inequality (see [21]).

Lemma 7.6. *Suppose that hcf$(b,q) = 1$. and $|\alpha - b/q| < 1/q^2$. Let $p(x) = \alpha x^k + \alpha_1 x^{k-1} + \ldots + \alpha_k$. Then*

$$\left|\sum_{n\leq t} e(p(n))\right| \leq Ct^{1+\epsilon}(q^{-1}+t^{-1}+qt^{-k})^{1/K},$$

where $K = 2^{k-1}$ and the constant C depends only on ϵ and k.

We extend this to the case where rather than the rational approximation of the leading coefficient, one considers the rational approximation of any of the non-constant coefficients.

Theorem 7.7. *Let $p(x) = a_k x^k + a_{k-1}x^{k-1} + \ldots + a_1 x + a_0$. If for some $0 < \sigma < 2^{-k^2}$, there is an l with $1 \leq l \leq k$ such that there is a rational approximation p_l/q_l to a_l such that $|a_l - p_l/q_l| < 1/(t^{l-\sigma}q_l)$ and $q_l > t^\sigma$, then for each $\epsilon > 0$, there is a constant C (dependent only on k and ϵ) such that*

$$\left|\sum_{n=1}^{t} e(p(n))\right| \leq Ct^{1+\epsilon-\eta},$$

where $\eta = \sigma/(12k \cdot 2^{k^2})$

Proof. Set $\alpha_i = 2^{k(l-i)}\sigma$. For each $i > l$, there is a rational approximation p_i/q_i to a_i satisfying $|a_i - p_i/q_i| < 1/(t^{i-\alpha_i}q_i)$ and $q_i < t^{i-\alpha_i}$ by Dirichlet's criterion. We note that $q_l > t^{\alpha_l}$ and take j to be the largest index for which the inequality $q_j > t^{\alpha_j}$ holds. If $j = k$, then the theorem holds trivially by Lemma 7.6 so we assume that $j < k$. For $i > j$, we therefore have $q_i \leq t^{\alpha_i}$, whereas for j, we have $t^{\alpha_j} < q_j < t^{j-\alpha_j}$. Writing $Q = q_{j+1}\cdots q_k$, we see that $Q < t^{\alpha_{j+1}+\ldots+\alpha_k} < t^{2\alpha_{j+1}}$. We observe also $\alpha_j - 2j\alpha_{j+1} > \alpha_j/2$ so that $q_j/Q^j > t^{\alpha_j/2}$.

Using Dirichlet's criterion again, there is a rational approximation c/d to $Q^j a_j$ satisfying $d \leq t^{j-\alpha_j}$ and

$$\left|Q^j a_j - \frac{c}{d}\right| \leq \frac{1}{t^{j-\alpha_j}d}. \tag{7.10}$$

Dividing through by Q^j, we see that $|a_j - c/(Q^j d)| \leq 1/(t^{j-\alpha_j}Q^j d)$. By the choice of q_j, we conclude that $Q^j d \geq q_j$, so that $t^{\alpha_j/2} < d < t^{j-\alpha_j}$.

We now modify the polynomial $p(n)$ to give a polynomial $p'(n)$, where the coefficients of n^i for $i > j$ are replaced by their rational approximations:

$$p'(x) = p_k/q_k x^k + \ldots + p_{j+1}/q_{j+1} x^{j+1} + a_j x^j + \ldots a_0.$$

For M between $t^{1-\eta}$ and t, set $A'(M) = \sum_{n \leq M} e(p'(n))$ We then bound $|A'(M)|$ by splitting the sum into residue classes modulo Q and applying Lemma 7.6. Write $p'(x) = p_0(x) + p_1(x)$, where $p_0(x)$ consists of the terms of degree greater than j (the perturbed terms) and $p_1(x)$ consists of the remaining terms. Then $e(p_0(n))$ is periodic with period Q so we see that

$$A'(M) = \sum_{n \leq M} e(p'(n)) = \sum_{r=0}^{Q-1} e(p_0(r)) \sum_{Qs+r \leq M} e(p_1(Qs+r)).$$

Note that the leading coefficient of $p_1(Qs+r)$ as a polynomial in s is $Q^j a_j$. From (7.10), we see that that we can apply Lemma 7.6 to conclude that there is a constant C depending on $\epsilon > 0$ and j such that

$$\begin{aligned}
|A'(M)| &\leq CQ(M/Q)^{1+\epsilon}(d^{-1} + (M/Q)^{-1} + d(Q/M)^j)^{1/2^{j-1}} \\
&\leq Ct^{1+\epsilon}\left(t^{-\alpha_j/2} + t^{2\alpha_{j+1}+\eta-1} t^{j-\alpha_j} (t^{1-\eta-2\alpha_{j+1}})^{-j}\right)^{1/2^{j-1}} \\
&= Ct^{1+\epsilon}\left(t^{-\alpha_j/2} + t^{-\alpha_j + j\eta + 2j\alpha_{j+1}}\right)^{1/2^{j-1}} \\
&\leq Ct^{1+\epsilon}\left(t^{-\alpha_j/2} + t^{-\alpha_j/2+\alpha_j/6}\right)^{1/2^{j-1}} \\
&\leq Ct^{1+\epsilon-\alpha_j/(3\cdot 2^{j-1})}
\end{aligned}$$

Now, we write $p(x) = p'(x) + \Delta(x)$, where $\Delta(x) = (a_k - p_k/q_k)x^k + \ldots + (a_{j+1} - p_{j+1}/q_{j+1})x^{j+1}$. We observe that $|\Delta(n) - \Delta(n-1)| \leq Ct^j/t^{j+1-\alpha_{j+1}} = Ct^{-(1-\alpha_{j+1})}$. We then estimate

$$\begin{aligned}
\left|\sum_{n \leq t} e(p(n))\right| &\leq t^{1-\eta} + \left|\sum_{n=t^{1-\eta}}^{t} e(p'(n)) e(\Delta(n))\right| \\
&= t^{1-\eta} + \left|\sum_{n=t^{1-\eta}}^{t} (A'(n+1) - A'(n)) e(\Delta(n))\right|.
\end{aligned}$$

Using summation by parts, we have

$$\begin{aligned}
&\left|\sum_{n=t^{1-\eta}}^{t} (A'(n+1) - A'(n)) e(\Delta(n))\right| \\
&\leq |A'(t)| + |A'(t^{1-\eta})| + \sum_{n=t^{1-\eta}}^{t} |A'(n)| |(e(\Delta(n+1)) - e(\Delta(n))| \\
&\leq Ct^{1+\epsilon-\alpha_j/2^j} + Ctt^{-(1-\alpha_{j+1})} t^{1+\epsilon-\alpha_j/(3\cdot 2^{j-1})} \\
&\leq Ct^{1+\epsilon+\alpha_{j+1}-\alpha_j/(3\cdot 2^{j-1})} \\
&\leq Ct^{1+\epsilon-\alpha_{j+1}/12} \leq Ct^{1+\epsilon-\eta}.
\end{aligned}$$

This completes the proof. □

We use this to prove the following general lemma.

Lemma 7.8. *Suppose that $p(x)$ is a polynomial of degree $k \geq 2$ with the property that two of its non-constant coefficients have a ratio which is badly approximable by rationals. Then there exist a $\sigma > 0$, an $\eta > 0$ and a $C > 0$ such that for $t^{-k+1/2} < |\beta| < t^\sigma$*

$$\frac{1}{t}\sum_{n\leq t} e(\beta p(n)) \leq Ct^{-\eta}.$$

Proof. Suppose $\alpha = a_i/a_j$ is badly approximable by rationals. That is there exists D and $m \geq 2$ such that $|\alpha - p/q| \geq D/q^m$ for all rationals p/q. Let $\sigma < 1/(3m+4)$. Then we show that for β in the range $t^{-\sigma} < |\beta| < t^\sigma$, one of the coefficients of $\beta p(n)$ satisfies the conditions of Theorem 7.7 above. Let $t^{-\sigma} < \beta < t^\sigma$. We use Dirichlet's criterion to approximate βa_i and βa_j as

$$\left|\beta a_i - \frac{p}{q}\right| < \frac{1}{t^{i-\sigma}q}$$
$$\left|\beta a_j - \frac{c}{d}\right| < \frac{1}{t^{j-\sigma}d},$$

where $q < t^{i-\sigma}$ and $d < t^{j-\sigma}$. The range of β ensures that c and p are non-zero. We show that either $q > t^\sigma$ or $d > t^\sigma$. Suppose for a contradiction that $q \leq t^\sigma$ and $d \leq t^\sigma$. It then follows that $p \leq Ct^{2\sigma}$ and $c \leq Ct^{2\sigma}$.

From the above equations, we see that

$$\frac{dp}{cq}\left(1 - \frac{1}{qt^{i-\sigma}}\right) \Big/ \left(1 + \frac{1}{dt^{j-\sigma}}\right) < \alpha < \frac{dp}{cq}\left(1 + \frac{1}{qt^{i-\sigma}}\right) \Big/ \left(1 - \frac{1}{dt^{j-\sigma}}\right).$$

In particular, we see that

$$\left|\alpha - \frac{dp}{cq}\right| \leq \frac{dp}{t^{1-\sigma}}$$

giving $dp/t^{1-\sigma} > D(cq)^{-m}$ by the badly approximable condition on α. We then see that $D < t^{(3m+4)\sigma-1}$ which is a contradiction for sufficiently large t. By Theorem 7.7, for $t^{-\sigma} < |\beta| < t^\sigma$, the conclusion of the lemma holds for such t, and hence by choosing a sufficiently large C for all t.

On the other hand, if $t^{-k+1/2} < |\beta| \leq t^{-\sigma}$, then we apply Van der Corput's lemma (Lemma 7.5) with $l = k$ to $\beta p(x)$ to show that a similar equation holds (with a different value of η). Taking the minimum of these values of η, we see that the lemma holds for β satisfying $t^{-k+1/2} < |\beta| < t^\sigma$. □

Finally, we apply this to prove Theorem 3.8 as follows:

Proof of Theorem 3.8. Suppose $p(x)$ is as in the statement of the theorem and $r(x) \prec x$. If $r(x) \preceq 1$ or $r(x) \succ (\log x)^n$ for sufficiently large n, then it follows from Theorem 3.5 that $\lfloor p(n)+r(n)\rfloor$ is good for pointwise convergence of L^2 functions. To prove the theorem, it is therefore necessary to consider the case $1 \prec r(x) \prec (\log x)^a$. As usual, we consider the case in which $r(x)$ is increasing, the decreasing case being similar. We can apply Lemma 7.8 to the polynomial $p(n)$ giving C, σ and η such that

$$\sum_{n\leq t} e(\beta p(n)) \leq Ct^{1-\eta} \tag{7.11}$$

for all t and $t^{-k+1/2} < |\beta| < t^\sigma$. We will denote the left side of the inequality by $A(t)$.

We note that

$$\left|\sum_{n=1}^{t} e(\beta a(n))\right| \le t^{1/2} + \left|\sum_{n=t^{1/2}}^{t} e(\beta a(n))\right|$$

We next partition the sum according to the value of $r(n)$ in a grid of width $t^{-\eta/2}$ and use the previous lemma to estimate the above sum. Let x_s be the first integer for which $r(x_s) > s/t^{\eta/2}$, $s_0 = t^{\eta/2}r(t^{1/2})$ and $s_1 = t^{\eta/2}r(t)$. Letting β satisfy $t^{-k+1/2} < |\beta| < t^{\min(\eta/4,\sigma)}$, we have

$$\begin{aligned}\left|\sum_{n\le t} e(\beta a(n))\right| &\le t^{1/2} + \sum_{s_0\le s<s_1}\left|\sum_{x_s<n\le x_{s+1}} e(\beta(p(n)+r(n)))\right| \\ &\le Ct^{1-\eta/4} + \sum_{s_0\le s<s_1}\left|\sum_{x_s<n\le x_{s+1}} e(\beta p(n))\right|,\end{aligned}$$

where the second inequality follows from the fact that for $n \in (x_s, x_{s+1}]$, $|e(\beta(p(n)+r(n))) - e(\beta(p(n)+s/t^{\eta/2}))| \le C|\beta|/t^{\eta/2} \le Ct^{-\eta/4}$. Since $\sum_{x_s<n\le x_{s+1}} e(\beta p(n)) = A(x_{s+1}) - A(x_s)$, we see from (7.11) that $|\sum_{x_s<n\le x_{s+1}} e(\beta p(n))| \le Ct^{1-\eta}$ for each t. It then follows that

$$\left|\sum_{n\le t} e(\beta a(n))\right| \le Ct^{1-\eta/4} + Ct^{\eta/2}r(t)t^{1-\eta} \le Ct^{1-\eta/4}.$$

This clearly satisfies the conditions of Theorem 7.1, completing the proof of the theorem. □

8. Pointwise convergence: necessary conditions

We recall from [17] that necessary and sufficient conditions for a sequence $(a(n))$ to be good for pointwise convergence of L^p functions are:

(1) For each measure-preserving transformation of the probability space Ω, there exists a dense subset of $L^p(\Omega)$ on which there is convergence;
(2) There exists a weak(p,p) maximal inequality of the form: For each measure-preserving system $(\Omega, \mathcal{B}, \mu, T)$, there exists a K such that

$$\mu\{\omega\colon M^*f(\omega) > \lambda\} \le \frac{K\|f\|_p^p}{\lambda^p}, \tag{8.1}$$

where $M^*f(\omega) = \sup_{t\ge 1}\left|\frac{1}{t}\sum_{n\le t} f(T^{a(n)}\omega)\right|$.

In particular, the constant K in the maximal inequality satisfies for each measure-preserving system and each function $f \in L^p$,

$$K \ge \frac{\lambda^p \mu\{\omega\colon M^*f(\omega) > \lambda\}}{\|f\|_p^p} \tag{8.2}$$

To establish necessary conditions, we show that if the function a differs from a rational polynomial by a function which grows too slowly, then (8.1) does not hold.

Proof of Theorem 3.7. Let $p = cq$ be a constant multiple of a polynomial with integer coefficients whose lowest degree term is of order $k \ge 3$. Let $1 \prec r(x) \preceq (\log x)^m$ where $m < 2k/(k+2)$. For any integer N (which we may assume to be large), consider the system $T(y) = y+1 \bmod cN^k$ and the function $f(y) = \chi_{[0,2]}(y)$.

Then $\|f\|_2^2 = 2/(cN^k)$. We assume as before that $r(x)$ is increasing (the case where $r(x)$ is decreasing being similar) and observe that $q(jN) \equiv 0 \pmod{N^k}$. Given $y \in [0, cN^k)$, let s_1 be the first real number such that $y + r(s_1) = 1 \pmod{cN^k}$ and s_2 be the first real number beyond s_1 such that $y + r(s_2) \equiv 2 \pmod{cN^k}$. For $y > 2$ and sufficiently large N, we have $s_1 = r^{-1}(cN^k - y + 1)$ and $s_2 = r^{-1}(cN^k - y + 2)$. We then have $y + a(Nj) \bmod cN^k \in [1, 2]$ for $s_1 \le Nj \le s_2$ so $T^{\lfloor a(Nj) \rfloor}(y) \in [0, 2]$ for $s_1 \le Nj \le s_2$. It follows that

$$M^* f(y) \ge \frac{1}{s_2} \left\lfloor \frac{s_2 - s_1}{N} \right\rfloor.$$

Using Lemma 4.3 for N large, we see that there is a constant $\kappa > 0$ such that for all but a small set of y,

$$M^* f(y) \ge \frac{\kappa (cN^k)^{\frac{1}{m} - 1}}{N}.$$

Taking $\lambda = \kappa (cN^k)^{\frac{1}{m} - 1}/N$, we see that $\mu\{y \colon M^* f(y) \ge \lambda\}$ is close to 1 for large N. From (8.2), we see that

$$K \ge \frac{CN^{2k(\frac{1}{m} - 1) - 2}}{2/(cN^k)} = CN^{\frac{2k}{m} - (k+2)}.$$

Since $m < 2k/(k+2)$, we see that the exponent is positive, so that there does not exist a finite constant K in the maximal inequality.

This completes the proof of the theorem. □

In the proof of Theorem 3.6, we will need to make use of the following result which is a compilation of results which appeared in [7] (Theorems 2 and 5 and Lemma 6)

Theorem 8.1. *Suppose that $q \in \mathbb{Z}[x]$ is a polynomial of degree $n > 1$. If for any prime p, the range of q modulo p is not all of the residue classes of p, then the cardinality of the range of q modulo p is at most $\alpha p + O(p^{1/2})$, where $\alpha = 1 - 1/n!$.*

We will also make use of the following result from the book of Lidl and Niederreiter ([15], Corollary 7.5).

Theorem 8.2. *Let $q \in \mathbb{Z}[x]$ be a polynomial of degree $n > 1$. If $p \equiv 1 \pmod{n}$, then the range of q modulo p is not all of the residue classes of p.*

Proof of Theorem 3.6. Let $p(x) = cq(x)$, where $q(x) \in \mathbb{Z}[x]$. Let $p_1, p_2, \ldots$ be the sequence of primes congruent to 1 modulo n. Let $1 - 1/n! < \beta < 1$. By Theorems 8.2 and 8.1, the range of $q(n) \bmod p_i$ has cardinality less than βp_i for sufficiently large i. Forming $P_N = p_1 p_2 \cdots p_N$, we see that the range of $q(n) \bmod P_N$ has cardinality less that $\beta^N P_N$. We will then use this in the construction of the maximal function showing that the constant in (8.2) is unbounded.

Let R_N be the range of $q(n) \bmod P_N$ and consider the system $T(y) = y + 1 \bmod cP_N$. Set $S = \bigcup_{j \in R_N} [cj - 1, cj + 1)$, $S' = \bigcup_{j \in R_N} [cj, cj + 1)$ and $f(y) = \chi_S(y)$. We then recall that we are considering the case $a(x) = p(x) + r(x)$, where $1 \prec r(x) \preceq \log x (\log \log x)^m$. We consider the case where $r(x)$ is (eventually) increasing, the case where $r(x)$ is decreasing being similar. As above, let y be any point in $[0, cP_N)$. Let s_1 be the first real number such that $y + r(s_1) \equiv 1 \pmod{cP_N}$ and s_2 be the first real number beyond s_1 such that $y + r(s_2) \equiv 2 \pmod{cP_N}$. We then have that $y + p(n) + r(n) \bmod cP_N \in S'$ for $s_1 < n < s_2$ so $T^{\lfloor a(n) \rfloor}(x) \in S$

for $s_1 < n < s_2$. It follows that $M^* f(y) \geq (s_2 - s_1)/s_2 \geq C/\exp((\log P_N)^m)$ by Lemma 4.3. Also, $\|f\|_2^2 \leq 2|R_N|/P_N \leq C\beta^N$ so we see that the constant in (8.2) satisfies

$$K \geq \frac{C}{\beta^N \exp(2(\log P_N)^m)}$$

By the prime number theorem for arithmetic progressions [11] (Corollary 1, Ch. IX, §3), we see that $\log P_N \leq CN \log N$. Since $m < 1$, we see that the denominator converges to 0 as N approaches ∞ so we then see that there is no constant as required. □

Proof of Theorem 3.9. Let $p \in C\mathbb{Q}[x]$ be of degree $n > 1$. Then let $1 \leq m \leq n$. Write $p(x) = cq(x)$ where $q \in \mathbb{Q}[x]$. We will perturb inductively the x^m coefficient of p so as to give a strictly irrational polynomial $\tilde{p}$ which differs from p only in the x^m coefficient and by an arbitrarily small amount, but for which the badness of $(\lfloor \tilde{p}(n) + \log n \rfloor)$ for L^2 functions persists. Write $p = p_0$.

Suppose that we have found a sequence of perturbations $p_1, p_2, \ldots, p_k \in c\mathbb{Q}[x]$ along with numbers $n_0 < n_1 < \ldots < n_{k-1}$ satisfying the following properties:

(1) $\lfloor p_{i+1}(n) + \log n \rfloor = \lfloor p_i(n) + \log n \rfloor$ for $1 \leq n \leq n_i$;
(2) For $i < k$, there exists a measure-preserving system Ω_i, a function f_i and a constant λ_i such that

$$(8.3) \qquad \frac{\lambda_i^2 \mu\{\omega\colon \sup_{1 \leq t < n_i} 1/t \sum_{n \leq t} f_i(T^{\lfloor p_i(n) + \log n \rfloor}(\omega)) > \lambda_i\}}{\|f_i\|_2^2} \geq i.$$

(3) Writing the x^m coefficient of p_i as ca_i/b_i, we have $0 < a_{i+1}/b_{i+1} - a_i/b_i \leq e^{-b_i}$ and $a_{i+1}/b_{i+1} - a_i/b_i < (a_i/b_i - a_{i-1}/b_{i-1})/2$.

Then we inductively extend the sequence as follows: Since $p_k \in c\mathbb{Q}[x]$, by the proof of Theorem 3.6, there exists a system Ω_k, a function f_k and a constant λ_k such that

$$\frac{\lambda_k^2 \mu\{\omega\colon M^* f_k(\omega) > \lambda_k\}}{\|f_k\|_2^2} > i.$$

By the monotone convergence theorem, there is an $n_k > n_{k-1}$ such that

$$(8.4) \qquad \frac{\lambda_i^2 \mu\{\omega\colon \sup_{1 < t < n_k} 1/t \sum_{n \leq t} f_k(T^{\lfloor p_k(n) + \log n \rfloor}(\omega)) > \lambda_k\}}{\|f_k\|_2^2} \geq k.$$

Since the fractional parts of $p_k(n) + \log n$ for $n \leq n_k$ have a maximum value which is strictly less than 1 ($1 - \epsilon_k$, say), it follows that increasing the x^m coefficient of p_k by a sufficiently small amount (which we can assume to be less than $\epsilon_k/(2(n_k)^m)$) one can find a polynomial $p_{k+1} \in c\mathbb{Q}[x]$ satisfying conditions 1, 2 and 3.

We then take a limit of these polynomials to get a polynomial $\tilde{p}$. The control of the amount of the increment of the x^m coefficient ensures that $\lfloor \tilde{p}(n) + \log n \rfloor = \lfloor p_i(n) + \log n \rfloor$ for $n \leq n_i$. It follows from equation (8.4) that there is no maximal inequality for $(\lfloor \tilde{p}(n) + \log n \rfloor)$.

We note that condition 3 implies that the x^m coefficient of $\tilde{p}$ is given by a number $c\alpha$ where α is transcendental. This completes the inductive step and hence the proof of the theorem. □

9. Notes on L^p, $p > 1$

Here we just state the "L^p" analogs of Theorems 3.4, 3.5 and 3.8. The proofs easily obtained by interpolating between the L^2 bounds and the trivial L^1 bound for $A_t - V_t$ (cf. (7.4)).

Theorem 9.1. *Let $a \in U$ be subpolynomial and have decomposition $p + r$. If there is an ϵ such that $r(x) \succ x^\epsilon$, then $(\lfloor a(n) \rfloor)$ is pointwise good for L^p for every $p > 1$.*

Theorem 9.2. *Let $a \in U$ be subpolynomial and have a decomposition $p + r$ where $\partial p = n \geq 2$. If for every positive m, $r(x) \succ (\log x)^m$ then $(\lfloor a(n) \rfloor)$ is pointwise good for L^p for every $p > 1$.*

Theorem 9.3. *Let $a \in U$ be subpolynomial, satisfy $a(x) \succ x$ and have a decomposition $p+r$ where p is a polynomial with the property that the ratio of two of its non-constant coefficients is badly approximable by rationals (e.g. $p \in C\mathbb{A}[x] \setminus C\mathbb{Q}[x]$). Then $(\lfloor a(n) \rfloor)$ is pointwise good for L^p for every $p > 1$.*

10. Open problems

Question 10.1. *Is the sequence $(n^2 + \lfloor \log^2 n \rfloor)$ a good sequence for pointwise convergence?*

Question 10.2. *In this question, we consider the structure of set of 'bad perturbations' of a polynomial. Let $p(x)$ be a real polynomial of degree at least 2.*

(1) *(Disk of convergence) Suppose that $r_1(x)$ and $r_2(x)$ belong to a Hardy field and $1 \prec r_1(x) \prec r_2(x) \prec x$. If $\lfloor p(n) + r_1(n) \rfloor$ is good for pointwise convergence, does it follow that $\lfloor p(n) + r_2(n) \rfloor$ is good for pointwise convergence?*
(2) *(Radius of convergence) Does there exist an $r^*(x) \succeq 1$ belonging to a Hardy field so that if $1 \prec r(x) \prec r^*(x)$, then $\lfloor p(n) + r(n) \rfloor$ is bad for pointwise convergence, whereas if $r^*(x) \prec r(x) \prec x$ then $\lfloor p(n) + r(n) \rfloor$ is good for pointwise convergence. (Note that the case $r^*(x) = 1$ corresponds to radius of convergence 0: the entire neighborhood of $p(x)$ is good for pointwise convergence).*
(3) *(Uniform upper bound on radii) Does there exist a q such that if $r(x)$ belongs to a Hardy field and $(\log x)^q \prec r(x) \prec x$ then $\lfloor p(n) + r(n) \rfloor$ is good for pointwise convergence?*

Question 10.3. *Let $a(x)$ be from a Hardy field and assume the sequence $(a(n))$ is good for pointwise convergence. Is $a(x)$ necessarily a polynomial?*

Let us point out that it is known that (n^δ) is bad for pointwise convergence for any nonzero rational number δ (cf. [19]), but it is not known if $(n^{\sqrt{2}})$ is bad. On the other hand, [10, Theorem 2.16] can be used to show that if $a(x)$ is from a Hardy field, and it satisfies $a(x) \to \infty$ but $a(x)/x^\epsilon \to 0$ for every positive ϵ, then $a(x)$ is not good for pointwise convergence.

Question 10.4. *Is $(\lfloor \pi x^2 + x + \log x \rfloor)$ good for pointwise convergence?*

11. Notes

Announcement of results: Some of the results in the present paper were announced in [5], but the results on almost everywhere convergence have been greatly extended.

Polynomial sequences: Bourgain's main paper is [6]. In it he proves that polynomials and their integer parts are good for L^p, $p > 1$. It is unknown what happens for $p = 1$.

Logarithmico-exponential functions: The book of Hardy [9] is a wonderful read.

Hardy fields: All the background material as well as further references can be found in [1, 2].

Hardy fields and uniform distribution: The main motivation to formulate all our theorems in terms of Hardy fields comes from [3].

$\sqrt{n}$ **is bad:** is proved in [19] using Bourgain's entropy method. An elementary, simple proof can be found in [10].

Estimates on trigonometric sums: Our idea to tie estimates on trigonometric sums along a real sequence to estimates on trigonometric sums along the *integer part* of the sequence is based on [8].

Good $\to$ integer part is good: The fact that if (a_n) is pointwise good then $(\lfloor a_n \rfloor)$ is also pointwise good is proved in [4] and, independently, in [13].

References

[1] M. Boshernitzan, *An extension of Hardy's class of 'orders of infinity'*, J. d'Analyse Math. **39** (1981), 235–255.

[2] ———, *New 'orders of infinity'*, J. d'Analyse Math. **41** (1981), 130–167.

[3] ———, *Uniform distribution and Hardy fields*, J. d'Analyse Math. **62** (1994), 225–240.

[4] M. Boshernitzan, R. Jones, and M. Wierdl, *Integer and fractional parts of good averaging sequences in ergodic theory*, (1996), 117–132.

[5] M. Boshernitzan and M. Wierdl, *Ergodic theorems along sequences and Hardy fields*, Proc. Acad. Natl. Sci. **93** (1996), no. 16, 8205–8207.

[6] J. Bourgain, *Point-wise ergodic theorems for arithmetic sets (appendix: The return time theorem)*, Publ. Math. I.H.E.S. **69** (1989), 5–45.

[7] Stephen D. Cohen, *The distribution of polynomials over finite fields*, Acta Arith. **XVII** (1970), 255–271.

[8] J. M. Deshouillers, *Problème de Waring avec exposants non entiers*, Bull. Soc. Math. France **101** (1973), 285–295.

[9] G.H. Hardy, *Orders of Infinity*, 2nd ed., vol. 12, Cambridge tracts in Math. and Math. Phys., 1924.

[10] R. Jones and M. Wierdl, *Convergence and divergence of ergodic averages*, Ergodic Theory and Dynamical Systems **14** (1994), 515–535.

[11] A. Karatsuba, *Basic analytic number theory*, Springer-Verlag, 1993.

[12] Michael Lacey, *On an inequality due to Bourgain*, Illinois J. Math. **41** (1997), no. 2, 231–236. MR 98k:42005

[13] Emmanuel Lesigne, *On the sequence of integer parts of a good sequence for the ergodic theorem*, Comment. Math. Univ. Carolin. **36** (1995), no. 4, 737–743. MR 97c:28039

[14] H. Niederreiter, *On a paper of Blum, Eisenberg, and Hahn concerning ergodic theory and the distribution of sequences in the Bohr group*, Acta Sci. Math. (Szeged) **37** (1975), 103–108.

[15] H. Niederreiter R. Lidl, *Finite fields*, Cambridge Univ. Press, Cambridge, 1983.

[16] J. Rosenblatt and M. Wierdl, *Pointwise ergodic theorems via harmonic analysis*, Ergodic theory and its connections with harmonic analysis (K. Petersen and I. Salama, eds.), London Mathematical Society Lecture Note Series, no. 205, Cambridge Univ. Press, 1995, pp. 5–151.

[17] S. Sawyer, *Maximal inequalities of weak type*, Ann. Math. **84** (1966), 157–174.

[18] Jean-Paul Thouvenot, *La convergence presque sûre des moyennes ergodiques suivant certaines sous-suites d'entiers (d'après Jean Bourgain)*, Astérisque (1990), no. 189-190, Exp. No. 719, 133–153, Séminaire Bourbaki, Vol. 1989/90. MR 92c:11082

[19] J. Bourgain V. Bergelson, M. Boshernitzan, *Some results on nonlinear recurrence*, J. d'Analyse Math. **62** (1994), 30–46.

[20] J. G. van der Corput, *Neue zahlentheoretische Abschatzungen II*, Mathematische Zeitschrift **29** (1929), 397–426.
[21] R. C. Vaughan, *The Hardy-Littlewood method*, Cambridge Univ. Press, Cambridge, 1981.

(Michael Boshernitzan) Department of Mathematics - MS 136, 6100 Main Street, Rice University, Houston, Texas 77005-1892
E-mail address: michael@math.rice.edu

(Grigori Kolesnik) Department of Mathematics, California State University, Los Angeles, CA 90032
E-mail address: gkolesn@calstatela.edu

(Anthony Quas and Máté Wierdl) University of Memphis, Department of Mathematical Sciences, 373 Dunn Hall, Memphis, TN 38152-3240
E-mail address: aquas@memphis.edu

E-mail address: mw@csi.hu