The Furstenberg set and its random version

Aihua FAN, Hervé QUEFFÉLEC and Martine QUEFFÉLEC

Abstract. We study some number-theoretic, ergodic and harmonic analysis properties of the Furstenberg set of integers $S = \{2^j 3^k\}$ and compare them to those of its random analogue T. In this half-expository work, we show for example that S is "Khinchin distributed", is far from being Hartman uniformly distributed while T is, also that S is a $\Lambda(p)$-set for all $2 < p < \infty$ and that T is a p-Rider set for all p such that $4/3 < p < 2$. Measure-theoretic and probabilistic techniques, notably martingales, play an important role in this work.

Mathematics Subject Classification (2020). Primary 37A44; Secondary 43A46, 60G46.

Keywords. Furstenberg set, Sidon set, Khinchin class, uniform distribution, martingale.

1. Introduction

The goal of this paper, which is mainly a survey but also contains several original results, is the study of two different sets of integers: the Furstenberg set S and its randomized version denoted T.

Since various notions (from harmonic analysis, probability, ergodic theory, number theory and diophantine approximation, fractal geometry) are involved, we begin with a rather long introduction in which the necessary notions are described in a sketchy way. But we strive to add precise bibliographical references, or references to the forthcoming sections, where those notions are presented in more detail. Complements can be found on harmonic analysis, probability and Banach spaces in [48], on ergodic theory in [60] and on distribution of sequences in [45].

Notations. Here are summarized the notations we adopt: $\mathbb{N}$ (resp. $\mathbb{N}_0$) denotes the set of positive (resp. non-negative) integers, $\mathbb{Z}$ the set of integers, $\{x\}$ the fractional part of the real number x. The cardinality of a finite set A is denoted either by $|A|$, or by $\#A$. Also for estimations, we make use of the classical Hardy's notation $\ll$ (bounded up to a positive multiplicative constant). We will mainly (but not exclusively) consider the group $G = \mathbb{T} = \mathbb{R}/\mathbb{Z}$ with m as its Haar measure and we will denote $\|x\| = \inf_{n \in \mathbb{Z}} |x - n|$ for $x \in \mathbb{T}$. We put $L^p(\mathbb{T})$ for $L^p(\mathbb{T}, m)$. For every $f \in L^1(\mathbb{T})$

(integrable and 1-periodic), more generally every finite measure μ on $\mathbb{T}$ and every $n \in \mathbb{Z}$, we set:

$$\int_{\mathbb{T}} f \, dm := \int_0^1 f(x) \, dx, \quad \widehat{f}(n) = \int_{\mathbb{T}} f \overline{e}_n \, dm, \quad \widehat{\mu}(n) = \int_{\mathbb{T}} \overline{e}_n \, d\mu,$$

where $e_n(x) = e(nx)$ with $e(x) = \exp(2i\pi x)$. $\mathbb{E}$ (resp. V) denotes the expectation (resp. the variance) in a probability space. $\mathbb{E}(X|\mathcal{C})$ denotes the conditional expectation of an integrable random variable X given the sub-σ-field $\mathcal{C}$.

1.1. The set S.

Our main object is the Furstenberg set S of positive integers defined as

$$S = \{2^j 3^k : j, k \in \mathbb{N}_0\}.$$

This is the multiplicative semi-group of $\mathbb{N}$, generated by 2 and 3. More generally, for a finite set of coprime numbers $\{q_1, \ldots, q_s\}$, we denote by $S(q_1, \ldots, q_s)$ the multiplicative semi-group generated by $q_1, \ldots, q_s$. If $2 = p_1 < p_2 < \cdots < p_s < \cdots$ is the ordered sequence of the prime numbers, $S(p_1, \ldots, p_s)$ is nothing but

$$D_s := \{n : P^+(n) \leq p_s\},$$

where $P^+(n)$ denotes the greatest prime divisor of $n \geq 2$. This set D_s is called the set of *s-friable integers* and has been intensively studied in analytic number theory; see [1, 6, 32, 36, 55, 62, 73, 74].

When $s = 1$, $S(q_1)$ is the so-obtained Hadamard lacunary set $\{q_1^n, n \geq 1\}$, which enjoys various properties, as well in harmonic analysis as in ergodic theory, that we shall discuss below. It is quite natural to explore the same properties for less sparse sets, e.g. $S(q_1, \ldots, q_s)$ as soon as $s \geq 2$.

We particularly focus on the set $S = S(2, 3)$ because of its closeness to a famous conjecture of Furstenberg, the $\times 2 \times 3$-conjecture, which asserts that *every continuous probability measure on the circle $\mathbb{T}$ which is both 2-invariant and 3-invariant must be the Haar measure m of $\mathbb{T}$*; see Section 2.2. We make no progress on this conjecture in our paper, but this problem acted as a motivation for revisiting the set S, which is clearly invariant under the multiplication by 2 and 3. By rearrangement the Furstenberg set gives rise to an increasing sequence of integers, called the "Furstenberg sequence" (also denoted S, to ease notation):

$$S = \{1 = s_1 < s_2 < \cdots < s_n < \cdots\}.$$

From different viewpoints (harmonic analysis, arithmetic, and dynamical systems), the set S or the sequence (s_n) is quite original and mysterious. Its rate of growth is intermediate between the growth of a polynomial sequence (e.g. the sequence (n^2) of

squares) and that of a power sequence (e.g. the sequence (2^n)) and this rate is not so easy to handle with. Ramanujan gave an amazing statement, of which Hardy had to give a full proof ([37]).

The sumset $\Sigma = \{2^j + 3^k\}$ was well-understood from the uniform distribution and harmonic analysis viewpoints [9], for the reason that it appears as a projection of a nice product set. But the multiplicative structure of S prevents us from using the same projection technique.

We shall study three types of properties of S.

Rate of increase and lacunarity. We say that a sequence of integers $E = (u_n)$ is *weakly lacunary* if $u_{n+1} - u_n \to \infty$, or *Hadamard lacunary* of ratio ρ if $\rho := \liminf u_{n+1}/u_n > 1$ (in the sequel, we shall say "Hadamard set"). The set S is not a Hadamard set because $s_{n+1}/s_n \to 1$, which was exploited by Furstenberg in [33]. We shall look for asymptotic behaviors of s_n, $s_{n+1} - s_n$, $(s_{n+1}/s_n) - 1$. Notice that estimates of such quantities for a Hadamard set are obvious. For the set S, it is easy to prove that
$$I_S(N) := |S \cap [1, N]| \approx (\log N)^2,$$
whence $s_n \approx \exp(\sqrt{n})$. Going further needs sharp estimates and more work (Section 2.1).

Distribution issues. Consider a set of integers $E = \{\lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots\} \subset \mathbb{Z}$. As is well known (Hermann Weyl's theorem [45]), for almost every $x \in (0, 1)$ the sequence $(\lambda_n x)$ is *equidistributed modulo* 1, meaning that, for any $h \in \mathbb{Z} \setminus \{0\}$, we have

$$(1) \qquad \frac{1}{N} \sum_{n=1}^{N} e(h\lambda_n x) \to 0.$$

We also say more briefly: almost every x is E-*normal*. The exceptional set of those x such that $(\lambda_n x)$ is not equidistributed mod 1 is Haar-negligible. It is interesting to study its properties. This exceptional set can be reduced to the singleton $\{0\}$. In other words, in some extreme cases, it can happen that

$$(2) \qquad \frac{1}{N} \sum_{n=1}^{N} e(\lambda_n x) \to 0 \quad \text{for all } x \neq 0,$$

and then we say that the set E is *Hartman uniformly distributed*.

We introduce two measuring tools. Hausdorff dimension is usually used to describe the size of sets in a metric space. Consider a subset X of $\mathbb{R}^d$. For $s > 0$ and $\epsilon > 0$,

we define

$$H_\epsilon^s(X) = \inf\left\{\sum_i |U_i|^s : X \subset \bigcup_i U_i,\ |U_i| < \epsilon\right\},$$

where $\{U_i\}$ is a countable family of subsets, called an ϵ-cover of X, and $|U_i|$ denotes the diameter of U_i. The *s-dimensional Hausdorff measure* is then defined by $H^s(X) = \sup_{\epsilon>0} H_\epsilon^s(X)$. There is a critical exponent ρ, called *Hausdorff dimension* of X and denoted by $\dim_H(X)$, such that $H^s(X) = +\infty$ for $s < \rho$ and $H^s(X) = 0$ for $s > \rho$.

Observe that $0 \le \dim_H(X) \le d$ and that $\dim_H(X) = d$ as soon as X has positive Lebesgue measure. For example, if X is the Cantor middle-third set, $\dim_H(X) = \log 2/\log 3$; see [25, Chapter 2] for more details about the Hausdorff dimension.

Next, recall that a probability measure μ on $\mathbb{T}$ is called a *Rajchman measure* if $\hat{\mu}(n) \to 0$ at infinity.

The support of a Rajchman measure is not arbitrary and has special porosity. The rate of decay for its Fourier transform is relevant in this respect; see [54] for a background.

The Hausdorff dimension of the exceptional set $W(S)$ of those $x \in \mathbb{T}$ such that $(s_n x)$ is not equidistributed and the existence of Rajchman measure supported by $W(S)$ holds our attention in Section 2.2.

The above property (1) is equivalent to the following one:

$$(3) \qquad \frac{1}{N}\sum_{n=1}^N f(\lambda_n x) \to \int_{\mathbb{T}} f\, dm \quad \text{a.e.}$$

for every Riemann-integrable function f. But surprisingly, Marstrand [55], refuting a conjecture of Khinchin, proved that for $E = \{\lambda_n\} = \mathbb{N}$, there are functions $f \in L^\infty(\mathbb{T})$ for which (3) fails. This led us to coin the *Khinchin class* of a subset $E = \{\lambda_n\} \subset \mathbb{N}$ as

$$(4) \qquad \mathcal{K}_E := \left\{ f \in L^1(\mathbb{T}) : \frac{1}{N}\sum_{n=1}^N f(\lambda_n x) \to \int_{\mathbb{T}} f\, dm \text{ a.e.}\right\}.$$

In terms of this notation, the above result of Marstrand can be restated as $\mathcal{K}_\mathbb{N} \not\supset L^\infty(\mathbb{T})$. But for the set S, Marstrand [55] proved that $\mathcal{K}_S \supset L^\infty(\mathbb{T})$ and later Nair [56] proved that $\mathcal{K}_S = L^1(\mathbb{T})$. We will revisit these results in Section 2.5 and give a simple proof of Nair's result in a weaker form.

Harmonic analysis properties. Now consider S as a subset of $\mathbb{Z}$, the dual group of the unit circle $\mathbb{T}$. To better understand the involved notions, it is useful to consider more generally (see [70], notably Chapter 5, for a detailed exposition) a compact abelian group G equipped with its normalized Haar measure m, its dual $\Gamma = \hat{G}$, the Fourier

transform $\widehat{f} \colon \Gamma \to \mathbb{C}$ of a function $f \in L^1(G)$ or $\widehat{\mu} \colon \Gamma \to \mathbb{C}$ of a bounded measure μ being defined as

$$\widehat{f}(\gamma) := \int_G f(x)\gamma(-x)\,dm(x), \quad \widehat{\mu}(\gamma) := \int_G \gamma(-x)\,d\mu(x).$$

The spectrum of f is (with a similar definition for μ)

$$\mathrm{sp}(f) := \{\gamma \in \Gamma : \widehat{f}(\gamma) \neq 0\}.$$

For a subset E of Γ, and a Banach space $X \subset L^1(G)$, X_E denotes the closed subspace of X defined by

$$X_E = \{f \in X : \mathrm{sp}(f) \subset E\}.$$

This subset E is declared "sparse", in a vague sense depending on X, if any $f \in X_E$ behaves better than a generic function of X. More specifically:

(a) E is called a $\Lambda(p)$-*set* if $L_E^1 \subset L^p$ for some fixed $p > 1$;

(b) E is called a *Sidon set* (or 1-Sidon set) if $f \in L_E^\infty(G)$ implies $\widehat{f} \in \ell^1(\Gamma)$;

(c) E is called a *p-Sidon set* if $f \in L_E^\infty(G)$ implies $\widehat{f} \in \ell^p(\Gamma)$ $(1 \le p < 2)$;

(d) E is called a *p-Rider set* if $f \in X_E(G)$ implies $\widehat{f} \in \ell^p(\Gamma)$, where X is the space of randomly bounded functions $(1 \le p < 2)$.

See [69, 70] for the first two notions, [75] and [9, p. 181] for the third, and [49, 50, 68] for the fourth.

For example, the set $E = \{2^j, j \ge 0\} \subset \mathbb{Z}$ is Sidon and hence $\Lambda(p)$ for all $p < \infty$. The set $F = \{(2^j, 3^k)\} \subset \mathbb{Z}^2$ is p-Sidon exactly for $p \ge 4/3$ ([9]). The notion of p-Riderness coincides with that of p-Sidonicity when $p = 1$ ([48]) and is weaker when $1 < p < 2$, but reveals easier to handle with; see [68] for a full arithmetic characterization. For example, it provides a simple proof of the non-trivial fact that the union of two Sidon sets is again Sidon [20].

Such "sparse" sets E are lacunary in the sense that

$$I_E(N) := |E_N| \quad \text{with } E_N := E \cap [1, N]$$

is much smaller than N. Indeed, we have

- $I_E(N) \ll N^{2/p}$ if E is a $\Lambda(p)$-set $(p > 2)$;

- $I_E(N) \ll \log N$ if E is a Sidon set;

- $I_E(N) \ll (\log N)^{p/(2-p)}$ if E is a p-Rider set $(1 \le p < 2)$.

Here is another kind of lacunary sets. We say that $E \subset \mathbb{Z}$ is a *Rajchman set* if for any probability measure μ on $\mathbb{T}$, $\lim_{|n| \to \infty, \, n \in E^c} \widehat{\mu}(n) = 0$ implies that μ is a Rajchman
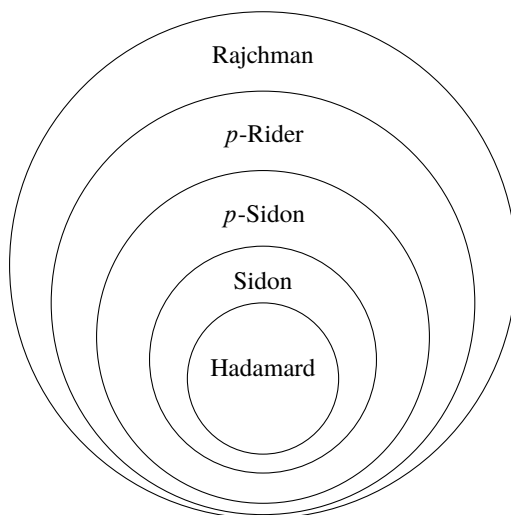
Figure 1
Harmonic classification of lacunary sets.

measure. The relations between these lacunary sets are shown in Figure 1 and the last inclusion is due to Lefèvre and Rodriguez-Piazza [47].

For the Furstenberg set S, we have $I_S(N) \approx (\log N)^2$, a property shared by any $\frac{4}{3}$-Rider set (NB $p/(2-p) = 2$ for $p = 4/3$). This leads naturally to the questions about $\Lambda(p)$-property or p-Sidonicity, p-Riderness for the set S and its random brother T, which is defined below. We will examine these properties for both S and T.

1.2. The set T. Motivated by the estimate $I_S(N) \approx (\log N)^2$ seen above, we define a random version T of S as follows in Section 5: We select the integer $n \geq 1$ with probability $\delta_n = (\log n)/n$ and we reject it with probability $1 - \delta_n$, our selections or rejections being independent. This approach is not new. It goes back to Erdős–Rényi who introduced random sets of integers in [21], it was then systematically developed by Bourgain [12], and became popular under the name "Selectors of Bourgain". See also [49], and the papers by Kahane and Katznelson [39, 40]. We obtain in this way a random set T of integers satisfying

$$\mathbb{E}(I_T(N)) = \sum_{n=1}^{N} \frac{\log n}{n} \approx (\log N)^2.$$

It can be proved that $I_T(N) \approx (\log N)^2$ almost surely (this follows from a non-standard law of large numbers; see [5, p. 276] for a proof). This random set T might be easier to deal with, and might give us some hint on what we could, or could not, hope for the

set S. All the properties we have previously discussed will be revisited with T in place of S.

1.3. Detailed content of the paper. Let us now be more specific on the content of this paper, which is organized as follows.

Section 1 is this introduction.

Section 2 investigates the first arithmetic or dynamical properties of S. We naturally start this section with an estimate of the cumulative function $I_S(N)$, leading to the asymptotic behaviors of $s_n, s_{n+1}/s_n$ and then of $s_{n+1} - s_n$ and $(s_{n+1} - s_n)/s_n$, improving previous inequalities due to Tijdeman [73,74] by providing effective constants. We obtain the following estimates (Corollary 2.4).

Theorem 1.1. *We have*

$$\frac{1}{(\log s_n)^\rho} \ll \frac{s_{n+1} - s_n}{s_n} \ll \frac{1}{(\log s_n)^{1/(\rho+1)}},$$

where the exponent ρ is explicit ($\rho = 4.117\ldots$).

Our estimations rely on a rational approximation of $\alpha = \log 2/\log 3$ due to Rhin [66] and Wu and Wang [76].

Then we switch to the study of the S-orbits, that is the sets $(s_n x)$ for $x \in \mathbb{T}$, which is relevant to the dynamics of the semi-group $\{2^j 3^k\}$. All infinite S-orbits $(s_n x)$ for $x \notin \mathbb{Q}$ are dense in $\mathbb{T}$ but not necessarily equidistributed (see [33]). We complete these assertions of Furstenberg by investigating in detail the set $W(S)$ of those $x \in \mathbb{T}$ such that $(s_n x)$ is not equidistributed. As observed above, $W(S)$ is negligible but it is uncountable and even has a positive Hausdorff dimension, for which we give estimates in Section 2 (Theorem 2.7). Moreover, we construct a probability measure supported on $W(S)$ with its Fourier transform vanishing at infinity (Theorem 2.9).

Theorem 1.2. *The set of x with a non-equidistributed S-orbit has a positive Hausdorff dimension ≥ 0.451621 and supports a* Rajchman *probability measure.*

Another question concerning the distribution of the orbits $(s_n x)$, related to a conjecture of Khinchin [43], has been solved in the '90s by Nair [56]. Recall that this conjecture, refuted by Marstrand, asked whether

$$\frac{1}{N} \sum_{n=1}^{N} f(nx) \to \int f \, dm \quad m\text{-a.e.}$$

for all $f \in L^1(\mathbb{T})$. For subsets of $\mathbb{N}$, we can speak of the problem of Khinchin. Nair's result solved this problem positively for the set S.

Theorem 1.3 (Nair). *For every* $f \in L^1(\mathbb{T})$,

$$\frac{1}{N} \sum_{n=1}^{N} f(s_n x) \to \int f \, dm \quad m\text{-a.e.}$$

We will give a very simple proof of this fact with a slight restriction on f (membership in $L \log^+ L$) and a release on the measure (Theorem 2.11).

Section 3 begins with reminders of notions of thin sets from harmonic analysis.

We recall here different notions of lacunarity: $\Lambda(p)$, q-Paley, p-Sidon and later p-Rider. The set S, thanks to its weakly lacunary behavior, is expected to share some of them. Actually, Gundy and Varopoulos [36] proved in the seventies that S is a $\Lambda(p)$-set for every $p > 2$ (and even q-Paley for every $1 \leq q < 2$) by a judicious interplay of the underlying martingales, and here probability comes into the matter.

Section 4 presents two probabilistic tools on which we will heavily rely, namely the Azuma inequalities for scalar martingales with bounded differences, which allow a unified presentation of some results (like Salem–Zygmund's theorem). A second probabilistic tool, the Burkholder inequalities on the square function of a scalar or Hilbert space-valued martingale, is recalled. Then, we prove in details the result of Gundy–Varopoulos [36] which deserves to be better known, the initial proof being rather abrupt.

Theorem 1.4 (Gundy–Varopoulos). *The set S is $\Lambda(p)$ for all $2 < p < \infty$ and even q-Paley for all $1 < q < 2$.*

We are not able to prove a p-Sidon or p-Rider property for S, although necessary conditions are shown to be satisfied for an interval of values of the parameter p. Whence our last Section 5.

In order to determine which is prevailing between arithmetic or density, we turn to a random version $T := T(\omega) = \{t_n\}$ of the set S, which a.s. shares with S the same cumulative function; in particular, $t_{n+1}/t_n \to 1$ a.s. and almost sure estimates of its growth can be performed (Theorem 5.1 and Theorem 5.2). Here are some results concerning this set T.

Theorem 1.5 (Bourgain). *The random set $T = \{t_1 < t_2 < \cdots\}$ is a.s. Hartman uniformly distributed i.e. a.s.* $\lim_{N \to \infty} \frac{1}{N} \sum_{n < N} e(t_n x) = 0$ *for every* $x \neq 0$.

This result, due to Bourgain [12], then revisited in [57] and [49], is confirming that the arithmetic property must be relevant in distribution problems since S itself is *not* Hartman uniformly distributed, as proved in Section 2.

Theorem 1.6. *The random set T admits a.s. a large Khinchin class X containing all $L^p(\mathbb{T})$ ($p > 1$): a.s. $\lim_{N\to\infty} \frac{1}{N} \sum_{n<N} f(t_n x) = 0$ a.e. when $f \in X$ satisfies the Khinchin conjecture for the positive integers.*

This second result, weaker than its deterministic analogue (Nair), seems new and could possibly be improved.

Theorem 1.7. *The random set T is a.s. a p-Rider set for $p > 4/3$.*

The above first two results are valid for more general random sets that we also consider in Section 5. The third result, not shared by all random sets (see [49]), acts as an incentive for pursuing with S itself. On the other hand, S is a $\Lambda(p)$-set for all $p < \infty$ (see Section 4, Theorem 4.12), but we do not know the answer as concerns T. However, it can be proved that T contains a relatively big subset which is a $\Lambda(p)$-set for all $p < \infty$. (We will come back to this question in a forthcoming paper.)

Addendum. We get a diagram by putting $2^k 3^j$ at the point (k, j) of the integer lattice of the plane. As pointed out in [61], this forms a basis for the generation of various kinds of musical scales (cf. [19]). An associated adic dynamical system is studied in [61].

2. First results on the Furstenberg set

The Furstenberg set, denoted by S in this paper, belongs to a large class of extensively studied multiplicative semi-groups of positive integers. As agreed in the introduction, we denote by $S(q_1, \ldots, q_s)$ the semi-group generated by some fixed coprime integers $q_1, \ldots, q_s$, and by $\{u_1 < u_2 < \cdots\}$ the sequence obtained from the set $S(q_1, \ldots, q_s)$ arranged in increasing order. So, the Furstenberg set $S = S(2, 3)$ is the special case of $s = 2$ and $q_1 = 2$, $q_2 = 3$. Let p_n be the n-th prime number. We know relatively few properties of the set

$$S(p_1, \ldots, p_s) =: \{u_1 < u_2 < \cdots\}.$$

The first observation that $u_{n+1} - u_n \to \infty$ is due to Thue [72], as a consequence of his work on algebraic numbers. Successive improvements have been obtained till the results of Tijdeman [73, 74], which state that

$$(5) \qquad \frac{u_n}{(\log u_n)^A} < u_{n+1} - u_n < \frac{u_n}{(\log u_n)^B}$$

holds with computable positive constants A, B (for n large enough). Thanks to a geometric consideration, Marstrand [55, p. 545] gave the following asymptotic estimate:

$$(6) \qquad |S(q_1, \ldots, q_s) \cap [1, \ldots, n]| = K_s (\log n)^s + r_n \quad \text{as } n \to \infty,$$

where

$$K_s = \frac{1}{s!} \prod_{j=1}^{s} \frac{1}{\log q_j}, \quad r_n = O((\log n)^{s-1}).$$

Marstrand stated the result for $q_j = p_j$, but his proof remains the same for general relatively coprime integers $q_1, \ldots, q_s$.

We will propose a more analytic approach to improve the remainder term r_n in (6) in our case $S := S(2, 3)$, and then deduce better constants A, B in (5) for $S(2, 3)$.

Dynamics studies the behavior of orbits under a transformation or a semi-group of transformations. We study here the action on $\mathbb{T}$ of the semi-group S generated by $x \mapsto 2x \bmod 1$ and $x \mapsto 3x \bmod 1$. The orbit of x under S is the set $\{s_n x \bmod 1\}$. More generally, given a subsequence of integers (k_n), we are interested in the orbit of $x \in \mathbb{T}$ defined by $(k_n x \bmod 1, n \geq 1)$. In the case of a Hadamard set, namely $k_n = q^n$ with $q \geq 2$, the orbits $O_q(x) := (\{q^n x\})$ can be described through the q-adic expansion of x; it is thus easy to construct uncountably many x such that $O_q(x)$ is not dense. Much better, for example when $q = 2$, there exist infinite closed orbits contained in a half-circle, called Sturmian orbits (cf. [16]). The other way round, almost all x give rise to a uniformly distributed orbit and the residual set $W(q)$ has Hausdorff dimension 1. Such questions are asked and some results are obtained in the weakly lacunary case described by the condition $\lim_{n \to \infty}(k_{n+1} - k_n) = +\infty$ (cf. [57]). Furstenberg [33] was interested in the distribution of the orbits $(s_n x)$, and showed that $(s_n x)$ is dense mod 1 for every irrational number x, a first notable difference from the Hadamard case.

We will prove that the set $W(S)$ of x such that $(s_n x)$ is not uniformly distributed is not contained in $\mathbb{Q}$ (in particular, $S = (s_n)$ is not Hartman uniformly distributed) and even has positive Hausdorff dimension. As we will show, if S were Hartman uniformly distributed, the famous $\times 2 \times 3$ Furstenberg conjecture would be true. On the other hand, we will prove that $W(S)$ is a M_0-set, meaning that $W(S)$ supports a probability measure whose Fourier coefficients tend to zero at infinity.

Related to the uniform distribution of the orbits $(s_n x)$ is a question of Khinchin. At the end of this section, we will give a simple proof of the fact, due to Marstrand [55], that bounded measurable functions are in the Khinchin class $\mathcal{K}_S$ of the Furstenberg set; see (4) for the definition of $\mathcal{K}_S$.

2.1. Asymptotic properties of the Furstenberg set $S = \{s_n\}$. Throughout this section, we put $\alpha = \log 2 / \log 3$. By a result of Gelfond [35, Theorem 12.2.2, p. 226], α is a transcendental number, but not Liouville (or else diophantine). Recall that a real number x is called ρ-diophantine, with $\rho > 0$, if for all $n \geq 1$ and a constant $c > 0$:

$$\|nx\| \geq cn^{-\rho}.$$

Any such ρ is called an irrationality exponent of x. We will need the following sharp estimate due to Rhin ([66, p. 160] with $\rho = 7.616$), which is improved by Wu and Wang ([76] with $\rho = 4.117$). The first study on logarithms of integers is due to Baker [4].

Theorem 2.1 ([66, 76]). *The transcendental number α is ρ-diophantine, with $\rho = 4.117 < \infty$.*

Let $S_A := S \cap [1, \ldots, A]$ (A being an integer). According to Hardy ([37, p. 69], notably the equation (5.4.1) there), Ramanujan made the claim

$$|S_A| = \frac{\log(2A)\log(3A)}{2\log 2 \log 3} + \cdots = \frac{1}{2\log 2 \log 3}\log^2 A + \frac{\log 6}{2\log 2 \log 3}\log A + \cdots$$

(where $\cdots$ is a remainder term) in his first letter to Hardy; the latter obtained

$$o(\log A / \log\log A)$$

for the error term (cf. [37, p. 74]) and this result stimulated the study of asymptotic estimates for the cumulative function $|S_A|$. We will prove the following theorems.

Theorem 2.2. *Let ρ be an irrationality exponent of α, and $\delta := \rho/(\rho + 1)$. As $A \to \infty$, we have*

$$(7) \qquad |S_A| = \frac{1}{2\log 2 \log 3}\log^2 A + \frac{\log 6}{2\log 2 \log 3}\log A + O((\log A)^\delta).$$

In particular, using Theorem 2.1, we can take $\rho = 4.117$ and $\delta = 0.80457299$.

This result improves Hardy's error term and the estimate (6) in the case of Furstenberg set S.

Our proof of (7) holds for $S(q_1, q_2)$ and the exponent δ in the error term $O((\log A)^\delta)$ depends on the irrationality exponent ρ of $(\log q_1)/\log q_2$.

From (7), we will deduce the following asymptotic expression for the n-th term s_n in S.

Theorem 2.3. *As $n \to \infty$,*

$$(8) \qquad s_n = \frac{1}{\sqrt{6}}\exp(C\sqrt{n} + \delta_n), \quad \text{where } C = \sqrt{2\log 2 \log 3} = 1.5230\ldots,$$

and the error term $\delta_n = O(n^{-r})$ with $r = 1/(2(\rho + 1)) = 0.0977\ldots$.

In particular, this result recovers a non-lacunarity property of S, namely $s_{n+1}/s_n \to 1$, which was exploited by Furstenberg [33]. We will finally deduce the inequalities (5) obtained by Tijdeman, with explicit exponents, in the Furstenberg case.

Theorem 2.4. *We have that (with $2r = 1/(\rho + 1)$)*

$$(9) \qquad \frac{1}{(\log s_n)^\rho} \ll \frac{s_{n+1} - s_n}{s_n} \ll \frac{1}{(\log s_n)^{2r}}$$

for n large enough. Recall that $\rho \sim 4.117$ and $2r \sim 0.1954$.

To prove Theorem 2.2, we need a Koksma's inequality and an Erdős–Turán's inequality. These two inequalities involve the notion of discrepancy. Let $x_1, \ldots, x_N$ be a finite set of numbers in the interval $[0, 1]$. Its discrepancy is defined by

$$D_N = D_N(x_1, \ldots, x_N) = \sup_{0 \le a < b \le 1} \left| \frac{\#\{1 \le n \le N : x_n \in [a, b]\}}{N} - (b - a) \right|.$$

We could say that the supremum is taken over intervals $[a, b]$. If we take the supremum over intervals $[0, b]$, we get a variant D_N^* of D_N. It is easy to see that $D_N^* \le D_N \le 2D_N^*$ (cf. [45, p. 91]).

Lemma 2.5 (Koksma inequality [45, p. 143]). *Let f be a function defined on the interval $[0, 1]$ of bounded variation $V(f)$. Let $x_1, \ldots, x_N$ be N given points in $[0, 1]$ of discrepancy D_N^*. Then*

$$(10) \qquad \left| \frac{1}{N} \sum_{j=1}^{N} f(x_j) - \int_0^1 f(t)\, dt \right| \le V(f) D_N^*.$$

Let us state the following special case of Erdős–Turán inequality, which says that it is possible to estimate the discrepancy D_N by the Weyl sums.

Lemma 2.6 (Erdős–Turán inequality [45, p. 114]). *There exists a constant $C > 0$ such that for any N given points $x_1, \ldots, x_N$ in $[0, 1]$ and for any positive integer m, we have*

$$(11) \qquad D_N \le C \left(\frac{1}{m} + \sum_{h=1}^{m} \frac{1}{h} \left| \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h x_n} \right| \right).$$

Proof of Theorem 2.2. Denote $N_A := \#\{(j, k) \ge 0 \; ; \; 2^j 3^k \le A\}$. Clearly,

$$N_A = \#\{(j, k) \ge 0 \; ; \; j \log 2 + k \log 3 \le \log A\} = \#\{(j, k) \ge 0 \; ; \; j\alpha + k \le B\alpha\},$$

with $B = \log A / \log 2$. Hence, we get

$$N_A = \sum_{k + j\alpha \le B\alpha} 1 = \sum_{0 \le j \le B} \left([(B - j)\alpha] + 1 \right).$$

Using the relation $[x] = x - \{x\}$, we get

$$(12) \qquad N_A = \sum_{0 \le j \le B} \left(1 + (B - j)\alpha\right) - \sum_{0 \le j \le B} \{(B - j)\alpha\}.$$

Let us first estimate the second sum on the right-hand side of (12), the first sum being easy to compute. Write

$$\sum_{0 \le j \le B} \{(B - j)\alpha\} = \sum_{j=1}^{N} f(x_j) + O(1)$$

with $f(x) = \{x\}$, $x_j = (B - j)\alpha$ and $N = [B]$. We claim that as $B \to \infty$, we have

$$(13) \qquad \sum_{0 \le j \le B} \{(B - j)\alpha\} = \frac{B}{2} + O(B^\delta), \quad \delta = \frac{\rho}{\rho + 1}.$$

Indeed, as $\int_0^1 f(x)\,dx = \frac{1}{2}$ and $V(f) = 1$, by the Koksma inequality (10), we get

$$(14) \qquad \left| \frac{1}{N} \sum_{j=1}^{N} f(x_j) - \frac{B}{2} \right| \le D_N.$$

Now, by the Erdős–Turán inequality (11), the discrepancy D_N of the finite sequence $x_j = (B - j)\alpha \bmod 1$ ($1 \le j \le N$) is, for every $m \in \mathbb{N}$, bounded above by

$$D_N \ll \frac{1}{m} + \sum_{h=1}^{m} \frac{1}{N\,h\|h\alpha\|}$$

(cf. [45, p. 123]); but Theorem 2.1 implies

$$\sum_{h=1}^{m} \frac{1}{h\|h\alpha\|} \ll \sum_{h=1}^{m} h^{\rho-1} \ll m^\rho,$$

which provides

$$D_N \ll \frac{1}{m} + \frac{m^\rho}{N}.$$

We optimize this quantity by taking $m = N^{1/(\rho+1)}$ and obtain $D_N \ll N^{-1/(\rho+1)}$; we conclude (13) with the aid of (14).

Back to (12). By using (13), we get

$$N_A = B + \alpha\left(B(N + 1) - \frac{N(N + 1)}{2}\right) - \frac{B}{2} + O(B^\delta),$$

whence

$$(15) \qquad N_A = \frac{\alpha + 1}{2} B + \alpha\left(BN - \frac{N^2}{2}\right) + O(B^\delta);$$

now, writing

$$BN - \frac{N^2}{2} = -\frac{1}{2}\big[(N - B)^2 - B^2\big] = \frac{B^2}{2} + O(1),$$

we deduce from (15) the asymptotic behavior:

$$(16) \qquad N_A = aB^2 + bB + O(B^\delta),$$

with $B = \log A / \log 2$, $\delta = \rho/(\rho + 1)$, $a = \alpha/2$ and $b = (1 + \alpha)/2$. This proves Theorem 2.2. ∎

Proof of Theorem 2.3. We deduce Theorem 2.3 from Theorem 2.2. Take $A = s_n$; thus $B = B_n = \log s_n / \log 2$ and $N_A = n$. The equality (16) gives first $B_n = O(\sqrt{n})$ and then

$$n = aB_n^2 + bB_n + \alpha_n \quad \text{with } \alpha_n = O(n^{\delta/2}).$$

Consider B_n as unknown of a second degree algebraic equation. Solving B_n gives $B_n = \frac{1}{2a}(-b + \sqrt{\Delta_n})$ with

$$\sqrt{\Delta_n} = \sqrt{b^2 + 4a(n - \alpha_n)} = 2\sqrt{an}\sqrt{1 + \frac{b^2}{4an} - \frac{\alpha_n}{n}}$$
$$= 2\sqrt{an}\left(1 - \frac{\alpha_n}{2n} + O(n^{-1})\right) = 2\sqrt{an} + O\big(n^{-(1-\delta)/2}\big).$$

Here we have used the fact that $2\delta > 1$. We then get

$$(17) \qquad B_n = \sqrt{\frac{n}{a}} - \frac{b}{2a} + \delta_n \quad \text{with } \delta_n = O\big(n^{-(1-\delta)/2}\big).$$

This gives Theorem 2.3, since $(1 - \delta)/2 = r$ and $B_n = \log s_n / \log 2$. ∎

Proof of Theorem 2.4. We easily deduce from Theorem 2.3 the right-hand inequality in (9). Since

$$\frac{s_{n+1} - s_n}{s_n} \ll C\big(\sqrt{n + 1} - \sqrt{n}\big) + |\delta_{n+1}| + |\delta_n|$$

with $\delta_n = O(1/n^r)$ and $r = 1/(2(\rho + 1)) \in (1/18, 1/17)$. With no more information on the sign of δ_n, we get at best the estimate $O(1)/n^r$.

For the reverse inequality, we write $s_n = 2^{c_n} 3^{d_n}$ and $s_{n+1} = 2^{c_{n+1}} 3^{d_{n+1}}$ and observe that $c_{n+1} - c_n \neq 0$ (which is indeed clear: if $c_{n+1} = c_n$, then $d_{n+1} \geq d_n + 1$ and $s_{n+1} \geq 3 s_n$, while clearly $s_{n+1} \leq 2 s_n$ for all n, since $2 s_n \in S$), obviously with the same sign as $d_n - d_{n+1} \neq 0$. Now, we bound

$$\frac{s_{n+1} - s_n}{s_n} > \log \frac{s_{n+1}}{s_n} = |(c_{n+1} - c_n)\alpha - (d_n - d_{n+1})| \cdot \log 3$$
$$= \big| |c_{n+1} - c_n|\alpha - |d_n - d_{n+1}| \big| \cdot \log 3$$
$$\geq \|(c_{n+1} - c_n)\alpha\| \cdot \log 3 \geq c|c_{n+1} - c_n|^{-\rho}.$$

(Once again, we make use of Rhin's or Wang–Wu's lower bounds.) Since

$$|c_{n+1} - c_n| \leq \max(c_n, c_{n+1}) \leq \max(\log_2 s_n, \log_2 s_{n+1}),$$

we obtain

$$\|(c_{n+1} - c_n)\alpha\| \geq C(\log s_{n+1})^{-\rho} \sim C(\log s_n)^{-\rho}.$$

This provides the left-hand inequality with exponent ρ; in particular, we deduce $s_{n+1} - s_n \to \infty$. ∎

Remark. We can add a small precision to (2.4): for infinitely many pairs (s_n, s_{n+1}), we have

$$\frac{s_{n+1} - s_n}{s_n} \leq \frac{2 \log 2 \log 3}{\log s_n}.$$

The proof goes as follows: due to the best approximation property of the convergents to α, one easily sees that $2^q, 3^p$ or $3^p, 2^q$ are consecutive terms in S for any convergent p/q to α (according to the parity of the indices) [6]. Now, assume for example $s_n = 2^q < 3^p = s_{n+1}$ to be consecutive in S. Since $|\alpha - p/q| < 1/q^2$, we get

$$\log \frac{s_{n+1}}{s_n} = \log \frac{3^p}{2^q} = \log 3 \times |q\alpha - p| < \frac{\log 3}{q} = \frac{\log 2 \log 3}{\log s_n}.$$

Hence, the result follows since

$$\frac{s_{n+1} - s_n}{s_n} \leq 2 \log \frac{s_{n+1}}{s_n}.$$

2.2. Hartman uniform distribution and Furstenberg conjecture.

It is amazing to notice that the Furstenberg conjecture would be implied by the assertion that the Furstenberg set is Hartman uniformly distributed. But the Furstenberg set is not Hartman uniformly distributed, as we will see.

We start with a few reminders on a notion of distribution for *sequences of integers*, studied by Hartman. Let $E := (k_n) \subset \mathbb{Z}$ and $E_N = E \cap [1, \ldots, N]$. Recall that $\mathbb{T}$ denotes the set of reals modulo one with $\mathbb{Z}$ as its dual group.

Definition 2.1. We say that the sequence E is Hartman uniformly distributed (H-ud in short) if one of the following equivalent conditions is satisfied:

(i) For every $x \in \mathbb{T} \setminus \{0\}$,

$$\frac{1}{N} \sum_{n=1}^{N} e(k_n x) \to 0$$

when $N \to \infty$.

(ii) The probability measure

$$\frac{1}{N} \sum_{n \leq N} \delta_{k_n}$$

is w*-convergent to the Haar measure of the Bohr compactification $\overline{\mathbb{Z}}$ of the group $\mathbb{Z}$.

In particular, such a sequence is dense in the Bohr compactification of $\mathbb{Z}$, the dual group of $\mathbb{T}$ equipped with the discrete topology.

Example 1. Of course, $E = \mathbb{N}$ is H-ud; but the set of squares is not. Take $x = 1/4$ and note that

$$\frac{1}{N} \sum_{n=1}^{N} e(n^2 x) \to \frac{1+i}{2}.$$

Example 2. Let us consider the Rudin–Shapiro sequence (r_n), which is defined by $r_n = (-1)^{u_n}$, where

$$u_n = \sum_{k \geq 0} \epsilon_k(n)\epsilon_{k+1}(n) \quad \text{for } n = \sum_{k \geq 0} \epsilon_k(n)2^k$$

(the dyadic expansion of n). It is known that

$$\| \sum_{n \leq N} r_n e_n \|_\infty \ll \sqrt{N}$$

(see [69]). We choose as $E = (n_k)$ the set of occurrences of 1 in the sequence, i.e. $r_{n_k} = 1$, $r_n = -1$ else. This set E is H-ud. Indeed, observe that

$$\sum_{n \leq N} r_n e(nx)) = \sum_{n \in E_N} e(nx) - \sum_{n \leq N, n \notin E_N} e(nx) =: S_N'(x) - S_N''(x).$$

In addition, we have

$$\sum_{n \leq N} e(nx) = S_N'(x) + S_N''(x).$$

It follows that

$$\sum_{n \in E_N} e(nx) = \frac{1}{2}\left(\sum_{n \leq N} e(nx) + \sum_{n \leq N} r_n e(nx)\right) = O(1) + O(\sqrt{N})$$

for $x \neq 0$. Then we conclude by using the relation (the above first equality when $x = 0$):

$$2|E_N| = N + \sum_{n \leq N} r_n \approx N.$$

Example 3. The Thue–Morse sequence (t_n) is defined by $t_n = (-1)^{s(n)}$, where

$$s(n) = \sum_{k \geq 0} \epsilon_k(n) \quad \text{for } n = \sum_{k \geq 0} \epsilon_k(n)2^k$$

(the dyadic expansion of n). It is known that

$$\left\|\sum_{n \leq N} t_n e_n\right\|_\infty \ll N^{\log 3/\log 4}$$

(see [34]). By the same argument as in Example 2, we can prove that the set $E = (n_k)$ of occurrences of 1 in the sequence (t_n) is H-ud. The result in [28] on the supremum norm of trigonometric polynomials with generalized Thue–Morse coefficients can be similarly used to get more H-ud sequences from generalized Thue–Morse sequences.

Furstenberg conjecture. Making use of the fact $s_{n+1}/s_n \to 1$, Furstenberg proved that *every closed infinite set of $\mathbb{T}$ which is S-invariant must be $\mathbb{T}$ itself* (see [11]) ("F is S-invariant" meaning $2F \subset F$ and $3F \subset F$). See [13] for effective versions of this result. Of course there are lots of closed infinite proper subsets only 2-invariant (or 3-invariant) and this deep difference led Furstenberg to his famous metric conjecture:

A continuous probability measure on $\mathbb{T}$ which is S-invariant must be the Lebesgue measure m.

(That μ is S-invariant or $\times 2 \times 3$-invariant means $\hat{\mu}(sn) = \hat{\mu}(n)$ for every $s \in S, n \in \mathbb{Z}$.) It is interesting to point out a connection between Hartman uniform distribution and this latter conjecture.

If $S = (s_n)$ *were Hartman uniformly distributed, in other terms, if we had*

$$(18) \qquad P_N(x) := \frac{1}{N}\sum_{n=1}^N e(s_n x) \to 0 \quad \text{for } x \in \mathbb{T} \setminus \{0\},$$

then the Furstenberg conjecture would be true.

Indeed, suppose that S is H-ud and that there exists a $\times 2 \times 3$-invariant continuous probability measure μ different from the Lebesgue measure. Then there exists a positive integer $a \neq 0$ such that $\hat{\mu}(a) \neq 0$. Consider the probability measure $\nu = \mu \circ \sigma_a^{-1}$, the image of μ under the transformation $\sigma_a \colon \mathbb{T} \to \mathbb{T}$ defined by $\sigma_a(x) = ax \bmod 1$, which is still continuous. Since $\hat{\nu}(n) = \hat{\mu}(an)$ for all $n \in \mathbb{Z}$, the measure ν is again $\times 2 \times 3$-invariant, with $\hat{\nu}(1) \neq 0$. Now, by the $\times 2 \times 3$-invariance of ν, we observe that

$$\int_{\mathbb{T}} P_N(-x) d\nu(x) = \frac{1}{N} \sum_{n=1}^{N} \hat{\nu}(s_n) = \hat{\nu}(1) \neq 0.$$

While, by (18) and Lebesgue's dominated convergence theorem, the left-hand side tends to $\nu(\{0\})$, a contradiction since ν is continuous.

We shall see that the Furstenberg set S is far from H-ud, which is a heuristic indication that Furstenberg's conjecture could be difficult to prove. We enlarge to the Weyl uniform distribution mod 1 of the orbits $(k_n x)$ for some sequence (k_n) of integers. Thanks to the classical Weyl's theorem, this means

$$\forall h \in \mathbb{Z} \setminus \{0\}, \quad \frac{1}{N} \sum_{n < N} e(hk_n x) \to 0 \quad \text{a.e.}$$

Observe the relation between the Weyl uniform distribution and the Hartman uniform distribution: *if (k_n) is H-ud, then the sequence $(k_n x)$ is uniformly distributed* mod 1 *for every irrational x.* The Furstenberg sequence $\{s_n\}$ is not H-ud, because there are uncountably many x such that $\{s_n x\}$ is not Weyl uniformly distributed, as we will see below.

2.3. The set $W(S)$ of non-S-normal numbers. We focus on the negligible set of non-S-normal numbers:

$$W(S) := \left\{ x \in \mathbb{T} : \{s_n x\} \text{ is not uniformly distributed} \right\}$$

At the end of his article [33], Furstenberg pointed out that $\{s_n x\}$ is not uniformly distributed mod 1 for some suitable Liouville number. As an immediate consequence, *the set S is not Hartman uniformly distributed.* We can say more than the above remark of Furstenberg. First, by constructing a lot of such Liouville numbers, we observe that the set $W(S)$ is uncountable; this was already noticed in [11] and recently, Badea and Grivaux [2] recover this fact in a non-explicit way (see also [3]). We can go further by examining the Hausdorff dimension of $W(S)$. In this respect, Erdős and Taylor [24, 26] proved that the negligible set of non-Λ-normal numbers has Hausdorff dimension 1 when $\Lambda \subset \mathbb{Z}$ is a Hadamard set; this result holds true in many other cases, for example

the set $W(\Lambda)$ for $\Lambda = \{3^j + 3^k\}_{j,k \geq 0}$ has the same dimension 1 although Λ is not Hadamard [57]. Boshernitzan [11] announced the result $\dim_H W(S) = 1$, due to Berend. However, Berend (personal communication) claims that he never thought he could prove it.

Theorem 2.7. *The Hausdorff dimension of $W(S)$ is ≥ 0.451621.*

Proof. We are going to construct inside $W(S)$ a kind of Cantor set, called a homogeneous Moran set [30, 31]. Here is a general construction (cf. [30, 31]). Let $(n_k)_{k \geq 1}$ be a sequence of positive integers with $n_k \geq 2$ and $(c_k)_{k \geq 1}$ be a sequence of positive numbers satisfying $0 < c_k < 1$ and $n_k c_k \leq 1$ ($k \geq 1$). For $k \geq 1$, let

$$D_k = \{(i_1, \dots, i_k) : 1 \leq i_j \leq n_j \text{ for } 1 \leq j \leq k\}$$

and

$$D = \bigcup_{k \geq 0} D_k \quad \text{with } D_0 = \emptyset.$$

For $\tau = (\tau_1, \dots, \tau_n) \in D_n$ and $\sigma = (\sigma_1, \dots, \sigma_m) \in D_m$, we define

$$\tau * \sigma = (\tau_1, \dots, \tau_n, \sigma_1, \dots, \sigma_m) \in D_{n+m}.$$

Suppose we are given an interval J of length 1. We can construct a family of subintervals $\mathcal{F} := \{J_\sigma : \sigma \in D\}$ of J as follows. The construction will be made for $\{J_\sigma : \sigma \in D_k\}$ inductively on k. First for $k = 0$, we choose $J_\emptyset = J$. Now suppose that $\{J_\sigma : \sigma \in D_{k-1}\}$ are constructed. For each $\sigma \in D_{k-1}$, we choose n_k subintervals $J_{\sigma*1}, \dots, J_{\sigma*n_k}$ in J_σ such that they have disjoint interiors and

$$\forall 1 \leq j \leq n_k, \quad \frac{|J_{\sigma*j}|}{|J_\sigma|} = c_k.$$

The set $E := \bigcap_{k \geq 1} \bigcup_{\sigma \in D_k} J_\sigma$ is called a *homogeneous Moran set* determined by $\mathcal{F}$. Notice that for given sequence (n_k) and (c_k), there are different constructions of $\mathcal{F}$ and then different Moran sets, because the positions $J_{\sigma*1}, \dots, J_{\sigma*n_k}$ in J_σ are arbitrary to some extent. There is a common lower bound for the Hausdorff dimensions of these Moran sets, given by the following proposition.

Lemma 2.8 ([30, 31]). *For the homogeneous Moran set defined above, we have*

$$\dim_H E \geq \liminf_{n \to \infty} \frac{\log n_1 n_2 \cdots n_k}{-\log c_1 c_2 \cdots c_{k+1} n_{k+1}}.$$

Now let us construct a Moran set in our set $W(S)$. Let us fix a sequence of integers (ℓ_k) which is quickly increasing in the sense that

$$\ell_0 = 0, \quad \frac{\ell_1 + \cdots + \ell_{k-1}}{\ell_k} \to 0.$$

For a small $\delta > 0$, define $m_k = \delta\ell_k$ so that $\ell_{k-1} < m_k < \ell_k$. We consider

$$E := \left\{ \sum_{j=1}^{\infty} \frac{a_j}{6^j} : a_j = 0 \text{ when } j \in (m_k, \ell_k) \text{ for some } k, 0 \leq a_j \leq 5 \text{ otherwise} \right\}.$$

The set E is a Moran set with

$$n_k = 6^{m_k - \ell_{k-1}}, \quad c_k = \frac{1}{6^{m_k - m_{k-1}}}.$$

Indeed, in the definition of

$$x = \sum_{j=1}^{\infty} a_j 6^{-j} \in E,$$

the digits a_j are free (meaning they can take any value between 0 and 5) for $j \in [\ell_{k-1}, m_k]$. So there are a number n_k of possible strings of digits (a_j) with $j \in [\ell_{k-1}, m_k]$ corresponding to subintervals of an $(k-1)$-level interval. We have then

$$\frac{\log n_1 n_2 \cdots n_k}{-\log c_1 c_2 \cdot c_{k+1} n_{k+1}} = \frac{\sum_{j=1}^{k}(m_j - \ell_{j-1})}{\sum_{j=1}^{k+1}(m_j - m_{j-1}) + \ell_k - m_{k+1}}$$

$$= \frac{\sum_{j=1}^{k}(m_j - \ell_{j-1})}{\ell_k}.$$

By Lemma 2.8,

$$\dim_H E \geq \liminf \frac{\sum_{j=1}^{k}(m_j - \ell_{j-1})}{\ell_k} = \liminf \frac{\sum_{j=1}^{k}(\delta\ell_j - \ell_{j-1})}{\ell_k} = \delta.$$

Next we show that $E \subset W(S)$ if $\delta > 0$ is small enough. Let us choose $p_k = (1-\epsilon)\ell_k$ with $\epsilon > 0$ small enough to have $m_k < p_k < \ell_k$. Observe that

$$\ell_k - p_k = \epsilon\ell_k \to \infty, \quad \frac{m_k}{p_k} \leq \frac{\delta}{1-\epsilon}.$$

Let $x = \sum_{j=1}^{\infty} a_j 6^{-j} \in E$. For every k, we have

$$(19) \qquad x = \sum_{j=1}^{m_k} \frac{a_j}{6^j} + \sum_{j=\ell_k}^{\infty} \frac{a_j}{6^j} =: a_k^* 6^{-m_k} + O(6^{-\ell_k}),$$

where a_k^* is an integer. Then we put $N_k = 6^{m_k}$, $N = 6^{p_k}$ and consider

$$(20) \qquad T_N = \sum_{2^i 3^j < N} e(2^i 3^j x) =: u_N + v_N + W_N - z_N,$$

with

$$u_N = \sum_{i < m_k,\, 2^i 3^j < N} e(2^i 3^j x), \qquad v_N = \sum_{j < m_k,\, 2^i 3^j < N} e(2^i 3^j x),$$

$$W_N = \sum_{i,j \geq m_k,\, 2^i 3^j < N} e(2^i 3^j x), \qquad z_N = \sum_{i < m_k,\, j < m_k} e(2^i 3^j x).$$

It is clear that (notice that $\log N = p_k \log 6$)

$$(21) \qquad |u_N| \leq \frac{\log 6}{\log 3} m_k p_k, \qquad |v_N| \leq \frac{\log 6}{\log 2} m_k p_k, \qquad z_N \leq m_k^2.$$

Now, if $i \geq m_k$, $j \geq m_k$ and $2^i 3^j < N$, from (19) we get

$$e(2^i 3^j x) = 1 + O\left(\frac{N}{6^{\ell_k}}\right) = 1 + O(6^{p_k - \ell_k}) = 1 + o(1),$$

because $\ell_k - p_k \to \infty$. It follows that

$$(22) \qquad W_N = C \log^2 N + o(\log^2 N) \quad \text{with } C = \frac{1}{2 \log 2 \log 3}.$$

On the other hand,

$$\frac{m_k p_k}{\log^2 N} \leq \frac{\delta}{(1 - \epsilon) \log^2 6}, \qquad \frac{m_k^2}{\log^2 N} \leq \frac{\delta^2}{(1 - \epsilon)^2 \log^2 6}.$$

Combining this, (21) and (22) results in

$$\liminf_{k \to \infty} \frac{|T_N|}{\log^2 N} \geq C - \left(\frac{\log 6}{\log 3} + \frac{\log 6}{\log 2}\right) \frac{\delta}{(1 - \epsilon) \log^2 6} - \frac{\delta^2}{(1 - \epsilon)^2 \log^2 6} > 0$$

if $\delta > 0$ is small enough. This proves $x \in W(S)$.

An explicit lower bound of the dimension can be obtained as the positive solution of

$$C - \frac{\delta}{\log 2 \log 3} - \frac{\delta^2}{\log^2 6} = 0,$$

i.e. $\delta^2 + 2a\delta - a = 0$ with $a = \frac{\log^2 6}{2 \log 2 \log 3}$, whence $\delta = 0.451621\ldots.$ ∎

2.4. $W(S)$ is an M_0-set. A M_0-set is a Borel subset of $\mathbb{T}$ that supports a non-trivial Rajchman measure. Recall that a Rajchman measure (or M_0-measure) on $\mathbb{T}$ is a measure μ such that

$$\lim_{|n| \to \infty} \widehat{\mu}(n) = 0.$$

The question "what does the support of a Rajchman measure look like?" was asked by Kahane and Salem [41, p. 59]. Russel Lyons [53] observed that a measure annihilating all non-normal sets $W(\Lambda)$ must be a Rajchman measure, but the converse is not true, since a Rajchman measure can be constructed on the set of non-normal numbers to base 2. The construction relies on the 2-adic expansion and the divisibility property of $\{2^n\}$. By using the 6-adic expansion as above, we can exhibit a Bernoulli-like probability measure which is Rajchman and concentrated on the set $W(S)$.

Theorem 2.9. *The set $W(S)$ supports a Rajchman measure, briefly, $W(S)$ is a M_0-set.*

Proof. The candidate will be the distribution $\mu := \mathbb{P}_X$ of a series of discrete random variables

$$X := \sum_{j=1}^{\infty} \varepsilon_j X_j,$$

where (ε_j) is a sequence of Bernoulli variables such that

$$\mathbb{P}(\varepsilon_j = 0) = 1/j \log j, \quad \mathbb{P}(\varepsilon_j = 1) = 1 - 1/j \log j,$$

and the variables X_j are to be defined, all the variables involved being independent. We fix a lacunary sequence of integers (m_k) so that $m_k/m_{k+1} \to 0$ and we put $n_k := m_{k+1} - m_k$; the variable X_k is equidistributed on the finite set

$$\left\{ \frac{a}{6^{m_{k+1}}} : 0 \le a \le 6^{n_k} - 1 \right\},$$

namely

$$\mathbb{P}(X_k = a/6^{m_{k+1}}) = \frac{1}{6^{n_k}}.$$

For every ω, $X(\omega)$ is a 6-adic expansion, with a large block of arbitrary digits. We have to check that μ is concentrated on $W(S)$ and that $\mu \in M_0$.

First, we prove that

$$\limsup \frac{1}{|S_N|} \left| \sum_{2^i 3^j \le N} e(2^i 3^j X) \right| > 0 \quad \text{a.s.},$$

which implies that X a.s. takes its values in $W(S)$ (recall that $S_N = S \cap [1, \ldots, N]$).

As in (20), we write

$$\sum_{2^i 3^j \le N} e(2^i 3^j X) = u_N + v_N + W_N - z_N$$

with now

(23) $$|u_N| \le \frac{\log N}{\log 3} m_k, \quad |v_N| \le \frac{\log N}{\log 2} m_k, \quad |z_N| \le m_k^2.$$

We then diverge from (20) and adopt the following notations:

$$M_k = 6^{m_k}, \quad N_k = 6^{m_{k+1}} 2^{-k-1}, \quad Q_k = M_{k+1}/M_k = 6^{n_k}.$$

Since $\sum P(\varepsilon_k = 0) = \infty$, almost surely $\varepsilon_k = 0$ for infinitely many k's (say for $\omega \in \Omega_0$, $\mathbb{P}(\Omega_0) = 1$). Let now $\omega \in \Omega_0$ and k large with $\varepsilon_k = \varepsilon_k(\omega) = 0$. We take $N = N_k$ in (23). Once more it holds that

$$e(2^i 3^j X) = e\left(2^i 3^j \sum_{n \geq k+1} \varepsilon_n X_n\right) \quad \text{if } i \geq m_k, \, j \geq m_k;$$

and, since $2^i 3^j < N$, for $n \geq k + 1$, we get that

$$2^i 3^j X_n \in \left\{0, \frac{2^i 3^j}{6^{m_{n+1}}}, \dots, 2^i 3^j \left(\frac{1}{6^{m_n}} - \frac{1}{6^{m_{n+1}}}\right)\right\}$$

with

$$2^i 3^j \sum_{n \geq k+1} \frac{1}{6^{m_n}} < \frac{2N}{6^{m_{k+1}}} = \frac{1}{2^k},$$

so that

$$(24) \qquad 2^i 3^j \left(\sum_{n \geq k+1} \varepsilon_n X_n\right) \in [0, 1/2^k) \quad \text{for such a } k.$$

We now focus on the main sum W_N. We begin by observing that

$$N = 6^{m_{k+1}} 2^{-k-1} \in \,]M_k, M_{k+1}[,$$

since $m_{k+1} - m_k \gg k$. We claim that

$$\limsup_N \frac{1}{|S_N|} \left| \sum_{i,j \geq m_k, \, 2^i 3^j < N} e(2^i 3^j X(\omega)) \right| > 0.$$

Indeed, by (24),

$$|W_{N_k}| = \left| \sum_{i,j \geq m_k, \, 2^i 3^j < N_k} e\left(2^i 3^j \sum_{n \geq k+1} \varepsilon_n X_n\right) \right| = \sum_{2^i 3^j < N_k, \, i,j > m_k} 1 + O\left(2^{-k} |S_{N_k}|\right).$$

Combined with (23), this gives for infinitely many k's, that almost surely,

$$\left| \sum_{2^i 3^j \leq N_k} e(2^i 3^j X) \right| = |S_{N_k}| + O\left(2^{-k} |S_{N_k}|\right) + O(m_k m_{k+1}),$$

whence

$$\frac{1}{|S_{N_k}|} \left| \sum_{2^i 3^j \leq N_k} e(2^i 3^j X) \right| = 1 + O(2^{-k}) + O(m_k/m_{k+1}),$$

since $|S_{N_k}| \approx \log^2 N_k \approx m_{k+1}^2$. The claim follows from $\lim_k m_k/m_{k+1} = 0$.

We now prove that $\widehat{\mu}(h) \to 0$ as $h \to \infty$. For $h \neq 0$, we have

$$\widehat{\mu}(h) = \prod_{j=1}^{\infty} \mathbb{E}(\exp 2i\pi h\epsilon_j X_j) = \prod_{j=1}^{\infty} [(1 - 1/j\log j)c_j + 1/j\log j],$$

where $c_j = \mathbb{E}(\exp 2\pi ihX_j)$. Clearly,

$$c_j = \frac{1}{Q_j} \sum_{a=0}^{Q_j-1} e^{2\pi ih(a/M_{j+1})} = \frac{1}{Q_j} \frac{e^{2\pi i(hQ_j/M_{j+1})} - 1}{e^{2\pi i(h/M_{j+1})} - 1}$$

and

$$|c_j| = \frac{M_j|\sin(\pi h/M_j)|}{M_{j+1}|\sin(\pi h/M_{j+1})|}.$$

Assuming $h > 0$, we fix the index k such that $M_k/2 < h \leq M_{k+1}/2$; then,

$$2h > M_k, \quad \pi h/M_{k+1} \leq \pi/2,$$

whence

$$M_{k+1}\sin(\pi h/M_{k+1}) \geq 2h$$

by concavity of the sine function on $[0, \pi/2]$. It ensues that

$$|c_{k-1}c_k| = M_{k-1}|\sin(\pi h/M_{k-1})|M_{k+1}|\sin(\pi h/(M_{k+1}))| \leq \frac{M_{k-1}}{2h} \leq \frac{M_{k-1}}{M_k} =: \delta_k;$$

this implies that one term in the product $\prod_j[(1 - 1/j\log j)c_j + 1/j\log j$ must be small when h becomes large: indeed, one among both coefficients $|c_k|$ and $|c_{k-1}|$, say $|c_k|$, must be $\leq \sqrt{\delta_k}$ so that

$$|\widehat{\mu}(h)| \leq |(1 - 1/k\log k)c_k + 1/k\log k|$$
$$\leq \sqrt{\delta_k} + \frac{1}{k\log k} = O\left(\frac{1}{k\log k}\right) = O\left(\frac{1}{\log\log h}\right)$$

by choosing $m_k = k!$ and using the Stirling formula. The exponent 1 in the last iterated logarithm can be shown to be optimal by the Davenport–Erdős–Levêque criterion [15, Chapter 1]. ∎

Comments. The property for a set to support a Rajchman measure says more on the *lack of porosity* of this set than on its size. This is perhaps an illustration of the uncertainty principle for measures: actually, $W(S)$ is rather big as it has positive Hausdorff dimension, but so is the porous triadic Cantor set K with Hausdorff dimension $\log 2/\log 3$ which however supports no Rajchman measure [42, p. 59]. We provide here a direct proof of the last fact based on the following property of K: The Cantor set K is

σ_3-invariant, in particular, for every $x \in K$ and $k \geq 1$, $3^k x \notin I :=]1/3, 2/3[$. Suppose that K supports a non-trivial M_0-measure μ. Without loss of generality we assume that $\hat{\mu}(0) \neq 0$ (otherwise we can consider the measure $e^{-2\pi i a t} \mu$ instead of μ, with a such that $\hat{\mu}(a) \neq 0$). Let now f be the sum of an absolutely summable trigonometric series ($f \in A(\mathbb{T})$), with its closed support inside I and $\hat{f}(0) = 1$; we then define $f_k \in A(\mathbb{T})$ by $f_k(t) = f(3^k t)$: for every $k \geq 0$, the spectrum of f_k is contained into $3^k \mathbb{Z}$ and its closed support stays away from K. It follows that

(25)
$$\forall n \neq 0, \quad \lim_{k \to \infty} \hat{f_k}(n) = 0$$

and

(26)
$$0 = \int f_k \, d\mu = \sum_n \hat{f_k}(n) \overline{\hat{\mu}(n)}.$$

Since $\hat{\mu}(n) \to 0$ as $|n| \to \infty$, given $\epsilon > 0$, there exists N such that $|\hat{\mu}(n)| \leq \epsilon$ for $|n| > N$, implying

$$\left| \sum_{|n|>N} \hat{f_k}(n) \overline{\hat{\mu}(n)} \right| \leq \epsilon \sum_n |\hat{f_k}(n)| = \epsilon \sum_n |\hat{f}(n)| =: C\epsilon,$$

whence, from (26),

$$\left| \sum_{|n|\leq N} \hat{f_k}(n) \overline{\hat{\mu}(n)} \right| = \left| \sum_{|n|>N} \hat{f_k}(n) \overline{\hat{\mu}(n)} \right| \leq C\epsilon.$$

However, $N = N_\epsilon$ being fixed, (25) leads to the contradiction (for ϵ small)

$$0 < |\hat{f}(0) \overline{\hat{\mu}(0)}| = \lim_{k \to \infty} \left| \sum_{|n|\leq N} |\hat{f_k}(n) \overline{\hat{\mu}(n)} \right| \leq C\epsilon.$$

Another striking example is the set of Liouville numbers which supports a Rajchman measure [10, 14], though of zero Hausdorff dimension.

2.5. Khinchin class of the Furstenberg set S. From the classical result of the equidistribution mod 1 of (nx), $x \notin \mathbb{Q}$, it follows that

$$\frac{1}{N} \sum_{n<N} f(nx) \to \int_{\mathbb{T}} f \, dm \quad \text{a.e.}$$

for every continuous, or even Riemann-integrable function f. Khinchin [43] conjectured that the result still holds with any function in $L^1(\mathbb{T})$. But, Marstrand [55] proved that it fails for some function in $L^\infty(\mathbb{T})$. However the class of L^1-functions satisfying

Khinchin's conjecture deserves to be explored and, many contributions took this path (see [8]). Let us cite one early result due to Koksma [44]: any L^2-function such that

$$\sum_{n=1}^{\infty} |\hat{f}(n)|^2 (\log \log n)^3 < \infty$$

satisfies Khinchin's conjecture.

We are interested in subsequences (k_n) of the sequence of positive integers, especially the Furstenberg sequence (s_n). We propose the following definition.

Definition 2.2. The Khinchin class of an increasing sequence $(k_n)_{n \geq 1}$ of integers is the class of functions $f \in L^1(\mathbb{T})$ satisfying

$$(27) \qquad \lim_{N \to \infty} \frac{1}{N} \sum_{n=2}^{N} f(k_n x) = \int_{\mathbb{T}} f \, dm \quad \text{a.e.}$$

The question of Khinchin, related to other sequences of integers, has been fruitfully studied by Marstrand in the above cited paper [55]. He proved the following (among others).

Theorem 2.10 (Marstrand). *Every function $f \in L^\infty(\mathbb{T})$ is in the Khinchin class of the Furstenberg sequence S.*

The assumption $f \in L^\infty(\mathbb{T})$ was later on dropped by Nair [56] in a difficult paper: assuming $f \in L^1(\mathbb{T})$ is enough. The Khinchin class of (s_n) is thus the whole of $L^1(\mathbb{T})$.

Using an ergodic argument, we give below a simple proof of Marstrand's result, which nearly recaptures Nair's generalization and holds in the more abstract context of a $\times 2 \times 3$-invariant probability measure, and S-ergodic (if any such measure different from m exists). A measurable function $f \colon \mathbb{T} \to \mathbb{C}$ is in the class $L \log^+ L$ with $L := L^1(\mu)$ if

$$\int_{\mathbb{T}} |f| \log^+ |f| \, d\mu < \infty.$$

Theorem 2.11. *Let μ be a S-invariant and S-ergodic probability measure. If $f \in L \log^+ L$, then,*

$$A_N f(x) := \frac{1}{N} \sum_{n \leq N} f(s_n x) \to \int f \, d\mu \quad \mu\text{-a.e.}$$

In particular,

$$\frac{1}{N} \sum_{n \leq N} \delta_{s_n x} \to \mu \text{ weakly*} \quad \mu\text{-a.e.}$$

Consequently, if Furstenberg's conjecture were not correct, there would exist an S-invariant probability measure carried by $W(S)$.

Observe that S-ergodicity means this: $h \circ \sigma_2 = h$ *and* $h \circ \sigma_3 = h$ imply h constant μ-a.e.; and this does not infer 2- or 3-ergodicity separately.

Proof. We denote by σ_q the q-shift: $x \mapsto qx$ on $\mathbb{T}$ and put $\sigma_q f(x) = f(qx)$. Now, assume that $f \in L^1(\mu)$ and $\int f \, d\mu = 0$; then we set $M = \log s_N / \log 2$ and we decompose, as in the proof of Proposition 2.2,

$$\sum_{n \leq N} f(s_n x) = \sum_{2^j 3^k \leq s_N} f(2^j 3^k x) = \sum_{j + k\alpha^{-1} \leq M} f(2^j 3^k x)$$

$$= \sum_{j \leq M} \sum_{k \leq (M-j)\alpha} f(2^j 3^k x) = \sum_{j \leq M} \sigma_2^j \Big(\sum_{k \leq (M-j)\alpha} f(3^k x) \Big).$$

Observing that $M \approx \sqrt{N}$, we get

$$|A_N f(x)| \lesssim \Big| \frac{1}{M^2} \sum_{j \leq M} \sigma_2^j \Big(\sum_{k \leq (M-j)\alpha} f(3^k x) \Big) \Big|$$

$$\lesssim \Big| \frac{1}{M} \sum_{j \leq M} \sigma_2^j \Big(\frac{M-j}{M} F_{M-j}(x) \Big) \Big|$$

where

$$F_K(x) := \frac{1}{K\alpha} \sum_{k \leq K\alpha} f(3^k x).$$

By Birkhoff's theorem applied to the dynamical system $(\mathbb{T}, \sigma_3, \mu)$, we know that $F_K(x) \to \ell(x)$ μ-a.e., where $\ell = \mathbb{E}_3(f)$ is the conditional expectation with respect to the 3-invariant sets. We thus decompose

$$T_M = \frac{1}{M} \sum_{j \leq M} \sigma_2^j \Big(\frac{M-j}{M} \big(F_{M-j}(x) - \ell(x) \big) \Big) + \frac{1}{M} \sum_{j \leq M} \frac{M-j}{M} \sigma_2^j \ell(x)$$

$$=: T'_M + T''_M.$$

It is easy to see that the second mean $T''_M \to 0$ by the following remark.

Lemma 2.12. *Let (z_j) be a sequence of complex numbers such that $Z_j / j \to 0$ with $Z_j = z_1 + \cdots + z_j$. Then*

$$\lim_{M \to \infty} \frac{1}{M} \sum_{j=1}^{M} \Big(1 - \frac{j}{M} \Big) z_j = 0.$$

The proof of this lemma is simple:

$$\frac{1}{M} \sum_{j=1}^{M} \Big(1 - \frac{j}{M} \Big) z_j = \frac{1}{M^2} \sum_{j=1}^{M-1} Z_j = o\Big(\sum_{j=1}^{M-1} j \Big) / M^2 = o(1).$$

Indeed, we just apply this lemma with $z_j = \sigma_2^j \ell(x)$: Birkhoff's theorem gives

$$Z_j/j \to \mathbb{E}_2(\ell) = \mathbb{E}_2 \mathbb{E}_3(f) \quad \mu\text{-a.e.},$$

since μ is 2-invariant. As $\mathbb{E}_2 \mathbb{E}_3(f)$ is S-invariant measurable function (observe that both projections $\mathbb{E}_2$ and $\mathbb{E}_3$ are commuting) and μ is assumed S-ergodic, $\mathbb{E}_2 \mathbb{E}_3(f)$ must be constant μ-a.e; so we get

$$\mathbb{E}_2 \mathbb{E}_3(f) = \int f \, d\mu = 0,$$

then $Z_j/j \to 0$ and finally $T_M'' \to 0$.

Now, we look at the first mean T_M', which satisfies

$$|T_M'| \leq \frac{1}{M} \sum_{j \leq M} \sigma_2^j \left| F_{M-j}(x) - \ell(x) \right|.$$

We will be done if we can prove that the right-hand side tends to zero μ-a.e.:

$$(28) \qquad \lim_{M \to \infty} \frac{1}{M} \sum_{j \leq M} \sigma_2^j |F_{M-j}(x) - l(x)| = 0 \quad \mu\text{-a.e.}$$

We will prove (28) by making use of the following improvement of Birkhoff's theorem [60, Chapter VI, p. 262)]. This improvement in [60] was not explicitly stated and was proved under the assumption of ergodicity. But this assumption is not necessary. For clarity we give a complete proof.

Lemma 2.13. *Let T be measure preserving transformation on $(X, \mathcal{B}, \mu)$ and let $(G_n) \subset L^1(\mu)$. Suppose that $G_n \to G$ a.e. and $G^* := \sup_n |G_n| \in L^1(\mu)$. Then*

$$(29) \qquad \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} |G_{n-k} - G| \circ T^k = 0 \quad \mu\text{-a.e.}$$

In particular, we have

$$(30) \qquad \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} G_{n-k} \circ T^k = \mathbb{E}(G|\mathcal{I}) \quad \mu\text{-a.e.},$$

where $\mathcal{I}$ is the σ-field of T-invariant sets.

To prove (29), we can assume $G = 0$ and $G_k \geq 0$ for all k. Let $g_n = \sup_{k \geq n} G_k$ for $n \geq 1$; by hypothesis, $g_n \leq g_1 = G^* \in L^1(\mu)$ for every n. Fix $1 \leq m < N$.

We decompose

$$\frac{1}{N} \sum_{k=0}^{N-1} G_{N-k} \circ T^k = \frac{1}{N} \sum_{k=0}^{N-m} G_{N-k} \circ T^k + \frac{1}{N} \sum_{k=N-m+1}^{N-1} G_{N-k} \circ T^k$$

$$\leq \frac{1}{N} \sum_{k=0}^{N-m} g_m \circ T^k + \frac{1}{N} \sum_{k=N-m+1}^{N-1} G_{N-k} \circ T^k =: A'_N + A''_N.$$

The term A''_N involves only $m-1$ functions and satisfies

$$0 \leq \limsup_{N\to\infty} A''_N = \limsup_{N\to\infty} \frac{1}{N} \sum_{j=1}^{m-1} G_j \circ T^{N-j} = 0 \quad \mu\text{-a.e.}$$

(by Borel–Cantelli's lemma, $\frac{1}{N}(h \circ T^N) \to 0$ μ-a.e. for every $h \in L^1(\mu)$). Switching to the first term A'_N, we apply Birkhoff's theorem to g_m and we get

$$\limsup_{N\to\infty} A'_N \leq \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-m} g_m \circ T^k = \mathbb{E}(g_m | \mathcal{I}) \quad \mu\text{-a.e.}$$

with $\mathcal{I}$ standing for the σ-field of T-invariant sets. Finally, we obtain

$$0 \leq L := \limsup_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} G_{N-k} \circ T^k \leq \mathbb{E}(g_m | \mathcal{I}) \quad \mu\text{-a.e.,}$$

and hence

$$0 \leq \mathbb{E}(L) \leq \mathbb{E}(g_m) \quad \text{for every } m.$$

Remember that $g_m \to 0$ μ-a.s. and $g_m \leq G^* \in L^1(\mu)$, so that the dominated convergence theorem applies: we get $\mathbb{E}(g_m) \to 0$, $\mathbb{E}(L) = 0$ and $L = 0$ μ-pp. We have thus proved (29), which, together with Birkhoff's theorem, implies (30):

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} G_{n-k} \circ T^k = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} G \circ T^k = \mathbb{E}(G | \mathcal{I}) \quad \mu\text{-a.e.}$$

Applied with $T = \sigma_2$ and $G_n = F_n$, $G = l$, the lemma gives the result. It remains to check that $\sup_K |F_K| \in L^1$ under our assumptions on f. And this is nothing but the classical maximal ergodic theorem, valid for any dynamical system $(X, \mathcal{B}, \mu, T)$, that we recall below (see [22, Chapter II, p. 38] and [59, Chapter IV, pp. 70–71] for the consequence).

Lemma 2.14. *Given any dynamical system $(X, \mathcal{B}, \mu, T)$, $g \in L^1(\mu)$ non-negative and $G_n = \frac{1}{n} \sum_{k<n} T^k g$, the maximal functions $G_N^* = \sup_{n \le N} G_n$ $(N \ge 1)$ satisfy*

$$\mu(G_N^* > t) \le \frac{1}{t} \int_X g \mathbf{1}_{\{G_N^* > t\}} \, d\mu \quad \text{for all } t > 0.$$

As a consequence, the maximal function $G^* = \sup_n G_n$ is in $L^1(\mu)$ as soon as g is in $L \log^+ L$.

The equality (28) is thus proved. Finally, if an S-invariant probability measure μ exists with $\mu \ne m$, since μ is a barycenter of S-invariant and S-ergodic probability measures, we can assume it to be S-ergodic as well, and there exists an integer $k \ne 0$ such that $\widehat{\mu}(k) \ne 0$. Applying Theorem 2.11 to $f = e_{-k}$, we get that, for μ almost all x:

$$\frac{1}{N} \sum_{n=1}^{N} e(-k s_n x) \to \int_{\mathbb{T}} e_{-k} \, d\mu = \widehat{\mu}(k) \ne 0,$$

so that $\mu(W(S)) = 1$. ∎

3. Thin sets in harmonic analysis

We first make a short reminder on Orlicz functions and spaces encountered just before, to be used later in this work.

(1) An Orlicz function $\varphi: [0, \infty[\to \mathbb{R}+$ is a non-negative, increasing and convex function with $\varphi(0) = 0$.

(2) If $(X, \mathcal{B}, \mu)$ is a probability space and φ an Orlicz function, the associated Orlicz space L^φ is defined as

$$L^\varphi = \left\{ f \text{ measurable} : \int_X \varphi\left(\frac{|f(x)|}{a}\right) d\mu(x) < \infty \text{ for some } a > 0 \right\}.$$

The Orlicz (Luxemburg) norm $\|f\|_{L^\varphi}$ of f is the infimum of those $a > 0$ such that

$$\int_X \varphi\left(\frac{|f(x)|}{a}\right) d\mu(x) \le 1.$$

(3) If φ moreover satisfies the Δ_2-condition, namely $\varphi(2x) \le C\varphi(x) \ \forall x > 0$, then

$$f \in L^\varphi \iff \int_X \varphi(|f(x)|) \, d\mu(x) < \infty.$$

Examples. (a) With $\varphi(x) = x^p$ $(1 \le p < \infty)$, we recover the usual Banach space L^p.

(b) Take $\varphi(x) = \varphi_\alpha(x) := x \log^\alpha(1 + x)$ $(0 < \alpha \le 1)$. This function is an Orlicz function satisfying Δ_2. Clearly, $L^{\varphi_\alpha} \hookrightarrow L^1$.

We next coin combinatorial tools which reveal adapted to spectral harmonic analysis [48].

3.1. Dissociate and quasi-independent sets.

Definition 3.1. A set $E \subset \mathbb{Z} \setminus \{0\}$ is said to be *quasi-independent* if for all *distinct* elements $x_1, \ldots, x_n \in E$ and for all $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, 0, 1\}$, the *relation* $\sum \varepsilon_k x_k = 0$ implies that all $\varepsilon_k = 0$. The quantity $\sum_k |\varepsilon_k|$ is called the *length* of the relation.

The set E is said to be dissociate if for all distinct elements $x_1, \ldots, x_n \in E$ and $\varepsilon_1, \ldots, \varepsilon_n \in \{-2, -1, 0, 1, 2\}$, $\sum \varepsilon_k x_k = 0$ implies all $\varepsilon_k = 0$.

A more comprehensible definition is the following: E is quasi-independent if every finite sum of elements of E admits a unique such decomposition; it is dissociate if every expansion $\sum \varepsilon_k x_k$ with $\varepsilon_k \in \{-1, 0, 1\}$ is unique.

A Hadamard set with ratio q is quasi-independent if $q \ge 2$, dissociate if $q \ge 3$. Every Hadamard set with ratio $q > 1$ is a finite union of dissociate sets.

3.2. p-Sidon sets, $\Lambda(p)$-sets.

1-periodic functions with spectrum in some Hadamard subset of $\mathbb{N}$ enjoy specific convergence and regularity properties. The most famous ones, that we recall here, have led to define the so-called "thin sets" in harmonic analysis.

(1) *A continuous function with spectrum in some Hadamard subset E possesses an absolutely convergent Fourier series.* In other terms, $f \in C_E \implies \hat{f} \in \ell^1$. The sets of integers enjoying this property have been called *Sidon sets* by Rudin [69] and more generally,

$$E \text{ is a } p\text{-Sidon set} \iff [f \in C_E \implies \hat{f} \in \ell^p]$$

for $1 \le p < 2$. (Of course, 1-Sidon = Sidon.) If E is p-Sidon, there exists a best constant $S_p(E)$, called the p-Sidon constant of E, such that

$$\|\hat{f}\|_p \le S_p(E)\|f\|_\infty \quad \text{for every } f \in C_E.$$

(2) *If $f \in L^1$ has a Hadamard spectrum, then f belongs to all the L^p, $p < \infty$.* The sets E enjoying this property for one $p > 2$ have been called $\Lambda(p)$-*sets* (again by Rudin) [69], and we denote by $\lambda_p(E)$ the best constant such that

$$\|f\|_p \le \lambda_p(E)\|f\|_2$$

for every polynomial with spectrum in E. Then, every $f \in L_E^2$ belongs to L^p and the above inequality still holds. See also [48, Vol. 2, pp. 144–151].

First examples. Quasi-independent sets are Sidon sets [69] with Sidon constant bounded by 8. A Hadamard set $E = \{n_j\} \subset \mathbb{N}$ is a $\Lambda(p)$-set for all finite p. More generally, a Sidon set E is a $\Lambda(p)$-set for all $p > 2$ with constant $\lambda_p(E) \leq 2S(E)\sqrt{p} = O(\sqrt{p})$ (also a result of Rudin, the converse being due to Pisier). Later on, Bourgain showed that true $\Lambda(p)$-sets do exist if $p > 2$, that means sets which are $\Lambda(p)$, but not $\Lambda(q)$ as soon as $q > p$. Whence, the restriction to the parameters $p > 2$. The question whether true $\Lambda(2)$-sets do exist is an open problem.

We present below characterizations of p-Sidon set and $\Lambda(p)$-sets.

- There exists a characterization of p-Sidon sets in terms of interpolation.

Theorem 3.1 ([21]). *Let $1 \leq p < 2$ and $E \subset \mathbb{Z}$ (or $\mathbb{N}$). Assume $\frac{1}{p} + \frac{1}{q} = 1$. The following assertions are equivalent:*

(1) *E is p-Sidon.*

(2) *$\sum |\widehat{f}(n)|^p < \infty$ for $f \in L_E^\infty$.*

(3) *If $b \in \ell^q(E)$, there exists $\mu \in M(\mathbb{T})$ such that $\widehat{\mu}_{|E} = b$.*

(4) *If $b \in \ell^q(E)$, there exists $f \in L^1(\mathbb{T})$ such that $\widehat{f}_{|E} = b$.*

Remark. We observe that the harmonic classification is in some sense complementary to the dynamical one detailed in [7] (see also [65, Section 4.3.2]): the implied set of integers E is viewed as a spectrum in the first classification and as a co-spectrum in the second one. In this respect, we can see that *a Sidon set cannot be Hartman-ud.* Indeed, if E is Hartman-ud, then C_E, the space of continuous functions with spectrum inside E, contains a copy of the Banach space c_0 of ultimately null sequences [52]; if E is Sidon, C_E is by definition isomorphic to ℓ^1, which does not contain c_0, hence E will never be Hartman uniformly distributed.

- The question whether every Sidon set is a finite union of quasi-independent sets is an open question, motivated by a characterization of a Sidon set due to Pisier [63]: $E \subset \mathbb{Z} \setminus \{0\}$ *is a Sidon set if and only if there exists $\delta > 0$ such that, from every finite subset $A \subset E$, a quasi-independent set B can be extracted with $|B| \geq \delta|A|$.* An analogue for p-Sidon sets will be considered later (Section 5.4).

- An arithmetical characterization of these "thin sets" – in terms of arithmetical progression or more generally mesh – is not yet complete (if it ever exists, cf. [48, Vol. 2, Ann. D, pp. 316–323]), but we have at our disposal necessary conditions, as this one, called *mesh condition.*

Definition 3.2. For $E \subset \mathbb{Z}$, we denote by $\alpha_E(N)$ the maximal number of elements of $E \cap \{a + b, a + 2b, \ldots, a + Nb\}$ over all arithmetical progressions of length N.

Proposition 3.2 ([69]). *The following hold:*

(1) *If E is $\Lambda(p)$ for $p > 2$, then $\alpha_E(N) \leq 4\lambda_p(E)^2 N^{2/p}$.*

(2) *If E is p-Sidon, then $\alpha_E(N) \leq c(\log N)^{p/(2-p)}$.*

Remarks. (1) The $\Lambda(q)$-sets with $2 \leq q < \infty$ can apparently be less lacunary than p-Sidon sets, $1 \leq p < 2$. We refer to [69] for the proof of the seminal result (1).

(2) For p-Sidon sets, we have nothing better than for p-Rider sets (see Proposition 4.8). When $p = 1$ i.e. E is Sidon, we can deduce from the result (1) the estimate of $\alpha_E(N)$ announced in (2). In fact, use the estimate $\lambda_p(E) = O(\sqrt{p})$ holding for a Sidon set E to get from (1) $\alpha_E(N) \leq cpN^{2/p}$ for every $p > 2$. Let us now choose $p = 2\log N$ with $N \geq 3$ so that $p > 2$, to get $\alpha_E(N) \leq c\log N$ as claimed in (2).

An example of a cartesian product. A classical example [75] is

$$E = \{(\lambda_j, \mu_k)\} =: A \times B,$$

the cartesian product of two Sidon sets in $\mathbb{Z}$. The set E is exactly $4/3$-Sidon in the product group $\mathbb{Z}^2$ [75]. In particular, $E = \{(2^j, 3^k)\}$ is $4/3$-Sidon in $\mathbb{Z}^2$. But the Furstenberg set is a Minkowski product. As we will see, it is a $\Lambda(p)$-set (Gundy–Varopoulos), but we do not know if it is a p-Sidon set for some p such that $4/3 \leq p < 2$.

We now present the notion of Paley set, which is related to that of $\Lambda(p)$-set. The Furstenberg set is such a set (Gundy–Varopoulos), as we will see.

3.3. Paley sets. The general framework consists of a set E of integers and a Banach space X of integrable functions on the circle. We say that *the set E is X-Paley if the Fourier transform of any function $f \in X$, once restricted to E, is square-summable* i.e.

$$\widehat{f}_{|E} \in \ell^2.$$

In this case, there exists a smallest constant $P(X, E) > 0$, *the Paley constant of the pair (X, E)*, such that the Paley inequality holds:

$$\|\widehat{f}_{|E}\|_2 \leq P(X, E)\|f\|_X, \quad \forall f \in X.$$

The terminology comes from a theorem of Paley, concerning specifically the case $X = H^1$, $E = \{2^n, n \geq 0\}$ (recall that $H^1 = \{f \in L^1(\mathbb{T}) ; \widehat{f}(n) = 0 \text{ when } n < 0\}$). It asserts that the powers of 2 constitute a H^1-Paley set.

Theorem 3.3 (Paley). *There exists $C > 0$ such that, for every $f \in H^1$,*

$$\left(\sum_{k \geq 0} |\widehat{f}(2^k)|^2\right)^{1/2} \leq C\|f\|_{H^1}.$$

Remarks. (1)   The example of Paley does not extend to $X = L^1(\mathbb{T})$. If K_N denotes the Fejér kernel of order N, we have

$$\sum_{n \geq 0} |\widehat{K_N}(2^n)|^2 \gtrsim \log N, \quad \text{but } \|K_N\|_1 = 1.$$

(2)   The Paley inequality holds for a space hardly smaller than $L^1(\mathbb{T})$, namely the Orlicz space $X = L^{\varphi_{1/2}}$ with our previous notations, and for any Sidon set E [77, Vol. II, Chapter XII, p. 132]. Hence, it holds for $X = L^q(\mathbb{T})$ $(q > 1)$ and Hadamard sets E.

(3)   The initial Paley inequality indeed tells that a Hadamard set is H^1-Paley; so is any finite union of Hadamard sets. *But this does not extend to arbitrary Sidon sets*; indeed, Sidon sets E exist for which $\sup_N |E \cap [N, 2N]| = \infty$, preventing them from being H^1-Paley [48, Vol. 2, Chapter VI, pp. 194–196].

We will be specially interested in the case $X = L^q(\mathbb{T}) =: L^q$ with $1 < q < 2$ and will speak of a *q-Paley set* in that context. An interesting duality exists between q-Paley sets and $\Lambda(p)$-sets, where $p > 2$ is the conjugate exponent of q.

Theorem 3.4. *Let E be a q-Paley set for some $1 < q < 2$. Then E is a $\Lambda(p)$-set and moreover, $\lambda_p(E) \leq P(L^q, E)$.*

Proof. Let $f \in L^2_E$ and $g \in L^q$ with $\|g\|_q = 1$. We have

$$\left| \int_{\mathbb{T}} f(x)g(-x)dm(x) \right| = \left| \sum_{n \in \mathbb{Z}} \hat{f}(n)\hat{g}(n) \right| = \left| \sum_{n \in E} \hat{f}(n)\hat{g}(n) \right|$$

$$\leq \|f\|_2 \left(\sum_{n \in E} |\hat{g}(n)|^2 \right)^{1/2}$$

$$\leq P(L^q, E)\|f\|_2\|g\|_q = P(L^q, E)\|f\|_2.$$

By taking the supremum over the g in the unit ball of L^q, we get

$$\|f\|_p \leq P(L^q, E)\|f\|_2, \tag{31}$$

hence the result.   ∎

Remark.   Note the gain when passing from "$\Lambda(p)$ to q-Paley": we enlarge the collection of f involved in the inequality since now the spectrum of the test function f is arbitrary.

Aiming to establish some properties of thin sets for the Furstenberg set S, we need a detour through probabilities, in particular through martingales.

4. From probability to harmonic analysis

In this paper, given a probability space $(\Omega, \mathcal{B}, \mathbb{P})$, we will need both *increasing* and *decreasing* sequences of sub-σ-fields $(\mathcal{F}_n)$ of $\mathcal{B}$. For an expert, there is no real difference (see [58, Chapter IV, p. 124] for a unified presentation), but we prefer to be more specific here.

4.1. Azuma's lemmas and subgaussianity of martingales. We consider martingales $(M_n)_{0 \leq n \leq N}$ adapted to a finite sequence of sub-σ-fields $(\mathcal{F}_n)_{0 \leq n \leq N}$ of $\mathcal{B}$. The conditional expectation given $\mathcal{F}_n$ is denoted $\mathbb{E}_n$, so that, in the increasing case, we have

$$\mathbb{E}_n(M_{n+1}) = M_n, \quad 0 \leq n < N.$$

We then write $M_N := \sum_{n=1}^N d_n$, where $d_n = M_n - M_{n-1}$, $\mathbb{E}_n(d_{n+1}) = 0$, and we assume $M_0 = 0$. In the case of a decreasing sequence $(\mathcal{F}_n)$, we will have

$$\mathbb{E}_{n+1}(M_n) = M_{n+1}, \quad 0 \leq n < N, \quad d_n = M_n - M_{n+1}, \quad \mathbb{E}_{n+1}(d_n) = 0.$$

We note the following seminal result.

Theorem 4.1 (Doob). *Let $(M_n)_{n \geq 0}$ be a complex martingale with respect to an increasing sequence $(\mathcal{F}_n)_{n \geq 0}$ of σ-fields. If (M_n) is bounded in L^1, it converges almost everywhere.*

Suppose now that (M_n) has bounded increments (d_n). In this context, Azuma proved the following "subgaussian" inequalities [46, Chapter I, pp. 31–32].

Lemma 4.2 (First inequality of Azuma). *Let $(M_n)_{0 \leq n \leq N}$ be a complex martingale with respect to $(\mathcal{F}_n)$, satisfying $|d_n| \leq C_n$ for $0 \leq n \leq N$. When $(\mathcal{F}_n)$ is increasing, for all $t > 0$, we have*

$$\mathbb{P}\big(|M_N| > t\big) \leq 4 \exp\left(-\frac{t^2}{4 \sum_{n=1}^N C_n^2}\right).$$

When $(\mathcal{F}_n)$ is decreasing, for all $t > 0$, we have

$$\mathbb{P}\big(|M_0| > t\big) \leq 4 \exp\left(-\frac{t^2}{4 \sum_{n=0}^{N-1} C_n^2}\right).$$

We will need a more precise version of Lemma 4.2, depending on two parameters.

Lemma 4.3 (Second inequality of Azuma). *Let $(M_n)_{0 \leq n \leq N}$ be a complex martingale with respect to an increasing sequence $(\mathcal{F}_n)$, which satisfies $|d_n| \leq a$ and*

$$\sum_{n=1}^N \mathbb{E}_{n-1}\big(|d_n|^2\big) \leq b^2,$$

where a and b are positive constants. Then, for all t > 0 we have

$$\mathbb{P}\big(|M_N| > t\big) \le 4\exp\Big[-\frac{t^2}{4b^2}\big(2 - e^{at/b^2}\big)\Big].$$

In particular, if $\exp(at/b^2) \le 3/2$, one has

$$\mathbb{P}\big(|M_N| > t\big) \le 4\exp\Big(-\frac{t^2}{8b^2}\Big).$$

Similarly, when $(\mathcal{F}_n)$ is decreasing, for all t > 0, we have

$$\mathbb{P}\big(|M_0| > t\big) \le 4\exp\Big[-\frac{t^2}{4b^2}\big(2 - e^{at/b^2}\big)\Big].$$

4.2. The Salem–Zygmund inequalities. The following theorem [71] plays a crucial role in Fourier analysis, and will be deduced here from the first inequality of Azuma. We provide short proofs.

Theorem 4.4 (Salem–Zygmund). *Let (r_n) be a Rademacher sequence, (a_n) a scalar sequence, and Q_N the random polynomial*

$$Q_N = \sum_{n=1}^{N} r_n a_n e_n$$

for all $N \ge 2$. Then

$$\mathbb{E}\big(\|Q_N\|_\infty\big) \le C\Big(\sum_{1}^{N} |a_n|^2\Big)^{1/2} \sqrt{\log N}.$$

Proof. It holds, using Bernstein's inequality for trigonometric polynomials of degree $\le N$ [64, Chapter V, p. 130]; see also [48, Vol. 1, Chapter VI, p. 241]:

$$\mathbb{E}\big(\|Q_N\|_\infty\big) \le C\mathbb{E}\Big(\sup_{x\in E}\Big|\sum_{n=1}^{N} r_n a_n e_n(x)\Big|\Big) =: C\mathbb{E}\big(\|Q_N\|_E\big),$$

where E denotes the set of $(4N)$-th roots of unity. The latter supremum is estimated as follows: let $A > 1$ and $\sigma_N^2 = \sum_{n=1}^{N} |a_n|^2$; then

$$\begin{aligned}
\mathbb{E}\big(\|Q_N\|_E\big) &= \int_0^\infty \mathbb{P}\big(\|Q_N\|_E > t\big)\,dt \le A + \int_A^\infty \mathbb{P}\big(\|Q_N\|_E > t\big)\,dt \\
&\le A + 4N\sup_{x\in E}\int_A^\infty \mathbb{P}\big(|Q_N(x)| > t\big)\,dt \\
&\le A + 4N\int_A^\infty \exp\Big(-\frac{t^2}{4\sigma_N^2}\Big)\,dt \le A + 8N\sigma_N\exp\Big(-\frac{A^2}{4\sigma_N^2}\Big)
\end{aligned}$$

(we used Lemma 4.2). Now, adjusting $A = \sqrt{4\log N}\,\sigma_N$ ends the proof. ∎

We will need a simple fact about the symmetrization [38]. Recall that the symmetrization $\widetilde{X}$ of a random variable X is defined (on $\Omega \times \Omega$) by

$$\widetilde{X}(\omega, \omega') = X(\omega) - X(\omega').$$

Proposition 4.5. *Let φ be an Orlicz function, and X a* centered *Banach space-valued random variable, with symmetrization $\widetilde{X}$. Then we have*

$$\mathbb{E}\big[\varphi(\|X\|)\big] \leq \mathbb{E}\big[\varphi(\|\widetilde{X}\|)\big] \leq \mathbb{E}\big[\varphi(2\|X\|)\big].$$

By symmetrization, we get the following more general Salem–Zygmund theorem.

Theorem 4.6. *Let $(X_n)_{1 \leq n \leq N}$ be a sequence of centered, square-integrable, complex-valued, independent random variables and P_N the associated random polynomial, $P_{N,\omega} = \sum_1^N X_n(\omega)e_n$; then*

$$\mathbb{E}\big(\|P_N\|_\infty\big) \leq C \sqrt{\sum_{1 \leq n \leq N} V(X_n)} \sqrt{\log N}.$$

Proof. Let $\widetilde{X}_n$ be a symmetrization of X_n. We write $\sum$ for $\sum_{1 \leq n \leq N}$. For every choice of signs (ε_n), Proposition 4.5 gives

$$\mathbb{E}\Big(\big\|\sum X_n e_n\big\|_\infty\Big) \leq \mathbb{E}\Big(\big\|\sum \widetilde{X}_n e_n\big\|_\infty\Big) = \mathbb{E}\Big(\big\|\sum \varepsilon_n \widetilde{X}_n e_n\big\|_\infty\Big).$$

We now take $\varepsilon_n = r_n(\omega')$, Rademacher variables independent from the $\widetilde{X}_n$, we integrate with respect to ω' and apply Fubini's theorem to get

$$\mathbb{E}\Big(\big\|\sum X_n e_n\big\|_\infty\Big) \leq \mathbb{E}_{\omega'}\mathbb{E}_\omega\Big(\big\|\sum r_n(\omega')\widetilde{X}_n(\omega)e_n\big\|_\infty\Big)$$
$$= \mathbb{E}_\omega\mathbb{E}_{\omega'}\Big(\big\|\sum r_n(\omega')\widetilde{X}_n(\omega)e_n\big\|_\infty\Big).$$

Salem–Zygmund's theorem now gives

$$\mathbb{E}\Big(\big\|\sum X_n e_n\big\|_\infty\Big) \leq C \sqrt{\log N}\,\mathbb{E}\Big(\sqrt{\sum |\widetilde{X}_n(\omega)|^2}\Big),$$

and by the Hölder inequality,

$$\mathbb{E}\Big(\sqrt{\sum |\widetilde{X}_n(\omega)|^2}\Big) \leq \Big(\mathbb{E}\sum |\widetilde{X}_n(\omega)|^2\Big)^{1/2} = \Big(\sum V(\widetilde{X}_n)\Big)^{1/2}.$$

Since $V(\widetilde{X}_n) = 2V(X_n)$, we can conclude, suitably changing the constant C. ∎

The following specialization will be useful in Section 5.

Theorem 4.7. *Let* $(\xi_n)_{1 \leq n \leq N}$ *be* $\{0, 1\}$*-valued independent variables with* $\mathbb{P}(\xi_n = 1) = \delta_n$ *and* $(\lambda_n)_{1 \leq n \leq N}$ *scalars of modulus* ≤ 1. *Set* $m_N = \sum_{n=1}^{N} \delta_n$. *Let* P_N *be the random trigonometric polynomial* $P_{N,\omega} = \sum_{1}^{N} (\xi_n(\omega) - \delta_n) \lambda_n \, e_n$. *Then*

$$\mathbb{E} \|P_N\|_\infty \leq C \sqrt{m_N \log N}.$$

Moreover, for some constant $C > 0$,

$$\mathbb{P}\big(\|P_N\|_\infty > t\big) \leq CN \exp\Big(-\frac{t^2}{8m_N}\Big) \quad \text{for all } t > 0 \text{ with } e^{t/m_N} \leq 3/2.$$

Proof. The first part is immediate from Theorem 4.6, since for the centered variables $X_n := \lambda_n(\xi_n - \delta_n)$, we have $V(X_n) \leq \delta_n - \delta_n^2 \leq \delta_n$. For the second part, first use Lemma 4.3 with $a = 1$ and $b^2 = m_N \geq \sum_{n=1}^{N} |\lambda_n|^2 (\delta_n - \delta_n^2)$ to get

$$\mathbb{P}\big(|P_N(x)| > t\big) \leq 4 \exp\Big(-\frac{t^2}{8m_N}\Big),$$

where x is a fixed $4N$-th root of unity and t as in the statement. So that, by Bernstein's inequality again,

$$\mathbb{P}\big(\|P_N\|_\infty > t\big) \leq CN \exp\Big(-\frac{t^2}{8m_N}\Big). \qquad \blacksquare$$

4.3. ***p*-Rider sets.** The functional $P \mapsto \mathbb{E}(\|P_\omega\|_\infty) =: [\![P]\!]$ which appeared in Theorem 4.4 (with $P_\omega = \sum r_n(\omega)\widehat{P}(n)e_n$, where (r_n) denotes a Rademacher sequence) is a norm, the "Pisier norm" on the space $\mathscr{P}$ of all trigonometric polynomials and the completion of $\mathscr{P}$ with respect to this norm is the so-called Pisier space of almost surely continuous functions [48, Vol. 2, Chapter VI]. This norm is much more flexible than the sup-norm because the characters e_n form an unconditional sequence for this norm, and motivates the following definition, a variant of p-Sidonicity. The set of polynomials with spectrum in a fixed set E is denoted by $\mathscr{P}_E$.

Definition 4.1. A set E of positive integers is called a p-Rider set (with $1 \leq p < 2$) if there exists a constant C such that

$$\|\widehat{f}\|_p \leq C[\![f]\!] \quad \text{for all } f \in \mathscr{P}_E.$$

The best C is called the p-Rider constant of E and is denoted $s_p(E)$.

Let us list a few properties of p-Rider sets.

- A p-Sidon set is a p-Rider set, and $s_p(E) \leq S_p(E)$ (obvious).

- For $p = 1$, we recover the notion of Sidon set, thanks to a result of Rider [67]. The analogue for $1 < p < 2$ is an open question.

- If $p < 4/3$, a p-Rider set is a q-Sidon set for each $q > p/(2 − p)$ [47].

As quoted in Proposition 3.2, p-Rider sets, in particular p-Sidon sets, satisfy an important mesh condition.

Proposition 4.8. *Let E be a p-Rider set. Then, there exists a constant C such that, for every integer $N \geq 2$:*

$$|E \cap [1, N]| \leq C(\log N)^\alpha \quad with \ \alpha = \frac{p}{2 − p}.$$

Proof. Let us write $E \cap [1, N] = \{\lambda_1 < \cdots < \lambda_n\}$; set $f = \sum_{j=1}^{n} e_{\lambda_j} \in C_E$. Using the Salem–Zygmund estimate of Theorem 4.6, we obtain

$$\llbracket f \rrbracket \leq C \sqrt{n} \sqrt{\log N}.$$

From our p-Riderness assumption, this gives

$$n^{1/p} \leq C s_p(E) n^{1/2} \sqrt{\log N},$$

and the result ensues. ∎

Remark. Proposition 4.8 already shows that the Furstenberg sequence S is at most $4/3$-Rider (in particular, $4/3$-Sidon). Indeed, we saw that $|S \cap [1, N]| \approx \log^2 N$, so that, if S is p-Rider, one must have $2 \leq p/(2 − p)$, or else $p \geq 4/3$. We were not able to show that S is p-Sidon or p-Rider for $p = 4/3$, or even for some $4/3 < p < 2$. But we shall show that its randomized version T is p-Rider for each $p > 4/3$.

4.4. Square function of a martingale. Here we present a fundamental inequality on the square function of a Hilbert space-valued martingale due to Burkholder obtained first in 1966 [17, Theorem 9], and later optimized in [18]. We will use it to prove that the Furstenberg set is a $\Lambda(p)$-set and even a Paley set, results due to Gundy and Varopoulos. And we will be more precise on the dependence on p.

We consider a martingale $M := (M_n)_{\geq 0}$ with respect to a *decreasing* sequence of sub-σ-fields $(\mathcal{F}_n)_{n \geq 0}$ of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Without loss of generality, we assume that $\mathcal{F}_0 = \mathcal{F}$ and $\mathcal{F}_n \downarrow \mathcal{F}_\infty = \{\emptyset, \Omega\}$. Recall that

(32) $$\mathbb{E}_{n+1}(d_n) = 0 \quad with \ d_n := M_n − M_{n+1}.$$

We also assume that M is a Hilbert space-valued martingale. The square function of the martingale M is the variable defined by

$$S = S(M) := \left(\sum_{n=0}^{\infty} \|d_n\|^2 \right)^{1/2},$$

where $\|d_n(\omega)\|$ denotes the Hilbert norm.

At this point, we recall a simple property of L^p-spaces ([51, p. 46], [48, Vol. 1, p. 189].

- If $p \geq 2$, the Banach lattice L^p is 2-convex, namely for any $(u_n) \subset L^p$:

$$(33) \qquad \left\| \left(\sum_n |u_n|^2 \right)^{1/2} \right\|_p \leq \left(\sum_n \|u_n\|_p^2 \right)^{1/2}.$$

- If $p \leq 2$, the Banach lattice L^p is 2-concave, namely

$$(34) \qquad \left\| \left(\sum_n |u_n|^2 \right)^{1/2} \right\|_p \geq \left(\sum_n \|u_n\|_p^2 \right)^{1/2}.$$

4.4.1. Burkholder's inequality. Let H be a Hilbert space. A given H-valued integrable function $f \in L^1(H)$ produces a special martingale $M(f) := (\mathbb{E}_n(f))$. We have the decomposition

$$(35) \qquad f - \mathbb{E}f = \sum_{n=0}^{\infty} d_n(f).$$

We state Burkholder's inequality for such martingales. In the following, q denotes the conjugate exponent of p and $p^* = \max(p, q)$. The following inequality is well known (see [18, Theorem 3.1 p. 87]).

Proposition 4.9 (Burkholder inequality). *Let H be a separable Hilbert space and let $f \in L^p(H)$ with zero mean $(1 < p < \infty)$. For the martingale $M_n = \mathbb{E}_n(f)$ with its martingales differences $d_n := d_n(f) = M_n - M_{n+1}$, we have*

$$(36) \qquad B_p^{-1} \|f\|_p \leq \|S(f)\|_p \leq B_p \|f\|_p,$$

where $S(f)^2 = \sum_{n=0}^{\infty} \|d_n\|^2$ and $B_p = p^* - 1$.

4.4.2. The Furstenberg set S is $\Lambda(p)$. We will first prove that the Furstenberg set S is a $\Lambda(p)$-set, a special case of the forthcoming Gundy–Varopoulos theorem.

The proof will use a special kind of scalar martingales. Let us first recall a basic fact. Let $m \geq 2$ be an integer. Denote by $\mathcal{F}^{(m)}$ the σ-field of m^{-1}-periodic Borel sets in $\mathbb{T}$. The following basic relations are easy to check.

Lemma 4.10. *For any $f \in L^1(\mathbb{T})$, we have*

$$(37) \qquad \mathbb{E}(f \,|\, \mathcal{F}^{(m)})(x) = \frac{1}{m} \sum_{k=0}^{m-1} f\left(x + \frac{k}{m} \right) = \sum_{m|n} \hat{f}(n) e_n(x).$$

We can also write $\mathbb{E}(f \,|\, \mathcal{F}^{(m)})(x) = f * \omega_m(x)$ with $\omega_m = m^{-1} \sum_{k=0}^{m-1} \delta_{k/m}$.

Observe that $\mathcal{F}^{(m)} = \sigma_m^{-1}(\mathcal{B}(\mathbb{T}))$, where σ_m is the map defined $\sigma_m(x) = mx \mod 1$. The first equality in (37) can be stated as

$$\mathbb{E}\big(f \,|\, \sigma_m^{-1}(\mathcal{B}(\mathbb{T}))\big)(x) = R \circ \sigma_m(x),$$

where R is the transfer operator

$$Rf(x) = \frac{1}{m} \sum_{k=0}^{m-1} f\left(\frac{x+k}{m}\right)$$

associated to the measure-preserving dynamical system $(\mathbb{T}, \mathcal{B}(\mathbb{T}), \lambda, \sigma_m)$, where λ is the Haar measure (cf. [27, Lemma 2.4]).

Now we fix an integer $a \geq 2$. For $n \geq 0$, let $\mathcal{F}_n = \mathcal{F}^{(a^n)}$ be the σ-field of a^{-n}-periodic Borel sets in $\mathbb{T}$. From the dynamical point of view, the iteration of σ_a is involved. Clearly $\mathcal{F}_0$ is the Borel field $\mathcal{B}(\mathbb{T})$ and $\mathcal{F}_n$ decreases to the trivial σ-field $\{\emptyset, \mathbb{T}\}$. When g is a trigonometric polynomial with zero mean, we can decompose g relatively to the σ-fields $\{\mathcal{F}_n\}$ according to (35). For simplicity, the corresponding martingale difference is still denoted by $d_n g$ and the corresponding square function by Sg.

Lemma 4.11. *Denote by L_a the set of integers not divisible by a. Then*

$$(38) \qquad d_n g(x) = \sum_{l \in L_a} \hat{g}(a^n l) e(a^n l x).$$

Hence,

$$(39) \qquad S^2(g)(x) = \sum_n \Big| \sum_{l \in L_a} \hat{g}(a^n l) e(a^n l x) \Big|^2.$$

Proof. By Lemma 4.10,

$$\mathbb{E}_n(g)(x) = g * \omega_{a^n}(x)$$

with $\hat{\omega}_n = \mathbf{1}_{a^n \mathbb{Z}}$. Since

$$d_n g = \mathbb{E}_n(g) - \mathbb{E}_{n+1}(g),$$

(38) follows from $\widehat{d_n g} = \hat{g} \cdot (\mathbf{1}_{a^n \mathbb{Z}} - \mathbf{1}_{a^{n+1} \mathbb{Z}}) = \hat{g} \cdot \mathbf{1}_{a^n L_a}$.   ∎

We denote by $S(a_1, \ldots, a_s)$ the semi-group generated by s integers $a_i \geq 2$. We say that $a_1, \ldots, a_s$ are *multiplicatively independent*, if the decomposition of any $d \in S(a_1, \ldots, a_s)$ as $d = a_1^{n_1} \cdots a_s^{n_s}$ is unique (e.g. $a_1 = 12$, $a_2 = 18$). Notice that $S(a_1, \ldots, a_s) \subset S(p_1, \ldots, p_t)$, where the p_j's are the primes involved in the decomposition of $a_1, \ldots, a_s$. It is possible that $t \neq s$ (e.g. for $a_1 = 24$, $a_2 = 30$, we have $s = 2, t = 3$).

Theorem 4.12 (Gundy–Varopoulos). *Suppose that $a_i \geq 2$ $(1 \leq i \leq s)$ are s coprime integers. The set $S(a_1, \ldots, a_s)$ is $\Lambda(p)$ for all $p > 2$, i.e.*

$$\|f\|_p \leq C_s(p)\|f\|_2$$

if $f \in \mathcal{P}_{S(a_1,\ldots,a_s)}$. Moreover, $C_s(p) \leq C p^{s-1/2}$ with C a constant.

Proof. We prove the result by induction on $s \geq 1$. The set $S(a_1) = \{a_1^n : n \geq 0\}$ is $\Lambda(p)$ because it is a Sidon set, and then we have $\lambda_p(S(a_1)) = O(\sqrt{p})$. Suppose that $S(a_1, \ldots, a_{s-1})$ is $\Lambda(p)$, namely for $f \in \mathcal{P}_{S(a_1,\ldots,a_{s-1})}$, we have

$$\text{(40)} \qquad \|f\|_p \leq C_{s-1}(p)\|f\|_2$$

for some constant $C_{s-1}(p)$. Assume $f \in \mathcal{P}_{S(a_1,\ldots,a_s)}$. We distinguish $a_s = a$ and write

$$\text{(41)} \quad f(x) := \sum_n \sum_{m_1,\ldots,m_{s-1}} a_{m_1,\ldots,m_{s-1},n} e(a_1^{m_1} \ldots a_{s-1}^{m_{s-1}} a^n x) = \sum_n f_n(a^n x),$$

with

$$f_n(y) = \sum_{m_1,\ldots,m_{s-1}} a_{m_1,\ldots,m_{s-1},n} e(a_1^{m_1} \ldots a_{s-1}^{m_{s-1}} y) \in \mathcal{P}_{S(a_1,\ldots,a_{s-1})}.$$

Since $a_1, \ldots, a_s$ are coprime, no number of $S(a_1, \ldots, a_{s-1})$ is divisible by a. Thus, applying Lemma 4.11, we get $d_n(f)(x) = f_n(a^n x)$. The associated square function is then equal to

$$Sf(x) := \left(\sum_n |d_n(x)|^2\right)^{1/2} = \left(\sum_n |f_n(a^n x)|^2\right)^{1/2}.$$

By the scalar Burkholder inequality (36),

$$\|f\|_p \leq B_p \|Sf\|_p.$$

It remains to estimate $\|Sf\|_p$. An important point is that the functions d_n have disjoint Fourier spectra (because $\mathrm{sp}(d_n) \subset a^n S(a_1, \ldots, a_{s-1})$), and hence are orthogonal in L^2. Firstly, the induction hypothesis (40) implies

$$\text{(42)} \qquad \|d_n\|_p = \|f_n\|_p \leq C_{s-1}(p)\|f_n\|_2 = C_{s-1}(p)\|d_n\|_2.$$

Then by (33) and since $p > 2$, we have

$$\begin{aligned}
\|Sf\|_p = \left\|\left(\sum |d_n|^2\right)^{1/2}\right\|_p &\leq \left(\sum \|d_n\|_p^2\right)^{1/2} \\
&\leq C_{s-1}(p)\left(\sum \|d_n\|_2^2\right)^{1/2} = C_{s-1}(p)\left(\left\|\sum d_n\right\|_2^2\right)^{1/2} \\
&= C_{s-1}(p)\|f\|_2.
\end{aligned}$$

We have used inequality (42), the orthogonality of d_n's and the decomposition $f = \sum d_n$. Finally,

$$\|f\|_p \leq B_p \|Sf\|_p \leq B_p C_{s-1}(p) \|f\|_2.$$

Thus, we have the claimed result with

$$C_s(p) = B_p C_{s-1}(p) = B_p^{s-1} C_1(p) = O(p^{s-1} \sqrt{p})$$

by Burkholder's inequality. This ends the proof of Theorem 4.12. ∎

4.5. The Gundy–Varopoulos theorem: S is Paley. Gundy and Varopoulos [36] proved the following theorem of *Paley-type* about the semi-group generated by s prime numbers $D_s := S(p_1, \ldots, p_s)$. This theorem goes beyond the Hadamard case considered by Paley (see Section 3.3) and beyond the previously proved $\Lambda(p)$-character of D_s (see Theorem 3.4). The proof will be essentially the same as for Theorem 4.12, except that we will need *two* martingales, one of them being vectorial, whereas only one scalar martingale was needed for proving the $\Lambda(p)$-character of D_s.

Theorem 4.13 (Gundy–Varopoulos). *The set D_s is a q-Paley set for all $1 < q < 2$, or else, for every function $f \in L^q$, we have*

$$(43) \qquad \left(\sum_{n \in D_s} |\hat{f}(n)|^2 \right)^{1/2} = \left(\sum_{n_j \in \mathbb{N}_0} |\hat{f}(p_1^{n_1} \cdots p_s^{n_s})|^2 \right)^{1/2} \leq B_s(q) \|f\|_q.$$

Moreover, $B_s(q) \leq (p-1)^s$, where $p = q/(q-1)$.

Proof. We will first assume that $s = 2$ to ease notations. So, we consider two distinct primes $a \geq 2$ and $b \geq 2$. We want to show that the multiplicative semi-group $S(a, b)$ generated by a and b is q-Paley. We will denote by $L := L_{a,b} \subset S(a, b)^c$ the set of integers ≥ 1 which are divisible neither by a nor by b. The proof is based on the following observations.

- We can assume that f is a polynomial and $\hat{f}(0) = 0$. Our aim is to dominate the quantity

$$(44) \qquad \sum_{m,n} |\hat{f}(a^m b^n)|^2 =: \sigma^2$$

by $\|f\|_q^2$ up to a multiplicative constant. Our first observation is that σ is bounded by $\|S\|_q$, where $S = (\sum_{m,n} |g_{m,n}|^2)^{1/2}$ with

$$g_{m,n}(x) = \sum_{l \in L} \hat{f}(a^m b^n l) e(a^m b^n l x) \quad (n \geq 0, m \geq 0).$$

Indeed, to let L^q-norms appear, we observe that

$$|\hat{f}(a^m b^n)|^2 = |\hat{g}_{m,n}(1)|^2 \leq \|g_{m,n}\|_q^2.$$

Summing over m, n and using (34) for the exponent $q < 2$, we get

$$(45) \qquad \sigma^2 \leq \sum_{m,n} \|g_{m,n}\|_q^2 \leq \|\mathcal{S}\|_q^2.$$

We will interpret $\mathcal{S}$ as the square function of a suitably chosen martingale, which will be handled thanks to Burkholder's inequalities.

• Let $\mathcal{F}^{(m)}$ (resp. $\mathcal{G}^{(n)}$) be the σ-field of a^{-m}-periodic (resp. b^{-n}- periodic) Borel sets, as before. We write $\mathbb{E}_m = \mathbb{E}(\cdot|\mathcal{F}^{(m)})$ and $\mathbb{E}'_n = \mathbb{E}(\cdot|\mathcal{G}^{(n)})$ the corresponding conditional expectations (which commute):

$$\mathbb{E}_m \mathbb{E}'_n = \mathbb{E}'_n \mathbb{E}_m.$$

Indeed, for $f \in L^1$, one has

$$\mathbb{E}_m \mathbb{E}'_n f(x) = \mathbb{E}'_n \mathbb{E}_m f(x) = f * \omega_{b^{-m}} * \sigma_{a^{-n}}(x) = f * \sigma_{a^{-n}} * \omega_{b^{-m}}(x),$$

by Lemma 4.10. We accordingly write $d_m = \mathbb{E}_m - \mathbb{E}_{m+1}$ and $d'_n = \mathbb{E}'_n - \mathbb{E}'_{n+1}$. Let $f \in L^1$ and consider its Fourier expansion

$$f(x) = \sum_{\mu,\nu,l \in L} a_{\mu,\nu,l} \, e(a^\mu b^\nu l x).$$

As Lemma 4.11 shows, passing from f to $d'_n(f)$ consists in freezing $\nu = n$ in this expansion, and then passing to $d_m d'_n(f)$ in freezing $\mu = m$ as well. So, the above commutativity leads to

$$(46) \qquad g_{m,n} = d_m(d'_n) = d'_n(d_m), \quad \mathcal{S}^2 = \sum_{m,n} |d_m(d'_n)|^2 = \sum_{m,n} |d'_n(d_m)|^2.$$

(Here and in what follows we omit the letter f).

• If F is H-valued, we will write $\|F\|_q$ for

$$\|F\|_{L^q(H)} = \left(\int \|F\|_H^q \, d\mathbb{P} \right)^{1/q}.$$

Let $H = \ell^2$. Consider the vector-valued function $F \in L^q(\ell^2)$ defined by

$$F = (d_0, d_1, \ldots, d_m, \ldots) : \Omega \to \ell^2.$$

Then consider the vector-valued martingale $M_n = \mathbb{E}'_n(F)$ with its martingale difference $d'M_n = \mathbb{E}'_n(F) - \mathbb{E}'_{n+1}(F)$. Observe that

$$d'M_n = \big(d'_n(d_0), d'_n(d_1), \ldots, d'_n(d_m), \ldots\big): \Omega \to \ell^2,$$

and that, by (46), the square function associated to (M_n) is equal to

$$(47) \qquad \Sigma^2 := \sum_n \|d'M_n\|^2 = \sum_{m,n} |d'_n(d_m)|^2, \quad \text{or else } \Sigma = S.$$

- On the other hand, the square function of the scalar martingale $(\mathbb{E}_m(f))$, which is defined by $Sf = (\sum_m |d_m|^2)^{1/2}$, is nothing but

$$(48) \qquad Sf(\omega) = \|F(\omega)\|_{\ell^2}.$$

Now, we end the proof as follows, using twice the Burkholder inequalities via (45), (47) and (48):

$$\sigma \le \|S\|_q = \|\Sigma\|_q \le B_q \|F\|_q = B_q \|Sf\|_q \le B_q^2 \|f\|_q \le (p-1)^2 \|f\|_q.$$

- The proof extends inductively to $s \ge 3$. To treat the general case of s distinct primes $p_1, \ldots, p_s$, we need some additional notations. If $\alpha := (n_1, \ldots, n_s) \in \mathbb{N}_0^s$ and $n = p_1^{n_1} \cdots p_s^{n_s}$, we write $n = p^\alpha$. We denote $E_n^{(k)}$ the conditional expectation with respect to the σ-field of p_k^{-n}-periodic sets, and $d_n^{(k)}$ the corresponding increment, as well as

$$\delta_\alpha = d_{n_s}^{(n_s)} \circ d_{n_{s-1}}^{(n_{s-1})} \circ \cdots \circ d_1^{(n_1)}.$$

Now, fix a trigonometric polynomial f, write for short $\delta_\alpha = \delta_\alpha(f)$; we set

$$\sigma_s = \Big(\sum_{\alpha \in \mathbb{N}_0^s} |\hat{f}(p^\alpha)|^2 \Big)^{1/2} \quad \text{and} \quad \Sigma_s = \Big(\sum_{\alpha \in \mathbb{N}_0^s} |\delta_\alpha(f)|^2 \Big)^{1/2}$$

the multivariate square function. We prove that $\sigma_s \le (p-1)^s \|f\|_q$ in two steps.

Step 1. We have (a) $\sigma_s \le \|\Sigma_s\|_q$ with the same proof as for $s = 2$.

Step 2. We have (b) $\|\Sigma_s\|_q \le (p-1)\|\Sigma_{s-1}\|_q$.

We just need to observe that Σ_s is the square function associated to the martingale $(M_{n_s}(F))$, where $F: \Omega \to \ell^2(\mathbb{N}_0^s)$ is defined by $F(\omega) = (\delta_\beta(\omega))_\beta$; here $\beta = (n_1, \ldots, n_{s-1})$ runs over $\mathbb{N}_0^{s-1}$. Then, Burkholder's inequality gives (b). Clearly, Theorem 4.13 follows from (a) and (b). ∎

The conclusion of Theorem 4.13 holds for $S(a_1, \ldots, a_s)$ when $a'_j s$ are multiplicatively independent, because $S(a_1, \ldots, a_s) \subset S(p_1, \ldots, p_t)$ where the p_j's are the prime factors of $a_1, \ldots, a_s$. One has then to abandon a precise estimate for the constant $B_s(q)$ in terms of s and q.

5. Random Furstenberg set

We begin by a definition, already coined by Erdős and Rényi [23], and then system-atically studied by Bourgain [12], which became popular under the name "Selectors of Bourgain". We mention in this respect the papers [49, 50, 57], and [3, 29, 39, 40]. Let $(\delta_k)_{k \geq 1}$ be a sequence of numbers such that $0 < \delta_k < 1$ for each k, and $(\xi_k)_{k \geq 1}$ be a sequence of independent Bernoulli variables, defined on some probability space $(\Omega, \mathbb{P})$ with

$$\mathbb{E}(\xi_k) = \mathbb{P}(\xi_k = 1) = \delta_k, \quad k = 1, 2, \dots,$$

and

$$R = R(\omega) = \{k \geq 1 \;;\; \xi_k(\omega) = 1\}.$$

The set R is called the random set of integers associated with the sequence (δ_k). We will always assume in this section that

(49) $\qquad\qquad (\delta_k)$ is non-increasing and $\sum_{k=1}^{\infty} \delta_k = \infty.$

Then, by the Borel–Cantelli lemma, R is almost surely infinite, and we write $R = \{u_1 < u_2 < \cdots < u_n < \cdots\}$. We will also set

$$m_N = \sum_{k=1}^{N} \delta_k = \mathbb{E}(\xi_1 + \xi_2 + \cdots + \xi_N) = \mathbb{E}\big(|R \cap [1, N]|\big) =: \mathbb{E}\big(|R_N|\big).$$

We will mostly assume that the "Bourgain condition" holds:

(50) $\qquad\qquad\qquad\qquad m_N / \log N \to \infty$

(this is the case when $k\delta_k \to \infty$). The question will next be: how to choose δ_k in order that $R(\omega)$ appears as a reasonable random version of the Furstenberg set S? And then what harmonic analysis and distribution properties does it almost surely possess?

The set $S = (s_n)_{n \geq 1}$ satisfies the lacunarity property $|S \cap [1, N]| \approx (\log N)^2$, or again $s_N \approx a e^{b\sqrt{N}}$. This motivates the choice (to which we will stick)

(51) $\qquad\qquad\qquad\qquad \delta_k = \dfrac{\log k}{k}.$

In fact, one then has

(52) $\qquad m_N = \sum_{k=1}^{N} \dfrac{\log k}{k} = \int_{1}^{N} \dfrac{\log t}{t}\, dt + O(1) = \dfrac{1}{2}(\log N)^2 + O(1).$

Notice that $\delta_k = O(1/k)$ corresponds to a random Sidon set [40, p. 364]. We will then reserve the letter T (instead of R) for this *random version* (51) *of the Furstenberg set* S.

The set S is quite rigid from the arithmetical point of view. We were partially able to deal with this arithmetic, and proved that S is $\Lambda(p)$ for all $p > 2$, is not Hartman uniformly distributed, and not p-Rider for $p < 4/3$. We will see that the random model $T(\omega)$, our main interest here, almost surely satisfies this p-Riderness property for $4/3 < p < 2$, among others.

5.1. T is almost surely weakly lacunary. We begin by a weak lacunarity property of general random sets R of integers. Then, we switch to our main case of interest: the random Furstenberg set T.

Theorem 5.1. *Let $R = (u_n)$ be the random set of integers corresponding to the sequence (δ_k) as in* (49). *Then*

(1) *R satisfies almost surely:* $\limsup_{n\to\infty}(u_{n+1} - u_n) = \infty$.

(2) *Assume moreover that $\sum_k \delta_k^2 < \infty$. Then, R is almost surely weakly lacunary, i.e.* $\lim_{n\to\infty}(u_{n+1} - u_n) = \infty$.

Proof. (1)   Fix an integer $q \geq 1$. Consider the events

$$A_k = \{\xi_{kq} = 1, \ \xi_{kq+j} = 0, \ 1 \leq j \leq q - 1\},$$

which satisfy

$$\mathbb{P}(A_k) = \delta_{kq} \prod_{j=1}^{q-1}(1 - \delta_{kq+j}) \geq \delta_{kq} \prod_{j=1}^{q-1}(1 - \delta_j) =: c_q \delta_{kq}.$$

Hence,

$$\sum_{k\geq 1} \mathbb{P}(A_k) = \infty.$$

Since the A_k's are independent, Borel–Cantelli's lemma gives $\mathbb{P}(A) = 1$, where $A = \limsup_k A_k$. Take $\omega \in A$, let k large with $\xi_{kq}(\omega) = 1$ and n the random (large) integer such that $u_n = kq$. Then, $u_{n+1} - u_n \geq q$. This shows that $\limsup(u_{n+1} - u_n) \geq q$ a.s., so that $\limsup(u_{n+1} - u_n) = \infty$ a.s.

(2)   Let (q_k) be a sequence of positive integers with $q_k \to \infty$ and

$$\sum_{k\geq 1} q_k \delta_k^2 < \infty.$$

For a pair of positive integers $1 \leq k < l$, consider the events

$$F_{k,l} = \{\xi_k = 1,\ \xi_l = 1\}, \quad G_k = \bigcup_{k+1 \leq l \leq k+q_k} F_{k,l}, \quad B = \liminf_k G_k^c.$$

We have

$$\mathbb{P}(F_{k,l}) = \delta_k \delta_l \leq \delta_k^2, \quad \mathbb{P}(G_k) \leq q_k \delta_k^2.$$

This implies $\mathbb{P}(B) = 1$; in other terms, we have almost surely for k large: $\xi_l = 0$ whenever $\xi_k = 1$ and $k < l \leq k + q_k$. Taking $k = u_n$ gives $u_{n+1} - u_n \geq q_k$ and the result. ∎

We now specialize to the random Furstenberg set $T =: \{t_n\}$, corresponding to the choice $\delta_k = \log k / k$, to get more quantitative results. Notably, T shares some precise gap properties with its deterministic relative S.

Theorem 5.2. *Almost surely, the difference $t_{n+1} - t_n$ satisfies:*

(1) $\limsup_{n \to \infty} \dfrac{t_{n+1} - t_n}{(t_n / \log t_n)\ \log \log t_n} \leq 2$;

(2) $\liminf_{n \to \infty} \dfrac{t_{n+1} - t_n}{t_n / (\log t_n)^{3+\delta}} \geq 1$ *for all $\delta > 0$.*

In particular, the weakly lacunary set T satisfies: almost surely $t_{n+1}/t_n \to 1$.

Proof. (1) Let (p_k) be a sequence of positive integers such that $p_k = o(k)$. Consider the event $E_k = \{\xi_k = 1,\ \xi_{k+1} = \cdots = \xi_{k+p_k} = 0\}$. Given $\varepsilon > 0$ and k large (depending on ε), we see that

$$\mathbb{P}(E_k) = \frac{\log k}{k} \prod_{j=k+1}^{k+p_k} \left(1 - \frac{\log j}{j}\right) \leq \frac{\log k}{k} \exp\left(-\sum_{j=k+1}^{k+p_k} \frac{\log j}{j}\right)$$

$$\leq \frac{\log k}{k} \exp\left(-p_k \frac{\log(k+p_k)}{k+p_k}\right) \leq \frac{\log k}{k} \exp\left(-(1-\varepsilon)p_k \frac{\log k}{k}\right).$$

We now take (ignoring the integer part issues)

$$p_k = a \frac{k}{\log k} \log \log k \quad \text{with } a = \frac{2+\varepsilon}{1-\varepsilon}$$

to get (for k large)

(53) $$\mathbb{P}(E_k) \leq \frac{\log k}{k} \frac{1}{(\log k)^{a(1-\varepsilon)}} = \frac{1}{k(\log k)^{1+\varepsilon}},$$

so that $\mathbb{P}(\liminf E_k^c) =: \mathbb{P}(A) = 1$. Now, fix $\omega \in A$. For $k \geq k_0(\omega)$, one has $\omega \in E_k^c$. Take $k = t_n$ with n large enough to ensure $k \geq k_0(\omega)$. Then, $\xi_k = 1$, and hence $\xi_j = 1$

for some $k < j \leq k + p_k$, meaning that $t_{n+1} \leq k + p_k = t_n + p_k$. It easily ensues for $\omega \in A$ that

$$\limsup_{n \to \infty} \frac{t_{n+1} - t_n}{(t_n / \log t_n) \log \log t_n} \leq a = \frac{2 + \varepsilon}{1 - \varepsilon}.$$

Letting ε tend to zero, we get the first assertion.

(2) The second one is similar: set $q_k = k/(\log k)^{3+\delta}$ where $\delta > 0$, and for $1 \leq k < l$ positive integers, we have

$$F_{k,l} = \{\xi_k = 1, \ \xi_l = 1\}, \quad F_k = \bigcup_{k+1 \leq l \leq k+q_k} F_{k,l}.$$

We get

$$\mathbb{P}(F_k) \leq c^2 q_k \left(\frac{\log k}{k}\right)^2 = \frac{c^2}{k(\log k)^{1+\delta}}.$$

Once again, $\mathbb{P}(\liminf_k F_k^c) = 1$, and almost surely, for n large enough

$$t_{n+1} - t_n \geq \frac{t_n}{(\log t_n)^{3+\delta}}. \qquad \blacksquare$$

5.2. R is almost surely Hartman uniformly distributed. The following random result appeared in [12] (see also [57]), with different terminology.

Theorem 5.3 (Bourgain). *Let (ξ_k) be a sequence of independent random $0 - 1$-valued variables as before, where (δ_k) satisfies the Bourgain condition* (50). *Then, a.s. the sequence $R := \{k \ ; \ \xi_k = 1\}$ is Hartman uniformly distributed.*

Proof. We give a proof relying on Section 4 (Salem–Zygmund theorem). To that effect, we decompose the sums

$$\sum_{n \in R_N} e(nx) := \sum_{n \leq N} \xi_n e(nx) = \sum_{n \leq N} (\xi_n - \delta_n) e(nx) + \sum_{n \leq N} \delta_n e(nx)$$

and the deterministic sums $\sum_{n \leq N} \delta_n e(nx)$, when $x \in \mathbb{T}$, $x \neq 0$, is fixed, are uniformly bounded in N (since (δ_n) is non-increasing) by a constant $C(x) < \infty$. As for the first random sum P_N, we resort to Theorem 4.7.

Choose $t = 6\sqrt{m_N \log N}$. We have $t/m_N = o(1)$ since $m_N/\log N \to \infty$, and thus $\exp(t/m_N) \leq 3/2$ for N large enough. Moreover, $t^2/(8m_N) \geq 4\log N$. Theorem 4.7, for N large enough, now gives

$$(54) \qquad \mathbb{P}\big(\|P_N\|_\infty > 6\sqrt{m_N \log N}\big) \ll N \exp(-4\log N) = N^{-3}.$$

In particular,

$$\mathbb{P}\big(\big||R_N| - m_N\big| > 6\sqrt{m_N \log N}\big) = \mathbb{P}\big(|P_N(0)| > 6\sqrt{m_N \log N}\big) \ll N^{-3}.$$

Borel–Cantelli's lemma and the relation $\sqrt{m_N \log N} = o(m_N)$ now imply the existence of $\Omega_0 \subset \Omega$, $\mathbb{P}(\Omega_0) = 1$, and for every $\omega \in \Omega_0$, of an integer $N(\omega)$ such that both inequalities below hold:

$$\frac{1}{2} m_N \leq |R_N| \leq 2m_N \quad \text{and} \quad \|P_N\|_\infty \ll \sqrt{m_N \log N}, \quad N \geq N(\omega).$$

We deduce, for $\omega \in \Omega_0$, $N \geq N(\omega)$ and any fixed $x \in \mathbb{T}$, $x \neq 0$:

$$\frac{1}{|R_N|} \Big| \sum_{n \leq N} \delta_n e(nx) \Big| \ll C(x)/m_N,$$

$$\sup_\tau \left(\frac{1}{|R_N|} \Big| \sum_{n \leq N} (\xi_n - \delta_n) e(n\tau) \Big| \right) \ll \sqrt{\log N / m_N},$$

whence the theorem, since $m_N \to \infty$ and $\log N / m_N \to 0$. $\blacksquare$

Corollary 5.4. *The random Furstenberg set T is a.s. Hartman uniformly distributed.*

Remarks. (1) The corollary is failing for the Furstenberg set S itself although $\sigma_N \approx (\log N)^2$! Indeed, we saw in Theorem 2.7 that, for uncountably many $x \in \mathbb{T}$, one has

$$\frac{1}{|S_N|} \sum_{n \in S_N} e(nx) \not\to 0.$$

(2) A question naturally arises: Is Bourgain's condition (50) necessary and sufficient?

5.3. Random integers and the Khinchin property. In this subsection, we will only assume that condition (49) holds for the sequence $(\delta_k)_{k \geq 1}$, and will not need Bourgain's condition (50). We begin by a complement on Orlicz functions (defined in Section 3).

Lemma 5.5. *Let φ be an Orlicz function such that $\psi := \varphi \circ u$ is concave, where $u(x) = \sqrt{x}$. Let $\varepsilon = (\varepsilon_k)_{1 \leq k \leq n}$ be a Rademacher sequence on a probability space Ω. If $a_1, \ldots, a_n$ are scalars, then*

$$\mathbb{E}\left[\varphi\Big(\Big| \sum_{k=1}^n \varepsilon_k a_k \Big| \Big) \right] \leq \sum_{k=1}^n \varphi(|a_k|).$$

As a corollary, if $X_1, \ldots, X_n$ are independent, centered, L^φ-integrable functions, we have

$$\mathbb{E}\left(\varphi\Big(\Big| \sum_{k=1}^n X_k \Big| \Big) \right) \leq \sum_{k=1}^n \mathbb{E}(\varphi(2|X_k|)).$$

Proof. Let $X = |\sum_{k=1}^{n} \varepsilon_k a_k|$. Jensen's inequality and the subadditivity of ψ give

$$\mathbb{E}[\varphi(|X|)] = \mathbb{E}[\psi(|X|^2)] \leq \psi(\mathbb{E}(|X|^2)) = \psi\left(\sum_{k=1}^{n} |a_k|^2\right)$$

$$\leq \sum_{k=1}^{n} \psi(|a_k|^2) = \sum_{k=1}^{n} \varphi(|a_k|).$$

And the corollary follows by symmetrization, noting that for a centered variable $Y \in L^{\varphi}$, it holds by Proposition 4.5 that

$$\mathbb{E}(\varphi(|Y|)) \leq \mathbb{E}(\varphi(|\tilde{Y}|)) \leq \mathbb{E}(\varphi(|2Y|)).$$

∎

Examples. (a) With $\varphi(x) = x^p$, $1 \leq p \leq 2$, we recover the fact that the Banach space L^p is of type p; see e.g. [48, Vol. 1, Chapter V, p. 188].

(b) We can also take $\varphi(x) = \varphi_\alpha(x) = x \log^\alpha(1 + x) =: xv(x)$, $0 < \alpha \leq 1$. Let us prove the concavity of ψ on $\mathbb{R}^+$. Since $-u''/u'^2 = 1/u$, we must show that

$$\varphi''(x)/\varphi'(x) \leq 1/x,$$

or else that

$$x^2 v''(x) + x v'(x) - v(x) \leq 0.$$

This is clear since $v'' \leq 0$ and

$$\frac{v'}{v} = \frac{\alpha}{(1+x)\log(1+x)} \leq \frac{1}{(1+x)\log(1+x)} \leq \frac{1}{x}.$$

We now recall that the *Khinchin class* of an increasing sequence $(u_n)_{n \geq 1}$ of positive integers is the class of L^1-functions f (here with mean zero) satisfying

(55) $$\tau_k(x) := \sum_{j=1}^{k} f(u_j x) = o(k) \quad \text{a.e.}$$

Marstrand [55] proved that there are L^∞-functions which are not in the Khinchin class of the set $\mathbb{N}$ of positive integers. Here is an interesting related result. Koksma [44] proved that if the function $f = \sum_{k \neq 0} a_k e_k \in L^2$ with

$$\sum_{|k| \geq 3} |a_k|^2 (\log \log |k|)^3 < \infty,$$

that is, $\hat{f}$ slightly better than ℓ^2, then f belongs to the Khinchin class of $\mathbb{N}$. We will prove here the following.

Theorem 5.6. *Let f be in the Khinchin class of the positive integers, with moreover $f \in L^{\varphi_\beta}$ for some $\beta > 1$. Then, f also belongs almost surely to the Khinchin class of the random set $R = \{k \; ; \; \xi_k = 1\} =: \{u_p\}$.*

Proof. Let $E \subset \mathbb{T}$ with $m(E) = 1$ and $\tau_k(x) = o(k)$ for $x \in E$; see (55). We begin by observing that

$$\frac{f(u_1 x) + \cdots + f(u_p x)}{p} = \frac{1}{S_N} \sum_{k=1}^{N} \xi_k f(kx) \quad \text{with } N = u_p,$$

where $S_N = \sum_{k=1}^{N} \xi_k = |R_N|$. The number N is random, but tends to ∞ almost surely, which is enough here. As we saw, almost surely, $m_N/2 \le S_N \le 2m_N$ for N large; hence, we can instead consider the quotient

$$A_N := \frac{1}{m_N} \sum_{k=1}^{N} \xi_k f(kx)$$

$$= \frac{1}{m_N} \sum_{k=1}^{N} (\xi_k - \delta_k) f(kx) + \frac{1}{m_N} \sum_{k=1}^{N} \delta_k f(kx) =: B_N + C_N.$$

We will show that, given $x \in E$, $A_N \to 0$ almost surely. The Fubini theorem will then give the result. We first observe that the deterministic term C_N tends to 0, because for any sequence (a_n),

$$\sum_{k=1}^{n} a_k = o(n) \quad \text{implies} \quad \sum_{k=1}^{n} \delta_k a_k = o\left(\sum_{k=1}^{n} \delta_k\right),$$

which is a consequence of (49) and an Abel summation.

For the random term B_N, we use a martingale argument. Let us consider the product space $(\Omega \times \mathbb{T}, \mathcal{A} \times \mathcal{B}(\mathbb{T}), \mathbb{P} \times m)$, on which is defined the martingale

$$M_n(\omega, x) = \sum_{k=1}^{n} \frac{(\xi_k(\omega) - \delta_k) f(kx)}{m_k}$$

relative to the filtration $\{\mathcal{F}_n\}$ defined by $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n) \times \mathcal{B}(\mathbb{T})$. Indeed,

$$\mathbb{E}_{\omega,x}\big[(\xi_n - \delta_n) f(nx) \mid \mathcal{F}_{n-1}\big] = f(nx) \mathbb{E}_{\omega,x}\big[(\xi_n - \delta_n) \mid \mathcal{F}_{n-1}\big] = 0.$$

Now, pick $0 < \alpha < \min(\beta - 1, 1)$ and put $\varphi = \varphi_\alpha$. We claim that M_n is bounded in L^φ, a fortiori in L^1. Then Doob's convergence theorem implies the "almost sure almost everywhere" convergence of the series

$$\sum_{k=1}^{\infty} \frac{(\xi_k(\omega) - \delta_k) f(kx)}{m_k}.$$

Applying Kronecker's lemma, we will conclude that $B_N \to 0$.

To prove our claim, we first note the easy estimation

(56) $$\int_{\mathbb{T}} |f| \log^{\alpha}\left(1 + \frac{|f|}{A}\right) dm(x) \ll \frac{1}{(\log A)^{\beta-\alpha}} \quad \text{as } A \to \infty.$$

Let us fix $a \geq 1$ and let

$$X_k(\omega) = \frac{(\xi_k(\omega) - \delta_k) f(kx)}{m_k}.$$

Note that

$$\varphi(|X_k|/a) \leq \frac{|\xi_k - \delta_k|}{m_k a} |f(kx)| \left[\log^{\alpha}\left(1 + \frac{|f(kx)|}{A}\right)\right] \quad \text{with } A = \frac{m_k}{|\xi_k - \delta_k|}.$$

We integrate with respect to x, using (56) and the invariance of m under $x \mapsto kx$, to get

$$\int \varphi(|X_k|/a) \, dx \leq C \frac{|\xi_k - \delta_k|}{m_k a} \frac{1}{\log^{\beta-\alpha} A} = C \frac{|\xi_k - \delta_k|}{m_k a} \frac{1}{\log^{\beta-\alpha}(m_k/(|\xi_k - \delta_k|))}.$$

We now take the expectation (with respect to ω) and obtain

(57) $$\mathbb{E}_{\omega,x}\big(\varphi(|X_k|/a)\big) \ll \frac{\delta_k}{m_k a (\log m_k)^{\beta-\alpha}}.$$

Noting that the assumptions of Lemma 5.5 are satisfied for fixed x, we get by permuting the order of integrations and changing in (57) a into $a/2$, $a \geq 2$,

$$\iint \varphi\left(\frac{|M_n(\omega, x)|}{a}\right) d\omega \, dx \leq \frac{C'}{a} \sum_{k=1}^{\infty} \frac{\delta_k}{m_k (\log m_k)^{\beta-\alpha}} =: \frac{C''}{a} < \infty.$$

Indeed, for any sequence (a_n) of positive numbers, we have

$$\sum_{n=1}^{\infty} \frac{a_n}{(a_1 + \cdots + a_n) \log^{\rho}(a_1 + \cdots + a_n)} < \infty$$

when $\rho > 1$. By definition of the Orlicz norm, we get the claim: $\|M_n\|_{\varphi} \leq C''$. ∎

5.4. T is almost surely p-Rider.

In [49] and [50], the authors constructed inside (a variant of) T a random subset *of positive relative density*, which is 4/3-Rider and $\Lambda(q)$ for each $q > 2$. (We will come back to this in a forthcoming work.) Here, with the idea of randomly mimicking the Furstenberg set S, we have to deal with the whole of T. We can prove the following.

Theorem 5.7. *Almost surely, the set* $T := T(\omega)$ *is* p-Rider for $p > 4/3$ and not p-Rider for all $p < 4/3$.

As we already showed in Proposition 4.8, the mesh condition implies that T is not p-Rider for $p < 4/3$. The proof of the other claim is more elaborate, and relies on the following result, due to L. Rodríguez-Piazza, which extends Pisier's result for Sidon sets [68, Teorema 2.3, pp. 85–86]. We refer to Section 3.1 for the notions involved.

Theorem 5.8 ([68]). *Let* E *be a set of* positive *integers and* $1 \leq p < 2$. *The following are equivalent:*

(1) E *is* p-Rider.

(2) *For every finite subset* A *of* E, *there exists a subset* B *of* A *such that*

$$B \text{ is quasi-independent and } |B| \geq \delta |A|^{\varepsilon},$$

where $\varepsilon = \frac{2}{p} - 1 < 1$ *and where* δ *is a positive constant.*

We now need the following lemma ([49, Lemma II.1, p. 115]).

Lemma 5.9 ([49]). *Let* $s \geq 2$ *and* A *be positive integers. Set*

$$\Omega_s(A) = \{\omega : T(\omega) \cap [A, \infty[\text{ contains at least a relation of length } s\}.$$

Then

$$\mathbb{P}(\Omega_s(A)) \leq \frac{C^s}{s^s} \sum_{j > A} \delta_j^2 m_j^{s-2} \quad \text{with } C = 4e.$$

We denote $T_k = T \cap [1, k]$, $k = 1, 2, \ldots$. Then the following holds.

Lemma 5.10. *Let* $A_n = (n \log^3 n)^n$; *then*

(1) $\sum_{n \geq 1} \mathbb{P}(\Omega_n(A_n)) < \infty$;

(2) *Almost surely,* $|T_{A_n}| \ll (n \log n)^2$ *for all integers* n *large enough.*

Lemma 5.10 means that with high probability, T does not contain any "short" relation at infinity, and contains relatively few elements near the origin.

Proof of Theorem 5.7. Let us take Lemma 5.10 for granted. Denote

$$\Omega' := \Omega \setminus \limsup_n \Omega_n(A_n), \quad \Omega'' := \{\omega : |T_{A_n}| \ll (n \log n)^2 \text{ for } n \text{ large}\}.$$

It holds that $\mathbb{P}(\Omega') = \mathbb{P}(\Omega'') = 1$, $\mathbb{P}(\Omega' \cap \Omega'') = 1$, by the Borel–Cantelli lemma. Let us fix $\omega \in \Omega' \cap \Omega''$ and then $T := T(\omega)$. Let $n_0(\omega) =: n_0 \geq 1$ such that $\omega \notin \Omega_n(A_n)$ for $n \geq n_0$. Then, by definition,

(58) $T \cap [A_n, \infty[$ contains no relation of length $\leq n$ for $n \geq n_0$.

Moreover,

(59) $$|T_{A_n}| \ll (n \log n)^2 \text{ for } n \geq n_1.$$

Now, we fix some finite subset E of T and we put $n = |E|^\varepsilon$. We can assume that $n \geq \max(n_0, n_1)$. We then observe that

$$|E \cap [A_n, \infty[| \geq |E| - |E \cap [1, A_n]| \geq n^{1/\varepsilon} - C(n \log n)^2 \geq \delta n^{1/\varepsilon} \geq n$$

for large n, provided that $1/\varepsilon > 2$, or else $\varepsilon = \frac{2}{p} - 1 < 1/2$. And this condition is equivalent to $p > 4/3$, which was our assumption. But then, if we take $F \subset E \cap [A_n, \infty[$ with $|F| = n$, this set is quasi-independent by (58), and has the right cardinality in order to apply Theorem 5.8.

All this shows that, for $\omega \in \Omega' \cap \Omega''$, $T(\omega)$ is p-Rider for all $p > 4/3$. This ends the proof of Theorem 5.7 *up to that of Lemma* 5.10. ■

A sketchy proof of Lemma 5.10. The proof is sketchy since it is similar to that of Lemma II.3 in [49]. First, Tchebycheff's inequality and (52) give

$$\mathbb{P}\left(\left||T_{A_n}| - \mathbb{E}\left(|T_{A_n}|\right)\right| > m_{A_n}\right) \leq \frac{m_{A_n}}{m_{A_n}^2} = \frac{1}{m_{A_n}} \ll \frac{1}{(n \log n)^2},$$

so that

$$\sum_{n \geq 1} \mathbb{P}\left(\left||T_{A_n}| - \mathbb{E}\left(|T_{A_n}|\right)\right| > (\log A_n)^2\right) < \infty.$$

In other terms, one has (59), giving the second assertion. Next, according to Lemma 5.9, we have

$$\mathbb{P}\left(\Omega_n(A_n)\right) \ll \frac{C^n}{n^n} \sum_{j > A_n} \frac{\log^2 j}{j^2} (\log^2 j)^{n-2}$$

$$\ll \frac{C^n}{n^n} \int_{A_n}^\infty \frac{(\log^2 t)^n}{t^2} \, dt \ll \frac{C^n}{n^n} \frac{(\log^2 A_n)^n}{A_n}$$

by using an integration by parts, or else

$$\mathbb{P}\left(\Omega_n(A_n)\right) \ll \frac{C^n}{n^n} \frac{(n \log n)^{2n}}{(n \log^3 n)^n} \ll \frac{C^n}{(\log n)^n},$$

which gives the first assertion. ■

References

[1] C. Aistleitner, I. Berkes and R. F. Tichy, On permutations of Hardy–Littlewood–Pólya sequences. *Trans. Amer. Math. Soc.* **363** (2011), no. 12, 6219–6244. Zbl 1270.42009 MR 2833551

[2] C. Badea and S. Grivaux, Kazhdan constants, continuous probability measures with large Fourier coefficients and rigidity sequences. *Comment. Math. Helv.* **95** (2020), no. 1, 99–127. Zbl 1451.43001 MR 4082895

[3] C. Badea, S. Grivaux and E. Matheron, Rigidity sequences, Kazhdan sets, and group topologies on the integers. 2018, arXiv:1812.09014v1

[4] A. Baker, Approximations to the logarithms of certain rational numbers. *Acta Arith.* **10** (1964), 315–323. Zbl 0201.37603 MR 171749

[5] J. Barral and A.-H. Fan, Covering numbers of different points in Dvoretzky covering. *Bull. Sci. Math.* **129** (2005), no. 4, 275–317. Zbl 1068.28005 MR 2134123

[6] A. Bérczes, A. Dujella and L. Hajdu, Some Diophantine properties of the sequence of S-units. *J. Number Theory* **138** (2014), 48–68. Zbl 1366.11055 MR 3168921

[7] V. Bergelson and E. Lesigne, Van der Corput sets in $\mathbb{Z}^d$. *Colloq. Math.* **110** (2008), no. 1, 1–49. Zbl 1177.37018 MR 2353898

[8] I. Berkes and M. Weber, On series $\sum c_k f(kx)$ and Khinchin's conjecture. *Israel J. Math.* **201** (2014), no. 2, 593–609. Zbl 1319.40001 MR 3265297

[9] R. Blei, *Analysis in integer and fractional dimensions*. Cambridge Stud. Adv. Math. 71, Cambridge Univ. Press, Cambridge, 2001. Zbl 1006.46001 MR 1853423

[10] C. Bluhm, On a theorem of Kaufman: Cantor-type construction of linear fractal Salem sets. *Ark. Mat.* **36** (1998), no. 2, 307–316. Zbl 1026.28004 MR 1650442

[11] M. D. Boshernitzan, Elementary proof of Furstenberg's Diophantine result. *Proc. Amer. Math. Soc.* **122** (1994), no. 1, 67–70. Zbl 0815.11036 MR 1195714

[12] J. Bourgain, On the maximal ergodic theorem for certain subsets of the integers. *Israel J. Math.* **61** (1988), no. 1, 39–72. Zbl 0642.28010 MR 937581

[13] J. Bourgain, E. Lindenstrauss, P. Michel and A. Venkatesh, Some effective results for $\times a \times b$. *Ergodic Theory Dynam. Systems* **29** (2009), no. 6, 1705–1722. Zbl 1237.37009 MR 2563089

[14] Y. Bugeaud, Nombres de Liouville et nombres normaux. *C. R. Math. Acad. Sci. Paris* **335** (2002), no. 2, 117–120. Zbl 1006.11043 MR 1920005

[15] — *Distribution modulo one and Diophantine approximation*. Cambridge Tracts in Math. 193, Cambridge Univ. Press, Cambridge, 2012. Zbl 1260.11001 MR 2953186

[16] S. Bullett and P. Sentenac, Ordered orbits of the shift, square roots, and the devil's staircase. *Math. Proc. Cambridge Philos. Soc.* **115** (1994), no. 3, 451–481. Zbl 0823.58012 MR 1269932

[17] D. Burkholder, Martingale transforms. *Ann. Math. Statist.* **37** (1966), 1494–1504. Zbl 0306.60030 MR 208647

[18] — Sharp inequalities for martingales and stochastic integrals. Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987). *Astérisque* 157–158 (1988), 75–94. Zbl 0656.60055 MR 976214

[19] N. Carey and D. Clampitt, Regions: A theory of tonal spaces in early medieval treatises. *J. Music Theory* **40** (1996), no. 1, 113–147.

[20] S. W. Drury, Sur les ensembles de Sidon. *C. R. Acad. Sci. Paris Sér. A-B* **271** (1970), A162–A163. Zbl 0194.42804 MR 271647

[21] R. E. Edwards and K. A. Ross, *p*-Sidon sets. *J. Functional Analysis* **15** (1974), 404–427. Zbl 0273.43007 MR 0358228

[22] M. Einsiedler and T. Ward, *Ergodic theory with a view towards number theory*. Grad. Texts in Math. 259, Springer, London, 2011. Zbl 1206.37001 MR 2723325

[23] P. Erdős and A. Rényi, Additive properties of random sequences of positive integers. *Acta Arith.* **6** (1960), 83–110. Zbl 0091.04401 MR 120213

[24] P. Erdős and S. J. Taylor, On the set of points of convergence of a lacunary trigonometric series and the equidistribution properties of related sequences. *Proc. London Math. Soc. (3)* **7** (1957), 598–615. Zbl 0111.26801 MR 92032

[25] K. Falconer, *Fractal geometry*. John Wiley and Sons, Chichester, 1990. Zbl 0689.28003 MR 1102677

[26] A. Fan, Lacunarité à la Hadamard et équirépartition. *Colloq. Math.* **66** (1993), no. 1, 151–163. Zbl 0826.11037 MR 1242653

[27] — Almost everywhere convergence of ergodic series. *Ergodic Theory Dynam. Systems* **37** (2017), no. 2, 490–511. Zbl 1378.37011 MR 3614035

[28] A. Fan, J. Schmeling and W. Shen, L^∞-estimation of generalized Thue–Morse trigonometric polynomials and ergodic maximization. *Discrete Contin. Dyn. Syst.* **41** (2021), no. 1, 297–327. Zbl 1469.11271 MR 4182323

[29] A. Fan and D. Schneider, Recurrence properties of sequences of integers. *Sci. China Math.* **53** (2010), no. 3, 641–656. Zbl 1193.28013 MR 2608321

[30] D. Feng, H. Rao and J. Wu, The net measure properties of symmetric Cantor sets and their applications. *Progr. Natur. Sci. (English Ed.)* **7** (1997), no. 2, 172–178. Zbl 0964.28004 MR 1459922

[31] D. Feng, Z. Wen and J. Wu, Some dimensional results for homogeneous Moran sets. *Sci. China Ser. A* **40** (1997), no. 5, 475–482. Zbl 0881.28003 MR 1461002

[32] É. Fouvry and G. Tenenbaum, Entiers sans grand facteur premier en progressions arithmetiques. *Proc. London Math. Soc. (3)* **63** (1991), no. 3, 449–494. Zbl 0745.11042 MR 1127146

[33] H. Furstenberg, Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1** (1967), 1–49. Zbl 0146.28502 MR 213508

[34] A. Gelfond, Sur les nombres qui ont des propriétés additives et multiplicatives données. *Acta Arith.* **13** (1967/68), 259–265. Zbl 0155.09003 MR 220693

[35] A. Gelfond and Y. Linnik, *Méthodes élémentaires dans la théorie analytique des nombres*. Monographies Internationales de Mathématiques Modernes 6, Gauthier-Villars, Paris, 1965. Zbl 0125.29604 MR 0188136

[36] R. F. Gundy and N. T. Varopoulos, A martingale that occurs in harmonic analysis. *Ark. Mat.* **14** (1976), no. 2, 179–187. Zbl 0371.60058 MR 448539

[37] G. H. Hardy, *Ramanujan. Twelve lectures on subjects suggested by his life and work*. Cambridge Univ. Press, Cambridge; The Macmillan Company, New York, 1940. Zbl 67.0002.09 MR 0004860

[38] J. Hoffmann-Jørgensen, Sums of independent Banach space valued random variables. *Studia Math.* **52** (1974), 159–186. Zbl 0265.60005 MR 356155

[39] J.-P. Kahane and Y. Katznelson, Distribution uniforme de certaines suites d'entiers aléatoires dans le groupe de Bohr. *J. Anal. Math.* **105** (2008), 379–382. Zbl 1153.43005 MR 2438431

[40] — Entiers aléatoires et analyse harmonique. *J. Anal. Math.* **105** (2008), 363–378. Zbl 1153.43004 MR 2438430

[41] J.-P. Kahane and R. Salem, Distribution modulo 1 and sets of uniqueness. *Bull. Amer. Math. Soc.* **70** (1964), 259–261. Zbl 0142.29604 MR 158216

[42] — *Ensembles parfaits et séries trigonométriques*. Second edn., Hermann, Paris, 1994. Zbl 0856.42001 MR 1303593

[43] A. Khinchin, Ein Satz über Kettenbrüche, mit arithmetischen Anwendungen. *Math. Z.* **18** (1923), no. 1, 289–306. Zbl 49.0159.03 MR 1544632

[44] J. F. Koksma, A Diophantine property of some summable functions. *J. Indian Math. Soc. (N.S.)* **15** (1951), 87–96 (1952). Zbl 0046.04802 MR 47094

[45] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*. Second edn., Dover Publications, New York, 2002 Zbl 0281.10001 MR 0419394

[46] M. Ledoux and M. Talagrand, *Probability in Banach spaces*. Ergeb. Math. Grenzgeb. (3) 23, Springer, Berlin, 1991. Zbl 0748.60004 MR 1102015

[47] P. Lefèvre and L. Rodríguez-Piazza, *p*-Rider sets are *q*-Sidon sets. *Proc. Amer. Math. Soc.* **131** (2003), no. 6, 1829–1838. Zbl 1010.43006 MR 1955271

[48] D. Li and H. Queffélec, *Introduction to Banach spaces: Analysis and probability. Vol. 1/2*. Cambridge Stud. Adv. Math. 166/167, Cambridge Univ. Press, Cambridge, 2018. Zbl 1404.46002 MR 3729312

[49] D. Li, H. Queffélec and L. Rodríguez-Piazza, Some new thin sets of integers in harmonic analysis. *J. Anal. Math.* **86** (2002), 105–138. Zbl 1018.43004 MR 1894479

[50] — On some random thin sets of integers. *Proc. Amer. Math. Soc.* **136** (2008), no. 1, 141–150. Zbl 1131.43007 MR 2350399

[51] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces. II*. Ergeb. Math. Grenzgeb. (3) 97, Springer, Berlin-New York, 1979. Zbl 0403.46022 MR 540367

[52] F. Lust-Piquard, Bohr local properties of $C_\Lambda(T)$. *Colloq. Math.* **58** (1989), no. 1, 29–38. Zbl 0694.43005 MR 1028158

[53] R. Lyons, The measure of nonnormal sets. *Invent. Math.* **83** (1986), no. 3, 605–616. Zbl 0585.10036 MR 827371

[54] — Seventy years of Rajchman measures. Proceedings of the Conference in Honor of Jean-Pierre Kahane (Orsay, 1993). *J. Fourier Anal. Appl.* (1995), Special Issue, 363–377. Zbl 0886.43001 MR 1364897

[55] J. M. Marstrand, On Khinchin's conjecture about strong uniform distribution. *Proc. London Math. Soc. (3)* **21** (1970), 540–556. Zbl 0208.31402 MR 291091

[56] R. Nair, On strong uniform distribution. *Acta Arith.* **56** (1990), no. 3, 183–193. Zbl 0716.11036 MR 1082999

[57] S. Neuwirth, Two random constructions inside lacunary sets. *Ann. Inst. Fourier (Grenoble)* **49** (1999), no. 6, 1853–1867. Zbl 0955.42009 MR 1738068

[58] J. Neveu, *Bases mathématiques du calcul des probabilités*. Masson et Cie, Paris, 1970. Zbl 0203.49901 MR 0272004

[59] — *Martingales à temps discret*. Masson et Cie, Paris, 1972. MR 0402914

[60] K. Petersen, *Ergodic theory*. Cambridge Stud. Adv. Math. 2, Cambridge Univ. Press, Cambridge, 1983. Zbl 0507.28010 MR 833286

[61] — An adic dynamical system related to the Delannoy numbers. *Ergodic Theory Dynam. Systems* **32** (2012), no. 2, 809–823. Zbl 1257.37006 MR 2901372

[62] W. Philipp, Empirical distribution functions and strong approximation theorems for dependent random variables. A problem of Baker in probabilistic number theory. *Trans. Amer. Math. Soc.* **345** (1994), no. 2, 705–727. Zbl 0812.11045 MR 1249469

[63] G. Pisier, De nouvelles caractérisations des ensembles de Sidon. In *Mathematical analysis and applications, Part B*, pp. 685–726, Adv. in Math. Suppl. Stud. 7, Academic Press, New York-London, 1981. Zbl 0468.43008 MR 634264

[64] H. Queffélec and M. Queffélec, *Diophantine approximation and Dirichlet series*. Texts Read. Math. 80, Hindustan Book Agency, New Delhi; Springer, Singapore, 2020. Zbl 1478.11002 MR 4241378

[65] M. Queffélec, *Substitution dynamical systems – spectral analysis*. Second edn., Lecture Notes in Math. 1294, Springer, Berlin, 2010. Zbl 1225.11001 MR 2590264

[66] G. Rhin, Approximants de Padé et mesures effectives d'irrationalité. In *Séminaire de Théorie des Nombres (Paris, 1985–86)*, pp. 155–164, Progr. Math. 71, Birkhäuser, Boston, MA, 1987. Zbl 0632.10034 MR 1017910

[67] D. Rider, Randomly continuous functions and Sidon sets. *Duke Math. J.* **42** (1975), no. 4, 759–764. Zbl 0345.43008 MR 387965

[68] L. Rodriguez Piazza, *Rango y propiedades de medidas vectoriales: Conjuntos p-Sidon p.s.* PhD thesis, Universidad de Sevilla, 1991

[69] W. Rudin, Trigonometric series with gaps. *J. Math. Mech.* **9** (1960), 203–227. Zbl 0091.05802 MR 0116177

[70] — *Fourier analysis on groups*. Reprint of the 1962 original. Wiley Classics Library, A Wiley-Interscience Publication, John Wiley and Sons, New York, 1990. Zbl 0698.43001 MR 1038803

[71] R. Salem and A. Zygmund, Some properties of trigonometric series whose terms have random signs. *Acta Math.* **91** (1954), 245–301. Zbl 0056.29001 MR 65679

[72] A. Thue, Über Annäherungswerte algebraischer Zahlen. *J. Reine Angew. Math.* **135** (1909), 284–305. Zbl 40.0265.01 MR 1580770

[73] R. Tijdeman, On integers with many small prime factors. *Compositio Math.* **26** (1973), 319–330. Zbl 0267.10056 MR 325549

[74] — On the maximal distance between integers composed of small primes. *Compositio Math.* **28** (1974), 159–162. Zbl 0283.10024 MR 345917

[75] G. S. Woodward, *p*-Sidon sets and a uniform property. *Indiana Univ. Math. J.* **25** (1976), no. 10, 995–1003. Zbl 0352.43004 MR 422997

[76] Q. Wu and L. Wang, On the irrationality measure of log 3. *J. Number Theory* **142** (2014), 264–273. Zbl 1303.11081 MR 3208402

[77] A. Zygmund, *Trigonometric series: Vols. I, II*. Sixth edn., Cambridge Univ. Press, London-New York, 1993. Zbl 0005.06303 MR 0236587

Aihua Fan, School of Mathematics and Statistics, Central China Normal University, 430079 Wuhan, China; and LAMFA, UMR 7352 CNRS, University of Picardie, 33 rue Saint Leu, 80039 Amiens, France; *e-mail:* ai-hua.fan@u-picardie.fr

Hervé Queffélec, CNRS, Laboratoire Paul Painlevé, UMR 8524, and Labex CEMPI (ANR-LABX-0007-01), Université de Lille, Cité Scientifique, Bât. M2, 59655 Villeneuve d'Ascq Cedex, France; *e-mail:* herve.queffelec@univ-lille.fr

Martine Queffélec, CNRS, Laboratoire Paul Painlevé, UMR 8524, and Labex CEMPI (ANR-LABX-0007-01), Université de Lille, Cité Scientifique, Bât. M2, 59655 Villeneuve d'Ascq Cedex, France; *e-mail:* martine.queffelec@univ-lille.fr