HAL open science

The Chowla and the Sarnak conjectures from ergodic theory point of view

El Houcein El Abdalaoui, Joanna Kulaga-Przymus, Mariusz Lemanczyk, Thierry de La Rue

To cite this version:

El Houcein El Abdalaoui, Joanna Kulaga-Przymus, Mariusz Lemanczyk, Thierry de La Rue. The Chowla and the Sarnak conjectures from ergodic theory point of view. 2014. hal-01070531v1

HAL Id: hal-01070531

https://hal.science/hal-01070531v1

Preprint submitted on 7 Oct 2014 (v1), last revised 27 Oct 2015 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Chowla and the Sarnak conjectures from ergodic theory point of view

H. El Abdalaoui, J. Kułaga-Przymus*, M. Lemańczyk*, T. de la Rue

October 7, 2014

Abstract

We rephrase the conditions from the Chowla and the Sarnak conjectures in abstract setting, that is, for sequences in $\{-1,0,1\}^{\mathbb{N}}$, and introduce several natural generalizations. We study the relationships between these properties and other notions from topological dynamics and ergodic theory.

Contents

1 Introduction **2**

2 Preliminaries **5**
2.1 Measure-theoretical dynamical systems 5
2.2 Topological dynamical systems 6

3 Ergodic theorem with Möbius weights **10**

4 The Chowla conjecture vs. the Sarnak conjecture – abstract approach **12**
4.1 Basic definitions 12
4.2 About (**Ch**) 13
4.3 (**Ch**) implies (**S**) 16
4.4 (**Ch**), (**S_0-strong**) and (**S-strong**) are equivalent 18
4.5 (**S_0**) and (**S**) are equivalent 20

5 (Ch) vs. various properties **26**
5.1 (**S**) does not imply (**Ch**) 26
5.2 (**Ch**) without genericity 27
5.3 The squares in (**Ch**) are necessary 28
5.4 (**Ch**) vs. recurrence 29
5.5 (**Ch**) vs. unique ergodicity 31

6 Entropy of sequences satisfying (Ch) **32**
6.1 Bounds for entropy 32
6.2 Two classes of sequences 40

*Research supported by Narodowe Centrum Nauki grant DEC-2011/03/B/ST1/00407.

7 Toeplitz sequences correlating with a given sequence, and their topological entropy **45**
7.1 Abstract setting 45
7.2 Applications 54

1 Introduction

A motivation for the present work comes from a dynamical point of view on some classical arithmetic functions taken up recently by Sarnak [21]. Namely, we consider the following two functions: the Möbius function $\boldsymbol{\mu}\colon \mathbb{N}^* := \mathbb{N}\backslash\{0\} \to \{-1,0,1\}$ given by $\boldsymbol{\mu}(1) = 1$ and

$$\boldsymbol{\mu}(n) = \begin{cases} (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes,} \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

and the Liouville function $\boldsymbol{\lambda}\colon \mathbb{N}^* \to \{-1,1\}$ defined by

$$\boldsymbol{\lambda}(n) = (-1)^{\Omega(n)},$$

where $\Omega(n)$ is the number of prime factors of n counted with multiplicities. An importance of these two functions in number theory is well known and may be illustrated by the following statement

$$\sum_{n\le N} \boldsymbol{\lambda}(n) = \mathrm{o}(N) = \sum_{n\le N} \boldsymbol{\mu}(n) \tag{2}$$

which is equivalent to the Prime Number Theorem, see e.g. [3], p. 91.

Recall also the classical connection of $\boldsymbol{\mu}$ with the Riemann zeta function, namely

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\boldsymbol{\mu}(n)}{n^s} \text{ for any } s \in \mathbb{C} \text{ with } \Re(s) > 1.$$

In [24], it is shown that the Riemann Hypothesis is equivalent to the following: for each $\varepsilon > 0$, we have

$$\sum_{n\le N} \boldsymbol{\mu}(n) = \mathrm{O}_\varepsilon\left(N^{\frac{1}{2}+\varepsilon}\right) \text{ as } N \to \infty.$$

In [5], Chowla formulated the following conjecture on the correlations of the Möbius function: for each choice of $1 \le a_1 < \cdots < a_r$, $r \ge 0$, with $i_s \in \{1,2\}$, not all equal to 2, we have

$$\sum_{n\le N} \boldsymbol{\mu}^{i_0}(n) \cdot \boldsymbol{\mu}^{i_1}(n+a_1) \cdot \ldots \cdot \boldsymbol{\mu}^{i_r}(n+a_r) = \mathrm{o}(N). \tag{3}$$

Recently, Sarnak [21] formulated the following conjecture: for any dynamical system (X,T), where X is a compact metric space and T is homeomorphism of zero topological entropy, for any $f \in C(X)$ and any $x \in X$, we have

$$\sum_{n\le N} f(T^n x)\boldsymbol{\mu}(n) = \mathrm{o}(N). \tag{4}$$

From now on, we refer to (4) as Sarnak's conjecture. Moreover, it is also noted in [21] that for any measurable dynamical system $(X, \mathcal{B}, \mu, T)$, for any $f \in L^2(X, \mathcal{B}, \mu)$, the condition (4) holds a.e. As can be shown, this a.e. version of (4) is a consequence of the following Davenport's estimation [6]: for each $\varepsilon > 0$, we have

$$\max_{z \in \mathbb{T}} \left| \sum_{n \leq N} z^n \boldsymbol{\mu}(n) \right| \leq \frac{N}{\log^{\varepsilon} N}, \tag{5}$$

combined with the spectral theorem (for a complete proof of the a.e. version of (4), see Section 3). Finally, Sarnak also proved that the Chowla conjecture (3) implies (4).

The aim of this paper is to deal with the Chowla conjecture (3) and the Sarnak conjecture (4) in a more abstract setting. In Section 4.1, we introduce conditions (**Ch**) and (**S**$_0$) in the context of arbitrary sequences $z \in \{-1, 0, 1\}^{\mathbb{N}}$. They are obtained from (3) and (4) by replacing $\boldsymbol{\mu}$ with z, respectively. In other words, we consider the sums of the form:

$$\sum_{n \leq N} z^{i_0}(n) z^{i_1}(n + a_1) \cdot \ldots \cdot z^{i_r}(n + a_r) \tag{6}$$

and

$$\sum_{n \leq N} f(T^n x) z(n), \tag{7}$$

and require that they are of order $\mathrm{o}(N)$ (a_s and i_s are as in (3), T, f and x are as in (4)). Finally, we define a new condition (**S**), formally stronger than (**S**$_0$), by requiring that the sum given by (7) is of order $\mathrm{o}(N)$ for any homeomorphism $T \colon X \to X$ of a compact metric space, any $f \in C(X)$ and any completely deterministic point $x \in X$.[1] Note that if $h_{top}(T) = 0$ then all points are completely deterministic.

We provide a detailed proof of the fact that (**Ch**) implies (**S**), see Theorem 4.10 below. Classical tools from ergodic theory, such as joinings (see Section 4.3), will be crucial for the proof. This approach (for $z = \boldsymbol{\mu}$ and (**S**$_0$) instead of (**S**)) was suggested in [21], together with a rough sketch of the proof.[2] Since (**S**) implies (**S**$_0$) directly from the definitions, we obtain the following:

$$(\textbf{Ch}) \implies (\textbf{S}) \implies (\textbf{S}_0).$$

By replacing (7) with the sums of the form

$$\sum_{n \leq N} f(T^n x) z^{i_0}(n) \cdot z^{i_1}(n + a_1) \cdot \ldots \cdot z^{i_r}(n + a_r) \tag{8}$$

in (**S**$_0$) and (**S**), we obtain conditions called (**S**$_0$**-strong**) and (**S-strong**), respectively. Notice that such sums generalize both (6) and (7). Clearly

$$(\textbf{S-strong}) \implies (\textbf{S}_0\textbf{-strong}) \implies (\textbf{Ch}).$$

[1] Recall that $x \in X$ is said to be completely deterministic if for any accumulation point ν of $(\frac{1}{N} \sum_{n \leq N} \delta_{T^n x})_{N \in \mathbb{N}}$, the system (X, ν, T) is of zero entropy.

[2] Sarnak also announced (and sent us a letter) a purely combinatorial proof of this result, see also [23].

In Section 4.4, we show that the above three properties are, in fact, equivalent:

$$(\mathbf{S\text{-}strong}) \iff (\mathbf{S}_0\mathbf{\text{-}strong}) \iff (\mathbf{Ch}).$$

Section 4.5 is devoted to the proof of Theorem 4.23 which says that although formally (**S**) is stronger than ($\mathbf{S}_0$), in fact, we have

$$(\mathbf{S}) \iff (\mathbf{S}_0).$$

Section 5 answers some natural questions about possible relations between the properties under discussion. First, in Section 5.1, we show that

$$(\mathbf{S}) \not\Longrightarrow (\mathbf{Ch}).$$

In Section 5.2, we show that a sequence $z \in \{-1,0,1\}^{\mathbb{N}}$ satisfying (**Ch**) need not be generic. In Section 5.3, we give an example of a sequence satisfying (**Ch**) with all $i_s = 1$, but failing to satisfy (**Ch**) in its full form. Finally, in Section 5.4 and Section 5.5, we discuss the properties of recurrence and unique ergodicity for sequences satisfying (**Ch**).

The remaining part of the paper deals with entropy from the point of view of condition (**Ch**). Section 6 is motivated by the question of describing the set

$$\{(h_{top}(z^2), h_{top}(z)) : \ z \in \{-1,0,1\}^{\mathbb{N}} \text{ satisfying } (\mathbf{Ch})\}. \tag{9}$$

For any sequence w satisfying (**Ch**) and such that $w^2 = \boldsymbol{\mu}^2$, we have (cf. [21] and Remark 6.22 below) $(h_{top}(w^2), h_{top}(w)) = (\frac{6}{\pi^2}\log 2, \frac{6}{\pi^2}\log 3)$, which might suggest that $h_{top}(z)$ depends linearly on $h_{top}(z^2)$ when (**Ch**) is satisfied; we show that there is no simple relationship between these numbers. However, there are some natural restrictions for the values of the pair $(h_{top}(z^2), h_{top}(z))$ which we describe in Section 6.1. Moreover, we show that the set of possible values is the same with the additional assumption that z satisfies (**Ch**) as it is without this assumption. In Section 6.2, we first show that given a sequence $u \in \{-1,0,1\}^{\mathbb{N}}$ that satisfies (**Ch**) and a sequence $\eta \in \{-1,0,1\}^{\mathbb{N}}$ that is completely deterministic, then $z := u \cdot \eta$ also satisfies (**Ch**). We apply this to u generic for some Bernoulli measures and η being a Sturmian sequence and for $z = u \cdot \ eta$ we calculate the value of the pair $(h_{top}(z^2), h_{top}(z))$. This shows that some elements of (9) arise in a natural way. Finally, we deal with the values of $(h_{top}(z^2), h_{top}(z))$ for sequences z satisfying (**Ch**) and such $z^2 = \boldsymbol{\mu}^2_{\mathscr{B}}$ ($\boldsymbol{\mu}_{\mathscr{B}}$ is a generalization of $\boldsymbol{\mu}$ defined in [2], see Section 2.2.5).

In Section 7, we deal with Toeplitz sequences [8], [13] over the alphabet $\{-1,0,1\}$. Although Toeplitz sequences are obtained as a certain limit of periodic sequences, their behavior differs from the behavior of periodic sequences. Given a sequence $z \in \{-1,0,1\}^{\mathbb{N}}$, satisfying some extra assumptions, see Theorems 7.1 and 7.3, we construct Toeplitz sequences t, which are not orthogonal to z and are of positive topological entropy, providing also more precise entropy estimates. We apply this to $z = \boldsymbol{\mu}$, $z = \boldsymbol{\mu}_{\mathscr{B}}$ and to sequences satisfying (**Ch**), defined in Section 6.2.3 and Section 6.2.4. For further motivations and related results see [1], [9].

2 Preliminaries

2.1 Measure-theoretical dynamical systems

2.1.1 Factors and extensions

Let $T\colon (X,\mathcal{B},\mu)\to(X,\mathcal{B},\mu)$ and $S\colon (Y,\mathcal{A},\nu)\to(Y,\mathcal{A},\nu)$ be automorphisms of probability standard Borel spaces.

Definition 2.1. We say that S is a *factor* of T (or T is an *extension* of S) if there exists $\pi\colon (X,\mathcal{B},\mu)\to(Y,\mathcal{A},\nu)$ such that $S\circ\pi=\pi\circ T$. To simplify notation, we will identify the factor S with the σ-algebra $\pi^{-1}(\mathcal{A})\subset\mathcal{B}$. Moreover, any sub-$\sigma$-algebra $\mathcal{A}\subset\mathcal{B}$ which is T-invariant, will be identified with the corresponding factor $T|_{\mathcal{A}}\colon (X/\mathcal{A},\mathcal{A},\mu|_{\mathcal{A}})\to(X/\mathcal{A},\mathcal{A},\mu|_{\mathcal{A}})$.

Let now $S_i\colon (Y_i,\mathcal{A}_i,\nu_i)\to(Y_i,\mathcal{A}_i,\nu_i)$, $i=1,2$, be factors of $T\colon (X,\mathcal{B},\mu)\to(X,\mathcal{B},\mu)$, with the factoring maps $\pi_i\colon X\to Y_i$, $i=1,2$. We will denote by $(Y_1,\mathcal{A}_1,\nu_1)\vee(Y_2,\mathcal{A}_2,\nu_2)$ the smallest factor of $\mathcal{B}$ containing both $(\pi_1)^{-1}(\mathcal{A}_1)$ and $(\pi_2)^{-1}(\mathcal{A}_2)$.[3]

2.1.2 Entropy

Let $T\colon (X,\mathcal{B},\mu)\to(X,\mathcal{B},\mu)$ be an automorphism of a probability standard Borel space. Recall that the measure-theoretic entropy of T is defined in the following way. Given a finite measurable partition $Q=\{Q_1,\dots,Q_k\}$ of X, we define

$$H(Q)=-\sum_{m=1}^{k}\mu(Q_m)\log\mu(Q_m).^{4}$$

(We may also write $H_\mu(Q)$ if we need to underline the role of μ.) The measure-theoretic entropy of T with respect to the partition Q is then defined as

$$h_\mu(T,Q)=\lim_{N\to\infty}\frac{1}{N}H\left(\bigvee_{n=0}^{N-1}T^{-n}Q\right),$$

where $\bigvee_{n=0}^{N-1}T^{-n}Q$ is the coarsest refinement of all partitions $T^{-n}Q$, $n=0,\dots,N-1$.

Definition 2.2 (Kolmogorov and Sinai)**.** The *measure-theoretic entropy* of T is given by

$$h(T,\mu)=\sup_{Q}h_\mu(T,Q),$$

where the supremum is taken over all finite measurable partitions.

Definition 2.3. We say that $T\colon (X,\mathcal{B},\mu)\to(X,\mathcal{B},\mu)$ is a *K-system* if any non-trivial factor of T has positive entropy.

Definition 2.4. Let $T_i\colon (X_i,\mathcal{B}_i,\mu_i)\to(X_i,\mathcal{B}_i,\mu_i)$, $i=1,2$, be such that T_2 is a factor of T_1.

[3] This factor can be viewed as a joining of S_1 and S_2, see Section 2.1.3.

[4] We consider 2 as the base of logarithm.

- The quantity $h(T_1, \mu_1) - h(T_2, \mu_2)$ is called the *relative entropy* of T_1 with respect to T_2.
- If, the extension $T_1 \to T_2$ is non trivial, and if for any intermediate factor $T_3 \colon (X_3, \mathcal{B}_3, \mu_3) \to (X_3, \mathcal{B}_3, \mu_3)$ between T_1 and T_2, the relative entropy of T_3 with respect to T_2 is positive, unless T_3 is equal to T_2, we say that the extension $T_1 \to T_2$ is *relatively K*.

2.1.3 Joinings

Definition 2.5. Given automorphisms of probability standard Borel spaces

$$T_i \colon (X_i, \mathcal{B}_i, \mu_i) \to (X_i, \mathcal{B}_i, \mu_i), \ i = 1, \dots, k,$$

let $J(T_1, \dots, T_k)$ be the set of all probability measures ρ on $(X_1 \times \cdots \times X_k, \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_k)$, invariant under $T_1 \times \cdots \times T_k$ and such that $(\pi_i)_*(\rho) = \mu_i$, where $\pi_i \colon X_1 \times \cdots \times X_k \to X_i$ is given by $\pi_i(x_1, \dots, x_k) = x_i$ for $1 \le i \le k$. Any $\rho \in J(T_1, \dots, T_k)$ is called a *joining*.

Definition 2.6. Following [11], we say that T_1 and T_2 are *disjoint* if $J(T_1, T_2) = \{\mu_1 \otimes \mu_2\}$. We then write $T_1 \perp T_2$.

Suppose now that $T_3 \colon (X_3, \mathcal{B}_3, \mu_3) \to (X_3, \mathcal{B}_3, \mu_3)$ is a common factor of T_1 and T_2. To keep the notation simple, we assume that $\mathcal{B}_3$ is a sub-σ-algebra of both $\mathcal{B}_1$ and $\mathcal{B}_2$. Given $\lambda \in J(T_3, T_3)$, we define the *relatively independent extension* of λ, i.e. $\widehat{\lambda} \in J(T_1, T_2)$, by setting for each $A_i \in \mathcal{B}_i$, $i = 1, 2$:

$$\widehat{\lambda}(A_1 \times A_2) := \int_{X_1/\mathcal{B}_3 \times X_2/\mathcal{B}_3} \mathbb{E}(\mathbf{1}_{A_1} | \mathcal{B}_3)(x) \mathbb{E}(\mathbf{1}_{A_2} | \mathcal{B}_3)(y) \ d\lambda(x, y).$$

Consider now these $\widetilde{\Delta} \in J(T_1, T_2)$ which project down to the diagonal joining $\Delta \in J(T_3, T_3)$ given by $\Delta(A \times B) := \mu(A \cap B)$. If $\widehat{\Delta}$ is the only such joining, we say that T_1 and T_2 are *relatively independent* over their common factor T_3. We then write $T_1 \perp_{T_3} T_2$.

Remark 2.7 (Theorem 1 in [12])**.** If the extension $T_1 \to T_3$ is of zero relative entropy and $T_2 \to T_3$ is relatively K then $T_1 \perp_{T_3} T_2$. In particular, (taking for T_3 the trivial one-point system), if T_1 has zero entropy and T_2 is K, then $T_1 \perp T_2$.

2.2 Topological dynamical systems

2.2.1 Invariant measures

Let $T \colon X \to X$ be a continuous map of a compact metric space. We denote by $\mathcal{P}_T(X)$ the set of T-invariant probability measures on $(X, \mathcal{B})$ with $\mathcal{B}$ standing for the σ-algebra of Borel sets. The space of probability measures on X is endowed with the (metrizable) weak topology:

$$\nu_n \xrightarrow[n \to \infty]{} \nu \iff \int_X f \ d\nu_n \xrightarrow[n \to \infty]{} \int_X f \ d\nu \text{ for each } f \in C(X),$$

where $C(X)$ denotes the space of continuous functions on X. The weak topology is compact.

By the Krylov-Bogolyubov theorem, $\mathcal{P}_T(X) \neq \varnothing$. In fact, for any $x \in X$, if we set

$$\delta_{N,x} := \frac{1}{N} \sum_{n \leq N} \delta_{T^n x},$$ [5]

and if, for some subsequence $(N_k)_{k\in\mathbb{N}}$ and some probability measure ν, $\delta_{N_k,x} \xrightarrow[k\to\infty]{} \nu$, then $\nu \in \mathcal{P}_T(X)$. In such a situation, we say that x is *quasi-generic* for ν along (N_k), and we set

$$\text{Q-gen}(x) := \left\{ \nu \in \mathcal{P}_T(X) : \delta_{N_k,x} \xrightarrow[k\to\infty]{} \nu \text{ for a subsequence } (N_k) \right\}.$$

If $\delta_{N,x} \xrightarrow[N\to\infty]{} \nu$, i.e. if $\text{Q-gen}(x) = \{\nu\}$, we say that x is *generic* for ν.

Definition 2.8 ([26], see also [14])**.** We say that $x \in X$ is *completely deterministic* if, for each $\nu \in \text{Q-gen}(x)$, we have $h(T,\nu) = 0$. We will then write

$$\text{Q-gen}(x) \subset [h = 0]. \tag{10}$$

2.2.2 Symbolic dynamical systems

Let A be a finite set and $\mathbb{I} = \mathbb{N}$ or $\mathbb{Z}$. Then $A^{\mathbb{I}}$ endowed with the product topology is a compact metric space. Coordinates of $w \in A^{\mathbb{I}}$ will be denoted either by w_n or by $w(n)$ for $n \in \mathbb{I}$.

Definition 2.9. The subsets of $A^{\mathbb{I}}$ of the form

$$C_t(a_0, \ldots, a_{k-1}) = \{w \in A^{\mathbb{I}} \colon w_{t+j} = a_j \text{ for } 0 \leq j \leq k-1\},$$

where $k \geq 1$, $t \in \mathbb{I}$ and $a_0, \ldots, a_{k-1} \in A$, are called *cylinders* and they form a basis for the product topology.

Definition 2.10. Any $C = (a_0, \ldots, a_{k-1}) \in A^k$, $k \geq 1$, is called a *block of length* k. For any $0 \leq i \leq k-1$, let $C(i) := a_i$.

We will identify blocks with the corresponding cylinders:

$$C = (a_0, \ldots, a_{k-1}) \in A^k \longleftrightarrow C_0(a_0, \ldots, a_{k-1}).$$

Definition 2.11. We say that a block $C = (a_0, \ldots, a_{k-1}) \in A^k$ *appears on* w if $w \in C_t(a_0, \ldots, a_{k-1})$ for some $t \in \mathbb{I}$.

On $A^{\mathbb{I}}$ there is a natural continuous action by the *left shift* (for $\mathbb{I} = \mathbb{Z}$, S is clearly invertible and it is a homeomorphism):

$$S \colon A^{\mathbb{I}} \to A^{\mathbb{I}},\ S((w_n)_{n\in\mathbb{I}}) = (w_{n+1})_{n\in\mathbb{I}} \text{ for } w = (w_n)_{n\in\mathbb{I}} \in A^{\mathbb{I}}.$$

Definition 2.12. Let $C = (a_0, \ldots, a_{k-1}) \in A^k$. The following quantity is called the *upper frequency* with which C appears on w:

$$\overline{\text{fr}}(C, w) := \limsup_{N\to\infty} \frac{1}{N} \sum_{n \leq N} \mathbf{1}_C(S^n w) = \limsup_{N\to\infty} \int \mathbf{1}_C \, d\delta_{N,w}.$$

[5] In what follows, we will also use notation $\delta_{T,N,x}$ if confusion could arise.

We will denote by the same letter S the action by the left shift restricted to any closed shift-invariant subset of $A^{\mathbb{I}}$ (such a subset is called a *subshift*). In particular, given $w \in A^{\mathbb{I}}$, we will consider the two following subshifts:

$$X_w := \{u \in A^{\mathbb{I}} \colon \text{ all blocks that appear on } u \text{ also appear on } w\}$$

and

$$X_w^+ := \{u \in A^{\mathbb{I}} \colon \text{ all blocks that appear on } u \text{ appear on } w \text{ with positive upper frequency}\}. \tag{11}$$

Finally, let $F \in C(A^{\mathbb{I}})$ be given by

$$F(w) := w(0) \text{ for } w \in A^{\mathbb{I}}. \tag{12}$$

We will use the same notation F, even if the domain of F changes, e.g. when we consider a subshift.

2.2.3 Topological entropy

Let T be a homeomorphism of a compact metric space (X, d). For $n \in \mathbb{N}$, let

$$d_n(x, y) := \max\{d(T^i x, T^i y) \colon 0 \leq i < n\}.$$

Given $\varepsilon > 0$ and $n \in \mathbb{N}$, let

$$N(\varepsilon, n) = \max\{|E| : E \subset X, d_n(x, y) \geq \varepsilon \text{ for all } x \neq y \text{ in } E\}.$$

Definition 2.13 (Bowen and Dinaburg)**.** The *topological entropy* $h_{top}(T)$ is defined as

$$h_{top}(T) = h_{top}(T, X) := \lim_{\varepsilon \to 0} \left(\limsup_{n \to \infty} \frac{1}{n} \log N(\varepsilon, n) \right).$$

We consider now the special case of a subshift, namely, $S \colon X_w \to X_w$, where $w \in A^{\mathbb{I}}$. Let

$$p_n(w) := |\{B \in A^n \colon B \text{ appears on } w\}|.$$

and put

$$h_{top}(w) := \lim_{n \to \infty} \frac{1}{n} \log p_n(w).$$

Then

$$h_{top}(w) = h_{top}(S, X_w). \tag{13}$$

In a similar way, given $\nu \in \mathcal{P}_S(A^{\mathbb{I}})$, we denote by $h_{top}(\mathrm{supp}(\nu))$ the following quantity:

$$h_{top}(\mathrm{supp}(\nu)) := \lim_{n \to \infty} \frac{1}{n} \log p_n(\mathrm{supp}(\nu)),$$

where

$$p_n(\mathrm{supp}(\nu)) := |\{B \in A^n \colon \nu(B) > 0\}|.$$ [6]

In particular, if Q-gen$(w) = \{\nu\}$, then $h_{top}(\mathrm{supp}(\nu)) = h_{top}(S, X_w^+)$ (see Lemma 5.11 below).

[6] It is not hard to see that $p_{m+n}(\mathrm{supp}(\nu)) \leq p_m(\mathrm{supp}(\nu)) \cdot p_n(\mathrm{supp}(\nu))$.

2.2.4 Invariant measures in symbolic dynamical systems

Remark 2.14. Any $\nu \in \mathcal{P}_S(A^{\mathbb{N}})$ is determined by the values it takes on blocks, so it can be extended to a measure in $\mathcal{P}_S(A^{\mathbb{Z}})$ taking the same value on each block as ν. This measure will be also denoted by ν.[7] Moreover, if $w \in A^{\mathbb{N}}$ is quasi-generic for $\nu \in \mathcal{P}_S(A^{\mathbb{N}})$ along (N_k) then for any $\overline{w} \in A^{\mathbb{Z}}$ such that $\overline{w}[0,\infty] = w$, the point $\overline{w}$ is quasi-generic for $\nu \in \mathcal{P}_S(A^{\mathbb{Z}})$ along (N_k).

For any probability distribution $(p_1, \dots, p_{|A|})$ on A, we denote by $B(p_1, \dots, p_{|A|})$ the corresponding Bernoulli measure on $A^{\mathbb{I}}$.

The cases $A = \{-1, 0, 1\}$ or $A = \{0, 1\}$ will be of special interest for us. Let $\pi : \{-1, 0, 1\}^{\mathbb{I}} \to \{0, 1\}^{\mathbb{I}}$ be the coordinate square map:

$$(\pi(w))_n := w_n^2,$$

which is clearly S-equivariant.

Given $\nu \in \mathcal{P}_S(\{0, 1\}^{\mathbb{I}})$, let $\widehat{\nu}$ denote the corresponding *relatively independent extension of* ν: for every block B, we set

$$\widehat{\nu}(B) := 2^{-|\mathrm{supp}(B)|}\nu(\pi(B)) = 2^{-|\mathrm{supp}(B)|}\nu(B^2), \tag{14}$$

where $\mathrm{supp}(B) := \{i : B(i) \neq 0\}$ and $B^2(i) := B(i)^2$. Clearly, $\widehat{\nu} \in \mathcal{P}_S(\{-1, 0, 1\}^{\mathbb{I}})$.

2.2.5 Möbius function and its generalizations

The following generalization of the Möbius function $\boldsymbol{\mu} \colon \mathbb{N}^* \to \{-1, 0, 1\}$ defined by (1) has been introduced in [2]. Let $\mathscr{B} = \{b_k \colon k \geq 1\} \subset \{2, 3, \dots\}$ be such that $b_k = a_k^2$ and a_k, $a_{k'}$ are relatively prime for $k \neq k'$. For $n \in \mathbb{N}$, let

$$\eta_{\mathscr{B}}(n) := \begin{cases} 0 & \text{if } b_k | n \text{ for some } k \geq 1, \\ 1 & \text{otherwise,} \end{cases}$$

$$\delta(n) := |\{k \geq 1 \colon a_k | n\}|$$

and

$$\boldsymbol{\mu}_{\mathscr{B}}(n) := (-1)^{\delta(n)} \cdot \eta_{\mathscr{B}}(n). \tag{15}$$

The classical case $\boldsymbol{\mu}$ corresponds to $\mathscr{B}$ being the set of squares of all primes.

2.2.6 Sturmian sequences

Usually Sturmian sequences are defined in the following way:

Definition 2.15. Let $w \in A^{\mathbb{I}}$, where A is a finite set. We say that the sequence w is *Sturmian* if $p_n(w) = n+1$ for all $n \in \mathbb{N}$ (in particular, $|A| = 2$, i.e. without loss of generality, $A = \{0, 1\}$).

Remark 2.16. Any Sturmian sequence can be obtained in the following way. Consider a line L with an irrational slope in the plane (see Figure 2 on page 42). We build w by considering the consecutive intersections of L with the integer grid, putting a 0 each time L intersects a horizontal line and a 1 each time it intersects a vertical line of the grid (if the line intersects a node, put either 0 or 1).

[7] The invertible dynamical system $S \colon (A^{\mathbb{Z}}, \nu) \to (A^{\mathbb{Z}}, \nu)$ is the natural extension of the non-invertible system $S \colon (A^{\mathbb{N}}, \nu) \to (A^{\mathbb{N}}, \nu)$.

Remark 2.17. Since we would like to use also some periodic sequences, we *extend* the notion of Sturmian sequences by allowing the slope of the line L to be *rational*, provided that the grid is chosen in such a way that L does not meet any node. For any sequence w obtained this way, $p_n(w) \leq n+1$.

Remark 2.18. Recall that any Sturmian sequence w is generic for a measure ν of zero entropy. Moreover, $\nu(B) > 0$ for any block B appearing on w.

For more information on Sturmian sequences, we refer the reader e.g. to [10].

2.2.7 Toeplitz sequences

Definition 2.19. Let $t \in A^{\mathbb{I}}$, where A is a finite set. We say that the sequence t is *Toeplitz* if for each $a \in \mathbb{I}$ there exists r_a such that $t(a) = t(a + kr_a)$ for each $k \in \mathbb{I}$.

Each Toeplitz sequence $t \in A^{\mathbb{I}}$ is obtained as a limit of some periodic sequences defined over the extended alphabet $A \cup \{*\}$. Namely, there exists an increasing sequence (p_n), $p_n | p_{n+1}$ such that for each $n \geq 1$,

$$t_n := T_n T_n \ldots, \quad \lim_{n\to\infty} t_n(j) = t(j) \text{ for each } j \in \mathbb{N},$$

where, for each $n \geq 1$, T_n is a block of length p_n over the alphabet $A \cup \{*\}$ and $*$ at position k at instance n means that $t(k)$ has not been defined at the stage n of the construction.

Whenever

$$(\text{the number of } * \text{ in } T_n)/p_n \to 0 \text{ when } n \to \infty,$$

we say that t is *regular*. The dynamical systems generated by regular Toeplitz sequences are uniquely ergodic and have zero entropy.

For non-regular Toeplitz sequences the entropy can be positive. Moreover, non-regular Toeplitz sequences can display extremely non-uniquely ergodic behavior.[8]

For more information about Toeplitz sequences, we refer the reader to [8], [13], [28].

3 Ergodic theorem with Möbius weights

Proposition 3.1. *Let T be an automorphism of a probability standard Borel space $(X, \mathcal{B}, \mu)$ and let $f \in L^1(X, \mathcal{B}, \mu)$. Then, for almost every $x \in X$, we have*

$$\frac{1}{N} \sum_{n \leq N} f(T^n x) \boldsymbol{\mu}(n) \xrightarrow[N\to\infty]{} 0.$$

Proof. We may assume without loss of generality that T is ergodic. Fix $f \in L^2(X, \mathcal{B}, \mu)$. By the Spectral Theorem, we have

$$\Big\| \frac{1}{N} \sum_{n \leq N} f(T^n x) \boldsymbol{\mu}(n) \Big\|_2 = \Big\| \frac{1}{N} \sum_{n \leq N} z^n \boldsymbol{\mu}(n) \Big\|_{L^2(\sigma_f)},$$

[8] Downarowicz proved that each abstract Choquet simplex can be realized as the simplex of invariant measures for a Toeplitz subshift.

where σ_f is the spectral measure of f.[9] Hence, by Davenport's estimation (5), for each $\varepsilon > 0$, we obtain

$$\Big\| \frac{1}{N} \sum_{n \le N} f(T^n x) \boldsymbol{\mu}(n) \Big\|_2 \le \frac{C}{\log^{\varepsilon} N}, \tag{16}$$

where C is a constant which depends only on ε. Take $\rho > 1$, then for $N = [\rho^m]$ for some $m \ge 1$, (16) takes the form

$$\Big\| \frac{1}{N} \sum_{n \le N} f(T^n x) \boldsymbol{\mu}(n) \Big\|_2 \le \frac{C}{(m \log(\rho))^{\varepsilon}} \text{ for any } \varepsilon > 0.$$

By choosing $\varepsilon = 2$, we obtain

$$\sum_{m \ge 1} \Big\| \frac{1}{[\rho^m]} \sum_{n \le [\rho^m]} f(T^n x) \boldsymbol{\mu}(n) \Big\|_2 < +\infty.$$

Hence, for almost every point $x \in X$, we have

$$\frac{1}{[\rho^m]} \sum_{n \le [\rho^m]} f(T^n x) \boldsymbol{\mu}(n) \xrightarrow[m \to \infty]{} 0. \tag{17}$$

Suppose additionally that $f \in L^\infty(X, \mathcal{B}, \mu)$. Then, if $[\rho^m] \le N < [\rho^{m+1}] + 1$, we obtain

$$\begin{aligned} \Big| \frac{1}{N} \sum_{n \le N} f(T^n x) \boldsymbol{\mu}(n) \Big| &= \Big| \frac{1}{N} \sum_{n \le [\rho^m]} f(T^n x) \boldsymbol{\mu}(n) + \frac{1}{N} \sum_{[\rho^m]+1 \le n \le N} f(T^n x) \boldsymbol{\mu}(n) \Big| \\ &\le \Big| \frac{1}{[\rho^m]} \sum_{n \le [\rho^m]} f(T^n x) \boldsymbol{\mu}(n) \Big| + \frac{\|f\|_\infty}{[\rho^m]} (N - [\rho^m]) \\ &\le \Big| \frac{1}{[\rho^m]} \sum_{n \le [\rho^m]} f(T^n x) \boldsymbol{\mu}(n) \Big| + \frac{\|f\|_\infty}{[\rho^m]} ([\rho^{m+1}] - [\rho^m]). \end{aligned}$$

Since $\frac{\|f\|_\infty}{[\rho^m]} ([\rho^{m+1}] - [\rho^m]) \xrightarrow[m \to +\infty]{} \|f\|_\infty (\rho - 1)$, using (17) and the fact that ρ can be taken arbitrarily close to 1, we obtain

$$\frac{1}{N} \sum_{n \le N} f(T^n x) \boldsymbol{\mu}(n) \xrightarrow[N \to \infty]{} 0 \text{ for a.e. } x \in X.$$

To finish the proof, notice that for any $f \in L^1(X, \mathcal{B}, \mu)$, and any $\varepsilon > 0$, there exists $g \in L^\infty(X, \mathcal{B}, \mu)$ such that $\|f - g\|_1 < \varepsilon$. It follows by the pointwise ergodic theorem that for almost all $x \in X$, we have

$$\lim_{N \longrightarrow \infty} \Big| \frac{1}{N} \sum_{n \le N} (f - g)(T^n x) \Big| < \varepsilon.$$

[9] Recall that σ_f is a finite Borel measure on the circle determined by its Fourier transform given by $\widehat{\sigma}_f(n) = \int f \circ T^n \cdot \overline{f} \, d\mu$, $n \in \mathbb{Z}$.

Hence,

$$\limsup_{N\longrightarrow\infty}\Big|\frac{1}{N}\sum_{n\leq N}f(T^nx)\boldsymbol{\mu}(n)\Big|\leq$$
$$\lim_{N\longrightarrow\infty}\Big|\frac{1}{N}\sum_{n\leq N}(f-g)(T^nx)\Big|+\limsup_{N\longrightarrow\infty}\Big|\frac{1}{N}\sum_{n\leq N}g(T^nx)\boldsymbol{\mu}(n)\Big|<\varepsilon.$$

Since $\varepsilon>0$ is arbitrary, the proof is complete. □

4 The Chowla conjecture vs. the Sarnak conjecture – abstract approach

4.1 Basic definitions

We will now introduce the necessary definitions concerning the Chowla conjecture and the Sarnak conjecture in the abstract setting, i.e. for arbitrary sequences, not only for $\boldsymbol{\mu}$.

Definition 4.1 (cf. [5], [21])**.** We say that $z\in\{-1,0,1\}^{\mathbb{I}}$ satisfies the condition (**Ch**) if

$$\textbf{(Ch)}\qquad \frac{1}{N}\sum_{n\leq N}z^{i_0}(n)\cdot z^{i_1}(n+a_1)\cdot\ldots\cdot z^{i_r}(n+a_r)\xrightarrow[N\to\infty]{}0$$

for each choice of $1\leq a_1<\ldots<a_r$, $r\geq 0$, $i_s\in\{1,2\}$ not all equal to 2.

Whenever (**Ch**) is satisfied for z, we will also say that z satisfies the Chowla conjecture.

Definition 4.2 (cf. [21])**.** We say that $z\in\{-1,0,1\}^{\mathbb{I}}$ satisfies the condition ($\mathbf{S}_0$) if, for any homeomorphism T of a compact metric space X with $h_{top}(T)=0$, and for any $x\in X$, we have

$$(\mathbf{S}_0)\qquad \frac{1}{N}\sum_{n\leq N}f(T^nx)z(n)\xrightarrow[N\to\infty]{}0.$$

Definition 4.3. We say that $z\in\{-1,0,1\}^{\mathbb{I}}$ satisfies the condition (**S**) if, for any homeomorphism T of a compact metric space X,

$$(\mathbf{S})\qquad \frac{1}{N}\sum_{n\leq N}f(T^nx)z(n)\xrightarrow[N\to\infty]{}0$$

for each $f\in C(X)$ and each $x\in X$ that is completely deterministic.

Whenever (**S**) is satisfied for z, we will also say that z satisfies the Sarnak conjecture.

Note that by the variational principle, see e.g. [25], if the topological entropy of T is zero, then all points are completely deterministic. Hence (**S**) implies ($\mathbf{S}_0$).

4.2 About (Ch)

Fix $z \in \{-1,0,1\}^{\mathbb{N}}$. Suppose that z^2 is quasi-generic for ν along (N_k), i.e. we have

$$\delta_{N_k,z^2} := \frac{1}{N_k} \sum_{n \le N_k} \delta_{S^n z^2} \xrightarrow[k\to\infty]{} \nu \in \mathcal{P}_S(X_{z^2}).^{10} \tag{18}$$

Remark 4.4. In the classical situation $z = \boldsymbol{\mu}$, z^2 is generic for the Mirsky measure [17], cf. [4], [21]. Moreover, the Mirsky measure on X_{z^2} has full topological support, cf. (51). In a more general framework, similar results hold for so called $\mathscr{B}$-free systems, see [2].

Recall that the function F was given by the formula (12), i.e. $F(w) = w(0)$.

Lemma 4.5. *Let $1 \le a_1 < \ldots < a_r$, $r \ge 0$ and $i_s \in \{1,2\}$, $0 \le s \le r$. Then the following equalities hold:*

$$\int_{\{-1,0,1\}^{\mathbb{Z}}} F^{i_0} \cdot F^{i_1} \circ S^{a_1} \cdot \ldots \cdot F^{i_r} \circ S^{a_r} \; d\widehat{\nu} = 0,$$

when not all i_s are equal to 2. Moreover,

$$\int_{\{-1,0,1\}^{\mathbb{Z}}} F^2 \cdot F^2 \circ S^{a_1} \cdot \ldots \cdot F^2 \circ S^{a_r} \; d\widehat{\nu} = \int_{\{0,1\}^{\mathbb{Z}}} F \cdot F \circ S^{a_1} \cdot \ldots \cdot F \circ S^{a_r} \; d\nu.$$

Proof. The assertion follows directly by the calculation:

$$\begin{aligned}
&\int_{\{-1,0,1\}^{\mathbb{Z}}} F^{i_0} \cdot F^{i_1} \circ S^{a_1} \cdot \ldots \cdot F^{i_r} \circ S^{a_r} \; d\widehat{\nu} \\
&\quad = \sum_{j_0,j_1,\ldots,j_r=\pm 1} j_0^{i_0} \cdot j_1^{i_1} \cdot \ldots \cdot j_r^{i_r} \\
&\qquad \cdot \widehat{\nu}\left(\left\{y \in \{-1,0,1\}^{\mathbb{N}} \colon (y(0), y(a_1), \ldots, y(a_r)) = (j_0, j_1, \ldots, j_r)\right\}\right) \\
&\quad = \Big(\sum_{j_0,j_1,\ldots,j_r=\pm 1} j_0^{i_0} \cdot j_1^{i_1} \cdot \ldots \cdot j_r^{i_r} \Big) \\
&\qquad \cdot \frac{1}{2^{r+1}} \nu\left(\left\{u \in \{0,1\}^{\mathbb{N}} \colon u(0) = u(a_1) = \ldots = u(a_r) = 1\right\}\right).
\end{aligned}$$

□

Lemma 4.6 (cf. [21] for $\boldsymbol{\mu}$)**.** *Let (N_k) be such that (18) holds. Then*

$$\frac{1}{N_k} \sum_{n \le N_k} z^{i_0}(n) \cdot z^{i_1}(n+a_1) \cdot \ldots \cdot z^{i_r}(n+a_r) \xrightarrow[k\to\infty]{} 0 \tag{19}$$

for each choice of $1 \le a_1 < \ldots < a_r$, $r \ge 0$, $i_s \in \{1,2\}$ not all equal to 2, if and only if

$$\delta_{N_k,z} \xrightarrow[k\to\infty]{} \widehat{\nu}. \tag{20}$$

[10] In view of Remark 2.14, we may consider $\mathcal{P}_S(X_{z^2})$ to be a subset of $\mathcal{P}_S(\{0,1\}^{\mathbb{Z}})$, i.e. $\nu \in \mathcal{P}_S(\{0,1\}^{\mathbb{Z}})$. In the same way, $\widehat{\nu} \in \mathcal{P}_S(\{-1,0,1\}^{\mathbb{Z}})$.

Proof. Note that, for each $k \geq 1$,

$$\frac{1}{N_k} \sum_{n \leq N_k} z^{i_0}(n) \cdot z^{i_1}(n+a_1) \cdot \ldots \cdot z^{i_r}(n+a_r) = \frac{1}{N_k} \sum_{n \leq N_k} \big(F^{i_0} \cdot F^{i_1} \circ S^{a_1} \cdot \ldots \cdot F^{i_r} \circ S^{a_r}\big)(S^n z). \tag{21}$$

Suppose that (20) holds. Then it follows from (21) that

$$\frac{1}{N_k} \sum_{n \leq N_k} z^{i_0}(n) \cdot z^{i_1}(n+a_1) \cdot \ldots \cdot z^{i_r}(n+a_r) \xrightarrow[k\to\infty]{} \int_{\{-1,0,1\}^{\mathbb{Z}}} F^{i_0} \cdot F^{i_1} \circ S^{a_1} \cdot \ldots \cdot F^{i_r} \circ S^{a_r} \, d\widehat{\nu}.$$

Therefore, in view of Lemma 4.5, we obtain (19).

Suppose now that (19) holds. Without loss of generality, we may assume that

$$\delta_{N_k,z} \xrightarrow[k\to\infty]{} \rho. \tag{22}$$

In view of (21), this implies

$$\frac{1}{N_k} \sum_{n \leq N_k} z^{i_0}(n) \cdot z^{i_1}(n+a_1) \cdot \ldots \cdot z^{i_r}(n+a_r) \xrightarrow[k\to\infty]{} \int_{\{-1,0,1\}^{\mathbb{Z}}} F^{i_0} \cdot F^{i_1} \circ S^{a_1} \cdot \ldots \cdot F^{i_r} \circ S^{a_r} \, d\rho.$$

It follows from (19) that

$$\int_{\{-1,0,1\}^{\mathbb{Z}}} F^{i_0} \cdot F^{i_1} \circ S^{a_1} \cdot \ldots \cdot F^{i_r} \circ S^{a_r} \, d\rho = 0, \tag{23}$$

whenever not all i_t are equal to 2. Moreover, since $F^2(u) = F(u^2)$ for any $u \in \{-1,0,1\}^{\mathbb{Z}}$, we deduce from (18) that

$$\int_{\{-1,0,1\}^{\mathbb{Z}}} F^2 \cdot F^2 \circ S^{a_1} \cdot \ldots \cdot F^2 \circ S^{a_r} \, d\rho = \int_{\{0,1\}^{\mathbb{Z}}} F \cdot F \circ S^{a_1} \cdot \ldots \cdot F \circ S^{a_r} \, d\nu. \tag{24}$$

In view of Lemma 4.5, (23) and (24), we have

$$\int_{\{-1,0,1\}^{\mathbb{Z}}} G \, d\widehat{\nu} = \int_{\{-1,0,1\}^{\mathbb{Z}}} G \, d\rho$$

for any

$$G \in \mathcal{A} := \{F^{i_0} \cdot F^{i_1} \circ S^{a_1} \cdot \ldots \cdot F^{i_r} \circ S^{a_r} : 1 \leq a_1 < \cdots < a_r, r \geq 0, i_s \in \mathbb{N}\}.$$

Since $\mathcal{A} \subset C(\{-1,0,1\}^{\mathbb{Z}})$ is closed under taking products and separates points, we only need to use the Stone-Weierstrass theorem to conclude that $\rho = \widehat{\nu}$. $\square$

The above lemma can be also viewed from the probabilistic point of view. Indeed, let $(X_n)_{n\geq 0}$ (or $(X_n)_{n\in\mathbb{Z}}$) be a stationary sequence of random variables taking values in $\{-1,0,1\}$. Notice that whenever

$$\mathbb{P}\big(\{X_0=j_0, X_{a_1}=j_1,\ldots,X_{a_r}=j_r\}\big) = \frac{1}{2^k}\mathbb{P}\big(\{X_0^2=j_0^2, X_{a_1}^2=j_1^2,\ldots,X_{a_r}^2=j_r^2\}\big), \tag{25}$$

for each choice of $1\leq a_1<\ldots<a_r$ and $j_s\in\{-1,0,1\}$, where $k:=|\{s\in\{0,1,\ldots,r\}: j_s\neq 0\}|$ then

$$\mathbb{E}(X_0^{i_0}\cdot X_{a_1}^{i_1}\cdot\ldots\cdot X_{a_r}^{i_r})=0 \tag{26}$$

for each choice of $1\leq a_1<\ldots<a_r$, $r\geq 0$, $i_s\in\{1,2\}$ not all equal to 2 (the proof is the same as the one of Lemma 4.5 with notational changes only).[11]

In fact, the following holds:

Lemma 4.7. *Conditions* (25) *and* (26) *are equivalent.*

Proof. We have already seen that (25) implies (26). Let us show the converse implication. In other words, we need to show that there exists at most one stationary process (that is, at most one S-invariant distribution on $\{-1,0,1\}^{\mathbb{N}}$) such that (26) holds. However, each stationary process (X_n) is entirely determined by the family

$$\{\mathbb{E}(\exp(i\sum_{j=0}^{n-1}t_jX_j)) : n\geq 1, (t_0,\ldots,t_{n-1})\in\mathbb{R}^n\}.$$

Since $\mathbb{E}(\exp(i\sum_{j=0}^{n-1}t_jX_j))=\sum_{k=0}^{\infty}\frac{i^k}{k!}\mathbb{E}\left(\sum_{j=0}^{n-1}t_jX_j\right)^k$, the result follows. □

As the proof shows, the above lemma can be proved in a more general framework, namely, for stationary processes having moments of all orders.

Remark 4.8. It follows immediately from Lemma 4.6 that each of the following conditions is equivalent to (**Ch**):

- Q-gen$(z)=\{\widehat{\nu}: \nu\in\text{Q-gen}(z^2)\}$;
- $\delta_{N_k,z^2}\xrightarrow[k\to\infty]{}\nu$ if and only if $\delta_{N_k,z}\xrightarrow[k\to\infty]{}\widehat{\nu}$.

Now, we can completely characterize sequences $z\in\{-1,1\}^{\mathbb{N}}$ satisfying (**Ch**).

Corollary 4.9. *The only sequences $u\in\{-1,1\}^{\mathbb{N}}$ satisfying* (**Ch**) *are generic points for the Bernoulli measure $B(1/2,1/2)$.*

Proof. Notice that u^2 is the generic point for the Dirac measure at $(1,1,...)$ and by Lemma 4.6, u is a generic point for the relatively independent extension of that Dirac measure, which is the Bernoulli measure $B(1/2,1/2)$. □

[11] Condition (25) means that the distribution of the process $(X_n)_{n\geq 0}$ is the relatively independent extension of the distribution of the (stationary) process $(X_n^2)_{n\geq 0}$.

4.3 (Ch) implies (S)

In this section, we will provide a dynamical proof of the following theorem:

Theorem 4.10 (Sarnak)**.** (**Ch**) *implies* (**S**).

Remark 4.11. In particular, (**Ch**) implies ($\mathbf{S}_0$) (see [21]), which has already been proved by Sarnak (in a letter; see also the proof by Tao [23]). The proof of the implication (**Ch**) $\implies$ (**S**) given below is to be compared with Sarnak's arguments on page 9 of [21]. Later, in Theorem 4.23, we show that (**S**) and ($\mathbf{S}_0$) are equivalent. Hence, another way to prove Theorem 4.10 is to use (**Ch**) $\implies$ ($\mathbf{S}_0$) and (**S**) $\iff$ ($\mathbf{S}_0$).

Fix some $\nu \in \mathcal{P}_S(\{0,1\}^{\mathbb{Z}})$.

Lemma 4.12. *The dynamical system* $(S, \{-1,0,1\}^{\mathbb{Z}}, \widehat{\nu})$ *is a factor of*

$$(S, \{0,1\}^{\mathbb{Z}}, \nu) \times (S, \{-1,1\}^{\mathbb{Z}}, B(1/2,1/2)).$$

Proof. It suffices to notice that, for $\xi\colon \{0,1\}^{\mathbb{Z}} \times \{-1,1\}^{\mathbb{Z}} \to \{-1,0,1\}^{\mathbb{Z}}$ given by

$$\xi(w,u)(n) := w(n) \cdot u(n),$$

we have

$$\xi_*(\nu \otimes B(1/2,1/2)) = \widehat{\nu},$$

which is straightforward by the definition of $\widehat{\nu}$. $\square$

Lemma 4.13. *The extension* $(S, \{-1,0,1\}^{\mathbb{Z}}, \widehat{\nu}) \xrightarrow{\pi} (S, \{0,1\}^{\mathbb{Z}}, \nu)$ *is either trivial (i.e. 1-1 a.e.) or relatively K.*

Proof. Notice that since the extension

$$(S, \{0,1\}^{\mathbb{Z}}, \nu) \times (S, \{-1,1\}^{\mathbb{Z}}, B(1/2,1/2)) \to (S, \{0,1\}^{\mathbb{Z}}, \nu)$$

is relatively K (see Theorem 1 in [12]), so is any nontrivial intermediate factor (over $(S, \{0,1\}^{\mathbb{Z}}, \nu)$). To see that $(S, \{-1,0,1\}^{\mathbb{Z}}, \widehat{\nu})$ is an intermediate factor, by the proof of Lemma 4.12, all we need to check is that $\pi \circ \xi$ equals to the projection on the first coordinate. The latter follows from the equality $w = (w \cdot u)^2$ which holds for $w \in \{0,1\}^{\mathbb{Z}}$ and $u \in \{-1,1\}^{\mathbb{Z}}$. $\square$

Remark 4.14. It is possible that the extension $(S, \{-1,0,1\}^{\mathbb{Z}}, \widehat{\nu}) \xrightarrow{\pi} (S, \{0,1\}^{\mathbb{Z}}, \nu)$ is trivial. In fact, it happens only if $\nu = \delta_{(\ldots,0,0,0,\ldots)}$. For, suppose that $Y \subset \{-1,0,1\}^{\mathbb{Z}}$, $\widehat{\nu}(Y) = 1$ is such that $\pi|_Y$ is 1-1. Fix a block $B \in \{0,1\}^k$ with $\nu(B) > 0$. Then the set $\{x \in \pi(Y) : x(n) = B(n),\ n = 0,1,\ldots,k-1\}$ is of positive ν-measure, and $|\pi^{-1}(x) \cap Y| \geq 2^{\mathrm{supp}(B)}$ as each block $C \in \{-1,0,1\}^k$, $C^2 = B$, has positive $\widehat{\nu}$-measure (whence $\widehat{\nu}(Y \cap C) > 0$). It follows immediately that the support of B has to be empty.

Lemma 4.15. $\mathbb{E}^{\widehat{\nu}}(F|\pi(w) = u) = 0$ *for* ν*-a.e.* $u \in \{0,1\}^{\mathbb{Z}}$.

Proof. We have

$$\mathbb{E}^{\widehat{\nu}}(F|\pi(w) = u) = \mathbb{E}^{\widehat{\nu}}(F|\{0,1\}^{\mathbb{Z}})(u) = \int_{\pi^{-1}(u)} F \, d\widehat{\nu}_u,$$

where $\widehat{\nu}_u$ is the relevant conditional measure in the disintegration of $\widehat{\nu}$ over ν. Notice that $\widehat{\nu}_u$ is the product measure $(1/2, 1/2)$ of all positions belonging to the support of u. If $u(0) = 0$ then the formula holds. If $u(0) = 1$ then F on $\pi^{-1}(u)$ takes two values ± 1 with the same probability, so the integral is still zero. $\square$

Consider now a homeomorphism T of a compact metric space X.

Lemma 4.16. *Let $x \in X$ be completely deterministic and suppose that z is a quasi-generic point for $\widehat{\nu}$ along the sequence (N_k). Assume that*

$$\delta_{T\times S, N_k, (x,z)} \to \rho \tag{27}$$

weakly in $\mathcal{P}_{T\times S}(X \times \{-1,0,1\}^{\mathbb{Z}})$. Then:

(a) ρ is a joining of (T, X, κ) and $(S, \{-1,0,1\}^{\mathbb{Z}}, \widehat{\nu})$ for some zero entropy measure $\kappa \in$ Q-gen(x);

(b) the factors $(T, X, \kappa) \vee (S, \{0,1\}^{\mathbb{Z}}, \nu)$ and $(S, \{-1,0,1\}^{\mathbb{Z}}, \widehat{\nu})$ are relatively independent over $(S, \{0,1\}^{\mathbb{Z}}, \nu)$ as factors of $(T \times S, X \times \{-1,0,1\}^{\mathbb{Z}}, \rho)$.

Proof. It follows from (27) that

$$\kappa := \rho|_X = \lim_{k\to\infty} \delta_{T, N_k, x},$$

and $h(T, \kappa) = 0$ since x is completely deterministic. Hence ρ is a joining of (T, X, κ) and $(S, \{-1,0,1\}^{\mathbb{Z}}, \widehat{\nu})$, and the extension

$$(T, X, \kappa) \vee (S, \{0,1\}^{\mathbb{Z}}, \nu) \to (S, \{0,1\}^{\mathbb{Z}}, \nu)$$

has relative entropy zero (by the Pinsker formula, see e.g. [19], Theorem 6.3). On the other hand, by Lemma 4.13, the extension

$$(S, \{-1,0,1\}^{\mathbb{Z}}, \widehat{\nu}) \to (S, \{0,1\}^{\mathbb{Z}}, \nu)$$

is relatively K. To complete the proof, we only need to use Remark 2.7. $\square$

Proof of Theorem 4.10. Assume that z satisfies (**Ch**). Fix (N_k) such that

$$\delta_{T\times S, N_k, (x,z)} \xrightarrow[k\to\infty]{} \rho \tag{28}$$

for some measure ρ. Then by Remark 4.8, the projection of ρ onto the second coordinate is of the form $\widehat{\nu}$ for some $\nu \in$ Q-gen(z^2). Take a function $f \in C(X)$. It follows from (28) that

$$\frac{1}{N_k} \sum_{n \le N_k} f(T^n x) z(n) = \frac{1}{N_k} \sum_{n \le N_k} f(T^n x) F(S^n z) \xrightarrow[k\to\infty]{} \int f \otimes F \, d\rho. \tag{29}$$

Using Lemma 4.15, we have

$$\mathbb{E}^{\rho}(F|\{0,1\}^{\mathbb{Z}}) = \mathbb{E}^{\widehat{\nu}}(F|\{0,1\}^{\mathbb{Z}}) = 0. \tag{30}$$

Therefore, and by Lemma 4.16 (b),

$$\mathbb{E}^{\rho}(f \otimes F|\{0,1\}^{\mathbb{Z}}) = \mathbb{E}^{\rho}(f|\{0,1\}^{\mathbb{Z}}) \, \mathbb{E}^{\rho}(F|\{0,1\}^{\mathbb{Z}}) = 0.$$

This yields $\int f \otimes F \, d\rho = 0$. $\square$

4.4 (Ch), (S_0-strong) and (S-strong) are equivalent

In this section, we will throw some more lights on Theorem 4.10, by considering some strengthening of properties of (**S**)-type.

Definition 4.17. A sequence $z \in \{-1,0,1\}^{\mathbb{I}}$ is said to satisfy the condition (**S_0-strong**) if for each homeomorphism T of a compact metric space X, with $h_{top}(T) = 0$, we have

$$(\mathbf{S_0\text{-strong}}) \qquad \frac{1}{N}\sum_{n\le N} f(T^n x) z^{i_0}(n)\cdot z^{i_1}(n+a_1)\cdot\ldots\cdot z^{i_r}(n+a_r) \xrightarrow[N\to\infty]{} 0$$

for each $f \in C(X)$, each $x \in X$ and each choice of $1 \le a_1 < \ldots < a_r$, $r \ge 0$, $i_s \in \{1,2\}$ not all equal to 2.

Definition 4.18. A sequence $z \in \{-1,0,1\}^{\mathbb{I}}$ is said to satisfy the condition (**S-strong**) if for each homeomorphism T of a compact metric space X, we have

$$(\mathbf{S\text{-strong}}) \qquad \frac{1}{N}\sum_{n\le N} f(T^n x) z^{i_0}(n)\cdot z^{i_1}(n+a_1)\cdot\ldots\cdot z^{i_r}(n+a_r) \xrightarrow[N\to\infty]{} 0$$

for each $f \in C(X)$, each completely deterministic $x \in X$ and each choice of $1 \le a_1 < \ldots < a_r$, $r \ge 0$, $i_s \in \{1,2\}$ not all equal to 2.

If the above holds, we will also say that z satisfies the strong Sarnak conjecture. In particular, for $z = \boldsymbol{\mu}$ the strong Sarnak conjecture takes the form

$$\frac{1}{N}\sum_{n\le N} f(T^n x) \boldsymbol{\mu}^{i_0}(n)\cdot \boldsymbol{\mu}^{i_1}(n+a_1)\cdot\ldots\cdot \boldsymbol{\mu}^{i_r}(n+a_r) \xrightarrow[N\to\infty]{} 0.$$

Proposition 4.19. *The conditions* (**Ch**), (**S_0-strong**) *and* (**S-strong**) *are equivalent.*

For the proof, we will need the following result.

Lemma 4.20. *Let $z \in \{-1,0,1\}^{\mathbb{I}}$ and let*

$$u(n) := z^{i_0}(n)\cdot z^{i_1}(n+a_1)\cdot\ldots\cdot z^{i_r}(n+a_r),\ n \in \mathbb{I},$$

for some natural numbers $1 \le a_1 < a_2 < \ldots < a_r$ and $i_s \in \{1,2\}$. Then the following holds:

(a) If z satisfies (**Ch**) *then u satisfies* (**Ch**) *provided that not all i_s are equal to 2.*

(b) If z is completely deterministic, then so is u.[12]

Proof. We write

$$\frac{1}{N}\sum_{n\le N} u^{j_0}(n)\cdot u^{j_1}(n+b_1)\cdot\ldots\cdot u^{j_t}(n+b_t)$$
$$= \frac{1}{N}\sum_{n\le N}\prod_{\alpha=0}^{r}\prod_{\beta=0}^{t} z^{i_\alpha j_\beta}(n+a_\alpha+b_\beta)$$

[12] In particular, this holds, if we replace z with z^2 and u with u^2.

with $a_0 = b_0 = 0$. Consider then the smallest α and β such that $i_\alpha = j_\beta = 1$. Since both sequences (a_i), (b_j) are strictly increasing, the sum $a_\alpha + b_\beta$ can be obtained only as $a_\gamma + b_\delta$ with either $a_\gamma < a_\alpha$ or $b_\delta < b_\beta$. It follows that, in the above sum, the term $z(n + a_\alpha + b_\beta)$ appears with the power

$$i_\alpha j_\beta + \text{ even number},$$

that is, an odd power, which completes the proof of part (a) of the lemma.

We will show now that the assertion (b) also holds. Suppose that

$$\delta_{S,N_k,u} \to \rho$$

and consider the following sequence of measures on $\{-1,0,1\}^{\mathbb{I}} \times \cdots \times \{-1,0,1\}^{\mathbb{I}}$:

$$\begin{aligned}\widetilde{\rho}_k :=& \frac{1}{N_k} \sum_{n \le N_k} \delta_{S^n z} \otimes \delta_{S^n S^{a_1} z} \otimes \cdots \otimes \delta_{S^n S^{a_r} z} \\ =& \frac{1}{N_k} \sum_{n \le N_k} \delta_{\left(\underbrace{S \times \ldots \times S}_{r+1}\right)^n (z, S^{a_1} z, \ldots, S^{a_r} z)}.\end{aligned}$$

Passing to a subsequence if necessary, we may assume that $\widetilde{\rho}_k$ converges to $\widetilde{\rho}$. Then $\widetilde{\rho}$ is a joining of $(S,\kappa_0), (S,\kappa_1), \ldots, (S,\kappa_r)$, where $\kappa_s \in$ Q-gen(z) for $0 \le s \le r$. Hence $h(S,\kappa_s) = 0$ for $0 \le s \le r$ and it follows that $h(S^{\times(r+1)}, \widetilde{\rho}) = 0$. Notice that $S\colon (\{-1,0,1\}^{\mathbb{I}}, \rho) \to (\{-1,0,1\}^{\mathbb{I}}, \rho)$ is a factor of

$$S^{\times(r+1)}\colon ((\{-1,0,1\}^{\mathbb{I}})^{\times(r+1)}, \widetilde{\rho}) \to ((\{-1,0,1\}^{\mathbb{I}})^{\times(r+1)}, \widetilde{\rho}),$$

with the factoring map $(x_0, \ldots, x_r) \mapsto x_0 \cdot \ldots \cdot x_r$. Therefore, we obtain $h(S,\rho) = 0$ and the assertion follows. □

Proof of Proposition 4.19. Since clearly (**S-strong**) implies (**S$_0$-strong**), which, in turn, implies (**Ch**), it suffices to show that (**Ch**) implies (**S-strong**). This however follows immediately from Theorem 4.10 and Lemma 4.20. □

Moreover, in view of Proposition 4.19 and Corollary 4.9, we immediately obtain the following:

Corollary 4.21. *If* (**Ch**) *holds for* $\boldsymbol{\lambda}$ *then for each homeomorphism* T *of a compact metric space* X*, we have*

$$\frac{1}{N} \sum_{n \le N} f(T^n x) \boldsymbol{\lambda}(n) \cdot \boldsymbol{\lambda}(n + a_1) \cdot \ldots \cdot \boldsymbol{\lambda}(n + a_r) \xrightarrow[N \to \infty]{} 0$$

for each $f \in C(X)$*, each completely deterministic* $x \in X$ *and for each choice of* $1 \le a_1 < a_2 < \ldots < a_r$*,* $r \ge 0$.[13]

Remark 4.22. Since the Bernoulli shifts are disjoint from all zero entropy transformations, arguments similar to those used in the proof of Theorem 4.10, together with Lemma 4.20, can be used to obtain another proof of Corollary 4.21.

[13] Clearly, since $\boldsymbol{\lambda}$ takes values in $\{-1,1\}$, we can remove the exponents i_s appearing originally in the condition (**Ch**).

4.5 ($\mathbf{S}_0$) and ($\mathbf{S}$) are equivalent

The purpose of this section is to prove the following result.

Theorem 4.23. *Properties* ($\mathbf{S}_0$) *and* ($\mathbf{S}$) *are equivalent.*

The first part of the proof deals with the symbolic case and shows that if a sequence u is quasi-generic for some shift-invariant measure of zero entropy, then u can be well approximated by a sequence that has zero topological entropy. In [27], the following characterization of completely deterministic points was stated without a proof:

> A sequence u is completely deterministic if and only if, for any $\varepsilon > 0$ there exists K such that, after removing from u a subset of density less than ε, what is left can be covered by a collection $\mathcal{C}$ of K-blocks such that $|C| < 2^{\varepsilon K}$.

The following lemma is a reformulation of this criterion in a language suitable for our needs.

Lemma 4.24. *Let A be finite nonempty set, and let $(N_k)_{k\geq 1}$ be an increasing sequence of integers, with $N_k | N_{k+1}$ for each k. Assume that $u \in A^{\mathbb{N}}$ satisfies*

$$\delta_{N_k,u} = \frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n u} \xrightarrow[k\to\infty]{} \nu, \tag{31}$$

where ν is such that $h(S,\nu) = 0$.

Then, for any $\varepsilon > 0$, we can find an arbitrarily large integer k and a map $\varphi : A^{N_k} \to A^{N_k}$, satisfying the following properties:

- $\left|\varphi\left(A^{N_k}\right)\right| < 2^{\varepsilon N_k}$;
- *the sequence $\overline{u}$ obtained from u by replacing, for each $j \geq 0$, the block $u|_{jN_k+1}^{(j+1)N_k}$ by its image by φ, is such that for each $s \geq 0$,*

$$\frac{1}{N_{k+s}} \left|\left\{1 \leq n \leq N_{k+s} : u_n \neq \overline{u}_n\right\}\right| < \varepsilon; \tag{32}$$

- *the first symbol occuring in $\overline{u}$ is the same as in u.*

We will need the following lemma taken from [22], (Lemma 1.5.4 p. 52).

Lemma 4.25. *For $0 < \delta < 1$, set*

$$H(\delta) := -\delta \log \delta - (1-\delta)\log(1-\delta).$$

Then, for any integer $N \geq 1$ and any $0 < \delta \leq 1/2$,

$$\sum_{k \leq \delta N} \binom{N}{k} \leq 2^{NH(\delta)}.$$

Proof of Lemma 4.24. Let P be the finite partition of $A^{\mathbb{N}}$ defined by the zero-th coordinate. Then $\bigvee_{j=0}^{n-1} S^{-j}P$ is the partition of $A^{\mathbb{N}}$ according to the n-block

appearing on coordinates from 0 to $n-1$. Since the entropy of (S,ν) vanishes, given an arbitrary $\delta > 0$, we can take n large enough so that

$$\frac{1}{n} H_\nu \left(\bigvee_{j=0}^{n-1} S^{-j} P \right) < \delta. \tag{33}$$

Now, let us say that an n-block is *heavy* if the ν-measure of the corresponding cylinder set is larger than $2^{-\varepsilon n}$, and say it is *light* otherwise. We claim that the ν-measure of the union of all light n-blocks is arbitrarily small whenever δ is chosen small enough. Indeed, for any light n-block B we have

$$\nu(B) \log \nu(B) \leq \nu(B) \log \frac{1}{2^{n\varepsilon}} = -\nu(B) \cdot n\varepsilon. \tag{34}$$

This and (33) imply

$$\varepsilon n \sum_{\text{light } n\text{-blocks } B} \nu(B) \leq - \sum_{\text{light } n\text{-blocks } B} \nu(B) \log(B) < \delta n,$$

which gives

$$\sum_{\text{light } n\text{-blocks } B} \nu(B) < \frac{\delta}{\varepsilon}.$$

Observe also that the number of heavy n-blocks cannot exceed $2^{\varepsilon n}$.

Say that an integer $j \geq 1$ is *good in* u if the n-block $u|_j^{j+n-1}$ is heavy. By (31) (applied to the characteristic function of the union of all light n-blocks), and assuming δ is small enough, we can take k large enough so that, for each $s \geq 0$,

$$\frac{1}{N_{k+s}} \left| \{ 1 \leq j \leq N_{k+s} : j \text{ is not good in } u \} \right| < \varepsilon^2/2. \tag{35}$$

We can also assume that k is large enough so that

$$\frac{n}{N_k} < \frac{\varepsilon}{2}. \tag{36}$$

Let us now define the map $\varphi \colon A^{N_k} \to A^{N_k}$. Let $W \in A^{N_k}$; we say that $j \in \{1, \ldots, N_k\}$ is *good in* W if $j + n - 1 \leq N_k$ and the n-block $W|_j^{j+n-1}$ is heavy. We say that W is *acceptable* if the proportion of $j \in \{1, \ldots, N_k\}$ which are good in W is larger than $1 - \varepsilon$. The definition of $\varphi(W)$ will depend on whether W is acceptable or not. If W is not acceptable, then we simply set $\varphi(W) := a^{N_k}$, where $a \in A$ is the first symbol occuring in the sequence u. If W is acceptable, then we run the following algorithm. Let j_1 be the first integer which is good in W, and inductively, define j_{i+1} as the smallest integer larger than or equal to $j_i + n$ which is good in W, provided such an integer exists. This algorithm outputs a finite list of integers $j_1, \ldots, j_r$ which are good in W, such that $j_i + n \leq j_{i+1}$, and such that the disjoint heavy n-blocks $W|_{j_i}^{j_i+n-1}$, $1 \leq j \leq r$, cover a proportion at least $(1 - \varepsilon)$ of W (because symbols which are not covered correspond to integers which are not good in W). Then, in W, replace by a all symbols which are not covered by these heavy n-blocks, and define $\varphi(W)$ as the resulting N_k-block.

The number of N_k-blocks which are images of some acceptable block W by this procedure is bounded by the number of choices for the subset of $\{1, \ldots, N_k\}$

where we put the letter a, times the number of choices for the heavy blocks. The former is bounded by the number of subsets of $\{1, \dots, N_k\}$ which have less than $\varepsilon\, N_k$ elements, which is at most $2^{H(\varepsilon)N_k}$ by Lemma 4.25. Since the number of heavy blocks is at most $2^{\varepsilon n}$, the latter is bounded by $(2^{\varepsilon n})^r$, which is less than $2^{\varepsilon N_k}$ (indeed, $nr \leq N_k$ because in W we see r disjoint heavy blocks of length n).

Observe that, by the construction of φ and by the choice of a, the first symbol in $\overline{u}$ is the same as in u.

Now, it only remains to show that (32) holds. Let $s \geq 0$. Each $m \in \{0, \dots, N_{k+s}/N_k - 1\}$ such that $u|_{mN_k+1}^{(m+1)N_k}$ is not acceptable gives rise in the corresponding subblock to at least εN_k integers j which are not good *in this subblock*. But there are two reasons why this could happen:

- either j is one of the last n positions of the subblock, which by (36) only concerns a number of integers at most $\varepsilon N_k/2$,
- or j is not good in u, which therefore concerns at least $\varepsilon N_k/2$ integers j in this subblock.

Then, (35) ensures that the proportion of integers $m \in \{0, \dots, N_{k+s}/N_k - 1\}$ such that $u|_{mN_k+1}^{(m+1)N_k}$ is not acceptable is less than ε. Moreover, observe that if W is an acceptable N_k-block, then $\varphi(W)$ differs from W in at most $\varepsilon\, N_k$ places. This concludes the proof of the lemma. □

Lemma 4.26. *Let k and $\overline{u}$ be produced as in Lemma 4.24. Let us consider $\overline{u}$ as a sequence in $(A^{N_k})^{\mathbb{N}}$, and denote by S_{N_k} the action of the shift map in this setting (that is, S_{N_k} shifts N_k letters in A at the same time). Set also, for each integer $s \geq 0$, $M_s := N_{k+s}/N_k$. Then there exists a subsequence $(s_\ell)_{\ell \geq 1}$, and an S_{N_k}-invariant probability $\overline{\nu}$ on $(A^{N_k})^{\mathbb{N}}$ such that*

- *we have the weak convergence*

$$\delta_{S_{N_k}, M_{s_\ell}, \overline{u}} = \frac{1}{M_{s_\ell}} \sum_{n \leq M_{s_\ell}} \delta_{S_{N_k}^n \overline{u}} \xrightarrow[\ell \to \infty]{} \overline{\nu},$$

- $h(S_{N_k}, \overline{\nu}) = 0$.

Proof. First, let μ be any weak limit of a subsequence of the form $\delta_{S^{N_k}, M_{s_\ell}, u}$, $\ell \geq 1$. Then μ is S^{N_k}-invariant. Moreover, we have

$$\nu = \frac{1}{N_k}(\mu + S_* \mu + \cdots + S_*^{N_k - 1} \mu).$$

Since $h(\nu, S_{N_k}) = 0$, we have also $h(\mu, S_{N_k}) = 0$. Let Φ be the continuous map defined by the N_k-block recoding φ from Lemma 4.24. We get the announced result with $\overline{\nu}$ the pushforward measure of μ by Φ. □

Lemma 4.27. *With the same assumptions as in Lemma 4.24, for any $\varepsilon > 0$, we can find a sequence $\overline{u} \in A^{\mathbb{N}}$ and a subsequence $(N_{k(\ell)})_{\ell \geq 1}$ such that:*

- $h_{top}(\overline{u}) = 0$;
- *for each $\ell \geq 1$, (32) with k replaced by $k(\ell)$, is satisfied.*

Proof. Let $\overline{u}^{(1)}$ be the sequence we obtain applying Lemma 4.24 with $\varepsilon/2$, and let $k(1)$ be the corresponding integer k. Then $\overline{u}^{(1)}$ can be viewed as an infinite concatenation of at most $2^{N_{k(1)}\varepsilon/2}$ different $N_{k(1)}$-blocks. By Lemma 4.26, and since all integers N_k, $k \geq k(1)$, are multiples of $N_{k(1)}$, we can apply Lemma 4.24 to the new sequence $\overline{u}^{(1)}$ itself, viewed as a sequence in $(A^{N_{k(1)}})^{\mathbb{N}}$. Doing this with $\varepsilon/4$, we obtain a new sequence $\overline{u}^{(2)}$ and an integer $k(2)$. If we consider both $\overline{u}^{(1)}$ and $\overline{u}^{(2)}$ as concatenation of $N_{k(1)}$-blocks, all blocks used in $\overline{u}^{(2)}$ are already used in $\overline{u}^{(1)}$, so that $\overline{u}^{(2)}$ is itself an infinite concatenation of at most $2^{N_{k(1)}\varepsilon/2}$ different $N_{k(1)}$-blocks. On the other hand, if we consider now both $\overline{u}^{(1)}$ and $\overline{u}^{(2)}$ as sequences in $A^{\mathbb{N}}$, they coincide on their first $N_{k(1)}$ symbols.

We go on in the same way by induction. At step ℓ, we have constructed a sequence $\overline{u}^{(\ell)}$ and we have an integer $k(\ell)$ satisfying

$$\frac{1}{N_{k(\ell)+s}} \left|\{1 \leq n \leq N_{k(\ell)+s} : u_n \neq \overline{u}^{(\ell)}_n\}\right| < \frac{\varepsilon}{2} + \cdots + \frac{\varepsilon}{2^\ell} \text{ for all } s \geq 0, \tag{37}$$

and for each $1 \leq j \leq \ell$,

(38) $\overline{u}^{(\ell)}$ is an infinite concatenation of at most $2^{N_{k(j)}\varepsilon/2^j}$ different $N_{k(j)}$-blocks.

Consider $\overline{u}^{(\ell)}$ as a sequence on the alphabet $A^{N_{k(\ell)}}$, which is quasi-generic for some $S_{N_{k(\ell)}}$-invariant probability with zero entropy along a subsequence of the original sequence (N_k). We apply on it Lemma 4.24 with $\varepsilon/2^{\ell+1}$ to get a new sequence $\overline{u}^{(\ell+1)}$ and an integer $k(\ell+1)$, satisfying the analog properties to (37) and (38) at level $\ell+1$, and such that $\overline{u}^{(\ell+1)}$ coincides with $\overline{u}^{(\ell)}$ on their first $N_{k(\ell)}$ symbols.

The sequence $(\overline{u}^{(\ell)})_{\ell\geq 1}$ which is obtained in this way, converges to a sequence $\overline{u}$, satisfying for all $\ell \geq 1$,

$$\overline{u}|_1^{N_{k(\ell)}} = \overline{u}^{(\ell)}|_1^{N_{k(\ell)}}.$$

By (37), this ensures that for each $\ell \geq 1$,

$$\frac{1}{N_{k(\ell)}} \left|\{1 \leq n \leq N_{k(\ell)} : u_n \neq \overline{u}_n\}\right| < \varepsilon.$$

Moreover, by (38), for each $\ell \geq 1$, $\overline{u}$ is an infinite concatenation of at most $2^{N_{k(\ell)}\varepsilon/2^\ell}$ different $N_{k(\ell)}$-blocks. Therefore, there are at most $N_{k(\ell)} \cdot 2^{N_{k(\ell)}\varepsilon/2^\ell}$ different $N_{k(\ell)}$-blocks which appear on $\overline{u}$. This implies that $h_{top}(\overline{u}) = 0$. □

To conclude the proof of the equivalence of (**S**) and (**S**$_0$), we need also some tool to pass from the continuous case of a general sequence $\big(f(T^n x)\big)_{n\in\mathbb{N}}$ to the discrete case of a symbolic sequence $x \in A^{\mathbb{N}}$ for some finite $A \subset \mathbb{R}$. This is the object of what follows.

For each finite subset $A \subset \mathbb{R}$, we denote by φ_A the function from $[\min A, +\infty)$ to A which maps $t \geq \min A$ to the largest element $a \in A$ satisfying $a \leq t$. We also denote by Φ_A the function from $[\min A, +\infty)^{\mathbb{N}}$ to $A^{\mathbb{N}}$ which maps each sequence $(y_n)_{n\in\mathbb{N}}$ to $\big(\varphi_A(y_n)\big)_{n\in\mathbb{N}}$.

Lemma 4.28. *Let $y = (y_n)_{n\in\mathbb{N}}$ be a bounded sequence of real numbers, with values in some compact interval $[\alpha, \beta]$. We assume that, along some subsequence of integers (N_k), the following weak convergence holds:*

$$\delta_{S,N_k,y} \xrightarrow[k\to\infty]{} \mu,$$

where μ is a shift-invariant probability on $[\alpha, \beta]^{\mathbb{N}}$. Then, for each $\varepsilon > 0$, we can find a finite subset $A \subset \mathbb{R}$ such that:

- $\forall t \in [\alpha, \beta]$, $|\varphi_A(t) - t| < \varepsilon$,
- *we have the weak convergence*

$$\delta_{S,N_k,\Phi_A y} \xrightarrow[k\to\infty]{} (\Phi_A)_*\mu. \tag{39}$$

Proof. The first condition required on A is easily satisfied: we just have to choose A so that

- $\min A < \alpha$,
- $\sup A \geq \beta$,
- the distance between two consecutive elements of A is always less than ε.

Then, for such an A, observe that φ_A is continuous on $[\min A, +\infty) \setminus A$ (A is the set of discontinuity points of φ_A), and that Φ_A is continuous on

$$[\min A, +\infty)^{\mathbb{N}} \setminus r(A),$$

where

$$r(A) := \{y \in [\alpha, \beta]^{\mathbb{N}} : y_n \in A \text{ for some } n \in \mathbb{N}\}.$$

Consider the pushforward measure of μ by the projection of $[\alpha, \beta]^{\mathbb{N}}$ to the first coordinate: This is a probability measure on the interval $[\alpha, \beta]$, with at most a countable number of atoms. Moreover, the pushforward measure of μ by the projection of $[\alpha, \beta]^{\mathbb{N}}$ to any other coordinate has the same atoms, since μ is shift-invariant. Choosing the elements of A from the complement of this set of atoms is always possible, and ensures that

$$\mu\big(r(A)\big) = 0. \tag{40}$$

Finally, note that for any k, the pushforward of $\delta_{S,N_k,y}$ by Φ_A is precisely $\delta_{S,N_k,\Phi_A y}$. Since by (40), the set of discontinuities of Φ_A has μ-measure 0, we get (39). □

Proof of Theorem 4.23. It is clear from the definitions that condition (**S**) implies (**S**$_0$).

Assume that $z \in \{-1, 0, 1\}^{\mathbb{N}}$ does not satisfy (**S**). Then there exist a homeomorphism T of a compact metric space X, a continuous function $f : X \to \mathbb{R}$, and a completely deterministic point $x \in X$ such that $\frac{1}{N}\sum_{n\leq N} f(T^n x)z_n$ does not converge to 0 as $N \to \infty$. We can thus find some increasing sequence of integers (N_k), and some $\theta \neq 0$ such that

$$\frac{1}{N_k}\sum_{n\leq N_k} f(T^n x)z_n \xrightarrow[k\to\infty]{} \theta. \tag{41}$$

Without loss of generality, we can further assume that $N_k | N_{k+1}$ for each k. Indeed, extracting a subsequence if necessary, we can always assume that

$$N_k/N_{k+1} \xrightarrow[k\to\infty]{} 0,$$

and then replace inductively each N_{k+1} by the closest multiple of N_k.

We can also assume that

$$\delta_{T,N_k,x} \xrightarrow[k\to\infty]{} \nu,$$

where ν is a T-invariant probability measure on X satisfying $h(T,\nu) = 0$ (because x is completely deterministic). Let $\alpha := \min f$, $\beta := \sup f$. If we set $y = (y_n)_{n\in\mathbb{N}} := \big(f(T^n x)\big)_{n\in\mathbb{N}} \in [\alpha,\beta]^{\mathbb{N}}$, then we also have

$$\delta_{S,N_k,y} \xrightarrow[k\to\infty]{} \mu,$$

where μ is the pushforward of ν to $[\alpha,\beta]^{\mathbb{N}}$ by the topological factor map defined by f. In particular, we have $h(S,\mu) = 0$.

Now, choose $\varepsilon > 0$, small enough so that

$$\varepsilon < |\theta|/4. \tag{42}$$

Let A be the finite set given by Lemma 4.28 applied to $y = f(T^n x)$ and ε, and set $u = (u_n)_{n\in\mathbb{N}} := \Phi_A(y)$. Then, we have

$$|u_n - f(T^n x)| < \varepsilon \text{ for all } n \in \mathbb{N}, \tag{43}$$

and

$$\delta_{S,N_k,u} \xrightarrow[k\to\infty]{} (\Phi_A)_*\mu.$$

Moreover, since $h(S,\mu) = 0$ and $\big(S, A^{\mathbb{N}}, (\Phi_A)_*\mu\big)$ is a measurable factor of $\big(S,[\alpha,\beta]^{\mathbb{N}},\mu\big)$, we also have $h\big(S,(\Phi_A)_*\mu\big) = 0$.

We apply now Lemma 4.27 to u and ε, obtaining a sequence $\overline{u}$ with $h_{top}(\overline{u}) = 0$ and a subsequence $(N_{k(\ell)})$ such that

$$\frac{1}{N_{k(\ell)}} \big|\big\{1 \le n \le N_{k(\ell)} : u_n \ne \overline{u}_n\big\}\big| < \varepsilon. \tag{44}$$

Then

$$\frac{1}{N_{k(\ell)}} \sum_{n\le N_{k(\ell)}} \overline{u}_n z_n = \frac{1}{N_{k(\ell)}} \sum_{n\le N_{k(\ell)}} (\overline{u}_n - u_n) z_n$$
$$+ \frac{1}{N_{k(\ell)}} \sum_{n\le N_{k(\ell)}} (u_n - f(T^n x)) z_n + \frac{1}{N_{k(\ell)}} \sum_{n\le N_{k(\ell)}} f(T^n x) z_n.$$

It follows from (44), by (43) and by (41) that for ℓ sufficiently large

$$\left| \frac{1}{N_{k(\ell)}} \sum_{n\le N_{k(\ell)}} \overline{u}_n z_n \right| \ge |\theta|/2 - 2\varepsilon.$$

Therefore (42) implies that

$$\frac{1}{N} \sum_{n\le N} \overline{u}_n z_n \not\to 0,$$

i.e. z does not satisfy ($\mathbf{S}_0$) and the assertion follows. □

5 (Ch) vs. various properties

5.1 (S) does not imply (Ch)

A natural question arises, whether it is possible to find a sequence which satisfies (**S**) and does not satisfy (**Ch**). We will provide now such an example.

Example 5.1 (z that satisfies (**S**) but not (**Ch**))**.** Consider the shift on $\{0,1,2,3\}^{\mathbb{N}}$ with the Bernoulli measure $B(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}) =: \kappa$ and let

$$\theta : \{0,1,2,3\}^{\mathbb{N}} \to \{-1,0,1\}^{\mathbb{N}}$$

be given by the code of length 2: $\theta(01) = \theta(12) = -1$, $\theta(02) = \theta(23) = 1$, all remaining blocks of length 2 sent to 0. Let $\omega \in \{0,1,2,3\}^{\mathbb{N}}$ be a generic point for κ (such a point exists by the ergodic theorem). Let $\nu := \theta_*(\kappa)$ be the measure for the subshift $Y := \theta(\{0,1,2,3\}^{\mathbb{N}} \subset \{-1,0,1\}^{\mathbb{N}}$. Then $z := \theta(\omega)$ is a generic point for ν and (S,Y,ν) is a Bernoulli automorphism [18]. Recalling that $F(u) := u(0)$ for $u \in Y$, we have

$$\frac{1}{N}\sum_{n\le N} z^{i_0}(n)\cdot z^{i_1}(n+a_1)\cdot\ldots\cdot z^{i_r}(n+a_r)$$
$$= \frac{1}{N}\sum_{n\le N}\big(F^{i_0}\cdot F^{i_1}\circ S^{a_1}\cdot\ldots\cdot F^{i_r}\circ S^{a_r}\big)(S^n z).$$

In particular (by genericity),

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N} z(n)z(n+1) = \int_Y F\cdot F\circ S\ d\nu. \tag{45}$$

Observe that

$$\int_Y F\ d\nu = 0. \tag{46}$$

However, the function $F\cdot F\circ S$ takes the value 1 with probability $2/4^3$ (given by the blocks 012 and 023), while the value -1 has probability $1/4^3$ (it is given by 123). It follows that the integral in (45) is not equal to zero, i.e. (**Ch**) does not hold. Note that, in this construction, z is a generic point (so the more z^2 is a generic point).

It remains to show that z satisfies (**S**). This is however clear: for any topological dynamical system (X,T) and any $x \in X$, each accumulation point, say ρ, of the sequence of empiric measures $\delta_{T\times S,N,(x,z)}$, $N \ge 1$, is a joining of $(T,X,\rho|_X)$ and (S,Y,ν) (the latter, since z is generic for ν). If x is completely deterministic, then $\rho|_X \in$ Q-gen(x) has zero entropy, hence $(X,T,\rho|_X)$ is disjoint from any K-system. In particular, $\rho = \rho_X \otimes \nu$ and (**S**) follows from (46).

Remark 5.2. The point z in the above example is clearly not completely deterministic. In fact, we have the following:

(47) if z satisfies (**S**) and is completely deterministic then z satisfies (**Ch**).

Indeed, since z is completely deterministic and satisfies (**S**), we have

$$\frac{1}{N}\sum_{n\le N} z^2(n) = \frac{1}{N}\sum_{n\le N} z(n)\cdot z(n) \to 0,$$

i.e. the support of z has zero density. It follows immediately that (**Ch**) also holds for z.

Remark 5.3. Example 5.1 can be seen as a starting point for a construction of sequences such that the convergence in (**Ch**) holds whenever $a_r < k_0$ ($k_0 \ge 2$), and fails for some choice of $1 \le a_1 < \dots < a_r = k_0$. Indeed, consider again the shift on $\{0,1,2,3\}^{\mathbb{N}}$ with the Bernoulli measure $B(1/4,1/4,1/4,1/4)$ and let

$$\theta : \{0,1,2,3\}^{\mathbb{N}} \to \{-1,0,1\}^{\mathbb{N}}$$

be given by the code of length k_0: $\theta(0*1) = \theta(1*2) = -1$, $\theta(0*2) = \theta(2*3) = 1$, where $*$ stands for any sequence of symbols from $\{0,1,2,3\}$ of length k_0-2 and all remaining blocks of length k_0 sent to 0. Let ν, ω and z be as in Example 5.1. By genericity

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N} z^{i_0}\cdot z^{i_1}(n+a_1)\cdot\ldots\cdot z^{i_r}(n+a_r)$$
$$= \int_Y \big(F^{i_0}\cdot F^{i_1}\circ S^{a_1}\cdot\ldots\cdot F^{i_r}\circ S^{a_r}\big)(S^n z)\, d\nu.$$

If $a_r < k_0$ then each of the functions $F\circ S^a$ in the above integral take the values 1 and -1 with probability $2/4^2$ and these events (as a varies from 0 to k_0-1) are independent. Therefore, whenever $a_r < k_0$, then the corresponding integral equals zero (when one of the i_s equals 1). However, the function $F\cdot F\circ S^{k_0}$ takes the value 1 with probability $2/4^3$ (given by the blocks $0*1*2$ and $0*2*3$) while the value -1 has probability $1/4^3$ (it is given by $1*2*3$), so the integral is not equal to zero. In other words, (**Ch**) fails for this sequence when $r=1$ and $a_1 = k_0$.

5.2 (Ch) without genericity

Example 5.4 (z that satisfies (**Ch**) with z^2 not generic)**.** Let $w_0 \in Y := \{-1,0,1\}^{\mathbb{N}}$ be a generic point for the Bernoulli measure $\kappa_0 := B(1/3,1/3,1/3)$, and $w_1 \in Y$ a generic point for the Bernoulli measure $\kappa_1 := B(1/2,0,1/2)$. Since the measures κ_0 and κ_1 are mutually singular, up to a set of $(\kappa_0+\kappa_1)$-measure zero, we can represent Y as a union $Y_0\cup Y_1$, $Y_0\cap Y_1 = \varnothing$ with Y_i being a set of full measure for κ_i, $i=0,1$.

Let $\mathbb{Z}\ni a_n \to \infty$ and set

$$M_1 := 1,\ M_{n+1} := a_{n+1}M_n. \tag{48}$$

We define a new sequence $w\in\{-1,0,1\}^{\mathbb{N}}$ by setting

$$w[M_{2i+1}, M_{2i+2}-1] := w_0[0, M_{2i+2}-M_{2i+1}-1],\ i\ge 0,$$

$$w[M_{2i}, M_{2i+1}-1] := w_1[0, M_{2i+1}-M_{2i}-1],\ i\ge 1.$$

Lemma 5.5. *We have Q-gen*$(w) = \{\alpha\kappa_0 + (1-\alpha)\kappa_1 : \alpha \in [0,1]\}$.

Proof. Suppose that, for some sequence (P_i), $\delta_{P_i,w} \to \nu$. Then, for each $i \geq 1$, there exists $s_i \geq 1$, so that $M_{s_i} \leq P_i < M_{s_i+1}$. By considering subsequences, if necessary, we can assume that $M_{s_i}/P_i \to \alpha$ (moreover, for any $\alpha \in [0,1]$ the sequence P_i can be chosen so that this convergence holds). Since $M_{s_i-1}/P_i = \frac{M_{s_i}}{a_{s_i}P_i} \to 0$, the sequence of measures $\frac{1}{P_i}\sum_{n<M_{s_i-1}} \delta_{S^n w}$ converges to 0 when $i \to \infty$. Moreover, since $(M_{s_i} - M_{s_i-1} - 1)/P_i \to \alpha$, the measure $\frac{1}{P_i}\sum_{M_{s_i-1}\leq n<M_{s_i}} \delta_{S^n w}$ is arbitrarily close to $\alpha\kappa_{j_i}$, where κ_{j_i} is either κ_0 or κ_1 depending on the parity of s_i (we pass again to a subsequence if necessary). In a similar way, $\frac{1}{P_i}\sum_{M_{s_i}\leq n<P_i} \delta_{S^n w} \to (1-\alpha)\kappa_{1-j_i}$ and the result follows. □

Clearly, w is not generic, and we can easily check that neither is w^2 (we obtain that w^2 is quasi-generic for all convex combinations of the Dirac measure at $(1,1,\ldots)$ and a Bernoulli measure).

Now, since (**Ch**) holds for w_0 and w_1, the integral of $F^{i_0}\cdot F^{i_1}\circ S^{a_1}\cdot\ldots\cdot F^{i_r}\circ S^{a_r}$ with respect to κ_i for $i = 0,1$ is equal to zero for any choice of $1 \leq a_1 < \ldots < a_r$, $r \geq 0$, $i_s \in \{1,2\}$ not all equal to 2. Therefore, for any such choice we also have

$$\int F^{i_0}\cdot F^{i_1}\circ S^{a_1}\cdot\ldots\cdot F^{i_r}\circ S^{a_r}\ d(\alpha\kappa_0 + (1-\alpha)\kappa_1) = 0,$$

which shows that (**Ch**) holds for w.

5.3 The squares in (Ch) are necessary

We will now show that the squares in (**Ch**) are necessary. In other words, we will show that (**Ch**) is not equivalent to the following condition:

$$\text{(Ch1)} \qquad \frac{1}{N}\sum_{n\leq N} z(n)\cdot z(n+a_1)\cdot\ldots\cdot z(n+a_r) \xrightarrow[N\to\infty]{} 0$$

for each choice of $1 \leq a_1 < \ldots < a_r$, $r \geq 0$. The example will be introduced in the probabilistic language (cf. the discussion on page 15). In order to obtain a sequence $z \in \{-1,0,1\}^{\mathbb{N}}$ satisfying (Ch1) and not satisfying (**Ch**) it suffices to take a generic point for the distribution of the process $(X_n)_{n\in\mathbb{N}}$ considered in the following example.

Example 5.6. Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables, taking values ± 1, each with probability $1/2$. Set

$$X_n := Y_n \mathbf{1}_{Y_{n+1}=1}.$$

Then for each choice of $0 \leq a_1 < \cdots < a_r$, we have

$$\mathbb{E}(X_{a_1}\cdot\ldots\cdot X_{a_r}) = \mathbb{E}(Y_{a_1}\mathbf{1}_{Y_{a_1+1}=1}Y_{a_2}\mathbf{1}_{Y_{a_2+1}=1}\ldots Y_{a_s}\mathbf{1}_{Y_{a_s+1}=1}) = \mathbb{E}(Y_{a_1}Z),$$

where Z is measurable with respect to the σ-algebra generated by $Y_{a_1+1}, Y_{a_1+2},\ldots$, hence is independent from Y_{a_1}. Since $\mathbb{E}(Y_{a_1}) = 0$, we get

$$\mathbb{E}(X_{a_1}\cdot\ldots\cdot X_{a_r}) = 0.$$

However, since $X_1^2\cdot Y_2 = \mathbf{1}_{Y_2=1}$, we have

$$\mathbb{E}(X_1^2 X_2) = \mathbb{E}(\mathbf{1}_{Y_2=1}\mathbf{1}_{Y_3=1}) = \frac{1}{4} \neq 0.$$

Remark 5.7. The dynamical system determined by $(X_n)_{n\in\mathbb{N}}$ is a non-trivial factor of the system determined by $(Y_n)_{n\in\mathbb{N}}$. Moreover, $(Y_n)_{n\in\mathbb{N}}$ is an independent process, so the associated dynamical system is K. Hence $h_{top}(z) > 0$ for any $z \in \{-1,0,1\}^{\mathbb{N}}$ generic for the distribution of $(X_n)_{n\in\mathbb{N}}$.

Question 5.8. Does there exist $z \in \{-1,0,1\}^{\mathbb{N}}$ with $h_{top}(z) = 0$, for which (**Ch**) fails but that satisfies (Ch1)?

Remark 5.9. Note that if $h_{top}(z) = 0$, and if the density of the support of z is positive, then (**Ch**) automatically fails because (**S**) fails.

5.4 (Ch) vs. recurrence

Let A be a nonempty finite set.

Definition 5.10. A sequence $w \in A^{\mathbb{N}}$ is said to be *recurrent* if each block B appearing on w appears on it infinitely often.

Note that, if

$$X_w^+ = X_w, \tag{49}$$

then obviously w is recurrent.

It is well-known (see, e.g. [7], pp. 189-190) that under the recurrence assumption, one can construct the topological natural extension of the one-sided subshift generated by w. More precisely, under the assumption of recurrence of w, there exists $\widetilde{w} \in A^{\mathbb{Z}}$ such that:

- $\widetilde{w}[0,\infty] = w$;
- each block appearing on $\widetilde{w}$ appears on w.

Lemma 5.11. *Let $w \in A^{\mathbb{N}}$ and consider the subshift $X_w \subset A^{\mathbb{N}}$. Then given a block $B \in A^r$ (for some $r \geq 1$), the following two conditions are equivalent:*

- *there exists $\nu \in$ Q-gen(w) such that $\nu(B) > 0$,*
- *B appears on w with positive upper frequency.*

In other words, $X_w^+ = \bigcup_{\nu\in Q\text{-}gen(w)} supp(\nu)$.

Proof. Let (N_k) and $\nu \in$ Q-gen(w) be such that $\delta_{N_k,w} \xrightarrow[k\to\infty]{} \nu$ and let $B \in A^r$, $r \geq 1$, be such that $\nu(B) > 0$. Since $\mathbf{1}_B \in C(A^{\mathbb{N}})$, it follows that

$$\overline{\mathrm{fr}}(B,w) \geq \lim_{k\to\infty} \int \mathbf{1}_B \, d\,\delta_{N_k,w} = \int \mathbf{1}_B \, d\nu = \nu(B) > 0.$$

Suppose now that B appears on w with positive upper frequency, i.e. we have

$$\lim_{k\to\infty} \int \mathbf{1}_B \, d\delta_{N_k,w} > 0, \tag{50}$$

for some sequence (N_k). Passing to a subsequence if necessary, we may assume that $\delta_{N_k,w} \xrightarrow[k\to\infty]{} \nu$ weakly; in particular, $\nu \in$ Q-gen(w). Moreover, by (50), $\nu(B) > 0$. $\square$

Fix $z \in \{-1,0,1\}^{\mathbb{N}}$, and set $X_N := \pi^{-1}(X_{z^2})$, $X_N^+ := \pi^{-1}(X_{z^2}^+)$.

Lemma 5.12. *If* (**Ch**) *holds for* z *then* $X_z^+ = X_N^+$.

Proof. Clearly, $X_z^+ \subset X_N^+$. Take $\widetilde{B} \in X_N^+$. Then $B := \pi(\widetilde{B}) \in X_{z^2}^+$ and by Lemma 5.11 there exists $\nu \in$ Q-gen(z^2) such that $\nu(B) > 0$. Therefore

$$\widehat{\nu}(\widetilde{B}) = \frac{1}{2^{|\mathrm{supp} B|}}\nu(B) > 0.$$

Since z satisfies (**Ch**), it follows from Lemma 4.6 that $\widehat{\nu} \in$ Q-gen(z). Lemma 5.11 implies now that $\widetilde{B} \in X_z^+$ and the assertion follows. □

Proposition 5.13. *Suppose that* (49) *holds for* z^2*, i.e.*

$$X_{z^2}^+ = X_{z^2} \tag{51}$$

and z *satisfies* (**Ch**)*. Then* (49) *holds for* z*; in fact,*

$$\bigcup_{\nu \in Q\text{-}gen(z^2)} supp(\widehat{\nu}) = X_z. \tag{52}$$

In particular, z *is recurrent.*

Proof. By Lemma 5.11 and Remark 4.8 (which can be applied since z satisfies (**Ch**)), we have

$$X_z^+ = \bigcup_{\widetilde{\nu} \in \text{Q-gen}(z)} \mathrm{supp}(\widetilde{\nu}) = \bigcup_{\nu \in \text{Q-gen}(z^2)} \mathrm{supp}(\widehat{\nu}). \tag{53}$$

It follows from (51) that $X_N^+ = X_N$. This and Lemma 5.12 imply

$$X_z \subset X_N = X_N^+ = X_z^+ \subset X_z$$

so (49) holds for z. Therefore and by (53), we conclude that (52) holds. □

Remark 5.14. Note that condition (51) is satisfied if Q-gen$(z^2) = \{\nu\}$ and supp $\nu = X_{z^2}$. This is the case for $z = \boldsymbol{\mu}$; in particular, $\boldsymbol{\mu}^2$ is recurrent. However, it is not known whether $\boldsymbol{\mu}$ is recurrent [21].

It is possible to have z satisfying (**Ch**) and non-recurrent with z^2 being recurrent. Consider the following two examples:

Example 5.15. For $i \geq 1$, let B_i be the block consisting of 10^i zeroes. Then set $A_1 := 1B_1$, $A_2 := A_1A_1B_2$ and in general $A_{s+1} := A_sA_sB_s$ for $s \geq 2$ to obtain in the limit the sequence z^2 which is recurrent. Replace first 1 by -1 without changing other positions to define z. Then z satisfies (**Ch**) and z is not recurrent. A "drawback" of this example is that the density of zeroes is equal to 1.

Example 5.16. We will use that same idea as in Example 5.15. Let (n_i) be an increasing sequence of natural numbers and let B_i be a block of length $2n_i$ of alternating ones and zeroes: $B_i = 1010\ldots10$. Then set $A_1 := 11B_1$, $A_2 := A_1A_1B_2$ and in general $A_{s+1} := A_sA_sB_s$. In the limit, we obtain an infinite sequence w. If (n_i) increases fast enough then w differs from $w' :=$

$(1,0,1,0,\ldots)$ on a set of density zero, whence w is generic for $\nu := \frac{1}{2}(\delta_{(1,0,1,0,\ldots)} + \delta_{(0,1,0,1,\ldots)})$. Let $u \in \{-1,1\}^{\mathbb{N}}$ be generic for $B(1/2,1/2)$ and let $z' := \xi(w',u)$, where $\xi(a,b)(n) := a(n)\cdot b(n)$ for $n \in \mathbb{N}$. Since w' is of zero entropy, (w',u) is generic for $\nu \otimes B(1/2,1/2)$. Hence z' is generic for $\xi_*(\nu \otimes B(1/2,1/2)) = \widehat{\nu}$. The sequence z' is a concatenation of blocks of length 3 of ± 1 separated by long blocks of $-1,0,1$ in which every second position is 0. We now modify z' to obtain z in the following way. The first block of 3 consecutive ± 1 (i.e. $z'(1)z'(2)z'(3)$) is replaced with $(-1,-1,-1)$ while all other 3-blocks of consecutive ± 1 are replaced with 111. Then z differs from z' on a subset of density zero, so z is still a generic point for $\widehat{\nu}$, i.e. (**Ch**) holds for z. Clearly, z is not recurrent, whereas $w = z^2$ has this property. Moreover, the density of 0's on z is equal to $1/2$.

5.5 (Ch) vs. unique ergodicity

Proposition 5.17. *Let $z \in \{-1,0,1\}^{\mathbb{N}}$ be such that (**Ch**) holds. Moreover, suppose that there exists a block B with*

$$\text{supp}(B) \neq \varnothing \text{ and } B \text{ appears on } z \text{ with positive upper frequency.} \tag{54}$$

Then the subshift X_z cannot be uniquely ergodic.

Proof. It suffices to show that the subshift X_z^+ is not uniquely ergodic. By Lemma 5.12, we have $X_z^+ = X_N^+$, whence $X_{z^2}^+ \subset X_z^+$. Therefore

$$\varnothing \neq \text{Q-gen}(z^2) \subset \mathcal{P}_S(X_z^+). \tag{55}$$

By Remark 4.8, we have $\text{Q-gen}(z) = \{\widehat{\nu} : \nu \in \text{Q-gen}(z^2)\}$. Let B be the block with non-empty support given by (54). Then it follows from Lemma 5.11 that there exists $\nu \in \text{Q-gen}(z^2)$ such that $\widehat{\nu}(B) > 0$. Moreover, for any block C with $B^2 = C^2$ we have $\widehat{\nu}(C) = \widehat{\nu}(B) > 0$. It follows that $\nu \neq \widehat{\nu}$, but $\{\nu, \widehat{\nu}\} \subset \mathcal{P}_S(X_z^+)$. □

For $z = \boldsymbol{\mu}$ or $z = \boldsymbol{\mu}_{\mathscr{B}}$ the fact that the subshift X_z is not uniquely ergodic "comes" from X_{z^2}. To see this, we need first to recall the following definition [15]:

Definition 5.18. A subshift $X \subset \{0,\ldots,k\}^{\mathbb{N}}$ is *hereditary* if for any $x \in X$ and $y \in \{0,\ldots,k\}^{\mathbb{N}}$ the condition $y(n) \leq x(n)$ satisfied for all n, implies that $y \in X$.

Remark 5.19. In view of [21] and [2], for $z = \boldsymbol{\mu}$ or $z = \boldsymbol{\mu}_{\mathscr{B}}$, the subshift X_{z^2} consists of all sequences $w \in \{0,1\}^{\mathbb{Z}}$ which are $\mathscr{B}$-admissible, i.e. such that

$$t(\text{supp}(w), b) < b \text{ for all } b \in \mathscr{B},$$

where for $A \subset \mathbb{Z}$ and $b \geq 1$, $t(A,b) := |\{c \in \mathbb{Z}/b\mathbb{Z} : \exists n \in A,\ n = c \bmod b\}|$ is the number of classes modulo b in A.

It follows immediately from the above remark that the subshift $X_{\boldsymbol{\mu}_{\mathscr{B}}^2}$ is hereditary. Now, each hereditary system of positive topological entropy (and such are $(S, X_{\boldsymbol{\mu}_{\mathscr{B}}^2})$ [2]) is not uniquely ergodic, e.g. [16].[14]

[14] A direct proof of non-unique ergodicity of $(S, X_{\boldsymbol{\mu}_{\mathscr{B}}^2})$ follows from the fact that each hereditary system has a fixed point and moreover, $X_{\boldsymbol{\mu}_{\mathscr{B}}^2}$ supports the Mirsky measure.

Remark 5.20. We can choose a generic point $z \in \{-1,1\}^{\mathbb{N}}$ for the Bernoulli measure $B(1/2,1/2)$ to obtain an example of z satisfying (**Ch**) and for which (S, X_z) is not uniquely ergodic while (S, X_{z^2}) has this property.

6 Entropy of sequences satisfying (Ch)

We will now concentrate on the following question:

> Which pairs of real numbers can be realized as $(h_{top}(z^2), h_{top}(z))$ for $z \in \{-1,0,1\}^{\mathbb{N}}$, under the additional assumption that z satisfies (**Ch**)?

6.1 Bounds for entropy

The authors would like to thank Benjamin Weiss for fruitful discussions which resulted in the material presented in this section.

First, we will address the question whether there are some natural restrictions for the value of the pair $(h_{top}(z^2), h_{top}(z))$ for sequences $z \in \{-1,0,1\}^{\mathbb{N}}$. Then we will impose the extra assumption that z satisfies (**Ch**). We begin by the following simple observation:

Lemma 6.1. *Let $z \in \{-1,0,1\}^{\mathbb{N}}$. Then*

$$h_{top}(z^2) \leq h_{top}(z) \leq \min(h_{top}(z^2) + 1, \log 3).$$

Proof. Notice that the maximal possible number of 1's in a block of length n is clearly equal to n and for each occurrence of 1 in such a block of z^2 there are at most two possible choices for this position in z: either -1 or 1. Therefore

$$p_z(n) \leq 2^n \cdot p_{z^2}(n).$$

It follows immediately that

$$\begin{aligned} h_{top}(z) = \lim_{n\to\infty} \frac{1}{n} \log p_z(n) &\leq \lim_{n\to\infty} \frac{1}{n} \log(2^n \cdot p_{z^2}(n)) \\ &= 1 + \lim_{n\to\infty} \frac{1}{n} \log p_{z^2}(n) = 1 + h_{top}(z^2). \end{aligned}$$

which completes the proof as clearly $h_{top}(z^2) \leq h_{top}(z) \leq \log 3$. □

6.1.1 Measure-theoretical setting

We will work first in the measure-theoretical setting, i.e. we will study $(h(\nu), h(\widehat{\nu}))$ for $\nu \in \mathcal{P}_S(\{0,1\}^{\mathbb{Z}})$. Recall from Lemma 4.25 the definition of H:

$$H(x) = -x \cdot \log x - (1-x) \cdot \log(1-x) \text{ for } x \in (0,1),$$

and consider

$$H_1 := H|_{(0,1/2]}, \; H_2 := H|_{[1/2,1)}.$$

Lemma 6.2. *Let $d \in [0,1]$, $n \geq 1$ and $\varepsilon > 0$. Let*

$$\mathcal{A}_{n,d,\varepsilon} := \left\{ C \in \{0,1\}^n : \left| \frac{|supp(C)|}{n} - d \right| < \varepsilon \right\}.$$

Then,

$$|\mathcal{A}_{n,d,\varepsilon}| \leq 2^{n \sup\{H(d') : d-\varepsilon \leq d' \leq d+\varepsilon\}}. \tag{56}$$

Moreover, if $\nu \in \mathcal{P}_S^e(\{0,1\}^{\mathbb{Z}})$ with $\nu([1]) = d$, then for n sufficiently large there exists $\mathcal{C}_n \subset \mathcal{A}_{n,d,\varepsilon}$ such that

$$\nu\left(\bigcup_{C \in \mathcal{C}_n} [C] \right) > 1 - \varepsilon,$$

and

$$|\mathcal{C}_n| \geq 2^{n(h(\nu)-\varepsilon)}.$$

Proof. If $d-\varepsilon \leq 1/2 \leq d+\varepsilon$, (56) is obvious because $H(1/2) = 1$. If $d+\varepsilon < 1/2$, this is a consequence of Lemma 4.25. Otherwise, we have $1 - d + \varepsilon < 1/2$, so we can apply Lemma 4.25 to $\mathcal{A}_{n,1-d,\varepsilon}$, and observe that H is symmetric, and that $|\mathcal{A}_{n,d,\varepsilon}| = |\mathcal{A}_{n,1-d,\varepsilon}|$.

Now, let $\nu \in \mathcal{P}_S^e(\{0,1\}^{\mathbb{Z}})$ with $\nu([1]) = d$. Set $h := h(\nu)$. It follows from the Shannon-McMillan theorem and from the mean ergodic theorem that for n sufficiently large there exists $\mathcal{C}_n \subset \mathcal{A}_{n,d,\varepsilon}$ such that

$$\nu\left(\bigcup_{C \in \mathcal{C}_n} [C] \right) > 1 - \varepsilon/2,$$

and, for all $C \in \mathcal{C}_n$,

$$2^{-(h+\varepsilon)n} < \nu([C]) < 2^{-(h-\varepsilon)n}.$$

It follows that

$$1 - \varepsilon/2 < \sum_{C \in \mathcal{C}_n} \nu([C]) < |\mathcal{C}_n| \cdot 2^{-(h-\varepsilon)n},$$

whence

$$|\mathcal{C}_n| > (1 - \varepsilon/2) 2^{n(h-\varepsilon)} \geq 2^{n(h-\varepsilon)}. \tag{57}$$

□

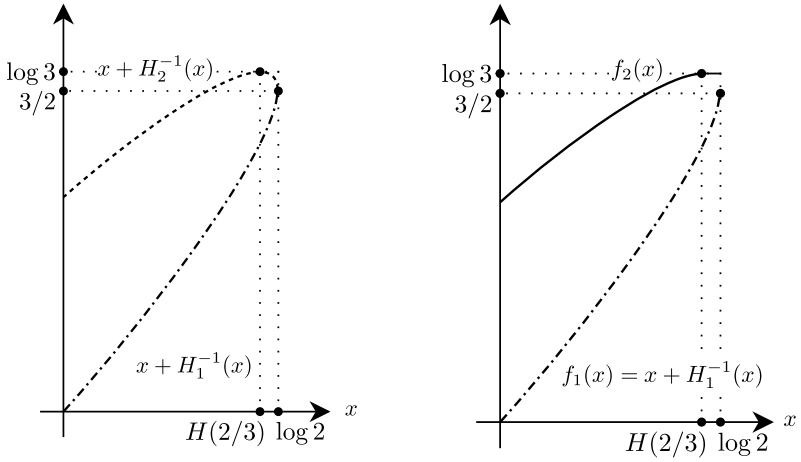
We will now use functions H_1 and H_2 to obtain bounds for the measure-theoretical entropy. The lower bound will be given by $x \mapsto x + H_1^{-1}(x)$ and the upper bound by $x \mapsto x + H_2^{-1}(x)$, see Figure 1a. More precisely, we have the following:

Proposition 6.3. *Let ν be in $\mathcal{P}_S^e(\{0,1\}^{\mathbb{Z}})$, and set $d := \nu([1])$. Then*

$$h(\widehat{\nu}) = h(\nu) + d, \text{ and } d \in [H_1^{-1}(h(\nu)), H_2^{-1}(h(\nu))].$$

Moreover, the following conditions are equivalent:

(a) $d = H_i^{-1}(h(\nu))$ for $i = 1$ or $i = 2$,



(a) Functions $x + H_1^{-1}(x)$ and $x + H_2^{-1}(x)$.

(b) Functions f_1 and f_2.

Figure 1: Graphs of the functions used in the bounds for measure-theoretic entropy (left) and topological entropy (right).

(b) $\nu = B(1-d, d)$.

Proof. Let $\varepsilon > 0$, set $h := h(\nu)$ and $d := \nu([1])$. Let $\mathcal{A}_{n,d,\varepsilon}$ be as in Lemma 6.2. Then

$$2^{(h-\varepsilon)n} \leq 2^{n \sup\{H(d'): d-\varepsilon \leq d' \leq d+\varepsilon\}}$$

for all n sufficiently large. Therefore, by the continuity of H, we obtain $h \leq H(d)$, which implies that

$$d \in [H_1^{-1}(h), H_2^{-1}(h)]. \tag{58}$$

We are now ready to estimate $h(\widehat{\nu})$. We have

$$\begin{aligned}
&\frac{1}{n} \sum_{\substack{B \in \{-1,0,1\}^n \\ \widehat{\nu}(B) > 0}} -\widehat{\nu}(B) \log \widehat{\nu}(B) \\
&\qquad = \frac{1}{n} \sum_{C \in \mathcal{A}_{n,d,\varepsilon}} \sum_{\substack{B \in \{-1,0,1\}^n \\ B^2 = C}} -\frac{\nu(C)}{2^{|\mathrm{supp}(C)|}} \cdot \log \frac{\nu(C)}{2^{|\mathrm{supp}(C)|}} \\
&\qquad\qquad + \frac{1}{n} \sum_{\substack{B \in \{-1,0,1\}^n \\ B^2 \in \mathcal{A}_{n,d,\varepsilon}^c \\ \nu(B^2) > 0}} -\widehat{\nu}(B) \log \widehat{\nu}(B).
\end{aligned}$$

Let $\delta := \widehat{\nu}\Big(\bigcup_{\substack{B\in\{-1,0,1\}^n\\ B^2\in\mathcal{A}^c_{n,d,\varepsilon}}} B\Big) = \nu\Big(\bigcup_{C\in\mathcal{A}^c_{n,d,\varepsilon}} C\Big) < \varepsilon$. Then

$$\sum_{\substack{B\in\{-1,0,1\}^n\\ B^2\in\mathcal{A}^c_{n,d,\varepsilon}\\ \nu(B^2)>0}} (-\widehat{\nu}(B))\log\widehat{\nu}(B) = \delta \sum_{\substack{B\in\{-1,0,1\}^n\\ B^2\in\mathcal{A}^c_{n,d,\varepsilon}\\ \nu(B^2)>0}} -\frac{\widehat{\nu}(B)}{\delta}\cdot\log\frac{\widehat{\nu}(B)}{\delta} - \delta\log\delta$$
$$\leq \delta n\log 3 - \delta\log\delta.$$

It follows that

$$h(\widehat{\nu}) \leq \lim_{n\to\infty}\frac{1}{n}\sum_{C\in\mathcal{A}_{n,d,\varepsilon}}\sum_{\substack{B\in\{-1,0,1\}^n\\ B^2=C}} -\frac{\nu(C)}{2^{|\mathrm{supp}(C)|}}\cdot\log\frac{\nu(C)}{2^{|\mathrm{supp}(C)|}} + \delta\log 3$$
$$\leq \lim_{n\to\infty}\left(\frac{1}{n}\sum_{C\in\mathcal{A}_{n,d,\varepsilon}}(-\nu(C)\log\nu(C)) + \frac{1}{n}\sum_{C\in\mathcal{A}_{n,d,\varepsilon}}\nu(C)|\mathrm{supp}(C)|\right) + \varepsilon\log 3$$
$$\leq h(\nu) + d + \varepsilon + \varepsilon\log 3.$$

Since ε is arbitrarily small, and remembering (58), we obtain

$$h(\widehat{\nu}) \leq h(\nu) + d \leq h(\nu) + H_2^{-1}(h(\nu)).$$

On the other hand, by similar arguments,

$$h(\widehat{\nu}) \geq h(\nu) + d \geq h(\nu) + H_1^{-1}(h(\nu)),$$

and the first part of the assertion follows.

For the second part of the assertion, consider the partition $\{[0],[1]\}$. This is a generating partition and it follows that $h(\nu) \leq H(d)$. Notice that the inequality is sharp, unless $\nu = B(1-d,d)$. In other words, (a) does not hold, unless (b) holds. Moreover, it is immediate that (b) implies (a). □

6.1.2 Topological setting

Using the results from the preceding section, we will pass now to the topological setting, i.e. we will study $(h_{top}(z^2), h_{top}(z))$. Recall from Proposition 6.3 that for any $\nu \in \mathcal{P}_S(\{0,1\}^{\mathbb{Z}})$, we have the inequalities

$$h(\nu) + H_1^{-1}(h(\nu)) \leq h(\widehat{\nu}) \leq h(\nu) + H_2^{-1}(h(\nu)).$$

We will provide similar bounds regarding the topological entropy. Let $f_1, f_2\colon [0,1] \to \mathbb{R}$ be given by

$$f_1(x) := x + H_1^{-1}(x),$$
$$f_2(x) := \begin{cases} x + H_2^{-1}(x), & \text{for } x < H(2/3),\\ \log 3, & \text{for } x \geq H(2/3).\end{cases}$$

The lower bound will be the same as in the measure-theoretical setting. The upper bound will be now given by f_2, see Figure 1b.

Remark 6.4. Elementary calculation shows that $x \mapsto x + H_2^{-1}(x)$ is an increasing function on $(0, H(2/3)]$.

Proposition 6.5. *Let $z \in \{-1,0,1\}^{\mathbb{N}}$. Then*

$$h_{top}(z) \le f_2(h_{top}(z^2)).$$

Moreover, $h_{top}(z) = f_2(h_{top}(z^2))$ if and only if $h_{top}(z) = \log 3$.

If additionally

$$X_z = \pi^{-1}(X_{z^2}), \tag{59}$$

then

$$f_1(h_{top}(z^2)) < h_{top}(z).$$

Lemma 6.6. *Let $\nu \in \mathcal{P}_S(\{0,1\}^{\mathbb{Z}})$. Then*

$$h(\widehat{\nu}) = \max\{h(\rho) : \rho \in \mathcal{P}_S(\{-1,0,1\}^{\mathbb{Z}}), \pi_*(\rho) = \nu\}.$$

Proof. For any $\rho \in \mathcal{P}_S(\{-1,0,1\}^{\mathbb{Z}})$ satisfying $\pi_*(\rho) = \nu$, we have

$$\begin{aligned} -h(\rho) &= \lim_{n\to\infty} \frac{1}{n} \sum_{B\in\{-1,0,1\}^n} \rho(B)\log\rho(B) = \lim_{n\to\infty} \frac{1}{n} \sum_{C\in\{0,1\}^n} \sum_{B^2=C} \rho(B)\log\rho(B) \\ &\ge \lim_{n\to\infty} \frac{1}{n} \sum_{C\in\{0,1\}^n} \sum_{B^2=C} \widehat{\nu}(B)\log\widehat{\nu}(B) = -h(\widehat{\nu}) \end{aligned}$$

(the inequality follows from the fact that $\sum_{i=1}^n a_i \log(a_i) \ge \sum_{i=1}^n \frac{a}{n}\log\frac{a}{n}$, where $a = \sum_{i=1}^n a_i$). □

Remark 6.7. Suppose that $z \in \{-1,0,1\}^{\mathbb{N}}$ satisfies (59), that is, for each B appearing on z^2 and each C such that $C^2 = B$, the block C appears on z. Then clearly

$$\nu \in \mathcal{P}_S(X_{z^2}) \iff \widehat{\nu} \in \mathcal{P}_S(X_z).$$

It follows that if $\rho \in \mathcal{P}_S(X_z)$ is such that $h_{top}(z) = h(\rho)$ then, by Lemma 6.6,

$$h_{top}(z) = h(\rho) \le h(\widehat{\pi_*(\rho)}) \le h_{top}(z),$$

i.e. $h_{top}(z) = h(\widehat{\pi_*(\rho)})$.

Lemma 6.8. *Let $z \in \{-1,0,1\}^{\mathbb{N}}$. Then*

$$h_{top}(z) \le \sup\{h(\widehat{\nu}) : \nu \in \mathcal{P}_S(X_{z^2})\}.$$

Moreover, the supremum is achieved by some measure in $\mathcal{P}_S(X_{z^2})$.

Proof. Let $X := \pi^{-1}(X_{z^2})$. Let $\rho \in \mathcal{P}_S(X)$ be such that $h(\rho) = h_{top}(X)$. Then by Lemma 6.6

$$h_{top}(z) \le h_{top}(X) = h(\rho) \le h(\widehat{\pi(\rho)}) \le h_{top}(X).$$

But $\pi(\rho) \in \mathcal{P}_S(X_{z^2})$, hence

$$h(\widehat{\pi(\rho)}) \le \sup\{h(\widehat{\nu}) : \nu \in \mathcal{P}_S(X_{z^2})\}.$$

On the other hand, if $\nu \in \mathcal{P}_S(X_{z^2})$, $\widehat{\nu} \in \mathcal{P}_S(X)$, and $h(\widehat{\nu}) \le h(\rho) = h_{top}(X)$. It follows that the supremum in the statement of the lemma is equal to $h(\widehat{\pi(\rho)})$, and the result is proved. □

Proof of Proposition 6.5. By Lemma 6.8,

$$h_{top}(z) \leq \max\{h(\widehat{\nu}) : \nu \in \mathcal{P}_S(X_{z^2})\}.$$

Let $\kappa \in \mathcal{P}_S(X_{z^2})$ be such that

$$h(\widehat{\kappa}) = \max\{h(\widehat{\nu}) : \nu \in \mathcal{P}_S(X_{z^2})\}.$$

By Proposition 6.3,

$$h(\widehat{\kappa}) \leq f_2(h(\kappa)) \leq \sup\{f_2(h(\nu)) : \nu \in \mathcal{P}_S(X_{z^2})\}.$$

By Remark 6.4, the latter expression is equal to $f_2(h_{top}(z^2))$, whenever $h_{top}(z^2) < H(2/3)$. But if $h_{top}(z^2) \geq H(2/3)$, then $f_2(h_{top}(z^2)) = \log 3 \geq f_2(h(\nu))$ for any $\nu \in \mathcal{P}_S(X_{z^2} h_{top}(z)$. Thus, we have obtained

$$\sup\{f_2(h(\nu)) : \nu \in \mathcal{P}_S(X_{z^2})\}(z) \leq f_2(h_{top}(z^2)).$$

If all of the above inequalities are equalities, then in particular

$$h(\widehat{\kappa}) = f_2(h(\kappa)),$$

which by Proposition 6.3 happens only if κ is Bernoulli. This implies

$$h_{top}(z^2) = \log 2 = 1.$$

But then

$$h_{top}(z) = f_2(h_{top}(z^2)) = f_2(1) = \log 3.$$

On the other hand, if $h_{top}(z) = \log 3$, then $h_{top}(z^2) = \log 2$ and we obtain

$$h_{top}(z) = \log 3 = f_2(h_{top}(z^2)).$$

Suppose now that (59) holds and let $\nu \in \mathcal{P}_S(X_{z^2})$ be such that $h_{top}(z^2) = h(\nu)$. Then, by Proposition 6.3,

$$h_{top}(z) \geq h(\widehat{\nu}) \overset{(*)}{\geq} f_1(h(\nu)) = f_1(h_{top}(z^2)).$$

We claim that the second inequality is sharp. Indeed, assume it is an equality. Then by Proposition 6.3, ν is Bernoulli. This, together with (59) gives $h_{top}(z) = \log 3$. Then $h_{top}(z^2) = \log 2$. But for these values, we have

$$\log 3 > \frac{3}{2} = f_1(\log 2).$$

□

We will show now that the bounds given by the above proposition are the best possible.

Proposition 6.9. *For any $h_{z^2} \in (0, \log 2]$ and $\varepsilon > 0$ there exists $z \in \{-1, 0, 1\}^{\mathbb{N}}$ such that*

- $h_{top}(z^2) \in [h_{z^2} - \varepsilon, h_{z^2} + \varepsilon]$,
- $h_{top}(z) \geq f_2(h_{z^2}) - \varepsilon$.

Moreover, there exists $z \in \{-1,0,1\}^{\mathbb{N}}$ satisfying condition (59) and such that

- $h_{top}(z^2) \in [h_{z^2} - \varepsilon, h_{z^2} + \varepsilon]$,
- $h_{top}(z) \leq f_1(h_{z^2}) + \varepsilon$.

Proof. The proof of both parts of the assertion goes along the same lines and we will provide the details only for the first part.

Let $d \geq 1/2$ be such that $H(d) = h_{z^2}$. Fix $\varepsilon > 0$ and for $n \in \mathbb{N}$ let $\mathcal{A}_{n,d,\varepsilon}$ be as in Lemma 6.2. Then,

$$|\mathcal{A}_{N,d,\varepsilon}| \leq 2^{N \sup\{(H(d'):d-\varepsilon \leq d' \leq d+\varepsilon\}}.$$

Applying the second part of Lemma 6.2 with $\nu := B(1-d,d)$, for N sufficiently large, we get

$$2^{(H(d)-\varepsilon)N} \leq |\mathcal{A}_{N,d,\varepsilon}|.$$

Fix such N and let $X \subset \{0,1\}^{\mathbb{N}}$ be the subshift consisting of these points x, for which any block appearing on x is a subword of a concatenation of some words form $A_{N,d,\varepsilon}$. Let

$$\mathcal{C}_{nN} := \{B \in X : |B| = nN\}.$$

Then

$$|\mathcal{A}_{N,d,\varepsilon}|^n \leq |\mathcal{C}_{nN}| \leq N \cdot |A_{N,d,\varepsilon}|^{n+1},$$

whence

$$H(d) - \varepsilon \leq h_{top}(X) \leq \sup\{(H(d') : d - \varepsilon \leq d' \leq d + \varepsilon\}.$$

Moreover, if n is sufficiently large, then, for $B \in \mathcal{C}_{nN}$ we have

$$\left| \frac{|\mathrm{supp}(B)|}{nN} - d \right| < 2\varepsilon.$$

For $n \in \mathbb{N}$, let

$$\mathcal{D}_{nN} := \{C \in \{-1,0,1\}^{nN} : C^2 \in \mathcal{C}_{nN}\}.$$

It follows that

$$|\mathcal{D}_{nN}| \geq 2^{(d-2\varepsilon)nN}|\mathcal{C}_{nN}| \geq 2^{(d-2\varepsilon)nN}|\mathcal{A}_{N,d,\varepsilon}|^n \geq 2^{(d-2\varepsilon)nN} \cdot 2^{(H(d)-\varepsilon)nN}.$$

Then

$$H(d) + d - \varepsilon - 2\varepsilon \leq h_{top}(\pi^{-1}(X)).$$

By the choice of d, $d = H_2^{-1}(h_{z^2})$. Then, by continuity of H, to complete the proof, it suffices to pick $z \in \pi^{-1}(X)$ such that any block appearing on $\pi^{-1}(X)$ also appears on z. □

6.1.3 Topological setting with (Ch) satisfied

Proposition 6.10. *Let $z, w \in \{-1,0,1\}^{\mathbb{N}^*}$. Then there exists $\overline{z} \in \{-1,0,1\}^{\mathbb{N}^*}$ such that:*

- $\lim_{k\to\infty} \frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n \overline{z}} = \lim_{k\to\infty} \frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n z}$ *for each subsequence (N_k) such that one of these limits exists,*

- $h_{top}(\overline{z}) = \max(h_{top}(z), h_{top}(w))$,
- $h_{top}(\overline{z}^2) = \max(h_{top}(z^2), h_{top}(w^2))$.

Proof. For a sequence x over a finite alphabet, we set

$$\mathcal{C}_n(x) := \{B : |B| = n, B \text{ appears on } x\}, \text{ so that } p_n(x) = |\mathcal{C}_n(x)| \text{ for } n \in \mathbb{N}.$$

The sequence $\overline{z}$ will be defined as a limit of sequences $\overline{z}_k$, which will be constructed inductively. Fix $\varepsilon_k \to 0$. Let $\overline{z}_1 := z$, and choose d_1 large enough so that $1/d_1 < \varepsilon_1$. Suppose that $d_1, \ldots, d_k$ and $\overline{z_1}, \ldots, \overline{z}_k$ are already chosen. Let d_{k+1} be large enough, so that

$$\frac{\min\{i : \overline{z}_k(j) = z(j) \text{ for } j \geq i\}}{d_{k+1}} < \varepsilon_{k+1}.$$

Let $B_{k+1} \in \{-1, 0, 1\}^{3d_{k+1}}$ be a block which appears on $\overline{z}_k$ infinitely many times. We define $\overline{z}_{k+1}$ by replacing some of the occurrences of B_{k+1} in $\overline{z}_k$ by blocks of the form

$$\underbrace{0\ldots0}_{d_{k+1}} B \underbrace{0\ldots0}_{d_{k+1}}, \text{ where } B \in \mathcal{C}_{d_{k+1}}(w)$$

in such a way that

- $\overline{z}_{k+1}[1, d_{k+1}] = \overline{z}_k[1, d_{k+1}]$,
- $\mathcal{C}_{d_{k+1}}(w) \cup \mathcal{C}_{d_{k+1}}(z) \subset \mathcal{C}_{d_{k+1}}(\overline{z}_{k+1})$.

It follows immediately that

$$h_{top}(\overline{z}) \geq \max(h_{top}(z), h_{top}(w)).$$

On the other hand,

$$p_{d_{k+1}}(\overline{z}) \leq \varepsilon_{k+1} d_{k+1} 3^{\varepsilon_{k+1} d_{k+1}} p_{d_{k+1}}(z) d_{k+1} + p_{d_{k+1}}(z) 2d_{k+1} + p_{d_{k+1}}(w) 2d_{k+1},$$

whence

$$\begin{aligned} h_{top}(\overline{z}) &\leq \max\left(\lim_{k\to\infty} \frac{1}{d_k} \log p_{d_k}(z), \lim_{k\to\infty} \frac{1}{d_k} \log p_{d_k}(w)\right) \\ &= \max(h_{top}(z), h_{top}(w)). \end{aligned}$$

Therefore

$$h_{top}(\overline{z}) = \max(h_{top}(z), h_{top}(w)).$$

In a similar way, we conclude that

$$h_{top}(\overline{z}^2) = \max(h_{top}(z^2), h_{top}(w^2)).$$

Moreover, if the replacement of blocks made in course of the construction is scarce enough, the resulting sequence $\overline{z}$ will be such that

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n \overline{z}} = \lim_{k\to\infty} \frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n z},$$

for any subsequence (N_k) such that one of the above limits exists. This completes the proof. $\square$

Applying the above proposition to $z = (0, 0, \dots)$, we obtain immediately the following corollary.

Corollary 6.11. *For any* $(h_{z^2}, h_z) \in [0, \log 2] \times [0, \log 3]$, *such that for some* $w \in \{-1, 0, 1\}^{\mathbb{N}}$, *we have*

$$(h_{z^2}, h_z) = (h_{top}(w^2), h_{top}(w)),$$

there exists $\overline{z}$ *such that* (**Ch**) *holds for* $\overline{z}$ *and*

$$(h_{z^2}, h_z) = (h_{top}(\overline{z}^2), h_{top}(\overline{z})).$$

6.2 Two classes of sequences

Notice that the result obtained in Corollary 6.11 is in some way "artificial". The sequence $\overline{z}$ satisfying (**Ch**) obtained this way has "a lot" of zeros (it is generic for the Dirac measure at the all zero sequence). Now, we will be interested in more natural examples of sequences z satisfying (**Ch**), such that z additionally satisfies (59) and

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\le N} z^2(n) > 0. \tag{60}$$

We will provide two classes of concrete examples, giving a positive answer to the following problem (cf. Corollary 6.5):

> Given $(h_{z^2}, h_z) \in [0, \log 2] \times [0, \log 3]$ such that $f_1(h_{z^2}) < h_z < f_2(h_{z^2})$,
> does there exist $z \in \{-1, 0, 1\}^{\mathbb{N}}$ satisfying (**Ch**), (60), (59)
> and such that $(h_{top}(z^2), h_{top}(z)) = (h_{z^2}, h_z)$?

in the following special cases:

- $(h_{z^2}, h_z) = (0, \delta)$, $\delta \in [0, 1]$,
- $(h_{z^2}, h_z) = (\delta, \delta \log 3)$, $\delta \in [0, 1]$.

6.2.1 Sequences satisfying (Ch)

Let $u \in \{-1, 1\}^{\mathbb{N}}$ be a generic point for the Bernoulli measure $B(1/2, 1/2)$.

Proposition 6.12. *If* $\eta \in \{0, 1\}^{\mathbb{N}}$ *is completely deterministic then* (**Ch**) *holds for* $z := \eta \cdot u$.

Proof. Let (N_k) be a subsequence such that

$$\delta_{S\times S, N_k, (\eta, u)} \xrightarrow[k\to\infty]{} \widetilde{\rho}.$$

Then

$$\delta_{S, N_k, \eta} \xrightarrow[k\to\infty]{} \rho,$$

where ρ is the projection of $\widetilde{\rho}$ onto the first coordinate and, by the assumption on η, $h(S, \rho) = 0$. Moreover, since u is generic for $B(1/2, 1/2)$, the measure $\widetilde{\rho}$

is a joining of (S,ρ) and $(S,B(1/2,1/2))$. Since $(S,\rho)\perp(S,B(1/2,1/2))$, this must be the product joining, i.e. $\widetilde{\rho}=\rho\otimes B(1/2,1/2)$.

It follows that $\eta\cdot u$ is quasi-generic along (N_k) for the image of $\rho\otimes B(1/2,1/2)$ via the map

$$m\colon \{0,1\}^{\mathbb{N}}\times\{-1,1\}^{\mathbb{N}}\to\{-1,0,1\}^{\mathbb{N}}, \tag{61}$$

given by $m(a,b)(n):=a(n)\cdot b(n)$. Clearly, $m_*(\rho\otimes B(1/2,1/2))=\widehat{\rho}$. The assertion follows from Lemma 4.6. □

Remark 6.13. Recall that $\boldsymbol{\mu}^2$ yields a system with discrete spectrum [17], so in particular $\boldsymbol{\mu}^2$ is completely deterministic.

Corollary 6.14. *Suppose that* (**Ch**) *holds for the Liouville function* $\boldsymbol{\lambda}$. *Then* (**Ch**) *holds for* $\boldsymbol{\mu}$.

Proof. The proof follows directly from the fact that $\boldsymbol{\mu}(n)=\boldsymbol{\lambda}(n)\cdot\boldsymbol{\mu}^2(n)$, Remark 6.13 and Proposition 6.12. □

Example 6.15. We will show that the converse of Proposition 6.12 is not true, i.e. there exists $\eta\in\{0,1\}^{\mathbb{N}}$ which is completely deterministic and $u\in\{-1,1\}^{\mathbb{N}}$ such that (**Ch**) holds for $z:=\eta\cdot u$ and (**Ch**) does not hold for u. Let $\eta=(0,1,0,1,\dots)$ and let $w\in\{-1,1\}^{\mathbb{N}}$ be a generic point for $B(1/2,1/2)$. Define

$$u(n):=\begin{cases} w(n) & \text{if } \eta(n)=1,\\ 1 & \text{if } \eta(n)=0.\end{cases}$$

Then $z:=\eta\cdot u=\eta\cdot w$, whence z satisfies (**Ch**) by Proposition 6.12. It remains to show that u does not satisfy (**Ch**). Indeed, we have $u=(1-\eta)+\eta\cdot w$, whence

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}u(n)=1-\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\eta(n)+\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\eta(n)\cdot w(n)=1/2,$$

in particular, (**Ch**) fails for u.

Proposition 6.12 turns out to be a particular case of the following result.

Proposition 6.16. *Suppose that* (**Ch**) *holds for* $u\in\{-1,0,1\}^{\mathbb{N}}$ *and* $\eta\in\{-1,0,1\}^{\mathbb{N}}$ *is completely deterministic. Then* (**Ch**) *holds for* $z:=\eta\cdot u$.

Proof. For each $1\le a_1<\ldots<a_k$, $k\ge 0$ and $i_s\in\{1,2\}$, $1\le s\le k$, we set

$$w(n)=\eta^{i_0}(n)\cdot\eta^{i_1}(n+a_1)\cdot\ldots\cdot\eta^{i_k}(n+a_k).$$

Then w is completely deterministic by Lemma 4.20 (b), and we obtain

$$\begin{aligned}\frac{1}{N}\sum_{n\le N}&z^{i_0}(n)\cdot z^{i_1}(n+a_1)\cdot\ldots\cdot z^{i_k}(n+a_k)\\ &=\frac{1}{N}\sum_{n\le N}w(n)u^{i_0}(n)\cdot u^{i_1}(n+a_1)\cdot\ldots\cdot u^{i_k}(n+a_k)\\ &=\frac{1}{N}\sum_{n\le N}F(S^nw)u^{i_0}(n)\cdot u^{i_1}(n+a_1)\cdot\ldots\cdot u^{i_k}(n+a_k)\xrightarrow[N\to\infty]{}0\end{aligned}$$

by Theorem 4.19. □

Remark 6.17. In Proposition 6.16, the condition (**Ch**) can be replaced by (**S**). The proof of the modified proposition goes along the same lines.

6.2.2 On Sturmian sequences

In this section, we will describe some properties of blocks appearing on Sturmian sequences.

Lemma 6.18. *For any $\delta \in [0,1]$ there exists a Sturmian sequence $\eta \in \{0,1\}^{\mathbb{N}}$ such that for n large enough*

$$\delta n - 3 < \#1(B_n) < \delta n + 3 \text{ for any } B_n \in \{0,1\}^n \text{ appearing in } \eta,$$

where $\#1(B_n) = |\{0 \le k \le n-1 : B_n(k) = 1\}|$. Moreover, for any Sturmian sequence η there exists $\delta \in [0,1]$ such that the above inequalities hold.

Proof. For $\delta \in \{0,1\}$ the proof is immediate, as it suffices to consider the sequences $(0,0,0,\dots)$ and $(1,1,1,\dots)$, respectively. Therefore, we may assume that $\delta \in (0,1)$.

Consider the integer grid and a line L in the plane[15] and build a sequence η by writing down 0 each time L intersects a horizontal line and 1 each time it intersects a vertical line of the grid (in case the line meets a node, we write either 0 or 1). Given $t > 0$, we fix a line segment L_t of the line L of length

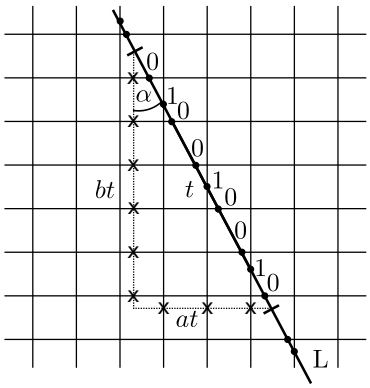


Figure 2: Sturmian sequence

t. Denote by a and b the absolute values of the sine and cosine of the angle at which the line intersects the vertical lines of the grid, respectively. We claim that L_t intersects the vertical lines of the grid either $[at]$ or $[at]+1$ times. This can be seen easily after drawing a picture (see Figure 2). Indeed, L_t intersects as many vertical lines of the grid as the side of the triangle opposed to angle α does. This side has length at and therefore indeed intersects either $[at]$ or $[at]+1$ vertical lines of the grid. This is the number of 1's in the corresponding block of η. In a similar way, the number of 0's is equal to $[bt]$ or $[bt]+1$. Denote the number of 1's in the block of η corresponding to L_t by $\#1(L_t)$ and the number of 0's by $\#0(L_t)$. Using this notation, we have

$$at < \#1(L_t) \le at + 1 \text{ and } bt < \#0(L_t) \le bt + 1. \tag{62}$$

[15] Recall that among lines L with a rational slope, we consider only those which do not intersect the nodes of the grid.

Therefore

$$(a+b)t-2 < n(L_t) \leq (a+b)t+2,$$

where $n(L_t) = \#1(L_t) + \#0(L_t)$. It follows that

$$\frac{n(L_t)-2}{a+b} \leq t < \frac{n(L_t)+2}{a+b}. \tag{63}$$

Now, fix $B_n \in \{0,1\}^n$. Then $n = n(L_t)$ for some $t > 0$ and for some line segment L_t of length t and it follows by (62) and (63) that

$$\begin{aligned} \#1(B_n) = \#1(B_{n(L_t)}) = \#1(L_t) &\leq at+1 \\ &< a\frac{n(L_t)+2}{a+b}+1 = \frac{a}{a+b}(n+2)+1 < \frac{a}{a+b}n+3 \end{aligned} \tag{64}$$

and in a similar way

$$\#1(B_n) > \frac{a}{a+b}(n-2)-1 > \frac{a}{a+b}n-3 \tag{65}$$

and the assertion follows as $\frac{a}{a+b} = \frac{|\tan\alpha|}{1+|\tan\alpha|}$ can take any value between 0 and 1. □

Remark 6.19. In particular, it follows from the above lemma that

$$\delta - \frac{3}{N} = \frac{1}{N}(\delta N - 3) \leq \frac{1}{N}\sum_{n\leq N}\eta(n) \leq \frac{1}{N}(\delta N+3) = \delta + \frac{3}{N},$$

whence

$$\frac{1}{N}\sum_{n\leq N}\eta(n) \to \delta. \tag{66}$$

Moreover,

$$\delta > 0 \text{ for any Sturmian sequence other than } (0,0,0,\dots). \tag{67}$$

6.2.3 Sequences z with $h_{top}(z^2) = 0$

Notice that using Proposition 6.12 and Sturmian sequences, we can easily obtain various sequences z satisfying (**Ch**) and such that $h_{top}(z^2) = 0$. It suffices to take $z = \eta \cdot u$, where u is a generic point for the $B(1/2,1/2)$ measure and η is a Sturmian sequence. We will now calculate the value of $h_{top}(z)$ for sequences obtained this way. Notice that by Lemma 6.1 and Remark 5.9, we have $h_{top}(z) \in (0,1]$.

Proposition 6.20. *For any Sturmian sequence $\eta \in \{0,1\}^{\mathbb{N}}$, and any $u \in \{-1,1\}^{\mathbb{N}}$ generic for the Bernoulli measure $B(1/2,1/2)$, the sequence $z := \eta \cdot u$ satisfies* (**Ch**)*, $h_{top}(z^2) = 0$ and $h_{top}(z) = \delta$, where δ is as in Lemma 6.18. Moreover, any $\delta \in [0,1]$ can be obtained in this way.*

Proof. Let $\eta \in \{0,1\}^{\mathbb{N}}$ be a Sturmian sequence, and δ be as in Lemma 6.18, i.e.

$$\delta n - 3 < \#1(B_n) < \delta n + 3 \tag{68}$$

for any $B_n \in \{0,1\}^n$ appearing in η. Fix a generic point $u \in \{-1,1\}^{\mathbb{N}}$ for the measure $B(1/2,1/2)$. Then by Proposition 6.12, $z = \eta \cdot u$ satisfies (**Ch**).

Since z satisfies (**Ch**), it is generic for the relatively independent extension of the measure given by the block frequencies in $z^2 = \eta$. Therefore, given a block C appearing on $\eta = z^2$, and B such that $B^2 = C$, B will appear on z. Hence

$$p_{z^2}(n) \cdot 2^{\delta n - 3} < p_z(n) < p_{z^2}(n) \cdot 2^{\delta n + 3}, \tag{69}$$

which yields $h_{top}(z) = \delta$. □

6.2.4 Sequences z with $h_{top}(z^2) > 0$

Proposition 6.21. *For any Sturmian sequence $\eta \in \{0,1\}^{\mathbb{N}}$ and any $u \in \{-1,0,1\}^{\mathbb{N}}$ generic for the Bernoulli measure $B(1/4,1/2,1/4)$, the sequence $z := \eta \cdot u$ satisfies (**Ch**), $h_{top}(z^2) = \delta$ and $h_{top}(z) = \delta \log 3$, where δ is as in Lemma 6.18. Moreover, any $\delta \in [0,1]$ can be obtained in this way.*

Proof. Take $u \in \{-1,0,1\}^{\mathbb{N}}$ generic for the Bernoulli measure $B(1/4,1/2,1/4)$. Notice that u^2 is generic for $B(1/2,1/2)$, and that $B(1/4,1/2,1/4)$ is the relatively independent extension of $B(1/2,1/2)$. By Lemma 4.6, u satisfies (**Ch**). Let $\eta \in \{0,1\}^{\mathbb{N}}$ be a Sturmian sequence, and let δ be as in Lemma 6.18. By Proposition 6.16, $z := u \cdot \eta$ also satisfies (**Ch**).

Notice that any block B appearing on z arises by replacing some of the 1's in a block C appearing on η by $0's$ or -1's. Moreover, all blocks of this form appear on z. Thus,

$$3^{\delta n - 3} \leq p_z(n) \leq (n+1) \cdot 3^{\delta n + 3},$$

whence $h_{top}(z) = \delta \log 3$. In a similar way, we obtain $h_{top}(z^2) = \delta$, which completes the proof. □

Remark 6.22. Suppose that $b_k = a_k^2$, $k \geq 1$ are pairwise relatively prime and let $\boldsymbol{\mu}_{\mathscr{B}}$ be given by formula (15). Let $z \in \{-1,0,1\}^{\mathbb{N}}$ be a sequence satisfying (**Ch**), such that $z^2 = \boldsymbol{\mu}_{\mathscr{B}}^2$. By Theorem 5.3. in [2], we have

$$h_{top}(z^2) = \prod_{k=1}^{\infty} \left(1 - \frac{1}{b_k}\right).$$

Moreover, the same arguments yield

$$h_{top}(z) = \log 3 \cdot \prod_{k=1}^{\infty} \left(1 - \frac{1}{b_k}\right).$$

Recall (cf. [21] and [20]) that in the classical case when $z = \boldsymbol{\mu}$, we have $\prod_{k=1}^{\infty} \left(1 - \frac{1}{b_k}\right) = \frac{6}{\pi^2}$.

7 Toeplitz sequences correlating with a given sequence, and their topological entropy

Since the Sarnak conjecture holds for periodic sequences, the following question arises:

> "How far" does (a_n) need to be from a periodic sequence so that the Sarnak conjecture "fails", i.e. $\frac{1}{N}\sum_{n\leq N} a_n \cdot \boldsymbol{\mu}(n) \not\to 0$?

Natural candidates for (a_n) are Toeplitz sequences, see Section 2.2.7, as they exhibit a strong periodic structure: any block appearing on a Toeplitz sequence, appears on it periodically (the period may vary, depending on the chosen block). We will work in an abstract setting, dealing, instead of $\boldsymbol{\mu}$, with a sequence $z \in \{-1,0,1\}^{\mathbb{N}}$ satisfying some additional assumptions. Under these assumptions, we will construct Toeplitz sequences $t = (t_n)$ such that

$$\frac{1}{N}\sum_{n\leq N} t_n \cdot z(n) \not\to 0 \tag{70}$$

and show that $h_{top}(t) > 0$, giving more precise entropy estimates.

The starting point for our constructions is the following simple observation: if the upper density of 1's in z^2 is positive then

$$\frac{1}{N}\sum_{n\leq N} z(n)\cdot z(n) \not\to 0.$$

The underlying idea of the constructions is to find a Toeplitz sequence t which has "as much as possible in common" with the sequence z under consideration.

We apply our results to the following two classes of sequences:

(a) sequences satisfying (**Ch**), related to Sturmian sequences (see Section 6.2.3 and Section 6.2.4);

(b) $z = \boldsymbol{\mu}$, $z = \boldsymbol{\mu}_{\mathscr{B}}$ and any sequence z such that $z^2 = w^2$, where w is as in (a).

Notice that in case (a), in view of Theorem 4.10, (70) clearly implies that t is not completely deterministic, so, in particular, $h_{top}(t) > 0$. Therefore, what we are really interested in, are the obtained entropy (lower) estimates. In case (b), we cannot refer to (**Ch**) anymore to show that $h_{top}(t) > 0$, it needs to be shown separately. Note however that our entropy estimates are not as precise as in case (a) (the reason is that we have less knowledge about z). It is also unclear whether the constructed Toeplitz sequences are not completely deterministic.

7.1 Abstract setting

Let $z \in \{-1,0,1\}^{\mathbb{N}^*}$ be such that

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{n\leq N} z^2(n) > 0. \tag{71}$$

We will show that there exists a Toeplitz sequence $t \in \{-1,0,1\}^{\mathbb{N}^*}$ which correlates with z, i.e.

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n\le N} t_n \cdot z(n) > 0 \tag{72}$$

(see Proposition 7.7 below). Moreover, under some additional assumptions on z, we will give estimates for $h_{top}(t)$. More precisely, we will prove the following.

Theorem 7.1. *Suppose that (71) holds. If, moreover, z is such that:*

(a) z is quasi-generic for some $\widetilde{\nu} \in \mathcal{P}_S(\{-1,0,1\}^{\mathbb{N}^*})$,

(b) $H := h_{top}(supp(\widetilde{\nu})) > 0$,

(c) there exist $q \ge 2$ and $b \ge 1$ such that, for all $m \ge 1$, the number of ergodic components of the action of S^{q^m} on $(\{-1,0,1\}^{\mathbb{N}^*}, \widetilde{\nu})$ is bounded by b,

then, for any $\varepsilon > 0$, t can be chosen so that $h_{top}(t) \ge (1-\varepsilon)H$ and (72) holds.

Remark 7.2. Note that condition (c) above holds for an arbitrary $q \ge 2$ whenever $(S, \widetilde{\nu})$ is ergodic and there exists $b \ge 1$ such that for any rational eigenvalue λ of $(S, \widetilde{\nu})$, $\lambda^c = 1$ for some $1 \le c \le b$. In particular, (c) holds if $(S, \widetilde{\nu})$ is totally ergodic.

Theorem 7.3. *Suppose that (71) holds. If moreover z is such that*

(a') z^2 is quasi-generic for ν,

(b') $H := h_{top}(supp(\nu)) > 0$,

(c') there exist $q \ge 2$ and $b \ge 1$ such that, for all $m \ge 1$, the number of ergodic components of the action of S^{q^m} on $(\{0,1\}^{\mathbb{N}^*}, \nu)$ is bounded by b,

then, for any $\varepsilon > 0$, t can be chosen so that $h_{top}(t) \ge (1-\varepsilon)H$ and (72) holds.

Remark 7.4. Although Theorem 7.1 seems to give a better lower entropy estimation than Theorem 7.3, it cannot be applied in many interesting cases (see Section 7.2.1) because of the assumption (c) which we are not able to verify. In such cases, we apply Theorem 7.3.

The proofs of Theorems 7.1 and 7.3 go along the same lines. Since they are quite technical, they will be split into several sections.

7.1.1 A Toeplitz sequence correlating with z

Fix some $q \ge 2$ and, for each $j \ge 1$, consider the arithmetic progression

$$A_j := \{j + nq^j : n \ge 0\} \subset \mathbb{N}^*.$$

Definition 7.5. We say that $j \in \mathbb{N}^*$ is *initial* if there is no $j' < j$ with $j \in A_{j'}$. Then,

$$\{A_j : j \text{ initial}\} \text{ is a partition of } \mathbb{N}^*.$$

When j is initial, we denote by A_j^* the set $A_j \setminus \{j\}$. Elements of A_j^* for some initial j are said to be *non-initial*. We denote the set of all non-initials by $\mathcal{N}$.

The Toeplitz sequence we are interested in, is the sequence $t = (t_n)_{t\in\mathbb{N}^*} \in \{-1,0,1\}^{\mathbb{N}^*}$ defined by

$$t_n := \begin{cases} z(n) \text{ if } n \text{ is initial,} \\ z(j) \text{ if } n \in A_j^* \text{ for some initial } j. \end{cases} \tag{73}$$

Lemma 7.6. *For any $N \geq 1$, we have*

$$\frac{\mathcal{N} \cap \{1,\dots,N\}}{N} \leq \frac{1}{q-1}.$$

Proof. Let j be initial. Since the difference of two consecutive terms in A_j^* is q^j, and since the first term of the arithmetic progression A_j is missing in A_j^*, we have $\left|A_j^* \cap \{1,\dots,N\}\right| \leq \frac{N}{q^j}$ for each $N \geq 1$. Hence,

$$\frac{|\mathcal{N} \cap \{1,\dots,N\}|}{N} \leq \sum_{j\notin\mathcal{N}} \frac{1}{q^j} \leq \sum_{j\geq 1} \frac{1}{q^j} = \frac{1}{q-1}.$$

□

Proposition 7.7. *Suppose that (71) holds. Then, for q sufficiently large, the Toeplitz sequence t defined by (73) correlates with z, i.e.*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n\leq N} t_n \cdot z(n) > 0.$$

Proof. We have

$$\frac{1}{N} \sum_{n\leq N} t_n \cdot z(n) = \frac{1}{N} \sum_{n\leq N, n\in\mathcal{N}} t_n \cdot z(n) + \frac{1}{N} \sum_{n\leq N, n\notin\mathcal{N}} t_n \cdot z(n), \tag{74}$$

where

$$\left| \frac{1}{N} \sum_{n\leq N, n\in\mathcal{N}} t_n \cdot z(n) \right| \leq \frac{|\mathcal{N} \cap \{1,\dots,N\}|}{N} \leq \frac{1}{q-1} \tag{75}$$

by Lemma 7.6. Moreover, using once more Lemma 7.6, we have

$$\begin{aligned} \frac{1}{N} \sum_{n\leq N, n\notin\mathcal{N}} t_n \cdot z(n) &= \frac{1}{N} \sum_{n\leq N, n\notin\mathcal{N}} z^2(n) \\ &= \frac{|\mathrm{supp}(z) \cap \mathcal{N}^c \cap \{1,\dots,N\}|}{N} = 1 - \frac{|((\mathrm{supp}(z))^c \cup \mathcal{N}) \cap \{1,\dots,N\}|}{N} \\ &\geq 1 - \frac{|(\mathrm{supp}(z))^c \cap \{0,\dots,N\}|}{N} - \frac{|\mathcal{N} \cap \{1,\dots,N\}|}{N} \\ &\geq \frac{|\mathrm{supp}(z) \cap \{1,\dots,N\}|}{N} - \frac{1}{q-1} = \frac{1}{N} \sum_{n\leq N} z^2(n) - \frac{1}{q-1}. \end{aligned}$$

Therefore

$$\frac{1}{N} \sum_{n\leq N} t_n \cdot z(n) \geq \frac{1}{N} \sum_{n\leq N} z^2(n) - \frac{2}{q-1}.$$

By (71), the latter expression is bounded below by a fixed positive number whenever q and N are large enough, which completes the proof. □

7.1.2 Two types of non-initial numbers

Fix an integer $m \geq 1$. For any integer $k \geq 0$, we consider the interval

$$I_{m,k} := \Big(kq^m,\ (k+1)q^m\Big] \cap \mathbb{N}^*.$$

We distinguish two types of non-initials in $I_{m,k}$:

Definition 7.8. A non-initial in $I_{m,k}$ is said to be:

- *of type 1* if it belongs to some A_j^* with $j \leq m$,
- *of type 2* if it belongs to some A_j^* with $j > m$.

Remark 7.9. Observe that, if for some $k \geq 1$ and some $1 \leq r \leq q^m$, $kq^m + r$ is a non-initial of type 1 in $I_{m,k}$, then for any other $k' \geq 1$, $k'q^m + r \in I_{m,k'}$ is also a non-initial of type 1 (since it belongs to the same A_j^*). Hence, the pattern formed by non-initials of type 1 inside $I_{m,k}$ does not depend on which $k \geq 1$ we consider.

On the other hand, consider A_{m+h}^* for some $h \geq 1$. This set of non-initial numbers intersects $I_{m,k}$ every q^h-th integer k, and when it does, the single non-initial point of type 2 in the intersection is always of the form $kq^m + r$ for some r depending on h but not on k.

7.1.3 The end of the interval

Fix additionally an integer $1 \leq \ell < m$, set $L := q^\ell$, and consider the last L elements of $I_{m,k}$:

$$I_{m,k,L} := \Big((k+1)q^m - L,\ (k+1)q^m\Big] \cap \mathbb{N}^*.$$

Definition 7.10. We say that the integer k is *good* if the only non-initial integers in $I_{m,k,L}$ are of type 1.

By Remark 7.9, for all good k's, the pattern formed by non-initial integers inside $I_{m,k,L}$ is always the same.

Proposition 7.11. *The upper density of the set*

$$\mathcal{M} := \{k \geq 1 : k \textit{ is not good}\}$$

is bounded by $1/q^{q^m - m - L}$.

Proof. Let $n \in I_{m,k,L}$ be a non-initial element of type 2. Then $n \in A_j^*$ for some initial $j > m$, and we have $n \equiv j \bmod q^j$, hence also $n \equiv j \bmod q^m$. This and the definition of $I_{m,k,L}$ imply $j > q^m - L$, i.e.

(76) the non-initials of type 2 inside $I_{m,k,L}$ belong to some A_j^* with $j > q^m - L$.

Now, fix an initial $j > m$ and let k_0 be such that $j \in I_{m,k_0}$. Then

$$\{k \geq 0 : I_{m,k} \cap A_j^* \neq \varnothing\} = \{k_0 + i \cdot q^{j-m} : i \geq 1\}.$$

Hence, for any $K \geq 1$, we have

$$\frac{1}{K}\left|\{0 \leq k < K : I_{m,k} \cap A_j^* \neq \varnothing\}\right| \leq \frac{1}{q^{j-m}}.$$

It follows that

$$\frac{1}{K}\left|\{0 \leq k < K : I_{m,k} \cap A_j^* \neq \varnothing \text{ for some } j > q^m - L\}\right| \\ \leq \sum_{j > q^m - L} \frac{1}{q^{j-m}} = \frac{1}{q^{q^m - m - L}} \sum_{h \geq 1} \frac{1}{q^h} < \frac{1}{q^{q^m - m - L}}.$$

In view of (76), this ends the proof. □

7.1.4 Density of non-initials of type 1 inside $I_{m,k,L}$

We want now to bound the density of non-initials of type 1 inside $I_{m,k,L}$ (which are the only non-initials in this interval when k is good).

Lemma 7.12. *Let $n \in I_{m,k,L}$ be a non-initial of type 1. Then $n \in A_j^*$, where j satisfies $j > q^j - L$ (*cf. (76)*).*

Proof. Let n be a non-initial of type 1 inside $I_{m,k,L}$. Then, by the definition of type 1, there exists an initial j with $j \leq m$ such that $n \in A_j^*$. Thus $n \equiv j \bmod q^j$, and also

$$n \equiv (k+1)q^m + j \bmod q^j.$$

Since $n \leq (k+1)q^m$, there exists an integer $s \geq 1$ with

$$n = (k+1)q^m + j - s\,q^j.$$

But $n > (k+1)q^m - L$, hence

$$(k+1)q^m - L < (k+1)q^m + j - s\,q^j \leq (k+1)q^m + j - q^j,$$

and the assertion follows. □

Proposition 7.13. *For $k \geq 1$, the proportion of non-initial elements of type 1 inside $I_{m,k,L}$ is equal to*

$$\frac{1}{q} + \frac{1}{q^2} + \cdots + \frac{1}{q^\ell}. \tag{77}$$

Proof. First, let us show that there are no non-initial elements of type 1 inside $I_{m,k,L}$ which are in some A_j^* with $j > \ell$. Indeed, suppose that such an element exists. Then, we can write $j = \ell + s$ for some integer $s \geq 1$, and Lemma 7.12 gives

$$\ell + s > q^{\ell+s} - q^\ell = q^\ell(q^s - 1).$$

If $q^s - 1 = 1$, then $q = 2$ and $s = 1$, and we get $\ell \geq 2^\ell$, which is impossible. Otherwise, using the inequality $\alpha\beta \geq \alpha + \beta$ for any $\alpha \geq 2$, $\beta \geq 2$, we obtain

$$\ell + s \geq q^\ell + q^s,$$

which is also impossible since $\ell < q^\ell$ and $s < q^s$.

It remains to estimate the contribution of non-initial elements of type 1 which are in some A_j^* with $j \le \ell$. For each such j, since q^j divides the length $L = q^\ell$ of $I_{m,k,L}$, we have

$$\frac{|A_j \cap I_{m,k,L}|}{L} = \frac{1}{q^j}.$$

Since j ranges from 1 to ℓ, (77) follows. □

7.1.5 Ergodic components

Proposition 7.14. *Let $\mathbb{A}$ be a finite alphabet, and let ν be a shift-invariant probability measure on $\mathbb{A}^{\mathbb{N}^*}$. Let $n \ge 1$ and suppose that (S^n, ν) has $b \ge 1$ ergodic components. If $\nu = \frac{1}{n}\sum_{s=0}^{n-1}(S^s)_*\eta$, where η is S^n-invariant then*

$$\nu = \frac{1}{n}\left(\left[\frac{n}{b!}\right]\sum_{s=0}^{b!-1}(S^s)_*(\eta) + \sum_{s=0}^{(n \bmod b!)-1}(S^s)_*(\eta)\right).$$

Proof. For $0 \le i \le b-1$, let $\rho^{(i)}$ be the ergodic components of (S^n, ν), i.e.

$$\nu = \alpha_0\rho^{(0)} + \cdots + \alpha_{b-1}\rho^{(b-1)}$$

for some $0 < \alpha_0, \ldots, \alpha_{b-1} < 1$, $\sum_{i=0}^{b-1}\alpha_i = 1$. Then

$$\nu = S_*(\nu) = \alpha_0 S_*(\rho^{(0)}) + \ldots \alpha_{b-1}S_*(\rho^{(b-1)}).$$

For any $0 \le i \le b-1$, $S_*(\rho^{(i)})$ is again an ergodic S^n-invariant measure. Since the ergodic decomposition of (S^n, ν) is unique, this yields a permutation $\pi\colon \{0, 1, \ldots, b-1\} \to \{0, 1, \ldots, b-1\}$,

$$\pi(i) = j \iff S_*(\rho^{(i)}) = \rho^{(j)}.$$

Clearly, $\pi^{b!} = \mathrm{Id}$. Now, since $\eta \ll \nu$,

$$\eta = \beta_0\rho^{(0)} + \cdots + \beta_{b-1}\rho^{(b-1)}$$

for some $0 \le \beta_i \le 1$, $\sum_{i=0}^{b-1}\beta_i = 1$, whence $(S^{b!})_*(\eta) = \eta$ and the assertion follows. □

Corollary 7.15. *Under the assumptions of Proposition 7.14, whenever $n \ge 2b!$, for any finite family of sets $\{A_i : i \in I\}$ with $\nu(A_i) > 0$, there exists $0 \le s \le b!-1$ such that*

$$\left|\left\{i \in I : (S^s)_*(\eta)(A_i) \ge \frac{1}{2}\nu(A_i)\right\}\right| \ge \frac{|I|}{b!}.$$

Proof. By Proposition 7.14, for every $i \in I$, there exists $0 \le s_i \le b!-1$ such that

$$\frac{1}{n}\left(\left[\frac{n}{b!}\right] + 1\right)(S^{s_i})_*(\eta)(A_i) \ge \frac{1}{b!}\nu(A_i).$$

Since $\frac{1}{n}\left(\left[\frac{n}{b!}\right] + 1\right) \le \frac{2}{b!}$, we have $(S^{s_i})_*(\eta)(A_i) \ge \frac{1}{2}\nu(A_i)$ and the result easily follows by Fubini's argument. □

Proposition 7.16. *Fix $\varepsilon > 0$. Let $\mathbb{A}$ be a finite alphabet, fix $w \in A^{\mathbb{N}^*}$ and suppose that the following conditions hold:*

(a) w is quasi-generic for some shift-invariant measure ν for which

(b) $H := h_{top}(supp(\nu)) > 0$,

(c) there exist $q \geq 2$ and $b \geq 1$ such that, for all $m \geq 1$, the number of ergodic components of the action of S^{q^m} on $(\mathbb{A}^{\mathbb{N}^*}, \nu)$ is bounded by b.

Then, for all $\ell \geq 1$ large enough, there exists $\tau_\ell > 0$ such that, for all $m > \ell$, we can find $0 \leq s \leq b! - 1$ satisfying (using as before the notation $L := q^\ell$)

$$\left|\left\{C \in \mathbb{A}^{L-b!} : \limsup_{K\to\infty} \frac{1}{K} \sum_{0\leq k<K} \mathbf{1}_C \left(S^{(k+1)q^m - L+1+s} w\right) \geq \tau_\ell\right\}\right| \geq 2^{H(1-\varepsilon)L}.$$

Proof. It follows by (b) that

$$\left|\left\{C \in \mathbb{A}^{L-b!} : \nu(C) > 0\right\}\right| \geq 2^{H(1-\varepsilon/3)(L-b!)} \geq 2^{H(1-2\varepsilon/3)L}, \tag{78}$$

when ℓ (and hence also L) is large enough. Fix such an ℓ, which additionally satisfies the following inequality:

$$\frac{2^{H(1-2\varepsilon/3)L}}{b!} \geq 2^{H(1-\varepsilon)L}. \tag{79}$$

Let

$$\tau_\ell := \frac{1}{2} \min\left\{\nu(C) : C \in \mathbb{A}^{L-b!},\ \nu(C) > 0\right\} \tag{80}$$

and take $m > \ell$.

By (a), we may find (N_j) such that

$$\nu = \lim_{j\to\infty} \delta_{S,N_j,w}.$$

Since $\left[\frac{N_j}{q_m}\right] q_m/N_j \to 1$, by replacing N_j with $\left[\frac{N_j}{q_m}\right] q_m$ if necessary, we can assume that $q_m | N_j$ for $j \geq 1$. Passing to a subsequence if necessary, we can further assume the existence of

$$\eta := \lim_{j\to\infty} \delta_{S^{q_m},N_j/q_m,w}.$$

Then $\nu = \frac{1}{q^m} \sum_{s=0}^{q^m-1} (S^s)_*(\eta)$. From (80), Corollary 7.15 applied to $\{C \in \mathbb{A}^{L-b!} : \nu(C) > 0\}$, (78) and (79), it follows that there exists $0 \leq s \leq b! - 1$ such that

$$\begin{aligned} \left|\left\{C \in \mathbb{A}^{L-b!} \colon (S^{q^m-L+1+s})_*(\eta)(C) \geq \tau_\ell\right\}\right| \\ \geq \left|\left\{C \in \mathbb{A}^{L-b!} \colon (S^{q^m-L+1+s})_*(\eta)(C) \geq \frac{1}{2}\nu(C)\right\}\right| \\ \geq \frac{2^{H(1-2\varepsilon/3)L}}{b!} \geq 2^{H(1-\varepsilon)L}. \end{aligned} \tag{81}$$

Notice that, if C is a cylinder such that $(S^{q^m-L+1+s})_*(\eta)(C) \geq a$ for some $a > 0$, then

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{0\leq k<K} \mathbf{1}_C(S^{(k+1)q^m-L+1+s}w) \geq a.$$

This and (81) imply

$$\left|\left\{C \in \mathbb{A}^{L-b!} \colon \limsup_{K\to\infty} \frac{1}{K} \sum_{0\leq k<K} \mathbf{1}_C(S^{(k+1)q^m-L+1+s}w) \geq \tau_\ell\right\}\right| \geq \frac{2^{H(1-2\varepsilon/3)L}}{b!} \geq 2^{H(1-\varepsilon)L},$$

which completes the proof. □

An immediate consequence of the above proposition are the following two corollaries.

Corollary 7.17. *Let $\varepsilon > 0$ and suppose that the assumptions (a'), (b') and (c') of Theorem 7.3 hold. Then for all $\ell \geq 1$ large enough, there exists $\tau_\ell > 0$ such that, for all $m > \ell$, we can find $0 \leq s \leq b! - 1$ satisfying*

$$\left|\left\{C \in \{0,1\}^{L-b!} \colon \limsup_{K\to\infty} \frac{1}{K} \sum_{0\leq k<K} \mathbf{1}_C\left(S^{(k+1)q^m-L+1+s}z^2\right) \geq \tau_\ell\right\}\right| \geq 2^{H(1-\varepsilon/2)L}.$$

Corollary 7.18. *Let $\varepsilon > 0$ and suppose that the assumptions (a), (b) and (c) of Theorem 7.1 hold. Then for all $\ell \geq 1$ large enough, there exists $\tau_\ell > 0$ such that, for all $m > \ell$, we can find $0 \leq s \leq b! - 1$ satisfying*

$$\left|\left\{C \in \{-1,0,1\}^{L-b!} \colon \limsup_{K\to\infty} \frac{1}{K} \sum_{0\leq k<K} \mathbf{1}_C\left(S^{(k+1)q^m-L+1+s}z\right) \geq \tau_\ell\right\}\right| \geq 2^{H(1-\varepsilon/2)L}.$$

7.1.6 Entropy estimates

Proof of Theorem 7.3. We will need the following notation: if $A = \{a_1 < a_2 < \cdots < a_r\}$ is a finite subset of $\mathbb{N}^*$, and if $x = (x(n))_{n\in\mathbb{N}^*}$ is a sequence in $\{0,1\}^{\mathbb{N}^*}$, we denote by $x(A)$ the finite sequence

$$x(A) := \big(x(a_1), \ldots, x(a_r)\big) \in \{0,1\}^r.$$

Fix $\varepsilon > 0$. Replacing q by q^r if necessary, for some large r (which does not alter the validity of (c')), we can assume that q is large enough to satisfy the assertion of Proposition 7.7, and also that

$$\frac{1}{q-1} < \frac{\varepsilon}{2}H. \tag{82}$$

Let ℓ be an integer large enough to satisfy the assertion of Corollary 7.17, and $L := q^\ell$. Then, by Proposition 7.11, we can take m large enough so that the upper density of the set of integers k which are not good is strictly less than τ_ℓ. Let $0 \le s \le b! - 1$ be given by Corollary 7.17. Then, for any $C \in \{0,1\}^{L-b!}$ satisfying

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{0\le k<K} \mathbf{1}_C \left(S^{(k+1)q^m - L + 1 + s} z^2 \right) \ge \tau_\ell,$$

there exist infinitely many good integers k such that the block corresponding to the cylinder set C appears at position s of $I_{m,k,L}$ on the sequence z^2. Since, by Corollary 7.17, the number of such cylinder sets is at least $2^{H(1-\varepsilon/2)L}$, we can deduce that

$$\left|\{z^2(I_{m,k,L}) : k \text{ good}\}\right| \ge 2^{H(1-\varepsilon/2)L}. \tag{83}$$

We will show now that for ℓ sufficiently large, the number of blocks of length L in t is at least $2^{H(1-\varepsilon)L}$, more precisely, we claim that

$$|\{t(I_{m,k,L}) : k \text{ is good}\}| \ge 2^{H(1-\varepsilon)L}.$$

We have

$$I_{m,k,L} = A_{m,k,L} \sqcup B_{m,k,L},$$

where

$$A_{m,k,L} := \{n \in I_{m,k,L} : n \in \mathcal{N}\},$$
$$B_{m,k,L} := \{n \in I_{m,k,L} : n \notin \mathcal{N}\}.$$

By Proposition 7.13,

$$|A_{m,k,L}| = L\left(\frac{1}{q} + \cdots + \frac{1}{q^l}\right) < \frac{L}{q-1} \qquad \text{whenever } k \text{ is good},$$

whence

$$|\{z^2(A_{m,k,L}) : k \text{ is good}\}| \le 2^{\frac{L}{q-1}}. \tag{84}$$

Observe also that, when k is good, the relative positions of $A_{m,k,L}$ and $B_{m,k,L}$ inside $I_{m,k,L}$ are always the same. Hence,

$$\begin{aligned} &\left|\{z^2(I_{m,k,L}) : k \text{ is good}\}\right| \\ &\qquad \le \left|\{z^2(A_{m,k,L}) : k \text{ is good}\}\right| \cdot \left|\{z^2(B_{m,k,L}) : k \text{ is good}\}\right|. \end{aligned} \tag{85}$$

Therefore, in view of (84), (83) and (82), we obtain

$$\begin{aligned} |\{z^2(B_{m,k,L}) : k \text{ is good}\}| &\ge \frac{|\{z^2(I_{m,k,L}) : k \text{ is good}\}|}{|\{z^2(A_{m,k,L}) : k \text{ is good}\}|} \\ &\ge \frac{1}{2^{\frac{L}{q-1}}} |\{z^2(I_{m,k,L}) : k \text{ is good}\}| \ge \frac{2^{H(1-\varepsilon/2)L}}{2^{\frac{L}{q-1}}} \ge 2^{H(1-\varepsilon)L}. \end{aligned} \tag{86}$$

Finally

$$\begin{aligned} &|\{t(I_{m,k,L}) : k \text{ is good}\}| \ge |\{t(B_{m,k,L}) : k \text{ is good}\}| \\ &\quad = |\{z(B_{m,k,L}) : k \text{ is good}\}| \ge |\{z^2(B_{m,k,L}) : k \text{ is good}\}| \ge 2^{H(1-\varepsilon)L} \end{aligned} \tag{87}$$

and the result follows. □

Proof of Theorem 7.1. The proof goes along the same lines as the proof of Theorem 7.3 (instead of $\{0,1\}$, we consider the alphabet $\{-1,0,1\}$). First, (using Corollary 7.18 instead of Corollary 7.17) we show that

$$\left|\{z(I_{m,k,L}) : k \text{ good}\}\right| \geq 2^{H(1-\varepsilon/2)L} \tag{88}$$

(cf. formula (83)). The formulas (85) and (86) are still valid, with z playing now the role of z^2. In (87) it suffices to remove the part involving z^2 to obtain the result. □

7.2 Applications

7.2.1 $\boldsymbol{\mu}$ and its generalizations: $h_{top}(z^2) > 0$

Let $\mathscr{B} = \{b_k : k \geq 1\}$ be a set of pairwise relatively prime numbers with $b_k = a_k^2$ and let $z(n) = \boldsymbol{\mu}_{\mathscr{B}}(n)$ be given by formula (15). Then the following is true:

(a) The point z^2 is generic for some measure ν. Moreover, for any block C appearing on z^2, $\nu(C) > 0$ (for $z = \boldsymbol{\mu}$, see [20] and for the general case, see [2]).

(b) $\frac{1}{N}\sum_{n\leq N} z^2(n) \xrightarrow[N\to\infty]{} h_{top}(\text{supp}(\nu)) = h_{top}(z^2) > 0$.

(c) (S,ν) has purely discrete spectrum. Moreover, for q prime:

- if $q \nmid b_k$ for all $k \geq 1$ then (S^q,ν) is totally ergodic,
- if $q \mid b_k$ for some $k \geq 1$ then such k is unique and for any $m \geq 1$, S^{q^m} has at most b_k ergodic components

(see Theorem 4.4. in [2]).

It follows immediately from the above that the assumptions of Theorem 7.3 hold for z. Thus we have obtained the following:

Corollary 7.19. *Fix $\varepsilon > 0$. For $z = \boldsymbol{\mu}_{\mathscr{B}}$ (including the case $z = \boldsymbol{\mu}$), there exists a Toeplitz sequence t which correlates with z, such that $h_{top}(t) \geq (1-\varepsilon)h_{top}(z^2)$.*

Remark 7.20. It would be interesting to know, whether we can find a Toeplitz sequence t so that t correlates with $\boldsymbol{\mu}_{\mathscr{B}}$ and, moreover, $h_{top}(t) \geq (1-\varepsilon)h_{top}(z)$ (cf. Remark 7.4).

Remark 7.21. Recall that $h_{top}(\boldsymbol{\mu}^2) = 6/\pi^2$. Therefore, in view of Proposition 7.7 and Theorem 7.3, in case $z = \boldsymbol{\mu}$, it suffices to take $q = 5$ in the construction of t, in order to obtain $h_{top}(t) > 0$.

7.2.2 Sturmian sequences

Let $\eta \in \{0,1\}^{\mathbb{N}^*}$ be a Sturmian sequence and let $u \in \{-1,0,1\}^{\mathbb{N}^*}$ be a generic point for some Bernoulli measure $\mathbb{B}$. Let $z := \eta \cdot u$. Let ν be the measure from Remark 2.18. Denote $\rho = m_*(\nu \otimes \mathbb{B})$, see (61).

Lemma 7.22. $h_{top}(supp(\rho)) = h_{top}(z) > 0$.

Proof. In view of Remark 5.9, it suffices to show that $h_{top}(\mathrm{supp}(\rho)) = h_{top}(z)$.

Clearly, whenever B is such that $\rho(B) > 0$, then B appears on z. Let now B be a block which appears on z. Then $B = B_1 \cdot B_2$ (the multiplication is to be understood coordinatewise) for some block B_1 which appears on η and some block B_2 which appears on u. Therefore

$$\rho(B) = \rho(B_1 \cdot B_2) = \nu \otimes \mathbb{B}(m^{-1}(B_1 \cdot B_2)) \\ \geq \nu \otimes \mathbb{B}(B_1 \times B_2) = \nu(B_1) \cdot \mathbb{B}(B_2) > 0$$

by Remark 2.18 and since $\mathbb{B}(C) > 0$ for any block C. This ends the proof. □

Lemma 7.23. $\lim_{N\to\infty} \frac{1}{N} \sum_{n\leq N} z^2(n) > 0$.

Proof. It follows from Remark 2.18 that $(S, X_\eta, \nu) \perp (S, X_u, \mathbb{B})$, whence (η, u) is generic for $\nu \otimes \mathbb{B}$. Therefore,

$$\begin{aligned}\frac{1}{N}\sum_{n\leq N} z^2(n) &= \frac{1}{N}\sum_{n\leq N} \eta^2(n)\cdot u^2(n) = \frac{1}{N}\sum_{n\leq N} \eta(n)\cdot u^2(n) \\ &= \frac{1}{N}\sum_{n\leq N} (\mathbf{1}_{\{w:w(0)=1\}} \otimes \mathbf{1}_{\{w:w(0)=\pm 1\}})((S\times S)^n(\eta, u)) \\ &\xrightarrow[N\to\infty]{} \int \mathbf{1}_{\{w:w(0)=1\}} \otimes \mathbf{1}_{\{w:w(0)=\pm 1\}}\, d(\nu\otimes\mathbb{B}) \\ &= \nu(\{w : w(0) = 1\}) \cdot \mathbb{B}(\{w : w(0) = \pm 1\}) > 0.\end{aligned}$$

□

Remark 7.24. Notice that (S, X_z, ρ) with $\rho = m_*(\nu \otimes \mathbb{B})$ is a factor of $(S, X_\eta, \nu) \times (S, X_u, \mathbb{B})$. Therefore, z is generic for ρ. Moreover, the eigenvalues of $(S, X_\eta, \nu) \times (S, X_u, \mathbb{B})$ and (S, X_η, ν) are the same and they are of the form λ^n, $n \in \mathbb{Z}$ for some $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. In particular, any eigenvalue of (S, X_z, ρ) is of the form λ^n for some $n \in \mathbb{Z}$.

Lemma 7.25. *There exists $b \geq 1$ such that (S^r, X_z, ρ) has at most b ergodic components for any $r \geq 1$.*

Proof. If (S, X_z, ρ) is totally ergodic, the assertion of the lemma is true. Assume now that (S, X_z, ρ) is not totally ergodic. Let $\lambda \in \mathbb{C}$ be as in Remark 7.24 and let $n_0 \geq 1$ be the smallest natural number such that $\lambda^{n_0} = 1$. It follows that (S^r, X_z, ρ) has at most n_0 ergodic components for any $r \geq 1$, which ends the proof. □

Corollary 7.26. *Fix $\varepsilon > 0$. For z defined as above, there exists a Toeplitz sequence t which correlates with z and such that*

$$h_{top}(t) \geq (1 - \varepsilon) h_{top}(z).$$

Proof. In view of Lemma 7.22, Lemma 7.23, Remark 7.24 and Lemma 7.25, the assumptions of Theorem 7.1 are satisfied for z and the assertion follows. □

References

[1] E. H. El Abdalaoui, S. Kasjan, and M. Lemańczyk. 0-1 sequences of the Thue-Morse type and Sarnak's conjecture. http://arxiv.org/abs/1304.3587, 04 2013.

[2] El H. El Abdalaoui, M. Lemańczyk, and T. de la Rue. A dynamical point of view on the set of $\mathcal{B}$-free integers. *International Mathematics Research Notices*, 2014. doi: 10.1093/imrn/rnu164

[3] T. M. Apostol. *Introduction to analytic number theory.* Springer-Verlag, New York-Heidelberg, 1976. Undergraduate Texts in Mathematics.

[4] F. Cellarosi and Y. G. Sinai. Ergodic properties of square-free numbers. *J. Eur. Math. Soc. (JEMS)*, 15(4):1343–1374, 2013.

[5] S. Chowla. *The Riemann hypothesis and Hilbert's tenth problem.* Mathematics and Its Applications, Vol. 4. Gordon and Breach Science Publishers, New York, 1965.

[6] H. Davenport. On some infinite series involving arithmetical functions. II. *Quart. J. Math. Oxford*, 8:313–320, 1937.

[7] T. Downarowicz. *Entropy in Dynamical Systems*, volume 18 of *New Mathematical Monographs.* Cambridge University Press, Cambridge, 2011.

[8] T. Downarowicz, *Survey of odometers and Toeplitz flows*, Contemporary Mathematics, Algebraic and Topological Dynamics (Kolyada, Manin, Ward eds), vol. 385, 2005, pp. 7-38.

[9] T. Downarowicz, S. Kasjan, *Odometers and Toeplitz subshifts revisited in the context of Sarnak's conjecture*, in preparation.

[10] N. P. Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics*, volume 1794 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[11] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory*, 1:1–49, 1967.

[12] E. Glasner, J.-P. Thouvenot, and B. Weiss. Entropy theory without a past. *Ergodic Theory Dynam. Systems*, 20(5):1355–1370, 2000.

[13] K. Jacobs, M. Keane, *0-1 sequences of Toeplitz type*, Z. Wahr. Verw. Geb. **13** (1969), 123-131.

[14] T. Kamae. Subsequences of normal sequences. *Israel J. Math.*, 16:121–149, 1973.

[15] D.G. Kerr, H. Li, *Independence in topological and C^*-dynamics*, Math. Ann. **338** (2007), 869-926.

[16] D. Kwietniak, *Topological entropy and distributional chaos in hereditary shifts with applications to spacing shifts and beta shifts*, Discrete Contin. Dyn Syst. **33** (2013), 2451-2467.

[17] L. Mirsky. Arithmetical pattern problems relating to divisibility by rth powers. *Proc. London Math. Soc. (2)*, 50:497–508, 1949.

[18] D. Ornstein. Factors of Bernoulli shifts are Bernoulli shifts. *Advances in Math.*, 5:349–364 (1970), 1970.

[19] W. Parry. *Entropy and generators in ergodic theory.* W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[20] R. Peckner. Uniqueness of the measure of maximal entropy for the squarefree flow. Arxiv, 05.2012.

[21] P. Sarnak. Three lectures on the Möbius function, randomness and dynamics. http://publications.ias.edu/sarnak/.

[22] P. Shields. *The ergodic theory of discrete sample paths*, volume 13 of *Graduate Sudies in Mathematics*. American Mathematical Society, 1996.

[23] T. Tao. *The Chowla conjecture and the Sarnak conjecture*. What's new (blog). http://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/.

[24] E. C. Titchmarsh. *The theory of the Riemann zeta-function*. The Clarendon Press Oxford University Press, New York, second edition, 1986. Edited and with a preface by D. R. Heath-Brown.

[25] P. Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982.

[26] B. Weiss. Normal sequences as collectives. *Proc. Symp. on Topological Dynamics and ergodic theory*, Univ. of Kentucky, 1971.

[27] B. Weiss. *Single orbit dynamics*, volume 95 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, Providence, RI, 2000.

[28] S. Williams, *Toeplitz minimal flows which are not uniquely ergodic*, Z. Wahr. Verw. Geb. **67** (1984), 95-107.

[29] R. J. Zimmer. Extensions of ergodic group actions. *Illinois J. Math.*, 20(3):373–409, 1976.

El Houcein El Abdalaoui
Laboratoire de Mathématiques Raphaël Salem, Normandie Université, Université de Rouen, CNRS – Avenue de l'Université – 76801 Saint Etienne du Rouvray, France
E-mail address: elhoucein.elabdalaoui@univ-rouen.fr

Joanna Kułaga-Przymus
Institute of Mathematics, Polish Acadamy of Sciences, Śniadeckich 8, 00-956 Warszawa, Poland
Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopina 12/18, 87-100 Toruń, Poland
E-mail address: joanna.kulaga@gmail.com

Mariusz Lemańczyk
Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopina 12/18, 87-100 Toruń, Poland
E-mail address: mlem@mat.umk.pl

Thierry de la Rue
Laboratoire de Mathématiques Raphaël Salem, Normandie Université, Université de Rouen, CNRS – Avenue de l'Université – 76801 Saint Etienne du Rouvray, France
E-mail address: Thierry.de-la-Rue@univ-rouen.fr