Odometers and Toeplitz systems revisited in the context of Sarnak's conjecture

by

Tomasz Downarowicz (Warszawa and Wrocław) and
Stanisław Kasjan (Toruń)

Abstract. Although Sarnak's conjecture holds for compact group rotations (irrational rotations, odometers), it is not even known whether it holds for all Jewett–Krieger models of such rotations ([1]). In this paper we show that it does, as long as the model is at the same a topological extension, via the same map that establishes the isomorphism, of an equicontinuous model. In particular, we recover (after [AKL]) that regular Toeplitz systems satisfy Sarnak's conjecture, and, as another consequence, so do all generalized Sturmian subshifts (not only the classical Sturmian subshift). We also give an example of an irregular Toeplitz subshift which meets our criterion. We give an example of a model of an odometer which is not even Toeplitz (it is weakly mixing), hence does not meet our criterion. However, for this example, we manage to produce a separate proof of Sarnak's conjecture. Next, we provide a class of Toeplitz sequences which fail Sarnak's conjecture (in a weak sense); all these examples have positive entropy. Finally, we examine the example of a Toeplitz sequence from [AKL] (which fails Sarnak's conjecture in the strong sense) and prove that it also has positive entropy (this proof has been announced in [AKL]).

This paper can be considered a sequel to [AKL], it also fills some gaps of [D].

1. Introduction. This work results from the discussions the authors held with Mariusz Lemańczyk about topological (in particular symbolic) models of odometers in the context of Sarnak's conjecture. We refer the readers to the recent paper of Lemańczyk et al. [AKL] for results concerning

([1]) Recently, H. El Abdalaoui, M. Lemańczyk and T. de la Rue [ALR2] proved that any topological model of an ergodic system with *irrational* discrete spectrum satisfies Sarnak's conjecture. So the problem remains open for odometers (pure rational spectrum) and mixed (rational and irrational) spectra. In this paper we focus mainly on models of odometers, so the interference with the above mentioned result is small.

Sarnak's conjecture for Morse systems, where some indispensable facts concerning regular Toeplitz subshifts were obtained partly independently and partly jointly, and inspired this work.

First of all, it has been discovered that although regular Toeplitz subshifts are the best known symbolic models (see next page for definition) of odometers, there are also other possibilities, the existence of which was not fully understood before. In the preceding work of the first author [D] there are some erroneous statements about regularity, and this paper fixes them; there exist *irregular* Toeplitz models of their underlying odometers. For completeness, we also give examples of models which are not even Toeplitz, some of them even topologically mixing. Sarnak's conjecture can be briefly proved for both regular and irregular Toeplitz models of odometers, as long as they are *isomorphic extensions* ([2]) of the odometer. The method uses the sole property of the Möbius function, that it is orthogonal to any periodic sequence, otherwise it relies on an easy spectral argument. As a digression, we apply a similar spectral method (but a different property of the Möbius function) to prove Sarnak's conjecture for isomorphic extensions of other equicontinuous systems, in particular for generalized Sturmian subshifts ([3]). The method applies also to a large class of models of rotations which are their topological extensions without being almost 1-1. The existence of this class (which turns out to be quite large) has not been known until a recent paper [DG].

Our method fails for models of equicontinuous systems which are not their topological extensions (e.g. topologically weakly mixing) and the case remains a challenge ([4]). Nonetheless, we are able to successfully apply it to our specific topologically mixing example. On this occasion we also give relatively simple examples of Toeplitz subshifts which massively fail Sarnak's conjecture (at many points of the system including those which are Toeplitz sequences), showing that just being a union of periodic sequences is insufficient. We also copy from [AKL] an example in which a Toeplitz sequence fails the conjecture so badly that the limit inferior of the absolute values of the averages is positive. As one might expect, the counterexamples have positive entropy, of which we give a detailed proof (which is intentionally skipped in [AKL] and left to be given in this paper).

([2]) A topological system (X, T) is an *isomorphic extension* of another, (Y, S), if both systems are uniquely ergodic, with invariant measures μ and ν, respectively, and there exists a topological factor map from X onto Y which is, at the same time, an isomorphism between (X, μ, T) and (Y, ν, S). See [DG] for more information on isomorphic extensions.

([3]) For generalized Sturmian subshifts this fact is covered by the still unpublished result of [ALR2]. Our argument was established earlier.

([4]) Now, by [ALR2], only when the spectrum contains a rational number.

2. Preliminaries on Toeplitz systems. The notation and terminology of this section is consistent with that of [D], to which we also refer for references to earlier papers. To make this paper selfcontained, the most crucial definitions will be repeated. By a *scale* we mean an increasing sequence $H = (p_k)_{k \geq 1}$ of positive integers such that $p_k \mid p_{k+1}$ for every k. The *adding machine* with scale H is the topological group G obtained as the inverse limit of the cyclic groups $\mathbb{Z}_{p_k} = \mathbb{Z}/p_k\mathbb{Z}$:

$$G = \varprojlim_{k} \mathbb{Z}_{p_k}.$$

By an *odometer* we will mean the topological dynamical system (G, τ) where τ is the homeomorphism $g \mapsto g + \mathbf{1}$ of G into itself, where $\mathbf{1} = (1, 1, \ldots)$ is the topological generator of G. The odometer is minimal, equicontinuous and zero-dimensional, and the conjunction of these three properties characterizes odometers among topological dynamical systems. Odometers are uniquely ergodic, with the Haar measure λ being the unique invariant measure. For us, an *odometer* also means the ergodic system (G, λ, τ). By a (*Jewett–Krieger*) *model* of an ergodic system we will understand any strictly ergodic (minimal and uniquely ergodic) topological dynamical system isomorphic (for its unique invariant measure ([5])) to the given ergodic system. Note that there may (and usually do) exist mutually nonconjugate models of the same ergodic system. For instance, an odometer is a model for itself, but there are other models as well, for example some symbolic systems over finite alphabets (which are never equicontinuous, so they cannot be conjugate to the odometer).

We will be mostly concerned with Toeplitz systems ([6]), understood as specified in the theorem below. Note that, contrary to the traditional terminology, we do not require that they are symbolic (nonetheless, all our examples will be symbolic). Toeplitz systems can be defined in a multitude of ways, which is captured in the following theorem. The relevant definitions and proofs can be found in [D] (for the reader's convenience, the definitions are also given in footnotes).

THEOREM 2.1. *The following conditions are equivalent for a topological dynamical system* (X, T). *A system satisfying them is called a* Toeplitz *system.*

(1) (X, T) *is the orbit closure of a regularly recurrent point* ([7]);

([5]) Convention: *isomorphism* is measure-theoretic while *conjugacy* is topological.

([6]) In the literature they are often called "Toeplitz flows"; we find this terminology confusing, as they are discrete time systems.

([7]) A point x is *regularly recurrent* if, for every open $U \ni x$, the set of return times to U contains an arithmetic progression $n\mathbb{Z}$.

(2) (X,T) *is a minimal almost 1-1 extension* ([8]) *of an odometer* (G,τ);

(3) (X,T) *is a semicocycle extension* ([9]) *of an odometer* (G,τ).

Conditions (1) and (2) are additionally related:

THEOREM 2.2. *If π is the almost 1-1 factor map from a Toeplitz system (X,T) to an odometer (G,τ) then $x \in X$ is regularly recurrent if and only if $\pi^{-1}(\pi(x)) = \{x\}$.*

In Theorem 2.1, the odometer (G,τ) appearing in (2) is the *maximal equicontinuous factor* ([10]) of (X,T). The adding machine G appearing in (3) (with the action of τ) is the maximal equicontinuous factor of (X,T) if and only if the semicocycle has the additional property of being *invariant under no rotations* ([11]). Otherwise, the maximal equicontinuous factor of (X,T) acts on a quotient group of G ([12]) and (G,τ) is not even a factor of (X,T) (let alone maximal equicontinuous). We can strengthen condition (3) as follows:

THEOREM 2.3. *Every Toeplitz system is conjugate to a semicocycle extension of its maximal equicontinuous factor.*

An important class of Toeplitz systems, called *regular* ([13]), is described by the conditions given below:

THEOREM 2.4. *Let π be the almost 1-1 factor map from a Toeplitz system (X,T) to an odometer (G,τ). The following conditions are equivalent:*

(1) *The set of points $g \in G$ such that $\#\pi^{-1}(g) = 1$ has full λ measure.*

(2) *The set of regularly recurrent points in (X,T) has full measure for every invariant measure on X.*

(3) *The set of discontinuities D_f of the relevant semicocycle f on G has λ measure zero.*

The following is obvious by (1):

THEOREM 2.5. *A regular Toeplitz system (X,T) is strictly ergodic and it is an isomorphic extension of its maximal equicontinuous factor (G,τ).*

([8]) The factor map $\pi : X \to G$ is *almost 1-1* if $\pi^{-1}(\pi(x)) = \{x\}$ holds on a residual (i.e., containing a dense G_δ) subset of X.

([9]) A *semicocycle* is a function $f : G \to K$ into a compact space, which is continuous on a residual subset of G. Let F denote the multifunction defined by the closure of the graph of f. Let $X_F = \{x \in K^{\mathbb{Z}} : (\exists g \in G)(\forall n \in \mathbb{Z})\ x(n) \in F(g+n)\}$. This set is closed and shift-invariant, and has a unique minimal subset which we denote by X_f; X_f with the action of the shift is called the *semicocycle extension* (associated with f).

([10]) Any other equicontinuous system which is a factor of (X,T) factors through (G,τ).

([11]) A semicocycle f on an adding machine G is *invariant under no rotations* if $F \circ (\cdot + g) = F \Rightarrow g = \mathbf{0}$.

([12]) This quotient group is G/H, where $H = \{g : F \circ (\cdot + g) = F\}$.

([13]) Note that the term "regular recurrence" was coined independently.

In particular, a regular Toeplitz system is a model for the underlying odometer. Unless π is 1-1 everywhere, (X, T) is not conjugate to (G, τ).

In [D, Theorem 13.1(5) and (6)] it is claimed that a strictly ergodic Toeplitz system isomorphic to its maximal equicontinuous factor is necessarily regular. Moreover, it is claimed that strict ergodicity need not be assumed if an isomorphism exists for some invariant measure. Unfortunately, the statement (even with strict ergodicity assumed) is false. A relevant counterexample is provided in the following sections.

Let us return to the general case. In symbolic systems regularly recurrent points are called *Toeplitz sequences* ([14]). Toeplitz subshifts (and Toeplitz sequences) were the first examples of almost 1-1 extensions of odometers and they are the most important. In this class we have an additional simplification:

THEOREM 2.6. *Every Toeplitz subshift over a finite alphabet Λ is conjugate to a semicocycle extension of its maximal equicontinuous factor with the semicocycle taking values in Λ.*

The advantage of having the finite-valued semicocycle is that its set of discontinuities $D_f \subset G$ is then closed. The elements of Toeplitz subshifts have specific structure, as described below.

DEFINITION 2.7. Let (X, T) be a Toeplitz subshift and let (G, τ) be the maximal equicontinuous factor of (X, T). For $x \in X$ and $p \in \mathbb{N}$ we denote

$$\mathsf{Per}_p(x) = \{n \in \mathbb{Z} : (\forall m \in \mathbb{Z})\, x(n) = x(n+mp)\}, \quad \mathsf{Aper}(x) = \mathbb{Z} \setminus \bigcup_{p \in \mathbb{N}} \mathsf{Per}_p(x),$$

and call these sets the *p-periodic part* and *aperiodic part*, respectively.

The union of periodic parts will not change if we unite over a scale $(p_k)_{k \in \mathbb{N}}$ of the adding machine G (then the union is increasing). It is important to know that $\mathsf{Per}_{p_k}(x)$ and $\mathsf{Aper}(x)$ depend only on $\pi(x)$. Clearly, x is a Toeplitz sequence if and only if its aperiodic part is empty.

Recall that for a set $A \subset \mathbb{Z}$ the *forward and backward densities* of A are defined as

$$\mathsf{dens}^+(A) = \lim_n \frac{1}{n} \#(A \cap [0, n-1]), \quad \mathsf{dens}^-(A) = \lim_n \frac{1}{n} \#(A \cap [-n, -1]),$$

respectively (provided the limits exist) and in case they coincide we call them the *density* of A and denote by $\mathsf{dens}(A)$.

THEOREM 2.8. *Let (X, T) be a Toeplitz subshift. Then, for every $p \in \mathbb{N}$, $\mathsf{dens}(\mathsf{Per}_p(x))$ exists and is constant throughout X.*

Let $d = 1 - \sup_k \mathsf{dens}(\mathsf{Per}_{p_k}(x))$. Then

---

([14]) Regular recurrence takes the form $(\forall n \in \mathbb{Z})(\exists p \in \mathbb{N})(\forall m \in \mathbb{Z})\, x(n) = x(n + mp)$.

(1) $d = \lambda(D_f)$;

(2) $\mathsf{dens}(\mathsf{Aper}(x)) \le d$ for every $x \in X$;

(3) $\mathsf{dens}(\mathsf{Aper}(x)) = d$ for μ-almost every $x \in X$, for every invariant measure μ on X.

In particular, regularity of (X, T) is equivalent to $d = 0$, and to the condition $\mathsf{dens}(\mathsf{Aper}(x)) = 0$ at every point x.

Let (X, T) be an irregular Toeplitz subshift. Consider the set of $x \in X$ such that $\mathsf{dens}(\mathsf{Aper}(x)) = d$ (by (3) above, this set has full invariant measure). Since $\mathsf{Aper}(x)$ is common for all points in the same fiber of π, our set equals $\pi^{-1}(E)$ for some $E \subset G$. For $x \in \pi^{-1}(E)$ we enumerate $\mathsf{Aper}(x) = \{n_i\}_{i \in \mathbb{Z}}$ assuming that the sequence (n_i) is increasing and letting n_0 be the smallest nonnegative element of the sequence. We let $y_x = (x(n_i))_{i \in \mathbb{Z}} \in \Lambda^{\mathbb{Z}}$ and call it the *aperiodic readout* of x. For $g \in E$ we define $Y_g = \{y_x : x \in \pi^{-1}(g)\}$.

DEFINITION 2.9. We say that the Toeplitz subshift (X, T) satisfies condition SAR (*same aperiodic readouts*) if Y_g is the same for every $g \in E$. We then denote the common space Y_g by Y.

It is easy to see that in this case Y is closed and shift invariant. The following theorem plays a crucial technical role in most of our examples (for proofs see [D]):

THEOREM 2.10. *Let (X, T) be an irregular Toeplitz subshift satisfying condition SAR. Let T_f be the skew product acting on $G \times Y$ given by*

$$T_f(g, y) = (\tau(g), S^g(y)),$$

where S^g equals the shift or the identity, depending on whether $g \in D_f$ or not (T_f is not continuous, but in what follows, we can ignore this fact). Then:

(1) *There is a bijection between invariant measures of (X, T) and invariant measures of $(G \times Y, T_f)$.*

(2) *Every invariant measure on (X, T) yields a system isomorphic to that given by its corresponding measure on the skew product.*

(3) *Every invariant measure on the skew product has marginals λ on G and some shift-invariant measure ν on Y.*

(4) *Every shift-invariant measure ν on Y appears as the marginal for at least one invariant measure on the skew product (for example for $\lambda \times \nu$).*

(5) *The entropy of the skew product with respect to an invariant measure equals d times the entropy of the corresponding marginal on Y.*

(6) *The topological entropy of the skew product (which equals the topological entropy of (X, T)) equals d times the topological entropy of the shift on Y.*

We will invoke the above facts several times.

3. Preliminaries on the Möbius function and Sarnak's conjecture. The *Möbius function* denoted by $\boldsymbol{\mu}$ is defined on the positive integers as follows:

$$\boldsymbol{\mu}(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n \text{ has a repeated prime factor}, \\ (-1)^r & \text{if } n \text{ is a product of } r \text{ distinct primes}. \end{cases}$$

This function was introduced by A. F. Möbius [M] to obtain inversion formulas for arithmetic functions [CDM]. It plays an important role in number theory. The reader is referred for more information to the rich literature in that area; let us only mention two fundamental monographs, [P] and [W].

DEFINITION 3.1. Let $\xi(n)$ and $\eta(n)$ be two bounded complex-valued sequences over $\mathbb{N}$. We say that these sequences are *uncorrelated* if

$$\lim_n \frac{1}{n} \sum_{i=1}^n \xi(i)\overline{\eta}(i) = 0.$$

One of the intriguing properties of the Möbius function is its apparent randomness in the distribution of its values. It is well known that the forward density of square-free numbers (i.e., of the set $\{n : |\boldsymbol{\mu}(n)| = 1\}$) exists and equals $6/\pi^2$ (see [N, Thm. 21.8 and Corollary]). On the other hand, the densities of positive and negative values are equal, implying that $\boldsymbol{\mu}$ is uncorrelated to the constant sequence (see [P, Thm. 5.1]).

Moreover, it is uncorrelated to any periodic function (an elementary proof can be found in [GL]), a fact which is connected with the laws of the distribution of primes along arithmetic progressions (see [Sa]). Let us remark that a more detailed analysis of this phenomenon, more precisely, of the behavior of the partial sums of the Möbius function, is an important area of study, connected with many fundamental number-theoretical problems: see e.g. the classical works [W], [Sch] and more recent papers [RR], [HS].

Sarnak [Sa] conjectures that $\boldsymbol{\mu}$ is uncorrelated to any sequence obtained by reading any continuous function along any orbit in any topological dynamical system with topological entropy zero:

CONJECTURE 3.2. *Let (X, T) be a topological dynamical system with topological entropy zero. Let $f : X \to \mathbb{C}$ be a continuous function. Fix an $x \in X$ and let $\xi(n) = f(T^n x)$ (for $n \geq 1$). Then ξ and $\boldsymbol{\mu}$ are uncorrelated.*

The conjecture is known to hold for relatively few types of dynamical systems, in particular for odometers, irrational rotations, nilsequences [GT] and horocycle flows [BSZ]. See also [B1], [B2], [ALR1], [G], [KL], [LS], [MR] for other results. We remark that the validity of Sarnak's conjecture for

odometers follows directly from the fact that the Möbius function is uncorrelated to any periodic sequence. The validity for irrational rotations can be proved by a criterion from [BSZ] or by a completely different property of the Möbius function, discovered by Davenport [Da]. It is a folklore fact that the conjecture holds for the classical Sturmian subshift; the proof uses heavily the fact that this subshift has complexity ([15]) $c(n) = n + 1$ ([MH]). Recently, in [AKL] Sarnak's conjecture has been proved for some Morse subshifts. Every continuous function on such a system decomposes as the sum of a function depending on the Toeplitz factor and an orthogonal one. Thus the method relies on two ingredients. For continuous functions orthogonal to the Toeplitz factor some specific spectral and disjointness methods are used. To handle the other ingredient, the authors simply prove the conjecture for regular Toeplitz subshifts.

In this note we extend the latter proof (in fact, we only notice that essentially the same proof applies) to a class slightly larger than regular Toeplitz systems, that of isomorphic extensions of minimal equicontinuous systems. This includes some not necessarily regular Toeplitz systems and generalized Sturmian subshifts.

4. Sarnak's conjecture for isomorphic extensions. The following fact has been observed jointly by the first author and M. Lemańczyk.

THEOREM 4.1. *Let (X, T) be an isomorphic extension of (Y, S) (i.e. the systems are uniquely ergodic and there exists a topological factor map $\pi : X \to Y$ which is, at the same time, an isomorphism for the corresponding ergodic systems). If Sarnak's conjecture holds for (Y, S) then it also holds for (X, T).*

Proof. Let $f : X \to \mathbb{C}$ be continuous. Then $f \in L^2(\mu)$ and $f' = f \circ \pi^{-1} \in L^2(\nu)$ (f' is defined ν-almost everywhere on Y). Since $C(Y)$ is dense in $L^2(\nu)$, there exists a continuous $g' : Y \to \mathbb{C}$ such that $\int |f' - g'|^2 \, d\nu < \epsilon^2$, hence $\int |f' - g'| \, d\nu < \epsilon$. The function $g = g' \circ \pi$ is continuous on X and $\int |f - g| \, d\mu < \epsilon$. Because in uniquely ergodic systems every point is generic ([16]), we have $\lim_n n^{-1} \sum_{i=1}^n |f(T^i x) - g(T^i x)| < \epsilon$ for any $x \in X$. Now, we write

$$\left| \frac{1}{n} \sum_{i=1}^n f(T^i x) \boldsymbol{\mu}(i) \right| \leq \left| \frac{1}{n} \sum_{i=1}^n g(T^i x) \boldsymbol{\mu}(i) \right| + \frac{1}{n} \sum_{i=1}^n |f(T^i x) - g(T^i x)| \, |\boldsymbol{\mu}(i)|.$$

The first average on the right hand side equals $|n^{-1} \sum_{i=1}^n g'(S^i y) \boldsymbol{\mu}(i)|$, where $y = \pi(x)$, and is small for large n, because Sarnak's conjecture holds on (Y, S). The last average does not exceed, for large n, the arbitrarily small ϵ.

([15]) The number of words of length n in the subshift.

([16]) That is, satisfies the ergodic theorem for every continuous function.

Thus the left hand side tends to zero with growing n. ∎

In [AKL] the reader will find a slightly different statement, in which (Y, S) is assumed to be *coalescent* ([17]) and the assumption that the isomorphism between (X, T) and (Y, S) is realized by the same topological factor map π is dropped (it is then fulfilled automatically). Recall that odometers and other ergodic group rotations are coalescent.

We now draw conclusions concerning particular types of topological dynamical systems. It seems that items (1b) ([18]), (2) and (3) below are new. Notice that for the classical Sturmian subshift we have obtained a new proof not relying on the exact complexity.

COROLLARY 4.2. *Sarnak's conjecture holds for all minimal isomorphic extensions of equicontinuous systems* ([19]), *in particular for:*

(1) *regular semicocycle extensions of any minimal equicontinuous system* ([20]), *in particular*

   (1a) *regular Toeplitz systems* (*see also* [AKL]);
   (1b) *generalized Sturmian subshifts* ([21]);

(2) *irregular Toeplitz subshifts which nevertheless are isomorphic extensions of odometers* ([22]) (*existence of such subshifts is proved in Example* 5.1 *below*);

(3) *isomorphic extensions of odometers* (*and other minimal equicontinuous systems*) *which are not their almost 1-1 extensions* (*existence of such systems is proved in* [DG]).

Proof. To be absolutely clear, let us argue why the conjecture holds for minimal equicontinuous systems. By the Halmos–von Neumann Theorem, every such system is uniquely ergodic and the space $L^2(\mu)$ is spanned by (at most countably many) continuous eigenfunctions. Thus every continuous function can be approximated in $L^2(\mu)$ (hence also in $L^1(\mu)$) by a finite sum

---

([17]) That is, such that every endomorphism of the system is an isomorphism.

([18]) This fact follows also from [ALR2].

([19]) In a recent paper [DG] these systems are characterized as minimal *mean equicontinuous systems* as defined by Li, Tu and Ye [LTY].

([20]) Notice that semicocycle extensions can be equally well defined on any strictly ergodic system, not necessarily an odometer. Regularity means that the set of discontinuities of the semicocycle has measure zero.

([21]) A classical Sturmian subshift is obtained as the semicocycle extension of the irrational rotation by an angle α, where the semicocycle is precisely the characteristic function of $[0, \alpha)$. In generalized Sturmian subshifts the semicocycle is allowed to be the characteristic function of any nondegenerate subinterval or even a finite union of intervals.

([22]) Note that if a system is an isomorphic extension of an equicontinuous system, then the latter is necessarily the maximal equicontinuous factor of the former. In the case of Toeplitz systems, the extension is then almost 1-1.

of continuous eigenfunctions. Now, by an argument as in the preceding proof, it suffices to verify the conjecture for continuous eigenfunctions. But every such function arises as a continuous function defined on either an odometer (if the eigenvalue is rational) or an irrational rotation (otherwise). ∎

5. Examples of models of odometers

EXAMPLE 5.1. *There exist* irregular *Toeplitz subshifts which, nevertheless, are isomorphic extensions of odometers* (*for these Sarnak's conjecture holds*).

Sketch of the construction. An explicit example of such a system is generated (as the shift-orbit closure) by the (unilateral) Toeplitz sequence described below.

Pick a block $B_1 = 000\ldots01000\ldots0$ of some length r_1 consisting of zeros but one symbol 1 (somewhere). We place this block p_1-periodically (for some $= p_1 > r_1$, we also let $q_1 = p_1$) (see Figure 1).

$0100***0100***0100***0100***0100***0100***0100***0100***0100***0100\ldots$

Fig. 1. In this figure $B_1 = 0100$, $r_1 = 4$ and $p_1 = 7$.

The unfilled places (the stars) come in blocks of length $Q_1 = q_1 - r_1$. We pick a block B_2 of some length $r_2 Q_1$, consisting of zeros but one symbol 1. We write this block into r_2 consecutive empty blocks and repeat p_2-periodically, where $p_2 = q_2 p_1$, for some $q_2 > r_2$ (see Figure 2).

$0100\mathbf{000}\,0100\mathbf{100}\,0100\mathbf{000}\,0100***\,0100***\,0100***\,0100\mathbf{000}\,0100\mathbf{100}\,0100\mathbf{000}\,0100\ldots$

Fig. 2. In this figure $B_2 = 000100000$, $r_2 = 3$ and $q_2 = 6$.

Now the unfilled positions come in clusters of $q_2 - r_2$ blocks of length $q_1 - r_1$.

We continue in this manner: in step $k+1$ we use a block B_{k+1} consisting of all zeros but one symbol 1, whose length equals r_{k+1} (a freely chosen number) times $Q_k = (p_1 - r_1)\ldots(p_k - r_k)$ (the number of unfilled positions in $[0, p_k - 1]$ in the sequence constructed so far); we use this block to fill all unfilled places in $[0, r_{k+1}p_k - 1]$; then we repeat it with a period $p_{k+1} = q_k p_k$ for some freely chosen $q_k > r_k$. Two more details must be taken care of: the products $\prod_{k=1}^{N}(1 - r_k/q_k)$, representing the density of unfilled positions after step N, must converge to a number $d > 0$. The second requirement is that for each k the symbol 1 appears in the future blocks $B_{k'}$ ($k' > k$) at positions whose remainders modulo Q_k take every possible value infinitely many times. With such an arrangement it is not hard to see that:

(1) the generated (bilateral) Toeplitz subshift (X, T) is an almost 1-1 extension of the odometer (G, τ) with scale $H = (p_k)$;

(2) (X, T) is irregular and satisfies condition SAR with the space Y of aperiodic readouts consisting of all $\{0, 1\}$-valued sequences having at most one symbol 1.

Clearly, Y supports only one invariant measure ν, the point mass at the fixpoint $(\ldots 000 \ldots)$. Theorem 2.10(1)–(3) implies that (X, T) is uniquely ergodic; the unique invariant measure μ is isomorphic to the only measure with marginals λ and ν, which is $\lambda \times \nu$. Since ν is supported by one point, the factor map π provides an isomorphism between μ and λ, as required. ∎

There exist strictly ergodic systems (also subshifts) isomorphic to an odometer (G, τ), yet whose maximal equicontinuous factor (G', τ') is a proper factor of (G, τ), and Theorem 4.1 does not apply to such systems. For G which is not simple (^{23}), such examples are easily obtained with G' being an adding machine and the system is an almost 1-1 extension of (G', τ'), while the remaining eigenvalues of G are realized by discontinuous eigenfunctions (see [DL]). There also exist models for which the maximal equicontinuous factor is trivial (hence all eigenvalues of (G, τ) are realized by discontinuous eigenfunctions and the system is topologically weakly mixing). Such examples can be produced for all odometers, including the simple ones. Below we give an example with the even stronger property of topological mixing (^{24}).

EXAMPLE 5.2. *Given an odometer (G, τ), there exists a strictly ergodic topologically mixing subshift (X, T) isomorphic to (G, τ).*

Sketch of the construction. Since we will be dealing with subshifts, T will always denote the shift transformation, regardless of the domain. Let $H = (p_k)$ denote the scale of the odometer. Let (X_0, T) be a regular Toeplitz subshift with maximal equicontinuous factor (G, τ), and let $x_0 \in X_0$ be a Toeplitz sequence. We will produce a sequence of topological conjugacies of (X_0, T), converging almost everywhere to an isomorphism with the desired subshift (X, T).

In Step 1, choose some $r_1 \in \mathbb{N}$ and find all periodic repetitions in x_0 of the central block $C_1 = x_0[-r_1, r_1]$. The period of the repetitions is some $p_1 \in H$. Choose q_1 such that $p_1 q_1 \in H$ and choose every q_1th periodic occurrence of C_1 in x_0 (avoiding the central one). Let us call these places *1-windows*. Now comes the modification: within each 1-window we shift the

---

(^{23}) An odometer is *simple* when its scale is (p^k) for a prime number p. Simple odometers have no infinite proper factors, others do.

(^{24}) It is known ([L]) that every aperiodic ergodic system has a topologically mixing strictly ergodic model. Here we provide a particular example.

contents one position to the left (sending the leftmost symbol to the right end). The modification passes over, in an obvious way, to all elements of X_0 and is invertible if q_1 is large enough ([25]). This is our conjugacy $\Phi_1 : X_0 \to X_1$ between two Toeplitz subshifts. We denote by x_1 the Toeplitz sequence $\Phi_1(x_0)$.

In Step 2 we choose some r_2 and we denote $C_2 = x_1[-r_2, r_2]$. We must take care of two details: C_2 must be long enough to include several 1-windows, and moreover its ends must fall *far* from the 1-windows, for instance, approximately in the middle between two of them. In the future this will prevent an accumulation of the end-irregularities. We find all periodic occurrences of C_2 in x_1, and their period $p_2 \in H$. Next we choose some large q_2 such that $p_2 q_2 \in H$ and we call every q_2th copy of C_2 in x_2 (avoiding the central one) a *2-window*. As before, we shift the contents of each 2-window one position to the left, sending the leftmost symbol to the right end. This modification spreads naturally to a conjugacy $\Phi_2 : X_1 \to X_2$ between Toeplitz subshifts. We let $x_2 = \Phi_2(x_1) = \Phi_2 \Phi_1(x_0)$.

We proceed in this manner infinitely many times, ensuring that the densities of the positions affected by consecutive modifications (i.e., the ratios $\rho_k = 2r_k + 1/(p_k q_k)$) are summable, and that the ends of the k-windows fall, for every $k' < k$, approximately in the middle between a pair of k'-windows (see Figure 3).

Fig. 3. The figure show the k-windows (i.e., areas affected by Φ_k): original—black, 1-windows—dark grey (above black), 2-windows—gray, 3-windows—light grey (the next 3-window is too far to be shown). The displacements by one position of dark grey within grey, and of grey and dark grey (by two positions) within light grey are too small to be seen.

Let x_∞ denote the sequence obtained as the limit of x_k (which exists because every position in x_0 is affected by the modifications at most finitely many times), and we let X_∞ be the shift orbit closure of x_∞ and finally we let X be a minimal subset of X_∞ ([26]). Notice that the maps $\Psi_k = \Phi_k \dots \Phi_2 \Phi_1$ converge at every point of X_0 whose coordinates are each affected at most finitely many times. We let Ψ be the (partially defined) limit map on X_0.

The following observations are standard and we skip their proofs:

(1) (X_∞, T) is topologically mixing (can be checked directly by the definition).

([25]) q_1 must be long enough so that the p_1-periodic part of any $x \in X_0$ can be determined by viewing any block of length $p_1 q_1 - 2r_1 - 1$.

([26]) This step allows us to avoid verifying whether X_∞ is minimal.

(2) Ψ is defined almost everywhere on X_0 and acts into X_∞.

(3) For each k, on X we can define the following map ψ_k: in each element of X we can easily identify and reverse the consequences of the modification $\Phi_1, \ldots, \Phi_k$ (although we "invert" Φ_k in a seemingly wrong order, these inverses commute). *Caution*: the maps ψ_k are not precisely the inverses of Ψ_k; they are defined on X_∞, not on X_k.

(4) The limit of the above maps ψ_k (call it ψ) is defined almost everywhere on X_∞ for every invariant measure.

(5) The range of ψ is contained in X_0 and ψ inverts Ψ wherever the latter is defined.

The last four facts imply that (X_∞, T) is uniquely ergodic and isomorphic to (X_0, T). Since the minimal subset $X \subset X_\infty$ supports an invariant measure, it supports the unique one, hence (X, T) is isomorphic to (X_0, T), and additionally minimal. Thus (X, T) is the desired example. ∎

Although we have no general proof of Sarnak's conjecture for all models of odometers, interestingly, it does hold in the above topologically mixing example. Since this example escapes any existing general method of proving Sarnak's conjecture (including the recent result of [ALR2]), the fact that we can still prove it (using a specific method) seems to be of particular interest. Again, we will only sketch the argument.

Sketch of proof of Sarnak's conjecture in Example 5.2. Since every continuous function on X can be approximated in $L^1(\mu)$ by a (continuous) linear combination of characteristic functions of cylinders corresponding to finite blocks, it suffices to verify the conjecture for such characteristic functions. We will do that for blocks of length 1, i.e., for the occurrences of the symbol 1. The argument for longer blocks is identical.

Fix an $\epsilon > 0$ and let k be such that the density of places where x_∞ differs from x_k (i.e., $\sum_{k'=k+1}^\infty \rho_{k'}$) is less than ϵ. Since x_k is regular Toeplitz, there are l and N such that the p_l-periodic part occupies at least a $1 - \epsilon$ fraction in every block of length N appearing in any element of X_k. The problem we are facing is that the (slightly perturbed) p_l-periodic part in $x \in X$ does not come from a continuous function on X any more, and we have no guarantee that *all* points realize the ergodic theorem for it. We need more subtle observations.

We can assume that N is large enough so that

$$\left| \sum_{i=1}^n \xi(i) \boldsymbol{\mu}(i) \right| < n\epsilon$$

for every $\{0, 1\}$-valued p_l-periodic sequence ξ and every $n \geq N$. Let $x \in X$ and consider the initial block $x[1, n]$ for $n \geq N/\epsilon$ (this is an arbitrary block B of length n appearing in x_∞). We will argue that there are (at most)

three blocks B_1, B_2, B_3 appearing in X_k, with lengths summing to n, such that $x[1, n]$ nearly equals the concatenation $B_1 B_2 B_3$ in the sense that the fraction of disagreements is at most 2ϵ. Once this is proved, we can write $B_1 = x'[1, m], B_2 = x''[m+1, m'], B_3 = x'''[m'+1, n]$ for some $x', x'', x''' \in X_k$ and $m, m' \in [1, n]$, $m \leq m'$, and then have the following estimation:

$$
\Big| \sum_{i=1}^{n} x(i)\boldsymbol{\mu}(i) \Big| \overset{2n\epsilon}{\approx} \Big| \sum_{i=1}^{m} x'(i)\boldsymbol{\mu}(i) + \sum_{i=m+1}^{m'} x''(i)\boldsymbol{\mu}(i) + \sum_{i=m'+1}^{n} x'''(i)\boldsymbol{\mu}(i) \Big|
$$

$$
= \Big| \sum_{i=1}^{m} x'(i)\boldsymbol{\mu}(i) + \sum_{i=1}^{m'} x''(i)\boldsymbol{\mu}(i) - \sum_{i=1}^{m} x''(i)\boldsymbol{\mu}(i)
$$

$$
+ \sum_{i=1}^{n} x'''(i)\boldsymbol{\mu}(i) - - \sum_{i=1}^{m'} x'''(i)\boldsymbol{\mu}(i) \Big| < 10n\epsilon,
$$

because

- the sums shorter than N contribute at most $sN < sn\epsilon$ ($s \leq 4$ is the number of such sums),
- in each sum not shorter than N:
    - x', x'' or x''' can be replaced by a p_l-periodic sequence and this will change the sum by less than ϵ times the summing length,
    - once the above replacement is done, the absolute value of the sum does not exceed ϵ times the summing length,
- the sum of the summing lengths equals $(5 - s)n$.

Clearly, this estimation ends the proof.

So, it remains to break a block B of length n appearing in x_∞ into at most three subblocks, as desired. We can think of B as of a block B_0 appearing somewhere in x_k, affected by some finitely many modifications $\Phi_{k+1}, \ldots, \Phi_K$. If B is entirely contained in a K-window without its endpoints, then Φ_K shifts the entire contents one position to the left, so that the result is the same as if we started from the block B_0' lying in x_k one position to the left with respect to B_0 and pretended that Φ_K did not affect it. In this manner, we can ignore all such cases and move on to the largest index (and denote it by K) for which only part of B is affected by Φ_K, that is, B contains an endpoint of a K-window. Now there are three possibilities:

(1) B intersects two or more K-windows,
(2) B contains one or both endpoints of just one K-window.

In case (1) the fraction of modifications introduced by Φ_K in B is at most $2\rho_K$ (and $\rho_{k'}$ for earlier modifications with $k' = k + 1, \ldots, K - 1$), so B differs from B_0 on at most 2ϵ places and there is no need to partition it (we

put $B_1 = B_0$, there is no B_2 or B_3). In case (2) we cut B at the endpoints of the K-window. This produces two or three subblocks B_1', B_2', B_3'. Notice that Φ_K affects only one of these subblocks and only by shifting it "in one piece". So, we only need to see how much each subblock is affected by the earlier modifications. Recall that for each $k' < K$ the endpoints of the k'-windows fall approximately $p_{k'}q_{k'}/2$ places away from the endpoints of the K-window. This implies that if a modification $\Phi_{k'}$ does affect a subblock B_i' ($i = 1, 2$ or 3), then the fraction of the modifications in this subblock is at most $2\rho_{k'}$. So each B_i' differs from a subblock of x_k on at most 2ϵ places. This completes the argument. ∎

6. Toeplitz sequences which fail Sarnak's conjecture. We will say that sequences ξ and η are *weakly* (resp. *strongly*) correlated if the upper (resp. lower) limit of $n^{-1} \sum_{i=1}^{n} \xi(i)\overline{\eta}(i)$ is positive.

6.1. Weak failure

EXAMPLE 6.1. *There exists an irregular $\{-1, 1\}$-valued Toeplitz subshift X such that some $x \in X$ are strongly correlated with $\boldsymbol{\mu}$. The set of such points x is dense in X. Moreover, weak correlation holds on a residual subset of X.*

The example is very simple, once the general construction of Toeplitz systems is understood: Consider an irregular Toeplitz subshift satisfying condition SAR and such that Y (the space of aperiodic readouts) is the full shift on two symbols $\{-1, 1\}$ (we skip the detailed construction of such a subshift; it is done by a "standard method", e.g., the Oxtoby technique, see [D]). Let $\pi : X \to G$ denote the maximal equicontinuous factor map onto the underlying adding machine and let λ be the Haar measure on G. Then, by Theorem 2.8 (and the explanations following that theorem), there is a set $G' \subset G$ with $\lambda(G') = 1$ which satisfies the following two conditions:

(1) all points x in the fiber $\pi^{-1}(g)$ agree along a common periodic part whose density equals $1 - d$,
(2) as x ranges over $\pi^{-1}(g)$, all possible $\{-1, 1\}$-valued sequences occur along the aperiodic part of x.

In particular, if we arrange that $d > 1 - 3/\pi^2$ (≈ 0.7) (which is easily done within the "standard method") then for every $g \in G'$ and $x \in \pi^{-1}(g)$ the set $\mathsf{A} = \mathsf{Aper}(x) \cap \{n \geq 0 : \boldsymbol{\mu}(n) \neq 0\}$ has lower forward density at least $d_0 = d + 6/\pi^2 - 1 > 3/\pi^2$. There exists $x_0 \in \pi^{-1}(g)$ such that for $n \in \mathsf{A}$, $x_0(n)$ equals $\boldsymbol{\mu}(n)$. It is obvious that (even in the "worst case scenario", when $x(i) = -\boldsymbol{\mu}(i)$ and $|\boldsymbol{\mu}(i)| = 1$ everywhere on the periodic part of x) we still

have

$$(6.1) \qquad \liminf_n \frac{1}{n} \sum_{i=1}^{n} x_0(i)\boldsymbol{\mu}(i) \geq d_0 - (1 - d) > 0.$$

Next, we will show that points x_0 as constructed above (satisfying (6.1)) lie densely in X. Consider a basic open set U in X, i.e., a cylinder corresponding to a block $B \in \{-1,1\}^{2k+1}$ occurring in X at the coordinates $[-k,k]$. By minimality, the same block occurs (perhaps at a different place) in the generating Toeplitz sequence, which implies that the same block occurs somewhere in the periodic part of every element of X, in particular in an element $x \in \pi^{-1}(G')$. Notice (directly from the definition) that the set G' is invariant (equivalently, $\pi^{-1}(G')$ is shift-invariant). Thus, by an appropriate shifting, we obtain a new point x such that B occurs in x at the coordinates $[-k,k]$ and still belongs to the periodic part of x, and $g = \pi^{-1}(x)$ belongs to G'. Using this particular g in the above construction of x_0 we produce the point x_0 such that $x_0[-k,k] = B$ (x_0 belongs to the same fiber as x and thus agrees with x along the periodic part, which includes the coordinates $[-k,k]$). In other words, we have constructed a point $x_0 \in U$ satisfying (6.1).

Next we observe that if we weaken (6.1) by requiring that the *upper* limit is larger than or equal to a positive $\epsilon < d_0 - (1 - d)$, then this holds on a residual set. Indeed, we can write

$$\left\{ x : \limsup_n \frac{1}{n} \sum_{i=1}^{n} x(i)\boldsymbol{\mu}(i) > \epsilon \right\} \subset \bigcap_{m \geq 1} \bigcup_{n \geq m} \left\{ x : \frac{1}{n} \sum_{i=1}^{n} x(i)\boldsymbol{\mu}(i) > \epsilon \right\}$$

$$\subset \left\{ x : \limsup_n \frac{1}{n} \sum_{i=1}^{n} x(i)\boldsymbol{\mu}(i) \geq \epsilon \right\}.$$

The first set contains the dense set of points satisfying (6.1), the middle set is of type G_δ (hence it is a dense G_δ), and thus the last set is residual. The example is thus constructed. ∎

Since the Toeplitz sequences form a residual subset inside a Toeplitz subshift, we obtain

COROLLARY 6.2. *There exist (irregular) Toeplitz sequences weakly correlated with the Möbius function.*

We remark that the Toeplitz subshift of the above example has positive entropy (equal to $d \ln 2$), hence it stands in no collision with Sarnak's conjecture.

6.2. Strong failure. The following example is replicated from [AKL].

EXAMPLE 6.3. *There exists a one-sided Toeplitz sequence strongly correlated with the Möbius function.*

We begin by describing a general scheme (used in [AKL]) of producing a one-sided Toeplitz sequence from another symbolic sequence. Let $y = (y_n)_{n \geq 1}$ be a one-sided sequence over a finite alphabet Λ. Let $H = (p_k)$ be a scale of an adding machine such that $p_1 \geq 3$. Since $p_{k+1} \geq 2p_k$ for every k, this condition implies that $\rho = \sum_{k \geq 1} 1/p_k < 1$. We define the associated one-sided Toeplitz sequence x as follows:

$$x_y(n) = \begin{cases} y_1, & n = 1 \bmod p_1, \\ y_2, & n = 2 \bmod p_2, \\ \vdots \\ y_{p_1}, & n = p_1 \bmod p_{p_1}, \\ y_{p_1+1}, & n = p_1 + 2 \bmod p_{p_1+1}, \\ y_{p_1+2}, & n = p_1 + 3 \bmod p_{p_1+2}, \\ \vdots \\ y_{2p_1-1}, & n = 2p_1 \bmod p_{2p_1-1}, \\ y_{2p_1}, & n = 2p_1 + 3 \bmod p_{2p_1}, \\ \vdots \end{cases}$$

We refrain from further detailed listing, as it becomes too complicated. The simple rule behind the scheme is that y_k is placed at the first position available after steps $1, \ldots, k-1$ and then it is repeated periodically with period p_k. This concludes the description of the scheme.

For further considerations, it will be convenient to highlight, for each k, the first placement of y_k in x_y (it is shown in boldface, while its further periodic repetitions are printed in normal font). While reading the following text and diagrams it is important to distinguish between boldface symbols $\mathbf{y}_k$ and normal font symbols y_k. The diagram below shows the filling scheme in case $p_k = 3^k$ with the boldface terms marked.

y₁y₂**y₃**y₁**y₄**y₅**y₁**y₆**y₇**y₁y₂**y₈**y₁**y₉**y₁₀**y₁y₁₁y₁₂**y₁y₂**y₁₃**y₁**y₁₄y₁₅**y₁**y₁₆y₁₇**y₁y₂y₃y₁**y₁₈y₁₉**y₁**y₂₀**···

The authors of [AKL] show that the lower density of the boldface symbols (which they call *initials*) is at least $1 - \rho$ (combining this with Lemma 7.1(3) below we see that in fact these symbols have density $1-\rho$), which can be made arbitrarily close to 1. They select y so that $x_y(n) = \boldsymbol{\mu}(n)$ whenever $x_y(n)$ is a boldface symbol. If $1 - \rho > 1 - 3/\pi^2$ then, for the same reasons as in (6.1) (with $1-\rho$ in the role of d), they find that x is strongly correlated with $\boldsymbol{\mu}$. Of course, in view of Sarnak's conjecture, one has to compute the topological entropy of the generated Toeplitz subshift (at least to check whether it is positive). We will do so in Section 8. ■

7. Properties of the scheme. We are interested in properties of Toeplitz sequences obtained through the above scheme for general sequences y. In particular, we would like to know whether positive entropy follows automatically from positive entropy of the orbit closure Y of y. As we soon show, the answer is negative. This is quite unfortunate, because it forces us to estimate the entropy of the example of [AKL] using tedious methods adapted to the particular example.

So, consider a general sequence y and the associated Toeplitz sequence x_y with the boldface symbols marked. The following lemma addresses the distribution of the boldface symbols in x_y. Statement (5) will be used immediately in Example 7.3, while statement (4) only in Section 8. Statements (1) and (2) are necessary to prove (4), while (3) is just a digression noted in passing.

LEMMA 7.1. *Let z be the $\{0,1\}$-valued sequence given by $z(n) = 1 \Leftrightarrow x_y(n)$ is a boldface symbol. Then:*

(1) *The frequency of zeros in the block $B_{k,0} = z[1, p_k]$ converges to ρ from below as $k \to \infty$.*

(2) *For every k, every block of the form*

$$B_{k,j} = B_{k,j}[1, p_k] = z[jp_k + 1, (j+1)p_k] \quad (n \geq 0)$$

can be obtained from $B_{k,0}$ by only replacing some 1's by 0's.

(3) *The symbols 0 in z have lower Banach density ρ.*

(4) *Given $\epsilon > 0$ there are $\delta > 0$ and $n_0 \in \mathbb{N}$ such that, for any $n \geq n_0$, the cardinality of different blocks B of length n, appearing in z and in which the frequency of 0's is at most $\rho + \delta$, does not exceed $2^{n\epsilon}$.*

(5) *For every natural m, z contains a block consisting of m single 1's separated by blocks of zeros of lengths at least m. In particular, the upper Banach density of zeros in z is 1.*

REMARK 7.2. (3) and (5) imply that the subshift generated by z is not uniquely ergodic; at least one invariant measure assigns to the cylinder of 0 the value ρ, and at least one – the value 1 (perhaps there are more possibilities). (4) implies that every measure of the first kind has entropy zero. We have not verified whether this subshift has topological entropy zero (regardless of the scale $H = (p_k)$).

Proof of Lemma 7.1. For (1) it suffices to observe that the frequency of zeros in $B_{k,0}$ equals

$$\frac{1}{p_k} \sum_{k' < k} \left(\frac{p_k}{p_{k'}} - 1 \right) = \sum_{k' < k} \frac{1}{p'_k} - \frac{k-1}{p_k}.$$

Indeed, in step $k' \leq k$, in $x_y[1, p_k]$ we have placed $p_k/p_{k'}$ symbols, of which one was boldface. Hence the formula.

For (2) note that $B_{k,0}(i) = 0$ if and only if $x_y(i)$ is a normal font (i.e., repeated) symbol $\mathrm{y}_{k'}$ for some k'. This is possible only when $p_{k'} < p_k$. But then p_k is a multiple of $p_{k'}$, which implies that $B_{k,j}(i) = 0$ for every j.

Clearly, (1) implies that the lower density (and lower Banach density) of zeros is at most ρ. By (2), the frequency of 0's in any $B_{k,j}$ may only be larger than that in $B_{k,0}$ (which is close to ρ). Every sufficiently long block B in z is a concatenation of the blocks $B_{k,j}$ and a negligibly small prefix and suffix, so the frequency of 0's in B is not less than ρ minus a negligibly small error term. This proves (3).

For (4) we argue as above: every block B of large length n, after removing a negligibly small prefix and suffix, becomes a concatenation of the blocks $B_{k,j}$ (with a large parameter k). By (2), this concatenation can be viewed as a periodic repetition of $B_{k,0}$ with some 1's replaced by 0's. But the number of replaced symbols 1 cannot substantially exceed $n\delta$, otherwise the overall frequency of 0's would be too large. Such a replacement can be performed in approximately $e^{n(-\delta \ln \delta - (1-\delta) \ln(1-\delta))}$ different ways, which, for an appropriately small δ, is smaller than $2^{n\epsilon}$.

For (5), we will need the following (somewhat lengthy)

CLAIM. *For every $m \geq 1$, after some number k_m of steps of filling x_y (i.e., after having placed the periodic repetitions of $y_1, \ldots, y_{k_m}$), x_y starts with the following configuration (later referred to as C): a continuous entirely filled block (with both boldface and normal font symbols) ending with the boldface $\mathbf{y}_{k_m}$ followed by a single unfilled coordinate, next a continuous block filled with normal font terms followed by a single unfilled position, next again a continuous block filled with normal font terms followed by a single unfilled position, and so on. The continuous filled blocks (including the first one) have strictly decreasing lengths and there are m of them (see the diagram below for $m = 6$, most of the indices are omitted). We do not require that the last unfilled position is single (it may be followed by more unfilled positions).*

$$\mathbf{y}_1 \mathbf{y} \, \mathbf{y} \, \mathrm{y} \, \mathbf{y} \, \mathbf{y} \, \mathrm{y} \, \mathbf{y}_{k_m} \square \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \square \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \square \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \square \, \mathrm{y} \, \mathrm{y} \, \mathrm{y} \square \, \mathrm{y} \, \mathrm{y} \square$$

Proof of the Claim. For $m = 1$ the condition is fulfilled after $k_1 = 1$ steps, so the induction starts. Suppose the claim holds for some $m \geq 1$. The pattern C is repeated periodically with period p_{k_m} and the repetitions cannot overlap (because the lengths of the filled blocks are all different). This implies that the pattern C is contained in $x_y[1, p_{k_m}]$. Let us turn to the first repetition of the pattern C further to the right (call it C'). It starts at position $p_{k_m} + 1$ and clearly, here all symbols are printed in normal font. Notice that at least two preceding positions: p_{k_m} and $p_{k_m} - 1$ are not occupied (because $p_{k_m} = 0 \bmod p_k$ and $p_{k_m} - 1 = p_k - 1 \bmod p_k$ for any $k \leq k_m$, while the positions filled with the symbol y_k have values modulo p_k positive and much smaller than p_k). Now we perform the construction steps $k_m + 1$,

$k_m + 2$, etc., of filling in x_y, until we fill position $p_{k_m} - 1$ (with a boldface symbol $\mathbf{y}_{k'}$ for some $k' > k_m$). Notice that the pattern C' extends at most to position $2p_{k_m}$, which is smaller than $p_{k_m+1} + 1$, so the repeated (normal font) symbols added in these new steps fall to the right of C' (i.e., they do not affect it). In this manner we fill all the unfilled positions within $x_y[1, p_{k_m} - 1]$ creating (together with C') a pattern as required for $m + 1$ in the induction (with $k_{m+1} = k'$). Note that the new initial continuously filled block $x_y[1, p_{k_m} - 1]$ has length $p_{k_m} - 1$, larger than or equal to the length of the pattern C' (perhaps without counting its last empty cell), in particular, for $m > 1$, it is strictly longer than the first completely filled block of C'. For $m = 1$ this also holds, because, in this case, the length of the first (and unique) filled block of C' is 1, while $p_1 - 1 > 1$. The claim is thus proved.

In the following construction steps, the unfilled positions in the pattern C are filled with the boldface symbols $\mathbf{y}_{k_m+1}, \ldots, \mathbf{y}_{k_m+m}$. Because the lengths of the separating normal font blocks strictly decrease, and there are $m - 1$ of them, the first one has length at least $m - 1$, the next one $m - 2$ and so on. Thus, taking for simplicity m to be even, we deduce that each of the boldface terms $\mathbf{y}_{k_m+1}, \ldots, \mathbf{y}_{k_m+m/2}$ is followed (and preceded) by a block of normal font symbols of length at least $m/2$. Renaming $m/2$ as m ends the proof. ∎

It follows from the construction (or can be easily arranged by choosing a subsequence) that $k_{m+1} > k_m + m$ for each $m \geq 1$.

EXAMPLE 7.3. *There exists a one-sided symbolic sequence y whose orbit closure Y has positive topological entropy and supports many invariant measures, yet the associated Toeplitz subshift X_y (the orbit closure of x_y) is strictly ergodic with entropy zero.*

We just need to decide about the contents of the sequence y. Denote $\Lambda = \{0, 1\}$. For each $m \geq 1$ we let $y[k_m + 1, k_m + m]$ be a block A_m and $y[k_m + m + 1, k_{m+1}]$ be the block consisting entirely of zeros. We arrange that the sequence $(A_m)_{m \geq 1}$ of blocks generates a positive entropy subshift Y_0 with many invariant measures (for example, the full shift on two symbols). It is clear that the orbit closure Y of y contains Y_0, hence has positive entropy and many invariant measures.

Let us ignore "accidental" periodic repetitions of symbols in x_y. That is, we will denote by $\mathrm{Per}_k(x_y)$ the set of positions of the $p_{k'}$-periodic repetitions of the symbols $\mathrm{y}_{k'}$ for $k' \leq k$. The density of $\mathrm{Per}_k(x_y)$ so defined is $\sum_{k'=1}^{k} 1/p_{k'}$.

It follows from the general facts concerning Toeplitz subshifts that if μ is an invariant measure on X_y, then μ-almost every $x \in X_y$ has the "non-accidental" periodic part of density $\rho < 1$ and the remaining part (which we

denote by Aper(x), although at the moment we only know it *contains* the true aperiodic part). Clearly, Aper(x) is infinite as it has density $1 - \rho$.

Suppose Aper(x) contains two positions n and $n + m$ such that $x(n) = x(n + m) = 1$. This implies that in x_y there are infinitely many positions n' such that $x(n') = x(n' + m) = 1$ and n' and $n' + m$ both belong to arbitrarily high periodic parts. That is, $x_y(n') = \mathrm{y}_{k'}$ (or $\mathbf{y}_{k'}$) and $x_y(n' + m) = \mathrm{y}_{k''}$ (or $\mathbf{y}_{k''}$), where k', k'' are arbitrarily large, for instance larger than both k and k_m and such that $p_{k'}, p_{k''} > m$. Suppose $k' < k''$ (the other case is symmetric). Shift (if necessary) the window $[n', n' + m]$ to the left by a multiple of $p_{k'}$ so that it starts with the boldface symbol $\mathbf{y}_{k'}$. Say, this is now $[n'', n'' + m]$. Position $n'' + m$ cannot be occupied by $\mathrm{y}_{k'''}$ with $k''' \leq k'$ because then $n' + m$ would also be occupied by the same $\mathrm{y}_{k'''}$ (while it is occupied by $\mathrm{y}_{k''}$ or $\mathbf{y}_{k''}$ with $k'' > k'$). This implies that $n'' + m$ is occupied by some $\mathrm{y}_{k'''}$ or $\mathbf{y}_{k'''}$ with $k''' > k'$. But in that case, since m is smaller than $p_{k'''}$, it must be the first occurrence, i.e., $\mathbf{y}_{k'''}$. Since $\mathbf{y}_{k'} = 1$, k' must belong to an interval $[k_{m'}, k_{m'} + m']$ for some m', and since $k' > k_m$, m' must be larger than or equal to m. This implies that the occurrence of $\mathbf{y}_{k'}$ in x_y (it occurs as $x_y(n'')$) is followed by a block of at least m normal font symbols. This is a contradiction since we have just shown that $x_y(n'' + m)$ is a boldface symbol.

We have proved that if x has an infinite aperiodic part, this part is filled with zeros except perhaps one 1. This immediately implies that X_y has entropy zero and is strictly ergodic (in fact, it is isomorphic to the odometer, like the system of Example 5.1).

Although the above statement already captures the most important properties of the Toeplitz subshift X_y, we have not yet guaranteed that

(1) the sequence x_y is irregular with the density of a "true" aperiodic part equal to $1 - \rho$,
(2) the odometer (G, τ) is a factor of X_y.

All these features must be arranged separately, by delicate modifications of y, yet without destroying what we have already achieved. We will only outline what needs to be done, skipping the tedious and not very interesting details.

We must realize that the construction steps of filling in x_y corresponds to successively defining the associated semicocycle as constant on some clopen subsets $C_k \subset G$. The positions of these subsets are determined by the scheme; they form a dense subset of G and have jointly Haar measure ρ.

For (1) we need to ensure that the cocycle is *discontinuous* at every point of the complementary set $D \subset G$. This can be done by making sure that we assign at least two different values in every neighborhood of every point of D, which can be achieved by modifying (if necessary) the values of y along a

very sparse subsequence, so sparse that it would not affect other properties. Notice that every neighborhood of every point in D contains infinitely many sets C_k, so we can choose an arbitrarily sparse subsequence of these sets which visits all such neighborhoods.

Likewise, for (2) we need the semicocycle to be invariant under no rotation. For this is suffices that we arrange a discontinuity point "unlike any other". This can also be done by very sparse modifications of y. ∎

8. Entropy of Example 6.3. We have eliminated the possibility of an "automatic" proof that the entropy in Example 6.3 is positive just based on the fact that $\boldsymbol{\mu}$ generates a subshift with positive topological entropy. On the other hand, in view of Sarnak's conjecture, we have to check the positivity of its topological entropy.

Proof of positivity of the entropy of Example 6.3. For short, we will call the squares of prime numbers *p-squares*. They will be denoted by $d_1, d_2, \ldots$.

Let B denote an arbitrary block appearing in the subshift generated by $|\boldsymbol{\mu}|$. Let n denote the length of B, which we assume is large. Our goal is to indicate a place (an interval of n consecutive coordinates) where B occurs in $|\boldsymbol{\mu}|$, and in the Toeplitz sequence x_y the number of boldface symbols is close to $n(1 - \rho)$ (i.e., nearly realizes the upper Banach density of such symbols). Recall that ρ is the sum of the inverses of the periods p_k, and is smaller than $3/\pi^2$. First we will argue that finding such places (for all long enough blocks B) suffices for positivity of the topological entropy of X_y.

Indeed, let C denote the block appearing in x_y over this interval. Then B can be reconstructed knowing C and two additional data: the positions of all the normal font symbols in x_y in the interval considered, and the contents of $|\boldsymbol{\mu}|$ at these positions. Since the number of normal font symbols is not larger than $n(\rho + \delta)$, Lemma 7.1 implies that there are at most $2^{n\epsilon}$ possibilities for the distribution of the normal font symbols, and then there are at most $2^{n(\rho+\delta)}$ possibilities for their contents in $|\boldsymbol{\mu}|$. This produces the estimate

$$\#\{B\} \leq \#\{C\} \cdot 2^{n\epsilon} \cdot 2^{n(\rho+\delta)},$$

where $\#\{B\}$ and $\#\{C\}$ denote the cardinalities of blocks of length n in $|\boldsymbol{\mu}|$ and in x_y, respectively. Since $|\boldsymbol{\mu}|$ generates a subshift of entropy $(6/\pi^2) \log 2$, i.e., $\#\{B\}$ is nearly $2^{6n/\pi^2}$, the cardinality $\#\{C\}$ is (ignoring the small terms) nearly $2^{n(6/\pi^2-\rho)}$, which yields positive topological entropy of X_y whenever $\rho < 6/\pi^2$ (while we have even assumed it is smaller than half of that number).

So, we focus on finding an interval as specified at the start of the proof. Find an interval I of length n where B occurs in $|\boldsymbol{\mu}|$ (such an interval exists). The positions of zeros in I can be divided into two classes: first class, coordinates divisible by any of the p-squares $d_1, \ldots, d_L$, where d_L is the largest

p-square smaller than n, and second class—the remaining ones (divisible by larger p-squares). Note that for each p-square $\geq n$ only one of its multiples can occur in I. Now, by the Chinese Remainder Theorem (see e.g. [N, Chap. I]) we can shift the interval I (and call the shifted interval I') so that

(1) the shift is by a multiple of $d_1 \cdots d_K$, where d_K is the largest p-square smaller than $4n^2$, and

(2) zeros of the second class become (after shifting) divisible by some a priori selected large p-squares $e_1, \ldots, e_q$ (to be specified shortly).

The zeros of the first class appearing in $|\boldsymbol{\mu}|$ over the interval I' are precisely the shifted zeros of the first class over I. All zeros of the second class occurring over I correspond (via the shift) to zeros of the second class occurring over I', but the latter interval can have more zeros of the second class (some new zeros divisible by p-squares larger than $4n^2$ and different from $e_1, \ldots, e_q$ can occur here). The configuration of zeros inherited from I is repeated in $|\boldsymbol{\mu}|$ in every interval I'' along an arithmetic progression starting with I' and advancing with step

$$M = d_1 \cdots d_K \cdot e_1 \cdots e_q$$

(this need not be the smallest step, just one which is sure). From now on we will observe only the intervals I'' appearing along this progression. In every such interval the additional zeros (if there are any) must be divisible by p-squares $\geq 4n^2$. Using Lemmas 8.2 and 8.3, provided at the end of the paper, one easily sees that the fraction (among the observed intervals) of intervals where there are additional zeros does not exceed

$$n \left(1 - \prod_{j>K} \left(1 - \frac{1}{d_j} \right) \right) \leq \frac{n}{\sqrt{d_{K+1}}} \leq \frac{1}{2}.$$

In other words, in at least around $1/2$ of the observed intervals in $|\boldsymbol{\mu}|$ there occurs precisely the block B.

It now suffices to arrange that most (a bit over $1/2$ is enough) of these intervals are such that in the Toeplitz sequence x_y there are nearly $n(1-\rho)$ boldface symbols.

Recall that $H = (p_k)$ is the scale used to construct x_y, and for each $k \geq 1$, p_{k+1} is a *proper* multiple of p_k (at least times 2). We are going to mark three "important points" k_1, k_2, k_3 on the axis of the parameter k (these points bear hidden dependence on n, not visible in the notation).

k_1: Let k_1 be the largest k such that $p_k < n$. Notice that $p_k \geq 2^k$ implies $k_1 < \ln n / \ln 2$, which (for large enough n) is smaller than $[n/\ln n]$.

Next "important points" require auxiliary functions. Let N_k be the number of prime factors (counted with multiplicities) of the largest common

divisor of p_k and $d_1 \cdots d_K$. The function $k \mapsto N_k$ is nondecreasing and becomes constant before it reaches $2K + 1$. This implies it eventually lies below the line $k/2$.

k_2: Let k_2 be the largest k (if any) for which $N_k \geq k/2$. Notice that $N_k \leq 2K$ implies $k_2 \leq 4K$, which (for n large) is small compared to n (equal to four times the number of primes smaller than $2n$, i.e., approximately $8n/\ln 2n$). At this point we assume that if $k_2 < 8n/\ln 2n$ (or does not exist at all), then we set $k_2 = [8n/\ln 2n]$. In particular, this guarantees that $k_2 > k_1$.

k_3: Let k_3 be some place, not smaller than k_2, where the function N_k has already reached its maximum. At this moment we define the numbers $e_1, \ldots, e_q$ to be relatively prime to p_{k_3+4q}. We can do it now, because the particular values of $e_1, \ldots, e_q$ have not been used in defining the preceding points.

Now consider a similar auxiliary function M_k defined analogously to N_k with M in place of $d_1 \cdots d_K$. Note that to the right of $k_3 + 4q$ the function N_k does not grow, while M_k may increase by at most $2q$ (possibly even in one jump); however, thanks to the specific choice of $e_1, \ldots, e_q$, this function will never again cross the line $k/2$.

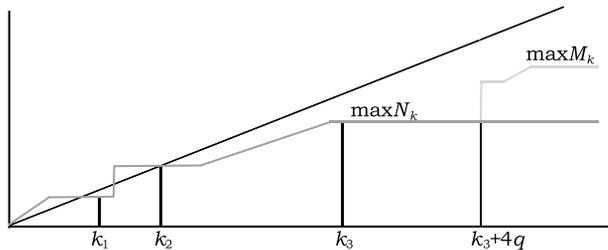The "important points" are shown in the figure below.

Fig. 4. The function N_k is shown in grey, and M_k in light grey.

We divide the positions of normal font symbols in x_y (equivalently, of zeros in z) into arithmetic progressions with periods p_k (along such a progression $x_y(i) = y_k$) and we classify these progressions in three groups:

(1) $k = 1, \ldots, k_1$,
(2) $k = k_1 + 1, \ldots, k_2$,
(3) $k > k_2$.

(The number k_3 is needed only to define $e_1, \ldots, e_q$ and is not used to separate the groups.) We will now analyze the progressions according to this classification.

1. Each progression from the first group occupies in every interval of length n approximately n/p_k positions (at most $2n/p_k$) for the largest k in the group. The union of these progressions occupies not more than $n(\rho + \epsilon)$ positions (for large enough n).

2. Every progression from the second group is represented in every interval of length n by at most one coordinate. Jointly these groups occupy at most $4K$ positions, which does not essentially exceed $8n/\ln 2n$, i.e., negligibly little compared to n.

So far we have been estimating the number of normal font symbols in an interval of length n, and so far it came out close to $n\rho$ (i.e., as we need it). From now on we will estimate the percentage of the "observed" intervals I'' which are disjoint from the progressions belonging to the third class.

3. Consider a k in the third group. The period p_k has with M no more than $k/2$ common prime divisors. However, p_k has at least k prime factors, each at least 2. This implies that $p_k/\gcd(p_k, M) \geq 2^{k/2}$. Applying Lemma 8.1 we conclude that the fraction of intervals I'' intersecting the progressions from the third group does not exceed $n\sum_{k>k_2} 2^{-k/2}$. Since $k_2 \geq 8n \log 2n$, this bound is arbitrarily small for large n.

To summarize, we can arrange that more than half of the observed intervals I'' do not intersect any of the progressions from the third group. In such intervals $|\boldsymbol{\mu}|$ reads B. Combining this with a previous estimate we conclude that there exist intervals in which both $|\boldsymbol{\mu}|$ reads B and in x_y there are nearly $n(1 - \rho)$ boldface symbols, as required. This completes the proof. ∎

Now the missing lemmas.

LEMMA 8.1. *Consider a collection of arithmetic progressions* $A_j = \{kp_j + r_j : k \in \mathbb{N}\}$ *of natural numbers, where* $r_j \geq 0$ *and* $p_j/j \to \infty$ *as* $j \to \infty$. *Let* M *be a natural number and* r *a nonnegative integer. The upper density of the set*

$$\left\{ k \in \mathbb{N} : kM + r \in \bigcup_{j \geq 1} A_j \right\}$$

is less than or equal to $\sum_{j \geq 1} 1/p_j'$, *where* $p_j' = p_j/\gcd(p_j, M)$ *for* $j \geq 1$.

Proof. Given n let J_n denote the maximal j such that $nM + r \geq p_j$. Note that for any n and j,

$$|\{k \in \mathbb{N} : k \leq n, kM + r \in A_j\}| \leq \frac{n}{p_j'} + 1$$

and the set on the left hand side is empty if $j > J_n$. It follows that

$$\frac{1}{n}\left|\left\{k \in \mathbb{N} : k \leq n, kM + r \in \bigcup_{j \geq 1} A_j\right\}\right|$$

$$= \frac{1}{n}\left|\left\{k \in \mathbb{N} : k \leq n, kM + r \in \bigcup_{j=1}^{J_n} A_j\right\}\right|$$

$$\leq \frac{1}{n}\sum_{j=1}^{J_n}|\{k \in \mathbb{N} : k \leq n, kM + r \in A_j\}|$$

$$\leq \frac{1}{n}\left(\sum_{j=1}^{J_n}\frac{n}{p'_j} + J_n\right).$$

Thanks to our assumption on p_j, $J_n/n \to 0$ as $n \to \infty$, thus the assertion follows. ∎

LEMMA 8.2. *In the family of all subsets of $\mathbb{N}$ (or of $\mathbb{Z}$) which have well defined density, the density can be viewed as a finitely additive probability measure. Then any finite collection of periodic sets with relatively prime periods is stochastically independent.*

Proof. Since every periodic set with period p decomposes as a disjoint union of finitely many arithmetic progressions with step p, it suffices to prove the lemma for arithmetic progressions (notice that the density of a progression with step p equals $1/p$). So, let $A_1, \ldots, A_k$ be arithmetic progressions with steps $p_1, \ldots, p_k$. We need to show that the density of their intersection equals $1/p_1 \ldots p_k$. This, however is obvious, because due to the relative primeness, this intersection is an arithmetic progression with step $p_1 \ldots p_k$. ∎

LEMMA 8.3. *For any $k \geq 1$ we have*

$$\prod_{j \geq k}\left(1 - \frac{1}{d_j}\right) \geq 1 - \frac{1}{\sqrt{d_k}}.$$

Proof. Recall that $d_j = q_j^2$, where q_j denote the consecutive primes. Thus

$$\prod_{j \geq k}\left(1 - \frac{1}{d_j}\right) = \prod_{j \geq k}\left(1 - \frac{1}{q_j^2}\right) = \lim_{N \to \infty}\prod_{j=k}^{N}\left(1 - \frac{1}{q_j^2}\right) \geq \lim_{N \to \infty}\prod_{n=q_k}^{N}\left(1 - \frac{1}{n^2}\right)$$

$$= \lim_{N \to \infty}\frac{q_k - 1}{q_k}\cdot\frac{N+1}{N} = 1 - \frac{1}{q_k} = 1 - \frac{1}{\sqrt{d_k}}. \quad\blacksquare$$

References

[AKL]   H. El Abdalaoui, S. Kasjan and M. Lemańczyk, *0-1 sequences of the Thue–Morse type and Sarnak's conjecture*, Proc. Amer. Math. Soc., to appear; arXiv:1304.3587v2.

[ALR1]   H. El Abdalaoui, M. Lemańczyk and T. de la Rue, *On spectral disjointness of powers for rank-one transformations and Möbius orthogonality*, J. Funct. Anal. 266 (2014), 284–317.

[ALR2]   H. El Abdalaoui, M. Lemańczyk and T. de la Rue, *Automorphisms with quasi-discrete spectrum, multiplicative functions and average orthogonality along short intervals*, arXiv:1507.04132 (2015).

[B1]   J. Bourgain, *Möbius–Walsh correlation bounds and an estimate of Mauduit and Rivat*, J. Anal. Math. 119 (2013), 147–163.

[B2]   J. Bourgain, *On the correlation of the Moebius function with rank-one systems*, J. Anal. Math. 120 (2013), 105–130.

[BSZ]   J. Bourgain, P. Sarnak and T. Ziegler, *Disjointness of Möbius from horocycle flows*, in: From Fourier and Number Theory to Radon Transforms and Geometry, in memory of Leon Ehrenpreiss, Developments Math. 28, Springer, 2012, 67–83.

[CDM]   H. Cohen, F. Dress and M. El Marraki, *Explicit estimates for summatory functions linked to the Möbius μ-function*, Funct. Approx. Comment. Math. 37 (2007), 51–63.

[Da]   H. Davenport, *On some infinite series involving arithmetical functions. II*, Quart. J. Math. Oxford 8 (1937), 313–320.

[D]   T. Downarowicz, *Survey of odometers and Toeplitz flows*, in: Algebraic and Topological Dynamics, Contemp. Math. 385, Amer. Math. Soc., 2005, 7–38.

[DG]   T. Downarowicz and E. Glasner, *Isomorphic extensions and applications*, Topol. Methods Nonlinear Anal., to appear; arXiv:1502.06999 (2015).

[DL]   T. Downarowicz and Y. Lacroix *Almost 1-1 extensions of Furstenberg–Weiss type*, Studia Math. 130 (1998), 149–170.

[GL]   A. O. Gelfond and Yu. V. Linnik, *Elementary Methods in the Analytic Theory of Numbers*, Int. Ser. Monogr. Pure Appl. Math. 92, Pergamon Press, 1966.

[G]   B. Green, *On (not) computing the Möbius function using bounded depth circuits*, Combin. Probab. Comput. 21 (2012), 942–951.

[GT]   B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. 175 (2012), 541–566.

[HS]   K. Halupczok and B. Suger. *Partial sums of the Möbius function in arithmetic progressions assuming GRH*, Funct. Approx. Comment. Math. 48 (2013), 61–90.

[KL]   J. Kułaga-Przymus and M. Lemańczyk, *The Möbius function and continuous extensions of rotations*, Monatsh. Math. (2015) (online).

[L]   E. Lehrer, *Topological mixing and uniquely ergodic systems*, Israel J. Math. 57 (1987), 239–255.

[LTY]   J. Li, S. Tu and X. Ye, *Mean equicontinuity and mean sensitivity*, Ergodic Theory Dynam. Systems (2014) (online).

[LS]   Y. Liu and P. Sarnak, *The Möbius function and distal flows*, Duke Math. J. 164 (2015), 1353–1399.

[MR]   C. Mauduit and J. Rivat, *Prime numbers along Rudin–Shapiro sequences*, J. Eur. Math. Soc. 17 (2015), 2595–2642.

[M]   A. F. Möbius, *Über eine besondere Art von Untersuchung des Reihen*, J. Reine Angew. Math. 9 (1832), 105–123.

[MH]   M. Morse and G. A. Hedlund, *Symbolic dynamics II: Sturmian trajectories*,
       Amer. J. Math. 62 (1940), 1–42.
[N]   W. Narkiewicz, *Number Theory*, World Sci., Singapore, 1977.
[P]   K. Prachar, *Primzahlverteilung*, Springer, 1957.
[RR]   O. Ramaré and R. Rumely, *Primes in arithmetic progressions*, Math. Comp. 65
       (1996), 397–425.
[Sa]   P. Sarnak, *Three Lectures on the Möbius Function Randomness and Dynamics*,
       http://publications.ias.edu/sites/default/files/MobiusFunctionsLectures(2).pdf.
[Sch] L. Schoenfeld, *An improved estimate for the summatory function of the Möbius
       function*, Acta Arith. 15 (1969), 221–233.
[W]   A. Walfisz, *Weylsche Exponentialsummen in der neueren Zahlentheorie,* Deutscher
       Verlag Wiss., Berlin, 1963.

Tomasz Downarowicz   Stanisław Kasjan
Institute of Mathematics   Faculty of Mathematics and Computer Science
Polish Academy of Sciences   Nicolaus Copernicus University
Śniadeckich 8   Chopina 12/18
00-656 Warszawa, Poland   87-100 Toruń, Poland
and   E-mail: skasjan@mat.umk.pl
Institute of Mathematics and Computer Science
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: downar@pwr.wroc.pl